

布尼亚病毒科全基因组序列比对分析*

刘雅婷^{1,2,3}, 张文超¹, 李正跃^{2**}, 李成云³, 朱有勇³, 李永忠⁴

(1. 云南农业大学农学与生物技术学院, 云南 昆明 650201; 2. 云南农业大学植物保护学院, 云南 昆明 650201;
3. 云南省植物病理重点实验室, 云南 昆明 650201; 4. 云南农业大学烟草学院, 云南 昆明 650201)

摘要: 布尼亚科病毒 (Bunyaviridae) 是一类医学和农业上的重要病毒。本文从 NCBI 数据库下载具有完整基因组序列的布尼亚科病毒的 5 个属 20 种病毒的序列, 用生物学软件 DNAMAN, DNASTAR 进行比对分析, 发现布尼亚科病毒属中的植物病毒在核酸序列及蛋白结构上与动物病毒有很大差异: (1) 只有植物病毒的 M 基因组能编码 NSm 运动蛋白; (2) 植物病毒和动物病毒在核酸序列和蛋白质序列长度均有差异, 表明该科病毒是进化速度较快的病毒; (3) 植物病毒核酸序列的 GC 含量低于动物病毒; (4) 通过 SMART 网络软件进行蛋白质拓扑结构分析发现植物病毒和动物病毒在糖蛋白 G_nG_c 的结构上存在显著差异; (5) 植物病毒糖蛋白结构较为复杂, 有较多的紊乱区域, 除 INSV 外, 其他病毒都具有 N 端信号肽。这将对以后分子生物学检测中引物的设计及病毒的鉴定等方面的研究起到积极的作用。

关键词: 布尼亚病毒科; 番茄斑萎病毒属; DNAMAN; DNASTAR; SMART; 序列比对

中图分类号: Q 939.47 **文献标识码:** A **文章编号:** 1004-390X (2008) 03-0315-06

Alignment and Analysis of Complete Genome of Bunyaviridae

LIU Ya-ting^{1,2,3}, ZHANG Wen-chao¹, LI Zheng-yue²,
LI Cheng-yun³, ZHU You-yong³, LI Yong-zhong⁴

(1. Faculty of Agricultural and Biotechnology, Yunnan Agricultural University, Kunming 650201, China;
2. Faculty of Plant Protection, Yunnan Agricultural University, Kunming 650201, China;
3. Phytopathology Laboratory of Yunnan Province, Yunnan Agricultural University, Kunming 650201, China;
4. Faculty of Tobacco Science, Yunnan Agricultural University, Kunming 650201, China)

Abstract: The Bunyaviridae family is a group of very serious medical and agricultural viruses. Complete genome sequences of 20 viruses in *Bunyaviridae* loaded from NCBI were analyzed by two biological softwares of DNAMAN and DNASTAR. The results indicated great variability in both nucleic acid sequence and protein structure between plant viruses and animal viruses: (1) only plant viruses could encoded NSm in M genome; (2) the length of nucleic acid sequence and protein sequence was different; (3) GC content of nucleic acid of animal viruses was higher than plant viruses of that; (4) protein topology analysis by online software SMART discovered that there was significant difference for the structure of glycoproteins (G_nG_c) between animal-infecting and plant-infecting viruses; (5) usually, protein of plant-infecting virus had more complicated structure, more low compositional complexity and had N-Signal peptides except INSV. The study would be helpful for designing primer to detect viruses in Bunyaviridae.

收稿日期: 2007-12-18

* 基金项目: 云南省自然科学基金面上项目 (2005C0036M; 2007C0036M); 云南省教育厅自然科学基金项目 (5Y0171B)

** 通讯作者 E-mail: lizhengyue@vip.km169.net

作者简介: 刘雅婷 (1971-), 女, 湖南祁东人, 硕士, 副教授, 主要从事植物病原微生物与寄主互作研究。
E-mail: liuyating999@yahoo.com.cn

