DOI:10.3969/j.issn.1002-2694.2021.00.081

·论　著·

结核分枝杆菌 Hrp1 蛋白的生物信息学分析

刘　静[1]，吴　芳[2]，张　杰[3]，董江涛[3]，柳小玲[2]，梁　粟[3]，王　菊[2]，
张春军[2]，吴江东[2]，章　乐[2]，赵海军[3]，张万江[2]

摘　要：目的　应用生物信息学预测分析 Hrp1 蛋白结构功能和生物学特性。**方法**　生物信息学软件 ORF Finder、SOSUI、Protpapram、ProtScale、Signal IP、Tmpred、TMHMM、NetNGlyc、NetPhos-3.1-Servera、PSORT 、WoLF PSORT、TargetP-2.0-CBS、NLStradamus、SOMPA、SWISS-MODEL、IEDB、SYFPEITHI、RANKPEP、BLAST、DNAStar、STRING 对 Hrp1 蛋白的理化性质、结构功能、生物学特性、抗原表位及蛋白相互作用进行预测分析。**结果**　Hrp1 蛋白原子总数 2 166，等电点 4.96，不稳定性指数 19.62，脂肪族氨基酸指数 100.35，平均疏水性 0.042，有跨膜区无信号肽，1 个糖基化位点 6 个磷酸化位点，二级结构以 α-螺旋为主。其亚细胞定位于细胞质，属于稳定的亲水性蛋白。经预测，Hrp1 蛋白含有 4 个 B 细胞表位，多个 T 细胞表位，且其相互作用蛋白分别为 Rv2627c、Rv2628、Rv1738、devR、Rv2624c、TB31.7、rip3、pfkB、guaA、hspX，并参与多种生物学过程。**结论**　生物信息学分析 Hrp1 蛋白为稳定亲水性胞浆蛋白，具有良好的抗原表位，为结核病诊断及治疗的潜在候选因子，是探索结核分枝杆菌持久性感染机制的重要研究靶标，可作为抗结核多肽疫苗的候选蛋白。

关键词：结核分枝杆菌；Hrp1 蛋白；生物信息学；结构

中图分类号：R378.91　　**文献标识码：**A　　**文章编号：**1002－2694(2021)08－0669－09

Bioinformatic analysis of the Hrp1 protein of *Mycobacterium tuberculosis*

LIU Jing[1], WU Fang[2], ZHANG Jie[3], DONG Jiang-tao[3], LIU Xiao-ling[2], LIANG Su[3], WANG Ju[2], ZHANG Chun-jun[2], WU Jiang-dong[2], ZHANG Le[2], ZHAO Hai-jun[3], ZHANG Wan-jiang[2]

(1. *Xi'an Jiaotong University City College, Shanxi* 710018, *China*;
2. *Department of Pathophysiology, Shihezi University School of Medicine/Xinjiang High Incidence of Local and Ethnic Diseases, Key Laboratory of the Ministry of Education, Shihezi* 832002, *China*;
3. *The First Affiliated Hospital of Medical College of Shihezi University, Shihezi* 832002, *China*)

Abstract: We applied bioinformatics to predict, analyze and understand the structure, function, and biological characteristics of Hrp1 protein. Bioinformatics software, including ORF Finder, SOSUI, Protpapram, ProtScale, Signal IP, Tmpred, TMHMM, NetNGlyc, NetPhos-3.1-Servera, PSORT, WoLF PSORT, TargetP-2.0-CBS, NLStradamus, SOMPA, SWISS-MODEL, IEDB, SYFPEITHI, RANKPEP, BLAST, DNAStar and STRING, were used to predict and analyze the physicochemical properties, structure, function, biological characteristics, and epitope and protein interaction of Hrp1 protein. The results showed that Hrp1 protein has a total number of atoms of 2 166, an isoelectric point of 4.96, an instability index of 19.62, an aliphatic amino acid index of 100.35 and an average hydrophobicity of 0.042. It additionally has a transmembrane region, no signal peptide, one glycosylation site and six phosphorylation sites. Its secondary structure is dominated by α-helices, and its subcellular localization is in the cytoplasm. It is thus a stable hydrophilic protein. Bioinformatic prediction and analysis indicated that Hrp1 protein contains four B cell epitopes and multiple T cell epitopes; its interacting proteins are Rv2627c, Rv2628, Rv1738, devR, Rv2624c, TB31.7, rip3, pfkB, guaA and hspX; and it participates in a variety of biological processes. Through bioinformatic analysis, this study reveals that Hrp1 is a stable hydrophilic cytoplasmic protein that has good epitopes and may be a potential candidate factor in the diagnosis and treatment of tuberculosis. Therefore, Hrp1 is an important research target in exploring the mechanism of persistent mycobac-

NSFC-新疆联合基金项目(No.U1803127)、国家自然基金项目(No.82060021)和兵团重点领域科技攻关计划项目(No.2018AB019)联合资助

通讯作者：张万江，Email：zwj1117@126.com；
ORCID:0000-0002-0294-473X
赵海军，Email：380585216@qq.com；
ORCID:0000-0002-9191-9607

作者单位：1.西安交通大学城市学院，陕西　710018；
2.石河子大学医学院病理生理学教研室/"新疆地方与民族高发病"教育部重点实验室，石河子　832002；
3.石河子大学医学院第一附属医院，石河子　832002

terium tuberculosis infection and may be used as a candidate protein for the development of an anti-tuberculosis polypeptide vaccine.

Keywords: *Mycobacterium tuberculosis*; Hrp1; bioinformatics; structure

Supported by the Grant from National Natural Science Foundation Project (No. U1803127), the Grant from National Natural Science Foundation Project (No. 82060021); Key Science and Technology Research Projects in Key Areas of the Corps (No. 2018AB019).

