



[12] 发明专利申请公开说明书

[21] 申请号 02829305.3

[43] 公开日 2005 年 7 月 13 日

[11] 公开号 CN 1639570A

[22] 申请日 2002.5.20 [21] 申请号 02829305.3

[30] 优先权

[32] 2002.5.17 [33] US [31] 10/147,299

[86] 国际申请 PCT/BR2002/000072 2002.5.20

[87] 国际公布 WO2003/098214 英 2003.11.27

[85] 进入国家阶段日期 2005.1.11

[71] 申请人 奥斯瓦尔多克鲁兹基金会

地址 巴西里约热内卢

[72] 发明人 A·I·科 M·加尔瓦雷斯

J·H·罗萨克罗达

I·C·斯奎拉 D·A·哈克

J·马苏纳加 L·W·赖利

M·巴罗奇 T·A·扬

[74] 专利代理机构 中国国际贸易促进委员会专利商
标事务所

代理人 刘晓东

权利要求书 3 页 说明书 31 页 序列表 10 页
附图 10 页

[54] 发明名称 钩端螺旋体内存在的具有重复性细菌 IG 样 (BIG) 结构域的蛋白质

[57] 摘要

本发明涉及编码钩端螺旋体细菌中具有重复细菌 Ig 样 (Big) 结构域的蛋白质 BigL1、BigL2 和 BigL3 的三种分离 DNA 分子以及它们在诊断、治疗和疫苗应用中的用途。依照本发明, 编码 BigL1、BigL2 和 BigL3 蛋白质的分离分子可用于诊断和预防能在人类和其它哺乳动物体内引起疾病的钩端螺旋体感染, 包括具有兽医学重要性的那些钩端螺旋体的感染。

1. 具有如SEQ ID NO: 2所示氨基酸序列或其功能等同序列的基本上纯的多肽。
2. 编码如SEQ ID NO: 2所示氨基酸序列或其功能等同序列的分离的多核苷酸片段。
3. 选自SEQ ID NO: 1或其功能等同序列的分离的多核苷酸。
4. 具有如SEQ ID NO: 4所示氨基酸序列或其功能等同序列的基本上纯的多肽。
5. 编码如SEQ ID NO: 4所示氨基酸序列或其功能等同序列的分离的多核苷酸。
6. 选自SEQ ID NO: 3或其功能等同序列的分离的多核苷酸。
7. 具有如SEQ ID NO: 6所示氨基酸序列或其功能等同序列的基本上纯的多肽。
8. 编码如SEQ ID NO: 6所示氨基酸序列或其功能等同序列的分离的多核苷酸。
9. 选自SEQ ID NO: 5或其功能等同序列的分离的多核苷酸。
10. 权利要求2、5或8的多核苷酸，其中所述多核苷酸序列来自钩端螺旋体 (*Leptospira*) 的种。
11. 选自SEQ ID NO: 8或其功能等同序列的基本纯的多肽序列。
12. 编码如SEQ ID NO: 8所示氨基酸序列或其功能等同序列的分离的多核苷酸片段。
13. 选自SEQ ID NO: 7或其功能等同序列的分离的多核苷酸。
14. 选自SEQ ID NO: 10或其功能等同序列的基本纯的多肽序列。
15. 编码如SEQ ID NO: 10所示氨基酸序列或其功能等同序列的分离的多核苷酸片段。
16. 选自SEQ ID NO: 9或其功能等同序列的分离的多核苷酸。
17. 可结合权利要求1、4、7、11和14中所述的基本纯的多肽或其功能等同序列的抗体。

18. 用于在哺乳动物受试者中诱发针对致病螺旋菌的免疫反应的药物组合物，它包括在药用可接受载体中的免疫原性有效量的权利要求1、4、7、11和14所述的基本纯的多肽或其功能等同序列。

19. 权利要求18的药物组合物，其中所述药用可接受载体包括佐剂。

20. 权利要求18的药物组合物，其中所述致病螺旋菌选自密螺旋体（*Treponema*）、疏螺旋体（*Borrelia*）和钩端螺旋体（*Leptospira*）。

21. 可用于在哺乳动物受试者中提供针对致病螺旋菌的免疫反应的药物组合物，它包括在药用可接受载体中的免疫学有效量的抗体，所述抗体能结合权利要求1、4、7、11和14所述的多肽或其功能等同序列。

22. 在哺乳动物受试者中诱发抗致病钩端螺旋体感染的免疫反应的方法，包括施用权利要求18的药物组合物。

23. 在哺乳动物受试者中诱发抗致病螺旋菌的免疫反应的方法，包括对受试者施用权利要求19的药物组合物。

24. 检测样品中病原体的方法，包括将疑似含致病螺旋菌的样品与可结合该病原体特异性的细胞组分的试剂接触并检测试剂与所述组分的结合。

25. 依照权利要求24的方法，其中可与病原体特异性的细胞组分相结合的试剂是用于鉴定bigL1、bigL2和bigL3多核苷酸的寡核苷酸。

26. 依照权利要求24的方法，其中与病原体特异性的细胞组分相结合的试剂是抗BigL1、BigL2或BigL3多肽或具有功能等同序列的多肽的抗体。

27. 权利要求22的方法，其中所述螺旋菌特异的组分是编码权利要求1、4、7、11和14所述多肽或其功能等同序列的多核苷酸。

28. 权利要求22的方法，其中所述螺旋菌特异的组分是权利要求1、4、7、11和14所述多肽或其功能等同序列。

29. 检测样品中抗权利要求1、4、7、11和14所述多肽或其功能等同序列的抗体的方法，它包括在允许该抗体与权利要求1、4、7、11和14所述多肽或其功能等同序列相结合的条件下将样品与权利要求1、4、

7、11和14所述多肽或其功能等同序列接触，并检测该抗体与权利要求1、4、7、11和14所述多肽或其功能等同序列的结合。

30. 检测样品中权利要求1、4、7、11和14所述多肽或其功能等同序列的方法，包括在允许样品中的所述多肽与抗体结合的条件下将样品与抗权利要求1、4、7、11和14所述多肽或其功能等同序列的抗体接触，并检测多肽与所说抗体的结合。

31. 检测样品中编码BigL1、BigL2、BigL3或其功能等同序列的核酸的方法，包括使用编码权利要求1、4、7、11和14所述多肽或其功能等同序列的多核苷酸。

32. 可用于检测BigL1、BigL2和BigL3多肽或具有功能等同序列的多肽；bigL1、bigL2和bigL3多核苷酸；或结合BigL1、BigL2、BigL3多肽或具有功能等同序列之多肽的抗体的试剂盒。

33. 可用于检测编码BigL1、BigL2、BigL3或其功能等同序列的核酸的试剂盒，所说的试剂盒包括含一个或多个容器的载体，所述容器中包括可含有与BigL1、BigL2、BigL3核酸或其功能等同序列杂交的多核苷酸的第一容器。

34. 可用于检测抗BigL1、BigL2、BigL3核酸或其功能等同序列的抗体的试剂盒，所说的试剂盒包括含一个或多个容器的载体，所述容器中包括含有BigL1、BigL2、BigL3多肽或其功能等同序列的第一容器。

钩端螺旋体内存在的具有重复性细菌Ig样 (BIG) 结构域的蛋白质

发明领域

本发明涉及编码钩端螺旋体 (*Leptospira* sp.) 细菌中的具有重复性细菌Ig样 (Big) 结构域的蛋白质BigL1、BigL2和BigL3的三种分离DNA分子以及它们在诊断、治疗和疫苗应用中的用途。依照本发明, 编码BigL1、BigL2和BigL3蛋白质的分离分子被用于诊断和预防能在人类和其它哺乳动物体内引起疾病的钩端螺旋体感染, 包括在兽医学方面的用途。

发明背景

螺旋菌是可运动的螺旋状细菌, 包括钩端螺旋体属 (*leptospira*)、疏螺旋体属 (*Borrelia*) 和密螺旋体属 (*Treponema*) 三个属, 它们是人类和其它动物的病原体。疏螺旋体和密螺旋体是包括莱姆病、回归热、梅毒和雅司病在内的疾病的病因。钩端螺旋体由8种致病的遗传相异类型和四种非致病的腐生种类组成 (1, 2)。钩端螺旋体还按血清变型情况进行了分类 - 已鉴定出200个以上的致病血清变型。脂多糖部分的结构异质性看上去似乎是血清变型中观察到的大量抗原变异的基础 (1, 2)。

钩端螺旋体病是一种动物传染性疾病: 通过接触家养的或野生的动物宿主或被它们尿液所污染的环境而传播给人。感染产生了各种类型的临床表现。疾病的早期特征为发热、寒战、头痛和严重的肌痛。在临床感染中有5-15%的病人病情发展至产生诸如黄疸、肾机能不全和出血现象等严重的多系统并发症 (1-4)。严重的钩端螺旋体病死亡率为5-40%。

钩端螺旋体病广泛的分布于世界各地。由于大范围的动物物种可作

为其宿主，所以它被认为是最普遍的人兽互传疾病（1）。钩端螺旋体病传统上是发生在诸如军人、农夫、矿工、污水和垃圾清理工、兽医和屠宰场工人等高危人群中的重要职业病（1-3）。不过，最近已出现了新的疾病传播模式，突出显示了钩端螺旋体病正成为公共健康问题的严重性。在发达国家，钩端螺旋体病已成为消遣性活动（1）和体育运动（1、4、5）相关性疾病发作的原因。在巴西和其它发展中国家，贫困状况已造成了高死亡率钩端螺旋体病在城市内广泛传播（4、5）。

除了其对公众健康的影响外，钩端螺旋体病作为牲畜和家养动物中疾病的成因而成为主要的经济负担（2）。钩端螺旋体病在牛、猪、绵羊、山羊、马和狗等动物中造成了流产、死产、不育、不生长、奶产量降低和死亡，在牲畜中诱发了致病钩端螺旋体的慢性感染和传播（2），并由于目前的国际和国内检疫规定而成为畜牧业经济损失的另一原因。

目前适当诊断工具的缺乏阻碍了对人和动物钩端螺旋体病的控制。标准的血清学检测、镜检凝集检测（MAT）不适于急性病例的鉴别，因为它只能在少数鉴定实验室中进行且需要做配对血清分析以达到足够的灵敏度（1、2）。正如在多项研究中所见到的，对MAT的依赖导致耽误了疾病发作起因的确定（1、2）。已发展了基于全细胞钩端螺旋体抗原制品的酶联免疫吸附检测法（ELISA）和其它快速血清学检测法用作筛选钩端螺旋体传染病的可选方法，尽管仍需要进行MAT以确认病例（1、2）。基于重组抗原的血清学检测法被广泛用于筛查莱姆病和梅毒等螺旋体传染病，但对于利用重组蛋白质进行钩端螺旋体病的血清诊断还未有广泛的研究。最近公开了一种用于牛钩端螺旋体病血清诊断的重组鞭毛抗原免疫捕捉检测法（6）。重组热休克蛋白Hsp58显示了与少数人类病例的血清样品有高度ELISA反应性（7）。不过，将重组抗原用于钩端螺旋体病的血清诊断的应用还未进行广泛的研究加以确认。

此外，目前尚未有可使用的有效干预方法可控制或预防钩端螺旋体病。环境调控的方法难以执行，因为致病钩端螺旋体在土壤和水中长期存活，并且野生及家养动物宿主众多（1、3）。研究方向集中于开发预防性的免疫接种法作为对钩端螺旋体病的干预。目前可供利用的疫苗基

于的是致病钩端螺旋体的灭活全细胞或膜制品，显示出通过诱导抗钩端螺旋体脂多糖的抗体而引起了保护性的反应（1、3）。不过，这些疫苗没有诱发针对感染的长期防护。此外，它们没有提供针对未包括在此疫苗制品中的钩端螺旋体血清变型的交叉保护免疫性。致病血清变型的数目众多（>200）和生产多血清变型疫苗的成本已成为开发以全细胞或膜制品为基础的有效疫苗这一策略的主要局限性。

如同在莱姆病和梅毒等螺旋体疾病中一样，钩端螺旋体病的发病机理依赖于感染早期病原体在宿主内广泛散布的能力（2）。膜相关性钩端螺旋体蛋白质被认为介导了使其能穿过并散布于宿主组织间的相互作用。推测的表面相关毒力因子是可用作疫苗开发的候选分子，诱发对这些因子的反应从而阻断在宿主内的散布。此外，膜相关蛋白质在宿主感染期间可引发免疫反应，因此构成了通过抗体依赖型吞噬作用和补体介导型致死作用等机制而进行免疫保护的靶分子。以重组蛋白质生产这些抗原靶分子是开发亚单位疫苗用于钩端螺旋体病保护性免疫接种的一种低廉方法。此外，选择致病钩端螺旋体中保守的表面相关靶分子可以避免目前获得的全细胞疫苗制品中所遇到的局限性。

钩端螺旋体病领域内的一个主要局限性在于用常规生化和分子方法鉴定表面相关性和宿主表达的蛋白质。从螺旋体布氏疏螺旋体的基因组序列中鉴别出了100种以上的表面相关脂蛋白。从基因组的大小和其生命周期的生物学推断，预计钩端螺旋体具有数目大得多的表面相关性靶分子。目前，通过膜提取物的分离、这些提取物中蛋白质的纯化和定性以及这些蛋白质靶分子的分子克隆，仅鉴定出不到10种表面相关蛋白质（8-14）（12）。用几种已鉴定靶分子的重组蛋白质LipL32、OmpL1和LipL41进行免疫接种，诱发了部分但不是完全的保护性反应（11、12）。

为了对钩端螺旋体蛋白质的表达进行更全面的了解，我们利用人类钩端螺旋体病中的体液免疫应答作为感染期间表达的蛋白质抗原的指示器。感染期间所表达的钩端螺旋体抗原的确定对于开发新的血清诊断和免疫保护方法具有潜在的重要意义。来自钩端螺旋体病患者的血清被用于鉴别来自钩端螺旋体基因组DNA噬菌体文库且表达免疫反应性多肽

的克隆。已发现这些克隆的一部分编码一个新的膜相关钩端螺旋体蛋白质家族。这些多核苷酸和多肽的鉴定以及它们在诊断钩端螺旋体病和诱发对致病螺旋体的免疫应答中的应用是本发明的基础。

发明概述

本发明涉及钩端螺旋体中的DNA分子以及它们所编码的具有重复细菌Ig样结构域的多肽。本发明描述了对最初来源于*L. kirschneri*和问号状钩端螺旋体的三种DNA分子的分离，这三种DNA分子分别编码此处称为“BigL1”、“BigL2”和“BigL3”的蛋白质基于这些多肽的预计氨基酸序列，它们的分子量分别为110、205和205kDa。这三种蛋白质具有12-13个约90个氨基酸长的串联重复序列。来自BigL1、BigL2和BigL3的重复序列相互间高度相关(>90%的氨基酸序列同一性)，且属于细菌病原体毒力因子中发现的细菌Ig样(Big)结构域家族。

编码具有Big结构域、本文称为“bigL1”、“bigL2”和“bigL3”的钩端螺旋体蛋白质的DNA分子可作为异源DNA插入用于生产肽和多肽的表达载体中。重组多肽可从用所说表达载体转化的代用宿主内纯化。BigL1、BigL2和BigL3衍生多肽是活跃感染和感染后的血清学标志，因为来自钩端螺旋体病患者和用致病钩端螺旋体感染或免疫的动物的血清识别分离的多肽。

而且，来自重组或天然抗原制品的BigL1、BigL2和BigL3多肽是有免疫原性的。从用纯化的重组BigL1、BigL2和BigL3多肽所免疫的实验动物中获得的抗体识别来自钩端螺旋体的天然抗原，可用于检测疑似感染的受试者样品中的致病螺旋菌。

此外，BigL1、BigL2和BigL3多肽会诱发抗致病螺旋菌的免疫反应。BigL1、BigL2和BigL3衍生多肽、抗这些多肽的抗体和编码BigL1、BigL2和BigL3的多核苷酸均可单独或与药用可接受载体联合用于治疗或预防钩端螺旋体的感染。由于Big结构域也存在于其它细菌病原体中与毒性相关的蛋白质中，所以上述部分也可用于治疗或预防与钩端螺旋体所引发疾病无关的感染。

在第一个实施方案中，本发明提供了分离的DNA分子bigL1、bigL2和bigL3以及由这些DNA分子编码或具有功能等同序列的多肽。此外，还提供了产生含有bigL1、bigL2和bigL3多核苷酸的表达载体和获得来源于这些序列的基本上纯化的多肽的方法。

本发明的第二个实施方案提供了用于在受试者中诱发针对致病螺旋菌的免疫反应的药物组合物，其中包括在药用可接受赋形剂中、选自BigL1、BigL2、BigL3和具有功能等同序列的多肽中的一种或多种免疫有效量的抗原。

在第三个实施方案中，本发明提供了用于鉴定结合BigL1、BigL2、BigL3多肽或含功能等同序列的多肽的化合物的方法，包括将含此化合物和BigL1、BigL2、BigL3多肽或具有功能等同序列的多肽在内的组分一起保温，保温条件要求足以允许这些组分相互作用，并检测所说的化合物与BigL1、BigL2、BigL3多肽或含功能等同序列的多肽的结合。优选的，本发明方法是利用来自有钩端螺旋体或其它相关细菌病原体活跃感染或感染后的受试者的血清进行血清诊断的方法。

第四个实施方案中，本发明提供了检测样品中病原体的方法，包括将怀疑含致病螺旋菌的样品与结合该病原体的特异性细胞组分的试剂接触，并检测所述试剂与所说组分的结合。在一方面，结合病原体特异性细胞组分的试剂是用于鉴定bigL1、bigL2和bigL3多核苷酸的寡核苷酸。在另一方面，结合病原体特异性细胞组分的试剂是抗BigL1、BigL2、BigL3多肽或其功能等同序列的多肽的抗体。

在第五个实施方案中，本发明提供了用于检测BigL1、BigL2、BigL3多肽或其功能等同序列的多肽；bigL1、bigL2和bigL3多核苷酸；或与BigL1、BigL2、BigL3多肽或其功能等同序列的多肽结合的抗体的试剂盒。

本发明的其它目的、特征和优势可以从以下详述中显而易见。不过，应当理解为，在简述本发明的优选实施方案时，详述和具体实施例都只是说明的方式，因为对于本领域技术熟练人员来说显而易见可以从此详述中在本发明的精神和范围内对其进行各种改变和修饰。

附图简述

图1A和B显示了钩端螺旋体中bigL基因序列Southern印迹分析结果。用NsiI消化来自问号状钩端螺旋体菌株Fiocruz L1-130(泳道1)、L. kirschneri菌株Rm52(泳道2)和L. biflexi菌株Patoc I(泳道3)的基因组DNA(3mcg/泳道)并进行琼脂糖凝胶电泳。转移至硝酸纤维素膜上后,分别用编码BigL重复结构域(BigL3的第四至第六个重复结构域,图1A)的DNA片段和bigL1、bigL2和bigL3的C末端区域探测印迹(图1B),所说的C末端区对于这些基因各自来说是独一无二的。

图2显示了编码L. kirschneri中BigL1和BigL3蛋白质的区域的基因组结构示意图。BigL1蛋白质会包含一信号肽(阴影框)和13个90个氨基酸长的细菌免疫球蛋白样结构域(实心框)。BigL3蛋白质包含一信号肽、12个90个氨基酸长的细菌免疫球蛋白样结构域和793个氨基酸的羧端(C-末端)结构域。显示了2156bp区域中有100%DNA序列同一性的位置。所述区域的构成在问号状钩端螺旋体和L. kirschneri中是保守的。

