

# 多角体蛋白的结构多样性和杆状病毒的进化研究\*

刘子夜 齐义鹏\*\* 王家旺 黄永秀

(武汉大学病毒研究所, 武汉 430072)

**摘要** 在测定 LsMNPV 多角体蛋白基因的全序列并推导出多角体蛋白氨基酸顺序的基础上, 与 22 种杆状病毒多角体蛋白的氨基酸顺序进行比较, 以氨基酸变异曲线研究蛋白质一级结构中氨基酸的保守性, 并推测出一个模式多角体蛋白(MPh)氨基酸顺序; 用 PROSIS 软件对 MPh 及其它 5 种代表性病毒的多角体蛋白进行了氨基酸亲水性分析, 还对 24 种多角体蛋白的二级结构作出了推测, 指出了特征性结构区与氨基酸高变区、亲水区多样性变化的相互关系; 通过多角体蛋白间氨基酸顺序最大同源性分析, 绘制了 23 种杆状病毒的系统进化树, 进一步讨论了 MPh 氨基酸顺序的典型性及多角体蛋白所体现的杆状病毒多样性和宿主依赖性进化。

**关键词** 杆状病毒, 多角体蛋白, 亲水性分析, 二级结构, 进化

**Structural comparisons of polyhedrin proteins and phylogeny of baculoviruses/Liu Ziye, Qi Yipeng, Wang jiaawang, Huang Yongxiu //CHINESE BIODIVERSITY. —1997, 5(1): 14~25**

We have sequenced the polyhedrin gene from LsMNPV and the deduced amino acid sequence was compared with other polyhedrin or granulin amino acid sequences from 22 baculoviruses. Conservation of amino acid sequences was counted using a curve and a model polyhedrin (MPh) was given reasonably. Hydrophilicity of MPh and polyhedrins from some representative baculoviruses was studied by the PROSIS software. Prediction of secondary structure of MPh and 23 polyhedrins was stated in computer. Relationship among the predicted structure in special amino acid regions and the variability, the hydrophilicity was found. Evolution of 23 baculoviruses was discussed according to the maximum homologous analysis of amino acid sequences between two polyhedrin separately.

**Key words** baculovirus, polyhedrin, hydrophilicity, secondary structure, evolution

**Author's address** Institute of Virology, Wuhan University, Wuhan 430072

杆状病毒一个最显著的形态特征是病毒粒子被包埋在结构致密的晶体状蛋白质基质(martrix)中而形成包涵体(occlusion body), 包涵体的 95% 由一种单一蛋白——多角体蛋白(polyhedrin)构成, 多角体蛋白在极晚期的表达量可达细胞可溶性蛋白的 1/3 以上。作为一个非必需结构基因, 近 10 年来, 人们对多角体蛋白基因(ocu)的高表达调控机制并利用其高表达启动了表达外源基因的研究特别感兴趣, 使 ocu 成为目前研究和应用得最多的昆虫杆状病毒基因<sup>[1,2]</sup>。迄今为止, 已测定了 20 几种多角体(包括同源颗粒体病毒的颗粒体)基因序列, 发现 ocu 是最保守的杆状病毒基因。

对杆状病毒多角体蛋白氨基酸顺序及其变异特性研究、氨基酸亲疏水区域分析, 特别是多

角体蛋白在结构和功能保守性的前提下,从分子进化的角度对杆状病毒系统进化的详细研究还未见报道。仅在 1986 年,以当时发表的不到 10 个 ocuDNA 序列为基础,Rohrmann<sup>[3]</sup>进行了 DNA 序列、氨基酸顺序同源性比较研究,并尝试描绘了一个较简单的进化树。但是,人们一直希望通过某个分子基础模型来阐明杆状病毒的系统进化问题,分析多角体蛋白一级结构的保守性、变异区及亲疏水区等结构特点,为研究多角体蛋白可能的其他生物功能打下基础,是另一个令人感兴趣的议题<sup>[4]</sup>。

生物多样性研究是一个正在兴起的课题,而有关微生物特别是病毒的多样性研究极少报道。感染昆虫的杆状病毒科病毒,特别是核型多角体病毒和颗粒体病毒家族,其多角体蛋白的结构和功能的高度保守性和可能的规律性变化似乎具备进行生物多样性研究的条件,本文拟对此作一有益的尝试。

本文在测定粘虫核型多角体病毒(LsMNPV)ocu DNA 序列的基础上预测多角体蛋白的氨基酸顺序,同时检索 22 种杆状病毒 ocuDNA 序列和蛋白氨基酸顺序;用计算机分析这 23 种杆状病毒的多角体蛋白氨基酸顺序的同源位置对氨基酸的利用情况,提出一个模式多角体蛋白(model polyhedrin, MPh)氨基酸顺序,并将其代入下面的研究;通过氨基酸变异区和氨基酸亲水区的对应分析,讨论多角体蛋白一级结构中保守性区域及高变的抗原决定位点的可能分布;还进行氨基酸组成分析和蛋白质二级结构预测;通过对上述 24 种多角体蛋白氨基酸最大同源性的相互比较,分析这些杆状病毒的亲缘关系并试绘进化树,进一步说明 MPh 的典型性。