Corresponding authors: Zhang Wan-jiang, Email: zwj1117@126.com; Zhao Hai-jun, Email: 380585216@qq.com

全世界约有 1/3 的人口感染了结核分枝杆菌(*Mycobacterium tuberculosis*,MTB),其中大多数是潜伏性感染(Latent Tuberculosis Infection,LTBI),目前无有效治疗 LTBI 患者的疫苗和药物,早发现结核杆菌潜伏感染患者,并对其进行诊断治疗,可以控制后期结核病患者人数[1]。Rv2626c 基因的表达是 MTB 处于休眠期的重要标志之一,Rv2626c 抗原受控于休眠调节子 DosR (dormancy regulon),编码缺氧反应蛋白 1(hypoxic response protein 1, Hrp1), Hrp1 蛋白具有较好的抗原性和免疫原性,能有效刺激机体产生一定的细胞免疫及较高的抗体水平,并且该蛋白易被 LTBI 患者所识别[2],多项研究提示其很有可能会成为一种新型诊断 MTB 潜伏感染的候选抗原和治疗潜伏性感染的候选疫苗。为此,本文对 Hrp1 蛋白进行生物信息学研究分析,对其结构功能进行全面分析预测,为进一步了解 MTB 长期潜伏感染机制提供理论依据,以期为筛选 MTB 诊断标志物和研制新型结核疫苗奠定基础。

1 材料与方法

1.1 一般资料 从美国国家生物技术信息中心 NCBI 网站(https://www.ncbi.nlm.nih.gov/)获取。在 GenBank 中检索关键词 Rv2626c,以获取 H37Rv 中 Rv2626c 全基因组信息,Rv2626c 基因位于 MTB 全基因组的 2952562-2952993 区域,基因 ID 为 888576;其编码蛋白为低氧反应蛋白 1(hypoxic response protein 1, Hrp1)属于分泌蛋白,在蛋白质库中的登录号为 NP-217142.1,在 uniprot 中的登录号为 P9WJA3。

通过 NCBI 网站查询,Hrp1 蛋白由 143 个氨基酸序列组成,检索的氨基酸序列如下:MTTARDIMNAGVTCVGEHETLTAAAQYMREHDIGALPICGDDDRLHGMLTDRDIVIKGLAAGLDPNTATAGELARDSIYYVDANASIQEMLNVMEEHQVRRVPVISEHRLVGIVTEADIARHLPEHAIVQFVKAICSPMALAS。

1.2 方 法

1.2.1 Rv2626c 基因开放阅读框架分析及其编码蛋白的理化性质及亲疏水性分析 分析 Rv2626c 基因的开放阅读框架,采用 NCBI 网站中的 ORF Finder 软件(http://www.ncbi.nlm.nih.gov/orffinder)。上传 Rv2626c 基因编码的氨基酸序列至 Protpapram 生信软件(https://web.expasy.org/protparam)及 ProtScale 生信软件(http://web.expasy.org/protscale),分析其编码蛋白 Hrp1 蛋白的理化性质,包括氨基酸数目、分子式、分子质量、氨基酸组成、等电点、亲(疏)水性等。

1.2.2 Hrp1 蛋白的信号肽、跨膜区、糖基化及磷酸化位点预测 运用信号肽分析系统 SOSUI(http://harrier.nagahama-i-bio.ac.jp/sosui/)在线软件及 Signal IP(http://www.cbs.dtu.dk/services/SignalP/)分析 Hrp1 的信号肽及其位置;运用跨膜区预测软件 Tmpred(https://embnet.vital-it.ch/software/TMPRED-form.ht)分析 Hrp1 蛋白跨膜区;TMHMM(http://www.cbs.dtu.dk/services/TMHMM/)软件输入 Hrp1 蛋白的氨基酸序列,对该蛋白的拓扑结构进行预测分析[3]。利用分析 Hrp1 的糖基化位点采用 NetNGlyc 软件(http://www.cbs.dtu.dk/services/NetNGlyc/);预测 Hrp1 蛋白的磷酸化位点采用 NetPhos 3.1 Server 磷酸化分析软件(http://www.cbs.dtu.dk/services/NetPhos/)。

1.2.3 Hrp1 蛋白的亚细胞定位分析 联合运用 3 种在线软件:PSORT(https://www.genscript.com/psort.html)、WoLF PSORT(https://wolfpsort.hgc.jp/)和 TargetP-2.0-CBS(http://www.cbs.dtu.dk/services/TargetP/)对 Hrp1 蛋白进行亚细胞定位分析;预测 Hrp1 蛋白的核定位信号,运用 NLStradamus 在线软件(http://www.moseslab.csb.utoronto.ca/NLStradamus/)。

1.2.4 Hrp1 蛋白结构预测 运用 SOPMA(https://npsa-prabi.ibcp.fr/cgi-bin/secpred_sopma.pl)预测 Hrp1 蛋白的二级结构;SWISS-MODEL 在

线软件（https://swissmodel.expasy.org）对 Hrp1 蛋白进行三级结构同源性建模分析。

1.2.5　Hrp1 蛋白与人类蛋白的同源性分析　采用 EXPASY 进入主页（https://web.expasy.org/blast/），将 Hrp1 蛋白的氨基酸序列输入后，点击 Homo sapiens，确定 Run BLAST 后，可对 Hrp1 蛋白氨基酸序列与人类蛋白的同源性进行 BLAST 分析。

1.2.6　Hrp1 蛋白的 B 细胞表位、辅助性 T 细胞表位、细胞毒性 T 淋巴细胞表位预测　在 IEDB（http://www.iedb.org/home-v3.php）网站，输入氨基酸序列，通过 Karplus&Schulz 分析柔韧性、Kolaskar&Tongaonkar 预测抗原性、Emini 预测抗原表面可及性、Parker 预测亲水性及 Chou&Fasman 法预测氨基酸编码蛋白质的 β 转角，对 B 细胞表位进行联合预测分析；β 转角与无规则卷曲多出现在蛋白表面，为凸出疏松结构，与抗体更有利结合，成为 B 细胞抗原表位可能性较大[4]；最后选取亲水性与表面可及性良好、柔韧性强、抗原性好、最好为卷曲和转角这种可能性大的区域作为候选 B 细胞表位，尽量避开 α 螺旋、β 折叠结构。通过 SYFPEITHI 在线生信软件（http://www.syfpeithi.de/0-Home.htm）对 Hrp1 抗原蛋白 Th 表位进行预测[5]，进入 SYFPEITHI 首页，选择 HLA-DRB1 * 0101、0401、0701、0801、1101、1501 这 6 种 HLA 基因分型进行预测辅助 T 细胞表位，选取得分为 22 分以上的氨基酸序列，综合分析后筛选出候选的辅助 T 细胞表位。利用 RANKPEP 在线生信软件进行预测，进入预测界面（http://imed.med.ucm.es/Tools/rankpep.html），将 Hrp1 氨基酸序列输入，选择 SYFPEITHI 所预测的 HLA 基因分型，找出在线软件红色标记的结果，筛选出结果较好的多肽序列作为候选的细胞毒性 T 淋巴细胞表位。最后采用 DNAStar 软件中 Protean 模板得到关于 Hrp1 蛋白表位预测的位置图，该图可直观看到 Hrp1 蛋白表位富集的区域，具体在图中表现为红色方块标记的区域。

