

## 黄病毒的分子生物学——进展与展望

陈火胜 郭辉玉

(中山医科大学微生物学与免疫学教研室, 广州)

### Molecular Biology of Flaviviridae——Progress and Perspective

Chen Huo-sheng & Guo Hui-yu

(Department of Microbiology and Immunology, Sun Yat-sen University of Medical Sciences, Guangzhou)

1984年国际病毒分类委员会(ICTV)建议将黄病毒属从披膜病毒科(Togaviridae)划出来成为新的病毒科——黄病毒科(Flaviviridae)。Westaway等<sup>[1]</sup>给黄病毒科作了如下的定义:由一组有包膜的单链RNA病毒组成,直径为40—50nm。包膜通常含糖蛋白, RNA有传染性, RNA的分子量约为 $4 \times 10^6$  Dalton, 3'末端缺乏多聚腺苷酸(PolyA)。体外实验中,病毒基因组RNA可翻译出结构蛋白,无亚基因组(Subgenomic)mRNA合成。成熟病毒颗粒的形态发生位于改变了的内浆网小池(cisternae)中。

黄病毒科目前仅由黄病毒属(Flavivirus)组成,黄病毒属大约包括70种病毒,其代表种是黄热病病毒(YFV)Asibi株<sup>[1]</sup>。按传播媒介,黄病毒可分为蜱传、蚊传及尚无已知传播媒介等三组,通过交叉中和试验可将大多数黄病毒划归七个抗原组<sup>[2,3]</sup>。

许多黄病毒是人类和动物的病原体。黄热病、登革热及多种脑炎都是由黄病毒引起的人类传染病,它们曾多次发生大流行,现在仍严重地危害人类的健康。绵羊跳跃病、Wesselsbron病和多种脑炎等则是人畜共患的黄病毒病<sup>[4,5]</sup>。黄病毒的特殊复制机理,一定程度上体现了生命过程的复杂性及多样性。因此,对黄病毒的分子生物学研究,不仅有助于诊断、防治由黄病毒感染所致的人类和动物疾患,还可能推进人类对这类RNA病毒的起源、进化、病毒与宿主细胞相互作用机理,甚至对生命本质的认识。

近几年来,黄病毒的分子生物学研究取得了一些重要进展。在黄病毒基因研究方面,Rice等<sup>[6]</sup>首先报道了黄热病病毒RNA的全序列分析结果。此外,其它几种黄病毒RNA的部分序列也已弄清楚。在黄病毒的蛋白质分析方面,对部分多肽进行了理化鉴定及氨基酸序列分析,对多肽的功能分析发现,某些多肽上存在免疫保护和抗体依赖的增强作用(ADE)抗原表位(epitope)。本文将综合介绍这些方面的研究进展并展望病毒分子生物学研究的前景。

本文于1987年8月10日收到。

## 一、黄病毒属与甲病毒属的区别<sup>(1)</sup>

提出黄病毒科的分类, 标志着病毒学家对黄病毒的认识不断向分子水平深入。黄病毒科的某些种(如黄热病病毒、登革病毒)是本世纪初发现的, 也是最早认识的与人类疾病有关的滤过性病毒<sup>(2,3)</sup>。在早期, 人们主要是根据其传播媒介, 划分为虫媒病毒; 后来则着眼于病毒本身的结构特性, 从B组虫媒病毒划分出披膜病毒科; 现在, 由于病毒分子生物学的进展, 又发现披膜病毒科中黄病毒属与其它病毒属(甲病毒属、风疹病毒属及痘病毒属), 特别是和甲病毒属相比, 存在很大的差异, 故将黄病毒属单独作为一个新的病毒科。

