

DOI:10.13350/j.cjpb.230103

• 论著 •

EB病毒 BZLF1 基因表位疫苗的预测及构建*

丁一¹, 李昆芳¹, 薛庆节¹, 屈艳琳^{2**}

(1. 济宁医学院基础医学院 山东济宁 272067; 2. 济宁医学院形态学实验室)

【摘要】 **目的** 采用生物信息学方法预测 EB 病毒 BZLF1 基因编码蛋白的结构、抗原表位及疫苗的初步构建, 为 EB 病毒相关疾病的疫苗的研制提供依据。 **方法** 通过 NCBI 获得 EB 病毒 BZLF1 蛋白的核苷酸序列和氨基酸序列, 采用 ORF Finder、ProtParam、SOPMA、DTU Health Tech、TMHMM-2.0、Cell-PLOC 2.0、NetPhos 3.1 Server、NetNGlyc、SWISS MODE、Immunomedicine Group、IEDB、BLAST、Uniprot 等生物信息学软件对 EB 病毒 BZLF1 蛋白的理化性质、亲疏水性、结构域、信号肽、跨膜区、亚细胞定位、磷酸化及糖基化位点、二级结构、三级结构、抗原决定簇、B 细胞与 T 细胞抗原表位、蛋白同源性及相互作用蛋白进行预测分析。在 T4DNA 连接酶作用下连接 BZLF1 基因与穿梭表达载体 pMV261, 对连接产物进行 PCR 及双酶切鉴定。 **结果** BZLF1 蛋白是由 245 个氨基酸组成的亲水蛋白, 分子式为 $C_{1185}H_{1850}N_{332}O_{366}S_8$, 相对分子质量为 26.860×10^3 , 理论等电点为 5.25, 脂肪族氨基酸指数为 71.35, 平均亲水系数为 -0.529; 该蛋白二级结构以无规则卷曲居多, 占 51.84%, 无信号肽和跨膜区, 存在 18 个磷酸化位点, 2 个糖基化位点, 含有 1 个碱性亮氨酸拉链结构。预测该蛋白有 9 个抗原决定簇, 8 个 B 细胞优势抗原表位, 9 个 CTL 细胞优势表位, 8 个 Th 细胞表位, 与人类环状 amp 反应元件结合蛋白相似度为 6.12%, 同源性较低。PCR 及双酶切鉴定含 BZLF1 基因的 pMV261 载体构建正确。 **结论** 生物信息学预测 BZLF1 属于亲水蛋白, 热稳定性较好。该蛋白含有多个磷酸化位点可参与相关信号通路, 对疾病的发展起到重要作用。BZLF1 含有多个 B、T 细胞抗原表位, 且与人体蛋白同源性低, 不易发生免疫交叉反应, 可作为 EB 病毒候选疫苗表位。

【关键词】 EB 病毒; BZLF1; 生物信息学; 抗原表位

【中图分类号】 R373

【文献标识码】 A

【文章编号】 1673-5234(2023)01-0012-07

[Journal of Pathogen Biology. 2023 Jan;18(1):12-18.]

Prediction and construction of Epstein-Barr virus BZLF1 gene epitope vaccine

DING Yi¹, LI Kun-fang¹, XUE Qing-jie¹, QU Yan-lin² (1. School of Basic Medicine, Jining Medical College, Jining, Shandong 272067; China; 2. Morphology laboratory, Jining Medical College) ***

【Abstract】 **Objective** The bioinformatics method was used to predict the structure and epitope of the protein encoded by the Epstein-Barr virus BZLF1 gene and the preliminary construction of the vaccine, so as to provide the basis for the development of the vaccine of Epstein-Barr virus related diseases. **Methods** The nucleotide and amino acid sequences of EB virus BZLF1 protein were obtained by NCBI. ORF Finder, ProtParam, SOPMA, DTU Health Tech, TMHMM-2.0, Cell-PLOC 2.0, NetPhos 3.1 Server, NetNGlyc, SWISS MODE, Immunomedicine Bioinformatics software such as Group, IEDB, BLAST and Uniprot analyzed the physicochemical properties, hydrophilicity and hydrophobicity, domain, signal peptide, transmembrane region, subcellular localization, phosphorylation and glycoylation sites, secondary structure, tertiary structure, antigen epitopes of B cell and T cell, protein homology and interaction of EB virus BZLF1 protein. Predictive analysis with protein. T4 DNA ligase was used to connect BZLF1 gene to shuttle expression vector pMV261, and PCR and double enzyme digestion were used to identify BZLF1 gene. **Results** BZLF1 protein is a hydrophilic protein composed of 245 amino acids, the molecular formula is $C_{1185}H_{1850}N_{332}O_{366}S_8$, the relative molecular weight is 26.860×10^3 , the theoretical isoelectric point is 5.25, the aliphatic amino acid index is 71.35, the average hydrophilic coefficient is -0.529. The majority of the secondary structure of the protein was random coil, accounting for 51.84%, no signal peptide and transmembrane region, there were 18 phosphorylation sites, 2 glycoylation sites, and 1 basic leucine zipper structure. It was predicted that the protein had 9 epitopes, 8 B-cell dominant epitopes, 9 CTL dominant epitopes and 8 Th cell epitopes. The similarity to human cyclic amp reaction element binding protein was 6.12%, indicating low