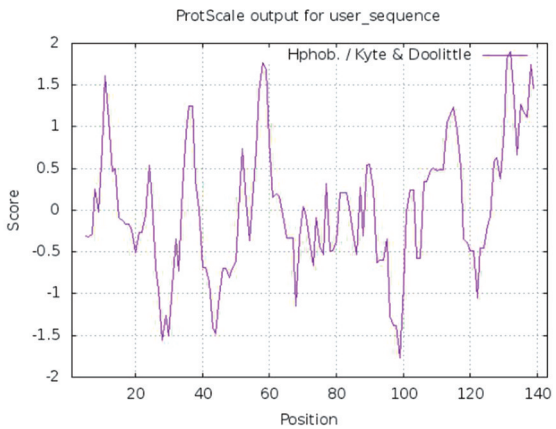
1.2.7　Hrp1 蛋白的相互作用蛋白预测及富集分析　将 Hrp1 全长氨基酸序列上传至 STRING 在线软件（http://string-db.org/cgi/input.pl），对 Hrp1 的相互作用蛋白进行预测，其生物过程进行功能富集分析。

2　结　果

2.1　Rv2626c 基因开放阅读框架分析及其编码蛋白的理化性质及亲疏水性分析结果

2.1.1　Rv2626c 基因开放阅读框架分析　结果显示 Rv2626c 基因（ID：888576）位于 MTB（H37Rv）的全基因组（NC-000962.3）的 2952562-2952993 区域，全长 432 bp，起始密码子为 ATG，终止密码子为 TAG，Rv2626c 基因含有 5 个开放阅读框，其中最长的一个开放阅读框显示 Rv2626c 基因基本被通读，全长 432 bp，与预测的编码 Hrp1 蛋白长度基本相似。

2.1.2　Rv2626c 基因编码蛋白质的理化性质及亲疏水性分析　根据 Protpapram 预测结果，该蛋白的理论等电点 PI 为 4.96，分子量为 15 517.72；该蛋白由 143 个氨基酸构成，包含 19 种氨基酸，其中以丙氨酸（A）（13.3%）、异亮氨酸（I）（8.4%）、缬氨酸（V）（8.4%）3 种氨基酸的组成比例较高，不包括色氨酸（W）、吡咯赖氨酸（O）和硒半胱氨酸（U）。带负电荷的残基（Asp＋Glu）总数为 21，带正电荷的残基（Arg＋Lys）总数为 11。Hrp1 蛋白分子式为 $C_{667}H_{1084}N_{194}O_{211}S_{10}$，原子总数 2 166，Rv2626c 基因编码蛋白的消光系数为 4 595，280 nm 处的吸光度为 0.296。当蛋白的 N 端残基为 Met 时，该蛋白在哺乳动物网状细胞中的半衰期为 30 h，在酵母体内的半衰期大于 20 h，在大肠杆菌体内的半衰期大于 10 h。Hrp1 蛋白不稳定性指数为 19.62，为稳定蛋白，脂肪族氨基酸指数为 100.35，平均亲水为 0.042。ProtScale 预测进一步分析显示 Rv2626c 基因编码蛋白为亲水性蛋白，第 99 位氨基酸亲水性得分最高为－1.767；第 132 为氨基酸疏水性得分最高为 1.9（图 1）。



Note：Abscissa for the amino acid position，the vertical axis for the hydrophobic score（pro，hydrophobic，respectively，with negative values，positive values）

图 1　Hrp1 蛋白的亲（疏）水性分析

Fig.1　Affinity analysis of Hrp1 protein

2.2 Hrp1 蛋白的信号肽、跨膜区、糖基化及磷酸化位点分析 运用信号肽分析系统 SOSUI 结果显示该蛋白为可溶性蛋白,无信号肽,该蛋白是具有一个跨膜区长度为 23 个氨基酸序列的跨膜螺旋,N 端氨基酸位于第 1 位,C 端位于第 16 位,跨膜区氨基酸序列为 MTTARDIMNAGVTCVG;Signal IP 结果显示该蛋白 max. C 值为 0.131,max. Y 值为 0.117,max. S 值为 0.143,mean S 值为 0.107,其中 mean S 值若大于 0.5,则预测为分泌蛋白,存在信号肽;故可知该蛋白没有信号肽,不是分泌蛋白而是膜蛋白。TMpred 分析 Hrp1 蛋白在 17-33、127-143 氨基酸之间可能存在跨膜结构,无信号肽(图 2)。

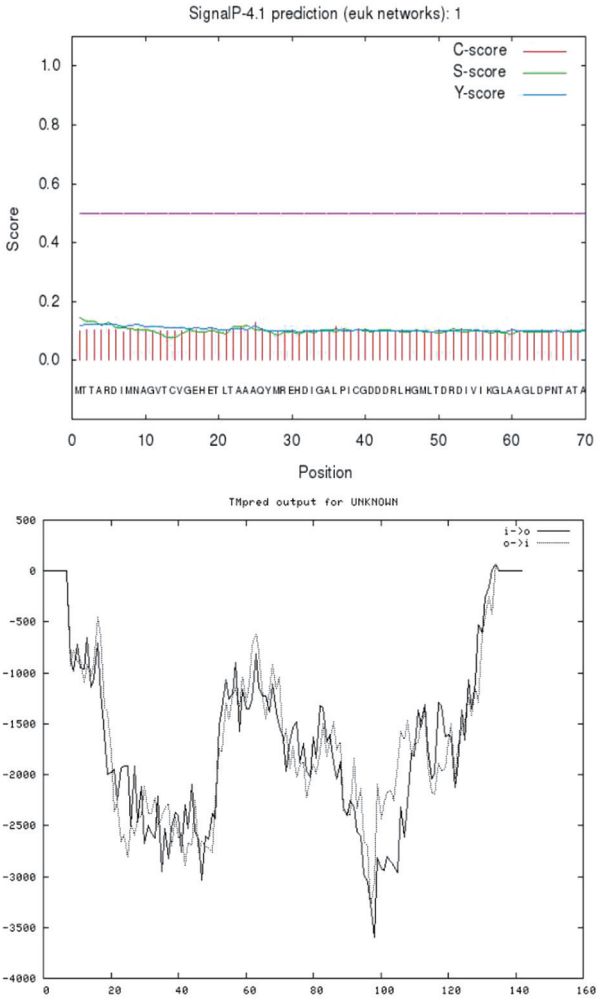
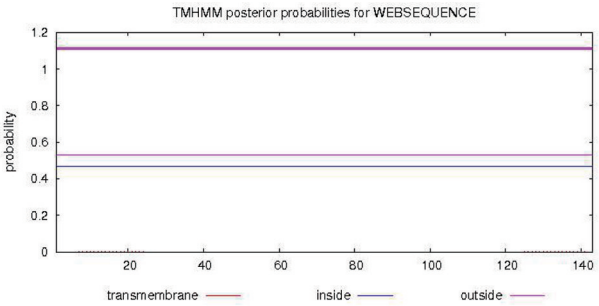


