

[19] 中华人民共和国国家知识产权局



[12] 发明专利说明书

专利号 ZL 03822208.6

[51] Int. Cl.

C12N 1/11 (2006.01)
C12N 15/30 (2006.01)
A61K 39/002 (2006.01)
A61K 39/005 (2006.01)
A61K 39/008 (2006.01)
A61K 39/012 (2006.01)

[45] 授权公告日 2008 年 6 月 11 日

[11] 授权公告号 CN 100393866C

[51] Int. Cl. (续)

A61K 39/015 (2006.01)

A61K 39/018 (2006.01)

A61P 33/02 (2006.01)

A61P 33/04 (2006.01)

A61P 33/06 (2006.01)

[22] 申请日 2003.9.19 [21] 申请号 03822208.6

[30] 优先权

[32] 2002. 9. 20 [33] EP [31] 02078953.3

[86] 国际申请 PCT/EP2003/010696 2003.9.19

[87] 国际公布 WO2004/026903 英 2004.4.1

[85] 进入国家阶段日期 2005.3.18

[73] 专利权人 英特威国际有限公司

地址 荷兰博克斯海尔

[72] 发明人 N·F·J·范泊佩尔

A·N·沃梅伦 T·C·沙普

[56] 参考文献

WO0066154A2 2000.11.9

WO9837185A2 1998.8.27

a low - background inducible promoter system
in Leishmaniadonovani. Yan shaofeng. Molecular
and biochemical parasitology, Vol. 119 No. 2.
2002

审查员 杨振宇

[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 罗菊华

权利要求书 2 页 说明书 46 页 附图 28 页

[54] 发明名称

减毒活寄生虫疫苗

[57] 摘要

本发明特别地涉及顶复门 (Apicomplexa) 和锥虫科 (Trypanosomatidae) 减毒活寄生虫, 涉及在疫苗中和这种疫苗制备中这种减毒活寄生虫的用途。而且本发明涉及包含这种减毒活寄生虫的疫苗, 和这种疫苗的产生方法。最后, 本发明涉及特异性 tet 阻抑蛋白融合蛋白和涉及包含这种 tet 阻抑蛋白融合蛋白的本发明减毒活寄生虫。

1、 顶复门 (*Apicomplexa*) 或锥虫科 (*Trypanosomatidae*) 减毒活寄生虫, 其特征在于: 所述寄生虫包含诱导型启动子控制下的核糖体蛋白质基因。

2、 根据权利要求 1 所述的减毒活寄生虫, 其特征在于: 所述寄生虫属于球虫亚纲 (*Coccidia*), 梨浆虫目 (*Piroplasmida*) 或血孢子虫目 (*Haemosporida*)。

3、 根据权利要求 2 所述的减毒活寄生虫, 其特征在于: 所述寄生虫属于艾美球虫科 (*Eimeridiidae*), 隐孢子虫科 (*Cryptosporidiidae*) 或肉孢子虫科 (*Sarcocystidae*)。

4、 根据权利要求 3 所述的减毒活寄生虫, 其特征在于: 所述寄生虫属于艾美球虫属 (*Eimeria*), 隐孢子虫属 (*Cryptosporidium*), 弓形体属 (*Toxoplasma*), 肉孢子虫属 (*Sarcocystis*) 或新孢子虫属 (*Neospora*)。

5、 根据权利要求 2 所述的减毒活寄生虫, 其特征在于: 所述寄生虫属于巴贝虫科 (*Babesiidae*) 或泰勒尔梨浆虫科 (*Theileriidae*)。

6、 根据权利要求 5 所述的减毒活寄生虫, 其特征在于: 所述寄生虫属于巴贝虫属 (*Babesia*) 或泰勒尔梨浆虫属 (*Theileria*)。

7、 根据权利要求 2 所述的减毒活寄生虫, 其特征在于: 所述寄生虫属于疟原虫属 (*Plasmodium*)。

8、 根据权利要求 1 所述的减毒活寄生虫, 其特征在于: 所述寄生虫属于锥虫属 (*Trypanosoma*) 或利什曼原虫属 (*Leishmania*)。

9、 根据权利要求 1 至 8 任一权利要求所述的减毒活寄生虫, 其特征在于: 所述诱导型启动子基于操纵基因位点和能可逆结合所述操纵基因位点的阻抑蛋白。

10、 根据权利要求 1 至 8 任一权利要求所述的减毒活寄生虫, 其特征在于: 所述诱导型启动子是可由抗生素诱导的。

11、 根据权利要求 10 所述的减毒活寄生虫, 其特征在于: 所述诱导型启动子是可由四环素或无水四环素或它们衍生物诱导的。

12、根据权利要求 11 所述的减毒活寄生虫，其特征在于：tetR 系统被用做诱导型启动子。

13、根据权利要求 1 至 8 任一权利要求所述的减毒活寄生虫，其特征在于：所述核糖体蛋白质基因是编码 L9，S3，质体-S9 或 S13 的基因。

14、根据权利要求 13 所述的减毒活寄生虫，其中所述核糖体蛋白质基因是编码地鼠弓形体 (*Toxoplasma gondii*) L9，S3，质体-S9 或 S13 的基因。

15、对抗寄生虫感染的疫苗，其特征在于：所述疫苗包含根据权利要求 1 至 13 任一权利要求所述的减毒活寄生虫和药物学上可接受的载体。

16、根据权利要求 1 至 13 任一权利要求所述的减毒活寄生虫用于制备对抗顶复门 (*Apicomplexa*) 或锥虫科 (*Trypanosomatidae*) 寄生虫所引起感染的疫苗的用途。

17、产生权利要求 15 所述疫苗的方法，所述方法包括混合根据权利要求 1 至 13 任一权利要求所述的减毒活寄生虫和药物学上可接受的载体。

18、编码包含 tet 阻抑蛋白和异源蛋白或其部分的 tet 阻抑蛋白融合蛋白的 DNA 片段，所述异源蛋白或其部分融合到 tet 阻抑蛋白 N 末端侧，其中单体形式的所述融合蛋白具有小于 60 kD 的分子量，并且无 GPI 锚，分泌/外泌信号和跨膜区。

19、权利要求 1 至 8 任一权利要求所述的减毒活寄生虫，其特征在于：所述寄生虫包含 tet 操纵基因位点和权利要求 18 所述编码 tet 阻抑蛋白融合蛋白的 DNA 片段。

20、权利要求 19 所述的减毒活寄生虫，其特征在于：所述寄生虫包含两个或多个 tet 操纵基因位点。

减毒活寄生虫疫苗

技术领域

本发明涉及顶复门 (*Apicomplexa*) 和动基体目 (*Kinetoplastida*) 减毒活寄生虫, 涉及在疫苗中和这种疫苗制备中这种减毒活寄生虫的用途, 涉及包含这种减毒活寄生虫的疫苗, 涉及这种疫苗的产生方法, 涉及特异性 tet-阻抑蛋白融合蛋白和涉及包含这种 tet-阻抑蛋白融合蛋白的减毒活寄生虫。

背景技术

在原生动物界中, 已知顶复门 (*Apicomplexa*) 和动基体目 (*Kinetoplastida*), 更具体地锥虫科 (*Trypanosomatidae*) 包含几种出名地致病寄生虫。

锥虫科 (*Trypanosomatidae*) 包含属于利什曼原虫属 (*Leishmania*) 和锥虫属 (*Trypanosoma*) 的寄生虫。

利什曼病是利什曼原虫属 (*Leishmania*) 引起的各种疾病表现的术语。在狗和人类中最常常发生这种疾病。该寄生虫由白蛉传播到哺乳动物宿主, 并且在世界上所有热带和亚热带地区流行。在宿主中, 巨噬细胞摄取寄生虫, 在巨噬细胞中寄生虫停留和繁殖, 引起慢性发炎过程。临床上, 狗中该疾病的特征是: 体重减轻, 贫血, 发热和淋巴结病。常常观察到皮肤损伤。在人类中, 多种利什曼原虫属 (*Leishmania*) 物种具有感染性, 它们中最致病性的是婴儿利什曼原虫 (*L. infantum*), 引起严重的内脏利什曼病 (也称为黑热病), 感染脾脏、肝脏和骨髓, 并且如果不治疗将是致命的。其它致病性利什曼原虫属 (*Leishmania*) 物种是即硕大利什曼原虫 (*L. major*) 和墨西哥利什曼原虫 (*L. mexicana*)。

已知多种锥虫物种, 其在人类和动物中引起各种不同的疾病。特别地, 已知两种锥虫物种是致病性的: 布氏锥虫 (*Trypanosoma brucei*) 和克氏

锥虫 (*Trypanosoma cruzi*)。

布氏锥虫 (*T. brucei*) 物种存在于非洲国家, 并且在人类中引起昏睡病和动物 (牛、马、猪) 中引起非洲锥虫病。布氏锥虫 (*T. brucei*) 由舌蝇传播, 将锥鞭毛体形式递送到宿主中。

克氏锥虫 (*T. cruzi*) 物种主要存在于南美, 该寄生虫具有广泛的宿主范围 (包括家畜和野生动物), 但是以在人类中引起查加斯病而著名。该寄生虫是由锥鼻虫 (象红猎蝽属 (*Rhodnius spp.*) 和锥蝽属 (*Triatoma spp.*)) 传播的。后循环的锥鞭毛体期感染宿主, 并且不同于布氏锥虫 (*T. brucei*), 将在不同细胞类型的宿主胞质内部繁殖。宿主细胞破裂后, 释放新的锥鞭毛体形式, 锥鼻虫可以再次摄取它。

顶复门 (*Apicomplexa*) 包含艾美球虫科 (*Eimeriidae*) 寄生虫。许多不同的艾美球虫属 (*Eimeria*) 物种存在于广泛种类的哺乳动物和鸟类中。感染鸡胃肠道的 7 种主要物种是禽艾美球虫 (*Eimeria tenella*), 尼氏艾美球虫 (*E. necatrix*), 波氏艾美球虫 (*E. brunetti*), 巨型艾美球虫 (*E. maxima*), 堆形艾美球虫 (*E. acervulina*), 塞前艾美球虫 (*E. praecox*) 和和缓艾美球虫 (*E. mitis.*)。这些艾美球虫属 (*Eimeria*) 物种都参与家禽球虫病。这使得艾美球虫属 (*Eimeria*) 成为家禽中最重要的寄生性疾病的原因, 给农民造成了巨大的经济损失。艾美球虫属 (*Eimeria*) 感染肠上皮细胞和粘膜下层, 引起严重的出血性肠炎, 这在年幼的鸟类中引起高死亡率。

这种疾病具有广泛的分布, 并且是现代家禽产业中最常记载的影响家禽的疾病。

还已知肉孢子虫科 (*Sarcocystidae*), 包括弓形体属 (*Toxoplasma*), 肉孢子虫属 (*Sarcocystis*) 和新孢子虫属 (*Neospora*), 有致病成员。

弓形体属 (*Toxoplasma*) 是广泛分布的寄生性感染, 存在于几乎所有哺乳动物中, 特别是山羊、绵羊和猪中, 也存在于人类中。在人类种群中流行可以高达总种群的 70%。常常由于食用寄生虫污染的煮得欠熟的肉类而发生这种感染, 也可能由于摄入分布在最终宿主猫粪便中的卵囊而发生感染。当动物或人类在怀孕期间被感染时, 它可以引起自发性流产或发

育中胎儿的先天性弓形体病。这可以引起神经病学后遗症或眼睛障碍。在免疫受损患者中可以发生慢性和致命的感染（脑炎）。

新孢子虫属（*Neospora*），特别是犬新孢子虫（*N. caninum*）是与弓形体属（*Toxoplasma*）非常相似的球虫寄生虫。然而，与弓形体属（*Toxoplasma*）相比，新孢子虫属（*Neospora*）有狗做为最终宿主。犬新孢子虫（*N. caninum*）诱导其中间宿主流产，并且可以在牛中引起严重的流产发作。怀疑另一种新孢子虫属（*Neospora*）物种，洪氏新孢子虫（*N. hughesi*）在马中引起马原虫脑脊髓炎。

许多肉孢子虫属（*Sarcocystis*）物种存在于牛、猪、绵羊、山羊和马中。经济地，认为 *Sarcocystis neurona* 是马中最常见的临床马原虫脑脊髓炎的原因。在美国，50% 马对 *S. neurona* 是血清反应阳性的。

疟原虫属（*Plasmodium*）属于血孢子虫目（*Haemosporida*），并且已知是疟疾的原因，其由蚊子传播。在人类中描述了 4 种疟原虫属（*Plasmodium*）物种，其中恶性疟原虫（*P. falciparum*）是最致病性和致命的。估测 4 亿人被感染，每年引起 2 百万人死亡。最初的临床症状是间歇发烧。最初感染后，疟原虫属（*Plasmodium*）寄生于红细胞，常常引起贫血。寄生的红细胞被隔离在内脏毛细管中，因此引起组织缺氧。这在大脑中特别严重，产生多发性点状出血，引起可能是致命的水肿和昏迷。尽管主要描述了人类中的疟原虫属（*Plasmodium*）物种，但是其它疟原虫属（*Plasmodium*）物种可以感染广泛种类的脊椎动物。

巴贝虫属（*Babesia*）和泰勒尔梨浆虫属（*Theileria*）都属于梨浆虫目，有寄生虫物种感染许多哺乳动物物种，并且引起各种不同的疾病。巴贝虫属（*Babesia*）物种由蜱传播，并且感染广泛的脊椎动物，引起称为巴贝虫病的疾病。该疾病以倦怠、贫血和寄生物血症为特征，在感染动物中引起多器官功能障碍。在晚期出现血红蛋白尿。牛中重要的巴贝虫属（*Babesia*）物种包括牛巴贝虫（*B. bovis*），分歧巴贝虫（*B. divergens*），大巴贝虫（*B. major*）和二联巴贝虫（*B. bigemina*）。在狗中，犬巴贝虫（*B. canis*），罗西巴贝虫（*B. rossi*），微小巴贝虫（*B. microti*）和吉氏巴贝虫（*B. gibsoni*）主要引起巴贝虫病，并且是常见的死亡原因。

已经报道有些巴贝虫属 (*Babesia*) 物种, 象分歧巴贝虫 (*B. divergens*) 和, 微小巴贝虫 (*B. microti*) 也感染人类。

泰勒尔梨浆虫属 (*Theileria*) 是蜱传播的疾病, 感染反刍动物, 并且是牛中的主要问题。泰勒尔梨浆虫属 (*Theileria*) 在白细胞和红细胞中感染和发育。病理学主要是由于白细胞内的阶段。在牛中应该区分两种主要的泰勒尔梨浆虫属 (*Theileria*) 物种, 小泰勒尔梨浆虫 (*T. parva*) 和环状泰勒尔梨浆虫 (*T. annulata*)。小泰勒尔梨浆虫 (*T. parva*) 引起东海岸热, 一种致死的牛疾病, 是各个非洲国家的地方性流行病。东海岸热以高烧、淋巴结病、严重肺水肿和消瘦为特征。环状泰勒尔梨浆虫 (*T. annulata*) 感染牛和水牛, 首先侵袭淋巴系统细胞, 后来做为内部红细胞形式出现在外周血中。环状泰勒尔梨浆虫 (*T. annulata*) 的感染通常称为热带泰勒尔梨浆虫病。这种疾病开始于高烧和淋巴结肿大, 接种倦怠、加速的脉搏、呼吸速率和厌食。在该疾病末期, 观察到贫血, 并且最后发生死亡。在马中, 马巴贝虫 (*Babesia equi*) (也被再命名为马泰勒尔梨浆虫 (*Theileria equi*)) 也是主要的病原体。

清楚的是, 这些年来已经研究了不同途径抗这些寄生虫的攻击。

与寄生性感染斗争的途径之一是药物学成分的应用, 诸如抗球虫剂 (anticoccidials) 的广泛应用, 今天抗球虫剂是家禽球虫病治疗中非常普通的治疗。另一种途径毫无怀疑地是接种。特别地, 在存在日益增加抗议抗生素应用的不满情况下, 需要新的和有效的疫苗, 特别是提供广泛保护的疫苗。

目前, 两种不同的方法被用在抗寄生性感染的接种中: 用活的减毒疫苗接种和用灭活 (杀死) 疫苗接种。如下面所概述的, 这两种方法都有它们的优点和缺点。

减毒疫苗的主要优点是它接近地模拟了天然感染: 它们激活所有阶段的免疫系统, 它们能诱导体液 IgG 和局部 IgA, 它们诱导了对许多保护性抗原的免疫应答, 它们提供了更持久的免疫性和更多的交叉反应。而且, 它们花费低, 并且在大多数情况下, 它们提供了快速免疫性。

减毒疫苗的缺点是发现正确水平减毒作用的困难和回复毒力的可能

性(这些是主要缺点),与受接种者接触的传播和免疫受损人和动物中的问题。

灭活疫苗的优点是如果提供加强,它们就提供了足够的体液免疫性,它们不显示突变和回复(大的优点),它们可以使用于免疫缺陷患者,并且原则上它们是安全的。

灭活疫苗的缺点是:它们常常不诱导(细胞)免疫性,需要加强,它们不提供局部免疫性(重要的),它们更昂贵,并且如果灭活低于100%,它们的应用是危险的。

然而,抗寄生虫疫苗的开发是复杂的,但愿是因为与其它微生物比较时,寄生虫的复杂性本身。接下来,即使在顶复门(*Apicomplexa*)和锥虫科(*Trypanosomatidae*)中的各种寄生虫,尽管它们相关,但是它们在其基因组成方面没有足够的相似性以配给同样地应用于所有这些寄生虫的共同减毒位点或失活方法。而且,为了制备减毒活疫苗,必需定位每个寄生虫的适合的减毒靶。为了产生死菌疫苗,人们需要知道通过灭活方法必须保持每个寄生虫的哪种抗原不变。此外,迄今为止,已经证明许多灭活的寄生虫疫苗无效。最后,存在各种不同的感染途径,不同的宿主,宿主中不同的宿主细胞,和生活周期期间甚至经常的宿主改变,这种宿主改变是大多数寄生虫的特征,并且一种寄生虫的宿主改变不同于另一种寄生虫的宿主改变。这也复杂化了疫苗的开发。

因此,迄今为止,抗寄生性感染的疫苗开发是困难的,费时和不是非常成功的。

发明内容

本发明的目的是提供抗顶复门(*Apicomplexa*)和锥虫科(*Trypanosomatidae*)寄生虫引起感染的疫苗,该疫苗联合了死疫苗和活减毒疫苗的大多数优点,并且几乎完全没有这些疫苗的缺点。而且,产生这种疫苗的方法普遍地可以应用于顶复门(*Apicomplexa*)和锥虫科(*Trypanosomatidae*)寄生虫。

所有顶复门(*Apicomplexa*)和锥虫科(*Trypanosomatidae*)寄生虫

的生活周期中,存在至少一个时刻,在该时刻中某阶段感染宿主细胞和开始分裂。现在惊奇地发现如果能在此感染时刻或在此感染时刻左右终止核糖体合成,寄生虫仍然进入宿主细胞,并且利用现有的核糖体库分裂几次,因此完美地模拟了天然感染。然而最后,几轮分裂后,后代寄生虫由于核糖体的缺乏而死亡。

其优点有:以最天然途径触发了感染后免疫应答的诱导,好象发生了强毒感染,但是与天然情况相反,该寄生虫将在一段时间后不可避免地灭绝。通过将核糖体蛋白质基因放置于诱导型启动子控制下达到了该目的。

诱导型启动子是可以被有意地开关的启动子。下面将给出这种启动子的例子。因为原则上所有核糖体蛋白质都是稳定、完全功能的核糖体合成所需要的,所以原则上,每种核糖体蛋白质基因都可以用做靶。所有顶复门 (*Apicomplexa*) 和锥虫科 (*Trypanosomatidae*) 寄生虫都有胞质核糖体,并且它们大多数都有质体核糖体和/或线粒体核糖体。所有这些核糖体都是寄生虫正常发育所必需的。因此,通过将核糖体蛋白质基因放置于诱导型启动子控制下可以获得本发明的活减毒寄生虫,而不论是否该核糖体蛋白质基因编码将掺入到质体、线粒体或胞质核糖体中的核糖体蛋白质。

各种寄生虫之间的核糖体蛋白质序列是高度保守的。因此,下面提供的核糖体序列的 DNA 探针可以用于检测顶复门 (*Apicomplexa*) 和锥虫科 (*Trypanosomatidae*) 每个寄生虫中类似核糖体蛋白质。此外,在 NCBI 蛋白质数据库 (<http://www.ncbi.nlm.nih.gov>) 中可以发现许多不同寄生虫的许多核糖体蛋白质基因序列。

