

(19) 中华人民共和国国家知识产权局



(12) 发明专利申请

(10) 申请公布号 CN 105622735 A

(43) 申请公布日 2016. 06. 01

(21) 申请号 201410594400. X

G01N 33/569(2006. 01)

(22) 申请日 2014. 10. 29

C12R 1/32(2006. 01)

(71) 申请人 中国科学院生物物理研究所

地址 100101 北京市朝阳区大屯路 15 号

申请人 上海交通大学

中国科学院武汉病毒研究所

广东体必康生物科技有限公司

(72) 发明人 张先恩 陶生策 朱国峰 毕利军

邓教宇 张治平 江何伟 周盈

谷晶 郭书娟 张鸿泰 王绪德

邓瑞萍 王宗秀 刘钟慧 顾佳

李阳 李文娟

(74) 专利代理机构 北京纪凯知识产权代理有限

公司 11245

代理人 关畅 陈晓庆

(51) Int. Cl.

C07K 14/35(2006. 01)

C12N 15/31(2006. 01)

G01N 33/68(2006. 01)

权利要求书3页 说明书14页

序列表11页 附图1页

(54) 发明名称

一组结核分枝杆菌蛋白及其编码基因与应用

(57) 摘要

本发明公开了一组结核分枝杆菌蛋白及其编码基因与应用。本发明公开一组蛋白,由如下(1)-(7)所示的蛋白组成:(1)SEQ ID No. 1所示的蛋白;(2)SEQ ID No. 3所示的蛋白;(3)SEQ ID No. 5所示的蛋白;(4)SEQ ID No. 7所示的蛋白;(5)SEQ ID No. 9所示的蛋白;(6)SEQ ID No. 11所示的蛋白;(7)SEQ ID No. 13所示的蛋白。本发明公开的蛋白质芯片对有效诊断活动性肺结核、有效监控肺结核病人的康复过程具有重要意义。

1. 一组蛋白,由如下(1)-(7)所示的蛋白组成:

- (1)SEQ ID No. 1 所示的蛋白;
- (2)SEQ ID No. 3 所示的蛋白;
- (3)SEQ ID No. 5 所示的蛋白;
- (4)SEQ ID No. 7 所示的蛋白;
- (5)SEQ ID No. 9 所示的蛋白;
- (6)SEQ ID No. 11 所示的蛋白;
- (7)SEQ ID No. 13 所示的蛋白。

2. 权利要求 1 所述蛋白的编码基因。

3. 根据权利要求 2 所述的编码基因,其特征在于:所述编码基因为如下(1)-(7)所示的 DNA 分子:

(1)DNA 分子 1:

- 1)SEQ ID No. 2 所示的 DNA 分子;
- 2) 在严格条件下与 1) 限定的 DNA 分子杂交且编码权利要求 1 中 SEQ ID No. 1 所示的蛋白的 DNA 分子;
- 3) 与 1) 或 2) 限定的 DNA 分子具有 90% 以上的同一性且编码权利要求 1 中 SEQ ID No. 1 所示的蛋白的 DNA 分子;

(2)DNA 分子 2:

- 1)SEQ ID No. 4 所示的 DNA 分子;
- 2) 在严格条件下与 1) 限定的 DNA 分子杂交且编码权利要求 1 中 SEQ ID No. 3 所示的蛋白的 DNA 分子;
- 3) 与 1) 或 2) 限定的 DNA 分子具有 90% 以上的同一性且编码权利要求 1 中 SEQ ID No. 3 所示的蛋白的 DNA 分子;

(3)DNA 分子 3:

- 1)SEQ ID No. 6 所示的 DNA 分子;
- 2) 在严格条件下与 1) 限定的 DNA 分子杂交且编码权利要求 1 中 SEQ ID No. 5 所示的蛋白的 DNA 分子;
- 3) 与 1) 或 2) 限定的 DNA 分子具有 90% 以上的同一性且编码权利要求 1 中 SEQ ID No. 5 所示的蛋白的 DNA 分子;

(4)DNA 分子 4:

- 1)SEQ ID No. 8 所示的 DNA 分子;
- 2) 在严格条件下与 1) 限定的 DNA 分子杂交且编码权利要求 1 中 SEQ ID No. 7 所示的蛋白的 DNA 分子;
- 3) 与 1) 或 2) 限定的 DNA 分子具有 90% 以上的同一性且编码权利要求 1 中 SEQ ID No. 7 所示的蛋白的 DNA 分子;

(5)DNA 分子 5:

- 1)SEQ ID No. 10 所示的 DNA 分子;
- 2) 在严格条件下与 1) 限定的 DNA 分子杂交且编码权利要求 1 中 SEQ ID No. 9 所示的蛋白的 DNA 分子;

3) 与 1) 或 2) 限定的 DNA 分子具有 90% 以上的同一性且编码权利要求 1 中 SEQ ID No. 9 所示的蛋白的 DNA 分子；

(6) DNA 分子 6：

1) SEQ ID No. 12 所示的 DNA 分子；

2) 在严格条件下与 1) 限定的 DNA 分子杂交且编码权利要求 1 中 SEQ ID No. 11 所示的蛋白的 DNA 分子；

3) 与 1) 或 2) 限定的 DNA 分子具有 90% 以上的同一性且编码权利要求 1 中 SEQ ID No. 11 所示的蛋白的 DNA 分子；

(7) DNA 分子 7：

1) SEQ ID No. 14 所示的 DNA 分子；

2) 在严格条件下与 1) 限定的 DNA 分子杂交且编码权利要求 1 中 SEQ ID No. 13 所示的蛋白的 DNA 分子；

3) 与 1) 或 2) 限定的 DNA 分子具有 90% 以上的同一性且编码权利要求 1 中 SEQ ID No. 13 所示的蛋白的 DNA 分子。

4. 含有权利要求 2 或 3 所述编码基因的重组载体、表达盒、转基因细胞系或重组菌。

5. 一种蛋白质芯片，包括基片和与其连接的权利要求 1 所述的蛋白质；且每种蛋白单独成立一个检测点。

6. 根据权利要求 5 所述的蛋白质芯片，其特征在于：所述基片为载体玻片。

7. 含有权利要求 1 所述蛋白的试剂盒。

8. 含有权利要求 5 或 6 所述蛋白质芯片的试剂盒。

9. 根据权利要求 8 所述的试剂盒，其特征在于：所述试剂盒还包含使用说明书，说明书中记载内容如下：

1) 将待检血清点样于权利要求 5 或 6 所述的蛋白质芯片进行检测，并用 Cy5 标记的抗人二抗进行显色，将蛋白质芯片在 Cy5 荧光的发射波长下进行检测，得到各检测点的信噪比，记作 SNR；

所述 SNR 的计算公式为
$$\text{SNR} = \frac{S - B}{\sigma}$$

所述公式中，S 和 B 分别为蛋白质芯片中检测点的原始的信号值和背景值， σ 代表检测点的原始信号的标准差；

所述原始的信号值为芯片扫描仪直接显示的数值；

所述背景值为蛋白质芯片上非样品点制区域的信号值；

所述标准差所对应的实验次数具体为 3 次；

2) 如果待检血清点样于所述蛋白质芯片上的权利要求 1 所述的 SEQ ID No. 1 所示的蛋白的所有检测点的 $\text{SNR} \geq 8.32$ 且 ≤ 10.21 ，则待检血清候选判定为权利要求 1 所述的 SEQ ID No. 1 所示的蛋白的抗体阳性，如果 $\text{SNR} \geq 6.94$ 且 ≤ 8.27 ，则待检血清候选判定为权利要求 1 所述的 SEQ ID No. 1 所示的蛋白的抗体阴性；

如果待检血清点样于所述蛋白质芯片上的权利要求 1 所述的 SEQ ID No. 3 所示的蛋白的所有检测点的 $\text{SNR} \geq 7.68$ 且 ≤ 9.95 ，则待检血清候选判定为权利要求 1 所述的 SEQ ID No. 3 所示的蛋白的抗体阳性，如果 $\text{SNR} \geq 6.31$ 且 ≤ 7.55 ，则待检血清候选判定为权利要求

1 所述的 SEQ ID No. 3 所示的蛋白的抗体阴性；

如果待检血清点样于所述蛋白质芯片上的权利要求 1 所述的 SEQ ID No. 5 所示的蛋白的所有检测点的 $SNR \geq 7.16$ 且 ≤ 9.75 , 则待检血清候选判定为权利要求 1 所述的 SEQ ID No. 5 所示的蛋白的抗体阳性, 如果 $SNR \geq 6.50$ 且 ≤ 7.03 , 则待检血清候选判定为权利要求 1 所述的 SEQ ID No. 5 所示的蛋白的抗体阴性；

如果待检血清点样于所述蛋白质芯片上的权利要求 1 所述的 SEQ ID No. 7 所示的蛋白的所有检测点的 $SNR \geq 7.44$ 且 ≤ 11.16 , 则待检血清候选判定为权利要求 1 所述的 SEQ ID No. 7 所示的蛋白的抗体阳性, 如果 $SNR \geq 6.04$ 且 ≤ 7.20 , 则待检血清候选判定为权利要求 1 所述的 SEQ ID No. 7 所示的蛋白的抗体阴性；

如果待检血清点样于所述蛋白质芯片上的权利要求 1 所述的 SEQ ID No. 9 所示的蛋白的所有检测点的 $SNR \geq 7.98$ 且 ≤ 9.83 , 则待检血清候选判定为权利要求 1 所述的 SEQ ID No. 9 所示的蛋白的抗体阳性, 如果 $SNR \geq 6.84$ 且 ≤ 7.70 , 则待检血清候选判定为权利要求 1 所述的 SEQ ID No. 9 所示的蛋白的抗体阴性；

