



(12)发明专利



(10)授权公告号 CN 103998604 B

(45)授权公告日 2016.12.28

(21)申请号 201280044383.1

(22)申请日 2012.07.12

(65)同一申请的已公布的文献号

申请公布号 CN 103998604 A

(43)申请公布日 2014.08.20

(30)优先权数据

11305909.1 2011.07.12 EP

12305450.4 2012.04.18 EP

(85)PCT国际申请进入国家阶段日

2014.03.12

(86)PCT国际申请的申请数据

PCT/EP2012/063640 2012.07.12

(87)PCT国际申请的公布数据

W02013/007772 EN 2013.01.17

(73)专利权人 特兰斯吉恩股份有限公司

地址 法国伊克什格拉芬斯特丹

(72)发明人 P·马丁 N·西尔韦斯特雷

J-B·马钱德

(74)专利代理机构 永新专利商标代理有限公司

72002

代理人 王健

(51)Int.Cl.

C12N 9/22(2006.01)

A61K 38/16(2006.01)

C07K 14/02(2006.01)

C12N 5/10(2006.01)

A61K 39/29(2006.01)

A61P 31/12(2006.01)

(56)对比文件

CN 1789279 A,2006.06.21,

RALF BARTENSCHLAGER et al.The P Gene Product of Hepatitis B Virus Is Required as a Structural Component for Genomic RNA Encapsidation.《 JOURNAL OF VIROLOGY》.1990,第64卷(第11期),5324-5332.

SUZANE KIOKO ONO-NITA et al.YMDD Motif in Hepatitis B Virus DNA Polymerase Influences on Replication and Lamivudine Resistance: A Study by In Vitro Full-Length Viral DNA Transfection.《HEPATOLOGY》.1999,第29卷(第3期),939-945.

审查员 滕文静

权利要求书5页 说明书36页

序列表48页 附图18页

(54)发明名称

HBV聚合酶突变体

(57)摘要

本发明涉及聚合酶HBV突变体多肽,其包含聚合酶活性被功能性破坏的突变的聚合酶结构域,以及涉及包含此种聚合酶突变体多肽的融合蛋白。本发明还涉及用于表达所述聚合酶突变体多肽的核酸分子和表达载体,以及涉及组合物,所述组合物可用于提高对HBV的免疫应答且目标在于提供针对HBV感染的保护性或治疗性作用。

1. 突变体多肽,其包含具有内部缺失的突变的聚合酶结构域,所述内部缺失功能性地破坏聚合酶活性,其中所述内部缺失是最少4个氨基酸残基以及最多30个氨基酸残基并且包含至少YMDD基序,所述YMDD基序天然存在于天然HBV聚合酶的聚合酶结构域中,其中所述突变的聚合酶结构域包含SEQ ID NO:1所示的氨基酸序列但缺少至少位置203处的Tyr残基、位置204处的Met残基、位置205处的Asp残基、位置206处的Asp残基、位置207处的Val残基、位置208处的Val残基和位置209处的Leu残基。

2. 权利要求1的突变体多肽,其中所述突变的聚合酶结构域包含SEQ ID NO:2所示氨基酸序列。

3. 权利要求1的突变体多肽,其还包含突变的RNaseH结构域,所述突变的RNaseH结构域包含功能性地破坏天然HBV聚合酶正常展现的RNaseH活性的一个或多个氨基酸残基的突变。

4. 权利要求3的突变体多肽,其中所述突变的RNaseH结构域中所包含的所述一个或多个突变选自:

-最少8个氨基酸和最多60个氨基酸的缺失,包括至少SEQ ID NO:3从位置39处的Glu(E)残基延伸至位置46处的Ala(A)残基的部分;

-用除了Asp(D)之外的氨基酸残基取代SEQ ID NO:3的位置10处的Asp(D)残基;

-用除了Val(V)之外的氨基酸残基取代SEQ ID NO:3的位置90处的Val(V)残基;

-用除了Thr(T)或Ala(A)之外的氨基酸残基取代SEQ ID NO:3的位置97处的Thr(T)或Ala(A)残基;

-用除了Asp(D)之外的氨基酸残基取代SEQ ID NO:3的位置98处的Asp(D)残基;和

-其任意组合。

5. 权利要求4的突变体多肽,其中SEQ ID NO:3的位置10、90、97或98处的所述取代的残基是用His(H)残基或用Tyr(Y)残基来置换。

6. 权利要求5的突变体多肽,其中SEQ ID NO:3的位置10处的所述残基是用His(H)残基取代,SEQ ID NO:3的位置90处的所述残基是用Tyr(Y)残基取代,SEQ ID NO:3的位置97处的所述残基是用Tyr(Y)残基取代和/或SEQ ID NO:3的位置98处的所述残基是用His(H)残基取代。

7. 权利要求4的突变体多肽,其中所述突变的RNase H结构域中所包含的缺失包含SEQ ID NO:3中从位置39处的Glu(E)残基延伸至位置63处的Leu(L)残基的至少25个氨基酸的部分。

8. 权利要求7的突变体多肽,其中所述突变的RNase H结构域中所包含的缺失包含SEQ ID NO:3中从位置31处的Xaa残基延伸至位置63处的Leu(L)残基的至少33个氨基酸的部分。

9. 权利要求8的突变体多肽,其中所述突变的RNaseH结构域包含SEQ ID NO:3中所示的氨基酸序列,但却(a)缺少从位置31处的残基Xaa延伸至位置63处的Leu(L)残基的33个氨基酸残基的部分并且包含(b)用His(H)残基取代位置10处的Asp(D)残基(D689H);(c)用Tyr(Y)残基取代位置90处的Val(V)残基(V769Y);(d)用Tyr(Y)残基取代位置97处的残基(T/A776Y)和(e)用His(H)残基取代位置98处的Asp(D)残基(D777H)。

10. 权利要求9的突变体多肽,其包含突变的RNaseH结构域,所述RNaseH结构域包含SEQ ID NO:4所示氨基酸序列。

11. 权利要求1的突变体多肽,其包含SEQ ID NO:5所示氨基酸序列。
12. 权利要求1的突变体多肽,与一个或多个另外的多肽或肽组合使用。
13. 权利要求12的突变体多肽,其中所述一个或多个另外的多肽或肽是选自HBc、HBs、X蛋白和其免疫原性片段的HBV多肽或肽。
14. 权利要求13的突变体多肽,其中所述另外的HBV多肽或肽源自基因型D HBV。
15. 融合蛋白,其包含权利要求1-14任一项的突变体多肽和HBV核心多肽。
16. 权利要求15的融合蛋白,其中所述HBV核心多肽是C-端截短的。
17. 权利要求16的融合蛋白,其中所述C-端截短的HBV核心多肽是在残基148处截短的。
18. 权利要求16的融合蛋白,其中所述C-端截短的HBV核心多肽是在残基149处截短的。
19. 权利要求15的融合蛋白,其中所述HBV核心多肽符合读框地融合至所述突变体多肽的N-端。
20. 权利要求19的融合蛋白,其中所述融合蛋白包含SEQ ID NO:6所示氨基酸序列。
21. 融合蛋白,其包含权利要求1-14任一项的突变体多肽和一个或多个HBsAg免疫原性结构域;或者权利要求15-20任一项的融合蛋白,其还包含一个或多个HBsAg免疫原性结构域。
22. 权利要求21的融合蛋白,在其N-端包含融合至所述突变体多肽的核心多肽和一个或多个HBsAg免疫原性结构域,所述一个或多个HBsAg免疫原性结构域融合至所述突变的聚合酶结构域中的内部缺失的位置中和/或融合至所述突变的RNaseH结构域中的缺失的位置中。
23. 权利要求22的融合蛋白,其包含SEQ ID NO:7-9任一项所示氨基酸序列。
24. 权利要求23的融合蛋白,其包含SEQ ID NO:8所示氨基酸序列。
25. 权利要求1-14任一项的突变体多肽或权利要求15-24任一项的融合蛋白,其符合读框地融合至信号肽和融合至跨膜肽。
26. 权利要求25的突变体多肽或融合蛋白,其包含SEQ ID NO:10-12任一项所示氨基酸序列。
27. 核酸分子,其编码权利要求1-14和25-26任一项的突变体多肽或权利要求15-26任一项的融合蛋白。
28. 权利要求27的核酸分子,其选自:
 - 核酸分子,其编码包含聚合酶结构域的突变体多肽,所述聚合酶结构域具有SEQ ID NO:1或2所示氨基酸序列;
 - 核酸分子,其编码包含RNaseH结构域的突变体多肽,所述RNaseH结构域具有SEQ ID NO:3或4所示氨基酸序列;
 - 核酸分子,其编码突变体多肽,所述突变体多肽包含SEQ ID NO:5所示氨基酸序列;或
 - 核酸分子,其编码融合蛋白,所述融合蛋白包含SEQ ID NO:6-12任一项所示氨基酸序列。
29. 权利要求27的核酸分子,其包含SEQ ID NO:13-17任一项所示核苷酸序列。
30. 载体,其包含权利要求27-29任一项的核酸分子。
31. 权利要求30的载体,其中所述载体是用于在高等真核细胞或生物体中表达的质粒或病毒载体。

32. 权利要求31的载体,其中所述载体是病毒载体,其源自反转录病毒、腺病毒、腺病毒相关病毒(AAV)、痘病毒、疱疹病毒、麻疹病毒、泡沫病毒、α病毒、或水疱性口炎病毒。

33. 权利要求32的载体,其中所述载体是复制缺陷腺病毒载体,源自人或黑猩猩腺病毒。

34. 权利要求33的载体,其中所述核酸分子被插入至腺病毒E1区并置于CMV启动子的控制之下。

35. 权利要求32的载体,其中所述载体是痘病毒载体,源自金丝雀痘、鸡痘或痘苗病毒。

36. 权利要求35的载体,其中所述痘苗病毒是Copenhagen毒株,Wyeth毒株和修饰的Ankara(MVA)毒株。

37. 权利要求36的载体,其中所述核酸分子被插入至MVA载体的缺失III处并且置于痘苗7.5K或pH5R启动子的控制下。

38. 权利要求32-37任一项的载体,其中所述载体选自:

-缺陷Ad载体,其包含插入至E1区位置的置于启动子控制之下的核酸分子,并且编码突变体多肽或融合蛋白,其中所述突变体多肽包含SEQ ID NO:5所示氨基酸序列,所述融合蛋白包含SEQ ID NO:6或SEQ ID NO:8所示氨基酸序列;

-复制缺陷Ad载体,其包含插入至E1区位置的置于启动子控制之下的核酸分子,并且包含SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:15所示的核苷酸序列;

-复制缺陷Ad载体,所述复制缺陷Ad载体包含插入至E1区位置的置于启动子控制之下的核酸分子,并且包含SEQ ID NO:16或SEQ ID NO:17所示的核苷酸序列;

-MVA载体,其包含置于痘苗启动子控制之下的核酸分子,并且编码突变体多肽或融合蛋白,其中所述突变体多肽包含SEQ ID NO:5或SEQ ID NO:10所示氨基酸序列,所述融合蛋白包含SEQ ID NO:6、SEQ ID NO:8或SEQ ID NO:12所示氨基酸序列;和

-MVA载体,其包含置于痘苗启动子控制下的核酸分子,并且包含SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:15所示的核苷酸序列。

39. 权利要求32的载体,其中所述载体是感染性病毒颗粒的形式。

40. 产生权利要求39的载体的方法,包括步骤:将病毒载体引入至适宜的细胞系,在适宜条件下培养所述细胞系从而使所述感染性病毒颗粒能够产生,从所述细胞系的培养物回收所产生的感染性病毒颗粒和任选地纯化所述病毒颗粒。

41. 宿主细胞,其包含权利要求27-29任一项的核酸分子或权利要求30-39任一项的载体。

42. 方法,用于重组产生权利要求1-14和25-26任一项的突变体多肽或权利要求15-26任一项的融合蛋白,包括步骤:将权利要求30-39任一项的载体引入适宜的宿主细胞以产生转染的或感染的宿主细胞,在适宜于所述宿主细胞生长的条件下体外培养所述转染的或感染的宿主细胞,回收细胞培养物并任选地纯化所产生的突变体多肽或融合蛋白。

43. 组合物,其包含至少权利要求1-14和25-26任一项的突变体多肽,权利要求15-26任一项的融合蛋白,权利要求27-29任一项的核酸分子,权利要求30-39任一项的载体,权利要求41的宿主细胞,或其任意组合。

44. 权利要求43的组合物,还包含药学可接受的载剂。

45. 权利要求43或44的组合物,其中所述组合物被配制用于肌内、皮下、皮内施用或划

破。

46. 权利要求43或44的组合物,其中所述组合物包含 5×10^8 、 10^9 、 5×10^9 、 10^{10} 、 5×10^{10} vp或 10^{11} vp剂量的腺病毒载体。

47. 权利要求43或44的组合物,其中所述组合物包含 5×10^6 、 10^7 、 5×10^7 、 10^8 或 5×10^8 pfu剂量的MVA载体。

48. 权利要求1-14和25-26任一项的突变体多肽,权利要求15-26任一项的融合蛋白,权利要求27-29任一项的核酸分子,权利要求30-39任一项的载体,权利要求41的宿主细胞,或权利要求43-47任一项的组合物在制备用于治疗或预防HBV感染或者HBV-相关疾病和病理状况的药物中的用途。

49. 权利要求1-14和25-26任一项的突变体多肽,权利要求15-26任一项的融合蛋白,权利要求27-29任一项的核酸分子,权利要求30-39任一项的载体,权利要求41的宿主细胞,或权利要求43-47任一项的组合物在制备用于治疗慢性HBV感染的药物中的用途。

50. 权利要求1-14和25-26任一项的突变体多肽,权利要求15-26任一项的融合蛋白,权利要求27-29任一项的核酸分子,权利要求30-39任一项的载体,权利要求41的宿主细胞,或权利要求43-47任一项的组合物在制备用于在所治疗的生物体中激发或刺激免疫应答的药物中的用途。

51. 权利要求50的用途,其中所述激发的或刺激的免疫应答是特异性的和/或非特异性的,体液的和/或细胞的。

52. 权利要求51的用途,其中所述免疫应答是针对HBV多肽/表位的CD4+或CD8+-介导的T细胞应答或者这两者。

53. 权利要求48-52任一项的用途,其中所述用途包括一次或多次施用治疗有效量的所述突变体多肽、融合蛋白、核酸分子、载体或组合物。

54. 权利要求48-52任一项的用途,其中所述载体是MVA载体并且所述用途包括彼此以一周间隔的三次皮下施用。

55. 权利要求48-52任一项的用途,其中所述载体是腺病毒载体并且所述用途包括一次或两次肌肉或皮下施用。

56. 权利要求48-52任一项的用途,其中所述用途与护理标准组合进行。

57. 权利要求56的用途,其中所述护理标准是核苷酸或核苷类似物(NUC)。

58. 权利要求57的用途,其中所述NUC选自拉米夫定、恩替卡韦、替比夫定、阿德福韦、阿德福韦酯或替诺福韦。

59. 权利要求48-52任一项的用途,其中根据初免-加强的疗法进行应用。

60. 权利要求59的用途,其中用MVA载体进行初免并用Ad载体进行加强。

61. 权利要求60的用途,其中所述MVA和/或所述Ad载体编码SEQ ID NO:8所示融合蛋白。

62. 权利要求60的用途,包括以3天到3个月不等的时间段间隔的至少3次皮下施用MVA载体,随后是腺病毒载体的肌肉或皮下加强。

63. 权利要求59的用途,其中以质粒DNA进行所述初免并以MVA载体进行所述加强。

64. 权利要求63的用途,其中所述质粒和/或所述MVA载体编码SEQ ID NO:8所示融合蛋白。

65. 权利要求63的用途,包括以2周到3个月不等的时间段间隔的至少3次肌肉施用DNA载体,和MVA载体的至少一次皮下加强。

66. 试剂盒,用于治疗HBV感染或用于在对象中激发免疫应答,其中所述试剂盒包含多种活性物质,所述活性物质选自:权利要求1-14和25-26任一项的突变体多肽,权利要求15-26任一项的融合蛋白,权利要求27-29任一项的核酸分子,权利要求30-39任一项的载体,权利要求41的宿主细胞,或权利要求43-47任一项的组合物。

HBV聚合酶突变体

发明领域

[0001] 本发明涉及聚合酶HBV突变体多肽,其包含聚合酶活性被功能性破坏的突变的聚合酶结构域,以及涉及包含此种聚合酶突变体多肽的融合蛋白。本发明还涉及用于表达所述聚合酶突变体多肽的核酸分子和表达载体,以及涉及组合物,所述组合物可用于提高对HBV的免疫应答且目标在于提供针对HBV感染的保护性或治疗性作用。本发明很特别的兴趣在于免疫治疗领域,更特别是用于治疗感染了HBV的患者,尤其是慢性感染的那些。

[0002] 发明背景

[0003] 乙型肝炎是重要的公众健康问题,全世界有超过3亿5千万人慢性感染,其中的20-40%有发展为慢性肝疾病、肝硬化和肝细胞癌的风险。尽管存在有效的保护性疫苗,但是乙型肝炎病毒(HBV)感染在许多国家(甚至是一些发达国家)依然泛滥,据估计全世界每年有四百五十万的新感染病例。与WHO所建议的进行全面疫苗接种不同的是,全疗程疫苗接种的覆盖为亚洲的25%至欧洲的75-90%不等。目前乙型肝炎是第十位的致死原因(每年大约一百万例死亡),而HBV相关的肝癌则排名第五位最常见的癌症。HBV感染的地理分布并不平衡,在西方国家的患病率低于1%,而在东南国家、非洲大部及赤道以南的美洲超过10%。

[0004] 乙型肝炎病毒是嗜肝DNA病毒科(hepadnaviridae)成员并主要感染肝,在肝细胞中复制。感染性颗粒是所谓的42-45nm的“丹氏粒(Dane particle)”,由外部的脂蛋白包膜(具有三个不同的表面蛋白(HB))和内部的核壳体组成,其主要结构蛋白是核心蛋白(HBcAg)。在核壳体内是连至病毒聚合酶蛋白(P)的单拷贝的HBV基因组。除42-45nm的病毒粒体外,HBV感染患者的血液还含有由HBsAg和宿主衍生的脂质所构成的20-nm的球体,其由感染的细胞释放。这些球体在数目上以 10^4 - 10^6 的量级超过病毒粒体。

[0005] 在病毒粒体进入肝细胞之后,通过目前仍未知的受体,核壳体将基因组HBV DNA转运至核,在此松弛环状DNA转变为共价闭合环状DNA(cccDNA)。所述cccDNA作为模板用于四个病毒RNA的转录,所述四个病毒RNA被输出至细胞质并被用作mRNA以用于HBV蛋白的翻译。最长的(前基因组)RNA也作为模板用于HBV复制,这在细胞质中于核壳体内发生。一些含有HBV DNA和聚合酶的壳体继而被转运回核,其在此释放新生成的松弛环状DNA以形成另外的cccDNA。cccDNA具有比肝细胞长的半衰期,其负责HBV的存续。其它壳体通过出芽至内质网内而被包膜并在穿过高尔基复合体之后被分泌。

[0006] HBV基因组的结构和功能组织已经被研究了超过30年。HBV基因组是松弛环状部分双链的DNA(大约3,200个核苷酸),其由全长的负链和较短的正链组成。其含有4个重叠的开放阅读框(ORF),C、S、P和X。C ORF编码核心蛋白(或HBcAg,构成核壳体的183个氨基酸长的蛋白)和病毒复制期间在患者血清中发现的第二种蛋白(已知为HBeAg,含有前核N-端延伸和部分的HBcAg)。核心蛋白的C-端很具碱性并且含有4个据推测结合核酸的Arg富集结构域以及许多磷酸化位点。S ORF编码三个表面蛋白,其全都具有相同的C端但在其N端由于存在三个符合读框的ATG启动子而有差异,所述三个符合读框的ATG启动子将S ORF分为三个区域,分别为S(226个氨基酸)、前-S2(55个氨基酸)和前-S1(108个氨基酸)。大表面抗原蛋白(L)在第一个ATG启动子处起始翻译之后产生,并且包含389个氨基酸残基(preS1-preS2-

S)。中等表面抗原蛋白(M)得自S区域和pre-S2区域的翻译(在第二个启动子ATG处起始),而226个氨基酸的小表面抗原蛋白(S,也命名为HBsAg)得自S区域的翻译(在第三个启动子ATG处起始)。HBV表面蛋白是糖蛋白,具有通过N-糖苷键附着的糖侧链。P ORF编码病毒聚合酶以及X ORF编码已知为X蛋白(其被认为是转录激活剂)的蛋白。

[0007] 病毒聚合酶根据HBV的基因型其长度约为832-845个氨基酸残基,并且其在长的开放阅读框(“P”)中编码,所述开放阅读框与核心基因和所有表面蛋白基因的3'端重叠。病毒聚合酶是由四个结构域组成的多功能蛋白,包括三个功能性结构域,分别是催化HBV复制中主要步骤(引发、DNA合成和RNA模板的去除)的末端蛋白、聚合酶和RNase H结构域,以及存在于末端蛋白和聚合酶结构域之间的非必要间隔结构域(参见例如Radziwill等,1990, J.Virol.64:613;Bartenschlager等,1990, J.Virol.64,5324)。负责酶活性的催化位点已经过表征。对于此点,形成保守的YMDD基序的四个残基(基于832个残基长的聚合酶编号的残基538-541)已经被显示为对DNA-和RNA-依赖性DNA聚合酶活性是必要的;而RNase H活性是基于DEDD基序,其涉及四个非保守氨基酸残基,分别为位置689处的Asp(D)、位置718处的Glu(E)、位置737处的Asp(D)和位置777处的Asp(D),以及其它几个氨基酸残基包括位置769处的Val(V)和位置776处的Thr(T)。现有技术中已描述了消除RT聚合酶和RNase H活性的不同突变(Chang等,1990, J.Virol.64:5553;Bartenschlager等,1990, J.Virol.64,5324, Radziwill等,1990, J.Virol.64:613和Chen等,1996, J.Virol.70:6151)。若干组已经成功地在各种宿主系统中表达了HBV聚合酶蛋白,但是已有报道其表达对表达细胞是毒性的,需要使用可诱导启动子(Choi等,2002, Antiviral Res.55:279;Karimi等,2002, J.Virol.76:8609)。

[0008] 许多临床前和临床研究着眼于CD4+和CD8+T细胞免疫应答对于有效抗病毒应答的重要性。确实观察到了已自然从乙型肝炎恢复的患者携有多特异性的和持续的由T辅助(T_H)和细胞毒T(CTL)淋巴细胞介导的应答,其可容易地在外周血液中监测到。抗-HBe和抗-HBs抗体的出现表明感染的良好结果。HBsAg-特异性抗体是中和性的,介导保护性免疫并且在临床恢复后终生保持。

[0009] 然而,慢性HBV感染很少由免疫系统解决。绝大多数慢性感染患者显示出弱的且短暂的CD4和CD8T细胞免疫应答,其在抗原方面受限并且对于病毒感染的清除无效。慢性乙型肝炎中细胞免疫应答作用功能的此种改变的原因目前还不是很清楚,虽然已经观察到其中涉及在HBV慢性感染患者中上调的不同的抑制性分子,如PD-1、CTLA4等。因此,需要能够诱导有效T-细胞应答的免疫调节策略。

[0010] 慢性乙型肝炎的常规治疗包括聚乙二醇化的干扰素-α(IFNa)和核苷/核苷酸类似物(NUC)如拉米夫定(lamivudine),以及最近的恩替卡韦(entecavir)、替比夫定(telbivudine)、阿德福韦(adeфовir)和替诺福韦(tenоfovir)。IFNa是有力的抗病毒分子,由此抑制病毒复制,然而,其在仅仅25-30%的患者中引起严重的副作用。NUC作用为HBV聚合酶的竞争性抑制剂,旨在抑制前基因组RNA逆转录成负DNA链及继而形成的双链病毒DNA。它们限制新病毒粒体的形成,但对于消除隐藏在感染的肝细胞的核中的超螺旋cccDNA无效,这构成了新的后代病毒的来源。这可以解释为何NUC的效力是短暂的并且在治疗停止之后立刻发生病毒反弹,需要患者终生接受治疗。另外,由于抗性HBV突变体的出现,长期效力也是有限的(据一些报道,拉米夫定治疗超过24%在一年后以及大约66%在四年后,虽然更新的NUC

显示少得多的药物抗性HBV突变体的出现)。现在已经分离出对抗病毒剂展现出降低的敏感性的许多HBV毒株,并且基因组测序反映出在聚合酶结构域中取代突变的显著性,包括在YMDD基序中(US2008-0233557;Zoulim and Locarnini,2009,Gastroenterology,137:1593)。

[0011] 在抗病毒治疗之外,目前还在努力尝试开发旨在改善宿主免疫应答的补充疗法,具体而言是那些由细胞毒T-和辅助T-淋巴细胞所介导的免疫应答。一些有前景的疫苗策略专注于HBV表面蛋白preS1和/或preS2(Zanetti等,2008,Vaccine26:6266;Mancini-Bourguine等,2006,Vaccine24:4482)以及专注于旨在同时靶向多个HBV抗原的多价免疫治疗方法。例如,用编码多个包膜、核心和聚合酶表位的多表位DNA疫苗免疫已显示出在临床前小鼠模型中激发CTL和T_H应答(Depla等,2008,J.Virol.82:435)。一种基于编码HBsAg、HBcAg和HBV聚合酶的DNA质粒的混合物的方法(WO2005/056051;WO2008/020656)在慢性乙型肝炎的转基因小鼠模型中证明了特异性抗-HBV细胞和体液应答(Chae Young Kim等,2008,Exp.Mol.Medicine40:669)。在韩国在HBV携带者中组合拉米夫定治疗开始了I期临床试验(Yang等,2006,Gene Ther.13:1110)。近来研究的另一方法涉及使用编码HBc和HBV聚合酶及Hbs免疫原性结构域的组合物载体化的治疗性疫苗(WO2011/01565)。用Ad-载体化的疫苗免疫的小鼠显示了针对所有表达的HBV抗原的T细胞应答,特别是针对聚合酶。

[0012] 由于感染的慢性和持续特性、其高发病率、HBV的连续传播以及相关疾病的显著死亡率,可能会预期HBV会持续许多年成为严重的全球健康威胁。因此,非常需要开发更有效的方法用以改善对HBV感染或HBV-相关疾病或病症的预防和治疗。特别是,还需要调和针对靶向的HBV抗原(特别是针对核心)的T细胞-介导的免疫的方法,且是低潜在毒性。此类方法对于治疗慢性感染HBV的对象尤其有用。

[0013] 此技术问题通过权利要求中所定义的实施方案而得到解决。

[0014] 通过如下对本发明目前优选的技术方案的描述,本发明的其它和另外的方面、特征及进展将会是显见的。提供这些技术方案是用于公开的目的。

[0015] 发明概述

[0016] 在一方面,本发明涉及突变体聚合酶多肽,其包含天然HBV聚合酶的至少500个氨基酸残基,其中所述突变体聚合酶多肽包含具有内部缺失的聚合酶结构域,所述内部缺失功能性地破坏聚合酶活性并且其中所述内部缺失包含至少YMDD基序,所述YMDD基序天然存在于天然HBV聚合酶的聚合酶结构域中。

[0017] 本发明还涉及编码所述突变体聚合酶多肽的核酸分子,包含所述核酸分子的载体,或者包含或编码所述突变体聚合酶多肽的组合物。

[0018] 本发明还涉及此突变体聚合酶多肽、核酸分子、载体或组合物的用途,特别是与另外的多肽组合(例如与一个或多个HBV多肽组合),以用于治疗、预防或抑制HBV感染或者改善与HBV感染相关的病况。

[0019] 本发明的另一方面包括用于在有需要的对象中治疗、预防或抑制HBV感染或者改善与HBV感染相关病况的方法,包括提供或施用此突变体聚合酶多肽、核酸分子、载体或组合物,最终与另外的多肽组合(例如与一个或多个HBV多肽组合)和/或与标准护理组合。

[0020] 本发明的另一方面涉及在有需要的对象中激发免疫应答的方法,包括提供或施用此突变体聚合酶多肽、核酸分子、载体或组合物,最终与另外的多肽组合(例如与一个或多

个HBV多肽组合)和/或与标准护理组合,以用于在此对象中诱导或刺激免疫应答用以治疗HBV感染或者改善与HBV感染相关的病况或症状。

[0021] 本发明的另一方面提供试剂盒,其包含多个容器和说明,以用于根据本文所述的组合物和方法向对象提供或施用此突变体聚合酶多肽、核酸分子、载体或组合物,最终与另外的多肽组合(例如与一个或多个HBV多肽组合)。

[0022] 发明详述

[0023] 本发明提供突变体聚合酶多肽,包含具有内部缺失的聚合酶结构域,所述内部缺失功能性地破坏聚合酶活性并且其中所述内部缺失包含至少YMDD基序,所述YMDD基序天然存在于天然HBV聚合酶的聚合酶结构域中。此种突变体聚合酶多肽或编码其的载体可以在用于治疗或预防HBV感染或与HBV感染相关的病况的组合物和方法中使用,最终与其它HBV多肽和/或护理标准组合。由于相关酶活性的破坏,本发明涉及在各种载体系统中表达和产生所述突变体聚合酶多肽。本发明也特别适于人类应用并且可以用于加强标准疗法(例如SOC)。用编码与HBc和Hbs免疫原性结构域融合的此突变体聚合酶多肽的载体免疫动物模型,激发了HBV特异性T细胞应答,并令人吃惊的观察到针对HBc和聚合酶二者的强免疫性。

[0024] 如下部分提供了本文所用的一些术语的更详细的解释。

[0025] 定义

[0026] 如本文遍及整个申请中所用,术语“a”和“an”所使用的含义在于“至少一个”、“至少第一个”、“一个或多个”或“多个”所提及的化合物或步骤,除非上下文另有说明。

[0027] 无论在本文中何处使用术语“和/或”,都包括“和”、“或”以及“所述术语所连接的元素的所有或任意的其它组合”的含义。

[0028] 术语“约”或“大约”如本文所用是指给定值或范围的上下10%之内,优选8%之内,更优选5%之内。

[0029] 术语“氨基酸”、“残基”和“氨基酸残基”是同义词并且涵盖了天然氨基酸以及氨基酸类似物(例如非天然的、合成的和修饰的氨基酸,包括D或L光学异构体)。

[0030] 术语“多肽”、“肽”和“蛋白”是指氨基酸残基的多聚体其包含至少9个或更多个经肽键连接的氨基酸。所述多聚体可以是线性的、分枝的或环状的,并且可以包含天然发生的氨基酸和/或氨基酸类似物,而且其可以被非氨基酸中断。作为一般表示,如果氨基酸多聚体超过50个氨基酸残基,其优选称作多肽或蛋白,而如果其长度为50个氨基酸或更短,则称作“肽”。

[0031] 如本文所用,当用于定义产品、组合物和方法时,术语“包含”(以及任何形式的包含,如“comprise”和“comprises”)、“具有”(以及任何形式的具有如“have”和“has”),“包括”(以及任何形式的包括如“includes”和“include”)或“含有”(以及任何形式的刚含有如“contains”和“contain”)是开放式的且不排除另外的、未提及的元素或方法步骤。因而,多肽“包含”氨基酸序列则该氨基酸序列可以是该多肽最终氨基酸序列的一部分。此种多肽可具有达数百个另外的氨基酸残基。“基本由……组成”是指排除任何有重要意义的其它组分或步骤。因而,基本由所提到组分组成的组合物将不会排除痕量的污染物和药学可接受的载体。多肽“基本由氨基酸序列组成”则此氨基酸序列最终仅存在几个另外的氨基酸残基。“由……组成”是指排除其它组分或步骤的痕量元素。例如,多肽“由氨基酸序列组成”则该多肽除了所提到的氨基酸序列之外不含有任何氨基酸。

[0032] 如本文所用,“HBV”和“乙肝病毒”可互换使用并且是指嗜肝DNA病毒科的任何成员(参见例如Ganem and Schneider在Hepadnaviridae(2001)中“The viruses and their replication”,pp2923-2969,Knipe DM等,eds.Fields Virology,4th ed.Philadelphia,Lippincott Williams&Wilkins或随后的版本)。氨基酸序列各种HBV多肽的和编码核苷酸序列可在专门的数据库找到(例如上文所述的那些)以及在文献中找到(参见例如Valenzuela等,1980,The amino acid sequences of the hepatitis B viral genome and the identification of the major viral genes(pp57-70)in“Animal Virus Genetics”;eds B.Fields,等;Academic Press Inc.,New York and Vaudin等,1988,J.Gen.Virol.69:1383)。

[0033] 如本文所用,术语“HBV聚合酶”是指保留至少500个天然HBV聚合酶蛋白中所包好的氨基酸残基的多肽。期望的是,此至少500个氨基酸残基分布于三个功能性结构域并优选分布于天然HBV聚合酶中天然存在的四个结构域。此术语涵盖可发现自、分离自、得自自然界中HBV来源的任意HBV毒株、分离物(isolate)或基因型的天然的(即自然发生的)聚合酶多肽,如上文与术语“HBV”相关联而提到的那些以及修饰的聚合酶(即突变体聚合酶多肽)及其片段。作为说明目的,本文所述的HBV聚合酶的氨基酸残基参照832个氨基酸长的聚合酶进行编号,其中基序Tyr Met Asp Asp(YMDD)中的残基Tyr是残基编号538。将氨基酸残基的编号依照其它聚合酶(例如843个或845个氨基酸长的)调整是本领域技术人员能够做到的。

[0034] 如本文所用,术语“天然”或“自然发生”在用于氨基酸序列(例如肽、多肽、蛋白等)或核苷酸序列(例如基因、核酸分子、多核苷酸等)时是指可发现自、分离自、得自自然界来源的氨基酸序列或核苷酸序列,其区别于实验室中人工修饰或由人所突变(即突变体)的那些。此种自然界来源包括从感染了或曾暴露于HBV的生物体收集的生物学样品(例如血液、血浆、血清、精液、唾液、组织切片、组织活检样本等)、培养的细胞(如HepG2.2.15、HuH6-C15(Sureau等,1986,Cell147:37;Sells等,1987,Proc.Natl.Acad.Sci.84(4):1005);HuH7.TA61或HuH7.TA62(Sun等,2006,J Hepatol.45(5):636),组织培养物以及重组材料。重组材料包括但不限于HBV分离物(例如可得自保藏机构的)、HBV基因组、研究组RNA或cDNA文库、含HBV基因组的载体或其片段或者已知包括此元素的任何现有技术载体)。

[0035] 作为说明目的,“天然HBV聚合酶”是指由任何现有技术中所述的自然发生的HBV基因型、毒株或分离物的ORF P所编码的HBV聚合酶(例如根据基因型832-845个氨基酸的多肽)或其片段。术语“天然”还涵盖代表具体基因型的HBV聚合酶多肽/肽,并因而包含对应于共有或接近共有序列的氨基酸序列(通常在各种特定基因型的HBV聚合酶的序列比对之后确定)。

[0036] 术语“突变体”如本文所用是指相对于天然对应物展现出一个或多个突变的多肽。作为说明目的,“突变体聚合酶多肽”是指如本文所述人工突变或实验室中由人改变之后源自天然聚合酶的聚合酶多肽。可涵盖任意突变,包括一个或多个核苷酸/氨基酸残基的取代、插入和/或缺失,非天然排列(例如与外援多肽/肽融合)以及这些可能性的任意组合。当考虑若干突变时,其可涉及连续的残基和/或非连续残基。可通过本领域技术人员已知的许多方法来产生突变,如定点诱变(例如使用Amersham的Sculptor™体外诱变系统,Les Ullis,France)、PCR诱变、DNA改组和通过化学合成技术(例如导致合成的核酸分子)。根据

优选的实施方案,本发明虑及的突变涵盖一个或多个氨基酸残基(连续或非连续)的缺失和/或取代,所述一个或多个氨基酸残基直接或间接参与天然HBV聚合酶所展现的至少一种酶活性,以旨在破坏所述至少一种酶活性如聚合酶活性和/或RNaseH活性。在本发明的上下文中,所得的突变体聚合酶多肽整体上保持与非突变部分中相应天然HBV聚合酶的高度相同性(例如至少80%)。

[0037] 如本文所用,术语“破坏”或其衍生词在与给定酶活性相关联时,是指“消除”(完全无残留活性)或者“显著降低”(残余活性低于天然聚合酶所展现活性的20%)。