缺少一种核糖体蛋白质已经可以干扰稳定核糖体形成的事实已经在各种植物、动物和微生物中得到了证明。只是做为例子:在果蝇属 (*Drosophila*) 中,已经证明 80 种核糖体蛋白质中一些核糖体蛋白质的突变产生典型表型,例如,细和短的刚毛,延迟的发育,和杂合子中雌性半不育性及纯合致死性。在例如核糖体蛋白质 S13 和 L9 中已经观察到伴随着突变的这种表型,称为微小表型 (Schmidt, A., Hollmann, M., Schäfer, U., Mol. Gen Genet. 251: 381-387 (1996), Sæbøe-Larssen

& S. , Lambertsson, A. , Genetics 143: 877-885 (1996))。另一个例子是酵母的核糖体蛋白质基因 YS3, 其编码酵母核糖体蛋白质 S3。它的破坏产生不能生存的酿酒酵母 (*Saccharomyces cerevisiae*) 单倍体孢子 (Finken-Eigen, M. , Domdey, H., Köhrer, K. , Biochemical and Biophysical research communications 223, 397-403 (1996))。这些研究证明下调单一核糖体蛋白质就已经能影响核糖体复合物的形成和/或正确的功能。

将要用于在本发明寄生虫中控制核糖体蛋白质基因转录的启动子仅仅需要满足一个先决条件。在寄生虫繁殖期间, 它们必需被打开。为了提供具有正常繁殖所必需天然数量核糖体的本发明寄生虫, 当然这是必需的。然而, 在接受寄生虫做为疫苗的受体宿主中必需关闭启动子。认为如果启动子起始它所控制基因的转录, 启动子被认为打开。在本发明中, 该基因将是核糖体蛋白质基因。如果启动子控制的基因转录低于在打开情况下基因转录的至少两倍, 启动子被关闭。优选地, 转录水平低至少 3、更优选地 4, 仍旧更优选地 5, 6 或者甚至 7 倍。要强调的是无论如何, 不需要转录的完全抑制。在寄生虫灭绝前, 低水平核糖体蛋白质转录最后将引起寄生虫延长的生命期间。因此它们将稍微更长期间地触发免疫系统。

原则上, 存在两种不同的可能: 如果不应用某些关闭启动子的条件, 启动子就打开, 或者如果不应用某些打开启动子的条件, 启动子就关闭。

优选地, 除非应用接受寄生虫做为疫苗的受体宿主中不存在的一些条件, 启动子处于关闭状态。

如果必需, 2 或更多种核糖体蛋白质基因可以置放于诱导型启动子控制下。如果启动子中所用的诱导型启动子不能被充分地关闭, 即, 如果诱导型启动子是渗漏启动子, 或者在例外的情况下, 一种特异性核糖体蛋白质的缺乏不足以去稳定该核糖体, 这将是优选的选择方案。

将通过下面的例子解释本发明。

鼠弓形体 (*Toxoplasma gondii*) 用猫做为终宿主, 并且分别使用食草动物, 杂食动物和食肉动物做为随后的中间宿主。在弓形体属 (*Toxoplasma*) 的情况下, 是寄生虫的卵囊/组织孢囊阶段最终感染人类。

人类和恒温动物是接种的靶哺乳动物，因此，弓形体属 (*Toxoplasma*) 速殖子是活减毒寄生虫所需要的寄生阶段。因此，速殖子是本发明的寄生阶段，在这个阶段，根据本发明核糖体蛋白质基因被置于诱导型启动子控制下。可以在打开启动子的条件下，用经典的方法繁殖由此制备的重组寄生虫，也称为减毒活寄生虫。在这些条件下，核糖体数量将相同于或接近于天然条件下的核糖体数量。如果为了疫苗目的培养了足够的寄生虫，收集该减毒寄生虫，并且做为疫苗施用。在待接种宿主中，不存在打开启动子的条件，因此启动子将保持在关闭状态。在接种时候，寄生虫和野生型寄生虫一样行动，因为核糖体库与天然状态相当。因此，优选地，感染过程和宿主细胞侵入过程将完美模拟天然感染过程。只要寄生虫在宿主细胞中开始分裂，它也分开其后代的核糖体库。因为，当在宿主细胞中时，（至少）一种核糖体蛋白质基因的启动子位于关闭位置，将存在减少的核糖体从头合成或者甚至不存在的核糖体从头合成。因此，后代将慢慢灭绝。然而，在感染过程的同时，免疫系统的触发已经如同在野生型寄生性感染中一样继续进行。因此，将建立最终的免疫性，好象曾经发生强毒野生型寄生虫的感染一样，而在一轮或几轮感染后，用于免疫性诱导的活减毒寄生虫已灭绝。下面的例子提供了进一步的细节。

除了新孢子虫属 (*Neospora*) 利用狗做为最终宿主，并且在例如牛、狗、绵羊和马中引起流产的事实外，犬新孢子虫 (*Neospora caninum*) 的生活周期与弓形体属 (*Toxoplasma*) 的生活周期相似。因此，新孢子虫属 (*Neospora*) 疫苗的方法紧密地相关于上述弓形体属 (*Toxoplasma*) 疫苗方法。关于弓形体属 (*Toxoplasma*)，速殖子是根据本发明核糖体蛋白质基因被置于诱导型启动子控制下的寄生阶段。Howe, D. K. 和 Sibley, L. D. METHODS: 13 (2): 123-33 (1997)) 描述了用于新孢子虫属 (*Neospora*) 的分子遗传学工具的进展。

为了产生活减毒艾美球虫属 (*Eimeria*) 寄生虫，裂殖子是根据本发明核糖体蛋白质基因被置于诱导型启动子控制下的寄生阶段。在这种情况下，尽管疫苗不包含裂殖子，但是包含芽胞化卵囊 (sporulated oocysts)。这是由于芽胞化卵囊是寄生虫被小鸡正常摄取的形式的事实。为了复制根

据本发明制备的最初重组裂殖子,将这些裂殖子引入到小鸡消化道中将足够了。然后小鸡将排出重组卵囊,并且可以分离重组卵囊和直接将其用做球虫病疫苗,例如施用到饮用水的口服疫苗中的活减毒寄生虫。从鸡粪分离卵囊是本领域熟知的标准方法。Kelleher, M. 和 Tomley, F. M. (Mol. Biochem. Parasitol. 97 (1-2): 21-31 (1998)) 等描述了艾美球虫属 (*Eimeria*) 的基因工程。

例如,从疟原虫属 (*Plasmodium*) 寄生虫红血球阶段开始,可以制备本发明的活减毒疟疾疫苗。疟原虫属 (*Plasmodium*) 重组孢子。孢子是由雌性蚊子注射到(人类)血流中的寄生虫阶段。注射后2分钟内,孢子感染肝脏,产生裂殖体和裂殖子。接着,裂殖子感染红细胞并且在其中复制。此时,核糖体库必须分配给大量后代寄生虫,并且此时是后代寄生虫逐渐灭绝的时候。此时已经充分地触发了整个免疫系统。这个例子再次示例了基于本发明重组寄生虫的疫苗的优点:它们共有活疫苗的所有优点和灭活疫苗的优点。优选地,用重组红细胞阶段的疟原虫属 (*Plasmodium*) 寄生虫或(不太好实践的)重组孢子进行接种。Crabb, B. S., 等, (Mol. Biochem. Parasitol. 90: 131-144 (1997)) 和 Wu, Y. 等, (Proc. Natl. Acad. Sci., 93: 1130-1134 (1996) 和 Proc. Natl. Acad. Sci., 92: 973-977 (1995)) 描述了用于疟原虫属 (*Plasmodium*) 的重组 DNA 技术。

而且,本发明的活减毒泰勒尔梨浆虫属 (*Theileria*) 疫苗可以基于重组裂殖子。可以在淋巴细胞中培养和维持这些裂殖子。正是在淋巴细胞中,裂殖子开始分裂,该分裂与淋巴细胞的分裂同步,而几个游离后代寄生虫将感染其它淋巴细胞,一起引起野生样免疫性的诱导,然而,如其它实例中所述,将产生由于核糖体慢慢增加的缺少而最后灭绝的后代。可以在原代淋巴细胞中繁殖和培养泰勒尔梨浆虫属 (*Theileria*)。参见,例如 Shkap V. 等, Vet. Parasitol. 65: 11-20 (1996) 和 Hulliger, L. J. Protozool. 12: 649-655 (1965)。

利用用于重组的裂殖子和/或滋养体,可以制备活减毒巴贝虫属 (*Babesia*) 疫苗。可以在红细胞中培养它们。整个方法相似于上述用于

泰勒尔梨浆虫属 (*Theileria*) 的方法。参见, I. a. Levy, M. G 和 Ristic, M. Science 207: 1218-1220 (1980)。

对于肉孢子虫属 (*Sarcocystis*) 物种, 诸如猪人肉孢子虫 (*S. suihominis*) 和 *S. neurona*, 子孢子和裂殖子都是本发明重组的靶。而且, 原理相同: 重组子孢子提供重组裂殖子, 并且在缺少从头核糖体蛋白质合成的情况下, 这些裂殖子由于缺少核糖体慢慢地灭绝。重组裂殖子可以直接用在疫苗中。参见, 例如 Murphy, A. J. 和 Mansfield, L. S. J. Parasitol. 85: 979-981 (1999) 和 Ellison, S. P. 等, Vet. Parasitol. 95: 251-261 (2001)。

就动基体目 (*Kinetoplastida*) 而言, 已经描述了布氏锥虫 (*Trypanosoma brucei*) (Wirtz, E. 和 Clayon, C. , Science 268: 1179-1183 (1995) 和 Biebinger, S. 等, Mol. & Biochem. Parasitol. 85: 99-112 (1997)), 刚果锥虫 (*Trypanosoma congolense*) (Inoue N. , 等, Mol. & Biochem. Parasitol. 120: 309-313 (2002)) 和杜氏利什曼原虫 (*Leishmania donovani*) (Yan, S. , 等, Mol. & Biochem. Parasitol. 112: 61-69 (2001)) 的四环素调节的基因表达, 并且如下所述, 可以调整四环素调节的基因表达以调节核糖体蛋白质基因转录: 简而言之, 寄生虫的前循环型是转染的靶。四环素阻抑蛋白整合到核糖体 RNA 重复的非转录间隔区中, 使得可以用依赖于四环素的方式调节异源基因 (在该参考文献中不是核糖体基因)。为了构建动基体目 (*Kinetoplastida*) 的本发明活减毒寄生虫, 首先, 一起插入额外拷贝的核糖体蛋白质基因和含有一个或多个四环素操作基因元件的启动子。随后, 从寄生虫基因组缺失内源基因拷贝。这可以容易地通过同源重组来完成, 优选地在重组标记存在情况下进行。这与下述用于顶复门 (*Apicomplexa*) 的方法相似。内源核糖体蛋白质基因的直接靶向对于利什曼原虫属 (*Leishmania*) 和锥虫属 (*Trypanosoma*) 不可行, 因为在利什曼原虫属 (*Leishmania*) 和锥虫属 (*Trypanosome*) 中大多数基因被组织为相同 DNA 链上邻近基因的大 (> 100-500 kb) 多顺反子簇。因此, 一种基因的抑制将引起位于下游所有基因的转录抑制 (Myler, P. J. 等, Med. Microbiol. Immunol. 190: 9-12

(2001))。

上面给出的例子实际上只是示例。它们决没有限制本发明的范围。在 *Encyclopaedic Reference of Parasitology*, Heinz Mehlhorn, Springer Verlag (2001) (ISBN 3-540-66829-2) 中可以发现各种各样顶复门 (*Apicomplexa*) 和锥虫科 (*Trypanosomatidae*) 寄生虫的例子和它们的生活周期。因此,对于顶复门(*Apicomplexa*)和锥虫科(*Trypanosomatidae*)的每种寄生虫,本领域技术人员利用上面给出的例子和 *Encyclopaedic Reference of Parasitology* 能正确地确定哪个阶段将是做为制备本发明活减毒寄生虫起点的优选阶段。

属于上述科的许多寄生虫都有各种不同的宿主。只是做为例子:存在巴贝虫属 (*Babesia*) 物种, 诸如感染狗的犬巴贝虫 (*B. canis*), 感染马、骡和驴的马巴贝虫 (*B. caballi*), 感染牛、野生反刍动物和人类的分歧巴贝虫 (*B. divergens*)。然而,在所有情况下,寄生虫的生活周期相似。因此,上文指出例如抗巴贝虫属 (*Babesia*) 的本发明疫苗可以基于重组裂殖子,这适用于所有的巴贝虫属 (*Babesia*) 物种。在上述 *Encyclopaedic Reference of Parasitology*, Heinz Mehlhorn, Springer Verlag (2001) (ISBN 3-540-66829-2) 中也可以发现关于一个科中各种物种生活周期的详细内容。

因此,本发明的一个实施方式涉及顶复门 (*Apicomplexa*) 和锥虫科 (*Trypanosomatidae*) 的减毒活寄生虫,并且它们的特征在于它们包含诱导型启动子控制下的核糖体蛋白质基因。

上面刚刚提到了诱导型启动子的概念。诱导型启动子是在外部因子作用下可以开或关的启动子。这种开关因子可以是生理因子,例如热;这是几十年以来本领域熟知的所有大量热激启动子的触发因素。这种因子也可以是化学性质的。而且,大量这种因子是本领域熟知的。存在了太多本领域已知的诱导型启动子,而不能全部提到它们。这里将提到几个例子。IPTG 诱导的 Lac 启动子可能是最常用的诱导型启动子之一。可以替代的诱导型启动子系统是例如四环素控制的反式激活系统 (Baron, U. 等, Oxford University Press 25: 2723-2729 (1995)) 和蜕皮激素诱导的表达系统

(Invitrogen) (Yao, T. P. 等, CELL 71: 63-72 (1992))。

原则上有两种诱导型启动子：在存在一定条件下打开的启动子，和在存在一定条件下关闭的启动子。该条件可以是化学物质的存在。

在本发明该实施方式优选的形式中，所用的启动子是在存在一定条件下被打开的启动子，所述一定条件在宿主中不天然存在。使用这种启动子的优点是一旦将它们施用给寄生虫的天然宿主，它们就自动地处于（或转换到）关闭的位置。这意味着优选地在“人工”条件即，天然宿主中不存在的条件下培养本发明的活减毒寄生虫以进行复制。

优选类型的诱导型启动子是基于操纵基因位点和阻抑蛋白的诱导型启动子类型，并且阻抑蛋白能可逆地结合所述操纵基因位点。然后，通过施用上述“条件”，即加热、化学试剂等存在或缺乏，可以调节阻抑蛋白的结合和分离。

可以非常有效地用在本发明减毒活寄生虫中的诱导型启动子的非常适合的例子，或者更精确地说启动子/操纵基因/阻抑蛋白复合物，是 tet 启动子/tet 操纵基因复合物，进一步也称为 tetR 系统。

已经描述了这种 tetR 系统，并且证明它在许多不同原生动物寄生虫中，诸如布氏锥虫(*T. brucei*) (Wirtz 等, Science 268: 1179-1183 (1995), Biebinger 等. (Mol. Biochem. Paras. 85: 99-112 (1997)) 中和溶组织内阿米巴(*Entamoeba histolytica*) (Hamann 等, Mol. Biochem. Paras. 84: 83-91 (1997)) 中起作用。tetR 系统也成功地用在弓形体属(*Toxoplasma*) 中以调节肌球蛋白 A 的表达(Meissner M, 等, Nucleic Acids Res. 29 (22): E115 (2001))。此外，在肠贾第鞭毛虫(*Giardia lamblia*) 和杜氏利什曼原虫(*Leishmania donovani*) 中也证明了四环素调节的表达，表明它在寄生虫中的普遍适用性(Yan S, 等, Mol Biochem Parasitol. 112 (1) : 61-9 (2001), Sun, C. H. 和 Tai, Mol. Biochem. Parasitol. 105 (1) : 51-60 (2000))。

这种复合物如下所述进行工作，并且实施例中更详尽地描述了这种复合物的操作。

原则上，必须进行两步以产生四环素调节的核糖体蛋白质表达：1.

四环素阻抑蛋白 (tetR) 基因的整合和表达和 2. 在转录起始附近的核糖体蛋白质基因的启动子中一种或多种四环素操纵基因元件的整合。

tet 阻抑蛋白基因是编码能结合 tet 操纵基因位点, 因此能阻断邻近基因转录的蛋白质的基因。现在, 将该基因置放于组成型启动子控制之下 (即, 重组寄生虫中组成型), 并且利用重组 DNA 技术将其引入到寄生虫中。因此, 重组寄生虫将合成 tet 阻抑蛋白。优选地, 将 tet 操纵基因引入到 STS 上游一个或多个核糖体蛋白质基因的转录起始位点附近, 优选地内源启动子中。因此, tet 阻抑蛋白将结合 tet 操纵基因, 因此阻断下游核糖体蛋白质基因的转录。然而, 在四环素存在的情况下, 阻抑蛋白将与 tet 操纵基因位点分离, 因此能使下游基因转录。因此, 在四环素存在的情况下, 如同在天然情况下一样, 重组寄生虫能复制。如果能在体外培养重组寄生虫 (对于许多寄生虫包括上述例子的大多数寄生虫都是这种情况), 可以容易地将四环素添加到生长培养基中。如果寄生虫生长需要在天然宿主中繁殖, 例如对于艾美球虫属 (*Eimeria*) 寄生虫就是这种情况, 可以容易地将四环素口服或注射地施用给宿主 (这种情况下为小鸡)。应该注意下面事项: 细胞外和细胞内寄生虫都摄取四环素。宿主细胞的细胞破裂对该药物对核糖体蛋白质表达调节产生作用是不必要的。

步骤 1, 可以如上述文献中所述, 获得四环素阻抑蛋白基因 (tetR) 的整合和表达。表明稳定转化和四环素阻抑蛋白基因可能整合的适合和熟知的选择性标记是例如 CAT 基因 (Kim, K., 等, Science 262 (5135): 911-4 (1993))。其它适合的稳定转染筛选标记也是本领域已知的, 诸如 DHFR-TS (Donald, R. G. 和 Roos, D. S., Proc. Natl. Acad. Sci. USA 90 (24): 11703-11707 (1993), Roos, D. S. 等, METHODS 13: 112-122 (1997)) 和 HXGPRT (Donald, R. G. 等, J. Biol. Chem. 271: 14010-14019 (1996), Donald, R. G. 和 Roos, D. S., Mol. Biochem. Parasitol. 91 (2): 295-305 (1998))。

Cre-lox 系统也提供了适合的筛选系统 (参见, Hardy, S. 等, Journ. Virol. 71: 1842-1849 (1997))。

如果 tetR 系统用做诱导型启动子系统, 核糖体蛋白质基因上游启动

子可以是例如内源启动子,现在通过在转录起始位点附近克隆 tet 操纵基因使得它可诱导(参见下面 tet 操纵基因序列和优选的插入位点的详细内容)。不用说,能提供足够高核糖体蛋白质基因转录水平的任何其它启动子也是适合的。

如果使用另一种诱导型启动子系统,可容易地利用该诱导型启动子,并且缺失内源启动子。然而,如果使用另一种调节元件,其原理相似于 tet 操纵基因,启动子本身同样可以是内源启动子。而且不用说,能提供足够高下游所克隆的核糖体蛋白质基因转录水平的任何其它启动子也是适合的。

步骤 2,用在 STS 附近含有一个或多个 tetO 位点(=tet 操纵基因位点)的核糖体蛋白质基因置换野生型核糖体蛋白质基因需要在选定的核糖体蛋白质基因启动子和该基因本身之间插入 tet 操纵基因位点。Yan S, 等. (Mol. Biochem. Parasitol. 112 (1): 61-9 (2001)), Wirtz, E 和 Clayon, C. (Science 268 (5214): 1179-83 (1995)) 和 Meissner M, 等. (Nucleic Acids Res. 29 (22): E115 (2001)) 描述了 tet 操纵基因。

单一 tet 操纵基因(tetO)位点的序列是:

5'-TCCCTATCAGTGATAGAGATC-3'。

原则上,选定的核糖体蛋白质基因前面单一 tet 操纵基因位点的插入足够了。然而, tetR 系统和所有生物系统一样不是可以诱导精确地从 0 % 到 100 % 活性,反之亦然。因此,如果需要更强的调节水平,优选地插入 2 个或多个操纵基因位点。tet 操纵基因干扰转录下游基因的 RNA 聚合酶的结合。因此,优选地,在相对于转录起始位点(这里称为 STS),从核苷酸-100 延伸到+3 的区域中的某处插入 tet 操纵基因。而且,在实施例中还描述了如何定位这种 STS。

可以用 Donald, R. G. 和 Roos, D. S. (Mol. Biochem. Parasitol. 91 (2): 295- 305 (1998)) 所述的打了就走(hit-and-run)策略进行用包含一个或多个 tet 操纵基因位点的重组基因置换野生型核糖体蛋白质基因的步骤。