与甲病毒属相比, 黄病毒属的病毒颗粒及其核心都较小, 核衣壳的对称性仍不十分清楚。甲病毒属的核衣壳则为二十面体对称。在核酸结构方面, 黄病毒的5'末端帽结构为I型, 3'末端缺乏PolyA, 结构蛋白的基因靠近5'末端, 而甲病毒5'帽属O型, 3'末端有PolyA, 结构蛋白的基因位于近3'末端。在核酸翻译成蛋白质的过程中, 黄病毒无亚基因组mRNA合成, 未检出多蛋白(Polyprotein)裂解及结构蛋白晚期合成的增多, 甲病毒则都能检出。在无细胞系统中, 黄病毒RNA翻译出结构蛋白, 而甲病毒RNA翻译出非结构蛋白。在结构蛋白方面, 黄病毒有三个即包膜糖蛋白(E)、核心蛋白(C)和膜样蛋白(M), 甲病毒则有两个包膜糖蛋白E<sub>1</sub>和E<sub>2</sub>, 而无膜样蛋白。在细胞内可见到并可分离到甲病毒的髓核或核衣壳, 但不能见黄病毒的核衣壳。甲病毒的成熟是在质膜, 通过出芽释放, 黄病毒的成熟及形态发生则不确定, 最近Hase等<sup>(7)</sup>报道登革病毒的成熟可能也是芽出。

## 二、黄病毒的基因组结构

通过分子克隆和核酸序列分析方法, 结合病毒蛋白质氨基酸序列分析的结果, 现在已经知道黄热病病毒17D RNA的全部核苷酸一级结构; 其它几种黄病毒RNA的部分一级结构也已清楚, 如西尼罗病毒(WNV)、澳洲墨莱谷病毒(MVEV)及登革病毒等等。根据黄病毒核酸的一级结构, 还推测出黄病毒RNA可能的二级结构。

### (一)黄热病病毒17D RNA的一级结构<sup>(6)</sup>

黄热病病毒17D是该病毒的减毒疫苗株。RNA分子量为 $3.75 \times 10^6$  Dalton, 含有10,862个核苷酸(nt), 5'末端118个核苷酸和3'末端511个核苷酸为非编码区, 其余10,233个核苷酸构成一特别长的开放读码区, 其起始密码为AUG, 终止密码为UGA, 可编码分子量为380,763 Dalton的多肽。

黄热病病毒17D RNA一级结构示意图见图1。

### (二)几种黄病毒RNA一级结构的比较

Castle等报道<sup>(8,9)</sup>, 西尼罗病毒全基因序列含10,960个核苷酸, 其开放读码区含

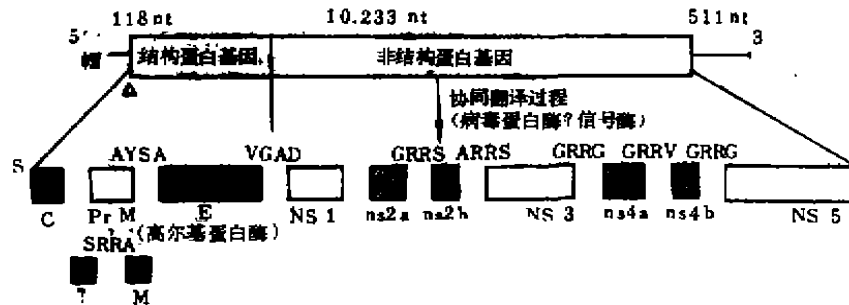


图 1: 黄热病病毒17D基因组 (10,862nt) 结构示意图

注: 1. 5'相属I型, 即m<sup>7</sup>G(5')ppp(5')AmpNp.....

2. nt为核苷酸缩写。

3. 结构蛋白及非结构蛋白的名称, 裂解点, 分子量及功能见表3。

Fig 1 Diagram illustrating gene order of YFV 17D (10,862 nt)

10,290个核苷酸, 可编码含3,430个氨基酸残基的多肽, 其中结构蛋白为787个氨基酸, 非结构蛋白为2,643个氨基酸, 多肽的分子量为379,670Dalton。Dalgarno等<sup>[10]</sup>分析了澳洲墨莱谷脑炎病毒基因组近5'末端的5,400个核苷酸的序列, 5'末端非编码区包括97个核苷酸, 这些研究结果都证实黄病毒有相同的基因排列顺序: 5'-C-prM(M)-E-NS1-ns2a-ns2b-NS3-3'。几种黄病毒RNA的碱基比例、双联核苷酸频率分别见表1及表2。

