

DOI: 10. 12138/j. issn. 1671—9638. 20233032

· 论 著 ·

新型冠状病毒 ORF3a 蛋白质的生物信息学分析

王 瑄^{1,2}, 李祥芳², 丁寿鹏², 吴利先²

(1. 南昌大学玛丽女王学院, 江西 南昌 330031; 2. 大理大学基础医学院, 云南 大理 671000)

[摘要] **目的** 用生物信息学分析严重急性呼吸综合征冠状病毒 2(SARS-CoV-2) 基因 ORF3a 蛋白质。**方法** 从 UniProt 数据库获取 ORF3a 蛋白质的氨基酸序列, 应用 ProtParam tool 分析其理化性质, SOPMA、SWISS-MODEL 预测二、三级结构, PortScale、Signal IP 4.0 Server、TMHMM Server 2.0 分析 ORF3a 蛋白质亲疏水性、信号肽、跨膜区。应用 ABCpred、SYFPEITHI 和 NetPhos 3.1 Server 预测 B、T 细胞表面抗原表位及磷酸化位点, VaxiJenv 2.0、AlgPred server 预测抗原性和过敏性。**结果** ORF3a 蛋白质含 275 个氨基酸, 为稳定、疏水性蛋白质, 有 3 个跨膜螺旋区、29 个磷酸化位点, 含多个 T、B 细胞抗原表位, 具有抗原性和非过敏性。**结论** ORF3a 蛋白质结构稳定, 含多个 T、B 细胞抗原表位, 可为研究 SARS-CoV-2 多表位疫苗提供基础。

[关键词] SARS-CoV-2; ORF3a; 生物信息学分析; 多表位疫苗

[中图分类号] R186 R511

Bioinformatics analysis of ORF3a protein of SARS-CoV-2

WANG Xuan^{1,2}, LI Xiang-fang², DING Shou-peng², WU Li-xian² (1. College of Nanchang University Queen Mary School, Nanchang 330031, China; 2. School of Basic Medical Sciences, Dali University, Dali 671000, China)

[Abstract] **Objective** To analyze ORF3a protein in severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by bioinformatics. **Methods** Amino acid sequence of ORF3a was obtained from UniProt. Physicochemical properties of the protein were analyzed by ProtParam tool. Secondary and tertiary structures were predicted by SOPMA and SWISS-MODEL. Hydrophobicity, signal peptide, and transmembrane region of ORF3a protein were analyzed by PortScale, Signal IP 4.0 Server, and TMHMM Server 2.0. B cell and T cell surface epitopes as well as phosphorylation sites were predicted by ABCpred, SYFPEITHI and NetPHos 3.1 Server. Antigenicity and allergenicity were predicted by VaxiJenv 2.0 and AlgPred server. **Results** ORF3a is a stable, hydrophobic protein with 3 transmembrane helical regions and 29 phosphorylation sites. It contains 275 amino acids and multiple T and B cell epitopes, and is antigenic and non-allergenic. **Conclusion** ORF3a protein is structurally stable. It contains multiple T and B cell epitopes, thus can provide basis for the study of SARS-CoV-2 multi-epitope vaccines.

[Key words] SARS-CoV-2; ORF3a; bioinformatics analysis; multi-epitope vaccine

严重急性呼吸综合征冠状病毒 2(SARS-CoV-2) 具有高度传染性, 已造成全球社会和经济的混乱。SARS-CoV-2 是单链 RNA 病毒, 大小约为 30 kb, 属于 β 冠状病毒属, 含至少 15 个开放阅读框(open reading frames, ORFs), 编码 4 个结构蛋白和 16 个非结构蛋白。非结构蛋白(non-structural proteins,

NSP) 包含由 ORF1ab 编码的 NSP1~16 和至少 9 个辅助蛋白 ORF3a、3b、6、7a、7b、8b、9b、9c 和 10^[1]。在严重急性呼吸综合征冠状病毒(SARS-CoV) 中, 结构蛋白之间的相互作用是其成熟所必需, 而非结构蛋白之间的结合保证了病毒复制^[2]的完成。辅助蛋白在体内和体外病毒复制中必不可少, 对病毒-

[收稿日期] 2022-06-16

[基金项目] 国家自然科学基金(81260456)

