

·研究简报·

WSSV-VAP1 基因克隆与序列分析

刘庆慧^{1,2}, 黄 健², 韩文君², 王清印²

(1. 中国海洋大学, 山东 青岛 266003; 2. 中国水产科学研究院 黄海水产研究所, 山东 青岛 266071)

摘要:根据已测序的 WSSV-VAP1 全基因序列设计 1 对引物, 经 PCR 扩增及琼脂糖凝胶电泳分析, 扩增出约 838 bp 的特定条带, 将其回收后克隆入 pGAPZα-A 载体中, 并进行序列测定与分析。结果表明, 扩增序列与 GeneBank 登录序列 (GenBank No. AF411634) 核苷酸同源性 100%。序列分析表明, WSSV-VAP1 基因编码蛋白无跨膜区域, 有多个蛋白激酶 C、酪蛋白激酶 II 磷酸化位点和糖基化位点, 并含有 RGD 序列, 预示其为 WSSV 侵染中的一个重要蛋白。

关键词: WSSV-VAP1; 基因; 克隆; 序列分析; 中国明对虾

中图分类号: S945.46 **文献标识码:** A **文章编号:** 1005-8737-(2005)03-0352-05

目前对虾白斑综合症病毒 (White spot syndrome virus, WSSV) 的病原学、流行病学、诊断学、感染机理等研究已深入到了分子水平^[1-4]。WSSV 的基因组全序列已先后被 3 个研究小组测定并在 GenBank 上公布, 已确定 WSSV 全基因 (约 300 kb) 含有 185 个开放阅读框^[2-4]。随着 WSSV 全基因公布, 推进了 WSSV 基因结构与功能的相关研究。由于 WSSV 基因组许多基因的功能是未知的, 因此对该病毒基因功能的预测大部分来自序列分析的结果。在 WSSV 的基因组内有許多基因与真核生物同源, 如 DNA 多聚酶、核糖核酸还原酶、蛋白激酶等, 这些基于序列分析的基因功能的预测结果为基因功能的研究奠定了基础。目前有 10 余个基因的功能得到初步确定, 有 36 个基因与已知的蛋白具有同源性, 尚有许多基因的功能未知。

生物信息学手段在现代生物学领域中发挥着越来越重要的作用, 计算机分析结果可以用来推测蛋白的功能, 辅助进一步的实验设计。本研究借助生物信息学手段, 对克隆 WSSV-VAP1 基因推测的蛋白进行蛋白理化特性、蛋白序列 (Motif) 特性以及二级结构预测, 为进一步研究其功能提供依据。

1 材料与方法

1.1 引物设计

根据已测序的 WSSV-VAP1 全基因序列, 采用 Primer 软件, 设计 WSSV-VAP1 正向和反向引物, 根据 pGAPZα-A 多克隆位点的特点以及 WSSV-VAP1 基因酶切特性, 分别在正向引物引入限制性内切酶 *EcoRI* 识别位点 (GAATTC), 在反向引物引入限制性内切酶 *XbaI* 识别位点 (TCTAGA), 合成引物 P₁: 5'-GGTGAATTTCATGGCGTAAACTTG-3', P₂: 5'-CGGCTCTAGAGGTGTCCAACAATTTAAAA-

3', 引物由宝生物工程 (大连) 有限公司合成。

1.2 PCR 扩增

从患白斑综合症的中国明对虾 (*Penaeus monodon*) 中提取并纯化 WSSV DNA, 作为 PCR 反应的模板。采用紫外分光光度法和琼脂糖凝胶电泳法测定模板 DNA 浓度后, 按照常规方法进行 PCR 反应。反应体系如下: 反应体积为 25 μL, 其中 10× *phi* DNA 聚合酶 buffer 含有 MgCl₂ (20 mmol/L), dATP, dTTP, dCTP, dGTP 各 0.25 mmol/L, 正向和反向引物浓度各为 50 pmol, *phi* Tag DNA 聚合酶 5 单位/μL。置 PCR 仪上按以下程序进行梯度 PCR 扩增: 94℃ 5 min 变性后进入循环 94℃ 1 min, 梯度 50.2℃ (51℃, 51.8℃, 52.6℃, 53.4℃, 54.2℃, 55℃, 55.8℃, 56.6℃, 57.4℃, 58.2℃, 59℃) 1 min, 72℃ 1 min, 32 个循环; 最后 72℃ 延伸 5 min, PCR 结束后, 进行琼脂糖凝胶电泳检测, 用 Winstar 小量胶回收试剂盒回收 PCR 产物。