图3显示了来自钩端螺旋体5个致病种菌株的DNA片段的聚合酶链式反应(PCR)扩增。简并引物是基于编码相应于第46-65位氨基酸的BigL3区域的L. kirschneri和问号状钩端螺旋体序列设计的。PCR反应是用来自5种致病菌(L. kirschneri、博氏钩端螺旋体(L. borg petersenii)、问号状钩端螺旋体(L. interrogans)、圣地罗斯钩端螺旋体(L. santarosai)和野口氏钩端螺旋体(L. noguchi))和两种非致病菌(双曲钩端螺旋体(L. biflexi)和沃尔氏钩端螺旋体(L. wolbachii))的纯化DNA进行的。

图4显示了用bigL1、bigL2和bigL3特异性引物对L. kirschneri的RNA提取物进行RT-PCR的扩增产物。用培养钩端螺旋体的RNA提取物进行逆转录反应(泳道“+”),然后用与bigL1、bigL2和bigL3中独特序列结合的引物进行聚合酶链式反应(PCR)扩增步骤。用基于lipL45内序列的引物进行扩增,作为对照反应,同样对未经过逆转录步骤的样

品进行的PCR反应作为对照。

图5显示了从感染了致病钩端螺旋体的病人和动物宿主以及用全问号状钩端螺旋体抗原制品免疫的实验室动物中收集的血清合并物对重组BigL3蛋白质 (rBigL3) 的免疫印迹反应性。用纯化的rBigL3 (1mcg/泳道, 泳道3) 进行蛋白质印迹分析。用来自钩端螺旋体病患者的血清 (泳道A)、健康个体的血清 (泳道B)、定居了问号状钩端螺旋体的被俘获大鼠 (泳道C)、未定居问号状钩端螺旋体的被俘获大鼠 (泳道D)、用体外培养问号状钩端螺旋体全抗原制品免疫的实验室大鼠 (泳道E) 和免疫接种前收集的实验室大鼠免疫前血清 (泳道F) 探测膜。对问号状钩端螺旋体全抗原制品 (泳道1) 和重组LipL32蛋白 (rLipL32, 泳道2) 的反应性均列出作为对照。左边的数字显示了参照分子量标准 (Invitrogen) 的位置和相对迁移率 (kDa)。

图6显示了在钩端螺旋体病急性发作阶段 (泳道A) 和康复阶段 (泳道B) 受试者病人对rBigL3的血清反应性的ELISA测定结果。将来自4个钩端螺旋体病患者 (实线) 和4个健康个体 (虚线)、稀释度为1:50、1:100和1:200的血清与RBigL3 (25 - 200ng/孔) 一起保温。分别用缀合了辣根过氧化物酶的Mu和 γ 链特异抗体测定IgM和IgG血清反应性。图中显示了平均吸收值 (OD450nm) 和标准偏差。

图7显示了来自29个钩端螺旋体病患者在病情急性发作阶段 (泳道2) 和康复阶段 (泳道3) 的血清以及28个健康个体的血清 (泳道1) 的rBigL3 IgM (A栏) 和IgG (B栏) 反应性。用血清 (1:50稀释度) 和缀合了辣根过氧化物酶的Mu和 γ 链特异抗体测定反应性。实心柱代表平均吸收值 (OD450nm)。

图8显示了钩端螺旋体病患者个体在病情急性发作 (泳道6-9) 和康复阶段 (泳道10-13) 对rBigL3的免疫印迹反应性。用纯化的rBigL3 (1mcg/泳道, 泳道3) 进行蛋白质印迹分析。用1:100稀释的血清探测膜。用缀合了碱性磷酸酶的 γ 链特异抗体测定对于BigL3区域1 (第2至第6个Big重复结构域) 的58kD重组蛋白的反应性。测定对于rLipL32 (1mcg/泳道) 的反应性作为对照。分别用Ponceau-S和考马斯蓝染色

后显示纯化rBigL32和rLipL32（泳道14）以及分子量标准（泳道15）的迁移率。

图9显示了大鼠抗rBigL3抗血清对rBigL3和来自问号状钩端螺旋体裂解产物的天然抗原的免疫印迹反应性。用纯化的rBigL3（1mg/泳道；泳道3、5、7、9）和来自人工培养的钩端螺旋体的全抗原制品（ 10^8 个钩端螺旋体/泳道；泳道2、4、6和8）进行免疫印迹。用从表达问号状钩端螺旋体 bigL3克隆DNA片段的大肠杆菌中得到的rBigL3免疫大鼠后收集的合并血清（稀释度1:500[泳道4和5]、1:100[泳道6和7]和1:2500[泳道8和9]）探测膜。免疫前血清是第一次免疫前获得的，用于免疫印迹分析中作为对照（泳道2和3）。分子量标准的迁移率（kDa）如图左侧所显示。

图10显示了兔抗rBigL3抗血清对来自钩端螺旋体菌株裂解产物的天然抗原的免疫印迹反应性。用以下培养菌株的全抗原制品（ 10^8 个钩端螺旋体/泳道）进行免疫印迹测定：泳道1，问号状钩端螺旋体波莫纳血清变型（kennewicki型）菌株RM211，低传代；泳道2，问号状钩端螺旋体犬血清变型菌株CDC Nic1808，低传代；泳道3，问号状钩端螺旋体波莫纳血清变型菌株PO-01，高传代；泳道4，问号状钩端螺旋体布拉迪斯拉发血清变型菌株AS-05，高传代；泳道5，L.kirschneri流感伤寒血清变型菌株RM52，低传代；泳道6，L.kirschneri流感伤寒血清变型菌株P8827-2，低传代；泳道7，L.kirschneri流感伤寒血清变型菌株86-89，低传代；泳道8，L.kirschneri流感伤寒血清变型菌株Moskva V，高传代；泳道9，L.kirschneri sv mozdok菌株5621，高传代；泳道10，L.kirschneri流感伤寒血清变型菌株RM52，高传代。利用从用表达L.kirschneri bigL3克隆DNA片段的大肠杆菌所表达的rBigL3免疫的兔子得到的血清探测膜，用从重组L.kirschneri GroEL蛋白免疫的兔子得到的血清作为对照检测样品。相应于BigL3和GroEL的天然抗原的位置和分子量标准的迁移率（kDa）分别如图左侧和右侧所示。

发明详述

为了方便理解，将说明书、实施例和附属权利要求书中所用的某些术语和词组的含义解释如下。除非另外说明，本文所用的所有技术和科学术语具有的含义与本发明所属领域内普通技术人员通常理解的含义相同。

BigL - 是具有串联重复序列的钩端螺旋体菌多肽，由于它们的序列同源性，每个所述重复序列都与细菌免疫球蛋白样 (Big) 结构域类似。Big结构域存在于细菌蛋白质中，并在诸如大肠杆菌、耶尔森氏菌属 (*Yersinia*) 和博德特氏菌 (*Bordetella*) 等细菌病原体中表达，具有宿主细胞粘附等毒性功能。

参照序列 - 是通过从天然生物体中分离或通过遗传工程获得的新序列，具有准确的生物学功能，所说的功能是本发明的特征。

功能等同序列 - 是与参照序列相关，并且由于变异性，即序列中所有的自发或诱发的修饰，包括核苷酸或氨基酸的取代和/或删除和/或插入、和/或在某一端序列的延伸和/或缩短所得到的、但未导致参照序列的特征性功能发生改变的序列。功能等同序列包括其片段和类似物。换言之，功能等同的序列是与参照序列“基本上相同”或“基本上同一”的序列，如与参照的氨基酸或核酸序列具有至少80%同源性的多肽或核酸序列。通常用进行序列分析的软件系统（遗传学计算机组序列分析软件包，Wisconsin大学生物工程中心，1710, University Avenue, Madison, Wis, 53705）测定同源性。

正如我们前文中所提及的，宿主感染期间表达的钩端螺旋体抗原在诊断测试的靶目标和疫苗的鉴定中是重要的。LipL32蛋白是这些靶目标中的一个，它是通过自然感染期间的免疫体液反应而鉴定为免疫显性抗原的。不过，基于检测钩端螺旋体病急性发作阶段患者血清中抗LipL32抗体的血清学测试的灵敏度是有限的（参阅，Flannery, B: “用于钩端螺旋体病血清诊断且以重组钩端螺旋体抗原为基础的酶联免疫吸附检测的评估 (Evaluation of recombinant *Leptospira* antigen-based Enzyme-linked Immunosorbent Assays for the serodiagnosis of Leptospirosis) J. Clin. Microbiology 2001; 39 (9) : 3303 - 3310;

W09942478)。

本发明基于的是与螺旋体菌种属、包括属于钩端螺旋体的那些种类相关的蛋白质BigL家族的鉴定。

依照本发明，BigL蛋白质家族被认为是宿主体液免疫反应的靶目标，它是在致病钩端螺旋体感染期间或用致病钩端螺旋体或重组BigL多肽免疫期间产生的。BigL多肽和编码这些多肽的多核苷酸可在诊断测试中用于鉴定包括人和动物宿主在内的不同物种中天然存在的感染。与标准诊断测试或已发表的文献中所用的方法相比，基于那些蛋白质的诊断测试具有更高的灵敏度和特异性。在初始阶段鉴定钩端螺旋体病。此外，在用于免疫接种的药物组合物中使用，BigL多肽可以诱发免疫反应。

在本发明中，三种BigL多肽的分子量是128.4kD、201.3kD和200.4kD，这依据的是根据编码这些多肽的分离多核苷酸bigL1、bigL2和bigL3所推断的氨基酸序列。BigL多肽的氨基酸序列具有信号序列和基本上符合螺旋体脂盒(lipbox)的推断的信号肽酶切割位点；因此BigL多肽是膜相关的脂蛋白。128.4kD、201.3kD和200.4kD的多肽分别被称为“BigL1”、“BigL2”和“BigL3”。

尽管本发明的BigL多肽最初分离自钩端螺旋体种属，但它们并非只可用于诱发抗致病生物钩端螺旋体的免疫反应，而是还可以诱发抗其它螺旋体菌和具有含Big结构域的因子的病原体的免疫反应。此外，BigL多肽可用于诊断由钩端螺旋体、其它致病螺旋菌和细菌病原体造成的感染。

结合了现有技术方法学的几种方法可用于获得编码BigL多肽的多核苷酸序列。这些方法包括但不局限于用基因组文库与探针的杂交探测同源核苷酸序列进行的DNA分离；筛选抗体表达文库来检测具有共同结构特征的克隆DNA片段；用能使目的DNA序列重组的起始子进行基因组DNA内的聚合酶链式反应(PCR)；以及以计算机为基础在序列数据库中搜索与bigL多核苷酸相似的序列。

本发明中抗原的鉴定是以对培养期间(体外)和宿主感染期间(体内)钩端螺旋体抗原差异表达的了解为基础的。钩端螺旋体抗原的差异

表达被认为在感染期间宿主适应性的改变中是重要的。我们使用一种策略来鉴定免疫反应性抗原以及相应地在宿主感染期间表达的抗原。致病钩端螺旋体感染患者的血清被用于从λ噬菌体内的基因组钩端螺旋体DNA文库中选择出编码免疫反应性多肽的多核苷酸序列。

本发明鉴定和分离了具有相应于SEQ ID No: 1、SEQ ID No: 2和SEQ ID No: 3的核苷酸序列的三种多核苷酸，以及这些多核苷酸编码的相应多肽BigL1、BigL2和BigL3的氨基酸序列。

步骤1 - 阳性克隆的筛选主要由以下步骤组成：

(a) 用适当的酶切割致病钩端螺旋体的DNA并连接入λ噬菌体基因组中的特异位点。用噬菌体感染宿主细菌，对产生的在IPTG诱导后表达重组多肽的克隆进行免疫印迹分析，其中将菌落溶解产物印迹膜与实验室证实患钩端螺旋体病的患者的血清一起保温，然后与缀合了辣根过氧化物酶的、可识别人Ig的二抗保温。通过指示剂反应探测阳性克隆，即以颜色的产生为基础检测抗原-抗体复合物。

(b) 用噬菌体载体特异序列作为测序反应的起始物测定克隆并分离的多核苷酸的序列。克隆序列的分析和引物步行策略的应用鉴定出了编码BigL1、BigL2和BigL3的基因的全部核苷酸序列。

(c) 大部分获得的阳性克隆包含编码热休克蛋白Hsp58和DanK以及外膜蛋白LipL41的基因，不过，根据与蛋白质家族(Pfam)数据库的比较，发现有编码含90个氨基酸的串联重复序列重复区的基因的克隆，所述重复区类似细菌型免疫球蛋白(Big)。通过克隆序列的分析，鉴定出三种基因，其中bigL1含12个串联重复片段，而bigL2和bigL3中含13个串联重复片段。

步骤2 - 蛋白质的亚克隆表达和纯化

- 根据两种 BigL 蛋白质序列确定两条寡核苷酸
- 从那些寡核苷酸 PCR 扩增编码重复区域部分的原始 BigL 部分
- 对扩增产物进行测序
- 对测序产物的编码区进行亚克隆
- 重组蛋白质的表达

- 重组蛋白质的纯化

免疫印迹分析证实了来自钩端螺旋体病患者和致病钩端螺旋体感染了啮齿动物宿主的血清产生主要针对BigL多肽BigL结构域重复片段的抗体，表明它们是感染期间所识别的主要抗原区。

相对于本发明的多肽而言，它们包含DNA、cDNA或RNA序列（和其互补核酸序列）以及它们的功能等同序列，即，编码完整或部分BigL1、BigL2和BigL3多肽、但由于变应性而并非完全相同的那些序列。

本发明的多肽和多核苷酸由BigL1、BigL2和BigL3以及编码这些多肽的多核苷酸组成；不过它们还另外包括具有功能等同序列的多肽和多核苷酸。

在本发明中，多核苷酸和多肽都可以是天然、合成或重组来源的，且具有保证它们生物学活性所必需的纯度。

本发明还涉及编码BigL1、BigL2和BigL3的多核苷酸，它们用于PCR反应中以获得bigL多核苷酸的完整或部分扩增DNA片段，其目的是为了检测样品中的钩端螺旋体或表达重组BigL多肽。在将起始子用于本发明中多核苷酸扩增的情况下，它们是由两个或多个脱氧核糖核苷酸或核糖核苷酸制成的天然或合成的寡核苷酸。

优选将各个起始子构建成与待扩增靶序列链的侧翼区基本上相似。在此意义上，如果相应的聚合物对于所考虑应用或用途来说能产生相同（并非完全一样）的作用，则该起始子就可称为其功能等同物。

本发明的多核苷酸序列还可插入诸如质粒、病毒或用于重组克隆的其它载体等表达载体中，所说的载体通过插入或掺入编码BigL1、BigL2和BigL3的完整或部分核苷酸序列或其功能等同序列而得以利用。这样的表达载体含有方便载体所插入的宿主中由遗传序列有效转录的启动子序列。所说的宿主可以包括原核或真核生物，包括酵母等微生物或昆虫和哺乳动物。这些用于表达载体构建和重组序列表达的方法，严格的说，是本技术领域内专家众所周知的。

本发明提供了产生与BigL1、BigL2和BigL3完整或部分多肽或其功能等同序列可结合的抗体的方法。这样的抗体可在一般螺旋菌感染的有

关研究和诊断中用作研究和诊断工具，更具体而言是用于开发钩端螺旋体病的治疗或预防性的诊断或治疗方法。这些抗体可以单独施用或作为药物组合物的一部分施用，所说的药物组合物利用这些抗体和制药学上可接受的载体作为抗螺旋体疗法的一部分。

本发明涉及利用BigL多肽或编码这些多肽的多核苷酸的药物组合物作为疫苗，或者作为预防疾病的疫苗，用于诱发对于致病螺旋菌感染或集落化的免疫保护性反应，或者作为提供有益效果的治疗疫苗，用于缩短致病螺旋菌感染患者临床病情发作的持续时间或严重性，或减弱诸如猪、牛、大鼠或狗等致病螺旋菌载体长期含有或排泄致病螺旋菌的宿主状态。所说的组合物可以用赋形剂和添加剂或辅助剂中的免疫有效量的分别抗BigL1、BigL2和BigL3的抗体制备，或用分离自钩端螺旋体病原体的BigL1、BigL2和BigL3中的一种或多种或重组BigL多肽或其功能等同序列制备。

本发明另一实施方案涉及用于在受试者中诱发对于致病螺旋菌，尤其是钩端螺旋体的免疫反应的药物组合物，包括在药用可接受载体中的免疫学有效量的BigL1、BigL2和BigL3或其功能等同序列。对于“受试者”一词，我们是指任何哺乳动物，包括人、啮齿动物、家养和实验室动物和牲畜。对于“免疫学有效量”一词，我们是指在受试者中诱发抗钩端螺旋体或任何其它致病螺旋菌或细菌病原体所必需的BigL多肽抗原的量。本发明进一步提供了试剂盒，用于：

1. - 检测BigL1、BigL2和BigL3多肽或其功能等同序列之一；
2. - 检测编码BigL1、BigL2和BigL3或其功能等同序列的核酸；
3. - 检测针对这些BigL1、BigL2和BigL3多肽或其功能等同序列的抗体。

用于检测BigL多肽的试剂盒包括使用含有一个或多个接收器的载体的类型，其中第一接收器包含与BigL1、BigL2和BigL3或其功能等同序列连接的试剂。

用于检测编码BigL多肽的多核苷酸的试剂盒包括使用含有一个或多个接收器的载体的类型，其中第一接收器包含与编码BigL1、BigL2和

BigL3或其功能等同序列的核酸序列杂交的多核苷酸。

用于检测抗BigL多肽的抗体的试剂盒包括使用含有一个或多个接收器的载体的类型，第一接收器包含BigL1、BigL2和BigL3多肽或其功能等同序列多肽。

现在将用实施例对本发明进行说明，它们不应被认为是对本发明的限制。

实施例 1:

实施例 1A: 细菌菌株、质粒和培养基

Leptospira kirschneri 流感伤寒血清变型菌株RM52是1983年在猪流产发作期间分离的。问号状钩端螺旋体 *copenhageni* 血清变型菌株 Fiocruz (L1-130) 分离自人钩端螺旋体病患者的血流。*L. kirschneri* 流感伤寒血清变型菌株RM52和其它钩端螺旋体菌株获自国家钩端螺旋体病鉴定中心(国家动物疾病中心, 农业研究局, 美国农业部, Ames, Iowa)。钩端螺旋体菌株在30℃培养于Johnson-Harris牛血清白蛋白-Tween80培养基(Bovuminar PLM-5微生物培养基, Intergen)(2)。将RM52分离物的低代数样本或者保存于液氮中, 或者在液体培养基中传代至少200次以得到高代数形式的菌株。高代数菌株不管任何剂量都不能在仓鼠内产生致命的感染, 只能以 10^7 的剂量通过腹膜内接种感染仓鼠。

