

DOI:10.13350/j.cjpb.220905

• 论著 •

鸟分枝杆菌 MAV-4563 蛋白的生物信息学分析*

马枫茜,陈瑞枫,丁寿鹏,李祥芳,高婧华,吴利先**

(大理大学基础医学院微生物与免疫学教研室,云南大理 671000)

【摘要】 目的 应用生物信息学方法预测分析鸟结核分枝杆菌 MAV-dalidaxue4563 基因编码蛋白的结构功能和生物特性。方法 采用 ProtParam、ProtScaleon Expasy, TMHMM Server v 2.0, SignalP4.1, PSORT, NetPhos 3.1, BLAST, SOPMA, Phyre2, ABCpred, SYFPEITHI 等生物信息学软件分析 MAV4563 蛋白的理化性质、亲疏水性、跨膜区、信号肽、亚细胞定位、磷酸化位点、保守域、二级结构及三级结构同源建模、B 细胞和 T 细胞表位等生物学特征。

结果 MAV4563 基因全长 756 bp, 编码 251 个氨基酸的蛋白。该蛋白分子质量为 28 ku, 等电点为 9.23, 分子式(formula)为 $C_{1264} H_{2013} N_{385} O_{353} S_7$, 为不稳定的疏水性蛋白。该蛋白无跨膜区, 无信号肽, 预测亚细胞定位于细胞质中, 有 13 个磷酸化位点; 蛋白二级结构中无规则卷曲占 41.83%; 预测该蛋白含多个 B 细胞和 T 细胞表位, 其中 MHC-I 型和 MHC-II 型表位集中在蛋白 C 端的后 90 个氨基酸区域。**结论** 预测 MAV-4563 蛋白含多个 B 细胞和 T 细胞表位, 其中 B 细胞表位和 MHC-I 型表位集中在 C 端的后 90 个氨基酸区域, 该区域可激活 NTM 特异性 CTL 和 Th 免疫应答, 可作为非结核病的靶向疫苗抗原, 为进一步研究该蛋白的治疗性 T 细胞疫苗奠定了基础。

【关键词】 鸟结核分枝杆菌; MAV-4563 蛋白; T 细胞疫苗

【中图分类号】 R373.911

【文献标识码】 A

【文章编号】 1673-5234(2022)09-1015-05

[*Journal of Pathogen Biology*. 2022 Sep.;17(9): 1015—1019.]

Bioinformatics analysis of *Mycobacterium avium* MAV-4563 protein

MA Feng-qian, CHEN Rui-feng, DING Shou-peng, LI Xiang-fang, Wu Li-xian (Department of Microbiology and Immunology, School of Basic Medical Sciences, Dali University, Dali, Yunnan 671000, China)

【Abstract】 **Objective** To apply bioinformatics methods to predict and analyze the structure-function and biological properties of the protein encoded by *Mycobacterium avium* tuberculosis MAV-dalidaxue4563 gene. **Methods** Bioinformatics software such as ProtParam, ProtScaleon Expasy, TMHMM Server v 2.0, SignalP4.1, PSORT, NetPhos 3.1, BLAST, SOPMA, Phyre2, ABCpred, SYFPEITHI was used to analyze MAV4563. The biological characteristics of the MAV4563 protein, including physicochemical properties, hydrophobicity, a transmembrane region, signal peptide, subcellular localization, phosphorylation sites, conserved domains, secondary and tertiary structure homology modeling, B- and T-cell epitopes, were analyzed. **Results** The MAV4563 gene is 756 bp long and encodes a protein of 251 amino acids. The protein has a molecular mass of 28ku, an isoelectric point of 9.23, and a molecular formula of $C_{1264} H_{2013} N_{385} O_{353} S_7$, which is an unstable hydrophobic protein. The protein has no transmembrane region, no signal peptide, and predicted subcellular localization in the cytoplasm with 13 phosphorylation sites; the irregular coiling in the secondary structure of the protein accounted for 41.83%; the protein was predicted to contain several B- and T-cell epitopes, among which the MHC-I and MHC-II epitopes were concentrated in the latter 90 amino acid region of the C-terminus of the protein. **Conclusion** The MAV-4563 protein is predicted to contain multiple B- and T-cell epitopes, among which the B-cell epitopes and MHC-I type epitopes are concentrated in the posterior 90 amino acid region of the C-terminus, which can activate NTM-specific CTL and Th immune responses and can be used as a non-tuberculosis targeted vaccine antigen, laying the foundation for further research on therapeutic T-cell vaccines for this protein.