1.3 重组质粒的限制性内切酶分析及序列测定

WSSV-VAP1 基因的 PCR 扩增片段经 *EcoRI* 和 *XbaI* 双酶切, 与载体 pGAPZα-A 经 *EcoRI* 和 *XbaI* 双酶切后进行连接, 将载体 pGAPZα-A 与 VAP1 的 PCR 产物连接产物转化 *E. coli* TOP10F' 感受态细胞, 涂布于含 25 μg/mL ZeocinTM 的 LB 低盐平板上, 待菌落长出后, 挑取白色菌落进行 PCR 筛选出重组菌。

将重组菌接种于含 25 μg/mL ZeocinTM 的 LB 低盐培养基上, 37℃ 振荡培养过夜, 抽提纯化重组质粒, 用限制性内切酶 *EcoRI*、*XbaI* 进行双酶消化, 取 5 μL 进行 0.8% 的琼脂糖凝胶电泳。

1.4 目的片段的序列测定及分析

该目的基因片段委托上海博亚生物工程有限公司进行序列测定, 并将测序结果与 VAP1 基因序列进行比较。

1.5 克隆片段的生物信息学分析

以克隆并测序得到的基因编码的氨基酸序列作为原始

收稿日期: 2004-07-27; 修订日期: 2004-10-25。

基金项目: 国家自然科学基金项目 (30271020); 国家重点基础研究项目 (G1999012002); 中国水产科学研究院科研基金项目 (2003-4-1)。

作者简介: 刘庆慧 (1962-), 女, 副研究员, 博士生, 主要从事生物化学与分子生物学方面的研究。E-mail: liqh@ysfri.ac.cn

通讯作者: 黄 健。E-mail: aquadist@public.qd.sd.cn

数据,选择生物学分析软件或网络服务器,有针对性的进行分析。对该基因编码的氨基酸结构特点、理化性质等进行分析。一级结构根据核酸序列采用 Omiga2.0 软件推导出其氨基酸序列,进行氨基酸组成、等当点及化学特性的分析。疏水性分析采用 Omiga2.0 软件,将氨基酸序列输入软件中,得到整个蛋白氨基酸序列的疏水分布图。跨膜特性采用在线分析软件 (<http://www.ch.embnet.org/software/TMPRED-form.html>)。蛋白序列模式采用 Omiga2.0 软件和 Prosite 在线分析软件。二级结构及特殊结构预测,采用 Omiga2.0 软件对 α -螺旋(α -helix)、 β -折叠(β -sheet)、 β -转角(β -turn)等主要结构进行预测。

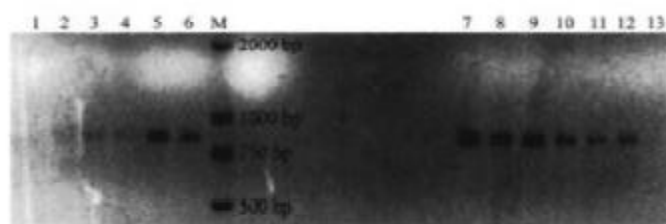


图1 梯度 PCR 扩增结果
Fig.1 Amplification by Gradient PCR

2 结果

2.1 PCR 扩增 WSSV-VAP1 基因的结果

根据 WSSV-VAP1 核苷酸序列,预测 PCR 所扩增的基因片段大小应为 838 bp 左右。梯度 PCR 结果(图 1)表明,当复性温度为 53.4℃、54.2℃、55℃、55.8℃、56.6℃、57.4℃、58.2℃、59℃ 时,PCR 扩增 WSSV-VAP1 基因的效果较好,选定 55℃ 为复性温度,进行大量扩增,将 PCR 反应产物进行琼脂糖凝胶电泳,结果显示,扩增出的片段大小与预测大小一致(图 2)。

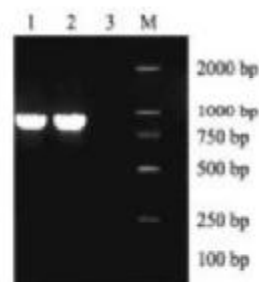


图2 PCR 扩增 WSSV-VAP1
Fig.2 Amplification WSSV-VAP1

2.2 pGAPZa-A-VAP1 的重组体筛选与测序

将 PCR 筛选阳性克隆,少量抽提重组质粒,重组质粒双酶切后产生一条分子量为 3 200 kb 左右的质粒 DNA 带与一条分子量为 830 bp 左右的 WSSV-VAP1 片段,进一步证明重

