

GII.4 型诺如病毒 VP2 蛋白生物信息学分析

林鹏, 黄玉雯, 李学鹏, 励建荣*

(渤海大学食品科学与工程学院, 渤海大学海洋研究院, 辽宁省食品安全重点实验室, 生鲜农产品贮藏加工及安全控制技术国家地方联合工程研究中心, 辽宁锦州 121013)

摘要: 为了解诺如病毒 VP2 蛋白的生物学特征, 该研究以当前诺如病毒 GII.4 型流行毒株 VP2 蛋白为研究对象, 应用生物信息学软件对 VP2 蛋白的理化性质、磷酸化位点、糖基化位点、跨膜区、信号肽、二级结构、三级结构、抗原决定簇和 B 细胞抗原表位进行预测分析。结果表明, 诺如病毒流行毒株 VP2 蛋白为稳定的亲水性蛋白, 平均等电点为 10.47、吸水系数为 -0.47、不稳定系数为 47.91, 碱性氨基酸约占含量的 11.60%; 该蛋白含有约 43 个潜在的磷酸化修饰位点和 2 个潜在的糖基化修饰位点; 大多数毒株不存在跨膜区和信号肽, 然而 GI.3 型诺如病毒 VP2 蛋白存在跨膜区和信号肽; 诺如病毒 VP2 蛋白二级结构中 α -螺旋平均占比 38.17%、延伸链平均占比 7.45%、 β -折叠平均占比 6.25%、无规则卷曲平均占比 48.13%; 三级结构预测模型的蛋白覆盖率为 56.70%; 平均存在 8 个潜在的蛋白质抗原决定簇和 25 个 B 细胞抗原表位。该研究运用生物信息学分析方法预测当前 GII.4 型诺如病毒流行毒株 VP2 蛋白的理化性质和结构等方面, 为进一步深入研究诺如病毒 VP2 蛋白功能奠定了理论基础。

关键词: 诺如病毒; VP2 蛋白; 生物信息学

文章编号: 1673-9078(2023)09-42-51

DOI: 10.13982/j.mfst.1673-9078.2023.9.1140

Bioinformatics Analysis of the VP2 Protein of GII.4 Norovirus

LIN Peng, HUANG Yuwen, LI Xuepeng, LI Jianrong*

(College of Food Science and Engineering, Institute of Ocean, Bohai University, Food Safety Key Laboratory of Liaoning Province, National Local Joint Engineering Research Center for Storage, Processing and Safety Control Technology of Fresh Agricultural and Aquatic Products, Jinzhou 121013, China)

Abstract: In order to understand the biological characteristics of the VP2 protein of norovirus (NoV), the VP2 protein of current NoV epidemic strains was used as the research object in this study. Bioinformatics softwares were used to predict and analyze the physico-chemical properties, phosphorylation sites, glycosylation sites, transmembrane regions, signal peptides, secondary structures, tertiary structure, antigenic determinants and B cell epitopes of NoV VP2 protein. The results showed that the VP2 protein of epidemic NoV strain was a stable hydrophilic protein with an average isoelectric point of 10.47, a water absorption coefficient of -0.47, and an instability coefficient of 47.91, with basic amino acids accounting for about 11.60%. The VP2 protein contains about 43 potential phosphorylation modification sites and 2 potential glycosylation modification sites. Most strains do not contain transmembrane domains and signal peptides, however, the GI.3 VP2 protein has the transmembrane domain and signal peptide. The average proportion of α -helix, extended chain, β -fold and irregular coil were 38.17%, 7.45%, 6.25% and 48.13%, respectively. The protein coverage of the tertiary structure prediction model was 56.70%. There were 8 potential protein antigenic determinants and 25 B cell epitopes on average. In this study, bioinformatics analysis methods were used to predict the physicochemical properties and structures of the VP2 protein of the currently epidemic NoV GII.4 strains, which has laid a theoretical foundation for further research on the function of the NoV VP2 protein.

Key words: norovirus; VP2 protein; bioinformatics

引文格式:

林鹏, 黄玉雯, 李学鹏, 等. GII.4 型诺如病毒 VP2 蛋白生物信息学分析[J]. 现代食品科技, 2023, 39(9): 42-51

LIN Peng, HUANG Yuwen, LI Xuepeng, et al. Bioinformatics analysis of the VP2 protein of GII.4 norovirus [J]. Modern Food Science and Technology, 2023, 39(9): 42-51

收稿日期: 2022-09-08

基金项目: 国家重点研发计划重点专项 (2019YFD0901702)

作者简介: 林鹏 (1990-), 男, 博士, 讲师, 研究方向: 食品质量与安全, E-mail: lphs1002@163.com

通讯作者: 励建荣 (1964-), 男, 博士, 教授, 研究方向: 水产品和果蔬贮藏加工及质量安全控制, E-mail: lijrg491@163.com

诺如病毒 (Norovirus, NoV) 曾称诺瓦克病毒 (Norwalk Viruses), 隶属于杯状病毒科 (Caliciviridae), 诺如病毒属 (Norovirus) 成员, 是一种单股正链 RNA 病毒。1968 年美国俄亥俄州诺沃克市暴发了 NoV 感染, 随后世界各地相继发现该病毒的其它亚型^[1]。NoV 基因组包含三个开放阅读框 (Open Reading Frames, ORFs), 其中 ORF1 编码非结构蛋白, ORF2 编码衣壳蛋白 (VP1), ORF3 编码次要结构蛋白 (VP2)^[2]。近年来由于新的流行毒株出现, 根据 VP1 蛋白的氨基酸序列和 ORF1 的 RNA 依赖 RNA 聚合酶 (RdRp) 区的核苷酸多样性, 分为不同的基因群和 P 群, 进一步将 NoV 分为 10 个基因群 (GI-GX) 和不同的 P 型, 其中 GI、GII、GIV、GVIII 和 GIX 基因群主要感染人类, 被通称为人源诺如病毒 (Human Norovirus, HuNoV)^[3]。根据流行病学研究表示, HuNoV 是引起流行性、自限性急性胃肠炎的主要病原体之一, 约占所有急性胃肠炎病例的五分之一^[4]。对健康的成人来说, NoV 引起一种急性自限性的胃肠道感染; 但在老年人、儿童和免疫功能低下的患者中 NoV 则会造成较高的发病率和死亡率^[5]。NoV 感染的主要传播为粪便-口途径, 主要由于食用含有 NoV 的食品, 以生食贝类或者食用加热不彻底的贝类最为常见^[6]。全年均可发生, 但多发生在秋冬季节^[7]。自 2013 年以来, 我国每年因 NoV 引起的急性胃肠炎病例逐年递增, 其主要临床表现为腹泻和呕吐等^[8]。NoV 具有较强的感染性, 且目前暂无特效药物。因此, 由 NoV 引起的食品安全问题亟待关注。