[0038] 术语“相同性”是指在两个多肽或核苷酸序列之间确切的氨基酸对氨基酸或核苷酸对核苷酸的对应性。两个序列之间相同性的百分比是序列共有的相同位置数目的函数,其中将最佳比对所需要引入的缺口的数目以及每个缺口的长度纳入考量。现有技术中有各种计算机程序和数学算法用于确定氨基酸序列间相同性的百分比,例如NCBI的Blast程序或Atlas of Protein Sequence and Structure中的ALIGN(Dayhoffed.,1981,Suppl.,3482-489)。在专美的数据库中也有用于确定核苷酸序列间同源性的程序(例如Genbank、Wisconsin Sequence Analysis Package,BESTFIT,FASTA和GAP程序)。作为说明目的,“至少80%序列相同性”如本文所用是指80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%。

[0039] 如本文所用,术语“分离的”是指蛋白、多肽、肽、多核苷酸、质粒载体、病毒载体、或宿主细胞从其天然环境移开(即,与其天然相关联的至少一种其它成分分开)。

[0040] HBV序列

[0041] 许多HBV序列适宜用于本文所述的实施方案,包括本领域研究人员轻易可得的此类序列,包括但不限于,在Genbank和PubMed中所述的HBV序列。作为说明目的,深入的系统进化分析已经将乙肝病毒分类为8种主要的基因型(A-H),它们显示出有区别的地理分布和临床结果,但是却展现出高度的序列保守性。各种HBV也可基于HBsAg-相关的血清学而分类为9种不同的亚型(ayw1、ayw2、ayw3、ayw4、ayr、adw2、adw4、adrq+和adrq-)(参见Mamum-Al Mahtab等的综述,2008,Hepatobiliary Pancreas Dis Int5:457;Schaeffer,2007,World Gastroenterol.7:14)。每种基因型和血清型涵盖不同的HBV毒株和分离物。分离物对应于分离自特定来源HBV的具体病毒(例如,患者样品或其它生物学HBV库存),而毒株涵盖各种在基因组序列方面彼此很接近的分离物。

[0042] 基因型A的示例性HBV包括但不限于分离物HB-JI444A和毒株HB-JI444A(登录号AP007263)。基因型B的示例性HBV包括但不限于克隆pJDW233(登录号D00329),分离物HBV/14611(登录号AF121243),HBV-B1(GenBank登录号AF282917.1),HBV毒株Whutj-37(GenBank登录号AY2933309.1),中国HBV毒株GDH1(GenBank登录号AY766463.1)和HBV分离物57-1亚型adw(GenBank登录号AY518556.1)。基因型C的示例性HBV包括但不限于分离物AH-1-ON980424(登录号AB113879),毒株HCC-3-TT(登录号AB113877),HBV分离物SWT3.3(GenBank登录号EU916241.1),HBV分离物H85(GenBank登录号AY306136.1),HBV毒株 C1248(GenBank登录号DQ975272.1),HBV分离物CHN-H155(GenBank登录号DQ478901.1)和HBV分离物GZ28-1(GenBank登录号EF688062)。基因型D的示例性HBV包括但不限于分离物KAMCHATKA27(登录号AB188243),ALTAY136(登录号AB188245)和Y07587(Genbank登录号Y07587及Stoll-Becker等,1997,J.Virol.71:5399)以及登录号AB267090所述的HBV分离物。基因型E的示例

性HBV包括但不限于分离物HB-JI411F和毒株HB-JI411(登录号AP007262)。基因型F的示例性HBV包括但不限于分离物HBV-BL597(登录号AB214516)和HBV-BL592(登录号AB166850)。基因型G的示例性HBV包括但不限于分离物HB-JI444GF和毒株HB-JI444G(登录号AP007264)。基因型H的示例性HBV包括但不限于分离物HBV-ST0404(登录号AB298362)和分离物HB-JI260F和毒株HB-JI260(登录号AP007261)。

[0043] 本发明的并不是要局限于这些示例性HBV序列。实际上,根据本发明所使用的任何或所有HBV多肽/肽的核苷酸和氨基酸序列可以在不同HBV分离物和基因型之间变换,并且此天然遗传变异包括在本发明的范围之内。另外,本发明中所使用的HBV多肽/肽可代表具体的基因型,并因而包含对应于共有或近共有序列的氨基酸序列。

[0044] 另外,每种HBV多肽/肽可独立源自任何目前已鉴别的HBV基因型、毒株或分离物,如上文所述的与术语“HBV”相关联的那些。此种设定可以允许提供针对更广泛的HBV基因型的保护,或者提供对具体地理区域(通过使用在此区域中风行的HBV基因型)或具体患者群体的适应。对于此点,基因型A和C是美国最流行的,基因型A和D是西欧国家最流行的,以及基因型D是地中海区域最流行的,而基因型B和C在中国最常见。来自印度的有限数据表明,基因型A和D在印度最流行。本领域技术人员能够依据要治疗的群体和/或地理区域来选择适宜的HBV基因型、血清型、毒株和/或分离物。

[0045] 根据优选的实施方案,本发明中所使用的HBV多肽/肽源自基因型D病毒,特别优选HBV分离物Y07587。

[0046] 突变体HBV聚合酶

[0047] 本发明的突变体聚合酶多肽包含具有内部缺失的突变的聚合酶结构域,所述内部缺失功能性地破坏聚合酶活性并且包含至少天然存在于天然聚合酶的聚合酶结构域中的YMDD基序。由所得的突变体聚合酶多肽所展现的对聚合酶活性的破坏,可以使用本领域已知的测定来进行评估(例如Radziwiłł等,1990,J Virol.64:613中所述的内源聚合酶测定)。

[0048] SEQ ID NO:1中提供了涵盖基因型B、C和D的天然HBV聚合酶的聚合酶结构域的通式氨基酸序列,其中位置7的Xaa是Thr(T)或Ala(A);位置13的Xaa是Asn(N),Arg(R)或His(H);位置16的Xaa是Ile(I)或Thr(T);位置38的Xaa是Thr(T)或Ala(A);位置53的Xaa是Ser(S)或Asn(N);位置54的Xaa是Thr(T)或Tyr(Y);位置55的Xaa是His(H)或Arg(R);位置91的Xaa是Ile(I)或Leu(L);位置109的Xaa是Pro(P)或Ser(S);位置118的Xaa是Thr(T)或Asn(N);位置121的Xaa是Asn(N)或Ile(I);位置122的Xaa是Ile(I)或Phe(F);位置124的Xaa是Tyr(Y)或Asn(N);位置127的Xaa是Gly(G)或Arg(R);位置131的Xaa是Asp(D)或Asn(N);位置134的Xaa是Asp(D)或Asn(N);位置145的Xaa是Leu(L)或Met(M);位置149的Xaa是Lys(K)或Gln(Q);位置151的Xaa是Phe(F)或Tyr(Y);位置221的Xaa是Phe(F)或Tyr(Y);位置222的Xaa是Thr(T)或Ala(A);位置223的Xaa是Ser(S)或Ala(A);位置224的Xaa是Ile(I)或Val(V);位置238的Xaa是Asn(N)或His(H);位置248的Xaa是Asn(N)或His(H);位置256的Xaa是Ser(S)或Cys(C);位置257的Xaa是Trp(W)或Tyr(Y);位置259的Xaa是Thr(T)或Ser(S);位置263的Xaa是Glu(E)或Asp(D);位置266的Xaa是Val(V)或Ile(I);位置267的Xaa是Leu(L)或Gln(Q);位置271的Xaa是Gln(Q),Met(M)或Glu(E);位置317的Xaa是Ser(S)或Ala(A);和位置332的Xaa是Cys(T)或Ser(S)。

[0049] 根据本发明,本发明的突变体聚合酶多肽中所包含的突变的聚合酶结构域缺少至少YMDD基序(存在于SEQ ID NO:1的此通式聚合酶结构域的位置203至位置206)。

[0050] 本发明还涵盖最少4个氨基酸残基和最多30个氨基酸残基的任意其它内部缺失(其包含至少此YMDD基序)。

[0051] 代表性的本发明的突变体聚合酶多肽包含突变的聚合酶结构域,所述突变的聚合酶结构域包含SEQ ID NO:1所示的氨基酸序列但缺少至少位置203处的Tyr残基,位置204处的Met残基、位置205处的Asp残基和位置 206处的Asp残基。

[0052] 在YMDD基序之外,优选所述内部缺失还包含所有的或部分的相邻VVL基序,所述VVL基序在天然HBV聚合酶结构域中存在所述YMDD基序的C端(对应于SEQ ID NO:1的位置207-209处的残基和对应于832个氨基酸的天然聚合酶的位置542-544的残基)。此种VVL基序实际上能够有助于形成“连接(junctional)”表位(例如共线合成的新表位),其有降低宿主的免疫应答(一个或多个HBV聚合酶-相关的表位)或使该免疫应答沉默的风险。

[0053] 优选地,本发明的突变体聚合酶多肽包含突变的聚合酶结构域,所述突变的聚合酶结构域具有SEQ ID NO:1所示的氨基酸序列但缺少至少位置203处的Tyr残基、位置204处的Met残基、位置205处的Asp残基、位置206处的Asp残基、位置207处的Val残基、位置208处的Val残基和位置209处的Leu残基。

[0054] 更优选地,本发明的突变体聚合酶多肽包含聚合酶结构域,所述聚合酶结构域包含这样的氨基酸序列、或者基本由该氨基酸序列组成、或者由该氨基酸序列组成,所述氨基酸序列与SEQ ID NO:2所示的氨基酸序列展现出至少80%的相同性,有利的是至少85%的相同性,优选至少90%的相同性,更优选至少95%的相同性,以及更优选至少100%相同性。更优选地,所述突变的聚合酶结构域包含SEQ ID NO:2所示的氨基酸序列。

[0055] 作为选择或者作为组合,本发明的突变体聚合酶多肽还包含突变的RNaseH结构域,所述突变的RNaseH结构域包含一个或多个氨基酸残基的突变,所述突变功能性地破坏天然HBV聚合酶中正常展现的RNaseH活性。

[0056] 如上文所述,RNase H活性中所涉及的功能性结构域已经在C-端部分内作图(map),更特别是天然的832个氨基酸长的HBV聚合酶的位置680至C-端(或者从天然的845个氨基酸长的HBV聚合酶的位置693至C-端),并且本发明涵盖了此结构域中与RNase H活性的破坏相关的任意突变(即最终导致低于天然RNaseH活性的20%的弱残留活性)。可以使用本领域已知的测定来评估由突变体聚合酶多肽所展现的RNase H活性的破坏(例如体外RNaseH活性测定或DNA-RNA串联分子分析,Radziwill等,1990,J Virol.64:613或Lee等,1997,Biochem.Bioph.Res.Commun.233(2):401中所述)。

[0057] SEQ ID NO:3中提供了涵盖基因型B、C和D的天然HBV聚合酶的 RNaseH结构域的通式氨基酸序列,其中位置2的Xaa是Ser(S)或Pro(P);位置19的Xaa是Ala(A)或Val(V);位置20的Xaa是Ile(I)或Met(M);位置30的Xaa是Val(V)或Leu(L);位置31的Xaa是Ala(A)或Ser(S);位置53的Xaa是Lys(K)或Asn(N);位置54的Xaa是Leu(L)或Ile(I);位置55的Xaa是Leu(L)或Ile(I);位置97的Xaa是Ala(A)或Thr(T);位置108的Xaa是Tyr(Y)或Ser(S);位置115的Xaa是Pro(P)或Leu(L);位置116的Xaa是Phe(F)或Tyr(Y);位置128的Xaa是Val(V)或Asp(D)。

[0058] 有利的是,本发明的突变体聚合酶多肽的RNaseH结构域中所包含的一个或多个突

变选自：

[0059] • 最少8个氨基酸和最多60个氨基酸的缺失,包括至少SEQ ID NO:3从大约位置39处的Glu(E)残基延伸至大约位置46处的Ala(A)残基的部分(de1 ELLAACFA);

[0060] • 用D之外的氨基酸残基取代SEQ ID NO:3的位置10处的Asp(D)残基;

[0061] • 用V之外的氨基酸残基取代SEQ ID NO:3的位置90处的Val(V)残基;

[0062] • 用T或A之外的氨基酸残基取代SEQ ID NO:3的位置97处的Thr(T)或Ala(A)残基;

[0063] • 用D之外的氨基酸残基取代SEQ ID NO:3的位置98处的Asp(D)残基;和

[0064] • 其任意组合。

[0065] 适宜组合的代表性例子包括但不限于(a)位置10、90、97和98处的氨基酸残基的取代;(b)包括GLLAACFA基序的8-60个氨基酸残基的缺失和任何所提到的位置处的氨基酸残基的取代;或(c)所有所列突变的组合。

[0066] 适宜的是,SEQ ID NO:3的位置10、90、97或98处的被取代的残基单独被置换为His(H)残基或Tyr(Y)残基,特别是优选His(H)残基取代SEQ ID NO:3的位置10处的残基(D689H),用Tyr(Y)残基取代SEQ ID NO:3的位置90处的残基(V769Y),用Tyr(Y)残基取代SEQ ID NO:3的位置97处的残基(T776Y或A776Y)和/或用His(H)残基取代SEQ ID NO:3的位置98处的残基(D777H)。

[0067] 适宜的是,在突变的RNase H结构域中所包含的缺失,包括SEQ ID NO:3中从大约位置39处的Glu(E)残基延伸至大约位置57处的Thr(T)残基的至少19个氨基酸残基的部分、优选SEQ ID NO:3中从大约位置39处的Glu(E)残基延伸至大约位置63处的Leu(L)残基的至少25个氨基酸残基的部分、更优选SEQ ID NO:3中从大约位置31处的残基Xaa(A或S)延伸至大约位置63处的Leu(L)残基的至少33个氨基酸残基的部分。

[0068] 优选地,本发明的突变体聚合酶多肽包含突变的RNaseH结构域,所述突变的RNaseH结构域包含SEQ ID NO:3中所示的氨基酸序列,但却(a)缺少从位置31处的残基Xaa(X)延伸至大约位置63处的Leu(L)残基的33个氨基酸残基的部分并且包含(b)用His(H)残基取代位置10处的Asp(D)残基残基(D689H);(c)用Tyr(Y)残基取代位置90处的Val(V)残基(V769Y);(d)用Tyr(Y)残基取代位置97处的残基(T/A776Y)和(e)用His(H)残基取代位置98处的Asp(D)残基(D777H)。

[0069] 更优选地,本发明的突变体聚合酶多肽包含突变的RNaseH结构域,所述突变的RNaseH结构域包含这样的氨基酸序列、或者基本由该氨基酸序列组成、或者由该氨基酸序列组成,所述氨基酸序列与SEQ ID NO:4所示的氨基酸序列展现出至少80%的相同性,有利的是至少85%的相同性,优选至少90%的相同性,更优选至少95%的相同性,以及更优选100%相同性。

[0070] 在更优选的实施方案中,本发明的突变体聚合酶多肽包含这样的氨基酸序列、或者基本由该氨基酸序列组成、或者由该氨基酸序列组成,所述氨基酸序列与SEQ ID NO:5所示的氨基酸序列展现出至少80%的相同性,有利的是至少85%的相同性,优选至少90%的相同性,更优选至少95%的相同性,以及更优选至少100%相同性。更优选的是包含SEQ ID NO:5所示氨基酸序列的突变体聚合酶多肽。

[0071] 在本发明的上下文中,本发明的突变体聚合酶多肽可包含另外的突变。然而,优选

避免可对免疫原活性有害的修饰,尤其是在富含B、CTL和/或T_H表位的部分。

[0072] 示例性的另外的修饰包括N-端截短。特别适宜的是正常存在于天然HBV聚合酶或SEQ ID NO:5的N-端的最少20个氨基酸残基和最多100个氨基酸残基的截短,特别优选从SEQ ID NO:5的位置1(Met起始区)或2延伸至大约位置47的截短。此修饰特别适宜于与天然HBV核心多肽组合使用的突变体聚合酶多肽,这是由于此种N-端截短有助于减少或者删除这两种多肽之间重叠的部分。然而,这也可以通过使用与C-端截短的HBV核心多肽组合的非截短的突变体聚合酶多肽来实现,如下文所述。

[0073] 期望的是,所得的突变体聚合酶多肽保留免疫原性质特别是刺激细胞介导的免疫应答的能力(在相同的范围内或者比天然聚合酶更高)。

[0074] 与其它多肽组合/融合

[0075] 在另一实施方案中,本发明的突变体聚合酶多肽可以与一个或多个另外的多肽或多肽组合使用。

[0076] 术语“组合”及变体如“组合使用”是指在相同宿主生物体中施用两个或更多个实体的行为,所述实体之一是本发明的物体。通常,该至少两种实体可以通过不同的途径以及根据不同的时间方案来施用。适宜的组合包括但不限于the combination of the本文所述的突变体聚合酶多肽(或编码其的载体)与抗病毒剂(SOC)的组合和/或与另外的多肽或编码此类另外的多肽的载体的组合。此种组合可以是(a)混合物的形式(例如两种或更多种多肽或载体的混合物),(b)该两种或更多种实体的融合物的形式或者(c)通过具体的表达设计(例如双顺反子或独立飘到)。例如,该两个或更多个实体可以使用不同的调节元件在相同载体内独立表达(例如用不同的启动子和终止序列)。独立表达设计特别适宜于for expression from质粒或麻疹载体。或者,所述两个或更多个实体可以双顺反子的方式表达,其在相同启动子和终止序列的控制下但需要使用另外的调节元件如IRES(代表内部核糖体进入位点)以允许两个或更多个顺反子从相同的mRNA翻译。本领域中有大量的IRES可选择,如源自脊髓灰质炎病毒、丙肝病毒和脑心肌炎病毒(EMCV)的那些(例如,参见W095/24485)。双顺反子设计特别适宜于从具有更有限的克隆能力的载体表达,如腺病毒载体。

[0077] 术语“融合”或“融合蛋白”如本文所用是指在单一多肽链中的两个或更多个多肽/肽的组合。优选地,通过遗传手段来进行融合,即通过符合读框地融合编码每个所述多肽/肽的核苷酸序列。对于“符合读框地融合”,其是指所融合的编码序列的表达导致单一的蛋白,在每个所融合的多肽/肽之间没有任何翻译终止子。融合可以是直接的(即之间没有任何另外的氨基酸残基)或通过接头(例如3-30个氨基酸长的肽,由重复的氨基酸残基组成如甘氨酸、丝氨酸、苏氨酸、天冬酰胺、丙氨酸和/或脯氨酸)。

[0078] 用于与本发明的突变体聚合酶多肽组合使用的另外的多肽或多肽优选是HBV基因组所编码的多肽或多肽,如任何天然HBV多肽、其修饰的衍生物和/或片段。此类HBV多肽或多肽的代表性例子包括但不限于HBc(核心)、HBs、X蛋白以及其任意免疫原性片段。

[0079] 根据本发明,如上文所述,本发明中所使用的任何另外的HBV多肽或多肽可源自与本发明的突变体聚合酶多肽所源自的HBV基因型、毒株或分离物不同或相同的HBV基因型、毒株或分离物。优选地,此类另外的HBV多肽源自基因型D HBV,且特别是源自Y07587分离物。

[0080] 优选的组合是本文所述突变体聚合酶多肽与另外的多肽之间融合蛋白的形式。此种融合优选是直接的,在融合的实体之间没有任何接头。

[0081] 与核心组合

[0082] 如本文所用,术语“核心多肽”是指这样的多肽,其保留天然HBV核心(HBc)蛋白中所包含的至少100个氨基酸残基。此术语涵盖了从自然界中HBV来源所发现、分离、获得的任何HBV毒株、分离物或基因型(如上文与术语“HBV”相关联的那些)的天然(即自然发生的)核心多肽,以及修饰的核心及其片段。

[0083] 本发明中所使用的HBV核心多肽可源自这样的HBV病毒,该HBV病毒与所述突变体聚合酶多肽所源自的HBV病毒具有相同的或不同的基因型。优选地,它们都源自源自基因型D病毒,更特别是源自Y07587分离物。核心多肽及其编码序列可以通过本领域技术人员已知的方式来产生,如通过编码序列的化学合成(例如导致合成的核酸分子)或者通过重组手段(例如相应核苷酸序列的定点诱变,PCR诱变,DNA改组)。

[0084] 可以涉及任何修饰,只要所得的核心在与本文所述的突变体聚合酶多肽组合或融合时保留显著的免疫原活性(优选在同样的范围或者比天然核心对应物更高)。

[0085] 适宜的修饰包括正常存在于天然核心多肽的C-端或在其C-端部分内的最少10个氨基酸残基和最多41个氨基酸残基的截短,特别优选从天然核心多肽的残基143、144、145、146、147、148或残基149延伸至C-端(残基183)的截短。其它适宜的修饰包括一个或多个氨基酸残基的内部缺失,特别是在表明暴露区域中,如位于残基80附近的区域,据推测其形成外部环(Argos等1988,EMBO J.7:819;Borisova等,1993,J.Virol.67:3696;Schodel等,1992,J.Virol.66:106;Yon等,1992,J.Gen Virol.73:2569;和Pumpens等,1995,Intervirology38:63)。

[0086] 在优选的实施方案中,所述组合是融合的形式。因此,本发明涉及融合蛋白,其包含本文所述的突变体聚合酶多肽和融合配偶体。优选地,所述融合配偶体是HBV核心多肽,特别优选C-端截短的核心多肽,尤其是在残基148处截短的。

[0087] 优选地,所述HBV核心多肽被符合读框地融合至本文所述的突变体聚合酶多肽的N-端,导致融合蛋白(始自起始区Met、无任何终止密码子的核心多肽(修饰的或天然的)、突变体聚合酶多肽(没有任何Met起始区)和终止密码子)。

[0088] 优选的融合蛋白包含这样的氨基酸序列、或者基本由该氨基酸序列组成、或者由该氨基酸序列组成,所述氨基酸序列与SEQ ID NO:6所示的氨基酸序列展现出至少80%的相同性,有利的是至少85%的相同性,优选至少90%的相同性,更优选至少95%的相同性,以及更优选100%相同性。更优选地,所述融合蛋白包含SEQ ID NO:6所示的氨基酸序列。

[0089] 与免疫原性HbsAg结构域组合

[0090] 作为选择或者与前文所述实施方案(与核心多肽组合)组合,本发明的突变体聚合酶多肽可以与HbsAg或其免疫原性片段/结构域组合使用。

[0091] 如本文所用,术语“免疫原性结构域”是指具有大约15至大约100个氨基酸残基、优选最少20个和最多60个连续的氨基酸的多肽,其包含至少一个正常存在于天然HBsAg蛋白中的特异于T辅助(TH)细胞和/或细胞毒T(CTL)细胞的B和/或T细胞表位。另外此类表位可以被局限为各种MHC I类和/或II类抗原(例如A2,A24,DR,DP等)。优选地,本发明中所使用的一个或多个HBsAg免疫原性结构域不包括HBV preS1和preS2多肽的任何部分。

[0092] 每个所述一个或多个免疫原性结构域可独立地源自相同的或不同的HBV病毒,该HBV病毒可以与本文所述的突变体聚合酶多肽(及最终的核心多肽)所源自的HBV病毒相同

或不同。优选地,所述一个或多个免疫原性结构域源自基因型D HBV,特别是源自Y07587分离物。

[0093] 可用于本发明的示例性免疫原性结构域在现有技术中有描述(例如W093/03764; W094/19011; Desombere等,2000, Clin. Exp. Immunol. 122:390; Loirat等,2000, J. Immunol. 165:4748; Schirmbeck等,2002, J. Immunol. 168:6253; Depla等,2008, J. Virol. 82:435和W02011/015656)。特别优选的免疫原性结构域包括W02011/015656中所述的env1和env2结构域。“Env1”对应于天然HBsAg从大约位置14至大约位置51的部分,而env2对应于HBsAg从大约位置165至大约位置194的部分。

[0094] 在优选的实施方案中,组合是以融合的形式,并且本发明涉及包含本文所述的突变体聚合酶多肽和一个或多个HBsAg免疫原性结构域的融合蛋白或者涉及还包含一个或多个HBsAg免疫原性结构域的上文所述的融合蛋白(包含至少本文所述的突变体聚合酶多肽和Hbc多肽)。所述一个或多个免疫原性结构域可以位于融合蛋白的N-端、C-端和/或内部,例如在突变体聚合酶多肽内(例如位于突变的聚合酶和/或RNaseH结构域中所缺少的部分处),或者位于核心和突变体聚合酶多肽之间。本领域技术人员能够根据需要确定翻译介导的调节元件的位置(例如起始区Met和密码子STOP位于该融合蛋白的N-和C-端)。

[0095] 特别感兴趣的融合蛋白包含本文所述的突变体聚合酶多肽、核心多肽和两个HBsAg免疫原性结构域,特别优选的融合是在其N-包含核心多肽(例如天然的183个残基或截短的148个残基,具有起始区Met)融合至突变体聚合酶多肽(无起始区Met),所述突变体聚合酶多肽包含融合于突变的聚合酶结构域中的内部缺失处(例如env1)和/或融合于突变的RNaseH结构域的缺失处(例如env2)的1个或2个HbsAg免疫原性结构域。

[0096] 在此实施方案的优选方面,本发明的融合蛋白包含这样的氨基酸序列、或者基本由该氨基酸序列组成、或者由该氨基酸序列组成,所述氨基酸序列与SEQ ID NO:7-9所示的任意氨基酸序列展现出至少80%的相同性,有利的是至少85%的相同性,优选至少90%的相同性,更优选至少95%的相同性,以及更优选100%相同性。更优选的实施方案是,包含SEQ ID NO:8所示的氨基酸序列的融合蛋白。

[0097] 在本发明的上下文中,本发明的突变体聚合酶多肽或本发明的融合蛋白可以还包含另外的结构特征。

[0098] 在一个实施方案中,其可包含旨在提高其于宿主生物体中免疫原活性另外的化合物(例如肽或多肽)。此类能够增强免疫原性的化合物在文献中已有描述,包括但不限于,钙网织蛋白(calreticulin)(Cheng等,2001, J. Clin. Invest. 108:669),结核分枝杆菌热休克蛋白70(HSP70)(Chen等,2000, Cancer Res. 60:1035),泛素(Rodriguez等,1997, J. Virol. 71:8497),细菌毒素如绿脓假单胞菌外毒素A的移位结构域(ETA(dIII))(Hung等,2001 Cancer Res. 61:3698)以及T辅助表位如Pan-Dr肽(Sidney等,1994, Immunity 1:751), pstS1GCG表位(Vordermeier等,1992, Eur. J. Immunol. 22:2631),破伤风类毒素肽P2TT(Panina-Bordignon等,1989, Eur. J. Immunol. 19:2237)和P30TT(Demotz等,1993, Eur. J. Immunol. 23:425),流感表位(Lamb等,1982, Nature 300:66)和血球凝集素表位(Rothbard等,1989, Int. Immunol. 1:479)。

[0099] 其它适宜的结构特征是有益于本发明的突变体聚合酶多肽或融合蛋白的合成、加工、稳定性和可溶性的那些(例如旨在修饰潜在切割位点、潜在糖基化位点和/或膜锚点以

改善呈递至细胞膜)。

[0100] 通过使用适宜的本领域熟知序列(如信号和/或跨膜肽)而将本文所述的突变体聚合酶多肽或融合蛋白的合成引至细胞表面对于免疫应答是有益的。简言之,信号肽通常存在于膜呈递或分泌多肽的N-端并且将其引至内质网(ER)。它们通常包含15-35个基本上疏水的氨基酸,其继而由特异性的位于ER的内肽酶去除以产生成熟多肽。跨膜肽在自然界中也是高度疏水的并且勇于将多肽锚定于细胞膜。可用于本发明背景中的跨膜和/或信号肽的选择很多。其可得自任何膜锚定的和/或分泌的多肽(例如细胞或病毒多肽),如免疫球蛋白、组织纤维蛋白溶酶原激活剂、胰岛素、狂犬病毒糖蛋白、HIV病毒包膜糖蛋白或麻疹病毒F蛋白或者可以是合成的。

[0101] 在一个实施方案中,本发明的突变体聚合酶多肽或融合蛋白符合读框地融合至信号肽(其插入至N-端翻译起始密码子的下游)。在另一实施方案中,本发明的突变体聚合酶多肽或融合蛋白符合读框地融合至信号肽(例如插入在其N-端)和融合至跨膜肽(例如插入在C-端,例如紧接在终止密码子的上游)。优选地,本发明中所使用的信号和跨膜肽源自狂犬病毒糖蛋白(参见例如W099/03885或W02008/138649)。优选的实施方案是这样的HBV聚合酶突变体多肽和融合蛋白,其包含这样的氨基酸序列、或者基本由该氨基酸序列组成、或者由该氨基酸序列组成,所述氨基酸序列与SEQ ID NO:10,SEQ ID NO:11或SEQ ID NO:12所示的氨基酸序列展现出至少80%的相同性,有利的是至少85%的相同性,优选至少90%的相同性,更优选至少95%的相同性,以及更优选100%相同性。

[0102] 核酸分子

[0103] 在另一方面,本发明提供编码本文所述的突变体聚合酶多肽和融合蛋白的分离的核酸分子。

[0104] 在本发明的背景中,术语“核酸”,“核酸分子”,“多核苷酸”和“核苷酸序列”可互换使用并且定义了任何长度的多聚脱氧核糖核苷酸(DNA)(例如,cDNA,基因组DNA,质粒,载体,病毒基因组,分离的DNA,探针,引物及其任意混合物)或多聚核糖核苷酸(RNA)(例如,mRNA,反义RNA)或混合的多聚核糖-多聚脱氧核糖核苷酸。其涵盖了单链或双链、线性或环状、天然的或合成的多核苷酸。另外,多核苷酸可包含非自然发生的核苷酸并且可以被非核苷酸成分所中断。

[0105] 本发明的核酸分子可使用本领域可得的序列数据和本文所提供的序列信息而产生自任何来源。例如,编码HBV聚合酶(及有需要时的核心多肽和HBsAg免疫原性结构域)的DNA序列可独立分离自含HBV的细胞、cDNA和基因组文库、病毒基因组或任何现有技术已知含有其的载体,并继而通过常规分子生物学或PCR技术而适宜地连接。或者,本发明的核酸分子还可产生自自动化过程中的化学合成(例如从重叠合成寡核苷酸合成基因组)。可通过本领域技术人员已知的许多方式来产生修饰,如化学合成、定点诱变、PCR诱变、DNA改组等。

[0106] 特别感兴趣的核酸分子选自:

[0107] • 编码突变体聚合酶多肽的核酸分子,所述突变体聚合酶多肽包含具有SEQ ID NO:1或2所示氨基酸序列的聚合酶结构域

[0108] • 编码突变体聚合酶多肽的核酸分子,所述突变体聚合酶多肽包含具有SEQ ID NO:3或4所示氨基酸序列的RNaseH结构域;

[0109] • 编码突变体聚合酶多肽的核酸分子,所述突变体聚合酶多肽包含与SEQ ID NO: 5所示氨基酸序列展现出至少80%的相同性(例如80%,85%,90%,95%,97%,100%)的氨基酸序列;或

[0110] • 编码融合蛋白的核酸分子,所述融合蛋白包含与SEQ ID NO:6-12所示的任意氨基酸序列展现出至少80%的相同性(例如80%,85%,90%,95%,97%,100%)的氨基酸序列。

[0111] 本发明并不局限于这些示例性的核苷酸序列,并包含任意旨在改善本发明中所使用的核酸分子的克隆、表达、稳定性的修饰(例如引入适宜的限制性酶切位点、将核苷酸序列简并化和/或优化以使宿主细胞中的翻译优化和/或抑制潜在的使核酸分子或其转录物不稳定的负面元素)。在虑及多个修饰时,它们可涉及连续残基和/或非连续残基。本发明所涉及的修饰涵盖了不改变所编码的多肽和融合蛋白的氨基酸序列的沉默修饰,以及翻译到所编码的多肽和融合蛋白中的修饰。

[0112] 在一个实施方案中,本发明的核酸分子可以在其全长核苷酸序列或其部分进行简并,从而降低本发明或宿主细胞中所使用的核酸分子之间的序列同源性。实际上可以对显示出高度的核苷酸序列相同性的核酸序列部分进行简并,并且本领域技术人员能够通过序列比对鉴别出这样的部分。例如如果载体携带编码本文所述突变体聚合酶多肽的核酸分子和编码另外的HBV多肽(由HBV基因组中重叠序列编码)的核酸分子,有利的是将所述核酸分子其中之一或者二者一起在重叠部分进行简并,从而避免在生产过程中的同源重组问题。

[0113] 作为选择或作为组合,可将本发明的核酸分子优化以提供在特定宿主细胞或生物体中的高水平表达。实际上已观察到,当多于一个密码子可用于编码给定氨基酸时,生物体的密码子使用模式是高度非随机的并且密码子使用在不同宿主间是有显著差异的。由于本发明所涵盖的核苷酸序列大部分是病毒来源(HBV),它们可能会具有不适宜于在宿主细胞(如细菌、低等或高等真核细胞)中有效表达的密码子使用模式。通常,可以通过如下方式进行密码子优化:将一个或多个“天然”(例如HBV)密码子(在感兴趣宿主细胞/生物体中不常使用的密码子)替换为一个或多个在感兴趣宿主细胞/生物体中更常使用的密码子。并不是必需要将所有的天然密码子(不常使用的密码子)替换,因为即便是部分替换也可实现提高的表达。另外,从严格遵循至优化的密码子使用的变化也可用于在所得的核酸分子中引入限制性酶切位点。

[0114] 另外,可以通过核苷酸序列的另外的修饰来改善在宿主细胞或生物体中的表达,所述修饰旨在防止罕见、非优化密码子的簇集和/或防止或修饰至少部分负面序列元件(其预期会对表达水平产生负面影响,例如富含AT或富含GC的序列区段;不稳定的直接的或插入的重复序列;RNA二级结构;和/或内部隐含调节元件如内部TATA-盒、chi-位点、核糖体进入位点、和/或剪接供体/受体位点)。

[0115] 本发明特别优选的实施方案是这样的核酸分子,其包含这样的核苷酸序列、或者基本由该核苷酸序列组成、或者由核苷酸序列组成,所述核苷酸序列与SEQ ID NO:13-17所示的任意氨基酸序列展现出至少80%的相同性,有利的是至少85%的相同性,优选至少90%的相同性,更优选至少95%的相同性,以及更优选100%相同性。

[0116] 本发明的另一实施方案涉及本发明的核酸分子的片段,例如限制性内切酶和PCR-生成的片段。此类片段可用作探针、引物或编码所编码免疫原性多肽的免疫原性部分的片段。

[0117] 载体

[0118] 在另一方面,本发明提供包含本发明核酸分子的载体。

[0119] 术语“载体”如本文所用是指载剂(vehicle),优选含有允许一个或多个核酸分子在宿主细胞或生物体内递送、传代和/或表达所需的元件的核酸分子或病毒颗粒。此术语涵盖了用于维持的载体(克隆载体)或用于在各种宿主细胞或生物体中表达的载体(表达载体),额外的染色体载体(例如多拷贝质粒)或整合载体(例如设计为用于整合至宿主细胞基因组并在宿主细胞复制时产生另外的拷贝的核酸分子)以及穿梭载体(例如在原核和/或真核宿主中都发挥作用)和转移载体(例如用于将核酸分子转移至病毒基因组内)。对于本发明的目的,所述载体可以是自然发生的遗传来源、合成的或人工的、或者是天然和人工元件的一些组合。

[0120] 在本发明的上下文中,术语“载体”需要广泛理解为包括质粒和病毒载体。“质粒载体”如本文所用是指可复制的DNA构建体。通常质粒载体含有可选择标记物基因,其允许在存在相应选择药物的情况下选择或排除携带有质粒载体的宿主细胞。各种阳性和阴性可选择标记物基因是本领域已知的。作为说明,抗生素抗性基因可用作阳性可选择标记物基因,其允许在存在相应选择抗生素的情况下选择宿主细胞。