* **【基金项目】** 山东省重点研发计划项目(No. 2018GSF118137); 济宁市重点研发计划项目(No. 2019SMNS020); 贺林院士工作站重点项目(No. JYHL2019ZD03); 济宁医学院 2022 年省级大学生创新训练计划项目(No. cx2022047z); 济宁医学院大学生创新训练计划项目(No. cx2021093)

** **【通讯作者】** 屈艳琳, E-mail: 2753167672@qq.com

【作者简介】 丁一(2002-), 女, 山东济南人, 研究方向: 微生物与肿瘤。学历: 临床医学专业本科在读 E-mail: 970092271@qq.com

homology. The pMV261 vector containing BZLF1 gene was correctly constructed by PCR and double enzyme digestion.

Conclusion Bioinformatics predicted that BZLF1 was a hydrophilic protein with good thermal stability. The protein contains multiple phosphorylation sites and is involved in related signaling pathways, which play an important role in the development of the disease. BZLF1 contains multiple B and T cell antigen epitopes, and has low homology with human proteins, so it is not easy to have immune cross-reaction, and can be used as a candidate Epstein-Barr virus vaccine epitopes.

【Key words】 Epstein-Barr virus(EBV); BZLF1; Bioinformatics; Antigen epitopes

EB病毒是1964年发现的第一个人类致癌病毒,属于 γ 型疱疹病毒,在人群中广泛传播,目前已潜伏感染约95%的世界人口。EB病毒主要感染B淋巴细胞和口咽部上皮细胞,具有潜伏期和裂解期,其中裂解期与多种肿瘤及癌症的发病机制相关,例如伯基特淋巴瘤、传染性单核淋巴增多症、鼻咽癌、EBV阳性胃癌等^[1]。多项研究证实BZLF1蛋白作为转录因子参与EBV复制起点的激活,该转录因子通过与CpG甲基化的病毒DNA结合^[2],从而促进裂解期蛋白的表达,潜伏在细胞中的病毒向裂解期转化,增强病毒对细胞的感染^[3]。此外,BZLF1编码蛋白与抑癌蛋白p53相互作用,通过泛素蛋白酶体途径降解p53,病毒在裂解期会激活DNA损伤反应,导致p53基因的C端磷酸化,进一步增加BZLF1与p53的亲合力,导致肿瘤的生长及转移^[4]。

目前,临床治疗EB病毒相关疾病主要为放化疗和免疫治疗,虽然这些疗法一定程度上可以缓解病情,但毒副作用较大,患者预后较差,复发率高,机体难以形成针对EB病毒的长期免疫反应^[5]。因此,寻找一种针对EB病毒蛋白的特异性疫苗很有必要。本研究对BZLF1蛋白进行生物信息学预测,并构建BZLF1的pMV261载体,为筛选有良好免疫原性的EB抗原蛋白疫苗奠定基础。

材料与方 法

1 BZLF1氨基酸序列及其编码基因序来源

登录NCBI数据库(<https://www.ncbi.nlm.nih.gov/>),搜索并下载EBV BZLF1蛋白的氨基酸序列和核苷酸序列。BZLF1基因登录号为3783744,编码蛋白Zta登录号为YP_401673.1。通过NCBI查询BZLF1蛋白由245个氨基酸组成,其序列如下:

MMDPNSTSEDVKFTDPDYQVPFVQAFDQA-
TRVYQDLGGPSQAPLPCVLPVLPPEPLPQGQLT-
AYHVSTAPTGSWFSA PQPAPENAYQAYAAPQL-
FPVSDITQNNQTNQAGGEAPQPGDNSTVQTAA-
AVVFACPGANQGQQLADIGVPQPAPVAAPARR-
TRKPQQPESLEECDSELEIKRYKNRVASRKCRA-
KFKQLLQHYREVAAAKSENDRRLRLLKQMCP-
SLDVDSIIPRTPDVLHEDLLNF.

2 方 法

2.1 BZLF1生物信息学预测

2.1.1 BZLF1基因开放阅读框预测 采用ORF Finder(<http://www.ncbi.nlm.nih.gov/orffinder>)在线网址预测BZLF1基因开放阅读框的数目。

2.1.2 BZLF1蛋白的理化性质及亲疏水性预测 运用protparam(<https://www.expasy.org/protparam>)预测BZLF1蛋白的分子质量、理论等电点、氨基酸组成、消光系数、半衰期等理化性质以及蛋白的亲疏水性。

2.1.3 信号肽、跨膜结构及结构域的预测 采用Conserved domains(<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)预测蛋白结构域;运用<http://services.healthtech.dtu.dk/service.php?SignalP-5.0>和<http://services.healthtech.dtu.dk/service.php?TMHMM-2.0>预测目的蛋白的信号肽和跨膜结构。

