

**Tadaiti Ozato Junior**

**Complementação do Seqüenciamento do Genoma do  
*Southern bean mosaic virus*, Isolado São Paulo, Expressão  
da Porção C-Terminal da Polimerase e Produção de Anti-  
soro Policlonal**

**Tadaiti Ozato Junior**

**Complementação do Seqüenciamento do Genoma do  
*Southern bean mosaic virus*, Isolado São Paulo, Expressão  
da Porção C-Terminal da Polimerase e Produção de Anti-  
soro Policlonal**

Dissertação apresentada para obtenção do grau de mestre do Programa de Pós-Graduação em Genética, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista – IBILCE/UNESP

Orientador: Prof. Dr. José Osmar Gaspar

São José do Rio Preto  
2007

Ozato Junior, Tadaiti.

Sequenciamento do genoma do *Southern bean mosaic virus*, isolado São Paulo, expressão da porção C-terminal da polimerase e produção de anti-soro policlonal / Tadaiti Ozato Junior. – São José do Rio Preto: [s.n.], 2007.

71 f. : il. ; 30 cm.

Orientador: José Osmar Gaspar

Dissertação (mestrado) – Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas

1. Fitopatologia. 2. Fitovirologia. 3. *Sobemovirus*. 4. Genomas - Sequenciamento. I. Gaspar, José Osmar. II. Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas III. Título.

CDU – 581.2

Este trabalho foi realizado no Laboratório de Fitovirologia, departamento de Zoologia-Botânica do Instituto de Biociências, Letras e Ciências Exatas de São José do Rio Preto (IBILCE), Universidade Estadual Paulista (UNESP), com auxílio-bolsa CAPES.

**Tadaiti Ozato Junior**

**Complementação do Seqüenciamento do Genoma do  
*Southern bean mosaic virus*, Isolado São Paulo, Expressão  
da Porção C-Terminal da Polimerase e Produção de Anti-  
soro Policlonal.**

COMISSÃO JULGADORA

DISSERTAÇÃO PARA OBTENÇÃO DO GRAU DE MESTRE

Presidente e Orientador: \_\_\_\_\_  
Prof. Dr. José Osmar Gaspar

2º Examinador: \_\_\_\_\_  
Profª. Dra. Cláudia Regina Bonini Domingos

3º examinador: \_\_\_\_\_  
Prof. Dr. Eliezer Rodrigues de Souto

São José do Rio Preto, 16 de fevereiro de 2007.

Dedico este trabalho:

Aos meus pais Tadaiti Ozato e Sonia Regina Ozato, a minha irmã Thamy Ozato que são a minha vida e o motivo de eu estar aqui.

Aos meus amigos quase irmãos

Newton Kenji Nagoshi e Daniel Kiyoshi Nagoshi por me apoiarem com suas amizades preciosas.

Ao meu orientador Prof. Dr. José Osmar Gaspar

Que me deu a oportunidade de pesquisar, me mostrando novos rumos. Sempre me deixou livre para fazer as opções, mas com a certeza de que confiava no meu trabalho e disposto a me corrigir sempre que necessário. O senhor representa o que é ser Orientador.

Obrigado

Vocês são muito especiais e preciosos.

Vencemos juntos!

## **AGRADECIMENTOS**

Agradeço em especial ao meu orientador Prof. Dr. José Osmar Gaspar pela oportunidade, orientação, imensa paciência e exemplo de profissionalismo.

Aos membros participantes da banca examinadora: Prof<sup>a</sup>. Dra. Claudia Regina Bonini Domingos e ao Prof. Dr. Eliezer Rodrigues de Souto.

Ao Prof. Dr. Jorge A. M. Rezende pela ajuda na obtenção do anti-soro.

Aos professores da pós-graduação em Genética que me ajudaram muito e colaboraram bastante para o meu crescimento acadêmico.

À coordenadora da pós em genética Maria Tercília Vilela de Azeredo Oliveira e à vice-coordenadora Claudia Regina Bonini Domingos.

À direção do Instituto de Biociências, Letras e Ciências Exatas (IBILCE) de São José do Rio Preto pelo fornecimento de recursos materiais.

Aos funcionários do IBILCE que de forma direta ou indireta me ajudaram na realização deste trabalho.

Aos meus amigos de laboratório: Ana Cecília Bergamim Pereira (Aninha), Lívia José Regatieri (Biston) e Ricardo Barros Mariuti (Tetchukine) que ajudaram no afazeres do laboratório e me agüentaram nas horas mais difíceis da minha pesquisa. Ao Luiz Gustavo Gardinassi (Bradock) que me ajudou muito no seqüenciamento.

À Priscila Belintani que por meio dela pude conhecer muitas dessas pessoas e graças a ela que me deu a oportunidade de iniciar na pesquisa. Minha eterna orientadora.

Aos meus amigos sempre fiéis Lisandra (Lis), Júlio Borges (Prof. Dr.) e Mário (Dr.) que sempre estiveram ao meu lado durante todo esse tempo e me auxiliaram nas horas mais difíceis com a amizade e opiniões.

À CAPES pelo auxílio financeiro concedido para o desenvolvimento d trabalho.

E a Deus que torna tudo possível!

## SUMÁRIO

<b>INTRODUÇÃO.....</b>	<b>9</b>
<b>OBJETIVOS.....</b>	<b>21</b>
<b>MATERIAL E MÉTODOS.....</b>	<b>22</b>
<b>RESULTADOS .....</b>	<b>32</b>
<b>DISCUSSÃO.....</b>	<b>56</b>
<b>CONCLUSÃO.....</b>	<b>60</b>
<b>RESUMO.....</b>	<b>61</b>
<b>ABSTRACT.....</b>	<b>62</b>
<b>BIBLIOGRAFIA.....</b>	<b>63</b>

## 1. INTRODUÇÃO

O feijão (*Phaseolus vulgaris* L.), leguminosa comum da família das fabáceas (Leguminosae) (CRONQUIST, 1981), constitui uma das culturas mais importantes existentes, sendo consumida por todo o mundo, devido principalmente ao seu valor nutritivo. A cultura do feijão no Brasil assume um grande valor social, pois constitui a base da alimentação da população, sendo fonte de proteína vegetal de baixo custo. A dupla arroz com feijão é o prato principal na alimentação da maioria da população brasileira (CULTURA, 2000). Dados da última safra (2004/2005) mostram que o Brasil produziu cerca de 2,9 milhões de toneladas de feijão, sendo a região sudeste a maior produtora (CONAB, 2006).

A produtividade da cultura é grandemente afetada pela ocorrência de doenças que diminuem sensivelmente a produção e, conseqüentemente, reduzem a sua oferta, provocando um aumento nos preços de mercado (CULTURA, 2000). Estas doenças têm exercido um papel relevante na baixa produtividade de feijoeiros no Brasil e em outros países latino-americanos (JAYASINGHE, 1982; COSTA & COSTA, 1983). As mais numerosas são ocasionadas por fungos seguindo-se as de etiologia viral, várias com expressiva importância econômica (GAMEZ, 1977).

No Brasil, foram descritas mais de dez viroses em feijoeiro (COSTA *et al.*, 1972; BIANCHINI *et al.*, 1997), citando-se aquela causada pelo *Southern bean mosaic virus* (SBMV), do gênero *Sobemovirus* (HULL & FARGETTE, 2005). Este vírus tem uma restrita gama de hospedeiras, confinada quase que exclusivamente a espécies da família das leguminosas, sendo algumas de interesse econômico como o feijoeiro comum e a soja. Muito embora o SBMV não seja no momento um problema fitossanitário, observações de campo no Estado do Paraná têm indicado a sua presença em infecções mistas principalmente com o *Bean golden mosaic virus* (Gênero *Begomovirus*, Família *Geminiviridae*), levando à possibilidade de possíveis problemas futuros para a cultura do feijoeiro (BIANCHINI, A. Comunicação pessoal).

## II GÊNERO *Sobemovirus*

### A. PROPRIEDADES BIOLÓGICAS

O nome *Sobemovirus* é originado das iniciais da espécie tipo do gênero *Southern bean mosaic virus* e o gênero compreende 13 espécies definitivas, sendo elas: *Blueberry shoestring virus* (BSSV), *Cocksfoot mottle virus* (CoMV), *Lucerne transient streak virus* (LTSV), *Rice yellow mottle virus* (RYMV), *Ryegrass mottle virus* (RGMoV), *Sesbania mosaic virus* (SeMV), *Solanum nodiflorum mottle virus* (SNMoV), *Southern bean mosaic virus* (SBMV), *Southern cowpea mosaic virus* (SCPMV), *Sowbane mosaic virus* (SoMV), *Subterranean clover mottle virus* (SCMoV), *Turnip rosette virus* (TRoV) e *Velvet tobacco mottle virus* (VTMoV) (HULL & FARGETTE, 2005). As espécies SBMV e SCPMV eram a princípio consideradas como isolados do SBMV (respectivamente isolados feijão e caupí) (FAUQUET & MAYO, 1999). Há ainda quatro possíveis espécies no gênero, cujas características são similares aos sobemovírus, mas que necessitam de informações adicionais para serem incluídas. São elas: *Cocksfoot mild mosaic virus* (CMMV), *Cynosurus mottle virus* (CnMoV), *Ginger chlorotic fleck virus* (GCFV) e *Rottboellia mottle virus* (RoMoV) (HULL & FARGETTE, 2005).

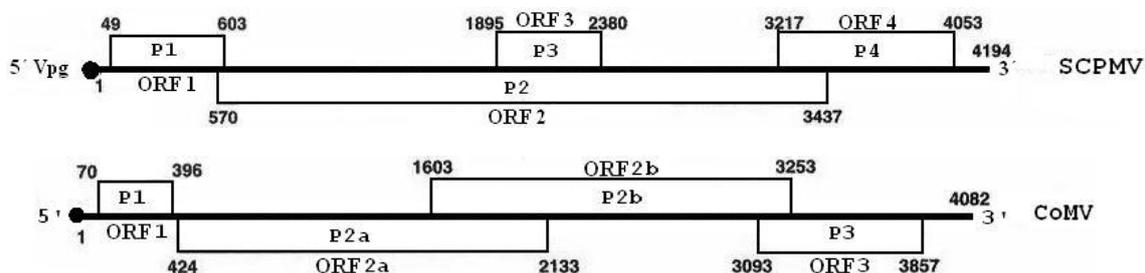
A maioria das espécies dos sobemovírus apresenta distribuição geográfica limitada, mas algumas espécies são encontradas em várias partes do mundo (HULL, 1988). Os sobemovírus infectam plantas tanto mono quanto dicotiledôneas, mas a gama de hospedeiras de cada espécie é relativamente restrita. Os sintomas induzidos são principalmente mosaico e mosqueado e, em geral, causam infecção sistêmica invadindo quase que todos os diferentes tipos de tecidos das plantas hospedeiras (HULL, 1988). São transmitidos mecanicamente, por vetores e por sementes. Para a maioria das espécies (CoMV, RYMV, SNMoV, SBMV, SCPMV, TRoV) a transmissão é feita por coleópteros enquanto o BSSV é transmitido por afídeos, o VTMoV por mirídeos e o SoMV por dípteros, homópteros e hemípteros (HULL, 1988; TAMM & TRUVE, 2000b).

## B. ORGANIZAÇÃO DO GENOMA

Os sobemovírus possuem partículas isométricas (28-30nm) contendo RNA genômico de 4-4,5kb com polaridade positiva e proteína capsidial com massa molecular de 29-30 kDa (SHEPHERD, 1971; SEHGAL, 1981). Uma pequena proteína denominada de proteína ligada ao genoma viral (“viral genome linked-protein”, VPg) é covalentemente ligada à extremidade 5' do RNA enquanto a extremidade 3' é destituída de cauda poliadenilada (GHOSH *et al.*, 1979). Existe também, um RNA subgenômico de aproximadamente 1,0 Kb responsável pela tradução da proteína capsidial (RUTGERS *et al.*, 1980; GHOSH *et al.*, 1981).

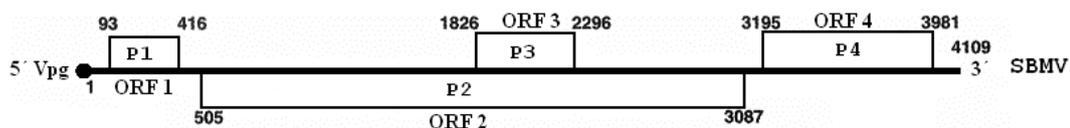
O genoma dos sobemovírus possui quatro cadeias abertas de leitura (“Open reading frame”, ORF) (Figura 1). Todos os sobemovírus seqüenciados contêm uma pequena ORF1 na extremidade 5'-terminal do genoma, que codifica a possível proteína do movimento célula-a-célula, e uma ORF4 na extremidade 3'-terminal que codifica a proteína capsidial. Com base nas diferenças de organização na parte central do genoma (que codifica uma poliproteína), os sobemovírus podem ser divididos em dois grupos: aqueles semelhantes ao SCPMV e aqueles semelhantes ao CoMV. A poliproteína do SCPMV é codificada por uma longa e contínua ORF2, contendo também uma região codificadora interna, ORF3. Organização semelhante do genoma já foi descrita para o isolado Arkansas do SBMV (SBMV-Ark) (LEE & ANDERSON, 1998), isolado “Ivory Coast” do RYMV (RYMV-CI) (YASSI *et al.*, 1994) e LTSV (FEFFRIES *et al.*, GenBank U31286).

Por outro lado, na organização genômica do CoMV falta a ORF2 contínua e uma região codificadora similar à ORF3 do SCPMV. Em vez disso, CoMV possui duas ORFs sobrepostas, ORF2a e ORF2b, e a poliproteína é expressa através de um mecanismo de mudança de fase (“frameshift”) ribossomal -1 (MÄKINEN *et al.*, 1995; RYABOV *et al.*, 1996).



**Figura 1:** Organização genômica do *Southern caupi mosaic virus* (SCPMV) e *Cocksfoot mottle virus* (CoMV); (●): VPg, proteína covalentemente ligada à extremidade 5'-terminal.

A única exceção com relação a esta subdivisão dos sobemovírus é o genoma do SBMV seqüenciado por Othman & Hull (1995). Neste, as ORFs 1, 2 e 4 não se sobrepõem como acontece com o SBMV-Ark e SCPMV (Figura 2).



**Figura 2:** Organização da genômica do *Southern bean mosaic virus* descrito por Hull & Fargette (2005), (●): VPg, proteína covalentemente ligada à extremidade 5'-terminal.

Diferentemente do SBMV, o isolado SBMV-Ark contém quatro ORFs que se sobrepõem, tornando-o mais similar em organização genômica (mas não em seqüência de aminoácidos de suas proteínas) ao SCPMV (LEE & ANDERSON, 1998). Foi proposto que essa diferença na organização genômica entre o SBMV e SBMV-Ark resulta ou de mutações ou de erro no seqüenciamento do SBMV.

### C. CITOPATOLOGIA

Os efeitos citopáticos mais evidentes causados por espécies de sobemovírus são a distribuição das partículas virais no citoplasma e a formação de grandes agregados cristalinos no citoplasma e vacúolos (FRANCKI *et al.*, 1985). As células infectadas mostram extensa vacuolização citoplasmática, sendo que algumas espécies invadem células do floema e xilema (SBMV, RYMV) e partículas virais podem também ser encontradas em membranas das pontuações das paredes primárias (RYMV) (DE ZOETEN & GAARD, 1969; OPALKA *et al.*, 1998; WEINTRAUB & RAGETLI, 1970). Partículas virais são também

encontradas no núcleo celular de plantas infectadas por sobemovírus. Yassi *et al.* (1994) verificaram que a parte N-terminal da proteína capsidial do RYMV contém uma seqüência que é idêntica ao “motivo de localização nuclear do tipo bipartido” (DINGWALL & LASKEY, 1991). Um motivo similar foi encontrado na região N-terminal da capa proteica de vários outros sobemovírus (MÄKINEN *et al.*, 1995). Estes resultados podem explicar a presença de partículas virais dos sobemovírus no núcleo celular durante o processo de infecção. Fora estas observações, nenhum determinante molecular tem sido associado com a localização tecido-específica ou subcelular das partículas dos sobemovírus. Nada é conhecido sobre a localização subcelular das proteínas não estruturais deste gênero (TAMM & TRUVE, 2000b).

#### **D. PRODUTOS GÊNICOS E SUAS POSSÍVEIS FUNÇÕES**

Os RNAs de vários sobemovírus já foram traduzidos *in vitro* em sistemas de “lisado de reticulócitos de coelho” e de “extrato de germe de trigo”. Os RNA do SBMV e SCPMV induziram a tradução de quatro proteínas nesses sistemas: 105 kDa, 75 kDa, 29 kDa e 14 kDa para o SBMV e 100 kDa, 70 kDa, 30 kDa e 20 kDa para o SCPMV (MANG *et al.*, 1982; SALERNO-RIFE *et al.*, 1980). Quatro polipeptídeos foram também traduzidos a partir dos RNA do CoMV (MÄKINEN *et al.*, 1995), LTSV (MORRIS-KRSINICK & FORSTER, 1983), SNMoV (KIBERSTIS & ZIMMERN, 1984) e TRoV (MORRIS-KRSINICK & HULL, 1981), apresentando somente ligeiras diferenças na massa molecular.