【Key words】 *Mycobacterium avium*; MAV-4563 protein; T cell vaccine

* ** 非结核分枝杆菌(*Nontuberculous mycobacterium*, NTM)引起的 NTM 病的发病率和流行率正在上升, 尤其是在结核病发病率下降的发达国家。NTM 感染既能影响免疫功能低下的患者, 也影响免疫功能正常的患者。引起人类疾病的 NTM 常见种类是缓慢生长的鸟分枝杆菌复合群(*Mycobacterium avium* complex, MAC), 其中鸟分枝杆菌亚种(*Mycobacteri-*

um subsp. Hominis, MAH)是与人类最相关的 MAC 生物体, 也是 T 细胞免疫缺陷个体感染的主要

* 【基金项目】 国家自然科学基金项目(No. 81760357)。

** 【通讯作者】 吴利先, E-mail: w_lixian@163.com

【作者简介】 马枫茜(1998-), 女, 贵州毕节人, 在读硕士研究生。研究方向: 感染与免疫。E-mail: mafengqianq@163.com

病原体^[1-2]。而关于病原体T细胞表位和人类基因组之间的交叉保护作用倍受关注^[3]。因此,了解MAH所致的NTM病的毒力决定因素和组分对抗非结核疫苗的研发和药物靶点的确定具有重要意义。MAV-4563蛋白属于MAH蛋白,是参与MAH致病的重要蛋白之一,参与鸟分枝杆菌在细胞与细胞间的扩散过程。本研究通过对MAV-4563基因及其编码蛋白的生物信息学分析,预测其二、三级结构和TB细胞表位等特性,为进一步了解MAV-4563蛋白引起机体抗TB免疫应答的机制及靶向疫苗研究提供重要线索。

材料与方法

1 鸟分枝杆菌 MAV4563 蛋白及其基因序列的获取

从NCBI全基因数据库(<https://www.ncbi.nlm.nih.gov>)中获取鸟分枝杆菌MAV-4563基因序列,从NCBI中通过BLAST获取该基因编码蛋白的氨基酸序列。

2 MAV-4563 蛋白的生物学信息分析

利用Protparam和ProtScaleon Expasy,等在线软件分析MAV-4563蛋白的氨基酸数,分子质量,等电点,分子式,稳定系数等理化性质;利用TMHMM Sever v 2.0,SignalP5.0,PSORT和Softberry等软件分析MAV-4563蛋白的跨膜区,信号肽该蛋白的分泌情况;利用Netphos 3.1 sever 和 Blast 等软件分析该蛋白的磷酸化位点和保守域;采用SOPMA和Phyre2预测MAV-4563的二级结构并建立其三维模型;利用IEDB,ABCpred和SYFPEITHI分别预测蛋白的B、T细胞表位,

结果

1 鸟分枝杆菌 MAV-4563 基因和编码蛋白基本信息

NCBI GenBank数据库显示鸟分枝杆菌MAV-4563基因全长756 bp,编码蛋白的氨基酸总数为251个,包含20种氨基酸。其中,Arg占13.1%,Leu占12.4%,Ala占10.0%。

2 MAV-4563 蛋白的理化性质

经Protparam软件分析MAV4563蛋白的分子式(formula)为C₁₂₆₄H₂₀₁₃N₃₈₅O₃₅₃S₇,由C、H、O、N、S等5种元素共4 022个原子组成(atomic composition),分子质量为2 847.568 ku,该蛋白携带正电荷残基数为(Asp+Glu)32个,携带负电荷残基数(Arg+Lys)37个,等电点(theoretical pI)为9.23,消光系数(280 nm)为31 190/(Mol·cm),吸光度(蛋白浓度为1 g/L)为1.095。当蛋白的N-末端是M(Met)的前提下,该蛋白哺乳动物网状细胞内的半衰期为30 h。MAV-4563蛋白的不稳定指数为42.57(>40),为不稳定蛋白。