## 1 材料和方法

### 1.1 23 种杆状病毒 ocuDNA 序列和/或蛋白质氨基酸顺序的来源

- (1)粘虫核型多角体病毒(多粒包埋)*Leucania (Pseudatetia) separata* MNPV (LsMNPV)<sup>[5]</sup>
- (2)苜蓿丫纹夜蛾核型多角体病毒(多粒包埋)*Autographa californica* MNPV (AcMNPV)<sup>[6,7]</sup>
- (3)Ag 夜蛾核型多角体病毒(多粒包埋)*Anticarsia gemmatilis* MNPV (AgMNPV)<sup>[8]</sup>
- (4)蓖麻蚕核型多角体病毒 *Attacus ricini* NPV (ArNPV)<sup>[9]</sup>
- (5)黄地老虎颗粒体病毒 *Agrotis segetum* GV (AsGV)<sup>[10]</sup>
- (6)黄地老虎核型多角体病毒(多粒包埋)*Agrotis segetum* MNPV (AsGV)<sup>[10]</sup>
- (7)家蚕核型多角体病毒 *Bombyx mori* NPV (BmNPV)<sup>[11]</sup>
- (8)油桐尺蠖核型多角体病毒(单粒包埋)*Buzura suppressaria* SNPV (BsSNPV)<sup>[12]</sup>
- (9)云杉卷叶蛾核型多角体病毒(多粒包埋)*Choristoneura fumiferana* MNPV (CfMNPV)<sup>[13]</sup>
- (10)大蜡螟核型多角体病毒(多粒包埋)*Galleria mellonella* MNPV (GmMNPV)<sup>[13]</sup>
- (11)美洲棉铃虫核型多角体病毒(多粒包埋)*Heliothis zea* MNPV (HzMNPV)<sup>[14]</sup>
- (12)舞毒蛾核型多角体病毒(多粒包埋)*Lymanlia dispar* MNPV (LdMNPV)<sup>[15]</sup>
- (13)甘兰夜蛾核型多角体病毒(多粒包埋)*Mamestra brassicae* MNPV (MbMNPV)<sup>[16]</sup>
- (14)欧洲松锈锯角夜蜂核型多角体病毒(单粒包埋)*Neodiprion sertifer* SNPV (NsSNPV)<sup>[17]</sup>
- (15)冷杉合毒蛾核型多角体病毒(单粒包埋)*Orgyia pseudotsugata* SNPV (OpSNPV)<sup>[18]</sup>
- (16)冷杉合毒蛾核型多角体病毒(单粒包埋)*Orgyia pseudotsugata* MNPV (OpMNPV)<sup>[18]</sup>
- (17)松夜蛾(小眼夜蛾)核型多角体病毒(多粒包埋)*Panolis flammea* MNPV (PfMNPV)<sup>[19]</sup>
- (18)甘兰菜粉蝶颗粒体病毒 *Pieris brassicae* GV (PbGV)<sup>[20]</sup>
- (19)甜菜夜蛾核型多角体病毒(多粒包埋)*Spodoptera exigua* MNPV (SeMNPV)<sup>[21]</sup>
- (20)草地贪夜蛾(秋粘虫)核型多角体病毒(多粒包埋)*Spodoptera frugiperda* MNPV (SfMNPV)<sup>[22]</sup>
- (21)海灰翅夜蛾核型多角体病毒(多粒包埋)*Spodoptera littoralis* MNPV (SiMNPV)<sup>[23]</sup>
- (22)欧洲大蚊核型多角体病毒(单粒包埋)*Tipula palludosa* SNPV (TpSNPV)<sup>[3]</sup>

(23)粉纹夜蛾颗粒体病毒 *Trichoplusia ni* GV (TnGV)<sup>[24]</sup>

注: MNPV: mullinucleocapsid nuclear polyhedrosis virus; SNPV: single nucleocapsid nuclear polyhedrosis virus; GV: granosis virus

1.2 多角体蛋白氨基酸分析

排列出 23 种杆状病毒多角体蛋白的氨基酸顺序, 以与 LsMNPV 多角体蛋白的同源大小为序; 分析每一位置上氨基酸的使用种类数, 以种类数相对氨酸位置作图, 描绘多角体蛋白一级结构上的氨基酸变异区域; 分析每一位置上共用次数相对最多的氨基酸, 推测出 MPh 的氨基酸顺序。以 MPh 及 LsMNPV、AcMNPV、HzMNPV、NsSNPV 和 AsGV 的多角体蛋白为例, 在计算机上用 PROSIS 软件以 Hopp 和 Wood<sup>[25]</sup> 的赋值进行氨基酸亲水性分析, 并比较亲疏区的同源性和特征。用 PROSIS 软件以 Chou 和 Fasman<sup>[26]</sup> 的赋值进行多角体蛋白二级结构推测, 比较 MPh 和 23 种多角体蛋白间特征性区域的异同。

1.3 多角体蛋白氨基酸最大同源性百分率及杆状病毒系统进化树

用 PROSIS 软件将 MPh 及 23 种多角体蛋白分别进行相互间的氨基酸最大同源性定量比较, 求出其百分率; 以此为依据用相邻参比法(neighbous-joiningmethod)<sup>[27]</sup> 绘出这些病毒的系统进化树。

2 结果

2.1 氨基酸顺序分析

23 种多角体蛋白及 MPh 氨基酸顺序排列比较(表 1)可以直观地分析氨基酸位置的同源性, 氨基酸的位置序号以 LsMNPV 及 MPh 为基准, AcMNPV 以下的多角体蛋白应注意“\*”计算氨基酸的实际位置和总数; 根据多角体蛋白氨基酸变异分析(图 1), 可见 3 个显著变异区, 氨基酸位置为 1~20、30~50、110~130, 其中 30~50 的变异最为显著。

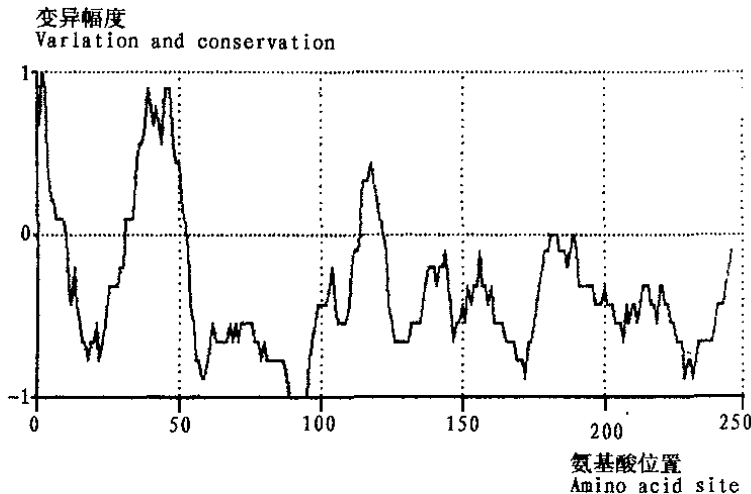


图 1 多角体蛋白一级结构氨基酸变异曲线  
Fig.1 Variation profile of amino acid in primary structure of polyhedrins

2.2 氨基酸亲水性分析

将分析曲线(图 2)直观比较发现: MPh 及鳞翅目昆虫病毒 AcMNPV、LsMNPV、HzMNPV 多角体蛋白存在 2 个高亲水区: 30~50、210~240; 在氨基酸位置 50 以后的亲疏区几乎完全一致, 30~50 的典型高亲水区同时也有较大的数值变化; 膜翅目昆虫病毒 NsSNPV 和颗粒体病