NoV 目前主要的流行毒株包括 GI.3、GII.2、GII.3、GII.4、GII.13 和 GII.17 型, 这些基因型很容易产生新的突变, 且新的突变型通常具有更强的传染性和变异性^[9]。在过去的 30 年中, GII.4 型占据主导地位, 导致了大约 70% NoV 的暴发^[10]。有证据表明, 每 2~3 年 GII.4 型都会出现一种新的变异株^[11]。目前 GII.4 型主要有六种流行变异株, 分别为 Grimsby 1995、Farmington Hills 2002、Hunter 2004、Den Haag 2006b、New Orleans 2009 和 Sydney 2012^[12]。

VP2 蛋白作为 NoV 次要结构蛋白, 可以和 VP1 蛋白组成 NoVs 病毒颗粒^[13]; 然而, VP2 蛋白对于病毒颗粒不是必需的, 主要对病毒的稳定性和衣壳化起到非常重要的作用^[14,15]。有研究表明, VP2 蛋白和 VP1 蛋白相互作用可促进衣壳蛋白的表达, 进而影响 NoV 的复制^[16,17]。NoV VP2 蛋白具有较高的突变性, 在不同毒株的 NoV 可能起到不同的生物学作用^[18]。

NoV 已被公认为是最主要的病毒性胃肠病原体之一, 影响着全世界各个年龄段的数亿人, 造成每年

全球直接经济损失为 42 亿美元, 社会经济损失 603 亿美元^[19,20]。因此, 研究 NoV 具有非常重要的现实意义。但目前关于 NoV 流行毒株 VP2 蛋白研究相对较少, 对其生物学作用了解不够全面, 不能对 NoV VP2 蛋白的理化性质、蛋白结构等提供有效的参考^[21]。本研究采用生物信息学对 NoV 流行毒株 VP2 蛋白结构与功能进行预测, 为进一步研究 VP2 蛋白功能及开发亚单位疫苗提供理论支持, 以此为解决由 NoV 引起的食品安全问题提供新的思路。

1 材料与方法

1.1 毒株序列

GII.4 LN-2021-18 毒株 VP2 基因序列由本实验室分离病毒测序。

1.2 参考毒株序列

综上所述根据目前主要的 NoV 流行毒株, 选取以下 12 株进行分析^[9-12]。NoV GI.3 Nashville 2016 毒株 (GenBank: MK073893.1); NoV GII.2 Japan 2020 毒株 (GenBank: LC646332.1); NoV GII.3 Ann Arbor 2016 毒株 (GenBank: MK073886.1); NoV GII.4 Tokyo 2021 毒株 (GenBank: LC699541.1); NoV GII.4 Sydney 2016 毒株 (GenBank: MK073885.1); NoV GII.4 Sydney 2012 毒株 (GenBank: KU678201.1); NoV GII.4 Johannesburg 2009 毒株 (GenBank: KP784691.1); NoV GII.4 Den Haag 2006b 毒株 (GenBank: KM245071.1); NoV GII.4 Hong Kong 2004 毒株 (GenBank: HM802551.1); NoV GII.4 Houston 2002 毒株 (GenBank: EU310927.1); NoV GII.13 USA 2010 毒株 (GenBank: JN899242.1); NoV GII.17 CN 2014 毒株 (GenBank: KU757047.1)。

1.3 实验方法

1.3.1 VP2 蛋白序列分析

使用 MEGA 6.05 软件对 VP2 蛋白的氨基酸序列进行分析比对, 采用 Neighbor-joining 法构建 VP2 蛋白进化树。

1.3.2 VP2 蛋白理化性质的预测

使用在线软件 ProtParam Tool (<https://web.expasy.org/protparam/>) 预测 VP2 蛋白的带电荷氨基酸数量、不稳定系数、平均吸水系数、相对分子质量、氨基酸组成及等电点; 进一步使用 ProtScale 在线软件 (<https://web.expasy.org/protscale/>) 预测 VP2 蛋白的亲疏水性。

1.3.3 VP2 蛋白磷酸化位点和糖基化位点的预测

在线软件 NetPhos 3.1 Server (<http://www.cbs.dtu.dk/services/NetPhos/>) 预测 VP2 蛋白的潜在磷酸化位点；在线软件 NetOGlyc 1.0 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>) 预测 VP2 蛋白的潜在糖基化位点。

1.3.4 VP2 蛋白跨膜区及信号肽的预测

在线软件 TMHMM Server v.2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) 预测 VP2 蛋白的跨膜区；在线软件 SignalP-5.0 Server (<http://www.cbs.dtu.dk/services/SignalP-5.0/>) 预测 VP2 蛋白的信号肽。

1.3.5 VP2 蛋白二级结构的预测

在线软件 PSIPRED Workbench (<http://bioinf.cs.ucl.ac.uk/psipred/>) 和 SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) 预测 VP2 蛋白的二级结构。

1.3.6 VP2 蛋白三级结构的预测

使用在线软件 Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) 建立 VP2 蛋白的空间构象, 并进行高级结构的预测。

1.3.7 VP2 蛋白潜在抗原表位的预测

在线软件 Predicting Anti-genic Peptides
(<http://imed.med.ucm.es/Tools/antigenic.pl#:~:text=Antigenic%20peptides%20are%20determined%20using%20the%20method%20of,The%20reported%20accuracy%20of%20method%20is%20about%2075%25.>) 预测 VP2

蛋白的抗原决定簇；使用在线软件 ABCpred (<https://webs.iitd.edu.in/raghava/abcpred/>) 预测 VP2 蛋白的潜在 B 细胞抗原表位。