表1: 几种黄病毒RNA的碱基比例 (%)  
Table 1. Base ratio of flavivirus RNA

黄病毒	A	G	U	C
YFV	27.3	28.4	23.0	21.3
MVEV	26.4	28.7	24.4	20.7
DV	30	26	22	22

注: YFV—黄热病病毒  
MVEV—澳洲墨莱谷脑炎病毒  
DV—登革病毒  
YFV—Yellow fever virus  
MVEV—Murray valley encephalitis virus  
DV—Dengue virus

表2: 黄病毒双联核苷酸频率 (%)  
Table 2. The frequency of double nucleotides in flavivirus

YFV, MVEV	A	G	U	C
A	7.94, 8.05	7.70, 7.51	6.04, 6.30	5.58, 6.22
G	8.95, 8.21	8.76, 7.60	5.36, 5.25	5.33, 5.76
U	2.81, 3.70	9.32, 8.95	5.54, 6.68	5.00, 4.49
C	7.55, 8.10	2.32, 2.76	6.04, 5.60	5.39, 2.79

比较黄热病病毒和澳洲墨莱谷脑炎病毒 RNA 近 3' 末端的核苷酸顺序, 发现 3' 末端非编码区中有两段序列 (约 20 个核苷酸) 是保守序列 (Conserved sequence-CS)<sup>[11]</sup>。其中一段保守序列澳洲墨莱谷脑炎病毒中还是重复序列 (repeated sequence), 但在黄热病病毒中未见到该重复序列 (见图 2), 尚不知道重复序列的功能, 可能与 RNA 复制及核衣壳形成有关。在黄热病病毒、西尼罗病毒和澳洲墨莱脑炎病毒的 5' 末端非编码区内没有颠倒重复序列, 故 5' 端不会形成特征性的稳定的主干——袢环 (stem-loop) 结构。

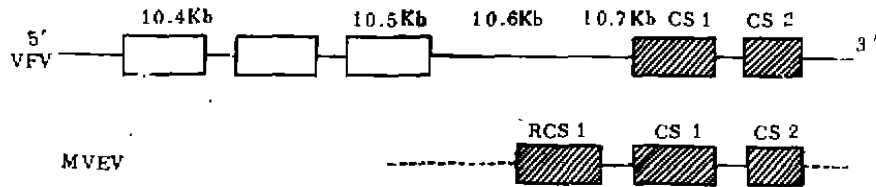


图 2 黄热病病毒 (YFV) 和澳洲墨莱谷脑炎病毒 (MVEV) RNA 的保守序列 (CS) 和重复保守序列 (RCS)