Estudos têm demonstrado que as proteínas de 100 e 70 kDa do SCPMV são relacionadas e traduzidas a partir do RNA genômico enquanto a proteína de 20 kDa é presumivelmente codificada pela ORF1 (MANG *et al.*, 1982; WU *et al.*, 1987). Já a poliproteína codificada pela ORF2 do SCPMV é processada por clivagem proteolítica para dar o produto de tradução de 70 kDa (WU *et al.*, 1987). A proteína de 30 kDa (proteína capsidial) é traduzida a partir de um pequeno RNA subgenômico (sgRNA) que já foi encontrado em tecidos infectados por sobemovírus bem como em partículas virais (RUTGERS *et al.*, 1980; WEBER & SEHGAL, 1982; RYABOV *et al.*, 1996; TAMM *et al.*, 1999). Estudos com o

CoMV, utilizando a síntese de proteínas *in vitro* e imunoprecipitação com anticorpos específicos, mostraram que as proteínas de 12 kDa, 71 kDa e 100 kDa são sintetizadas a partir do RNA genômico do vírus. A proteína 12 kDa é produzida a partir da ORF1, a proteína 71 kDa a partir da ORF2a e a 100 kDa é a poliproteína produzida a partir das ORFs 2a e 2b. A proteína de 34 kDa (proteína capsidial) é produzida a partir de um sgRNA. A seguir, os produtos gênicos individuais, seus mecanismos de expressão e possíveis funções são mais detalhados.

### **D.1. Proteína P1.**

Todos os sobemovírus caracterizados codificam uma pequena proteína a partir da ORF1, mas estas apresentam seqüências e massas moleculares diferentes (11,7 a 24,3 kDa) e não são relacionadas com qualquer outra proteína conhecida. Análises de mutantes incapazes de produzir P1 ou mutantes produtores de proteínas truncadas indicaram que a P1 do RYMV ou SCPMV não é necessária para a replicação viral, porém é necessária para o movimento célula-a-célula e movimento sistêmico (BONNEAU *et al.*, 1998; SIVAKUMARAN *et al.*, 1998). Proteína P1 recombinante do CoMV foi capaz de interagir com RNA de fita única (ssRNA) numa maneira independente da seqüência (TAMM & TRUVE, 2000a). Assim, levando-se em conta que nenhum outro produto gênico dos sobemovírus estaria relacionado com o movimento célula-para-célula ou movimento sistêmico do vírus, é proposto que a proteína P1 representa a proteína do movimento dos sobemovírus. Foi também descrito que a proteína P1 do RYMV funciona como supressora do mecanismo de silenciamento gênico (“posttranscriptional gene silencing”, PTGS) (VOINNET *et al.*, 1999).

### **D.2. Poliproteína**

Os sobemovírus caracterizados codificam uma proteína com massa molecular de cerca de 100 kDa. As estruturas genômicas do SBMV, SBMV-Ark, SCPMV, RYMV e LTSV permitem a síntese desta proteína a partir de uma única ORF contínua. Por outro lado, a metade C-terminal da poliproteína do CoMV e

RYMV é codificada por uma ORF2b separada e é traduzida como parte da poliproteína pelo mecanismo de mudança de fase (“frameshift”) ribossomal -1 (MÄKINEN *et al.*, 1995). A tradução da poliproteína dos sobemovírus não é iniciada a partir de RNA subgenômico (nunca detectado em tecidos infectados), mas sim a partir do próprio RNA genômico. Assim, as ORFs 1 e 2 são traduzidas a partir de seus respectivos códons de iniciação AUG. Produtos de tamanho esperado de ambas as ORFs foram observados após tradução dos RNAs genômicos do SCPMV (SIVAKUMARAN *et al.*, 1998) e do CoMV (TAMM *et al.*, 1999). Na poliproteína dos sobemovírus, pelo menos três domínios funcionais podem ser encontrados: um para serino protease, um para VPg e um para RNA polimerase dependente de RNA (RpRd) (GORBALENYA *et al.*, 1988; LEE & ANDERSON, 1998; MÄKINEN *et al.*, 1995; OTHMAN & HULL, 1995; RYABOV *et al.*, 1996; WU *et al.*, 1987; YASSI *et al.*, 1994). Não é ainda conhecido se estas são as únicas funções da poliproteína.

#### **D.2.a. Serino Protease**

A possível protease é similar às cisteína proteases encontradas nos picornavírus (Família *Picornaviridae*) e é característica de todos os sobemovírus, polerovírus (Família *Luteoviridae*), *Pea enation mosaic virus* (PEMV, Gênero *Enamovirus*) e *Mushroom bacilliform virus* (MBV, Gênero *Barnavirus*). O motivo serino protease está localizado no terço N-terminal da poliproteína codificada pela ORF2 do SBMV, SBMV-Ark, SCPMV, RYMV e LTSV e é codificada pela ORF2a no caso do CoMV (GORBALENYA *et al.*, 1988). O possível sítio catalítico H-D-S (Histidina, posição 181-Ácido Aspártico, posição 216-Serina, posição 284) encontrado no SCPMV é também conservado no SeMV e outros sobemovírus (LOKESH *et al.*, 2001). Mais recentemente, foi demonstrado que a serino protease do SeMV possui atividade proteolítica atuando em *cis* e *trans*, além de possuir possíveis hélices que estariam associadas à ligação com membranas (SATHESHKUMAR *et al.*, 2004). Com exceção do LTSV, todos os produtos da ORF2 dos sobemovírus mostraram a presença de tais hélices sugerindo que elas podem ter função de ligação às membranas. Em muitos vírus

de RNA de polaridade positiva, foi observado que a replicação e montagem das partículas virais ocorrem em associação com membranas (CARETTE *et al.*, 2002; DEN BOON *et al.*, 2001).

#### **D.2.b. Proteína ligada ao genoma (“viral genome-linked protein”, VPg)**

Uma proteína de 12 kDa foi descrita como covalentemente ligada à região 5'-terminal do genoma do SBMV (GHOSH *et al.*, 1981) e uma de 10 kDa ligada ao genoma do SCPMV (MANG *et al.*, 1982). A posição da região codificadora da VPg no genoma dos sobemovírus difere daquela que é característica de muitas famílias de vírus de RNA, incluindo-se *Picornaviridae* e *Comoviridae*, isto é, VPg-Protease-RpRd (KOONIN & DOLJA, 1983). No caso dos sobemovírus a seqüência dentro da poliproteína é Protease-VPg-RpRd, similar ao também encontrado no *Potato leaf-roll virus* (Gênero *Polerovirus*), PEMV e MBV (REVILL *et al.*, 1999; VAN DER WILK *et al.*, 1989; WOBUS *et al.*, 1998).

Em todas as espécies de vírus desse gênero, as proteínas VPg possuem uma seqüência conservada de aminoácidos (WAD ou WGD; W= Triptofano, A= Alanina, D, Ácido aspártico; G= Glicina) seguida de uma região rica em D ou E (E= Ácido glutâmico) (MÄKINEN *et al.*, 1995), sendo que este motivo é a única seqüência conservada entre as proteínas VPg dos sobemovírus (VAN DER WILK *et al.*, 1997; 1998).

#### **D.2.c. RNA polimerase dependente de RNA (RpRd)**

A região C-terminal da poliproteína dos espécies virais desse gênero codifica uma possível RpRd com base na presença do motivo GDD e de motivos conservados que a este circundam e que são característicos das RNA polimerases dependentes de RNA (KOONIN, 1991; KOONIN & DOLJA, 1983). As possíveis RpRd dos sobemovírus mostram extensivas similaridades com RdRp de vários outros vírus de RNA de polaridade positiva, tais como polerovírus, enamovírus e barnavírus (TAMM & TUVE, 2000b).

Relativamente pouco é conhecido sobre os sinais de replicação necessários para o início da síntese das fitas de polaridade positiva e negativa dos

sobemovírus. Seqüências similares foram encontradas na porção inicial da região 5'-terminal de vários sobemovírus, variando de ACAAAA, ACAA e CACAAA (HACKER & SIVAKUMARAN, 1997; LEE & ANDERSON, 1998; YASSI *et al.*, 1994). Estes resultados indicam que RNA genômicos e subgenômicos dos sobemovírus iniciam com seqüências muito similares e que provavelmente são importantes para a replicação viral. Seqüências similares foram também encontradas na região 5'-terminal do RNA genômico e subgenômico do *Red clover necrotic mosaic virus* (Gênero *Dianthovirus*, Família *Tombusviridae*) (XIONG & LONMEL, 1989; ZAVRIEV *et al.*, 1996) e MBV (REVIL *et al.*, 1994). A conservação desta seqüência na porção inicial da região 5'-terminal dos RNA genômicos e subgenômicos de diversos vírus pertencentes a diferentes gêneros e famílias sugerem que ela (ou sua seqüência complementar na fita de polaridade negativa) pode ter importante função na síntese do RNA viral. Entretanto, o papel dessas seqüências ainda não foi testado em laboratório.

### **D.3. Proteína P3**

Os genomas do SCPMV, SBMV-Ark, RYMV e LTSV possuem uma pequena ORF3 incluída dentro da ORF2 numa cadeia aberta de leitura -1. É possível que a ORF3 seja traduzida por um mecanismo de mudança de fase (“frameshift”) ribossomal -1 similar àquele proposto para a ORF2b do CoMV, mas isso ainda não foi testado experimentalmente (MÄKINEN *et al.*, 1995; TAMM *et al.*, 1999). Se verdadeiro, o produto de tradução *in vitro* de 70 kDa do SCPMV poderia representar uma proteína de fusão entre a ORF2 e a ORF3, uma vez que a massa molecular calculada para a proteína de fusão da ORF2-ORF3 do SCPMV seria 65,8 kDa (TAMM *et al.*, 1999). Assim, o mecanismo de tradução, bem como a função ou funções desta ORF são, até o presente, desconhecidas.

#### **D.4. Proteína capsidial**

As proteínas da capa dos sobemovírus são codificadas pelas ORFs localizadas nas ORFs 3'-terminais, seja a ORF4 no caso dos genomas do tipo SCPMV ou ORF3 para o tipo CoMV (HULL & FARGETTE, 2005). No genoma do SCPMV o sítio de ligação da proteína capsidial é localizado na região codificadora da protease na ORF2 (WU *et al.*, 1987). Nos sobemovírus, esta é a região mais conservada dentro da ORF2. Entretanto, estes dados não demonstram diretamente que a ligação da proteína capsidial nesta região serve para a montagem das partículas do SCPMV (HACKER, 1995).

Estudos com diversos sobemovírus têm indicado que a proteína capsidial é essencial tanto para o movimento célula-a-célula como para o movimento sistêmico do vírus (BRUGIDOU *et al.*, 1995; SIVAKUMARAN *et al.*, 1998). Além disso, evidência direta que a proteína capsidial determina a gama de hospedeiras sistêmicas dos sobemovírus foi obtida pela caracterização de um mutante de quebra de resistência do SBMV-Ark, o SBMV-S (LEE & ANDERSON, 1998). O SBMV-S é capaz de infectar sistemicamente cultivares de feijão onde o tipo selvagem SBMV-Ark infecta apenas localmente. Análises do genoma indicaram 7 diferenças nas seqüências dos genomas mas apenas 4 mudanças na seqüência de aminoácidos deduzidos. Três dessas mudanças foram identificadas na proteína capsidial podendo ter um efeito na interação RNA-proteína capsidial, necessária para a correta montagem e desmontagem das partículas e, desse modo, determinar o movimento a longa distância do vírus na planta hospedeira.

### **III. O *Southern bean mosaic virus* no Brasil**

O nome *Southern bean mosaic virus* deve-se ao fato de o vírus ter sido primeiramente descrito infectando feijoeiros na região sul dos Estados Unidos da América (ZAUMEYER & HARTER, 1942). Atualmente, sabe-se que o SBMV apresenta-se amplamente disseminado, sendo encontrado também em outros países como México, Nicarágua, Venezuela, Gana, Nigéria, Senegal, França,

Índia, Paquistão e Espanha (HULL & FARGETTE, 2005; TAMM & TRUVE, 2000b, VERHOEVEN *et al.*, 2003 e SEGUNDO *et al.*, 2004).

No Brasil, um isolado do SBMV foi detectado pela primeira vez no Distrito Federal (SBMV-DF) (CUPERTINO *et al.*, 1982) e, posteriormente, mas ainda nos anos 80, foi também encontrado no Estado de São Paulo (SBMV-SP) (COSTA, A. S., Dados não publicados).

Os besouros crisomelídeos *Diabrotica speciosa* e *Ceratoma arcuata* são os principais vetores do SBMV no Brasil (COSTA & BATISTA, 1979, SILVEIRA JR. *et al.*, 1983 e MEYER *et al.*, 1992), sendo que *C. arcuata* foi identificado como o vetor mais eficiente na transmissão desse vírus (SILVEIRA JR. *et al.*, 1983), porém as relações vírus/vetor ainda não são bem conhecidas (MEYER *et al.*, 1992). Também foi descrita a transmissão do SBMV pelas larvas de *C. arcuata*, do sistema radicular de plantas infectadas para o de plantas saudias, em uma taxa de 5,1% (MEYER *et al.*, 1993). Esta taxa é pequena se comparada com a taxa de transmissão pelo inseto adulto (58,3%) (SILVEIRA JR. *et al.*, 1983), no entanto, as larvas constituem população mais numerosa, o que torna a transmissão substancialmente importante (MEYER *et al.*, 1993). Isso evidencia a possibilidade de transmissão do vírus através do solo, fato importante para a epidemiologia do SBMV.

Em 2002, um isolado do SBMV foi encontrado no Estado do Paraná (Gasparin, 2002). No entanto, devido à facilidade de transmissão do SBMV por besouros crisomelídeos (SILVEIRA JR. *et al.*, 1983, MEYER *et al.*, 1992), pode-se aventar que a distribuição geográfica do vírus pelo Brasil seja muito mais ampla.

O isolado descrito no Estado de São Paulo (SBMV-SP) teve algumas de suas propriedades determinadas. Através da análise de RNA de fita dupla (dsRNA), evidenciou-se a existência de somente duas espécies de dsRNA de 4,2 kb e 1,0 kb que respectivamente corresponderiam ao RNA genômico e ao RNA subgenômico, como as formas replicativas do vírus (MOREIRA & GASPAR, 2002). O genoma do SBMV-SP foi parcialmente seqüenciado nas ORF 1 (ESPINHA *et al.*, 2004) e ORF 4 (ESPINHA *et al.*, 2005), mostrando através das

relações filogenéticas e dos dados sobre identidade e similaridade que o isolado de São Paulo está proximamente relacionado com o isolado americano de Arkansas (SBMV-Ark) (LEE & ANDERSON, 1998).

## **2. OBJETIVOS**

Os objetivos deste trabalho foram o seqüenciamento e a caracterização molecular das ORF 2 e 3 do genoma do SBMV-SP, completando o seqüenciamento de todo o genoma do isolado de São Paulo. Além disso, também constituiu o objetivo deste projeto a expressão da porção C-terminal da RNA Polimerase RNA Dependente para a obtenção de anti-soro policlonal para essa proteína.

### **3. MATERIAL E MÉTODOS**

#### **3.1 Inoculação de Plantas.**

O inóculo para transmissão do vírus para feijoeiro (*Phaseolus vulgaris* L.) “Jalo” foi preparado através da maceração de plantas infectadas em tampão fosfato de potássio 0,02M pH 7,4 contendo sulfito de sódio 0,02M. O homogeneizado foi manualmente aplicado com um abrasivo (Cellite, Sigma) nas folhas de feijoeiro.

#### **3.2. Purificação do Vírus.**

O SBMV-SP foi purificado seguindo-se o protocolo descrito por Moreira e Gaspar (2002) em que foram utilizadas folhas de *Phaseolus vulgaris*, L. frescas e que apresentaram sintomas do SBMV-SP. Folhas foram trituradas em liquidificador tamponadas com tampão fosfato 0,1M pH 7,2 com 0,75% de sulfito de sódio. O homogeneizado foi filtrado em gaze e o suco clarificado com clorofórmio (0,5ml/g). A fase tamponada foi coletada e seu pH ajustado para 4,8 com ácido acético glacial para precipitação dos contaminantes e então submetida a uma centrifugação de 12.000 xg por 5 min. O vírus foi precipitado da fase tamponada com a adição de polietileno glicol 8000 (PEG) 10% e cloreto de sódio 1% sob agitação de 2 horas à 4°C. Nova centrifugação foi realizada a 16.000 xg por 10 min., o sedimento foi ressuscitado em tampão KH<sub>2</sub>PO<sub>4</sub> 0,01M, pH 7,2 sob agitação de 30 min. a 4°C. Após centrifugação realizada a 16.000 xg por 10 min. e o sobrenadante foi coletado e ultracentrifugado por 1h e 45 min. a 100.000 xg. O sedimento foi ressuscitado em 1ml de tampão KH<sub>2</sub>PO<sub>4</sub> 0,01M, pH 7,2 e centrifugado novamente por 2h e 30 min. a 150.000 xg em gradiente linear de sacarose (10-50%) em tampão KH<sub>2</sub>PO<sub>4</sub> 0,01M pH 7,2. A banda obtida após a centrifugação foi coletada e diluída no mesmo tampão anterior e centrifugada por 3h a 35.000 rpm. O sedimento foi recuperado em 1ml de KH<sub>2</sub>PO<sub>4</sub> 0,01M, pH 7,2.

### 3.3. Espectrofotometria.

O vírus purificado foi submetido à leitura em espectrofotômetro sob comprimento de onda de 260 e 280nm (luz ultravioleta) para determinação da concentração viral, utilizando o coeficiente de extinção do SBMV cujo valor é 6,0 em leitura a 260nm (SEHGAL, 1981).

### 3.4. Extração de RNA das Partículas Virais.

A extração do RNA das partículas virais foi realizada através da utilização do “RNeasy Plant Mini kit” de acordo com instruções do fabricante (Qiagen).