图 2 Hrp1 蛋白跨膜结构预测(左)与信号肽预测(右)

Fig.2 Predicted protein transmembrane structure (left) and signal peptide (right)

TMHMM 提交蛋白序列的全部氨基酸残基共 143 个,结果预测分析得到该蛋白(图 3)跨膜螺旋氨基酸数为 0.007 25(该指标>18 提示很有可能为跨膜蛋白),N 端前 60 个氨基酸中跨膜氨基酸数为 0.003 25,N 端位于膜胞质侧的总概率为 0.469 45。综合分析表明 Hrp1 蛋白位于膜外,为非跨膜蛋白。



The axis of represents the amino acid residue number (1-210) corresponding to the submitted protein sequence. The vertical axis values are the probabilities that each amino acid is inside the membrane, outside the membrane, and the helix region.

图 3 TMHMM Server.2.0 预测 Hrp1 蛋白跨膜螺旋

Fig.3 TMHMM Server. 2.0 prediction of the Hrp1 protein transmembrane helix

在线软件 NetNGlyc 分析 Hrp1 的糖基化位点存在于 84 位(图 4);采用 NetPhos 3.1 Servera 在线软件预测 Hrp1 蛋白有 6 个磷酸化位点,其中磷酸化丝氨酸位点有 2 个,分别位于 77、106 位氨基酸;磷酸化苏氨酸位点有 2 个,分别位于 3、50 位氨基酸;2 个磷酸化酪氨酸位点,分别位于 79、80 位氨基酸(图 5)。

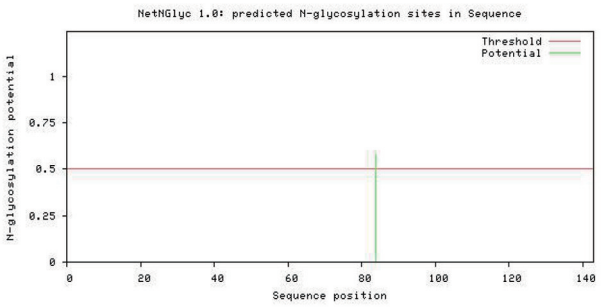


图 4 Hrp1 蛋白亚基糖基化位点分析

Fig.4 Predicted O-glycosylation sites of Hrp1 protein

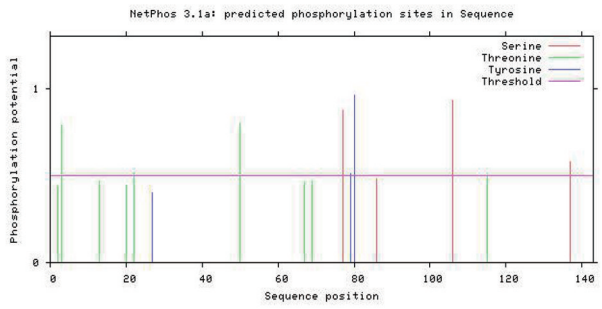
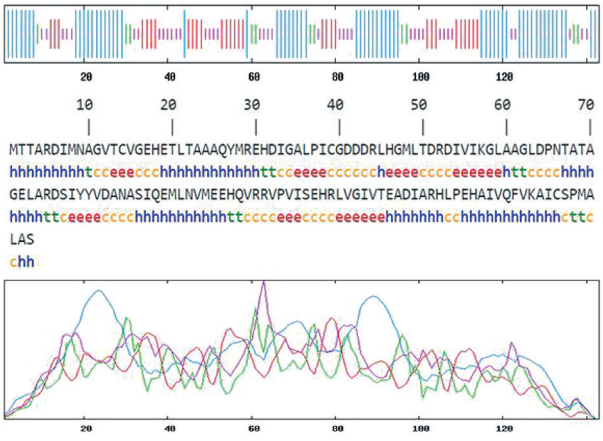


图 5 Hrp1 蛋白磷酸化位点分析

Fig.5 Predicted phosphorylation sites of Hrp1 protein

2.3　Hrp1 蛋白的亚细胞定位分析　PSORT 结果显示该蛋白其亚细胞定位于细胞质，可靠性为 94.1%；在线软件 WoLF PSORT 分析结果得出 Cysk 为 12，cyto 为 9.5，mito 为 4，nucl 为 1.5，cyto-nucl 为 6，表明 Hrp1 蛋白可能定位于细胞质；TargetP-2.0-CBS 结果显示该蛋白属于分泌蛋白的可能性为 0.000 1，定位在其他位置的可能性为 0.999 7；NLStradamus 在线软件预测 Hrp1 蛋白无 NLS，可推测此蛋白位于细胞核外。综合 4 种在线软件分析可得 Hrp1 蛋白亚细胞定位于细胞质。

2.4　Hrp1 蛋白结构预测分析　利用在线分析软件 SOPMA 预测 Hrp1 蛋白的二级结构，α-螺旋（Hh）63 个，占 44.06%；β-转角（Tt）11 个，占 7.69%；β-折叠（Ee）30 个，占 20.98%；无规则卷曲（Cc）39 个，占 27.27%（图 6）。该蛋白结构较稳定。



Blue red green and purple stand for alpha helix extended strand beta turn and random coil respectively