组质粒是阳性。将阳性克隆进行测序,测序结果表明所有碱基与测序的 WSSV-VAP1 基因序列完全相同(图 3),证实克隆到载体 pGAPZaA 上的外源 DNA 片段确为 VAP1 基因。

1	ctcggatga	acttggatga	tggtcttgig	aatatcaca	acagagatga	agatcttaca
			K D E D	L T K	L V S E	A I K
61	aaactcgat	ccgagccat	aaagcggcga	gciaaactig	tatttgacac	taaaatcaa
	R R A	K T V F	D T K	N Q A	G F D M	R R Q
121	gcaggggttg	acatggagcg	tcaggttgaa	gcigcattat	atgagcgaal	atccaaaagc
	V E A	A L Y E	A I S	K K K	E K A I	K A F
181	aaagaaaagc	ccatnaagcc	attcgatagc	ctctatcang	aaagcggcga	tgaaattaca
	D E L	I Q E R	G D E	I T P	L T T M	Q Y E
241	cttttgacta	caatgcagta	tgaagagatga	gtaacccgta	caatbaactcc	ctcatggacg
	E W V	N R T I	T P S	L T T	E N L L	G D V
301	actgaaatit	tattaggtga	tggttgagcac	gccgattiti	tactggaccg	aatgacaccc
	E H A	D F L L	D R M	T P V	S E E D	I E G
361	gttagcgagc	aaatattga	agattitgci	gccttctact	tttagggaggt	atcagattca
	F A A	S T F K	E V S	D S K	T A T V	I V K
421	aaactgcga	cagtcatagt	taagggcagat	tgtgnaacgc	ggagatctcg	tgaaattgat
	A D C	E T G D	I D E	V Y N	L A P S	F G V
481	aatctgcac	caicattggc	cgtcactcaa	gaatttanaa	tatataggtc	aaacatttcc
	T Q E	I K I Y	R S N	N S	S E L D	N V A
541	tcggaaattg	atantgtcgc	agattcttgc	catatttata	aaatttctgc	aacagatgc
	D S F	H I Y K	I S A	T D S	D S G N	T K K
601	gcacgtacga	atctnaaaa	attcttgat	ggagtttagga	atnaaaagc	agattttacg
	L L Y	G L R N	K K A	G Y T	C L C R	I F A
661	tgttitgta	gaatttttgc	agaatttgaa	tcagatggga	ttatggccaa	tacaaatgc
	E I E	S D G I	M A N	T N I	G V A E	N N R
721	gggttcgcig	aaacaacng	agatgaatit	gatgaanacg	aaagaggtga	atgtggttit
	D E I	D E N E	E G K	Y G F	L I P K	Q P A
781	taantacca	aacaccagc	tggtgcnaaa	tigatcatct	acttctttit	aaattgttgg
	G A K	L I I Y	F F L	N C W	T	
841	acgtga					

图3 WSSV-VAP1 基因核苷酸序列和推导的氨基酸序列
Fig.3 Nucleotide sequence and predicted amino acid sequence

2.3 生物信息学分析

以 Omiga2.0 软件对 WSSV-VAP1 基因编码氨基酸序列的基本特征进行分析,结果显示(表1),其编码蛋白分子量为 31.56 kD,等当点(PI)4.422,氨基酸组成中带正电荷氨基酸(Arg、Lys)25 个,占 8.89%,带负电荷氨基酸(Asp、Glu)49 个,占 17.44%,整个蛋白带负电荷,疏水性氨基酸 116 个,占 41.28%,亲水性氨基酸 127 个,占 45.19%。疏水性分析结果显示(图4),N 末端氨基酸以亲水性为主,其疏水性区域位于 C 末端。