[作者简介] 王瑄(2001-), 男(白族), 云南省大理市人, 临床医学本科生, 主要从事感染与免疫研究。

[通信作者] 吴利先 E-mail: w_lixian@163.com

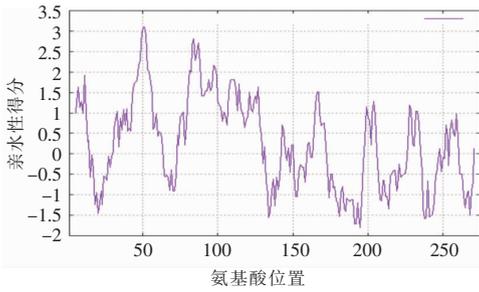


图 3 ORF3a 蛋白质的亲疏水性预测

Figure 3 Prediction of hydrophobicity of ORF3a protein

2.5 ORF3a 蛋白质跨膜区和信号肽预测 ORF3a 蛋白质跨膜螺旋数为 3。见图 4。跨膜氨基酸数为 66。207(>18 表示很可能为跨膜蛋白质),在胞质内的概率为 0.006,说明 ORF3a 蛋白质存在跨膜区,为跨膜蛋白质。应用 Signal IP 4.0 Server 预测 ORF3a 蛋白质共 275 个氨基酸,D 值为 0.181,阈值 0.500,无信号肽。见图 5。

2.6 ORF3a 蛋白质磷酸化位点 ORF3a 蛋白质含 29 个磷酸化位点,其中丝氨酸磷酸化位点 10 个(氨基酸位置为 74、92、135、171、177、180、195、209、220、253),苏氨酸磷酸化位点 12 个(氨基酸位置为 9、12、14、24、64、170、190、208、223、269、270、271),酪氨酸磷酸化位点 7 个(氨基酸位置为 156、160、184、189、211、212、264)。见图 6。

