

DOI:10.13350/j.cjpb.220511

• 论著 •

结核分枝杆菌 Mce3R 蛋白的生物信息学分析*

宋娜娜¹,付玉荣²,伊正君^{1**}

(1. 潍坊医学院医学检验学院,山东潍坊 261053;2. 潍坊医学院基础医学院)

【摘要】 目的 采用生物信息学方法分析结核分枝杆菌的 Rv1963c 基因及其编码的 Mce3R 蛋白结构和功能。 方法 自 NCBI 网站获取 Rv1963c 基因及 Mce3R 蛋白的基本信息;使用 ProtParam 和 ProtScale 预测 Mce3R 蛋白的理化性质和亲疏水性;使用 SignalIP4.1 Server、TMHMM ServerV.2.0 和 NetPhos3.1 软件预测 Mce3R 蛋白的信号肽、跨膜结构及磷酸化位点;使用 SOPMA 预测 Mce3R 蛋白的二级结构,通过 SWISS MODEL 软件构建该蛋白的三级结构模型;分别使用 IEDB 和 SYFPEITHI 预测 Mce3R 蛋白的 B 和 T 细胞表位;通过 STRING 数据库预测 Mce3R 的相互作用蛋白。

结果 Rv1963c 基因全长 1 221 bp,编码的 Mce3R 蛋白氨基酸数为 406,分子式为 $C_{1948} H_{3177} N_{587} O_{581} S_7$;亲水性系数 -0.132,预测其为亲水性蛋白。Mce3R 蛋白含有磷酸化位点和抗原表位。**结论** 生物信息学方法分析 Mce3R 蛋白含有一个保守的结构域,其结构域与结核分枝杆菌耐药性的产生密切相关,为抗结核药物靶点的选择提供了新方向。预测 Mce3R 蛋白含有大量抗原表位,可能成为结核病疫苗的候选蛋白。

【关键词】 结核分枝杆菌;Rv1963;Mce3R 蛋白;生物信息学分析

【中图分类号】 R378.911

【文献标识码】 A

【文章编号】 1673-5234(2022)05-0550-04

[Journal of Pathogen Biology. 2022 May;17(5):550-553, 573.]

Bioinformatics analysis of MCE3R protein of *Mycobacterium tuberculosis*

SONG Na-na¹, FU Yu-rong², YI Zheng-jun¹ (1. School of Medical Laboratory, Weifang Medical University, Weifang 261053, Shandong, China; 2. School of Basic Medicine Sciences, Weifang Medical University)

【Abstract】 **Objective** To analyze the structure and function of *Mycobacterium tuberculosis* Rv1963c gene and its encoding Mce3R protein by bioinformatics method. **Methods** The basic information of Rv1963c gene and Mce3R protein was obtained from NCBI website. The physicochemical properties and hydrophilicity of Mce3R protein were predicted by ProtParam and ProtScale. SignalIP4.1 Server, TMHMM ServerV.2.0 and NetPhos3.1 software were used to predict the signal peptide, transmembrane structure and phosphorylation site of Mce3R protein. SOPMA was used to predict the secondary structure of Mce3R protein, and SWISS MODEL was used to construct the tertiary structure MODEL of Mce3R protein. The B and T cell epitopes of Mce3R protein were predicted by IEDB and SYFPEITHI, respectively. The interacting proteins of Mce3R were predicted using STRING database. **Results** The total length of Rv1963c gene was 1 221 bp, and the number of amino acids encoding Mce3R protein was 406, and the molecular formula was $C_{1948} H_{3177} N_{587} O_{581} S_7$. The hydrophilic coefficient was -0.132, which was predicted to be hydrophilic protein. Mce3R contains phosphorylation sites and epitopes. **Conclusion** The Mce3R protein contains a conserved domain identified by bioinformatics methods, which is closely related to the development of drug resistance in *Mycobacterium tuberculosis*, providing a new direction for the selection of anti-tuberculosis drug targets. It is predicted that Mce3R protein contains a large number of epitopes and may be a candidate protein for tuberculosis vaccine.