Key words: Bunyaviridae; Tospovirus; DNAMAN; DNASTAR; SMART; alignment

布尼亚科病毒 (Bunyaviridae) 具有 300 多个病毒成员, 分为 5 个属, 包括布尼亚病毒属 (*Bunyavirus*)、汉坦病毒属 (*Hantavirus*)、内罗华病毒属 (*Nairovirus*)、白蛉热病毒属 (*Phlebovirus*) 和番茄斑萎病毒属 (*Tospovirus*)^[1]。布尼亚科病毒粒子通常为具有包膜的球形, 直径约 80 ~ 120 nm, 表面具有长 5 ~ 10 nm 的糖蛋白突起, 这些突起被包埋在厚度约为 5 nm 的双层脂质膜中。该科病毒的所有成员为三分基因组, 单个病毒粒子包裹有大 (L)、中 (M)、小 (S) 3 种负义或双义线形 ssRNA, 编码 4 种结构蛋白, 包括两种外源糖蛋白 (G_nG_c)、一种外壳蛋白 (N) 和一种超大的多聚酶 (L)。每个基因组 RNA 片段的末端核苷酸碱基配对, 形成非共价闭环 RNA。尽管布尼亚科病毒 5 个成员属之间在结构上有很多相似的地方, 但在它们之间仍存在很多差异: 该科 4 个属主要感染脊椎动物和节肢动物, 通常对脊椎动物寄主是溶细胞性的, 而对无脊椎动物仅引起不显著的细胞病理变化; 番茄斑萎病毒属是布尼亚科病毒中唯一能感染植物的一个属, 它通过蓟马的繁殖传播, 其寄主范围很广, 可感染 80 个科, 1 000 多种植物, 引起植物广泛褪绿、坏死、矮化和耳突等症状^[2]。番茄斑萎病毒 (TSWV) 是番茄斑萎病毒属的代表成员, 最早由 Samule, Bald, Pittman 于 1930 年记述, 该病毒在多种粮经作物上导致严重病害, 是重要的农作物病毒病原之一, 在国际上深受重视^[3]。

近年来, 随着生物技术和信息技术的发展, 大量的基因组序列被测定并公开。布尼亚病毒的动物病毒属是导致人和动物脑膜炎或出血热等疾病的重要病原体^[4]; 番茄斑萎病毒属在世界各地广泛分布, 引起多种栽培作物和观赏植物病害, 造成了巨大的经济损失^[5]。近年来其病毒成员被测序, 基因组序列被存储在基因库中。以此同时, 出现了一些非常有用的生物学软件, DNAMAN 是高度集成化的分子生物学应用软件, 可用于多重序列比对、蛋白分析, 几乎囊括了所有日常核酸蛋白质序列的分析工作。DNASTAR 是序列格式转换、基因寻找、序列比较的常用工具。本文利用 DNASTAR, DNAMAN, SMART 生物信息学软件对从 NCBI 网站上下载的布尼亚科病毒具有基因组全序

列的 20 个病毒的序列进行比对分析, 找出布尼亚科病毒中番茄斑萎病毒属与其它 4 个属病毒之间核酸序列及蛋白质结构上的差异, 为以后研究更多关于布尼亚病毒科病毒提供基本信息。

1 材料与方法

1.1 材料

本文从 NCBI (National Center for Biotechnology Information <http://www.ncbi.nlm.nih.gov/>) 数据库下载布尼亚科病毒的 63 个基因组全序列, 经 DNASTAR 进行一致性 (Identity) 和同源关系 (Homology) 比较分析, HTNVZ37 和 HTNVZ10 具有很高的 consistency, 很可能是一个病毒, 因此仅保留了 HTNVZ10, 整理后得到 5 个属 20 种病毒, 共 60 个基因组全序列 (见表 1)。

1.2 分析方法

将从 NCBI 数据库中下载的病毒基因组序列经人工分析整理归类得到 5 个属 20 种病毒 (见表 1), 用 DNASTAR 对不同病毒核酸序列的 GC 含量进行比较, 并对 20 种病毒核酸和蛋白质序列比对, 获得系统进化树; 用软件 DNAMAN 作核酸和蛋白质序列的比对, 比较了不同病毒之间的一致性和相同性; 用网络软件 SMART (<http://smart.embl-heidelberg.de/>) 分析 20 种病毒的 L 蛋白、糖蛋白和核壳体蛋白的拓扑结构。根据软件分析的结果, 结合人工分析比较布尼亚科病毒中番茄斑萎病毒属与其它动物病毒属之间的异同。

