植物单链RNA病毒中的正链病毒的基因结构和表达调控

陆惠华　　龚祖埙

（中国科学院上海生物化学研究所，上海）

PLANT POSITIVE-STRANDED RNA VIRUSES: STRUCTURE, EXPRESSION AND REGULATION OF THE VIRAL GENES.

Lu Hui-hua　　Gong Zu-xun

(*Shanghai Institute of Biochemistry, Academia Sinica. Shanghai*)

植物病毒可以按其遗传物质的不同分为四大类[1]，即双链DNA病毒、单链DNA病毒、双链RNA病毒和单链RNA病毒。其中单链RNA病毒又可以按基因RNA复制和表达方式的不同而分为正链RNA病毒和负链RNA病毒两种，大约70%的植物病毒属于正链RNA病毒。越来越多的实验证据表明，植物病毒，尤其是植物正链RNA病毒，具有某些不同于其它生命形式的特性，这些特性主要表现在基因结构、表达和调节等方面。本文就近年来植物正链RNA病毒基因结构和表达调控方面的研究进展作一个绍介。

一、植物正链RNA病毒的基因结构

1. 基因在RNA上的排列方式

病毒基因组RNA上各个基因之间的排列方式如何？这种有机的排列又怎样协调了各个基因的表达？这是研究基因结构的中心问题。

有些植物正链RNA病毒的基因是一条单链RNA。例如烟草花叶病毒(TMV)、芜菁黄花叶病毒(TYMV)、马铃薯X病毒(PVX)等等，都是只含一条单链RNA的病毒，通常把这类病毒称为单分体病毒。还有些植物正链RNA病毒的基因是由几条不同的RNA组成，如烟草脆裂病毒(TRV)、雀麦花叶病毒(BMV)、黄瓜花叶病毒(CMV)等。TRV的基因为两条分子量不同的RNA片段，分别装配在两种长度不同的棒状病毒颗粒中[2]。BMV含有四种RNA分子，分别包装在三种外观相同密度不一的病毒颗粒中[3]。这类以分散的RNA片段作为基因组的病毒称为多分体病毒。据推测，多分体病毒是由单分体病毒演化而来的[4]。

一般认为，每一病毒基因组编码的蛋白质至少在三种以上[5]。所以，所有单分体病毒的基因 RNA 都是多顺反子，几个基因依次排列在一条 RNA 链上。多分体病毒的基因 RNA 有些是多顺反子，有些是单顺反子。但是有趣的是几乎所有植物正链 RNA病毒的基因 RNA 在功能上都只是单顺反子，就是说，不管 RNA 上排列有多少个顺反子，只有 5′ 端的那个顺反子能直接翻译蛋白质。其它顺反子则要通过另外的途径才能作为翻译的模板。这种独特的现象具有什么生物学意义？正是许多研究者共同关心的问题。

病毒 RNA 上基因的有机组织为基因的表达调控提供了重要的结构基础。

2. 5′端非编码区的结构

迄今的研究表明，所有植物正链 RNA 病毒的 RNA 的5′末端结构可归结为两类：一类是帽子结构（$m^7G_{5'ppp5'}Xp$）另一类是与 RNA 共价连接的蛋白质，该蛋白质系由病毒编码，称为 VPg（Viral protein genome-linked）。从5′末端起至翻译起始信号（通常是 AUG）之间的核苷酸顺序称为 5′ 端非编码顺序。通常认为 5′ 端的这些结构对病毒 RNA 的翻译表达具有调控意义。

（1）5′端帽子结构

几乎所有真核细胞 mRNA 5′ 端都是帽子结构。关于真核 mRNA 帽子结构和功能已经有较详尽的论述[6]。已知帽子结构至少有两种功能：一是促进 mRNA 与核糖体的识别，有利于翻译起始复合物的形成，二是防止核酸酶对 mRNA 的降解破坏。所以帽子结构的存在提高了 mRNA 的翻译能力。大多数植物正链 RNA 病毒的基因 RNA 5′ 端是帽子结构。但是一个明显的差别是，病毒 RNA 帽子后面紧接着的两个核苷不被甲基化[7]（图 1）。人为地将 TMV、BMV 的RNA 上的帽子除去则大大地影响了 RNA的体外翻译能力。同时病毒 RNA 的致病能力也消失。所以，一般认为病毒 RNA 帽子结构的功能与真核 mRNA 的相似。但是豇豆花叶病毒（CPMV）、烟草坏死病毒（TNV）、