【Key words】 *Mycobacterium tuberculosis*; Rv1963; Mce3R protein; bioinformatics

*** 结核病(Tuberculosis, TB)主要是由结核分枝杆菌(*Mycobacterium tuberculosis*, Mtb)引起的传染病。HIV/AIDS 大流行、发展中国家公共卫生系统的恶化以及多重耐药性和极耐药性 TB^[1]的出现进一步助长了 TB 的发展。至少对两种治疗 TB 的一线药物异烟肼(INH)和利福平产生耐药性的耐多药菌株(MDR-TB)的出现使 TB 的防治难度增大。这些感染需要额外的二线药物和更长的治疗时间,治疗成功率也大幅下降^[2,3],迫切需要开发针对 MDR-TB 的新策略。Mce3R 是含有 TetR 结构域一种转录调节因子, TetR

结构域是一种环境传感器,而含有 TetR 结构域的蛋白大多能够参与抗生素耐药性的产生、渗透压力、外排泵、多药耐药性、代谢调节和发病机制等^[4]。有研究报告

* 【基金项目】 山东省自然科学基金面上项目和重大基础研究资助项目(No. R2018MH001, ZR2018ZC1054)。

** 【通讯作者】 伊正君, E-mail: fuyizhengjun@163.com

【作者简介】 宋娜娜(1996-),女,山东青岛人,硕士研究生,主要研究方向为临床检验诊断学。

E-mail: 1445260658@qq.com

道 Mce3R 诱导菌能够增强 INH 杀菌效果,这为抗结核药物靶点的选择提供新方向^[5,6]。Mce3R 参与胆固醇调节的抗压通路,该通路有助于提高药物耐受性^[7]。但 Mtb 的 Rv1963c 基因及其编码的 Mce3R 蛋白的具体结构和功能不清楚。本研究利用生物信息学方法对 Mce3R 蛋白的结构和功能进行分析,为抗结核药物靶点和新型疫苗的研发提供理论依据。

材料与方法

1 目的基因及相关编码信息获取

从 GenBank 数据库中获取 Mce3R (Rv1963c) 核酸序列及其编码 Mce3R 蛋白的氨基酸序列等信息。Mtb H37Rv 标准株基因组序列号为 NC_000962., Rv1963c 的 Gene ID 为 885950, 其基因及编码的 Mce3R 蛋白序列为。

2 方法

2.1 Rv1963c 基因相关信息及开放阅读框分析 从 NCBI 获取 Rv1963c 基因序列信息, 使用 ORF Finger 软件分析其开放阅读框。

2.2 Mce3R 蛋白基本理化性质分析 使用 ExPasy 的 ProtParam 分析 Mce3R 蛋白的物理特征及化学特性,如分子质量、分子式、氨基酸构成、等电点、不稳定指数等。

2.3 Mce3R 蛋白的亲疏水性及信号肽、跨膜区、磷酸化位点分析 使用 ProtScale 预测 Mce3R 蛋白的亲疏水性和可溶性;使用 SignalIP4.1 Server 软件预测分析 Mce3R 蛋白有无信号肽及数量;使用 PSORTB 进行 Mce3R 蛋白的亚细胞定位分析;使用 TMHMM Server v. 2.0 分析蛋白的跨膜区域;使用 NetPhos3.1 分析软件预测 Mce3R 蛋白的磷酸化位点。

2.4 Mce3R 蛋白二级结构分析与三级结构建模 采用 SOPMA 软件预测分析 Mce3R 蛋白的二级结构;通过 SWISS-MODEL 分析软件建立 Mce3R 蛋白序列的三级结构模型。

2.5 蛋白质抗原表位分析 分别使用 IEDB、SYFPEITHI 预测 Mce3R 蛋白的 B、T 细胞抗原表位。

2.6 相互作用蛋白预测 通过 STRING 数据库预测 Mce3R 的相互作用蛋白。

结 果

1 Rv1963c 基因相关信息及序列分析

以序列号 NC_000962.3 从 NCBI 的 GenBank 数据库中获得结核分枝杆菌 H37Rv 标准株的全基因组序列。Rv1963c 基因在全基因组中所处的位置为 2205582-2206802, 基因全长 1 221 bp, 起始子、终止子分别为 ATG、TAG, CG 含量约占碱基组成的

66.34%。Mce3R 蛋白的氨基酸序列号为,由 406 个氨基酸组成。Rv1963c 基因含有 6 个 ORF, 最长 ORF 1 221 bp, 编码 406 个氨基酸(图 1)。