图 6　Hrp1 抗原蛋白二级结构预测

Fig.6　Prediction of the secondary structure of Hrp1 antigen protein

利用 SWISS-MODEL 软件，对 Hrp1 蛋白的三级结构进行预测分析，选择同源性最高的 1xkf.1 为模板，基于 Hrp1 蛋白的 X 射线分析进行同源三级建模（图 7）。其 GMQE 评分为 0.91，QMEAN 评分为 1.47，手势向上，说明该模型结果可靠，质量较好。

2.5　同源性分析结果　BLAST 结果得出，Hrp1 中氨基酸序列在第 11-70 位，与人类蛋白质锌指蛋白 440（ZNF440）的 96-153 位氨基酸相似性较高，但二者的相似度仅为 25%，因此，Hrp1 与人类蛋白相比，同源性较低，交叉反应发生的概率较低，可以作为候选的新型结核疫苗和结核病诊断靶点，该抗原蛋白的表位可以进行生信分析，对结核病诊断与候选表位疫苗有一定的参考研究价值。在研究分析中，即在预测和筛选表位时，应注意避免选择位于 11-70 的氨基酸序列，以降低交叉反应的概率[6]。



图 7　Hrp1 蛋白三级结构的预测

Fig.7　Prediction of the tertiary structure of Hrp1 protein

2.6　Hrp1 蛋白的 B 细胞表位、辅助性 T 细胞表位、细胞毒性 T 淋巴细胞表位预测

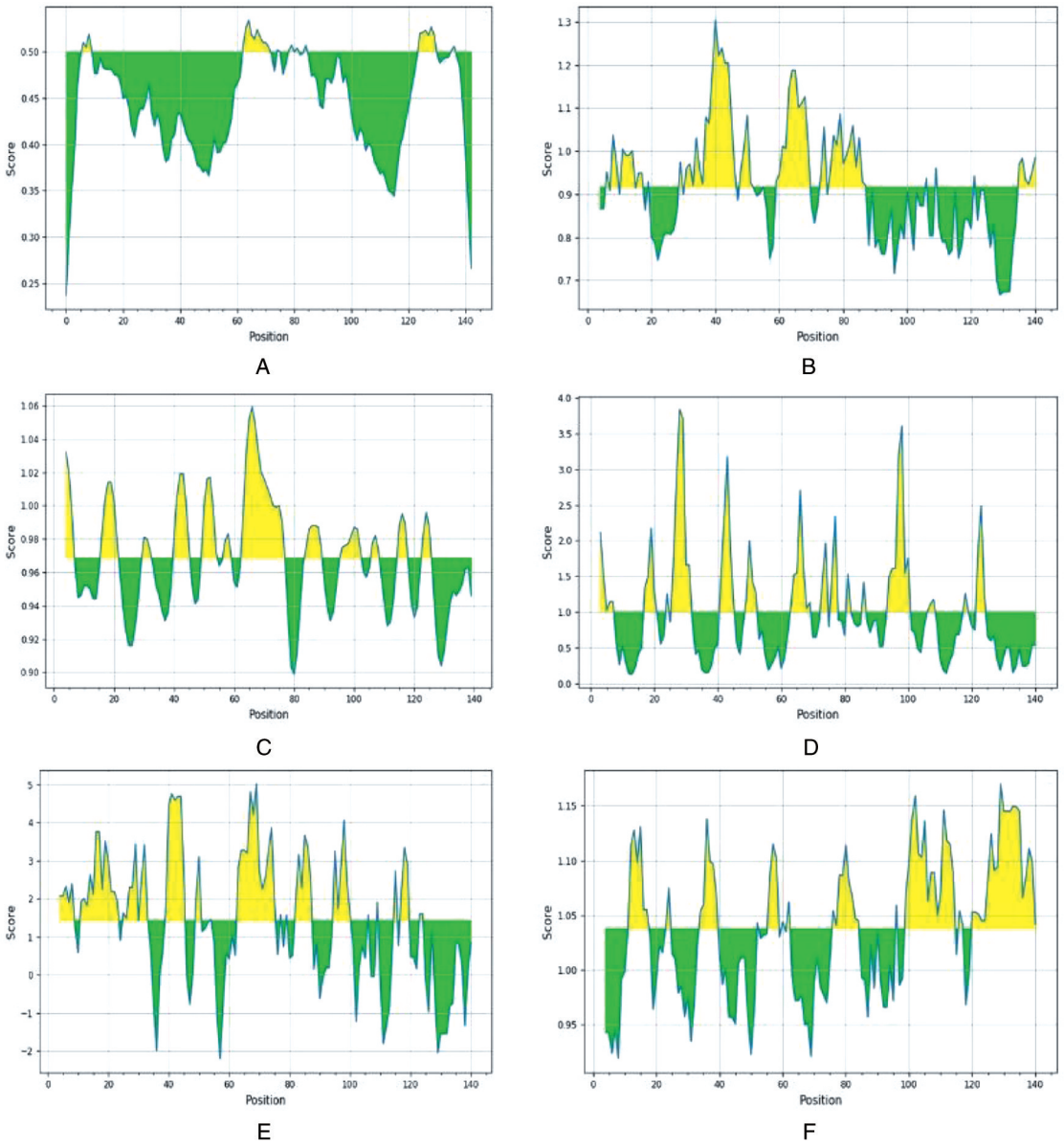
2.6.1　Hrp1 抗原蛋白的 B 细胞表位预测　通过 IEDB 软件对 Hrp1 蛋白的 B 细胞表位进行预测，结果显示 Hrp1 蛋白的可塑性、表面可及性、线性表位、β-转角、亲疏水性、抗原 6 个方面（图 8），柔韧性越强越有利与抗体结合，柔韧性参数以基线 1.0 为参照，高于 1.0 的氨基酸区段柔韧性强，是易形成抗原表位的区域，可及性高于基线 1.0 的区段易形成 B 细胞表位，再通过结合线性表位、β-转角、抗原性这 3 个方面筛选，筛选得出 B 细胞表位（表 1）。

表 1　Hrp1 蛋白 B 细胞抗原表位预测结果

Tab.1　Prediction of B cell epitopes of Hrp1 protein

序号	序列	起始氨基酸位置
1	VTCVGE	12
2	GALPIC	34
3	DSIYYVDAN	76
4	VRRVPVISEHRLVGIV	99

2.6.2　预测 Hrp1 蛋白的辅助性 T 细胞表位结果　使用辅助性 T 细胞表位分析软件预测 Hrp1 蛋白的限制性 Th 细胞表位，对候选意义的每一种 HLA 亚型相应的表位数目情况进行统计，其中 HLA-DRB1 * 0401、HLA-DRB1 * 0701 表型的表位数量最多，结果如表 2 所示；对于某个表型预测分值较高的氨基酸序列进行下一步分析选择，以作为候选 Th 表位，如表 3 为最终结果。