2.1.4 空间结构预测 运用SOPMA软件(http://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl)预测BZLF1蛋白的二级结构;运用SWISS MODEL(<https://swissmodel.expasy.org/>)预测BZLF1蛋白的三级结构。

2.1.5 磷酸化和糖基化位点预测及亚细胞定位 运用NetPhos 3.1 Server(<http://www.cbs.dtu.dk/services/NetPhos/>)预测BZLF1蛋白的磷酸化位点;NetNGlyc(<http://www.cbs.dtu.dk/services/NetNGlyc/>)预测BZLF1蛋白的糖基化位点;运用Cell-Ploc 2.0(<http://www.csbio.sjtu.edu.cn/bioinf/Cell-Ploc-2/>)和Uniport预测BZLF1蛋白在宿主细胞的定位。

2.1.6 B细胞表位预测 通过IEDB在线网站(https://www.iedb.org/home_v3.php)对BZLF1蛋白的B细胞线性表位,可塑性,抗原性, β 转角等性质进行预测,进一步筛选出可塑性强、抗原性强、 β 转角较多的序列作为优势B细胞表位。

2.1.7 T细胞表位预测 通过IEDB在线网站(https://www.iedb.org/home_v3.php)对BZLF1蛋白的T细胞表位进行预测,利用数据库中的MHC I

和 MHC II 分子进行分析,筛选出与样本 MHC 分子回应值 100% 的序列作为 T 细胞优势表位。

2.1.8 同源性分析及抗原决定簇预测 利用 EXPASY(<https://web.expasy.org/blast>)在线网站,输入 BZLF1 蛋白的氨基酸序列,选择 Homo Sapiens 进行查询。利用 Immunomedicine Group (<http://imed.ucm.es/Tools/antigenic.pl>)在线网站对 BZLF1 蛋白的抗原决定簇进行分析。

2.1.9 相互作用蛋白预测 运用 Uniprot 生信软件对 BZLF1 的相互作用蛋白进行预测,输入 BZLF1 名称,选择 UniProtKB 数据库进行查询预测。

2.2 重组载体的构建与鉴定

2.2.1 BZLF1 与载体 pMV261 的连接 根据 BZLF1 碱基序列,由上海生物工程有限公司(青岛分公司)合成引物。BZLF-F: 5'-GCCGGATCCATGATGGACCCAACTCGACTTC-3' (引入 BamHI 酶切位点); BZLF-R: 5'-GCGGTCGACTAAGAAATTTAAGAGATCCTCGTG-3' (引入 SaII 酶切位点)。BZLF1 片段扩增条件均为:PCR 反应体系 25 μl,其成分为酶 1 μl、dNTP 2 μl、Buffer 5 μl、模板 0.5 μl、上游引物和下游引物分别 0.1 μl、水 16.8 μl。反应条件:98 °C 预变性 4 min;98 °C 变性 10 s,61 °C 退火 30 s,72 °C 延伸 70 s,共 35 个循环,72 °C 再延伸 5 min。将酶切 pMV261 和 BZLF1 基因片段进行连接。连接反应体系(40 μl):BZLF1 25 μl,pMV261 5 μl,T4 连接酶 5 μl,Buffer 5 μl。混匀后 16 °C 连接过夜。

2.2.2 重组载体的鉴定 采用 PCR 和双酶切法两种方法进行重组质粒的鉴定。PCR 反应体系(25 μl):5 × PrimerSTAR GXL Buffer 5 μl,dNTP Mixture 2 μl,PrimerSTAR GXL DNA Polymerase 0.5 μl,BZLF-R 0.1 μl,模板 0.5 μl,ddH₂O 补充至 25 μl。PCR 反应程序:98 °C 预变性 4 min;98 °C 变性 10 s,60 °C 退火 15 s,68 °C 延伸 60 s,重复 34 个循环;68 °C 再延伸 5 min。取 PCR 产物进行 1.5% 琼脂糖凝胶电泳,回收目的片段并测序。

重组质粒双酶切反应体系:重组质粒 1 μg,BamHI 1 μl,SaII 1 μl,10 × T Buffer 3 μl,加 ddH₂O 补充至 20 μl,混匀后置 30 °C 水浴锅 8 h,取酶切产物进行 1% 琼脂糖凝胶电泳,回收后测序。

结 果

1 预测 BZLF1 基因开放阅读框的数目

BZLF1 基因位于全基因组(NC_007605.1)的 89838-90943 位,总长度为 1106nt。BZLF1 基因共有 7 个开放阅读框,其中 ORF1 为最长的开放阅读框,起始密码子为 ATG,终止密码子为 TAG,编码 200 个氨