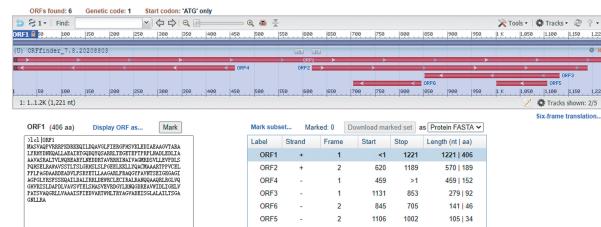


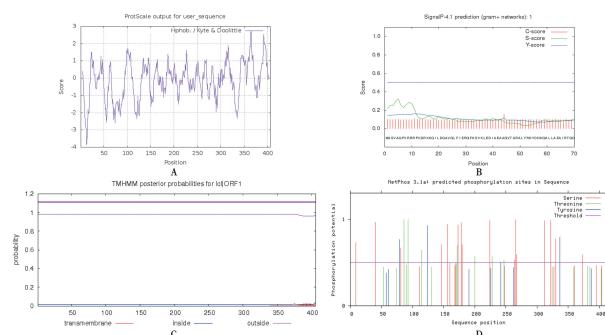
图 1 Rv1963c 基因的开放阅读框分析
Fig. 1 Open reading frame analysis of Rv1963c gene

2 Mce3R 蛋白的理化性质

使用 predictprotein 软件对自 GenBank 中获取的 Mce3R 氨基酸序列进行分析,Mce3R 蛋白由 406 个氨基酸构成,分子质量为 44.341 66 ku, 等电点 8.85。Mce3R 分子式为 C₁₉₄₈H₃₁₇₇N₅₈₇O₅₈₁S₇, 包含 6 300 个原子。预测 Mce3R 蛋白脂肪指数为 103.23, 不稳定指数 46.67, 可能为不稳定蛋白。

3 Mce3R 蛋白的亲疏水性及信号肽、跨膜区、磷酸化位点分析

使用 ProtScale 软件预测 Mce3R 蛋白的亲疏水性(图 2A), 其中亲水性系数最大值为 2.711, 最小值为 -3.867, 平均亲水系数为 -0.132, 为亲水性蛋白。



A Mce3R 蛋白的亲疏水性分析 B Mce3R 蛋白信号肽预测
C TMHMM server 预测 Mce3R 蛋白跨膜螺旋 D NetPhos3.1
Mce3R 蛋白磷酸化位点预测

图 2 Mce3R 蛋白的亲疏水性及信号肽、跨膜区、磷酸化位点分析

A Hydrophilicity analysis of Mce3R protein B Prediction of Mce3R protein signal peptide C prediction of Mce3R protein transmembrane helix by TMHMM Server D Prediction of NetPhos3.1 Mce3R protein phosphorylation site

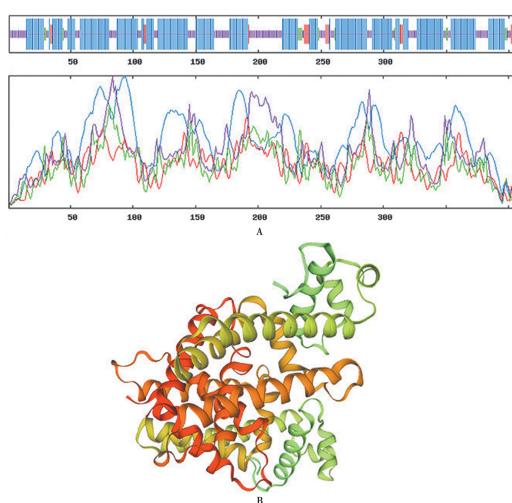
Fig. 2 Analysis of the hydrophilicity, signal peptide, transmembrane region and phosphorylation site of Mce3R protein

利用 PSORTB 进行亚细胞定位分析,结果显示 Mce3R 蛋白定位于胞质。使用 SignalIP4.1 Server 软件预测 Mce3R 蛋白信号肽(图 2B), 标准值 D 值为 0.202, 小于阈值, 表明 Mce3R 蛋白无信号肽端。

TMHMM 分析 Mce3R 蛋白不存在跨膜螺旋(图 2C),且氨基酸残基均未突破 inside 阈值限(蓝色横线),表明该蛋白氨基酸序列全长均属于膜内侧。NetPhos 3.1 软件预测 Mce3R 蛋白包含的磷酸化位点数目为 27 个,其中包括 15 个 Serine,7 个 Threonine,4 个 tyrosine(图 2D)。