如果待检血清点样于所述蛋白质芯片上的权利要求 1 所述的 SEQ ID No. 11 所示的蛋白的所有检测点的 $SNR \geq 7.81$ 且 ≤ 10.62 , 则待检血清候选判定为权利要求 1 所述的 SEQ ID No. 11 所示的蛋白的抗体阳性, 如果 $SNR \geq 6.29$ 且 ≤ 7.63 , 则待检血清候选判定为权利要求 1 所述的 SEQ ID No. 11 所示的蛋白的抗体阴性；

如果待检血清点样于所述蛋白质芯片上的权利要求 1 所述的 SEQ ID No. 13 所示的蛋白的所有检测点的 $SNR \geq 8.05$ 且 ≤ 10.87 , 则待检血清候选判定为权利要求 1 所述的 SEQ ID No. 13 所示的蛋白的抗体阳性, 如果 $SNR \geq 6.45$ 且 ≤ 7.84 , 则待检血清候选判定为权利要求 1 所述的 SEQ ID No. 13 所示的蛋白的抗体阴性；

3) 根据步骤 2), 在待检血清的检测结果中, 如果该待检血清对权利要求 1 所述的 SEQ ID No. 1 所示的蛋白、权利要求 1 所述的 SEQ ID No. 3 所示的蛋白、权利要求 1 所述的 SEQ ID No. 5 所示的蛋白、权利要求 1 所述的 SEQ ID No. 7 所示的蛋白、权利要求 1 所述的 SEQ ID No. 9 所示的蛋白、权利要求 1 所述的 SEQ ID No. 11 所示的蛋白、权利要求 1 所述的 SEQ ID No. 13 所示的蛋白中的至少 4 种蛋白的抗体为阳性, 则候选判定该待检血清是活动性肺结核阳性血清, 该血清来源的患者为活动性肺结核患者, 否则该待检血清为活动性肺结核阴性血清, 该血清来源的患者为非活动性肺结核患者；

所述蛋白质芯片制备时所用的各点样液中的权利要求 1 所述的蛋白的浓度均为 $50 \mu\text{g/ml}$ ；

所述待检血清的稀释倍数为 100 倍。

10. 权利要求 1 所述的蛋白、权利要求 5 或 6 所述的蛋白质芯片或权利要求 7-9 任一所述的试剂盒在制备诊断或辅助诊断活动性肺结核的产品中的应用。

一组结核分枝杆菌蛋白及其编码基因与应用

技术领域

[0001] 本发明涉及一组蛋白及其编码基因与应用 ;特别涉及一组结核分枝杆菌蛋白及其编码基因与应用,属于生物医药领域。

背景技术

[0002] 多个世纪以来,结核病持续在全球成为一个不容忽视的公共卫生问题。目前全球已有三分之一的人口携带结核分枝杆菌,仅 2010 年一年就新增结核病病例 880 万,死亡 145 万,平均不到 22 秒即有一人死于肺结核病,结核病高居传染病死亡人数之首。而我国是全球 22 个结核病高负担国家之一,结核病患者人数高居全球第二位,受感染人数超过 5 亿,仅 2010 年一年就有新发病例 90 ~ 120 万,占据了全球总新增病例的约 12%,如不及时采取有效措施,在未来十年内可能有 3000 万人发病,将会导致严重的公共卫生问题和社会问题,所以从国家战略层面必需尽快实现对肺结核的有效控制。

[0003] 科学表明由于活动性肺结核的极易传播性,对活动性肺结核的诊断及尽快消除传染源,结核病的治疗效果评估及控制结核病流行有非常重大的意义。然而肺结核的根治目前还不能完全做到,据研究表明活动性结核病患者经不同的短程化疗治愈后,有 1% -9% 的患者将会复发,而在一些特殊人群中,高达 20% (Hong Kong Chest Service/British Medical Research Council. Am Rev Respir Dis, 1991, 143:700-706),而且耐药性结核病进一步传播与复发未及时诊断的结核患者有很大的关系。所以关于结核治愈患者的活动性肺结核诊断对于防治结核的整体战略具有很大的意义。另外,从临床治疗角度看,结核病的常规治疗方案一般需要 6 个月的时间,而常规结核病的治疗药物有一定的毒副作用,病人的依从性并不好。由于个体间的差异,理论上不同病人实际需要的用药时间可能会有很大的不同,但对于如何细分病人,目前仍缺乏客观的判断指标。通过系统地比较活动性结核病人和结核康复人群的血清,有可能发现一批可以准确区分这两类人群的生物标识物,这些标识物将可用于康复过程的监控,从而有可能指导活动性结核病人的个体化用药,缩短用药周期,提高病人对治疗的依从性,降低治疗费用,提高最终的治疗效果。

[0004] 目前活动性判断应综合临床、X 线表现和痰菌决定,而主要依据是痰菌和 X 线。痰涂片检查简便易行,准确性较高,是活动性肺结核确诊的金标准,但某些研究中其检出率甚至低于 10%。基于核酸扩增的诊断方法,如实时定量 PCR 以及 DNA 芯片的优点是灵敏度高且耗时短,其问题在于特异性较难保证,容易导致假阳性结果。而基于抗原抗体反应的血清学诊断,由于其简便性、快速性,日益成为重要的活动性肺结核临床辅助性诊断手段,并且能满足区分结核康复人群与活动性肺结核的要求。因此,从长远角度综合来看,应致力于发现敏感性和特异性均较好的结核标志物。

[0005] 理想的结核诊断标志物应符合以下条件:(1) 敏感性高;(2) 特异性高;(3) 存在于体液,特别是血液中,易于检测。但是目前肺结核分枝杆菌血清学诊断所针对的抗原如抗原 5、38KD 抗原、30/31KD 抗原以及抗原 60 等,均存在阳性检出率低,与其它分支杆菌的交叉免疫性等问题,虽然在疗效监测、提示复发、判断预后和高危人群的普查中具有一定的价

值,但目前仍不能用于活动性肺结核的确诊及对病人康复过程的监控。为了实现对活动性肺结核的高灵敏性,高特异性的诊断,及对康复过程的有效监控,急需在分子水平上寻找到更敏感、更特异的活动性肺结核生物标志物。

发明内容

[0006] 本发明的目的是提供一组结核分枝杆菌蛋白及其编码基因与应用。

[0007] 本发明提供一组蛋白,由如下(1)-(7)所示的蛋白组成:

[0008] (1)SEQ ID No. 1 所示的蛋白;

[0009] (2)SEQ ID No. 3 所示的蛋白;

[0010] (3)SEQ ID No. 5 所示的蛋白;

[0011] (4)SEQ ID No. 7 所示的蛋白;

[0012] (5)SEQ ID No. 9 所示的蛋白;

[0013] (6)SEQ ID No. 11 所示的蛋白;

[0014] (7)SEQ ID No. 13 所示的蛋白。

[0015] 上述蛋白是用于诊断或辅助诊断活动性肺结核的一组蛋白。

[0016] 上述蛋白的编码基因也属于本发明的保护范围。

[0017] 上述编码基因中,所述编码基因为如下(1)-(7)所示的DNA分子:

[0018] (1)DNA分子1:

[0019] 1)SEQ ID No. 2 所示的DNA分子;

[0020] 2) 在严格条件下与1)限定的DNA分子杂交且编码上述蛋白中SEQ ID No. 1 所示的蛋白的DNA分子;

[0021] 3) 与1)或2)限定的DNA分子具有90%以上的同一性且编码上述蛋白中SEQ ID No. 1 所示的蛋白的DNA分子;

[0022] (2)DNA分子2:

[0023] 1)SEQ ID No. 4 所示的DNA分子;

[0024] 2) 在严格条件下与1)限定的DNA分子杂交且编码上述蛋白中SEQ ID No. 3 所示的蛋白的DNA分子;

[0025] 3) 与1)或2)限定的DNA分子具有90%以上的同一性且编码上述蛋白中SEQ ID No. 3 所示的蛋白的DNA分子;

[0026] (3)DNA分子3:

[0027] 1)SEQ ID No. 6 所示的DNA分子;

[0028] 2) 在严格条件下与1)限定的DNA分子杂交且编码上述蛋白中SEQ ID No. 5 所示的蛋白的DNA分子;

[0029] 3) 与1)或2)限定的DNA分子具有90%以上的同一性且编码上述蛋白中SEQ ID No. 5 所示的蛋白的DNA分子;

[0030] (4)DNA分子4:

[0031] 1)SEQ ID No. 8 所示的DNA分子;

[0032] 2) 在严格条件下与1)限定的DNA分子杂交且编码上述蛋白中SEQ ID No. 7 所示的蛋白的DNA分子;