基酸,与目的蛋白全长几乎一致,表明 BZLF1 蛋白全长可基本被翻译(图 1)。



图 1 BZLF1 基因开放阅读框分析
Fig. 1 Open reading frame analysis of BZLF1 gene

2 BZLF1 蛋白的理化性质及亲疏水性

ProtParam (<https://www.expasy.org/prot-param>)预测 BZLF1 蛋白由 245 个氨基酸组成,分子式为 C₁₁₈₅H₁₈₅₀N₃₃₂O₃₆₆S₈,相对分子质量为 26.860 × 10³,理论等电点为 5.25;当蛋白的 N 端序列为 Met 时,该蛋白在哺乳动物网织红细胞体外实验中的半衰期为 30 h,在大肠埃希菌中的半衰期 > 10 h,其中以脯氨酸(11.8%)、甘氨酸(11.4%)、谷氨酰胺(10.2%)、亮氨酸(8.2%)所占比例较高,无吡咯赖氨酸和硒代半胱氨酸,带负电荷(天冬氨酸+谷氨酸)的残基总数为 26,带正电荷(精氨酸+赖氨酸)的残基总数为 21。该蛋白的消光系数为 21 680,280 nm 处的吸光度为 0.87。脂肪族氨基酸指数为 71.35,平均亲水系数为 -0.529,为亲水蛋白。

3 信号肽、跨膜结构及结构域

Conserved domains 预测 BZLF1 蛋白在 175-226 氨基酸序列存在碱性亮氨酸拉链结构(图 2)。通过在线网站 Signal-5.0 和 TMHMM-2.0 预测 BZLF1 蛋白无信号肽和跨膜结构(图 3)。

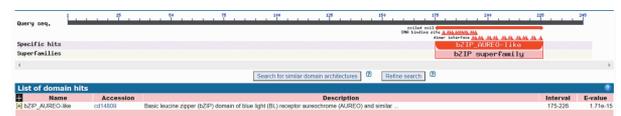


图 2 BZLF1 蛋白结构域预测
Fig. 2 Prediction of BZLF1 protein domain

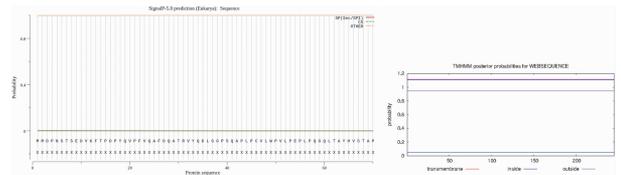


图 3 BZLF1 蛋白信号肽(左)及跨膜结构(右)预测
Fig. 3 Prediction of BZLF1 protein signal peptide (Left) and transmembrane structure (Right)

4 BZLF1 蛋白的空间结构

通过 SOMPA 预测蛋白的二级结构,无规则卷曲占 51.84%,α 螺旋占 35.92%,延伸片段占 7.76%,β 转角占 4.49%(图 4)。无规则卷曲占比高表明该蛋白有较好的抗体嵌合性。采用 SWISS MODEL 预测的三级结构与数据库中的 2c9n.1.C 模板相似度 100%(图 5)。

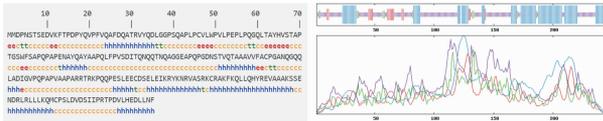
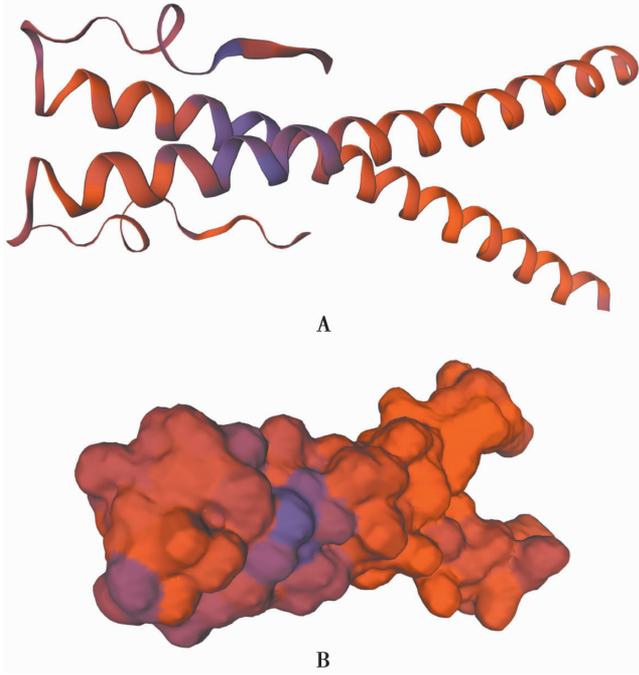


图 4 BZLF1 蛋白二级结构
Fig. 4 The Secondary structure of BZLF1 protein



A 飘带模型 B 表面模型
图 5 BZLF1 蛋白三级结构
A Ribbon model B Surface model
Fig. 5 Tertiary structure of BZLF1 protein

5 BZLF1 蛋白磷酸化和糖基化位点及亚细胞定位

阈值设定为 0.5 时,运用 NetPhos 3.1Server 预测 BZLF1 共有 18 个潜在磷酸化位点,其中丝氨酸位点有 9 个,分别在 6、8、40、73、167、173、186、208、209 位,苏氨酸位点有 7 个,分别在 7、14、71、101、124、159、234 位,酪氨酸有 2 个在 86 和 89 位(图 6)。NetNGlyc 预测 BZLF1 蛋含有白糖基化位点 2 个,在第 5 和 1 定 19 位,概率分别为 72.82%和 69.47%(图 7)。通过 Cell-Ploc 2.0 和 Uniport 分析,BZLF1 蛋白主要位于宿主细胞核(图 8)。