4 Mce3R 蛋白的二、三级结构

使用 SOPMA 进预测 Mce3R 蛋白二级结构(图 3A),其中包含 255 个(占 62.81%) α -螺旋(Hh),19 个(占 4.68%)伸展链(Ee),13 个(占 3.20%) β -转角(Tt),119 个(占 29.31%)无规卷曲(Cc)。使用 SWISS MODEL 软件构建的 Mce3R 蛋白三级结构模型见图 3B。



A SOPMA 预测 Mce3R 蛋白二级结构 B Mce3R 蛋白三级结构同源建模

图 3 Mce3R 蛋白的二、三级结构预测

A Prediction of Mce3R protein secondary structure by SOPMA
B Homologous modeling of Mce3R protein tertiary structure.

Fig. 3 Prediction of secondary and tertiary structure of Mce3R protein

5 Mce3R 蛋白相关抗原表位

应用 IEDB 软件预测 Mce3R 蛋白的 B 细胞表位,其中序列 170-221,322-341,117-13 等为优势抗原表位。采用 SYFPEITHI 软件预测分析 T 细胞抗原表位,HLA-A * 02,01 其中>21 分以上的有 30 个。

6 Mce3R 相互作用蛋白

通过 STRING 数据库预测 Mce3R 的相互作用蛋白,结果见图 4。包括 Rv1936, Rv1937, ephB (Rv1938), vapB35 和 mce3A 等可与 Mce3R 蛋白相互作用;包含 Rv1936, Rv1937, ephB(Rv1938)的 MEL2 操纵子被认为参与氧化应激,有利于 Mtb 对抗环境阻力^[8]。MEL2 具有促进活性氧(ROS)的解毒作用。相对于野生型菌株,在 Mtb MEL2 突变体中细胞壁毒力脂质、硫油蜡二霉菌酸(PDIM)的产生增加。vapB35

是一种 II 型毒素-抗毒素(TA)系统的抗毒素成分,这可能有利于细菌持续处于非复制性、耐药性状态,从而逃避宿主免疫系统和抗生素的杀伤。Mce3A 是由 mce3 操纵子编码的一种侵袭素/黏附素样蛋白,在感染 mce3 缺失突变体和野生菌株小鼠的器官中检测到细菌负荷减少,存活时间增加,说明突变体减弱了细菌毒力^[9]。Mce3R 及其相互作用蛋白可富集于解毒的生物学过程。

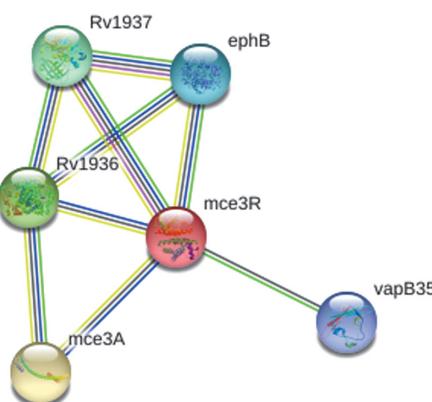


图 4 Mce3R 蛋白的相互作用蛋白预测
Fig. 4 Protein interaction prediction of Mce3r protein

讨 论

在一些发展中国家,通常有一半被诊断患有 TB 的人将在诊断后不久死亡。为了有效治愈 TB,治疗中几乎总是包括同时使用多种抗生素。在为期 2 个月的初始治疗阶段,联合使用两种一线药物,即利福平和异烟肼,通常还会使用额外的抗生素,如吡嗪酰胺和乙胺丁醇。根据疾病进展的严重程度,治疗的延续阶段会额外持续 4 到 7 个月,并包括二线抗生素。尽管进行了长时间的联合药物治疗,但耐药性的出现越来越多,使得患 MDR-TB 的人数增多,因此需要开发具有新靶点的新药以对抗 Mtb 耐药菌株。