白。

3 MAV-4563 蛋白的亲疏水性

Protparam预测MAV4563蛋白的脂肪指数(aliphatic index)为95.22,疏水性平均值(GRAVY)为-0.280,为疏水性蛋白。ProtScaleon Expasy分析该蛋白亲水性得分最大值为1.556,最小值为-2.656(图1)。正值表示该氨基酸疏水,负值表示该氨基酸亲水。

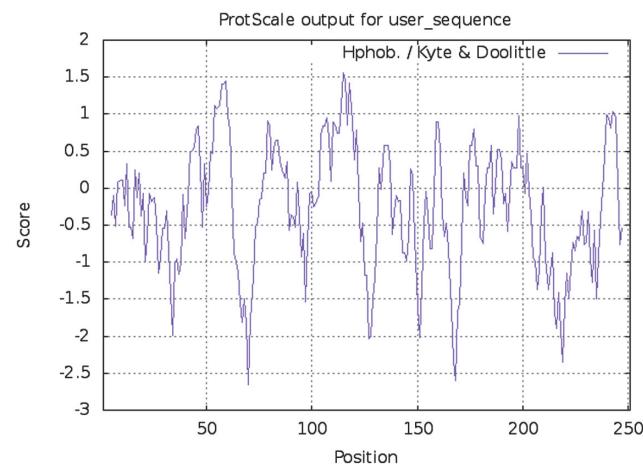


图1 MAV4563蛋白亲疏水性分析

Fig. 1 Hydropathicity of MAV4563

4 MAV-4563 蛋白的跨膜区和信号肽预测及亚细胞定位

经TMHMM Server v. 2.0预测MAV4563蛋白无跨膜区(图2)。SignalP4.1在线预测该蛋白无信号肽(图3)。前100个氨基酸中C_{max}评分为0.123;Y_{max}在第65位氨基酸,评分为0.109;S_{max}评分为0.133。可能的信号肽(1-64位氨基酸)平均S评分为0.098,可能的信号肽平均D评分为0.103,即MAV-4563无信号肽,说明该蛋白不是分泌蛋白。Softberry预测该蛋白的亚细胞定位在细胞质中的分数为7.3分,与PSORT软件预测该蛋白的亚细胞定位在细胞质中的结果一致。

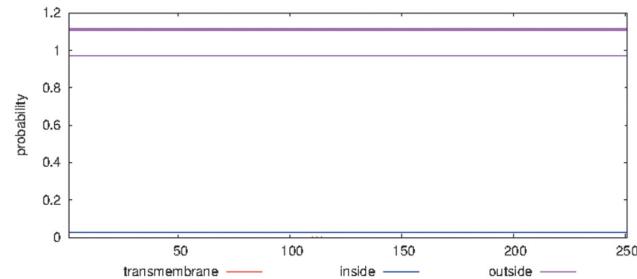


图2 MAV-4563蛋白跨膜区预测

Fig. 2 Predicted transmembrane region of MAV-4563 protein

5 MAV-4563 蛋白的磷酸化位点以及保守域预测

NetPhos3.1Servera在线软件预测MAV4563蛋

白有13个磷酸化位点。其中,丝氨酸(serine)磷酸化位点4个,分别位于24、61、125、137位氨基酸;苏氨酸(threonine)磷酸化位点6个,分别位于38、89、146、158、182、197位氨基酸;酪氨酸(tyrosine)磷酸化位点2个,分别位于7、13位氨基酸(图4)。利用BLAST工具分析预测MAV4563有两个结构域,属于泛素样折叠超家族,Mut7-C核糖核酸酶超家族(图5)。