Fig. 2 Conserved and repeated conserved sequences between YFV and MVEV

根据相邻核酸序列对 AUG 起动翻译效率的影响, 最有效的起动序列为 ACC

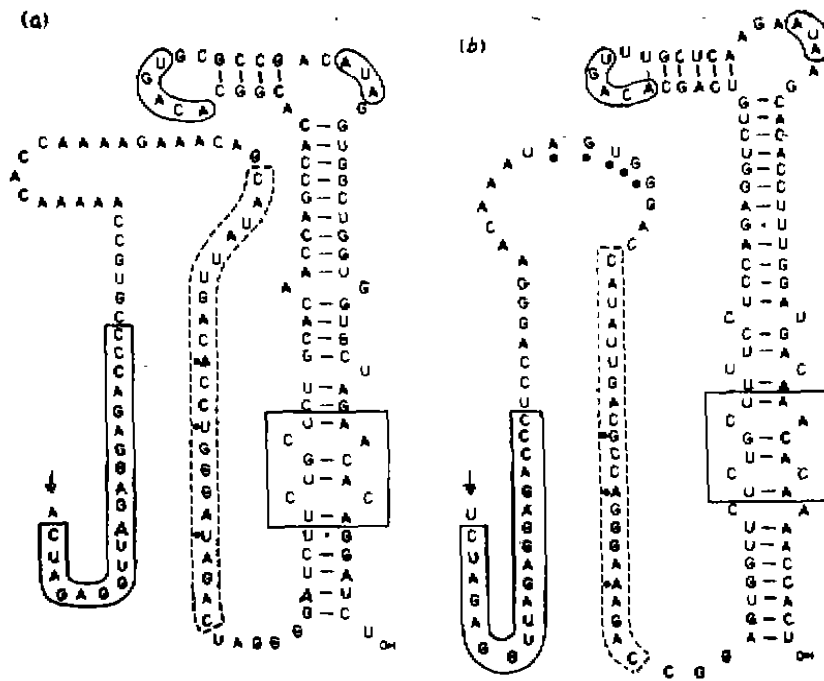


图 3 西尼罗病毒 (a) 和黄热病病毒 (b) RNA 3' 端二级结构。  
注: 方框内为保守序列。

Fig. 3 Secondary structure model of 3'-RNA in WNV (a) and YFV (b)

AUGG<sup>[12]</sup>。西尼罗病毒近5'端的第一个AUG位于UCGAUGU序列中其起动效率极低；第二个AUG位于AAUAUGG序列中，较第一个起动序列更有效。但实际上，西尼罗病毒通常使用第一个起动序列，这可能是由于黄病毒的结构蛋白和非结构蛋白都对宿主细胞有毒性作用，短期大量积累会导致宿主细胞死亡，故西尼罗病毒选择了长时间慢速增殖的策略<sup>[9]</sup>。此外，圣路易脑炎病毒(SLEV)、登革病毒基因组的部分序列也有报道<sup>[18,14,16]</sup>。

### (三)黄病毒 RNA 的二级结构

黄病毒 RNA 的二级结构是在一级结构基本弄清的前提下，由一级结构通过计算机程序推测出来的<sup>[16,17,18]</sup>。基本原理是根据 RNA 分子内部核苷酸最大可能地按碱基互补原则配对，形成双股区(主干)，不能配对的碱基成环凸出，这样就可构成自由能最小的稳定的二级结构。目前仅提出了黄病毒 RNA 3'末端的二级结构模型(见图3)。

## 三、黄病毒的蛋白质分析

黄病毒的蛋白质大体可分为结构蛋白和非结构蛋白，前者存在于完整病毒的颗粒中，后者仅存在于病毒复制的细胞之中。由于各种蛋白电泳技术的建立(如聚丙烯酰胺凝胶电泳、免疫印迹等)和氨基酸序列分析技术以及单克隆抗体的应用，已经检出黄病毒包含11—15种多肽，部分多肽的理化性质及部分氨基酸序列也已清楚。在黄病毒蛋白质的功能研究方面，近年来也积累了一些资料<sup>[19-21]</sup>，主要集中于起免疫保护作用 and 抗体依赖增强作用的表位在多肽中的分布。

### (一)黄病毒蛋白质

表3列举了黄病毒蛋白质的名称、裂解位置、理化性质及某些生物学功能<sup>[6]</sup>。

### (二)黄病毒蛋白质的氨基酸序列分析

Bell等<sup>[24]</sup>分析了黄热病病毒、圣路易脑炎病毒、登革病毒等三种黄病毒的结构蛋白的N末端氨基酸序列。Rice等<sup>[6,21]</sup>分析了黄热病病毒及澳洲墨莱谷脑炎病毒的蛋白质。结果表明，糖蛋白E的N末端33个氨基酸残基是保守序列，33个残基中13个残基(40%)相同，而且对形成二级结构起重要作用的氨基酸残基都是保守的。然而，核衣壳蛋白C的同源性则显著较低，仅25%，且含有大量的碱性氨基酸(赖氨酸、精氨酸)，黄热病病毒和圣路易脑炎病毒分别含1/4及1/3的碱性氨基酸。对三个最大的非结构蛋白的部分氨基酸序列分析表明，糖蛋白NS1由信号酶裂解产生，NS3和NS5的N端是在双碱性氨基酸残基之后裂解形成的。这些结果，反过来又证明了核酸序列分析的正确性。氨基酸保守序列对于研究黄病毒的起源、进化及疫苗研制都相当重要。图4是几种黄病毒E糖蛋白N端的高度保守区序列。