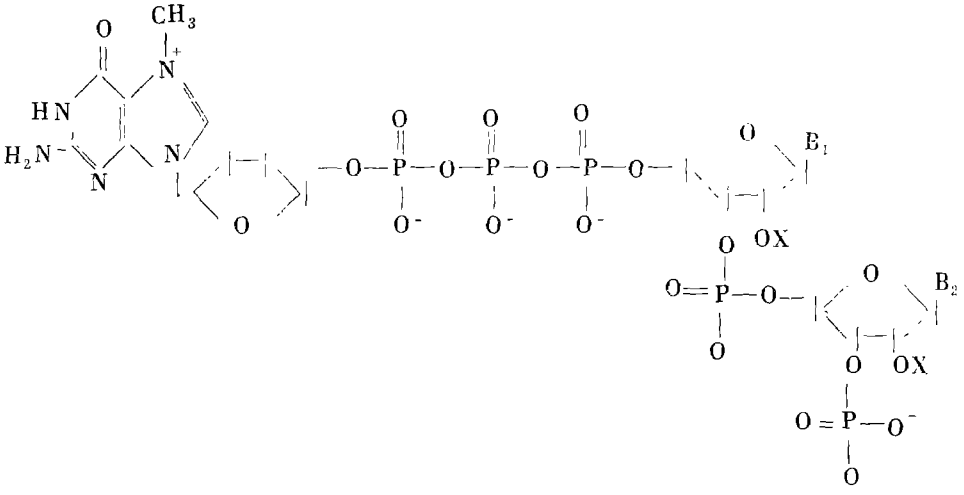


图1. RNA5′端帽子结构
真核细胞mRNA：X＝CH_3
病毒RNA： X＝H

Fig 1. Cap structure at the 5′-terminal of plus sense RNAs. Eukaryotic mRNA X＝CH_3; Viral RNA: X＝H

烟草坏死病毒卫星病毒（STNV）等RNA的5′端没有帽子结构，却也是蛋白质翻译的有效模板。人为地在STNV 5′端加上帽子也并不能提高它的翻译活力[6]。此外，真核mRNA和病毒RNA的帽子结构甲基化的差别说明什么意义，这些问题都不很清楚。

（2）VPg结构

VPg是一些病毒RNA 5′端独有的结构[8]。从1977年起陆续报道在某些DNA病毒和RNA病毒核酸的5′端发现这种与核酸共价连接的蛋白质分子。VPg的分子量都较小，脊髓灰质炎病毒、乙型肝炎病毒、腺病毒等动物病毒具有VPg结构，植物病毒目前只在一些正链RNA病毒发现有VPg存在，这些病毒分布于马铃薯Y病毒组、蠕传多角病毒组、黄症病毒组和豇豆花叶病毒组，另外，南方菜豆花叶病毒（SBMV）RNA的5′端也找到VPg。根据一些直接的和间接的实验证据，人们推测VPg可能与病毒RNA的复制有关。

（3）5′端非编码顺序

RNA 5′端非编码顺序的研究是最令人感兴趣的问题之一。由于核酸一级结构测定技术的改进，有关植物病毒RNA顺序的报道逐渐增多，为5′端非编码顺序的比较研究提供了方便。

病毒RNA的5′端非编码顺序长度不一，一般为10—100个核苷酸。编码病毒外壳蛋白的亚基因RNA上这段顺序都很短。帽子结构后往往紧跟着GUA或GUU。在体外或体内蛋白质翻译研究中，TMV、CMV、BMV，苜蓿花叶病毒（AlMV）等许多植物病毒RNA都表现出一个共同的现象，即每一病毒中翻译效率最高的大都是那些合成病毒外壳蛋白的RNA[4]。图二所示的就是几种高效表达RNA的5′端非编码顺序。可以看到这类高效表达RNA具有如下特点：一是5′端非编码区很短；二是5′端都是帽子结构；再就是5′端非编码区中A、U两种碱基出现的频率非常高，这种特征顺序在真核生物mRNA的5′端非编码区很少见[9]。有人将植物病毒RNA5′端的这种富含A、U顺序称为Butler Boxes[10]。有些Butler Boxes还能够自身配对折叠成紧密的发夹结构。这种特殊的顺序和结构与RNA的高效表达有何联系，它们是不是核糖体或者其他调控因子识别的信号？深入研究这类高效表达RNA的5′端结构，对于植物基因工程无疑是很有意义的。