### 3.5. Síntese do DNA Complementar (cDNA).

Os cDNAs foram sintetizados para a amplificação dos fragmentos da ORF 2 e foram produzidos a partir do RNA viral, utilizando um volume de cada reação de 25µL que continha tampão (Tris-HCl 50mM pH 8,3; KCl 75mM); DTT 10mM; MgCl<sub>2</sub> 3mM; 0,2mM de cada dNTP (dGTP, dATP, dCTP e dTTP); 1 unidade de inibidor de RNase, oligonucleotídeos anti-senso descritos abaixo (3R e 2R), 6µg de RNA viral e foi desnaturado à 95°C por 5 minutos; 200 unidades de transcriptase reversa “Superscript” (Invitrogen). A reação foi incubada por 2 horas à 42°C.

### 3.6. Amplificação por Reação em Cadeia de Polimerase (“Polimerase Chain Reaction” – PCR).

Para amplificar por PCR os 5 fragmentos para seqüenciamento da ORF 2 foram utilizados os seguintes oligonucleotídeos senso (F) e anti-senso (R):

1F	5'-GTGGTAGACGGCAAGATGGT-3'	1R	5'-TGGCTTTCTCGTTGATTTC-3'
2F	5'-TACTCCAAGCAGGAAAGT-3'	2R	5'-AATAAGCTCAGCCATAAG-3'
3F	5'-TCGGATGGCACACTCATAGA-3'	3R	5'-TCAGGACTGTCACACCTCCA-3'
4F	5'-GTATCATCCAGGCCGCTCAC-3'	4R	5'-TGGAATTGTTCCAAACATGG-3'
5F	5'-CTCAGACCATCGGTTTTGGT-3'	5R	5'-GGCTGCTTCTCCTTCTTTGA-3'

\* Oligonucleotídeos desenhados a partir da seqüência do SBMV-Ark depositada no Genbank (nº acesso: NC004060).

Os tamanhos esperados dos fragmentos amplificados a partir dos oligonucleotídeos acima são:

**1F/1R:** 498 pares de base (pb)

**2F/2R:** 870 pb

**3R/3F:** 693 pb

**4R/4F:** 598 pb

**5R/5F:** 834 pb

Para amplificar por PCR o fragmento que corresponde a porção C-terminal da polimerase, que foi utilizado para o processo de expressão, o seguinte par de oligonucleotídeos foi utilizado:

- Exp-F: 5'-GGATCC GCC CAG AAT AGG T-3'
- Exp-R: 5'-AAGCTTTC ACA CCT CCA GCC-3'

Este par de oligonucleotídeos apresenta sítios de restrição para as endonucleases *Bam* HI (GGATCC) e *Hind* III (AAGCTT) e o produto de amplificação utilizando esse par de oligonucleotídeos tem o tamanho esperado de 686 pb.

O volume de cada reação da PCR foi de 25µL consistindo de tampão (Tris-HCl 20mM pH 8,4; KCl 50mM); 5mM de MgCl<sub>2</sub>; 0,2mM cada dCTP, dGTP, dATP e DTTP; 10µg/mL de cada oligonucleotídeo (senso e anti-senso); 1µL da mistura da reação de cDNA e 2 unidades de *Taq* DNA-polimerase (Invitrogen). O programa da PCR foi realizado em um termociclador MJ Research que consiste em uma desnaturação inicial de 94°C por 5 minutos e ciclos de 1 minuto de desnaturação à 94°C, 1 minuto de anelamento à 50°C e 1 minuto de extensão à 72°C. Após 35 ciclos e uma extensão final por 10 minutos à 72°C. Os fragmentos de DNA amplificados foram visualizados em gel de agarose 1%, sob luz ultravioleta, em tampão TAE (Tris-HCl 0,04M; acetato de sódio 0,02M e EDTA 0,001M pH 8,3) a uma voltagem de 80V e corado com brometo de etídio (0,05µg/mL) (SAMBROOK *et al.*, 1989).

### 3.7. Clonagem dos Fragmentos Amplificados da ORF 2 (Seqüenciamento).

Após a purificação dos fragmentos amplificados foi realizada a reação de ligação ao vetor pDrive (Qiagen). Para a reação de ligação foram utilizados o produto do PCR, solução salina e vetor pDrive (Qiagen) conforme orientação do fabricante.

Após a reação de ligação procedeu-se a transformação do recombinante em *Escherichia coli* linhagem TOP 10 (Invitrogen), utilizando-se 50µL de células competentes e 3µL de vetor ligado. Essa mistura permaneceu no gelo durante 1 hora e então, foi submetida ao choque de temperatura à 42°C por 60 segundos. Em seguida, foram adicionados 250µL de meio líquido SOC (meio SOB (peptona, extrato de levedura, KCl e NaCl), 10mM de glicose e 10mM de Mg<sup>2+</sup>) e as bactérias foram submetidas a regeneração (1 hora à 37°C com agitação a 200 rpm). Posteriormente, as bactérias foram plaqueadas em meio sólido seletivo (LB + ampicilina 100µg/mL), contendo ainda 40µL de X-Gal 20mg/mL (5-bromo-4-cloro-3-indolil-β-D-Galactopiranosídio), utilizado para a detecção colorimétrica da atividade da enzima β-galactosidase. Após crescimento a 37°C por 16 horas, as colônias brancas foram isoladas e multiplicadas em outra placa de meio sólido por incubação durante aproximadamente 20 horas. Estas colônias foram testadas por PCR para comprovar a existência do inserto. As colônias que apresentaram o inserto foram inoculadas em tubos individuais de 5mL de meio líquido seletivo (LB + ampicilina 100µg/mL) sob agitação constante à 37°C para posterior extração do DNA plasmidial.

A extração do DNA recombinante foi feita utilizando-se o “PureLink Quick Plasmid Miniprep kit” (Invitrogen), seguindo instruções do fabricante. Em linhas gerais, 5 ml de cultura líquida foram centrifugados por 1 minuto a 12.000 xg e o sobrenadante descartado. O precipitado foi ressuspenso em tampão de suspensão e, em seguida foi acrescentado tampão de lise. O tubo foi manualmente invertido 5 vezes e incubado à temperatura ambiente por 5 minutos. A esta mistura foi adicionado tampão de precipitação e centrifugado a 12.000 xg por 10 minutos.

O sobrenadante resultante da centrifugação foi transferido para coluna de afinidade e centrifugado novamente por 12.000 xg durante 1 minuto. O efluente foi descartado e tampão de lavagem adicionado à coluna. Seguiu-se outra centrifugação a 12.000 xg por 1 minuto e o efluente foi novamente descartado. A coluna foi colocada em um novo tubo e o DNA foi eluído em tampão TE (10 mM Tris-HCl (pH 8,0), 0,1 mM EDTA) aquecido (65-70°C), com incubação da coluna por 1 minuto a temperatura ambiente e centrifugação por 2 minutos a 12.000 xg.

Após a extração do DNA plasmidial dos clones selecionados, foi feita digestão com enzima de restrição (*EcoRI*). A comprovação da formação do DNA recombinante foi verificada através da liberação do fragmento esperado. A análise dos produtos de digestão foi feita em gel de agarose 1%.

### **3.8. Seqüenciamento dos Clones.**

O seqüenciamento de nucleotídeos foi realizado em ambas as direções (senso e anti-senso) utilizando-se os oligonucleotídeos universais M13 e o "DNA sequencing Kit Big-Dye Terminator Cycle Sequencing Ready Reaction" (Perkin Elmer) em um seqüenciador ABI Prism 377 (Applied Biosystems, Inc.), através do método de terminação de cadeia por dideoxinucleotídeos ("dideoxynucleotide chain termination method") (SANGER *et al.*, 1977).

### **3.9. Análise do Seqüenciamento.**

A análise comparativa das seqüências foi feita através do algoritmo BLAST (<http://www.ncbi.nlm.nih.gov>) e o alinhamento das seqüências consenso pelo programa Multalin (<http://prodes.toulouse.inra.fr/multalin/multalin.html>). As seqüências de aminoácidos deduzidas foram determinadas pelo programa Translate (<http://www.expasy.org/tolls/dna.html>). As relações filogenéticas foram determinadas pelo método de máxima parcimônia, com auxílio do programa PAUP v 4.0b10 (SWOFFORD, 1997) e com teste de confiança de topologia (*bootstrap*) de pelo menos 1000 réplicas.

### **3.10. Subclonagem do Fragmento Amplificado da Polimerase do SBMV-SP (Expressão).**

A subclonagem dos fragmentos amplificados por PCR foi feita utilizando-se o vetor pGEM<sup>®</sup>-T, de acordo com instruções do fabricante (Promega). Após a extração do DNA plasmidial dos clones selecionados, foi feita digestão com enzimas de restrição (*Hind* III e *Bam* HI). A comprovação da formação do DNA recombinante foi verificada através da liberação do fragmento esperado. A análise dos produtos de digestão foi feita em gel de agarose 1%.

### **3.11. Clonagem dos Fragmentos da Polimerase (Expressão).**

O vetor recombinante (pGEM + inserto) foi purificado e submetido à dupla digestão com enzimas de restrição *Bam* HI e *Hind* III. O fragmento liberado após a digestão foi extraído e purificado do gel utilizando MinElute Gel Extraction Kit (Qiagen) conforme instruções do fabricante. Em seguida este fragmento foi ligado em vetores de expressão pela enzima T4 ligase (Promega) conforme instrução do fabricante e em seguida transformação em *E. coli*. O primeiro sistema utilizado foi o pMALc-2x (New England Biolabs) em que foi utilizada a linhagem de *E. coli* DH5 $\alpha$ . O mesmo procedimento de foi utilizado para o sistema pET 28a (Novagen) em que foi utilizada a linhagem bacteriana BL21.

### **3.12 Seqüenciamento do fragmento clonado em vetor de expressão.**

Os vetores recombinantes foram seqüenciados, primeiramente o vetor recombinante no sistema pMAL-c2x (New England Biolabs) e posteriormente o ou outro vetor recombinante no sistema pET 28 a (Novagen). Ambos foram seqüenciados pelo método Sanger utilizando o equipamento ABI Prism 377 (Applied Biosystems, Inc). Foram utilizados os oligonucleotídeos iniciadores 5'-senso (malE primer) para o vetor PMALc-2x indicado pelo fabricante (New England Biolabs) e o M13 para o vetor pET 28a. As seqüências foram obtidas e alinhadas através do algoritmo BLAST (<http://www.ncbi.nlm.nih.gov>).

### **3.13. Expressão em *E. coli*.**

O pré-inóculo da colônia transformante selecionada consistiu de 7 mL de meio LB líquido juntamente com 100µg/mL de ampicilina. Após incubação a 37°C sob agitação de 200 rpm por aproximadamente 16 horas, o pré-inóculo foi adicionado a um novo meio de cultura (250 mL de meio LB, 100µg/mL de ampicilina e 2mg/mL de Glicose) que foi incubado sob as mesmas condições. Quando a cultura de células atingiu a densidade óptica de 0,6 em 550nm, a indução de expressão foi feita com 1 mM de isopropil-β-D-tiogalactopiranosida (IPTG), a 37°C por 4 horas sob agitação de 200 rpm. As células foram separadas do meio por centrifugação a 4.000 x g durante 20 minutos a 4°C e posteriormente congeladas a -20°C.

### **3.14. Purificação da proteína de fusão em coluna de afinidade na forma nativa (sistema pMAL c2-x).**

A lise celular foi realizada através da adição de 5mL de Tampão de Lise (Fosfato de sódio 20mM, NaCl 200mM, EDTA 1mM, pH 7,2) seguida de sonicação (8 vezes de 15 segundos). Após a sonicação, o lisado foi centrifugado por 20 minutos a 4.000 x g a 4°C.

O sobrenadante do lisado foi aplicado em coluna de amilose (2,0 mL de resina), previamente equilibrada com 8 volumes de Tampão de Coluna (Fosfato de sódio 20mM e NaCl 200 mM, pH 7,2) e permaneceu sob agitação leve por aproximadamente 4 horas a 4°C. Em seguida a coluna foi lavada com 12 volumes de Tampão de Coluna e a proteína eluída com 5 volumes de Tampão de Coluna acrescido de 100mM de maltose.

As amostras resultantes da eluição foram verificadas em SDS-PAGE 10% e as frações contendo a proteína referente à porção C-terminal da polimerase do SBMV-SP foram reunidas e concentradas utilizando-se concentradores Centricon-30 (Millipore) por centrifugação.

### **3.15. Purificação da proteína expressa em coluna de afinidade na forma desnaturada (sistema pET 28 a).**

A lise celular foi realizada através da adição de 5mL de Tampão de Lise (Fosfato de sódio 20mM pH 7,8, NaCl 500mM, 6M de guanidina e 20mM  $\beta$ -mercaptoetanol) seguida de sonicação (8 vezes de 15 segundos). Após a sonicação, o lisado foi centrifugado por 20 minutos a 4.000 x g a 4°C.

O sobrenadante do lisado foi aplicado em coluna de níquel (2,0 mL de resina), previamente equilibrada com 8 volumes de Tampão de Ligação (Fosfato de sódio 20mM pH 7,8, NaCl 500mM, 6M de guanidina 20mM,  $\beta$ -mercaptoetanol e 8M de uréia) e permaneceu sob agitação leve por aproximadamente 4 horas a 4°C. Em seguida a coluna foi lavada com 10 volumes de Tampão de Lavagem 1 (Fosfato de sódio 20mM pH 6,0, NaCl 500mM, 6M de guanidina 20mM,  $\beta$ -mercaptoetanol e 8M de uréia), em seguida uma nova lavagem foi feita com o tampão de lavagem 2 (Fosfato de sódio 20mM pH 5,3, NaCl 500mM, 6M de guanidina 20mM,  $\beta$ -mercaptoetanol e 8M de uréia) e a proteína eluída com 5 volumes de Tampão de Eluição (Fosfato de sódio 20mM pH 4,0, NaCl 500mM, 6M de guanidina 20mM,  $\beta$ -mercaptoetanol e 4M de uréia).

As amostras resultantes da eluição foram verificadas em SDS-PAGE 10% e as frações contendo a proteína referente à porção C-terminal da polimerase do SBMV-SP foram reunidas e concentradas utilizando-se concentradores Centricon-30 (Millipore) por centrifugação.

### **3.16. Eletroforese**

As amostras de proteínas obtidas das purificações foram analisadas em SDS-PAGE conforme Laemmli (1970). O gel de empacotamento foi feito a uma concentração de 5% e os de resolução de 10 ou 12%. Para a aplicação das amostras no SDS-PAGE as proteínas foram desnaturadas por fervura a 100°C por 3 minutos em tampão de amostra (Tris. HCl 0,125 M, pH 6,8, contendo 4% de dodecil sulfato de sódio - SDS, glicerol 20%,  $\beta$ -Mercaptoetanol 10% e azul de bromofenol 0,001M), utilizando-se o aparelho Mini Protean II (BioRad). A fixação das proteínas foi realizada utilizando uma mistura de metanol, água e

ácido acético (50:50:100) por 60 minutos. Para corar o gel foi utilizada a coloração com Comassie 0,25% na mistura fixadora e em seguida o gel foi descorado com uma mistura de metanol e água (1:1) por 6 horas.

### **3.17. Produção de Anti-soro**

Foram inoculados coelhos da raça Nova Zelândia com o produto da purificação (*Maltose binding protein* - MBP + C-terminal da polimerase). Cinco injeções intramusculares foram aplicadas na região da coxa (100µg de proteína cada aplicação), com intervalo de 7 dias entre elas. A primeira injeção continha uma mistura da proteína purificada mais adjuvante completo de Freund (Sigma) na proporção de 1:1. As aplicações seguintes foram feitas uma mistura da proteína purificada mais adjuvante incompleto de Freund (Sigma) também na proporção de 1:1. Após 7 dias após a aplicação da última injeção foi coletado o sangue do coelho e incubado a 4°C durante 12 horas e em seguida realizou-se uma centrifugação a 3000 rpm por 20 minutos para separar o plasma sanguíneo das hemácias. Feita a separação coletou-se a fração do plasma sanguíneo onde se encontrava os anticorpos policlonais.

### **3.18. Imunolocalização da porção C-Terminal da Polimerase**

A membrana de celulose foi incubada por 2 horas à temperatura ambiente em tampão PBS (fosfato de sódio 0,01 M pH 7,2, NaCl 0,15 M) contendo soroalbumina bovina (BSA) 5% e Tween-20 0,3% (PBS-T). Em seguida a membrana foi incubada por 1 hora com o anti-soro primário (anti - C-terminal da polimerase) diluído 1: 2500 com tampão PBS-T contendo BSA 0,5% (DIETZGEN & FRANCKI, 1987). Foram feitas três lavagens com PBS-T por 15 minutos cada. Em seguida a membrana foi incubada à temperatura ambiente com anti-soro anti-IgG de coelho (Promega) conjugado com fosfatase alcalina diluído 1:7500 em PBS-T contendo 5% de BSA. Após a incubação com o anti-soro secundário foram feitas 3 lavagens com PBS por 5 minutos cada. Após as lavagens, a membrana foi incubada por 20 minutos a temperatura ambiente com

mistura de 5-bromo-4-cloro.3-indolil fosfato/nitro-blue tetrazolio (BCIP/NBT - substrato para fosfatase alcalina).

### **3.19. Western blot**

Foi realizada uma eletroforese das proteínas conforme o item 3.15 para fazer a transferência em membrana de nitrocelulose em um eletrotransferidor (Mini Transblot Eletroforetic Cell – BioRad) operando a 100V por 1 hora em tampão de transferência (Tris-Glicina- Tris 0,025 M e Glicina 0,192M) contendo metanol 20%. A imunolocalização foi realizada conforme o item 3.18.