生物信息学分析显示 Mce3R 蛋白氨基酸组成中排名前三位的都是脂肪族氨基酸,包括丙氨酸(15%)、亮氨酸(12.8%)、精氨酸(10.3%),说明 Mce3R 蛋白可能与脂质转运或者代谢相关。有研究也证明 Mce3R 可能负调控参与脂质代谢的相关基因且参与胆固醇调节的抗压通路^[7,10]。胆固醇代谢在 Mtb 的持久性、毒力和细胞内存活中起重要作用^[11]。胆固醇在 Mtb 代谢中的关键作用不仅是提供碳和能量,而且还涉及转录组重塑程序,该程序有助于在特定的压力诱导吞噬体信号发生之前对不利的宿主细胞环境产生耐受性。

Mce3R 蛋白含有磷酸化位点表明其可能参与细胞间的信号传导。本研究预测 Mce3R 蛋白含有

62.81%的 α -螺旋,说明该蛋白相对保守。Mce3R蛋白无规卷曲占29.31%,说明其具有重要功能。Mce3R是一种TetR转录抑制因子。TetR结构域包含保守的螺旋-转角-螺旋DNA结合域和C端配体调节域,广泛分布在细菌,主要参与抗生素抗性和有毒的化学化合物耐受性的调节。Mce3R对于一线抗结核药物INH杀伤结核具有增效作用。

预测Mce3R的可能相互作用蛋白为Rv1936,Rv1937,ephB(Rv1938),vapB35和mce3A,并且可能参与解毒的生物学过程。这些分子大多和氧化还原或者脂代谢相关,从而影响Mtb自身存活或是宿主免疫机制(图5)。

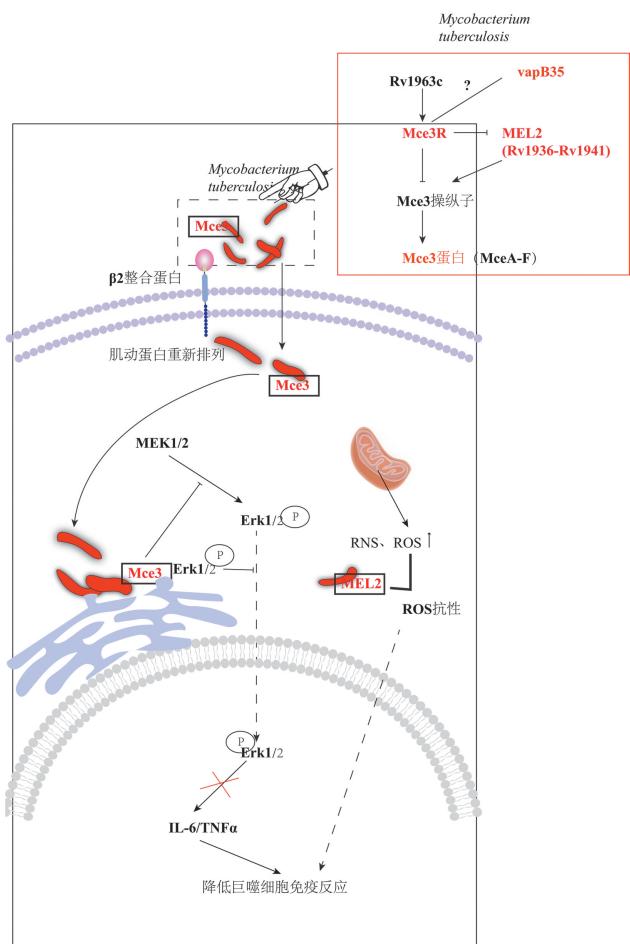


图5 Mce3R及相互作用蛋白调控
Fig. 5 Regulation of Mce3R and interacting proteins

Mtb的mce3操纵子的表达直接受Mce3R和两个转录单Rv1936-Rv1941和Rv1933c-Rv1935c的调控^[12],从而表明3个操纵子的编码产物之间存在功能关系。mce3操纵子基因由mce3R(Rv1963)基因调控。此外,有证据表明mce3操纵子与致病性Mtb的毒力有关,Mce3操纵子基因可能在Mtb存活和适应不利条件(如静止期)开始时起作用^[13,14]。整个Mce3R基因座提供对活性氧和活性氮的抗性,其

mce3操纵子所编码的mce3蛋白在操纵宿主细胞信号传导中具有重要作用。MEL2(Rv1936-Rv1941)被认为参与氧化应激,这与该操纵子编码催化机制,即以产生脂肪酸的醛可以从细胞环境除去氧化物质阻力的观点一致^[15]。vapB35是一种II型毒素-抗毒素(TA)系统的抗毒素成分,但具体作用机制尚不清楚。Mce3R调节子能逆向调节脂质代谢,Mce3R调节子的存在表明胆固醇不仅仅是Mtb的营养物质,同时可能在Mtb的其他途径中发挥作用,例如抗逆性^[16]。这一调控编码了胆固醇调控的抗应激通路,因此认为这一通路在发病机制和药物耐受性方面具有重要意义。Mce3R重要的生物学功能为药物靶点的选择提供了依据。