BMV RNA 4：[11] m GpppGUAUUAAUAAUAAUG·····
TMV CP mRNA：[5] m GpppGUUUUAAAUAUG·····
CCMV RNA 4：[11] m GpppGUAAUUUAUCAUG····
ALMV RNA 4：[5] m GpppGUUUUUAUUUUUAAUUUUCUUUCAAAUACUU
CCAUCAUG·····
CMV RNA 4：[12] m GpppGUUUAGUUGUUGUUCACCUGAGUCGUGUUUUCUUUGUUUU
GCGUCUCAGUGUGCCUAUG·····

图2. 一些高效表达植物病毒RNA的5′端非编码顺序（Butler Boxes）

Fig. 2. The 5′-terminal non-coding sequences of some viral RNAs having efficient translating capacity.

3. 3′端非编码顺序

就目前的研究来看，植物正链RNA病毒的3′端有三种结构，即tRNA状结构、poly（A）结构以及无规则顺序。

（1）3′端tRNA状结构

许多植物正链RNA病毒的RNA 3′端能在体外和体内专一地被氨酰tRNA合成酶识别，结合专一的氨基酸（表1）。人们很自然地提出问题：这些病毒的RNA 3′端是否与tRNA一样，具有特殊的三叶草结构？目前BMV、TYMV等病毒组的RNA 3′端核

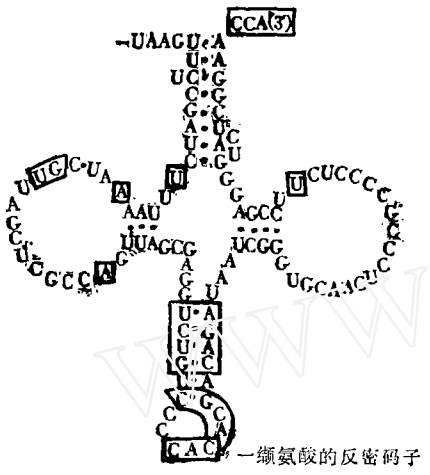


图3. TYMV RNA3′端的tRNA状结构
Fig. 3. tRNA-like structure at the 3′-terminal of TYMV RNA

表1. tRNA状结构的氨酰化[13]
Table 1. Aminoacylation of tRNA-like structures of some plant viral RNAs.

病毒RNA	专一的氨基酸
芜菁黄花叶病毒RNA	缬氨酸
烟草花叶病毒RNA(普通株)	组氨酸
烟草花叶病毒RNA(豇豆株)	缬氨酸
黄瓜花叶病毒RNA	酪氨酸
雀麦花叶病毒RNA	酪氨酸
豇豆褪绿斑驳病毒RNA	酪氨酸
蚕豆斑驳病毒RNA	酪氨酸

苷酸顺序已经测定，发现这些病毒RNA 3′端的确能自身折叠形成类似于tRNA的三叶草结构，虽然类似的程度略有不同，但是这些RNA 3′端结尾的三个核苷酸都是CCA。图三所示的是TYMV RNA 3′端108个核苷酸形成的三叶草结构。

tRNA状结构有些什么功能？对此主要有下列三种假设[13]：（i）作为氨基酸供体参与蛋白质合成；（ii）作为病毒RNA复制的调节因素；（iii）通过与氨酰tRNA合成酶的作用增强病毒RNA的翻译能力。

因为tRNA状结构能与氨酰tRNA合成酶、核苷酸/tRNA转核苷酸酶、EF—Tu等蛋白质合成有关因子相互作用，并且具有专一的氨酰化反应[14]，有人推测它在细胞内可能充当氨基酸供体。但是体外实验中病毒RNA的tRNA状结构并不能象tRNA那样作为氨基酸供体参与蛋白质合成。