1.4 数据分析

试验数据使用 GraphPad Prism 7.0 软件进行分析, 结果以平均值表示。

3 结果分析

3.1 VP2 蛋白进化树构建

进化树构建结果显示 G II.4 LN-2021-18 与 G II.4 Sydney 2016 毒株位于同一个小分支, 亲缘关系最近; G I.3 Nashville 2016、G II.2 Japan 2020、G II.3 AnnArbor 2016、G II.13 USA 2010 和 G II.17 CN 2014 同属于一个大分支 (图 1)。

3.2 VP2 蛋白的理化性质

使用 ProtScale 在线软件预测 NoV 流行毒株 VP2 蛋白的理化性质, 结果显示 NoV VP2 蛋白均为亲水蛋白 (图 2); NoV VP2 蛋白的平均等电点为 10.47, 平均吸水系数为-0.47, 平均不稳定系数为 47.91, 碱性氨基酸平均含量为 11.60%; 其中 GL.3 Nashville 2016 毒株 VP2 蛋白的平均吸水系数、不稳定系数和碱性氨基酸含量较低, 分别为-0.27、42.13 和 9.30%; G II.17 CN 2014 毒株 VP2 蛋白的不稳定系数较高为 56.74。NoV 流行毒株 VP2 蛋白具体理化性质 (表 1)。

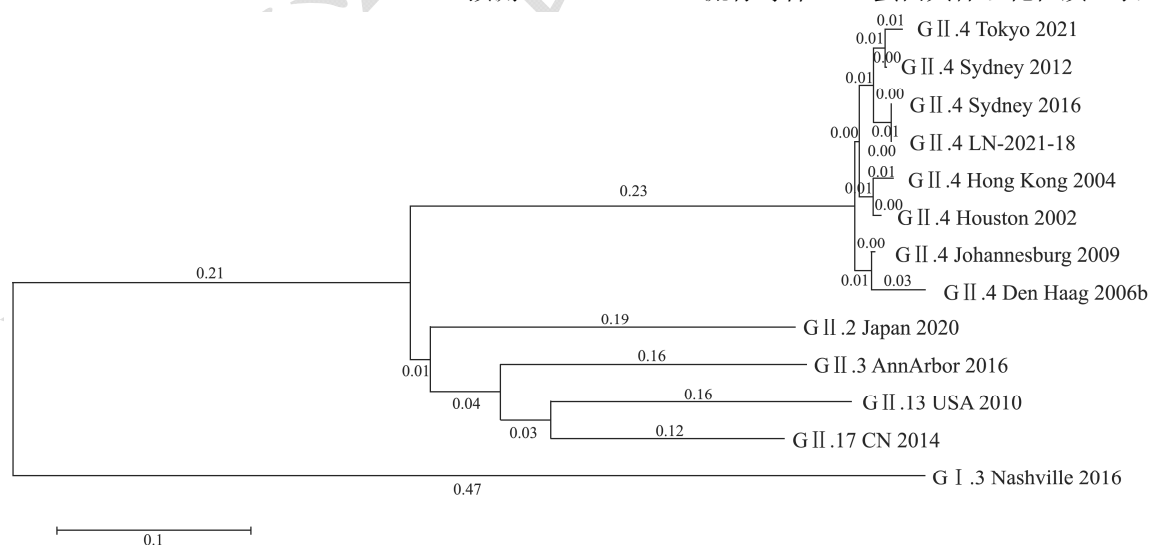


图 1 NoV VP2 蛋白的进化树分析结果

表 1 NoV 流行毒株 VP2 蛋白的理化性质

Table 1 Physicochemical properties of VP2 protein of NoV epidemic strains				
流行毒株	等电点	平均吸水系数	不稳定系数	碱性氨基酸比例/%
GII.4 LN-2021-18	10.47	-0.48	47.84	12.30
GI.3 Nashville 2016	10.21	-0.27	42.13	9.30
GII.2 Japan 2020	10.35	-0.56	45.18	11.20
GII.3 AnnArbor 2016	10.69	-0.44	46.58	11.00
GII.4 Tokyo 2021	10.71	-0.53	50.98	12.70
GII.4 Sydney 2016	10.47	-0.48	47.84	12.30
GII.4 Sydney 2012	10.58	-0.50	51.98	12.30
GII.4 Johannesburg 2009	10.58	-0.52	42.38	11.90
GII.4 Den Haag 2006b	10.67	-0.52	46.02	11.90
GII.4 Hong Kong 2004	10.58	-0.53	43.22	12.30
GII.4 Houston 2002	10.40	-0.50	47.16	11.20
GII.13 USA 2010	10.27	-0.42	54.88	10.60
GII.17 CN 2014	10.16	-0.40	56.74	10.80
平均值	10.47	-0.47	47.91	11.60

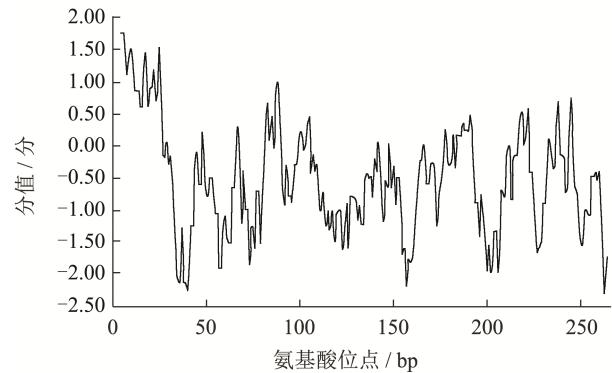


图 2 GII.4 LN-2021-18 VP2 蛋白的亲疏水性预测结果

Fig.2 Hydrophobicity prediction results of GII.4 LN-2021-18 VP2 protein

3.3 VP2 蛋白磷酸化位点和糖基化位点

NetPhos 3.1 Server 和 NetOGlyc 1.0 Server 分别预测 VP2 蛋白潜在的磷酸化位点和糖基化位点。结果显示, NoV VP2 蛋白含有多个磷酸化位点(图 3)和糖基化位点(图 4)。NoV VP2 蛋白潜在的磷酸化位点平均为 43 个和糖基化位点平均存在 2 个(表 2); 8 种 GII.4 亚型毒株在第 20、92、132、160、163、167、175、180、184、190、193、199、209、219、225、226、228、229、230、235、244 和 266 位存在共同丝氨酸位点, 第 71、94、114、124、176、194、222、236 和 251 位存在共同苏氨酸位点, 第 116 位存在共同酪氨酸位点; GII.2、GII.3、GII.13、GII.17 型毒株在第 20、81、168 和 171 位存在共同丝氨酸位点, 第 103 和 121 位存在共同苏氨酸位点; GI.3、GII.2、GII.3、