Trent等<sup>[15]</sup>比较了几种黄病毒核衣壳蛋白的氨基酸序列，提出了黄病毒的进化树，见图5。

表3 黄病毒的蛋白质与多肽  
Table 3. The proteins and polypeptides of flavivirus

新名称	旧称	N端裂解点	分子量 $kD$	糖化	功能
结构蛋白					
C	V <sub>2</sub> (NV <sub>1</sub> $\frac{1}{2}$ )	M ↓ S	13~16	-	核衣壳蛋白
prM	(NV <sub>2</sub> )(NV <sub>2</sub> $\frac{1}{2}$ )	?	19~23	+	M的前体
M	V <sub>1</sub>	SRR ↓ A	8~8.5	-	毒粒包膜蛋白
E	V <sub>3</sub>	AYS ↓ A	51~60	- , +	主要毒粒膜蛋白
非结构蛋白					
NS <sub>1</sub>	NV <sub>3</sub>	VGA ↓ D	44~49	+	可溶性补结抗原
ns2a	(NV <sub>2</sub> $\frac{1}{2}$ )(NV <sub>2</sub> )	(TVA ↓ V)	16~21	-	疏水性, 功能不清
ns2b	(NV <sub>1</sub> $\frac{1}{2}$ )	(GRR ↓ S)	12~15	-	疏水性, 功能不清
NS <sub>3</sub>	NV <sub>4</sub>	ARR ↓ S	67~76	-	复制酶成份
ns4a	(NVX)(NV <sub>2</sub> $\frac{1}{2}$ )	(GRR ↓ S)	24~32	-	疏水性, 功能不清
ns4b	(NV <sub>1</sub> )	(QRR ↓ V)	10~11	-	疏水性, 功能不清
NS <sub>5</sub>	NV <sub>5</sub>	GRR ↓ C	91~98	-	复制酶成份?

注: E有糖化及非糖化两种类型。裂解点是据序列分析资料推测的。  
氨基酸缩写符号: A-丙氨酸, C-胱氨酸, D-天冬氨酸, E-谷氨酸, F-苯丙氨酸,  
G-甘氨酸, H-组氨酸, I-异亮氨酸, K-赖氨酸, L-亮氨酸,  
M-蛋氨酸, N-天冬酰胺, P-脯氨酸, Q-谷氨酰胺, R-精氨酸,  
S-丝氨酸, T-苏氨酸, V-缬氨酸, W-色氨酸, Y-酪氨酸。

note: E protein has two types eg, glycolated and non-glycolated.  
The cleavage sites are putative.

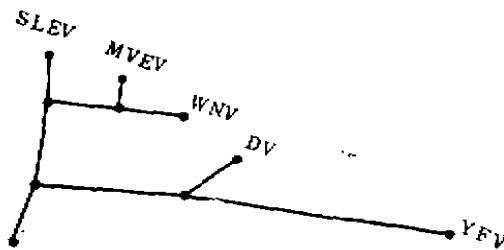


图4 根据C蛋白序列分析提出的黄病毒进化树。

注: 从图中可知, 所有黄病毒起源一共同祖先; 圣路易脑炎病毒(SLEV), 澳洲墨莱谷脑炎病毒(MVEV)和西尼罗病毒(WNV)亲缘关系较近; 而登革病毒(DV)和黄热病病毒(YFV)则关系密切。

Fig. 4. Flavivirus evolutionary tree based on sequence analyses of core protein

note: All flaviviruses have a common ancestor, There are close relationship among SLEV, MVEV and WNV while DV and YFV are more related.

YFV	P CCA	P CCU	P UUU	G GGA	D GAC	S AGC	Y UAC	I AUU
WNV	P CCC	P CCG	F UUU	G GGU	D GAC	S UCU	Y UAC	I AUC
MVEV	P CCA	P CCC	F UUC	G GGA	D GAC	S UCA	Y UAC	I AUU
DV	P CCU	P CCA	F UUC	G GGA	D GAC	S AGC	Y UAC	I AUC

图 5: 几种黄病毒E糖蛋白N端高度保守区序列。

注: 每排上行是氨基酸, 下行是密码子, 从图中可见氨基酸比核苷酸序列更保守, 氨基酸缩写符号见表 3 注释。

Fig. 5. Highly conserved sequence of E protein among several flaviviruses

note: Amino acid are above and condons are below in each rank. The sequence of amino acids is much more conserved than that of nucleotides.