[0121] 术语“病毒载体”如本文所用是指包括至少一个病毒基因组元件并且可以包装至病毒颗粒或包装到病毒颗粒内的核酸载体。术语“病毒”、“病毒粒体”、“病毒颗粒”和“病毒载体颗粒”可互换使用,是指核酸载体依据适宜的条件(允许感染性病毒颗粒的产生)转导至适宜的细胞或细胞系时所形成的病毒颗粒。在本发明的背景中,术语“病毒载体”必需宽泛地理解为包括诱导核酸载体(例如DNA病毒载体)以及其产生的病毒颗粒。术语“感染性”是指病毒载体感染并进入宿主细胞或生物体的能力。病毒载体可以是复制-活性的或-选择性的(例如被改造为在特定的宿主细胞中更好地或选择性的复制),或者可以是无遗传能力的从而是复制-缺陷的或复制-受损的。

[0122] 适宜于本发明的背景的载体包括但不限于,用于在原核宿主细胞如细菌中表达(例如大肠杆菌(*E.coli*)、枯草芽孢杆菌(*Bacillus subtilis*)或李斯特氏菌(*Listeria*))的噬菌体、质粒或粘粒载体;用于在酵母(例如酿酒酵母(*Saccharomyces cerevisiae*)、裂殖酵母(*Saccharomyces pombe*)、毕赤氏酵母(*Pichia pastoris*))中表达的载体;用于在昆虫细胞系统(例如Sf9细胞)中表达的杆状病毒载体;用于在植物细胞系统(例如Ti质粒,花椰菜花叶病毒CaMV;烟草花叶病毒TMV)中表达的病毒和质粒载体;以及用于在高等真核细胞或生物体中表达的病毒和质粒载体。

[0123] 通常,此类载体是商业可购买的(例如Invitrogen、Stratagene、Amersham Biosciences、Promega等)或者可得自保藏机构的如American Type Culture Collection(ATCC, Rockville, Md.)或者在许多出版物中有描述其序列、组织和生产方法从而使得技术人员能够应用。

[0124] 适宜的质粒载体的代表性例子包括但不限于,pREP4、pCEP4(Invitrogen)、pCI(Promega)、pVAX(Invitrogen)和pgWiz(Gene Therapy System Inc)。

[0125] 适宜的病毒载体的代表性例子可产生至各种不同的病毒(例如反转录病毒、腺病毒、腺病毒-相关病毒(AAV)、痘病毒、疱疹病毒、麻疹病毒、泡沫病毒、 α 病毒,水疱性口炎病毒等)。如上文所述,术语“病毒载体”涵盖载体DNA、基因组DNA以及其所产生的病毒颗粒。

[0126] 本发明还涵盖与脂质或多聚物复合以形成颗粒结构如脂质体、脂复合物或纳米颗粒的载体(例如质粒DNA)。

[0127] 在一个实施方案中,本发明的载体是腺病毒载体。其可衍生自各种人类或动物(例如犬、羊、猿等)腺病毒。可使用任何血清型。期望的是,所述腺病毒载体是复制缺陷的并源自人类Ad,特别是源自稀有血清型的人类Ad,或源自黑猩猩Ad。人类腺病毒的代表性例子包括亚属C Ad2Ad5和Ad6,亚属B Ad11、Ad34和Ad35以及亚属D Ad19、Ad24、Ad48和Ad49。Chimp Ad的代表性例子包括但不限于AdCh3(Peruzzi等,2009,Vaccine27:1293)、AdCh63(Dudareva等,2009,vaccine27:3501)以及任何现有技术中描述的那些(参见例如W003/000283;W003/046124;W02005/071093;W02009/073103;W02009/073104;W02009/105084;W02009/136977和W02010/086189)。

[0128] 复制缺陷腺病毒载体可以如现有技术中所述那样获得,例如通过缺失病毒复制所必需的腺病毒基因组或其部分的至少一个区域,特别优选缺失包含E1编码序列的E1区(例如参照人腺病毒5型的序列从大约位置459延伸至3510,人腺病毒5型在GeneBank以登录号M73260、以及在Chroboczek等,1992,Viol.186:280中有公开)。本发明还涵盖在腺病毒基因组内具有另外的缺失/修饰(非必需的E3区的全部或部分或者其它必需的E2、E4区的全部或部分,如W094/28152;Lusky等,1998,J.Viol.172:2022中所述)的载体。

[0129] 本发明的核酸分子可插入到腺病毒基因组的任何位置,并且可位于相对于所涉及区域自然转录方向而言的有义或反义朝向。优选地,本发明的核酸分子被插入至取代腺病毒E1区并置于CMV启动子的控制下。

[0130] 适宜于本发明背景的其他病毒载体有痘病毒载体,其可得自痘病毒属的任何成员,特别优选源自金丝雀痘、鸡痘或牛痘(痘苗vaccinia)病毒的痘病毒载体,优选后者。适宜的痘苗病毒包括但不限于Copenhagen毒株(Goebel等,1990,Viol.179:247;Johnson等,1993,Viol.196:381),Wyeth毒株以及特别是修饰的Ankara(MVA)毒株(Antoine等,1998,Viol.244:365)。构建重组痘病毒的一般条件是本领域熟知的。本发明的核酸分子优选被插入至痘病毒基因组内的非必需基因座中。胸苷激酶基因特别适宜用于插入到Copenhagen痘苗载体中,以及缺失II(deletion II)或缺失III(deletion III)适宜用于插入到MVA载体中。优选地,本发明的核酸分子被插入到MVA载体的缺失III并置于痘苗7.5K或pH5R启动子的控制下。

[0131] 适宜于本发明背景的病毒载体有麻疹病毒属,其可得自副粘液病毒科,特别优选麻疹病毒。现有技术中有各种减毒的毒株(Brandler et al,2008,CIMID,31:271;Singh等,1999,J.Viol.73(6):4823),例如但不限于,Edmonston A和B毒株(Griffin等,2001,Field's in Virology,1401-1441)、Schwartz毒株(Schwarz A,1962,Am J Dis Child,103:216)、S-191或C-47毒株(Zhang等,2009,J Med Virol.81(8):1477)。特别适宜在P和M基因之间插入。

[0132] 根据本发明,本发明的载体中所包含的核酸分子是适宜于在宿主细胞或生物体中表达的形式,这意味着该核酸分子被置于适宜的调节序列的控制下。如本文所用,术语“调节元件”是指允许、有助于或调节核酸分子在给定宿主细胞或生物体中的表达的任何元件,包括核酸或其衍生物(即mRNA)的复制(replication)、重复(duplication)、转录、剪接、翻译、稳定性和/或转移。

[0133] 本领域技术人员会理解,调节序列的选择可取决于这样的因素,如载体本身、宿主细胞、所期望的表达水平等。启动子特别重要。在本发明的上下文中,其对于引导核酸分子在许多类型的宿主细胞中的表达可以是组成型的(constitutive)或者特异于特定的宿主细胞(例如肝特异性调节序列)或者应对特定事件或外源因素而进行调节(例如通过温度、营养添加物、激素等)或者根据病毒循环的阶段调节(例如晚期或早期)。为了优化载体的生产并设法防止所表达的多肽的毒性,也可以使用在生产步骤期间应对特定事件或外源因素而受阻抑的启动子。

[0134] 适宜于哺乳动物细胞中组成型表达(constitutive expression)的启动子包括但不限于巨细胞病毒(CMV)即时早期启动子(Boshart等,1985,Cell141:521)、RSV启动子、腺病毒主要晚期启动子、磷酸甘油激酶(PGK)启动子(Adra等,1987,Gene60:65)、单纯疱疹病毒(HSV)-1的胸苷激酶(TK)启动子和T7聚合酶启动子。痘苗病毒启动子特别适宜于痘病毒载体的表达。代表性例子包括但不限于痘苗7.5K、H5R、11K7.5(Erbs等,2008,Cancer Gene Ther.15:18)、TK、p28、p11和K1L启动子,以及合成启动子如Chakrabarti等(1997,Biotechniques23:1094)、Hammond等(1997,J.Virological Methods66:135)和Kumar and Boyle(1990,Virology179:151)中所述的那些以及早期/晚期嵌合启动子。适宜于麻疹介导的表达的启动子包括但不限于引导麻疹转录单元表达的任何启动子(Brandler and Tangy,2008,CIMID31: 271)。肝特异性启动子包括但不限于以下的那些启动:HMG-CoA还原酶(Luskey,1987,Mol.Cell.Biol.7:1881);固醇调节元件1(SRE-1;Smith等,1990,J.Biol.Chem.265:2306);白蛋白(Pinkert等,1987,Genes Dev.1:268);磷酸烯醇丙酮酸羧基激酶(PEPCK)(Eisenberger等,1992,Mol.Cell Biol.12:1396); α -1抗胰蛋白酶(Ciliberto等,1985,Cell141:531);人转铁蛋白(Mendelzon等,1990,Nucleic Acids Res.18:5717);和FIX(US5,814,716)基因。

[0135] 本领域技术人员会理解,本发明的核酸分子表达的调节元件还可以包含另外的元件,以用于:转录的适宜起始、调节和/或终止(例如polyA转录终止序列),mRNA的转移(例如核定位信号序列)、加工(例如剪接信号)、和稳定性(例如内含子和非编码5'和3'序列),翻译(例如起始区Met、三联前导序列、IRES核糖体结合位点,Shine-Dalgarno序列等)至宿主细胞或生物体内,以及纯化步骤(例如标签)。

[0136] 本发明特别优选的实施方案涉及选自如下的载体(或病毒颗粒):

[0137] • 缺陷Ad载体,其包含插入至E1区位置的置于启动子(如CMV启动子)控制之下的核酸分子,并且编码突变体聚合酶多肽或融合蛋白,其中所述突变体聚合酶多肽包含SEQ ID NO:5所示氨基酸序列,所述融合蛋白包含SEQ ID NO:6或SEQ ID NO:8所示氨基酸序列;

[0138] • 缺陷Ad载体,其包含插入至E1区位置的置于启动子(如CMV启动子)控制之下的核酸分子,并且包含SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:15所示的核苷酸序列;

[0139] • 复制缺陷Ad载体特别是缺陷AdCh3,其包含插入至E1区位置的置于启动子(如CMV启动子)控制之下的核酸分子,并且包含SEQ ID NO:16或SEQ ID NO:17所示的核苷酸序列;

[0140] • MVA载体,其包含置于痘苗启动子(如7.5K或pH5R启动子)控制之下的核酸分子,并且编码突变体聚合酶多肽或融合蛋白,其中所述突变体聚合酶多肽包含SEQ ID NO:5或SEQ ID NO:10所示氨基酸序列,所述融合蛋白包含SEQ ID NO:6、SEQ ID NO:8或SEQ ID

NO:12所示氨基酸序列;和

[0141] • MVA载体,其包含置于痘苗启动子(如7.5K或pH5R启动子)控制之下的核酸分子,并且包含SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:15所示的核苷酸序列。优选地,所述核酸分子被插入至MVA基因组的缺失III。

[0142] 如果需要,本发明的载体可以还包含一个或多个转基因,例如与本发明的核酸分子一起在宿主细胞或生物体中表达的感兴趣的基,以旨在改善对HBV感染或者任何由HBV感染引起的或与其相关的疾病或病况的治疗或保护活性。适宜的转基因包括但不限于免疫调节剂如细胞因子和源自潜在共感染生物体的任何其它抗原(例如HIV、结核分枝杆菌(*Tuberculosis mycobacterium*)等)。如果使用转基因,其可由本发明的载体表达或者由组合使用的独立载体(其可与本发明载体相同或不同)表达。例如,可以设想组合使用表达聚合酶多肽或本发明的融合蛋白的腺病毒以及表达免疫调节剂的腺病毒。

[0143] 根据优选的实施方案,本发明的载体是感染性病毒颗粒的形式。通常,通过包括如下步骤的方法来产生此类病毒颗粒:(a)将本发明的病毒载体引入到适宜的细胞系中,(b)在适宜的条件下培养所述细胞系从而使所述感染性病毒颗粒能够产生,(c)从所述细胞系的培养物回收所产生的病毒颗粒,和(d)任选纯化所述回收的病毒颗粒。

[0144] 当病毒载体是缺陷型时,所述颗粒通常在补充细胞系中产生或者经使用辅助病毒产生,其反式(*in trans*)提供非功能性病毒基因。例如,用于补充E1-缺失的腺病毒载体的适宜细胞系包括293细胞(Graham等,1997,*J.Gen.Virol.*36:59-72)以及HER-96和PER-C6细胞(例如Fallaux等,1998,*Human Gene Ther.*9:1909-1917;W097/00326)或这些细胞系的任意衍生物。但是本领域中所描述的任何细胞系也都可用于本发明的背景中,特别是产生用于人类用途产物的任何细胞系,如Vero细胞、HeLa细胞核禽类细胞特别适宜于传代痘病毒载体。适宜的禽类细胞包括但不限于从得自孵化的卵的鸡胚胎制备的原代鸡胚胎成纤维细胞(CEF),和鸭细胞系(例如W003/076601,W02009/004016,W02010/130756和US2011-008872中所述的)。

[0145] 感染性病毒颗粒可回收自细胞上清和/或裂解后的细胞。可以根据标准技术将其进一步纯化(层析、在氯化铯梯度中超离心,如例如W096/27677,W098/00524,W098/22588,W098/26048,W000/40702,EP1016700和W000/50573中所述)。

[0146] 本发明还涵盖已经过修饰以允许优先靶向特定宿主细胞的载体或病毒颗粒。靶向的载体的特征是在其表面存在配体,所述配体能够识别并结合至细胞和表面暴露的成分如细胞特异性标记物(例如HBV-感染的细胞)、组织特异性标记物(例如肝特异性标记物)、以及病毒(例如HBV)抗原。适宜配体的例子包括针对HBV抗原结构域的抗体或其片段。可以通过将配体遗传插入至病毒表面上存在的多肽(例如腺病毒纤维、五位体(penton)、pIX或痘苗p14基因产物)来实现靶向。

[0147] 宿主细胞

[0148] 在另一方面,本发明还涉及包含本发明的核酸分子或载体(或病毒颗粒)的宿主细胞。

[0149] 如本文所用,术语“宿主细胞”应宽泛地理解,而不应受到涉及组织、器官或分离的细胞中特定组成的任何局限。此类细胞可以是单一类型的细胞或者不同类型细胞的组,如培养的细胞系、原代细胞和增殖细胞。在本发明的上下文中,术语“宿主细胞”包括原核细

胞、低等真核细胞如酵母、和其它真核细胞如昆虫细胞、植物和哺乳动物(例如人或非人)细胞以及能够产生本发明载体的细胞(例如293、HER96、PERC.6细胞、Vero、HeLa、CEF、鸭细胞系等)。此术语包括可以或者已经接受了本文所述载体的细胞,以及此类细胞的后代。可以在常规发酵生物反应器、烧瓶、和培养皿中培养宿主细胞。可以在适宜于给定宿主细胞的温度、pH和氧气含量下进行培养。本文不会详细描述各种原核和真核宿主细胞以及已知用于产生本发明所使用的多肽和载体的方法。

[0150] 根据本发明具体的实施方案,宿主细胞可进一步被包裹。细胞包裹技术是本领域熟知的。

[0151] 本发明的另一方面是用于重组产生本发明的突变体聚合酶多肽或融合蛋白的方法,其中使用本发明的载体(或感染性病毒颗粒)和/或宿主细胞。通常,所述方法包括(a)将载体引入适宜的宿主细胞以产生转染的或感染的宿主细胞,(b)在适宜于所述宿主细胞生长非人条件下体外培养所述转染的或感染的宿主细胞,(c)回收细胞培养物,和(d)任选地,从所回收的细胞和/或培养物上纯化所述突变体聚合酶多肽或融合蛋白。

[0152] 本领域技术人员应能够知晓本领域中可用的许多表达系统以用于在适宜的宿主细胞(如上文所述的那些)中表达突变体聚合酶多肽或融合蛋白,以及知晓用于将载体引入宿主细胞的方法。此类方法包括但不限于,显微注射、CaPO₄-介导的转染、DEAE-葡聚糖-介导的转染、电穿孔、脂染/脂质融合、基因枪、转导、病毒感染以及通过各种手段直接施用至宿主生物体内。本发明的载体可与转染试剂组合使用以便促进引入宿主细胞,如多阳离子聚合物(例如壳聚糖、聚甲基丙烯酸酯、PEI等)和阳离子脂质(例如DC-Chol/DOPE、目前可得自Promega的转染阳离子脂质体(transfectam lipofectin))。

[0153] 可以在常规发酵生物反应器、烧瓶、和培养皿中培养宿主细胞。可以在适宜于给定宿主细胞的温度、pH和氧气含量下进行培养。继而可以通过熟知的纯化方法来纯化所述突变体聚合酶多肽或所述融合蛋白,包括硫酸铵沉淀、酸萃取、凝胶电泳;过滤和层析方法(例如反相、大小排阻、离子交换、亲和性、疏水相互作用、羟基磷灰石、高效液相层析等)。要使用的条件和技术取决于如下因素:例如净电荷、分子量、疏水性、亲水性,并且对于本领域技术人员而言是显见的。另外,纯化的水平将取决于用途。

[0154] 组合物

[0155] 在另一方面,本发明提供了组合物,其包含至少本文所述的突变体聚合酶多肽或融合蛋白、核酸分子、载体(例如感染性病毒颗粒)、或宿主细胞(本文也称作“活性物质”)或其任意组合(例如本文所述的不同多肽或载体/病毒颗粒的组合或者不同基因型的组合)。优选地,所述组合物是药物组合物,其治疗有效量的活性物质之外还包含药学可接受的载剂。

[0156] 如本文所用,“药学可接受的载剂”包括任何和所有的载体(carrier)、溶剂、稀释剂、赋形剂、佐剂、分散介质、包被、抗菌和抗真菌剂、和吸收减缓剂,及类似的等,其适宜于施用至宿主生物体以及特别是人类。

[0157] 如本文所用,“治疗有效量”是足以减缓一种或多种通常与HBV感染相关的症状的剂量或者足以减缓由HBV感染引起的或与其相关的任何疾病或病况的剂量。当涉及预防应用时,此术语是指足以防止或延缓HBV感染的建立的剂量。“治疗性”组合物被设计和施用于已经感染HBV的宿主生物体,目的在于缓解或改善至少一种由所述HBV感染引起或与其相关

的疾病或病况,最终与一个或多个本文所述的常规治疗手段组合(例如用核苷、核苷酸类似物和/或基于IFN的疗法治疗)。例如,诱导免疫应答的治疗有效量可以是引起免疫系统激活所需的量(例如导致发展出抗-HBV应答)。

[0158] 术语“宿主生物体”一般是指需要或者可得益自本发明的任意产品和方法的脊椎动物,特别是许多哺乳动物物种以及尤其是家养动物、农牧动物、体育动物、和灵长类包括人类,如已经诊断为感染HBV或有感染HBV风险的生物体(其因而疑似患有或者有风险会患有由HBV感染引起的或与其相关的疾病或病况)。在优选的实施方案中,所述宿主生物体是慢性感染了HBV病毒或者共感染了HBV病毒和其它病毒(例如人类免疫缺陷病毒HIV)的人类患者。感染性HBV可以来自与所述突变体聚合酶多肽或本发明使用的任何其它HBV多肽/肽所源自的HBV(例如基因型D)相同的基因型、毒株或分离物,或者其可以来自不同的基因型(例如基因型B或C)。本发明人最近研究了基于基因型D的疫苗组合物的交叉反应潜力(参见美国申请13/423,193)。大规模的计算机研究显示出,HBV聚合酶、核心和Env蛋白的氨基酸序列在基因型B、C和D之间于蛋白整体水平高度保守,而在T细胞表位水平也是如此。在适宜动物模型中的体内免疫支持了诱导交叉反应T细胞应答(识别基因型B和C表位)的能力。即使将此研究局限于HLA-A2表位,其也证明了基于基因型D抗原的疫苗组合物能够诱导与其它HBV基因型交叉反应的T细胞应答。

[0159] 本发明的组合物适宜被缓冲,以便适宜以生理或微碱性的pH来用于人类用途(例如大约pH7至大约pH9)。适宜的缓冲剂包括但不限于磷酸盐缓冲剂(例如PBS)、碳酸氢盐缓冲剂和/或Tris缓冲剂。

[0160] 本发明的组合物还可包含适用于人类或动物用途的稀释剂。其优选是等渗的、低渗的或微弱高渗的,并具有相对低的离子强度。代表性的例子包括蒸馏水、生理盐水(例如氯化钠)、Ringer's溶液、葡萄糖、海藻糖或蔗糖溶液、Hank's溶液、以及其它水成的生理平衡盐溶液(参见例如最新版的Remington:The Science and Practice of Pharmacy, A.Gennaro,Lippincott,Williams&Wilkins)。

[0161] 本发明的组合物中所包括的药学可接受的载剂必须也允许在生产和长期储存的条件下于冷冻(例如-70℃,-20℃)、冷藏(例如4℃)或室温中保持其稳定性(即至少一个月,优选至少一年)。对于这一方面,formulations which are particularly adapted to特别适于本发明的组合物的配制物包括(a)1M蔗糖,150mM NaCl,1mM MgCl₂,54mg/l Tween80,10mM Tris pH8.5(尤其在活性物质是腺病毒载体时);(b)10mg/ml甘露醇,1mg/ml HSA,20mMTris,pH7.2,和150mM NaCl;以及(c)生理盐水。

[0162] 另外的药学可接受的赋形剂可用于提供所期望的药学或药代动力学性质,包括例如修饰或保持配制物的pH、渗透性、粘度、澄清、颜色、无菌、稳定性、溶出率(rate of dissolution),修饰或保持释放或吸收至日人或动物生物体中,促进穿越血液屏障的转移或进入特定的器官(例如肝)。

[0163] 另外,本发明的组合物可包含适宜于在人类中全身性或粘膜应用的一种或多种佐剂。优选地,所述佐剂能够刺激对本发明组合物的免疫性,特别是T细胞介导的免疫性,例如通过Toll样受体(TLR)如TLR-7、TLR-8和TLR-9。可使用的佐剂的代表性例子包括但不限于铝、矿物油乳剂如弗氏完全佐剂和弗氏不完全佐剂(IFA)、脂多糖或其衍生物(Ribi等,1986,Immunology and Immunopharmacology of Bacterial Endotoxins,Plenum

Publ. Corp., NY, p407-419), 皂苷类如QS21(Sumino等, 1998, J. Virol. 72:4931; W098/56415)、咪唑-喹啉化合物如咪喹莫特(Imiquimod)(Suader, 2000, J. Am Acad Dermatol. 43:S6), S-27609(Smorlesi, 2005, Gene Ther. 12:1324)和相关的化合物如W02007/147529中所述的, 胞嘧啶磷酸酯鸟苷寡聚脱氧核苷酸如CpG(Chu等, 1997, J. Exp. Med. 186:1623; Tritel等, 2003, J. Immunol. 171:2358)以及阳离子肽如IC-31(Kritsch等, 2005, J. Chromatogr Anal. Technol Biomed Life Sci 822:263)。

[0164] 本发明的组合物适宜于各种使用方式。

[0165] 术语“施用”(以及施用的任何形式如“施用了”)如本文所用是指将治疗剂如本文所述的突变体聚合酶多肽、融合蛋白、核酸分子、载体递送至宿主细胞或生物体。现有技术中有许多方法和手段如直接施用至宿主生物体。

[0166] 直接施用可通过全身、局部或粘膜途径进行。全身施用包括例如皮下、皮内、肌肉、静脉内(例如注射至供应肝脏的静脉如门静脉)、腹膜内、瘤内、血管内、动脉内注射(例如通过肝动脉灌注)以及划破(scarification)。可以使用常规针筒和针头进行注射, 或使用本领域中任意其它适宜的设备(例如电穿孔)。或者, 本发明的组合物可经粘膜途径施用, 如口服/消化、鼻内、气管内、肺内、阴道内或直肠内途径。can可通过小迪的喷雾或烟雾、喷雾剂、或干粉状的组合使用适宜的投放器来进行呼吸道中的施用。也可以使用经皮手段来进行局部施用(例如贴片等)。

[0167] 在本发明的上下文中, 所述组合物优选配制为用于肌肉、皮下、皮内施用或划破。

[0168] 本发明的组合物可以是各种形式, 例如固态、液态或冷冻的。固态(例如干粉化的或冻干的)组合物可以通过涉及真空干燥和冷冻干燥的方法来获得。对于粘膜施用, 可将所述组合物配制成胃抗性胶囊和颗粒以用于口服施用, 配制成栓剂以用于直肠或阴道施用, 最终与有助于提高粘膜膜的孔大小的吸收增强剂组合。此类吸收增强剂通常是这样的物质, 其与粘膜膜的磷脂结构域具有结构相似性, 如脱氧胆酸钠、甘胆酸钠、二甲基- β -环糊精、十二烷基-1-溶血磷脂胆碱)。

[0169] 适宜的剂量可以作为各种参数的函数来调节, 特别是施用方式; 所使用的组合物; 宿主生物体的年龄、健康和体重; 症状的性质和程度; 同时进行治疗的类型; 治疗的频率; 和/或预防或治疗的需要。基于相关的情况, 施行者能够常规地确定所需计算的其它改善。

[0170] 作为一般指引, 含载体组合物的适宜剂量从 10^5 至约 10^{13} vp(病毒颗粒)、iu(感染性单位)或pfu(噬斑形成单位)不等, 取决于所使用的载体和定量技术。可用于评估样品中存在的vp、iu和pfu的量的技术是本领域内常规的。例如, 腺病毒颗粒(vp)的数目通常是通过测量A260光吸收或通过HPLC来确定的, iu滴度是通过定量DBP免疫荧光确定的, 而pfu则是通过计数许可的细胞感染后的噬斑数目来确定的。根据FDA指引, 优选vp/iu比率低于100。优选的剂量含有约 10^5 至约 10^{12} 的vp, 特别优选的剂量为约 5×10^8 、约 10^9 、约 5×10^9 、约 10^{10} 、约 5×10^{10} vp或约 10^{11} vp的本发明的腺病毒载体。约 5×10^5 至约 10^9 pfu的剂量优选用于基于痘苗(MVA)的组合物, 特别优选对的剂量为约 5×10^6 、约 10^7 、约 5×10^7 、约 10^8 或约 5×10^8 pfu。基于载体质粒的组合物可以10 μ g-20mg之间的剂量施用, 有利的是在100 μ g-2mg之间。蛋白组合物可以10ng-20mg之间的一个或多个剂量施用, 特别优选的剂量为约0.1 μ g至约2mg的免疫原性多肽每kg体重。可以单次剂量或在一定时间间隔之后的一或多次重复剂量来进行施用。

[0171] 在另一具体的实施方案中, 本发明的突变体聚合酶多肽、融合蛋白、核酸分子、载

体、宿主细胞或组合物,可以与另外的多肽或肽或编码该另外的多肽或肽的载体组合使用。优选地,所述另外的多肽或肽是HBV抗原,特别优选如本文所述的Hbc多肽和/或一个或多个HBs免疫原性结构域。所述HBV多肽或肽可以从载体表达,特别是选自如下的载体:质粒DNA、腺病毒(例如Ad5、AdCh3、AdCh63等)、痘病毒(例如牛痘如MVA)和麻疹载体。因此本发明还涉及这样的组合物,其包含编码本发明的突变体聚合酶多肽或融合蛋白的载体以及编码至少一种另外的多肽/肽(如本文所述的HBV核心和/或HBs免疫原性结构域)的载体的混合物。

[0172] 本发明的突变体聚合酶多肽、融合蛋白、核酸分子、载体、宿主细胞或组合物可用于治疗各种疾病和病理状况的方法,特别是那些由HBV感染引起的或与其相关的。因此,本发明还涵盖了,本发明的突变体聚合酶多肽、融合蛋白、核酸分子、载体、宿主细胞或组合物用于根据本文所述的疗法治疗或预防HBV感染、HBV-相关疾病和病理状况的用途,特别是慢性HBV感染。本发明还涉及在有需要的生物体中进行治疗的方法,包括根据本发明的疗法向所述生物体施用至少一次此类活性物质,其中所述施用是以足以治疗或预防HBV感染(例如特别是慢性HBV感染)的量或者足以改善一个或多个与HBV-相关疾病和病理状况相关的症状的量。在特定的实施方案中,可以根据本文所述的疗法利用本发明的活性物质和方法来打破HBV慢性对象中通常会遇到的HBV-特异性免疫耐受。

[0173] 术语“治疗(treating)”(以及任何形式的治疗如“治疗(treatment)”)如本文所用是指预防(例如对有感染HBV的风险的对象进行预防)和/或医疗(therapy)(例如在诊断为感染了HBV的对象中)。其对于治疗HBV慢性感染和/或HBV感染患者中的肝损伤(包括肝硬化和肝癌)特别有用。治疗需要向宿主细胞或生物体外部施用或内部施用医疗剂如本文所述的突变体聚合酶多肽,最终与其它HBV多肽组合或者与with the护理标准(SOC)(例如用核苷或核苷酸类似物治疗)组合。

[0174] 通常,在根据本文所述的疗法施用至宿主生物体内之后,本发明的突变体聚合酶多肽、融合蛋白、核酸分子、载体、宿主细胞和/或组合物对被治疗的宿主生物体提供医疗益处(相对于基准状态或相对于不进行治疗所预期的状态)。所述医疗益处可以由医师或其它技术护理人员通常使用的相关的临床测试来证实,包括例如,经治疗的生物体的血液、血浆或血清中所定量的HBV病毒负载的降低,和/或肝酶活性水平的降低(例如丙氨酸氨基转移酶(ALT)和/或天冬氨酸氨基转移酶(AST)),和/或疾病稳定了的(未恶化)状态(例如通常与HBV感染相关的病况的稳定或减轻,如肝炎症/脂肪化/纤维化),和/或血清标记物水平的降低如HBeAg或HBsAg(例如HBe或HBs血清转化),和/或所治疗的生物体对常规疗法反应的改善和/或和/或相对于不治疗所预期的存活而言存活延长。

[0175] 在本发明的上下文中,所述医疗益处可以是暂时的(在中断施用之后持续f一个或两个月)或者是持续的(几个月或几年)。由于临床状态的自然进程在不同患者之间可以有很大不同,并不需要在每个经治疗的生物体中均观察到所述医疗益处,而只需要在显著数目的生物体中观察到即可(例如可通过本领域已知的任何统计学检验来确定两个组之间的统计学显著差异,例如Tukey参数检验、Kruskal-Wallis检验、根据Mann和Whitney的U检验、Student's t-检验、Wilcoxon检验等)。

[0176] 此类测量可以在本文所述的施用之前进行(基准)以及在治疗期间的各个时间点进行并且在治疗终止后至少进行12周。作为一般指引,可以使用PCR测定或本领域可接受的

任何其它方法来测定病毒负载(例如Roche Ampli Prep/Cobas Taqman测定v2.0, Abbott实时乙肝病毒表现测定)。在优选的实施方案中,施用本发明的突变体聚合酶多肽、融合蛋白、核酸分子、载体、宿主细胞和/或组合物会导致病毒负载的降低(暂时的或持续至少一个 $\log_{10}$ 、优选至少1.5个 $\log_{10}$ 、更优选至少2个 $\log_{10}$),这是相较于基准时的测量的病毒负载或者相较于对照组(未治疗对象)。施用本文所述的活性物质可导致与基准或对照组相比至少暂时回归正常的ALT和/或AST值。肝酶活性的水平可以常规地在医学实验室或医院进行评估。或者,施用本文所述的活性物质会导致血清标记物HBe和/或HBs至少暂时的降低至少一个 $\log_{10}$ 、优选至少1.5个 $\log_{10}$ 更优选至少2个 $\log_{10}$ (相较于基准时测量的血清标记物水平或者相较于对照组(未治疗对象))。HBV血清标记物的水平可以常规地在医学实验室或医院进行评估,并且大量的试剂盒也是商业可购买的(例如Abbott Laboratories, Organon Technika开发的免疫测定)。

[0177] 优选地,使用或施用本发明的突变体聚合酶多肽、融合蛋白、核酸分子、载体、宿主细胞和/或组合物以用于在经治疗的生物体中激发或刺激免疫应答。因此,本发明还涵盖通过在宿主生物体中施用本发明的突变体聚合酶多肽、融合蛋白、核酸分子、载体、宿主细胞和/或组合物,以用于激发或刺激针对HBV的免疫应答的方法。

[0178] 所激发的或刺激的免疫应答可以是特异性的(即针对HBV表位/抗原)和/或非特异性的(先天的)、体液的和/或细胞的。在本发明的上下文中,优选所述免疫应答是针对HBV多肽/表位的T细胞应答CD4+或CD8+-介导的,或者是这两者。

[0179] 本文所述活性物质激发免疫应答的能力可以使用各种直接或间接测定在体外或体内进行评估,这是本领域内的标准技术。下文中的实施例部分也对测试和验证进行了说明。

[0180] 对于可用于评估免疫应答的发起和激活的技术的一般描述,参见例如Coligan等(1992和1994, *Current Protocols in Immunology*; ed J Wiley&Sons Inc, National Institute of Health或随后的版本)。刺激体液应答的能力可以通过抗体结合和/或竞争结合来确定(参见例如Harlow, 1989, *Antibodies*, Cold Spring Harbor Press)。

[0181] 对于非特异性免疫的评估,可以例如通过测量NK/NKT-细胞(例如激活的代表活性和水平),以及IFN相关细胞因子和/或趋化因子产生盒, TLR和其它先天免疫标记物的激活(Scott-Algara等, 2010 *PLoS One* 5(1), e8761; Zhou等, 2006, *Blood* 107, 2461-2469; Chan, 2008, *Eur. J. Immunol.* 38, 2964-2968)。

[0182] 评估细胞免疫可以例如通过如下方式:使用常规生物测定对激活的T细胞(包括那些衍生自CD4+和CD8+的T细胞)所产生的细胞因子进行定量(例如通过ELISpot、通过多参数流式细胞术或ICS、通过使用多元技术或ELISA进行的细胞因子谱分析等来对T细胞进行表征和定量);确定T细胞的增殖能力(例如通过 $[^3\text{H}]$ 胸腺嘧啶掺入测定来进行T细胞增殖测定);测定抗原特异性T淋巴细胞在致敏化的对象中的细胞毒性能力;或通过适宜的动物模型的免疫。

[0183] 本发明的突变体聚合酶多肽、融合蛋白、核酸分子、载体、宿主细胞和/或组合物的免疫原性能力也可在动物模型中确认,可以用适宜的感染性或肿瘤诱导物质来攻击所述动物模型(例如表达HBV基因产物的痘苗病毒或单核细胞增生李斯特氏菌(*Listeria Monocytogenes*))或者用编码全长HBV基因组的DNA来注射所述动物模型(如Huan等, 2010,

Proc.Natl.Acad.Sci.107:9340中所述),以确定所述感染性或肿瘤诱导物质的中和以及最终对相关症状的部分抗性,反映出抗-HBV免疫应答的诱导或增强。示例性的动物模型包括但不限于实施例中所所述的HLA-A2.1转基因小鼠、和Chisari等(1996, Curr.Top.MicroBiol.Immunol.,206:149)和Halverscheid等(2008,J.Med.Virol.80:583)中所所述的HBV转基因小鼠。

[0184] 所述用途或方法包括一次或多次施用(1、2、3、4、5、6、7、8、9、10等)治疗有效量的所述活性物质,所述施用彼此之间由适宜的时间段所间隔并且由相同的施用途径或不同的施用途径在相同部位或不同部位来进行。在本发明背景中特别优选肌肉和皮下途径。彼此间隔大约一周的三次皮下施用特别适宜于基于MVA的组合物和载体,而一次或两次肌肉(彼此可间隔大约一个月或更久)或皮下施用特别适宜于基于Ad的组合物和载体。第一系列的施用两个月或若干个月之后可接着使用相同的活性物质进行一次或多次随后的施用,以便唤起抗-HBV免疫应答。