抗原表位能够与淋巴细胞表面受体结合从而发挥免疫作用,是疫苗开发的重要参考依据。通过生物信息学分析发现Mce3R蛋白含有大量抗原表位,为新型抗结核药物提供了理论依据。

【参考文献】

- [1] 乔甫,黄文治,宗志勇.多重耐药菌感染防控的意义与发展趋势[J].中国护理管理,2019,19(8):1129-1134.
- [2] Puckett S, Trujillo C, Wang Z, et al. Glyoxylate detoxification is an essential function of malate synthase required for carbon assimilation in *Mycobacterium tuberculosis*[J]. Proc Natl Acad Sci USA, 2017,114(11):E2225-E2232.
- [3] 杨雨昕,付玉荣,伊正君.结核分枝杆菌PPE32蛋白的生物信息学分析[J].中国病原生物学杂志,2019,14(2):131-136.
- [4] Ramos JL, Martínez-Bueno M, Molina-Henares AJ, et al. The TetR family of transcriptional repressors[J]. Microbiol Mol Biol Rev, 2005,69(2):326-356.
- [5] Ma S, Morrison R, Hobbs SJ, et al. Transcriptional regulator-induced phenotype screen reveals drug potentiators in *Mycobacterium tuberculosis*[J]. Nat Microbiol, 2020,6(1):44-50.
- [6] Korkegian A, O'Malley T, Xia Y, et al. The 7-phenyl benzoxaborole series is active against *Mycobacterium tuberculosis*[J]. Tuberculosis, 2018(108):96-98.
- [7] Yang X, Yuan T, Ma R, et al. Mce3R Stress-resistance pathway is vulnerable to small-molecule targeting that improves tuberculosis drug activities[J]. ACS Infect Dis, 2019,5(7):1239-1251.
- [8] Janagama HK, Tounkang S, Cirillo SL, et al. Molecular analysis of the *Mycobacterium tuberculosis* lux-like mel2 operon [J]. Tuberculosis, 2013(93):S83-S87.
- [9] Senaratne RH, Sidders B, Sequeira P, et al. *Mycobacterium tuberculosis* strains disrupted in mce3 and mce4 operons are attenuated in mice[J]. J Med Microbiol, 2008,57(2):164-170.
- [10] Santangelo MP, Klepp L, Nunez-Garcia J, et al. Mce3R, a TetR-type transcriptional repressor, controls the expression of a regulon involved in lipid metabolism in *Mycobacterium tuberculosis*[J]. Microbiology, 2009,155(7):2245-2255.
- [11] Rameshwaram NR, Singh P, Ghosh S, et al. Lipid metabolism and intracellular bacterial virulence: key to next-generation therapeutics[J]. Future Microbiol, 2018(13):1301-1328.

(下转573页)