延长因子EF-Tu和EF-TS是Qβ噬菌体RNA复制酶的亚基[15]，于是有人根据tRNA状结构与EF-Tu的关系，推测tRNA状结构与EF-Tu、EF-1的结合可能激活了RNA复制酶。

C. Florentz等分析了TYMV、BMV、TMV和CMV等RNA的高级结构，认为这些RNA的5′端顺序和3′端tRNA状结构相互配对成"封闭状"结构[16]。在"封闭状"RNA中5′端翻译起始部位和起始密码都不暴露在外，当3′端tRNA状结构结合氨酰

tRNA 合成酶之后，封闭状态破坏，5′端顺序暴露，所以 tRNA 状结构可能具有调节病毒 RNA 翻译活力的功能。

总之，3′端 tRNA 状结构的功能有待进一步研究。但是不难推测，这种经过长期进化保存下来的特殊结构很可能对病毒的生存具有重要意义。

（2）3′端的 poly（A）

虽然poly（A）是真核细胞 mRNA 3′端的特征性结构，但这样的结构在植物病毒 RNA 中并不多见。仅马铃薯Y病毒组、蟠传多角体病毒组和豇豆花叶病毒组的成员在基因 RNA3′端有poly（A）顺序。各种病毒 RNA 的poly（A）长度不一，通常为15至200个核苷酸。在某些病毒 RNA 上，顺反子之间也存在长度不同的寡聚腺苷酸顺序。已知真核细胞 mRNA 3′端 polyA 的功能主要是维持 mRNA 的稳定性[17]。推测病毒 RNA的poly(A)也有类似的作用。令人感兴趣的是，植物病毒中尚未找到同时具有帽子结构和poly(A)顺序的 RNA 分子，而这两个结构几乎是所有真核细胞 mRNA 都具备的共同特征！另外，具有3′端poly (A)的病毒 RNA，其5′端都是 VPg 结构。这究竟是偶然的巧合，抑或有某种内在的关联？植物病毒 RNA 的3′端 poly(A)到底有哪些功能，目前还不十分清楚。

（3）3′端无规则顺序

植物正链 RNA 病毒中还有一类成员，诸如 PVX、AlMV、TRV 和烟草条斑病毒(TSV)等等，基因 RNA 的3′端是无规则排列的核苷酸链，一级结构和高级结构都没有规律性。这些病毒不但 RNA 3′端呈无规则状态，而且在基因结构、病毒形态上也无规律可循。就基因结构而言，PVX 是单分体病毒，TRV 是两分体病毒、AlMV是多分体病毒；就病毒形态而言，PVX、AlMV 和TSV 分别为线状、杆状和球状病毒。但是这类病毒有一个重要性质，当外壳蛋白不存在时，病毒 RNA 便不能复制。外壳蛋白能与 RNA 的某一区域专一识别，激活 RNA 的复制[18][19]。

二、 植物正链RNA病毒的基因表达和调控

基因的表达和调控是分子生物学研究的核心问题之一。植物正链 RNA 病毒的基因表达和调控采用了一些独特的方式。

1.亚基因化：

亚基因化是植物正链 RNA 病毒基因表达的最常见模式。已知大多数病毒基因 RNA 在功能上是单顺反子，只有5′端的那个顺反子能够直接表达。那么 RNA 上的其他顺反子又如何表达呢？亚基因化就是病毒表达非 5′端顺反子的重要途径。

亚基因是在病毒感染细胞后复制产生的一系列病毒 mRNA，它们的分子量均小于基因 RNA，亚基因的 5′端分别是那些在基因 RNA 中关闭的非5′端顺反子。图 3 是几种植物病毒基因表达的简图。可以看到，病毒通过亚基因化表达了携带的所有遗传信息。一些亚基因 RNA，比如 BMV 的RNA_4、CMV的RNA_4等等，可以被包入成熟的病毒质粒；也有些亚基因 RNA 不出现在成熟病毒中，例如 TMV 的外壳蛋白亚基因等等。这主要取决于亚基因 RNA 上是否存在病毒装配时为外壳蛋白识别的顺序。