本领域技术人员将能发现利用其它阳性和阴性选择性标记的联合的

可替代的方法。例如, HSV 胸苷激酶可以用做阴性选择性标记 (LeBowitz, J. H. 等, Mol. Biochem. Parasitol. 51 (2): 321-5 (1992), Fox, B. A. 等, Mol. Biochem. Parasitol. 116 (1): 85-8 (2001))。

用于本发明重组弓形体属 (*Toxoplasma*) 寄生虫构建的分子工具在新孢子虫属 (*Neospora*) 中同样可以很好地工作 (Howe, D. K. 和 Sibley, L. D. METHODS 13 (2): 123-133 (1997))。

在艾美球虫属 (*Eimeria*) 中, 相同的方法同样适用。只是做为例子: Kelleher, M. 和 Tomley, F. M. (Mol Biochem Parasitol. 97 (1-2): 21-31 (1998)) 证明 β -半乳糖苷酶可以在禽艾美球虫 (*E. tenella*) 中瞬时表达。

对于泰勒尔梨浆虫属 (*Theileria*), Adamson, R. 等. (Mol. Biochem. Parasitol. 114 (1): 53-61 (2001)) 已经开发了瞬时转染感染性单核环形泰勒尔梨浆虫 (*Theileria annulata*) 子孢子的方法。

在疟原虫属 (*Plasmodium*) 中, 被突变以赋予对乙胺嘧啶抗性的二氢叶酸还原酶-胸苷酸合酶 (dhfr-ts) 编码序列, 或嘌呤霉素-N-乙酰基转移酶, 或曲霉属 (*Aspergillus*) 杀稻瘟菌素 S 脱氨基酶 (BSD) 基因, 或来源于转座子 Tn5 的新霉素磷酸转移酶 II (NEO) 基因已经被描述为选择性标记 (Wu, Y., 等, Proc. Natl. Acad. Sci. USA. 93 (3): 1130-4 (1996), Wang, P., 等, Mol. Biochem. Parasitol. 123 (1): 1-10 (2002), de Koning-Ward, T. F., 等. (Mol. Biochem. Parasitol. 117 (2): 155-60. (2001))。

相似的选择性标记在巴贝虫属 (*Babesia*) 中也起作用。

因此, 本领域技术人员能将本发明应用于属于顶复门 (*Apicomplexa*) 和锥虫科 (*Trypanosomatidae*) 的全部范围的寄生虫。

该实施方式优选的形式涉及属于球虫亚纲 (*Coccidia*), 梨浆虫目 (*Piroplasmida*) 或血孢子虫目 (*Haemosporida*) 的本发明减毒活寄生虫。

在该实施方式更优选的形式中, 减毒活寄生虫属于艾美球虫科 (*Eimeridiidae*), 隐孢子虫科 (*Cryptosporidiidae*) 或肉孢子虫科 (*Sarcocystidae*)。

在该实施方式甚至更优选的形式中, 减毒活寄生虫属于艾美球虫属

(*Eimeria*), 隐孢子虫属(*Cryptosporidium*), 弓形体属(*Toxoplasma*), 肉孢子虫属(*Sarcocystis*) 和新孢子虫属(*Neospora*)。

在该实施方式另一个更优选的形式中, 减毒活寄生虫属于巴贝虫科(*Babesiidae*) 或泰勒尔梨浆虫科(*Theileriidae*)。

在该实施方式甚至更优选的形式中, 减毒活寄生虫属于巴贝虫属(*Babesia*) 或泰勒尔梨浆虫属(*Theileria*)。

在该实施方式另一个更优选的形式中, 减毒活寄生虫属于疟原虫属(*Plasmodium*)。

还是在实施方式另一个更优选的形式中, 减毒活寄生虫属于锥虫属(*Trypanosoma*) 或利什曼原虫属(*Leishmania*)。

在甚至更优选的形式中, 减毒寄生虫属于墨西哥利什曼原虫(*Leishmania mexicana*), 婴儿利什曼原虫(*L. infantum*) 或硕大利什曼原虫(*L. major*) 物种或布氏锥虫(*Trypanosoma brucei*) 或克氏锥虫(*T. cruzi*) 物种。

在该实施方式另一个优选的形式中, 本发明活减毒寄生虫核糖体蛋白质基因受可被抗生素诱导的诱导型启动子控制。

更优选地, 这些抗生素是四环素或无水四环素, 或它们的衍生物。

在该实施方式另一个优选的形式中, 选定的核糖体蛋白质基因是编码 L9, S3, 质体-S9 或 S13, 优选地鼠弓形体(*Toxoplasma gondii*) L9, S3, 质体-S9 或 S13 的基因。

SEQ ID NO: 1 中描述了编码 9 号大亚基核糖体蛋白质(L9) 的基因的核苷酸序列, 及包含启动子区的上游序列。

区域	1	2296	启动子	启动子区
区域	2297	2461	e	外显子 1
区域	2416	2418	atg	atg 起始密码子
基因	2416	4831	cds	编码序列
区域	2462	3838	i	内含子 1
区域	3839	4260	e	外显子 2

区域	4261	4727	i	内含子 2
区域	4728	4834	e	外显子 3
区域	4829	4831	终止	TAA 终止密码子

SEQ ID NO: 2 中描述了编码质体 9 号小亚基核糖体蛋白质 (S9) 的基因的核苷酸序列及包含启动子区的上游序列。

区域	1	3076	启动子	启动子区
区域	3077	3616	e	外显子 1
区域	3156	3158	atg	ATG 起始密码子
基因	3156	4325	cds	编码序列
区域	3617	3874	i	内含子 1
区域	3875	4034	e	外显子 2
区域	4035	4130	i	内含子 2
区域	4131	4338	e	外显子 3
区域	4323	4325	终止	TAA 终止密码子
区域	4326	4338	3' utr	3' UTR

SEQ ID NO: 3 中描述了编码 13 号小亚基核糖体蛋白质 (S13) 的基因的核苷酸序列及包含启动子区的上游序列。

区域	1	1289	启动子	启动子区
区域	1290	1594	e	外显子 1
区域	1448	1450	atg	ATG 起始密码子
基因	1448	3639	cds	编码序列
区域	1595	2527	i	内含子 1
区域	2528	2615	e	外显子 2
区域	2616	3489	i	内含子 2
区域	3490	3639	e	外显子 3

SEQ ID NO: 4 中描述了编码 3 号小亚基核糖体蛋白质 (S3) 的基

因的核苷酸序列及包含启动子区的上游序列。

区域	1	1177	启动子	启动子区
区域	1178	1308	e	外显子 1
区域	1291	1293	atg	ATG 起始密码子
基因	1291	2651	cds	编码序列
区域	1309	1752	i	内含子 1
区域	1753	2137	e	外显子 2
区域	2138	2249	i	内含子 2
区域	2250	2389	e	外显子 3
区域	2390	2486	i	内含子 3
基因	2487	2748	e	外显子 4
区域	2649	2651	终止	TAA 终止密码子
区域	2652	2748	3' utr	3' UTR

本发明减毒活寄生虫非常适于用在疫苗中。如上所详尽阐述的，这是由于它们联合了活减毒疫苗和灭活疫苗的优点，而没有它们缺点的事实。因此，本发明的另一个实施方式涉及用在疫苗中的本发明减毒活寄生虫。

本发明另一个实施方式涉及抗寄生虫感染，并且包含本发明活减毒寄生虫和药理学可接受载体的疫苗。

药理学可接受的载体可以是例如无菌水或无菌生理盐水溶液。在更复杂的形式中，载体可以是例如本领域熟知的缓冲液诸如 PBS。

在优选的形式中，本发明疫苗也可以含有免疫刺激物质，所谓的佐剂。佐剂一般包含以非特异性方式加强宿主中免疫应答的物质。大量不同的佐剂是本领域已知的。在母牛疫苗中常常使用的佐剂例子是胞壁酰二肽，脂多糖，几种葡聚糖和聚糖和 Carbopol®（均聚物）。

疫苗也可以包含所谓的“运载体（vehicle）”。运载体是蛋白质附着但不共价连接它的化合物。这种运载体是例如脂质载体，ISCOMs®, dendromers, Niosomes, 微粒，特别是基于脱乙酰壳多糖的微粒，多糖基质等，生物微囊，微藻酸盐，脂质体和 macrosols，所有这些载体都是本

领域已知的。特别用于口服接种中的微粒，更具体地是基于脱乙酰壳多糖的微粒非常适于用做疫苗运载体。

其中抗原部分包埋在运载体中的该运载体的一种特殊形式是所谓的 ISCOM® (EP 109.942, EP 180.564, EP 242.380)。

此外，疫苗可以包含一种或多种适合的表面活性化合物或乳化剂，例如 Span®或 Tween®。而且，疫苗可以包含一种或多种免疫刺激物诸如细胞因子，例如干扰素。

当与灭活寄生虫比较时，可以施用相对低量的基于上述活减毒重组寄生虫的疫苗，因为在感染期间它们会繁殖自身。因此，每个剂量中非常合适的量范围是 10^2 到 10^7 个寄生虫。每个剂量中低于 10^2 个寄生虫的数量不能总是确保所有免疫动物中足够水平的保护。要是只从经济观点考虑，每个剂量中范围是 10^7 到 10^8 个寄生虫，尽管适合，但不是非常切实可行。

本发明另一个实施方式涉及本发明减毒活寄生虫用于对抗顶复门 (*Apicomplexa*) 或锥虫科 (*Trypanosomatidae*) 寄生虫所引起感染的疫苗制备的用途。

本发明另一个实施方式涉及本发明疫苗的生产方法，该方法包括本发明活减毒寄生虫和药理学可接受载体的混合。

例如可以皮内、皮下、肌内、腹膜内、静脉内或在粘膜表面诸如口服或鼻内施用本发明的疫苗。

tet 阻抑蛋白基因是原核生物来源的基因。因此，该基因的密码子使用在真核生物体诸如本发明涉及的活减毒寄生虫中是次最佳的。因此，本领域技术人员能受到启发而修改 tet 阻抑蛋白基因的编码序列，使得它对应于真核生物细胞的密码子使用，因此获得合成的 tet 阻抑蛋白基因。这已经被 Meissner M, 等. (Nucleic Acids Res. 29 (22): E115 (2001)) 完成。当然，人们将预期因为这种合成的 tet 阻抑蛋白基因已经完全适于真核生物细胞，所以它不能被进一步优化了。而且，人们将预期这种“合成的” tet 阻抑蛋白将是真核生物细胞中最适合的阻抑蛋白。原则上，这种蛋白质是与天然蛋白质相同的蛋白质，因此根据定义其最适合与 tet 操纵基因位点相互作用。

然而，现在惊奇地发现包含融合到天然即原核生物 tet 阻抑蛋白 N 末端部分的异源基因（一部分）的重组基因编码的融合蛋白，甚至比完全适应真核生物的“合成的”tet 阻抑蛋白提供更显著优越的 tet 操纵基因调节。

因此，这种融合蛋白将是优选用在本发明活减毒寄生虫中的阻抑蛋白。这甚至是更出人意料地发现，因为 tet 阻抑蛋白 3D 结构研究预测 N-末端融合将不利地干扰 DNA 结合。然而，在实践中惊奇地发现不是这种情况。

异源基因是编码除了 tet 阻抑蛋白外蛋白质的任何基因。异源蛋白质是除了 tet 阻抑蛋白外的任何蛋白质。重组基因是任何人工制备的基因，该基因包含融合到编码 tet 阻抑蛋白质 N 末端的 tet 阻抑蛋白基因一侧的异源基因（部分）。

融合蛋白必须能到达核以与 tet 操纵基因相互作用。因此，tet 阻抑融合蛋白存在一些要满足的先决条件：单体 tet 阻抑蛋白融合蛋白的最终分子量必须是 <60 kD，融合蛋白质的异源部分必须在 tet 阻抑蛋白 N 末端侧，并且融合蛋白必须无 GPI 锚（anchors），分泌/外泌信号（Secretion/excretion signals）和跨膜区。原则上，满足这些先决条件和（因此）能靶向核的每个蛋白质或其部分都可以用于和 tet 阻抑蛋白进行 N 末端融合。

不需要利用全长异源蛋白质进行融合。利用这种异源蛋白的部分就足够了。认为该部分是做为异源融合蛋白的至少 10 个氨基酸，优选地至少 20 个氨基酸的片段。优选地，该部分来源于异源蛋白质 N 末端侧。例如，选定的异源蛋白质是绿色、红色和黄色荧光蛋白质和 CAT 蛋白质。

因此，本发明的另一个实施方式涉及编码 tet 阻抑蛋白融合蛋白的 DNA 片段，其特征在于：它包含 tet 阻抑蛋白和在 tet 阻抑蛋白 N 末端侧融合的异源蛋白或其部分，其中单体形式融合蛋白具有 <60 kD 的大小，并且融合蛋白质无 GPI 锚，分泌/外泌信号和跨膜区。

本发明另一个实施方式涉及这种 tet 阻抑蛋白融合蛋白，其特征在于：它包含 tet 阻抑蛋白和在 tet 阻抑蛋白 N 末端侧融合的异源蛋白或其

部分,其中单体形式融合蛋白具有<60 kD的大小,并且融合蛋白质无 GPI 锚,分泌/外泌信号和跨膜区。

用语“跨膜区”提到的膜是位于细胞胞质和外界之间的这些膜。具体地,这些膜排除了核和胞质之间的膜。优选地,本发明 tet 阻抑融合蛋白的确有特异地将融合蛋白引导到核的一些信号。这点是清楚的,因为 tet 阻抑蛋白融合蛋白(如同天然 tet 阻抑蛋白基因所需要的)必须进入核以便能够调节它所控制基因的转录。

由于它的普遍性特征, tetR 系统和 tet 阻抑蛋白融合蛋白的联合不仅可以用在本发明活减毒寄生虫中,当然也可以用在其它寄生虫中和其它真核生物细胞和生物体中。对于任何基因之表达的调节,这在真核细胞中是广泛适用的。

因此,当这种寄生虫包含与上述 tet 阻抑融合蛋白(编码遗传信息)联合的 tet 操纵基因时,本发明减毒活寄生虫更适于做为疫苗的基础。这能够更好地阻断和诱导核糖体基因的转录。

因此,在更优选形式的本发明减毒活寄生虫中,其中通过四环素,无水四环素或它们的衍生物调节基因诱导的减毒活寄生虫包含 tet 操纵基因和编码上述 tet 阻抑蛋白融合蛋白的遗传信息。

如实施例中将要说明的,如果串连地克隆一个或多个 tet 操纵基因位点,上述 tet 阻抑蛋白融合蛋白的意外特征更显著。用语“串连地”应该在以下意义上做广义地解释:可以直接相互邻近地克隆 tet 操纵基因位点,或者在两个或多个 tet 操纵基因位点之间中具有间隔区序列。如上所述,优选地, tet 操纵基因位点被克隆在相对于 STS 的-100 到+3 之间的区域。

因此,在更优选的形式中,本发明的这种减毒活寄生虫不仅包含上述 tetR 系统和 tet 阻抑蛋白融合蛋白,也包含 2 个或多个,而不是 1 个 tet 操纵基因位点。

实施例

实施例 1

试验过程期间所用的引物:

插入的限制位点被加有下划线。

SEQ ID NO	#	名称	序列 5' → 3'
5	1	SAG3-FW	CGATAAGCTTCGAATCTCTGAACGGATGTGT
6	2	TUB5-RV	CGAGATCTGGGAATTCAAGAAAAATGCCAACG
7	3	TETAVR5-FW	CGATCCTAGGATGTCTAGATTAGATAAAAG
8	4	TETPST3-RV	CGTCTGCAGTTAAGACCCACTTTTACATTTAAG
9	5	T3	ATTAACCCTCACTAAAGGGAA
10	6	SAG1/1634-RV	CGATAAGCTTTCGGGGGGGCAAGAATTGTGT
11	7	REV 13A	GCGCCCCATGGTGACGGAGAAAAATCG
12	8	REV 13B (嵌套引物)	GGGAACCGCAAGGTGGGAGCGGAGAAC
13	9	S13PROMFUS FW	GCATAAGCTTCCTCGCAGAGATTGTCAGTG
14	10	S13PROMFUS RV	GCATTCTAGAGGCAGACATGCCCTTTCCAGG
15	11	LACZ-AVR11 FW	CGATCCTAGGATGACCATGATTACGGATTCACT
16	12	LACZ-PST1 RV	CGATCTGCAGTTATTTTGGACACCAGACCAA
17	13	S13INSTETO+3FW	GGTTCTCCCCCTCAATCCCTATCAGTGATAGAGATCTCTCTCTTCTCT
18	14	S13INSTETO+3RV	AGAGAAAGGAAGAGAGATCTCTATCACTGATAGGGATTGAGGGGAGAACC
19	15	S13SUBTETO-23FW	CTACGCGGCCGACGGTCCCTATCAGTGATAGAGATCTTCCTCGACGGGTTC
20	16	S13SUBTETO-23RV	GAACCCGTCGAGGAAGATCTCTATCACTGATAGGGACCGTCGGCCCGTAG
21		S13NOT1-FW	CGATGCGGCCGCGTCAGTGCATGACACAACCG
22		S13SAC1-RV	GCTAGAGCTCCTGTAAAGTCGCCAGAGAAGCAC
23		M13-REV	AACAGCTATGACCATGATTACGC
24		S13CL FW3	CGATAGTGTGCAATAACAGG
25		HRCHECK II 5 S13-FW	GTCGAGTCCTGTAGGTTTCATC
26		HRCHECK II S13-RV	CTCCGAAGGAGTCTCTCAGTG
27		T7	AATACGACTCACTATAG
28		HXGPRT/BGLII-FW	CGATAGATCTAAAATGGCGTCCAAACCCATTG
29		HXGPRT/PSTI-RV	CGATCTGCAGTTACTTCTCGAACTTTTTCGG

TubYFP/TR-sagCAT (9332bp) 的构建

如下所述,逐步改造质粒 *ptubYFP/TR-sagCAT*。首先,利用引物 SAG3FW (#1, SEQ ID NO: 5) 和 TUB5RV (#2, SEQ ID NO: 6), 通过从 *ptubYFP/YFP-sagCAT* 构建体(Llopis, J.等, PNAS 97 (8): 4363-4368 (2000))扩增鼠弓形体 (*Toxoplasma gondii*) 微管蛋白 A (tub) 启动子, 制备构建体 *ptubCAT/GFP*。用 HindIII 和 BglIII 消化 PCR 产物及质粒 *pdhfrCAT/GFP* (Striepen, B. 等, Molecular and Biochemical Parasitology 92: 325-338 (1998))。并且将它们互相连接。这产生了 *ptubCAT/GFP*, 其中已经用 tub 启动子置换了 dhfr 启动子。所得到的质粒是基于 Bluescript pKS+®(Stratagene, La Jolla, CA), 并且含有 α -微管蛋白启动子, 通过 BglIII 位点将氯霉素乙酰转移酶 (CAT) 编码序列和绿色荧光蛋白的融合物与该 α -微管蛋白启动子分开。

为了获得 *ptubYFP/TR* 构建体, 将 CAT 编码序列交换为黄色荧光蛋白 (YFP), GFP 编码序列交换为 tet 阻抑蛋白编码序列 (tetR)。通过 BglIII 和 AvrII, 从 *ptubYFP/YFP-sagCAT* 切下 YFP 基因, 并且连接到 *ptubCAT/GFP* 构建体 BglIII 和 AvrII 位点之间, 置换 CAT 编码序列。利用引物 TETAVR5-FW (#3, SEQ ID NO: 7) 和 TETPST3-RV (#4, SEQ ID NO: 8), 通过 PCR 从大肠杆菌 (*E. coli*) Tn10 扩增 tetR 编码序列 (Hillen, W. 和 Berens, C., Annu. Rev. Microbiol. 48: 345-369 (1994)), 用 AvrII 和 PstI 消化, 并且在构建体中通过将 GFP 编码序列交换为 tetR 编码序列进行连接。所得到的质粒命名为 *ptubYFP/TR*。

最后, CAT 筛选盒被插入到 tub 启动子上游, 产生 *ptubYFP/TR-sagCAT* 质粒。这是通过利用引物 T3 (#5, SEQ ID NO: 9) 和 SAG1/1634 RV (#6, SEQ ID NO: 10), 从前述 *ptubYFP/YFP-sagCAT* 构建体扩增 CAT 盒, 用 HindIII 消化, 并且连接到 *ptubYFP/TR* 构建体唯一的 HindIII 位点来完成的。