[0033] 3) 与 1) 或 2) 限定的 DNA 分子具有 90% 以上的同一性且编码上述蛋白中 SEQ ID No. 7 所示的蛋白的 DNA 分子；

[0034] (5) DNA 分子 5；

[0035] 1) SEQ ID No. 10 所示的 DNA 分子；

[0036] 2) 在严格条件下与 1) 限定的 DNA 分子杂交且编码上述蛋白中 SEQ ID No. 9 所示的蛋白的 DNA 分子；

[0037] 3) 与 1) 或 2) 限定的 DNA 分子具有 90% 以上的同一性且编码上述蛋白中 SEQ ID No. 9 所示的蛋白的 DNA 分子；

[0038] (6) DNA 分子 6；

[0039] 1) SEQ ID No. 12 所示的 DNA 分子；

[0040] 2) 在严格条件下与 1) 限定的 DNA 分子杂交且编码上述蛋白中 SEQ ID No. 11 所示的蛋白的 DNA 分子；

[0041] 3) 与 1) 或 2) 限定的 DNA 分子具有 90% 以上的同一性且编码上述蛋白中 SEQ ID No. 11 所示的蛋白的 DNA 分子；

[0042] (7) DNA 分子 7；

[0043] 1) SEQ ID No. 14 所示的 DNA 分子；

[0044] 2) 在严格条件下与 1) 限定的 DNA 分子杂交且编码上述蛋白中 SEQ ID No. 13 所示的蛋白的 DNA 分子；

[0045] 3) 与 1) 或 2) 限定的 DNA 分子具有 90% 以上的同一性且编码上述蛋白中 SEQ ID No. 13 所示的蛋白的 DNA 分子；

[0046] 含有上述任一所述编码基因的重组载体、表达盒、转基因细胞系或重组菌也属于本发明的保护范围。

[0047] 一种蛋白质芯片也属于本发明的保护范围，包括基片和与其连接的上述的蛋白质；且每种蛋白单独成立一个检测点；

[0048] 所述蛋白质芯片上，每种蛋白点样形成的检测点的数量 ≥ 2 。

[0049] 上述蛋白质芯片中，所述基片为载体玻片；

[0050] 所述蛋白质芯片是分别将所述的蛋白质点样于所述载体玻片上制成的；

[0051] 所述蛋白质芯片还包括 IgG 标准品或 Cy5 标记羊抗人抗体二抗形成的检测点作为阳性对照；

[0052] 所述 IgG 标准品具体购自 YEASEN，产品目录号为 36110ES10；

[0053] 所述 Cy5 标记羊抗人抗体二抗具体购自 lifeholder，产品目录号为 LH-G11345。

[0054] 含有上述蛋白的试剂盒也属于本发明的保护范围。

[0055] 含有上述任一所述的蛋白质芯片的试剂盒也属于本发明的保护范围。

[0056] 上述试剂盒中，所述试剂盒还包含使用说明书，说明书中记载内容如下：

[0057] 1) 将待检血清点样于上述任一所述的蛋白质芯片进行检测，并用 Cy5 标记的抗人二抗进行显色，将蛋白质芯片在 Cy5 荧光的发射波长下进行检测，得到各检测点的信噪比，记作 SNR；

[0058] 所述 SNR 的计算公式为
$$\text{SNR} = \frac{S - B}{\sigma}$$

[0059] 所述公式中,S和B分别为蛋白质芯片中检测点的原始的信号值和背景值, σ 代表检测点的原始信号的标准差;

[0060] 所述原始的信号值为芯片扫描仪直接显示的数值;

[0061] 所述背景值为蛋白质芯片上非样品点制区域的信号值;

[0062] 所述标准差所对应的实验次数具体为3次;

[0063] 2) 如果待检血清点样于所述蛋白质芯片上的上述蛋白中所述的SEQ ID No. 1所示的蛋白的所有检测点的 $SNR \geq 8.32$ 且 ≤ 10.21 ,则待检血清候选判定为上述蛋白中所述的SEQ ID No. 1所示的蛋白的抗体阳性,如果 $SNR \geq 6.94$ 且 ≤ 8.27 ,则待检血清候选判定为上述蛋白中所述的SEQ ID No. 1所示的蛋白的抗体阴性;

[0064] 如果待检血清点样于所述蛋白质芯片上的上述蛋白中所述的SEQ ID No. 3所示的蛋白的所有检测点的 $SNR \geq 7.68$ 且 ≤ 9.95 ,则待检血清候选判定为上述蛋白中所述的SEQ ID No. 3所示的蛋白的抗体阳性,如果 $SNR \geq 6.31$ 且 ≤ 7.55 ,则待检血清候选判定为上述蛋白中所述的SEQ ID No. 3所示的蛋白的抗体阴性;

[0065] 如果待检血清点样于所述蛋白质芯片上的上述蛋白中所述的SEQ ID No. 5所示的蛋白的所有检测点的 $SNR \geq 7.16$ 且 ≤ 9.75 ,则待检血清候选判定为上述蛋白中所述的SEQ ID No. 5所示的蛋白的抗体阳性,如果 $SNR \geq 6.50$ 且 ≤ 7.03 ,则待检血清候选判定为上述蛋白中所述的SEQ ID No. 5所示的蛋白的抗体阴性;

[0066] 如果待检血清点样于所述蛋白质芯片上的上述蛋白中所述的SEQ ID No. 7所示的蛋白的所有检测点的 $SNR \geq 7.44$ 且 ≤ 11.16 ,则待检血清候选判定为上述蛋白中所述的SEQ ID No. 7所示的蛋白的抗体阳性,如果 $SNR \geq 6.04$ 且 ≤ 7.20 ,则待检血清候选判定为上述蛋白中所述的SEQ ID No. 7所示的蛋白的抗体阴性;

[0067] 如果待检血清点样于所述蛋白质芯片上的上述蛋白中所述的SEQ ID No. 9所示的蛋白的所有检测点的 $SNR \geq 7.98$ 且 ≤ 9.83 ,则待检血清候选判定为上述蛋白中所述的SEQ ID No. 9所示的蛋白的抗体阳性,如果 $SNR \geq 6.84$ 且 ≤ 7.70 ,则待检血清候选判定为上述蛋白中所述的SEQ ID No. 9所示的蛋白的抗体阴性;

[0068] 如果待检血清点样于所述蛋白质芯片上的上述蛋白中所述的SEQ ID No. 11所示的蛋白的所有检测点的 $SNR \geq 7.81$ 且 ≤ 10.62 ,则待检血清候选判定为上述蛋白中所述的SEQ ID No. 11所示的蛋白的抗体阳性,如果 $SNR \geq 6.29$ 且 ≤ 7.63 ,则待检血清候选判定为上述蛋白中所述的SEQ ID No. 11所示的蛋白的抗体阴性;

[0069] 如果待检血清点样于所述蛋白质芯片上的上述蛋白中所述的SEQ ID No. 13所示的蛋白的所有检测点的 $SNR \geq 8.05$ 且 ≤ 10.87 ,则待检血清候选判定为上述蛋白中所述的SEQ ID No. 13所示的蛋白的抗体阳性,如果 $SNR \geq 6.45$ 且 ≤ 7.84 ,则待检血清候选判定为上述蛋白中所述的SEQ ID No. 13所示的蛋白的抗体阴性;

[0070] 3) 根据步骤2),在待检血清的检测结果中,如果该待检血清对上述蛋白中所述的SEQ ID No. 1所示的蛋白、上述蛋白中所述的SEQ ID No. 3所示的蛋白、上述蛋白中所述的SEQ ID No. 5所示的蛋白、上述蛋白所述中的SEQ ID No. 7所示的蛋白、上述蛋白中所述的SEQ ID No. 9所示的蛋白、上述蛋白中所述的SEQ ID No. 11所示的蛋白、上述蛋白中所述的SEQ ID No. 13所示的蛋白中的至少4种蛋白的抗体为阳性,则候选判定该待检血清是活动性肺结核阳性血清,该血清来源的患者为活动性肺结核患者,否则该待检血清为活动性肺

结核阴性血清,该血清来源的患者为非活动性肺结核患者;

[0071] 所述蛋白质芯片制备时所用的各点样液中的所述的蛋白的浓度均为 50 μ g/ml;

[0072] 所述待检血清的稀释倍数为 100 倍;

[0073] 所述 Cy5 标记的抗人二抗的浓度为 1 μ g/ml;

[0074] 所述 Cy5 标记的抗人二抗具体的商品名称为 Cy5 标记羊抗人抗体二抗,购自 lifeholder,产品目录号为 LH-G11345;

[0075] 所述发射波长的检测通道具体为 635nm 通道;

[0076] 所述各点样液的点样体积为 1nL;

[0077] 所述待检血清稀释时所使用的样品稀释液由溶剂和溶质组成,溶剂为水,溶质为 NaCl、 KH_2PO_4 、 $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$ 和 KCl;所述 NaCl 在所述样品稀释液中的浓度为 8.0g/L,所述 KH_2PO_4 在所述样品稀释液中的浓度为 0.2g/L,所述 $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$ 在所述样品稀释液中的浓度为 2.9g/L,所述 KCl 在所述样品稀释液中的浓度为 0.2g/L。

[0078] 上述蛋白、上述任一所述的蛋白质芯片或上述任一所述的试剂盒在制备诊断或辅助诊断活动性肺结核的产品中的应用也属于本发明的保护范围。

[0079] 通过本发明提供的蛋白质芯片,检测活动性肺结核患者及正常人血清中 7 个蛋白抗原各自相应的 IgG 抗体的水平,联合分析这 7 个蛋白质的相应抗体的检测结果,判断被检人是否为活动性肺结核,检测结果表明,本发明提供的蛋白质芯片辅助诊断活动性肺结核的最佳工作点的特异性为 79%,敏感性为 80.2%,均高于现有技术中活动性肺结核诊断的指标。

[0080] 本发明提供的结核分枝杆菌蛋白质芯片具备快速分析的优势,可以应用于检测活动性肺结核病人相关血清 200 份(活动性肺结核病人 100 份、结核康复病人血清 100 份),在较短时间内比较了活动性肺结核患者和肺结核康复患者样品中的不同。本发明提供的蛋白质芯片对有效诊断活动性肺结核、有效监控肺结核病人的康复过程具有重要意义。

附图说明

[0081] 图 1 为所制备的 7 个抗原蛋白 MT1571、Rv0324、Rv0350、Rv1860、Rv3874、Rv3881c 和 Rv3899c 的银染定量结果图。

[0082] 图 2 为所制备的 7 个抗原蛋白 MT1571、Rv0324、Rv0350、Rv1860、Rv3874、Rv3881c 和 Rv3899c 鉴定结果图。