## RESULTADOS

### 4.1. Sintomatologia.

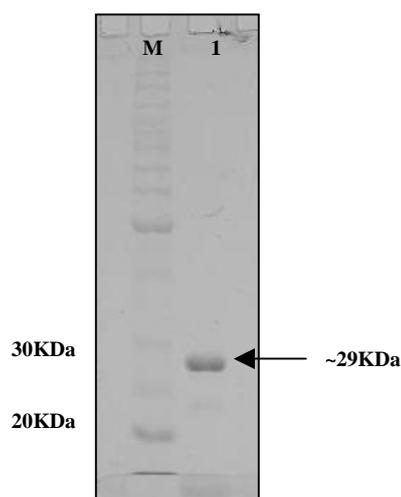
O SBMV-SP, quando inoculado em *P. vulgaris* L. “Jalo”, induziu infecção sistêmica com mosaico, mosqueado clorótico nas folhas inoculadas e distorção foliar nas folhas mais jovens (Figura 1).



**Figura 1:** Comparação de folhas de feijoeiro sadia (esquerda) e infectada (direita);  
**Abaixo:** Visualização da planta infectada.

#### 4.2. Obtenção do Vírus Purificado.

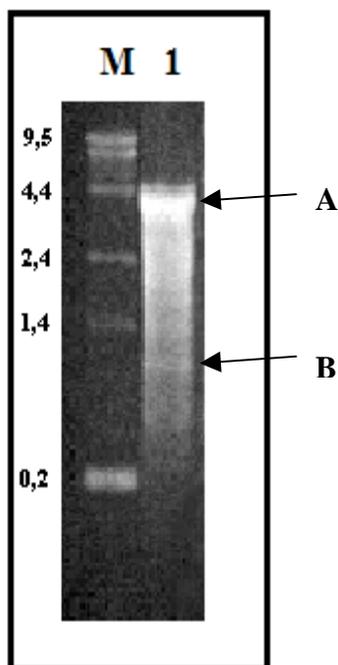
Levando-se em consideração o coeficiente de extinção de 6,0mg/ml/cm estabelecido para vários sobemovírus, determinou-se que a concentração do SBMV-SP purificado foi de 12,8 mg/100g de folha. Através da análise de SDS-PAGE 10% foi possível visualizar uma banda de 29kDa correspondente à proteína capsidial (figura 2).



**Figura 2: SDS-PAGE 12%** -(M) Marcador de massa molecular (kDa) BenchMark (Invitrogen); 1: Proteína capsidial de ~ 29 KDa (seta).

### 4.3. Extração do RNA Viral.

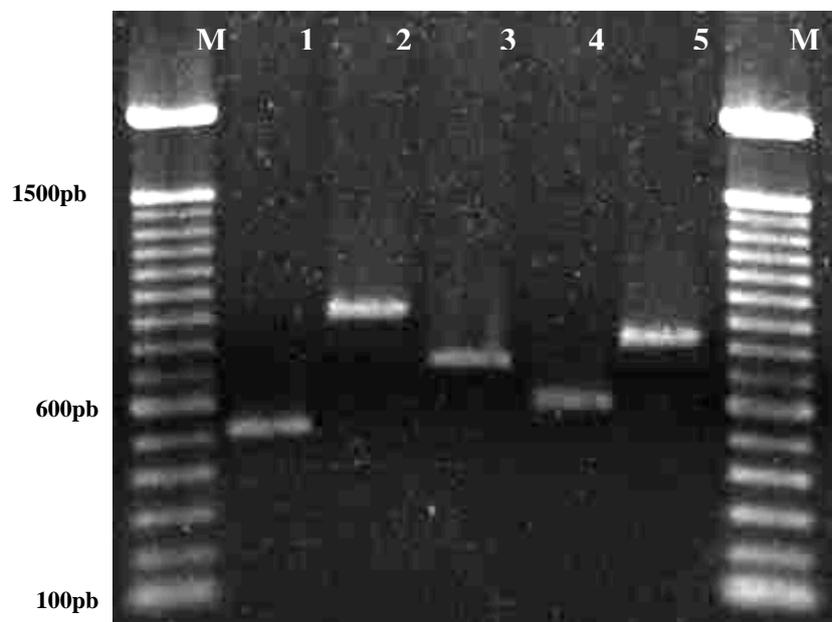
A análise do RNA viral em gel de agarose 1% evidenciou uma banda de aproximadamente 4,2Kb referente ao RNA genômico e uma outra banda de aproximadamente 1,0 Kb referente ao RNA subgenômico (Figura 3).



**Figura 3:** Análise do RNA viral extraído de partículas purificadas de SBMV-SP em gel de agarose 1%. (**M**): Marcador de tamanho molecular 0,25-9.5 Kb RNA Ladder (Invitrogen); (**1**): (A) RNA genômico com ~ 4,2kb (B) RNA subgenômico com ~ 1,0kb.

#### 4.4. Amplificação por PCR dos Fragmentos da ORF 2 (Seqüenciamento).

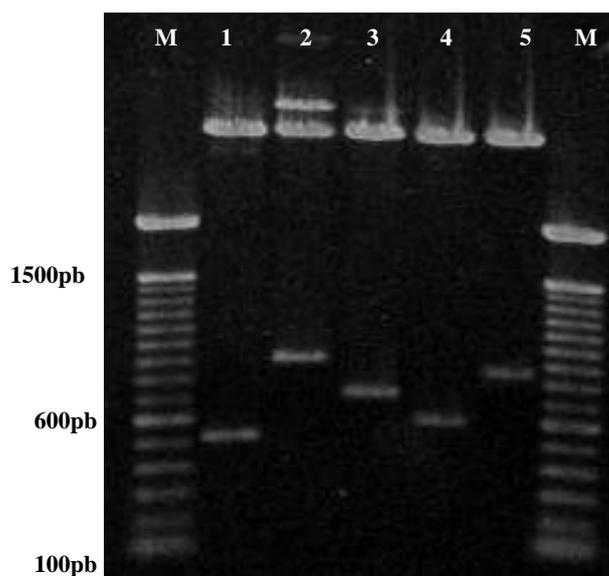
Os produtos da reação de PCR foram analisados em gel nativo de agarose 1% evidenciando bandas de tamanhos esperados conforme a figura 4.



**Figura 4.** Gel nativo de agarose 1% (M): Marcador de tamanho molecular 100 bp DNA Ladder (Invitrogen); **1-5:** Produtos amplificados. (1): 498pb; (2): 870 pb; (3): 693; (4): 598pb; (5): 834pb.

#### 4.5. Digestão com Enzima de Restrição com *Eco* RI.

A análise em gel de agarose 1% dos vetores recombinantes, após digestão com enzima de restrição *Eco* RI, evidenciou a liberação dos fragmentos esperados como mostra a figura 5.



**Figura 5:** Gel nativo de agarose 1%; (M): marcador de tamanho molecular de 100 bp DNA Ladder (Invitrogen); 1-5: Vetores cortados e os fragmentos de interesse (1): 498pb; (2): 870 pb; (3): 693; (4): 598pb; (5): 834pb.

#### 4.6. Sequenciamento.

A figura 6 mostra a seqüência completa de nucleotídeos das ORF 2 e 3 depositada no GenBank (acesso: DQ875594) com suas respectivas seqüências de aminoácidos deduzidas.

```

1   atgtatcatccaggcgcgtcaccatcattcctgataacgtagcaaatggtatctgcgcg 60
    M Y H P G R S P S F L I T L A N V I C A

61  gcaatcttgtacgacatccgatgggggggtaccaaccaggatcactagttccaatagtg 120
    A I L Y D I R M G G Y Q P G S L V P I V
                                     SP

121 gcttggatgaccccgcttcgttacgctgctcgttgagcgcgtcattcgtgacatacctt 180
    A W M T P F V T L L W L S A S F V T Y L
                                     SP

181 tataggtatgctcgaactcgaactcgttccagaggagaaagtggccagagtttattatagc 240
    Y R Y A R T R L L P E E K V A R V Y Y T
                                     SP

241 gcgcaatctgcgccttactttgaccggccctgggtgtcatgatgcaatttgcccctagc 300
    A Q S A P Y F D P A L G V M M Q F A P S
                                     SP

301 catggcggcgcagcatagaagtgcaagttaacccttcgtggattagtctcttagtgggc 360
    H G G A S I E V Q V N P S W I S L L G G

361 tctctcaagataaacggagatgacgcctctaataagtgctgctgtgttggaagcttttat 420
    S L K I N G D D A S N E S A V L G S F Y

421 tcttctgtgaaacctgggtgatgaaccagctagtttggtagctattaagagtggtcctcag 480
    S S V K P G D E P A S L V A I K S G P Q

481 accatcggtttttggtttagaactaagatcgacgggtgatgactgcctcttcacagacaat 540
    T I G F G C R T K I D G D D C L F T D N

541 cacgtttggaataattccatgcgccctaccgcttttagcgaagctggcaagcaggttgcg 600
    H V W N N S M R P T A L A K A G K Q V A

601 attgaagattgggacacccctctctcttgtagaccataagatgcttgatttcgtagtggtg 660
    I E D W D T P L S C D H K M L D F V V V

661 cgtgtgccaaaaacggtgtggtccaaactaggagtgaagcgactcaactggtttgtoca 720
    R V P K H V W S K L G V K A T Q L V C P

721 tctgataaggacgctgtaacctgttatggtagcttagttctgatagcttgttgcgggg 780
    S D K D A V T C Y G G S S S D S L L S G

781 acggcgtttttagtaaggttgatttctcttggaaagtaaccactcatgccccacggca 840
    T G V C S K V D F S W K L T H S C P T A
                                     SP

841 gctggctggagcggaaactccaatttactctagcagaggtgtggttaggaatgcacgttggg 900
    A G W S G T P I Y S S R G V V G M H V G
                                     SP

901 tttgaagatatcgaaaactcaaccgtggcgtgaacgctttctacgtgtctaactacttg 960
    F E D I G K L N R G V N A F Y V S N Y L

961 ttgaggtctcaagagactctacctcctgatctgtccgtcatcgaatacctttcgaggac 1020
    L R S Q E T L P P D L S V I E I P F E D
                                     Vpg

1021 gtagaaaccggagttatgagttcattgaggtcgagattaaaggaagaggttaaggctaaa 1080
    V E T R S Y E F I E V E I K G R G K A K
                                     Vpg

```

1081 cttggtaagcgtgagttcgccttgattccggaatcaggaaaatactgggcccgatgatgat 1140  
 L G K R E F A W I P E S G K Y W A D D D  
 Vpg  
 1141 gacgactctttgccccaccgcccgaaggtggttagacggcaagatggtgtggacttcagct 1200  
 D D S L P P P P K V V D G K M V W T S A  
 Vpg  
 1201 caggaaaccgtcgagagcctttaaactaccagagggcgccaggggtcaagggccttgccg 1260  
 Q E | T V A E P L N Y Q R A A G S R P L P  
 RpRd  
 1261 ccctttctcaacttgcaggctacaacttcaaagaaggagaagcagcctctacaagaggaa 1320  
 P F L N L Q A T T S K K E K Q P L Q E E  
 ORF 3  
 1321 tgccccttagatttgttgggcagtcggccttgcaagtttagagagctgtgtagaaaagata 1380  
 C P L D L L G S R L A S L E S C V E K I  
 M P L R F V G Q S A C K F R E L C R K D  
 1381 ctccagatgaagtccttagagctactagagcttctcccgaattgtcagacttctcctggc 1440  
 L Q M K S L E L L E S S R N C Q T S P G  
 T P D E V L R A T R V F P E L S D F S W  
 1441 ctgagcggaggtccaaagcagagcttcaactcctgtactccaagcaggaaagttaatc 1500  
 L S E A S P K Q S F T P C Y S K Q E S L I  
 P E R G A S K A E L H S L L L Q A G K F N  
 1501 ccaccgcaatccaaggaatctgaaggagcttgtcaaacctccttgagcgcctaccocg 1560  
 P P Q S Q G I L K E L V K T S L S A T P  
 P T A I P R N L E G A C Q N L L E R Y P  
 1561 cctccaaatcctgttactgcctcgtggagaagcctggccttcgacgcagctctacgaag 1620  
 P P N P V T A F V E K P G P S T Q S T K  
 A S K S C Y C L R G E A W S F D A V Y E  
 1621 aagtctgcaagaaggcgcaatcggcggaatcaacgagaaagccagtcagggaggtcccc 1680  
 K S A R R R R N R R K S T R K P V Q G V P  
 E V C K K A Q S A E I N E K A S P G S P  
 1681 ctctcccgctctgcctccaccaacaagacctccttaagaggcacttagaattagttgct 1740  
 L S R L A S T N K D L L K R H L E L V A  
 P L P S R L H Q Q R P P •  
 1741 ttgtgtgttaccgagagattgttcttactcagtgaaagctgaggacctgcacaataaatcc 1800  
 L C V T E R L F L L S E A E D L H N K S  
 1801 cctgtggacctagttcagatggggttgtgacaccagttcggctgtttgtcaagcaggag 1860  
 P V D L V Q M G L C D P V R L F V K Q E  
 RpRd  
 1861 ccccatgcttcccgaaggtgaaggagggttagatttcgcttgatttcatccgtttcgctg 1920  
 P H A S R K V K E G R F R L I S S V S L  
 RpRd  
 1921 gtggatcagctttagagcgaatgcttttcgggccccaaaaccagcttgagatcgctgag 1980  
 V D Q L V E R M L F G P Q N Q L E I A E  
 RdRp  
 1981 tgggaacatattccctcaaacctggatggcctttcgctgcaacgacaagccaaaagc 2040  
 W E H I P S K P G M G L S L Q R Q A K S  
 RpRd  
 2041 ttgttcgacgatttgagagtcacaacattctcgttgcctcgcggctgaagccgacatc 2100  
 L F D D L R V K H S R C P A A E A D I S  
 RpRd  
 2101 ggttttgactggtctgttcaagactgggagttgtgggctgatgtagagatgagaatagtt 2160  
 G F D W S V Q D W E L W A D V E M R I V  
 RpRd  
 2161 ttaggaggttttgacagaagttgtccatagccgctagaaacaggttttcgtgtttcatg 2220  
 L G G F G Q K L S I A A R N R F S C F M  
 2221 aactcagctctccagctctcggatggcacactcatagaacagcaactgcctggtattatg 2280  
 N S V F Q L S D G T L I E Q Q L P G I M

**2281** aagtctggttcttactgcacttctcaacaaactccagaatacgttgcccttatggctgaa **2340**  
 K S G S Y C T S S T N S R I R C L M A E  
*RpRd*

**2341** cttattggttccccatgggtgatcgccatgggtgatgattctggtgaggggtgggtgat **2400**  
 L I G S P W C I A M G D D S V E G W V D  
*RpRd*

**2401** ggtgcgaaagacaagtacatgagattaggccacacgtgcaaggattataaacctgtgca **2460**  
 G A K D K Y M R L G H T C K D Y K P C A

**2461** acatccatttccggctgcttatacagaggtagagttttgctctcacgttataaggaagat **2520**  
 T S I S G R L Y E V E F C S H V I R E D

**2521** cgatggttggtggcgtcgtggcccaaaactctgtataaatacttgtctgagggcaagtgg **2580**  
 R C W L A S W P K T L Y K Y L S E G K W

**2581** ttctttgaggatcttgagcgagagcttgggtcttctccccactggcccagaatcagacac **2640**  
 F F E D L E R E L G S S P H W P R I R H

**2641** tacgtagtcgggaatactccatcgcccgcacaaaactagattagaaaattcaagtccgagc **2700**  
 Y V V G N T P S P D K T R L E N S S P S

**2701** tatggcgaagaggtgacaaaacaacagtttagccaaggctatagcgaacactctggaagc **2760**  
 Y G E E A D K T T V S Q G Y S E H S G S

**2761** cccggccactcaatcgaggaggccaggaaccggagacggcgccgttctgctgcaaggca **2820**  
 P G H S I E E A Q E P E T A P F C C K A

**2821** gcctcagtctatccaggctggggcatccatggcccctattgctcagggggctatggttcg **2880**  
 A S V Y P G W G I H G P Y C S G G Y G S

**2881** cttacgtga **2889**  
 L T •

**Figura 6:** Seqüência de nucleotídeos e a aminoácidos deduzidos das ORFs 2 e 3 (nº de acesso DQ875594); os aminoácidos em negrito indicam o início das regiões referentes aos produtos *SP* (Serino Protease), *Vpg* (proteína ligada ao genoma viral) e *RpRd* (RNA polimerase RNA-dependente); | - indica o sítio de clivagem para a serino protease nos aminoácidos E e T (E: glutamato e T: Treonina) ; marcados em cinza estão indicados os domínios e motivos da poliproteína; os nucleotídeos (atg) sublinhados e em negrito referem-se ao códon de iniciação das ORF 2 e ORF 3; em azul a seqüência deduzida de aminoácidos da ORF3; • indica o códon de terminação; sublinhado em vermelho a seqüência heptanucleotídeo referente ao sinal de *frameshift* -1 para a ORF 3.