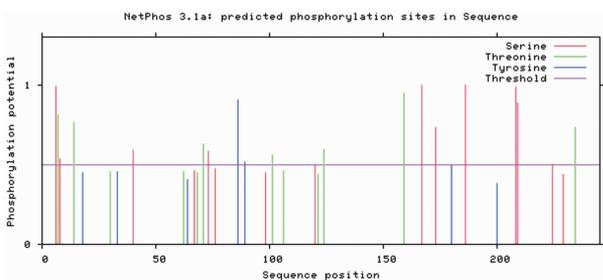


图 6 BZLF1 蛋白磷酸化位点分析
Fig. 6 Predicted phosphorylated sites of BZLF1 protein

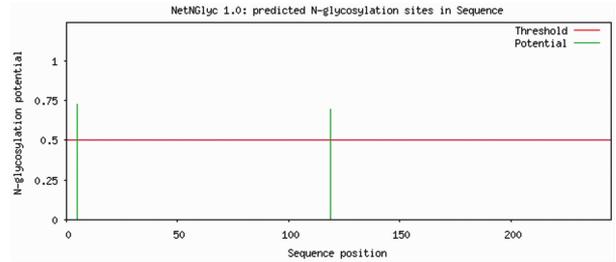


图 7 BZLF1 蛋白糖基化位点分析
Fig. 7 Predicted glycosylation sites of BZLF1 protein

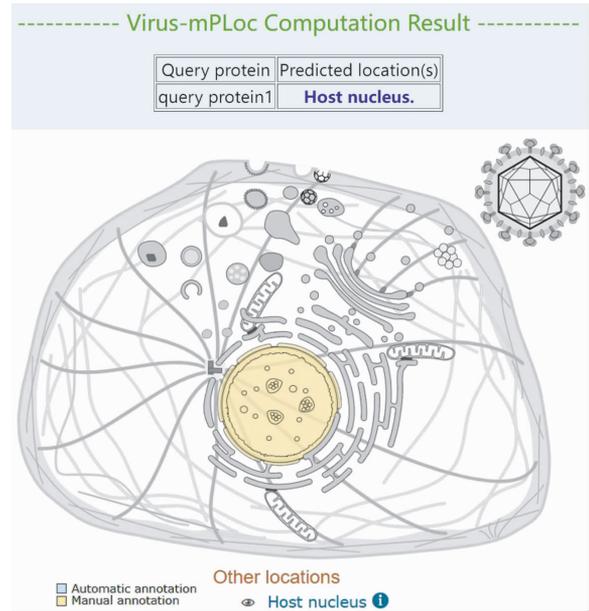


图 8 BZLF1 蛋白亚细胞定位分析
Fig. 8 Subcellular localization analysis of BZLF1 protein

6 BZLF1 抗原蛋白的 B 细胞表位

通过 IEDB 在线网站 (https://www.iedb.org/home_v3.php) 预测 BZLF1 蛋白的优势 B 细胞表位。蛋白的可塑性与抗体的结合度相关,选择可塑性大于 1 的序列,再结合线性表位、抗原性、β 转角三个性质进行筛选,均选取评分高于平均值的序列,最终得出 B 细胞的优势表位有 8 个(表 1)。

表 1 BZLF1 蛋白 B 细胞优势表位预测
Table 1 Prediction of dominant B cell epitopes of BZLF1 protein

| 序号 No. | 起始氨基酸位置 Epitope Starting Position | 氨基酸序列 Epitope Sequence | 终止氨基酸位置 Epitope Ending Position |
|-----------|---|------------------------------|---------------------------------------|
| 1 | 2 | MDPNSTS | 8 |
| 2 | 12 | KFTDPDY | 18 |
| 3 | 53 | PEPLPQG | 59 |
| 4 | 105 | QTNQAGG | 111 |
| 5 | 143 | DIGVPQP | 149 |
| 6 | 158 | RTRKPQQ | 164 |
| 7 | 208 | SSENDRL | 214 |
| 8 | 222 | CPSLDVD | 228 |

7 BZLF1 抗原蛋白的 T 细胞表位

运用 IEDB 筛选出 BZLF1 蛋白的优势 T 细胞表位(图 9)。其中 CTL 细胞优势表位有 9 个(表 2), Th 细胞优势表位有 8 个(表 3)。综合分析 BZLF1 蛋白 T 细胞的优势表位在 54-64、172-183、190-198、208-214 位点。



图 9 BZLF1 蛋白优势 T 细胞表位分析
Fig.9 Analysis of dominant T cell epitopes of BZLF1 protein