[0083] 图 3 为应用蛋白质芯片辅助诊断活动性肺结核的方法的特异性和敏感性统计结果图。

具体实施方式

[0084] 下述实施例中所使用的实验方法如无特殊说明,均为常规方法。

[0085] 下述实施例中所用的材料、试剂等,如无特殊说明,均可从商业途径得到。

[0086] 结核分枝杆菌 (*Mycobacterium M. tuberculosis*)H37Rv 菌株(北京株)在文献“Nature 1998Nov 12 ;396(6707):190.”中公开过,公众可从中国科学院生物物理研究所,上海交通大学,中国科学院武汉病毒研究所,广东体必康生物科技有限公司获得。

[0087] 结核分枝杆菌 (*Mycobacterium M. tuberculosis*)CDC1551 菌株在文献“J

Bacteriol. 2002 October ;184(19) :5479 - 5490. ”中公开过, 公众可从中国科学院生物物理研究所, 上海交通大学, 中国科学院武汉病毒研究所, 广东体必康生物科技有限公司获得。

[0088] pDONR221 载体购自 Invitrogen。

[0089] pEGH-A 载体在文献“Jian Zhu, Heng Zhu, et al :J. Virol. May 2009vol. 83no. 10 5219-5231”中公开过, 公众可从中国科学院生物物理研究所, 上海交通大学, 中国科学院武汉病毒研究所, 广东体必康生物科技有限公司获得。

[0090] Pep4 酿酒酵母 (*Saccharomyces cerevisiae*) 在文献“Heng Zhu, Michael Snyder, et al :Nature Genetics 26, 283 - 289 (2000) doi:10. 1038/81576”中公开过, 公众可从广东体必康生物科技有限公司获得。

[0091] 谷胱甘肽琼脂糖亲和介质 (谷胱甘肽 beads) 购自国家生化工程技术研究中心。

[0092] Anti-GST 鼠抗购自 NovaGen, 产品目录号为 71097-3。

[0093] 羊鼠抗荧光 800 通道购自 LI-COR, 产品目录号为 926-32212。

[0094] 商品化的三维 H 载体玻片购自北京博奥生物芯片有限责任公司, 产品目录号为 420042。

[0095] IgG 标准品购自 YEASEN, 产品目录号为 36110ES10。

[0096] Cy5 标记羊抗人抗体二抗购自 lifeholder, 产品目录号为 LH-G11345。

[0097] 实施例 1、蛋白质芯片的制备

[0098] 一、7 种结核分枝杆菌抗原蛋白的制备

[0099] (一) 蛋白的表达

[0100] 1、从分枝杆菌属结核分枝杆菌 (*Mycobacterium M. tuberculosis*) H37Rv 菌株 (北京株) 和 CDC1551 菌株中, 分别通过 PCR 克隆获得 MT1571、Rv0324、Rv0350、Rv1860、Rv3874、Rv3881c 和 Rv3899c 七个蛋白的编码基因片段, 分别将各编码基因片段连接到 pDONR221 载体上, 得到各重组质粒。

[0101] 2、通过 LR 酶 (Invitrogen) 将编码基因的片段换至经过改造的能够表达 GST 标签的 pEGH-A 载体上, 得到各重组质粒。

[0102] 3、将步骤 2 得到的各重组质粒转化到 Pep4 酿酒酵母中, 获得表达 MT1571、Rv0324、Rv0350、Rv1860、Rv3874、Rv3881c 和 Rv3899c 7 种蛋白的酿酒酵母菌株, 在诱导培养基中培养, 待其 OD600 为 0.6-0.8 时, 加入终浓度为 2g/L 的半乳糖, 诱导 6h, 然后 4000rpm 离心收集菌, -80℃ 保存。

[0103] 每 1L 诱导培养基的组分如表 1 所示。

[0104] 表 1 诱导培养基 (1L)

[0105]

成分	货号及厂家	质量 (g/L)
YNB	Y1251 SIGMA	1.7
Ammonium Sulfate	A4418 SIGMA	5

[0106]

SC-URA mix	SIGMA	2
葡萄糖	G0643 SIGMA	20

[0107] 表 1 所示的诱导培养基的余量为水。

[0108] (二) 蛋白的纯化

[0109] 1、裂解液的制备

[0110] 50ml 裂解液 (如表 2 所示) 中加 50 μ l β -巯基乙醇, 125 μ l PMSF 及两片 Roche 蛋白酶抑制剂;

[0111] 2、取 120ml 步骤 (一) 的诱导培养基培养收集到的菌体, 加 400 μ l 氧化锆珠和 400 μ l 裂解液, 4 $^{\circ}$ C 环境中震荡 30s, 然后置冰上 2min, 重复四次;

[0112] 3、11, 000rpm 离心 2min, 取上清于一个新的 15ml 离心管中;

[0113] 4、重复 2 和 3 四次, 将上清收集于同一离心管中;

[0114] 5、添加裂解液至 12ml 即原始诱导培养基体积的 1/10, 同时用没加抑制剂的裂解液将谷胱甘肽 beads 清洗 3 次, 将 300 μ l 的 beads 加入所述溶液中;

[0115] 6、将溶液 4 $^{\circ}$ C 孵育 2h;

[0116] 7、11, 000rpm 离心 2min 后取上清保存于 4 $^{\circ}$ C。Beads 用清洗液 I 和清洗液 II 各洗 3 次;

[0117] 8、加 300 μ l 洗脱缓冲液孵育 beads 15min 后, 离心取上清收集于一新的离心管中, 重复一次, 得到的洗脱液即溶有目的蛋白。

[0118] 上述纯化过程中所用到的裂解液、清洗液 I、清洗液 II 和洗脱缓冲液的组份分别如表 2- 表 5 所示。

[0119] 表 2 裂解液 (1L)

[0120]

成分	厂家	g/L(质量/体积)
Tris	SIGMA	6.057g
NaCl	SIGMA	5.844g
EGTA	SIGMA	0.38g
丙三醇	SIGMA	100mL

[0121] 表 3 清洗液 I (1L)

[0122]

成分	厂家	g/L(质量/体积)
Tris	SIGMA	6.057g
NaCl	SIGMA	5.844g
EGTA	SIGMA	0.38g

[0123]

丙三醇	SIGMA	100mL
TritonX-100 (100%)	SIGMA	1ml
β -巯基乙醇	SIGMA	1ml (即用即加)
PMSF	SIGMA	0.087g

[0124] 表 4 清洗液 II (1L)

[0125]

成分	厂家	g/L(质量/体积)
HEPES	SIGMA	11.916g
NaCl	SIGMA	5.844g
EGTA	SIGMA	0.38g
丙三醇	SIGMA	100mL
TritonX-100 (100%)	SIGMA	1ml
β -巯基乙醇	SIGMA	1ml (现用现加)
PMSF	SIGMA	0.087g

[0126] 表 5 洗脱缓冲液 (1L)

[0127]

成分	厂家	g/L(质量/体积)
HEPES	SIGMA	11.916g
NaCl	SIGMA	5.844g
Glycerol	SIGMA	100mL
谷胱甘肽	SIGMA	3.070g

[0128] 以上裂解液、清洗液 I、清洗液 II 和洗脱缓冲液的余量为水。

[0129] 二、蛋白的鉴定

[0130] 下述实验过程中,所述的溶剂或液体的百分数为体积百分数。

[0131] (一) 材料的准备

[0132] 配置两块 12% 的 SDS-PAGE 胶,一块用于银染定量,一块用于 Western Blotting。

[0133] (二) 将步骤一制备的各目的蛋白溶液各取 20 μ l,进行 SDS-PAGE 用于银染或 Western Blotting。

[0134] 其中,银染时同时制备规格浓度梯度的 BSA 样品作为银染定量的对照,Western Blotting 时所用的第一抗体是 Anti-GST 鼠抗,第二抗体是羊鼠抗荧光 800 通道。

[0135] 具体步骤如下:

- [0136] 1、跑胶
- [0137] 依次每孔加入 12 μ l 步骤一制备的各目的蛋白溶液样品, BSA 梯度样品, 2.5 μ Marker (购自 Takara) 记录次序。80V 30min, 140V 1h。
- [0138] 2、银染操作步骤:
- [0139] 固定: 30min 或者更长时间 40% 乙醇 10% 冰醋酸加水到 250ml ;
- [0140] 致敏: 30min 75ml 乙醇 30% 乙醇 17g 乙酸钠 28.2g 三水乙酸钠 0.5g 硫代硫酸钠 (大苏打) 加水到终体积 250ml ;
- [0141] 水洗: 3x10min ;
- [0142] 银染: 20min 0.625g AgNO₃ 100 μ l 37% 甲醛 (在使用前加入) 加水到终体积 250ml ;
- [0143] 水洗: 2x1min ;
- [0144] 显色: 时间视情况而定 6.25g Na₂CO₃ 50 μ l 37% 甲醛 (在使用前加入) 加水到终体积 250ml ;
- [0145] 终止: 10min 1g 甘氨酸加水到终体积 250ml ;
- [0146] 保存: 1% 冰醋酸, 4 $^{\circ}$ C。
- [0147] 3、Western-Blotting 步骤:
- [0148] 转膜: 15V 40min (半干转, Bio-Rad)。转膜缓冲液: 甘氨酸 2.9g ; Tris 5.8g ; SDS 0.37g ; 甲醇 200ml ; 加 ddH₂O 定容至 1000ml
- [0149] 封闭: 5% 脱脂牛奶 (Bio-Rad) 1h。
- [0150] 第一抗体孵育: Anti-GST 鼠抗终浓度 1 μ g/ml 1h
- [0151] 第二抗体孵育: 羊鼠抗荧光 800 通道终浓度 1 μ g/ml 1h
- [0152] Odyssey 扫描仪扫描, 保存图片。
- [0153] 进行银染定量和 Western-Blotting 鉴定的结果分别如图 1 和图 2 所示。
- [0154] 图 1 结果表明, 所制备的 MT1571、Rv0324、Rv0350、Rv1860、Rv3874、Rv3881c 和 Rv3899c 蛋白的量均为 50 μ g/ml ; 图 2 结果证明, 所制备的 MT1571、Rv0324、Rv0350、Rv1860、Rv3874、Rv3881c 和 Rv3899c 蛋白正确。
- [0155] 经测序, 所制备的各目的蛋白的序列如下:
- [0156] MT1571 的氨基酸序列如 SEQ ID No. 1 所示, 其编码基因序列如 SEQ ID No. 2 所示。
- [0157] Rv0324 的氨基酸序列如 SEQ ID No. 3 所示, 其编码基因序列如 SEQ ID No. 4 所示。
- [0158] Rv0350 的氨基酸序列如 SEQ ID No. 5 所示, 其编码基因序列如 SEQ ID No. 6 所示。
- [0159] Rv1860 的氨基酸序列如 SEQ ID No. 7 所示, 其编码基因序列如 SEQ ID No. 8 所示。
- [0160] Rv3874 的氨基酸序列如 SEQ ID No. 9 所示, 其编码基因序列如 SEQ ID No. 10 所示。
- [0161] Rv3881c 的氨基酸序列如 SEQ ID No. 11 所示, 其编码基因序列如 SEQ ID No. 12 所示。