Note: A:Linear Epitope Prediction 2.0 Results; B:Beta-Turn Prediction Result; C:Fexibility Prediction Results; D:Surface Accessibility Prediction Results; E:Hydrophlicity Prediction; F:Antigenicity Results

图 8 IEDB 预测 Hrp1 蛋白的 B 细胞表位

Fig.8 Prediction of B cell epitopes of Hrp1 protein in IEDB

表 2 SYFPEITHI 软件预测的 Hrp1 蛋白 Th 细胞表位信息

Tab.2 Information on Hrp1 protein Th cell epitopes predicted by SYFPEITHI software

MHC-Ⅱ类分子亚型	SYFPEITHI 表位数(得分)
HLA-DRB1 * 0101	0
HLA-DRB1 * 0401	24(18-28)
HLA-DRB1 * 0701	14(18-28)
HLA-DRB1 * 0801	0
HLA-DRB1 * 0901	0
HLA-DRB1 * 1101	7(18-21)
HLA-DRB1 * 1501	0

表 3 Hrp1 蛋白 Th 细胞表位综合分析

Tab.3 Comprehensive analysis of Th cell epitopes of Hrp1 protein

序号	起始氨基酸位置	氨基酸序列	SYFPEITHI 得分
1	4	ARDIMNAGVTCVGEH	22
2	24	AAQYMREHDIGALPI	22
3	45	LHGMLTDRDIVIKGL	22
4	76	DSIYYVDANASIQEM	28
5	88	QEMLNVMEEHQVRRV	22
6	99	VRRVPVISEHRLVGI	28
7	108	HRLVGIVTEADIARH	22
8	129	VQFVKAICSPMALAS	26

2.6.3　Hrp1 蛋白的细胞毒性 T 淋巴细胞预测分析结果　根据预测软件结果，选择出每一种表型分值较高的多肽序列，其可成为候选 CTL 表位，具体结果见表 4。

表 4　Hrp1 蛋白 CTL 表位综合分析

Tab.4　Comprehensive analysis of Hrp1 protein CTL epitopes

序号	起始氨基酸位置	氨基酸序列	RANKPEP 得分
1	4	ARDIMNAGV	12.893
2	7	IMNAGVTCV	15.36
3	19	ETLTAAAQY	9.289
4	49	LTDRDIVI	12.424
5	72	ELARDSIYYV	11.411
6	90	MLNVMEEHQV	12.234
7	122	HLPEHAIV	11.37

结合 Hrp1 蛋白的 B 细胞表位、辅助性 T 细胞表位、细胞毒性 T 淋巴细胞表位预测结果，根据 DNAStar 软件中 Protean 模板得到关于 Hrp1 蛋白表位预测的位置图，该图可直观看到 Hrp1 蛋白表位富集的区域，具体的表位富集表现在红色方块标记的区域（图 9）。

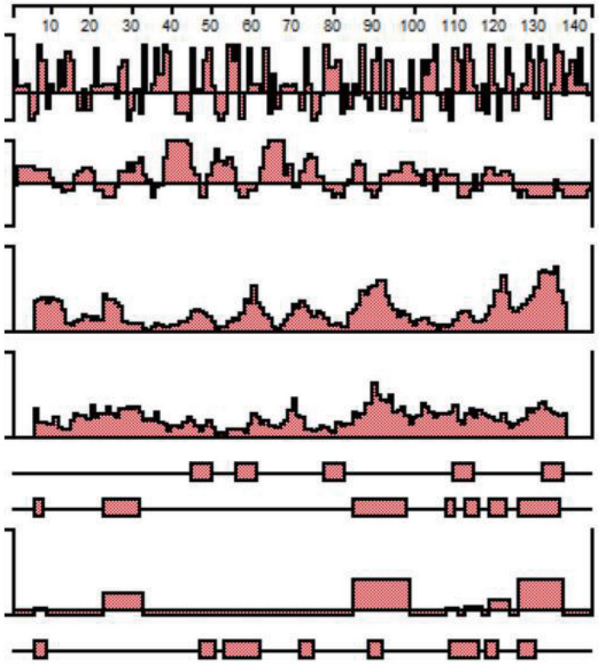


图 9　Hrp1 蛋白表位预测位置图

Fig.9　Predicted Hrp1 protein epitope position map

2.7　Hrp1 蛋白的相互作用网络预测结果　STRING 在线分析软件预测了与 Hrp1 蛋白相互作用的蛋白有 Rv2627c、Rv2628、Rv1738、devR、Rv2624c、TB31.7、rip3、pfkB、guaA、hspX（图 10）。通过分析 Hrp1 的相互作用蛋白，可得 Hrp1 可能在活化的巨噬细胞内产生大量 NO，然后作为细胞间的信号分子杀伤被吞噬的微生物，调节体内细胞一系列重要的功能；该蛋白可能与抑制蛋白质合成有关，刺激巨噬细胞和外周血单个核细胞分泌重要的细胞因子，然后减少细菌生长，该蛋白的相互作用网络主要涉及 NF-kB 信号通路。