GII.13、GII.17 型毒株在第 81 位有共同丝氨酸位点。除 GII.4 Hong Kong 2004, 其余 7 种 GII.4 型毒株均存在第 171 位存在糖基化位点。

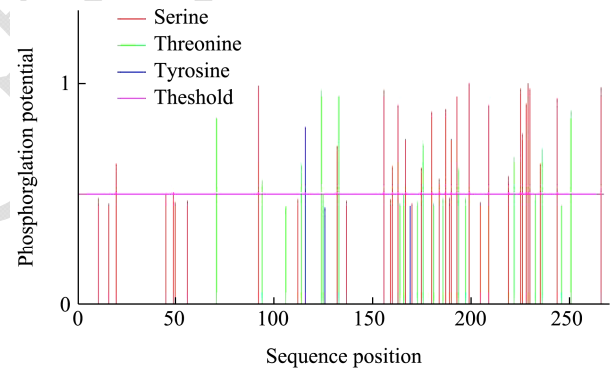


图 3 GII.4 LN-2021-18 VP2 蛋白磷酸化位点预测结果

Fig.3 Prediction results of phosphorylation sites of GII.4 LN-2021-18 VP2 protein

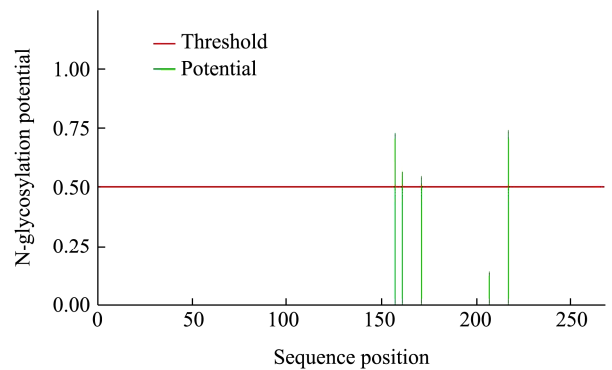


图 4 GII.4 LN-2021-18 VP2 蛋白糖基化位点预测结果

Fig.4 Prediction results of glycosylation sites of GII.4 LN-2021-18 VP2 protein

表 2 NoV 流行毒株 VP2 蛋白的磷酸化位点和糖基化位点

流行毒株	磷酸化位点			糖基化位点/个
	丝氨酸位点/个	苏氨酸位点/个	酪氨酸位点/个	
GII.4 LN-2021-18	25	10	1	4
GI.3 Nashville 2016	20	12	2	1
GII.2 Japan 2020	27	21	2	2
GII.3 AnnArbor 2016	33	14	1	2
GII.4 Tokyo 2021	29	11	1	2
GII.4 Sydney 2016	25	10	1	4
GII.4 Sydney 2012	30	11	1	2
GII.4 Johannesburg 2009	28	13	1	2
GII.4 Den Haag 2006b	28	12	1	2
GII.4 Hong Kong 2004	27	13	1	1
GII.4 Houston 2002	28	11	1	2
GII.13 USA 2010	35	16	3	2
GII.17 CN 2014	37	15	2	3
平均值	29	13	1	2