#### 4.7. Análise do Seqüenciamento.

As tabelas 1 e 2 mostram dados de identidade para as seqüências de aminoácidos deduzidos e nucleotídeos da ORF 2 e 3 do SBMV-SP com outros sobemovírus, indicando os valores de identidade. A figura 7 mostra um alinhamento entre as seqüências de aminoácidos de algumas espécies de sobemovírus e isolados de SBMV, referentes à região do sítio de ligação da proteína capsidial (*Coat Protein Biding Site* - CPBS), localizado internamente à região da serino protease da ORF 2. A figura 8 mostra a organização genômica do SBMV-SP, evidenciando a região da Vpg com sua seqüência de aminoácidos e algumas regiões importantes, como os sítios de clivagem da poliproteína (E/T) e algumas regiões conservadas na porção N-terminal da Vpg (20 primeiros aminoácidos) e os aminoácidos WAD. O alinhamento mostrado na figura 9 é da região N-terminal da Vpg (20 primeiros aminoácidos), comparando-se à mesma região de outros sobemovírus.

#### ORF-2

	SBMV	SBMV-Ark	SBMV-SP	SeMV	SCPMV	RYMV	LTSV	CoMV
SBMV		76,6	71,9	61,3	49,3	45,6	38,6	40,5
SBMV-Ark	60,2		91,4	75,7	58,2	43,5	44,4	45,0
SBMV-SP	57,9	95,0		76,0	58,8	43,5	44,7	46,0
SeMV	49,9	77,0	77,7		58,9	43,3	43,7	45,6
SCPMV	36,5	52,7	53,1	55,5		42,9	43,3	50,5
RYMV	24,4	31,3	31,4	32,0	30,8		40,5	44,3
LTSV	27,7	37,4	36,9	36,9	35,4	30,6		40,3
CoMV	24,9	31,0	30,6	31,8	30,8	39,3	31,2	

**Tabela 1:** Comparação (porcentagem de identidade) entre a seqüência de nucleotídeos (acima da diagonal) e aminoácidos deduzidos (abaixo da diagonal) da ORF 2 de isolados do *Southern bean mosaic virus* e espécies do gênero *Sobemovirus*. O número de acesso ao GenBank para os vírus citados são: SeMV (NC002568), SBMV (L34672), SBMV-ARK (NC004060), SBMV-SP (DQ875594) RYMV (NC001575), SCPMV (NC001625), CoMV (NC002618), LTSV (NC001696).

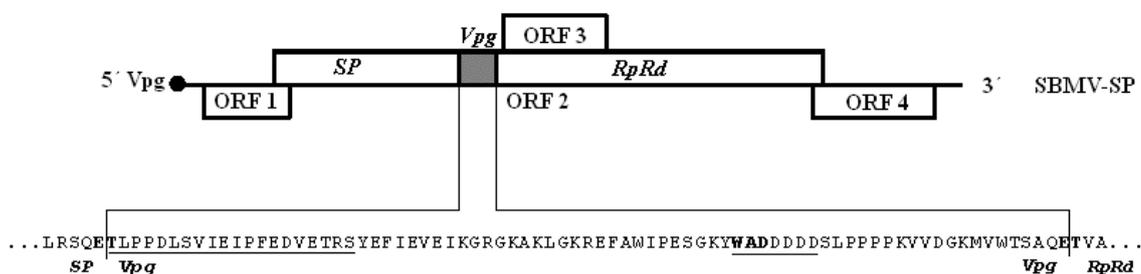
## ORF-3

	SBMV-Ark	SBMV-SP	SeMV	SCPMV	RYMV	LTSV
SBMV-Ark		97,0	76,1	54,4	45,1	38,5
SBMV-SP	97,7		76,1	54,8	45,4	39,1
SeMV	73,6	73,6		60,5	42,2	41,9
SCPMV	48,8	48,8	52,5		41,0	38,6
RYMV	23,5	22,0	18,8	19,1		41,1
LTSV	36,5	35,9	37,1	37,1	17,6	

**Tabela 2:** Comparação (porcentagem de identidade) entre a seqüência de nucleotídeos (acima da diagonal) e aminoácidos deduzidos (abaixo da diagonal) da ORF 3 de isolados do *Southern bean mosaic virus* e espécies do gênero *Sobemovirus*. O número de acesso ao GenBank para os vírus citados são: SeMV (NC002568), SBMV-ARK (NC004060), SBMV-SP (DQ875594), RYMV (NC001575), SCPMV (NC001625), LTSV (NC001696).

<b>SBMV-SP</b>	281	<b>A</b>	<b>G</b>	<b>W</b>	<b>S</b>	<b>G</b>	<b>T</b>	<b>P</b>	<b>I</b>	<b>Y</b>	<b>S</b>	<b>S</b>	291
<b>SBMV-Ark</b>	281	.	.	.	.	.	.	.	.	.	.	.	291
<b>SeMV</b>	281	.	.	.	.	.	.	.	<b>L</b>	.	.	.	291
<b>SCPMV</b>	281	<b>K</b>	.	.	.	.	.	.	<b>L</b>	.	<b>T</b>	<b>R</b>	291
<b>RYMV</b>	282	<b>S</b>	.	.	.	.	<b>S</b>	.	<b>L</b>	.	<b>H</b>	<b>K</b>	292
<b>CfMV</b>	273	<b>G</b>	.	.	.	.	<b>S</b>	.	<b>L</b>	.	<b>H</b>	<b>K</b>	283
<b>LTSV</b>	217	<b>K</b>	.	.	.	.	.	.	<b>L</b>	.	<b>V</b>	<b>G</b>	227

**Figura 7:** Comparação entre as seqüências de aminoácidos da região do sítio de ligação da proteína capsidial (*Coat Protein Biding Site - CPBS*) de isolados do SBMV e espécies do gênero *Sobemovirus*. Os números à esquerda da seqüência referem-se à posição do primeiro aminoácido e a direita a do último aminoácido deste sítio em relação à posição na seqüência de aminoácidos da poliproteína da ORF 2. Em destaque está o aminoácido serina (S) presente no sítio catalítico da serino protease (SP).



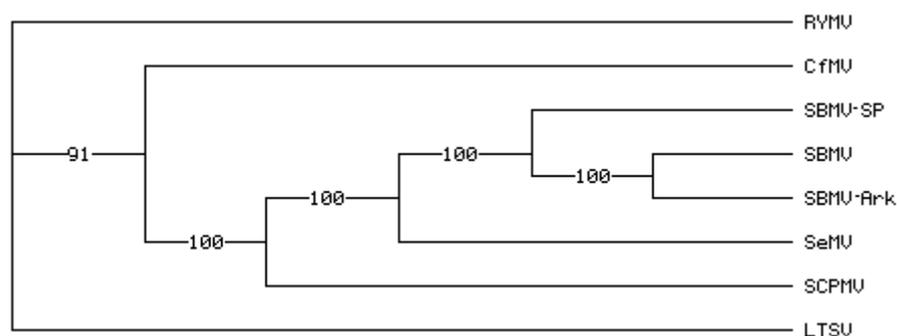
**Figura 8:** Esquema da organização do genoma do SBMV-SP, evidenciando a região da Vpg (em cinza) com sua respectiva seqüência completa de aminoácidos. E|T: indica o sítio de clivagem para a serino protease (SP) na poliproteína, localizados entre as proteínas SP e Vpg e entre Vpg e RpRd. As seqüências sublinhadas são regiões conservadas da Vpg para os sobemovírus.

<b>SBMV-SP</b>	<b>326</b>	<b>TLPPDLSVIEIIPFEDVETRS</b>	<b>345</b>
<b>SBMV-Ark</b>	<b>326</b>	<b>...E.....</b>	<b>345</b>
<b>SeMV</b>	<b>326</b>	<b>...E..I....D.....</b>	<b>345</b>
<b>SCPMV</b>	<b>326</b>	<b>...E.GYR..SL...GL..</b>	<b>345</b>

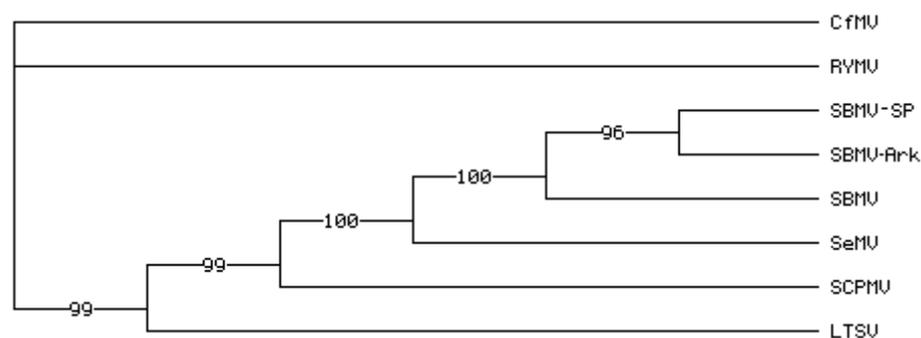
**Figura 9:** Comparação entre as seqüências de aminoácidos da região N-terminal das Vpgs de isolados do SBMV e espécies do gênero *Sobemovirus*. Os números à esquerda da seqüência referem-se à posição do primeiro aminoácido e a direita a do último aminoácido desta seqüência em relação à posição na seqüência de aminoácidos da poliproteína da ORF 2.

#### 4.8. Análise Filogenética.

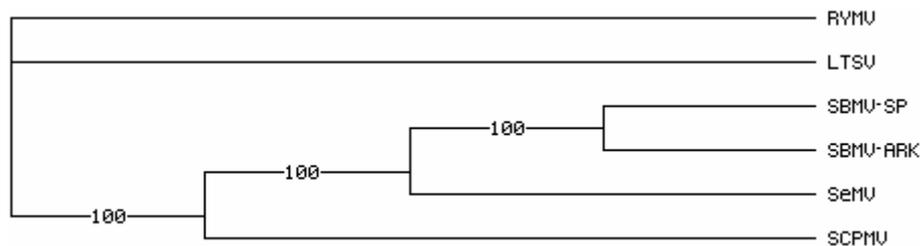
A análise filogenética pelo método de máxima parcimônia para as seqüências de nucleotídeos e aminoácidos deduzidos da ORFs 2 (figuras 10 e 11), as seqüências de nucleotídeos e aminoácidos deduzidos da ORF 3 (figuras 12 e 13). A figura 14 refere-se a árvore para a seqüência de nucleotídeos dos genomas de espécies e isolados de sobemovírus .



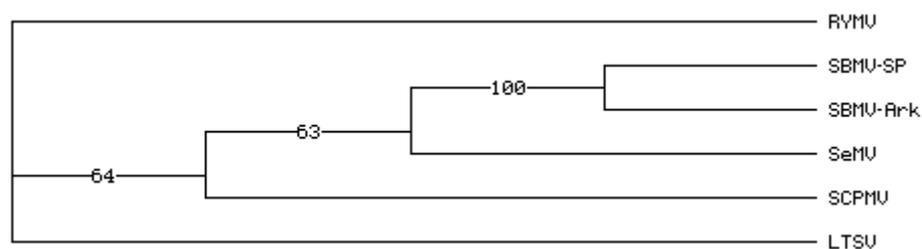
**Figura 10.** Análise filogenética com base na seqüência completa de nucleotídeos da **ORF-2** do SBMV-SP e outros sobemovírus, pelo método de máxima parcimônia com 1000 réplicas. As 8 espécies estão identificadas à direita da árvore e os valores de “bootstrap” (em porcentagem) estão localizados à esquerda.



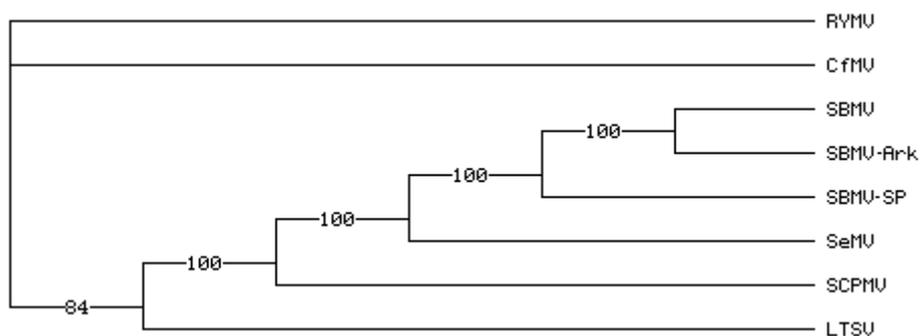
**Figura 11.** Análise filogenética com base na seqüência completa de aminoácidos deduzidos da **ORF-2** do SBMV-SP e outros sobemovírus, pelo método de máxima parcimônia com 1000 réplicas. As 8 espécies estão identificadas à direita da árvore e os valores de “bootstrap” (em porcentagem) estão localizados à esquerda.



**Figura 12.** Análise filogenética com base na seqüência completa de nucleotídeos da **ORF-3** do SBMV-SP e outros sobemovírus, pelo método de máxima parcimônia com 1000 réplicas. As 6 espécies estão identificadas à direita da árvore e os valores de “bootstrap” (em porcentagem) estão localizados à esquerda.



**Figura 13.** Análise filogenética com base na seqüência completa de aminoácidos deduzidos da **ORF-3** do SBMV-SP e outros sobemovírus, pelo método de máxima parcimônia com 1000 réplicas. As 6 espécies estão identificadas à direita da árvore e os valores de “bootstrap” (em porcentagem) estão localizados à esquerda.



**Figura 14.** Análise filogenética com base na seqüência completa de nucleotídeos dos genomas dos *Sobemovirus*, pelo método de máxima parcimônia com 1000 réplicas. As 8 espécies estão identificadas à direita da árvore e os valores de “bootstrap” (em porcentagem) estão localizados à esquerda.

#### 4.9. Complementação do Sequenciamento do Genoma do SBMV-SP.

As seqüências das ORF 1, 2, 3 e 4 do SBMV-SP foram reagrupadas para montagem do genoma completo do vírus (figura 15). As seqüências das ORFs 1 e 4 foram obtidas do GenBank (ORF1 n° de acesso: AY340587 e ORF4 n° de acesso: AY340586). De acordo com Espinha *et al.* (2004) há sobreposição de 34 nucleotídeos entre as ORFs 1 e 2 e entre as ORFs 2 e 4 também há sobreposição de 188 nucleotídeos (ESPINHA *et al.*, 2005).

```

5'  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
      5      15      25      35      45      55
CACAAAATAT AAGAAGGAAA GCTGGATTTC CTACCTTTGT GTTTCCATTG TCGAAGCATT

....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
      65      75      85      95      105     115
GGTCAATACT TATCAATTGG TGCATTGTTT GCATGAGCTA CCGATTCTTA GTAGTCAAAG
                        *ORF1

....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
      125     135     145     155     165     175
CCGTTGGTTT TCTTGGTTTC CATTGAGACG CTACTCGCAT TCTGTCAGAG ACTGAGATCG

....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
      185     195     205     215     225     235
TAGACGTTCC TTCGTCCATT GATTTCGTCG GTGAAACCGA GTTACGCCTA GAAAACGCTT

....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
      245     255     265     275     285     295
GGCCCCAAGG TGGTGAGAGA TACACTATCC TACCTAGGTT CAACGTTTCC ATTGACTTCA

....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
      305     315     325     335     345     355
CGTACCATCC AGTGCCTGTC GAGATCATCT GTAGGGTTTG TGCTACTTCC CTTACTGTTG

....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
      365     375     385     395     405     415
TCTTTAGCAA GTGGAACTTC CATTGCGAAA GGAAGGGCCA TTTTGTGCCA GTAGACCAGA

....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
      425     435     445     455     465     475
ACGGGAATCT GTTTAGGGTT GGAACGCTCC GGGAGACGGG AGAGAAATAC TTCTACTTCT

....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
      485     495     505     515     525     535
GTGAGAAATC TATCTGCAGA CAATGATATCA TCCAGGCCGC TCACCATCAT TCCTGATAAC
                        *ORF2                                •ORF1

....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
      545     555     565     575     585     595
GTTAGCAAAT GTTATCTGCG CGGCAATCTT GTACGACATC CGTATGGGGG GGTACCAACC

....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
      605     615     625     635     645     655
AGGATCACTA GTTCCAATAG TGGCCTGGAT GACCCCGTTC GTTACGCTGC TCTGTTGAG

```

```

.....|.....| .....|.....| .....|.....| .....|.....|
  665      675      685      695      705      715
CGCGTCATTC GTGACATACC TTTATAGGTA TGCTCGAACT CGACTGCTTC CAGAGGAGAA

.....|.....| .....|.....| .....|.....| .....|.....|
  725      735      745      755      765      775
AGTGGCCAGA GTTTATTATA CGGCGCAATC TGCGCCTTAC TTTGACCCGG CCCTGGGTGT

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
  785      795      805      815      825      835
CATGATGCAA TTTGCCCTA GCCATGGCGG CGCCAGCATA GAAGTGCAAG TTAACCCCTC

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
  845      855      865      875      885      895
GTGGATTAGT CTCTTAGGTG GCTCTCTCAA GATAAACGGA GATGACGCCT CTAATGAGTC

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
  905      915      925      935      945      955
TGCTGTGTTG GGAAGCTTTT ATTCTTCTGT GAAACCTGGT GATGAACCAG CTAGTTTGGT

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
  965      975      985      995     1005     1015
AGCTATTAAG AGTGGTCCTC AGACCATCGG TTTTGGTTGT AGAACTAAGA TCGACGGTGA

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 1025     1035     1045     1055     1065     1075
TGACTGCCTC TTCACAGACA ATCACGTTTG GAATAATTCC ATGCGCCCTA CCGCTTTAGC

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 1085     1095     1105     1115     1125     1135
GAAAGCTGGC AAGCAGGTTG CGATTGAAGA TTGGGACACC CCTCTCTCTT GTGACCATAA

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 1145     1155     1165     1175     1185     1195
GATGCTTGAT TTCGTAGTGG TCGGTGTGCC AAAACACGTG TGGTCCAAAC TAGGAGTGAA

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 1205     1215     1225     1235     1245     1255
AGCGACTCAA CTGGTTTGTC CATCTGATAA GGACGCTGTA ACCTGTTATG GTGGATCTAG

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 1265     1275     1285     1295     1305     1315
TTCTGATAGC TTGTTGTCGG GGACGGGCGT TTGTAGTAAG GTTGATTCTT CTTGGAAGTT

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 1325     1335     1345     1355     1365     1375
AACCCACTCA TGCCCCACGG CAGCTGGCTG GAGCGGAACT CCAATTTACT CTAGCAGAGG

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 1385     1395     1405     1415     1425     1435
TGTGGTAGGA ATGCACGTTG GGTTTGAAGA TATCGGAAAA CTCAACCGTG GCGTGAACGC

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 1445     1455     1465     1475     1485     1495
TTTCTACGTG TCTAACTACT TGTTGAGGTC TCAAGAGACT CTACCTCTCG ATCTGTCCGT

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 1505     1515     1525     1535     1545     1555
CATCGAAATA CCTTTCGAGG ACGTAGAAAC CCGGAGTTAT GAGTTCATTG AGGTCGAGAT