2 结果与分析

2.1 蛋白序列的长度

用 DNASTAR 对下载的基因组序列进行格式转换后, 计算病毒蛋白序列的长度。结果表明同一个属的病毒的基因组长度差异不大, 其中以汉坦病毒属 (*Hantavirus*) 的 L 基因组尤为显著: 10 个病毒 L 蛋白序列长度相差不超过 5 bp。 G_nG_c 蛋白除了布尼亚病毒属 (*Orthobunyavirus*) (1420 ~ 1441 bp) 和内罗华病毒属 (*Nairovirus*) (1551 ~ 1684 bp) 稍大, 其它的病毒都很接近。该科 5 个属的 20 种病毒都编码 L, G_nG_c , NC 蛋白, 但只有植物病毒番茄斑萎病毒属 (*Tospovirus*) 成员的 M 基因组能编码 Nsm 蛋白, 而其它动物病毒属的

M 基因组中则没有 NSm 蛋白; 番茄斑萎病毒属 (*Tospovirus*)、布尼亚病毒属 (*Orthobunyavirus*) 和白蛉热病毒属 (*Phlebovirus*) 的 S 基因组能编码 NSs 蛋白, 汉坦病毒属 (*Hantavirus*) 和内罗华病毒属 (*Nairovirus*) 两个属的病毒则不能, 但是它们的 NC 蛋白序列比前三者长。内罗华病毒属 (*Nairovirus*) L 基因组所编码的 L 蛋白特别长

(3945 ~ 4036 bp), 而番茄斑萎病毒属 (*Tospovirus*) L 基因组所编码的蛋白长度次之 (2865 ~ 2878 bp), 且 4 种病毒所编码的蛋白序列长度比较接近; 与布尼亚病毒属 (*Orthobunyaviru*)、汉坦病毒属 (*Hantavirus*)、白蛉热病毒属 (*Phlebovirus*) 相比则相对较长。

表 1 文中所涉及的布尼亚病毒科全基因组的主要特征

Tab. 1 Features of complete genome sequences of *Bunyaviruses* used in the study

属 genera	病毒名称 name of virus	缩写 Abr.	RNA	GC 含量 GC / %	序列号 accession number
<i>Orthobunyavirus</i> 布尼亚病毒属	<i>LaCrosse virus</i>	LACV	L	34	AF528165
			M	38	AF528166
			S	41	AF528167
	<i>Oropouche virus</i>	OROV	L	34	AF484424
			M	35	AF441119
			S	46	AY237111
	<i>Bunyamvera virus</i>	BUNV	L	33	X14383
			M	36	M11852
			S	41	D00353
<i>Hantavirus</i> 汉坦病毒属	<i>Dobrava-Belgrade virus</i>	DOBV	L	37	AJ410617
			M	40	AJ410616
			S	43	AJ410615
	<i>Hantaan virus</i>	HTNV	L	36	X55901
			M	39	M14627
			S	42	M14626
	<i>Seoul virus</i>	SEOV	L	37	X56492
			M	40	S47716
			S	42	AY273791
	<i>Hantavirus Z10</i>	HTNVZ10	L	37	AF189155
			M	40	AF276987
			S	43	AF184987
	<i>Sin Nombre virus</i>	SINV	L	37	L37901
			M	39	L25783
			S	38	L25784
<i>Tula virus</i>	TULV	L	36	AJ005637	
		M	38	Z69993	
		S	41	Z69991	
<i>Puumala virus</i>	PUUV	L	35	Z66548	
		M	38	X61034	
		S	38	X61035	
<i>Andes virus</i>	ANDV	L	36	AF291704	
		M	40	AF291703	
		S	41	AF291702	

续表

属 genera	病毒名称 name of virus	缩写 Abr.	RNA	GC 含量 GC /%	序列号 accession number		
<i>Nairovirus</i>	<i>Crimean-Congo</i>	CHFV	L	41	AY389361		
内罗华病毒属	<i>hemorrhagic fever virus</i>		M	43	AF467768		
			S	45	U88410		
		<i>Dugbe virus</i>	DUGV	L	39	U15018	
			M	41	M94133		
			S	43	AF434161		
<i>Phlebovirus</i>	<i>Sandfly fever</i>	SFNV	L	43	X68414		
白蛉热病毒属	<i>Naples virus</i>		M	44	X89628		
			S	47	X53794		
		<i>Uukuniemi virus</i>	UUKV	L	46	D10759	
			M	47	M17417		
			S	50	M33551		
		<i>Rift Valley fever virus</i>	RVFV	L	44	X56464	
			M	45	M25276		
			S	49	X53771		
		<i>Tospovirus</i>	<i>Impatiens necrotic</i>	INSV	L	32	X93218
		番茄斑萎病毒属	<i>spot virus</i>		M	33	M74904
	S			32	X66972		
<i>Watermelon silver</i>	WSMoV			L	34	AF133128	
<i>mottle virus</i>				M	34	U75379	
	S			32	U78734		
<i>Tomato spotted</i>	TSWV			L	33	D10066	
<i>wilt virus</i>				M	35	S48091	
	S			35	D00645		
<i>Groundnut bud</i>	GBNV			L	33	AF025538	
<i>necrosis virus</i>				M	34	U42555	
	S	34	U27809				