蛋白位点和序列模式分析表明(表2),其编码蛋白含有 N-糖基化位点 3 个,酪氨酸硫酸化位点 1 个,蛋白激酶 C 磷酸化位点 3 个,酪蛋白激酶 II 磷酸化位点 6 个,酪氨酸激酶磷酸化位点 1 个,肉豆蔻基(MYRISTYL)位点 2 个,NUCLEAR 位点 1 个,黏附蛋白 RGD 位点 1 个。图 5 示出各序列模式(motif)在氨基酸全序列中的位置,从图中可以看出,黏附蛋白 RGD 序列附近各有一个 N-糖基化位点和蛋白激酶 C 磷酸化位点,该 2 位点的糖基化和磷酸化对 RGD 序列行使黏附宿主细胞,进而进入细胞具有重要作用。跨膜性分析表明(图6),WSSV-VAP1 基因编码蛋白无明显跨膜区域。

二级结构分析显示(图7),WSSV-VAP1 基因编码的蛋白主要以 α -螺旋和 β -折叠为主,间或有 β -转角。

表 1 WSSV-VAP1 基因编码蛋白氨基酸含量

Tab.1 Amino acid content of WSSV-VAP1 gene encoded protein

氨基酸种类 Amino acid	所含个数 No.	百分含量/% Percentage
丙氨酸 Ala(A)*	23	8.19
精氨酸 Arg(R)≠	11	3.91
胱氨酸 Cys	4	1.42
甘氨酸 Gly(G)	14	4.98
组氨酸 His(H)*	2	0.71
异亮氨酸 Ile(I)*	22	7.83
亮氨酸 Leu(L)*	21	7.47
赖氨酸 Lys(K)*	22	7.83
蛋氨酸 Met(M)*	5	1.78
苯丙氨酸 Phe(F)*	12	4.27
脯氨酸 Pro(P)*	6	2.14
丝氨酸 Ser(S)*	16	5.69
苏氨酸 Thr(T)*	21	7.47
色氨酸 Trp(W)*	2	0.71
酪氨酸 Tyr(Y)*	9	3.20
缬氨酸 Val(V)*	16	5.69
天冬酰胺 Asn(N)	20	7.12
天冬氨酸 Asp(D)*	22	7.83
谷氨酰胺 Gln(Q)*	6	2.14
谷氨酸 Glu(E)*	27	9.62
合计 Total	281	100

注: * 疏水氨基酸; ≠ 亲水氨基酸

Note: * Hydrophobicity; ≠ Hydrophilic

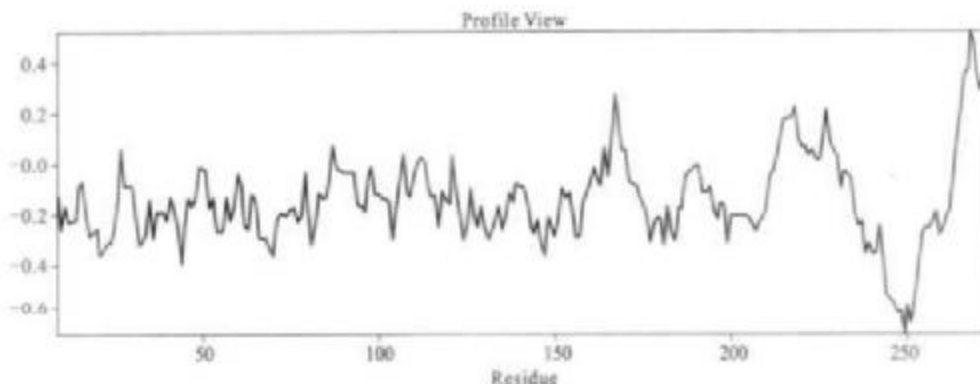


图4 WSSV-VAP1 编码蛋白疏水特性

Fig.4 Hydrophobicity of WSSV-VAP1 gene encoded protein

3 讨论

利用生物信息学分析蛋白的序列模式对鉴别未知蛋白的功能及其归属具有重要的意义。以 OMEGA2.0 等生物学软件对 VAP1 基因进行分析,结果表明该基因编码蛋白质亲水性区域占优势,是一个亲水性蛋白质。利用 Omiga 2.0 软件对该蛋白质可能的修饰位点进行分析,发现该蛋白有多个位点可以发生磷酸化修饰,个别位点可以发生糖基化修饰。蛋白质的磷酸化和脱磷酸化在细胞信号传导中起重要作用,