```

```

.....|.....| .....|.....| .....|.....| .....|.....|
 1565      1575      1585      1595      1605      1615
TAAAGGAAGA GGTAAAGGCTA AACTTGGTAA GCGTGAGTTC GCTTGGATTTC CGGAATCAGG

.....|.....| .....|.....| .....|.....| .....|.....|
 1625      1635      1645      1655      1665      1675
AAAATACTGG GCCGATGATG ATGACGACTC TTTGCCCCCA CCGCCGAAGG TGGTAGACGG
.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 1685      1695      1705      1715      1725      1735
CAAGATGGTG TGGACTTCAG CTCAGGAAAC CGTCGCAGAG CCTTTAAACT ACCAGAGGGC

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 1745      1755      1765      1775      1785      1795
GGCAGGGTCA AGGCCCTTGC CGCCCTTCT CAACTTGCAG GCTACAACCT CAAAGAAGGA

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 1805      1815      1825      1835      1845      1855
GAAGCAGCCT CTACAAGAGG AATGCCCCCTT AGATTTGTTG GGCAGTCGGC TTGCAAGTTT
                                *ORF3

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 1865      1875      1885      1895      1905      1915
AGAGAGCTGT GTAGAAAAGA TACTCCAGAT GAAGTCCTTA GAGCTACTAG AGTCTTCCCC

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 1925      1935      1945      1955      1965      1975
GAATTGTCAG ACTTCTCCTG GCCTGAGCGA GGCTCCAAAG CAGAGCTTCA CTCCCTGCTA

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 1985      1995      2005      2015      2025      2035
CTCCAAGCAG GAAAGTTTAA TCCCACCGCA ATCCAAGGA ATCTTGAAGG AGCTTGTCAA

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 2045      2055      2065      2075      2085      2095
AACCTCCTTG AGCGCTACCC CGCCTCCAAA TCCTGTACT GCCTTCGTGG AGAAGCCTGG

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 2105      2115      2125      2135      2145      2155
TCCTTCGACG CAGTCTACGA AGAAGTCTGC AAGAAGGCGC AATCGGCGGA AATCAACGAG

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 2165      2175      2185      2195      2205      2215
AAAGCCAGTC CAGGGAGTCC CCCTCTCCCG TCTCGCCTCC ACCAACAAAG ACCTCCTTAA
                                                ●ORF3

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 2225      2235      2245      2255      2265      2275
GAGGCACTTA GAATTAGTTG CTTTGTGTGT TACCGAGAGA TTGTTCTTAC TCAGTGAAGC

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 2285      2295      2305      2315      2325      2335
TGAGGACCTG CACAATAAAT CCCCTGTGGA CCTAGTTCAG ATGGGGTTGT GCGACCCAGT

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 2345      2355      2365      2375      2385      2395
TCGGCTGTTT GTCAAGCAGG AGCCCCATGC TTCCCGAAAG GTGAAGGAGG GTAGATTTCG

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 2405      2415      2425      2435      2445      2455
CTTGATTTCA TCCGTTTCGC TGGTGGATCA GCTTGTAGAG CGAATGCTTT TCGGGCCCCA

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 2465      2475      2485      2495      2505      2515
AAACCAGCTT GAGATCGCTG AGTGGAACA TATTCCCTCA AAACCTGGTA TGGGCCTTTC

```

```

.....|.....| .....|.....| .....|.....| .....|.....|
 2525      2535      2545      2555      2565      2575
GCTGCAACGA CAAGCCAAAA GCTTGTTCGA CGATTTGAGA GTCAAACATT CTCGTTGTCC

.....|.....| .....|.....| .....|.....| .....|.....|
 2585      2595      2605      2615      2625      2635
TGCGGCTGAA GCCGACATAT CGGGTTTTGA CTGGTCTGTT CAAGACTGGG AGTTGTGGGC

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 2645      2655      2665      2675      2685      2695
TGATGTAGAG ATGAGAATAG TTTTAGGAGG TTTTGGACAG AAGTTGTCCA TAGCCGCTAG

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 2705      2715      2725      2735      2745      2755
AAACAGGTTT TCGTGTTTCA TGAACTCAGT CTTCCAGCTC TCGGATGGCA CACTCATAGA

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 2765      2775      2785      2795      2805      2815
ACAGCAACTG CCTGGTATTA TGAAGTCTGG TTCTTACTGC ACTTCCTCAA CAAACTCCAG

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 2825      2835      2845      2855      2865      2875
AATACGTTGC CTTATGGCTG AACTTATTGG TTCCCATGG TGTATCGCCA TGGGTGATGA

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 2885      2895      2905      2915      2925      2935
TTCTGTTGAG GGTGGGTTG ATGGTGCGAA AGACAAGTAC ATGAGATTAG GCCACACGTG

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 2945      2955      2965      2975      2985      2995
CAAGGATTAT AAACCCTGTG CAACATCCAT TTCCGGTCGC TTATACGAGG TAGAGTTTTG

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 3005      3015      3025      3035      3045      3055
CTCTCACGTT ATAAGGGAAG ATCGATGTTG GTTGGCGTCG TGGCCCAAAA CTCTGTATAA

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 3065      3075      3085      3095      3105      3115
ATACTTGTCT GAGGGCAAGT GGTTCTTTGA GGATCTTGAG CGAGAGCTTG GGTCTTCTCC

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 3125      3135      3145      3155      3165      3175
CCACTGGCCC AGAATCAGAC ACTACGTAGT CGGGAATACT CCATCGCCCG ACAAACCTAG

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 3185      3195      3205      3215      3225      3235
ATTAGAAAAT TCAAGTCCGA GCTATGGCGA AGAGGCTGAC AAAACAACAG TTAGCCAAGG

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 3245      3255      3265      3275      3285      3295
CTATAGCGAA CACTCTGGAA GCCCCGGCCA CTCAATCGAG GAGGCCAGG AACCGGAGAC

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 3305      3315      3325      3335      3345      3355
GGCGCCGTTT TGCTGCAAGG CAGCCTCAGT CTATCCAGGC TGGGGCATCC ATGGCCCCTA

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
 3365      3375      3385      3395      3405      3415
TTGCTCAGGG GGCTATGGTT CGCTTACGTG AACCATCGCT TAGAACGGCT GGAGGAGTGA

```

●ORF2

```

.....|.....| .....|.....| .....|.....| .....|.....|
 3425      3435      3445      3455      3465      3475
CTGTCTGAC GCACTCTGAG CTCTCAACAG AGCTTGCTGT GACGAATGCG ATAGTTGTCT

.....|.....| .....|.....| .....|.....| .....|.....|
 3485      3495      3505      3515      3525      3535
CTTCTGAGCT TGTCATGCCC TTCACAATGG GCACTTGGCT TCGAGGCGTG GCGGCCAATT

.....|.....| .....|.....| .....|.....| .....|.....|
 3545      3555      3565      3575      3585      3595
GGTCGAAGTA CAGTCTGTTT TCAGTGAGGT ATRACTTACCT CCCCTCTTGT CCTTCAACGA

.....|.....| .....|.....| .....|.....| .....|.....|
 3605      3615      3625      3635      3645      3655
CATCTGGGTC CATTCAATATG GGTTCCTCAAT ATGATATGGC TGACACTCTT CCCGTATCCG

.....|.....| .....|.....| .....|.....| .....|.....|
 3665      3675      3685      3695      3705      3715
TTAACCAGTT ATCCAACCTT AGAGGCTATG TGTCAGGGCA GGTCTGGTCT GGATCCTCTG

.....|.....| .....|.....| .....|.....| .....|.....|
 3725      3735      3745      3755      3765      3775
GATTGTGCTA TATAAATGGC ACGAGGTGTT CTGACACCGC CAACGCTATC ACGACCACTT

.....|.....| .....|.....| .....|.....| .....|.....|
 3785      3795      3805      3815      3825      3835
TGGACGTTGC AAAGCTTGGT AAGAAGTGGT ATCCTTTCAA GACCAGTACG GACTTCACTG

.....|.....| .....|.....| .....|.....| .....|.....|
 3845      3855      3865      3875      3885      3895
CCGCTGTTGG CGTAATAGTC AACATTGCTA CTCCCCTGGT CCCGGCTAGG CTAGTGATAG

.....|.....| .....|.....| .....|.....| .....|.....|
 3905      3915      3925      3935      3945      3955
CCATGCTGGA TGGGTCGAGT TCTACGGCTG TGAGTACTGG ACGCCTATAC GTGTCGTACA

.....|.....| .....|.....| .....|.....| .....|.....|
 3965      3975      3985      3995      4005      4015
CTGTACAGCT AATTGAGCCG ACTGCCTTGG CCTTAAACAA CTGAGGAGTT GTATAATAAT
                                     ●ORF4

.....|.....| .....|.....| .....|.....| .....|.....|
 4025      4035      4045      4055      4065      4075
ACCTGCACCC TTCTCTTTGG CAGGGAGGGT GTTTCGTTTT CACAATGCCA CGCGCTTGAG

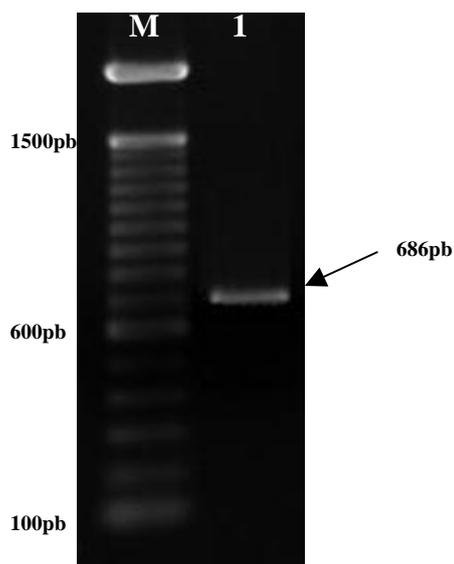
.....|.....| .....|.....| .....|.....| .....|.....| ... 3'
 4085      4095      4105      4115      4125
GGAGAATGCA CGTTAATCAT CCCTCCGCTA GTGATGGAGC GTAATCCAAA AGT

```

**Figura 15:** Sequência de nucleotídeos do Genoma do SBMV-SP. \* indica os códons de iniciação (ATG) das ORFs 1, 2, 3 e 4; (●) indica o códon de terminação (TGA para as ORFs 1, 2 e 4 e TAA para a ORF 3). Os nucleotídeos coloridos marcados em negrito e sublinhados indicam os códons de iniciação e terminação das ORFs: em vermelho ORF 1, em preto ORF 2, em verde ORF 3 e em azul ORF 4.

#### 4.10. Amplificação por PCR do Fragmento da Porção C- Terminal da RpRd.

O produto da reação de PCR foi analisado em gel nativo de agarose 1% evidenciando bandas de tamanho esperado de aproximadamente 686 pb (Figura 16).



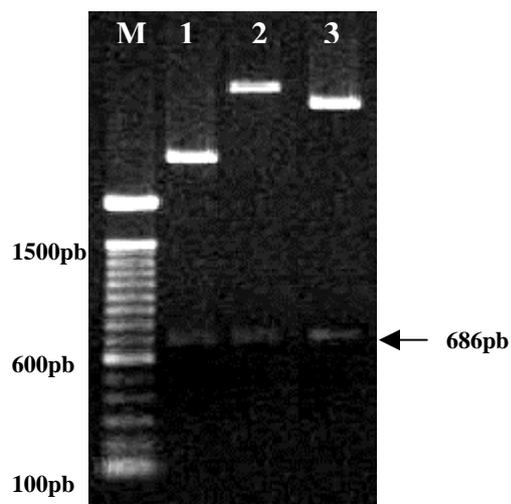
**Figura 16.** Gel nativo de agarose 1% (M): Marcador de tamanho molecular 100 bp DNA Ladder (Invitrogen); (1) Produto amplificado: 686pb (seta).

#### **4.11. Clonagem, Expressão e Purificação do Peptídeo da Porção C-Terminal da RpRd.**

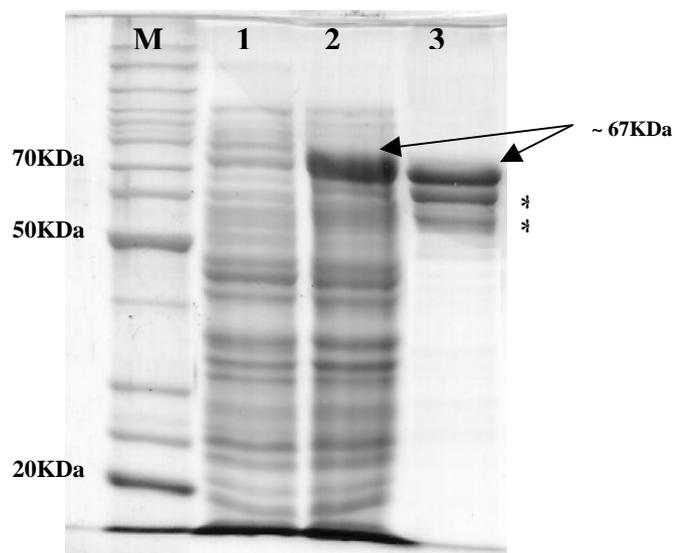
A figura 17 mostra o corte com enzimas de restrição dos vetores recombinantes (pGEM-T + inserto, pMALc-2x + inserto e pET 28 a + inserto), indicando a liberação do fragmento tamanho esperado de aproximadamente 686 pb, confirmando a presença do inserto. Através do seqüenciamento realizado dos vetores recombinantes, confirmou-se a presença do fragmento da região 3' do gene da polimerase, possibilitando iniciar o processo de expressão da proteína.

A proteína foi expressa na forma solúvel em células de *E. coli* induzidas com 1mM de IPTG por 4 horas. A análise da expressão proteica foi realizada através de SDS-PAGE 10% e revelou a produção de uma proteína de massa molecular aproximada de 67 kDa (Figura 18), conforme o esperado, uma vez que a proteína de fusão ("Maltose Binding Protein" – MBP) do vetor tem aproximadamente 43kDa e a porção C-Terminal da RpRd do SBMV-SP tem massa molecular estimada em 24kDa. As amostras que foram eluídas da resina de amilose confirmaram os resultados de expressão. A análise em SDS-PAGE 10% evidenciou uma banda com massa molecular aproximada de 67kDa, o que representa a proteína fusionada esperada (Figura 16).

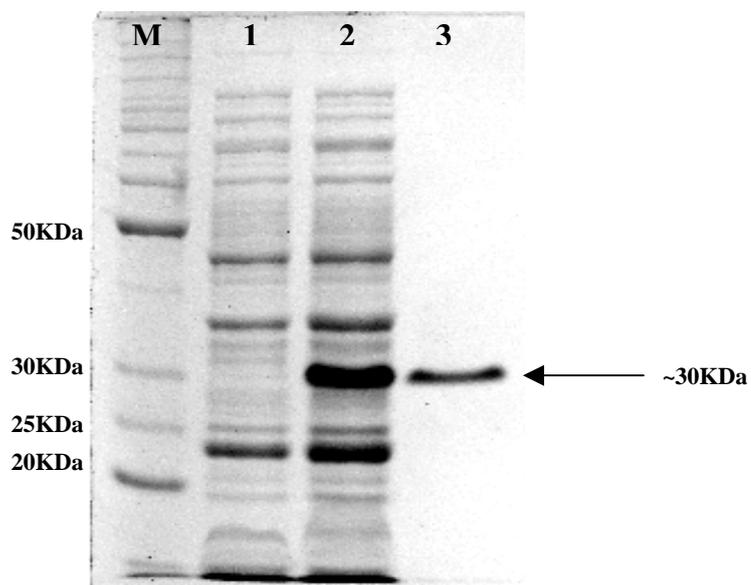
No outro sistema de expressão em *E. coli* (pET 28 a) a cultura foi submetida a uma indução com 1mM IPTG por 3 horas. A análise da expressão também foi realizada através de um SDS PAGE 10%. A proteína também foi obtida na forma solúvel, com um tamanho aproximado de 30kDa. As amostras obtidas da eluição da resina de níquel (Invitrogen) foram aplicadas em SDS-PAGE 10% e analisadas, A purificação da proteína confirma a expressão da proteína recombinante conforme a figura 19.