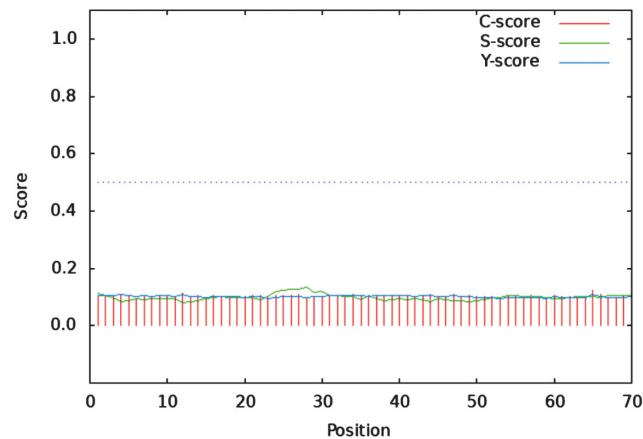
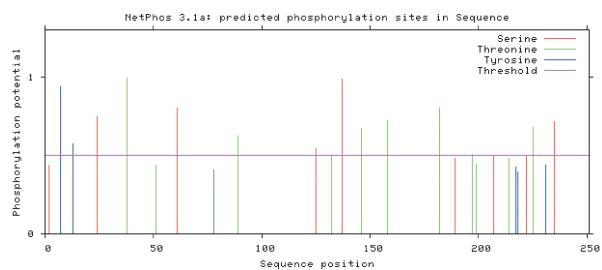


图3 MAV-4563蛋白信号肽预测
Fig. 3 MAV-4563 protein signal peptide prediction



注:serine 丝氨酸;threonine 苏氨酸;tyrosine 酪氨酸;Threshold 阈值。
图4 MAV-4563蛋白磷酸化位点分析
Fig. 4 Analysis of MAV-4563 protein phosphorylation site



图5 BLAST软件预测MAV-4563蛋白的结构域
Fig. 5 BLAST software predicts the structural domain of MAV-4563 protein

6 MAV4563蛋白的二级结构预测与三级同源建模

SOPMA在线分析软件预测MAV4563蛋白由 α -螺旋(Ah)、延伸链(Ee)、 β -转角(Tt)和无规则卷曲结构(Cc)组成,其比例分别为37.45%、15.94%、4.78%、41.83%(图6)。利用Phyre2软件分析MAV4563蛋白的三级结构,选择预测可信度为99%的d1vjka为模板进行三级同源建模,结果如图7。

7 MAV4563蛋白的B、T细胞表位

MAV4563蛋白经ABCpred在线工具预测得到多个可能B细胞抗原表位,其中得分>0.51的有28

个,得分>0.80(0.81-0.98)的有10个(表1)。

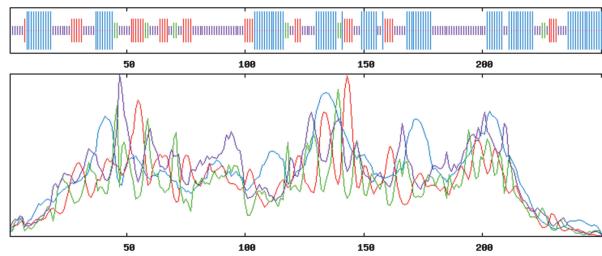


图6 MAV-4563蛋白的二级结构
Fig. 6 Secondary structure of MAV-4563 protein

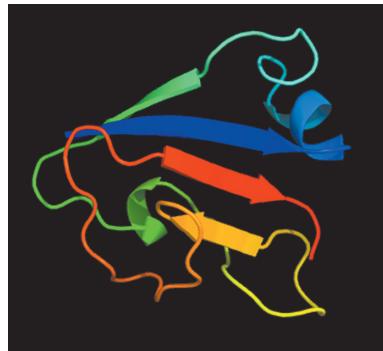


图7 MAV-4563蛋白的三级结构
Fig. 7 Tertiary structure of MAV-4563 protein

表1 MAV-4563蛋白B细胞抗原表位预测
Table 1 Predicted B-cell antigenic epitopes of MAV-4563 protein

排序Rank	表位序列Epitope sequence	起始位点Starting site	评分Rating
1	CGRIYWPMSGHHAKLLA	227	0.98
2	KRRAVTHGLFVRPDDP	153	0.92
3	DIGPTARLRPVPLRDP	85	0.91
4	AGDRIAAYPMFEALDI	71	0.87
5	SRGAAVRRPFRAHQTV	24	0.84
6	VSDQLLPLTRRYYDDF	206	0.83
7	EAMGIPHTEVDLIVVN	44	0.82
8	PMFEALDIGPTARLRP	79	0.81
9	TLTAVAKDEVSDLQLPP	197	0.81
10	GRLAWLLRLFGFDVWW	109	0.81