[0185] 如果需要,本发明的方法或用途可以与一个或多个常规疗法组合进行(例如放疗、化疗和/或手术)。优选地,本发明的方法或用途与一种或多种药物关联,所述药物可用于治疗或预防HBV感染、HBV相关疾病和病理状况。它们的施用可以在本发明活性物质的施用之前、同时或之后。适宜药物的代表性的例子包括但不限于聚合酶抑制剂、RNase H抑制剂、核苷类似物、核苷酸类似物、TLR激动剂、IFN、N-糖基化抑制剂、siRNA、反义寡核苷酸、抗-HBV抗体、免疫调节剂、治疗疫苗和通常用于治疗HBV相关肝癌的抗肿瘤物质(例如阿霉素、阿霉素(adriamycin)与碘油或索拉菲尼(sorafenib))。另外,所述活性物质可以与其它治疗性疫苗组合使用,如合成肽、重组抗原、VLP、编码HBV蛋白(核心、preS1、PreS2、S和/或聚合酶)的载体,其特别适宜于引发抗-HBV体液应答。本发明特别适宜的方法或用途是,与护理标准组合,以及特别是与细胞因子治疗组合(例如IFN α ,聚乙二醇化的IFN α 2a或2b如Pegasys(Roche)、Pegintron(Schering Plough)或IntronA(Schering Plough))和/或与核苷酸或核苷类似物(NUC)如拉米夫定、恩替卡韦、替比夫定、阿德福韦、阿德福韦酯(dipivoxil)或替诺福韦组合。用NUC治疗仅仅是部分有效的(1年的治疗之后仅在3-5%的对象中观察到感染的解决)并且需要长期医疗(可能是终生的)。预期本发明的活性物质和方法会带来能补充NUC对病毒复制的作用的免疫维度,由此导致对此类治疗的改善(例如降低实现医疗益处所需的NUC的剂量和缩短所需NUC治疗的时间)或解决感染百分比的提高(超过5%)。

[0186] 在具体的实施方案中,可以根据初免-加强(prime boost)疗法来进行本发明的方法或用途,其包括顺序施用一次或多次初免组合物以及一次或多次加强组合物。通常,初免和加强组合物使用包含或编码至少共同的抗原结构域的不同载体。另外,初免和加强组合物可以由相同的途径或不同的途径施用于相同部位或不同部位。例如,基于多肽的组合物可通过粘膜途径施用,而基于载体的组合物优选注射,例如皮下注射MVA载体、肌肉注射DNA质粒以及皮下或肌肉注射腺病毒载体。

[0187] 在一个实施方案中,用MVA载体进行初免以及用Ad载体进行加强,特别优选编码本文所述突变体聚合酶蛋白或融合蛋白(例如SEQ ID NO:8所示的融合蛋白)的MVA和/或Ad载体。将MVA载体施用至生物体一次或多次,随后施用腺病毒载体一次或多次,特别优选至少3次皮下施用MVA载体(以3天至3个月不等的时间段间隔)随后肌肉或皮下加强腺病毒载体(例如MVA初免之后大约1个月至1年)。

[0188] 在另一实施方案中,用质粒DNA载体进行初免以及用MVA载体进行加强,特别优选编码本文所述突变体聚合酶蛋白或融合蛋白(例如SEQ ID NO:8所示的融合蛋白)的质粒和/或MVA载体。将DNA载体施用至生物体一次或多次,随后施用MVA载体一次或多次,特别优选至少3次肌肉施用DNA载体(以2周至3个月不等的时间段间隔)和至少一次皮下加强MVA载体(例如DNA初免之后大约1个月至1年)。优选地,通过电穿孔施用所述DNA载体。

[0189] 本发明还涉及试剂盒,其用于HBV感染或者HBV相关疾病或病理状况的治疗,其中所述试剂盒包含多个选自如下的活性物质:本发明的突变体聚合酶多肽、融合蛋白、核酸分子、载体、宿主细胞和/或组合物以及用于将所述多个活性物质施用至有需要的生物体的说明。更优选地,所述生物体是慢性感染了HBV的患者。

[0190] 上文所引用的所有专利公开、出版物和数据库条目都具体地以其整体援引加入本文,就如同具体且独立地指出将这些单独的专利、出版物或条目援引加入本文。

[0191] 附图描述

[0192] 图1说明了本发明的突变体聚合酶多肽、融合蛋白和抗原组合。

[0193] 图2说明了在用质粒pTG18188(Core-Pol-Env1-Pol-Env2-Pol)、pTG18194(Core-Pol)或pTG13135(空)免疫HLA-A2转基因小鼠后进行的Elispot IFN γ 测定。结果表示为 10^6 细胞的斑点数,对应特异于每种HBV HLA-A2肽的产IFN γ 细胞的频率,在经免疫小鼠的 10^6 脾细胞的实验中评估。每个条代表用一种或其它质粒接种的个体小鼠并由其参考编号表示(1.1或2.3……等),并且用相同质粒免疫的所有小鼠的平均也在每组中标出(图中标为“平均”的条)。对于个体小鼠和平均,将特异于不同测试肽的产IFN γ 细胞的频率堆叠。条由不同的符号填充,每种符号代表针对一种具体HBV肽的应答,如图中标注所示。

[0194] 图3A-F说明了在用质粒pTG18188(Core-P-E1-P-E2-P或Core-Pol-Env1-Pol-Env2-Pol)、pTG18194(Core-Pol)或pTG13135(空)免疫HLA-A2转基因小鼠后进行的ICS测定。结果表示为产IFN γ (单独或与TNF α 组合)的CD8(图3A、3B、3D继而3F)、或CD4(图3C和3E)T细胞的百分比,特异于每种HBV HLA-A2肽(图3A)或涵盖HBV核心的肽汇集(图3B和3C)、聚合酶(图3D和3E)和env(图3F)抗原。每个条代表用一种或其它质粒接种的个体小鼠并由其参考编号表示(1.1或2.3……等),并且用相同质粒免疫的所有小鼠的平均也在每组中标出(图中标为“平均”的条)。对于个体小鼠和平均,将特异于不同测试肽的CD8或CD4T细胞的频率堆叠。条由不同的符号填充,每种符号代表针对一种具体HBV肽的应答,如图中标注所示。

[0195] 图4说明了在用腺病毒AdTG18201(Core-Pol-Env Ad)、AdTG18202(Core-Pol Ad)、AdTG18203(Pol Ad)和AdTG15149(空Ad)免疫HLA-A2转基因小鼠后进行的Elispot IFN γ 测定。结果表示为 10^6 细胞的斑点数,对应特异于每种涵盖感兴趣抗原的肽汇集的产IFN γ 细胞的频率,在经免疫小鼠的 10^6 脾细胞的实验中评估。每个条代表用一种或其它质粒接种的个体小鼠并由其参考编号表示(1.1或2.3……等),并且用相同质粒免疫的所有小鼠的平均也在每组中标出(图中标为“平均”的条)。对于个体小鼠和平均,将特异于不同测试肽的产IFN γ 细胞的频率堆叠。条由不同的符号填充,每种符号代表针对一种具体HBV肽的应答,如图中标注所示。

[0196] 图5说明了在用AdTG18201(Core-Pol-Env Ad)免疫HLA-A2转基因小鼠后进行的Elispots IFN γ 测定。结果表示为 10^6 细胞的斑点数,对应特异于每种HBV HLA-A2表位、或不相关肽的产IFN γ 细胞的频率,在经免疫小鼠的 10^6 脾细胞的实验中评估。每个条代表用

AdTG18201接种的个体小鼠(用其参考号码表示,3.1to3.8),并且所有这些用AdTG18201免疫的小鼠的中位值也标出(图中标为“中位值”的条)。对于个体小鼠和中位值,将特异于不同HBV HLA-A2表位的产IFNg细胞的频率堆叠,而在存在不相关肽时所观测到的产IFNg细胞的频率在另外的条中标出(图中标为“Irrel”的条)。条由不同的符号填充,每种符号代表针对一种具体HBV肽的应答,如图中标注所示。

[0197] 图6说明了在用AdTG18201(Core-Pol-Env Ad)或AdTG15149(空Ad)免疫HLA-A2小鼠后进行的ICS测定。结果表示为单独或与TNFa组合产IFNg的CD8T细胞的百分比,特异于两种选择的肽汇集,称作PP8和PC1,并分别涵盖了HBV聚合酶的部分(氨基酸725-835)和HBV核心蛋白的部分(氨基酸1-100)。每个条代表用一种或其它腺病毒接种的个体小鼠并由其参考编号表示(1.1或3.2……等),并且用相同腺病毒免疫的所有小鼠的中位值也在每组中标出(图中标为“中位值”的条)。条由不同的符号填充,每种符号代表所检测的CD8T细胞产生的细胞因子,如图中标注所示。

[0198] 图7说明了在用AdTG18201(Core-Pol-Env Ad)或AdTG15149(空Ad)免疫HLA-A2小鼠后进行的体内CTL测定。结果表示为体内特异性裂解的百分比,也即特异于所测试的HBV HLA-A2表位的裂解。每个正方形或三角形符号代表个体小鼠,并且用相同腺病毒免疫的所有小鼠的平均也在每组中标出(图中以粗条符号表示)。

[0199] 图8说明了在用AdTG18201(Core-Pol-Env Ad)或AdTG15149(空Ad)免疫HBV转基因小鼠后进行的ICS测定。结果表示为产IFNg和特异于表位的TNFa的CD8T细胞的百分比,所述表位来自HBV聚合酶(2种肽的混合,称作VSA和N13F)或来自HBV包膜(1种肽F称作13L),并且在经接种的小鼠的脾和肝中发现了二者。每个条代表用一种或其它腺病毒接种的个体小鼠并由其参考编号表示(1.1或3.2……等),并且用相同腺病毒免疫的所有小鼠的中位值也在每组中标出(图中标为“中位值”的条)。图中标出的虚线对应实验的截断值,也即阈值,在其上观测到的CD8T细胞百分比被认 为是阳性免疫应答。

[0200] 图9说明了在用AdTG18201(Core-Pol-Env Ad)以不同剂量(从 10^5 iu至 10^9 iu)免疫HLA-A2转基因小鼠后进行的Elispots IFNg测定。结果表示为 10^6 细胞的斑点数,对应特异于每种HBV HLA-A2表位、或不相关肽、或仅存在培养基的产IFNg细胞的频率,在经免疫小鼠的 10^6 脾细胞的实验中评估。每个实线或虚线条代表用AdTG18201接种的个体小鼠并且所有用AdTG18201以一种具体剂量免疫的小鼠的平均也标出(阴影线的条)。不同剂量由不同颜色或符号表示,如图中标注所示。虚线表示截断值,如材料和方法部分中所定义,在其上所观测到的T细胞应答被认为是阳性的。

[0201] 图10说明了在用AdTG18202(Core-Pol Ad)根据不同的注射日程免疫HLA-A2转基因小鼠后进行的Elispots IFNg测定。小鼠被免疫一次(正方形)或者以一周的间隔接受3次注射(三角形)或一周的间隔接受6次注射(圆形)。结果表示为 10^6 细胞的斑点数,对应特异于每种HBV HLA-A2表位、或腺病毒载体、或不相关肽、或仅存在培养基的产IFNg细胞的频率,在经免疫小鼠的 10^6 脾细胞的实验中评估。每个符号(正方形,三角形或圆形)代表用AdTG18201接种的个体小鼠并且所有用AdTG18202以一种测试的日程免疫的小鼠的平均由粗实线表示。虚线表示截断值,如材料和方法部分中所定义,在其上所观测到的T细胞应答被认为是阳性的。

[0202] 图11说明了在用AdTG18202(Core-Pol Ad)根据不同的注射日程免疫HLA-A2转基

因小鼠后进行的Elispots IFNg测定。小鼠在监测T细胞应答之前2周被免疫一次(组1,由正方形表示)、或在监测T细胞应答之前20周被免疫一次(组2,由三角形表示)、或以2个月的间隔被免疫两次(组3,由圆形表示)、或以4个月的间隔被免疫两次(组4,由叉表示)、或以2个月的间隔被免疫三次(组5,由菱形表示)。对于除了组2之外的所有组,在最后一次注射2周之后监测T细胞应答。结果表示为 10^6 细胞的斑点数,对应特异于每种HBV HLA-A2表位、或不相关肽、或仅存在培养基的产IFNg细胞的频率,在经免疫小鼠的 10^6 脾细胞的实验中评估。每个符号(正方形、三角形、圆形、叉、菱形)代表用AdTG18202接种的个体小鼠并且所有用AdTG18201以一种测试的日程免疫的小鼠的平均由粗实线表示。虚线表示截断值,如材料和方法部分中所定义,在其上所观测到的T细胞应答被认为是阳性的。

实施例

[0203] 1.材料和方法

[0204] 如下所示的构建(参见图1)是根据如Maniatis等(1989,Laboratory Manual,Cold Spring Harbor Laboratory Press,Cold Spring Harbor NY或subsequent editions)中所述的一般遗传改造和分子克隆技术来进行的,或者在使用商业试剂盒时是根据厂商推荐进行的。PCR扩增技术是本领域技术人员熟知的(参见例如PCR protocols-A guide to methods and applications,1990,由Innis,Gelfand,Sninsky andWhite,Academic Press出版)。携带有氨苄青霉素抗性基因的重组质粒在大肠杆菌C600(Stratagene)中于补加了 $100\mu\text{g/ml}$ 抗生素的琼脂或液体培养基中复制。

[0205] MVA载体的构建由穿梭质粒和MVA基因组之间的同源重组来产生,如之前所述Erbs等(2008,Cancer gene Ther.15:18)。该“基本的”穿梭质粒含有多克隆位点、由缺失III的旁侧序列包围的痘苗病毒(VV)启动子以及在p11K7.5痘苗启动子(Falkner and Moss,1988)控制之下的大肠杆菌黄嘌呤鸟嘌呤磷酸核糖基转移酶(GPT)选择基因。简言之,用没有任何插入转基因的基因组MVA(MVA-null)感染了CEF细胞,并继而通过 CaCl_2 沉淀用穿梭质粒(携带感兴趣的基因,克隆于VV启动子的下游)转染。同源重组在MVA-null和所述穿梭质粒之间发生,并且通过多个步骤的霉酚酸选择分离了重组病毒。重组MVA病毒由PCR控制,在CEF中扩增,并通过噬斑测定在CEF上对病毒储备(stock)进行滴定。

[0206] 对于腺病毒载体构建,首先通过将感兴趣的核酸分子插入至所述“基本的”穿梭质粒pTG13135中而构建了腺病毒穿梭质粒。通常,将核酸分子插入至pTG13135的NheI和NotI限制酶且位点内(pTG13135含有CMV驱动的表达盒,其被腺病毒序列所包围(分别是腺病毒核苷酸1454及核苷酸35135781))以允许通过同源重组进一步产生载体基因组(Charrier等,1996,J.Virol.70:4805)。继而通过重组穿梭载体(经Bst1107I和PacI消化)与经ClaI消化线性化的pTG15375(编码完整的腺病毒基因组)之间的同源重组来获得腺病毒载体。所得的腺病毒载体是E3(核苷酸2786730743)和E1(核苷酸4553512)缺失的,其中E1区域置换为表达盒,其从5'至3'含有:CMV即时早起增强子/启动子、嵌合的人 β 球蛋白/IgG内含子(如可得自Promega的pCI载体中所发现的)、感兴趣的核酸分子和SV40晚期多腺苷酸化信号。继而通过将经PacI线性化的病毒基因组转染至E1补充细胞系中而获得腺病毒颗粒。病毒传代、纯化和滴定如此前所述那样进行(Erbs等,2000,Cancer Res.60:3813)。

[0207] 1.1.载体构建和产生

[0208] 下文说明的载体已经过改造以表达突变体聚合酶多肽(最终融合至核心多肽和/或包膜蛋白的免疫原性结构域)。所有HBV序列源自HBV毒株Y07587,其序列如国际数据库中(Genbank Y07587)及不同出版物中所述。其是血清型ayw的基因型D病毒。

[0209] 如下实施例说明了截短的核心多肽(aa1-148)与突变的聚合酶多肽(命名为Pol*)的融合,所述突变的聚合酶多肽包含两个内部缺失(从位置538至544以及从位置710至742)和4个氨基酸取代(分别为D689H、V769Y、V776Y和D777H),如SEQ ID NO:6中所示,以及更长的版本还包含两个免疫原性Env结构域(分别为Env1和Env2,从HBs蛋白的氨基酸14延伸至51,和从氨基酸165延伸至194)插入到缺失的pol区域的位置中(如SEQ ID NO:8所示)。

[0210] 1.1.1.表达截短的Core-Pol*-Env1-Env2(或Core-Pol-Env1-Pol-Env2-Pol)融合的质粒和腺病毒载体的构建和产生

[0211] 编码截短的Core-Pol*-Env1-Env2融合蛋白(氨基酸序列如SEQ ID NO:8所示)的合成基因(3024个核苷酸,如SEQ ID NO:15中所示)由GENEART(Regensburg, Germany)合成。将该合成的片段插入至pTG13135穿梭质粒的NheI和NotI限制性酶切位点中,提供了pTG18188。继而通过在经Bst1107I和PacI消化的pTG18188与经ClaI消化线性化了的pTG15375之间的同源重组而获得了腺病毒载体。所得的腺病毒载体pTG18201是E3和E1缺失的,其中E1区域被置换为表达盒,该表达盒含有由CMV启动子驱动的编码截短的Core-Pol*-Env1-Env2的合成序列。通过将经PacI线性化的病毒基因组转染至E1补充细胞系而获得腺病毒颗粒(AdTG18201)。

[0212] 1.1.2.表达截短的Core-Pol*的质粒和腺病毒载体的构建和产生

[0213] 编码截短的Core-Pol*融合蛋白的合成基因(2820个核苷酸,如SEQ ID NO:14中所示)由GENEART(Regensburg, Germany)合成。将该合成的片段插入至pTG13135穿梭质粒的NheI和NotI限制性酶切位点中,提供了pTG18194。继而通过在经Bst1107I和PacI消化的pTG18194与经ClaI消化线性化了的pTG15375之间的同源重组而获得了腺病毒载体。所得的腺病毒载体pTG18202是E3和E1缺失的,其中E1区域被置换为表达盒,该表达盒含有由CMV启动子驱动的编码截短的Core-Pol*的合成序列。通过将经PacI线性化的病毒基因组转染至E1补充细胞系而获得腺病毒颗粒(AdTG18202)。

[0214] 1.1.3.表达Pol*的质粒和腺病毒载体的构建和产生

[0215] 编码截短的Pol突变体多肽的合成基因(2379个核苷酸,如SEQ ID NO:13中所示)由GENEART(Regensburg, Germany)合成。将该合成的片段插入至pTG13135穿梭质粒的NheI和NotI限制性酶切位点中,提供了TG18195。继而通过在经Bst1107I和PacI消化的pTG18195与经ClaI消化线性化了的pTG15375之间的同源重组而获得了腺病毒载体。所得的腺病毒载体pTG18203是E3和E1缺失的,其中E1区域被置换为表达盒,该表达盒含有由CMV启动子驱动的编码Pol*的合成序列。通过将经PacI线性化的病毒基因组转染至E1补充细胞系而获得腺病毒颗粒(AdTG18203)。

[0216] 1.2.小鼠模型中的免疫原性评估

[0217] 通过在HLA转基因小鼠的免疫之后的Elispot IFN γ 测定和细胞内细胞因子染色(ICS)而对抗原免疫原性进行了体内评估。

[0218] 1.2.1.小鼠模型

[0219] 研究中使用的HLA-A2.1转基因小鼠由Pascolo等描述(1997, J. Exp. Med. 185:

2043)。这些小鼠的H-2D^b和鼠β₂-微球蛋白基因被敲除,并且表达转基因单链组织相容性I类分子(HHD分子),其中人β₂m的C-端共价连接至嵌合重链的N-端(HLA-A*0201a1-a2,H-2D^ba3跨膜和胞质内结构域)。免疫了7-10周龄的小鼠(雄性和雌性)。小鼠的平均重量约25-30g。

[0220] 研究中使用的HBV转基因小鼠由Halverscheid等描述(2008,J.Med.Virol.80:583-590)并承蒙Reinhold Schirmbeck惠赠。这些小鼠有C57Bl/6J遗传背景并且转基因了HBV基因组(1.4拷贝的HBV基因组,在位置1438有突变(T至C)其避免了HBsAg蛋白的小形式的表达并且抑制了HBV感染性颗粒的形成)。免疫了10-16周龄的小鼠(雄性和雌性)。小鼠的平均重量约25-30g。

[0221] 1.2.2.免疫方案

[0222] 1.2.2.1DNA免疫方案

[0223] 进行DNA免疫方案以便评估实施例1.1中所述质粒编码的不同融合蛋白的免疫原性。用于免疫的DNA在无内毒素的条件下产生。以15天的间隔用100μg/注射的每种测试质粒免疫小鼠两次,所述注射是经肌内途径注射至胫骨前肌。在第一次DNA注射前继续拧心脏毒素注射以有利于DNA免疫原性。在最后一次DNA注射的15天后评估了细胞免疫应答。

[0224] 1.2.2.2腺病毒免疫方案

[0225] 进行了腺病毒免疫方案以便比较有Ad载体(如实施例1.1中所述产生)编码的不同融合蛋白的免疫原性。用编码不同融合蛋白的腺病毒免疫小鼠一次(10⁸iu/小鼠/注射),其中所述注射经皮下途径注射至尾根。在最后一次腺病毒注射的15天后评估了细胞免疫应答。

[0226] 也用AdTG18201评估了不同剂量的腺病毒。用10⁵、10⁶、10⁷、10⁸或10⁹iu的AdTG18201通过皮下途径于尾根处免疫小鼠一次。

[0227] 也用AdTG18202测试了不同的免疫日程,以不同的时间间隔对小鼠注射了一次、两次、三次或六次。每次注射均使用10⁸iu/小鼠经皮下途径于尾根处进行。对时间间隔1周的一次、三次或六次注射一起进行了比较。也对如下日程一起进行了比较:检测诱导的T细胞应答之前2周或20周注射一次、2injections at以2或4个月的间隔注射2次、和以2个月的间隔注射3次。

[0228] 1.2.3肽

[0229] 用于体外细胞刺激的肽是9-10个氨基酸的短肽(被描述或预测为 HLA-A2限制性表位)或15个氨基酸的长肽(包括在涵盖所有感兴趣抗原的肽文库中)。

[0230] 对应于所描述或预测为核心蛋白、Pol蛋白或Env结构域的HLA-A2限制性表位的短肽由Eurogentec(Belgium)合成并以10mM的浓度溶于100%DMSO(sigma,D2650)。

[0231] 涵盖全长核心、Pol和包膜结构域的肽文库由ProImmune(Oxford,United Kingdom)合成。该核心、Pol和Env文库由15聚体的肽(重叠11个氨基酸)组成。将每种粗肽以50mM的浓度溶于100%DMSO(sigma,D2650)。对每个文库,将肽汇集成每种肽2mg/ml的浓度:

[0232] -HBV核心蛋白被涵盖为21和22个肽的2个汇集(汇集1(PC1):22个肽涵盖了核心的残基1-100;汇集2(PC2):21个肽涵盖了核心的残基89-183);

[0233] -HBV Pol蛋白被涵盖为24个肽的8个汇集(汇集1(PP1):24个肽涵盖了aa45-151;汇集2(PP2):24个肽涵盖了aa141-251(从aa205-219的肽被排除因为在100%DMSO或DMSO+Tris100mM pH9中不可溶;从aa221-235的肽被溶于DMSO+Tris100mM pH9因为在100%DMSO中

不可溶);汇集3(PP3):24个肽涵盖了aa241-347;汇集4(PP4):24个肽涵盖了aa337-447(从373-387的肽被排除因为在100%DMSO或DMSO+Tris100mM pH9中不可溶);Poo15(PP5):24个肽涵盖了aa437-543;汇集6(PP6):24个肽涵盖了aa533-639;汇集7(PP7):24个肽涵盖了aa629-735;汇集8(PP8):24个肽涵盖了aa725-835);

[0234] -Env结构域被涵盖为9个和10个肽的2个汇集(汇集1(PE1):10个肽涵盖了HBs残基9-59;汇集2(PE2):9个肽涵盖了HBs残基157to194)。

[0235] 对于在具有C57BL/6J遗传背景的HBV转基因小鼠中进行的实验,文献中描述的或此前实验中鉴别的在具有C57BL/6J遗传背景的小鼠中具有活性的HBV肽被用于体外细胞刺激。它们是短肽(VSAAFYHLPL,聚合酶;SEQ ID NO:24)或长肽(聚合酶的NLNVSIPTWTHKVGNF称作N13F(SEQ ID NO:25),和包膜蛋白的FLWEWASARFSWLSL称作F13L(SEQ ID NO:26))。它们由Eurogentec(Belgium)或ProImmune(Oxford,United Kingdom)合成。每种肽以10mM的浓度被溶于100%DMSO(sigma D2650)。ICS测定期间以10mM的浓度使用(即便是测试两种肽的混合物的时候)。

[0236] 1.2.4. IFNg Elispot测定

[0237] 收集了来自经免疫小鼠的脾细胞,并裂解了红血细胞(Sigma,R7757)。在Multiscreen平板(Millipore,MSHA S4510)中每孔 2×10^5 细胞一式三份培养40h,所述平板包被有抗-小鼠IFN单克隆抗体(BD Biosciences;10 μ g/ml,551216),培养是在MEM培养基(Gibco,22571)中进行,其中补加了10%FCS(JRH,12003-100M)、80U/mL青霉素/80 μ g/mL链霉素(PAN,P06-07-100)、2mM L-谷氨酰胺(Gibco,25030)、1x非必需氨基酸(Gibco,11140)、10mM Hepes(Gibco,15630)、1mM丙酮酸钠(Gibco,31350)和50 μ M β -巯基乙醇(Gibco,31350),并且其中存在10单位/ml的重组鼠IL2(Peprotech,212-12),单独存在则作为阴性对照,或者还有with:

[0238] -10 μ M的存在于HBV抗原中的选择的HLA-A2限制性肽,由SEQ ID NO:18-23所示的质粒编码(FLP、ILC用于核心,VLQ、FLG和GLS用于Env以及SLY用于Po1)或者不相关的肽;

[0239] -每种肽终浓度5 μ g/ml的肽的汇集

[0240] -5 μ g/ml的Concanavalin A(Sigma,C5275)作为阳性对照。

[0241] 通过如前所述(Himoudi等,2002,J.Virol.76:12735)的Elispot(细胞因子-特异性酶联免疫斑点)测定来对产IFNg的T细胞进行定量。将阴性对照孔中斑点的数目(对应于产IFNg的T细胞)从含有HBV肽的实验孔所检测的斑点数目中减去。结果显示为得自一式三份孔的平均值。所观测应答的阳性的实验阈值(或截断值)是通过计算这样的阈值确定的:其对应于仅有培养基所观测的斑点的平均值+2个标准偏差,报道为 10^6 细胞。与CTLElispot读取器关联的技术截断值也被定义为50斑点/ 10^6 细胞(在此值之上则读取器的CV(差异系数)系统性地低于20%)。The highest value between the 技术截断值和每个实验计算的实验阈值之间的最高值被用于定义每个实验的截断值。通过使用Mann-Whitney检验来进行Elispot应答的统计学分析。P值等于或小于0.05则被认为是显著的。

[0242] 1.2.5. 细胞内细胞因子染色(ICS)测定

[0243] 在来自每个组的每个动物的脾细胞上进行ICS。在用裂解缓冲剂(Sigma,R7757)裂解红血细胞之后,将平底96-孔平板中的每孔 2×10^6 细胞在完全的MEM培养基(Gibco BRL,22571)中温育,其中存在10单位/ml的鼠重组IL-2(Peprotech,212-12),单独存在则作为阴

性对照,或还有10 μ M的特异性HBV肽或还有每种肽终浓度5 μ g/ml的肽汇集或还有10 μ M的不相关肽。立即以1 μ l/ml的终浓度加入GolgiPlug(BD Biosciences,555029)5h。接着,在V-底96-孔平板中收集细胞并用1%FCS-PBS洗涤。使用针对CD3的单克隆抗体(仓鼠MAb抗-CD3e-PE,稀释1/200)、针对CD8的单克隆抗体(大鼠MAb抗-CD8a-APC,稀释1/600)和针对CD4的单克隆抗体(大鼠MAb抗-CD4-PerCP,稀释1/600)来进行染色(所有抗体均来自BDBiosciences,分别为553063、553035和553052),染色是在50 μ l的1%FCS-PBS中于室温进行15min。在洗涤后,用Cytotfix/Cytoperm将细胞固定并渗透(permeabilize)以及用Perm/Wash溶液(BD Biosciences,554714)洗涤。接着,早室温加入抗-小鼠IFN γ -PE抗体(BD Biosciences,554412557724)和抗-小鼠TNF α -Alexa488抗体(BD Biosciences,557719)或抗-小鼠IFN γ -PE抗体(BD Biosciences,554412557724)15min,并在用Perm/Wash洗涤后,将细胞在1%FCS-PBS中重悬并使用FacsCalibur(Becton Dickinson)通过流式细胞术来分析。对CD3e $^{+}$ 、CD8a $^{+}$ 细胞或CD3e $^{+}$ 、CD4 $^{+}$ 细胞进行门控以确定IFN γ +CD8 $^{+}$ 或IFN γ +CD4 $^{+}$ T或TNF α +CD8 $^{+}$ 或TNF α +CD4 $^{+}$ T或IFN γ +TNF α +CD8 $^{+}$ 或IFN γ +TNF α +CD4 $^{+}$ T细胞群体的百分比。从仅有培养基所获得的百分比被认为是背景。

[0244] 对于在HBV转基因小鼠中进行的实验,也在每个组的每只动物的肝上进行ICS。在处死小鼠后,通过肝门静脉向肝原位灌注冷的PBS直至该器官变苍白。收集该肝脏,将其置于PBS+FCS2%溶液中,切成小片,轻柔地过压70 μ m的细胞滤网并接着悬于冷的PBS+2%FCS溶液中。离心后,用冷的PBS+2%FCS溶液在此洗涤细胞。新的离心之后,将含细胞的沉淀重悬于10mL的Percoll溶液,700g室温离心12分钟,并再次用PBS+2%FCS溶液洗涤。继而如前所述的脾细胞那样将红血细胞裂解,所有后续步骤均按照前文脾细胞的部分所述的那样进行。需要注意,对于肝脏,由于所收集的T淋巴细胞的量是有限的,每孔的细胞数目是变化的:将所获得的所有细胞以这样的方式培养:所有孔中分派等量的细胞。

[0245] 1.2.6体内CTL测定

[0246] 如Fournillier等(2007,Vaccine,25:7339-53)所述在HLA-A2转基因小鼠中进行体内CTL测定。脾细胞悬液得自同基因型的小鼠脾脏并在红血细胞的裂解之后被调节至20 $\times 10^6$ 细胞/mL。将一半细胞与感兴趣的HBV肽(SLY,FLP或ILC)以10 μ M的终浓度在37 $^{\circ}$ C温育1h,而另一半细胞则未处理。接着将5(6)-羧基荧光素双乙酸琥珀酰亚胺酯(CFSE)(Molecular Probes,C1157)加至细胞,其中10 μ M(CFSE-高)加至未处理的细胞而1 μ M(CFSE-低)被加至HBV-肽脉冲的细胞10min。在用PBS洗涤后,将所有群体混合并将20 $\times 10^6$ 的总细胞经眼眶后静脉窦注射至麻醉的小鼠,小鼠之前曾经由AdTG18201或AdTG15149免疫(2周前)。因而,CFSE-低群体代表了特异性的靶标(其应当可通过接种而被细胞毒T细胞所裂解),而CFSE-高群体是内部参照,使得测定能够归一化。来自受体小鼠的脾细胞在24h后经流式细胞术分析,以检测CFSE-标记的细胞。对每只动物,计算了肽-脉冲的靶标和未经处理的靶标之间的比例($R = \text{CFSE-低细胞的数目} / \text{CFSE-高细胞的数目}$)。每只动物的特异性裂解的百分比由下式确定: $\% \text{ lysis} = (1 - R_{\text{小鼠}} / R_{\text{参照}}) \times 100$ 其中 $R_{\text{参照}}$ 是得自2只未经其它处理的(naïve)HLA-A2小鼠的R的平均,该两只小鼠也注射了CFSE-标记靶标的相同悬液。如果特异性裂解的百分比高于10%,则认为应答是阳性的。

[0247] 1.3通过电子显微镜体外分析AdTG18201

[0248] 用AdTG18201以不同的MOI(25-100)在悬液中且在降低的培养基体积条件下感染

A549细胞(人类肺腺癌上皮细胞系),并继而在收集分析前培养16H、24H或48H。在这些不同的时间点收集细胞并继而用戊二醛(2%稀释在0.2M二甲胂酸钠缓冲剂中)固定。继而干燥细胞,将其包括在树脂中封闭并继而切成超薄切片。继而用乙酸铀酰和柠檬酸铅将所获得的栅格(grid)染色并通过电子显微镜观测。

[0249] 2.结果

[0250] 2.1.DNA质粒pTG18188和pTG18194表达的HBV融合蛋白的免疫原性

[0251] 在HLA-A2转基因小鼠中评估了由DNA质粒表达的HBV融合蛋白的免疫原性。在两次肌肉注射pTG18188(tCore-Pol*-Env1-Env2)或pTG18194(tCore-Pol*)或pTG13135作为阴性对照(空质粒)之后,通过Elispot IFN γ 和ICS(IFN γ /TNF α)来评估特异性T细胞应答,其中使用已知的HLA-A2表位,所述表位存在于聚合酶、核心或包膜结构域和/或涵盖感兴趣HBV抗原的重叠肽的汇集。

[0252] 2.1.1.通过Elispot测定评估HBV特异性产IFN γ 细胞

[0253] 如图2所示,用编码HBV融合蛋白“tCore-Pol*”的质粒pTG18194免疫诱导了特异于HLA-A2限制性SLY表位(SEQ ID NO:23,位于HBV聚合酶(位置816-824)内)的产IFN γ 细胞。用质粒pTG18194免疫还导致了诱导高频率的产IFN γ 细胞,所述细胞特异于2个核心HLA-A2限制性表位FLP(SEQ ID NO:18,位于HBV核心蛋白的位置18-27)和ILC(SEQ ID NO:19,位于HBV核心蛋白的位置99-108)。在8只经接种小鼠中的4只观测到阳性应答。

[0254] 如图2所示,编码HBV融合蛋白“Core-Pol*-Env1-Env2”的质粒pTG18188也诱导产IFN γ 细胞,所述细胞特异于pol HLA-A2表位SLY和两个2核心HLA-A2限制性表位FLP和ILC。在8只经接种小鼠中的4只观测到阳性应答。另外,也检测到了特异于HLA-A2GLS表位(SEQ ID NO:22,位于HBsAg的Env2位置185-194)的产IFN γ 细胞,虽然是低频率且在一只经接种小鼠中。

[0255] 2.1.2.通过细胞内染色测定来评估诱导的HBV特异性产IFN γ T CD8+和CD4+细胞

[0256] 2.1.2.1.特异于HLA-A2限制性表位的CD8T细胞应答

[0257] 通过ICS测定评估了仅产生IFN γ 或者组合产生IFN γ 与TNF α (靶向HLA-A2限制性表位)的CD8T细胞的百分比,所述表位包括在聚合酶(SLY)、核心(FLP和ILC)和包膜结构域(VLQ、FLG和GLS)之内。结果如图3A所示为特异于这些表位的CD8+T细胞和产IFN γ 的CD8+T细胞(单独产IFN γ 细胞或同时产IFN γ 和TNF α 的细胞的总和)的百分比。用pTG18194(表达tCore-Pol*)免疫的8只动物中的4只携带有产IFN γ 的CD8+T细胞,其特异于分别位于核心和Pol抗原中的FLP、ILC和SLY HLA-A2限制性表位。类似地,用pTG18188(表达tCore-Pol*-Env1-Env2)免疫的8只动物中的4只也携带有产IFN γ 的T CD8+T细胞,其特异于FLP、ILC和SLY表位。另外,如ELISPOT测定中已经观测到的那样,用质粒pTG18188免疫的8只动物中的4只显示了由产IFN γ 的CD8+T细胞介导的特异于GLSHLA-A2限制性表位的应答,所述限制性表位Env2结构域内。用pTG13135免疫并未如预期那样诱导任何特异性的应答。

[0258] 2.1.2.2特异于涵盖了核心蛋白、聚合酶蛋白和Env结构域的肽的汇集的CD8和CD4T细胞应答。

[0259] 特异于涵盖了核心蛋白的肽的汇集的应答

[0260] 通过ICS测定评估了能够单独产生IFN γ 或组合产生IFN γ 和TNF α (应答于涵盖了核心蛋白(PC)的肽的汇集)的CD8和CD4T细胞的百分比。结果表示为特异于这些肽的汇集的

CD8+或CD4+T细胞和产生IFN γ 的CD8+或CD4+T细胞(单独产IFN γ 细胞或同时产IFN γ 和TNF α 的细胞的总和)的百分比。

[0261] 如图3B所示,针对涵盖核心蛋白的肽的2个汇集(PC1和PC2),检测到了产IFN γ 的CD8+T细胞的阳性百分比,其中CD8+T细胞应答主要集中于汇集核心1的肽。用pTG18194或pTG18188接种的小鼠中观测到的反应性CD8+T细胞的百分比与用阴性对照(pTG13135)接种的小鼠中观测到的百分比有显著差异($p < 0.05$, Mann Withney检验),对于两个肽汇集(1和2)都是如此。