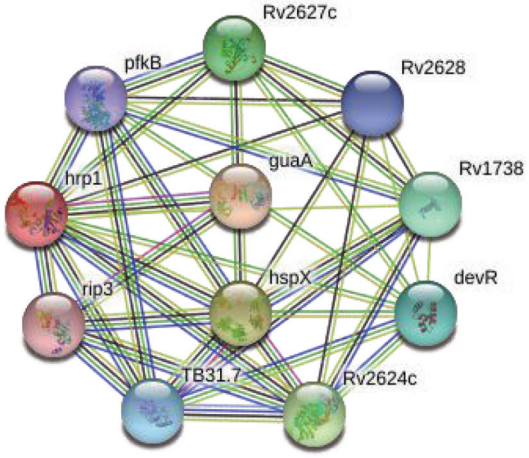


图 10　MTB H37Rv Hrp1 相互作用蛋白

Fig.10　Protein interactions of MTB H37Rv Hrp1 protein

3　讨　论

目前在临床上唯一预防结核病的疫苗是卡介苗（BCG），但卡介苗基因组发生的众多突变导致相当一部分人群丧失保护效力，其对不同人群的保护效率差异较大，机体保护水平范围从 0%～80%不等，且对潜伏感染 MTB 没有保护效果[7]。在临床上约有 90%的结核杆菌感染者，其身体基本上处于健康状态，但是如果机体免疫系统缺乏抵抗力，结核杆菌会重新复制，病灶被打破，随之大量的细菌释放，首先进入肺部，随后至全身，发展为活动性肺结核，故结核病疫情控制工作者一直关注结核潜伏期感染诊断试剂和治疗性疫苗的研究开发。结核杆菌可以在细胞内潜伏数十年，因此需要一个可以诱导具有长期记忆的综合免疫保护应答疫苗，在持续感染人群中，检测到对持续感染期抗原的具有一定的特异性反应，而没有在卡介苗免疫人群检测到，因此猜想以潜伏感染相关蛋白设计相关疫苗，有希望研究出针对持续感染期抗原的保护性免疫反应，这有助于在潜伏感染期间清除 MTB 的复发。对结核分枝杆菌相关潜伏感染蛋白以生信软件为基础进行结构功能

的生物信息学分析,可为结核病潜伏感染致病机制的阐述及新型药物结核疫苗的开发提供一定的研究理论基础,而关于抗原表位的研究也成为最近 MTB 诊断预防治疗的重点方向。在本文研究中,联合运用生物信息学的技术方法对 Hrp1 蛋白的相关信息进行分析、处理、模拟和预测,本文预测抗原表位的模式,可提高该蛋白用于表位疫苗的研发及诊断和实验试剂开发的水平,避免耗费大量的人力和物力,对免疫学的发展非常有益,使免疫学数据更加丰富。

缺氧反应蛋白 1 是由 Rv2626c 基因编码,Bashir 等在研究中发现实验鼠的巨噬细胞表面结合重组 Hrp1 蛋白后可以引起小鼠体内发生Ⅰ型免疫反应,并且小鼠体内的 NO 及 iNOS 分泌量增加,小鼠巨噬细胞 B7-1、B7-2 和 CD40 的表达随之增强,经重组 Hrp1 蛋白刺激后的小鼠巨噬细胞和活动性结核病患者的外周血单个核细胞可分泌高水平的 IL-12 及 TNF-α[8]。Azzurri 等应用 ELISA 方法对结核病患者血清中抗 Hrp1 抗体水平进行检测,结果发现未经治疗的活动性结核病患者高于结核病密切接触者及健康社区对照人群,结核病患者化疗后抗 Hrp1 抗体水平下降[9]。Roupie 等将 Hrp1 DNA 疫苗免疫小鼠,发现小鼠 IFN-γ 和 IL-2 分泌增加,Hrp1 DNA 疫苗可以诱导较强的体液免疫反应及 Th1 型细胞免疫反应的效力优于其他潜伏感染抗原 Rv1733c、Rv1738、Rv2029c、Rv2031c、Rv2023、Rv2627c 及 Rv2628[10]。由此可知 Hrp1 具有较好的免疫原性,其通过固有免疫及获得性免疫两种方式来调节巨噬细胞的效应能力,且该抗原蛋白很容易被结核分枝杆菌潜伏患者所识别,多项研究提示其很有可能会成为一种新型诊断 MTB 潜伏感染的候选抗原和潜伏性感染的候选疫苗。

本研究利用多种工具和生物信息学资源对结核分枝杆菌 Hrp1 蛋白的性质、结构、功能等进行分析。结核分枝杆菌 Hrp1 由 143 个氨基酸组成,该蛋白分子式为 C667 H 1084 N194 O211 S10,理论等电点 PI 为 4.96,分子量为 15 517.72,不稳定性指数为 19.62,为稳定疏水性蛋白。ORF Finder 发现 Rv2626c 基因全长 432 bp,含有 5 个开放阅读框,其中最长的一个开放阅读框显示 Rv2626c 基因基本被通读,全长 432 bp,与预测的编码 Hrp1 蛋白长度基本相似。综合利用 Protpapram、ProtScale、SOSUI 工具预测 Hrp1 蛋白无信号肽,并通过在线生信软件可知该蛋白不属于分泌蛋白,主要定位在结核分枝杆菌的细胞质,在细胞质内发挥功能[11]。其二级结构主要为 α-螺旋和无规则卷曲,共占 71.33%,生物信息学分析 Hrp1 蛋白含有 6 个磷酸化位点,具有良好的抗原优势表位和氨基酸磷酸化位点,推断猜想以该蛋白设计疫苗,有助于在潜伏感染期间清除 MTB 的复发,是探索结核分枝杆菌持久性感染机制的重要研究靶标。该蛋白无规则卷曲占 27.27%,表明其容易与抗体嵌合,并预测 Hrp1 含有多个 T 细胞的抗原优势表位。蛋白抗原的二级结构与 B 细胞表位关系密切[12],对 Hrp1 二级结构、亲水性、表面可能性、柔韧性和抗原性等进行全面分析,最后结果可知 12-18、34-40、76-85、99-114 位等 4 个 B 细胞表位。目前预测辅助性 T 细胞表位效果较好且常用的在线生信软件是 SYFPEITHI,本研究筛选中国人群常见的 MHC-Ⅱ类分子亚型 DRB1 * 0101、0401、0701、0801、1101、1501[13],综合分析得到的结果对于研究中国人群具有一定的实际意义。经过结果分析发现,本研究对 Hrp1 抗原蛋白的 Th 细胞、CTL 和 B 细胞抗原表位进行生物信息学预测分析,并应用生物信息学预测分析 Hrp1 蛋白的结构和功能,发现 Hrp1 蛋白含有较丰富的 B、T 细胞抗原表位,为进一步研究该蛋白的功能和明确其在 MTB 潜伏感染机制中作用提供理论依据,以期为筛选 MTB 诊断标志物和研制新型表位结核疫苗奠定基础[14]。STRING 数据库分析显示 Hrp1 蛋白与 Rv2627c、Rv2628、Rv1738、devR、Rv2624c、TB31.7、rip3、pfkB、guaA 和 hspX 有相互作用,该蛋白可能刺激产生 NO 及细胞因子的过程受转录因子 NF-kB 调控,NO 可通过 MTB 的休眠调节子启动 Rv2626c 基因在内的多个下游基因的表达,使 MTB 进入休眠状态,抵御宿主免疫系统的杀伤。与抑制蛋白质合成有关,还可能会刺激巨噬细胞和外周血单个核细胞分泌重要的细胞因子[15]。