表1. 23种杆状病毒多角体蛋白及模式多角体蛋白氨基酸序列

Table 1. Amino acid sequences of 23 baculovirus polyhedrins and MFn

	10	20	30	40	50	60	70	80	90	100	110	120
1. MPh	MY***TRSYNPT	LGRTVYDVK	YKNGLSVIK	NKRKKGHLIE	HEEEKTLDP	LDRYLVADP	FLGPGKNQKL	TLFKEIRNVK	PDTMKLVNV	SGKEFLRETW	TRFMDSFPI	VNDQEVMDVF
2. Lm	MY***TRSYNPS	LGRTVYDVK	YKNGLSVIK	NANRKGHIE	HEEEKTLDP	LDRYLVADP	FLGPGKNQKL	TLFKEIRNVK	PDTMKLVNV	SGKEFLRETW	TRFMDSFPI	VNDQEVMDVF
3. Mm	.....S.....	.....R.Y.....	.....K.....	.....V.R.....	.....K.....	.....N.....	.....N.....	.....N.....	.....N.....	.....N.....	.....I.....	.....I.....
4. Bm	.....K.....	.....A.....	.....A.....	.....ND.....	.....Q.....	.....K.....	.....N.....	.....N.....	.....N.....	.....N.....	.....I.....	.....I.....
5. Pm	.....FP.....	.....A.....	.....A.....	.....Q.....	.....A.H.....	.....K.....	.....N.....	.....N.....	.....N.....	.....N.....	.....I.....	.....I.....
6. OpS	.....F.....	.....F.....	.....K.E.LAL	.....I.R.....	.....E.V.....	.....E.TY.....	.....I.....	.....I.....	.....I.....	.....I.....	.....I.....	.....I.....
7. SfM	.....A.....	.....F.....	.....K.E.LLQ	.....I.R.....	.....E.TY.....	.....I.....	.....I.....	.....I.....	.....I.....	.....I.....	.....I.....	.....I.....
8. Sem	.....S.T.K.....	.....F.....	.....K.....	.....LE.....	.....H.N.S.....	.....K.....	.....S.....	.....S.....	.....S.....	.....R.....	.....K.....	.....I.....
9. Hzm	.....PD.....	.....R.T.I.....	.....A.....	.....K.....	.....A.....	.....N.....	.....N.....	.....N.....	.....N.....	.....G.....	.....K.....	.....I.....
10. AcM	.....SRYSA.N.S.H.....	.....K.....	.....H.....	.....K.HDAL	.....R.AD.RE.H.....	.....K.....	.....M.....	.....N.....	.....N.....	.....I.....	.....V.....	.....Y.....
11. SfM	.....PD.....	.....R.T.I.....	.....LL.....	.....ED.H.....	.....H.M.....	.....Y.....	.....V.N.....	.....I.....	.....I.....	.....V.....	.....V.....	.....Y.....
12. OpM	.....PN.....	.....T.I.....	.....GL.....	.....K.L.....	.....K.QW.....	.....N.M.....	.....Y.....	.....V.N.....	.....I.....	.....V.....	.....V.....	.....Y.....
13. Bm	.....PD.....	.....R.T.I.....	.....K.....	.....LL.....	.....Q.E.....	.....S.G.....	.....H.I.....	.....N.....	.....I.....	.....V.....	.....V.....	.....Y.....
14. AgM	.....PD.....	.....R.T.I.....	.....K.....	.....LL.....	.....ED.H.....	.....H.M.....	.....I.....	.....N.....	.....I.....	.....V.....	.....V.....	.....Y.....
15. CfM	.....HNF.N.F.A.....	.....K.....	.....E.....	.....Q.K.Q.....	.....LEQ.....	.....I.RS.H.....	.....N.....	.....N.....	.....N.....	.....N.....	.....N.....	.....IY.....
16. Ldm	.....PN.....	.....RPT.I.....	.....A.....	.....K.....	.....LXX	.....XTXX.ZGXV	.....ZLL.....	.....N.....	.....N.....	.....K.....	.....Y.....	.....Zxx.....
17. GmM	.....PD.....	.....R.T.I.....	.....KA.....	.....IV.....	.....E.....	.....HW.....	.....N.M.....	.....N.....	.....I.....	.....V.....	.....V.....	.....I.....
18. A1	.....NF.....	.....T.I.....	.....F.....	.....K.Q.L.....	.....LK.KQ.....	.....TFM.....	.....V.N.....	.....V.N.....	.....V.N.....	.....V.N.....	.....V.N.....	.....V.I.....
19. AsM	.....PNAQ.GT.....	.....AKS*I.....	.....G.DI.N.S.KGR.....	.....DQD.W.KHQRRA	.....NGFILP.G.....	.....RT.....	.....HV.M.VM.Q.V.N.I.AN.....	.....AI.....	.....S.Y.....	.....V.....	.....T.I.....	.....N.Y.EI.....
20. NsS	.....G*YKNSLR.SRH	.....N.T.C.I.....	.....HL.Y.....	.....LG.DVRH.BEL.K	.....EAQFDPIK.I	.....ANQ.M.T.....	.....R.....	.....V.I.....	.....IR.....	.....C.A.....	.....ISEE.....	.....TT.I.I.IW
21. TnG	.....G*YHL.AL.SKH	.....E.T.C.I.....	.....Q.H.S.A.L	.....DVKN.DRLR	.....EAEI.PV.I	.....A.Q.M.T.....	.....R.....	.....VRI.....	.....C.R.E.....	.....L.C.....	.....M.....	.....TT.QI.NMW
22. Pbc	.....G*YKNSLR.SRH	.....A.TSC.I.....	.....Q.H.QIA.NG	.....DVR.DRR.S	.....EAKYAPLK.L	.....ANQ.M.T.....	.....R.....	.....VRI.....	.....R.E.....	.....H.IC.....	.....ISEE.....	.....TT.QI.NIL
23. AsG	.....QYGVND.G.E.N	.....VDYFNASHAG	.....L.DRSKQPYV	.....DHD?YP	.....	.....	.....	.....	.....	.....	.....	.....
24. Tps	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....