[0262] 如图3C所示,针对涵盖核心蛋白的肽的1个汇集,也检测到了产IFN γ 的CD4+T细胞的阳性百分比,即用pTG18194或pTG18188接种的两组小鼠中的汇集核心2。用pTG18188接种的小鼠中观测到的反应性CD4+T细胞的百分比与用阴性对照(pTG13135)接种的小鼠中观测到的百分比有显著差异($p < 0.05$, Mann Withney检验),这是对于汇集核心2而言。

[0263] 特异于涵盖了聚合酶蛋白的肽的汇集的应答

[0264] 通过ICS测定评估了能够单独产生IFN γ 或组合产生IFN γ 和TNF α (应答于涵盖了聚合酶蛋白的肽的汇集)的CD8和CD4T细胞的百分比。结果表示为特异于这些肽的汇集的CD8+或CD4+T细胞和产生IFN γ 的CD8+或CD4+T细胞(单独产IFN γ 细胞或同时产IFN γ 和TNF α 的细胞的总和)的百分比。

[0265] 如图3D所示,主要针对肽的一个汇集PP8检测到了产IFN γ 的CD8+T细胞的阳性百分比。具体到PP8而言,用pTG18194或pTG18188接种的小鼠中观测到的反应性CD8+T细胞的百分比与用阴性对照(pTG13135)接种的小鼠中观测到的百分比有显著差异($p < 0.05$, Mann Withney检验)。需要注意的是,用pTG18188接种的组中的一只小鼠也显示了针对汇集4、汇集5和汇集6的产IFN γ 的CD8+T细胞的阳性百分比。

[0266] 如图3E所示,针对涵盖Po1蛋白的肽的4个汇集,检测到了产IFN γ 的CD4+T细胞的弱的但却是阳性的百分比,也即在用pTG18194或pTG18188接种的两组小鼠中的汇集Po11、汇集Po14、汇集Po15和汇集Po16,每组8只经测试小鼠中的至少3只显示出了应答。

[0267] 特异于涵盖了包膜结构域的肽的汇集的应答

[0268] 通过ICS测定评估了能够单独产生IFN γ 或组合产生IFN γ 和TNF α (应答于涵盖了包膜结构域Env1和Env2的肽的汇集)的CD8和CD4T细胞的百分比。在此实验中未检测到特异性的CD4+T细胞应答。CD8+T细胞应答的结果示于图3F,表示为特异于这些肽的汇集的CD8+T细胞和产生IFN γ 的CD8+T细胞(单独产IFN γ 细胞或同时产IFN γ 和TNF α 的细胞的总和)的百分比。具体而言,在用pTG18188接种的1只小鼠中针对肽的一个汇集(汇集Env2)检测到了弱的但却是阳性的产IFN γ 的CD8+T细胞的百分比。

[0269] 2.2.腺病毒AdTG18201、AdTG18202和AdTG18203表达的HBV融合蛋白的免疫原性

[0270] 2.2.1.通过使用重叠肽的汇集的Elispots IFN γ 来对HBV特异性产IFN γ 的T细胞进行评估

[0271] 在用AdTG18201或AdTG18202或AdTG18203或AdTG15149(空腺病毒用作阴性对照)免疫的HLA-A2转基因小鼠中评价了由人腺病毒5表达的HBV Po1突变体和融合蛋白的免疫原性。通过Elispot IFN γ 评估了在一次皮下注射腺病毒之后诱导的特异性T细胞应答,其中使用涵盖了HBV感兴趣抗原核心(PC1-2)、聚合酶(PP1-8)和Env(PE1-2)结构域的重叠肽的汇集。

[0272] 如图4中所示,编码HBV突变体聚合酶多肽的AdTG18203自己能够诱导特异于聚合酶肽汇集4、5、6和8的产IFN γ 细胞。所有免疫的小鼠显示出特异性T细胞应答,其中高频率的产IFN γ 细胞主要针对聚合酶肽汇集4和8。

[0273] 如图4所示,编码HBV融合蛋白“tCore-Pol*”的AdTG18202诱导了特异于肽汇集PP2、PP3、PP4、PP5和PP8的产IFN γ 细胞,该聚合酶-特异性应答主要集中于针对PP2、PP3和PP8。用AdTG18202免疫还导致了诱导高频率的产IFN γ 细胞,所述细胞特应于2个核心肽汇集PC1和PC2,其中靶向PC1的T细胞有更高的频率。靶向聚合酶和核心抗原的阳性应答在5只经免疫小鼠中的5只均有观测到。

[0274] 编码HBV融合蛋白“Core-Pol*-Env1-Env2”的AdTG18201也被发现了免疫原性,如图4所示。更具体地,在所有接种的小鼠中都诱导了特异于聚合酶肽汇集PP2、PP3和PP8的产IFN γ 细胞,以及针对PP4和PP5也是如此虽然其斑点较弱且应答小鼠的频率较低。用AdTG18201免疫还导致了诱导高频率的产IFN γ 细胞,所述细胞特应于2个核心肽汇集PC1和PC2,这是在所有免疫的小鼠中(5/5)。用AdTG18201免疫还诱导了针对Env结构域的特异性T细胞应答,虽然那些应答有些弱且零散,其中5只小鼠中的1只显示靶向PE1的应答,以及其中5只小鼠中的2只显示靶向PE2的应答。

[0275] 2.2.2.在用AdTG18201免疫后通过Elispots IFN γ 使用HLA-A2肽来评估HBV特异性的产IFN γ T细胞

[0276] 在HLA-A2转基因小鼠中评估了由AdTG18201表达的HBV融合蛋白的免疫原性。通过一次皮下注射AdTG18201或AdTG15149(空腺病毒用作阴性对照)免疫动物。通过Elispot IFN γ 评估了特异性T细胞应答,其中使用包含在聚合酶(SLY)、核心(FLP和ILC)和包膜(VLQ和GLS)的HLA-A2 限制性表位。

[0277] 如图5所示,编码HBV融合蛋白《Core-Pol*-Env1-Env2》的AdTG18201被发现有免疫原性。更具体地,在所有AdTG18201-接种的小鼠中,均诱导了特异于聚合酶HLA-A2表位(SLY)的产IFN γ 细胞。同时,AdTG18201也诱导了特异于2个核心蛋白的HLA-A2表位(FLP和ILC)的产IFN γ 细胞,且具有高频率。用AdTG18201免疫还诱导了针对Env结构域的特异性T细胞应答,虽然频率和应答小鼠的数目较低,其中8只测试小鼠中的2只显示出针对VLQ肽的阳性T细胞应答,以及8只测试小鼠中的5只显示出针对GLSP肽的阳性T细胞应答。

[0278] 2.2.3.在由AdTG18201并使用选择的肽的汇集免疫HLA-A2小鼠之后通过细胞内染色测定评估HBV特异性产IFN γ 和/或TNF α 的CD8+T细胞

[0279] 通过ICS测定评估了能够单独产生IFN γ 或组合产生IFN γ 和TNF α 的CD8+T细胞的百分比,其中所述细胞是应对于选择肽汇集,所述汇集涵盖了部分的聚合酶蛋白(PP8,氨基酸725-835)和部分的HBV核心蛋白(PC1,氨基酸1-100)。结果表示为特异于这些肽汇集和单独产生IFN γ 及组合产生IFN γ 和TNF α 的CD8+T细胞的百分比。

[0280] 如图6所述,AdTG18201能够特异性地诱导高百分比的单独产IFN γ 的CD8+T细胞以及组合产IFN γ 和TNF α 的CD8+T细胞(识别PP8和PC1汇集的肽)。所有接种的小鼠显示出高百分比的单一产生(仅IFN γ)和双产生(IFN γ 和TNF α)的特异性CD8+T细胞。

[0281] 需要注意的是,在另一种小鼠模型C57B16小鼠中进行的类似的实验,显示了AdTG18201免疫原性的类似结果(未示出)。

[0282] 2.2.4.在用AdTG18201免疫HLA-A2小鼠之后用体内CTL测定评估体内功能性CD8+T

细胞的诱导

[0283] 在用AdTG18201或AdTG15149(作为阳性对照)免疫后,通过体内细胞溶解(或CTL)测定在HLA-A2小鼠中评估了AdTG18201诱导体内功能性CD8T细胞(显示细胞溶解活性)的能力,并且其中使用3个已知被产IFN γ 的CD8+T细胞所靶向HLA-A2的表位(SLY,FL和ILC)。

[0284] 如图7所示,AdTG18201能够诱导高百分比的针对聚合酶表位(SLY)的体内特异性裂解,其中在所有经免疫的小鼠中均检测到裂解并且百分比范围在42%–75%(图7a)。也显示出AdTG18201能够诱导高百分比的针对核心蛋白两个测试的HLA-A2表位FLP(图7a)和ILC(图7b)的体内特异性裂解,其中百分比范围分别在32%–69%和3%–64%。针对env表位的应答可以检测到但却是低水平。

[0285] 这些数据清楚地证明了AdTG18201诱导体内功能性CD8+T细胞(其显示细胞溶解活性并靶向HBV聚合酶和HBV核心蛋白)的能力。

[0286] 2.2.5. 在用AdTG18201免疫后并使用ICS测定来评估功能性CD8+T细胞在HBV转基因小鼠中的诱导

[0287] 在HBV转基因小鼠中评估了AdTG18201在耐受小鼠模型中诱导功能性T细胞的能力。实际上,这些小鼠对于HBV基因组是转基因的,并因此对HBV抗原耐受,在一定程度上模拟了HBV慢性患者中遇到的耐受。通过一次皮下注射AdTG18201(10^8 iu)或AdTG15149(作为阴性对照)免疫HBV转基因小鼠。在接种小鼠的脾和肝中均通过ICS监测诱导的T细胞(检测产IFN γ 和TNF α 的CD8+T细胞)。在此特定模型中,经鉴别在C57B1/6J小鼠中具有反应性的肽被用于筛选诱导的T细胞应答:对于聚合酶是VSA和N13F肽的汇集而对于包膜则是F13L肽。

[0288] 如图8所述,在AdTG18201接种的小鼠的脾和肝中检测到了产IFN γ 和TNF α 的功能性CD8+T细胞,其中在两个器官中,5只测试小鼠中的4只显示特异于聚合酶的功能性产IFN γ /TNF α 的CD8+T细胞,以及5只测试小鼠中的2只显示特异于包膜的功能性产IFN γ /TNF α 的CD8+T细胞。如预期的那样,在用空AdTG15149免疫的小鼠中或者当使用不相关肽刺激是,没有检测到应答。

[0289] 综上,这些数据证明了表达融合蛋白(其含有RNaseH缺陷和YMDD-缺失的pol突变体、env结构域和核心)的病毒载体AdTG18201在HBV耐受模型中诱导功能性产IFN γ 和TNF α 的CD8+T细胞的能力。

[0290] 2.3对用AdTG18201或AdTG18202进行不同剂量和日程的免疫进行评估。

[0291] 2.3.1. 腺病毒剂量评估

[0292] 在HLA-A2转基因小鼠以不同的剂量评价了AdTG18201所表达的HBV融合蛋白的免疫原性。通过一次皮下注射AdTG18201(以 10^5 iu或 10^6 iu或 10^7 iu或 10^8 iu或 10^9 iu的剂量)或AdTG15149(以 10^9 iu的剂量,空腺病毒用作阴性对照)来免疫动物。were evaluated by通过Elispot IFN γ 使用HLA-A2限制性表位(包含在聚合酶(SLY)、核心(FLP和ILC)和编码(VLQ和GLS)中)来评估特异性T细胞应答。

[0293] 如图9所示,编码HBV融合蛋白《Core-Pol*-Env1-Env2》的AdTG18201在以 10^7 iu、 10^8 iu和 10^9 iu的剂量注射时被发现具有免疫原性。更具体地,使用 10^5 和 10^6 iu的剂量,未检测到特异于所测试的核心、聚合酶或Env结构域的HLA-A2表位的产IFN γ 细胞。对于剂量 10^7 、 10^8 和 10^9 iu,则检测到了特异性的产IFN γ 细胞,其靶向于2个所测试的核心表位以及所测试的Pol表位。对3个表位(SLY、FLP和ILC)观测到了剂量效应。对于Env结构域的2个表位(VLQ

和GLS),产IFNg细胞的频率在剂量 10^7 和 10^8 低而该频率在 10^9 iu的剂量明显提高。

[0294] 2.3.2. 评估短时间间隔的多次免疫日程

[0295] 在HLA-A2转基因小鼠中根据不同的免疫日程评价由AdTG18202表达的HBV融合蛋白的免疫原性。将AdTG18202施用一次或三次(三周期间1注射/周)或六次(六周期间1注射/周),并在最后一次注射2周之后通过Elispots IFNg测定评价所诱导的免疫T细胞应答,并且其中使用HLA-A2限制性表位,SLY(Po1)以及FLP和ILC(核心)。一些小鼠用空腺病毒免疫六次作为阴性对照(未显示)。

[0296] 如图10所示,无论使用何种测试日程,编码融合蛋白"Core-Po1*"的AdTG18202均被发现具有免疫原性。更具体地,仅使用培养基和不相关蛋白则未检测到特异性T细胞应答,而在存在3种测试的HBV表位SLY、FLP和ILC时,则检测到高且类似频率的产IFNg细胞。所检测到的产IFNg T细胞的频率在以1周的间隔注射一次、三次或六次的组之间表现相当,而在IFNg产生水平方面并未表现出T细胞耗竭,这是因为在短的时间间隔中进行了很高数目的免疫免疫。当小鼠被注射六次时其腺病毒特异性T细胞应答表现地高于注射一次或三次的时候。如所预期的那样,在用空腺病毒免疫后未检测到HBV-特异性T细胞应答。

[0297] 2.3.3. 评估长时间间隔的多次免疫日程

[0298] 在HLA-A2转基因小鼠中根据不同的免疫日程评价由AdTG18202表达的HBV融合蛋白的免疫原性。

[0299] 将AdTG18202施用一次(在监测T细胞应答之前2周(组1)或20周(组2))或两次(以2个月(组3)或4个月(组4)的间隔注射两次,最后一次注射的2周后监测T细胞应答)或三次(以2个月的间隔(组5),最后一次注射的2周后监测T细胞应答)。通过Elispots IFNg测定并使用HLA-A2限制性表位SLY(Po1)以及FLP和ILC(核心)来监测所诱导的T细胞应答。一些小鼠用空腺病毒免疫一次或以2个月的间隔免疫3次作为阴性对照(未显示)。

[0300] 如图11所示,无论使用何种测试日程,编码融合蛋白"Core-Po1*"的AdTG18202均被发现具有免疫原性,而在以AdTG18202免疫后在仅存在培养基或使用不相关肽时则未检测到特异性T细胞应答。更具体地,组2中所观测到的特异性T细胞应答显示,尽管其低于组1中一次免疫之后2周所观测到的,但是一次注射AdTG18202之后的T细胞应答在注射20周之后仍存在。在组3和4中所观测到的T细胞应答显示,第一次免疫之后2或4个月进行的第二次免疫能够唤起特异于HBV表位的T细胞应答,且水平至少达到组1中观测到的初次免疫应答,对于SLY表位而言甚至稍微更高。在以2个月的间隔免疫了三次的小鼠(组5)中观测到了类似的结果,通过第二次和第三次注射将所诱导的T细胞应答唤起至类似于组1中所观测的水平。A如所预期的那样,在用空腺病毒免疫后未检测到HBV-特异性T细胞应答。

[0301] 2.4 电子显微镜观察

[0302] 用AdTG18201在MOI25、50或100体外感染A54细胞,并在感染后16h、24h或48h收集细胞。继而处理所收集的细胞以通过电子显微镜观察。

[0303] 在AdTG18201感染的细胞的核和细胞质中观察到一些病毒样颗粒(VLP),而在空腺病毒感染的细胞中没有观察到这些结构。这些VLP主要位于核内。在一些细胞中观察到了蛋白团聚体(aggregate)和VLP。

序列表

<110> 特兰斯吉恩股份有限公司
P·马丁
J-B·马钱德
N·西尔韦斯特雷

<120> HBV 聚合酶突变体

<130> 361194D31275

<150> EP11305909.1
<151> 2011-07-12

<150> EP12305450.4
<151> 2012-04-18

<160> 26

<170> PatentIn version 3.3

<210> 1
<211> 344
<212> PRT
<213> Hepatitis B virus

<220>
<221> MISC_FEATURE
<222> (7)..(7)
<223> Xaa (7) can be Thr (T) or Ala (A)

[0001] <220>
<221> MISC_FEATURE
<222> (13)..(13)
<223> Xaa (13) can be Asn (N), Arg (R) or His (H)

<220>
<221> MISC_FEATURE
<222> (16)..(16)
<223> Xaa (16) can be Ile (I) or Thr (T)

<220>
<221> MISC_FEATURE
<222> (38)..(38)
<223> Xaa (38) can be Thr (T) or Ala (A)

<220>
<221> MISC_FEATURE
<222> (53)..(53)
<223> Xaa (53) can be Ser (S) or Asn (N)

<220>
<221> MISC_FEATURE
<222> (54)..(54)
<223> Xaa (54) can be Thr (T) or Tyr (Y)

<220>
<221> MISC_FEATURE
<222> (55)..(55)
<223> Xaa (55) can be His (H) or Arg (R)

<220>
<221> MISC_FEATURE
<222> (91)..(91)
<223> Xaa (91) can be Ile (I) or Leu (L)

<220>
<221> MISC_FEATURE

<222> (109)..(109)
<223> Xaa (109) can be Pro (P) or Ser (S)

<220>
<221> MISC_FEATURE
<222> (118)..(118)
<223> Xaa (118) can be Thr (T) or Asn (N)

<220>
<221> MISC_FEATURE
<222> (121)..(121)
<223> Xaa (121) can be Asn (N) or Ile (I)

<220>
<221> MISC_FEATURE
<222> (122)..(122)
<223> Xaa (122) can be Ile (I) or Phe (F)

<220>
<221> MISC_FEATURE
<222> (124)..(124)
<223> Xaa (124) can be Tyr (Y) or Asn (N)

<220>
<221> MISC_FEATURE
<222> (127)..(127)
<223> Xaa (127) can be Gly (G) or Arg (R)

<220>
<221> MISC_FEATURE
<222> (131)..(131)
<223> Xaa (131) can be Asp (D) or Asn (N)

[0002] <220>
<221> MISC_FEATURE
<222> (134)..(134)
<223> Xaa (134) can be Asp (D) or Asn (N)

<220>
<221> MISC_FEATURE
<222> (145)..(145)
<223> Xaa (145) can be Leu (L) or Met (M)

<220>
<221> MISC_FEATURE
<222> (149)..(149)
<223> Xaa (149) can be Lys (K) or Gln (Q)

<220>
<221> MISC_FEATURE
<222> (151)..(151)
<223> Xaa (151) can be Phe (F) or Tyr (Y)

<220>
<221> MISC_FEATURE
<222> (221)..(221)
<223> Xaa (221) can be Phe (F) or Tyr (Y)

<220>
<221> MISC_FEATURE
<222> (222)..(222)
<223> Xaa (222) can be Thr (T) or Ala (A)

<220>
<221> MISC_FEATURE
<222> (223)..(223)
<223> Xaa (223) can be Ser (S) or Ala (A)

<220>

[0003]

```

<221> MISC FEATURE
<222> (224)..(224)
<223> Xaa (224) can be Ile (I) or Val (V)

<220>
<221> MISC FEATURE
<222> (238)..(238)
<223> Xaa (238) can be Asn (N) or His (H)

<220>
<221> MISC FEATURE
<222> (248)..(248)
<223> Xaa (248) can be Asn (N) or His (H)

<220>
<221> MISC FEATURE
<222> (256)..(256)
<223> Xaa (256) can be Ser (S) or Cys (C)

<220>
<221> MISC FEATURE
<222> (257)..(257)
<223> Xaa (257) can be Trp (W) or Tyr (Y)

<220>
<221> MISC FEATURE
<222> (259)..(259)
<223> Xaa (259) can be Thr (T) or Ser (S)

<220>
<221> MISC FEATURE
<222> (263)..(263)
<223> Xaa (263) can be Glu (E) or Asp (D)

<220>
<221> MISC FEATURE
<222> (266)..(266)
<223> Xaa (266) can be Val (V) or Ile (I)

<220>
<221> MISC FEATURE
<222> (267)..(267)
<223> Xaa (267) can be Leu (L) or Gln (Q)

<220>
<221> MISC FEATURE
<222> (271)..(271)
<223> Xaa (271) can be Gln (Q), Met (M) or Glu (E)

<220>
<221> MISC FEATURE
<222> (317)..(317)
<223> Xaa (317) can be Ser (S) or Ala (A)

<220>
<221> MISC FEATURE
<222> (317)..(317)
<223> Xaa (317) can be Ser (S) or Ala (A)

<220>
<221> MISC FEATURE
<222> (332)..(332)
<223> Xaa (332) can be Cys (C) or Ser (S)

<400> 1
Glu Asp Trp Gly Pro Cys Xaa Glu His Gly Glu His Xaa Ile Arg Xaa
1          5          10          15

```

	Pro	Arg	Thr	Pro	Ala	Arg	Val	Thr	Gly	Gly	Val	Phe	Leu	Val	Asp	Lys	
				20					25					30			
	Asn	Pro	His	Asn	Thr	Xaa	Glu	Ser	Arg	Leu	Val	Val	Asp	Phe	Ser	Gln	
		35						40					45				
	Phe	Ser	Arg	Gly	Xaa	Xaa	Xaa	Val	Ser	Trp	Pro	Lys	Phe	Ala	Val	Pro	
		50					55					60					
	Asn	Leu	Gln	Ser	Leu	Thr	Asn	Leu	Leu	Ser	Ser	Asn	Leu	Ser	Trp	Leu	
	65					70					75					80	
	Ser	Leu	Asp	Val	Ser	Ala	Ala	Phe	Tyr	His	Xaa	Pro	Leu	His	Pro	Ala	
					85					90					95		
	Ala	Met	Pro	His	Leu	Leu	Val	Gly	Ser	Ser	Gly	Leu	Xaa	Arg	Tyr	Val	
				100					105					110			
	Ala	Arg	Leu	Ser	Ser	Xaa	Ser	Arg	Xaa	Xaa	Asn	Xaa	Gln	His	Xaa	Thr	
		115						120					125				
	Met	Gln	Xaa	Leu	His	Xaa	Ser	Cys	Ser	Arg	Asn	Leu	Tyr	Val	Ser	Leu	
		130					135					140					
[0004]	Xaa	Leu	Leu	Tyr	Xaa	Thr	Xaa	Gly	Arg	Lys	Leu	His	Leu	Tyr	Ser	His	
	145					150					155					160	
	Pro	Ile	Ile	Leu	Gly	Phe	Arg	Lys	Ile	Pro	Met	Gly	Val	Gly	Leu	Ser	
					165					170					175		
	Pro	Phe	Leu	Leu	Ala	Gln	Phe	Thr	Ser	Ala	Ile	Cys	Ser	Val	Val	Arg	
					180					185				190			
	Arg	Ala	Phe	Pro	His	Cys	Leu	Ala	Phe	Ser	Tyr	Met	Asp	Asp	Val	Val	
		195						200					205				
	Leu	Gly	Ala	Lys	Ser	Val	Gln	His	Leu	Glu	Ser	Leu	Xaa	Xaa	Xaa	Xaa	
		210					215					220					
	Thr	Asn	Phe	Leu	Leu	Ser	Leu	Gly	Ile	His	Leu	Asn	Pro	Xaa	Lys	Thr	
	225					230					235					240	
	Lys	Arg	Trp	Gly	Tyr	Ser	Leu	Xaa	Phe	Met	Gly	Tyr	Val	Ile	Gly	Xaa	
					245					250					255		
	Xaa	Gly	Xaa	Leu	Pro	Gln	Xaa	His	Ile	Xaa	Xaa	Lys	Ile	Lys	Xaa	Cys	
				260					265					270			
	Phe	Arg	Lys	Leu	Pro	Val	Asn	Arg	Pro	Ile	Asp	Trp	Lys	Val	Cys	Gln	
		275						280					285				

Arg Ile Val Gly Leu Leu Gly Phe Ala Ala Pro Phe Thr Gln Cys Gly
290 295 300

Tyr Pro Ala Leu Met Pro Leu Tyr Ala Cys Ile Gln Xaa Lys Gln Ala
305 310 315 320

Phe Thr Phe Ser Pro Thr Tyr Lys Ala Phe Leu Xaa Lys Gln Tyr Leu
325 330 335

Asn Leu Tyr Pro Val Ala Arg Gln
340

<210> 2

<211> 337

<212> PRT

<213> artificial sequence

<220>

<223> mutated Y07587 polymerase domain (del YMDDVVL)

<400> 2

Glu Asp Trp Gly Pro Cys Ala Glu His Gly Glu His His Ile Arg Ile
1 5 10 15

Pro Arg Thr Pro Ala Arg Val Thr Gly Gly Val Phe Leu Val Asp Lys
20 25 30

[0005]

Asn Pro His Asn Thr Ala Glu Ser Arg Leu Val Val Asp Phe Ser Gln
35 40 45

Phe Ser Arg Gly Asn Tyr Arg Val Ser Trp Pro Lys Phe Ala Val Pro
50 55 60

Asn Leu Gln Ser Leu Thr Asn Leu Leu Ser Ser Asn Leu Ser Trp Leu
65 70 75 80

Ser Leu Asp Val Ser Ala Ala Phe Tyr His Leu Pro Leu His Pro Ala
85 90 95

Ala Met Pro His Leu Leu Val Gly Ser Ser Gly Leu Ser Arg Tyr Val
100 105 110

Ala Arg Leu Ser Ser Asn Ser Arg Ile Phe Asn Tyr Gln His Gly Thr
115 120 125

Met Gln Asn Leu His Asp Ser Cys Ser Arg Asn Leu Tyr Val Ser Leu
130 135 140

Leu Leu Leu Tyr Gln Thr Phe Gly Arg Lys Leu His Leu Tyr Ser His
145 150 155 160

Pro Ile Ile Leu Gly Phe Arg Lys Ile Pro Met Gly Val Gly Leu Ser
165 170 175

Pro Phe Leu Leu Ala Gln Phe Thr Ser Ala Ile Cys Ser Val Val Arg
180 185 190

Arg Ala Phe Pro His Cys Leu Ala Phe Ser Gly Ala Lys Ser Val Gln
195 200 205

His Leu Glu Ser Leu Phe Thr Ala Val Thr Asn Phe Leu Leu Ser Leu
210 215 220

Gly Ile His Leu Asn Pro Asn Lys Thr Lys Arg Trp Gly Tyr Ser Leu
225 230 235 240

His Phe Met Gly Tyr Val Ile Gly Cys Tyr Gly Ser Leu Pro Gln Asp
245 250 255

His Ile Ile Gln Lys Ile Lys Glu Cys Phe Arg Lys Leu Pro Val Asn
260 265 270

Arg Pro Ile Asp Trp Lys Val Cys Gln Arg Ile Val Gly Leu Leu Gly
275 280 285

Phe Ala Ala Pro Phe Thr Gln Cys Gly Tyr Pro Ala Leu Met Pro Leu
290 295 300

Tyr Ala Cys Ile Gln Ser Lys Gln Ala Phe Thr Phe Ser Pro Thr Tyr
305 310 315 320

Lys Ala Phe Leu Cys Lys Gln Tyr Leu Asn Leu Tyr Pro Val Ala Arg
325 330 335

[0006]

Gln

<210> 3
<211> 153
<212> PRT
<213> Hepatitis B virus

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> Xaa (2) can be Ser (S) or Pro (P)

<220>
<221> MISC_FEATURE
<222> (19)..(19)
<223> Xaa (19) can be Ala (A) or Val (V)

<220>
<221> MISC_FEATURE
<222> (20)..(20)
<223> Xaa (20) can be Ile (I) or Met (M)

<220>
<221> MISC_FEATURE
<222> (30)..(30)

<223> Xaa (30) can be Val (V) or Leu (L)
 <220>
 <221> MISC_FEATURE
 <222> (31)..(31)
 <223> Xaa (31) can be Ala (A) or Ser (S)
 <220>
 <221> MISC_FEATURE
 <222> (53)..(53)
 <223> Xaa (53) can be Lys (K) or Asn (N)
 <220>
 <221> MISC_FEATURE
 <222> (54)..(54)
 <223> Xaa (54) can be Leu (L) or Ile (I)
 <220>
 <221> MISC_FEATURE
 <222> (55)..(55)
 <223> Xaa (55) can be Leu (L) or Ile (I)
 <220>
 <221> MISC_FEATURE
 <222> (97)..(97)
 <223> Xaa (97) can be Ala (A) or Thr (T)
 <220>
 <221> MISC_FEATURE
 <222> (108)..(108)
 <223> Xaa (108) can be Tyr (Y) or Ser (S)
 [0007] <220>
 <221> MISC_FEATURE
 <222> (115)..(115)
 <223> Xaa (115) can be Pro (P) or Leu (L)
 <220>
 <221> MISC_FEATURE
 <222> (116)..(116)
 <223> Xaa (116) can be Phe (F) or Tyr (Y)
 <220>
 <221> MISC_FEATURE
 <222> (128)..(128)
 <223> Xaa (128) can be Val (V) or Asp (D)
 <400> 3
 Arg Xaa Gly Leu Cys Gln Val Phe Ala Asp Ala Thr Pro Thr Gly Trp
 1 5 10 15
 Gly Leu Xaa Xaa Gly His Gln Arg Met Arg Gly Thr Phe Xaa Xaa Pro
 20 25 30
 Leu Pro Ile His Thr Ala Glu Leu Leu Ala Ala Cys Phe Ala Arg Ser
 35 40 45
 Arg Ser Gly Ala Xaa Xaa Xaa Gly Thr Asp Asn Ser Val Val Leu Ser
 50 55 60
 Arg Lys Tyr Thr Ser Phe Pro Trp Leu Leu Gly Cys Ala Ala Asn Trp
 65 70 75 80

Ile Leu Arg Gly Thr Ser Phe Val Tyr Val Pro Ser Ala Leu Asn Pro
85 90 95

Xaa Asp Asp Pro Ser Arg Gly Arg Leu Gly Leu Xaa Arg Pro Leu Leu
100 105 110

Arg Leu Xaa Xaa Arg Pro Thr Thr Gly Arg Thr Ser Leu Tyr Ala Xaa
115 120 125

Ser Pro Ser Val Pro Ser His Leu Pro Asp Arg Val His Phe Ala Ser
130 135 140

Pro Leu His Val Ala Trp Arg Pro Pro
145 150

<210> 4

<211> 120

<212> PRT

<213> artificial sequence

<220>

<223> mutated Y07587 RNaseH domain (deletion 710-742 ; subsittution D689H V769Y V776Y D777H)

<400> 4

Arg Pro Gly Leu Cys Gln Val Phe Ala His Ala Thr Pro Thr Gly Trp
1 5 10 15

[0008]

Gly Leu Val Met Gly His Gln Arg Met Arg Gly Thr Phe Leu Ser Arg
20 25 30

Lys Tyr Thr Ser Phe Pro Trp Leu Leu Gly Cys Ala Ala Asn Trp Ile
35 40 45

Leu Arg Gly Thr Ser Phe Val Tyr Tyr Pro Ser Ala Leu Asn Pro Tyr
50 55 60

His Asp Pro Ser Arg Gly Arg Leu Gly Leu Ser Arg Pro Leu Leu Arg
65 70 75 80

Leu Pro Phe Arg Pro Thr Thr Gly Arg Thr Ser Leu Tyr Ala Asp Ser
85 90 95

Pro Ser Val Pro Ser His Leu Pro Asp Arg Val His Phe Ala Ser Pro
100 105 110

Leu His Val Ala Trp Arg Pro Pro
115 120

<210> 5

<211> 792

<212> PRT

<213> artificial sequence

<220>

<223> Y07587 polymerase with mutated polymerase and RNaseH domains

<400> 5

Met Pro Leu Ser Tyr Gln His Phe Arg Arg Leu Leu Leu Leu Asp Asp
1 5 10 15

Glu Ala Gly Pro Leu Glu Glu Glu Leu Pro Arg Leu Ala Asp Glu Gly
20 25 30

Leu Asn Arg Arg Val Ala Glu Asp Leu Asn Leu Gly Asn Leu Asn Val
35 40 45

Ser Ile Pro Trp Thr His Lys Val Gly Asn Phe Thr Gly Leu Tyr Ser
50 55 60

Ser Thr Val Pro Val Phe Asn Pro His Trp Lys Thr Pro Ser Phe Pro
65 70 75 80

Asn Ile His Leu His Gln Asp Ile Ile Lys Lys Cys Glu Gln Phe Val
85 90 95

Gly Pro Leu Thr Val Asn Glu Lys Arg Arg Leu Gln Leu Ile Met Pro
100 105 110

Ala Arg Phe Tyr Pro Asn Val Thr Lys Tyr Leu Pro Leu Asp Lys Gly
115 120 125

[0009]

Ile Lys Pro Tyr Tyr Pro Glu His Leu Val Asn His Tyr Phe Gln Thr
130 135 140

Arg His Tyr Leu His Thr Leu Trp Lys Ala Gly Ile Leu Tyr Lys Arg
145 150 155 160

Glu Thr Thr His Ser Ala Ser Phe Cys Gly Ser Pro Tyr Ser Trp Glu
165 170 175

Gln Lys Leu Gln His Gly Ala Glu Ser Phe His Gln Gln Ser Ser Gly
180 185 190

Ile Leu Ser Arg Pro Pro Val Gly Ser Ser Leu Gln Ser Lys His Arg
195 200 205

Lys Ser Arg Leu Gly Leu Gln Ser Gln Gln Gly His Leu Ala Arg Arg
210 215 220

Gln Gln Gly Arg Ser Trp Ser Ile Arg Ala Gly Ile His Pro Thr Ala
225 230 235 240

Arg Arg Ser Phe Gly Val Glu Pro Ser Gly Ser Gly His Ser Thr Asn
245 250 255

Leu Ala Ser Lys Ser Ala Ser Cys Leu Tyr Gln Ser Pro Val Arg Lys

[0010]

Leu Glu Ser Leu Phe Thr Ala Val Thr Asn Phe Leu Leu Ser Leu Gly
545 550 555 560

Ile His Leu Asn Pro Asn Lys Thr Lys Arg Trp Gly Tyr Ser Leu His
565 570 575

Phe Met Gly Tyr Val Ile Gly Cys Tyr Gly Ser Leu Pro Gln Asp His
580 585 590

Ile Ile Gln Lys Ile Lys Glu Cys Phe Arg Lys Leu Pro Val Asn Arg
595 600 605

Pro Ile Asp Trp Lys Val Cys Gln Arg Ile Val Gly Leu Leu Gly Phe
610 615 620

Ala Ala Pro Phe Thr Gln Cys Gly Tyr Pro Ala Leu Met Pro Leu Tyr
625 630 635 640

Ala Cys Ile Gln Ser Lys Gln Ala Phe Thr Phe Ser Pro Thr Tyr Lys
645 650 655

Ala Phe Leu Cys Lys Gln Tyr Leu Asn Leu Tyr Pro Val Ala Arg Gln
660 665 670

[0011]

Arg Pro Gly Leu Cys Gln Val Phe Ala His Ala Thr Pro Thr Gly Trp
675 680 685

Gly Leu Val Met Gly His Gln Arg Met Arg Gly Thr Phe Leu Ser Arg
690 695 700

Lys Tyr Thr Ser Phe Pro Trp Leu Leu Gly Cys Ala Ala Asn Trp Ile
705 710 715 720

Leu Arg Gly Thr Ser Phe Val Tyr Tyr Pro Ser Ala Leu Asn Pro Tyr
725 730 735

His Asp Pro Ser Arg Gly Arg Leu Gly Leu Ser Arg Pro Leu Leu Arg
740 745 750

Leu Pro Phe Arg Pro Thr Thr Gly Arg Thr Ser Leu Tyr Ala Asp Ser
755 760 765

Pro Ser Val Pro Ser His Leu Pro Asp Arg Val His Phe Ala Ser Pro
770 775 780

Leu His Val Ala Trp Arg Pro Pro
785 790

<210> 6
<211> 939
<212> PRT

<213> artificial sequence

<220>

<223> fusion between a truncated version of HBc (1-148) and the mutated Y07587 polymerase