大肠杆菌XL1-Blue MRF' #mcrA(183#mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac [F' proAB lacI^qZ #M15 Tn10 (Tetr)] (Stratagene) 和大肠杆菌PLK-F' (endA1 gyrA96 hsdR17 lac⁻ recA1 relA1 supE44 thi-1 [F' lacI^qZ #M15]) 被用作分别以λ ZapII (Stratagene) 和λ TriplEx (Clontech) 载体感染的宿主细胞。大肠杆菌SOLR (e14⁻ [mcrA], #[mcrCB-hsdSMR-mrr]171 sbcC recB recJ umuC: : Tn5 [Kan^r] uvrC lac gyrA96 relA1 thi-1 endA1 λ', [F' proAB lacI^qZ #M15], Su⁻ [非抑制性的]) 和大肠杆菌BM25.8 (supE44 thi # lac-proAB [F' traD36 proAB⁺ lacI^qZ #M15] λ imm434 (Kan^r) P1 (cam^r)

hsdR ($r^{K12^-} m^{K12^-}$) 分别用于pBluescript和pTriplEx噬菌粒的体内切割。BLR (DE3) pLysS [F⁻ompT hsdS_B (r_B - m_B -) gal dcm-(srl-recA) 306::Tn10 (TcR) (DE3) pLysS (CmR)] (Novagen) 被用作pRSET表达载体 (Invitrogen) 的宿主菌株。将大肠杆菌菌株培养于补充了100 μg/ml 氯苄青霉素、100 μg/ml羧苄青霉素或25 μg/ml氯霉素的LB中, 根据需 要选择合适 的抗生素。抗生素购自Sigma。

实施例 1B: bigL 基因 的分离和定性:

此 实施例描述了bigL基因的鉴定和分离。用Yelton和Charon的方法 (15) 从毒性、低传代L. kirschneri流感伤寒血清变型菌株RM52制备基因组DNA。用基因组DNA专用试剂盒 (Qiagen) 从问号状钩端螺旋体血清变型copenhageni菌株Fiocruz L1 - 130的临床分离物中制备基因组DNA。用QIAquick PCR纯化试剂盒 (Qiagen) 获取纯化的DNA。用Tsp509I 部分消化基因组DNA, 并按照所提供的说明书 (Clontech) 连接入λ TriplEx臂中。用Gigapack III Gold Packaging Extract (Stratagene) 试剂盒将已消化并连接的基因组DNA包装入λ噬菌体的头部。通过感染大肠杆菌XL1 - Blue测定文库的噬菌体滴度。

为了筛选基因组文库, 将约 10^3 pfu滴度的噬菌体铺于大肠杆菌XL1 - Blue菌苔上, 然后转移到硝酸纤维素膜上 (Schleicher&Schuell), 用IPTG致敏并按文献所推荐的方法处理 (Schleicher&Schuell)。用配制于含0.05% Tween - 20的Tris - 缓冲盐水 (pH7.8) (TBST) 或含0.05% Tween - 20的磷酸缓冲盐水 (pH7.4) (PBST) 中的5%脱脂奶封闭硝酸纤维素滤膜, 与来自经实验室证实患钩端螺旋体病的患者并1: 50稀释的合并血清保温1小时。自1996年和1999年之间巴西城市流行病中所鉴定的患者在康复阶段收集血清, 并在使用前与大肠杆菌溶解产物一起预吸附以去除抗大肠杆菌的抗体。用TBST或PBST洗膜三次, 然后与缀合了碱性磷酸酶 (Sigma) 且1: 1000稀释的兔或山羊抗人免疫球蛋白抗体保温1小时以上。用NBT (0.3mg/ml) 和BCIP (0.15mg/ml) 检测或用ECL蛋白质印迹检测试剂 (Amersham) 呈色并对Hyperfilm (Amersham) 曝光, 由此鉴定含抗原 - 抗体复合物的噬菌斑。

将各阳性噬菌斑于4℃保存于1ml SM (0.1M NaCl, 8mM MgSO₄, 50mM Tris-HCl pH7.5; 0.01%明胶, 加1-2滴氯仿) 中。对与合并血清反应的λ噬菌斑克隆再另外进行两步纯化。如供应商(分别为Stratagene和Clontech)所述通过用λ克隆感染大肠杆菌SOLR或BM25.8菌株来切下插入λ噬菌体的基因组DNA片段。

用连接到插入片段相邻区的载体特异引物获得了插入片段的头500-700个核苷酸序列。对于131个克隆的核苷酸序列分析鉴定出其中13个具有编码约90个氨基酸长的串联重复区的DNA插入片段。随后在Pfam6.6 (<http://pfam.wustl.edu/>) 中鉴定出各重复序列属于细菌免疫球蛋白样(Big)结构域的Big2家族。

为了鉴定符合预期氨基酸序列的全长蛋白质的编码序列, 将通过引物步行法和嵌套缺失体测序的组合获得的受试者序列装配成克隆的核苷酸序列。通过去除从插入片段内部延伸至插入片段侧翼的多克隆位点内的限制性片段, 由质粒克隆得到缺失体。寡核苷酸合成并获自GIBCO BRL或Operon。进行反向PCR(iPCR), 以获得含剩余基因和侧翼DNA的序列。UCLA核心测序研究室、耶鲁/Keck核心DNA测序研究室和加州大学伯克利分校测序研究室均进行此测序反应。

已发现两种*L. kirschneri*克隆和四种问号状钩端螺旋体克隆编码我们称为细菌免疫球蛋白样钩端螺旋体蛋白1即bigL1的基因。*L. kirschneri* bigL1的完整核苷酸序列和基因产物的预期氨基酸序列如SEQ ID NO: 1和SEQ ID NO: 2中所示。发现6种*L. kirschneri*克隆编码我们称为bigL2的第二种基因。*L. kirschneri* bigL2的完整核苷酸序列如SEQ ID NO: 3中所示。*L. kirschneri* bigL2看来似乎是假基因, 在第1011位核苷酸处有一个额外的腺嘌呤残基, 导致了移码突变和下游的TAG终止密码子。不过, 用合并的病人血清进行抗体筛选能鉴定到含有编码bigL2基因产物的DNA片段的λ克隆, 大概是克隆片段没有移码突变并以允许表达可被病人血清识别的产物的方向插入所致没有移码突变的*L. kirschneri* bigL2基因产物的预计氨基酸序列如SEQ ID NO: 4所示。发现第五个问号状钩端螺旋体克隆编码最初被认为属于BigL1的几个

Big重复片段。不过发现此第五个问号状钩端螺旋体克隆的上游编码DNA与bigL1的上游序列不同。对bigL1基因侧翼区的测序显示此第五个问号状钩端螺旋体克隆相应于bigL1下游的第三种基因,被称为bigL3(图2)。自*L. kirschneri* DNA获得bigL3的完整核苷酸序列并如SEQ ID NO: 5中所示。*L. kirschneri* bigL3基因产物的预计氨基酸序列如SEQ ID NO: 6中所示。

正如前文所述,所有三种bigL基因均编码信号肽和基本上符合螺旋体脂盒(lipobox)的推定的信号肽酶切割位点(Haake, D. A. 2000年,螺旋体的脂蛋白以及发病机理 *Spirochetal lipoproteins and pathogenesis. Microbiology. 146: 1491 - 1504*)。已知螺旋体脂蛋白的序列比较显示对螺旋体脂盒的限制比大肠杆菌脂盒要宽松的多。例如,大部分大肠杆菌脂蛋白在相对于半胱氨酸的-3位置是亮氨酸,而螺旋体的脂蛋白在此位置却还可以具有许多其它的疏水氨基酸,包括缬氨酸、苯丙氨酸和异亮氨酸。包括对半胱氨酸之后的氨基酸进行定位诱变的大肠杆菌实验显示酸性残基引起了脂蛋白分类至细胞质膜上。钩端螺旋体脂蛋白的序列分析表明相似的分类信号存在于这些细菌中。例如,LipL31是唯一一种在半胱氨酸之后的头两个氨基酸中具有非对立性负电荷的脂蛋白,也是唯一一种被专门分拣至细胞质膜的脂蛋白。象外膜脂蛋白LipL32和LipL41一样,BigL蛋白在+2和+3位置具有不带电荷的氨基酸,表明它们会被分拣至外膜上。

在它们的信号肽之后,所有三种蛋白质都包含一系列约90个氨基酸长的串联重复片段。成熟的BigL1蛋白质几乎完全由13个重复片段组成,而与之大不相同的是,BigL2和BigL3包含12个重复片段及随后的大羧基端结构域。尽管在三种蛋白质中发现的31种独一无二的重复片段具有高度序列差异,所有这些重复片段都通过Pfam数据库的比对被鉴定为细菌免疫球蛋白样Big蛋白质家族,其E-值低至 4×10^{-30} 。

bigL1、bigL2和bigL3的问号状钩端螺旋体和*L. kirschneri*形式高度相关,DNA和氨基酸序列同一性大于90%。在两个物种中都具有包括bigL1和bigL3 5'末端在内的DNA序列同一区(图2)。在两种基因中,

序列同一区都从ATG起始密码子开始延伸至1890bp的位置。bigL1和bigL3之间大段的DNA序列同一区导致了BigL1 (SEQ ID NO: 2) 和BigL3 (SEQ ID NO: 6) 头630个氨基酸 (第1-630位) 的氨基酸序列完全相同。此相同区域相应于头6个BigL结构域重复片段。

实施例 2

实施例 2A: bigL 基因的定性和 bigL DNA 和 RNA 的检测

此实施例描述了钩端螺旋体属各种中多拷贝bigL基因的分布和用于检测样品中bigL DNA和RNA的方法。

DNA 印迹分析

进行DNA印迹分析以鉴定问号状钩端螺旋体菌株Fiocruz L1-130、L. kirschneri菌株RM52和双曲钩端螺旋体菌株Patoc I基因组DNA中的多拷贝bigL基因。DNA限制和修饰酶购自New England Biolabs。用血液和细胞培养物试剂盒 (Qiagen, Valencia, Calif.) 从培养7天的500ml钩端螺旋体细胞培养物提取基因组DNA。用5-20个单位的NsiI消化约3mcg DNA过夜, 终体积50mc1。然后用酚: 氯仿: 异戊醇纯化DNA并用100%冷乙醇和3M乙酸钠沉淀, 用70%乙醇洗涤。然后用5-20个单位的PacI再消化经纯化的DNA过夜, 终体积为25mc1。在0.8%琼脂糖凝胶中20V分离双消化的DNA过夜。然后将此胶在变性缓冲液 (1.5M NaCl, 0.5N NaOH) 中保温2次30分钟, 在中和缓冲液 (1M Tris (pH7.4) 1.5M NaCl) 中保温两次30分钟。按照Southern所述的方法将基因组DNA转移至带正电的尼龙膜 (Roche Molecular Biochemicals, Indianapolis, Ind.) 上。

用PCR Dig探针合成试剂盒 (Roche, Mannheim, 德国) 合成探针。按制造商说明书在终体积50mc1中进行反应。扩增的温度循环是94℃5分钟, 94℃30秒、57℃30秒和72℃1分钟进行35个循环, 最终延伸时间为7分钟。探针序列如下: 为了扩增编码BigL重复结构域的bigL DNA片段, 选择相应于BigL3 重复结构域4-6编码区的bigL3 DNA序列, BigL3-395 gat-ttt-aaa-gtt-aca-caa-gc 和 BigL3-573

aaa-ccg-gac-tac-tta-cct-ttc-c; 而为了扩增特异于各bigL基因的bigL DNA片段, 选择编码BigL基因产物C-末端区的序列: BigL1.2078p, tta-cgg-cta-cag-gta-ttt-tta-cg 和 BigL1.2691p att-gga-aga-ttt-cca-agt-aac-c, BigL2.5121p tat-cta-cgc-tgc-aaa-tgg 和 BigL2.5865p ttg-ttg-gcg-ata-cgt-ccg, BigL3.5071p cat-aac-tct-cct-cat-aac-a和BigL3.5548p tat-gta-gag-ata-aga-tcc。

将UV交联膜在Dig简易杂交溶液 (Roche) 中42℃预杂交1小时。在杂交前, 将Dig标记探针煮沸10分钟并迅速冰置5分钟。将变性探针与杂交溶液混合并与膜一起在42℃保温过夜。杂交后, 在室温用2xSSC(NaCl, 柠檬酸钠)、0.1%SDS洗膜两次5分钟。然后在42℃用0.1xSSC、0.1%SDS洗膜两次30分钟。将膜对Biomax ML膜 (Eastman Kodak, Rochester, 纽约) 曝光1-3分钟, 以检测化学发光产物。

图1A和B显示了Southern印迹的结果。相应于编码BigL重复片段的DNA序列的探针与*L. kirschneri*和问号状钩端螺旋体中的多个DNA片段杂交 (图1A)。相反, 用来自非致病双曲钩端螺旋体的经消化基因组DNA则未鉴定到杂交。基于问号状钩端螺旋体各bigL基因产物特异C-末端区编码序列的探针与已消化的问号状钩端螺旋体基因组DNA中的一个独特片段杂交, 因此证实了存在各一拷贝的实施例1中所鉴定的三种bigL基因 (图1B)。这些结果阐明了以检测非致病钩端螺旋体中未发现的DNA片段为基础特异鉴定致病钩端螺旋体的方法。

实施例 2B: 钩端螺旋体基因组 DNA 中 bigL 基因序列的 PCR 检测

此实施例说明了bigL基因在致病钩端螺旋体中的分布。为了检测在其它钩端螺旋体属种中的bigL基因, 以实施例1中所鉴定的来自*L. kirschneri*菌株RM52和问号状钩端螺旋体菌株Fiocruz L1-130的bigL基因的序列对比为基础设计简并引物。被称为BigL-1up的“上游”引物的序列是相应于bigL1和bigL3 (SEQ ID NO: 1和5) 中相对于起始密码子的A来说为第46-65位的5'- (GC) AAAGTTG (TC) (AG) (TC) G (TG) CTTGGCC-3'。被称为BigL-2dn的“下游”引物的序列是相应于bigL1和bigL3 (SEQ ID NO: 1和5) 中相对于

起始密码子的A来说为第506 - 487位的5' - (GC) (AT) ACC (AG) TC (CT) GAAAA (AG) AT (AT) CC-3'。各引物长为20个核苷酸。将这些引物设计成与bigL2中相对于bigL2起始密码子的A来说为第97 - 116和590 - 571位的位点 (SEQ ID NO: 3) 退火。

用来自高传代和低传代钩端螺旋体菌株的纯基因组DNA进行PCR反应。图3中, 经PCR反应鉴定了扩增的DNA片段, 并对所有4种致病种属菌株的基因组DNA进行了评估。片段具有基于bigL1/bigL3(461bp)和bigL2(494bp)序列预计的电泳迁移率。在被评估的两种非致病钩端螺旋体种中未鉴定到扩增的DNA片段。因此此实施例举例说明了此PCR方法在特异性鉴定样品内来自致病钩端螺旋体的DNA中的应用。

实施例 2C: 钩端螺旋体 bigL RNA 的逆转录酶 - 聚合酶链式反应 (RT - PCR) 检测法

此实施例举例说明了样品中bigL RNA的检测法。将*L. kirschneri* 菌株RM52培养至晚指数生长期, 用热 - 酚方法从 1×10^{10} 个钩端螺旋体细胞中提取总RNA, 酒精沉淀后重悬浮于水中 (ref)。用6个单位的DNase I (Ambion) 在 $70 \mu\text{l}$ DNase I缓冲液 (10mM Tris - HCl pH7.5, 25mM MgCl_2 , 1mM CaCl_2 , 在来自Ambion的1xRNA保护液中) 中 37°C 消化约 $2 \mu\text{g}$ 钩端螺旋体RNA 30分钟。为了灭活DNase I, 加入25mM EDTA $1.75 \mu\text{l}$ 终止反应, 在 70°C 加热5分钟使酶热失活。如说明书所述 (Qiagen) 用约200ng钩端螺旋体RNA和Omniscrypt逆转录酶进行RT - PCR。以下引物用于引发逆转录酶反应:

bigL1, 5' -CGCAGAAATTTTAGAGGAACCTACAG-3'
bigL2, 5' -TTTGACTCCAAGACGCAGAGGATGAT-3'
bigL3, 5' -ATTTTCAAGATTTGTTCTCCAGATTT-3';
lipL45, 5' -ATTACTTCTTGAACATCTGCTTGAT-3'.

lipL45, 5' -ATTACTTCTTGAACATCTGCTTGAT-3'

用Taq聚合酶 (Qiagen) 进行逆转录反应产物的DNA PCR。在PCR前, 将以下引物加入反应液中:

bigL1, 5'-CTGCTACGCTTGTTGACATAGAAGTA-3'
bigL2, 5'-TAGAACCAACACGAAATGGCACAACA-3'
bigL3, 5'-ATCCGAAGTGGCATAACTCTCCTCAT-3'
lipL45, 5'-TGAAAAGAACATTACCAGCGTTGTA-3'.

根据为逆转录反应添加的引物, 预期得到500bp、479bp、440bp和438bp的PCR产物。为了进行PCR, 将反应混合物放入Techne Progene热循环仪中。起始变性步骤是95℃1分钟, 然后进行30个循环的反应, 每个循环为: 95℃变性30秒, 53℃退火30秒, 72℃延伸30秒。最后在72℃再保温30秒。

图4中的结果显示RT-PCR方法可检测BigL3转录物和对照LipL46转录物。BigL1和BigL2转录物未能检测到, 这表明BigL3表达于钩端螺旋体中, 而BigL1和BigL2可能不在其中表达。此外, 这些结果证明了RT-PCR方法在鉴定样品内特异BigL基因转录物中的应用。

实施例 3

重组 BigL 蛋白质的表达和纯化

此实施例举例说明了*bigL*基因的DNA序列在表达和纯化重组BigL多肽中的用途。设计两对寡核苷酸用于表达问号状钩端螺旋体 BigL3的两个区域。第一个区域是BigL3内相应于第2至第6重复结构域的区域, 相当于*L. kirschneri* BigL3 DNA 序列中SEQ ID NO: 6第131-649位。基于实施例1所鉴定的λ问号状钩端螺旋体 BigL3克隆的序列设计寡核苷酸, 它们的序列是:

45B-1 5'-ATGGGACTCGAGATTACCGTTACACCAGCCATT-3'
 45B-2 5'-ATTCCATGGTTATCCTGGAGTGAGTGTATTTGT-3'

用寡核苷酸45B-1和45B-2以及纯化的问号状钩端螺旋体基因组DNA进行PCR扩增, 获得DNA片段。用XhoI和NcoI酶(New Biolabs)消化这些片段, 然后连接入pRSETA表达载体(Invitrogen)(16)。用载体特异引物和引物步行法对克隆产物进行测序, 1557bp长的产物序列如SEQ ID NO: 7中所示。所编码的519个氨基酸长且被称为BigL3区域1多肽的预计序列如SEQ ID NO: 8中所示。

选择第二个用于表达的区域, 其中包含问号状钩端螺旋体 BigL3 C

- 末端区域最后的200个氨基酸。此区域相应于*L. kirschneri* BigL3中 SEQ ID NO: 6的第1687-1886位氨基酸。用于克隆此区域的寡核苷酸是:

```
BIGLCTERM1 5' aac-ctc-gag-cat-aac-tct-cct-cat-aac 3'
BIGLCTERM2 5' ttc-gaa-ttc-tta-ttg-att-ctg-ttg-tct-g 3'
```

用寡核苷酸BIGLCTERM1和BIGLCTERM2以及纯化的问号状钩端螺旋体基因组DNA进行PCR扩增, 获得DNA片段。用XhoI和EcoRI酶(New Biolabs)消化这些片段, 然后连接入pRSETA表达载体(Invitrogen)(16)。用载体特异引物和引物步行法对克隆产物进行测序, 600bp长的产物的核苷酸序列如SEQ ID NO: 9中所示。所编码的200个氨基酸长且被称为BigL3区域2多肽的预计序列如SEQ ID NO: 10中所示。