(未完待续)  
(To be continued)

(续表1)

(Continued from Table 1)

	130	140	150	160	170	180	190	200	210	220	230	240	
1. MPH	LVINRTRP	NRCYKFLAQH	ALRCDEDPVP	HEFIRIVEPS	YVGSNNERYI	SLAKGGGCP	VANLHSEYTN	SPEEFINRVI	WENFYKPIVY	VGTDASAEIE	ILLEVSLVFK	IKGFAPDAPL YTGPA	
2. LsM	LVINRTRP	NRCYKFLAQH	ALRCDEDPVP	HEVIRIVERS	CVGTNNERYV	SLAKRGGCP	VANLHSEYTN	SPEEFINRVI	WENFYKPIVY	VGTDASAEIE	ILLEVSLVFK	IKGFAPDAPL YNGPA	
3. MbM				F.....	P. Y. S.....	I.....						E..... T.....	
4. BsS		YR.....		F.....	P. Y. S.....	I.....						V.E..... T.....	
5. PfM		Y.....	E.....	F.....	P. Y. S.....							E..... T.....	
6. OpS		R.....	E.....	F.....	P. Y. S.....	I.....	A.....	H.....				E..... S.K.....	
7. SfM		R.....		F.....	PV Y. N.....	I.....	K.....	H.....				E.....	
8. SeM		R.....		F.....	PV Y.....	I.....	K.....					E.....	
9. HzM		Y.....	I.....	F.....	P. Y. S.....	I.....	KY.....	A.....	D. TN.....			E..... T.....	
10. AcM		V.....		D.....	P. W. S.....	I.....	K.....	I.....	Q. D.....	I.....		V.E..... FT.....	
11. SlM		R.....	E.....	D.....	P. Y.....	I.....	A.....	T.....	S.....	Y.....		E..... T.....	
12. OpM		V.....	W.C.....	F.....	P. Y. M.....	I.....	K.....	I. I. A.....	S.V.....	I.....	S.....	V.E..... FT.....	
13. Bm		A.LK.....	W.E.....	F.....	M.P. Y. M.....	I.....	K.....	I. I.....	S.V.....	I.....	I.....	E..... FT.....	
14. AgM		L.....	SHC.....	F.....	P. Y. M.....	I.....	K.....	I. I.....	S.V.....	I.....	S.....	V.E..... FT.....	
15. CfM		V.....	W.C.....	F.....	P. Y. M.....	I.....	K.....	I. I. A.....	S.V.....	I.....	G.....	M.I..... V.E..... A. FT.....	
16. LdM		Y.V.....	EG.....	F.....	P. T.E*.....	I.....	IR.....	A. T.....	H.I.S.....	D.....	T.....	E..... FQ.....	
17. GmM		A.LK.....	D.....	F.....	W. SXZ.....	I.....	K.....	I.....	X.....	22. D.....	ZK.....	I... TS... Z II..... V.E..... FT.....	
18. Ar		L.....	V.W.C.....	F.....	P. H. M.....	I.....	P. K.....	I.....	S.V.....	XELLOA.....	I.....	G..... I..... V.E.V..... FT.....	
19. AsM		EA.LK.....		F.....	PD Y. V.G.....	I.....			I.....	N. N.....	S.E.....	RNN..... I..... L.L.V.E..... I..... S.....	
20. NsS		E.RCT.NKS		G. M. M <sup>1</sup> ILC	LPFALSNI	ILRETDPLS	V. L. D.....	KPQFNEI	LD. I. IL	CH. HR.....	I.....	G..... VMI. A. T. I..... E..... FT.....	
21. TnG		FELOL. MQ.		Y. TM. Y	.....	DF. QHDPY	Y. PD. RE. I	N. S. ...LAF	LTC. Q. I. NE	N. D. FDQ. L	PY. HR. L.....	T..... I..... VMI. A. L..... E..... V..... FT.....	
22. PbG		FEIQV. MQ.		Y. HM. Y	.....	DF. AQDPY	YI. P. IE. I	N. * K. FAF	L. C. Q. V. ND	N. T. FED. L	PY. HR. L.....	I. T. S. T..... FL..... E..... V..... T.....	
23. AsG		FEIQL. MQ.		Y. T. Y	.....	GAH.....	A	DF. QGDPY	Y. P. QLE. I	N. S. K. YAY	LTC. Q. I. NE	N. DF. FDEHL	PY. HR. L..... MN..... I..... MI..... VI..... E..... V..... T.....
24.													

按与LsMNPV的同源性大小排列, TpsNPV只有N端顺序, ^ 表示模式多角体蛋白与LsMNPV的不同氨基酸位置, . 表示该氨基酸与LsMNPV相同, \* 表示不存在该位置。

The sequences are arranged to indicated their homology in decrement compared with LsMNPV; the N-terminal sequence is given only for TpsNPV;

^ indicates the different amino acid site between MPH and LsMNPV; . represents the same amino acid site to LsMNPV; \* represents the empty site in amino acid sequences of other polyhedrins