根据IEedb分析和ABCpred的预测结果,推测第31-37、50-71、82-95、111-125、153-164、206-220、229-239位氨基酸为MAV4563蛋白的B细胞表位可能性大。经SYFPEITHI软件预测MAV4563蛋白具有多个可能的CTL抗原表位,其中MHC类型为HLA-A*02:01限制性CTL表位,评分>20(21-28)分以上的有14个(表2),但是MHC类型为HLA-DRB1*0101、HLA-DRB1*0301、HLA-DRB1*0401、HLA-DRB1*0701限制性Th表位的检出数均为0。通过IEedb在线软件预测获得MAV4563蛋白的多个MHC-1和MHC-2型表位(表2,表3)。MHC-1型抗原表位,即HLA-A*01:01,HLA-A*02:01,HLA-A*03:01,Consensus Percentile Rank分值(该分值越

低,抗原表位的可能性越大) <1(0.28-0.925)的有4个,这4个肽段均来自C末端。其中分值最低的肽段氨基酸,即HLA-A*03:01,位于229-237位,其Consensus Percentile Rank分值为0.28;HLA-A*02:01肽段氨基酸的分值为0.8,位于239-247位,与SYFPEITHI软件预测的HLA-A*02:01的结果一致。MHC-2型表位,即HLA-DRB1*0301、HLA-DRB1*0401,adjusted-ran k分值(该分值越低,抗原表位的可能性越大) <1(0.38-0.89)的有10个,均来自C端的后60个氨基酸(表4)。

表2 SYFPEITHI软件预测MAV4563蛋白的CTL抗原表位
Table 2 CTL antigenic epitopes of MAV4563 protein predicted by SYFPEITHI software

等级 Rank	序列 Sequence	起始位点 Starting site	评分 Score
1	KLLALVERL	239	28
2	DLTGRLAPL	180	26
3	ILLTRDRGL	143	25
4	RLAPLSRCV	184	25
5	AMGIPHTEV	45	23
5	DLIVVNGSV	54	23
5	VIDVNLGRL	103	23
5	LLRLFGFDV	114	23
6	QALGVIRRL	171	22
7	GIPHTEVDL	47	21
7	LDIGPTARL	84	21
7	VNLGRLAWL	106	21
7	NLGRLAWLL	107	21
7	LLTRDRGLL	144	21

表3 IEDB预测MAV-4563蛋白的MHC-I型抗原表位
Table 3 MHC-I antigenic epitopes of MAV-4563 protein predicted by IEDB

等级 Rank	序列 Sequence	肽段位置 Peptide position	得分 Score
1	RIYWPGSHH	229-237	0.28
2	LTRDDGLLK	145-153	0.315
3	KLLALVERL	239-247	0.8
4	DADDQTLAA	127-135	0.925

表4 IEDB预测MAV-4563蛋白的MHC-II型抗原表位
Table 4 IEDB predicted MHC-II antigenic epitopes of MAV-4563 protein

等级 Rank	序列 Sequence	肽段位置 Peptide position	得分 Score
1	RDPRFVIDVNGLRLA	98-112	0.38
2	PRFVIDVNGLRLAWL	100-114	0.46
3	LRDPRFVIDVNGLGRL	97-111	0.60
4	HRILLTRDRGLLKRR	141-155	0.67
4	RILLTRDRGLLKRRRA	142-156	0.67
5	QHRILLTRDRGLLKRR	140-154	0.68
6	ILLTRDRGLLKRRRAV	143-157	0.69
7	EQHRILLTRDRGLLK	139-153	0.70
8	RFVIDVNGLRLAWLL	101-115	0.89

讨 论

目前尚无针对MAV感染的疫苗,并且在预防免

疫缺陷病引发的一系列疾病中缺乏一种成功的、高效的和长期免疫力的预防性疫苗^[4-5],因此研发治疗性疫苗通过增强患者的免疫反应治愈MAV感染成为免疫治疗的主要方向之一。研究表明,治疗性疫苗可以加速疾病缓解,缩短治疗时间,并可能提高治疗成功率^[6]。