在BL21(DE3)pLysogen(Invitrogen)中表达重组蛋白质rBigL区域1和2。将异丙基-β-D-硫代半乳糖苷(IPTG; 终浓度2mM, Life Technologies)加入用编码钩端螺旋体DNA片段的pRSET质粒转化且培养至对数期的大肠杆菌BLR(DE3)pLysS(Novagen)培养物中, 以表达His6-融合蛋白。用6M盐酸胍溶解培养物沉淀并用Ni²⁺-次氨基三乙酸-琼脂糖(Qiagen和Pharmacia)通过亲和层析纯化His6-融合蛋白。通过凝胶电泳并用考马斯亮蓝染色来评估洗脱的His6融合蛋白的纯度。用PBS、10%(体积比)甘油、0.025%(重量/体积)叠氮化钠透析蛋白质。透析后, 用bicinchoninic acid测定蛋白质的浓度(42)。纯化的BigL3区域1转移后的Ponceau-S(Sigma Chem Co)-染色硝酸纤维素膜如图7中所示。纯化BigL3的相对迁移率与约58kD的预计分子量相似, 所说的分子量是根据重组蛋白质的预计氨基酸序列计算的。

实施例 4

实施例 4A: 针对重组 BigL 蛋白质的抗体的检测

此实施例举例说明了用BigL多肽检测受试者样品中抗体的几种方法中的两种。此外, 此实施例提供了用血清诊断试剂盒确定疑似感染的受试者中感染存在情况的方法。

在受感染测试者的样品中免疫印迹检测抗 BigL 多肽抗体

如前所述(17),将纯化的重组BigL3区域2多肽(1mcg/泳道)(实施例3)在不连续缓冲体系中进行十二烷基硫酸钠-聚丙烯酰胺12%凝胶电泳(SDS-PAGE)并转移至硝酸纤维素膜(Osmomics)上。用含5%脱脂奶的TBST封闭硝酸纤维素滤膜,并与以下类型的合并血清一起保温1小时以上:实验室证实患钩端螺旋体病的患者血清、其尿和肾培养物对致病钩端螺旋体呈阳性且为钩端螺旋体宿主的俘获大鼠(*Rattus norvegicus*)血清、和用问号状钩端螺旋体血清变型copenhageni菌株Fiocruz L1-130总裂解物免疫的实验性实验室大鼠和兔子血清。作为对照实验,将封闭后的膜与以下来源的血清一起保温:健康巴西受试者、无培养物或血清学证据显示存在钩端螺旋体感染的俘获大鼠以及免疫接种前的实验室大鼠和兔子。临用前将血清1:100稀释。漂洗后,将膜与缀合了碱性磷酸酶且1:1000稀释的山羊抗人 γ 链抗体(Sigma)一起保温1小时以上。通过NBT(0.3mg/ml)和BCIP(0.15mg/ml)进行颜色反应来检测抗原-抗体复合物。来自钩端螺旋体病患者和感染了致病钩端螺旋体的俘获大鼠的合并血清强烈地识别纯化的重组BigL3区域1蛋白质。不过,用完整钩端螺旋体裂解物免疫的大鼠则没有与BigL3多肽发生可观察到的结合,表明尽管BigL3于人工培养的钩端螺旋体中表达(实施例2,图4),但可能存在bigL3基因的差异表达。体外可能不存在足够量的天然BigL3蛋白质,而在天然感染期间,钩端螺旋体在体内产生了足够量的BigL3,从而诱发强免疫反应。此外,本实施例说明了一系列的动物在感染期间产生抗BigL3的免疫反应,以及对此免疫反应的检测和对针对重组BigL3多肽的抗体的检测可用作鉴定受试者中感染存在与否的方法。

为了进一步说明抗重组BigL3多肽抗体检测方法的用途,用实验室证实患有钩端螺旋体病的患者、巴西和美国健康个体以及住院或门诊诊断非钩端螺旋体病患者的血清进行免疫印迹评估。用微量凝集试验和培养物分离法证实临床疑似患者中钩端螺旋体病的诊断(5)。在巴西城市Salvador进行的5年钩端螺旋体病监测期间采集来自钩端螺旋体病患者的血清。对照个体的血清采集获自巴西Salvador住院和门诊及健康个

体的预存血清库以及通过美国疾病控制和预防中心赠送。所用血清的清单如表1中所示。按上述方法分析1: 100稀释的血清。在免疫印迹中发现1mcg重组BigL3区域1多肽的任何可见色带都被认为是阳性反应。

图8说明了来自受试钩端螺旋体病患者的血清与重组BigL3反应。表1总结了证明钩端螺旋体病患者中90%以上的住院病人和约70%的门诊病人在活跃感染期与rBigL3反应的有关发现。所有的(100%)钩端螺旋体病患者在其康复期都与rBigL3反应。表2用标准诊断试验比较了对rBigL3的血清反应性。在疾病的始发阶段, rBigL3血清反应性比标准诊断试验所观察到的高。美国健康受试者和88%的巴西健康受试者不与rBigL3反应, 证明了此对于rBigL3的反应是特异的。当以健康巴西受试者中IgM血清反应性的频率为基础进行计算时, 反应的特异性增加到了100%。总而言之, 这些发现说明, 此方法可作为活性感染的血清学标志, 而且是可用于钩端螺旋体病诊断的试剂盒的基础。

表1还总结了在钩端螺旋体病高危的地方性病区内rBigL3血清反应性的有关发现。居住于这些区域的人口中25%被证实为rBigL3 IgG血清阳性, 表明此反应可能是鉴定已过去的感染的有效标志。在证实患钩端螺旋体病的患者中, 56%的病人在感染钩端螺旋体病后的两年期间仍呈现抗rBigL3的血清反应性(表2)。在感染钩端螺旋体病后2-4年期间, 18%的病人被证实有rBigL3血清反应性。总而言之, 这些发现说明基于免疫印迹法的试剂盒可检测已过去的钩端螺旋体病感染。

实施例 4B: 已感染受试者样品中抗 BigL 多肽抗体的基于 ELISA 的检测

本实施例说明了ELISA方法可用于检测抗BigL多肽抗体以及确认疑似感染人群中的钩端螺旋体病患者。用悬浮于0.05M 碳酸钠, pH9.6中的His6-融合rBigL3在4℃包被平底聚苯乙烯微量滴定板(Corning)过夜, 0.5-100ng/孔。用蒸馏水洗平板两次, 用PBS、0.05%(v/v) Tween20 (PBST)洗3次。平板用封闭溶液(PBST/1% [重量/体积]牛血清白蛋白)室温保温2小时, 用PBST漂洗4次后, -20℃保存待用。将孔与50-200倍稀释于封闭液中的50 μl血清室温搅拌保温1小时。用PBST洗4次后,

将孔与缀合了辣根过氧化物酶且5000 - 20000倍稀释的抗人 μ 或 γ -链山羊抗体(Sigma) 50 μ l室温边搅拌边保温1小时。然后, 平板用PBST洗两次, 用PBS洗三次, 并在室温黑暗的条件下与底物缓冲液(0.03% [v/v] 过氧化氢, 25mM 柠檬酸, 50mM Na_2HPO_4 , pH5.0) 中的0.01% (重量/体积) 3,3',5,5'-N-四甲联苯胺以50 μ l/孔一起保温20分钟。加入25 μ l 2N H_2SO_4 终止颜色反应并在Emax微量平板读数仪(Molecular Devices, Sunnyvale, CA) 中测定在405nm波长的吸收值。

进行初始试验, 以确定能最好地区别来自实验室证实患钩端螺旋体病病例($n=4$)和来自巴西钩端螺旋体病地方性病区健康受试者($n=4$)的血清样品ELISA反应的抗原浓度(mcgg/孔)。用50、100或200倍血清稀释液和25、50、100和200ng/孔的抗原浓度进行棋盘滴定。图6说明了与对照个体相比, 对于钩端螺旋体病患者来说在所有血清稀释度和rBigL3多肽浓度都观察到了显著增高的吸收值。

在随后确定灵敏度和特异性的检测中, 用50ng rBigL3包被平板。分别用50和10000倍稀释的初始血清和二抗缀合物进行保温。测试双份平行试样的受试者血清样品, 计算两份测定的平均值供分析。对配对测定值之间的差异超过10%的样品重新测定。在各平板上均包括与所有重组抗原均反应的阳性对照血清样品和阴性对照血清样品, 均为双份, 作为测定的质量对照。图7说明, 与对照受试者相比, 急性发作期的钩端螺旋体病患者的IgM和IgG血清反应性吸收值显著增加(图7)。在比较疾病康复期病人的吸收值时, 这些差异有所增大。这些实验说明用基于ELISA的方法检测抗rBigL3多肽抗体对于鉴定钩端螺旋体病的感染是有效的, 可以用作诊断试剂盒。

实施例 5

在受试者中诱导抗钩端螺旋体的免疫反应

本实施例阐明了可通过用重组BigL蛋白质的免疫接种来诱发抗BigL蛋白质的免疫反应。用实施例3中所述的方法获得来源于问号状钩端螺旋体的纯化重组BigL3多肽。用配制于弗氏佐剂(Sigma) 中的40mcg

rBigL3皮下接种免疫实验室大鼠(Wistar株)。在第3周和第6周用20mcg rBigL3进行补充免疫。初次免疫后7周采集血液并处理成血清。按实施例4所述用rBigL3(1mcg/泳道)进行免疫印迹测定。图9显示了rBigL3-免疫大鼠的血清反应性。rBigL3是一种有效免疫原,在共3次免疫后,可以高于1:2500的滴度诱发免疫印迹rBigL3血清反应性。此外,针对rBigL3多肽的抗体能识别完整钩端螺旋体裂解物(10^8 钩端螺旋体/泳道)中的天然抗原(图9)。在免疫印迹中相对迁移率为200kD处的带呈弱染色,而较低相对迁移率处的带则染色较强,可能代表了200kD或高分子量BigL蛋白质的降解。针对这些天然抗原的血清反应性是特异的,因为在免疫前血清中未观察到有反应。

用来自*L. kirschneri*的纯化重组BigL多肽进行免疫原性实验。将纯化的重组蛋白质加样于制备型12% SDS-PAGE凝胶上,并通过电泳使其迁移进入分离胶。从胶上切下含100-200smcg重组蛋白质的带,干燥脱水,碾成粉末,溶于1ml水中,与1ml完全弗氏佐剂(Sigma)混合,皮下和肌肉内注射接种无钩端螺旋体抗体的新西兰白兔(Harlan Sprague Dawley)。此初次免疫后第4和第8周用粉末形式聚丙烯酰胺凝胶中相似量的融合蛋白混合不完全弗氏佐剂(Sigma)进行额外免疫接种。初次免疫后10周采集兔子的血液并处理成血清(Harlow, 1988)。如前所述(Guerreiro等, *Infect Immun* 2001年)用 10^8 个钩端螺旋体/泳道的浓度进行免疫印迹测定。

图10说明了用来自*L. kirschneri*的rBigL3免疫接种诱发了针对*L. kirschneri*和诸如问号状钩端螺旋体等其它致病钩端螺旋体中天然BigL3多肽的高抗体滴度。总而言之,这些发现说明用rBigL多肽进行免疫接种诱发了除了针对用于设计此重组rBigL多肽的种类外的其它致病螺旋菌的免疫反应。而且,此免疫接种方法产生的抗体可用于检测样品中的致病螺旋菌。

最后,本实施例证实了在有毒力的低传代株系中观察到天然BigL多肽的存在,而在无毒力的减毒高传代株系中则未观察到(图10)。来自rBigL3-免疫兔子的血清可识别有毒力株系钩端螺旋体完整裂解物中

相当于BigL3的预计200kDa的片段，而在无毒力的减毒株系中则没有。此实施例说明BigL蛋白质是毒性的标志，且抗BigL蛋白质的抗体可用于鉴定致病力强的株系。既然BigL本身可能是一毒性因子，如实施例中所证实的，诱发针对BigL蛋白质的免疫反应将可作为疫苗应用。

表 1. 按蛋白质印迹法进行的钩端螺旋体病患者血清和对照组血清中抗 rBigL 和 rLipL32 的 IgG 和 IgM 抗体的检测

研究组	测试 数目	rBigL3血清反应性			rLipL32血清反应性		
		IgM	IgG	IgM或IgG	IgM	IgG	IgM或IgG
阳性反应数目 (%)							
已证实患钩端螺旋体病的住院病例							
急性期	52	37 (71)	46 (88)	48 (92)	22 (42)	21 (50)	38 (73)
康复期	52	19 (37)	52 (100)	52 (100)	21 (40)	45 (86)	46 (88)
已证实患钩端螺旋体病的门诊病例							
急性期	14	6 (42)	8 (57)	9 (64)	2 (14)	2 (14)	3 (21)
康复期	14	7 (50)	14 (100)	14 (100)	6 (42)	5 (36)	8 (57)
健康受试者对照组							
非地方性病区 (美国)	30	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
地方性病区 (巴西)	40	0 (0)	5 (12)	5 (12)	2 (6)	0 (0)	2 (6)
高危地方性病区 (巴西)	40	0 (0)	10 (25)	10 (25)	4 (10)	5 (12)	8 (20)
病人对照组							
登革热	15	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
莱姆病	15	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
VDRL - 阳性	20	0 (0)	1 (5)	1 (5)	0 (0)	1 (5)	1 (5)

表 2. 针对钩端螺旋体病的以 rBigL3 和 rLipL32 为基础的蛋白质印迹测定与标准诊断测试的比较

疾病始发后的时间	测试数目	标准诊断评估			rBigL蛋白质印迹血清反应性			rLipL32蛋白质印迹血清反应性		
		MAT滴度倒数最大中值(范围)	MAT滴度倒数(x100)	ELISA-IgM	IgM	IgG	IgM 或 IgG	IgM	IgG	IgM或IgG
		阳性反应数目(%)								
急性期(N=52) ^a										
2-6天	21	200(0-1600)	12(57)	11(52)	12(57)	16(76)	17(81)	8(38)	8(38)	12(57)
7-15天	31	400(0-3200)	17(55)	20(91)	25(81)	30(97)	31(100)	14(45)	23(74)	26(84)
康复早期(N=52)										
16-21天	21	800(200-12800)	21(100)	15(100)	7(33)	21(100)	21(100)	8(38)	18(86)	19(90)
21-30天	31	1600(0-6400)	31(100)	21(100)	12(39)	31(100)	31(100)	13(42)	27(87)	27(87)
康复晚期(N=59)										
0-23个月	25	400(0-800)	21(84)	24(96)	0(0)	14(56)	14(56)	2(8)	2(8)	3(12)
24-47个月	17	400(100-1600)	17(100)	7(41)	0(0)	3(18)	3(18)	2(12)	2(12)	3(18)
48-78个月	17	200(0-800)	15(88)	5(29)	0(0)	3(18)	3(18)	2(12)	1(6)	3(18)

^a急性期血清样品在得到医院准许的前提下采集。

对本领域技术熟练人员来说显而易见的是，可以对本发明的化合物和方法进行各种修饰和改变。所以，本发明意图涵盖这些修饰和改变，只要它们在本发明权利要求和其等同内容的范围内。因而，本发明只受以下权利要求的限制。

参考文献

1. Levett PN., Leptospirosis即钩端螺旋体病, Clin Microbiol Rev., 2001, 14 (2) : 296 - 326.
2. Faine SB、Adler B、Bolin C、Perolat P, Leptospira and leptospirosis即钩端螺旋体和钩端螺旋体病, 第2版, 墨尔本, 澳大利亚, MediSci, 1999.
3. Farr RW, Leptospirosis即钩端螺旋体病, Clin Infect Dis, 1995, 21 (1) : 1 - 6, 检查7 - 8.
4. Lomar AV、Diament D、Torres JR, Leptospirosis in Latin America即拉丁美洲的钩端螺旋体病, Infect Dis Clin North Am, 2000, 14 (1) : 23 - 39, vii - viii.
5. Ko AI、Galvao Reis M、Ribeiro Dourado CM、Johnson WD、Jr. Riley LW, Urban epidemic of severe leptospirosis in Brazil即严重钩端螺旋体病在巴西的城市流行, Salvador Leptospirosis Study Group即萨尔瓦多钩端螺旋体病研究小组, Lancet, 1999, 354 (9181) : 820 - 5.
6. Bughio NI、Lin M、Surujballi OP, Use of recombinant flagellin protein as a tracer antigen in a fluorescence polarization assay for diagnosis of leptospirosis即重组鞭毛蛋白在用于诊断钩端螺旋体病的荧光极化测定法中作为示踪抗原的用途, Clin Diagn Lab Immunol, 1999, 6 (4) : 599 - 605.
7. Park SH、Ann BY、Kim MJ, Expression and immunologic characterization of recombinant heat shock protein 58 of leptospira species: a major target antigen of the humoral immune response即钩端螺旋体物种的重组热休克蛋白58的表达和免疫学鉴定: 体液免疫应答的主要靶抗原, DNA Cell Biol, 1999, 18 (12) : 903 - 10.
8. Haake DA、Walker EM、Blanco DR、Bolin CA、Miller MN、Lovett MA, Changes in the surface of 问号状钩端螺旋体 serovar

grippotyphosa during in vitro cultivation即体外培养过程中问号状钩端螺旋体流感伤寒血清变型的表面变化, *Infect Immun*, 1991, 59 (3): 1131-40.

9. Haake DA、Champion CI、Martinich C等人, Molecular cloning and sequence analysis of the gene encoding OmpL1, a transmembrane outer membrane protein of pathogenic *Leptospira* spp即编码OmpL1 (致病性钩端螺旋体物种的跨膜外膜蛋白)的基因的分子克隆和序列分析, *J Bacteriol*, 1993, 175 (13): 4225-34.

10. Haake DA、Martinich C、Summers TA等人, Characterization of leptospiral outer membrane lipoprotein LipL36: downregulation associated with late-log-phase growth and mammalian infection即钩端螺旋体外膜脂蛋白LipL36的鉴定: 与对数晚期生长和哺乳动物感染有关的下调, *Infect Immun*, 1998, 66 (4): 1579-87.

11. Haake DA、Mazel MK、McCoy AM等人, Leptospiral outer membrane proteins OmpL1 and LipL41 exhibit synergistic immunoprotection即钩端螺旋体外膜蛋白OmpL1和LipL41展示协同免疫保护作用, *Infect Immun*, 1999, 67 (12): 6572-82.

12. Haake DA、Chao G、Zuerner RL等人, The leptospiral major outer membrane protein LipL32 is a lipoprotein expressed during mammalian infection即钩端螺旋体主要外膜蛋白LipL32是哺乳动物感染过程中表达的脂蛋白, *Infect Immun*, 2000, 68 (4): 2276-85.

13. Shang ES、Exner MM、Summers TA等人, The rare outer membrane protein, OmpL1 of pathogenic *Leptospira* species is a heat-modifiable porin即致病性钩端螺旋体物种的稀有外膜蛋白OmpL1是热可变孔道蛋白, *Infect Immun*, 1995, 63 (8): 3174-81.

14. Shang ES、Summers TA、Haake DA, Molecular cloning and sequence analysis of the gene encoding LipL41, a surface-exposed lipoprotein of pathogenic *Leptospira* species即编码LipL41 (制备型钩端螺旋体物种的表面暴露脂蛋白)的基因的分子克隆和序列分析,

Infect Immun, 1996, 64 (6) : 2322 - 30.

15. Yelton DB、Charon NW, Cloning of a gene required for tryptophan biosynthesis from *Leptospira biflexa* serovar patoc into *Escherichia coli*即将双曲钩端螺旋体血清变型patoc中的色氨酸生物合成所需要的基因克隆到大肠杆菌中, *Gene*, 1984, 28 (2) : 147 - 52.