2.7 ORF3a 蛋白质抗原表位 ORF3a 蛋白质有多个 B 细胞抗原表位,其中得分 >0.51(临界参考值 0.51)的可能 B 细胞抗原表位 24 个,>0.85 的 9 个。见表 1。应用 SYFPEITHI 软件预测获得多个

可能的 T 细胞抗原表位,MHC 类型为 HLA-A × 02:01。见表 2。

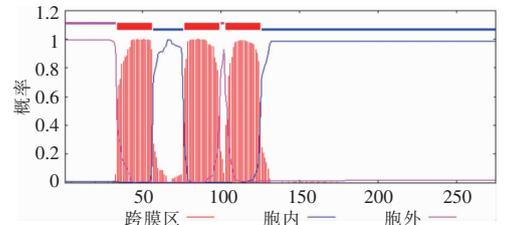


图 4 ORF3a 蛋白质跨膜区预测

Figure 4 Prediction of transmembrane region of ORF3a protein

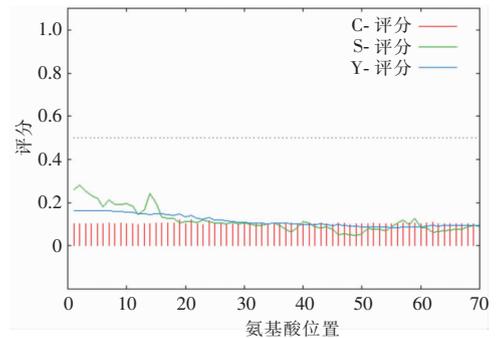


图 5 ORF3a 蛋白质的信号肽预测

Figure 5 Prediction of signal peptide of ORF3a protein

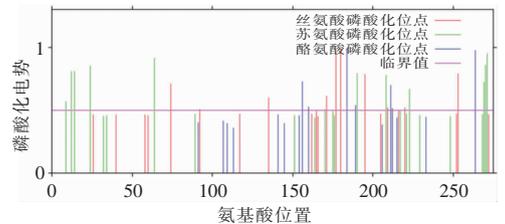


图 6 ORF3a 蛋白质磷酸化位点预测

Figure 6 Prediction of phosphorylation site of ORF3a protein

表 1 ORF3a 蛋白 B 细胞抗原表位预测

Table 1 Prediction of B cell epitopes of ORF3a protein

等级	序列	起始氨基酸位置	得分	等级	序列	起始氨基酸位置	得分
1	QGEIKDATPSDFVRAT	17	0.94	11	TDTGVEHVTFIYNKI	221	0.81
2	TSPISEHDYQIGGYTE	176	0.93	11	LHSYFTSDYYQLYSTQ	203	0.81
3	HVQIHTIDGSSGVVNP	243	0.91	11	TNCYDYCIPYNSVTSS	151	0.81
4	IGTVTLKQGEIKDASP	10	0.90	12	GSSGVVNPVMEPIYDE	251	0.79
5	TIPIQASLPFGWLIVG	34	0.89	13	MEPIYDEPTTTTSVPL	260	0.77
5	YSTQLSTDTGVEHVTF	215	0.89	13	TSSIVITSGDGTTSPI	164	0.77
6	NKIVDEPEEHVQIHTI	234	0.88	14	PLLYDANYFLCWHNTNC	138	0.73
6	FVRIIMRLWLCWKCRS	120	0.88	15	FVRATATIPIQASLPF	28	0.71
7	KCRSKNPLLYDANYFL	132	0.87	16	AGLEAPFLYLYALVYF	99	0.69
8	TEKWESGVKDCVVLHS	190	0.85	16	ASKIITLKKRWQLALS	59	0.69
9	YFLCWHNTNCYDYCIPY	145	0.84	17	KRWQLALS KGVHFVCN	67	0.65
10	YQIGGYTEKWESGVKD	184	0.83	18	FVTVYSHLLLVAAGLE	87	0.64

表 2 ORF3a 蛋白质 T 细胞抗原表位预测

Table 2 Prediction of T cell epitopes of ORF3a protein

等级	序列	起始氨基酸位置	得分	等级	序列	起始氨基酸位置	得分
1	WLIVGVALL	45	30	7	IVGVALLAV	47	22
2	ALSKGVHVV	72	29	7	GLEAPFLYL	100	22
2	YLYALVYFL	107	29	7	STDTGVEHV	220	22
3	NLLLLFVTV	82	27	8	ALLAVFQSA	51	20
4	ALVYFLQSI	110	26	8	LLFVTVYSH	85	20
5	HLLLVAAGL	93	25	8	TVYSHLLL	89	20
5	LLYDANYFL	139	25	8	APFLYLYAL	103	20
6	DLFMRIFTI	2	24	8	HTIDGSSGV	247	20
6	ATIPIQASL	33	24	8	TIDGSSGVV	248	20
6	SLPFGWLIV	40	24				

2.8 ORF3a 蛋白质亚细胞定位 应用 ProtComp Version 分析该蛋白质最可能定位在宿主细胞的细胞膜,不可能存在于细胞质或细胞膜外。见表 3。

表 3 ORF3a 蛋白质的亚细胞定位

Table 3 Subcellular localization of ORF3a protein

位置权重	LocDB	PotLocDB	NNets	Pentamers	得分
细胞质	0	0	0.09	0	0
细胞膜	0	0	2.68	3.03	9.97
细胞膜外	0	0	0.05	0	0

2.9 ORF3a 蛋白质抗原性和过敏原性预测 基于蛋白质的主要化学特征,应用 VaxiJenv 2.0 预测其抗原性,使用针对各种目标生物体(细菌、病毒、肿瘤、寄生虫和真菌)的非对齐方法,准确率为 70%~89%。本研究选择“病毒”为目标生物体(阈值 0.4),总体预测抗原可能性为 0.4945,具有抗原性。应用 AlgPred server 预测 ORF3a 蛋白质为非过敏原性,不诱导过敏原特异性抗体的产生。

3 讨论

SARS-CoV-2 感染是 21 世纪第三次传染病大流行,病毒传播速度快、生存能力强、致死率高,已致全球 600 多万人丧失生命。虽然目前疫苗在预防 SARS-CoV-2 感染方面有效,但新的病毒突变体可能会削弱现有疫苗的有效性,所以迫切需要找出突变体相关疫苗,推进 SARS-CoV-2 感染预防方案的合理设计。SARS-CoV-2 是一种高度复杂和多样化的 RNA 病毒,具有大基因组、复杂的基因库,以及转录和翻译机制,感染后可引起高发病率和高病死