- Acad Sci U S A,2018,115(52):E12353-e12362.
- [11] Patricia JS, Steven M, Karen CC. Understanding the promises and hurdles of metagenomic next-generation sequencing as a diagnostic tool for infectious diseases[J]. Clin Infect Dis,2018, 66(5):778-788.
- [12] Charles YC, Steven AM. Clinical metagenomics. Nat Rev Genet,2019,20(6):341-355.
- [13] Tony L, Placide MK, Samia NN, et al. Metagenomic next-generation sequencing of the 2014 ebola virus disease outbreak in the democratic Republic of the Congo[J]. J Clin Microbiol, 2019,57(9):e00827-19.
- [14] Silvia IS, Sneha S, Samia NN, et al. Coinfections of zika and chikungunya viruses in bahia, brazil, identified by metagenomic next-generation sequencing[J]. J Clin Microbiol,2016,54(9): 2348-53.
- [15] Julien T, Tony L, Louis DP, et al. Genomic epidemiology reconstructs the introduction and spread of zika virus in central America and Mexico[J]. Cell Host Microbe,2018,23(6):855-864.
- [16] Han DS, Li ZY, Li R, et al. mNGS in clinical microbiology laboratories:on the road to maturity[J]. Crit Rev Microbiol, 2019,45(5-6):668-685.
- [17] Miao Q, Ma YY, Wang QQ, et al. Microbiological diagnostic performance of metagenomic next-generation sequencing when applied to clinical practice[J]. Clin Infect Dis,2018,67(suppl_2):S231-s240.
- [18] Stephanie LM ,Patricia JS. Next-generation sequencing in clinical microbiology:Are we there yet? [J] Clin Lab Med,2019, 39 (3):405-418.
- [19] Steve M, Samia NN, Erik S, et al. Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid[J]. Genome Res,2019,29(5):831-842.
- [20] Meredith LC, Susanna KT, Thomas W, et al. Metagenomic Next-Generation Sequencing for Identification and Quantitation of Transplant-Related DNA Viruses[J]. J Clin Microbiol, 2019,57(12):e01113-19.
- [21] Liangjun C,Weiyong L, Qi Z, et al. RNA based mNGS approach identifies a novel human coronavirus from two individual pneumonia cases in 2019 Wuhan outbreak[J]. Emerg Microbes Infect,2020,9(1):313-319.
- [22] Xing XW,Zhang JT,Ma YB,et al. Evaluation of next-generation sequencing for the diagnosis of infections of the central nervous system caused by the neurotropic herpes viruses:A pilot study [J]. Eur Neurol,2018,80(5-6):283-288.
- [23] S Dusko E , MetaHIT C. Metagenomics of the intestinal microbiota:potential applications[J]. Gastroenterol Clin Biol, 2010,34(Suppl 1):S23-8.
- [24] Greninger AL. The challenge of diagnostic metagenomics[J]. Expert Rev Mol Diagn,2018,18(7):605-615.
- [25] Jie H, Erlie J, Yang DL, et al. Metagenomic next-generation sequencing versus traditional pathogen detection in the diagnosis of peripheral pulmonary infectious lesions[J]. Infect Drug Resist,2020(13):567-576.
- [26] 朱美利,张剑青,赵芝焕.宏基因组测序在感染性疾病诊治中的应用进展[J].实用医学杂志,2020,2(36):131-135.
- [27] Gu W, Deng XD, Lee M, et al. Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids [J]. Nat Med,2021,27(1):115-124.
- [28] 陈洋,仇秦威,尚潇潇,等.肠道菌群研究存在的问题及解决方案 [J].中华炎性肠病杂志,2019,3(3):182-188.
- [29] Lankelma JM, van Vught LA, Belzer C, et al. Critically ill patients demonstrate large interpersonal variation in intestinal microbiota dysregulation: a pilot study [J]. Intensive Care Med,2017,43(1):59-68.

【收稿日期】 2022-02-24 【修回日期】 2022-05-13

(上接 553 页)

- [12] Hemati Z,Derakhshandeh A,Haghkhah M, et al. Mammalian cell entry operons;novel and major subset candidates for diagnostics with special reference to *Mycobacterium avium* subspecies paratuberculosis infection[J]. Vet Q,2019,39(1):65-75.
- [13] Soto-Ramirez MD, Aguilar-Ayala DA, Garcia-Morales L, et al. Cholesterol plays a larger role during *Mycobacterium tuberculosis* *in vitro* dormancy and reactivation than previously suspected[J]. Tuberculosis,2017(103):1-9.
- [14] Fenn K,Wong CT,Darbari VC. *Mycobacterium tuberculosis* uses

- Mce proteins to interfere with host cell signaling[J]. Front Mol Bioscs,2019(6):149.
- [15] Moule MG,Cirillo JD. *Mycobacterium tuberculosis* dissemination plays a critical role in pathogenesis [J]. Front Cell Infect Microbiol,2020(10):65.
- [16] Mavi PS, Singh S, Kumar A. Reductive stress: New insights in physiology and drug tolerance of *Mycobacterium* [J]. Antioxid Redox Signal,2020,32(18):1348-1366.

【收稿日期】 2021-11-26 【修回日期】 2022-02-22