表2. 多角体蛋白二级结构分析

Table 2. Predicted analysis of secondary structure of polyhedrins

	10	20	30	40	50	60	70	80	90	100	110	120
1. MPH	SS***TTTTTTT	TTTTSTTTT	TTSTTTSSC	HHHHHHHHH	HHHHHSHHT	TTSSSCCCT	TTTTTTTHH	HHHHHSCCT	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
2. LsM	SS***TTTTTTT	TTTTSTTTT	TTSTTTSSC	TTTTTHHHH	HHHHHSHHT	TTSSSCCCT	TTTTTTTHH	HHHHHSSST	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
3. MbM	SS***TTTTTTT	TTTTSTTTT	TTSTTTTTC	TTTTTTTHH	HHHHHSHHT	TTSSSCCCT	TTTTTTTHH	HHHHHSCCT	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
4. Bss	SS***TTTTTTT	TTTTSTTTT	TTSSSSSSHH	HHHHHHHHH	HHHHHSHHT	TTTTTHHHT	TTTTTTTHH	HHHHHSCCT	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
5. PfM	SS***TTTTTTT	TTTTSTTTT	TTSSSSSSHH	TTTTTTTHH	HHHHHSHHT	TTSSSCCCT	TTTTTTTHH	HHHHHSCCT	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
6. Ops	SS***TTTTTTT	TTTTSTTTT	TTSSSSSSHH	HHHHHHHHH	HHHHHSHHT	TTTTTHHHT	TTTTTTTHH	HHHHHSCCT	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
7. SfM	SS***TTTTTTT	TTTTSTTTT	SSSTTTSSC	HHHHHHHHH	HHHHHSHHT	TTSSSHHHT	TTTTTTTHH	HHHHHSHST	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
8. SeM	SS***TTTTTTT	TTTTSTTTT	SSSTTTSSC	HHHHHHHHH	HHHHHSHHT	TTTTTHHHT	TTTTTTTHH	HHHHHSHST	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
9. HzM	SS***TTTTTTT	TTTTSTTTT	TTSSSSSSHH	HHHHHHHHH	HHHHHSHHT	TTTTTHHHT	TTTTTTTHH	HHHHHSCCT	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
10. AcM	C***TTTTTTSS	STTTSTTTT	TTSSSSSSHH	HHHHHHHHH	HHHHHSHHT	TTTTTCCTT	TTTTTTTHH	HHHHHSCCT	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
11. SlM	STTTTSTTTT	TTTTSTTTT	TTSSSSSSHH	HHHHHHHHH	HHHHHSHHT	TTTTTHHHT	TTTTTTTHH	HHHHHSCCT	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
12. OpM	C***TTTTTTSS	STTTSTTTT	TTSTTTSSC	TTTTTTTHH	HHHHHSHHT	TTSSSHHHT	TTTTTTTHH	HHHHHSCCT	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
13. Bm	C***TTTTTTT	TTTTSTTTT	TTSTTTSSC	HHHHHHHHH	HHHHHSHHT	TTTTTCCTT	TTTTTTTHH	HHHHHSCCT	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
14. AgM	C***TTTTTTSS	STTTSTTTT	TTSTTTSSC	HHHHHHHHH	HHHHHSHHT	TTSSSHHHT	TTTTTTTHH	HHHHHSCCT	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
15. CfM	C***TTTTTTSS	STTTSTTTT	TTSTTTSSC	HHHHHHHHH	HHHHHSHHT	TTSSSHHHT	TTTTTTTHH	HHHHHSCCT	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
16. LdM	S***SSSSSSSS	STTTSTTTT	TTSSSSSSSS	HHHHHHHHH	HHHHHSHHT	SSSSSHHHT	TTTTTTTHH	HHHHHSCCT	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
17. GmM	C***TTTTTTSS	STTTSTTTT	TTSSSSSSHH	HHHHHHHHH	HHHHHSHHT	HHHHHSHHT	TTTTTTTHH	HHHHHSCCT	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	SSSSSSSSSS
18. Ar	C***TTTTTTSS	STTTSTTTT	TTSTTTSSC	HHHHHHHHH	HHHHHSHHT	TTTTTCCTT	TTTTTTTHH	HHHHHSCCT	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTSSSSSS
19. AsM	SS***TTTTTTT	TTTTSTTTT	SSSTTTSSC	HHHHHHHHH	HHHHHSHHT	TTTTTHHHT	TTTTTTTHH	HHHHHSCCT	TTTTSSSSST	TTTTSSSSSS	SSSHHTTTTS	TTTTHHHHHH
20. Nss	T***TTTTTSHH	*HCC*CTTT	TTTTSSSCC	CCCHHHTTT	THHHHTTTTS	SSSSSTTTT	TTTTTHSSS	SSSSSSSSSS	SSSSSSSSST	TTTTSSSSSS	SSSSSSSSSS	TTTTHHHHHH
21. ThG	CT*TTTTTCCTT	TTTTSTTTT	TTSSHSSSSS	SHHHHHHHH	HHHHHSCCS	SSSSSHHHT	TTTTTTTSS	SSSSSSHHHH	HHSSSSSTT	TTTTHHHHHH	HHHHHHHSS	TTTTHHHHHH
22. PbG	ST*TTTSSSCCT	TTTTSTTTT	TTSSSSSSSS	TTTTTTTHH	HHSSSSSSSS	SSSSSSSHHT	TTTTTTTSH	HHHHHHHHH	HHSSSSHTT	TTTTHHHHHH	HHHHHHHSS	TTTTSSSSSS
23. AsG	CT*TTTTTCCTT	TTTTSTTTT	TTSSSHTTT	TTTTTTTCC	CCGCCCCCHS	SSSSSSSHHT	TTTTTTTSH	HHHHHHHHH	HHHHHHHTT	TTTTHHHHHH	HHHHHHHSS	TTTTHHHHHH
24. Tps	SSSTTTTTTTTT	TTTTTTTHH	HTTTTSSSS	CCC?CC								

(未完待续)

(To be continued)

(续表2)

(Continued from Table 2)

	130	140	150	160	170	180	190	200	210	220	230	240
1. Mph	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
2. Lsm	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
3. MbM	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
4. Bss	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
5. Pfm	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
6. Ops	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
7. sfM	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
8. Sem	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
9. Hzm	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
10. AcM	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
11. sLM	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
12. OpM	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
13. Bm	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
14. AgM	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
15. CfM	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
16. LdM	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
17. GmM	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
18. Ar	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
19. AsM	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
20. Nss	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
21. Tng	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
22. Pbg	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
23. AsG	SSSTTTTT	TTTTNNNN	HNCCTTTT	NNNNNN	TTTTTTTT	SSSTTTTT	SSSTTTTT	TTTTNNNN	NNNNSSSS	STTTNNNN	SSSSSSSS	NNNNCCCC
24.												

H:  $\alpha$  螺旋, S:  $\beta$  折叠, T:  $\beta$  转角, C: 无规则卷曲. 位置序号以 Mph 或 Lsm 为准, \* 表示不存在该位置H:  $\alpha$ -helix, S:  $\beta$ -sheet, T:  $\beta$ -turn, C: random coil. The number of amino acid site is in order to the sequences of Mph and/or Lsm

polyhedrin, \* represents the empty site in amino acid sequences of other polyhedrins