[0162] Rv3899c 的氨基酸序列如 SEQ ID No. 13 所示，其编码基因序列如 SEQ ID No. 14 所示。

[0163] 三、预点抗原的蛋白质芯片的制备

[0164] 在含上述制备的 50 μ g/ml 的 MT1571 抗原蛋白的洗脱液溶液中，加入终浓度为 25%（体积百分比）的甘油、0.02%（体积百分比）的 Tween20、0.05mg/ml 的 BSA 和 0.1g/L 的 NaN_3 。采用生物芯片点样仪将上述混合液点于载体玻片（载体玻片作为基片，载体玻片为商品化的三维 H 载体玻片）上，每点点样约 1nL、2 个平行点。采用生物芯片点样仪（购自北京博奥生物芯片有限责任公司）点制。

[0165] 将上述 MT1571 抗原蛋白分别替换为 Rv0324、Rv0350、Rv1860、Rv3874、Rv3881c 和 Rv3899c、作为阳性对照的 IgG(1mg/ml) 标准品或 Cy5 标记羊抗人抗体二抗 (20 μ g/ml)，混合液中其它组分和终浓度不变，制备出分别含 Rv0324、Rv0350、Rv1860、Rv3874、Rv3881c、Rv3899c 抗原蛋白、IgG 标准品或 Cy5 标记羊抗人抗体二抗的混合液。

[0166] 以上述相同的点样方法，将分别含 Rv0324、Rv0350、Rv1860、Rv3874、Rv3881c、Rv3899c、IgG 标准品或 Cy5 标记羊抗人抗体二抗的混合液点于上述同一载体玻片上，形成微阵列，每玻片可点 12 个重复平行封闭区间。点样结束后贴上 12 个孔的塑料围栏，并保持在 35% RH 湿度 4 $^{\circ}$ C 环境中 16h 后，将玻片放置于塑料盒中封口 -80 $^{\circ}$ C 低温保存。

[0167] 实施例 2、应用蛋白质芯片辅助诊断活动性肺结核

[0168] 一、待测血清样品的准备

[0169] 将全血标本（由广东省结核病控制中心提供，经过体检者的知情同意）于室温放置 2 小时或 4 $^{\circ}$ C 过夜后于 1000g 离心 20 分钟左右，取上清即可立即检测；或进行分装，并将标本放于 -20 $^{\circ}$ C 或 -80 $^{\circ}$ C 保存，但应避免反复冻融。4 $^{\circ}$ C 解冻后的样品应再次离心，然后检测。

[0170] 二、诊断过程中所涉及的各种缓冲液及试剂的配制

[0171] (1) 样品稀释液（见表 6）：pH 7.4PBS 溶液

[0172] 表 6

[0173]

样品稀释液	质量 (g)
NaCl	8.0
KH_2PO_4	0.2
$\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$	2.9
KCl	0.2
ddH ₂ O	加至 1000 mL

[0174] (2) 洗涤液（见表 7）：pH 7.4 的 PBST 溶液

[0176] 表 7

[0177]

洗涤液	质量 (g)
NaCl	8.0
KH ₂ PO ₄	0.2
Na ₂ HPO ₄ · 12H ₂ O	2.9
KCl	0.2
Tween-20	0.5 mL
ddH ₂ O	加至 1000 mL

[0178] (3) 芯片封闭液 (见表 8): 含 BSA 的 pH 7.4PBS 溶液

[0179] 表 8

[0180]

封闭液	质量 (g)
BSA	10
NaCl	8.0
KH ₂ PO ₄	0.2
Na ₂ HPO ₄ · 12H ₂ O	2.9
KCl	0.2
ddH ₂ O	加至 1000 mL

[0181] (4) Cy5 标记抗人第二抗体的浓缩液: 使用市售的 Cy5 标记羊抗人抗体二抗, 稀释至浓度为 1mg/ml, 分装于避光小管。

[0182] 三、应用蛋白质芯片辅助诊断活动性肺结核

[0183] 使用实施例 1 制备的蛋白质芯片可以检测被检人血液样品中 7 个蛋白抗原 MT1571、Rv0324、Rv0350、Rv1860、Rv3874、Rv3881c 和 Rv3899c 各自相应的 IgG 抗体的水平, 联合分析这 7 个蛋白质的相应抗体的检测结果, 可以判断被检人是否为活动性肺结核。

[0184] (一) 具体操作步骤

[0185] 1) 将实施例 1 点制好密封的蛋白质芯片从 -80℃ 取出, 室温复温 10 分钟。

[0186] 2) 封闭: 将芯片放入洗涤盒, 加入约 50ml 芯片封闭液 (表 8), 摇床 50rpm, 室温 1h。

[0187] 3) 快速甩掉芯片上多余的液体, 置于湿盒内。

[0188] 4) 待测样品的稀释与加样: 将待测血清样品按体积比 1:100 用上述配制的样品稀释液 (表 6) 稀释, 取 30 μL 稀释后的含待测血清的溶液加入到封闭的围栏封闭空间中。反应 1h, 室温。待检测样品临用前 15 分钟内配制。

[0189] 5) 将芯片从湿盒中取出, 置于洗涤盒, 加入约 50ml 上述配制的洗涤液 (见表 7), 摇床 50rpm, 室温 5min, 重复 3 次。

[0190] 6) 快速甩掉芯片上多余的液体, 置于湿盒内。

[0191] 7) 每个封闭空间加入 30 μL 已经稀释至终浓度 1 μg/ml 的 Cy5 标记抗人第二抗体于芯片上, 避光反应 1h。

[0192] 8) 将芯片从湿盒中取出,置于洗涤盒,加入约 50ml 上述配制的洗涤液(见表 7),摇床 50rpm,室温 5min,重复 3 次。再用约 50ml 超纯水洗一次,5min。

[0193] 9) 离心干燥,用 Genepix 扫描仪在 635nm 通道下读取数据。

[0194] (二) 待测样品分别对 7 种抗原蛋白抗性的判定

[0195] 用 Genepix 扫描仪在 635nm 通道(该通道的选择是 Cy5 荧光标记决定的)扫描得到的结果是通过点样点的信噪比(SNR)来进行衡量,点样点的信噪比(SNR)计算公式为:

$$[0196] \quad \text{SNR} = \frac{S - B}{\sigma} \quad (\text{公式 1})$$

[0197] 公式 1 中, S 和 B 分别代表芯片中原始的信号值(是扫描仪直接显示的数值)和背景值(非样品点制区域的信号值),而 σ 则是代表该点原始信号的标准差(σ 对应的实验次数为 3 次)。

[0198] 将蛋白质芯片上点样点的信噪比值与下述表 9-表 15 的信噪比区间(所述信噪比区间包括所给出的信噪比区间的两端值)进行对比,以进行待测样品对相应抗原抗性的阳性、阴性结果的判定。

[0199] 抗 MT1571(见表 9):

[0200] 表 9

[0201]

信噪比区间	判定
(8.32, 10.21)	阳性
(6.94, 8.27)	阴性

[0202] 抗 Rv0324(见表 10):

[0203] 表 10

[0204]

信噪比区间	判定
(7.68, 9.95)	阳性
(6.31, 7.55)	阴性

[0205] 抗 Rv0350(见表 11):

[0206] 表 11

[0207]

信噪比区间	判定
(7.16, 9.75)	阳性
(6.50, 7.03)	阴性

[0208] 抗 Rv1860(见表 12):

[0209] 表 12

[0210]

信噪比区间	判定
(7.44, 11.16)	阳性
(6.04, 7.20)	阴性

[0211] 抗 Rv3874 (见表 13) :

[0212] 表 13

[0213]

信噪比区间	判定
(7.98, 9.83)	阳性
(6.84, 7.70)	阴性

[0214] 抗 Rv3881c (见表 14) :

[0215] 表 14

[0216]

信噪比区间	判定
(7.81, 10.62)	阳性
(6.29, 7.63)	阴性

[0217] 抗 Rv3899c (见表 15) :

[0218] 表 15

[0219]

信噪比区间	判定
(8.05, 10.87)	阳性
(6.45, 7.84)	阴性

[0221] (三) 待测样品检验结果的判定

[0222] 判定标准 : 如果待测样品的检测结果中, 至少有 4 种抗体呈阳性 (判断抗体是否为阳性的标准 : 实施例 1 制备的蛋白质芯片上同一抗原的 2 个平行点样点都判定为阳性时, 该抗体为阳性), 则判定该待测样品是活动性肺结核阳性, 否则为活动性肺结核阴性。