16. Flannery B、Costa D、Carvalho FP等人, Evaluation of recombinant *Leptospira* antigen-based enzyme-linked immunosorbent assays for the serodiagnosis of leptospirosis即对基于重组钩端螺旋体抗原的酶联免疫吸附测定法用于钩端螺旋体病血清诊断的评估, *Journal of Clinical Microbiology*, 2001, 39 (9) : 3303 - 3310.

17. Guerreiro H、Croda J、Flannery B等人, Leptospiral proteins recognized during the humoral immune response to leptospirosis in humans即人体中针对钩端螺旋体病的体液免疫应答过程中识别的钩端螺旋体蛋白质, *Infect Immun*, 2001, 69 (8) : 4958 - 68.

1) 一般信息

I. a) 申请人: FIOCRUZ和KO, Albert I.; HAAKE, David A.; REIS, Mitermayer Galvao; MATSUNAGA, James; CRODA, Julio Henrique Rosa; SIQUNEIRA, Isadora Cristina; RILEY, Lee W.; BAROCCHI, Michele; YOUNG, Tracy Ann.

I. b) 地址:

Centro de Pesquisas Goncalo Moniz

Fundacao Oswaldo Cruz/MS

Rua Waldemar Falcao, 121

Salvador, Bahia 40295 - 001

巴西

II) 发明名称: “钩端螺旋体种中存在的具有重复性细菌Ig样 (Big) 结构域的蛋白质家族”

III) 序列数目: 10

IV) 计算机可读形式:

IV. a) 介质类型: 软盘

IV. b) 计算机: IBM PC兼容机

IV. c) 操作相同: PC-DOS/MS-DOS

2) 序列的一般信息

I. a) 序列编号: SEQ ID NO: 1

II) 序列特征:

II. a) 长度: 3672碱基对

II. b) 类型: DNA

II. c) 链: 单链

II. d) 拓扑结构: 线性

III) 在基因组中的位置:

III. a) 生物体: *Leptospira kirschneri*

III. b) 名称/关键词: CDS

III. c) 在图谱中的位置: (0) - (3672)

ATGAAGAGAACATTTTGTATTTTCGATTCTTCTTTTCGATGTTTTTTCAAAGTTGTATGTCTTG
GCCACTTTTAACCAAGTCTCGCGGGTTTAGCAGCTGGTAAAAAAGTAATGGGCTGCCCTTTT
TCCACCTTCTATTAAGTAACTCTGATCCAGTTATTACAAGGATCGAGCTCAGTTATCAAAT
TCTTCCATCGCAAAAGGTACAAGTACAACCTCTCGAAGTCACCGCAATCTTTGATAACGGAAC
AAATCAGAATATTACGGATTTCGACATCTATCGTTTTCCGATGCCCAATCAATCGTTGACATTC
AAGGTAACAGAGTCAGAGGAATCGCTTCTGGTCTTCCATTATAAAAGCTGAATACAACGGG
ATGTATTCTGAACAAAAATTACGGTTACACCAGCCACGATAAACTCAATTCAAGTTACGAG
TTTAGATGACGGTATATTACCTAAAGGTACAAATCGTCAATTTGCTGCCATCGGTATCTTTT
CGGATGGTTCTCATCAAGATATTTCCAACGATCCATTGATCGTTTGGTCTTCCAGTAATATA
GATTTAGTTCGAGTAGATGATTCCGGTTTTGGCCTCAGGTATCAATTTAGGAACGGCTCATAT
TCGTGCATCCTTTCAATCAAAACAAGCCTCCGAAGAGATAACTGTTGGTGACGCTGTTCTTT
CTTCTATCCAAGTAACTTCCAACAGTCCAATATTCTCTCGGAAAAAACAACAACTCACA
GCTACTGGAATTTATTCGGATAACTCTAACAGGGATATTTCTCTTCTGTTATCTGGAATTC
TTCTAATTCCACTATCGCTAATATTAGAATAACGGAATATTAGAAACAGCTGATACTGGAA
TTGTTACTGTTTTCTGCTTCTAGAGGTAATATAAATGGTTCATAAACTAATCGTCACTCCT
GCTGCCTTAGTTTTCTATTTCTGTTTCTCCTACAAATTCTGCAGTAGCAAAAGGTTTACAAGA
AAACTTTAAAGCTACAGGGATCTTACAGATAATTCGAACTCAGATATTACAGATCAAGTTA
CTTGGGATTCCTTAATCCGGATATTTCTTCCATTTCCAATGCAAGTGATAGCCACGGGTTA
GCTTCCACACTCAACCAAGGAAATGTTAAGGTCACCGCTTCCATCGGTGGAATACAAGGATC
CACTGATTTTAAAGTTACACAAGAGGTATTAACCTTCCATCGAAGTTTCTCCAGTTTACCTT
CAATTGCAAAAGGACTAACTCAGAAATTTACGGCGATCGGGATTTTTACGGATAACTCCAAA
AAAGATATTACAAATCAAGTCACTTGGAAATCTTCTTCAGCAATCGCAAGCGTGTCTAACTT
AGATGATAATAAAGGTCTGGGAAAAGCTCACGCTGTTGGAGACACGACTATTACCGCTACTT
TAGGAAAAGTTTACGGTAAAACCTTGGTTTACTGAGTTCCTGCGGTTCTCACTTCTATTCAA
ATCAATCCTGTAAATCCTTCTTTCGAAAAGGGTTAACTCAAAAATTTACGGCTACTGGGAT
CTACTCTGACAACCTAACAAGGACATTACTTCCCTCCGTTACTTGGTTCTCATCCGATTCTT
CAATCGCAACAATTTCAAACGCCAAAAAATCAAGGAACTCTTACGGAGCAGCTACAGGA
GCAACGGATATTAAAGCCACATTCGGAAAGGTAAGTAGTCCAGTTTCTACGTTATCCGTTAC
TGCTGCAAACTTGTGAAATACAAATCACACCGGCCGCTGCTTCCAAAGCAAAGGGAATTT
CCGAAAGATTTAAAGCAACCGGTATTTTTACAGACAACCTAATTCCGATATTACAAATCAG
GTCACCTGGAGTTCATCTAATACAGATATTCTTACCGTTTTCAATACAAACGCCAAACGCGG
GTTAGGTTCCACTTTAAAACAAGGAACTGTTAAAGTTATCGCTTCCATGGGTGGAATCGAAA
GTTCTGTAGATTTTACCGTACACAGGCTAATTTGACTTCGATCGAAGTCTCTCCAACCTCGC
TCTTCGATTGCAAAAGGACTAACTCAAAAATTTACCGCTATAGGTATTTTTACGGATCATT
TAAGAAGGATATTACAGAGCAAGTTACTTGGAAAGTCTTCTTCGAAAGTATTAAATATGTTGA
ATGCATCCGGTGAAGAAGGAAGAGGTAAGGCAATTTACAGTCGGGAAAGCGACCATTACTGCA
ACCTTAGAAAACTTTCCGGGAAAGCTGATATTACAGTTACTCCCGCGGTTCTTACTTCAAT
TCAAATCAGTCCTGTGAAACCTTCTCTTGTAAAAGGTTAACAGAAAATTTTTCTGCTACAG
GTATCTACTCTGATAATTCCAGCAAGGACATAACTTCCCTCCGTTACATGGCATTGTTCAAC
AACTCTGTTGCAACGATCTCGAACACGAAAAATTACCATGGACAAGCTCACGCAACCGGTAC
AGGGATAGTGGGTATTAAAGCGACATTGGGAAATGTAAGCAGCCAGTTTCCAAATTATCCG
TTACCGCAGCAGAAGTGGTTGAGATTGTGTTAAATCCTACTTTATCTCACAAGGCCAAGGGA
CTTACTGAAAATTTTAAAGCGACCGGCGTATTTACGGACAATTCGACAAAAGATATTACCGA
CCAGGTTACTTGGAAATCTTCCAATACTGCCTACGCAGAAATTTCAAACGCAACTGGAAGTA
AAGGGGTTGTTAATGCACTCTCGAAGGGAACGAGTCACATTTCCGCTACCTTAGGTTCAATT
TCAAGTGCAATGCGACATTCCAAGTTACTCCAGCAAAAATAGCTTCGATCGAAATAACACC
AAATAATTTCTTCTTGATCAAAAACCTTAGTTATCCATTTAAAGCAATTGGAATCTATACGG
ATAATACAAAGACACACATTACAAAACAAGTTTCTGGTCTTCTCTGATCCGAATGTTGCA

TCGATCGATAACACATTTTCATTGGCTGGCTCAGCTACCGCAATCGATGATGGAAAAACGAA
 CATCACTGCAACGTTATCCGACTCTATGTCCGCTTCCACTACTTTGTATGTCACCTCTGCTA
 CGCTTGTTGACATAGAAGTAAAACCTAGTATCTTCGTTCTGAGTGAAGGTCTTACACTACAA
 CTGACCGCTACCGGCATCTATTCGGATTACTCTACCTATGATTTGACTCAGGTTGTAACGTG
 GACTTCCAGCGAACCATCCAACATTTTCGATCGAAAATACAGCCGGTAAAAAAGGTAAAGTAA
 CGGCTCTTGCATTTGGAGCTTCAGAATTTACGGCAACCTACGATTCTATTGAAAGTAATCGA
 GCTTGGATATTTGTCAATGACGAGAAATTTGTAAACATAACCATTAGTTCTTCTCAAGTTTT
 GACAGACAAGGGCTTGACTCAACAATCAAAGCAATCGGAACTTTGAAAAAGGTAGCGAAC
 TTGACCTTACGGATCTTGTAACTGGAAGTCCTCTGATTCTAAGGTAGCTTCTATCGGTAAC
 TCTAATGATGACAGAGGTTAATAACACCGCTTTCTGTAGGTTCTCTAAAAATTTCTGCGAC
 TTACAATTCTATCCATAGTAACTCTATTGATTTTGAAGTAACTCCAGAAATATTAGCCTCTA
 TTAAAACGAAGCCG

I. b) 序列编号: SEQ ID NO: 2

II) 序列特征:

II. a) 长度: 1224

II. b) 类型: 氨基酸

II. c) 拓扑结构: 线性

III) 类型: 肽

IV) 来源: N-末端至C-末端部分

MKRTFCISILLSMFFQSCMSWPLLTSLAGLAAGKKSNGLPFFHLLLSNSDPVITRIELSYQNSSIAKGTSTTLEVTAFD
 NGTNQNTDSTSIIVSDAQSIVDIQGNRVRGIASGSSIIKAEYNGMYSEQKITVTPATINSIQVTSLDDGILPKGTNRQFA
 AIGIFSDGSHQDISNDPLIVWSSNIDLVRVDDSGLASGINLGTAHIRASFQSKQASEEITVGDVAVLSSIQVTSNSPNIP
 LGKKQKLTATGIYSDNSNRDISSSVIWNSSNSTIANIQNNGILETADTGIVTVSASRGNINGSIKLIVTPAALVSI SVSP
 TNSAVAKGLQENFKATGIFTDNSNSDITDQVTWSSNPDILSISNASDSHGLASTLNQGNVKTASIGGIQGSTDFKVTQ
 EVLTSIEVSPVLPISIAKGLTQKFTAIGIFTDNSKKDITNQVTWSSSAIASVSNLDDNKGLGKAHAVGDTTITATLGKVS
 GKTWFTVVPVAVLTSIQINPVNPSLAKGLTQKFTATGIYSDNSNKDITSSVTWFS SDSSSIATISNAKKNQNSYGAATGAT
 DIKATFGKVSSPVSTLSVTAAKLVEIQITPAAASKAKGISERFKATGIFTDNSNSDITNQVTWSSSNTDILTVSNTNAKR
 GLGSTLKQGTVKVIASMGGIESVDFVTQANLTSIEVSPTRSSIAKGLTQKFTAIGIFTDHSKKDITEQVTWKSSSKVL
 NMLNASGEEGRGKAISVGKATITATLEKLSGKADITVTPAVLTSIQISPVKPSLVKGLTENFSATGIYSDNSSKDITSSV
 TWHSFNNSVATISNTKNYHGQAHATGTGIVGIKATLGNVSSPVSKLSVTAELVEIVLNPTLSHKAKGLTENFKATGVFT
 DNSTKDITDQVTWKSSNTAYAEISNATGSKGVNALS KGTSHISATLGSISSANATFQVTPAKIASIEITPNNFFLIKKL
 SYPFKAIGIYTDNKTDTIKQVSWSSSDPNVASIDNTFSLAGSATAIDDGKTNITATLSDSMSASTLYVTSATLVDIEV
 KPSIFVLSEGLTLQLTATGIYSDYSTYDLTQVVTWTSSEPSNISIENTAGKKGKVTALAFGASEFTATYDSIESNRAWIF
 VNDEKFNITISSSQVLTDKGLTQQFKAIGTFEKGSELDTDLVTVWKSSDSKVASIGNSNDDRGLITPLSVGSSKISATY
 NSIHSNSIDFEVTPPEILASIKTKP

I. c) 序列编号: SEQ ID NO: 3

II) 序列特征:

II. a) 长度: 5863个碱基对

II. b) 类型: 核酸

II. c) 链: 单链

II. d) 拓扑结构: 线性

III) 在基因组中的位置

III.a) 生物体: *Leptospira kirschneri*

III.b) 名称/关键词: CDS

III.c) 在图谱中的位置: (0) - (5863)

ATGCCTAAACATATCAACAACTCAGAGATAAAAAACGTGGCCTTTTCTTCAGTTATTTTTATTCTTTTTCTAACATT
CAGCCTATTTTTTTGGAAAGTTGCGCGCTTGGCCAAATTTTTCAGGCACACCTGGTTTATTAGCAGGTAAGGAAAGCG
GAGCAACAATTCACCTTTGGATGCTTTTTTAGGAATAGATAATCCGCTCGAATCGGAGCCATCCGAAGCAGAGTTAGAT
CGGATCGAAATTTCCGTACCGAACTCAAATTTAGCTCGAGGACTACTTTACATCTAAACGCCACAGCCATCTATAAGA
CAATACTCACCGAGATATTTCTCGGAAGGATCCTGGTCTTACGGATTTCGAGCATTCTCAAGCTATTAACACAATCTC
AATTCAAAGGAATGAATCTAGGTTCTGGAAACGTTAATGTATCCTTTCAAGGAAAAACGCAACTACAACGTTAACCGTT
ACATCCGCTGTTTTGTCCGATCTGACCGTAACTTGTGTGAACCAAGGTAGTCCATTACCTGTGGAATCGATCGTCAATG
TAAATTAGAAGGAATTTTTCGGACGGTAGTACTCAGGTTTAACTTCGGATCCAAGCGCGTCTGGAACGTAACCCAT
CTTCTATTGCAAGGTGTAACACACCAGGTTAGTTTTCCGGACTTTCAGGTAACACTTTTTATTACCCTCTTATGGA
AGTAAAACTCCAGTTTGAATGTGACCGTAAAGTGGCGCAACCCTAGCTCGATCTCAGTGACTCTGCCAATCAAGTTA
TCCTCTGGCAAGGTCCAACAGTACACAGCAATCGGAACCTACAGCAATCAGTCCACTCAAGATTTAACAAATCAGGTTT
CCTGGGCTTCTTTAAATACTTCGTTGCTACGATCGATAATTCTACATCCGCCAAAGGTATGCTTACTACTCAATCAACC
GGTTCAGCAACATCACGGCAACGTTAGCGGAAATACCGGACAGACTACTAGTAAACGTCACCTCCGCGATTCTTACTA
GTATTAGCATCACTCCTGCAAAATCCAAGCGTAGCCAAATGGAAGCATTATATCTTACCGCCACCGGAGTTTTTCGGAT
GGTACAGTTTTCCGACATTACCAACCAAGTAACTGGTCCAGTTCCCTTAAACAAGTGTAGCTACCAGGATAACTCAGGCCG
TTTTTCCGGAAGAATTTCCGGAGTCCGAGTTGGTAGTACGAATATCACCGCCGATCGGTGGAGTAGATATACGGTTT
CTTTAAATGTTACCAACGCCACTTTAGAAATCGATTCAGTGGTTTTCCGATTCCCATTCGATAGCTCGAGGTACGTTACG
TTTGTACAAGCATAGGAGTCTACTCGGACGGTTCTTCTCAAACATAAGTGATCAAGTTGCCTGGAACGCTCTAATTC
TTCAATATTACAAATATCTAATTTAAATGCAGTCCCAAAGAGAAATACAATCTCCTTCTCCGGAGGCTTAGGTACAG
CAAGGATCACCGCAACTTTAGAGCAaTCTCCTCATATACCACATCTCGGTCAATGCAGCAACTTTAGTTTCTATCGAA
GTGTACCCACAAATCCTTCCGATCTTCCAGGACTTACCGTTCCCTTTACGGCGACCGGgTATTATACGGATGGAAGTAA
TCAAATCTGACTTCTCAAGTAACTTGGAAATCTCCAACACGAACAGAGCTACAATCAGCAACGCAACCGAACTCAAG
GAATGCCTTGGGCTCTTCTGTGGAACCTACGAACATATCaGCAACGTTAGGTGCGGTTACTTCTCCGTTACCCTCTT
ACGGTCACAAACGCGGTTTTTAAATTCGATCAGGATTAATCCTGCTCTTCTTCCGTTAGCAGTAGGAAGAAGTCTGAACCT
TACTGCAACCGGAACTTATTCTGACGGAAGTAACCAAGATTTAATACCTCCGTCGCTTGGACGAGTACGGATTCTTCCA
TCGTTTCCGTAGACAACGCCCTCAGGTAGACAGGGGCAGACAGGTTGTTGCACAGGTAACACTCAGATCAGTGCACACA
TTAGCGGAACTTCTTCTGCTATCAATTTACGGTAAGTGCAGCGGTTTTAGATTCAATTCAGTAATCTGGAAGATTC
TCCGATTGCAAAGGAACTTCTACAAGAGCAATCGCGACGGGTGTTTTTTCAGACGGAAGCAATTTGAATATTAGTGATC
AAGTTATTGGGATAGTTACAAACAAACGTTGATCCAGCTAGGAGTTTTAGAAACCGGTCCTAAAAGAACTGATGAAT
TCTCCCGCAATGGAAACAGTACCCTGGAACTCAAGGATCACTGCAACGTTAGGAGGTGTGAGCGGATACGCCGATCT
TACAGTAATCGCTCAAGTTTAAACAGCATTCAAATCAGCATCCAGACATCCGAGCGTTGCCAACGGTACTGACTCAAAT
TTACTGCAACCGGAGTTTACTCAGATGGTAGCAATCAGAATCTAACCGATTCCGTTACTTGGGCGTCTTCCAATCTGCT
GTTGCCACGATCAGCAACGCTTCCGGAACCAACGGTAAAGCTACTACTCTTCAAACCTGGATCCACCAATATCAGCGGAG
TCTGGGCGCCACTACTTCTGATCCAAGTGTATTAACGGTTACAACGCAACCTTAAACAAGTATCAGGATCGCTCCACCT
CTTCTTCAACATCGCAAAGGATTAATCAAGACTTTGTAGCGACCGGTTATTATACAGATGGTTCTCTAGAGACCTG
ACCCTCAAGTCACTTGGAAATCTTCCAATACTTACCCTACGCTACGCAATGCAAACGGAACTCAAGGAAGATGGC
CGCGTTCGATCTGTTCTACAAATATCTCCGCTCTTAGGAGAACGTATAGTCAGACCACAACCTACCGGTTACTAT
CTGCGGTTCTGAATTCGATCCAGGTTTTCTCCAGCGGACTTAGTGTAGCCAAAGGAAACCAAGGCCTACCCGCGATC
GGAGTATATTAGATTTAGCACGTTAGACGTTACTTCTCAGGTTACCTGGACTTCTTCCAGCGTTTCGATCGCTACGAT
CAGCAATGCAAGCGGACACGAAGGTTTAGCTACGGCTGTAGGCACGGAACTTCCACAATTACCGCAACTCTTGAGGAA
TTTCTAATTTACAGGTTTACCGGTTACGGCCCGGCTATTGGTTTTCTTTCGGTAGGTCCTACCAATAGTTTTGTTTTAT
ATGACACAAACAAAAATTTATGGCTACTGGAACGTTATCTGACGGAACGATGCAGGATCTTACAACCTCAAGTCAACCTG
GACTTCTTCCGATACAACCTTGGGAACAATCAGCAACGCTTCCGAAATAGAAGGTAGGGCTACAGGAATTTGCTGCGGTTG
CCATAACGATCACTGCGACTTTGGGAAGTATCAGCGGAAACACTTCTTTGACTATAATCTTTTTAGATACGATAGCACCT
GCGATCACAACGTTAGTCGCTTAACCTCTACTACTTTAAGAATTACATATTCGAAAACGTAACGAAACCCAGGCAAA
AACCGCGGCAATACAAACGCTCTTACTTCTCCGTAACCGGAAGTTGTTTCAGATAACAGCAACTTTACTTCTACCT
CTTCTGTGATTACTGTTTCTCAGTGAAGCGGATCTGTGTTGTTCTAATCTAGGTTCTTCCAAACGCTTAAAC
GCACCTATACGATTTAGTGAATAAATCGGAAATACAAGATCTTCTACAACCCCAACAATTTGGGTTGTGCAAACTA
CGAGACTTCTTAGGACAGGAACAATCAAATCGTATCCGCTTCCGAAATAGAAGGTAGGGCTACAGGAATTTGATTTGCT
CTAAGGCTCCTAATCTGGAAACAATGTCCGCGGTTCCGCAGAATGTACCGGTTCTGCAGAATGTTCTAATCGTTACAAA
ATTTCCGGAGCAAGCGATCTGGAACAATTAACAGCGTAAAGGTGTAGATGGAATTTTGTAAACGGAGCAACTGCAGA
TTCCGCAAAAGTATGCGTAATTCATAATTTAGTACAACCGGACACAATATACAATCACTGCGGATTCGGTAGACG
GAGACGGATTTGACAACCTCAAGTGGGGATCAATCCGAAATCTTTGGATACAGAGAATCTTCAATCTTCTCCAGAGAC
AGGCTTCTTTTTAGGATGTGGAACGTTCCGGTCACTTTGCGACGGACCGATTCCATCGATCCAAACTCATCCAC
GTTCCGTTATCTAATCGATTTAACTCTAAGATCTATTCCAGGACAAACAATTCGGGAACGGAGCGGTTTCGATTTGCCT
ATGATGGAAGTGTCCAGAAATCAGTTCATTTCTCCTTTGAAAAGACACAACCGTTCAAGACGGGTGACGCGACTAACGTA
AGTTCAAACCTCAGCTTCTCCAGAGAGAATTCGATCTCGGTTCCGCTTACGTTACATTAGGACACTCCGGATGTACTAC
AAACAACGGAACTCTTCTCTAGGATGTGGTCCGGATAACGAAACGGAAAGAGGAGTATTCGCTACTGGAATCTTTCCA
CGGCTCCTATCTATTTGTTGCAGCTGCAAAAACCGTAGCGGACGGCCTGGGgCAATACTTATTTGATTATCTGTATTAC