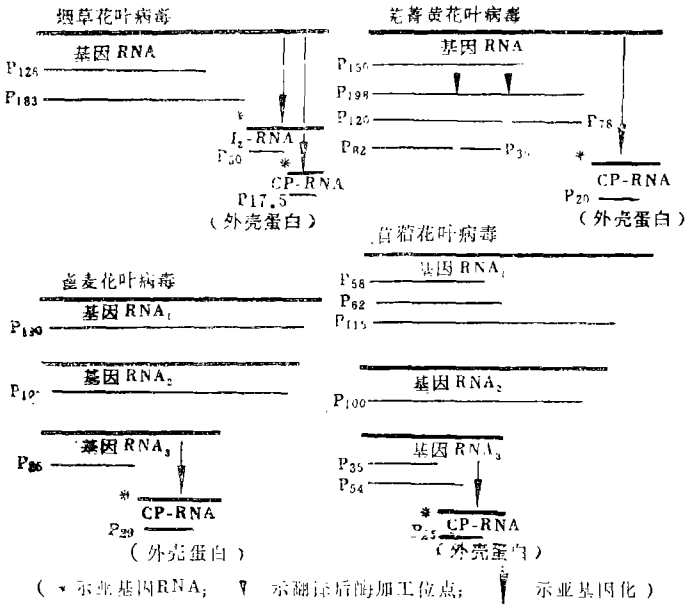
亚基因化的分子机制是什么？对此曾有过三种假设：（i）从双链RNA复制型中的负链RNA的某些特定起始点局部转录产生亚基因。（ii）基因RNA转录负链时不成熟中止，然后这些不成熟的负链RNA又作为模板复制出亚基因RNA。（iii）全长的基因RNA经专一的RNA酶酶切形成亚基因RNA。

W.A.Miller 等通过分子克隆制备了大量的BMV 负链RNA_3作为研究亚基因化机制的对象。他们的实验证据证明，BMV RNA_3的负链内存在一个特定的序列，体外实验中复制酶能专一地识别这个序列，局部转录出BMV的亚基因RNA——RNA_4[20]。

亚基因化的生物学意义在于调控病毒基因的表达，这种调控体现在复制和翻译两个层次。由于各个RNA复制起点对复制酶的亲和力不同，基因RNA和各种亚基因 RNA 复制的拷贝数是不同的；又由于各RNA的5′端对核糖体的识别能力不同而具有不同的翻译活力。这样，同一病毒基因RNA上的几个顺反子表达的量可以十分悬殊。而悬殊的表达对于病毒的复制是十分必要的。就TMV而言，复制一个TMV RNA分子可能仅需一个复制酶分子，而将这条RNA装配成病毒颗粒却需要2130个外壳蛋白分子，可见不同病毒蛋白的需求量差异很大。

亚基因化使得所有顺反子的核糖体结合部位、翻译起始密码都暴露在外，而且大多数都加上了帽子结构，因此亚基因化提高了每一顺反子的翻译效率。

另外，亚基因化也是植物病毒与寄主细胞长期相互适应、选择的结果。原核生物的mRNA是多顺反子结构，其翻译的调控相对比较简单，随着生物进化，生命体的翻译表达系统也逐步完善，真核细胞的mRNA结构也相应发生变化。已知高等植物的mRNA都是单顺反子，可以想象植物细胞的翻译表达系统与单顺反子 mRNA 是高度契合的。植



图四 几种植物正链RNA的基因表达简图

Fig. 4. Strategies used in the gene expression of some plant positive-stranded RNA viruses.

物病毒为了利用寄主细胞的翻译表达系统，必须通过某种机制使本身的多顺反子RNA与之适应，这种选择、适应关系可能是亚基因化机制形成的进化动力。

2.翻译后酶加工

有些病毒RNA翻译时，最先合成的是一个分子量很大的蛋白质前体，然后由专一的蛋白酶加工裂解，产生各种成熟的病毒蛋白质，这就是翻译后酶加工现象。目前已知，豇豆花叶病毒组[21]、蠕传多角体病毒组[22]、芜菁黄花叶病毒[23]等都有翻译后的酶加工。如图五所示CPMVRNA$_1$最终表达产物多达十种，动力学实验说明这十种蛋白质都是从一个分子量为200k的前体产生的[24]。200k前体中包含了两个蛋白水解酶，分子量分别为24k和32k。24k蛋白酶能使200k前体自身水解，32k蛋白酶则专一地对RNA$_2$翻译的蛋白质前体进行酶加工[25]。

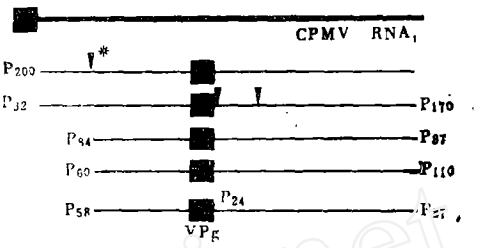


图5. 豇豆花叶病毒RNA$_1$的翻译后酶加工
Fig. 5. Post-translational cleavage of the gene product of CPMV RNA1.