图 1 中示出了 *TubYFP/TR-sagCAT* 的构建和它的全长序列。

实施例 2

鼠弓形体 (*Toxoplasma gondii*) 核糖体蛋白质基因 S13 起始转录位

点的测定

为了测定核糖体蛋白质基因 S13 转录的起始, 从在 Vero 细胞中生长的鼠弓形体 (*Toxoplasma gondii*) RH Δ HXGPRT 速殖子分离 RNA。利用 GeneRacer®试剂盒 (Invitrogen), 从总 RNA 获得基因特异性全长 cDNA。利用该试剂盒, 将寡 RNA 连接到全长 mRNA 末端, 用寡 dT 发生逆转录后, 用结合寡 RNA 的 GeneRacer 引物及基因特异性引物, 通过 PCR 进行扩增产生产物。然后, 可以测定转录起始 (STS)。对于核糖体蛋白质基因 S13, 这可以利用下面引物: REV13A (#7, SEQ ID NO: 11) 和 REV13B (#8, SEQ ID NO: 12) 来完成。引物#7 与 GeneRacer 引物一起使用以获得产物, 随后, 引物#8 用于嵌套式 PCR。PCR 产物显示出 3 条带; 2 条弱带和 1 条强带。分离了显示最大量产物的条带, 并且测定 STS, 表明为位置 0。在图 3A 和 3B 中, 相对于起始密码子表示 STS。

实施例 3

S13/LZ 构建体

为了检测 tet 阻抑蛋白的诱导表达, 利用受 S13 启动子控制, 有或无单一 tetO 位点存在的 lacZ 基因制备几种报道构建体。首先, 制备质粒 S13/lacZ (参见图 2 最终构建体的结构和序列), 随后如下所述, 该质粒用于插入或置换 tetO 位点序列。

用引物 S13PROMFUS FW (#9, SEQ ID NO: 13) 和 S13PROMFUS RV (#10, SEQ ID NO: 14), 用 PCR 从鼠弓形体 (*Toxoplasma gondii*) RH/ Δ HXGPRT 株系的基因组 DNA 扩增 S13 启动子区。利用引物 LACZ-AVR11 FW (#11, SEQ ID NO: 15) 和 LACZ-PST1 RV (#12, SEQ ID NO: 16), 用 PCR 从 BL21 细菌的基因组 DNA 扩增 lacZ 编码序列。随后, 用 HindIII 和 XbaI 消化 S13 PCR 产物, 而用 AvrII 和 PstI 消化 lacZ PCR 产物。质粒 ptubYFP/YFP-sagCAT 用于将 ptubYFP 部分及 CAT 筛选盒交换成 S13 启动子部分。剩余的 YFP 基因被交换为 lacZ 基因, 产生 S13/lacZ 质粒。通过定点诱变, S13/lacZ 质粒用于插入或置换单一 tet 操纵基因 (tetO) 位点序列 (5'-TCCCTATCAGTGATAGAGATC-3')。这是利用 QuickChange®定点诱

变试剂盒(Stratagene)来完成的。在测定的 STS 附近插入或置换 tet0。引物 S131NSTETO+3 FW (#13, SEQ ID NO: 17) 和 S131NSTETO+3 RV (#14, SEQ ID NO: 18) 用于在相对于 STS 的+3 位置插入 tet0 位点, STS 表示为 0。引物 S13SUBTETO-23 FW (#15, SEQ ID NO: 19) 和 S13SUBTETO-23 RV (#16, SEQ ID NO: 20) 用于置换相对于 STS -43 到-23 之间的序列为 tet0 序列。在不存在或存在 tetR 和 (无水) 四环素的情况下, 用 CPRG 测定法 (Seeber, F. 等, Gene 169: 39-45 (1996)), 在鼠弓形体 (*Toxoplasma gondii*) 株系 RH Δ HXGPRT, REP1/2 (Meissner, M. 等, Nucleic Acids Research 29 (22), E115 (2001)) 和 tubYFP/TR 中检测了这两个构建体 S13instet0+3/lacZ 和 S13subtet0-23/lacZ 及 S13/lacZ。

图 2 中示出了 S13/lacZ 构建体, 并且图 3A 中示出了 S13/lacZ 构建体中 tet 操纵基因的置换或插入位点。

L9/LZ 构建体

图 3B 中示出了 tet0 插入/置换到 rp-L9 启动子中。

实施例 4

带有 pTub-YFP-TR-sagCAT 的稳定转染子弓形体属 (*Toxoplasma*) 寄生虫的筛选

如 Roos, D. S. 等. ("Methods in Microbial Pathogenesis" In Methods in Cell Biology (1994), D. G. Russell, 编者)所述进行弓形体属 (*Toxoplasma*) 寄生虫的电穿孔。

根据 Kim, K. , 等. (Science 262 (5135): 911-4 (1993)) 进行稳定转染子的筛选。

根据 Roos, D. S. 等. (1994, 上文), 再一次进行 S13/LZ, S13i+3/lacZ 和 S13s-23/lacZ 的电穿孔。

实施例 4 的结果

测定电穿孔到 tub-YFP-TR 株系中、包含单一 tet 操纵基因的 S13 启动子驱动的 lacZ 表达

检测了下面的构建体

a) S13/LZ: 这是 tub-YFP-TR 转染子弓形体属 (*Toxoplasma*) 株系, 用 S13 核糖体蛋白质基因启动子控制下的 LacZ 基因瞬时转染。在该构建体中不存在 tet 操纵基因位点

b) S13i+3/lacZ: 这是 tub-YFP-TR 转染子弓形体属 (*Toxoplasma*) 株系, 用 S13 核糖体蛋白质基因启动子控制下的 LacZ 基因瞬时转染, 该启动子其还带有在相对于 STS 的位点+3 插入的 tet 操纵基因位点 (参见图 3A)。

c) S13s-23/lacZ: 这是 tub-YFP-TR 转染子弓形体属 (*Toxoplasma*) 株系, 用 S13 核糖体蛋白质基因启动子控制下的 LacZ 基因瞬时转染, 该启动子还带有在相对于 STS 的位点-23 被置换的 tet 操纵基因位点 (参见图 3A)。

图 3B 中示出了 tetO 插入/置换到 rp-L9 启动子中的相似构建体。

如图 4 中所观察到的, 在无水四环素和四环素存在和不存在的情况下, 如所期望的, tub-YFP-TR 产生了相同水平的 LacZ。

在不存在无水四环素和四环素的情况下, 用构建体 S13i+3/lacZ 转染引起了一定数量 LacZ 的产生, 该一定数量 LacZ 是在无水四环素和四环素存在情况下产生的 LacZ 量的一半。

这清楚地表明在该株系中 LacZ 转录的可诱导性。

在不存在无水四环素和四环素的情况下, 用构建体 S13s-23/lacZ 转染引起了一定数量 LacZ 的产生, 该一定数量 LacZ 是在无水四环素和四环素存在情况下产生的 LacZ 量的 1/3。

这清楚地表明在该株系中 LacZ 转录的可诱导性。

而且, 这些结果证明了 tet 操纵基因位点相对于 STS 所位于的位点不是非常至关重要的。它也证明了可以通过插入和置换引入 tet 操纵基因位点。

用构建体电穿孔的瞬时转染子 CPRG 测定, 该构建体包含 S13 启动子

驱动的 LacZ 基因, 该 S13 启动子包含单一 tet 操纵基因或双 tet 操纵基因。

在该测定中, 比较了下面构建体:

a) 上述 S13/LZ,

b) 上述 S13s-23/lacZ (I) (= S13s-23/lacZ),

c) 除了另一个 tet 操纵基因位点已经被克隆到紧邻第一个 tet 操纵基因下游的事实外, 等同于 S13s-23/lacZ 的 S13s-23/lacZ (II)。利用与 S13s-23/lacZ (I) 相似的技术装配该构建体。

如图 5 中所示, 在不存在四环素的情况下, 上述合成的 tet 阻抑蛋白基因 (Meissner) 和本发明的融合 tet 阻抑蛋白基因 (tub-YFP-TR) 都能阻断 LacZ 转录。更惊人地, 可以清楚地看出, 当使用两个邻近 tet 操纵基因位点时, 和单一 tet 操纵基因位点的使用相比, 表达的阻断好 3 至 4 倍。

在包含上述合成的 tet 阻抑蛋白基因 (Meissner) 的株系中和包含本发明融合 tet 阻抑蛋白基因的株系中, CPCR 测定瞬时转染子, 比较 lacZ 表达。

意外地, 从图 5 中可以看出, 当与用上述合成的 tet 阻抑蛋白 (Meissner) 发现的阻断比较时, 本发明融合 tet 阻抑蛋白产生了显著更好的 LacZ 转录阻断。而且, 意外地, 当与用上述合成的 tet 阻抑蛋白基因 (Meissner) 发现的诱导比较时, 用本发明融合 tet 阻抑蛋白基因发现了好得多的 LacZ 转录诱导。

实施例 5

利用同源重组, 利用打了就走 (hit-and-run) 诱变方法, 在核糖体蛋白质 S13 基因座中插入 tet 操纵基因元件

在核糖体蛋白质 S13 基因座 (S13) 的情况下, 为了在基因组上特定基因座整合 tet 操纵基因位点, 需要同源重组。为了进行同源重组, 需要整合位点上游和下游大的序列部分 (在这种情况下是 ~1200bp) 以获得同

源重组，而不是非同源重组。如 Donald 等. (Mol. Biochem. Paras. 91: 295-305 (1998)) 所述，可以利用次黄嘌呤-黄嘌呤-鸟嘌呤磷酸核糖基转移酶 (HXGPRT) 基因做为选择性标记，分两步在特定基因座整合序列元件。

详细地，含有 HXGPRT 盒前面整合位点附近 S13 基因座一部分的转染质粒将与同源基因组 DNA S13 基因座同源重组一次，产生假二倍体型 I 和 II (图 6)。如所述 (Donald 等 1998, 上文)，用霉酚酸和黄嘌呤，在 HXGPRT 正选择下进行这步。随后，用抗 HXGPRT 的 6-硫黄嘌呤筛选第二次同源重组，引起假二倍体型丧失和速殖子产生，该速殖子有或无整合在 S13 基因座的 tet 操纵基因位点 (~1: 1 比率)。这个方法称为打了就走 (hit-and-run) 诱变。

为了进行这个方法，首先制备含有 DHFR 启动子控制下 HXGPRT 筛选盒的质粒。从鼠弓形体 (*Toxoplasma gondii*) RH 速殖子分离 RNA。利用 SUPERScript™ II RnaseH-逆转录酶 (Gibco BRL) 和标准分子生物学方法 (Sambrook & Russell: "Molecular cloning: a laboratory manual" (2001), Cold Spring Harbor Laboratory Press; ISBN: 0879695773)，用该 RNA 制备 cDNA。利用引物 HXGPRT/BGLII-FW (SEQ ID NO: 28) 和 HXGPRT/PSTI-RV (SEQ ID NO: 29)，从鼠弓形体 (*T. gondii*) RH 速殖子 cDNA 扩增 HXGPRT 编码序列，并且筛选剪接变异体 I 以进一步使用 (Donald 等, J. Biol. Chem. 271: 14010-14019 (1996))。用 BglII 和 PstI 消化 PCR 产物和质粒 pdhfrCAT/GFP (Striepen, B. 等, Mol. Biochem. Paras. 92: 325-338 (1998))，随后，CAT/GFP 编码序列与 HXGPRT 编码序列交换，产生 dhfrHXGPRT 构建体，命名为 pminiHXGPRT。

随后，利用引物 S13NOTI-FW (SEQ ID NO: 21) 和 S13SACI-RV (SEQ ID NO: 22)，从鼠弓形体 (*T. gondii*) RH 速殖子基因组 DNA PCR 扩增含有 tet 操纵基因整合位点 (相对于 STS 的 -43/-23) 上游 (~1200bp) 和下游 (~1200bp) 区域的 DNA 部分。用 NotI 和 SacI 消化该 PCR 产物和 pminiHXGPRT，随后，PCR 产物连接到 HXGPRT 盒下游。最后，如实施例 3 所述，利用定点诱变，用引物 S13SUBTETO-23FW (SEQ ID NO: 19) 和引物 S13SUBTETO-23RV (SEQ ID NO: 20)，通过置换插入 tet 操纵基因，产

生 pS13s- 23/pminiHXGPRT。

如前所述（实施例 4），将环形 pS13s- 23/pminiHXGPRT 质粒电穿孔到 RH Δ HXGPRT 速殖子中。如 Donald 等. (1998, 上文) 所述，感染到 Vero 细胞单层后，开始霉酚酸/黄嘌呤筛选。

根据 Kim, K., 等. (Science 262 (5135): 911-4 (1993)) 产生稳定转染子后，挑选几个克隆的寄生虫株系。从这些克隆的每个克隆分离基因组 DNA。利用引物 M13-REV (SEQ ID NO: 23), S13CL FW3 (SEQ ID NO: 24), HRCHECK II 5 S13-FW (SEQ ID NO: 25), HRCHECK II S13-RV (SEQ ID NO: 26), 和 T7 (SEQ ID NO: 27), 对这些基因组 DNA 样品进行 PCR 分析以检查这些转染子中假二倍体形式的存在或不存在。详细地分析了 4 个克隆 (c4, c5, c6 和 c9), 并且株系 RH Δ HXGPRT 和 Vero 细胞的基因组 DNA 用做阴性对照。使用不同的引物组合 (图 6) 以通过 PCR 扩增这些样品的基因组 DNA, 这些引物列为: 23/24, 25/26, 23/26, 和 25/27, 表示 SEQ ID NO: 23 和 24 引物组合, 等等。图 7 中示出了该结果。

引物组合 23/24 显示不同克隆中质粒的存在。引物 M13-REV (SEQ ID NO: 23) 与非转染寄生虫 (RH Δ HXGPRT) 缺少的载体部分退火。所有转染克隆都显示了正确大小的条带 (2.8kb), 表明电穿孔后, 所有稳定的转染子都摄取了质粒, 并且在筛选期间保留了该质粒。随后, 引物组合 25/26 显示是否假二倍体形式存在于克隆中。在基因组上, 两种引物都位于载体中存在的 S13 部分的上游 (引物 HRCHECK II 5 S13-FW (SEQ ID NO: 25)) 或下游 (引物 HRCHECK II S13-RV (SEQ ID NO: 26))。如果假二倍体不存在, 将 PCR 扩增 “野生型” S13 位置, 这产生 ~2.6kb 的产物, 正如用克隆 c4 和野生型寄生虫 RH Δ HXGPRT 可以观察到的。这表明克隆 c4 是没有假二倍体的稳定转染子, 表明发生了非同源重组。克隆 c5, c6 和 c9 不存在 2.6kb PCR 产物, 表明这些克隆的确含有假二倍体形式。此外, 对于克隆 c5 和 c9, 可以观察到当假二倍体存在时所预期的约 10kb 产物。对于克隆 c6 没有检测到 10kb 产物。用引物组合 23/26 和引物组合 25/27 进行 PCRs 以证明 p13s-23/pminiHXGPRT 载体并置在 S13 基因座两侧。引物 M13-REV (SEQ ID NO: 23) 位于载体序列中, 引物 HRCHECK II S13-RV (SEQ

ID NO: 26)位于该载体同源 S13 部分下游基因组上。引物 T7 (SEQ ID NO: 27)位于载体序列中, 引物 HRCHECK II 5 S13-FW (SEQ ID NO: 25)位于该载体同源 S13 部分上游基因组上。在野生型的情况下, 引物组合 23/27 不与 DNA 退火, 因此没有 PCR 产物被扩增。在假二倍体情况下, 引物组合 23/26 产生 4.6kb 产物, 并且组合 25/27 产生 2.6kb PCR 产物。图 7 中所示的数据证明实际上, 阳性克隆显示了两种组合的正确条带, 而阴性样品没有观察到产物。

因此, 该 PCR 分析用于证实成功地进行了用打了就走 (hit-and-run) 诱变方法, 同源重组到例如 S13 基因座中。

附图说明

图 1: TubYFP/TR-sagCAT 构建体的描述:

图 1A: 全部序列: 相关区域标在序列下面; 限制性酶识别位点标在序列上面。

图 1B: TubYFP/TR-sagCAT 构建体相关特征和区域的列表。

图 1C: TubYFP/TR-sagCAT 构建体的图谱。

图 2: S13/lacZ 构建体的描述:

图 2A: 全部序列: 相关区域标在序列下面; 限制性酶识别位点标在序列上面。

图 2B: S13/lacZ 构建体相关特征和区域的列表。

图 2C: S13/lacZ 构建体的图谱。

图 3:

图 3A: rp-S13 启动子中 tetO 插入/置换:

核糖体蛋白质 S13 启动子部分序列, 也标明了相对于 STS 的+3 插入和-23 置换位点。也标明了编码区域的最初 3 个氨基酸。

图 3B: rp-L9 启动子中 tetO 插入/置换。

图 4: 在 1 μ g/ml 无水四环素 (Atc) 存在或 1 μ g/ml 四环素 (Tc) 存在的情况下, 没有抗生素, 测定用构建体 S13/LZ, S13i+3/lacZ 和 S13s-23/lacZ 电穿孔的 tubYFP/TR 稳定转染子的 LacZ 表达水平。OD 是

LacZ 表达的指标。横轴的标记表明使用了 1.25×10^6 个速殖子（原始制备数量的 50%）。

图 5: 测定用构建体 S13/LZ, S13s-23/lacZ (I) 和 S13s-23/lacZ (II) 电穿孔的不同株系 (RH, REP, tubYFP/TR) 中 LacZ 表达水平。

RH 表示没有 tet 阻抑蛋白基因的株系。REP 表示带有合成的 tet 阻抑蛋白基因 (Meissner) 的株系。TYT 表示带有融合 tet 阻抑蛋白基因 (tub-YFP-TR) 的株系。在这些比较试验中使用相等数量的细胞。如图中所标明的, 在四环素不存在或存在的情况下进行所述试验。

图 6: 第一步打了就走 (hit-and-run) 诱变后, 类型 I 和 II 假二倍体形式的形成。

图 7: 不同克隆基因组 DNA 上的 PCR 以测定假二倍体形式的存在。

序 列 表

<110> AKZO Nobel N.V.