磷酸化、糖基化位点可能是 WSSV-VAP1 蛋白完成其生理功能的结构基础。特别是该蛋白序列中酪蛋白激酶 II 磷酸化位点有 6 个,表明酪蛋白激酶 II (CK II) 对该蛋白生物功能的发挥具有重要的调节作用。序列分析表明该蛋白在 75~77 位置含有黏附蛋白基序 RGD,大量的研究表明,许多病毒和病原体通过黏附于宿主细胞而进一步完成其复制循环,黏附蛋白中的 RGD 基序是蛋白相互作用的必需基团,其在调节细胞识别及病原体的感染过程中起非常重要的作用^[9-11]。因此该蛋白的生物学功能值得关注。

表2 WSSV-VAP1 编码蛋白序列模式分析结果

Tab.2 Analysis motif of predicted amino acid

模块名称 Motif	特征序列 Specific sequence	位置 Position
ASN_GLYCOSYLATION(N-糖基化位点)	NRTI	92-95
ASN_GLYCOSYLATION(N-糖基化位点)	NNSS	178-181
ASN_GLYCOSYLATION(N-糖基化位点)	NNSS	179-182
SULFATION(酪氨酸硫酸化位点)	tglddevYnlpsfg	153-167
PKC_PHOSPHO_SITE(蛋白激酶C磷酸化位点)	SskK	58-60
PKC_PHOSPHO_SITE(蛋白激酶C磷酸化位点)	iK	133-135
PKC_PHOSPHO_SITE(蛋白激酶C磷酸化位点)	TskK	205-207
CK2_PHOSPHO_SITE(酪蛋白激酶 II 磷酸化位点)	TvfiD	33-36
CK2_PHOSPHO_SITE(酪蛋白激酶 II 磷酸化位点)	SeeD	122-125
CK2_PHOSPHO_SITE(酪蛋白激酶 II 磷酸化位点)	Tfke	133-136
CK2_PHOSPHO_SITE(酪蛋白激酶 II 磷酸化位点)	SeiD	181-184
CK2_PHOSPHO_SITE(酪蛋白激酶 II 磷酸化位点)	SotD	196-199
CK2_PHOSPHO_SITE(酪蛋白激酶 II 磷酸化位点)	TdsD	198-201
TYR_PHOSPHO_SITE(酪氨酸激酶磷酸化位点)	RrqpEaaiY	46-54
MYRISTYL(N-)	GlmaNT	233-238
MYRISTYL(N-)	GVaeNN	241-246
NUCLEAR	RRtpenalycaiskkke	46-62
RGD	RGD	75-77