[0223] 四、应用蛋白质芯片辅助诊断活动性肺结核的特异性和敏感性

[0224] 采用 200 份活动性肺结核相关患者的血清 (临床诊断为活动性肺结核的患者 100 份, 肺结核康复患者 100 份; 血清样本由广东省结核病控制中心提供, 血清样本的取得均经过患者及体检者的知情同意。) 对实施例 1 的蛋白质芯片进行了特异性和敏感性检测, 诊断阈值为阳性抗体数大于或等于 4 即被检人阳性抗体数大于或等于 4, 即被判断为活动性肺结核阳性。

[0225] 检测结果如图 3 所示。

[0226] 图 3 中, 横坐标假阳性率代表特异性, 纵坐标真阳性率代表敏感性。根据阳性似然

比即敏感性 / 特异性计算得出实施例 1 的蛋白质芯片辅助诊断活动性肺结核的最佳工作点的特异性为 79%，敏感性为 80.2%，均大大高于现有技术中活动性肺结核诊断的指标。

[0227] 100 份活动性肺结核患者中，有 80 份被检测出阳性；100 份肺结核康复患者中，有 21 份被检测出阳性。

[0001]

序列表

<110>中国科学院生物物理研究所 上海交通大学 中国科学院武汉病毒研究所 广东体必康生物科技有限公司
 <120>一组结核分枝杆菌蛋白及其编码基因与应用
 <160>14

<210>1

<211>66

<212> PRT

<213>结核分枝杆菌 (*Mycobacterium M. tuberculosis*)

<400>1

```
Met Pro Ser Ser Gly Thr Ala Leu Ala His Pro Asp Gln Ser Leu Glu
1           5           10          15
Asp Ile Arg Thr Gly Glu Leu Thr Asp Leu Lys Asp Gly Pro Gln Gly
          20          25          30
Tyr Leu Met Ala Leu Ser Val Val Glu Ser Ala Ala Thr Gly Pro Ile
          35          40          45
Pro His Ala Pro Ser Ile Glu Ala Arg Arg Ala Ile Tyr Gln Asp Leu
          50          55          60
Gly Met
65
```

<210>2

<211>201

<212>DNA

<213>结核分枝杆菌 (*Mycobacterium M. tuberculosis*)

<400>2

```
atgccgagtt cgggaaccgc actagctcat ccgcaccaga gcotggagga catcaggacc      60
ggcgagttga cagatctcaa agatgggccc caaggetatt tgatggctct ttgggtgctc      120
gagtggcgg caaccggtcc gataccgcac gctccatcca tagaggccag acgggggatc      180
tatcaagate taggcattgtg a                                201
```

<210>3

<211>226

<212> PRT

<213>结核分枝杆菌 (*Mycobacterium M. tuberculosis*)

<400>3

```
Met Ala Gly Gln Ser Asp Arg Lys Ala Ala Leu Leu Asp Gln Val Ala
1           5           10          15
Arg Val Gly Lys Ala Leu Ala Asn Gly Arg Arg Leu Gln Ile Leu Asp
          20          25          30
Leu Leu Ala Gln Gly Glu Arg Ala Val Glu Ala Ile Ala Thr Ala Thr
          35          40          45
Gly Met Asn Leu Thr Thr Ala Ser Ala Asn Leu Gln Ala Leu Lys Ser
          50          55          60
Gly Gly Leu Val Glu Ala Arg Arg Glu Gly Thr Arg Gln Tyr Tyr Arg
```

[0002]

Ser Val Leu Glu Gly Gly Asp Pro Val Val Val Ala Asn Ser Glu Gly
 20 25 30
 Ser Arg Thr Thr Pro Ser Ile Val Ala Phe Ala Arg Asn Gly Glu Val
 35 40 45
 Leu Val Gly Gln Pro Ala Lys Asn Gln Ala Val Thr Asn Val Asp Arg
 50 55 60
 Thr Val Arg Ser Val Lys Arg His Met Gly Ser Asp Trp Ser Ile Glu
 65 70 75 80
 Ile Asp Gly Lys Lys Tyr Thr Ala Pro Glu Ile Ser Ala Arg Ile Leu
 85 90 95
 Met Lys Leu Lys Arg Asp Ala Glu Ala Tyr Leu Gly Glu Asp Ile Thr
 100 105 110
 Asp Ala Val Ile Thr Thr Pro Ala Tyr Phe Asn Asp Ala Gln Arg Gln
 115 120 125
 Ala Thr Lys Asp Ala Gly Gln Ile Ala Gly Leu Asn Val Leu Arg Ile
 130 135 140
 Val Asn Glu Pro Thr Ala Ala Ala Leu Ala Tyr Gly Leu Asp Lys Gly
 145 150 155 160
 Glu Lys Glu Gln Arg Ile Leu Val Phe Asp Leu Gly Gly Gly Thr Phe
 165 170 175
 Asp Val Ser Leu Leu Glu Ile Gly Glu Gly Val Val Glu Val Arg Ala
 180 185 190
 Thr Ser Gly Asp Asn His Leu Gly Gly Asp Asp Trp Asp Gln Arg Val
 195 200 205
 Val Asp Trp Leu Val Asp Lys Phe Lys Gly Thr Ser Gly Ile Asp Leu
 210 215 220
 Thr Lys Asp Lys Met Ala Met Gln Arg Leu Arg Glu Ala Ala Glu Lys
 225 230 235 240
 Ala Lys Ile Glu Leu Ser Ser Ser Gln Ser Thr Ser Ile Asn Leu Pro
 245 250 255
 Tyr Ile Thr Val Asp Ala Asp Lys Asn Pro Leu Phe Leu Asp Glu Gln
 260 265 270
 Leu Thr Arg Ala Glu Phe Gln Arg Ile Thr Gln Asp Leu Leu Asp Arg
 275 280 285
 Thr Arg Lys Pro Phe Gln Ser Val Ile Ala Asp Thr Gly Ile Ser Val
 290 295 300
 Ser Glu Ile Asp His Val Val Leu Val Gly Gly Ser Thr Arg Met Pro
 305 310 315 320
 Ala Val Thr Asp Leu Val Lys Glu Leu Thr Gly Gly Lys Glu Pro Asn
 325 330 335
 Lys Gly Val Asn Pro Asp Glu Val Val Ala Val Gly Ala Ala Leu Gln
 340 345 350
 Ala Gly Val Leu Lys Gly Glu Val Lys Asp Val Leu Leu Leu Asp Val
 355 360 365
 Thr Pro Leu Ser Leu Gly Ile Glu Thr Lys Gly Gly Val Met Thr Arg
 370 375 380

[0004]

Leu Ile Glu Arg Asn Thr Thr Ile Pro Thr Lys Arg Ser Glu Thr Phe
 385 390 395 400
 Thr Thr Ala Asp Asp Asn Gln Pro Ser Val Gln Ile Gln Val Tyr Gln
 405 410 415
 Gly Glu Arg Glu Ile Ala Ala His Asn Lys Leu Leu Gly Ser Phe Glu
 420 425 430
 Leu Thr Gly Ile Pro Pro Ala Pro Arg Gly Ile Pro Gln Ile Glu Val
 435 440 445
 Thr Phe Asp Ile Asp Ala Asn Gly Ile Val His Val Thr Ala Lys Asp
 450 455 460
 Lys Gly Thr Gly Lys Glu Asn Thr Ile Arg Ile Gln Glu Gly Ser Gly
 465 470 475 480
 Leu Ser Lys Glu Asp Ile Asp Arg Met Ile Lys Asp Ala Glu Ala His
 485 490 495
 Ala Glu Glu Asp Arg Lys Arg Arg Glu Glu Ala Asp Val Arg Asn Gln
 500 505 510
 Ala Glu Thr Leu Val Tyr Gln Thr Glu Lys Phe Val Lys Glu Gln Arg
 515 520 525
 Glu Ala Glu Gly Gly Ser Lys Val Pro Glu Asp Thr Leu Asn Lys Val
 530 535 540
 Asp Ala Ala Val Ala Glu Ala Lys Ala Ala Leu Gly Gly Ser Asp Ile
 545 550 555 560
 Ser Ala Ile Lys Ser Ala Met Glu Lys Leu Gly Gln Glu Ser Gln Ala
 565 570 575
 Leu Gly Gln Ala Ile Tyr Glu Ala Ala Gln Ala Ala Ser Gln Ala Thr
 580 585 590
 Gly Ala Ala His Pro Gly Gly Glu Pro Gly Gly Ala His Pro Gly Ser
 595 600 605
 Ala Asp Asp Val Val Asp Ala Glu Val Val Asp Asp Gly Arg Glu Ala
 610 615 620

Lys

625

<210>6

<211>1878

<212>DNA

<213>结核分枝杆菌 (*Mycobacterium M. tuberculosis*)

<400>6

```

atggtctgtg cggtcgggat cgacctcggg accaccaact ccgttgtctc ggttotggaa      60
ggtggcgacc cggtcgtcgt cgccaactcc gagggctcca ggaccacccc gtcaattgtc      120
gcgttcgccc gcaacggtga ggtgctggtc ggccagcccg ccaagaacca ggcagtgacc      180
aacgtcgatc gcaccgtgcg ctcggtcaag cgacacatgg gcagcgactg gtccatagag      240
attgacggca agaaatacac cgcgcgggag atcagcgccc gcattotgat gaagctgaag      300
cgcgacgceg aggcctacct cggtgaggac attaccgacg cggttatcac gacgcccgcc      360
tacttcaatg acgcccagcg tcaggccacc aaggacgcgc gccagatcgc cggcctcaac      420
gtgctgcgga tcgtcaacga gccgaccgog gccgcgctgg cctacggcct cgaacagggc      480
gagaaggagc agcgaatcct ggtottcgac ttgggtggtg gcaetttoga cgtttccctg      540