率的严重呼吸道疾病。与其他冠状病毒一样,SARS-CoV-2 独特的辅助蛋白可以破坏人体先天免疫机制,可能与病毒高致病性有关,并在病毒复制中发挥重要作用。因此,研究这些蛋白质可能阐明 SARS-CoV-2 破坏细胞功能的致病机制。近年来,生物信息学以及新的计算和预测工具在现代结构疫苗学中的应用极大地促进了新型疫苗的开发。在此,通过生物信息学研究展示 ORF3a 作为 SARS-CoV-2 疫苗靶点的潜在意义。

ORF3a 是 SARS-CoV-2 最重要的辅助蛋白质之一,也是 SARS-CoV-2 基因组编码的最大辅助蛋白质。ORF3a 区域含有 G-四联体结构,该结构在人类和其他生物的基因转录中发挥调节作用,可促进 SARS-CoV-2 的转录和复制^[5]。ORF3a 基因位于病毒基因组中的刺突蛋白和包膜蛋白基因之间,其蛋白的主要作用包括毒性、传染性、离子通道活性、形态发生和病毒释放。频率最高的突变是 Q57H^[6],可能会改变病毒的传播和发病机制,故了解其复杂的突变特征可以指导疫苗的设计和开发。此外,ORF3a 蛋白质的主要特征之一是存在一个富含半胱氨酸的结构域,其参与酶的亲核取代反应和蛋白-蛋白相互作用^[7]。研究^[8]发现,SARS-CoV-2 ORF3a 与 SARS-CoV ORF3a 具有 72% 的同源性,可形成同型四聚体钾敏感离子通道,诱导细胞周期阻滞和细胞凋亡,并影响病毒的摄取和释放,但 SARS-CoV-2 ORF3a 的凋亡活性低于 SARS-CoV ORF3a,故 SARS-CoV-2 感染者在早期阶段症状较轻甚至无症状^[9]。感染后期,ORF3a 激活 IL-1 β 前基因表达和 IL-1 β 分泌,可导致严重的肺损伤^[10]。研究^[11]通过高通量筛选试验发现,ORF3a 蛋白质与人血红素加氧酶 1(heme oxygenase 1, HMOX1)

蛋白高度相互作用,并抑制 HMOX1 的抗炎功能,从而加重炎症和组织损伤。此外,ORF3a 也能阻止自噬体与溶酶体融合,保护病毒不被破坏^[12]。

本研究采用生物信息学方法预测 ORF3a 蛋白质含有若干 T、B 细胞抗原表位,具有抗原性和非过敏原性,可用于多表位疫苗的设计,使其有效地识别和组装更能刺激免疫系统的 B 细胞和 T 细胞表位,诱导免疫反应。ORF3a 蛋白质的不稳定性指数 < 40,被认为是稳定的蛋白质,故所设计的肽疫苗是稳定的。高脂肪指数(103.42)表明蛋白质在较宽温度范围内具有较高的稳定性。本研究预测 ORF3a 蛋白质主要位于宿主细胞的细胞膜,与 Issa 等^[8]研究一致。ORF3a 也被证明定位于质膜、核周区域和高尔基体^[13],并诱导细胞凋亡、病毒颗粒释放、细胞毒性^[14]、高尔基体碎片和细胞内囊泡的形成^[15]。ORF3a 具有 3 个跨膜螺旋区(TM1-TM3,残基 41-132),具有促凋亡功能,目前已分离出 17 个突变体,其中 10 个突变位点出现在 ORF3a 的跨膜结构域内,并与中心孔或离子通道接触。其他变异位点位于 ORF3a 结构的不同位置,最常见的突变是 Q57H。尽管该突变点位于跨膜通道,但通道的几何形状没有发生显著改变,表明通道特性没有改变^[16],ORF3a 结构域的相对保守性提示了疫苗设计的可能性。研究预测 ORF3a 是疏水性蛋白质,能够在宿主细胞膜上形成孔,被称为“病毒孔蛋白”。病毒孔蛋白定位于宿主细胞膜结构,如内质网、高尔基体,可以促进离子的运输,调节离子稳态,为病毒感染创造有利环境,也促进病毒免疫逃逸。SARS-CoV-2 感染给医疗行业带来重大压力,如何进行有效的预防和治疗成为亟待解决的问题。对 SARS-CoV-2 的蛋白进行生信分析有助于了解病毒自身结构和侵染过程,根据蛋白的抗原表位进行疫苗研发,使疾病预防更有针对性。

利益冲突:所有作者均声明不存在利益冲突。

[参考文献]

- [1] Chan JFW, Kok KH, Zhu Z, et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan[J]. *Emerg Microbes Infect*, 2020, 9(1): 221-236.
- [2] Snijder EJ, Decroly E, Ziebuhr J. The nonstructural proteins directing coronavirus RNA synthesis and processing[J]. *Adv Virus Res*, 2016, 96: 59-126.
- [3] Rota PA, Oberste MS, Monroe SS, et al. Characterization of a novel coronavirus associated with severe acute respiratory

syndrome[J]. *Science*, 2003, 300(5624): 1394-1399.