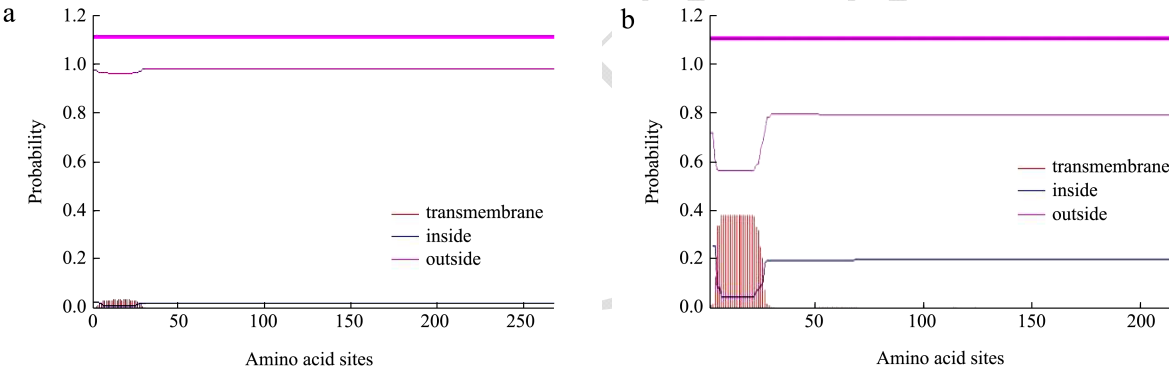


图 5 VP2 蛋白跨膜结构域预测结果

Fig.5 Prediction results of transmembrane domain of VP2 protein

注：a 为 GII.4 LN-2021-18 毒株，b 为 GI.3 Nashville 2016 毒株。图 6 同。

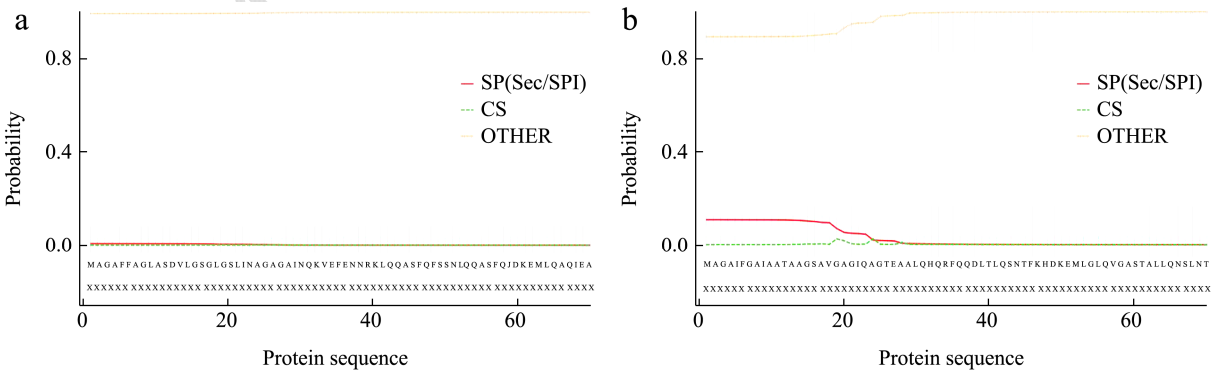


图 6 VP2 蛋白信号肽预测结果

Fig.6 Prediction results of signal peptide of VP2 protein

3.4 VP2 蛋白跨膜区及信号肽的预测

TMHMM Server v.2.0 和 SignalP5.0 Server 在线软件分别预测 VP2 蛋白是否存在跨膜区和信号肽。结果

显示，大多数 NoV 流行毒株 VP2 蛋白无跨膜区和信号肽（图 5a 和图 6a）；然而 GI.3 毒株 VP2 蛋白有跨膜区和信号肽（图 5b 和图 6b）。

3.5 VP2 蛋白二级结构的预测

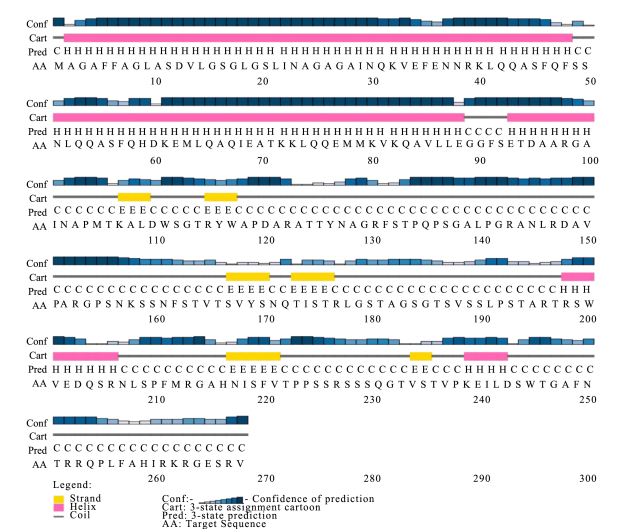


图 7 PSIPRED 软件预测的 G II.4 LN-2021-18 VP2 蛋白的二级结构

Fig.7 Secondary structure results of G II.4 LN-2021-18 VP2 protein measured on PSIPRED software

采用 PSIPRED 软件预测 NoV 流行毒株 VP2 蛋白二级结构主要形式，结果分析发现所有 VP2 蛋白结构

表 3 NoV 流行毒株 VP2 蛋白的二级结构

Table 3 Secondary structure results of VP2 protein of NoV epidemic strains				
流行毒株	α -螺旋(蓝色)/%	延伸链(红色)/%	β 转角(绿色)/%	无规则卷曲(紫色)/%
G II.4 LN-2021-18	36.94	7.09	6.72	49.25
G I.3 Nashville 2016	44.65	7.44	8.37	39.53
G II.2 Japan 2020	35.14	6.56	6.18	52.12
G II.3 AnnArbor 2016	37.40	8.27	5.51	48.82
G II.4 Tokyo 2021	37.69	8.96	6.72	46.64
G II.4 Sydney 2016	36.94	7.09	6.72	49.25
G II.4 Sydney 2012	36.94	8.96	5.60	48.51
G II.4 Johannesburg 2009	37.31	7.84	5.97	48.88
G II.4 Den Haag 2006b	38.06	7.46	5.60	48.88
G II.4 Hong Kong 2004	37.31	7.09	5.97	49.63
G II.4 Houston 2002	38.81	8.21	5.60	47.39
G II.13 USA 2010	41.56	4.53	6.17	47.74
G II.17 CN 2014	37.45	7.34	6.18	49.03
平均值	38.17	7.45	6.25	48.13

3.6 VP2 蛋白三级结构的预测

通过在线网站 Phyre2 对 NoV 流行毒株 VP2 蛋白的三级结构分析。预测结果发现 VP2 蛋白最高的置信度为 56.70%，蛋白质的覆盖率为 10%。其中 G II.3 毒株、G II.4 Tokyo 2021 毒株、G II.4 Hong Kong 2004

相似（图 7）；同时利用 SOPMA 预测 VP2 蛋白二级结构主要形式的含量，预测结果显示 NoV 流行毒株 VP2 蛋白中 α -螺旋（蓝色）平均占比 38.17%、延伸链（红色）平均占比 7.45%、 β 转角（绿色）平均占比 6.25%、无规则卷曲（紫色）平均占比 48.13%（图 8 和表 3）；进一步分析发现 G I.3 Nashville 2016 毒株 VP2 蛋白以 α -螺旋为主，其余 12 种蛋白均以无规则卷曲为主。无规则卷曲较为松散且易发生扭曲盘旋并暴露在蛋白质表面，较易成为抗原表位。

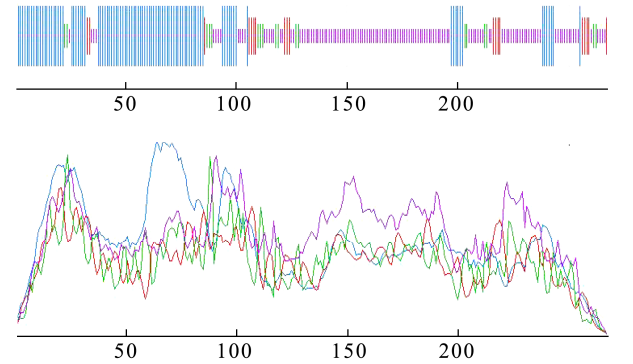


图 8 SOPMA 软件预测的 G II.4 LN-2021-18 VP2 蛋白的二级结构

Fig.8 Secondary structure results of G II.4 LN-2021-18 VP2 protein measured on SOPMA software

毒株、G II.4 Sydney 2016 毒株和 G II.4 LN-2021-18 毒株相似（图 9a）；G I.3 毒株、G II.13 毒株和 G II.17 毒株相似（图 9b）；G II.2 毒株（图 9c）；G II.4 Houston 2002 毒株和 G II.4 Johannesburg 2009 毒株相似（图 9d）；G II.4 Den Haag 2006b 毒株和 G II.4 Sydney 2012 毒株相似（图 9e）。