```

[0005]

```

ctggagatcg gegaggggtgt ggttgaggtc cgtgccactt cgggtgacaa ccacctcggc 600
ggcgacgaact gggaccagcg ggtcgtcgat tggetgggtgg acaagttcaa gggcaccagc 660
ggcatcgatc tgaccaagga caagatggcg atgcagcggc tgcgggaagc cgcgagaag 720
gcaaagatcg agctgagttc gagtcagtc accctcgatca acctgcccta catcacgctc 780
gacgcgcgaca agaaccggtt gttccttagac gageagctga cccgcgcgga gttcoaacgg 840
atcaactcagg accctgctgga ccgcaactgc aagccggttc agtcgggtgat cgtgacacc 900
ggcatttcgg tgtcggagat cgatcacgtt gtgctcgtgg gtggttcgac ccggtgccc 960
ggcgtgaccg atctgggtcaa ggaactcacc ggcggcaagg aaccaacaa gggcgtcaac 1020
cccgatgagg ttgtcgcggc gggagcgcct ctgcaggcgc gcgtcctcaa gggcgagggtg 1080
aaagacgttc tgctgcttga tgttaccocg ctgagcctgg gtatcgagac caaggcggg 1140
gtgatgacca ggtcctatga gcgcaacacc acgatcccca ccaagcggtc ggagactttc 1200
accaccgcgc acgacaacca accgtcgggtg cagatccagg tctatcaggg ggagcgtgag 1260
atcgcgcgcg acaacaagtt gctcgggtcc ttcgagctga ccgcatccc gccggcgcgc 1320
cgggggattc cgcagatcga ggtaactttc gacatcgacg ccaacggcat tgtgcaactc 1380
accgcacaagg acaagggeac cggcaaggag aacaacgatcc gaatccagga aggctcgggc 1440
ctgtccaagg aagacattga ccgcatgatc aaggacgcgc aagcgcacgc cgaggaggat 1500
cgcaagcgtc gcgaggaggc cgatgttcgt aatcaagccg agacattggt ctaccagacg 1560
gagaagtctg tcaaaagaaca gcgtgaggcc gaggggtggtt cgaaggtaac tgaagacacg 1620
ctgaacaagg ttgatgcgcg ggtggcggaa gcgaaggcgg cacttgccgg atcggatatt 1680
tcggccatca agtcggcgat ggagaagctg gccaggagat cgcaggctct ggggcaagcg 1740
atctaogaag cagctcaggc tgcgteacag gccactggcg ctgcccacc cggcggcgag 1800
cggggcggtg cccaccocgg ctcggtgat gacgttgtgg acgcggagggt ggtcgaacgc 1860
ggcggggagg ccaagtga 1978

```

<210>7

<211>325

<212> PRT

<213>结核分枝杆菌 (*Mycobacterium M. tuberculosis*)

<400>7

```

Met His Gln Val Asp Pro Asn Leu Thr Arg Arg Lys Gly Arg Leu Ala
1           5           10           15
Ala Leu Ala Ile Ala Ala Met Ala Ser Ala Ser Leu Val Thr Val Ala
20           25           30
Val Pro Ala Thr Ala Asn Ala Asp Pro Glu Pro Ala Pro Pro Val Pro
35           40           45
Thr Thr Ala Ala Ser Pro Pro Ser Thr Ala Ala Ala Pro Pro Ala Pro
50           55           60
Ala Thr Pro Val Ala Pro Pro Pro Pro Ala Ala Ala Asn Thr Pro Asn
65           70           75           80
Ala Gln Pro Gly Asp Pro Asn Ala Ala Pro Pro Ala Asp Pro Asn
85           90           95
Ala Pro Pro Pro Pro Val Ile Ala Pro Asn Ala Pro Gln Pro Val Arg
100          105          110
Ile Asp Asn Pro Val Gly Gly Phe Ser Phe Ala Leu Pro Ala Gly Trp
115          120          125
Val Glu Ser Asp Ala Ala His Phe Asp Tyr Gly Ser Ala Leu Leu Ser

```

[0006]

130 135 140
 Lys Thr Thr Gly Asp Pro Pro Phe Pro Gly Gln Pro Pro Pro Val Ala
 145 150 155 160
 Asn Asp Thr Arg Ile Val Leu Gly Arg Leu Asp Gln Lys Leu Tyr Ala
 165 170 175
 Ser Ala Glu Ala Thr Asp Ser Lys Ala Ala Ala Arg Leu Gly Ser Asp
 180 185 190
 Met Gly Glu Phe Tyr Met Pro Tyr Pro Gly Thr Arg Ile Asn Gln Glu
 195 200 205
 Thr Val Ser Leu Asp Ala Asn Gly Val Ser Gly Ser Ala Ser Tyr Tyr
 210 215 220
 Glu Val Lys Phe Ser Asp Pro Ser Lys Pro Asn Gly Gln Ile Trp Thr
 225 230 235 240
 Gly Val Ile Gly Ser Pro Ala Ala Asn Ala Pro Asp Ala Gly Pro Pro
 245 250 255
 Gln Arg Trp Phe Val Val Trp Leu Gly Thr Ala Asn Asn Pro Val Asp
 260 265 270
 Lys Gly Ala Ala Lys Ala Leu Ala Glu Ser Ile Arg Pro Leu Val Ala
 275 280 285
 Pro Pro Pro Ala Pro Ala Pro Ala Pro Ala Glu Pro Ala Pro Ala Pro
 290 295 300
 Ala Pro Ala Gly Glu Val Ala Pro Thr Pro Thr Thr Pro Thr Pro Gln
 305 310 315 320
 Arg Thr Leu Pro Ala
 325

<210>8

<211>978

<212>DNA

<213>结核分枝杆菌 (*Mycobacterium M. tuberculosis*)

<400>8

```

atgcatcagg tggaccccaa cttgacacgt cgcaagggac gattggcggc actggctatc      60
ggggcgatgg ccagcgccag cctggtgacc gttgcggtgc ccgcgaccgc caacgcgatc      120
cgggagccag cgccccggt acccaaacg gcgcctcgc cgccgtgcac cgtcgcagcg      180
ccaccgcgac cggcgacacc tgttgcccc ccaccaccg ccgcccgcga caogccgaat      240
gcccagccgg gcgataccaa cgcagcacct ccgcccggcg acccgaacgc accgcccgca      300
cctgtcattg ccccaaacgc accccaacct gtccggatcg acaaccgggt tggaggattc      360
agottcgcgc tgctcgtcgg ctgggtggag tctgacgcgc cccacttoga ctacggttca      420
gcactcctca gcaaaaaccac cggggaccgc ccatttcccg gacagccgcc gccggtggcc      480
aatgacacce gtatcgtgct cggccggcta gaccaaaagc tttacgccag cgcggaagcc      540
acgactcca aggcgcgggc ccggttgggc tcggacatgg gtgagttota tatgcctac      600
ccgggcaccc ggatcaacca ggaaaccgtc tcgctcgacg ccaaccgggt gtctggaagc      660
gogtcgtatt acgaagtoaa gttcagcgat ccgagtaagc cgaacggcca gatctggacg      720
ggcgtaatcg gctcgcggc ggcaaacgca ccggaccccg ggccccctca gcgctggttt      780
gtggtatggc tcgggaccgc caacaaccog gtggacaagc gcgcgcccaa ggcgctggcc      840
gaatcgatcc ggcctttggt cgcgcccgcg ccggcgcggc caccggctcc tgcagagccc      900
gtcccggcgc cggcgccggc cggggaagtc gctcctacc cgcgcgaccc gacaccgcag      960
    
```

[0007]


```

                420                425                430
Asp Glu Ala Leu Tyr Thr Glu Asp Arg Ala Trp Thr Glu Ala Val Ile
                435                440                445
Gly Asn Arg Arg Arg Gln Asp Ser Lys Glu Ser Lys
                450                455                460
<210>12
<211>1383
<212>DNA
<213>结核分枝杆菌 (Mycobacterium M. tuberculosis)
<400>12
atgacgcagt cgcagaccgt gacggtggat cagcaagaga ttttgaacag ggccaacgag      60
gtggaggccc cgatggcggg cccaccgact gatgtcccca tcacaccgtg cgaactcacg    120
gcggtataaaa acgcgcacca acagctggta ttgtccgcgc acaacatggg ggaatacctg    180
gcggcgcggtg ccaaagagcg gcagcgtctg gcgacctcgc tgcgcaacgc ggccaaggcg    240
tatggcgagg ttgatgagga ggcctgcgacc gcgctggaca acgacggcga aggaactgtg    300
caggsagaat cggccggggc cgtcggaggg gacagtccgg ccgaaactaac cgatacgcgcg    360
agggtggcca cggccggtga acccaacttc atggtatcca aagaagcggc aaggaagctc    420
gaaaacggcg accaaggcgc atcgcctcgc cactttgcgg atgggtggaa cactttcaac    480
ctgacgctgc aaggcgacgt caagcgttcc cgggggtttg acaactggga aggcgatgcg    540
gctaaccgctt gcgagccttc gctcgatcaa caacggcaat ggatactcca catggccaaa    600
ttgagcgctg cgatggccaa gcaggctcaa tatgtcgcgc agctgcaact gtgggctagg    660
cgggaacatc cgaactatga agacatagtc gggetcgaac ggetttacgc ggaaaaacct    720
tcggcccgcg accaaattct cccggtgtac gcggagtatc agcagaggtc ggagaagggtg    780
ctgaccgaat acaacaacaa ggcagcctcg gaaccggtaa accgcgcgaa gcctccccc    840
gccatcaaga tcgaccgcgc ccgcctccg caagagcagg gattgatccc tggcttcctg    900
atgcgcgctg ctgaccgctc cgggtgtgact ccggtaccg ggatgccagc cgcaccgatg    960
gttccgccta ccgatcgcgc ggggtgtggc ctccggctg acacggcggc gcagctgacg   1020
tcggctgggc gggaaagcgc agcctgtgct ggcgacgtgg cggtaaaagc ggcacgcctc   1080
ggctggcggtg gaggcgcgcg ggtgcctcgc gcgcctttgg gatccgcgat cggggcgccc   1140
gaatcggtgc ggcccctcgc cgtcgtgtgac attgccggct taggccaggg aagggcggc   1200
ggcgcgcgcg cgtcggcgcg cgggtgcatg ggaatgcgga tgggtgcgcg gcacagggga   1260
caagggggcg ccaagtccaa gggttctcag caggaagacg aggcgctota cacccaggat   1320
cgggcacgga ccgagccgct catttgtaac cgtcggcgcc agaacagtaa ggagtogaag   1380
tga
                1383