表 2 BZLF1 蛋白 CTL 细胞优势表位分析
Table 2 Analysis of dominant epitopes of BZLF1 protein CTL

| 序号 No. | MHC I 类分子亚型 MHC class I molecular subtype | 起始氨基酸位置 Epitope Starting Position | 氨基酸序列 Epitope Sequence | 终止氨基酸位置 Epitope Ending Position |
|--------|---|-----------------------------------|------------------------|---------------------------------|
| 1 | HLA-B | 54 | EPLPQQQLTAY | 64 |
| 2 | | 77 | APQPAPENAY | 86 |
| 3 | | 172 | DSELEIKRYKNR | 183 |
| 4 | | 44 | LPCVLWPVL | 52 |
| 5 | | 190 | RAKFKQLLQ | 198 |
| 6 | | 209 | SENDRLRL | 217 |
| 7 | HLA-C | 187 | RKCRAKFKQLLQH | 199 |
| 8 | | 197 | LQHYREVAA | 205 |
| 9 | HLA-A | 195 | QLLQHYREV | 203 |

表 3 BZLF1 蛋白 Th 细胞优势表位分析
Table 3 Analysis of dominant epitopes of BZLF1 protein in Th cells

| 序号 No. | 起始氨基酸位置 Epitope Starting Position | 氨基酸序列 Epitope Sequence | 终止氨基酸位置 Epitope Ending Position |
|--------|-----------------------------------|------------------------|---------------------------------|
| 1 | 2 | MDPNSTS | 8 |
| 2 | 12 | KFTPPDY | 18 |
| 3 | 53 | PEPLPQG | 59 |
| 4 | 105 | QTNQAGG | 111 |
| 5 | 143 | DIGVPQP | 149 |
| 6 | 158 | RTRKPQQ | 164 |
| 7 | 208 | SENDRL | 214 |
| 8 | 222 | CPSLDVD | 228 |

8 BZLF1 蛋白的同源性分析及抗原决定簇预测

Immunomedicine Group 在线网站分析 BZLF1 蛋白的平均抗原倾向性为 1.0399, 蛋白表面可形成 9 个抗原决定簇, 分别位于氨基酸序列的 9-26、28-35、40-68、85-99、121-134、137-154、167-173、185-205、213-241 位点(图 10), 显示该蛋白抗原倾向指数较高, 易形成抗原决定簇的位点较多, 是一种免疫原性较好的蛋白。BLAST 预测 BZLF1 蛋白氨基酸序列的第 162-194 位点和人类环状 amp 反应元件结合蛋白(C9J896)的 221-251 序列相似性较高, 为 45%(15/33), 占总序列的 6.12%(15/245)(图 11)。利用核苷酸序列测出与

人类的多聚免疫球蛋白受体(P01833)有 41%的相似性, 占总序列的 1.8%(图 12)。BZLF1 蛋白与人类蛋白的同源性较低, 不易发生免疫交叉反应, 免疫原性较好, 可作为设计疫苗的潜在蛋白。

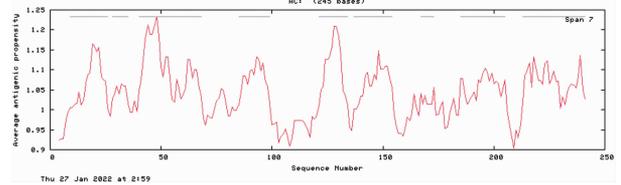


图 10 BZLF1 蛋白抗原决定簇预测
Fig.10 Prediction of BZLF1 protein epitopes

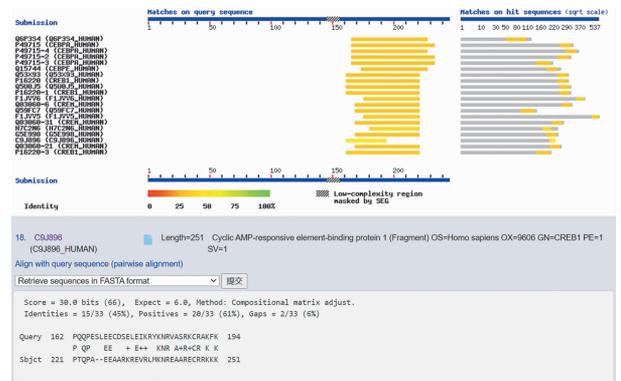


图 11 基于氨基酸序列预测 BZLF1 蛋白同源性
Fig.11 Amino acid sequence predicts homology of BZLF1 protein

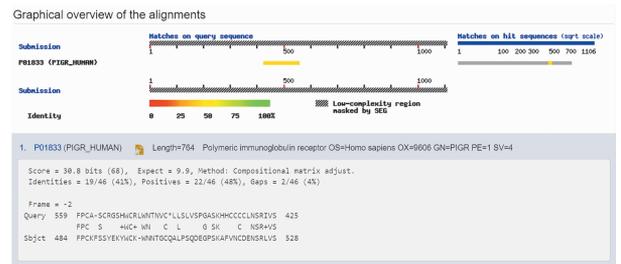


图 12 基于核苷酸序列预测 BZLF1 蛋白同源性
Fig.12 Nucleotide sequence predicts homology of BZLF1 protein