<400> 6

Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu
1 5 10 15

Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp
20 25 30

Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys
35 40 45

Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu
50 55 60

Leu Met Thr Leu Ala Thr Trp Val Gly Gly Asn Leu Glu Asp Pro Ile
65 70 75 80

Ser Arg Asp Leu Val Val Ser Tyr Val Asn Thr Asn Met Gly Leu Lys
85 90 95

Phe Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg
100 105 110

[0012]

Glu Thr Val Ile Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr
115 120 125

Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro
130 135 140

Glu Thr Thr Val Pro Leu Ser Tyr Gln His Phe Arg Arg Leu Leu Leu
145 150 155 160

Leu Asp Asp Glu Ala Gly Pro Leu Glu Glu Glu Leu Pro Arg Leu Ala
165 170 175

Asp Glu Gly Leu Asn Arg Arg Val Ala Glu Asp Leu Asn Leu Gly Asn
180 185 190

Leu Asn Val Ser Ile Pro Trp Thr His Lys Val Gly Asn Phe Thr Gly
195 200 205

Leu Tyr Ser Ser Thr Val Pro Val Phe Asn Pro His Trp Lys Thr Pro
210 215 220

Ser Phe Pro Asn Ile His Leu His Gln Asp Ile Ile Lys Lys Cys Glu
225 230 235 240

Gln Phe Val Gly Pro Leu Thr Val Asn Glu Lys Arg Arg Leu Gln Leu

	245	250	255
Ile Met Pro Ala Arg Phe Tyr Pro Asn Val Thr Lys Tyr Leu Pro Leu	260	265	270
Asp Lys Gly Ile Lys Pro Tyr Tyr Pro Glu His Leu Val Asn His Tyr	275	280	285
Phe Gln Thr Arg His Tyr Leu His Thr Leu Trp Lys Ala Gly Ile Leu	290	295	300
Tyr Lys Arg Glu Thr Thr His Ser Ala Ser Phe Cys Gly Ser Pro Tyr	305	310	315
Ser Trp Glu Gln Lys Leu Gln His Gly Ala Glu Ser Phe His Gln Gln	325	330	335
Ser Ser Gly Ile Leu Ser Arg Pro Pro Val Gly Ser Ser Leu Gln Ser	340	345	350
Lys His Arg Lys Ser Arg Leu Gly Leu Gln Ser Gln Gln Gly His Leu	355	360	365
Ala Arg Arg Gln Gln Gly Arg Ser Trp Ser Ile Arg Ala Gly Ile His	370	375	380
Pro Thr Ala Arg Arg Ser Phe Gly Val Glu Pro Ser Gly Ser Gly His	385	390	395
Ser Thr Asn Leu Ala Ser Lys Ser Ala Ser Cys Leu Tyr Gln Ser Pro	405	410	415
Val Arg Lys Ala Ala Tyr Pro Ala Val Ser Thr Phe Glu Lys His Ser	420	425	430
Ser Ser Gly His Ala Val Glu Leu His Asn Leu Pro Pro Asn Ser Ala	435	440	445
Arg Ser Gln Ser Glu Arg Pro Val Phe Pro Cys Trp Trp Leu Gln Phe	450	455	460
Arg Asn Ser Lys Pro Cys Ser Asp Tyr Cys Leu Ser His Ile Val Asn	465	470	475
Leu Leu Glu Asp Trp Gly Pro Cys Ala Glu His Gly Glu His His Ile	485	490	495
Arg Ile Pro Arg Thr Pro Ala Arg Val Thr Gly Gly Val Phe Leu Val	500	505	510
Asp Lys Asn Pro His Asn Thr Ala Glu Ser Arg Leu Val Val Asp Phe	515	520	525

[0013]

	Ser	Gln	Phe	Ser	Arg	Gly	Asn	Tyr	Arg	Val	Ser	Trp	Pro	Lys	Phe	Ala
	530						535					540				
	Val	Pro	Asn	Leu	Gln	Ser	Leu	Thr	Asn	Leu	Leu	Ser	Ser	Asn	Leu	Ser
	545					550					555					560
	Trp	Leu	Ser	Leu	Asp	Val	Ser	Ala	Ala	Phe	Tyr	His	Leu	Pro	Leu	His
					565					570					575	
	Pro	Ala	Ala	Met	Pro	His	Leu	Leu	Val	Gly	Ser	Ser	Gly	Leu	Ser	Arg
				580					585					590		
	Tyr	Val	Ala	Arg	Leu	Ser	Ser	Asn	Ser	Arg	Ile	Phe	Asn	Tyr	Gln	His
		595						600					605			
	Gly	Thr	Met	Gln	Asn	Leu	His	Asp	Ser	Cys	Ser	Arg	Asn	Leu	Tyr	Val
	610					615						620				
	Ser	Leu	Leu	Leu	Leu	Tyr	Gln	Thr	Phe	Gly	Arg	Lys	Leu	His	Leu	Tyr
	625					630					635					640
	Ser	His	Pro	Ile	Ile	Leu	Gly	Phe	Arg	Lys	Ile	Pro	Met	Gly	Val	Gly
				645					650						655	
[0014]	Leu	Ser	Pro	Phe	Leu	Leu	Ala	Gln	Phe	Thr	Ser	Ala	Ile	Cys	Ser	Val
				660					665					670		
	Val	Arg	Arg	Ala	Phe	Pro	His	Cys	Leu	Ala	Phe	Ser	Gly	Ala	Lys	Ser
		675						680					685			
	Val	Gln	His	Leu	Glu	Ser	Leu	Phe	Thr	Ala	Val	Thr	Asn	Phe	Leu	Leu
	690						695					700				
	Ser	Leu	Gly	Ile	His	Leu	Asn	Pro	Asn	Lys	Thr	Lys	Arg	Trp	Gly	Tyr
	705					710					715					720
	Ser	Leu	His	Phe	Met	Gly	Tyr	Val	Ile	Gly	Cys	Tyr	Gly	Ser	Leu	Pro
				725					730					735		
	Gln	Asp	His	Ile	Ile	Gln	Lys	Ile	Lys	Glu	Cys	Phe	Arg	Lys	Leu	Pro
			740						745					750		
	Val	Asn	Arg	Pro	Ile	Asp	Trp	Lys	Val	Cys	Gln	Arg	Ile	Val	Gly	Leu
		755					760						765			
	Leu	Gly	Phe	Ala	Ala	Pro	Phe	Thr	Gln	Cys	Gly	Tyr	Pro	Ala	Leu	Met
	770					775						780				
	Pro	Leu	Tyr	Ala	Cys	Ile	Gln	Ser	Lys	Gln	Ala	Phe	Thr	Phe	Ser	Pro
	785					790					795					800

Thr Tyr Lys Ala Phe Leu Cys Lys Gln Tyr Leu Asn Leu Tyr Pro Val
805 810 815

Ala Arg Gln Arg Pro Gly Leu Cys Gln Val Phe Ala His Ala Thr Pro
820 825 830

Thr Gly Trp Gly Leu Val Met Gly His Gln Arg Met Arg Gly Thr Phe
835 840 845

Leu Ser Arg Lys Tyr Thr Ser Phe Pro Trp Leu Leu Gly Cys Ala Ala
850 855 860

Asn Trp Ile Leu Arg Gly Thr Ser Phe Val Tyr Tyr Pro Ser Ala Leu
865 870 875 880

Asn Pro Tyr His Asp Pro Ser Arg Gly Arg Leu Gly Leu Ser Arg Pro
885 890 895

Leu Leu Arg Leu Pro Phe Arg Pro Thr Thr Gly Arg Thr Ser Leu Tyr
900 905 910

Ala Asp Ser Pro Ser Val Pro Ser His Leu Pro Asp Arg Val His Phe
915 920 925

Ala Ser Pro Leu His Val Ala Trp Arg Pro Pro
930 935

[0015]

<210> 7

<211> 860

<212> PRT

<213> artificial sequence

<220>

<223> fusion mutated Y07587 polymerase with env1 and env2 immunogenic domains

<400> 7

Met Pro Leu Ser Tyr Gln His Phe Arg Arg Leu Leu Leu Asp Asp
1 5 10 15

Glu Ala Gly Pro Leu Glu Glu Glu Leu Pro Arg Leu Ala Asp Glu Gly
20 25 30

Leu Asn Arg Arg Val Ala Glu Asp Leu Asn Leu Gly Asn Leu Asn Val
35 40 45

Ser Ile Pro Trp Thr His Lys Val Gly Asn Phe Thr Gly Leu Tyr Ser
50 55 60

Ser Thr Val Pro Val Phe Asn Pro His Trp Lys Thr Pro Ser Phe Pro
65 70 75 80

Asn Ile His Leu His Gln Asp Ile Ile Lys Lys Cys Glu Gln Phe Val

	85	90	95
	Gly Pro Leu Thr Val Asn Glu Lys Arg Arg Leu Gln Leu Ile Met Pro 100 105 110		
	Ala Arg Phe Tyr Pro Asn Val Thr Lys Tyr Leu Pro Leu Asp Lys Gly 115 120 125		
	Ile Lys Pro Tyr Tyr Pro Glu His Leu Val Asn His Tyr Phe Gln Thr 130 135 140		
	Arg His Tyr Leu His Thr Leu Trp Lys Ala Gly Ile Leu Tyr Lys Arg 145 150 155 160		
	Glu Thr Thr His Ser Ala Ser Phe Cys Gly Ser Pro Tyr Ser Trp Glu 165 170 175		
	Gln Lys Leu Gln His Gly Ala Glu Ser Phe His Gln Gln Ser Ser Gly 180 185 190		
	Ile Leu Ser Arg Pro Pro Val Gly Ser Ser Leu Gln Ser Lys His Arg 195 200 205		
	Lys Ser Arg Leu Gly Leu Gln Ser Gln Gln Gly His Leu Ala Arg Arg 210 215 220		
[0016]	Gln Gln Gly Arg Ser Trp Ser Ile Arg Ala Gly Ile His Pro Thr Ala 225 230 235 240		
	Arg Arg Ser Phe Gly Val Glu Pro Ser Gly Ser Gly His Ser Thr Asn 245 250 255		
	Leu Ala Ser Lys Ser Ala Ser Cys Leu Tyr Gln Ser Pro Val Arg Lys 260 265 270		
	Ala Ala Tyr Pro Ala Val Ser Thr Phe Glu Lys His Ser Ser Ser Gly 275 280 285		
	His Ala Val Glu Leu His Asn Leu Pro Pro Asn Ser Ala Arg Ser Gln 290 295 300		
	Ser Glu Arg Pro Val Phe Pro Cys Trp Trp Leu Gln Phe Arg Asn Ser 305 310 315 320		
	Lys Pro Cys Ser Asp Tyr Cys Leu Ser His Ile Val Asn Leu Leu Glu 325 330 335		
	Asp Trp Gly Pro Cys Ala Glu His Gly Glu His His Ile Arg Ile Pro 340 345 350		
	Arg Thr Pro Ala Arg Val Thr Gly Gly Val Phe Leu Val Asp Lys Asn 355 360 365		

	Pro	His	Asn	Thr	Ala	Glu	Ser	Arg	Leu	Val	Val	Asp	Phe	Ser	Gln	Phe	
	370						375					380					
	Ser	Arg	Gly	Asn	Tyr	Arg	Val	Ser	Trp	Pro	Lys	Phe	Ala	Val	Pro	Asn	
	385					390					395					400	
	Leu	Gln	Ser	Leu	Thr	Asn	Leu	Leu	Ser	Ser	Asn	Leu	Ser	Trp	Leu	Ser	
					405				410						415		
	Leu	Asp	Val	Ser	Ala	Ala	Phe	Tyr	His	Leu	Pro	Leu	His	Pro	Ala	Ala	
				420					425					430			
	Met	Pro	His	Leu	Leu	Val	Gly	Ser	Ser	Gly	Leu	Ser	Arg	Tyr	Val	Ala	
		435						440					445				
	Arg	Leu	Ser	Ser	Asn	Ser	Arg	Ile	Phe	Asn	Tyr	Gln	His	Gly	Thr	Met	
	450						455						460				
	Gln	Asn	Leu	His	Asp	Ser	Cys	Ser	Arg	Asn	Leu	Tyr	Val	Ser	Leu	Leu	
	465					470					475					480	
	Leu	Leu	Tyr	Gln	Thr	Phe	Gly	Arg	Lys	Leu	His	Leu	Tyr	Ser	His	Pro	
					485				490						495		
[0017]	Ile	Ile	Leu	Gly	Phe	Arg	Lys	Ile	Pro	Met	Gly	Val	Gly	Leu	Ser	Pro	
			500					505						510			
	Phe	Leu	Leu	Ala	Gln	Phe	Thr	Ser	Ala	Ile	Cys	Ser	Val	Val	Arg	Arg	
		515						520					525				
	Ala	Phe	Pro	His	Cys	Leu	Ala	Phe	Ser	Val	Leu	Gln	Ala	Gly	Phe	Phe	
		530					535						540				
	Leu	Leu	Thr	Arg	Ile	Leu	Thr	Ile	Pro	Gln	Ser	Leu	Asp	Ser	Trp	Trp	
	545					550					555					560	
	Thr	Ser	Leu	Asn	Phe	Leu	Gly	Gly	Thr	Thr	Val	Cys	Leu	Gly	Gln	Gly	
				565						570					575		
	Ala	Lys	Ser	Val	Gln	His	Leu	Glu	Ser	Leu	Phe	Thr	Ala	Val	Thr	Asn	
			580						585					590			
	Phe	Leu	Leu	Ser	Leu	Gly	Ile	His	Leu	Asn	Pro	Asn	Lys	Thr	Lys	Arg	
		595					600						605				
	Trp	Gly	Tyr	Ser	Leu	His	Phe	Met	Gly	Tyr	Val	Ile	Gly	Cys	Tyr	Gly	
	610						615					620					
	Ser	Leu	Pro	Gln	Asp	His	Ile	Ile	Gln	Lys	Ile	Lys	Glu	Cys	Phe	Arg	
	625					630					635					640	

Lys Leu Pro Val Asn Arg Pro Ile Asp Trp Lys Val Cys Gln Arg Ile
 645 650 655
 Val Gly Leu Leu Gly Phe Ala Ala Pro Phe Thr Gln Cys Gly Tyr Pro
 660 665 670
 Ala Leu Met Pro Leu Tyr Ala Cys Ile Gln Ser Lys Gln Ala Phe Thr
 675 680 685
 Phe Ser Pro Thr Tyr Lys Ala Phe Leu Cys Lys Gln Tyr Leu Asn Leu
 690 695 700
 Tyr Pro Val Ala Arg Gln Arg Pro Gly Leu Cys Gln Val Phe Ala His
 705 710 715 720
 Ala Thr Pro Thr Gly Trp Gly Leu Val Met Gly His Gln Arg Met Arg
 725 730 735
 Gly Thr Phe Leu Trp Ala Ser Ala Arg Phe Ser Trp Leu Ser Leu Leu
 740 745 750
 Val Pro Phe Val Gln Trp Phe Val Gly Leu Ser Pro Thr Val Trp Leu
 755 760 765
 Ser Val Ser Arg Lys Tyr Thr Ser Phe Pro Trp Leu Leu Gly Cys Ala
 770 775 780
 Ala Asn Trp Ile Leu Arg Gly Thr Ser Phe Val Tyr Tyr Pro Ser Ala
 785 790 795 800
 Leu Asn Pro Tyr His Asp Pro Ser Arg Gly Arg Leu Gly Leu Ser Arg
 805 810 815
 Pro Leu Leu Arg Leu Pro Phe Arg Pro Thr Thr Gly Arg Thr Ser Leu
 820 825 830
 Tyr Ala Asp Ser Pro Ser Val Pro Ser His Leu Pro Asp Arg Val His
 835 840 845
 Phe Ala Ser Pro Leu His Val Ala Trp Arg Pro Pro
 850 855 860
 <210> 8
 <211> 1007
 <212> PRT
 <213> artificial sequence
 <220>
 <223> fusion between truncated core (1-148)- mutated Y07587 polymerase
 - env1 and env2
 <400> 8
 Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu

1	5	10	15
Ser Phe Leu	Pro Ser Asp Phe Phe	Pro Ser Val Arg Asp	Leu Leu Asp
	20	25	30
Thr Ala Ser	Ala Leu Tyr Arg Glu	Ala Leu Glu Ser	Pro Glu His Cys
	35	40	45
Ser Pro His	His Thr Ala Leu Arg Gln	Ala Ile Leu Cys Trp	Gly Glu
	50	55	60
Leu Met Thr	Leu Ala Thr Trp Val Gly	Gly Asn Leu Glu Asp	Pro Ile
65	70	75	80
Ser Arg Asp	Leu Val Val Ser Tyr Val	Asn Thr Asn Met Gly	Leu Lys
	85	90	95
Phe Arg Gln	Leu Leu Trp Phe His Ile	Ser Cys Leu Thr Phe	Gly Arg
	100	105	110
Glu Thr Val	Ile Glu Tyr Leu Val Ser	Phe Gly Val Trp Ile	Arg Thr
	115	120	125
Pro Pro Ala	Tyr Arg Pro Pro Asn	Ala Pro Ile Leu Ser	Thr Leu Pro
	130	135	140
Glu Thr Thr	Val Pro Leu Ser Tyr Gln	His Phe Arg Arg	Leu Leu Leu
145	150	155	160
Leu Asp Asp	Glu Ala Gly Pro Leu Glu	Glu Glu Leu Pro Arg	Leu Ala
	165	170	175
Asp Glu Gly	Leu Asn Arg Arg Val	Ala Glu Asp Leu Asn	Leu Gly Asn
	180	185	190
Leu Asn Val	Ser Ile Pro Trp Thr His	Lys Val Gly Asn Phe	Thr Gly
	195	200	205
Leu Tyr Ser	Ser Thr Val Pro Val Phe	Asn Pro His Trp Lys	Thr Pro
	210	215	220
Ser Phe Pro	Asn Ile His Leu His Gln	Asp Ile Ile Lys Lys	Cys Glu
225	230	235	240
Gln Phe Val	Gly Pro Leu Thr Val Asn	Glu Lys Arg Arg Leu	Gln Leu
	245	250	255
Ile Met Pro	Ala Arg Phe Tyr Pro Asn	Val Thr Lys Tyr Leu	Pro Leu
	260	265	270
Asp Lys Gly	Ile Lys Pro Tyr Tyr	Pro Glu His Leu Val	Asn His Tyr
	275	280	285

[0019]

Phe Gln Thr Arg His Tyr Leu His Thr Leu Trp Lys Ala Gly Ile Leu
 290 295 300

Tyr Lys Arg Glu Thr Thr His Ser Ala Ser Phe Cys Gly Ser Pro Tyr
 305 310 315 320

Ser Trp Glu Gln Lys Leu Gln His Gly Ala Glu Ser Phe His Gln Gln
 325 330 335

Ser Ser Gly Ile Leu Ser Arg Pro Pro Val Gly Ser Ser Leu Gln Ser
 340 345 350

Lys His Arg Lys Ser Arg Leu Gly Leu Gln Ser Gln Gln Gly His Leu
 355 360 365

Ala Arg Arg Gln Gln Gly Arg Ser Trp Ser Ile Arg Ala Gly Ile His
 370 375 380

Pro Thr Ala Arg Arg Ser Phe Gly Val Glu Pro Ser Gly Ser Gly His
 385 390 395 400

Ser Thr Asn Leu Ala Ser Lys Ser Ala Ser Cys Leu Tyr Gln Ser Pro
 405 410 415

[0020] Val Arg Lys Ala Ala Tyr Pro Ala Val Ser Thr Phe Glu Lys His Ser
 420 425 430

Ser Ser Gly His Ala Val Glu Leu His Asn Leu Pro Pro Asn Ser Ala
 435 440 445

Arg Ser Gln Ser Glu Arg Pro Val Phe Pro Cys Trp Trp Leu Gln Phe
 450 455 460

Arg Asn Ser Lys Pro Cys Ser Asp Tyr Cys Leu Ser His Ile Val Asn
 465 470 475 480

Leu Leu Glu Asp Trp Gly Pro Cys Ala Glu His Gly Glu His His Ile
 485 490 495

Arg Ile Pro Arg Thr Pro Ala Arg Val Thr Gly Gly Val Phe Leu Val
 500 505 510

Asp Lys Asn Pro His Asn Thr Ala Glu Ser Arg Leu Val Val Asp Phe
 515 520 525

Ser Gln Phe Ser Arg Gly Asn Tyr Arg Val Ser Trp Pro Lys Phe Ala
 530 535 540

Val Pro Asn Leu Gln Ser Leu Thr Asn Leu Leu Ser Ser Asn Leu Ser
 545 550 555 560

	Trp	Leu	Ser	Leu	Asp	Val	Ser	Ala	Ala	Phe	Tyr	His	Leu	Pro	Leu	His	
					565					570						575	
	Pro	Ala	Ala	Met	Pro	His	Leu	Leu	Val	Gly	Ser	Ser	Gly	Leu	Ser	Arg	
				580					585					590			
	Tyr	Val	Ala	Arg	Leu	Ser	Ser	Asn	Ser	Arg	Ile	Phe	Asn	Tyr	Gln	His	
		595						600					605				
	Gly	Thr	Met	Gln	Asn	Leu	His	Asp	Ser	Cys	Ser	Arg	Asn	Leu	Tyr	Val	
		610					615						620				
	Ser	Leu	Leu	Leu	Leu	Tyr	Gln	Thr	Phe	Gly	Arg	Lys	Leu	His	Leu	Tyr	
	625					630					635					640	
	Ser	His	Pro	Ile	Ile	Leu	Gly	Phe	Arg	Lys	Ile	Pro	Met	Gly	Val	Gly	
				645						650					655		
	Leu	Ser	Pro	Phe	Leu	Leu	Ala	Gln	Phe	Thr	Ser	Ala	Ile	Cys	Ser	Val	
				660					665						670		
	Val	Arg	Arg	Ala	Phe	Pro	His	Cys	Leu	Ala	Phe	Ser	Val	Leu	Gln	Ala	
		675						680					685				
[0021]	Gly	Phe	Phe	Leu	Leu	Thr	Arg	Ile	Leu	Thr	Ile	Pro	Gln	Ser	Leu	Asp	
		690					695					700					
	Ser	Trp	Trp	Thr	Ser	Leu	Asn	Phe	Leu	Gly	Gly	Thr	Thr	Val	Cys	Leu	
	705					710					715				720		
	Gly	Gln	Gly	Ala	Lys	Ser	Val	Gln	His	Leu	Glu	Ser	Leu	Phe	Thr	Ala	
				725						730					735		
	Val	Thr	Asn	Phe	Leu	Leu	Ser	Leu	Gly	Ile	His	Leu	Asn	Pro	Asn	Lys	
				740					745					750			
	Thr	Lys	Arg	Trp	Gly	Tyr	Ser	Leu	His	Phe	Met	Gly	Tyr	Val	Ile	Gly	
		755						760					765				
	Cys	Tyr	Gly	Ser	Leu	Pro	Gln	Asp	His	Ile	Ile	Gln	Lys	Ile	Lys	Glu	
		770					775						780				
	Cys	Phe	Arg	Lys	Leu	Pro	Val	Asn	Arg	Pro	Ile	Asp	Trp	Lys	Val	Cys	
	785					790					795					800	
	Gln	Arg	Ile	Val	Gly	Leu	Leu	Gly	Phe	Ala	Ala	Pro	Phe	Thr	Gln	Cys	
				805						810					815		
	Gly	Tyr	Pro	Ala	Leu	Met	Pro	Leu	Tyr	Ala	Cys	Ile	Gln	Ser	Lys	Gln	
				820					825						830		

Ala Phe Thr Phe Ser Pro Thr Tyr Lys Ala Phe Leu Cys Lys Gln Tyr
 835 840 845
 Leu Asn Leu Tyr Pro Val Ala Arg Gln Arg Pro Gly Leu Cys Gln Val
 850 855 860
 Phe Ala His Ala Thr Pro Thr Gly Trp Gly Leu Val Met Gly His Gln
 865 870 875 880
 Arg Met Arg Gly Thr Phe Leu Trp Ala Ser Ala Arg Phe Ser Trp Leu
 885 890 895
 Ser Leu Leu Val Pro Phe Val Gln Trp Phe Val Gly Leu Ser Pro Thr
 900 905 910
 Val Trp Leu Ser Val Ser Arg Lys Tyr Thr Ser Phe Pro Trp Leu Leu
 915 920 925
 Gly Cys Ala Ala Asn Trp Ile Leu Arg Gly Thr Ser Phe Val Tyr Tyr
 930 935 940
 Pro Ser Ala Leu Asn Pro Tyr His Asp Pro Ser Arg Gly Arg Leu Gly
 945 950 955 960
 Leu Ser Arg Pro Leu Leu Arg Leu Pro Phe Arg Pro Thr Thr Gly Arg
 965 970 975
 Thr Ser Leu Tyr Ala Asp Ser Pro Ser Val Pro Ser His Leu Pro Asp
 980 985 990
 Arg Val His Phe Ala Ser Pro Leu His Val Ala Trp Arg Pro Pro
 995 1000 1005
 <210> 9
 <211> 1042
 <212> PRT
 <213> artificial sequence
 <220>
 <223> fusion between native core (1-183) - mutated Y07587 polymerase -
 env1 and env2
 <400> 9
 Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu
 1 5 10 15
 Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp
 20 25 30
 Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys
 35 40 45
 Ser Pro His His Thr Ala Leu Arg Glu Ala Ile Leu Cys Trp Gly Glu
 50 55 60

[0022]

	Leu	Met	Thr	Leu	Ala	Thr	Trp	Val	Gly	Gly	Asn	Leu	Glu	Asp	Pro	Ile	
	65					70					75					80	
	Ser	Arg	Asp	Leu	Val	Val	Ser	Tyr	Val	Asn	Thr	Asn	Met	Gly	Leu	Lys	
				85						90					95		
	Phe	Arg	Gln	Leu	Leu	Trp	Phe	His	Ile	Ser	Cys	Leu	Thr	Phe	Gly	Arg	
				100					105						110		
	Glu	Thr	Val	Ile	Glu	Tyr	Leu	Val	Ser	Phe	Gly	Val	Trp	Ile	Arg	Thr	
			115					120					125				
	Pro	Pro	Ala	Tyr	Arg	Pro	Pro	Asn	Ala	Pro	Ile	Leu	Ser	Thr	Leu	Pro	
			130					135					140				
	Glu	Thr	Thr	Val	Val	Arg	Arg	Arg	Gly	Arg	Ser	Pro	Arg	Arg	Arg	Thr	
	145					150					155					160	
	Pro	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Ser	Pro	Arg	Arg	Arg	Arg	Ser	
					165					170						175	
	Gln	Ser	Arg	Glu	Ser	Gln	Cys	Pro	Leu	Ser	Tyr	Gln	His	Phe	Arg	Arg	
				180						185					190		
[0023]	Leu	Leu	Leu	Leu	Asp	Asp	Glu	Ala	Gly	Pro	Leu	Glu	Glu	Glu	Leu	Pro	
			195					200					205				
	Arg	Leu	Ala	Asp	Glu	Gly	Leu	Asn	Arg	Arg	Val	Ala	Glu	Asp	Leu	Asn	
		210					215						220				
	Leu	Gly	Asn	Leu	Asn	Val	Ser	Ile	Pro	Trp	Thr	His	Lys	Val	Gly	Asn	
	225					230					235					240	
	Phe	Thr	Gly	Leu	Tyr	Ser	Ser	Thr	Val	Pro	Val	Phe	Asn	Pro	His	Trp	
				245						250					255		
	Lys	Thr	Pro	Ser	Phe	Pro	Asn	Ile	His	Leu	His	Gln	Asp	Ile	Ile	Lys	
				260				265						270			
	Lys	Cys	Glu	Gln	Phe	Val	Gly	Pro	Leu	Thr	Val	Asn	Glu	Lys	Arg	Arg	
			275					280					285				
	Leu	Gln	Leu	Ile	Met	Pro	Ala	Arg	Phe	Tyr	Pro	Asn	Val	Thr	Lys	Tyr	
		290					295						300				
	Leu	Pro	Leu	Asp	Lys	Gly	Ile	Lys	Pro	Tyr	Tyr	Pro	Glu	His	Leu	Val	
	305					310					315					320	
	Asn	His	Tyr	Phe	Gln	Thr	Arg	His	Tyr	Leu	His	Thr	Leu	Trp	Lys	Ala	
					325					330					335		

	Gly	Ile	Leu	Tyr	Lys	Arg	Glu	Thr	Thr	His	Ser	Ala	Ser	Phe	Cys	Gly
				340					345						350	
	Ser	Pro	Tyr	Ser	Trp	Glu	Gln	Lys	Leu	Gln	His	Gly	Ala	Glu	Ser	Phe
			355					360					365			
	His	Gln	Gln	Ser	Ser	Gly	Ile	Leu	Ser	Arg	Pro	Pro	Val	Gly	Ser	Ser
		370					375					380				
	Leu	Gln	Ser	Lys	His	Arg	Lys	Ser	Arg	Leu	Gly	Leu	Gln	Ser	Gln	Gln
	385					390					395					400
	Gly	His	Leu	Ala	Arg	Arg	Gln	Gln	Gly	Arg	Ser	Trp	Ser	Ile	Arg	Ala
					405					410					415	
	Gly	Ile	His	Pro	Thr	Ala	Arg	Arg	Ser	Phe	Gly	Val	Glu	Pro	Ser	Gly
				420						425				430		
	Ser	Gly	His	Ser	Thr	Asn	Leu	Ala	Ser	Lys	Ser	Ala	Ser	Cys	Leu	Tyr
			435					440						445		
	Gln	Ser	Pro	Val	Arg	Lys	Ala	Ala	Tyr	Pro	Ala	Val	Ser	Thr	Phe	Glu
		450					455					460				
[0024]	Lys	His	Ser	Ser	Ser	Gly	His	Ala	Val	Glu	Leu	His	Asn	Leu	Pro	Pro
	465					470					475				480	
	Asn	Ser	Ala	Arg	Ser	Gln	Ser	Glu	Arg	Pro	Val	Phe	Pro	Cys	Trp	Trp
					485					490					495	
	Leu	Gln	Phe	Arg	Asn	Ser	Lys	Pro	Cys	Ser	Asp	Tyr	Cys	Leu	Ser	His
			500						505					510		
	Ile	Val	Asn	Leu	Leu	Glu	Asp	Trp	Gly	Pro	Cys	Ala	Glu	His	Gly	Glu
		515						520					525			
	His	His	Ile	Arg	Ile	Pro	Arg	Thr	Pro	Ala	Arg	Val	Thr	Gly	Gly	Val
		530				535						540				
	Phe	Leu	Val	Asp	Lys	Asn	Pro	His	Asn	Thr	Ala	Glu	Ser	Arg	Leu	Val
	545					550					555				560	
	Val	Asp	Phe	Ser	Gln	Phe	Ser	Arg	Gly	Asn	Tyr	Arg	Val	Ser	Trp	Pro
					565					570					575	
	Lys	Phe	Ala	Val	Pro	Asn	Leu	Gln	Ser	Leu	Thr	Asn	Leu	Leu	Ser	Ser
			580						585					590		
	Asn	Leu	Ser	Trp	Leu	Ser	Leu	Asp	Val	Ser	Ala	Ala	Phe	Tyr	His	Leu
		595						600					605			

```

Pro Leu His Pro Ala Ala Met Pro His Leu Leu Val Gly Ser Ser Gly
610                               615                               620

Leu Ser Arg Tyr Val Ala Arg Leu Ser Ser Asn Ser Arg Ile Phe Asn
625                               630                               635                               640

Tyr Gln His Gly Thr Met Gln Asn Leu His Asp Ser Cys Ser Arg Asn
645                               650                               655

Leu Tyr Val Ser Leu Leu Leu Leu Tyr Gln Thr Phe Gly Arg Lys Leu
660                               665                               670

His Leu Tyr Ser His Pro Ile Ile Leu Gly Phe Arg Lys Ile Pro Met
675                               680                               685

Gly Val Gly Leu Ser Pro Phe Leu Leu Ala Gln Phe Thr Ser Ala Ile
690                               695                               700

Cys Ser Val Val Arg Arg Ala Phe Pro His Cys Leu Ala Phe Ser Val
705                               710                               715                               720

Leu Gln Ala Gly Phe Phe Leu Leu Thr Arg Ile Leu Thr Ile Pro Gln
725                               730                               735

Ser Leu Asp Ser Trp Trp Thr Ser Leu Asn Phe Leu Gly Gly Thr Thr
740                               745                               750

Val Cys Leu Gly Gln Gly Ala Lys Ser Val Gln His Leu Glu Ser Leu
755                               760                               765

Phe Thr Ala Val Thr Asn Phe Leu Leu Ser Leu Gly Ile His Leu Asn
770                               775                               780

Pro Asn Lys Thr Lys Arg Trp Gly Tyr Ser Leu His Phe Met Gly Tyr
785                               790                               795                               800

Val Ile Gly Cys Tyr Gly Ser Leu Pro Gln Asp His Ile Ile Gln Lys
805                               810                               815

Ile Lys Glu Cys Phe Arg Lys Leu Pro Val Asn Arg Pro Ile Asp Trp
820                               825                               830

Lys Val Cys Gln Arg Ile Val Gly Leu Leu Gly Phe Ala Ala Pro Phe
835                               840                               845

Thr Gln Cys Gly Tyr Pro Ala Leu Met Pro Leu Tyr Ala Cys Ile Gln
850                               855                               860

Ser Lys Gln Ala Phe Thr Phe Ser Pro Thr Tyr Lys Ala Phe Leu Cys
865                               870                               875                               880

Lys Gln Tyr Leu Asn Leu Tyr Pro Val Ala Arg Gln Arg Pro Gly Leu

```

[0025]

[0026]

68

	Leu	Gly	Asn	Leu	Asn	Val	Ser	Ile	Pro	Trp	Thr	His	Lys	Val	Gly	Asn	
	65					70					75					80	
	Phe	Thr	Gly	Leu	Tyr	Ser	Ser	Thr	Val	Pro	Val	Phe	Asn	Pro	His	Trp	
				85					90						95		
	Lys	Thr	Pro	Ser	Phe	Pro	Asn	Ile	His	Leu	His	Gln	Asp	Ile	Ile	Lys	
				100				105						110			
	Lys	Cys	Glu	Gln	Phe	Val	Gly	Pro	Leu	Thr	Val	Asn	Glu	Lys	Arg	Arg	
			115					120					125				
	Leu	Gln	Leu	Ile	Met	Pro	Ala	Arg	Phe	Tyr	Pro	Asn	Val	Thr	Lys	Tyr	
		130					135					140					
	Leu	Pro	Leu	Asp	Lys	Gly	Ile	Lys	Pro	Tyr	Tyr	Pro	Glu	His	Leu	Val	
	145				150					155					160		
	Asn	His	Tyr	Phe	Gln	Thr	Arg	His	Tyr	Leu	His	Thr	Leu	Trp	Lys	Ala	
				165					170						175		
	Gly	Ile	Leu	Tyr	Lys	Arg	Glu	Thr	Thr	His	Ser	Ala	Ser	Phe	Cys	Gly	
			180						185					190			
[0027]	Ser	Pro	Tyr	Ser	Trp	Glu	Gln	Lys	Leu	Gln	His	Gly	Ala	Glu	Ser	Phe	
			195					200					205				
	His	Gln	Gln	Ser	Ser	Gly	Ile	Leu	Ser	Arg	Pro	Pro	Val	Gly	Ser	Ser	
		210					215					220					
	Leu	Gln	Ser	Lys	His	Arg	Lys	Ser	Arg	Leu	Gly	Leu	Gln	Ser	Gln	Gln	
	225					230					235				240		
	Gly	His	Leu	Ala	Arg	Arg	Gln	Gln	Gly	Arg	Ser	Trp	Ser	Ile	Arg	Ala	
				245					250					255			
	Gly	Ile	His	Pro	Thr	Ala	Arg	Arg	Ser	Phe	Gly	Val	Glu	Pro	Ser	Gly	
			260						265				270				
	Ser	Gly	His	Ser	Thr	Asn	Leu	Ala	Ser	Lys	Ser	Ala	Ser	Cys	Leu	Tyr	
		275						280					285				
	Gln	Ser	Pro	Val	Arg	Lys	Ala	Ala	Tyr	Pro	Ala	Val	Ser	Thr	Phe	Glu	
		290					295					300					
	Lys	His	Ser	Ser	Ser	Gly	His	Ala	Val	Glu	Leu	His	Asn	Leu	Pro	Pro	
	305					310					315				320		
	Asn	Ser	Ala	Arg	Ser	Gln	Ser	Glu	Arg	Pro	Val	Phe	Pro	Cys	Trp	Trp	
				325					330					335			