(三)黄病毒蛋白质的免疫保护和抗体依赖增强表位 应用单克隆抗体分析黄病毒蛋白质的抗原表位是近几年的研究进展之一。人们希望通过表位分析, 找到主要起免疫保护作用的多肽, 从而为制备疫苗奠定基础。最初注意到黄病毒结构蛋白 E 能诱导中和抗体, 最近人们将注意力逐渐转向非结构蛋白 NS1<sup>(25-27)</sup>。NS1 是含量最多的非结构蛋白, 性质属糖蛋白, 用纯化的 NS1 免疫小鼠可完全保护黄热病脑炎及部分保护登革热脑炎的发病, 抗 NS1 的单克隆抗体也有被动免疫保护作用。NS1 较 E 优越之处在于, 用 NS1 免疫无明显的抗病毒颗粒抗体产生, 因而发生抗体增强效应的可能性较小。

抗体依赖的增强效应在登革出血热及其它几种黄病毒病的发病中可能起重要作用。用单克隆抗体分析表明, 同型登革病毒诱导的单克隆抗体对同型病毒既有中和作用又有抗体依赖的增强效应, 异型病毒(如登革病毒 4 型)免疫制备的单克隆抗体对登革病毒乙型也有明显的抗体依赖增强效应<sup>(28)</sup>。结合临床资料发现, 与登革出血热有关的登革病毒乙型往往含有多个抗体增强表位。过去认为, 抗体增强表位只与包膜糖蛋白有关, 但 Morens 等发现, 与包膜不发生反应的单克隆抗体也有抗体增强效应<sup>(29)</sup>。所谓抗体增强效应, 就是将一定量的病毒与适当稀释的抗体(可以是单克隆或多克隆的抗体)混合感染单核细胞系(如 P388D<sub>1</sub>), 感染并培养一段时间后, 检测培养上清中的病毒量, 并与对照组比较, 如果病毒量增多(往往以蚀斑形成单位表示), 则说明该抗体有增强(病毒繁殖)的效应。

#### 四、展 望

黄病毒的分子生物学研究目前主要集中于核酸和蛋白质两方面, 今后将继续完成其它黄病毒的核苷酸及氨基酸的序列分析, 在得到更多的一级结构资料后, 通过比较分

析,可能更加明确黄病毒种属间的亲缘关系程度,黄病毒的起源及进化。研究黄病毒 RNA 的二级结构以及 RNA 多聚酶的性质和作用,将会更好地阐明黄病毒的复制过程。

研究黄病毒核酸和蛋白质的目的之一是防治黄病毒病。这方面的工作已经开始, Henchal 等将制备的登革病毒 cDNA 探针做 RNA-cDNA 杂交,检测培养上清中的登革病毒 RNA 获得成功,进一步研究可望将该探针应用于临床登革热的快速诊断及现场流行病学调查<sup>[30]</sup>。Asphani 等<sup>[31]</sup>报道了生物素光标记的黄病毒核酸探针用于斑点杂交成功。已有实验室开始登革病毒基因工程疫苗的研制<sup>[32]</sup>。

黄病毒的致病机理仍未阐明。有些黄病毒嗜神经组织(如多种脑炎病毒),有些嗜特异脏器(如黄热病病毒嗜肝脏),也有些嗜单核——巨噬细胞(如登革病毒)。黄病毒的这些亲嗜性是否与表面的特殊识别结构有关?黄病毒识别结构与宿主细胞受体如何相互作用?都是值得深入探讨的问题。近年来的研究表明,登革病毒上的抗体依赖增强表位可能与登革出血热的发病机理有关。已发现小鼠体内有黄病毒的抵抗基因<sup>[33]</sup>,而人类则对多种黄病毒普遍敏感,尚不清楚小鼠抵抗黄病毒感染的基因产物。