**Figura 17:** Gel nativo de agarose 1% corte enzimático com *Bam* HI e *Hind* III dos vetores recobinantes que contém o fragmento da polimerase. **(M)** Marcador de tamanho molecular (100 pb DNA Ladder - Invitrogen). **(1)** DNA plasmidial (pGEM) recombinante digerido com *Bam* HI e *Hind* III. **(2)** DNA plasmidial (pMAL) recombinante digerido com *Bam* HI e *Hind* III. **(3)** DNA plasmidial (pET 28a) recombinante digerido com *Bam* HI e *Hind* III. **(Seta)** indicando o fragmento de interesse liberado de 686 pb.



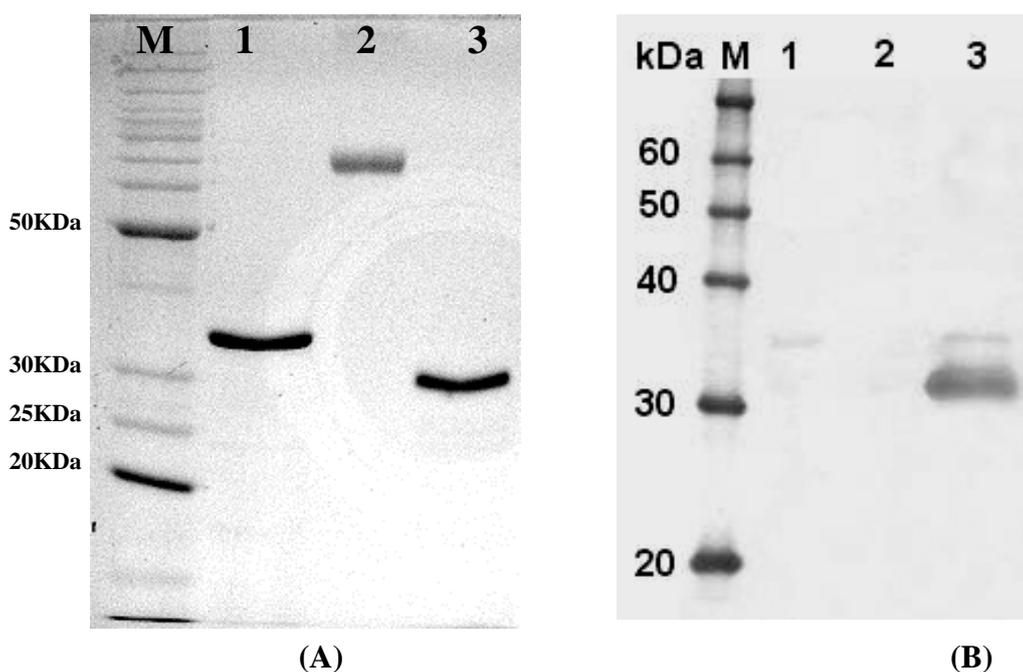
**Figura 18:** SDS-PAGE 10% mostrando a expressão em *E. coli* e purificação parcial da proteína referente a porção C-Terminal da RdRp fusionada a MBP em coluna de amilose. **(M)** Marcador de massa molecular (kDa) BenchMark (Invitrogen). **(1)** Extrato de células antes da indução da expressão. **(2)** Extrato de células após 4 horas de indução com 1mM de IPTG. **(3)** Proteína referente a porção C-terminal da RpRd fusionada a MBP parcialmente purificada com massa molecular estimada em ~67 kDa (seta); (\*) contaminantes.



**Figura 19:** SDS-PAGE 10% mostrando a expressão em *E. coli* e purificação parcial da proteína referente a porção C-Terminal da RdRp fusionada a cauda de histidina em resina de níquel (Invitrogen). **(M)** Marcador de massa molecular (kDa) BenchMark (Invitrogen). **(1)** Extrato de células antes da indução da expressão. **(2)** Extrato de células após 3 horas de indução com 1mM de IPTG. **(3)** Proteína referente a porção C-terminal da RpRd fusionada a cauda histidina purificada com massa molecular estimada em ~ 30 kDa (seta).

#### 4.12. Western blot.

A especificidade foi confirmada ao reagir o anti-soro anti-C-terminal da polimerase com a proteína recombinante (C-terminal da polimerase + cauda de histidina) expressa no sistema pET 28 a. Uma banda de aproximadamente 30 kDa mostra a reação.



**Figura 20:** (A) SDS PAGE 12%: (M) marcador de massa molecular (BenchMark™ Protein Ladder, Invitrogen); (1) proteína do movimento do *Pepper ringspot virus*; (2) BSA e (3) Porção C-terminal da Polimerase do SBMV. (B) Western blot: (M) marcador de massa molecular (Magic Mark XP, Invitrogen); (1) Proteína do movimento do *Pepper ringspot virus*; (2) BSA e (3) Porção C-Terminal da Polimerase do SBMV.

## 5. DISCUSSÃO

Os resultados obtidos neste trabalho mostram que a ORF 2 do SBMV-SP contém 2889 nucleotídeos, incluindo-se o códon de terminação UGA, com 962 aminoácidos deduzidos e massa molecular estimada de aproximadamente 105 kDa. Dentro desta ORF foi localizada a seqüência da ORF3 contendo 398 nucleotídeos, incluindo-se o códon de terminação UAA, com 132 aminoácidos deduzidos e massa molecular estimada de aproximadamente 15 kDa.

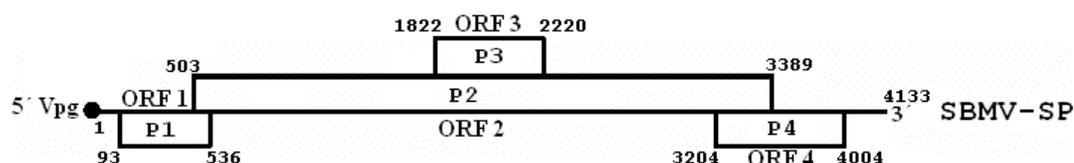
A análise feita a partir do seqüenciamento das ORFs 2 e 3 do genoma do SBMV-SP, quando comparadas com outras espécies do mesmo gênero (*Sobemovirus*) e isolados do SBMV, depositadas no GenBank, mostrou que a ORF 2 possui maior identidade (91,4% na seqüência de nucleotídeos e 95,0% na seqüência de aminoácidos deduzidos) com o isolado Arkanas do SBMV (Tabela 1). O resultado foi similar em relação à ORF3 com valores de identidade de 97,0% para a seqüência de nucleotídeos e 97,7% para a de aminoácidos deduzidos (Tabela 2).

Os dados obtidos por Espinha *et al.* (2004) através do seqüenciamento da ORF 1 do SBMV-SP mostram que as seqüências de nucleotídeos e aminoácidos deste vírus apresentam maior identidade (79%) com o isolado americano (SBMV-Ark) e pouca identidade com as demais espécies de vírus do gênero *Sobemovirus*. A extremidade 3' terminal da ORF1 do SBMV-SP se sobrepõe em 34 nucleotídeos com a ORF 2 (ESPINHA *et al.*, 2004), sendo que este dado foi também confirmado no seqüenciamento da ORF 2 do SBMV-SP (Figura 15).

O seqüenciamento realizado para a ORF 4 do isolado São Paulo do SBMV (ESPINHA *et al.*, 2005) mostrou que há sobreposição de 129 nucleotídeos com a ORF 2, resultado também obtido no presente trabalho. As seqüências de nucleotídeos e aminoácidos da ORF 4 também apresentam maior identidade com as seqüências de nucleotídeos e aminoácidos do SBMV-Ark, 91% e 93% respectivamente. Estes dados de identidade das ORFs 1 e 4 confirmam os dados de identidade das ORFs 2 e 3 deste trabalho.

Os dados obtidos pela análise filogenética pelo método da máxima parcimônia para as seqüências de aminoácidos das ORFs 2 e 3, corroboram os dados de identidade. Para a ORF 3, os isolados SBMV-SP e SBMV-Ark foram agrupados em um mesmo ramo em 100% das réplicas, valor próximo dos 96% obtidos para a ORF 2. Deve-se notar, entretanto, que na análise filogenética das seqüências de nucleotídeos dos genomas dos *Sobemovirus*, os isolados SBMV e o SBMV-Ark foram agrupados num mesmo ramo em 100% das réplicas, o mesmo acontecendo quando da análise das ORF 2. Este fato pode ser explicado pelo fato de que o SBMV e o SBMV-Ark são isolados encontrados nos Estados Unidos da América do Norte e o SBMV-SP encontrado no Brasil, indicando haver alguma diversidade genética entre eles mas que não redundam, entretanto, em mudanças na seqüência de aminoácidos. A análise filogenética para a ORF3 não foi feita visto que no seqüenciamento do SBMV inexistente esta ORF (OTHMAN & HULL, 1995). Mais recentemente, Hull & Fargette (2005) mostram a existência de uma ORF 3 no genoma do SBMV, iniciando na posição 1826 e terminando na posição 2296. Entretanto, esta seqüência não codifica uma proteína, havendo vários códons de terminação na seqüência. Já foi proposto que as diferenças na organização genômica do SBMV e SBMV-Ark, resultam de mutações ou erro no seqüenciamento do SBMV (LEE & ANDERSON, 1998).

Reunindo os resultados obtidos de seqüenciamento deste trabalho com os dados de seqüenciamento obtidos por Espinha *et al.* (2004, 2005) das ORFs 1 e 4 do SBMV-SP, foi possível concluir que o genoma do *Southern bean mosaic virus* encontrado no Estado de São Paulo (SBMV-SP) possui 4133 nucleotídeos, compreendendo três ORFs (1, 2 e 4) sobrepostas e uma (ORF3) inserida na ORF2, apresentando organização genômica similar àquela descrita para o SBMV-Ark (LEE & ANDERSON, 1998) e dentro do padrão de organização representado pelo *Southern cowpea mosaic virus* (SCPMV). Um esquema da organização genômica do SBMV-SP é mostrado a seguir.



Analisando a região referente a RpRd foi possível identificar uma seqüência de heptanucleotídeos (TTTAAAC) com início na posição 1723 e término na posição 1729 (figura 6). Esta seqüência possivelmente atua como uma região sinalizadora para o mecanismo ribossomal de *frameshift* -1 para a tradução da ORF 3, semelhante ao que ocorre com o CoMV para a ORF 2b (LEE & ANDERSON, 1998; TAMM *et al.*, 1999). Assim, a ORF3 seria traduzida por um mecanismo de *frameshift* -1, o que produziria uma proteína de fusão de aproximadamente 70 kDa. Tradução *in vitro* do RNA genômico do SCPMV produziu uma proteína de aproximadamente 70 kDa, o que poderia ser a proteína de fusão da ORF2-ORF3 (WU *et al.*, 1987; TAMM *et al.*, 1999). Entretanto, ainda são necessários mais dados experimentais para comprovação de tal hipótese.

Através da análise das seqüências de aminoácidos, foi possível identificar outras regiões conservadas, localizadas na ORF 2, citando-se o sítio de ligação da capa proteica (*Coat Protein Binding Site* - CPBS) e a porção N-terminal da seqüência da VPg. A CPBS, inserida na região da serino protease, mostrou ser uma seqüência muito conservada entre os sobemovírus (Figura 9). Esta região é importante para a formação do complexo ribonucleoproteico (RNPC) propiciando estabilização dos vírions (HACKER, 1995). O aminoácido serina (S) em destaque na Figura 7 está presente em todas as seqüências dos sobemovírus analisadas, mostrando ser de grande importância, já que faz parte do sítio catalítico da serino protease. Também foi possível identificar a posição da tríade catalítica H-D-S (Histidina, posição 181-Ácido Aspártico, posição 216-Serina, posição 284) presente nas serino proteases. Esta tríade catalítica também é encontrada no SCPMV, SeMV e outros sobemovírus (LOKESH *et al.* 2001).

O alinhamento da porção N-terminal da VPg mostrou que esta região é altamente conservada nas espécies SBMV-SP; SBMV-Ark; SeMV e SCPMV, o que não se verifica para as espécies CoMV (MÄKINEN *et al.*, 1995) e RYMV (NGON *et al.*, 1994). Os sítios de clivagem para a serino protease E/T (E- Glutamato e T- Treonina) também foram identificados, estando localizados entre as seqüências SP-VPg e VPg-RpRd (Figura 8), de maneira similar ao descrito para espécies do gênero *Luteovirus* (VAN DER WILK *et al.*, 1997). Uma outra

seqüência altamente conservada nas VPg dos sobemovírus, representada pelos aminoácidos WAD/WGD seguidos por um domínio rico em D/E (MÄKINEN *et al.*, 1995), foi também identificada no SBMV-SP (Figura 8), onde a seqüência WAD é seguida de um domínio rico em D. Motivos similares a este também estão presentes nas VPg do *Potato leafroll virus* (PLRV) e *Beet western yellows virus* (BWYV).

A expressão da porção C-terminal do polimerase do SBMV-SP em *E. coli* no sistema pMAL c2-x produziu uma proteína de fusão de aproximadamente 67 kDa em forma nativa. Entretanto, mesmo após purificação em coluna de afinidade de amilose, alguns contaminantes ainda estavam presentes na preparação (Figura 18). Mesmo assim, alíquotas foram utilizadas para imunização de coelhos visando a produção de anti-soro específico. Utilizando-se a expressão em sistema pET 28a, houve a produção de uma proteína não fusionada, desnaturada, de aproximadamente 30 kDa (Figura 19). O anti-soro obtido, quando testado em *Western Blot*, mostrou-se específico para a porção C-terminal da polimerase, não reagindo com outras proteínas utilizadas como controle (Figura 20).

Este anti-soro será de grande valia em estudos futuros, possibilitando melhor entender o ciclo de multiplicação do *Southern bean mosaic virus* na célula hospedeira, uma vez que ainda pouco é conhecido sobre aspectos da replicação dos sobemovírus.

## **6. CONCLUSÃO**

- O isolado do SBMV encontrado no Estado de São Paulo é idêntico ao isolado do mesmo vírus descrito no Estado de Arkansas nos Estados Unidos da América.
- O anti-soro produzido para a proteína de fusão (C-terminal da polimerase + MBP) reagiu eficientemente com a proteína recombinante purificada expressa no sistema pET 28a (C-terminal da polimerase + cauda de histidina), mostrando ser específico para essa proteína.

## 7. RESUMO

O presente trabalho consistiu no seqüenciamento e caracterização molecular das Cadeias Abertas de Leitura (*Open Reading Frames* - ORFs) 2 e 3 do genoma do isolado São Paulo do *Southern bean mosaic virus* (SBMV-SP), completando-se o seqüenciamento de todo o genoma desse isolado. A ORF 2 codifica uma poliproteína (serino protease – VPg – RNA polimerase RNA-dependente) e a ORF 3 um produto com função desconhecida.

O seqüenciamento da ORF 2 apresentou 2889 nucleotídeos, incluindo-se o códon de terminação UGA, com 962 aminoácidos deduzidos e massa molecular estimada de aproximadamente 105 kDa. Dentro da ORF 2, localiza-se a ORF 3 contendo 398 nucleotídeos, incluindo-se o códon de terminação UAA, com 132 aminoácidos e massa molecular estimada de aproximadamente 15 KDa.

A análise feita a partir das seqüências das ORFS 2 e 3 do genoma do SBMV-SP, quando comparadas com outras espécies do mesmo gênero e isolados do SBMV, depositadas no GenBank, mostrou que a ORF 2 apresenta maior identidade (91,4% na seqüência de nucleotídeos e 95,0% na seqüência de aminoácidos deduzidos) com o isolado de Arkansas. Resultado similar foi obtido em relação à ORF 3 com valores de identidade de 97,0% tanto para as seqüências de nucleotídeos e aminoácidos deduzidos. Dados de filogenia corroboram os dados de identidade.

Regiões conservadas do gênero *Sobemovirus* também foram identificadas, tais como Sítio de Ligação à Capa Protéica (CBPS), a tríade catalítica da serino protease (H-D-S) e a seqüência de heptanucleotídeos (TTTAAAC).

A expressão em *Escherichia coli* da porção C-terminal da RNA Polimerase RNA Dependente (RpRd) produziu uma proteína de fusão de aproximadamente 67 kDa no sistema pMAL c2-x e de 30 kDa no sistema pET 28a. Quando a proteína de fusão foi injetada em coelhos houve a produção de anti-soro específico para a proteína recombinante.

## 8. ABSTRACT

The present work consisted of the sequencing and molecular characterization of the Open Reading Frames (ORFs) 2 and 3 from the São Paulo isolate genome of *Southern bean mosaic virus* (SBMV-SP), completing the sequencing of all the genome of this isolate. The ORF 2 encodes a polyprotein (serine protease - VPg - RNA dependent RNA polimerase) and the ORF 3 one product with unknown function.

The sequencing of the ORF 2 reveals 2889 nucleotides, including the stop codon (UGA), with 962 deduced amino acids and estimated molecular weight of approximately 105 kDa. Nested in the ORF 2 were found the ORF 3 with 398 nucleotides, including the stop codon (UAA), with 132 deduced amino acids and estimated molecular weight of approximately 15 kDa.

The analysis made from the sequences of the ORFs 2 and 3 from the SBMV-SP genome, when compared with other species of the same gender and isolates of SBMV, deposited in the GenBank, showed that the ORF 2 presents higher identity (91,4% in the nucleotide sequence and 95,0% in the deduced amino acids sequence) with Arkansas isolate. Similar result was obtained in relation to the ORF 3 with identity values of 97,0% for the nucleotides and deduced amino acids sequences. Phylogeny data corroborate the identity data.

Conserved regions of the *Sobemovirus* gender had also been identified such as Coat Protein Binding Site (CPBS), serine protease catalytic triad (H-D-S) and heptanucleotide sequence (TTTAAAC).