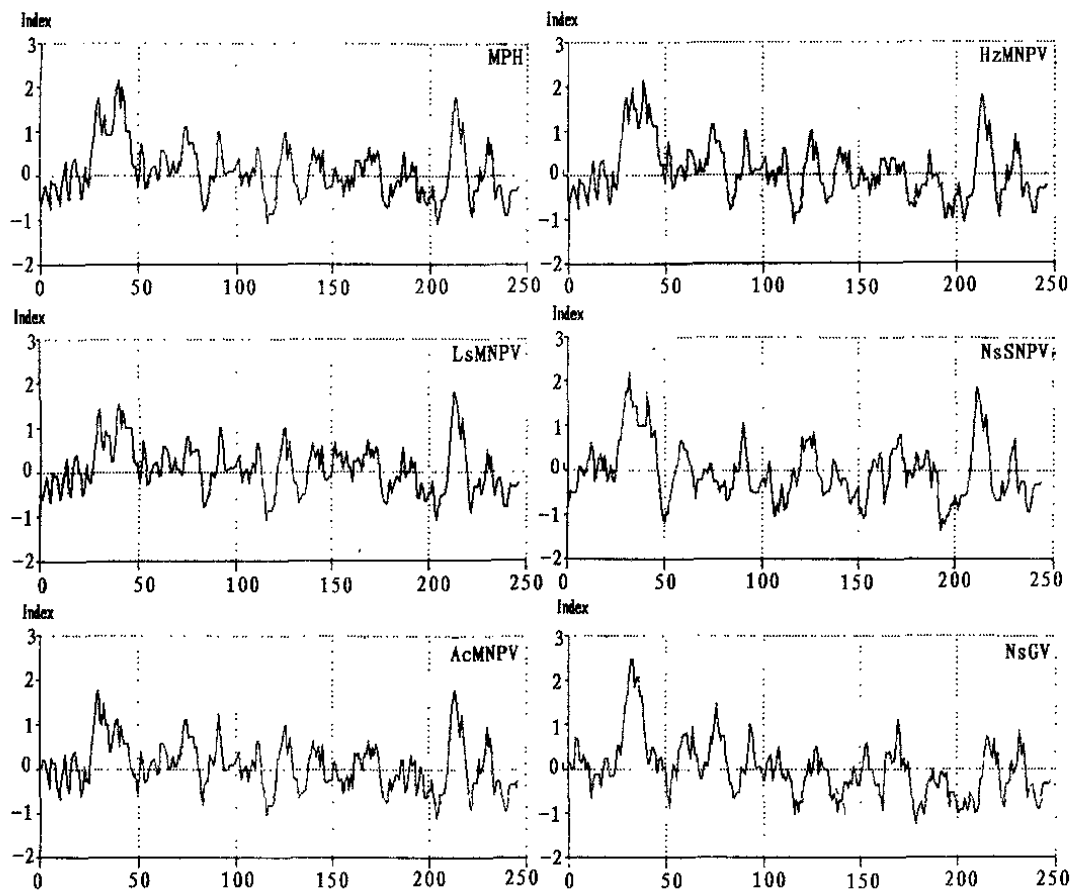


图 2 几种多角体蛋白的氨基酸亲水性分析曲线  
Fig.2 Hydrophilicity profile of amino acid of polyhedrins from selected baculoviruses

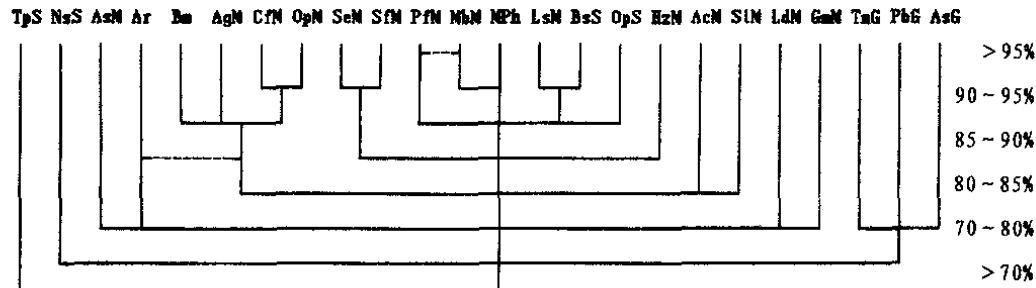


图 3 23 种杆状病毒相对模式多角体蛋白的进化关系  
Fig.3 Phylogenetic relationship among 23 baculoviruses based on the polyhedrins compared with MPh  
百分率为该层次上的最大同源性水平;——表示相互间具有一定的同源性;---表示与相邻一个或一组间具有较明显的同源性  
Percentage represents the range of maxium homology; ——indicates the homology in a range; --- indicates the special homology of one-to-one of groups

毒 AsGV 同样存在 30~50 高亲水区, NsSNPV 还有 210~240 高亲水区, 虽然有较大的数值变化, 但说明这两个亲水区不同程度的保守性。

2.3 蛋白质二级结构分析

根据推测(表 2): 在 AsMNPV 以上的多角体蛋白具有相似的二级结构; 有 6 个明显同源的  $\alpha$  螺旋区, 在 MPh 上位置为 36~48、68~76、134~157、193~204、215~220、231~235, 其中第一个  $\alpha$  螺旋区最典型; 还存在较多的  $\beta$  转角。





## 2.4 最大同源性分析及进化树

根据 24 种多角体蛋白氨基酸顺序的最大同源性(表 3)绘制的系统进化树(图 3)表明:与 MPh 氨基酸顺序有 70% 以上同源性的包括从 AsMNPV 到 GmMNPV 所有鳞翅目昆虫病毒的多角体蛋白,仅有 50% 以上同源性的为 NsSNPV 和 3 个 GV 的多角体蛋白,从双翅目昆虫病毒 TpSNPV 的 N 端分析结论可看出其远缘性。

## 3 讨论

从前,人们用 SDS-PAGE 电泳<sup>[28]</sup>和免疫学方法<sup>[29]</sup>来研究并鉴别不同的杆状病毒种或株,后来发现不同来源的包涵体蛋白间广泛存在抗原交叉反应,据分析这是包涵体蛋白结构的高度保守性所致<sup>[30]</sup>。在放弃了免疫学方法以后,多角体蛋白基因 DNA 和蛋白质氨基酸一级结构研究成为新的热点。本文从多角体蛋白氨基酸亲水性分析发现 2 个较明显的亲水区,位于 MPh 的 30~50 和 210~240,据认为该区域可能是多角体蛋白的主要抗原决定簇<sup>[25]</sup>。但是,根据氨基酸变异曲线分析,第一个亲水区 30~50 同时也是最显著的氨基酸变异区,其亲水数值的变化也较大。据分析这个区域可能被二级结构所掩盖而使不同的多角体蛋白间并不表现出不同的血清学性质,因此,作为更主要的抗原决定簇的则可能是第二亲水区即 210~240。多角体蛋白的一级结构是相当保守的,因而赋予了多角体蛋白的保守功能。大量存在的  $\beta$  转角说明多角体蛋白是高度卷曲的。有研究认为多角体蛋白高级结构的形成有赖于离子键、氢键和疏水作用,本文不讨论前两种作用。沿用亲水区即抗原决定簇的观点,高变的 30~50 亲水区应该使不同的多角体蛋白具有种属特异的血清学性质,但事实正好相反,可以认为这一区域在二级结构中是被掩盖在内部的,因此推测疏水作用在二级结构中是不起主导作用的。从 MPh 的一级结构看,3 个 Cys 位于 133、144 和 179,可能在折叠过程中形成了分子内二硫键,这是结构相对紧密的共价键。而且,相似的一级结构导致相似的二级结构,根据熵增原理范德华力也参与维系了分子内的高级结构。正是这种相似的致密结构使得二级以上的高级结构即多角体结构具有高度稳定性,不排除分子间二硫键在此间的作用。