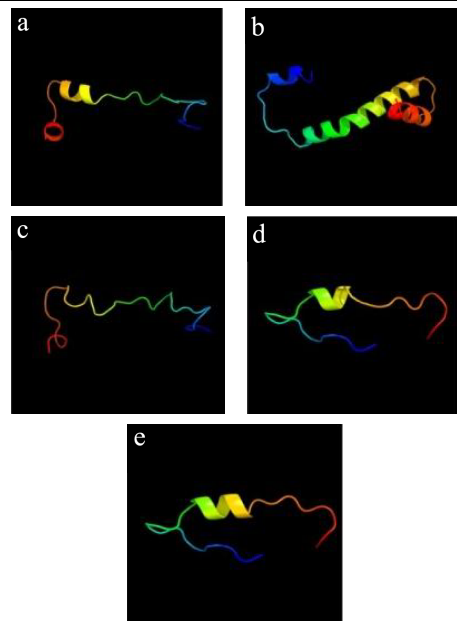


图9 VP2 蛋白三级结构空间模型预测结果

Fig.9 Prediction results of tertiary structure spatial model of VP2 protein

注: a 为 G II.3 毒株、G II.4 Tokyo 2021 毒株、G II.4 Hong Kong 2004 毒株、G II.4 Sydney 2016 毒株、G II.4 LN-2021-18 毒株; b 为 G I.3 毒株、G II.13 毒株、G II.17 毒株; c 为 G II.2 毒株; d 为 G II.4 Houston 2002 毒株、G II.4 Johannesburg 2009 毒株; e 为 G II.4 Den Haag 2006b 毒株、G II.4 Sydney 2012 毒株。

3.7 VP2 蛋白潜在抗原表位的预测

表4 G II.4 LN-2021-18 VP2 蛋白抗原决定簇的预测结果

Table 4 Prediction results of antigen determinants of G II.4 LN-2021-18 VP2 protein

条数	开始位置/位	结束位置/位	序列
1	4	23	5'-AFFAGLASDVLGSGLGSLIN-3'
2	41	57	5'-LQQASFQFSSNLQQASF-3'
3	81	88	5'-VKQAVLLE-3'
4	146	152	5'-LRDAVPA-3'
5	161	171	5'-NFSTVTSVYSN-3'
6	184	193	5'-SGTSVSSLPS-3'
7	216	223	5'-HNISFVTP-3'
8	230	240	5'-SQGTVSTVPKE-3'
9	253	259	5'-RQPLFAH-3'

抗原决定簇是蛋白质表面部分可以使免疫系统产生抗体的区域，一般由 6~12 个氨基酸或碳水基团组成。采用 Predicting Anti-genic Peptides 软件预测 NoV VP2 蛋白的抗原表位，结果显示 VP2 蛋白平均存在 8 个抗原决定簇；其中 G II.4 LN-2021-18 毒株 VP2 蛋白存在 9 个抗原决定簇，即第 4~23 位、第 41~57 位、

第 81~88 位、第 146~152 位、第 161~171 位、第 184~193 位、第 216~223 位、第 230~240 位和第 253~259 位，见表 4。

进一步采用 ABCpred 预测结果显示，NoV 流行毒株 VP2 蛋白平均存在 25 个 B 细胞抗原表位；其中 G II.4 LN-2021-18 存在 26 个 B 细胞抗原表位，分别为第 110~125 位、第 171~186 位、第 240~255 位等，具体表位及序列，见表 5。

表5 G II.4 LN-2021-18 VP2 蛋白潜在 B 细胞抗原表位的预测结果

Table 5 Prediction results of potential B cell epitopes of G II.4 LN-2021-18 VP2 protein

条数	开始位置/位	结束位置/位	序列
1	110	125	5'-DWSGTRYWAPDARATT-3'
2	171	186	5'-NQTISTRLGSTAGSGT-3'
3	240	255	5'-EILDSWTGAFNTRRQP-3'
4	61	76	5'-KEMLQAQIEATKKLQQ-3'
5	98	113	5'-RGAINAPMTKALDWSG-3'
6	89	104	5'-GGFSETDAARGAINAP-3'
7	39	54	5'-RKLQQASFQFSSNLQQ-3'
8	186	201	5'-TSVSSLPSTARTRSWV-3'
9	158	173	5'-KSSNFSTVTSVYSNQT-3'
10	50	65	5'-SNLQQASFQHDKEMLQ-3'
11	195	210	5'-ARTRSWVEDQSRNLSP-3'
12	147	162	5'-RDAVPARGPSNKSSNF-3'
13	26	41	5'-AGAINQKVEFENNRKL-3'
14	208	223	5'-LSPFMRGAHNISFVTP-3'
15	7	22	5'-AGLASDVLGSGLGSLI-3'
16	138	153	5'-GALPGRANLRDAVPA-3'
17	128	143	5'-AGRFSTPQPSGALPGR-3'
18	231	246	5'-QGTVSTVPKEILDSWT-3'
19	214	229	5'-GAHNISFVTPSSRSS-3'
20	117	132	5'-WAPDARATTYNAGRFS-3'
21	178	193	5'-LGSTAGSGTSVSSLPS-3'
22	69	84	5'-EATKKLQQEMMKVKQA-3'
23	220	235	5'-FVTPSSRSSSQGTVS-3'
24	19	34	5'-GSLINAGAGAINQKVE-3'
25	79	94	5'-MKVKQAVLLEGGFSET-3'
26	253	268	5'-RQPLFAHIRKRGESRV-3'

4 讨论

本研究通过生物信息学方法对 NoV G II.4 型流行毒株 VP2 蛋白的等电点进行预测，结果显示：VP2 蛋白的平均等电点为 10.47。等电点可以用来分离沉淀蛋

白,相比于传统的等电点测定方法,生物信息学方法更加简便、快速^[22]。VP2 蛋白中含量最高的氨基酸是丝氨酸,其次是丙氨酸;预测所获得的 VP2 蛋白分子质量可用于 VP2 蛋白的分离和提纯。