<120> 减毒活寄生虫疫苗

<130> 2002-017-EP

<150> EP 02078953

<151> 2002-09-20

<160> 29

<170> PatentIn version 3.2

<210> 1

<211> 4834

<212> DNA

<213> 鼠弓形体

<400> 1

```

cctagttgtg ttcgcaacag tacaccgtcc tgagtgagtc gagaacatca gagatgagca      60
cacgcaatag cgggtccgccagggtgcatt tgtccacatc gggatgcaca gagtggcacg      120
agtcgcacaa aagcagatac tagagacaag gagagagtgc ggcctaacca gaattcgact      180
cagtttcttg acccattcgt tagggtcggt ctcagcctcc ttcaggattt ccgtcaagac      240
atctttgcta gcttcccgtc gcagacatga aaggcagtgt cacgcataaa gagccgattg      300
aaacgcagtc acagagatac gaagaaatca aagcccgtgg aaagcgaacg gctgggatgt      360
agctgagaaa gcaaattcac tggcgggtgca aagagccaat gaaatcaggg tcgcgtagag      420
gaactataaa acgtgaaaaa cgtgccttcc gagtctcgca aagggtgcgca tcgatcccac      480
atttgagaga aggttgcgag gcagtaataa gggcagggga gaggataaaa tccgatagac      540
ccagttcttg gtctcccaga acggggacag gaccggacgc ctgcaagggt ggatcacaaac      600
tccagaggca aagccgccac ggaggaacgg aatccatgac cgagtggaat tataacgaag      660
aggtgtttgt cgtcggaatg gtgccaagac aaaaaaaaaa aaatgttttag acgctcgact      720
gtgcactagc ggggggcggg gtgcaaaagg gacgagtgtg ctcagtgggt cgagagtaac      780
tgaaaaaacg gtgcaaaata tggagcctta cgtggagccg cagggggcag aacagatgtc      840
tcagaagaaa gtccgagaga acagaagaaa aacgagaaaa gtgatgggcg actcatgcag      900
agtggcgcgga cgagtctgtc tctcagacga gcttaccagt gctgggcgga ggtaaaggaa      960
agaagtcaag acgcggacct tgaggggggt ggacagcatg atgaatcgct gatgtatgta     1020
ctttagaagc gcaggagtta agagtcgagc ggcattggcag gacgaccagt tgtcctttat     1080
gcttcgcaga taggcaatat atctgctgct gagggcctca tttctggaga gttgcgttgt     1140
ccctgtcgtc gcctcatcct ttatctccgt gtttgtctct tccagggcag ccttctgact     1200

```

accgccaac	gggcttcctt	cttcggattc	catttgagat	agccgtagaa	gcagaggaag	1260
agccgtcaga	acgcttgccg	cggcagaaaa	acacttaaag	ggcgtcacaa	gattgatata	1320
ggcaagagga	atggacgtca	acaggctgat	tcataagtga	cgctcccagt	aagtggcgga	1380
cagccatgaa	aatgagcggc	cgagtttgca	gaaacagaga	aagaggctctg	catcctggcg	1440
aagagccgcc	ggacaccctg	cttctctttc	acagttcgta	ggtgccaaga	ccaggaccaa	1500
attatcgccc	ttcttagcaa	accttgagcc	gagttaccgg	agaggttagc	cgaaaaagaa	1560
tcgaaacgaa	gacgccattt	tttgtctcca	ttgcacacgg	acggaccgta	gcttgtctct	1620
cagcatatct	tacgacgttt	tgcggctggt	atcgctaaca	caccacaaag	agaaatgggt	1680
tatcgaaaaa	cttgttagcc	ggatggtaaa	gagatgcaga	aggcagtcgg	cagtaattcg	1740
gttttcgtca	gttgtggcgt	gctggcacac	tcacgttttt	ccagcgtcac	atgctgcctg	1800
attcacgcag	aaactgcatg	tgcgtgcgt	gtctcgctg	cctcaggatg	cccttgctgg	1860
ccgatagtga	ggaaggaaaa	acggctccag	caaaatgttg	gttctattcg	gcgagtgcgg	1920
gtattccttc	cacaaggctg	agacaccgtc	gagtgttttc	cttcggact	gaaccccgga	1980
aaagtcactt	tgcaccgtag	attccacgtg	ctccagcgcg	gctgtcaatt	ttcgacactg	2040
cgcgaaacggc	ttgccaacaa	gaccaggctc	gcgcgcccgg	cttttcacat	tcccgacggc	2100
ttatatacgg	aaggctttgc	caggcgtatt	ctggccgcgt	ggggtcgaaa	gaaagtcgaa	2160
aaagagcatg	cttgtcaagt	gcatgcggcc	atgtagggtg	ctaggacccc	tgttaaat	2220
ccagggtgcg	gggcaactaa	gtggcctctc	ttcgcgctcg	cttcggactg	ttctctgggt	2280
tggcctcgct	tcgccacaga	cacttgtcga	cgcgtctcag	ggagtctgag	cccgttgtat	2340
ttttttcgct	gtctttttgg	cggttcccg	ttccctctga	ctgccgactc	tccctctctc	2400
cgcctcgctg	ccaccatgaa	gtctgtttat	gcctgtgaga	ctatcaccat	ccctgcggga	2460
ggtaagtttc	tcgacctacg	agagggtgaa	ctgcggagaa	gacgaatgaa	acattgcccc	2520
gcttgatctt	tgaggagag	ttgccagatt	ctgcggctcc	acagccctcg	ttttttttcc	2580
tcccgcatgt	gttagatgtg	tcccgacccc	gagggaagcg	atcgacacgc	tgggaaggaa	2640
cgcccgatga	gcggaaaagt	tttggaattc	aggccccgat	gcgcaaagtg	gcaagtgtct	2700
tggacccac	tgaggaagcc	gaacagcagc	attttacaga	tcttcgccac	tgaggagggg	2760
ggcggatctg	ggaggtgaag	aggcgcgga	ccgtgttcca	cttggctttt	ctccgcattc	2820
gctgtgtctg	ctctgcgtca	aaaatccgca	tgctttgttg	tcattcaaag	aggcatctg	2880
ggcgcttgt	tctttgttct	gcgcgatcca	caacagtctg	acccgccaga	gaatacggtc	2940

tgttctgtcc ggtgactggc gatggggaaa tgggggaaac tgtgtcgtca gcgagtgaag	3000
gcgtttttta gtggaatttc tacattgtgc aagcacacag aagggtgtccc gtgctaatat	3060
ctggaacagt agattatgat taggtagtgg aacagggaga gcgtctgttg tacatcactg	3120
tctgcactcg tttgtactac aacgaagttg ttgatgcgt gaactgggtg tcgattgcat	3180
agacatagcg tggaaaagta gaagacaggg ttgtatgcga ggctctgtgt gcacctgttt	3240
catgtggaca agaccaccgg gcatatgctg gctgttgctt caacacgctg ccgaaacatg	3300
tcacggcggt gcgggggaaa ggagtcgttg tagaaacat agagagagtt gaggtagctc	3360
ttgatgtctc gcaaaaatgg gactggcacc tgttgtctgt gtcttcgatt aacacgagcg	3420
ccgccactgc gtttgatgct cgctaactgg gcagcgtcgt gtacgtacag ctggaatagc	3480
gtaatttgtgt gtttttgtac tctttctgtt tgagtttcat caaagagggc cagccacaaa	3540
atgggagcag ggggatattg gagggcatat gataagtgc ccctgtgtgc atacttctag	3600
aatagacagg aattcgagag cgaagctgtc tgaacagaga tctgcaggtt ctggtttgac	3660
tgtgtaggca ggtttctgta gcgacgggag tcgcaatgca gaggccgct tggatttgg	3720
tgtttcaaga tgtttgcatc cctctgagag caatcgcttt ttgtcctgtt ttgcgtgtct	3780
ctcggctgtg tgccctctga aagaaaatgt tgcattcggt tgcggttttc tgctgcagtc	3840
acgggtggatg tgaagtcgcg ggtggtgact gtgaagggca agtaggcgaa atcacgcgtg	3900
cattccgcca cctccctgtc gacatccaga agaccaagtc tggaaaccga ctgaaggctg	3960
agatgtggta tggaacctgc acagacctca gctgcatccg cacgctgtgc tctcacatca	4020
agaacatgtt cactgggtgtg atgaagaagt tccagtacaa gatgcgcttc gtgtatgcac	4080
attttcccat caacgtgaac atcagcggca acggaactgt cgtcgaaatc cgcaacttct	4140
tgggcgagga gcgtgtgcgg atcgtcaaga tgcttccggg agttaagtgc gagaaggcca	4200
caaacgtcaa ggatgaaatc gcgctcactg gaactgacgt cgagctcgtc tctcgatcag	4260
gtagaggctc cgaggaactg aaaaggggag tgggtgtccg gtatgcgcgc atctaatatg	4320
agttttggag gtgcggaggc agcaaggaag cgtatagatg tgggcattta tgaatgtgca	4380
tctatgttgg tgtaattctg tgtatgcctg attgcgacgt gcccacaacc acctccaggt	4440
tggaagaaga gagaaactga taacgggtgga ccccgagagc gggattaccg ggaactctcg	4500
gacggtcgtt gtatactcat ctacgtggg cgagggggag gtggtttgtc cttcgatgtt	4560
gccacagatt tggagggtgag gtgtcttcat atcctgcatg tgtgtctgca ccagcagata	4620
tgtaactgc caggtgagac acgttgtcga gccacaggta tttttgtgta tctgcattgc	4680
attaacatgg tttgtattct totgttctgt atgcttctct tcttcagcgg ctctcatcca	4740

tcaatcgact ttggtcagga gaaaggatat ccgtaaattt ttggatggca tctacgtgtc 4800
agagaccagc actgtcgaac aggacgcgtaaatg 4834

<210> 2
<211> 4338
<212> DNA
<213> 鼠弓形体

<220>
<221> misc_feature
<222> (565)..(566)
<223> n是a, c, g或t

<220>
<221> misc_feature
<222> (625)..(626)
<223> n是a, c, g或t

<220>
<221> misc_feature
<222> (639)..(640)
<223> n是a, c, g或t

<400> 2
atcgcatgac ctgatcacgc acggaagaa acgaatagtc gccatttgaa gtgagatccg 60
tgtgaacagg tcagatcacc gaatcggacg atatatgcac tgaaggcagc ggagccagct 120
gtaaacaaag aactggacag cagctaccgt agctgtagac ggacgcgact tcgagagcgt 180
ccacgtcaaa cctcaccgat ctgcacctca taaagaggca tgtgggctgg gagatacagg 240
ggtgaagaag gagagacaat ttgcgtaagg aggcgaagct ttcgatttcc aggtgcgatt 300
ggagtcgccg ccacaggaga cgcgaactcc tcaaaaacgg acacggagaa gccctgtgca 360
gagacaacgg aaagaatgtc ctgacgagag agttgcaaaa gaatgttgaa caattaaagc 420
aatgatgcag actcgaagat ctaacgcctc gcaggtctca acggttgctg tgatcgccca 480
tttacacagt ccttaagttg agtgcattgag aggcctttgca gctcaaggca acgctgtaaa 540
cagcagtgtt atgaatcggg tgccnntatt gaggcgtctg cgtctggctg gtccatcaag 600
ccaaaagacg cttgtaaaca ggatnntcca ttccaatggn gacagacagt ttggcaactg 660
tcatcacacg tgacgttaaa aggcaccgtt aagcgcattga caaggaaagg tcacccgcga 720
tttacacaca ccaggtgccg tagctgtcga tgaatgcgaa ttccagagtt tttctctccg 780
acactacata agctgtaaat gctcattctg tcattcgttg accgtgttta ctacggggaa 840
tcgagaaacg gaatatcaag aacacaggct gtcaaaagac accgcgaaac ctgcttgccg 900
aatctaacgg ttgcctctgg ccatttatgt gtttctcgcc tgtgccttgt tcgctgcaga 960

cacagcctga gtcgcagcg aggtaaatac gaagaaaaac ctgacgagct ctgtcagatc	1020
tgtacaagcg acagaagcgg attgacagag gagagtgcgc gacggtgacg agagtgagag	1080
tcgactacga agttagagga caccaggggtg gcgaatgtgc caatacgag cttgaaaggg	1140
tcgagatcga caatcgaaac tcacttcact cgttaaaciaa tcgagcggtt gctgcagggt	1200
ttgtttgggg caccgccct ttgcctccc acccatcgga ttcagccgc agtactccac	1260
cagcaaaaca gcatcgaggc cgtatgcctc gaagaagtct ccaacctgca aaagaaaggc	1320
accacgtcag agcaaggaaa tcaagactca accaggtgta cagacaccgc atacgtcgc	1380
caggaacccc ggtctaggac aactttgcta gtgctgtca aaaggtggaa cggagaaggc	1440
gagacagcag actggcgggtg ccagttcaaa tcaccactgc ctgaagcgcg gggaccgaga	1500
cagtttgca tgttggaatt cctgtgacgc acacactttg gaacttgctt gaataatcag	1560
aactttgtcg ggccgaagtc gttttgttc tcgtacgaag acgaggagag gaggcataatg	1620
cagggagtca ggaaccatcg acggatggtc gaaggaaaga aaagagagct gccgccggga	1680
agcgggacgg gaaagaagcg gcagccttgc caagaacgtc ggggtgtactg ccaccgaggg	1740
aggcgggcag ctttctgcag acgaacgcag aacggagcag ttttcttcg ctcttcgact	1800
cggcttcggt cccagcaggt tgctgcgcgt cggcgcttcc ggacgcttct gcgtgtggaa	1860
gaagcggccg gacaggcggg atgcgtttga agggaaatga gtctgcctgt ctgaaatcgc	1920
gtgaggcaac tgaatggctg gacgtgcgag gcgtcggctc ggatggacaa caaaagcgac	1980
gcgtcgagaa gaggcaagaa agggcaagct ggcgcgacag ccacgcaact ggcctgctgc	2040
gtcccttgca gctgccgaaa agaagtccag cagctgatgc tgcccaagag tggcccaccc	2100
tccgcagctg ctgtcgcgtt tcgcgacctt ctgtaaaaaa ctcttgagc tgccggcgtg	2160
catcttgctt gcggaaagca aggcattgat cgaccagtac gcttgccgaa cggagcaggc	2220
taaacctgc cggcaaccgc gctgtgcgg aaacatttcc atcagcagtc tcgcgttgac	2280
ctccactgag atgcacacaa gtgaaaagag aatcgtctgc agtttccagc gatagcgcgc	2340
gaaagaagcg gagccgcgga gaggcgccc cggaccgaa gcgggcccga gagcgcgaga	2400
ggcaatggag caaaaaaatt cctgcctaga ggagaatcgc cattcctggc cagggtcact	2460
gcatacagaa tccctccca tcgtgggtct gatgaaaag agaaaaggac accattgtgg	2520
tgtggagccg cggtgccgtc tcggctttcc cgttgggtcaa tgcagaagcg cgtccccagt	2580
gactagagcc gaaccaccgg cgactccaat aaaggggcct tctactccat tcaggggtgt	2640
cgcattgtaa aactgagttc tctaatacat ggcacacctc ggaaaaacac tactcacagt	2700
cgtggaatat cttaaatagt cgagcgtcc agagtaccaa gctgcaatgc ccagctgcct	2760

```

ctcggacaac accggtggac taggcagcgc atcaaggaga cccccccag aggcggcctt 2820
actccagaac atgcgaagga tgtcactgtt gatgcaaat cagggtactg accatcaagt 2880
gaacagtttc tagagtatcc acgtggctcc aaaaagtaaa gtatggcccc acatcaaacc 2940
gagacyagcc tgtcttcgtt agaaacgctc gtgccgatga cgtcagtgca tgcacgctgc 3000
agacacgacc caaagtagct gatacggatga ggaatgagca tgctgccgat tcaaaatcgt 3060
cagtcgcggc tacagcagtg ttttatgtac atcctcgttt tctttttatt caagaaaccc 3120
aggtaacatg tgcttgaagg gcggacttgg gtagaatggc cctcgaacgt tggcgccgag 3180
gggttcgggtg cagccttcca tctgggctg taatgtgtct tgcgtatttt ctctgcgggc 3240
ctgcagtcgg aactgccttc atagtccctc aacgcaacct gcatagcttt accaggaatt 3300
tccgtgtcct tcccttgggt gggaagcgat ggtttgcagg agatgcagct gtgcaacact 3360
gtaaaggctg gtttcagcaa acgggactct cccctcggtc actcccaggg agtgtcggcc 3420
aggtgacggc ctggtctgcg gtcacagaag gatgtccttt cgaagatgca gtccatgaaa 3480
cttcgcttcc tagtgactgg gtggagagga gtgaaccac agctctaggg acggatgcga 3540
cactctcaag cattccatcc gagctacctg agcctggaac gccgcggaac agcttctcat 3600
ggggcacccg gaaaaggtag tgacgagaga atgctttctc agaattgac cttcgccttc 3660
tgtgagaaac actgtcacga ggactgcaat cttttcaggt gtatcttttc tgtacagctg 3720
tgcagcgttc acgcttttga ccggagagta cccgtgatgt gacttgcggg ttgggtagag 3780
cctagtccag ctgtctgtca tcttctgaac gggagttagg aagccgtaac aactaatatg 3840
tgcacgcggg tttctccggt gtgcgcgtgg gcaggtctta tgcaagggtt cggttcccc 3900
ctggttctgg aaagctgacg atcaacaaca gagacgcagc cgactatctt caagacaatc 3960
catggtggat tcataattgc atcgccac tcatggaact gcagctggag aatgaatttg 4020
atatcattgc ggaggtaagc cacgttagct agcttcgggt gacagtcgat atgtttcgca 4080
atgatttctt acagcgtctt tctacggtt ttgcttgctt tcgacatcag gccacgggg 4140
gaggcctcgg cgggcagctt ggagcaatca tgcttgctgt cgctcgggag attgtgcgac 4200
agcgaccgga actgcgaccc cctcttcggc gacgaggtt tctgaccgta gacgcaagga 4260
aggtcgagag gaagaaattt ggtcttcgga aggcgagaaa gaaggaacag tacagcaaac 4320
ggtaggggtg gtaggtca 4338

```

<210> 3

<211> 3639

<212> DNA

<213> 鼠弓形体


```

<400> 3
cctcgagag attgtcagt catgacacaa ccgcgaaaag ccggcagccg cggtaatac 60
gggacgagga aaacgactga gcgtcacaac agaagcagcc gagtaaacgg cgaaggaaat 120
ggaaaggacc caagtaaaat ttcttgaaga atttcagcgc aacaactctg cgggttcttg 180
cgaatagagg aatttcactt cctcatcgtc tgatttatgc tttcatcatc tgccgctcaa 240
cagccgaata aacggttctc ggtogettcc ttaaaactcta cttcagtagt tgaaactctt 300
ttgcttcacg agccttcgtc tcagccctca ccgtcctgag ttctgtcttt gttgaggaaa 360
gctcccgctg aaaaaacagg actttgtttg cagattttca tgtgtactgg aaagtgagat 420
gtgacttggg gaagtccgct ttaaaatttc cattgttttc tcaaaatgaa aagtctaaaa 480
aatcgaagtg cgtgccccgc gaggaattcc cctctgcaga ttgttttgc atttatatgt 540
cgtttttacg gagaaaagtc ccaagctgct gtccttctc taactagatg ttgaacgcta 600
gcacatatgc accagatgct tctgaagtat acctaaacgc accttgggaa caactgtgct 660
cccattcata aaactcatal aagtcaccaa gcattgcata ccggtgagac ataacaacgg 720
aagctagact actccccct gttattgcac actatcgaaa aggattccta ggtttctatc 780
ctctgccttt tcctggggca cactgcagag aaactaccgt gcgcgctacc tcccgcgtg 840
cgaggcgata gcaaaacgct ttggaaggaa aaagtcgaga aatcgacgac tgcgtctctt 900
gaatccgaga gagggatcca accaccgag ttctctgcat gtgcagcatc tgcaagaacg 960
tgataatgca tgaactcgat catcgctta totgtgtgca tgcattttcg aaaaagaaag 1020
gcgttttctg cgcggagact cgcgcggagg caagacgaga ctttctctc ttccaaactg 1080
cgagccacgg gggcgcatgc aatttgaaca tcacgcaaaa tcccaaacg ggtggggtgg 1140
agccgcacaa ttttttggca tgcagcgtt agcctgagct gcggtggggg cttttgtcgc 1200
gagcgtgggg tgccgcgaga gagcaacgcg gcgctacgcg gccgacgggt cttctgggaa 1260
gcctcgcatc tcctcgacgg gttctccct caattctctt ctttctctg cgtcttctc 1320
aggtggcttc gtcacgggtt tttctctcgc cgttcgtgct ccgctgtgtg tccggagtgc 1380
cgcgacagat cgagggcggt ctccgctccc acctgcggt tcccaatttc gatttttctc 1440
cgtcaccatg gggcgcatgt acggtcctgg aaagggcagc tctgccagcg ctctccctg 1500
gcgcagaaa ccccgacat ggctgaaaat caagccgtcg gacgtcgaag agcacattgc 1560
caagcttgca aagaagggcc agacccctc ccaggtaact tggcggggaa gaaggagaa 1620
aaacgacgga gttgccgcg ctgcggtctg ggaacaacg ggaagtgc caggaaaaat 1680
acgcgttctc caggtcatcg ggggaaaacg ctgcgaatc ccagccctc gactctccgc 1740

```

agttgttttt cgcagtcctt tcgccctcgg accgttcagc ggagatgggg acgagaatcc	1800
tctccctccc tctgctgagt ttcccccgt ctctcgtgtc tcaaaaaagg ctgcagaaat	1860
gctcgtttgc ctcacagccg gacctctctg ttgaccaaag cgcctcagtcg ggatttcttg	1920
cgcgggggatg ggctgaagca acgaaacggg atggacgttt gtcggtttcc ccgctgtcgt	1980
gtgacttgct cgaggaacca aagaacggga acgagcggga ggaagccggg ggaaaatttg	2040
cgttcctccc cgcaaacct atccgcaaaa atgcctcgtt tcgcgaactg tggacggggg	2100
ggaacaacgc gttgtgttgt ttgtgcatat ctgactgaca cggtcctgcg cgtgggtggc	2160
tgggatccga gcggtcgaa gacagtctct cggaaattcg cgagcggacc tactgttgtt	2220
cactgaacgg ttgcgttctc ctgtgaaatc aacatggttt cttgtgcagt ttaccgaaag	2280
tgaggacgac atgttttttg tgaccgtggc ggggcgttc cgcgtcggcg gcaaacggga	2340
cgatccaatg cggcgaaacc gggggagagt cattgcacac gatgaagtcg cgaacgacca	2400
aggcaatttt ttccggaggtc aaatcacatc ttccgagaag tatgaatgca tgtggaaggt	2460
cgtctgtttt tccattctcc acgttcttct tgcttttggc tgcttgtgct tctctggcga	2520
cttacagatc ggcgtcacac tgagagactc cttcggagtg ccgcaggtta aatccgtcac	2580
tggcaacaag atcctccgta tcctcaagct ccaaggtagg ttacgcgtt aagggaaaca	2640
ctagtttcaa tcttctcgag aacactggag gggcggagat cggggcgag ctccttccag	2700
gctcaagaag gtctgggagt gaaaaacgaa cgcaaatgca tggacgttgt atatatgtat	2760
gcctggatac ggtgtgagg tagcgcctt ggcaggagca agtgtgaagt ttgcgtctgt	2820
ttggagaagg aatgacgccg cgtcgtcgcg agggcgttct ctgccttccc ggttctctgt	2880
tctttgagaa agaacgtttt tcgcgtttct ccgcagttag aggttccctt cgaagaggca	2940
cctagatcag tcgactcgtt cttgaggagg ctggccttcg tcagtgtgtc tgctgcttct	3000
ctcactgcaa cactgtctcc cttgaagaga tttagcgag atgctgattt tctggcgttc	3060
agctctctgc cgtcgccctc tccaacatgt caagaagcac gttgcttgct tccctctttg	3120
tccagcaaag tggagttttt gtatgcgtgc aatctatgca atcgagagct tgctgaagcg	3180
acgttgctct cctctctccc aagtgtatgc tctccgcgtt tcttcgtctg gttaaaaaga	3240
aacggcgtcg ctgtccttcc cttcgtggcg aactcggat tgtttctcaa atccgatcta	3300
ctgtcagccg tccaagtgcc tgtctgacct ctttctcga ctcctcagtg cacatttaga	3360
gcgcgtggaa gcgactgttc aagtcctcct ctcctctgtt tctctagggt ctgaagagcg	3420
ccaagtgcgt ttccgaggtc ctccagaccc ggcgccacca gtgttctccc gactgttctt	3480
ttttttcagg tcttgcccc gagctgcctg aggacttgta ctacttgatc aagaaagccg	3540

tgagcgtgcg aaagcacttg gagagaaaca ggaaggacaa ggacgccaaag ttccgtctga 3600
 ttcttgttga gtctcgaatt caccgtcttg ctcgctaca 3639