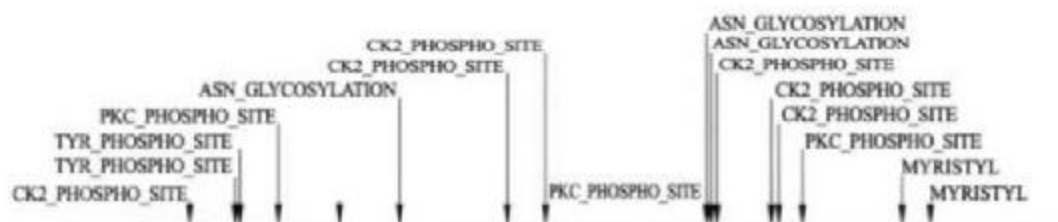


图5 WSSV-VAP1 编码蛋白模式位置图

Fig.5 Motif pattern of predicted protein

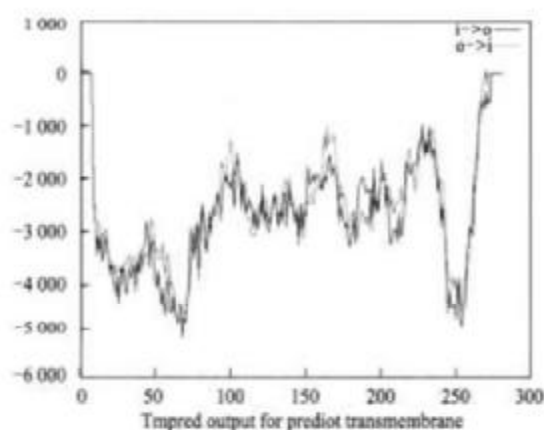


图6 WSSV-VAP1 编码蛋白跨膜分析

Fig.6 Transmembrane analysis of WSSV-VAP1

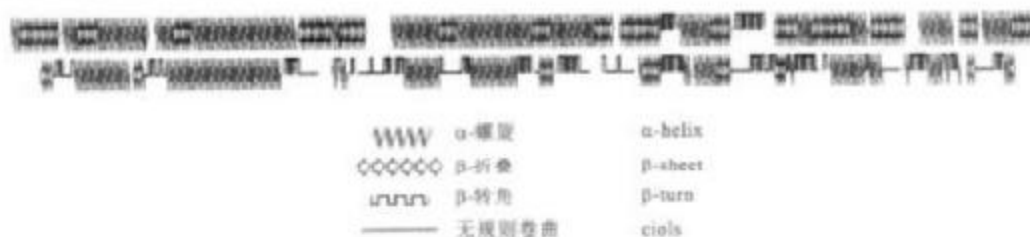


图7 WSSV-VAP1 二级结构

Fig.7 Second structure of WSSV-VAP1

参考文献:

- [1] 黄 健,宋晓玲,于 佳,等.杆状病毒性的皮下及造血组织坏死——对虾暴发性流行病的病原和病理学[J].海洋水产研究,1995,16(1):1-10.
- [2] 黄 健,蔡生力,宋晓玲,等.对虾暴发性流行病病原的人工感染研究[J].海洋水产研究,1995,16(1):51-57.
- [3] 战文斌,俞开康,孟庆星.中国对虾(*Penaeus chinensis*)杆状病毒的研究[J].中国水产科学,1995,2(3):22-28.
- [4] Wang C H, Lo C F, Liu J H, et al. Purification and genomic analysis of baculovirus associated with white spot syndrome (WSBV) of *Penaeus monodon* [J]. Dis Aquat Org, 1995, 23: 239-242.
- [5] Chou H Y, Huang C Y, Wang C H, et al. Pathogenicity of a baculovirus infection causing white spot syndrome in cultured penaeid shrimp in Taiwan [J]. Dis Aquat Org, 1995, 23: 165-173.
- [6] Lo C F, Liu J H, Ho C H, et al. Detection of baculovirus associated with white spot syndrome (WSBV) in penaeid shrimps using polymerase chain reaction [J]. Dis Aquat Org, 1997, 25: 133-141.
- [7] Yang F, He J, Lin X, et al. Complete genome sequence of the shrimp white spot baculiform virus [J]. Journal of Virology, 2001, 75(23): 11 811-11 820.
- [8] Tsai M F, Yu H T, Tseng H F. Identification and characterization of a shrimp white spot syndrome virus (WSSV) gene that encodes a novel chimeric polypeptide of cellular-type thymidine kinase and thymidylate kinase [J]. Virology, 2000, 277(1): 100-110.
- [9] van Hulten M C, Wittereids J, Peter S. The white spot syndrome virus DNA genome sequence [J]. Virology, 2001, 286: 7-22.
- [10] Boonakiat Y, Hughes P J, Ghazi F & Starway G. Arginine-glycine-aspartic acid motif is critical for human parvovirus 1 entry [J]. Journal of Virology, 2001, 75: 10 000-10 004.
- [11] Mateu M G, Valero M L, Andreu D & Domingo E. Systematic replacement of amino acid residues within an Arg-Gly-Asp containing loop of foot-and-mouth disease virus and effect on cell recognition [J]. Journal of Biological Chemistry, 1996, 271: 12 814-12 819.

Clone and sequence analysis of white spot syndrome virus (WSSV)-VAP1 gene

LIU Qing-hui^{1,2}, HUANG Jie², HAN Wen-jun², WANG Qing-yin²

(1. Ocean University of China, Qingdao 266003, China; 2. Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao 266071, China)

Abstract: PCR primer was designed according to WSSV-VAP1 nucleotide sequence from GeneBank. After PCR reaction and electrophoresis analysis, a specific 838 bp DNA band was amplified and cloned into vector of pGAPZa-A. Nucleotide sequence analysis showed the sequence amplified was 100% identity with the sequence from GeneBank (GenBank No. AF411634) of WSSV-VAP1. The predicted protein from nucleotide sequence of WSSV-VAP1 had no notable transmembrane region and it contained much PKC-PHOSPHO-SITE, CK2-PHOSPHO-SITE and ASNGLYCOSYLATION-SITE. The cell attachment RGD motif in WSSV-VAP1 indicated that this protein might play an important role in WSSV infectivity.

Key words: WSSV-VAP1; gene; clone; sequence analysis; *Penaeus chinensis*

Corresponding author: HUANG Jie. E-mail: aquadis@public.sd.cn