蛋白质磷酸化可以将外源刺激转化为内源的信号,是调节和控制蛋白质活力和功能的最基本、最普遍,也是最重要的机制^[23]。随着蛋白质磷酸化研究的深入,生物信息学产生的数据已经成为磷酸化蛋白质组研究中重要的基础。与传统的磷酸化位点测定方法相比,生物信息学可以更快更方便的对蛋白磷酸化位点预测。本文结果发现 NoV 流行毒株 VP2 蛋白平均存在 43 个磷酸化位点(表 2)。根据预测获得的 VP2 蛋白磷酸化位点可用于更深入的研究 VP2 蛋白对 NoV 的作用。此外有研究表明,病毒蛋白的磷酸化位点可能与其核定位信号和/或出核信号有关,这为研究 NoV VP2 蛋白的入核和出核提供了理论基础^[24]。

之前的研究表明 α -螺旋和 β -折叠位于蛋白质内部,较难形成抗原决定簇^[25]。其中 α -螺旋还可以与刺突蛋白结合,从而进行病毒检测、中和抗体分析和治疗^[26]。无规则卷曲多位于蛋白质外侧,可以折叠和扭转,因此无规则卷曲区域产生抗原决定簇的概率更高^[27]。NoV 流行毒株 VP2 蛋白二级结构预测结果显示, α -螺旋占比 38.17%、延伸链占比 7.45%、 β 转角占比 6.25%、无规则卷曲占比 48.13%。VP2 蛋白无规则卷曲所占比例最高,这表示其产生抗原决定簇区域的概率也比较高。根据预测获得的 VP2 蛋白二级结构和三级结构可为 NoV 疫苗的开发提供基础。

B 细胞主要介导体液免疫和细胞免疫,可以识别多种抗原和抗原裂解后的空间构象^[28]。研究表明 NoV 疫苗可以诱导 B 细胞免疫反应,因此 B 细胞抗原表位的预测显得尤为重要^[29]。使用 ABCpred 在线软件进行预测,得出 NoV 流行毒株 VP2 蛋白最优潜在 B 细胞抗原表位平均为 25 个。在目前的研究中,从 NoV 结构蛋白预测到与 B 细胞结合的抗原表位,发现可能与 B 细胞相互作用并启动免疫反应的潜在表位,以此诱导体液免疫和细胞免疫^[30]。根据预测获得的 B 细胞抗原表位可为开发基于 NoV VP2 蛋白的疫苗提供依据。

NoV VP2 蛋白的基本性质和其在病毒颗粒内部的位置决定了它能够与 VP1 蛋白相互作用而稳定病毒样颗粒,同时发现 G II.4 VP2 蛋白会对病毒 RdRp 活性造成影响^[31,32]。早期研究发现,VP2 蛋白可以提高 VP1 蛋白的稳定性及其表达水平,并且通过免疫共沉淀等技术验证了 VP2 蛋白上第 108~152 个氨基酸可能是与 VP1 蛋白相互作用的位点^[33,34]。利用酵母双杂交技术等实验表明了 G II.4 VP2 蛋白与 VP1 蛋白在病

毒进化中共存,且 VP2 蛋白序列的可变性可能是功能驱动的^[35]。这些研究表明 VP2 蛋白在 NoV 中发挥着重要作用,对其进行生物信息学分析可以对 NoV 的预防提供一种新思路。

5 结论

本研究基于软件分析预测,当前 NoV 流行毒株 G II.4 VP2 蛋白有如下特征:(1)为稳定的亲水性蛋白,平均等电点 10.47、吸水系数-0.47、不稳定系数 47.91、碱性氨基酸含量 11.60%。(2)平均存在 43 个蛋白质磷酸化位点和 2 个糖基化位点;除 G I.3 型 NoV 外,大多数毒株不存在跨膜蛋白和信号肽。(3)VP2 蛋白二级结构主要形式平均含量分别为 α -螺旋 38.17%、延伸链 7.45%、 β 转角 6.25%、无规则卷曲 48.13%;三级结构预测模型的蛋白覆盖率为 56.70%;平均存在 8 个潜在的蛋白质抗原决定簇和 25 个 B 细胞抗原表位。本文预测的 NoV 流行毒株 G II.4 VP2 蛋白的理化性质为深入研究 VP2 蛋白功能提供理论参考,同时二级、三级结构和抗原表位的预测也为研制基于 VP2 蛋白的 NoV 疫苗提供理论基础。这为解决由 NoV 引起的食品安全问题提供了新的思路。

参考文献

- [1] Kapikia A Z. The discovery of the 27-nm Norwalk virus: an historic perspective [J]. Journal of Infectious Diseases, 2000, 181(Suppl 2): S295-302.
- [2] Ruoff K, Devant J M, Hansman G. Natural extracts, honey, and propolis as human norovirus inhibitors [J]. Scientific Reports, 2022, 12(1): 8116.
- [3] Chhabra P, De G M, Parra G I, et al. Updated classification of norovirus genogroups and genotypes [J]. Journal of General Virology, 2019, 100(10): 1393-1406.
- [4] Ahmed S M, Hall A J, Robinson A E, et al. Global prevalence of norovirus in cases of gastroenteritis: a systematic review and meta-analysis [J]. Lancet Infectious Diseases, 2014, 14(8): 725-730.
- [5] Lindsay L, Wolter J, Coster I D, et al. A decade of norovirus disease risk among older adults in upper-middle and high income countries: a systematic review [J]. BMC Infectious Diseases, 2015, 15: 425.
- [6] 廖小艳,陈丽丽,白亚龙.食品中诺如病毒检测技术研究进展[J].食品与机械,2021,37(4):200-206,32.
- [7] Bok K, Green K Y. Norovirus gastroenteritis in immunocompromised patients [J]. New England Journal of Medicine, 2012, 367(22): 2126-2132.