如果说通过亚基因化病毒实现了由多顺反子向单顺反子的转化，那么相反，翻译后酶加工却使一个单顺反子RNA最终给出一系列成熟的、功能各异的病毒蛋白。如果不考虑病毒蛋白在细胞中的稳定性，那么从同一前体产生的各个蛋白质，无论它们具有怎样的功能，都应是等量的。从生物学角度来看似乎不够经济合理，所以翻译后酶加工很可能含有另一层意义。

除TYMV之外，采用翻译后酶加工的病毒RNA都有类似的结构，即5′端是VPg，3′端是poly(A)。这种结构特征和翻译后酶加工现象有无内在的联系，目前还难以作出推断。

3.超读翻译：

植物病毒RNA上的某些翻译终止密码比较弱，翻译时有部分新合成的肽链能通过这些终止密码而继续延伸至下一个终止密码，这就是所谓超读翻译。因为超读翻译是体外和体内实验中都存在的现象，而且对某一病毒来说超读的比例是相对稳定的。所以基本上可以排除人为因素导致超读翻译的可能性。超读翻译在植物正链RNA病毒中较常见，TMV、TYMV、TRV、AlMV、SBMV等等均有超读现象。例如TMV基因RNA翻译126k和183k两种蛋白质，183k是超读产物，比126k蛋白多了C端的一段肽链[26]。有人认为两种蛋白质可能具有不同的功能[27]。TYMV基因RNA首先翻译出150k、198k两种蛋白质，其中198k是超读产物，能够自身酶加工成120k、78k、82k和35k四种病毒蛋白。由于这些酶切点是150k、198k两种蛋白质共有的，而150k蛋白质并不被酶加工，可以推测，蛋白酶活性部位在198k蛋白的C端，而且该酶对底物可能有构象上的要求。TYMV RNA通过超读翻译结合翻译后酶加工，使一个阅读框架编码了6种蛋白质。

体外翻译体系中加入校正tRNA时超读比例显著增加，加入适量的多胺化合物超读也会有所上升[7]。烟叶、麦胚、麦叶中的tRNA^{Trp}、兔网状红细胞中的tRNA^{Trp}[28]都能超读TMV RNA上的第一个终止密码AUG。另外，当TMV RNA浓度较低时超读比

例有所增加[29]，提示在病毒感染早期超读翻译可能是重要的。

4.非成熟终止及其它

除了上述几种表达调控方式之外，一些植物病毒还采用其它的策略。TYMV、AlMV RNA 翻译时除预期的翻译产物之外还给出分子量小一些的产物，这些产物是核糖体复合物在病毒RNA的某些特定位点非成熟中止形成的。移码翻译目前只有AlMV RNA的翻译中观察到。非成熟中止和移码翻译在植物病毒中并不多见，这种现象的本质和机理尚不明瞭。

此外，还有几个植物正链RNA病毒的基因RNA以多顺反子形式翻译表达。例如香石竹斑驳病毒（CarMV）、烟草坏死病毒（TNV）等。

植物病毒能以其有限的遗传信息影响细胞乃至整个植株的代谢，必然要充分利用每一个RNA分子的编码潜力。正如在其他生物中发现的基因重叠现象一样，植物病毒RNA的超读翻译、非成熟中止以及移码突变能从一个阅读框架翻译几种不同的蛋白质，对于小小的病毒来说显然是很有必要、很有意义的。

表2总结了几种主要的植物正链RNA病毒的基因结构和表达调控方式。可以看到，不同组的病毒表达模式各不相同，而同一病毒又往往采用多种调控方式，精巧复杂的调控贯穿于病毒的整个复制周期。

表2. 一些植物正链RNA病毒的基因结构与表达调控[5,7]

Table 2. Gene structure, expression and regulation of some plant positive stranded RNA viruses.