The Expression in *Escherichia coli* of the C-terminal region of the RNA dependent RNA polymerase produced a fusion protein of approximately 67 kDa in pMAL c2-x system and a 30 kDa protein in pET 28a system. When the fusion protein was injected in rabbits it had the production of specific anti-serum for the recombinant protein.

## 9. BIBLIOGRAFIA

BIANCHINI, A.; MARINGONI, A. C.; CARNEIRO, S. M. T. P. G. Doenças de feijoeiro. In: Kimati, H., Amorin, L. Bergamin Filho, A., Camargo, L. E. A., Rezende, J. A. M. (Eds.) Manual de fitopatologia, 3ª ed., Agronômica Ceres, 1997. Pp.376-399.

BONNEAU, C; BRUGIDOU, C.; CHEN, L.; BEACHY, R.; FAUQUET, C. Expression of the *Rice yellow mottle virus* P1 *in vitro* translation and *in vivo* and its involvement in virus spread. *Virology*, v.244, p. 79-86, 1998.

BRUGIDOU, C.; HOLT, C.; NGON A YASSI, M.; ZHANG, S.; BEACHY, R.; FAUQUET, C. Synthesis of an infectious full-length cDNA clone of *Rice yellow mottle virus* and mutagenesis of the coat protein. *Virology*, v.206, p. 108-115, 1995.

CARETTE, J. E.; VAN LENT, J.; MAC FARLANE, S. A.; WELLINK, J. VAN KAMMEN, A. Cowpea mosaic virus 32-and 60 kilodalton replication proteins target and change the morphology of endoplasmatic reticulum membranes. *Journal of Virology*, v. 76 (12), p. 2313-2320, 2002.

CONAB (Companhia Nacional de Abastecimento), 2006. Disponível em <<http://www.conab.gov.br/>> Acesso em: 28 de julho de 2006.

COSTA, A. S.; KITAJIMA, E. W.; MIYASAKA, S; ALMEIDA, L. D. Moléstias de feijoeiros causadas por vírus. In: I Simpósio Brasileiro do Feijão/Viçosa, Anais (Vieira, C., ed). Imprensa Univ., p. 342-384, 1972.

COSTA, C. L.; BATISTA, M. F. Viroses transmitidas por coleópteros no Brasil. *Fitopatologia Brasileira*, v. 4, p. 177-179, 1979.

COSTA, A. F.; COSTA, C. L. Efeitos de infecções simples e duplas de quatro vírus sobre a produção e algumas características da planta do feijoeiro (*Phaseolus vulgaris* L.). Fitopatologia Brasileira, v. 8, p. 616, 1983.

CRONQUIST, A. An integrated system of classification of flowering plants. Columbia University Press, New York, p.1262, 1981.

CULTURA-FEIJÃO; Bahia, 2000. Disponível em <<http://www.seagri.ba.gov.br/feijão1.htm>> Acesso em: 24 de janeiro 2001.

CUPERTINO, F. P.; LIN, M. T.; KITAJIMA, E. W.; COSTA, C. L. Occurrence of southern bean mosaic virus in central Brazil. Plant Disease, v. 66, p. 742-743, 1982.

DE ZOETEN, G. A.; GAARD, G. Possibilities for inter-and intracellular translocation of some icosahedral plant virus. Journal of Cell Biology, v. 40, p. 814-823, 1969.

DEN BOON, J. A.; CHEN, J. AHLQUIST, P. Identification of sequences in *Brome mosaic virus* replicase protein 1a that mediate association with endoplasmatic reticulum membranes. Journal of Virology, v. 75 (24), p. 12370-12381, 2001.

DINGWALL, C.; LASKEY, R. A. Nuclear targeting sequence – A consensus. Trends Biochemistry Science, v. 16, p. 478-481, 1991.

ESPINHA, L. M.; GASPAR, J. O.; WARD, R. J.; RULLER, R., CAMARGO, L. E. A. Caracterização da região 5' - Terminal de um isolado brasileiro do *Southern bean mosaic vírus*. Fitopatologia Brasileira. V. 29, p. 328-331, 2004.

ESPINHA, L. M.; GASPAR, J. O.; MOREIRA, A. E.; PEREIRA, A. C. B.; BELINTANI, P.; CAMARGO, L. E. A. Caracterização da região 3' - Terminal de um isolado brasileiro do *Southern bean mosaic vírus*. Fitopatologia Brasileira. V. 30, p. 546-549, 2005.

FAUQUET, M. C.; MAYO, M. A. Abbreviations for plant names. Archives of Virology, v. 144, p. 1249-1273, 1999.

FICHA TÉCNICA-FEIJÃO, 2000. Disponível em <<http://www.consulcoop.com.br.agr.produtosfeijão.htm>>.

FRANCKI, R. I. B., MILNE, R. G.; HATTA, T. Sobemovirus group. In: Atlas of Plant Viruses. CRC Press, Boca raton, Fl., v.1. p. 153-169, 1985.

GAMEZ, R. Lãs enfermedades virales como factores limitantes em la produccion de frijol (*Phaseolus vulgaris* L.) em América latina. Fitopatologia, v. 12, p. 24-26, 1977.

GASPARIN, M. D. G. Detecção e caracterização parcial de um isolado do Vírus do Mosaico do Feijoeiro do Sul do Paraná. Maringá, 2002. 41f. Dissertação (Mestrado) - Universidade Estadual do Maringá.

GORBALENYA, A. E.; KOONIN, E.V.; BINOV, V. M.; DONCHENKO, A. P. Sobemovirus genome appears to encode a serine protease related to cysteine proteases of picornaviruses. FEBS Letter, v. 236, p. 287-290, 1988.

GOSH, A.; DASGUPTA, R.; SALERNO-RIFE, T.; RUTGERS, T.; KAESBERG, P. *Southern bean mosaic virus* has a 5'-linked protein but lacks 3' terminal poly(A). Nucleic Acids Research, v. 7, p. 2137-2146, 1979.

GOSH, A.; RUTGERS, T.; AMNG, K. Q.; KAESBERG, P. Characterization of coat protein mRNA of *Southern bean mosaic virus* and its relationship to the genomic RNA. *Journal of Virology*, v. 39, p. 87-92, 1981

HACKER, D. L. Identification of coat protein binding site on *Southern bean mosaic virus* RNA. *Virology*, v. 207, p. 562-565, 1995.

HACKER, D. L.; SIVAKUMARAN, K., Mapping and expression of *Southern bean mosaic virus* genomic and subgenomic RNAs. *Virology*, v. 234, p. 317-327, 1997.

HULL, R. The Sobemovirus group. In: *The Plant Viruses. Polyhedral virions with monopartite genomes* (Koenig R., ed.), Plenum Press, v.3, p. 113-146, 1988.

HULL, R. AND FARGUETTE, D. Genus *Sobemovirus*. In: *Virus Taxonomy: eighth report of the International Committee on Taxonomy of Viruses*. Academic Press, California, USA, 2005.

JAYASINGHE, W. U. Chlorotic mottle of bean (*Phaseolus vulgaris* L.). CIAT series 09EB (2), Cali, Colombia, v. 82, p. 156, 1982.

KIBERSTIS, P. A. AND ZIMMERN, D. Translational strategy of *Solanum nodiflorum mottle virus* RNA: synthesis of coat protein precursor in vitro and in vivo. *Nucleic Acids Research*, v, 12, p. 933-943, 1984.

KOONIN, E. V. The phylogeny of RNA-dependent RNA-polymerases positive-strand RNA viruses, *Journal of General Virology*, v. 72, p. 2197-2206, 1991.

KOONIN, E., V. AND DOLJA, V., V. Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Critical Reviews in Biochemistry and Molecular Biology*, v 28, p. 375-430, 1983.

LAEMMLI, U. K.. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, v. 227, p. 680-685, 1970.

LEE, L.; ANDERSON, J. Nucleotide sequence of a resistance breaking mutant of *Southern bean mosaic virus*. *Archives of Virology*, v. 143, p. 2189-2201, 1998.

LOKESH, G. L.; GOPINATH K., SATHESHKUMAR P. S.; SAVITRI H. S. Complete nucleotide sequence of *Sesbania mosaic virus*: a new virus species of the genus Sobemovirus. *Archives of Virology*, v. 146, p. 209-223, 2001.

MÄKINEN, K., TAMM, T., NAESS, V., TRUVE, E. PUURAND, U. MUNTHE, T. AND SAARNA, M. Characterization of *Cocksfoot mottle virus* genomic RNA and sequence comparisons with related viruses. *Journal of General Virology*, v. 76, p. 2817- 2825, 1995.

MANG, K. Q.; GOSH, A.; KAESBERG. P. A comparative study of the cowpea and bean strains of *Southern bean mosaic virus*. *Virology*, v. 116, p264-274, 1982.

MEYER, M. C.; COSTA, C. L.; SANTOS, O. R. Inter-relações de *Ceratoma arcuata* (Coleoptera: Chysomelidae) com quatro vírus de leguminosas. *Fitopatologia Brasileira*, v. 17, p. 314-320, 1992.

MEYER, M. C.; COSTA, C. L.; SANTOS, O. R. Transmissão de quatro vírus de leguminosas pelas larvas de *Ceratoma arcuata* (Coleoptera: Chysomelidae). *Fotopatologia Brasileira*, v. 18, p. 97-101, 1993.

MOREIRA, A. E.; GASPAR, J. O. Propriedades moleculares de um isolado brasileiro do *Southern bean mosaic virus*. *Fitopatologia Brasileira*, v. 27, p. 292-297, 2002.

MOORIS-KRSINICH, B. A. M. AND HULL, R. Translation of *Turnip rosette virus* RNA in rabbit reticulocyte lysate and wheat germ extract. *Virology*, v. 114, p. 98-112, 1981.

MOORIS-KRSINICH, B. A. M. AND FOSTER, R. L. S. *Lucerne transient streak virus* RNA and its translation in rabbit reticulocyte lysate and wheat germ extract. *Virology*, v. 128, p. 176-185, 1983.

NGON A YASSI, M.; RITZENTHALER, C.; BRUGIDOU, C.; FAUQUET, C.; BEACHY, R. N. Nucleotide sequence and genome characterization of *Rice yellow mottle virus* RNA. *Journal of General Virology*, v. 75, p. 249-257, 1994.

OPALKA, N.; BRUGIDOU, C.; BONNEAU, C.; NICOLE, M.; BEACHY, R. N.; YAGER, M. FAUQUET, C. Movement of *Rice yellow mottle virus* between xylem cells through pit membranes. *Proceedings of National Academy of Science, USA*, v. 95, p. 3323-3328, 1998.

OTHMAN, Y.; HULL, R. Nucleotide sequence of the bean strain of *Southern bean mosaic virus*. *Virology*, v. 206, p. 287-297, 1995.

REVIL, P. A.; DAVIDSON, A. D. AND WEIGHT, P. J. The nucleotide sequence and genome organization of *Mushroom bacilliform virus*: a single-stranded RNA virus of *Agaricus bisporus* (Lange). *Virology*, v. 202, p. 904-911, 1994.

REVIL, P. A.; DAVIDSON, A. D. AND WEIGHT, P. J. Identification of subgenomic mRNA encoding the capsid protein of *Mushroom bacilliform virus*, a single-stranded RNA mycovirus. *Virology*, v. 260, p. 273-276, 1999.

RUTGERS, T.; SALERNO-RIFE, T.; KAESBERG, P. Messenger RNA for the coat protein of *Southern bean mosaic virus*. *Virology*, v. 104, p.506-509, 1980

RYABOV, E. V. KRUTOV, A. A.; NOVIKOV, V. K.; ZHELEZNIKOVA, O. V.; MOROZOV, S. Y.; ZAVRIEV, S. K. Nucleotide sequence of RNA from the sobemovirus found in infected cocksfoot shows a luteovirus-like arrangement of the putative replicase and protease genes. *Phytopathology*, v. 86, p. 391-397, 1996.

SALERNO-RIFE, T.; RUTGERS, T., KAESBERG, P. Translation of southern bean mosaic virus RNA in wheat germ embryo and rabbit reticulocyte extracts. *Journal of Virology*, v. 34, p. 51-58, 1980.

SAMBROOK, J.; FRITSCH E. F.; MANIATS, T. *Molecular cloning: a laboratory manual* 2 ed. New York Cold Spring Harbor Laboratory, v. 3, 1989.

SANGER, F.; NICKLEN, S. AND COULSON, A. R. DNA sequencing with chain terminating inhibitors. *Proceedings of National Academy of the Science*, v. 74, p. 5463-5467, 1977.

SATHESHKUMAR, P. S.; LOKESH, G. L.; SAVITRI. H. S., Polyprotein processing: cis and trans proteolytic activities of *Sesbania mosaic virus* serine protease. *Virology*, v. 318 (1), p. 429-438, 2004.

SEGHAL, O. P. Southern Bean Mosaic Virus Group. In: *Handbook of plant Virus Infections and Comparative Diagnosis* (Kurstak, P., ed.) p. 91-121. Elsevier/North Holland, 1981.

SEGUNDO, E., GIL-SALAS, F. M., JANSSEN, D., MARTIN, G. E CUDRADO, I. M. First Report of *Southern bean mosaic virus* Infecting French Bean in Morocco. *Plant Disease*, v. 88, p. 1162, 2004.

SHEPHERD, R. J., Southern bean mosaic virus. In: Description of Plant Viruses. Commonwealth Mycological Institute/Association of Applied Biologists Kew, Surrey, England, nº 57, 1971.

SILVEIRA JR., W. G.; COSTA, C. L.; COSTA, A. F.; LIN, M. T. *Diabrotica speciosa* e *Cerotoma arcuata*, vetores do vírus do mosaico do feijoeiro no Distrito Federal. Fitopatologia Brasileira, v. 8, p. 625, 1983.

SIVAKUMARAN, K.; BENJAMIN, C. F., HACKER, D. L. Identification of viral genes required for cell-to-cell movement of *Southern bean mosaic virus*. Virology, v. 252, p. 376-386, 1998.

SWOFFORD, D. PAUP: Phylogenetic analysis using parcimony. Version 4.0b10. Smithsonian Institution, Washington, D. C., 1997.

TAMM, T; MÄKINEM, K. AND TRUVE, E. Identification of genes encoding for the *Cocksfoot mottle virus* proteins, Archives of Virology, v. 144, p. 1557-1567, 1999.

TAMM, T.; TRUVE, E. RNA-binding activities of Cocksfoot mottle sobemovirus protein. Virus Research, v. 66, p. 197-207, 2000a.

TAMM, T. AND TRUVE, E. Sobemoviruses. Journal of Virology, v. 74, p. 6231-6241, 2000b.

VAN DER WILK, F.; HUISMAN, M. J.; CORNELISSEN, B. J.; HUTTINGA, H. AND GOLDBACH, R. Nucleotide sequence and organization of *Potato leafroll virus* genomic RNA. FEBS Letter, v. 245, p. 51-56, 1989.

VAN DER WILK, F.; VERBEEK, M.; DULLEMANS, A. M. AND VAN DEN HEUVEL, J. F. J. M. The genome linked-protein of potato leafroll virus is located

downstream of the putative protease domain of the ORF 1 product. *Virology*, v. 234, p. 300-303, 1997.

VAN DER WILK, F.; VERBEEK, M.; DULLEMANS, A.; VAN DEN HEUVEL, J. F. J. M.; The genome-linked protein (VPg) of Southern bean mosaic virus is encoded by the ORF2. *Virus Genes*, v. 17, p. 21-24, 1998.

VERHOEVEN, J. Th.J., ROENHORST, J. W., LESEMANN, D. -E., SEGUNDO, E., VELASCO L., RUIZ L., JANSSEN D. E CUADRADO, I. M. *Southern bean mosaic virus* the Causal Agent of a New Disease of *Phaseolus vulgaris* Beans in Spain. *Journal of Plant Pathology* v. 109, p. 935-941, 2003.

VOINNET, O.; PINTO, Y. M. AND BAULACOMBE, D. C. Supression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. *Proceedings of National Academy of the Science*, v. 96, p. 14147-14152, 1999.

WEBBER, K. A.; SEHGAL, O. P. Subgenomic RNAs in virions of *Southern bean mosaic virus*. *Phytopathology*, v. 72, p. 909-913, 1982.

WEINTRAUB, M.; RAGETLI, H. W. J. Electron microscopy of the bean and cowpea strains of Southern bean mosaic virus within leaf cells. *Journal of Ultrastructural Research*, v. 32, p. 167-189, 1970.

WOBUS, C. E.; SKAF, J. S.; SCHULTZ, M. AND DE ZOETEN, G. A. Sequencing genomic localization and initial characterization of the VPg of *Pea enation mosaic anamovirus*. *Journal of General Virology*, v. 79, p. 2023-2025, 1998.

WU, S.; RINEHART, C. A.; KAESBERG, P. Sequence and organization of Southern bean mosaic virus genomic RNA. *Virology*, v. 161, p. 73-80, 1987.

XIONG, Z. AND LOMMEL, S. A. The complete nucleotide sequence and genome organization of *Red clover necrotic mosaic virus* RNA-1. *Virology*, v. 171, p. 543-554, 1989.

YASSI, M. N.; RITZENTHALER, C.; BRUGIDOU, C., FAUQUET, C. AND BEACHY, R. N. Nucleotide sequence and genome characterization of *Rice yellow mottle virus* RNA. *Journal of General Virology*, v.75, p. 249-257, 1994.

ZAUMEYER, W. J.; HARTER, L. L. A new virus diseases of bean. *Phytopathology*, v. 32, p. 438, 1942.

ZAVRIEV, S. K.; HICKEY, C. M. AND LOMMEL, S. A. Mapping of the *Red clover necrotic mosaic virus* subgenomic RNA. *Virology*, v. 216, p. 407-410, 1996.