9 BZLF1 相互作用蛋白

运用 Uniprot 软件预测 BZLF1 可分别与人类 CREB 转录共激活因子 2 和活化蛋白 C 激酶 1 的受体相互作用(图 13)。CREB 转录共激活因子 2 活化状态为脱磷酸化, 通过识别 cAMP 反应元件(CRE)位点激活转录, 在 SIK/TORC 信号通路中作为共激活剂, 并且在 LKB1/AMPK/TORC2 信号通路中调节糖异生。活化蛋白 C 激酶 1 的受体为 40S 核糖体亚基的组成部分, 参与核糖体质量控制(RQC)的启动, 结合并稳定活化的蛋白激酶 C(PKC), 招募 PKC 到核糖体中, 导致 EIF6 的磷酸化, 还能通过延长细胞周期的 G0/G1 来抑制细胞生长。因此, BZLF1 可通过蛋白相互作用介导疾病进展。

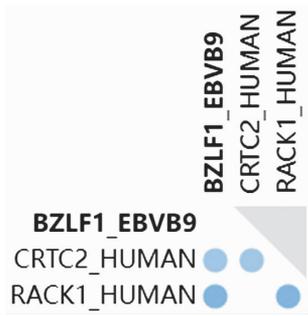
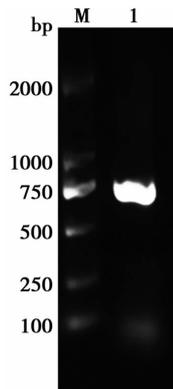


图 13 BZLF1 相互作用蛋白预测
Fig. 13 BZLF1 interaction protein prediction

10 重组质粒的构建及鉴定

PCR 扩增 BZLF1 基因, 扩增片段大小为 738 bp (图 14), 与预期相符。将 BZLF1 基因片段和载体经酶切后连接, 连接产物进行双酶切鉴定, 结果如图 15。酶切目的片段与预期相符。测序碱基序列结果与生物公司购买目的片段附带的序列列表比对, 同源性 100%。

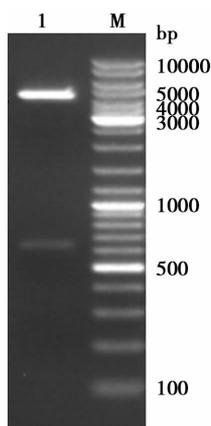


M DNA 标志物(DL2000) 1 BZLF1 DNA PCR 产物

图 14 BZLF1 DNA PCR 产物 1%琼脂糖凝胶电泳分析

M DNA markers(DL2000) 1 DNA PCR product of BZLF1 was amplified by PCR

Fig. 14 PCR electrophoresis identification



M DNA 标志物(DL10000) 1 重组质粒双酶切

图 15 重组质粒双酶切鉴定

M DNA markers(DL10000) 1 Double digestion product

Fig. 15 Identification of recombinant plasmid by double enzyme digestion

讨论

近年来生物信息学技术不断发展, 蛋白质与基因数据库不断更新, 因此可通过大数据对基因组学及蛋白组学的相关信息进行分析、整合, 从而研究疾病进展的分子机制、设计免疫原性强的疫苗、研发特异性好的靶向药物等, 为疾病的诊断及治疗提供更多元、更精准的方法^[6]。

EB 病毒主要以潜伏或裂解的方式感染宿主细胞, 其中病毒的裂解诱导淋巴瘤及癌症的进展。BZLF1 在病毒潜伏期的建立和裂解期的激活方面发挥重要作用, 因此有必要对该蛋白进行研究^[7]。EBV 即刻早期基因 BZLF1 的产物是一种多功能蛋白, 其 N 端为反式激活域, C 端存在碱性亮氨酸拉链(bZIP)区域^[8], 其中 bZIP 区域中的 RRTRK 氨基酸序列与甲基化的病毒 DNA 亲和力高^[9]。病毒潜伏期间, EBV 基因组的 CpG 序列逐渐甲基化, BZLF1 的表达激活病毒转录因子 Rta。在 Rta 的辅助作用下, BZLF1 蛋白特异性识别甲基化病毒 DNA 的 TGAGTCA 序列来激活裂解期相关蛋白基因的转录, 从而促进病毒的复制和对宿主细胞的感染^[10]。

BZLF1 不仅在诱导病毒裂解期基因复制中发挥重要作用, 而且还能通过多种信号通路改变宿主环境, 利于病毒逃避免疫系统的识别及转移。相关研究已证实 BZLF1 可抑制先天免疫反应中相关信号通路, 如 NF- κ B 和 TNF 信号通路。其中 BZLF1 与 TNF- α 启动子结合以抑制 NF- κ B 的激活, 进而抑制相关促炎症因子的表达; BZLF1 通过降低 TNFR1 蛋白启动子的活性来抑制其表达, 从而阻断 TNF 信号通路^[11-12]。