人和昆虫在进化树上的跨度是如此之大,黄病毒是如何适应这两类亲缘关系甚远的宿主呢?有人发现,黄病毒的 G—C 双联核苷酸与人的 G—C 双联核苷酸,都较由碱基比例预期的数目低得多,如黄病毒 G+C 为 48%,预期 G—C 频率为 6.1%,实际为 2.4%,人 G+C 为 40%,预期 4% 实际 G—C 为 0.8%,而昆虫及其它昆虫病毒的 G—C 双联核苷酸频率与由碱基成份推算的频率较接近。从进化论的观点看,黄病毒对其主要宿主——昆虫的生存似应有某些有益作用,否则没有宿主,它本身无法生存。RNA 病毒较 DNA 病毒的变异快得多<sup>[34]</sup>,黄病毒 RNA 可能促进昆虫的适应性变异<sup>[35]</sup>,黄病毒对人类除了致病外,是否也有某些进化上的意义呢?

除了核酸与蛋白质外(在黄病毒中的含量分别为 6% 及 66% 左右),还有脂类(17%)及糖类(9%)等生物大分子尚无人问津。这些生物大分子在黄病毒中的功能仍不清楚。

总之,对于黄病毒,还有许多问题需要人们从分子水平去探讨。

### 参 考 文 献

- [1] Westaway EG et al., 1985, *Intervirolgy* 24: 83.
- [2] Porterfield JS, 1980, in *Togaviridae* Schlessinger RW ed, Academic Press New York pp. 13
- [3] Porterfield JS, 1986, in *Togaviridae and Flaviviridae* Schlessinger S et al. ed, Plenum Press New York pp. 1.
- [4] Shope RE, 1980, in *Togaviridae* Schlessinger RW ed, Academic Press, New York, pp. 47.
- [5] 陈火胜等, 1987, 广东省登革热研讨会论文集 第 1 页。
- [6] Rice CM et al., 1985, *Science* 229: 721.
- [7] Huse T, 1987, *Arch Virol* 92(3—4): 273.
- [8] Castle E et al., 1986, *Virology* 149: 10.
- [9] Castle E et al., 1987, *Arch virol* 92: 309



- [10] Dalgarno L et al., 1986, *J Mol Biol* 187: 309.
- [11] Rice CM et al., 1987 in *Molecular Trends in Virology* Bauer H et al., ed, in press.
- [12] Kozak M. 1986, *Cell* 44: 283.
- [13] Trent DW et al., 1987, *Virology* 156: 293.
- [14] Bangti Zhao et al., 1986, *Virology* 155: 77.
- [15] Deubel V et al., 1986, *Virology* 155: 365.
- [16] Wengler G et al., 1986, *J gen Virol* 67: 1183.
- [17] Brinton MA et al., 1985, *Virology* 153: 113.
- [18] Grange T et al., 1985, *FEBS* 188: 159.
- [19] Smith GW et al., 1985, *J gen Virol* 66: 559.
- [20] Verndam AV et al., 1984, *J Virol Methods* 9: 283.
- [21] Rice CM et al., 1986, *Virology* 151: 1.
- [22] Wengler G et al., 1985, *Virology* 147: 264.
- [23] Castle E et al., 1985, *Virology* 145: 227.
- [24] Bell JR et al., 1985, *Virology* 143: 224.
- [25] Schlesinger JS et al., 1986, *J Virol* 60: 1153.
- [26] Schlesinger JS et al., 1987, *J gen Virol* 68: 853.
- [27] Henchal EH et al., 1987, *J gen Virol* 68: 845.
- [28] Morens DM et al., 1987, *J gen Virol* 68: 91.
- [29] Halstead SB et al., 1987, *J gen Virol* 68: 91.
- [30] Henchal EH et al., 1987, *J Virol Methods* 15: 187.
- [31] Khan AM et al., 1987, *J Virol Methods* 15: 121.
- [32] Young SH et al., 1987, *CALTECH Biology Annual Report*, pp 129.
- [33] Brinton MA, 1981, *J Virol* 39: 413.
- [34] Holland J et al., 1982, *Science* 215: 1577.
- [35] Strauss EC et al., 1983, *Curr Top Microbiol Immunol* 105: 1.