<210> 4
 <211> 2748
 <212> DNA
 <213> 鼠弓形体

<400> 4
 ctgccgcttc cttaacctct ggaaggggtt gaagcttttt cgactgaaa cgcgagagac 60
 acgaatgagc tgaccacttt tcttgcattc gctggccctg taccggccgc atttcttact 120
 tcgtacacct tcaactgtact cacacccgaa aacttcagaa gtccgggcttt gctgcaggcg 180
 actcagggca gaggagaagg tgaacatgcc tcccccaatt ttgcccgcag tttgcctcgg 240
 gttgctctct tccccaacct agatgaaatg aaaccactcg atcagccatg tatttccgcg 300
 aaaagggttg ctaccaagcc cacatttggc aagcactctg gaagatgcgg cgccacggaa 360
 gcaaccctcg accgcaccgc tgtcgcgggg tctcgacgtc gtccgcgtta acatgatgtc 420
 ccgccagagc ttccactggt ctgacgagta catggatgca ataagcaagt cgtctccact 480
 cgtacacgaa tgtccagcac cagcaagctg catcacttca atcgcttggc agcccaacgc 540
 ttcttttgtc tgctttctgt ttttgctgcg agactgaccg agcacgaagc ttacacgaaa 600
 cgcgaagctc taatcgcta catccactc cttcttcaga aatggcaggg aagctacact 660
 cttgttaacg tctctagggt taaaggggtt gtgcaccgag ccgatcacgt tacggacacg 720
 ggctcggggc tcaactcgtt acaaggaata aacaactaac gcacaatcct ggttacattc 780
 gggcccacag catataaccc ctccggaggc ttgtattcca gctatcgaaa aaaaagcaat 840
 tcgatgtaat tccctccaat agccccgagc gyatgtcatc tacaagtggc agccttggta 900
 gggacccctc ttgggtatgc ccggacagat gcgccagtga gatctttaac ctccgcgtaa 960
 agataggtgt ctgctgtagc ggtgcgcttt tttgtgttgc atgcatgcat caagctggcc 1020
 gggacaaggc ggttgccccg atggatggat ggcaaaacca ctgtgctgca ggcagcagcc 1080
 ctccctcgga ggctcctctt gtggggcacg gcgcatgcc cagcacaacg tagcggcctt 1140
 gtccagcatg gacgaaaagc agtgggggag actcccagag gaaagcggtt ccatgcaaag 1200
 gggaaacagg ggatttttgt cacgacgagc tgtttgtcca cttttgtgag gtggtaattt 1260
 gacacagttc tcatccctgt tttgtccaag atggcgctga ccggcaacgt aaggggattc 1320
 tccgcactgc attgctctct tgagggagaa aaggccgcg tgaaaacatg ctgttctcca 1380
 ccagttggca gattgaagca gtctacggga agatgcacgg tttaatcatt gttgaatttc 1440
 gtctgcagtc tgactcttcc gtattggagc acgcgatgtc tcgttgcggt tgaaatcgta 1500

```

ctgtctggat ctctttgagt gaagaacacg cggcaggccg cagttttttt gcagggcctt 1560
ggtaccaaat gcttgtttac atatttgcc tgcgagttc ttttgcattg ttttagatat 1620
gcgtggagac tgttaaata acaaccgctc ggagatattg tgcgcggccc agcaatgctc 1680
tggttccact cccgtcgtga gcaggggaacg catgggtgggc ttttgtggct tctgtgtgta 1740
tgccgtctgc agacttgcaa aaagagaaag ttcgtcaagg atggtgtctt ccaggcggag 1800
ctcaatgagt tcctctcctg cacactgtcc gaggatgggt actcgggagt tgaagtccgt 1860
gtgactccca tccgcacaga gatcatcatc cgcgccacca ggactaggga agtgctcggc 1920
gacaagggaa ggcgtatccg cgaattgacg tcggtcgttc agaagcgatt cggcttcgcg 1980
cccgaactcg ttgagctctt cgcgcgagct gtggagaacc gtggtctgtg cgccatggca 2040
caggcagaat cgttgccgta caagcttcta aaggggcttg cagtcagacg cgcttctat 2100
ggtgtcctcc gccacatcat ggagtccgga gctaaagggt agtgctgaca aagtgccatg 2160
tattgtatga ggtaacttga atttagagtg tgaacaaaaa gcattagtcg actgtcacac 2220
gtatcttcgc cggacttttt tcttttcagg ttgcgaggtc gtcgtgtccg gtaaacttcg 2280
cgctcagcgt gccaaagagca tgaagttcaa ggatgggttac ctgatctcta ctggagagcc 2340
ctcgaagatg ttcgtcgacc aagcaatccg ctccgtgcaa ctccgacaag taagtttcaa 2400
attattaagc ctcaattacg tagtaaaggg caatttgtgt aggagctagt atgtacagag 2460
gcagtgtatg tgtgtttttt ttgcaggggt ttcttgggtg tagagtcaag atcatgctgc 2520
cgcatgaccc ggagggcaaa cgtggccccc cgaaccgcgt gccggatact attatcgtga 2580
tggatcccaa gccagagatc cccgttgtgc agcctgagga gatggacgag ggagtgtctg 2640
gtccaatgta atgagtgatt cgtgcgtgac tgttgattta tgggaggagg gtgtccacat 2700
gtgcgtgacc gtggagcagc cgcttaacga aattcgcattg ctcttcg 2748

```

<210> 5
 <211> 31
 <212> DNA
 <213> 人工

<220>
 <223> 引物: SAG3-FW

<400> 5
 cgataagctt cgaatctctg aacggatgtg t

31

<210> 6
 <211> 33
 <212> DNA
 <213> 人工

<220>
<223> 引物: TUB5-RV

<400> 6
cgagatctgg gaattcaaga aaaaatgccca acg 33

<210> 7
<211> 30
<212> DNA
<213> 人工

<220>
<223> 引物: TETAVR5-FW

<400> 7
cgatcctagg atgtctagat tagataaaag 30

<210> 8
<211> 33
<212> DNA
<213> 人工

<220>
<223> 引物: TETPST3-RV

<400> 8
cgtctgcagt taagaccac ttccacattt aag 33

<210> 9
<211> 21
<212> DNA
<213> 人工

<220>
<223> 引物: T3

<400> 9
attaaccctc actaaaggga a 21

<210> 10
<211> 31
<212> DNA
<213> 人工

<220>
<223> 引物: SAG1/1634-RV

<400> 10
cgataagctt tcgggggggc aagaattgtg t 31

<210> 11
<211> 27
<212> DNA
<213> 人工

<220>
<223> 引物: REV 13A

<400> 11
gcgccccatg gtgacggaga aaaatcg 27

<210> 12
<211> 27
<212> DNA
<213> 人工

<220>
<223> 引物: REV 13B (嵌套引物)

<400> 12
gggaaccgca aggtgggagc ggagaac 27

<210> 13
<211> 30
<212> DNA
<213> 人工

<220>
<223> 引物: S13PROMFUS FW

<400> 13
gcataagctt cctcgcagag attgtcagtg 30

<210> 14
<211> 31
<212> DNA
<213> 人工

<220>
<223> 引物: S13PROMFUS RV

<400> 14
gcattctaga ggcagacatg ccctttccag g 31

<210> 15
<211> 33
<212> DNA
<213> 人工

<220>
<223> 引物: LACZ-AVR11 FW

<400> 15
cgatcctagg atgacctga ttacggattc act 33

<210> 16
<211> 31
<212> DNA
<213> 人工

<220>
 <223> 引物: LACZ-PSTI RV

 <400> 16
 cgatctgcag ttatTTTTga caccagacca a 31

 <210> 17
 <211> 50
 <212> DNA
 <213> 人工

 <220>
 <223> 引物: S13INSTETO+3FW

 <400> 17
 ggttctcccc tcaatcccta tcagtgatag agatctctct tcctttctct 50

 <210> 18
 <211> 50
 <212> DNA
 <213> 人工

 <220>
 <223> 引物: S13INSTETO+3RV

 <400> 18
 agagaaagga agagagatct ctatcactga tagggattga ggggagaacc 50

 <210> 19
 <211> 51
 <212> DNA
 <213> 人工

 <220>
 <223> 引物: S13SUBTETO-23FW

 <400> 19
 ctacgcggcc gacggtccct atcagtgata gagatcttcc tcgacgggtt c 51

 <210> 20
 <211> 51
 <212> DNA
 <213> 人工

 <220>
 <223> 引物: S13SUBTETO-23RV

 <400> 20
 gaaccggtcg aggaagatct ctatcactga tagggaccgt cgccgcgta g 51

 <210> 21
 <211> 32
 <212> DNA
 <213> 人工

<220>
 <223> 引物: S13NOTI-FW

 <400> 21
 cgatgcggcc gcgtcagtgc atgacacaac cg 32

 <210> 22
 <211> 32
 <212> DNA
 <213> 人工

 <220>
 <223> 引物: S13SACI-RV

 <400> 22
 gctagagctc ctgtaagtcg ccagagaagc ac 32

 <210> 23
 <211> 23
 <212> DNA
 <213> 人工

 <220>
 <223> 引物: M13-REV

 <400> 23
 aacagctatg accatgatta cgc 23

 <210> 24
 <211> 20
 <212> DNA
 <213> 人工

 <220>
 <223> 引物: S13CL FW3

 <400> 24
 cgatagtgtg caataacagg 20

 <210> 25
 <211> 21
 <212> DNA
 <213> 人工

 <220>
 <223> 引物: HRCHECK II 5 S13-FW

 <400> 25
 gtcgagtcct gtaggttcat c 21

 <210> 26
 <211> 21
 <212> DNA
 <213> 人工

<220>
<223> 引物: HRCHECK II S13-RV

<400> 26
ctccgaagga gtctctcagt g 21

<210> 27
<211> 17
<212> DNA
<213> 人工

<220>
<223> 引物: T7

<400> 27
aatacgactc actatag 17

<210> 28
<211> 32
<212> DNA
<213> 人工

<220>
<223> 引物: HXGPRT/BGLII-FW

<400> 28
cgatagatct aaaatggcgt ccaaaccat tg 32

<210> 29
<211> 31
<212> DNA
<213> 人工

<220>
<223> 引物: HXGPRT/PSTI-RV

<400> 29
cgatctgcag ttacttctcg aactttttgc g 31

图1A

1 gtggcacttt tcggggaaat gtgcgggaa cccctatttg tttatttttc taaatacatt
61 caaatatgta tcgcctcatg agacaataac cctgataaat gcttcaataa tattgaaaa
121 ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt
181 gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt
241 tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt
301 ttgcgccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcg
361 tattatcccg tattgaogcc gggcaagagc aactcggctg ccgcatacac tattctcaga
421 atgacttggg tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa
481 gagaattatg cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga
541 caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa
601 ctgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca
661 ccacgatgcc tgtagcaatg gcaacaacgt tcgcgaaact attaactggc gaactactta
721 ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac
781 ttctgcgctc ggcccttcg gctggctggg ttattgctga taaatctgga gcoggtgagc
841 gtgggtctcg cggatcatt gcagcactgg gccagatgg taagccctcc cgtatcgtag
901 ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga
961 taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt
1021 agattgattt aaaacttcac ttttaattta aaaggatcta ggtgaagatc ctttttgata
1081 atctcatgac caaatccct taacgtgagt ttctgttcca ctgagcgtca gaccccgtag
1141 aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa

```

1201 caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt
1261 ttccgaagggt aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc
1321 cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa
1381 tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa
1441 gacgatagtt accggataag gcgcagcggg cgggctgaac ggggggttcg tgcacacagc
1501 ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa
1561 gcgccacgct tccgaagggt agaaaggcgg acaggatatc ggtaagcggc agggtcggaa
1621 caggagagcg cagcaggagc cttccagggg gaaacgcctg gtatctttat agtcctgtcg
1681 ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctgcgcaggg gggcggagcc
1741 tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc tggccttttg
1801 ctccatggtt ctttctgctg ttatccctcg attctgtgga taaccgtatt accgcctttg
1861 agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcaggtca gtgagcaggg
1921 aagcgggaaga gcgcccaata cgcaaaccgc ctctccccgc gcgttgggcg attcattaat
1981 gcagctggca cgacagggtt ccgcactgga aagcgggcag tgagcgcaac gcaattaatg
2041 tgagtttagc cactcattag gcaccccagg ctttacactt tatgcttcgg gctcgtatgt
2101 tgtgtggaat tgtgagcggg taacaatttc acacaggaaa cagctatgac catgattacg
                                     KpnI
                                     -----
2161 ccaagcgcgc aattaaccct cactaaaggg aacaaaagct gggtagcggg cccccctcg
      >>.....T3 .....>>
                                     HindIII
                                     -----
                                     PstI
                                     -----
2221 aggtcgacgg tatcgataag cttgatatcg aattcctgca gcccccgaga cgcgtgttct
      >>.pSAG1.>
2281 aaccacaaac cttgagacgc gtgttccaac cagcaccctt gacacgcgtg ttccaaccac
      >.....pSAG1.....>
2341 gcaccctgag acgogtgttc tawccacgca cctgagacg cgtgtttctaa ccacgcaccc
      >.....pSAG1.....>

```

```

2401  tgagacgcgt gttctgcgcg acaatgtgca cctgtaggaa gctgtagtca ctgctgattc
>.....pSAG1.....>

2461  tcactgttct eggcaagggc cgacgacggg agtacagttt ttgtgggcag agcggttgtg
>.....pSAG1.....>

2521  cagcttttcg ttcttctcgg ttgtgtcaca tgtgtcattg tcgtgtaa acacggttgt
>.....pSAG1.....>

2581  atgtcggttt cgctgcacca cttcattatt tcttctggtt ttttgacgag tatgcatgag
>>>
>>.....SAG1 CDS.....>>
>>>
>>>

2641  aaaaaaatca ctggatatac cacggttgat atatccaat ggcctogtaa agaacatttt
>.....CAT CDS.....>

2701  gaggcatttc agtcagttgc tcaatgtacc tataaccaga ccgttcagct ggatattacg
>.....CAT CDS.....>

2761  gcctttttta agacgtaaa gaaaaataag cacaagtttt atccggcctt tattcacatt
>.....CAT CDS.....>

2821  cttgcgcgcc tgatgaatgc tcatccggaa ttccgtatgg caatgaaaga cggtgagctg
>.....CAT CDS.....>

2881  gtgatatggg atagtgttca cccttgttac accgttttcc atgagcaa ac tgaaacgttt
>.....CAT CDS.....>

2941  tcatcgctct ggagtgaata ccacgacgat ttccggcagt ttctacacat atattcgaa
>.....CAT CDS.....>

3001  gatgtggcgt gttacggtga aaacctggcc tatttcctta aagggtttat tgagaatatg
>.....CAT CDS.....>

3061  tttttcgtct cagccaatcc ctgggtgagt ttcaccagtt ttgatttaaa cgtggccaat
>.....CAT CDS.....>

3121  atggacaact tcttcgcccc cgttttcacc atgggcaaat attatacgca aggcgacaag
>.....CAT CDS.....>

3181  gtgctgatgc cgctggcgat tcaggttcat catgcggtt gtgatggctt ccatgtcggc
>.....CAT CDS.....>

3241  agaatgctta atgaattaca acagtactgc gatgagtggc agggcggggc ktaaktrac
>.....CAT CDS.....>
>>...>

3301  accgttgtgc tcacttttca aatcgacaaa ggaacacac ttogtgcagc atgtgcccc
>.....SAG1-I.....>

```

```

3361 ttataaagaa actgagttgt tccgctgttg cttgcagggtg tcacatccac aaaaaccggc
>.....SAG1-I.....>

3421 cgactctaaa taggagtgtt tccgagcaag cagcgaaagt ttatgactgg gtccgaatct
>.....SAG1-I.....>

3481 ctgaacggat gtgtggcgga cctggctgat gttgatcgcc gtcgacacac gcgccacatg
>.....SAG1-I.....>

3541 ggtcaataca caagacagct atcagttgtt ttagtcgaac cggtaacac aattcttgcc
>.....SAG1-I.....>
      HindIII
      -----
3601 cccccgaaag cttcgaatct ctgaacggat gtgtggcgga cctggctgat gttgatcgcc
>....>>
      >>.....SAG1-II.....>

3661 gtcgacacac gcgccacatg ggtcaataca caagacagct atcagttgtt ttagtcgaac
>.....SAG1-II.....>

3721 cggtaacac aattcttgcc cccccgaggg ggatccacta gttctagagc ggctctgagg
>.....SAG1-II.....>>
      >>.....vector.....>

3781 tcgacggtat cgataagcta gagcttcagc atcatctctg gaagcatccc ctgaactgcc
>.....vector.....>>
      >>.....pTUB1.....>

3841 tgagtctacc aagagcactg gogaaggctg tgagtagtcg gacaggcacg gtgactcatg
>.....pTUB1.....>

3901 ttgttggaca gtagcgagct ctgggttaac cgcattattca ctaactggct ccgtctgtgt
>.....pTUB1.....>

3961 gtcattagat actgaatcag gtaacgatac atgagcagca tcctcgtgtt ccaggcgcat
>.....pTUB1.....>

4021 gtctctgccc ggcttgcaac caaggacccg tgggtcatct ttgggtctct ccgtactggg
>.....pTUB1.....>

4081 tggtagaggt gaaactgtcg acgtggatgc agctgccctg cttagagtac ggacgaagtg
>.....pTUB1.....>

4141 aatagctcgg tctgcatgaa caaggggctc tgaggcccg cgtgatacga aaggtttgct
>.....pTUB1.....>

4201 ggctactgaa cataggtctc gcagtgcggg ggcatactcc agtcggcctt cgcgaaactt
>.....pTUB1.....>

4261 cgtgaccagg cgatcaacaa cggcgctcac gcgagttaa tcgtccaga gaaagccacc
>.....pTUB1.....>

4321 atagtgcacc atatactgoc ggcacatctt gctgaaagtc gtggcgtgtc gatcagctaa
>.....pTUB1.....>