病毒名称	病毒的基因结构			病毒基因的表达调控方式				
	基因的组织	5′端结构	3′端结构	亚基因化	超读翻译	非成熟中止	翻译后酶加工	移码翻译
烟草花叶病毒	单分体	帽子	tRNA状	+	+	/	/	/
马铃薯x病毒		帽子	无规则顺序	?	+	/	/	/
番茄蚀纹病毒		VPg	poly（A）	?	?	/	?	/
芜菁黄花叶病毒		帽子	tRNA状	+	+	/	+	/
菽麻花叶病毒		帽子	tRNA状	+	+	+	/	/
烟草脆裂病毒	两分体	帽子	无规则顺序	+	+	/	+	/
豇豆花叶病毒		VPg	poly（A）	/	/	/	/	/
番茄黑环斑病毒		VPg	poly(A)	/	/	/	+	/
大麦条斑花叶病毒	两/多分体	帽子	tRNA状	+	/	/	/	/
烟草条斑病毒	多分体	帽子	无规则顺序	+	/	/	/	/
黄瓜花叶病毒		帽子	tRNA状	+	/	/	/	/
雀麦花叶病毒		帽子	tRNA状	+	/	/	/	/
苜蓿花叶病毒		帽子	无规则顺序	+	/	+	/	?
豇豆褪绿斑驳病毒		帽子	tRNA状	+	/	/	/	/

参 考 文 献

[1] 龚祖埙，1978，《生物化学与生物物理学报》10(2) 83-102。
[2] Robinson, D. J., 1983, *J.Gen.Virol.*64(3):657-665.
[3] Lane, L.C. et al. 1971, *Nature New Biol.* 232:40-43.
[4] Atabekov, J.G. et al.1979, *Adv.Virus Res.*25:1-91.
[5] Joshi, S.et al., 1984, *FEBS Lett.*177(2):163-174.
[6] Filipowicz, W. 1978, *FEBS Lett.* 96(2):1-11.
[7] Davies, J.W. 1982, *J.Gen.Virol.* 61:1-14.
[8] Wimmer, E., 1982, *Cell* 28:199-201.
[9] Kozak, M., 1984, *Nucleic Acids Res.* 12(2):857-872.
]10] Goelet, P. et al., 1982, *Proc. Natl.Acad.Sci.* U.S.A. 79:5818-5822.
[11] Dasgupta, R. et al., 1982, *Nucleic Acids Res.* 10(2):703-713.
[12] Gould, A.L. et al., 1982, *Eur. J. Biochem.* 126(2):217-226.
[13] Haenni, A.L. et al., 1982, *Progr.Nucl. Acid Res. Mol. Biol.*27:85-103.
[14] Joshi, R.L.et al., 1984 *Nucleic Acids Res.*12(19):7467-7478.
[15] Blumanthal, T. et al., 1979, *Ann. Rev.Biochem.* 48:525-548.
[16] Florentz, C. et al., 1984, *FEBS Lett.* 176(2):295-300.
[17] Brawerman, G., 1981, *Biochem.J.* 197:1-6.
[18] Zuidema, D.et al. 1985, *Virol.* 140:342-350.
[19] Zuidma, D. et al., 1983, *Virol.* 129:255-260.
[20] Miller, W.A. et al., 1984, *Nature* 313:68-70.
[21] Pelham, H.R.B., 1979, *Virol.* 96:463-477.
[22] Bret, A.M.K. et al., 1983, *Virol.* 130:523-526.
[23] Dougherty, W.G.et al., 1980, *Virol.* 104:183-194.
[24] Franssen, H. et al.1984, *Virus Res.* 1:39-49.
[25] Franssen, H.et al., 1982, *J.Virol.* 41:8-17.
[26] Pelham, H.R.B., 1978, *Nature* 272:469-471.
[27] Van Telgen, H.J., 1985, *Virol.* 143:612-616.
[28] Abraham, A.K., 1983, *Progr.Nucl.Acid Res.Mol.Biol.* 28:82-100.
[29] Joshi, S. et al., 1983, *Virol.*127:100-111.

www.cnki.net