TCCGCAGACACTTCTACTAATAACAAGTTTCAAATATATAGATCTAGGATCGATCACCGGAACTTTAACC GCCGGAAC TTC
TTCGTTACTGTACTCAATAATAGAGTGTTCAGGTTTTGCAAAGTCAAGCAACGACGGAAATCGGATTGTTCCGAGGAC
TTAATGCACCCGATTTTGGATTTGTAACGTTTAACTCAGCGGACTCAGGAACCTGGATTTTGTACTCCAGGCTCCAAC TGC
GACCGTTTTGACGGAACCAAAGGAAAAAGAATCCGGATCGATTTCCTTCCTTACTTCGGAGGACCGTCCACCGGTTTTATT
AGGAATTAATAATAATGCACATCCAAACTGGGCCATATATATCGGAGTCGATTCATGTTTCGTATTTAAAAATCGTATCT
ATGCCGCAACCGGAGGATTACACGCGGTAGGACATAACGGTTCATAATACGTTCTACAACCTGCAGATCCAACCGCGGCT
TGTACCCGGACCGGACTCTTGTCTAACTGGGTGGAAATTGGACCTAGAACCAACACGAAATGGCACAAACAGTCCCACAAA
CAACTGGTCTCTTTAGAGTTAAATCAATTTTACAATCTGATCCGGGAGATAAGGCGTTTGACAAATTTGCCGAGTTCA
ACAATAACCTTTATGTAAC TAGAACCATTGTATTCAAAGTTCTCAAGCGACTGGAATCAGAACCAATCCAGGAACCGTA
ACAGGATGTACAGACGGAACTACAAATCGAAGGGCACAACCTTTGAAATGTGATCCTACAATTT CAGGAAACACGAG
CGAATGTGATGCAGCGGATTGGTCGGTCGTAGGCGACGACGGAACCGGAATCACAACATGGGAGATTTCAAAACCGAA
CGATCACCATTGGTGATGAAAACGGATCCTATCTTTACATAGGATATGATAATCCAACCGGAATCAGAAATTTATAGAACC
AACGTAGCCAACCCGGGATCATCCTCTGCGTCTTGGAGTCAAATCGCCGGGAAACGGTCTCACAGATGCGACTAACGTTCA
ACAAATTTACTCGGCCGTATCCGTACCTCCGGAAGTATCAATTATATCTACGTAAGCGCTGGAAAAAGTAAACGTTTCTG
TTCGGACGTATCGTCAACAAAAT

I. d) 序列编号: SEQ ID NO: 4

II) 序列特征:

II. a) 长度: 1954

II. b) 类型: 氨基酸

II. c) 拓扑结构: 线性

III) 类型: 肽

IV) 来源: N-末端至C-末端部分

MPKHINKLRDKKTWPFLLQFIFILFLTFFSLFFLES CAAWPIFSGTPGLLAGKKSGANNSLWMLFLGIDNPLESEPESEALD
RIEISVPNSNLARGTTLHLNATAIYKDNTHRDISSEGSWSSTDSSILKLLTQSQFKGMNLGSGNVNVSFQGNATTTTLTV
TSAVLSDLTVTCVNVQGSPLPVGIDRQCKLEGISFDGSTQVLTSDPSASWNVTSQSSIAGVNTTGLVSGLSPGNFTFITTSYG
SKTSSLNVTVSAATLSSISVTPANSSYPLGKVQOYTAIGTYSNQSTQDLTNQVSWASLNTSVATIDNSTSAKMLTTQST
GSANITATLGGITGQTTVNVTSAVLTSITITPANPSVANGRTLYLATGVFSDGTVSDITNQVTWSSSLTSVATADNSGG
LSGRISGVGVGSTNITAAIGGVDITVSLNVTNATLESIQVVS DSHSIARGTSTFVQAIQVYSDGSSQNISDQVAVNNSNS
SILQISNLNAVPKREIQSPSSGGLGTARITATLEAISSYTDISVNAATLVSVIEVSPTNPSVSSGLTVPFFTATGVYTDGSN
QNLTSQVTWNSNTRATISNANGTQGIALGSSVGTNI SATLGAVTSSATTLVTNAVLSITITPSPVAVGRSLNL
TATGYS DGSNQDLTTSVAWTS DSSIVSDNASGRQQTGVAQNTQISATLGGTSSAINFVSAAVLDSIQVTTLEDS
PIAKGTSTRAIATGVFSDGSNLNIDQVIWDS SQTNVIQLGVLETGPKKLMNSPANGNSTTGTSRITATLGGVSGYADL
TVIAPSLTSIQIDPTHPSVANGLTQNFTATGVYSDGSNQNLTD SVTWASSNPVAVATISNAGTNGKATTLQGTSTNISAS
LGATSDPSVLTVTNATLTSITIAPTSSFNIAKGLNQDFVATGYT DGSRRDLTQVTWNSNNTSTATISNANGTQGRMA
AVDTGSTNISASLGGTYSQTTNVTVTSAVLNSIQVSPADISVAKGNTKAYTAIGVYSD FSTLDVTSQVTWTS SSVSIATI
SNASGHEGLAVTGTSTITATLGGISNSTSLVTAAVLVLSVSGPTNSFVYMTQTKNFMTGYS DGTMQDLTQVTW
TSSDTLGTISNAFGIEGRATGIAAGAITITATLGSISGNTSLTIIFLDTIAPAITNVVALTPPTLRITYSENVNETQAK
TAANYKALATSSVTGSCDNSNFTSTSSVITVSSVSGSGSVFLTLGSSQTSNAPYTIILVNKSGIQDLSTTPNNLGCANY
GDFLQEQIKIVSASCANSNSVILNFSKAPKSGNNVAGSAECTGSAECSNRYKISGASDLGTINSVKVLDGICNGATAD
SAKVCVIHNLVQTGAQYTIITADSV DGDGFDNSSWGSIRNSLDTENLQSSPRDRASFLGCGTSPVNFADGPI SIDPNSST
FGYLIDFNSKIYSGPNNSNGALRFAYDGSVPESVQFSFEKDTTVQGDATNVSSNSASSRENSISVPPYVTLGHSGCTT
NNGTSLSLGCGPDNENGRGVFATGILSSVSYLFAAAKTVADGLGQYLFDYLYSADTSTNTSFKYIDLGSITGTLTAGTS
SLTVLNNRVFAGFAKNSNDGIGLFGGLNAPDFGFVTFNSADSGTGCTPGSNCDADFDTGKGRIRIDFLPYFGGPSTGLL
GINNNAHPNWAYYIGVDSMFVFNRIYAANGGLHAVGHNGSIIRSTADPTAAGTGPDCSNWVEIGPRTNTKWHNSPTN
NWFSLNQLFYNLIPGDKAFAQFAEFNNNLYVTRTICIQSSQATGIRTNPGTPTVGTGCTDGTTRRRAQLWKCDPTISGNTS
ECDAADWSVVGDDGTGINTMGDSTNRTITMVMKNGSYLYIGYDNPNGIRIYRTNVANPGSSASWSQIAGNGLTDATNVQ
QIYSAVSVPSSGSIYIYVSAGKSNVSVRTRYRQQN

I. e) 序列编号: SEQ ID NO: 5

II) 序列特征:

II. a) 长度: 5658个碱基对

II. b) 类型: 核酸

II. c) 链: 单链

II. d) 拓扑结构: 线性

III) 在基因组中的位置

III. a) 生物体: *Leptospira kirschneri*

III. b) 名称/关键词: CDS

III. c) 在图谱中的位置: (0) - (5658)

```

ATGAAGAGAACATTTTGTATTTTCGATCTCTCTTCGATGTTTTTCAAAGTTGTATGTCTTGGCCACTTTTAACCGTCT
CGCGGGTTTAGCAGCTGGTAAAAAAGTAATGGGCTGCCCTTTTCCACCTTCTATTAAGTAACCTGATCCAGTTATTA
CAAGGATCGAGCTCAGTTATCAAAATCTTCCATCGCAAAGGTACAAGTACAACCTCGAAGTCACCGCAATCTTTGAT
AACGGAACAAATCAGAATATTA,CGGATTTCGACATCTATCGTTTCCGATGCCCAATCAATCGTTGACATTCAAGGtAACAG
AGTCAGAGGAATCGCTTCTGGTCTTCCATTATAAAAGCTGAATACAACGGGATGTATTCTGAACAAAAAATTACGGTTA
CACCAGCCACGATAAACTCAATTCAAGTTACGAGTTTAGATGACGGtATATTACCTAAAGGTACAATCGTCAATTTGCT
GCCATCGGTATCTTTTCGGATGGTTCATCAAGATATTTCCAACGATCCATTGATCGTTTGGTCTTCCAGTAATATAGA
TTTAGTTTCGAGTAGATGATTCCGGTTTGGCTCAGGTATCAATTTAGGAACGGCTCATATTCGtGCATCCTTTCAATCAA
AACAGCCTCCGAAGAGATAACTGTTGGTGACGCTgTTCTTTCTTCTATCCAAGTAACTTCCAACAGTCCAATATTCCT
CTCGGAAAAAACAAAACTCACAGCTACTGGAAATTTATTCGGATAACTCTAACAGGGATATTTCTCTTCTGtTATCTG
GAATTTCTTAATTCCTACTATCGCTAATATTCAGAATAACGGAATATTAGAAACAGCTGATACTGGAATTTGTTACTGTTT
CTGCTTCTAGAGGTAATATAAATGGTTCCATAAAACTAATCGTCACTCCTGCTGCCTTAGTTTCTATTTCTGTTCTCCT
ACAAATCTGCAGTAGCAAAGGTTTACAAGAAAACCTTAAAGCTACAGGGATCTTTACAGATAAATCGAATCAGATAT
TACAGATCAAGTTACTTGGGATTTCTAATCCGGATATTCCTTCCATTCCAATGCAAGTGATAGCCACGGTTAGCTT
CCACACTCAACCAAGGAAATGTTAAGGTCACCGCTTCCATCGGTGGAATACAAGGATCCACTGATTTTAAAGTTACACAA
GAGGTATTAACCTCCATCGAAGTTTCTCCAGTTTACCTTCAATTCGAAAAGGACTAACTCAGAAATTTACGGcGATCGG
GATTTTACGGATAACTCCAAAAAGATATTCAAATCAAGTCACTTGGAATTTCTTTCAGCAATCGCAAGCGTGTCTA
ACTTAGATGATAATAAAGGCTGGGAAAAGCTCACGCTGTTGGAGACACGACTATTACCGCTACTTTAGGAAAAGTTTCA
GGTAAAACCTTGTTTACTGTAGTTCTGCGGTTCTCACTTCTATTCAAATCAATCCTGTAATCCTTCTCTTGCAAAGG
GTTAACTCAAAAATTTACGGCTACTGGGATCTACTGTACAACCTAACCAAGGACATTACTTCTCCGTTACTTGGTTCT
CATCCGATCTTCAATCGCAACAAATTTCAAACGCCAAAAAATCAAGGAACTCTTACGGAGCAGCTACAGGAGCAACG
GATATTAAGCCACATTCGGAAAGGTAAGTAGTCCAGTTTCTACGTTATCCGTTACTGCTGCAAACCTTGTGAAATACA
AATCACACCGCGCGCTGCTTCCAFAGCAAAGGAAATTTCCGAAAGATTAAAGCAACCGGATTTTTTACAGACAACTCTA
ATTCGGATATTACAAATCAGGTCACCTGGAGTTCATCTAATACAGATATTGCTGAAATTACAAATACCAGAGGAAGCAAA
GGTATTACAAATACACTCACTCCCGGATCGAGTGAAATATCCGcCGCTCTCGGTTCAATCAAAAGTTCTAAAGTAAATTT
GAAGTAACTCCGGCACAATTGATTTCCATTGCAGTAACACCTACAATCCATCAGTTGCAAAGGTTCAATACGACAAT
TTAAGCCACCGGAACATATACGGATCATTCGGTACAAGACGTGACTGCCCTAGCTACCTGGTCTTCTTCCAATCCCAGA
AAAGCAATGGTTAAACAACGTTACAGGTTCCGTTACAACAGTGGCTACCGGAAATACAAATATTAAGCAACGATAGACTC
CATATCCGGATCTTCCGTTTTGAAATGCTACTCCTGCCTTCTTACTTCTATCGAGATAACACCGACGATTAATCTATCA
CTCACGGTCTTACAAAACAATTTAAAGCGACTGGTATCTTTTCAGATAAATCTACTCAAATTTGACTCAGCTTGTAACT
TGGATTTCTCCGATCCCTCCAAGATCAAGATCGAAAATAACTCCGGTATAGCAACAGCTTCTGCATTAGGAAGTTCGAA
TATTACGGCCATCTACAAATTTGTCCAAAGTTCCCAATTCGATCACAGTCACTGACTTAAAAGTAAAAGTATAACTA
TCAGTCTTCTCAAGTTCAATAGCCAAAGGATTGACCCAACAATTTAAAGCGATCGGAACTTTTATAGATGGTCTGAA
CAAGAAATTACGAATCTTGTGACCTGGTATTCCTCCAATCCGATATTGTTCCATCAATAATTTCTGCGGGTAAAAAAGG
TTTAGCGaCCGCaCTCTCAATAGGTTCTTCCAACATCTCCGCAATTTACAATCTATAAGCAGTAATAAAATAAATTTTA
ATGTAAGCGCCGCGCTAGATTCCATTAATCAATCCAGTACAATAACATCGCCAAGGACTTACCAACAATAT
ACTGCGCTTGGCGTTTATTTCAGACTCCACCATTTCAGGACATCAGCGATTTAGTTACATGGTCCAGTTCCAATTTCTGACTC
GATCAGCATCTCCAATTCGACCGGAACCAAGGGAAAAGCGACCGCTTTACAGATTGGAAAGAGCAAAATTTACCGGACTT
ACAATTCATTTGAAAACATAAATCTAATGTCAGCGCAGCACTCTCTCTCGATTTTTATATCTCTACCAATACA
AATATAAACACCCCGTATCAAAAATTTCTTTGCAATGGGAACGTTATTCGGACGGAACCAACCGGATTTAATCTTCTC
GGTTACATGGTCCAGTTTCAAGCTCAAGCTCAAGCAAGGTTGAGTAACGCATCTGAAACGAAAGGATTGGTTACAGGGATTA
CTTTCGAAATCCTAATCACAGCGACCTACGGCTCAGTGTGGGAAATACAATTTCTCACAGTAAACAAAACCGACAG
ATAGCTCCGACGGTTAATCCGTAAGTTCTTATCACTACTACCATCCAAGTTGTATATTCAGAATCCAATAACATCA
GGAAGCCCTTGATTTATCCAATTAACAAATAAATTAATAGTTCCAATTTTTACGGACATTTGTTTCGGATAATACGGACTTCA
ATTCCAATTTCAAACCGCAGATTTTCTCTTAGTAGTATCAAAGGAAGTAAAAATACTTTTACGATTACACTTTCACAT