```

```

4381 gtcctgtgtg tacttggaag tgagcagctc ttccagttcg gtccagtcac attgccatgc
>.....pTUB1.....>

4441 cctagccacc gtttcaccag cgtccgatgc aggtcacct agccgatacc tgtgaccgtt
>.....pTUB1.....>

4501 catatatgag ttgctattac atctgtoggc cgtaacgaca caaggagatg cggctggagg
>.....pTUB1.....>

4561 caacggggtt gtgagcacca ttttgactct gtcacggagt aaaaaacatt tattcgaact
>.....pTUB1.....>

4621 ttgtacgagc gcagtcagta gaagtcaacc acgcgtatca actacgctgc aattacagag
>.....pTUB1.....>

4681 gacagggaca gggaaaaaaa gccgaagagg ttgccgggtg taggagatga cgagacgttt
>.....pTUB1.....>

4741 aaacggccgc tggctattgt tcgggcctcg ttgagccaca agcgatcaag gtgaaaacaa
>.....pTUB1.....>

4801 agttaaatag ttatgctgga gcgattgccc tgctgaatct tcagatcgga cgaacagtca
>.....pTUB1.....>

4861 ttctggccca ctgtacttga tgtgttcgat gtaaatcaca ctcagtcgcy tgctcggtag
>.....pTUB1.....>

4921 caatcaagt gctcttttct ctcttttcta gacacggtaa gaacgcttat gaacacgcat
>.....pTUB1.....>

4981 acacgcatag tttttgctag aatgcagoga ccagatgtcg caaggctcgc tccccctcg
>.....pTUB1.....>

5041 actggagaat caagaaaaac ctgcgttgat cccaaacgta ctctgtggtt ggtgcaacca
>.....pTUB1.....>

5101 gaagtttcat actgatcaaa agccagtga cagctggggg acattgcagg tctggtgctt
>.....pTUB1.....>

5161 caagaagcgc tgaagaagaa agtggcgaaa cctcggcag ttgccttga agaggcgccg
>.....pTUB1.....>

5221 tgcattgaac ttctgaagtg cgtagtacct ggtctctaact gctgttttgt gtttcgctgt
>.....pTUB1.....>

5281 ctgggcagca gtagaatgct gtgccagaat tagccactat ttagacatt tatttacaca
>.....pTUB1.....>

5341 ttttttttct gatgaacttg gcttattcat tttttcaagt cttgccactg ggtggtggca

```

```
>.....pTUB1.....>

5401  tgagactcgc ttagatgtat gtgggtgttg caatcacgct ggatgctcgg cttatttctg
>.....pTUB1.....>

5461  agttttttgt ggttttgaca atgggaacga tttcagagct actatttcac gtggtaagggt
>.....pTUB1.....>

5521  tatgagccac taaaaaacg aagaaaaacg ctgtttgag aaacaatagc aaactgtttt
>.....pTUB1.....>

5581  togtcatagt aactcagcgc ccccttgccc cccccagca gtgagatgca agacaatcct
>.....pTUB1.....>

5641  ctctaccaca gcttttggtg cgtctgtttc aaattttcag cgctcgcaa aggcatcacg
>.....pTUB1.....>

5701  aacaacatta tgagagggag caggttgtg gggctggcg gtgcaggaat gtgttcctgc
>.....pTUB1.....>

5761  gaaaaaggcc tctgctgaga aggtcgtggc gtttgaana tatccgaggt agcaagact
>.....pTUB1.....>

5821  tgtttcagtg ctcccctttt gaagacctgc ggcggcagt cactgaagag taactccaaa
>.....pTUB1.....>

5881  tcacgcgggt gagacttggg tttttcgtt atccttcaga agagtgtgtt ttogtttaat
>.....pTUB1.....>

5941  togtcacaga ccacgaaaaa cgaaccatcg aagacgatca ctgcgtccgc gtgcatctgg
>.....pTUB1.....>

6001  atggatgacc cacatctgtt gcagcgtcg cagacatgca tgtcccgct tcgtgaaatt
>.....pTUB1.....>

6061  ctctgcatca gggagtgat caggaaatc cgtctcagcg ggtgacgtt ggggagcagg
>.....pTUB1.....>

6121  ccggctcgcg ggggcagtca gatgccgaag gcgtaactca ggaagcttgc gctcatcgca
>.....pTUB1.....>

6181  gaacaggggt ggtgcctgca ttgggtgcgg ttggtgatcc tggttggacc ggtggagatg
>.....pTUB1.....>

6241  cgcgcgcacg aaggggatgt gtcagaaaca tttgtttgt tctctgtgaa cttttagatg
>.....pTUB1.....>

6301  tggttaaaggc ggcgaatatt agcagagagt cctccttgtt ggattctctc ttgaatttcg
>.....pTUB1.....>
```

```

6361 ccctttcctt ctctttgcga gtctcgtaga gaacaagcac tcgttcgccg tccctgacga
>.....PTUB1.....>

6421 cgcaaccogc gcagaagaca tccaccaaac ggtgttacac aatcaccttg tgtgaagttc
>.....PTUB1.....>

                                     BglII
                                     -----
6481 ttgcggaaaa ctactcgttg gcattttttc ttgaattccc agatctaaaa tggtgagcaa
>.....PTUB1.....>>
                                     >>.YFP CDS.>
                                     >>>>

6541 gggcgaggag ctgttcacog ggggtgtgcc catcctggtc gagctggacg gcgacgtaaa
>.....YFP CDS.....>

6601 cggccacaag ttcagcgtgt cggcgaggcg cgaggcgcat gccacctacg gcaagctgac
>.....YFP CDS.....>

6661 cctgaagttc atctgcaoca cgggcaagct gccctgccc tggcccaccc tcgtgaccac
>.....YFP CDS.....>
                                     PstI
                                     -----
6721 cttcggctac ggcctgcagt gcttcgccc ctaccccgac cacatgaagc agcacgactt
>.....YFP CDS.....>

6781 cttcaagtcc gccatgcccg aaggctacgt ccaggagcgc accatcttct tcaaggacga
>.....YFP CDS.....>

6841 cggcaactac aagaccgcg cggaggtgaa gttcgagggc gacaccctgg ygaaccgcat
>.....YFP CDS.....>

6901 cgagctgaag ggcctcgact tcaaggagga cggcaacatc ctggggcaca agctggagta
>.....YFP CDS.....>

6961 caactacaac agccacaacg tctatatcat ggccgacaag cagaagaacg gcacaaaggt
>.....YFP CDS.....>

7021 gaacttcaag atccgccaca acatcgagga cggcagcgtg cagctcgccg accactacca
>.....YFP CDS.....>

7081 gcagaacacc cccatcggcg acggccccgt gctgctgccc gacaaccact acctgagcta
>.....YFP CDS.....>

7141 ccagtcgcc ctgagcaaag accccaacga gaagcgcgat cacatggtec tgctggagtt
>.....YFP CDS.....>
                                     AvrII
                                     -----
7201 cgtgaccgcc gccgggatca ctctcgcat ggacgagctg tacaagccta ggatgtctag
>.....YFP CDS.....>>
                                     >>.....>
                                     >>>>

7261 attagataaa agtaaagtga ttaacagcgc attagagctg cttaatgagg tcggaatoga
>.....TR CDS.....>

```



```

7321 aggtttaaca acccgtaaac tegcccagaa gctaggtgta gagcagccta cattgtattg
>.....TR CDS.....>

7381 gcatgtaaaa aataagcggg ctttgctcga cgccttagcc attgagatgt tagataggca
>.....TR CDS.....>

7441 ccatactcac ttttgccctt tagaaggga aagctggcaa gattttttac gtaataacgc
>.....TR CDS.....>

7501 taaaagtttt agatgtgctt tactaagtca tcgcgatgga gcaaaagtac atttaggtac
>.....TR CDS.....>

7561 acggcctaca gaaaaacagt atgaaactct cgaatatcaa ttagcctttt tatgccaaca
>.....TR CDS.....>

7621 aggtttttca ctagagaatg cattatatgc actcagcgt gtggggcatt ttactttagg
>.....TR CDS.....>

7681 ttgcgtattg gaagatcaag agcatcaagt cgctaaagaa gaaagggaac cacctactac
>.....TR CDS.....>

7741 tgatagtatg ccgccattat tacgacaagc tatogaatta ttgatcacc aaggtcaga
>.....TR CDS.....>

7801 gccagccttc ttattcggcc ttgaattgat catatcgga ttagaaaaac aacttaaatg
>.....TR CDS.....>
      PstI
      -----
7861 tgaaagtggg tcttaactgc agccacaga agctgccgt ctctcgtttt cctctctttt
>....TR CDS....>>
                                >>.....DHFR.....>

7921 cggagggatc agggagagtg cctcgggtcg gagagagctg acgaggggt gccagagacc
>.....DHFR.....>

7981 cctgtgtcct ttatcgaaga aaaggatga ctcttcattg gccatttcac acagtctcac
>.....DHFR.....>

8041 ctgccttgt tttcttttg tcaatcagaa cgaagcgag ttgogggtga cgcagatgtg
>.....DHFR.....>

8101 cgtgtatcca ctogtgaatg cgttatcgtt ctgtatgccg ctagagtgtt ggactgttc
>.....DHFR.....>

8161 tgtctgcccc cgacagcaga caactttcct tctatgcact tgcaggatgg tgcagcgaa
>.....DHFR.....>

8221 acgacggaga gaaaggagca ccctctcagt ttccctacga tgtgctgtca gtttcgactc
>.....DHFR.....>

8281 ttcaccgcga acgattggcg atacgtctct gttgacttgt taggctccga ccacgaagct
>.....DHFR.....>

```

```

8341 cccttaacta rataagccgc gacacctaag tgtacaccat ttgcagatcg ataattctgcg
>.....DHFR.....>

8401 accgctgaat ccgtccagat cagtaaaacc gcaccaccta agtgtaaacc ttgttttaggt
>.....DHFR.....>

8461 cgataaaatg ctaccaaccc ccaccacaaa togagccttg agcgtttctg cgcacgogtt
>.....DHFR.....>

8521 ggcctacgtg acttgctgat gcctgcctct gccattcat gccagtcagt gcgcataaaa
>.....DHFR.....>

8581 atgtggacac agtcggttga caagtgttct ggcaggctac agtgacaccg cggtaggggg
>.....DHFR.....>
      NotI
      -----
8641 gatccactag tctagagcgg ccgccaccgc ggtggagctc caattcgccc tatagtgagt
>.....DHFR.....>>
                                     <<...T7....<

8701 cgtattacgc gcgctcactg gccgtgttt tacaacgtcg tgactgggaa aaccttgccg
<...T7....<<

8761 ttaccaact taatgcctt gcagcacatc cccctttcgc cagctggcgt aatagcgaag

8821 agggccgcac cgatcgccct tcccaacagt tgccagcct gaatggcgaa tgggacgcgc

8881 cctgtagcgg cgcattaagc gggcggggtg tgggtggtac gcgcagcgtg accgctacac

8941 ttgccagcgc cctagcgcgc gctcctttcg ctttcttccc ttcttttctc gccacgttcg

9001 ccggtcttcc cgtcaagct ctaaateggg ggctcccttt agggttccga tttagtgcct

9061 tacggcacct cgaccccaaa aaacttgatt aggggtgatgg ttcacgtagt gggccatcgc

9121 cctgatagac ggtttttcgc cctttgacgt tggagtccac gtctcttaat agtggactct

9181 tgttccaaac tggaacaaca ctcaacccta tctcgggtcta ttcttttgat ttataagga

9241 ttttgccgat ttggcctat tggtaaaaa atgagctgat ttaacaaaaa tttaacgcga

9301 attttaacaa aatattaacg cttacaattt ag

```

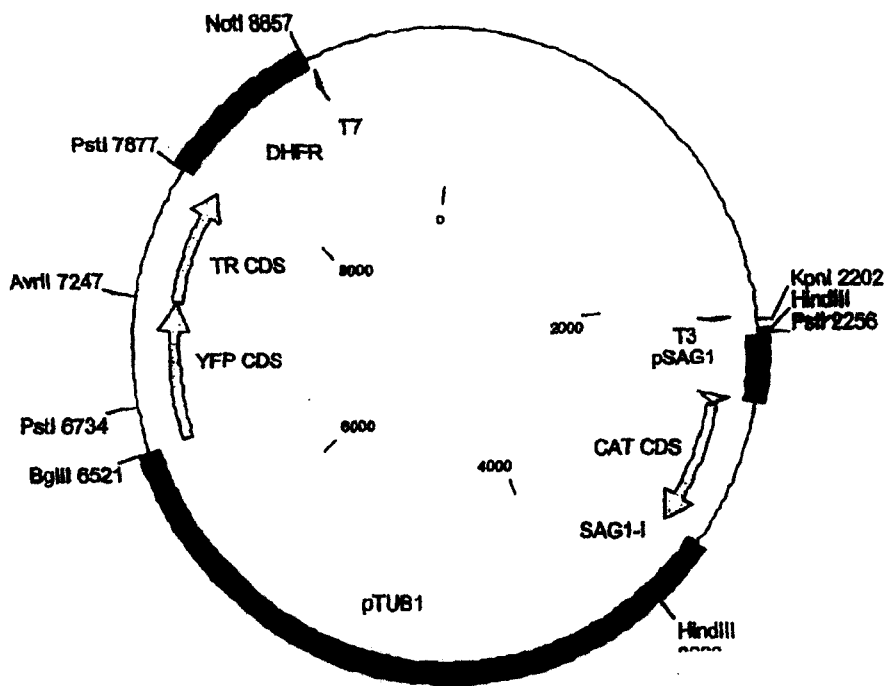
图 1B

TubYFP/TR-sagCAT

分子特征:

类型	起点	终点	名称	说明
基因	2172	2192	T3	用于测序的T3引物
区域	2271	2580	pSAG1	包含启动子的SAG1 5' 区
区域	2581	2583	SAG1 ATG-I	第一ATG
基因	2581	2634	SAG1 CDS	SAG1编码序列
区域	2632	2634	SAG1 ATG-II	第二ATG
基因	2638	3294	CAT CDS	氯霉素乙酰转移酶编码序列
区域	3295	3607	SAG1-I	SAG1 3' 非翻译区
区域	3614	3747	SAG1-II	用于起始tub启动子的3' 非翻译区 重复部分
区域	3748	3799	vector	未知载体部分
区域	3800	6520	pTUB1	包含启动子的TUB1 5' 部分
基因	6530	7246	YFP CDS	黄色荧光蛋白编码序列
区域	6530	6532	YFP ATG-I	第一ATG
基因	7253	7876	TR CDS	Tet阻抑蛋白编码序列
区域	7253	7255	TR ATG-I	第一ATG
区域	7886	8656	DHFR	DHFR3' 非翻译区
基因	8710	8690	C T7	用于测序的T7引物

图 1C



ptub^{YFP}TR_sagCAT

9332 bp

图2A

```

1  agcttctctcg cagagattgt cagtgcattg cacaaccgcg aaaagccggc agccgcggta
   >.....'启动子.....>

61  atacggggac gaggaaaacg actgagcgtc acaacagaag cagccgagta aacggcgaag
   >.....'启动子.....>

121 gaaatggaaa ggacccaagt aaaatttctt gaagaatttc agcgcaacaa ctctgcgggt
   >.....'启动子.....>

181 tcttgcaat agaggaattt cacttctca tcgtctgatt tatgctttca tcatctgcg
   >.....'启动子.....>

241 ctcaacagcc gaataaacg ttctcggtc cttccttaa ctctacttca gtagttgaa
   >.....'启动子.....>

301 ctcttttgc tcaagagcct tcgtctcagc cctcacgctc ctgagttctg tctttgttg
   >.....'启动子.....>

361 ggaaagctcc cgctgaaaaa acaggacttt gtttcagat tttcatgtgt actggaaagt
   >.....'启动子.....>

421 gagatgtgac ttggggaagt cgcctttaa atttccattg ttttctcaa atgaaaagtc
   >.....'启动子.....>

                                     PstI
                                     -----
481 taaaaaatcg aagtgcgtgc cccgcgagga attccctct gcagatttgt tttgcattta
   >.....'启动子.....>

541 tatgtcgttt ttacggagaa aagtccaag ctgctgctcc ttctctaact agatgttgaa
   >.....'启动子.....>

601 cgctagcaca tatgcaccag atgcttctga agtataccta aacgcacctt ggaacaact
   >.....'启动子.....>

661 gtgctcccat tcataaaact catacaagtc accaagcatg ccataccggt gagacataac
   >.....'启动子.....>

721 aacggaagct agactactcc ccctgttat tgcacactat cgaaaaggat tcttaggttt
   >.....'启动子.....>

                                     PstI
                                     -----
781 ctatcctctg ccttttctcg gggcacactg cagagaaact accgtgcgcg ctacctccg
   >.....'启动子.....>

841 acgtgcgagg cgatagcaaa acgcttttga aggaaaaagt cgagaaatcg acgactgcgt
   >.....'启动子.....>

```

```

901 ctcttgaatc cgagagaggg atccaaccca ccgagttctc tgcattgtca gcattctgcaa
>.....'启动子'.....>

961 gaacgtgata atgcatgaac tgcattcatcg ccttatctgt gtgcatgcat ttctgaaaaa
>.....'启动子'.....>

1021 gaaaggcgtt ttctgcgcgg agactcgcgc ggaggcaaga cgagactttc tctcttcca
>.....'启动子'.....>

1081 aactcgcgagc caggggggcg catgcaattt gaacatcacg caaatccca aaacgggtgg
>.....'启动子'.....>

1141 ggtggagccg caaacttttt tggcatgcag cgttgagcct gagctgcgtt ggggctttt
>.....'启动子'.....>

1201 gtcgcgagcg tgggtgcgcg cgagagagca acgcgcgcgt acgcgcgcga cgggtctctt
>.....'启动子'.....>

1261 gggaaacctc gcatttctc gacgggttct cccctcaatt ctcttcttt ctctgcgtct
>.....'启动子'.....>
>>.....e.....>

1321 tcttcaggtg gcttcgtcac cggttttct cctcgcgttc gtgctcgcgt gtgtgtccgg
>.....e.....>

1381 agtgcgcgca cagatcgagg gcgttctccg ctcccacctt ggggttccca atttcgattt
>.....e.....>

1441 ttctcgtca ccatggggcg catgtacggt cctggaaaag gcattgtctgc ctctaggatg
>.....e.....>
>>>
>>.....cds.....>
>>...>

1501 accatgatta cggattcact ggccgtcgtt ttacaacgtc gtgactggga aaacctggc
>.....'LacZ'.....>

1561 gttaccaaac ttaatgcct tgcagacat cccctttcg ccagctggcg taatagcgaa
>.....'LacZ'.....>

1621 gaggcccgca ccgacgccc ttccaacag ttgcgcagcc tgaatggcg atggcgcttt
>.....'LacZ'.....>

1681 gcctggtttc cggcaccaga agcgggtgcg gaaagctggc tggagtggca tcttctgag
>.....'LacZ'.....>

1741 gccgatactg tctcgtccc ctcaactgg cagatgcacg gttacgatgc gccatctac
>.....'LacZ'.....>

1801 accaagtgta cctatccat tacgggtcaat ccgcggttg ttccaacgga gaatccgacg
>.....'LacZ'.....>

```

```

1861 gggtgttact cgctcacatt taatgttgat gaaagctggc tacaggaagg ccagacgcga
>.....'LacZ'.....>

1921 attattttttg atggcgtaa ctggcggtt catctgtggt gcaacgggcy ctgggtcggt
>.....'LacZ'.....>

1981 tacggccagg acagtogttt gccgtctgaa tttgacctga ggcattttt acgcgccgga
>.....'LacZ'.....>

2041 gaaaaccgcc tcgcggtgat ggtgctgcgc tggagtgcg gcagtatatc ggaagatcag
>.....'LacZ'.....>

2101 gatattgtggc ggatgagcgg cattttccgt gacgtctcgt tgctgcataa accgactaca
>.....'LacZ'.....>

2161 caaatcagcg atttccatgt tgccactcgc tttaatgatg atttcagccg cgctgtactg
>.....'LacZ'.....>

2221 gaggtgaag ttcagatgtg cggcgagttg cgtgactacc tacgggtaac agtttcttta
>.....'LacZ'.....>

2281 tggcagggtg aaacgcaggc cgcagcggc accgcgcctt tcggcggtga aattratgat
>.....'LacZ'.....>

2341 gagcgtggtg gttatgcga tcgcgtcaca ctacgtctga acgtcgaaaa ccgaaactg
>.....'LacZ'.....>

2401 tggagcgccg aaatcccgaa tctctatcgt gcggtggtg aactgcacac cgcgcagcg
>.....'LacZ'.....>

2461 acgctgattg aagcagaagc ctgcgatgtc ggtttccgcg aggtgcggat tgaaatggt
>.....'LacZ'.....>

2521 ctgctgctgc tgaacggcaa gccgttgctg attcgaggcg ttaacogtca cgagcatcat
>.....'LacZ'.....>

2581 cctctgcatg gtcaggatcat ggatgagcag acgatggtgc aggatatcct gctgatgaag
>.....'LacZ'.....>

2641 cagaacaact ttaacgcgct gcgctgttcg cattatccga accatccgct gtggtacag
>.....'LacZ'.....>

2701 ctgtgcgacc gctacggcct gtatgtggtg gatgaaggca atattgaaac ccacggcatg
>.....'LacZ'.....>

2761 gtccaatga atcgtctgac cgatgatccg cgctggctac cggcgatgag cgaacgggta
>.....'LacZ'.....>

2821 acgcgaatgg tgcagcgca tcgtaatcac ccgagtgtga tcctctggtc gctggggaat
>.....'LacZ'.....>