2.2 布尼亚科 20 种病毒 GC 含量

用软件 DNASTar 对不同病毒核酸序列的 GC 含量进行分析发现布尼亚病毒科 20 种病毒中白蛉热病毒属 (*Phlebovirus*) 病毒核酸的 GC 含量比其他 4 个属稍高 (最低 43.88%), 且每个基因组不同病毒之间含量差异较大。植物病毒 *Tospovirus* 属 GC 含量略比其他 4 个动物病毒属略低 (最高 40.15%), 含量也相对稳定, 不同病毒的同个基因组相差不超过 3 个百分点。

2.3 病毒的 L, NC 蛋白相似性

通过 DNAMAN 对布尼亚科 20 种病毒共有的 L 基因组及编码 NC 蛋白的序列进行比对分析, 得到 L 相似矩阵 NC 矩阵。病毒属内的相似性明显要高于属间的相似性; 汉坦病毒属 (*Hantavirus*)

的相似程度明显高于其它属: L 基因组 97.3% ~ 86.3%, 这与分类学相吻合。NC 相似程度 97.2% ~ 62.4%; 从整体上看 L 相似程度要大于 NC。除白蛉热病毒属 (*Phlebovirus*) 外, 动物病毒属内 L 蛋白的相似性高于植物病毒。植物病毒的 NC 蛋白除了 GBNV 与 WSMoV 的相似性到 80.5% 外其它都很低 (30.5% ~ 55.8%)。

2.4 糖蛋白拓扑结构预测

本研究通过网络软件 SMART 对 20 种病毒的糖蛋白 (G_nG_c) 预测分析, 发现同一个属的病毒糖蛋白的拓扑结构具有一定相似性, 但不同的病毒间又存在着差异。不同属的病毒的糖蛋白在拓扑结构上存在着差异, 植物病毒与动物病毒之间的差异尤为明显。布尼亚病毒属 (*Orthobunyavir-*

us) 病毒的糖蛋白具有较多的跨膜结构, 汉坦病毒属 (*Hantavirus*) 的糖蛋白拓扑结构相对简单, SINV, TULV, PUUV 都只有一个跨膜结构。植物病毒——*Tospovirus* 属 4 个病毒的结构较为复杂有较多的紊乱区域, 在 COOH - 端的相同位置具有跨膜结构, INSV 除外, WSMoV, TSWV, GBNV 的 NH - 端都具有短信号肽, 从结构上来看 TSWV 与 INSV, GBNV 与 WSMoV 非常接近, 这与病毒

系统进化树的分析结果相吻合。而动物病毒中只有 BUNV, HTNVZ10, CHFV, DUGV, SFNV 含有信号肽。

2.5 基因组 Identity, Homology 的比较

利用生物学软件 DNAMAN 对 20 种病毒的蛋白质和核酸序列进行比对分析, 除 NSm 外, 无论是 Identity 值, 还是 Homology 值都很低 (见表 2)。

表 2 利用 DNAMAN 分析获得 Bunyaviridae 科病毒的核酸和蛋白质序列比对

Tab. 2 Alignment of nucleic acid and protein sequence by the DNAMAN

		RNA - L		RNA - M	RNA - S	
		L	NSm	G ₁ G ₂	NC	NSs
蛋白质 protein	Identity	25.03%	65.26%	25.45%	34.70%	20.79%
	Homology	3.00%	43.00%	9.0%	10.00%	11.00%
核酸 nucleic acid	Identity	33.15%	69.92%	37.79%	39.66%	35.55%
	Homology	30.00%	53.00%	30.00%	28.00%	29.00%