```

TCACAAATCTTAAACAAATCATACACACTTGTAGTCAACAAACAAGGAATTCACGATCTTTCTCCATTCCAAATTCCTT
AAGTTGTCCAAATAACTCTGATTTTATAGGAAAAGAAACAACCTCAAACCTACAAGTGCAGTTTGTAAATTCCTTAAACCAAG
TGATCGTTCTTTTTCCAAACCTTTATATCTGGAAAGGAAGTAAACAAAATCCGTGGAATGTTCAAATCCGTCCCAATGT
GAATCCAGATATAAATTTGCAGGTGTGTCTTCATTTGGGAAGTATTACGAGCGTTAGAAATTTAGATGGAAAAGTATGCGG
TGGAGCACCCGGCAGACTCCTCGAAAATATGTTTAAACACACTCCCTTCTCAATCAGGTGGTCAATATACGATCATCGCCG
CAAATGATTTGAACGGAGACGGCTTTGACAACAAATCCTGGGGAGCAATTCGAGATTCATTTCGATCAAGAAAACCTACAA
CCTTCTCCGAAAGATAGAATCAACTTTATAGGTTGTGGAATTCGCCCTCAACTTTATGGATGGCCCGATCGTGTGAGA
TCCTTTGGAGACGGTTCGATTTCCGGCTCTCTGTAGATTACAACAATCAAATCTATCTAGGACCGAATGTAAAAGGAA
ACCAAGCAGCTCGATTCAATTACGACGGAACCTTTCCGGAATCTATTTCTTTTCTTTTACCCAAGATAAAAATGCCACT
AACCGTCTCTTCAAGAGATGGAGGAATCCGGTTCGGAATACGTTACGATCGGTCATACCGGTTGTAATCTCAATAG
TGCAGACATCACTACTGGATGTGGTCCAGATAACGAAGATGGACGTGGGGTTTTTGGCCACCGGATCATTAGACAAAAAT
CTCATATTTTATAGCAGGTTCAAACCAAGGAGATTCAACTATCTCTATTATTCTCAGATACCGATACAAACCTTAAT
TTTTAAATATATCAGTATGGGAAAAATTAAGGATTTGGCGACTGCAGGAACCTCATCTATCGAGTTCTAGACGATCGGAT
CCATGTAGGTTTTGCAAAAAAAATCAAATCTAAACGCACCTGATTTCCGTAATAACCTTTAATACATCCGAGCACA
ATCGATGTGCAATTTAAACAACCTGTGAAGCCTCTGACGGATACCGCGTAATCGTTTGTAGAATCGATAGAATGCCATTAC
TTTGGCGCGGCTCCGTGGATGCAGTCAATTATAAACTCATAAATCTGATAATTCCTCGATCAACTGGGGTTATATGT
GGGAATAGATTCTCTATTCTGTTTTAAAGAAAACTTTACGCCGAAACGGAGGATTTCAAATTCATTACATAATGGAA
GTATAATACACTCTACAGTGCAAATCCTAGTCTTGTGAAGGAATCAATCGTTGTTCCAGTTGGAAGACACAGCACCT
AGATCCAATCCGAAGTGGCATACTCTCTCTATACCAATGGTTTTCACTGGAGCTTACAAAGTATCGAGATTTAATTC
GGCGGATAAAGCATTCTCTCAATTCGCAGAAATTAACGGAAGATTGTATGTAACAAGAACGATCTGTGTAACGAAAGAAG
ATCACTCCGGACTCAGACAAAGTTACAAACTTTGAAAGGTTGTACAGACGGAAGTATACAAATCGAAGACCTCAACTT
TGGAAATGTGATCCGACTCTAACCGCGGATACAACAACCTGCGAAGCAAAGATTGGTCTTTAGTAGGAGATAATGGAAC
CGGGTTACGAATTCGGAGACGATTCCAATCACAGTATGACGATGGTAGTTGCAAGTGGATCTATCTCTACGTAGGTT
TTGACAACGAAACGGAATTCAAATCTGGAGAACAAATCTGAAAATCCTGGAAGTTCATCACAGACTGGGAGCCTATA
GGAATAGGCGGATTAAGAGACGTTACCAATCGTCAAATTTATTCGGCTATATCCGGAATGAATTTGGTGTAAATTCGT
ATATATAAGCGTAGGAAATAAGATCAACCGGTTAAAATTTACAGACAACAGAACCAA

I. f) 序列标识编号: SEQ ID NO: 6

II) 序列特征:

II. a) 长度: 1886

II. b) 类型: 氨基酸

II. c) 拓扑结构: 线性

III) 类型: 肽

IV) 来源: N-末端至C-末端部分

MKRTFCISILLSMFFQSCMSWPLLTSLAGLAAGKKSNGLPFFHLLLSNSDPVITRIELSYQNSSIAKGTSTTLEVTAFD
NGTNQNITDSTSIVSDAQSIVDIQGNRVIRGASGSSIIKAEYNGMYSEQKITVTPATINSIQVTSLDDGILPKGTNRQFA
AIGIFSDGSHQDISNDPLIVWSSSNI DLVRVDDSGLASGINLGTAHIRASFQSKQASEEITVGDAVLSSIQVTSNSPNIP
LGKKQKLTATGIYSDNSNRDISSSVIWNSSNSTIANIQNNGIETADTGI VTVSASRGNINGSIKLIVTPALVSI SVSP
TNSAVAKGLQENFKATGIFTDMSNSDITDQVTWSDSSNPILSISNASDSHGLASTLNQGNVKTASIGGIQGSTDFKVTQ
EVLTSIEVSPVLPISIAKGLTQKFTAIGIFTDMSKKDITNQVTWSSSAIASVSNLDDNKGLGKAHAVGDTTITATLGKVS
GKTWFTVVPVAVLTSIQINPVNPSLAKGLTQKFTATGIYSDNSNKDITSSVTWFSDDSSIATI SNAKKNQGN SYGAATGAT
DIKATFGKVSSPVSTLSVTAAKLVEIQITPAAASKAKGISERFKATGIFTDMSNSDITNQVTWSSSNTDIAEITNTRGSK
GITNTLTPGSSEISAALGSIKSSKVIKVT PAQLISIAVTPNPSVAKGLIRQFKATGTYTDHSVQDV TALATWSSSNPR
KAMVNNVTGSVTTVATGNTNIKATIDSI SGSSVLNVT PALLTSIEITPTINSITHGLTKQFKATGIFSDKSTQNLTLQVLT
WISSDPSKIKIENNSGIATASALGSSNITAIYKFVQSSPIPIVTDLKLKSI TIS P SSSSIAKGLTQQFKAIGTFIDGSE
QEITNLVTWYSSKSDIVPINNSAGKKGLATALSIGSSNISAIYNSISSNKNFVNSAATLDSIKINPVNNSIAKGLTQQY
TALGVYSDSTIQDISDLVTWSSSNSDSISISNSTGTKGKATALQIGKSKI TATYNSISKNINLTVSAAATLSSIFISPTNT
NIAINTVSKQFFAMGTYS DGTKDLTSSVTWSSSNQAQAKVNSAETKGLVTGITSGNPIITATYGSVSGNTILT VNKDIT
IAPTQSVVLSPTTIQVLYSESI NNQEA LDLSNYKI INSSNFYGHCSNDNTDFNSNSQTADFSLSSIKGSKNFTFTL SH
SQILNKSYTLVVNQGHLDLSSI P NSLSCPNNSDFI GKEKLTSAVCNSLNQVIVSFSKPLYSGKEVTKSVECSNPSQC
ESRYKFAGVSSLSITSVRILDGKVC GGAPADSSKICLTHSLQSGGQY TIIAANDLNGDGF DNKSGWAIRDSFDQENLQ
PSPKDRINFICGNSPLNFM DGPVSDPFGDGSDFGLVDYNNQI YLGPVNVKGNQAARFN YDGF P P E S I F F S F T Q D K N A T
NRASSRDGGI PVPNYVTI GHTGCTLNSADI TTGCGPDNEDGRGVFATGSLDKKSHI FIAGSKP RRFNYLYSSD TD TNLN
FKYI SMGKITGLTAKTSSIAVLDDRIHVGFAKNQNLNAPDFGKITFNTSEHNRCAI VNNCEASDGYRGNRFRIDRMPY
FGGSDVAVNYKTHKSDNSINWGYVYGDLSL FVFEKLYAANGFPNLSHNHSIIHSTSANPSPCEGINRCSWSKDTAP
RSNPKWHNSPHTNWFSLTKYRDLI PADKAFSQFAEFNGRLYVTRTI CVTKEDHSGLRQSLQTLKCTDGSYTNRRPQL
WKCDPTLTGDTTTC EAKDWSLVGDNGTGFTNFGDSDNSHSM TMVVASGSYLVGFDNENGIQIWRTNLENPGSSSHDWEPI
GIGGLRDVTNRQIYSAISGMNFGVNFVYISVGNKQDPVKIYRQONQ

I. g) 序列的编号: SEQ ID NO: 7

II) 序列特征:

II. a) 长度: 1557个碱基对

II. b) 类型: 核酸

II. c) 链: 单链

II. d) 拓扑结构: 线性

III) 在基因组中的位置

III. a) 生物体: 问号状钩端螺旋体

III. b) 名称/关键词: CDS

III. c) 在图谱中的位置: (0) - (1557)

```
ATTACCGTTACACCAGCCATTCTTAACCTCAATTCAAGTTACGAGTTTAGAGTCAGGTATACTACCTAAAGGT
ACTAATCGTCAATTCTCAGCCATCGGTATCTTTTCGGATGGTTCATCAGGATATTTCCAACGAACCACTG
ATCGTTTGGTCTTCCAGTAATCCTGATTTGGTTCGAGTAGATGATTCAGGGTTGGCATCAGGGATCAATTTA
GGAACAGCTCATATTCGTGCATCCTTTCAATCAAACAAGGGGCTGAAGAAATGACCGTTGGAGATGCTGTT
CTCTCTCAAATCCAAGTAACTTCAAACGATCTGAATATTCCTCTCGGAAAAAACAACAACTAACAGCTACG
GGAATCTATTCGGATAACTCTAACAGGGATATTCCTCTCTGTTATTTGGAATTCTTCTAATTCCTACTATC
GCTAATATTCAAACAACGGAATATTAGAAACAGCTGATACTGGTATTGTCAGTGTCTTGCTTCTAGCGAG
AATATAATCGGATCCGTAAAACCTAATCGTTACTCCAGCAGCCTTAGTTTCTATTTCTGTTTCTCCGACAAAT
TCTACAGTTGCAAAAGGTTTACAAGAAAACCTTAAAGCTACAGGGATCTTTACAGATAATTCAAACTCGGAT
ATTACCGACCAAGTTACTTGGGATTCTTCTAATACCGATATTTCTCTCAATTTCCAATGCAAGTGATAGCCAC
GGATTAGCTTCCACACTCAACCAAGGGAATGTTAAAGTCACTGCTTCCATCGGTGGAATACAAGGATCCACT
GATTTTAAAGTTACACAAGCTGCATTGACTTCCATCGAAGTCTCTCCAACCTCGCACTTCCATTGCAAAAGGA
CTAACTCAAAGTTTACTGCGATCGGGATTTTTACGGATAACTCTAAGAAGGATATTACGGATCAAGTCACT
TGGAATTTCTTTCAGCAATCGTAAGCGTGTCTAACTTAGACAACAATAAAGGTCTGGGAAAACCAACTCA
GTTGGAAAACAGCACTATTACCGCAACCTTAGGAAAAGTTTCAGGTAACACTTGGTTTTACTGTAGTTTCTGCG
GTTCTCACTTCTATTCAAATCAATCCTGTAAATCCTTCTCTGCAAAAGGGTAACTCAAAAATTTACGGCT
ACTGGGATCTACTCTGACAACCTAACAAGGACATTACTCCGCTGTACGTGGTTCTCATCCGATTCTTCA
ATCGCGACGATTTCAAACGCCCAAAAAAATCAAGGAAACGCTTACGGAGCAGCTACAGGAGCAACGGATATT
AAAGCCACATTTCGAAAGGTAAGTAGTCCGTTTCTACGTTATCTGTTACAGCTGCAAAGCTTGTGAAATC
CAAATCACACCGGCTGCTGCTTCCAAAGCAAAGGGACTCACAGAAAGATTCAAGGCTACTGGTATCTTTACG
GATAACTCAAATTCGATATTACAAATCAAGTTACCTGGAATTCCTCTAATACGGATATTGCTGAAATTTAA
AATACCAGTGAAGTAAAGGTATTACAAATACACTCACTCCAGGA
```

I. h) 序列编号: SEQ ID NO: 8

II) 序列特征:

II. a) 长度: 519

II. b) 类型: 氨基酸

II. c) 拓扑结构: 线性

III) 类型: 肽

IV) 来源: N-末端至C-末端部分(0) - (519)

ITVTPAILNSIQVTSLESGILPKGTNRQFSAIGIFSDGSHQDISNEPLIVWSSSNPDLVRVDDSGLASGINL
 GTAHIRASFQSKQGAEMTVGDAVLSQIQVTSNDLNIPLGKKQKLTATGIYSDNSNRDISSSVIWNSSNSTI
 ANIQNNGILETADTGIVTVSASSENIIGSVKLIIVTPAALVSI SVSPTNSTVAKGLQENFKATGIFTDNSNSD
 ITDQVTWSSNTDILSISNASDSHGLASTLNQGNVKVTASIGGIQGSTDFKVTQAALTSIEVSPTRTSIAGK
 LTQKFTAIGIFTDNSKKDITDQVTWNSSSAIVSVSNLDNNKGLGKTNSVGNTTITATLGKVSNTWFTVVP
 VLTSIQINPVNPSLAKGLTQKFTATGIYSDNSNKDITSAVTFWSSDSSIATISNAQKNQGNAYGAATGATDI
 KATFGKVSSPVSTLSVTAAKLVEIQITPAAASKAKGLTERFKATGIFTDNSNSDITNQVTWNSSNTDIAEIK
 NTSGSKGITNTLTPG

I. i) 序列编号: SEQ ID NO: 9

II) 序列特征:

II. a) 长度: 600个碱基对

II. b) 类型: 核酸

II. c) 链: 单链

II. d) 拓扑结构: 线性

III) 在基因组中的位置

III. a) 生物体: 问号状钩端螺旋体

III. b) 名称/关键词: CDS

III. c) 在图谱中的位置: (0) - (600)

CATAACTCTCCTCATAACAATTGGTTTTCACTGGAGCTTACAAAGTATCGGAATTTAATTCGGCGGATAAA
 GCATTCTCTCAATTCGCAGAATTTAACGGAAGATTGTATGTAACAAGAACGATCTGCGTAACGAAAGAAGAT
 CACTCCGGACTCAGACAAAGTTTACAAACTGTGGAAGGTTGTACGGACGGAAGTTATACAAATCGAAGACCC
 CAACTTTGGAAATGTGATCCGACTCTAACC GCGATACAACAACCTGCGAAGCAGAAGATTGGTCTTTAGTA
 GGAGATAACGGAACCGGATTTACAAACTTTGGAGACAAATCCAATCACAGTATGACGATGATGGTTGCAAGT
 GGATCTTATCTCTACATAGGTTTTGATAACGAAAACGGAATTCAAATCTGGAGAACAATCTTGAAAATCCT
 GGAAGTTCATCACACAACTGGGAACCTATAGGAATAGCGGATTAAGAGACGTTACCAATCGTCAAATTTAT
 TCGGCTATATCCGGAATGAATTTGGTGTAATTTTCGTATATATAAGCGTAGGAAACAAAAATAAACCGGTC
 AAAATTTACAGACAACAGAATCAA

I. j) 序列编号: SEQ ID NO: 10

II) 序列特征:

II. a) 长度: 200

II. b) 类型: 氨基酸

II. c) 拓扑结构: 线性

III) 类型: 肽

IV) 来源: N-末端至C-末端部分(0) - (200)

HNSPHNNWFSLELTKYRNLI PADKAFSQFAEFNGRLYVTRTICVTKEDHSGLRQSLQTVEGCTDGSYTNRRP
QLWKCDPTLTGDTTCEAEDWSLVGDNGTGFTNFGDNSNHSMTMMVASGSYLYIGFDNENGIQIWRTNLENP
GSSSHNWEPIGIGGLRDVTNRQIYSAISGMNFGVNFVYISVGKNKPKVKIYRQONQ

图1

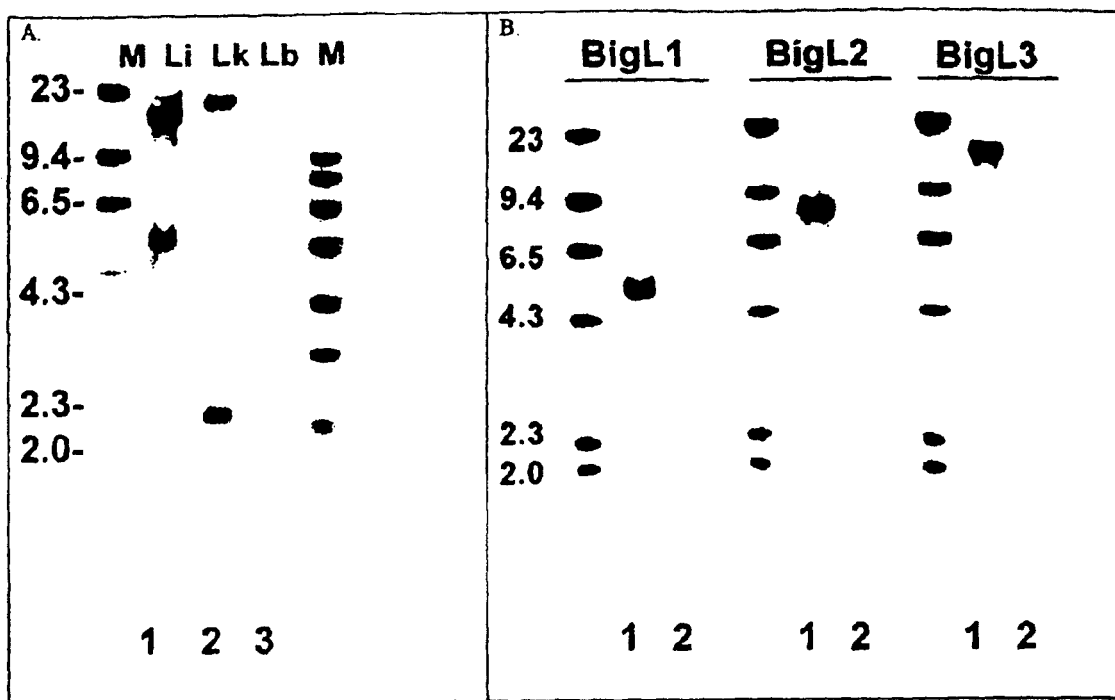


图2

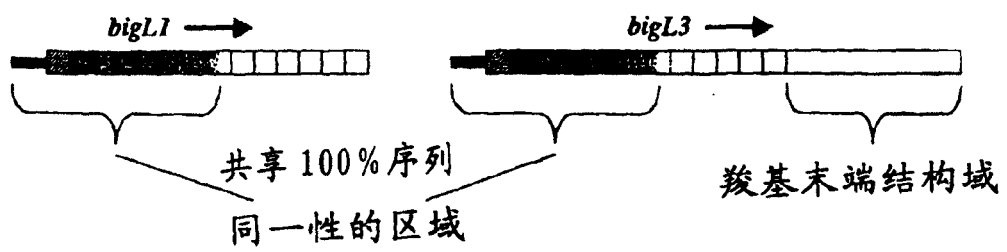


图3

1kb梯度(NE Biolabs)
 用 HaeIII 消化的 Φ X174RF
 L. kirschneri 流感伤寒血清型(RM52), 低传代
 问号状钩端螺旋体 sv lai (L391)
 圣地罗斯钩端螺旋体 sv bakeri (LT79)
 问号状钩端螺旋体布拉迪斯发血清型(AS-05)
 沃尔氏钩端螺旋体 sv biflexa(codice)
 L. kirschneri 流感伤寒血清型(RM52), 高传代
 博氏钩端螺旋体 sv hardjo (HB-15B/93U)
 L. kirschneri sv mozdok(5621)
 双曲钩端螺旋体 sv patoc(PatocI)
 问号状钩端螺旋体波莫纳血清型(RZ11)
 野口氏钩端螺旋体 sv proechymis(LT796)
 博氏钩端螺旋体 sv tarassovi (第11号)
 无模板