Leu Gln Phe Arg Asn Ser Lys Pro Cys Ser Asp Tyr Cys Leu Ser His
 340 345 350
 Ile Val Asn Leu Leu Glu Asp Trp Gly Pro Cys Ala Glu His Gly Glu
 355 360 365
 His His Ile Arg Ile Pro Arg Thr Pro Ala Arg Val Thr Gly Gly Val
 370 375 380
 Phe Leu Val Asp Lys Asn Pro His Asn Thr Ala Glu Ser Arg Leu Val
 385 390 395 400
 Val Asp Phe Ser Gln Phe Ser Arg Gly Asn Tyr Arg Val Ser Trp Pro
 405 410 415
 Lys Phe Ala Val Pro Asn Leu Gln Ser Leu Thr Asn Leu Leu Ser Ser
 420 425 430
 Asn Leu Ser Trp Leu Ser Leu Asp Val Ser Ala Ala Phe Tyr His Leu
 435 440 445
 Pro Leu His Pro Ala Ala Met Pro His Leu Leu Val Gly Ser Ser Gly
 450 455 460
 Leu Ser Arg Tyr Val Ala Arg Leu Ser Ser Asn Ser Arg Ile Phe Asn
 465 470 475 480
 Tyr Gln His Gly Thr Met Gln Asn Leu His Asp Ser Cys Ser Arg Asn
 485 490 495
 Leu Tyr Val Ser Leu Leu Leu Leu Tyr Gln Thr Phe Gly Arg Lys Leu
 500 505 510
 His Leu Tyr Ser His Pro Ile Ile Leu Gly Phe Arg Lys Ile Pro Met
 515 520 525
 Gly Val Gly Leu Ser Pro Phe Leu Leu Ala Gln Phe Thr Ser Ala Ile
 530 535 540
 Cys Ser Val Val Arg Arg Ala Phe Pro His Cys Leu Ala Phe Ser Gly
 545 550 555 560
 Ala Lys Ser Val Gln His Leu Glu Ser Leu Phe Thr Ala Val Thr Asn
 565 570 575
 Phe Leu Leu Ser Leu Gly Ile His Leu Asn Pro Asn Lys Thr Lys Arg
 580 585 590
 Trp Gly Tyr Ser Leu His Phe Met Gly Tyr Val Ile Gly Cys Tyr Gly
 595 600 605
 Ser Leu Pro Gln Asp His Ile Ile Gln Lys Ile Lys Glu Cys Phe Arg

[0028]

610	615	620
Lys Leu Pro Val Asn Arg Pro Ile Asp Trp Lys Val Cys Gln Arg Ile 625 630 635 640		
Val Gly Leu Leu Gly Phe Ala Ala Pro Phe Thr Gln Cys Gly Tyr Pro 645 650 655		
Ala Leu Met Pro Leu Tyr Ala Cys Ile Gln Ser Lys Gln Ala Phe Thr 660 665 670		
Phe Ser Pro Thr Tyr Lys Ala Phe Leu Cys Lys Gln Tyr Leu Asn Leu 675 680 685		
Tyr Pro Val Ala Arg Gln Arg Pro Gly Leu Cys Gln Val Phe Ala His 690 695 700		
Ala Thr Pro Thr Gly Trp Gly Leu Val Met Gly His Gln Arg Met Arg 705 710 715 720		
Gly Thr Phe Leu Ser Arg Lys Tyr Thr Ser Phe Pro Trp Leu Leu Gly 725 730 735		
Cys Ala Ala Asn Trp Ile Leu Arg Gly Thr Ser Phe Val Tyr Tyr Pro 740 745 750		
[0029]		
Ser Ala Leu Asn Pro Tyr His Asp Pro Ser Arg Gly Arg Leu Gly Leu 755 760 765		
Ser Arg Pro Leu Leu Arg Leu Pro Phe Arg Pro Thr Thr Gly Arg Thr 770 775 780		
Ser Leu Tyr Ala Asp Ser Pro Ser Val Pro Ser His Leu Pro Asp Arg 785 790 795 800		
Val His Phe Ala Ser Pro Leu His Val Ala Trp Arg Pro Pro Tyr Val 805 810 815		
Leu Leu Ser Ala Gly Ala Leu Thr Ala Leu Met Leu Ile Ile Phe Leu 820 825 830		
Met Thr Cys Cys Arg Arg Val Asn Arg Ser Glu Pro Thr Gln His Asn 835 840 845		
Leu Arg Gly Thr Gly Arg Glu Val Ser Val Thr Pro Gln Ser Gly Lys 850 855 860		
Ile Ile Ser Ser Trp Glu Ser His Lys Ser Gly Gly Glu Thr Arg Leu 865 870 875 880		
<210> 11		
<211> 1027		

<212> PRT
 <213> artificial sequence
 <220>
 <223> fusion between sequence signal - truncated core (1-148) -mutated
 Y07587 polymerase - sequence TM

<400> 11

Met Val Pro Gln Ala Leu Leu Phe Val Pro Leu Leu Val Phe Pro Leu
 1 5 10 15

Cys Phe Gly Lys Phe Pro Ile Asp Ile Asp Pro Tyr Lys Glu Phe Gly
 20 25 30

Ala Thr Val Glu Leu Leu Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser
 35 40 45

Val Arg Asp Leu Leu Asp Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu
 50 55 60

Glu Ser Pro Glu His Cys Ser Pro His His Thr Ala Leu Arg Gln Ala
 65 70 75 80

Ile Leu Cys Trp Gly Glu Leu Met Thr Leu Ala Thr Trp Val Gly Gly
 85 90 95

[0030] Asn Leu Glu Asp Pro Ile Ser Arg Asp Leu Val Val Ser Tyr Val Asn
 100 105 110

Thr Asn Met Gly Leu Lys Phe Arg Gln Leu Leu Trp Phe His Ile Ser
 115 120 125

Cys Leu Thr Phe Gly Arg Glu Thr Val Ile Glu Tyr Leu Val Ser Phe
 130 135 140

Gly Val Trp Ile Arg Thr Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro
 145 150 155 160

Ile Leu Ser Thr Leu Pro Glu Thr Thr Val Pro Leu Ser Tyr Gln His
 165 170 175

Phe Arg Arg Leu Leu Leu Leu Asp Asp Glu Ala Gly Pro Leu Glu Glu
 180 185 190

Glu Leu Pro Arg Leu Ala Asp Glu Gly Leu Asn Arg Arg Val Ala Glu
 195 200 205

Asp Leu Asn Leu Gly Asn Leu Asn Val Ser Ile Pro Trp Thr His Lys
 210 215 220

Val Gly Asn Phe Thr Gly Leu Tyr Ser Ser Thr Val Pro Val Phe Asn
 225 230 235 240

Pro His Trp Lys Thr Pro Ser Phe Pro Asn Ile His Leu His Gln Asp
 245 250 255
 Ile Ile Lys Lys Cys Glu Gln Phe Val Gly Pro Leu Thr Val Asn Glu
 260 265 270
 Lys Arg Arg Leu Gln Leu Ile Met Pro Ala Arg Phe Tyr Pro Asn Val
 275 280 285
 Thr Lys Tyr Leu Pro Leu Asp Lys Gly Ile Lys Pro Tyr Tyr Pro Glu
 290 295 300
 His Leu Val Asn His Tyr Phe Gln Thr Arg His Tyr Leu His Thr Leu
 305 310 315 320
 Trp Lys Ala Gly Ile Leu Tyr Lys Arg Glu Thr Thr His Ser Ala Ser
 325 330 335
 Phe Cys Gly Ser Pro Tyr Ser Trp Glu Gln Lys Leu Gln His Gly Ala
 340 345 350
 Glu Ser Phe His Gln Gln Ser Ser Gly Ile Leu Ser Arg Pro Pro Val
 355 360 365
 Gly Ser Ser Leu Gln Ser Lys His Arg Lys Ser Arg Leu Gly Leu Gln
 370 375 380
 Ser Gln Gln Gly His Leu Ala Arg Arg Gln Gln Gly Arg Ser Trp Ser
 385 390 395 400
 Ile Arg Ala Gly Ile His Pro Thr Ala Arg Arg Ser Phe Gly Val Glu
 405 410 415
 Pro Ser Gly Ser Gly His Ser Thr Asn Leu Ala Ser Lys Ser Ala Ser
 420 425 430
 Cys Leu Tyr Gln Ser Pro Val Arg Lys Ala Ala Tyr Pro Ala Val Ser
 435 440 445
 Thr Phe Glu Lys His Ser Ser Ser Gly His Ala Val Glu Leu His Asn
 450 455 460
 Leu Pro Pro Asn Ser Ala Arg Ser Gln Ser Glu Arg Pro Val Phe Pro
 465 470 475 480
 Cys Trp Trp Leu Gln Phe Arg Asn Ser Lys Pro Cys Ser Asp Tyr Cys
 485 490 495
 Leu Ser His Ile Val Asn Leu Leu Glu Asp Trp Gly Pro Cys Ala Glu
 500 505 510
 His Gly Glu His His Ile Arg Ile Pro Arg Thr Pro Ala Arg Val Thr

[0031]

515	520	525
Gly Gly Val Phe Leu Val Asp Lys Asn Pro His Asn Thr Ala Glu Ser		
530	535	540
Arg Leu Val Val Asp Phe Ser Gln Phe Ser Arg Gly Asn Tyr Arg Val		
545	550	555
Ser Trp Pro Lys Phe Ala Val Pro Asn Leu Gln Ser Leu Thr Asn Leu		
565	570	575
Leu Ser Ser Asn Leu Ser Trp Leu Ser Leu Asp Val Ser Ala Ala Phe		
580	585	590
Tyr His Leu Pro Leu His Pro Ala Ala Met Pro His Leu Leu Val Gly		
595	600	605
Ser Ser Gly Leu Ser Arg Tyr Val Ala Arg Leu Ser Ser Asn Ser Arg		
610	615	620
Ile Phe Asn Tyr Gln His Gly Thr Met Gln Asn Leu His Asp Ser Cys		
625	630	635
Ser Arg Asn Leu Tyr Val Ser Leu Leu Leu Leu Tyr Gln Thr Phe Gly		
645	650	655
Arg Lys Leu His Leu Tyr Ser His Pro Ile Ile Leu Gly Phe Arg Lys		
660	665	670
Ile Pro Met Gly Val Gly Leu Ser Pro Phe Leu Leu Ala Gln Phe Thr		
675	680	685
Ser Ala Ile Cys Ser Val Val Arg Arg Ala Phe Pro His Cys Leu Ala		
690	695	700
Phe Ser Gly Ala Lys Ser Val Gln His Leu Glu Ser Leu Phe Thr Ala		
705	710	715
Val Thr Asn Phe Leu Leu Ser Leu Gly Ile His Leu Asn Pro Asn Lys		
725	730	735
Thr Lys Arg Trp Gly Tyr Ser Leu His Phe Met Gly Tyr Val Ile Gly		
740	745	750
Cys Tyr Gly Ser Leu Pro Gln Asp His Ile Ile Gln Lys Ile Lys Glu		
755	760	765
Cys Phe Arg Lys Leu Pro Val Asn Arg Pro Ile Asp Trp Lys Val Cys		
770	775	780
Gln Arg Ile Val Gly Leu Leu Gly Phe Ala Ala Pro Phe Thr Gln Cys		
785	790	795
		800

[0032]

	Gly Tyr Pro Ala Leu Met Pro Leu Tyr Ala Cys Ile Gln Ser Lys Gln	
	805	810 815
	Ala Phe Thr Phe Ser Pro Thr Tyr Lys Ala Phe Leu Cys Lys Gln Tyr	
	820	825 830
	Leu Asn Leu Tyr Pro Val Ala Arg Gln Arg Pro Gly Leu Cys Gln Val	
	835	840 845
	Phe Ala His Ala Thr Pro Thr Gly Trp Gly Leu Val Met Gly His Gln	
	850	855 860
	Arg Met Arg Gly Thr Phe Leu Ser Arg Lys Tyr Thr Ser Phe Pro Trp	
	865	870 875 880
	Leu Leu Gly Cys Ala Ala Asn Trp Ile Leu Arg Gly Thr Ser Phe Val	
	885	890 895
	Tyr Tyr Pro Ser Ala Leu Asn Pro Tyr His Asp Pro Ser Arg Gly Arg	
	900	905 910
	Leu Gly Leu Ser Arg Pro Leu Leu Arg Leu Pro Phe Arg Pro Thr Thr	
	915	920 925
[0033]	Gly Arg Thr Ser Leu Tyr Ala Asp Ser Pro Ser Val Pro Ser His Leu	
	930	935 940
	Pro Asp Arg Val His Phe Ala Ser Pro Leu His Val Ala Trp Arg Pro	
	945	950 955 960
	Pro Tyr Val Leu Leu Ser Ala Gly Ala Leu Thr Ala Leu Met Leu Ile	
	965	970 975
	Ile Phe Leu Met Thr Cys Cys Arg Arg Val Asn Arg Ser Glu Pro Thr	
	980	985 990
	Gln His Asn Leu Arg Gly Thr Gly Arg Glu Val Ser Val Thr Pro Gln	
	995	1000 1005
	Ser Gly Lys Ile Ile Ser Ser Trp Glu Ser His Lys Ser Gly Gly	
	1010	1015 1020
	Glu Thr Arg Leu	
	1025	
	<210> 12	
	<211> 1095	
	<212> PRT	
	<213> artificial sequence	
	<220>	
	<223> fusion between signal sequence -core (1'148)- mutated Y07587	

polymerase -env1-env2 and TM sequence

<400> 12

Met Val Pro Gln Ala Leu Leu Phe Val Pro Leu Leu Val Phe Pro Leu
1 5 10 15

Cys Phe Gly Lys Phe Pro Ile Asp Ile Asp Pro Tyr Lys Glu Phe Gly
20 25 30

Ala Thr Val Glu Leu Leu Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser
35 40 45

Val Arg Asp Leu Leu Asp Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu
50 55 60

Glu Ser Pro Glu His Cys Ser Pro His His Thr Ala Leu Arg Gln Ala
65 70 75 80

Ile Leu Cys Trp Gly Glu Leu Met Thr Leu Ala Thr Trp Val Gly Gly
85 90 95

Asn Leu Glu Asp Pro Ile Ser Arg Asp Leu Val Val Ser Tyr Val Asn
100 105 110

Thr Asn Met Gly Leu Lys Phe Arg Gln Leu Leu Trp Phe His Ile Ser
115 120 125

[0034]

Cys Leu Thr Phe Gly Arg Glu Thr Val Ile Glu Tyr Leu Val Ser Phe
130 135 140

Gly Val Trp Ile Arg Thr Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro
145 150 155 160

Ile Leu Ser Thr Leu Pro Glu Thr Thr Val Pro Leu Ser Tyr Gln His
165 170 175

Phe Arg Arg Leu Leu Leu Leu Asp Asp Glu Ala Gly Pro Leu Glu Glu
180 185 190

Glu Leu Pro Arg Leu Ala Asp Glu Gly Leu Asn Arg Arg Val Ala Glu
195 200 205

Asp Leu Asn Leu Gly Asn Leu Asn Val Ser Ile Pro Trp Thr His Lys
210 215 220

Val Gly Asn Phe Thr Gly Leu Tyr Ser Ser Thr Val Pro Val Phe Asn
225 230 235 240

Pro His Trp Lys Thr Pro Ser Phe Pro Asn Ile His Leu His Gln Asp
245 250 255

Ile Ile Lys Lys Cys Glu Gln Phe Val Gly Pro Leu Thr Val Asn Glu

	260	265	270
	Lys Arg Arg Leu Gln Leu Ile Met Pro Ala Arg Phe Tyr Pro Asn Val 275 280 285		
	Thr Lys Tyr Leu Pro Leu Asp Lys Gly Ile Lys Pro Tyr Tyr Pro Glu 290 295 300		
	His Leu Val Asn His Tyr Phe Gln Thr Arg His Tyr Leu His Thr Leu 305 310 315 320		
	Trp Lys Ala Gly Ile Leu Tyr Lys Arg Glu Thr Thr His Ser Ala Ser 325 330 335		
	Phe Cys Gly Ser Pro Tyr Ser Trp Glu Gln Lys Leu Gln His Gly Ala 340 345 350		
	Glu Ser Phe His Gln Gln Ser Ser Gly Ile Leu Ser Arg Pro Pro Val 355 360 365		
	Gly Ser Ser Leu Gln Ser Lys His Arg Lys Ser Arg Leu Gly Leu Gln 370 375 380		
	Ser Gln Gln Gly His Leu Ala Arg Arg Gln Gln Gly Arg Ser Trp Ser 385 390 395 400		
[0035]	Ile Arg Ala Gly Ile His Pro Thr Ala Arg Arg Ser Phe Gly Val Glu 405 410 415		
	Pro Ser Gly Ser Gly His Ser Thr Asn Leu Ala Ser Lys Ser Ala Ser 420 425 430		
	Cys Leu Tyr Gln Ser Pro Val Arg Lys Ala Ala Tyr Pro Ala Val Ser 435 440 445		
	Thr Phe Glu Lys His Ser Ser Ser Gly His Ala Val Glu Leu His Asn 450 455 460		
	Leu Pro Pro Asn Ser Ala Arg Ser Gln Ser Glu Arg Pro Val Phe Pro 465 470 475 480		
	Cys Trp Trp Leu Gln Phe Arg Asn Ser Lys Pro Cys Ser Asp Tyr Cys 485 490 495		
	Leu Ser His Ile Val Asn Leu Leu Glu Asp Trp Gly Pro Cys Ala Glu 500 505 510		
	His Gly Glu His His Ile Arg Ile Pro Arg Thr Pro Ala Arg Val Thr 515 520 525		
	Gly Gly Val Phe Leu Val Asp Lys Asn Pro His Asn Thr Ala Glu Ser 530 535 540		

	Arg	Leu	Val	Val	Asp	Phe	Ser	Gln	Phe	Ser	Arg	Gly	Asn	Tyr	Arg	Val	
	545					550					555					560	
	Ser	Trp	Pro	Lys	Phe	Ala	Val	Pro	Asn	Leu	Gln	Ser	Leu	Thr	Asn	Leu	
					565					570					575		
	Leu	Ser	Ser	Asn	Leu	Ser	Trp	Leu	Ser	Leu	Asp	Val	Ser	Ala	Ala	Phe	
				580					585					590			
	Tyr	His	Leu	Pro	Leu	His	Pro	Ala	Ala	Met	Pro	His	Leu	Leu	Val	Gly	
		595						600					605				
	Ser	Ser	Gly	Leu	Ser	Arg	Tyr	Val	Ala	Arg	Leu	Ser	Ser	Asn	Ser	Arg	
	610						615					620					
	Ile	Phe	Asn	Tyr	Gln	His	Gly	Thr	Met	Gln	Asn	Leu	His	Asp	Ser	Cys	
	625					630					635					640	
	Ser	Arg	Asn	Leu	Tyr	Val	Ser	Leu	Leu	Leu	Leu	Tyr	Gln	Thr	Phe	Gly	
				645					650						655		
	Arg	Lys	Leu	His	Leu	Tyr	Ser	His	Pro	Ile	Ile	Leu	Gly	Phe	Arg	Lys	
				660					665					670			
[0036]	Ile	Pro	Met	Gly	Val	Gly	Leu	Ser	Pro	Phe	Leu	Leu	Ala	Gln	Phe	Thr	
			675					680					685				
	Ser	Ala	Ile	Cys	Ser	Val	Val	Arg	Arg	Ala	Phe	Pro	His	Cys	Leu	Ala	
	690							695				700					
	Phe	Ser	Val	Leu	Gln	Ala	Gly	Phe	Phe	Leu	Leu	Thr	Arg	Ile	Leu	Thr	
	705					710					715					720	
	Ile	Pro	Gln	Ser	Leu	Asp	Ser	Trp	Trp	Thr	Ser	Leu	Asn	Phe	Leu	Gly	
					725					730					735		
	Gly	Thr	Thr	Val	Cys	Leu	Gly	Gln	Gly	Ala	Lys	Ser	Val	Gln	His	Leu	
				740					745					750			
	Glu	Ser	Leu	Phe	Thr	Ala	Val	Thr	Asn	Phe	Leu	Leu	Ser	Leu	Gly	Ile	
		755						760					765				
	His	Leu	Asn	Pro	Asn	Lys	Thr	Lys	Arg	Trp	Gly	Tyr	Ser	Leu	His	Phe	
	770						775					780					
	Met	Gly	Tyr	Val	Ile	Gly	Cys	Tyr	Gly	Ser	Leu	Pro	Gln	Asp	His	Ile	
	785					790					795					800	
	Ile	Gln	Lys	Ile	Lys	Glu	Cys	Phe	Arg	Lys	Leu	Pro	Val	Asn	Arg	Pro	
					805					810					815		

	Ile	Asp	Trp	Lys	Val	Cys	Gln	Arg	Ile	Val	Gly	Leu	Leu	Gly	Phe	Ala	
				820					825					830			
	Ala	Pro	Phe	Thr	Gln	Cys	Gly	Tyr	Pro	Ala	Leu	Met	Pro	Leu	Tyr	Ala	
			835					840					845				
	Cys	Ile	Gln	Ser	Lys	Gln	Ala	Phe	Thr	Phe	Ser	Pro	Thr	Tyr	Lys	Ala	
		850					855					860					
	Phe	Leu	Cys	Lys	Gln	Tyr	Leu	Asn	Leu	Tyr	Pro	Val	Ala	Arg	Gln	Arg	
	865					870				875					880		
	Pro	Gly	Leu	Cys	Gln	Val	Phe	Ala	His	Ala	Thr	Pro	Thr	Gly	Trp	Gly	
					885				890						895		
	Leu	Val	Met	Gly	His	Gln	Arg	Met	Arg	Gly	Thr	Phe	Leu	Trp	Ala	Ser	
			900					905						910			
	Ala	Arg	Phe	Ser	Trp	Leu	Ser	Leu	Leu	Val	Pro	Phe	Val	Gln	Trp	Phe	
		915					920						925				
	Val	Gly	Leu	Ser	Pro	Thr	Val	Trp	Leu	Ser	Val	Ser	Arg	Lys	Tyr	Thr	
		930					935					940					
[0037]	Ser	Phe	Pro	Trp	Leu	Leu	Gly	Cys	Ala	Ala	Asn	Trp	Ile	Leu	Arg	Gly	
	945				950						955					960	
	Thr	Ser	Phe	Val	Tyr	Tyr	Pro	Ser	Ala	Leu	Asn	Pro	Tyr	His	Asp	Pro	
				965					970					975			
	Ser	Arg	Gly	Arg	Leu	Gly	Leu	Ser	Arg	Pro	Leu	Leu	Arg	Leu	Pro	Phe	
				980					985					990			
	Arg	Pro	Thr	Thr	Gly	Arg	Thr	Ser	Leu	Tyr	Ala	Asp	Ser	Pro	Ser	Val	
		995					1000						1005				
	Pro	Ser	His	Leu	Pro	Asp	Arg	Val	His	Phe	Ala	Ser	Pro	Leu	His		
		1010					1015					1020					
	Val	Ala	Trp	Arg	Pro	Pro	Tyr	Val	Leu	Leu	Ser	Ala	Gly	Ala	Leu		
		1025					1030					1035					
	Thr	Ala	Leu	Met	Leu	Ile	Ile	Phe	Leu	Met	Thr	Cys	Cys	Arg	Arg		
		1040					1045					1050					
	Val	Asn	Arg	Ser	Glu	Pro	Thr	Gln	His	Asn	Leu	Arg	Gly	Thr	Gly		
		1055					1060					1065					
	Arg	Glu	Val	Ser	Val	Thr	Pro	Gln	Ser	Gly	Lys	Ile	Ile	Ser	Ser		
		1070					1075					1080					

Trp Glu Ser His Lys Ser Gly Gly Glu Thr Arg Leu
1085 1090 1095

<210> 13
<211> 2379
<212> DNA
<213> artificial sequence

<220>
<223> nucleotide sequence encoding the mutated Y07587 polymerase

<400> 13
atgccactga gctaccagea ctttcgcagg ctgttaetgc tggatgatga agctggaccg 60
ctggaggagg agctgccacg tcctggctgat gagggactga accgtcgtgt gctgaggac 120
ctgaacctgg gcaacctgaa cgtgagcatt ccttggactc ataaggctgg aaacittacg 180
ggactttatt ctctacigt aectgtcttt aacctcatt ggaaaacacc ctcttttctt 240
aatatacatt tacaccaaga cattatcaag aaatgtgaac aattttagg accactcaca 300
gtcaatgaga aaagaagact gcaattgatt atgcttcta ggtttatcc aaatgttacc 360
aaatatctgc catggataa gggatataaa ccttattatc cagaacatct agttaatcat 420
tacttccaaa ccagacatta ttacacact ctatggaagg caggtatatt atataagaga 480
gaaacaacac atagtgcctc attttgggg tcaccatatt ctgggaaca aaagctacag 540
catggagcag aatctttcca ccagcaatcc tctgggatcc ttcccgacc accagttgga 600
[0038] tccagcttc agagcaaca ccgcaatcc agattgggac ttcaatcca acaaggacac 660
ctggccagac gccacaagg taggagctgg agcattcgag ctgggatcca cccaccgca 720
cggaggctct ttgggttga gccctcagge tcagggcatt ctacaaacct tgccagcaaa 780
tcagctcctt gcctctacca atcgccagtc aggaaggcag cctaccctgc tgtctccacc 840
tttgagaaac actcatcttc agcccatgca gtggaactcc acaaccttcc accaaactct 900
gcaagatccc agagtgagag gctgttatt cctgtctgtt ggtccagtt caggaaacgt 960
aaacctgttt ccgactactg tctctccat atgttcaat ttctcgagga ttggggaccc 1020
tgcgtgaac acggtgagca ccatattcgc atcccgagaa cgcagcacg cgtgaccgtt 1080
ggcgtgttcc tggtgataa gaaccacat aacacggtg aaagccgtct ggttgttgac 1140
tttagccagt tcagccgtg caattatcgc gttagctgga ctaagtttgc ggtgccgaat 1200
ctgcagagcc tgacgaatct gtaagcagc aatttaaget ggtgagett agacgttagc 1260
gcagccttct accacctgcc actgcacca gcagcgatgc cacacctgt ggtgggcagc 1320
agcggcttga gccgttacgt ggcacgctg agragcaaca gccgtatatt taattatcaa 1380
catggcacga tgcaaaattt acatgatagc tgtagccgta atctgtact gagcctgtta 1440
ctgttatatc agacgtttgg tcgcaagctg catatatata gccaccgat tattttaggg 1500
ttccgcaaga tcccgatggg tgttggtctg tctccattcc tgttagcgca attcaccage 1560
gcaatctgca gctgtgtgc cagagcgttt ccgatttgc tggcgtttag eggtgcgaaa 1620
agcgtgcaac acctggaaag cctgttcacg gcagtgcga acttctctgt gagcctgggc 1680


```

attcacctga accctaacaa aacaagaga tggggttact ctttacattt catgggetat 1740
gtcattggat gttatgggtc attgccacaa gatcacatca tacaaaagat caaagaatgt 1800
tttcgaaaac ttctgttaa cagacctatt gattggaaag tctgtcaacg tattgtgggt 1860
cttttgggtt ttgtgtctcc ttttacacaa tgtggttalc ctgttttaat gccittgtat 1920
gcatgtattc agtcgaagca ggccttttact ttctcgccaa cttacaagge ctttcigtgt 1980
aaacaatacc tgaaccttia cctgtttgtt cggcaaagac caggctctgtg ccaagtgttt 2040
gtcacgcaa cccctactgg ctggggattg gtcattggac atcagcgcat gcgtggaacc 2100
tttctgtcca gaaaatatac atcgtttcca tggctgctag gtgtgtctgc caactggata 2160
ctgagaggca ccagcttctg gtattatecc agcgtctcca acccttacca tgatccttct 2220
cgaggtagac tgggccttag cagacctctg ctgagactcc ccttcgacc cacaaccgga 2280
agaacaagcc tgatgcca laggccatgc gcccagcc accctctga tagagtcac 2340
tttgccagcc cactccatgt ggcctggagg cctccctaa 2379

```

<210> 14
 <211> 2820
 <212> DNA
 <213> artificial sequence

<220>
 <223> nucleotide sequence encoding the fusion between truncated core
 (1-148) and mutated Y07587 polymerase

[0039]

```

<400> 14
atggatatcg atcctacaa ggagttcgtt gccaccgtcg aactgtctag ctctctgccc 60
agcgatttct tcccaagct acgtgacctt ctatatacag cctcagctct gtatcgggaa 120
gccttagagt ctctgagca ttgttaact caccatactg ctctcaggca agcaattctg 180
tgtgtgggag aactaatgac tetagctacc tgggtgggtg gtaatttgga agatccaata 240
tcaggggacc tagtagtcag ttatgtcaac actaatatgg gactaaagtt ccgacaacta 300
ttgtggtttc acattttctt tctacatttt ggaagagaaa cagttataga atatttggtg 360
tctttcggag tgtgattcgc cactctctca gcttatagac caccaaacgc accgatactg 420
agcaccttgc cagaaaccac cgtgccactg agctaccagc acttttcag gctgttactg 480
ctggatgatg aagctggacc gtgaggag gagctgccac gtctggetga tgagggaactg 540
aaccgtctgt tggctgagga cctgaacctg ggcaacctga acgtgagcat tctttggact 600
cataaggttg gaaactttac gggactttat tctttactg tactgtctt taaccttcac 660
tggaanaaac cctcttttcc taatatacat ttacaccaag acattatcaa gaaatgtgaa 720
caattttagt gaccactcac agtcanttag aanaagagac tgcanttgat tatgcctgtt 780
aggttttate caaatgttac caaatatttg cacttgata aggttattaa acctatttat 840
ccagaacatc tagttaatca ttacttccaa accagacatt atttacacac tctatggaag 900
gcaggatatat tatataagag agaaacaaca catagtgcct cattttgtgg gtcaccatat 960
tcttgggaac aaaagctaca gcatggagca gaatctttcc accagcaatc ctctgggatt 1020

```

	ctttcccgac caccagttgg atccagcctt cagagcaaac accgcaaate cagattggga	1080
	cttcaatccc aacaaggaca ctggccaga cgccaacaag giaggagctg gagcattcga	1140
	gttgggattc aceccaccgc acggagggtt ttggggigg agccctcagg ctcagggcct	1200
	tctacaaacc ttgccageaa atcagcctcc tgcctctacc aatcgceagt caggaaggca	1260
	gcciacctcg ctgtctccac ctctgagaaa cactcactct caggccatgc agtggaaactc	1320
	cacaaccctc caccaaaactc tgcaagatcc cagagigaga ggctgtatt tceetgctgg	1380
	tggctccagt tcaggaacag taaacctgt tccgactact gtctctccca tategtcaat	1440
	cttctcgagg attggggacc ctgcctgaa cacggtgagc accatattcg catcccgaga	1500
	acgccagcac gcgtgaccgg tggcgtgttc ctggtggata agaaccaca taacacggct	1560
	gaaagccgtc tggttgttga ctttagccag ttcagccgtg gcaattatcg cgttagctgg	1620
	cctaagtttg cgggtgccaa tctgcagagc ctgacgaatc tgtaagcag caatttaagc	1680
	tggttagctt tagacgttag cgcagccttc taccacctgc cactgcaccc agcagcgatg	1740
	ccacaccctc tggtagggcag cagcggctcg agccgttacg tggcacgctt gacagcaac	1800
	agccgtatct ttaattatca acatggcagc atgcaaaatt tacatgatag ctgtagccgt	1860
	aatctgtacg tgagcctggt actgttatat cagacgtttg gtgcgaagct gcatttatac	1920
	agccaccgca ttattttagg gttecgcaag atcccgatgg gtgttggtct gtctccattc	1980
	ctgttagcgc aattcaccag cgaatctgc agcgttgtgc gcagagcgtt tccgcattgc	2040
[0040]	ctggcgltta gcgtgtcgaa aagcglgcaa caccctggaaa gcctglitac ggcagtgacg	2100
	aaattctctc tgagcctggg cattcacctg aaccttaaca aaacaaagag atggggttac	2160
	tctttacatt tcatgggcta tgtcattgga tgttatgggt cattgccaca agatcacatc	2220
	atacaaaaga tcaaagaatg ttttcgaaaa ctctctgtta acagacctat tgattggaaa	2280
	gtctgtcaac gtattgtggg tcttttgggt ttgtctgtc ctittacaca atgtggitat	2340
	ctgtctttaa tgcctttgta tgcattgatt cagtcgaagc aggcctttac tttctcgcca	2400
	acttacaagg cctttctgtg taaacaatac ctgaaccttt accctgttgc tcggcaaaaga	2460
	ccaggctctg gccaaagtgt tgcctacgca acccctactg gctggggatt ggtcatggga	2520
	catacagcga tgcgtggaa ctctctgtcc agaaaatata catcglttcc atggctgcta	2580
	ggctgtgctg ccaactggat actgagagcc accagcttcg tgtattatcc cagcgtcttc	2640
	aacctttacc atgatecttc tcgaggtaga ctgggcctga gcagacctct gctgagactc	2700
	cccttccgac ccacaaccgg aagaacaagc ctgtaigccg atagccctag cgtcccccgc	2760
	cacctccctg atagagtcca ttttgcacgc ccactccatg tggcctggag gcctccctaa	2820
<210>	15	
<211>	3024	
<212>	DNA	
<213>	artificial sequence	
<220>		
<223>	nucleotide sequence encoding the fusion between truncated core (1-148)-mutated Y07587 polymerase-env1 and env2	