目前对 MTB 缺氧相关潜伏期抗原蛋白 Hrp1 的机制尚不十分清晰明确[16],深入研究 Hrp1 的免疫学特性及结构功能特征,对控制结核潜伏感染、预防结核发病和结核免疫治疗具有一定意义。本研究对 Hrp1 蛋白进行了全面的生物信息学分析,采用多种方式在多个角度进行了分析和预测,有效地规避了单一软件预测的缺陷,为进一步研究 Hrp1 蛋白的结构、功能、理化性质提供参考,为进一步揭示结核分枝杆菌的致病机制奠定理论基础。

利益冲突:无

引用本文格式:刘静,吴芳,张杰,等. 结核分枝杆菌 Hrp1 蛋白的生物信息学分析[J]. 中国人兽共患病学报,2021,37(8):669-677. DOI:10.3969/j.

issn.1002－2694.2021.00.081

参考文献：

[1] 赵飞，杜昕，李涛，等.基于世界卫生组织公共数据库的中国结核病流行趋势与预测[J]. 临床药物治疗杂志，2018，16(4)：1-3. DOI：10.3969/j.issn.1672-3384.2018.04.001

[2] 曹旭东，陈创夫，张辉，等.GFP-Rv2626c 融合蛋白表达载体的构建及表达纯化[J].中国人兽共患病学报，2009，25(4)：361-364. DOI：10.3969/j.issn.1002-2694.2009.04.017

[3] 刘相群，伊正君，付玉荣.结核分枝杆菌 PpiA 蛋白的生物信息学分析及功能预测[J].中国病原生物学杂志，2020，15(6)：661-666. DOI：10.13350/j.cjpb.200608.

[4] 曾江勇. 口蹄疫病毒 Hankou/99 株结构蛋白的二级结构及其 B 淋巴细胞表位的预测[J]. 黑龙江畜牧兽医，2015，7(5)：54-56. DOI：10.13881/j.cnki.hljxmsy.2015.0415.

[5] 崔瑞娜，白雪娟，阳幼荣，等. 结核分枝杆菌 Rv3134c 蛋白抗原表位的预测[J]. 中国人兽共患病学报，2015，12(6)：51-56. DOI：10.3969/j.issn.1002-2694.2015.06.010.

[6] 卞茂红，沈际佳，刘淼，等. 人 Rh(D)血型抗原模拟表位的筛选和鉴定[J]. 中华检验医学杂志，2008，31(3)：305-308. DOI：10.3321/j.issn：1009-9158.2008.03.014

[7] White RG，Houben RM. Towards elimination in industrialised countries：expanding diagnosis and treatment of LTBI among immigrants [J]. Int J Tuberculosis Lung Dis，2014，18(4)：380-380. DOI：10.5588/ijtld.14.0156

[8] Bashir N，Kounsar F，Mukhopadhyay S，et al. *Mycobacterium tuberculosis* conserved hypothetical protein rRv2626c modulates macrophage effector functions[J]. Immunology，2010，130(1)：34-45. DOI：10.1111/j.1365-2567.2009.03196.x

[9] Azzurri A，Kanaujia GV，Sow OY，et al. Serological markers of pulmonary tuberculosis and of response to anti-tuberculosis treatment in a patient population in Guinea.[J]. Int J Immunopathol Pharmacol，2006，19(1)：199-208. DOI：10.1038/s41598-020-64428-z

[10] Roupie V，Viart S，Leroy B，et al. Immunogenicity of eight *Mycobacterium avium* subsp. paratuberculosis specific antigens in DNA vaccinated and map infected mice[J]. Veter Immunol Immunopathol，2012，145(1/2)：74-85. DOI：10.1016/j.vetimm.2011.10.012

[11] 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. DOI：10.3969/j.issn：1002-2694.2018.00.214

[12] 杨赛丽，谢婉莹，刘仪，等.结核分枝杆菌 FabG4 的生物信息学分析及功能初探[J].石河子大学学报(自然科学版)，2020，38(1)：96-103. DOI：10.13880/j.cnki.65-1174/n.2020.22.001

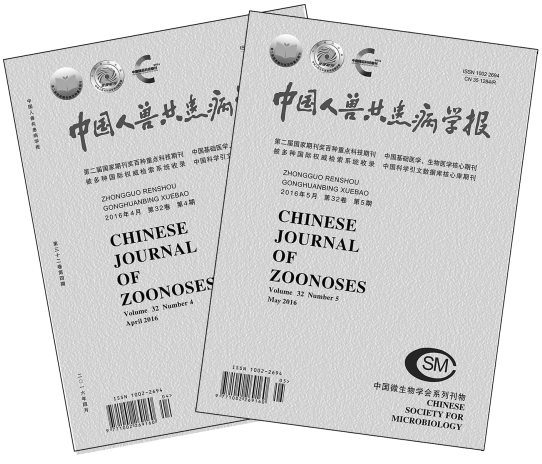
[13] Roupie V，Romano M，Zhang L，et al. Immunogenicity of eight dormancy regulon-encoded proteins of *Mycobacterium tuberculosis* in DNA-vaccinated and tuberculosis-infected mice [J]. Infect Immun，2007，75(2)：941-949. DOI：10.1128/IAI.01137-06

[14] 张薇，柏银兰，康健，等. 结核分枝杆菌持续感染期抗原 Rv2626c 的表达、纯化和鉴定[J]. 现代生物医学进展，2012，12(10)：78-81.DOI：10.13241/j.cnki.pmb.2012.10.006

[15] 段瑞芳，武翠玲，王牛牛，等. 人 ACE2 蛋白结构和功能的生物信息学分析[J]. 中国人兽共患病学报，2020，36(9)：712-719. DOI：10.13621/j.1001-5949.2020.11.0969

[16] 吉萍，范小勇，吴康，等. 结核分枝杆菌潜伏性感染相关抗原的研究进展[J]. 中华微生物学和免疫学杂志，2015，1(13)：59-64. DOI：10.3760/cma.j.issn.0254-5101.2015.01.013.

收稿日期：2020-10-12　编辑：王晓欢



欢迎订阅

《中国人兽共患病学报》

邮发代号：34-46