图 4

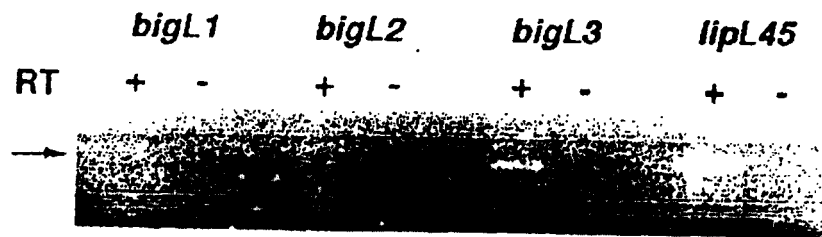
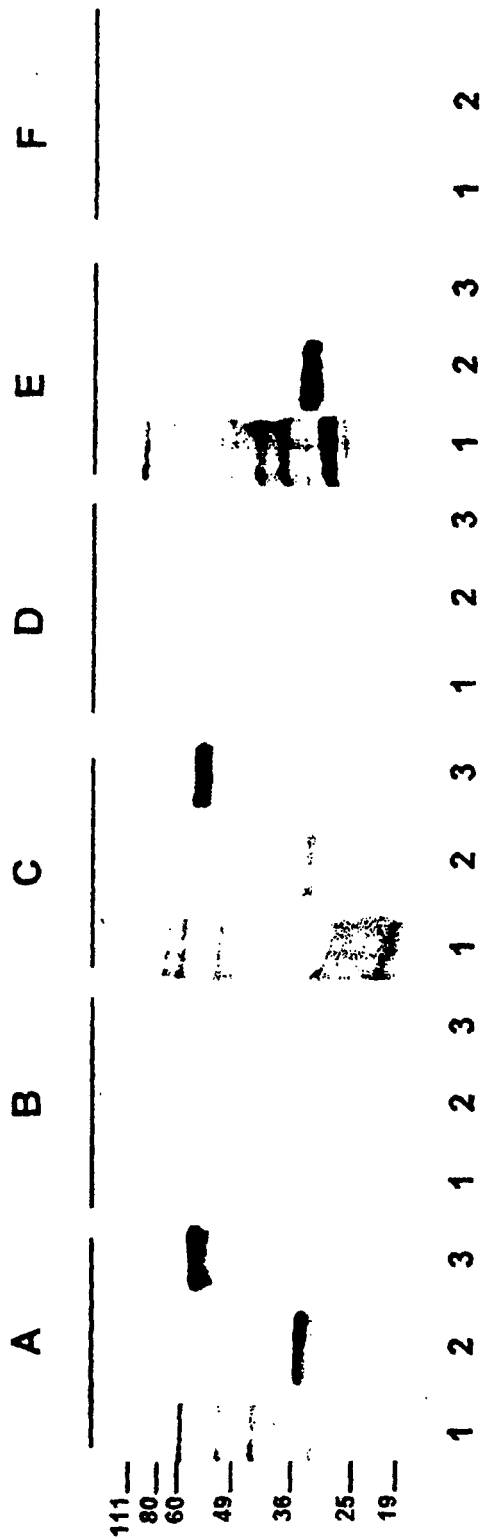


图5



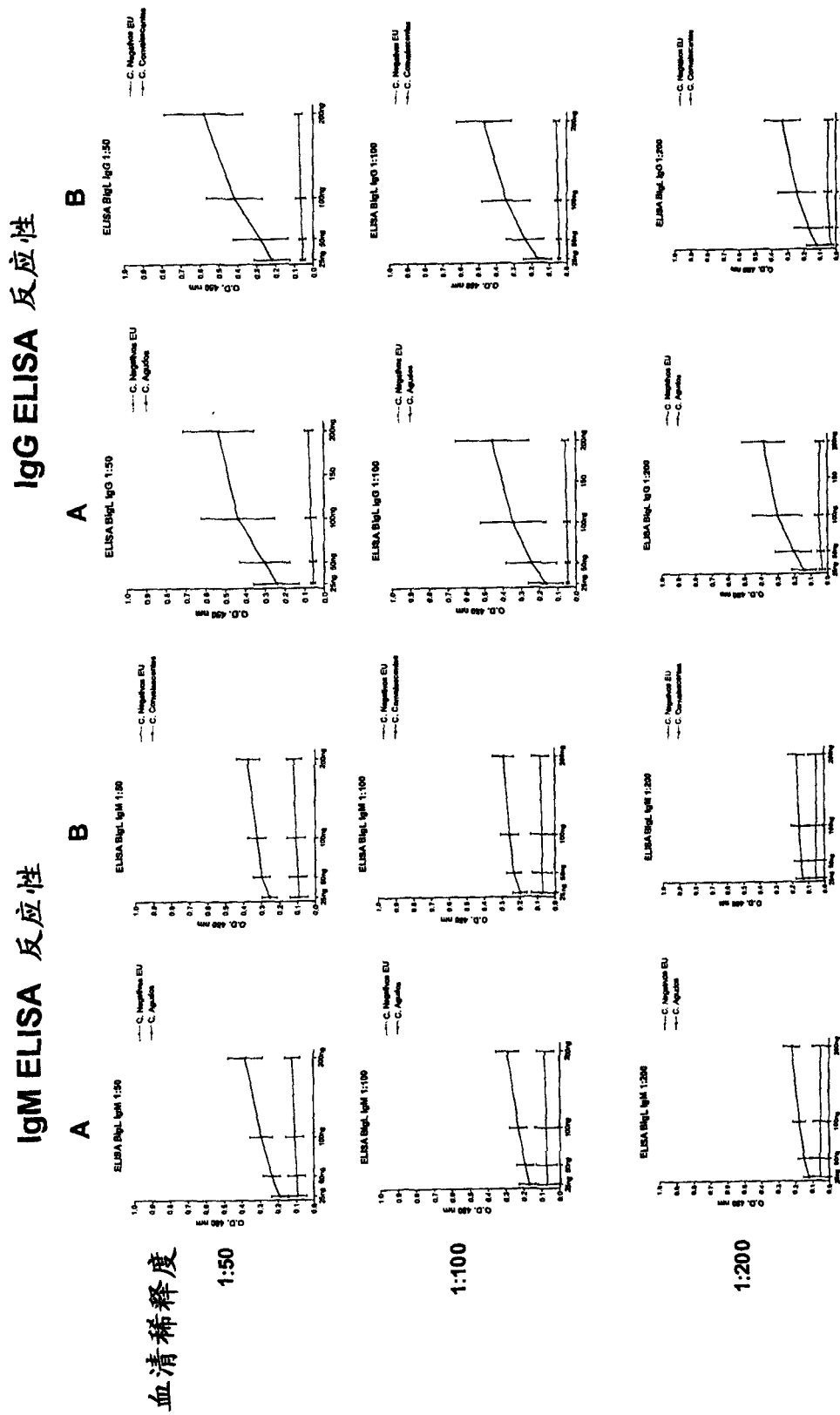


图6

图7

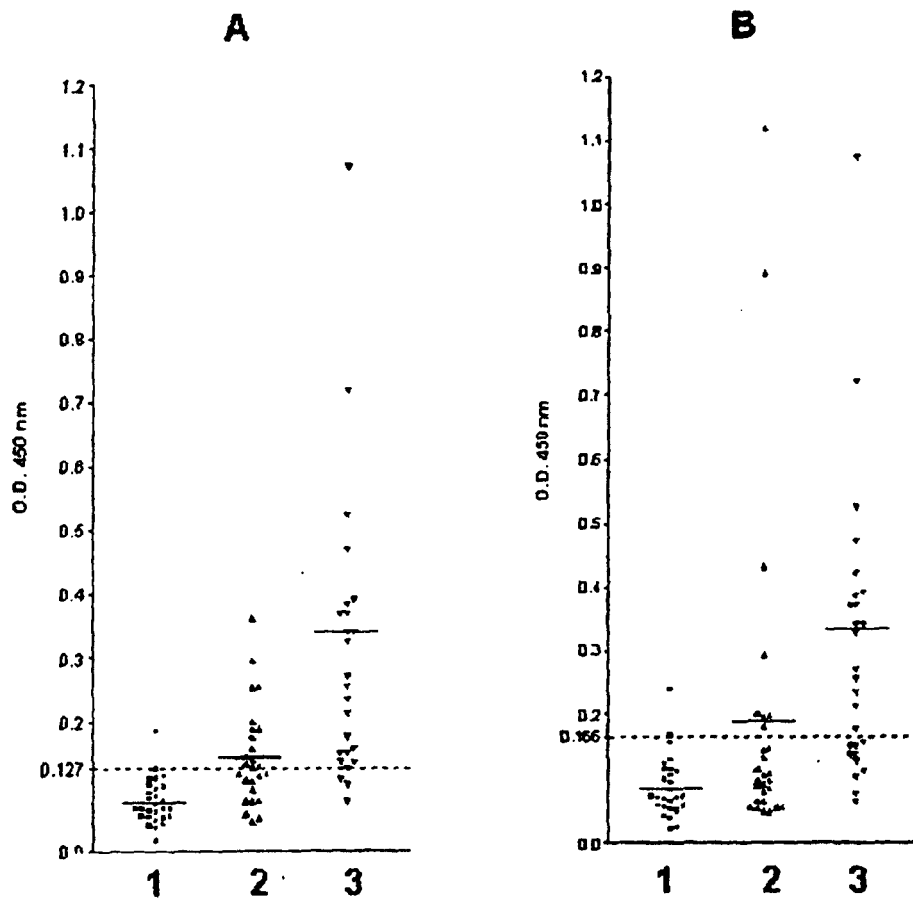
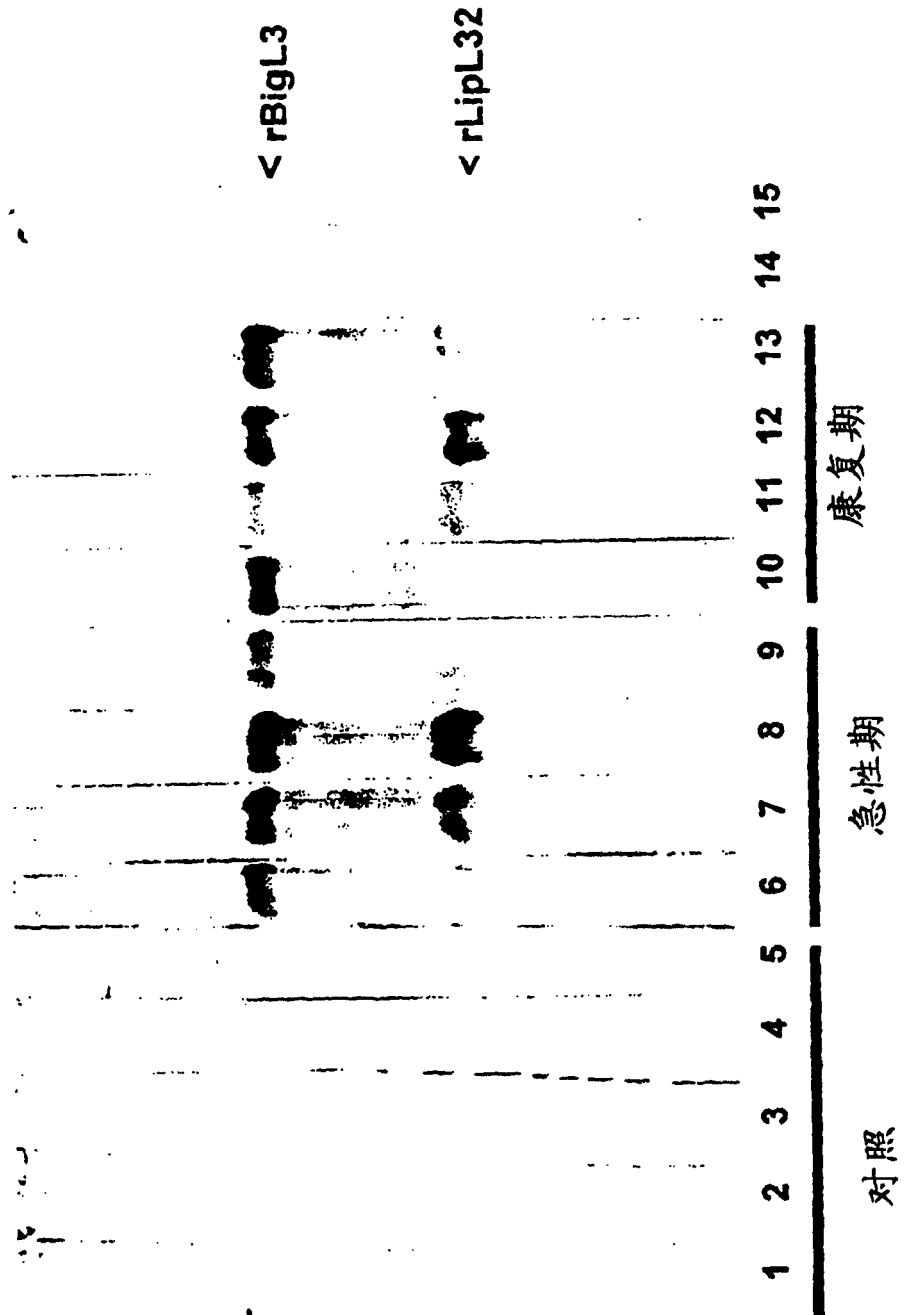


图 8



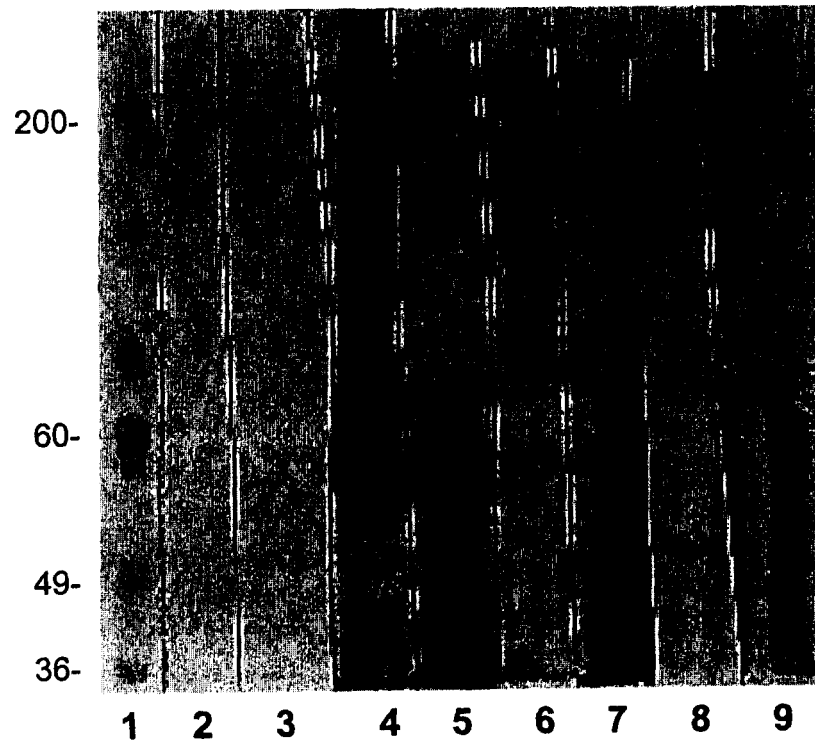
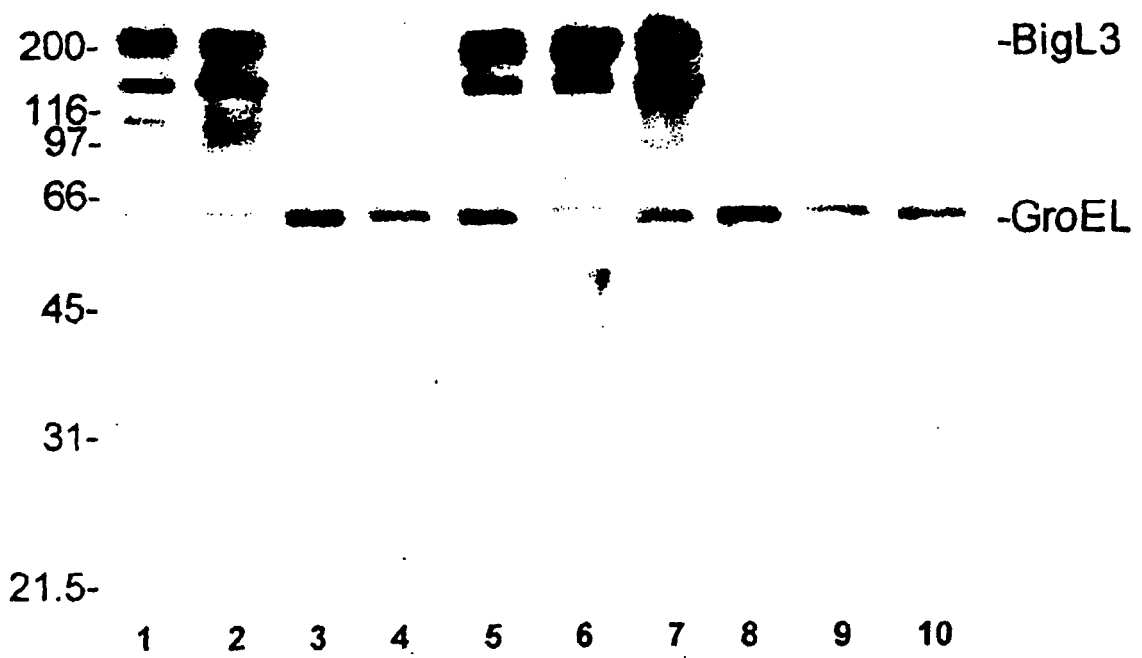


图9

图10



专利名称(译)	钩端螺旋体内存在的具有重复性细菌Ig样(BIG)结构域的蛋白质		
公开(公告)号	CN1639570A	公开(公告)日	2005-07-13
申请号	CN02829305.3	申请日	2002-05-20
[标]申请(专利权)人(译)	奥斯瓦尔多克鲁兹基金会		
申请(专利权)人(译)	奥斯瓦尔多克鲁兹基金会		
当前申请(专利权)人(译)	奥斯瓦尔多克鲁兹基金会		
[标]发明人	A I 科 M 加尔瓦雷斯 J H 罗萨克罗达 I C 斯奎拉 D A 哈克 J 马苏纳加 L W 赖利 M巴罗奇 TA扬		
发明人	A·I·科 M·加尔瓦雷斯 J·H·罗萨克罗达 I·C·斯奎拉 D·A·哈克 J·马苏纳加 L·W·赖利 M·巴罗奇 T·A·扬		
IPC分类号	C07H21/04 C07K14/20 C12N1/21 G01N33/53 C07K17/00		
CPC分类号	A61P31/04 A61P37/04 C07H21/04 C07K14/20 C12Q1/689 Y02A50/56 A61K2039/505 C07K16/1207 C12Q1/68 G01N33/56911 G01N2469/20		
代理人(译)	刘晓东		
优先权	10/147299 2002-05-17 US		
外部链接	Espacenet SIPO		

摘要(译)

本发明涉及编码钩端螺旋体细菌中具有重复细菌Ig样(Big)结构域的蛋白质BigL1、BigL2和BigL3的三种分离DNA分子以及它们在诊断、治疗和疫苗应用中的用途。依照本发明，编码BigL1、BigL2和BigL3蛋白质的分离分子可用于诊断和预防能在人类和其它哺乳动物体内引起疾病的钩端螺旋体感染，包括具有兽医学重要性的那些钩端螺旋体的感染。