有关另一大类昆虫包涵体病毒即属呼肠孤病毒科(Reoviridae)的质型多角体病毒(*Cytoplasmic polyhedrosis virus*, CPV)结构蛋白研究,Galinski 等<sup>[31]</sup>报道了几种 CPV 多角体蛋白基因核苷酸序列的相互比较及与同一宿主来源的 NPV 多角体蛋白氨基酸的比较,指出:分类第 5 型的 OpCPV、HaCPV 和 EsCPV,虽然来自不同宿主并分离自不同地区,其核苷酸序列具有相当高的同源性;比较 OpCPV 与第 1 型的 BmCPV,其多角体蛋白氨基酸顺序同源性不到 50%;将来源于同一宿主的 OpCPV 与 OpSNPV、OpMNPV 的多角体蛋白氨基酸顺序进行比较,没有发现前者与后者的明显同源性;而且,BmCPV、OpCPV 和 OpNPV 的氨基酸亲水性分析结果也是不同的;但是,在其一级结构中芳香族氨基酸(Phe、Trp、Tyr)和碱性氨基酸(Lys、Arg)的存在是比较保守的,这一结构特点赋予了所有 CPV 和 NPV 的多角体相同的基本功能,即形成包涵体保护病毒不受环境的破坏,但在昆虫肠的碱性环境又能被溶解释放出病毒粒子了。

本文直接用氨基酸顺序作为多角体蛋白同源性研究的指标,没有采用 DNA 序列资料,这是因为核苷酸的差异并不能完全代表蛋白质中氨基酸的差异。ocu 中的大部分氨基酸对密码子的使用是完全的,如 Leu 用了 TTA、TTG、CTT、CTC、CTA 和 CTG 等全部 6 个密码子。虽然存在普遍的密码子偏好(bias)<sup>[5,32]</sup>,如 Tyr 对 TAC 的选用比 TAT 高得多,但是不同多角体蛋白的同一个氨基酸,在 DNA 序列上常可能是不同的三个核苷酸。因此,氨基酸顺序比 DNA

序列更能准确地反映蛋白质的同源性及其他结构特征,特别是一级结构保守性所决定的高级结构及功能的相似性。用多角体蛋白氨基酸顺序作为杆状病毒系统进化的指标,比用 *ocu* 的调控区序列如 5'端的 12 bp 的恒致序列(consensus sequence 即 Rohrmann box)<sup>[3,19]</sup>更直接可靠。调控区的变异性影响转录或基因表达的程度,是能或不能、高水平或低水平、早期或晚期的问题,在相同的代谢系统中并不影响蛋白质的结构和功能。多角体蛋白的作为结构蛋白的功能是不变的,作为所谓正蛋白(orthologous protein),其变异可以反映杆状病毒的进化或宿主依赖性进化的某些信息<sup>[30]</sup>。本研究首先表明,多角体蛋白的一级结构是相当保守的,氨基酸顺序的部分变异也不致影响高级结构的相似性和功能的稳定性,因而这种变异和不变异正好符合种属依赖的特异性,成为杆状病毒系统发生、多样性变化和病毒宿主范围及其进化评价指标的最佳选择。

van Doorn 等<sup>[33]</sup>在研究丙型肝炎病毒(*Hepatitis C virus*, HCV)变异时,描绘了 HCV core/E<sub>1</sub> 区(nt477~924)核苷酸序列的进化树,并由此揭示了某地区 HCV 多个新亚型的存在。在最近有关分子进化的研究中,作为慢性疲劳综合征(chronic fatigue syndrome, CFS)病原体的内源性病毒(entroviruses)<sup>[34]</sup>、甲型流感病毒(*Influenza virus A*, H1N1)的血凝素(HA<sub>1</sub>)<sup>[35]</sup>,仍通过比较不同病毒某一段核苷酸或蛋白质氨基酸一级结构的同源性百分率,并绘制系统进化树评价其遗传学相关性。所谓宿主依赖性进化在乳多空病毒(Papovaviruses)的研究中就有过报道<sup>[36]</sup>,这种进化方式是在长期自然选择压力下形成的,造成了病毒感染的种间壁垒(species barrier),这个论点为作为杀虫剂的杆状病毒对远缘生物(包括人类)的安全性提供了理论依据<sup>[30]</sup>。有人<sup>[37]</sup>报道了感染完全不同的宿主如昆虫和对虾的杆状病毒的多角体蛋白存在着一定的同源性,这类个案可认为是生物多样性的特例。本文以多角体蛋白氨基酸顺序的最大同源性比较分析为基础推测的系统进化模型,完全支持了杆状病毒进化的宿主依赖性观点,表 2 中排列从 LsMNPV 到 AsMNPV 均来自鳞翅目昆虫,其多角体蛋白同源性高达 80% 以上,并且并无多粒包埋与单粒包埋之间的区别,而来自膜翅目昆虫的 NsSNPV 及来自双翅目昆虫的 TpSNPV 与之则少有同源性,GV 则明显地自成一族。本文提出的 MPh 氨基酸顺序在 70% 以上的同源性程度上涵盖了所有参与研究的来自鳞翅目昆虫病毒的多角体蛋白,说明 MPh 在一定程度上具有典型性。本文进一步认为,既然多角体蛋白的进化方式表现出如此明显的宿主依赖性,那么,昆虫杆状病毒多角体蛋白的同源性则可能作为研究昆虫本身系统进化和分类的分子生物学依据之一,也是病毒多样性研究的良好范例。

## 参 考 文 献

- 1 Luckow V A, Summers M D. Trends in the development of baculovirus expression vectors. *Bio-Technology*, 1988, **6**:47~55
- 2 Miller L K. Baculoviruses as gene expression vectors. *Annual Review of Microbiology*, 1988, **42**:177~199
- 3 Rohrmann G F. Polyhedrin structure. *Journal of General Virology*, 1986, **67**:1499~1513
- 4 刘子夜,齐义鹏. 粘虫核型多角体病毒包涵体蛋白的性质及其对几种肿瘤细胞生长的抑制作用. *中国病毒学*, 1993, **8**:357~365
- 5 王家旺,黄永秀,齐义鹏. 银染色测定粘虫核型多角体病毒多角体基因序列. *微生物学报*, 1995, **35**:155~160
- 6 Hoof van Iddekinge B J L, G E Smith, M D Summers. Nucleotide sequence of the polyhedrin gene of *Autographa californica* nuclear polyhedrosis virus. *Virology*, 1983, **131**:561~565
- 7 Ayres M D, S C Howard, R D Possee. The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. *Virology*, 1994, **202**:585~605
- 8 Zanotto de A P M, M J Sampaio, J E Maruniak. The *Anticarsia gemmatilis* nuclear polyhedrosis gene region: sequence analysis, gene product and structural comparisons. *Journal of General Virology*, 1992, **73**:1049~1056