- [4] Wu SQ, Tian C, Liu PP, et al. Effects of SARS-CoV-2 mutations on protein structures and intraviral protein-protein interactions[J]. *J Med Virol*, 2021, 93(4): 2132-2140.
- [5] Ji DY, Juhás M, Tsang CM, et al. Discovery of G-quadruplex-forming sequences in SARS-CoV-2[J]. *Brief Bioinform*, 2021, 22(2): 1150-1160.
- [6] Hassan SS, Attrish D, Ghosh S, et al. Pathogenic perspective of missense mutations of ORF3a protein of SARS-CoV-2[J]. *Virus Res*, 2021, 300: 198441.
- [7] To J, Torres J. Beyond channel activity: protein-protein interactions involving viroporins[J]. *Subcell Biochem*, 2018, 88: 329-377.
- [8] Issa E, Merhi G, Panossian B, et al. SARS-CoV-2 and ORF3a: nonsynonymous mutations, functional domains, and viral pathogenesis[J]. *mSystems*, 2020, 5(3): e00266-20.
- [9] Ren YJ, Shu T, Wu D, et al. The ORF3a protein of SARS-CoV-2 induces apoptosis in cells[J]. *Cell Mol Immunol*, 2020, 17(8): 881-883.
- [10] Siu KL, Yuen KS, Castaño-Rodríguez C, et al. Severe acute respiratory syndrome coronavirus ORF3a protein activates the NLRP3 inflammasome by promoting TRAF3-dependent ubiquitination of ASC[J]. *FASEB J*, 2019, 33(8): 8865-8877.
- [11] Batra N, De Souza C, Batra J, et al. The HMOX1 pathway as a promising target for the treatment and prevention of SARS-CoV-2 of 2019 (COVID-19)[J]. *Int J Mol Sci*, 2020, 21(17): 6412.
- [12] Miao GY, Zhao HY, Li Y, et al. ORF3a of the COVID-19 virus SARS-CoV-2 blocks HOPS complex-mediated assembly of the SNARE complex required for autolysosome formation[J]. *Dev Cell*, 2021, 56(4): 427-442. e5.
- [13] Yuan XL, Li JY, Shan YJ, et al. Subcellular localization and membrane association of SARS-CoV 3a protein[J]. *Virus Res*, 2005, 109(2): 191-202.
- [14] Freundt EC, Yu L, Goldsmith CS, et al. The open reading frame 3a protein of severe acute respiratory syndrome-associated coronavirus promotes membrane rearrangement and cell death[J]. *J Virol*, 2010, 84(2): 1097-1109.
- [15] Yang S, Tian MJ, Dai YL, et al. Infection and chronic disease activate a brain-muscle signaling axis that regulates muscle performance[J]. *bioRxiv*, 2022. DOI: 10.1101/2020.12.20.423533. Epub ahead of print.
- [16] Bianchi M, Borsetti A, Ciccozzi M, et al. SARS-CoV-2 ORF3a: mutability and function[J]. *Int J Biol Macromol*, 2021, 170: 820-826.

(本文编辑:翟若南、左双燕)

本文引用格式:王瑄,李祥芳,丁寿鹏,等.新型冠状病毒 ORF3a 蛋白质的生物信息学分析[J]. *中国感染控制杂志*, 2023, 22(2): 224-228. DOI:10.12138/j.issn.1671-9638.20233032.

Cite this article as: WANG Xuan, LI Xiang-fang, DING Shou-peng, et al. Bioinformatics analysis of ORF3a protein of SARS-CoV-2[J]. *Chin J Infect Control*, 2023, 22(2): 224-228. DOI: 10.12138/j.issn.1671-9638.20233032.