本研究通过 NCBI 数据库获取 BZLF1 基因序列及其编码蛋白氨基酸序列, 利用生物信息学分析软件对蛋白的理化性质及亲疏水性进行预测。BZLF1 蛋白为亲水性蛋白, 脂肪族氨基酸指数较高, 说明其热稳定性较高, 利于疫苗的储存和运输。该蛋白的 175-226 氨基酸序列存在碱性亮氨酸拉链结构, 可与病毒 DNA 结合促进其向裂解期转化, 但不存在信号肽及跨膜区。预测该蛋白含有 18 个磷酸化位点, 主要为丝氨酸、苏氨酸、酪氨酸残基, 提示 BZLF1 可能通过磷酸化和去磷酸化介导相关信号通路。预测该蛋白含有 2 个糖基化位点, 表明 BZLF1 蛋白可作为诊断 EB 病毒相关疾病的潜在标志物。该蛋白主要定位于宿主细胞核发挥作用, 与文献^[11]报道一致。其二级结构中无规则卷曲占比高, 为 51.84%, 与抗体嵌合性好, 容易形成潜在抗原表位。采用 IEDB 在线网站分析 BZLF1 蛋白的 B 细胞表位, 筛选可塑性好、抗原性强、二级结构较多的表位作为优势 B 细胞表位。病毒感染期间,

B细胞介导的体液免疫不可或缺,其通过分泌多种细胞因子阻碍病毒感染^[14]。预测结果显示,BZLF1有8个优势表位,可通过良好的免疫原性来诱导B细胞发生免疫反应,使机体获得长期免疫保护。通过IEDB在线网站筛选出BZLF1的CTL细胞优势表位有9个,Th细胞优势表位有8个,综合分析54-64、172-183、190-198、208-214表位序列可引起较强的T细胞反应。BZLF1蛋白与人类蛋白的同源性较低,不易发生免疫交叉反应,可作为设计疫苗的潜在蛋白。

之前针对EB病毒疫苗的研究主要集中于gp350包膜蛋白,该疫苗不能有效预防EBV感染,因此尚未普及接种^[15]。本研究采用生物信息学方法预测BZLF1蛋白含有多个T、B细胞表位,免疫原性良好,且与人体蛋白同源性低,不易发生免疫交叉反应。PMV261质粒是一种工程菌——大肠埃希菌的穿梭质粒,有多个限制性内切酶酶切位点,能稳定携带并表达目的基因,是一种理想的基因表达载体。本研究探索了pMV261质粒和BZLF1的构建条件,并成功构建出BZLF1的重组pMV261质粒,为针对EB病毒BZLF1的表位疫苗的研发奠定了实验基础。

【参考文献】

- [1] Yin H, Qu J, Peng Q, et al. Molecular mechanisms of EBV-driven cell cycle progression and oncogenesis [J]. *Med Microbiol Immunol*, 2019, 208(5): 573-583.
- [2] Murata T, Sugimoto A, Inagaki T, et al. Molecular basis of Epstein-Barr virus latency establishment and lytic reactivation [J]. *Viruses*, 2021, 13(12): 2344.
- [3] Yang J, Deng W, Hau PM, et al. Epstein-Barr virus BZLF1 protein impairs accumulation of host DNA damage proteins at damage sites in response to DNA damage[J]. *Lab Invest*, 2015, 95(8): 937-950.
- [4] Tornesello ML, Annunziata C, Tornesello AL, et al. Human Oncoviruses and p53 Tumor suppressor pathway deregulation at the origin of human cancers[J]. *Cancers (Basel)*, 2018, 10(7): 213.
- [5] Hartlage AS, Liu T, Patton JT, et al. The Epstein-Barr Virus lytic protein BZLF1 as a candidate target antigen for vaccine development[J]. *Cancer Immunol Res*, 2015, 3(7): 787-794.
- [6] Canzonieri R, Lacunza E, Abba MC. Genomics and bioinformatics as pillars of precision medicine in oncology [J]. *Medicina (B Aires)*, 2019, 79(Spec 6/1): 587-592.
- [7] Van Sciver N, Ohashi M, Pauly NP, et al. Hippo signaling effectors YAP and TAZ induce Epstein-Barr Virus (EBV) lytic reactivation through TEADs in epithelial cells[J]. *PLoS Pathog*, 2021, 17(8): e1009783.
- [8] Kusano S, Ikeda M. Interaction of phospholipid scramblase 1 with the Epstein-Barr virus protein BZLF1 represses BZLF1-mediated lytic gene transcription[J]. *J Biol Chem*, 2019, 294(41): 15104-15116.
- [9] Weber E, Buzovetsky O, Heston L, et al. A noncanonical basic motif of Epstein-Barr Virus ZEBRA protein facilitates recognition of methylated DNA, high-affinity dna binding, and lytic activation [J]. *J Virol*, 2019, 93(14): e00724-19.
- [10] Bernaudat F, Gustems M, Gunther J, et al. Structural basis of DNA methylation-dependent site selectivity of the Epstein-Barr virus lytic switch protein ZEBRA/Zta/BZLF1[J]. *Nucleic Acids Res*, 2022, 50(1): 490-511.
- [11] Kraus RJ, Cordes BA, Sathiamoorthi S, et al. Reactivation of Epstein-Barr Virus by HIF-1 α requires p53[J]. *J Virol*, 2020, 94(18): e00722-20.

【收稿日期】 2022-08-08 【修回日期】 2022-10-25