在MAC中,MAH表现出广泛的基因突变导致该菌能够在各种环境和宿主中生存^[7]。有研究表明,MAH形成的生物膜和微团聚体可以穿过宿主的呼吸道或肠道黏膜与宿主相互作用,成为引起MAH持续感染的第一关键因素^[8-10]。微团聚体结合蛋白(Mbp-1)和微团聚体侵袭蛋白(Mip-1)诱导MAH生物膜和微团聚体的形成,使其附着于宿主上皮细胞并通过mip-1与宿主蛋白相互作用从而侵入宿主上皮细胞^[11]。在入侵过程中,涉及细菌表面与宿主细胞膜之间的相互作用。MAH表面暴露的蛋白质可促进其粘附以利于穿过粘膜屏障,并且可能在克服宿主免疫反应而在感染早期的宿主巨噬细胞中存活和复制方面发挥重要作用^[12-13]。因此,通过识别和阻断生物膜形成过程中必需的细胞蛋白,就可以减少细菌负荷并潜在地干扰细菌感染,从而起到一定的预防作用。MAV-4563蛋白属于MAH蛋白,可能参与MAH表面蛋白形成生物膜过程,并且在MAH在宿主中的存活也发挥重要作用^[14]。BLAST分析显示MAV-4563蛋白含有泛素样折叠超家族,该家族是一个小而保守,是功能独特的重要蛋白质储存库。该家族成员与许多癌症、炎症、免疫缺陷病等疾病有关^[15],除了参与经典的蛋白质降解途径外,还参与DNA修复、激酶调控和信号转导。泛素结合域参与人类疾病的发病机制,使得Ub-UBD相互作用成为开发新型治疗药物的潜在靶点^[16]。运用在线工具分析MAV-4563蛋白属于疏水性不稳定蛋白,无跨膜区,无信号肽,位于细胞质内,且含有13个磷酸化位点。其二级结构主要为无规则卷曲,占41.83%,表明其容易与抗体嵌合。预测该蛋白含有多个B、T细胞的抗原优势表位,通过IEDB综合分析确定31-37、50-71、82-95、111-125、153-164、206-220、229-239位氨基酸为MAV4563蛋白的B细胞表位的可能性大。经SYFPEITHI和IEDB在线软件分析,筛选出我国HLA基因分布频率较高的等位基因HLA-A*01:01,HLA-A*02:01,HLA-A*03:01限制性CTL表位,MHC-II类型为HLA-DRB1*0101、HLA-DRB1*0301、HLA-DRB1*0401、HLA-DRB1*0701限制性Th表位。通过分析,MAV-4563蛋白的CTL表位集中在145-237氨基酸残基区域内,其Th细胞表位主要集中在97-157氨基酸残基区域内。MAV-4563蛋白的C端后90个氨基酸区域含有丰富

的B细胞和CTL细胞表位,证明MAV-4563蛋白通过其C端后90位的肽链在激活特异性T、B细胞免疫应答中至关重要,可作为T细胞疫苗候选抗原。

有研究认为T细胞疫苗可以重新激活细胞毒性T细胞,具有清除病原体的能力,从而提供促炎环境,即是免疫系统抵御病原体的重要驱动力^[17]。MAV感染的适应性免疫主要是由Th1细胞介导的,抗结核分枝杆菌疫苗在具有免疫能力的个体中可能对非结核分枝杆菌感染表现出交叉保护,且治疗性结核病疫苗的主要目标是增强针对结核分枝杆菌的Th1型细胞免疫反应^[4-6,18]。因此,研发治疗性T细胞疫苗成为非结核分枝杆菌免疫治疗的理想策略。T细胞疫苗是基于细胞毒性T细胞或辅助性T细胞表位的识别而合成,与整体的免疫原性潜力有关,因而T细胞表位含量越高,就越有可能引起免疫反应,从而达到免疫治疗的目的^[3]。本研究预测分析MAV-4563蛋白的T细胞表位含量丰富且位于C端的后90个氨基酸区域,易于激活细胞免疫。因此表明该蛋白是一种具有潜力的T细胞疫苗候选抗原,为非结核分枝杆菌新型疫苗的研发奠定了基础。