- 9 胡建新, 丁红珍, 呈祥甫. 蓖麻蚕核型多角体病毒多角体蛋白基因的克隆和核苷酸序列分析. 遗传学报, 1993, **20**:300~303
- 10 Kozlov E A, N V Rodnin, L J Palchikovskaya. The amino acid sequence determination of a granulin and polyhedrin from two baculoviruses infecting *Agrotis segetum*. *Virology*, 1992, **189**:320~323
- 11 Iatrou K, K Ito, I I J Witkiewicz. Polyhedrin gene of *Bombyx mori* nuclear polyhedrosis virus. *Journal of Virology*, 1985, **54**:436~445
- 12 Hu ZH, Liu MF, Xie TE. Nucleotide sequence of the *Buzura suppressaria* single nucleocapsid nuclear polyhedrosis virus polyhedrin gene. *Journal of General Virology*, 1993, **74**:1617~1620
- 13 Hill J E, J Kuzio, P Faulkner. Nucleotide sequence of the p74 gene of a baculovirus pathogenic to the spruce budworm, *Choristoneura fumiferana* multicapsid nuclear polyhedrosis virus. *Biochim. Biophys. Acta*, 1993, **1172**:107~109
- 14 Cowan P, D M Bulach, D E Tribe. Nucleotide sequence of the polyhedrin gene region of *Helicoverpa zea* single nucleocapsid nuclear polyhedrosis virus; placement of the virus in Lepidopteran nuclear polyhedrosis virus group II. *Journal of General Virology*, 1994, **75**:3211~3218
- 15 Chang M T, C Lanner-Herrera, M Fikes. Nucleotide sequence of *Lymantria dispar* nuclear polyhedrosis virus polyhedrin gene. *Journal of Invertebrate Pathology*, 1989, **53**:241~246
- 16 Cameron I R, R D Possee. Conservation of polyhedrin gene promoter function between *Autographa californica* and *Mamestra brassicae* nuclear polyhedrosis viruses. *Virus Research*, 1989, **12**:183~199
- 17 Rohrmann G F. Baculovirus structural proteins. *Journal of General Virology*, 1992, **73**:749~761
- 18 Leisy D J, M Nesson, G S Beaudreau. Location and nucleotide sequence of the *Orgyia pseudotsugata* single nucleocapsid nuclear polyhedrosis virus polyhedrin gene. *Journal of General Virology*, 1986, **67**:1073~1079
- 19 Oakley R, I R Cameron, R D Possee. Analysis of transcription initiation in the *Panolis flammea* nuclear polyhedrosis virus polyhedrin gene. *Journal of General Virology*, 1989, **70**:769~775
- 20 Chakerian R, G F Rohrmann, G S Beaudreau. The nucleotide sequence of the *Pieris brassicae* granulosus virus granulin gene. *Journal of General Virology*, 1985, **66**:1263~1269
- 21 van Strien E A, D Zuidema, J M Vlask. Nucleotide sequence and transcription analysis of the polyhedrin gene of *Spodoptera exigua* nuclear polyhedrosis virus. *Journal of General Virology*, 1992, **73**:2813~2821
- 22 Gonzalez M A, G E Smith, M D Summers. Insertion of the SfMNP polyhedrin gene into an AcMNPV polyhedrin deletion mutant during viral infection. *Virology*, 1990, **170**:160~175
- 23 Croizier L, G Croizier. Nucleotide sequence of the polyhedrin gene of *Spodoptera littoralis* multiple nucleocapsid nuclear polyhedrosis virus. *Biochim. Biophys. Acta*, 1994, **1218**:457~459
- 24 Akiyoshi D, R Chakerian, G S Beaudreau. Cloning and sequencing of the granulin gene from the *Trichoplusia ni* granulosus. *Virology*, 1985, **141**:328~332
- 25 Hopp T P, K R Woods. Prediction of protein antigenic determinants from amino acid sequences. Proceedings of the National Academy of Science, U.S.A., 1981, **78**:3824~3828
- 26 Chou P Y, G D Fasman. Prediction of the secondary structure of proteins from their amino acid sequence. *Advances in Enzymology*, 1978, **47**:145~148
- 27 Saitou N, M Nei. The neighbour joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 1987, **4**:406~425
- 28 Summers M D, G E Smith. Baculovirus structural polypeptides. *Virology*, 1978, **84**:390~420
- 29 Rohrmann G F. Characterization of N-polyhedrin of two baculovirus strains pathogenic for *Orgyia pseudotsugata*. *Biochemistry*, 1977, **16**:1631~1634
- 30 Rohrmann G F, D J Leisy, G S Beaudreau. Identification, cloning, and R-loop mapping of the polyhedrin gene from the multicapsid nuclear polyhedrosis virus of *Orgyia pseudotsugata*. *Virology*, 1982, **121**:51~60
- 31 Galinski M S, Y Yu, G S Beaudreau. Analysis of the C-polyhedrin genes from different geographical isolates of a type 5 *Cytoplasmic polyhedrosis virus*. *Journal of General Virology*, 1994, **75**:1969~1974
- 32 Bennetzen J L, B D Hall. Codon selection in yeast. *Journal of Biological Chemistry*, 1982, **257**:3026~3031
- 33 van Doorn L J, G E M Kleter, W G V Quint. Sequence analysis of *Hepatitis C virus* genotypes 1 to 5 reveals multiple novel subtypes in the *Benelux* countries. *Journal of Virology*, 1995, **69**:3893~3896
- 34 Galbraith D N, C Nairn, G B Clements. Phylogenetic analysis of short enteroviral sequences from patients with chronic fatigue syndrome. *Journal of General Virology*, 1995, **76**:1701~1707
- 35 Pyhala R, N I Koneu, L Kinnuneu. Evolution of the HA<sub>1</sub> domain of human *Influenza A* (H1N1) virus; loss of glycosylation sites and occurrence of herald and conserved strain. *Journal of General Virology*, 1995, **76**:205~210
- 36 Soeda E, T Maruyama, B E Griffin. Host-dependent evolution of three *Papova* viruses. *Nature*, 1980, **285**:165~167
- 37 Vickers J E, R J G Lester, J M Pemberton. Evidence for homology between polyhedrin genes of baculoviruses of a prawn and an insect. *Journal of Invertebrate Pathology*, 1994, **63**:207~208