- [8] Bruggink L D, Oluwatoyin O, Sameer R, et al. Molecular and epidemiological features of gastroenteritis outbreaks involving genogroup I norovirus in Victoria, Australia, 2002-2010 [J]. *Journal of Medical Virology*, 2012, 84(9): 1437-1448.
- [9] Cannon J L, Barclay L, Collins N R, et al. Genetic and epidemiologic trends of norovirus outbreaks in the United States from 2013 to 2016 demonstrated emergence of novel G II. 4 recombinant viruses [J]. *Journal of Clinical Microbiology*, 2017, 55(7): 2208-2221.
- [10] Liao Y, Wang L, Hong X, et al. The VP2 protein exhibits cross-interaction to the VP1 protein in norovirus G II. 17 [J]. *Infection Genetics and Evolution*, 2022, 100: 105265.
- [11] Tohma K, Lepore C J, Gao Y, et al. Population genomics of G II. 4 noroviruses reveal complex diversification and new antigenic sites involved in the emergence of pandemic strains [J]. *MBIO*, 2019, 10(5): e02202-19.
- [12] Kendra J A, Tohma K, Ford-siltz L A, et al. Antigenic cartography reveals complexities of genetic determinants that lead to antigenic differences among pandemic G II. 4 noroviruses [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2021, 118(11): e2015874118.
- [13] 宋杰,董少忠.人诺如病毒最新研究进展[J].*病毒学报*,2017, 33(5):785-790.
- [14] Jiang X, Wang M, Graham D Y, et al. Expression, self-assembly, and antigenicity of the Norwalk virus capsid protein [J]. *Journal of Virology*, 1992, 66(11): 6527-6532.
- [15] Mohamed F F, Ktob G K F, Ismaeil M E A, et al. Phylogeny of bovine norovirus in Egypt based on VP2 gene [J]. *International Journal of Veterinary Science and Medicine*, 2018, 6(1): 48-52.
- [16] Liu Z, Zhang M, Shen Z, et al. The coordinating role of the human norovirus minor capsid protein VP2 is essential to functional change and nuclear localization of the major capsid protein VP1 [J]. *Archives of Virology*, 2019, 164(4): 1173-1180.
- [17] Campillay-Véliz C P, Carvajal J J, Avellaneda A M, et al. Human norovirus proteins: implications in the replicative cycle, pathogenesis, and the host immune response [J]. *Frontiers in Immunology*, 2020, 11: 961.
- [18] Seah E L, Gunesekeere I C, Marshall J A, et al. Variation in ORF3 of genogroup 2 norwalk-like viruses [J]. *Archives of Virology*, 1999, 144(5): 1007-1014.
- [19] Eetienney M, Tarris G, Abou-hamad N, et al. Epidemiological impact of G II. 17 human noroviruses associated with attachment to enterocytes [J]. *Frontiers in Microbiology*, 2022, 13: 858245.
- [20] Wang N, Pan G, Liu P, et al. Advances and future perspective on detection technology of human norovirus [J]. *Pathogens*, 2021, 10(11): 1383.
- [21] Zhou R, Yin J, Lu J, et al. Genome sequences of five microviruses and a murine norovirus [J]. *Microbiology Resource Announcement*, 2021, 10(21): e0023221.
- [22] Andersen T, Pepaj M, Trones R, et al. Isoelectric point separation of proteins by capillary pH-gradient ion-exchange chromatography [J]. *Journal of Chromatography A*, 2004, 1025(2): 217-226.
- [23] Ardito F, Giuliani M, Perrone D, et al. The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review) [J]. *International Journal of Molecular Medicine*, 2017, 40(2): 271-280.
- [24] Zhan J, Watts E, Brice A M, et al. Molecular basis of functional effects of phosphorylation of the C-terminal domain of the rabies virus P protein [J]. *Journal of Virology*, 2022, 96(9): e0011122.
- [25] Fujiwara K, Toda H, Ikeguchi M. Dependence of alpha-helical and beta-sheet amino acid propensities on the overall protein fold type [J]. *BMC Structural Biology*, 2012, 12: 18.
- [26] Fernandes L A, Gomes A A, Guimaraes B G, et al. Engineering defensin alpha-helix to produce high-affinity SARS-CoV-2 spike protein binding ligands [J]. *Protein Science*, 2022, 31(6): e4355.
- [27] 李智伟,张峰波,卢佩佩,等.羊种布鲁氏菌 OMP16 蛋白结构的生物信息学分析[J].*中国人兽共患病学报*,2019,35(1):21-27,33.
- [28] Padula L, Fisher E, Rivas K, et al. Secreted heat shock protein gp96-Ig and OX40L-Fc combination vaccine enhances SARS-CoV-2 Spike (S) protein-specific B and T cell immune responses [J]. *Vaccine-X*, 2022, 12: 100202.
- [29] Ramani S, Neill F H, Ferreira J, et al. B-Cell responses to intramuscular administration of a bivalent virus-like particle human norovirus vaccine [J]. *Clinical and Vaccine Immunology*, 2017, 24(5): e00571-16.
- [30] Eltilib E, Almofti Y A, Nouri M, et al. Modeling and in silico analysis for prediction of epitopes vaccine against Norwalk virus from capsid protein (VP1) through reverse vaccinology [J]. *American Journal of Infectious Diseases and Microbiology*, 2020, 8(1): 29-44.

- [31] Vongpunsawad S, Venkataram P B V, Estes M K. Norwalk virus minor capsid protein VP2 associates within the VP1 shell domain [J]. *Journal of Virology*, 2013, 87(9): 4818-4825.
- [32] Subba-Reddy C V, Goodfellow I, Kao C C. VPg-primed RNA synthesis of norovirus RNA-dependent RNA polymerases by using a novel cell-based assay [J]. *Journal of Virology*, 2011, 85(24): 13027-13037.
- [33] Bertolotti Ciarlet A, Crawford S E, Hutson A M, et al. The 3' end of Norwalk virus mRNA contains determinants that regulate the expression and stability of the viral capsid protein VP1: a novel function for the VP2 protein [J]. *Journal of Virology*, 2003, 77(21): 11603-11615.
- [34] Glass P J, Zeng C Q, Estes M K. Two nonoverlapping domains on the Norwalk virus open reading frame 3 (ORF3) protein are involved in the formation of the phosphorylated 35K protein and in ORF3-capsid protein interactions [J]. *Journal of Virology*, 2003, 77(6): 3569-3577.
- [35] Chan M C, Lee N, Ho W S, et al. Covariation of major and minor viral capsid proteins in norovirus genogroup II genotype 4 strains [J]. *Journal of Virology*, 2012, 86(2): 1227-1232.