【参考文献】

- [1] Koh WJ. Nontuberculous mycobacteria-overview[J]. Microbiol Spectr, 2017, 5(1): TNM17-0032-2016.
- [2] To K, Cao R, Yegiazaryan A, et al. General overview of Nontuberculous mycobacteria opportunistic pathogens: *Mycobacterium avium* and *Mycobacterium abscessus*[J]. J Clin Med, 2020, 9(8): 2541.
- [3] De Groot AS, Moise L, Terry F, et al. Better epitope discovery, precision immune engineering, and accelerated vaccine design using immunoinformatics tools. [J]. Front Immunol, 2020(11): 442.
- [4] Kannan N, Haug M, Steigedal M, et al. *Mycobacterium smegmatis* vaccine vector elicits CD4⁺ Th17 and CD8⁺ Tc17 T cells with therapeutic potential to infections with *Mycobacterium avium* [J]. Front Immunol, 2020(11): 1116.
- [5] Larijani MS, Ramezani A, Sadat S M. Updated Studies on the development of HIV therapeutic vaccine[J]. Curr HIV Res, 2019, 17(2): 75-84.
- [6] Afkhami S, Villela AD, D'Agostino MR, et al. Advancing immunotherapeutic vaccine strategies against pulmonary tuberculosis. [J]. Front Immunol, 2020(11): 557809.
- [7] Jeong-Ih S, Sung J S, Min-Kyoung S. Differential genotyping of *Mycobacterium avium* complex and its implications in clinical and environmental epidemiology[J]. Microorganisms, 2020, 8(1): 98.
- [8] Babrak L, Danelishvili L, Rose SJ, et al. The environment of " *Mycobacterium avium* subsp. *Hominis*" microaggregates induces synthesis of small proteins associated with efficient infection of respiratory epithelial cells. [J]. Infect Immun, 2015, 83(2): 625-636.
- [9] Mullis SN, Falkinham JO. Adherence and biofilm formation of *Mycobacterium avium*, *Mycobacterium intracellulare*, and *Mycobacterium abscessus* to household plumbing materials[J]. J Appl Microbiol, 2013, 115(3): 908-914.
- [10] Rindi L, Garzelli C. Genetic diversity and phylogeny of *Mycobacterium avium*[J]. Infect Genet Evol, 2014(21): 375-383.
- [11] Babrak L, Danelishvili L, Rose SJ, et al. Microaggregate-associated protein involved in the invasion of epithelial cells by *Mycobacterium avium* subsp. *hominis*[J]. Virulence, 2015, 6(7): 694-703.
- [12] Harriff MJ, Danelishvili L, Wu M, et al. *Mycobacterium avium* genes MAV_5138, and MAV_3679 are transcriptional regulators that play a role in the invasion of epithelial cells, in part by their regulation of CipA, a putative surface protein interacting with host cell signaling pathways. [J]. J Bacteriol, 2009, 191(4): 1132-1142.
- [13] McNamara M, Tzeng S, Maier C, et al. Surface proteome of " *Mycobacterium avium* subsp. *Hominis*" during the early stages of macrophage infection. [J]. Infect Immun, 2012, 80(5): 1860-1868.
- [14] Bermudez LE, Danelishvili L, Babrak L, et al. Evidence for genes associated with the ability of *Mycobacterium avium* subsp. *hominis* to escape apoptotic macrophages[J]. Front Cell Infect Microbiol, 2015(5): 63.
- [15] Randles L, Walters KJ. Ubiquitin and its binding domains. [J]. Front Biosci (Landmark edition), 2012, 17 (6): 2140-2157.
- [16] Sokratous K, Hadjisavvas A, Diamandis EP, et al. The role of ubiquitin-binding domains in human pathophysiology[J]. Crit Rev Clin Lab Sci, 2014, 51(5): 280-290.
- [17] Tye GJ, Lew MH, Choong YS, et al. Vaccines for TB: Lessons from the past translating into future potentials[J]. J Immunol Res, 2015(2015): 916780.
- [18] Schijns V, Fernandez-Tejada A, Barjaktarovic, et al. Modulation of immune responses using adjuvants to facilitate therapeutic vaccination[J]. Immunol Rev, 2020, 296(1): 169-190.

【收稿日期】 2022-04-18 【修回日期】 2022-07-02