2.6 Bunyaviridae 科属中具有全序列的病毒的核酸保守区域

通过生物信息学软件 DNAMAN 对布尼亚整个科的病毒基因组进行分析, 从整个科来看, 核酸的保守序列很少; 针对 NSm 来说, 则具有较高的保守性。

2.7 DNASTAR 的系统进化分析

利用生物信息学软件 DNASTAR 对 L, G_nG_c, NSm, NC, NSs 蛋白进行比对分析, 做出系统进化树。以 RNA L 蛋白序列 (合成依赖于 RNA 的 RNA 复制酶, RdRp) 为基础得出的系统进化树 (见图 1), 在进化值接近 500 时分为 5 个族, 第 1 个族包括 HTNV, HTNVZ10, SEOV, DOBV, PUUV, TULV, SINV, ANDV; 第 2 个族包括 LACV, BUNV, OROV; 第 3 个族包括 UUKV, RVFV, SFNV; 第 4 个族包括 INSV, TSWV; 第 5 个族包括 CHFV, DUGV。这与国际病毒分类委员会 2003 年发布的 Bunyaviridae 病毒分类有一定的差异, *Tospovirus* 属病毒在进化树值 749.9 时才归到一个族。以 NC 蛋白序列为基础得出的系统进化树, 与 NSs 的略有不同, 在进化树值 400 ~ 500 之间时分成 4 个族, 即 *Hantavirus*, *Orthobunyavirus* 一个族, CHFV, DUGV, SFNV 一个族, WSMoV, GBNV, INSV, TSWV, UUKV 一个族, RVFV 一个族。与国际病毒分类委员会分类系统比较, 进化值为 500 左右时 *Phlebovirus* (SFNV,

UUKV) 聚为一族, 而 *Phlebovirus* 属的病毒 (SFNV, UUKV, RVFV) 没有归到一个族。

3 讨论

无论是 L 相似矩阵还是 NC 矩阵中都可以看出, 属内病毒的相似性要大于属间, 这与系统进化树相吻合。尽管有学者认为 *Tospovirus* 是由 Bunyaviridae 科其它属由于长期进化而形成的依赖寄主植物和蓟马的一类病毒, 但是布尼亚科 20 种病毒的 Identity 和 Homology 值都很低, 表明该科病毒是进化较为迅速的。L, NSs 的序列差异较大。

布尼亚科病毒 5 个属中, 只有植物病毒 *Tospovirus* 属的基因组 M 编码 NSm 蛋白。这可能是 *Tospovirus* 属病毒能够适应植物寄主的一种表现^[6]。NSm 在 *Tospovirus* 病毒的传播中扮演着一个重要角色, 这种蛋白有修饰限制胞间连丝的能力^[7], 它可能是通过管状结构穿过胞间连丝来实现病毒核蛋白 (RNP) 的细胞间运动, NSm 还可能与植物寄主的范围及发病症状有着联系^[6]。在病害诊断中, 无论是从寄主植物组织中, 还是从媒介昆虫体内检测该类病毒时, 利用该属特有的 NSm 蛋白编码区来设计引物, 是比较理想的。

布尼亚病毒属、汉坦病毒属、番茄斑萎病毒属中都有 NSs 蛋白, 番茄斑萎病毒属中的 NSs 序列比较长, 汉坦病毒属次之、布尼亚病毒属较小。但 GC 含量则不同, 番茄斑萎病毒属稍低, 不到

38%；而两个动物病毒属均超过 44%。不同属的 NSs 蛋白在大小和氨基酸序列上都有很大变化，编码机制也可能不同^[7]。NSs 对病毒的转录和复制是非必要的，它是一种附属蛋白，作为病毒性因子，抑制寄主细胞的蛋白质合成及 IFN（干扰因子）的功能^[4]，增强病毒的致病性。

研究认为糖蛋白是决定昆虫媒介能传播 *Tospo-*

virus 属病毒的重要因素，它能与昆虫中肠细胞表面受体相互作用，从而促进蓟马对病毒粒子的吸收^[8]，同时与病毒在媒介体内的吸收及稳定性有关，但对植物系统的感染却并非是必须的^[9]。*Tospovirus* 属病毒除了 INSV，载糖蛋白的 NH-端都含有信号肽分子，这对寄主细胞的识别是重要的。