```

```

<210>13
<211>410
<212> PRT
<213>结核分枝杆菌 (Mycobacterium M. tuberculosis)
<400>13

```

```

Met Val Thr Gly Gln Pro Ala Ala Ala Gly Ala His Ser Leu Ser Glu
1                5                10                15
Gly Ala Met Thr Ala Met Gln Ser Gly Ser Val Pro Pro Pro Gln Ala
                20                25                30
Thr Pro Pro Ile Thr Thr Pro Pro Val Val Ser Ala Pro Thr Met Ala
                35                40                45

```

[0010]

Ala Gly Ile Glu Ala Thr His Gly Pro Val Asp Thr Pro Ala Asn Thr
50 55 60
Ser Gly Ala Pro Pro Ala Ser Thr Gly Thr Thr Gly Pro Val Ala Pro
65 70 75 80
Thr Val Val Thr Ala Gly Pro Val Ala Ala Pro Ala Ala Pro Val Val
85 90 95
Gly Gly Ser Ala Val Pro Ala Gly Pro Leu Pro Ala Tyr Gly Ser Asp
100 105 110
Leu Arg Pro Pro Val Val Ala Ala Pro Ala Val Pro Ser Val Pro Thr
115 120 125
Ala Pro Val Ser Gly Ala Pro Val Ala Pro Ser Ala Ser Ser Ala Pro
130 135 140
Ser Ala Gly Gly Ala Leu Val Ser Pro Val Glu Arg Ala Ala Ser Lys
145 150 155 160
Ala Val Ala Gly Gln Ala Gly Ala Ser Ser Ser Thr Met Ala Gly Ala
165 170 175
Ser Ala Leu Ser Ala Thr Ala Gly Ala Thr Ala Gly Ala Val Ser Ala
180 185 190
Arg Ala Ala Glu Gln Gln Arg Leu Gln Arg Ile Val Asp Ala Val Ala
195 200 205
Arg Gln Glu Pro Arg Ile Ser Trp Ala Ala Gly Leu Arg Asp Asp Gly
210 215 220
Thr Thr Thr Leu Leu Val Thr Asp Leu Ala Gly Gly Trp Ile Pro Pro
225 230 235 240
His Val Arg Leu Pro Ala Asn Val Thr Leu Leu Glu Pro Thr Ala Arg
245 250 255
Arg Arg Asp Ala Asp Val Ile Asp Leu Leu Gly Ala Val Val Ala Val
260 265 270
Ala Ala His Glu Ser Asn Thr Tyr Val Ala Glu Pro Gly Pro Asp Ala
275 280 285
Pro Ala Leu Thr Gly Asp Arg Ser Ala Arg Ser Ala Ile Pro Lys Val
290 295 300
Asp Glu Phe Gly Pro Thr Leu Val Glu Ala Val Arg Arg Arg Asp Ser
305 310 315 320
Leu Pro Arg Ile Ala Gln Ala Ile Ala Leu Pro Ala Val Arg Lys Thr
325 330 335
Gly Val Leu Glu Asn Glu Ala Glu Leu Leu His Gly Cys Ile Thr Ala
340 345 350
Val Lys Glu Ser Val Leu Lys Ala Tyr Pro Ser His Glu Leu Thr Ala
355 360 365
Val Gly Asp Trp Met Leu Leu Ala Ala Ile Glu Ala Leu Ile Asp Glu
370 375 380
Gln Asp Tyr Leu Ala Asn Tyr His Leu Ala Trp Tyr Ala Val Thr Thr
385 390 395 400
Arg Arg Gly Gly Ser Arg Gly Phe Ala Ala
405 410

[0011]

<210>14

<211>1233

<212>DNA

<213>结核分枝杆菌 (*Mycobacterium M. tuberculosis*)

<400>14

```

atggtgacgg ggcaaccggc cggggtggc ggcattgcg tgtcggaagg ggcaatgacg      60
gogatgcagt gggggtcggg tccccgcgc caggccactc etccgataac gacaccgccc      120
gtggtgtctg cggcgaacct ggctgggggc atcgaagcca cacaccgggc agttgacacg      180
cggcggaaca cctcgggcgc tccaccggcg tcgaccggta ctaccgggca ggtcggcgcg      240
accgtggtga ccgcggggcc ggtgggggca ccgctgcgc cggtggttgg tggctcggct      300
gtccccgchg gaccgctgcc ggcttacggc tctgatctac ggcccccgct cgtggcagcc      360
cccgcctgtc cctcggttcc taaggcgccc gtatccggcg cgcgggtggc gccctcggcg      420
tcaccggccc catcggcggg tggggcgctg gttctcccg tggagcgcg agcctcgaaa      480
gctgtggctg gacaggetgg tgcgagctcg tcgacaatgg ccggcgccctc ggcactgtcg      540
gccaccgccc ggcggaacgc gggcggggta tcggctcggg cggctgagca gcaaccgcta      600
cagcgaatcg tggatgccgt ggcgcgccag gaccgcgcaa tctcatgggc ggccgggctg      660
cggcagcagc gcaccaccac cctgctggtc accgatttgg ccggcgggtg gattccgccc      720
cacgtcgggc tgcccgcgaa cgtgacgctg ctggagocaa ccgcgcgacg ccgtgatgcc      780
gacgtgatcg acttgctggg cgcctcgtc gccgtggcag cccaccagtc caaccacctc      840
gttgccgagc cggggccaga tgcgcctgcg ctgaccgggtg atcggtcagc gcctcggcg      900
atacccaagg tggacgagtt cgggcgcacc ttggtcgaag ctgtgcgccc ccgcgatagc      960
ctgcgcggga tcgcgcaggc gatcgcgctg ccggcggtac gcaaaaaccg cgtgctggaa      1020
aacgaagccc agctgctgca cggctgcctc accgcggtea aggagtcggg gctcaaggcc      1080
tatcccagtc acgagctcac cgtctcggg gattggatgc tgcctggggc gatcggggca      1140
ctgatcgacg agcaggacta cctcgcacaac taccacctgg cctggtatgc cgtaacccac      1200
aggcgtggcg gctcacgagg gttcctgccc tga.

```

1233



图 1

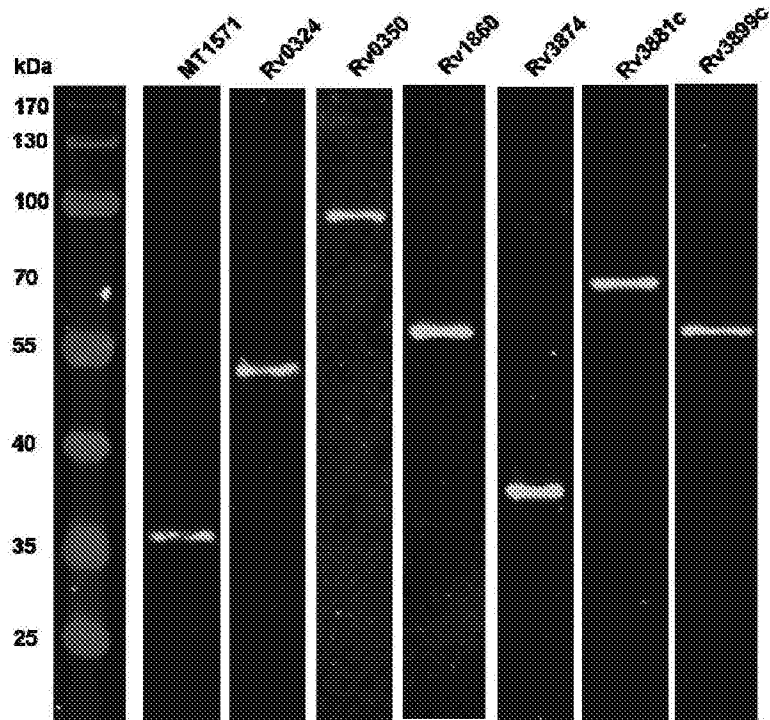


图 2

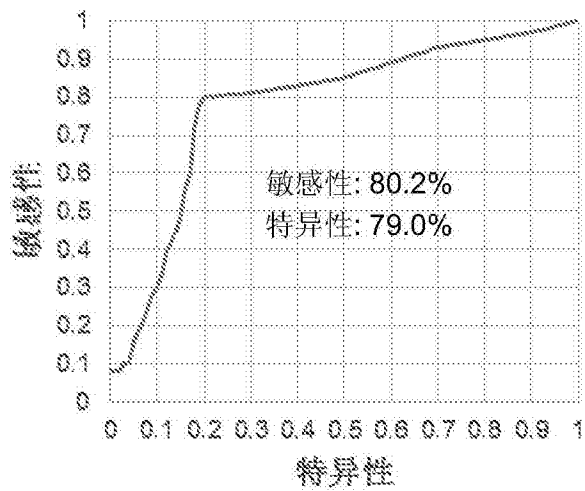


图 3