```

```

2881 gaatcaggcc acggcgctaa tcacgacgcg ctgtatcgct ggatcaaata tgcgatccct
>.....'LacZ'.....>

2941 tccccgccgg tgcagtatga agggggcgga gccgacacca oggccaccga tattatttgc
>.....'LacZ'.....>

3001 ccgatgtacg cgcgcgtgga tgaagaccag ccttcccggt ctgtgcccga atggtccatc
>.....'LacZ'.....>

3061 aaaaaatggc ttccgctaac tggagagacg cgcgcgtgga tcctttgoga atacgcccac
>.....'LacZ'.....>

3121 gcgatgggta acagtcttgg cgttttcgct aaatactggc aggcgtttcg tcagtatccc
>.....'LacZ'.....>

3181 cgtttacagg ggcgcttcgt ctgggactgg gtggatcagt cgctgattaa atatgatgaa
>.....'LacZ'.....>

3241 aacggcaacc cgtggtcggc ttacggcggt gattttggcg atacgccgaa cgatcgccag
>.....'LacZ'.....>

3301 ttctgtatga acggtctggt ctttgccgac cgcacgcgc atccagcgct gacggaagca
>.....'LacZ'.....>

3361 aaacaccagc agcagttttt ccagttccgt ttatccgggc aaaccatoga agtgaccagc
>.....'LacZ'.....>

3421 gaatacctgt tccgtcatag cgataacgag ctccctgcact ggatggtggc gctggatggt
>.....'LacZ'.....>

3481 aagccgctgg caagcgggta agtgccctct gatgtcgctc cacaaggtaa acagttgatt
>.....'LacZ'.....>

3541 gaactgcctg aactaccgca gcggagagc gccgggcaac tctggctcac agtacggta
>.....'LacZ'.....>

3601 gtgcaaccga acgcgaccgc atggtcagaa gccgggcaca tcagcgctg gcagcagtg
>.....'LacZ'.....>

3661 cgtctggcgg aaaacctcag tgtgacgctc ccgcgcgct cccacgccat cccgcatctg
>.....'LacZ'.....>

3721 accaccagcg aaatggattt ttgcacgag ctgggtaata agcgttggca atttaaccgc
>.....'LacZ'.....>

3781 cagtcaggct ttctttcaca gatgtggatt ggcgataaaa aacaactgct gacgccgctg
>.....'LacZ'.....>

3841 cgcgatcagt tcaccgctgc accgctggat aacgacattg gcgtaagtga agcgaccgc
>.....'LacZ'.....>

```



```

3901 attgacccta acgcctgggt cgaacgctgg aaggcggcgg gccattacca ggccgaagca
>.....'LacZ'.....>

3961 gcgttggtgc agtgcacggc agatacactt gctgatgcgg tgctgattac gaccgctcac
>.....'LacZ'.....>

4021 gcgtggcagc atcaggggaa aaccttattt atcagcggga aaacctaccg gattgatggt
>.....'LacZ'.....>

4081 agtgggtcaa tggcgattac cgttgatggt gaagtggcga gcgatacacc gcacccggcg
>.....'LacZ'.....>

4141 cggattggcc tgaactgcc a gctggcgag gtagcagagc gggtaaaactg gctcggatta
>.....'LacZ'.....>

4201 gggccgcaag aaaactatcc cgaccgctt actgccgctt gttttgaccg ctgggatctg
>.....'LacZ'.....>

4261 ccattgtcag acatgtatac cccgtaogtc ttcccagagc aaaacggctc gcgctgcggg
>.....'LacZ'.....>

4321 acgcgcgaat tgaattatgg cccacaccag tggcgcgcg acttccagtt caacatcagc
>.....'LacZ'.....>

4381 cgctacagtc aacagcaact gatggaaacc agccatogcc atctgctgca cgcggaagaa
>.....'LacZ'.....>

4441 ggcacatggc tgaatatcga cggtttccat atggggattg gtggcgacga ctccctggagc
>.....'LacZ'.....>

4501 ccgtcagtat cggcggaatt ccagctgagc gccggctcgt accattacca gttggtctgg
>.....'LacZ'.....>

          PstI
          -----
4561 tgtcaaaaat aactgcagc cacagaagct gcccgctctc cgttttctc tcttttggga
>.....'LacZ'.....>
          >>.....DHFR .....>

4621 gggatcaggg agagtgcctc gggtcggaga gagctgaoga gggggtgcca gagaccctg
>.....DHFR .....>

4681 tgtcctttat cgaagaaaag ggatgactct tcattgtggc tttcacacag tctcactcg
>.....DHFR .....>

4741 ccttggtttc tttttgtcaa tcagaacgaa agcgagttgc gggtgacgca gatgtgcgtg
>.....DHFR .....>

4801 tatccactcg tgaatgcgtt atcgttctgt atgcogctag agtgctggac tgttgctgtc
>.....DHFR .....>

```

```

4861  tgcccacgac agcagacaac tttccttcta tgcacttgca ggatgggtgca gcgcaaacga
>.....DHFR .....>

4921  cggagagaaa ggagcacccct ctacagtttcc ctacgatgtg ctgtoagttt cgaactctca
>.....DHFR .....>

4981  ccgogaacga ttggogatac gtctctgttg acttggttagg ctccgaccac gaagctccct
>.....DHFR .....>

5041  taactarata agcccgacac cctaagtgtc caccatttgc agatcgataa tctgcgaccg
>.....DHFR .....>

5101  ctgaatccgt ccagatcagt aaaaccgcac cacctaagtg taaaccttgt ttaggtcgat
>.....DHFR .....>

5161  aaaatgctac caacccccac ccacaatcga gccttgagcg tttctgcgca cggttggcc
>.....DHFR .....>

5221  tacgtgactt gctgatgcct gcctctggcc attcatgcca gtcagtgcgc ataaaaatgt
>.....DHFR .....>

5281  ggacacagtc ggttgacaag tggtctggca ggctacagtg acaccgcggt ggagggggat
>.....DHFR .....>

                                NotI
                                -----
5341  ccactagttc tagagcggcc gccaccgcgg tggagctcca attcgcccta tagtgagtcg
>...DHFR .....>

                                <<....T7.....<
5401  tattacgcgc gctcactggc cgtcgtttta caacgtcgtg actgggaaaa ccctggcggt
<.T7..<<
>>.....pKS+.....>

5461  acccaactta atcgcccttc agcacatccc cttttcgcca gctggcgtaa tagcgaagag
>.....pKS+.....>

5521  gcccgacccg atcgcccttc ccaacagttg cgcagcctga atggcgaatg ggacgcgcc
>.....pKS+.....>

5581  tgtagcgccg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt
>.....pKS+.....>

5641  gccagcgccc tagcgccgcg tcctttcgct ttcttccctt cttttctcgc cactgtcgcc
>.....pKS+.....>

5701  ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt tagtgcttta
>.....pKS+.....>

5761  cggcacctcg accccaaaaa acttgattag ggtgatggtt cactagtggt gccatcgccc
>.....pKS+.....>

```

```
5821  tgatagacgg tttttcgccc tttagcgttg gagtcacgt tctttaatag tggactcttg
>.....PKS+.....>

5881  ttccaaactg gaacaacact caaccctatc toggctctatt cttttgattt ataagggatt
>.....PKS+.....>

5941  ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt taacgcgaat
>.....PKS+.....>

6001  ttttaacaaa tattaacgct tacaatttag gtggcacttt tcggggaaat gtgcgoggaa
>.....PKS+.....>

6061  cccctatttg tttatttttc taaatacatt caaatatgta tcogctcatg agacaataac
>.....PKS+.....>

6121  cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catttcogtg
>.....PKS+.....>

6181  tcgcccttat tccctttttt gcggcatttt gccttcctgt ttttgctcac ccagaaacgc
>.....PKS+.....>

6241  tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg
>.....PKS+.....>

6301  atctcaacag cggtaagatc cttgagagtt ttcccccga agaacgtttt ccaatgatga
>.....PKS+.....>

6361  gcacttttaa agttctgcta tgtggcgcgg tattatcccg tattgaagcc gggcaagagc
>.....PKS+.....>

6421  aactcggtcg ccgcatacac tattctcaga atgacttggg tgagtactca ccagtcacag
>.....PKS+.....>

6481  aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgcctgc ataaccatga
>.....PKS+.....>

6541  gtgataacac tgcggccaac ttactcttga caacgatcgg aggaccgaag gagctaaccg
>.....PKS+.....>

6601  cttttttgca caacatgggg gatcatgtaa ctgccttga tcgttgggaa ccggagctga
>.....PKS+.....>

6661  atgaagccat accaaacgac gagcgtgaca ccacgatgcc tntagcaatg gcaacaacgt
>.....PKS+.....>

6721  tgcgcaaact attactggc gaactactta ctctagcttc ccggcaacaa ttaatagact
>.....PKS+.....>

6781  ggatggaggg ggataaagtt gcaggaccac ttctgcctc gcccttccg gctggctggg
>.....PKS+.....>
```

6841 ttattgctga taamtctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg
>.....PKS+.....>

6901 ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta
>.....PKS+.....>

6961 tggatgaacg aaatagacag atcgctgaga taggtgcctc actgattaag cattggtaac
>.....PKS+.....>

7021 tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat ttttaattta
>.....PKS+.....>

7081 aaaggatcta ggtgaagatc ctttttgata atctcatgac caaatccct taacgtgagt
>.....PKS+.....>

7141 tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt
>.....PKS+.....>

7201 tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtgggtt
>.....PKS+.....>

7261 gtttgccgga tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc
>.....PKS+.....>

7321 agataccaaa tactgtcctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg
>.....PKS+.....>

7381 tagcaccgcc tacatacctc gctctgctaa tcctgttacc agtggtgct gccagtggcg
>.....PKS+.....>

7441 ataagtcgtg tcttacggg ttggactcaa gacgatagtt accggataag gcgcagcgg
>.....PKS+.....>

7501 cgggctgaac ggggggttcg tgcacacagc ccagcttga gcaacgacc tacaccgaac
>.....PKS+.....>

7561 tgagatacct acagcgtgag ctatgagaaa ggcgccgct tcccgaagg agaaaggcgg
>.....PKS+.....>

7621 acaggtatcc ggtaagcggc agggtcggaa caggagagcg cacgaggag cttccagggg
>.....PKS+.....>

7681 gaaacgcctg gtatctttat agtctgtcg ggtttcgcca cctctgactt gagcgtcgat
>.....PKS+.....>

7741 ttttgtgatg ctcgtaggg gggcggagcc tatggaaaa cgcagcaac gcggcctttt
>.....PKS+.....>

7801 tacggttcct ggccttttgc tggccttttg ctacatggt ctttctgcg ttatccctg
>.....PKS+.....>

```

7861 attctgtgga taacogtatt accgcctttg agtgagctga taccgctcgc cgcagccgaa
>.....pKS+.....>

7921 cgacogagcg cagcgagtca gtgagcgagg aagcgggaaga gogcccaata cgcaaaccgc
>.....pKS+.....>

7981 ctctccccgc gogttggcgg attcattaat gcagctggca cgacagggtt ccogactgga
>.....pKS+.....>

8041 aagcgggcag tgagcgcaac gcaattaatg tgagttagct cactcattag gcaccccgag
>.....pKS+.....>

8101 ctttacactt tatgcttcgg gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc
>.....pKS+.....>

8161 acacaggaaa cagctatgac catgattacg ccaagcgcg c aattaaccct cactaaaggg
>.....pKS+.....>>
>>.....T3 .....>

          KpnI                      HindIII
          -----                      -----
8221 aacaaaagct gggtagcggg cccccctcg aggtcgacgg tatcgata
>>

```

图 2B

分子: S13/lacZ, 8268 bps 环状DNA

分子特征:

类型	起点	终点	名称	说明
区域	1	1294	' 启动子	启动子区
区域	1295	1492	e	外显子1
区域	1453	1455	atg	ATG起始
基因	1453	1492	cds	基因
基因	1495	4578	'LacZ	来源于大肠杆菌的lacZ基因
区域	4582	5354	DHFR	DHFR3' 非翻译区
基因	5408	5388 C	T7	用于测序的T7引物
区域	5408	8202	pKS+	pKS+载体
基因	8202	8222	T3	用于测序的T3引物

图 2C

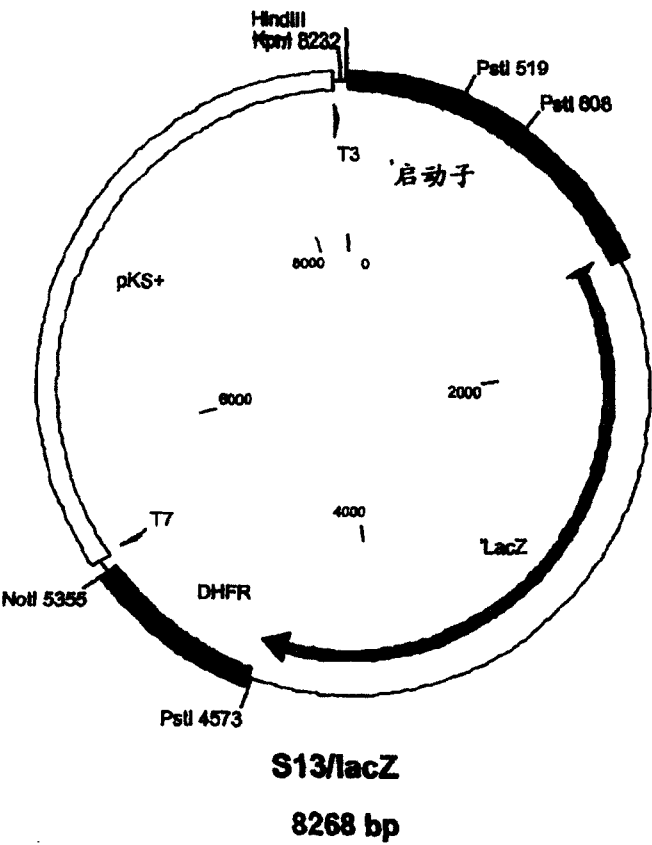


图 3A

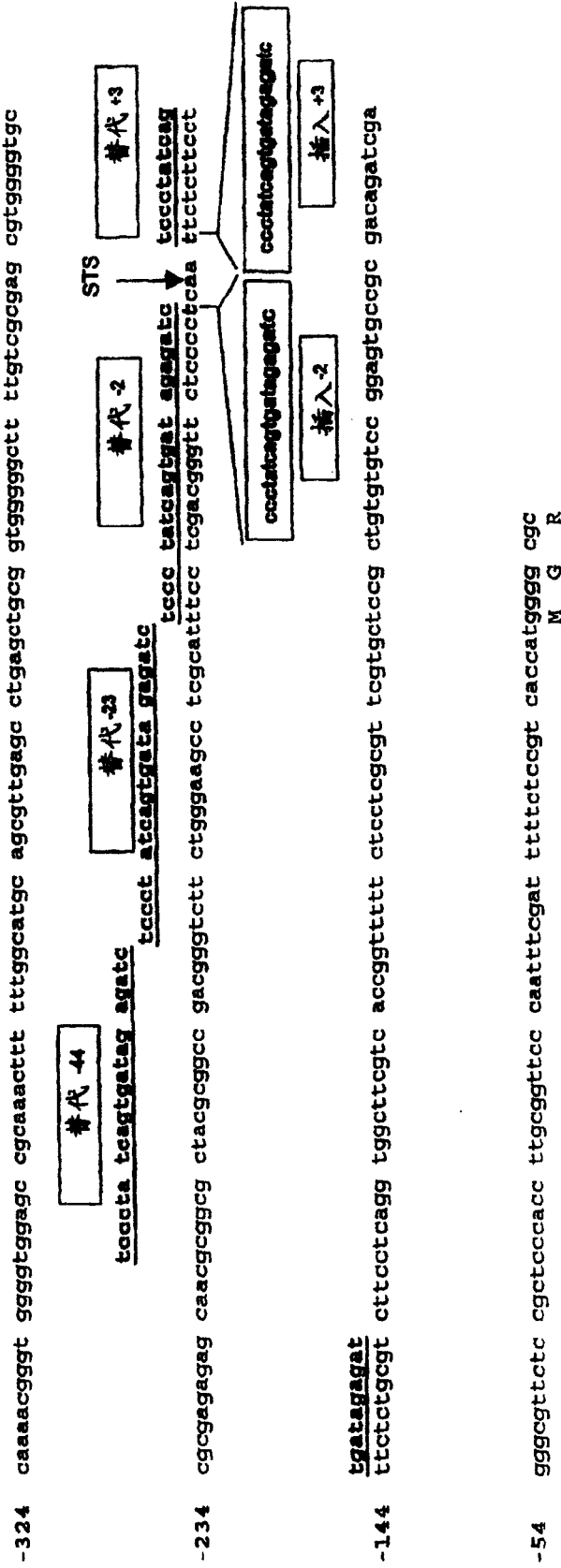


图4

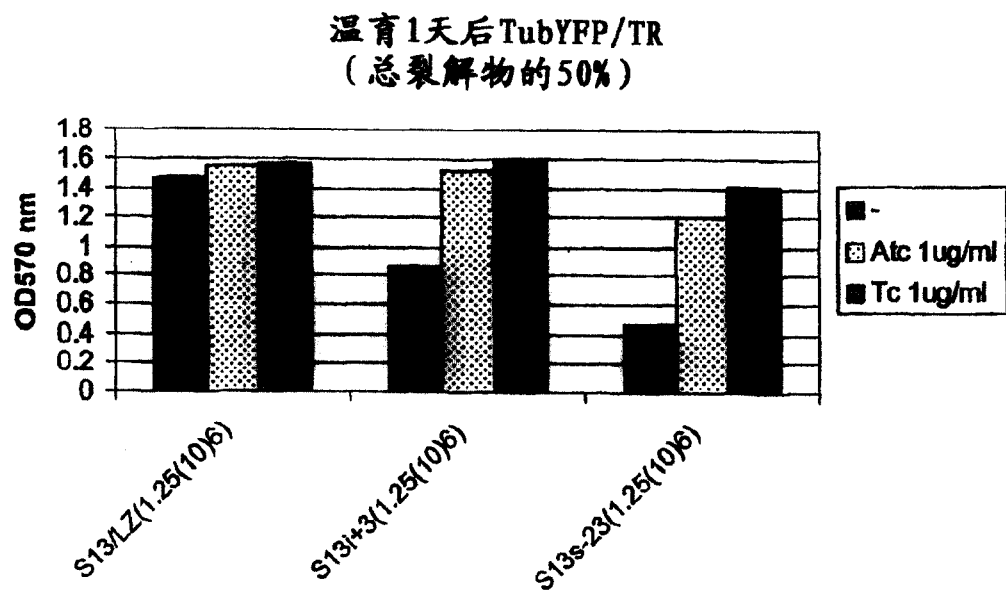


图5

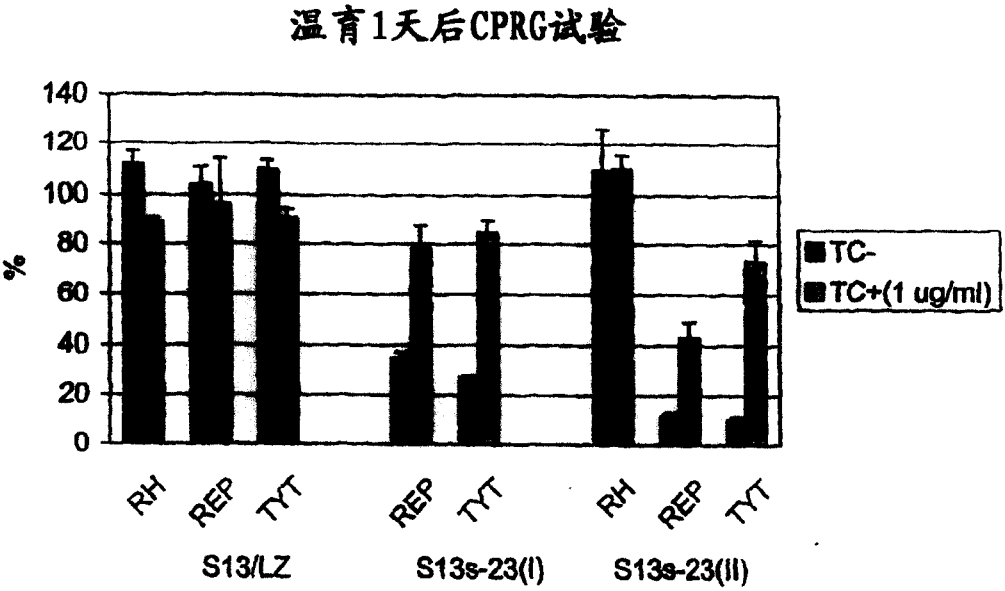


图6