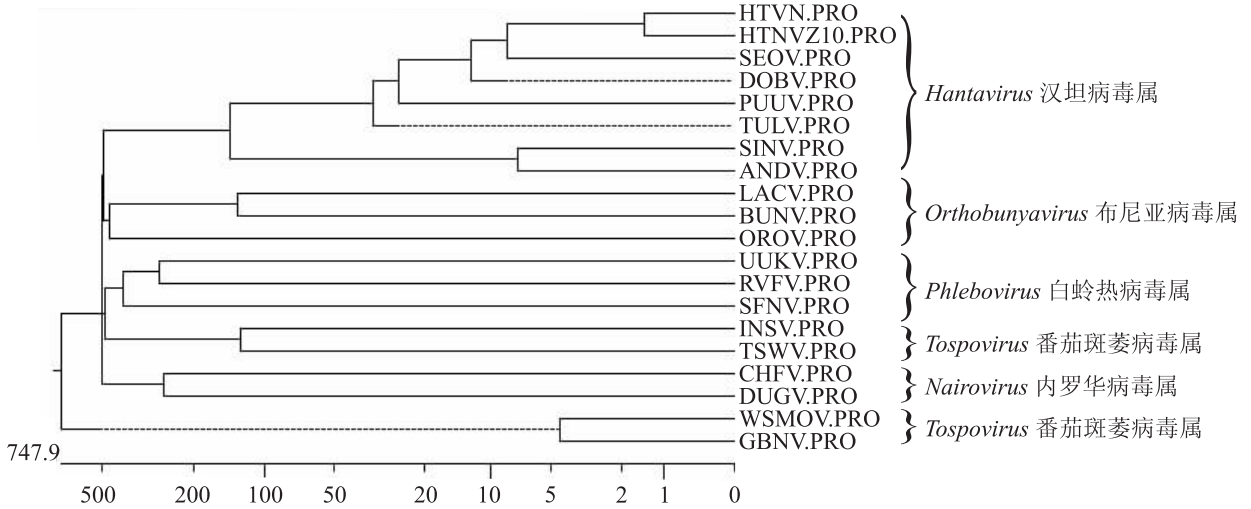


图 1 依赖于 RNA 的 RNA 复制酶 (RdRp) 系统进化树

Fig. 1 Phylogenetic tree of RNA-dependent RNA polymerase (RdRp)

[参考文献]

- [1] 徐耀先, 周晓峰, 刘立德. 分子病毒 [M]. 武汉: 湖北科学技术出版社, 2000.
- [2] M T MOMOL, J E FUNDERBURK, S OLSON. Management of TSWV on tomato with UV - reflective mulch and acibenzolar - S - methyl [C]. in Proceedings of the 7th International Symposium on Thysanoptera, Australian National Insect Collection, Canberra (eds Marullo R, Mound LA). 2002; 111 - 116
- [3] 张仲凯, 李毅. 云南植物病毒 [M]. 北京: 科学出版社, 2001.
- [4] ALAIN KOHL, REGINALD F CLAYTON, FRIEDEMANA WEBER, et al. Bunyamwera Virus Nonstructural Protein NSs Counteracts Interferon Regulatory Factor 3 - Mediated Induction of Early Cell Death [J]. Journal of Virology, 2003, 77 (14): 7999 - 8008.
- [5] PETER D, PETERS D, GOLDBACH R. (eds.). Recent Progress in Tospovirus and Thrips Research [M]. Wageningen, The Netherlands, 1998.
- [6] I. CORTEZ, A. AIRES, A. -M. PEREIRA. Genetic organization of iris yellow spot virus M RNA: indications for functional homology between the Gc glycoproteins of tospoviruses and animal-infecting bunyaviruses [J]. Arch Virol, 2002, 147: 2320
- [7] ANNE BRIDGEN, FRIEDEMAN WEBER, JOHN K. Fazakerley, Bunyamwera bunyavirus nonstructural product that contributes to viral pathogenesis [J]. PNSA, 2001, 98: 668.
- [8] I. CROTEZ, A. AIRES, A. -M. PEREIRA. Genetic organization of iris yellow spot virus M RNA: indications for functional homology between the Gc glycoproteins of tospoviruses and animal-infecting bunyaviruses [J]. Arch Virol, 2002, 147: 2321
- [9] GERNANDA ANTINOLFI LOVATO, TATSUYA NAGATA, RENATO DEOLIVEIRA RESENDE. Sequence Analysis of the Glycoproteins of tomato Chlorotic Spot Virus and Groundnut Ringspot virus and Comparison with other Tospoviruses [J]. Virus Genes, 2004, 29: 322.