[0041]	<400> 15	
	atggatatcg atocctacaa ggagttcggg gccaccgtcg aactgctgag ctctctgccc	60
	agcgatttct tccaaagcgt acgtgacctt ctagatacag cctcagctct gtatcgggaa	120
	gccttagagt ctcttgagca ttgttcacct caaccatactg ctctcaggca agcaattctg	180
	tgttggggag aactaatgac tctagctacc tgggtgggtg gtaatttga agatecaata	240
	tccagggacc tagtagtcag ttatgtcaac actaataagg gactaaagtt ccgacaacta	300
	ttgtgttttc acatttcttg tctcactttt ggaagagaaa cagttataga atatttgggtg	360
	tccttcggag tgiggattcg cactctccca gcttatagac caccaaacgc accgatactg	420
	agcaccctgc cagaaaccac cgtgccactg agctaccagc actttcgcag gctgttactg	480
	ctggatgatg aagctggacc gctggaggag gagctgccac gtctggctga tgagggactg	540
	aaacctgctg tggttgagga cctgaacctg ggcaacctga acgtgagcat tccttggact	600
	cataaggtgg gaaaatttac gggactttat tcttctactg tacctgtctt taacctctat	660
	tggaaaacac cctcttttcc taatatatat ttacaccaag acattatcaa gaaatgtgaa	720
	caattttagt gaccactcac agtcaatgag aaaagaagac tgcaattgat tatgcctgct	780
	aggttttata caaatgttac caaatatttg ccatttgata aggttatata acctatttat	840
	ccagaacatc tagttaatca ttacttccaa accagacatt atttacacac tctatggaag	900
	gcaggtatat tatataagag agaaacaaca catagtgctt catttttggtg gtcaccatat	960
	tcttgggaac aaaagctaca gcatggagca gaattcttcc accagcaatc ctctgggatt	1020
	cttctccgac caccagtigg atccagcctt cagagcaaac accgcaaatc cagattggga	1080
	cticaatccc aacaaggaca cctggccaga cgccaacaag gtaggagctg gagcattcga	1140
	gctgggattc accccaccgc accgaggtct ttgggggtgg agccctcagg ctccaggcat	1200
	tctacaaacc ttgccagcaa atcagcctcc tgcctctacc aatgccagc caggaaggca	1260
	gcctaccctg ctgtctccac ctctgagaaa cactcactct caggccatgc agtggaaactc	1320
	cacaaccttc caccaaactc tgcaagatcc cagagtgaga ggctgtatt tccctgctgg	1380
	tggtctcagt tcaggaacag taaacctgt tccgactact gtctctccca tatctcaat	1440
	cttctcgagg attggggacc ctgcctgaa caeggtgagc accatattcg catcccgaga	1500
	acgccagcac gcgtgaccgg tggcgtgttc ctggtggata agaaccacaa taacacggt	1560
	gaaagccgtc tggttgttga cttagccag ttcagccgtg gcaattatcg cgttagctgg	1620
	cctaagtitt cggtgccgaa tctgcagagc ctgacgaate tgtaagcag caatttaage	1680
	tggttgagct tagacgttag cgcagccttc taccacctgc cactgcaccc agcagcgatg	1740
	ccacacctgc tgggtggcag cagcgtctg agccgttacg tggcacgctt gagcgcaac	1800
	agccgtatat ttaattatca acatggcacg atgcaaaatt tacatgatag ctgtagccgt	1860
	aatctgtacg tgagcctgtt actgttatat cagacgtttg gtgcgaagct gcatttatac	1920
	agccaccgca ttatttttagg gtcccgcaag atcccgatgg gtgttggtct gtctecattc	1980
	ctgttagcgc aattcaccag cgcaatctgc agcgttgtgc gcagagcgtt tccgcattgc	2040

```

ctggcggttta gcgtcctgca agcaggcttc ttectgctga cccgtattct gaccattcca 2100
caaagcctgg atagctggtg gaccagcctg aacttctctg gtggcaccac gggttgccctg 2160
ggtcagggtg cgaagaagct gcaacacctg gaaagcctgt tcacggcagt gacgaacttc 2220
ctgtcgagcc tgggcattca ctgaacctt aacaaaacaa agagatgggg ttactcttta 2280
catttcatgg gctatgtcat tggatgitat gggtcattgc cacaagatca catcatacaa 2340
aagatcaaa g aatgttttcg aaaacttctt gtaacagac ctattgattg gaaagtctgt 2400
caacgtattg tgggtctttt gggttttgct gctcctttta cacaatgtgg ttatcctgct 2460
ttaatgcctt tgtatgcatt tattcagtcg aagcaggctt ttactttctc gccaaacttac 2520
aaggcctttc tgtgtaaaca atacctgaac ctttaccctg ttgctcgcca aagaccaggt 2580
ctgtgccaa g tgtttgctca cgcaacctt aciggtctgg gattggctat gggacatcag 2640
cgcatgcgig gaacctttct ggggcaagc gcacgttla gclggclgag ctgclggll 2700
ccgttctgac aatgtttgtt gggctcgagc ccaaccgtgt ggctgagcgt gtcagaaaa 2760
tatacatctt ttccatggct gctaggtctg gctgccaact ggatactgag aggcaccagc 2820
ttctgtatt atccagegc tctcaacct taccatgac ctctcagg tagactgggc 2880
ctgagcagac ctctgctgag actcccttc cgaccacaa ccggaagaac aagcctgtat 2940
gccgatagcc ctagcgtcc cagccacctt cctgatagag tccattttgc cagccactc 3000
catgtggcct ggaggcctcc ctaa 3024

```

[0042]

```

<210> 16
<211> 3723
<212> DNA
<213> artificial sequence

<220>
<223> nucleotide sequence encoding core (1-183) IRES and fusion
between mutated Y07587 polymerase-env1-env2

<400> 16
atggatatg atccctacaa ggagttcgtt gccaccgtcg aactgctgag ctccctgccc 60
agegatttct tcccaagcgt acgtgacctt ctgatacag cctcagctct gtatcgggaa 120
gccttagagt ctctcgagca ttgttaacct caccatactg ctctcaggca agcaattctg 180
tgctggggag aactaatgac tctagctacc tgggtgggtg gtaatttga agatecaata 240
tecagggacc tagtagtcag ttatgtcaac actaatatgg gactaaagt cgcacaacta 300
ttgtggttgc acatttcttg tctactttt ggaagagaaa cagttataga atatttggtg 360
tctttcggag tgtggattcg cactcteca gcttatagac caccaaacgc accgatactg 420
agcaccctgc cagaaaccac cgtggtgctg cgtctgggtc gcagccncg caggcgtacc 480
ccaagccac gtcgtcgag aagccagagc ccacgacgtc gcaggagcca gagecgtgag 540
agccagtgtt aaaagcttga tatcgaattc acgcgtcccc cccctaact tactggccga 600
agccgttttg aataaggcgg gtgtgcgttt gtctatatgt tattttccac catattgccc 660
tcttttgcca atgtgagggc ccgaaacct ggcctgtct tcttgacgag caittctagg 720

```

ggcttttccc cttctgccaa aggaatgcaa ggtctgttga atgtcgtgaa ggaagcagtt	780
ccctctggaag cttcttgaag acaacaacg tctgtagcga ccctttgcag gcagcggaac	840
cccccaactg ggcacaggtg cctctgcggc caaaagccac gtgtataaga tacacctgca	900
aaggcggcac aaccccagtg ccacgttgtg agttggatag ttgtggaaag agtcaaatgg	960
ctctctecaa gcgtattcaa caaggggctg aaggatgccc agaaggtaac ccattgtatg	1020
ggatctgate tggggccctg gtgcacatgc ttacatgtg tttagtcgag gttaaaaaac	1080
gtctagggccc ccgaaccac ggggacgtgg ttttctttg aaaaacacga tgataatacc	1140
atgccactga gtaaccagea cttegcagg ctgttactgc tggatgatga agctggaccg	1200
ctggaggagg agctgccacg tetggctgat gagggaacta accgtcgtgt ggctgaggac	1260
ctgaacctgg gcaacctgaa cgtgagcatt ccttggaact ataagggtgg aaactttacg	1320
ggactttatt etttctactgt accgtctttt aaccttcatt ggaacacace ctcttttctt	1380
aatafacatt tacaccaaga cattatcaag aaatgtgaac aatttgtagg accactcaca	1440
gtcaatgaga aaagaagact gcaattgatt atgcctgcia ggittttatcc aaatgttacc	1500
aaataattgc catiggataa gggattataa ccttattatc cagaacatct agttaatcat	1560
tacttccaaa ccagacatta ttacacact ctatggaagg caggtatatt atataagaga	1620
gaaacaacac atagtgcctc attttgtggg icaccatatt ctgggaaca aaagctacag	1680
catggagcag aatcttttca ccagcaatcc tetgggattc ttcccgacc accagttgga	1740
tccagcttc agagcaacaa ccgcaatcc agattgggac ttcaatccca acaaggacac	1800
ctggccagac gccancaagg taggagctgg agcattcgag ctgggattca ccccaccga	1860
cggaggtctt ttggggtgga gccctcagge tcagggcatt ctacaaacct tgccagcaaa	1920
tcagctcctt gcctctacca atcgccagtc aggaaggcag cctaccctgc tgtctccacc	1980
tttgagaaac actcactctc aggccatgca gtggaactcc acaaccttcc accaaactct	2040
gcaagatccc agagtgaag gccgttatct cctgtctggt ggtccagtt caggaacagt	2100
aaacctgtt ccgactactg tctctcccat atcgtaact ttctcgagga ttggggaccc	2160
tgcgtgaaac acgttgagca ccatattgc atcccagaa cgcagcacg cgtgaccgt	2220
ggcgtgttcc tggtgataa gaaccacat aacacggctg aaagccgtct ggttgttgac	2280
tttagccagt tcagccgtgg caattatgc gttagctggc etaagtttgc ggtgccgaat	2340
ctgcagacc tgacgaatct gtttagcgc aatttaagct ggtgagctt agacgtagc	2400
gcagcttct accacctgcc actgcacca gcagcgatgc cacacctgct ggtgggcagc	2460
agcggcttga gccgttactg ggcacgctg agcagcaaca gccgtatatt taattatcaa	2520
catggcacga tgcaaaattt acatgatagc ttagccgta atctgtactg gacctgtta	2580
ctgttatate agacgtttg tcgaagctg catttatata gccacccgat tattttaggg	2640
ttccgeaaga tcccgatggg tttggctg tetccattcc tgttagcga attcaccage	2700
gcaatctgea gcgttgtgc cagagcgttt ccgcatggc tggcgtttag cgtccigcaa	2760

[0043]

```

gcaggcttct tctgctgac ccgtattctg accattccac aaagccctgga tagctggtgg 2820
accagcctga acttccctggg tggcaccacg gtttgcttgg gtcagggtgc gaaaagcgtg 2880
caacacctgg aaagcctgtt cagggcagtg acgaacttcc tgcctgagcct gggcattcac 2940
ctgaacccta acaaaacaaa gagatggggt tactctttac atttcatggg ctatgtcaat 3000
ggatgttatg ggtcattgcc acaagatcac atcatacaaa agatcaaaaga atgttttcga 3060
aaacttcttg ttaacagacc tattgattgg aaagtcctgc aacgtattgt gggctttttg 3120
ggttttgctg ctctctttac acaatgtgtt tatectgctt taatgccttt gtatgcatgt 3180
attcagtcga agrcagcttt tactttctcg ccaacttaca aggcctttct gtgtaacaa 3240
tacctgaacc tttaacctgt tgetcggcaa agaccaggtc tgtgccaagt gtttctcac 3300
gcaaccctta ctggctgggg atgggtcatg ggacatcagc gcatgcgtgg aacctttctg 3360
tgggcaagcg cagcctttag ctgctgagc ctgctggttc cgttcgtgca atggtttgig 3420
ggtctgagcc caacctgtgt gctgagcgtg tccagaaaat atacatcgtt tccatggctg 3480
ctaggtctgt ctgccaactg gatactgaga ggcaccagct tcgigtatta tcccagcgt 3540
ctcaaccctt aecatgatcc ttctcaggt agactgggcc tgagcagacc tctgctgaga 3600
ctccccttcc gaccacaaac cggaagaaca agcctgtatg ccgatagccc tagcgtcccc 3660
agccacctcc ctgatagagt ccattttgcc agcccactcc atgtggcctg gaggcctccc 3720
taa 3723

```

[0044]

```

<210> 17
<211> 3723
<212> DNA
<213> artificial sequence

<220>
<223> nucleotide sequence encoding fusion mutated Y07587
polymerase-env1-env2 IRES core (1-183)

```

```

<400> 17
atgccactga gctaccagca ctttcgagg ctgttactgc tggatgatga agctggaccg 60
ctggaggagg agctgccacg tctggctgat gagggactga accgtcgtgt ggctgaggac 120
ctgaacctgg gcaacctgaa cgtgagcatt ccttggactc ataaggtggg aaactttacg 180
ggactttatt ctctactgt acctgtcttt aacctctatt ggaanacacc ctcttttctt 240
aatatacatt tacaccaaga cattatcaag aaatgtgaac aatttgtagg accactcaca 300
gtcaatgaga aaagaagact gcaattgatt atgcctgcta ggttttatcc aaatgttacc 360
aaatatattg cattggataa gggatataaa ccttattatc cagaacatct agttaatcat 420
tacttccaaa ccagacatta ttacacact ctatggaagg caggtatatt atataagaga 480
gaaacaacac atagtgcctc attttgggg tcacatatt ctgggaaca aaagctacag 540
catggagcag aatctttcca ccagcaatcc tctgggatc ttcccgcacc accagttgga 600
tccagccttc agagcaaaac ccgcaaatcc agattgggac ttcaatccca acaaggacac 660
ctggccagac gccacaaggt taggagctgg agcattcgag ctgggattca cccaccgca 720

```

	cggagggtctt ttgggggtgga gccctcaggc tcagggcatt ctacaaacct tgccagcaaa	780
	tcagccctctt gctctlacca ategccagtc aggaaggcag cctacctgc tctctccacc	840
	tttgagaaac actcactctc aggecatgca gtggaaactec acaaccttcc accaaactct	900
	gcaagatccc agagttagag gctgtatatt cctgtctgtt ggtctcagtt caggaacagt	960
	aaacctgttt ccgactactg tctctcccat atcgtaaat tctctagaga ttggggaccc	1020
	tgcgtgaac accgttagca ccatattgc atcccgagaa cgcagcagc cgtgaccgtt	1080
	ggcgtgttcc tgggtgataa gaaccacat aacacggctg aaagccgtct ggttgtgac	1140
	tttagccagt tcagccgttg caattatgc gttagctggc cttaagttgc ggtgcccgaat	1200
	ctgcagagcc tgacgaatct gttaagcagc aatttaagct ggtgagctt agacgttagc	1260
	gcagccttct accacctgcc actgcaccac gcagcagatgc cacacctgtt ggtgggcagc	1320
	agcggcttga gccgttaagt ggcacgcctg agcagcaaca gccgtatatt taattatcaa	1380
	catggcacga tgcaaaattt acatgatagc ttagccgta atctgtactg gagccigtta	1440
	ctgttatatc agacgtttgg tcgcaagctg catttataca gccaccgat tattttaggg	1500
	ttccgcaaga tcccgatggg tgttggtctg tctccattcc tgttagcgca attcaccagc	1560
	gcaatctgca gcgttgtgcg cagagcgttt ccgcattgcc tggcgtttag cgtctctcaa	1620
	gcagccttct tctgtctgac ccgtattctg accattccac aaagcctgga tagctgtgg	1680
	accagcctga acttctctgg ttgcaccacg gtttgcctgg gtcagggtgc gaaaagcgtg	1740
[0045]	caacacctgg aaagcctgtt cagcgacgtg acgaacttcc tctgagcct gggaattcac	1800
	ctgaacccta acaaaacaaa gagatgggtt tactctttac attcatggg ctatgtcatt	1860
	ggatgttatg ggtcattgcc acaagatcac atcatacaaa agatcaaaga atgttttcca	1920
	aaacttctct ttaacagacc tattgattgg aaagtctgtc aacgtattgt gggctttttg	1980
	ggttttctgt ctccttttac acaatgttgt tatcctgctt taatgccttt gtatgcatgt	2040
	attcagtega agcagctttt tactttctcg ccaacttaca aggcctttct gtgtaaacaa	2100
	tacctgaacc ttacctctgt tctctggcaa agaccaggtc tgtccaagt gttgtctac	2160
	gcaaccctca ctggttggg attgtctatg ggacatcagc gcattgcgtg aacctttctg	2220
	tgggcaagcg cacgttttag ctggtcagac ctgctgttcc cgttctgtca atgttttctg	2280
	ggtctgagcc caaccgtgtg gctgagcgtg tccagaaaat atacatcgtt tccatggctg	2340
	ctaggctgtg ctgcacactg gatactgaga ggcaccagct tctgtatta tccagcgtt	2400
	ctcaaccttt accatgatcc ttctcgaggt agactgggcc tgagcagacc tctgtgaga	2460
	ctccccttcc gaaccacac ccgaagaaca agcctgtatg ccgatagccc tagcgtcccc	2520
	agccacctcc ctgatagagt ccattttgcc agcccactcc atgtggcctg gaggcctccc	2580
	taaaagcttg atatgaatt cagcgttccc cccctaaag ttaactggccg aagcctgtt	2640
	gaataaggcc ggtgtgctgt tgtctatatg ttattttcca ccatattgcc gtcttttggc	2700
	aatgtgaggc cccggaaccc tggccctgtc ttcttgacga gcaatcttag ggtcttttcc	2760
	ctctctgcca aaggaatgca aggtctgtg aatgtctgta aggaagcagt tctcttgaa	2820

```

gtttcttgaa gacaaacaac gictgiagcg accctttgca ggcagcggaa cccccacct 2880
ggcgacaggt gcctetgcgg ccaaaagcca cgtgtataag atacacctgc aaaggcggca 2940
caacccceagt gccacgttgt gagttggata gtgttggaan gagtcaaatg getctectca 3000
agcgtatcca acaaggggct gaaggatgcc cagaaggtac ccatttgtat gggatctgat 3060
ctggggcctc ggtgcacatg ctttacatgt gtttagtcga ggttaaaaaa cgtctaggcc 3120
ccccgaacca cggggagctg gttttccttt gaaaaacacg atgataatac catggataac 3180
gatccctaca aggagttcgg tgccaccgtc gaactgctga gcttctgcc cagcgatttc 3240
ttccaagcg tacgtgacct tctagataca gccctagctc tgtatcggga agccttagag 3300
tctctgagc attgttacc tcaccatact gctctcagc aagcaattct gtgctgggga 3360
gaactaatga ctctagctac ctgggigggt ggttaatttg aagatccaat atccagggac 3420
ctagtagtca gtatgtcaa cactaataat ggactaaagt tcgacaact attgtggttt 3480
cacattcttt gtctcacttt tsgaagagaa acagttaga aatatttgg gtcttctgga 3540
gtgtggattc gcactctctc agcttataga ccacaaacg caccgatact gagcacctg 3600
ccagaaacca ccgtggtcgc tegtctgtgt cgcagccac gcagcgtac cccaagecca 3660
cgtcgtcga gaagccagag cccacgaagt cgcaggagcc agagcgtga gagccagtgc 3720
taa 3723

```

[0046]

```

<210> 18
<211> 10
<212> PRT
<213> artificial sequence

<220>
<223> FLP peptide (core 18-27)

<400> 18
Phe Leu Pro Ser Asp Phe Phe Pro Ser Val
1             5             10

<210> 19
<211> 9
<212> PRT
<213> artificial sequence

<220>
<223> ILC peptide (core 99-108)

<400> 19
Ile Leu Cys Trp Gly Glu Leu Met Thr
1             5

<210> 20
<211> 9
<212> PRT
<213> artificial sequence

<220>
<223> VLQ peptide (env1)

```


<400> 20

Val Leu Gln Ala Gly Phe Phe Leu Leu
1 5

<210> 21

<211> 9

<212> PRT

<213> artificial sequence

<220>

<223> FLG peptide (env1 41-49)

<400> 21

Phe Leu Gly Gly Thr Thr Val Cys Leu
1 5

<210> 22

<211> 10

<212> PRT

<213> artificial sequence

<220>

<223> GLS peptide (env2)

<400> 22

Gly Leu Ser Pro Thr Val Trp Leu Ser Val
1 5 10

[0047]

<210> 23

<211> 9

<212> PRT

<213> artificial sequence

<220>

<223> SLY peptide (polymerase 816-824)

<400> 23

Ser Leu Tyr Ala Asp Ser Pro Ser Val
1 5

<210> 24

<211> 10

<212> PRT

<213> artificial sequence

<220>

<223> VSA peptide (polymerase)

<400> 24

Val Ser Ala Ala Phe Tyr His Leu Pro Leu
1 5 10

<210> 25

<211> 15

<212> PRT

<213> artificial sequence

<220>

[0048]

<223> N13F peptide (polymerase)

<400> 25

Asn	Leu	Asn	Val	Ser	Ile	Pro	Trp	Thr	His	Lys	Val	Gly	Asn	Phe
1				5					10				15	

<210> 26

<211> 15

<212> PRT

<213> artificial sequence

<220>

<223> F13L peptide (env)

<400> 26

Phe	Leu	Trp	Glu	Trp	Ala	Ser	Ala	Arg	Phe	Ser	Trp	Leu	Ser	Leu
1				5					10				15	

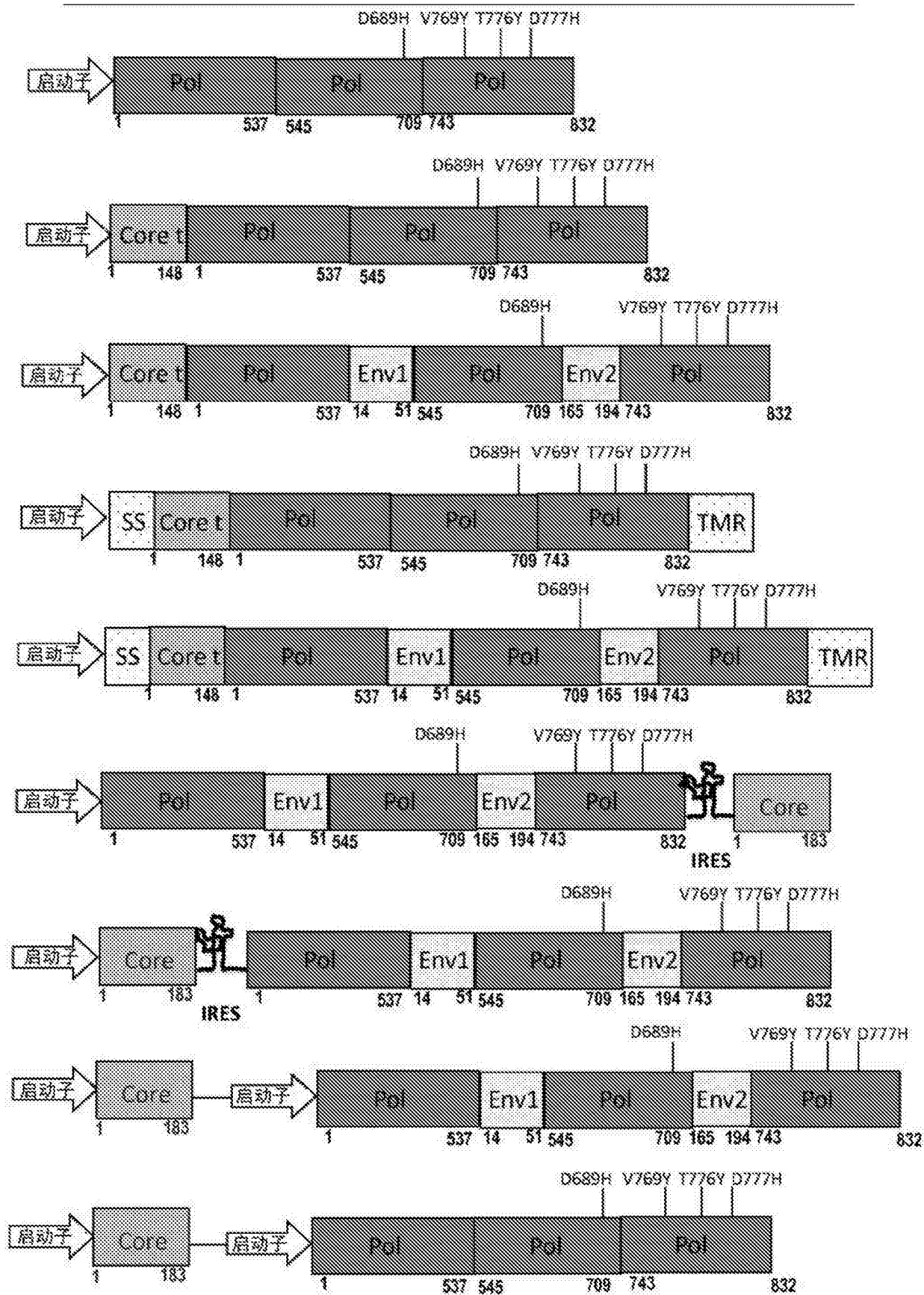


图1

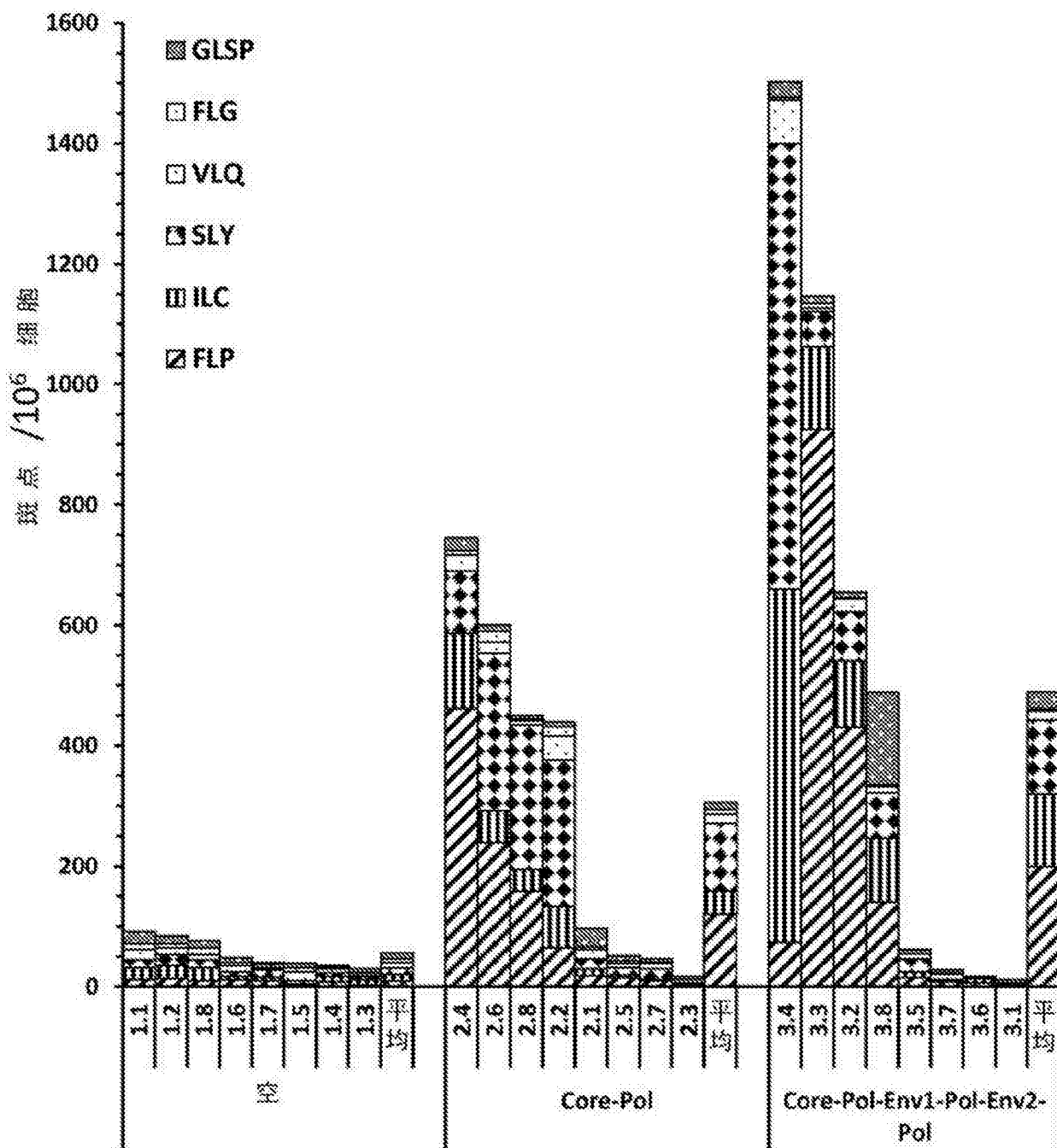


图2

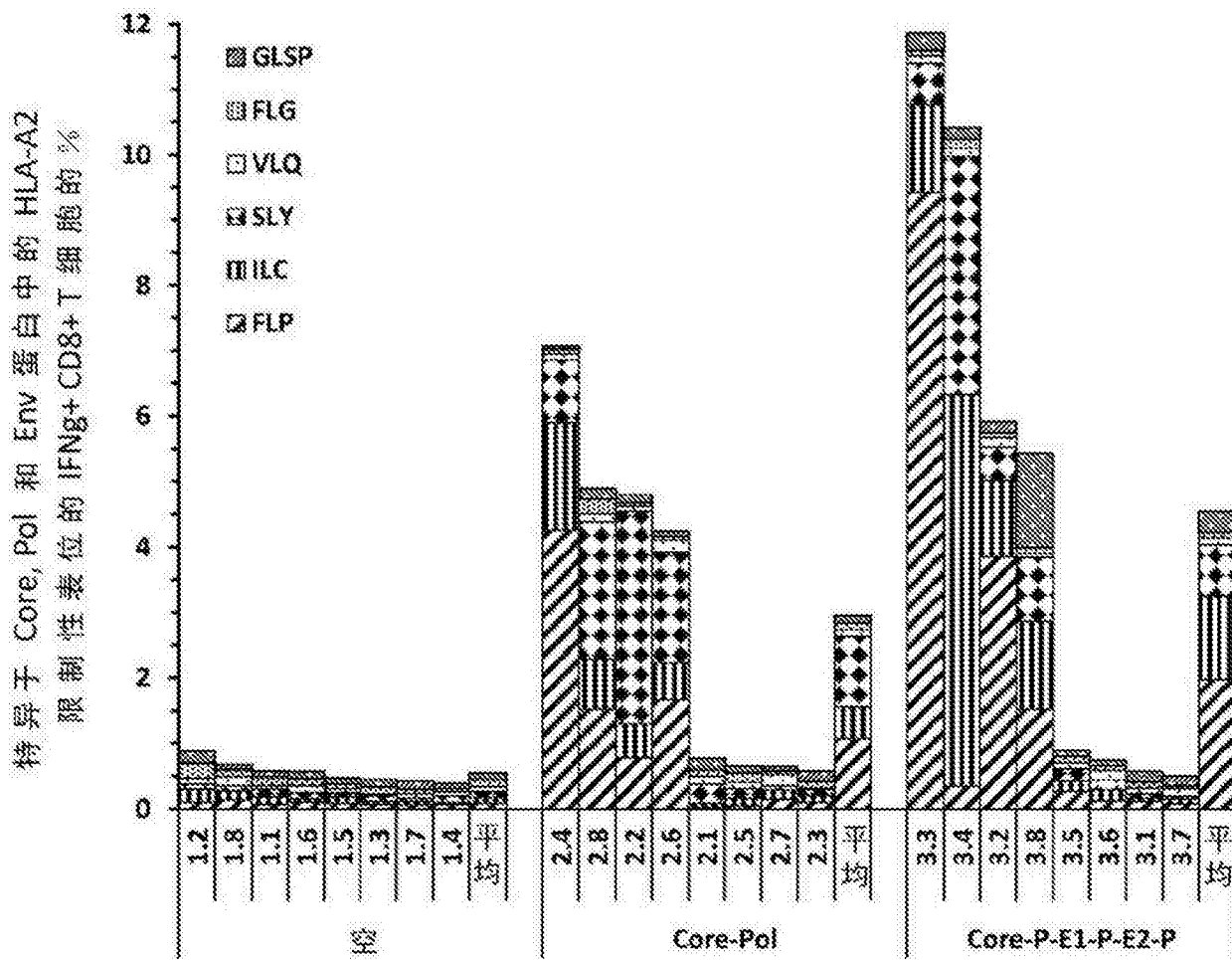


图3A

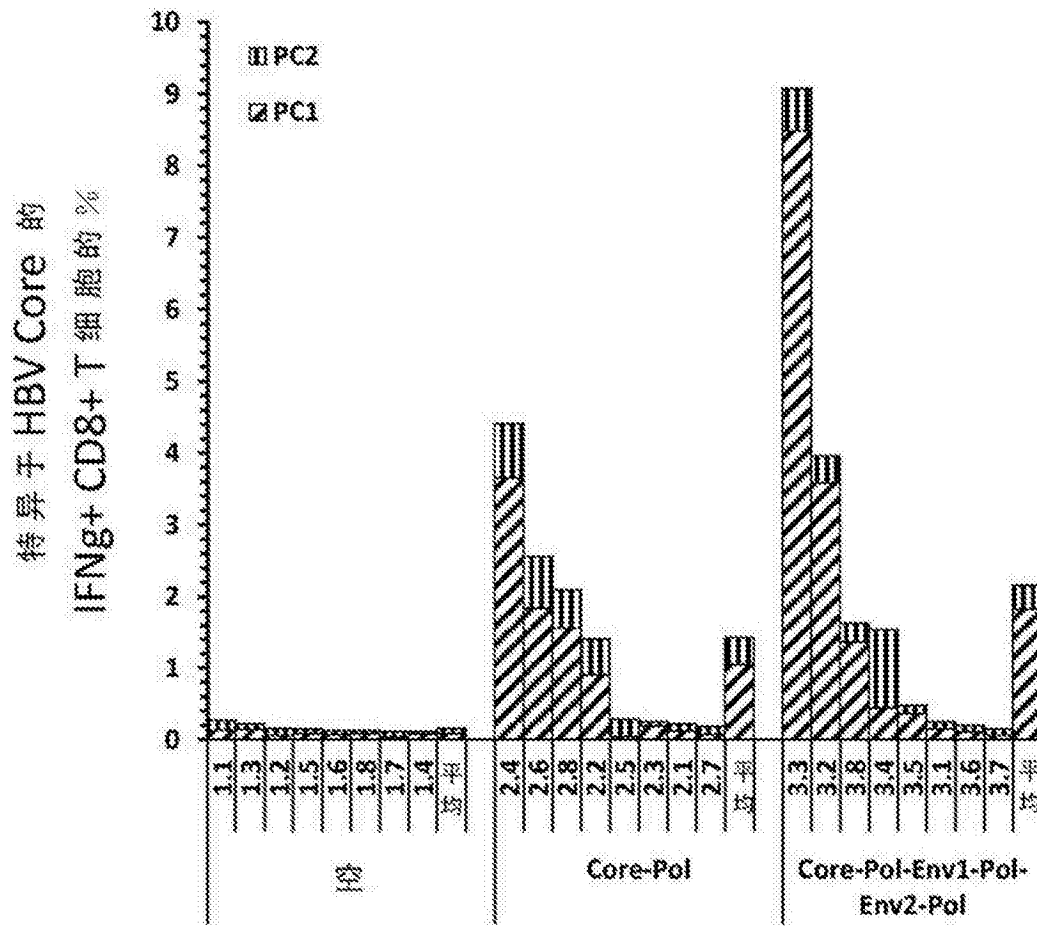


图3B

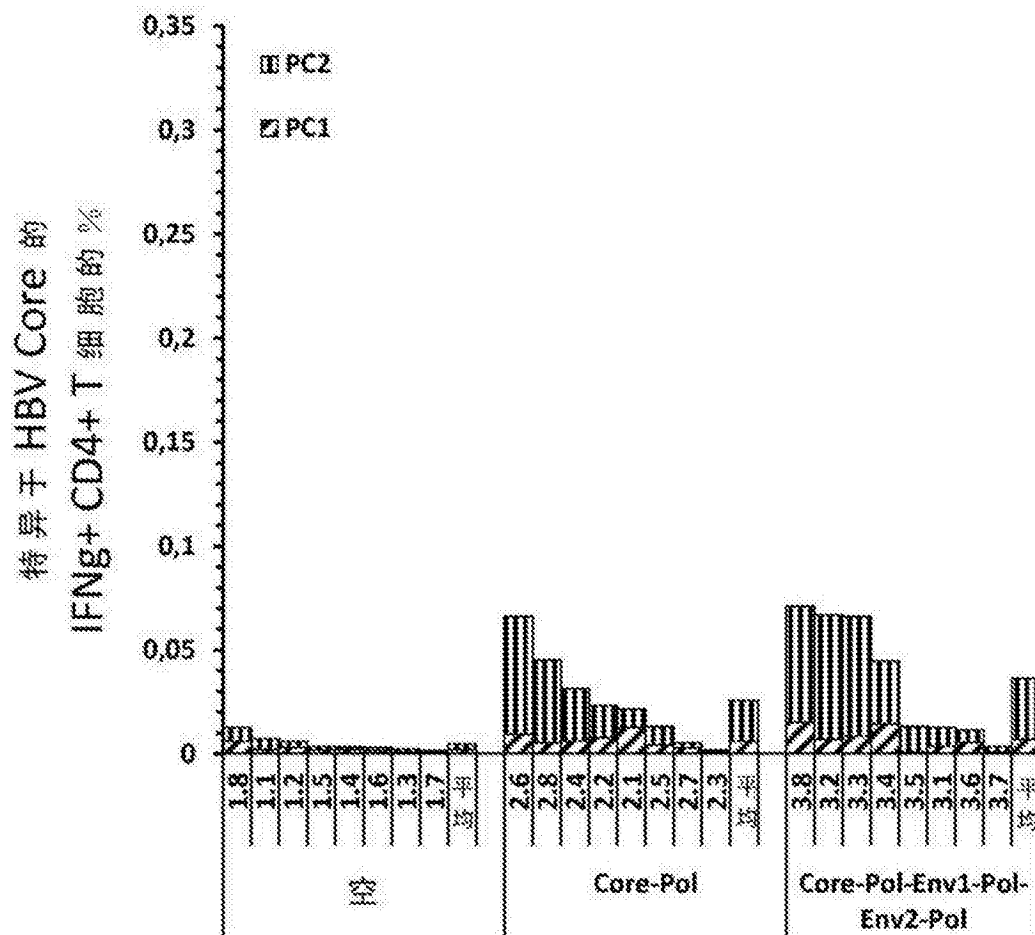


图3C

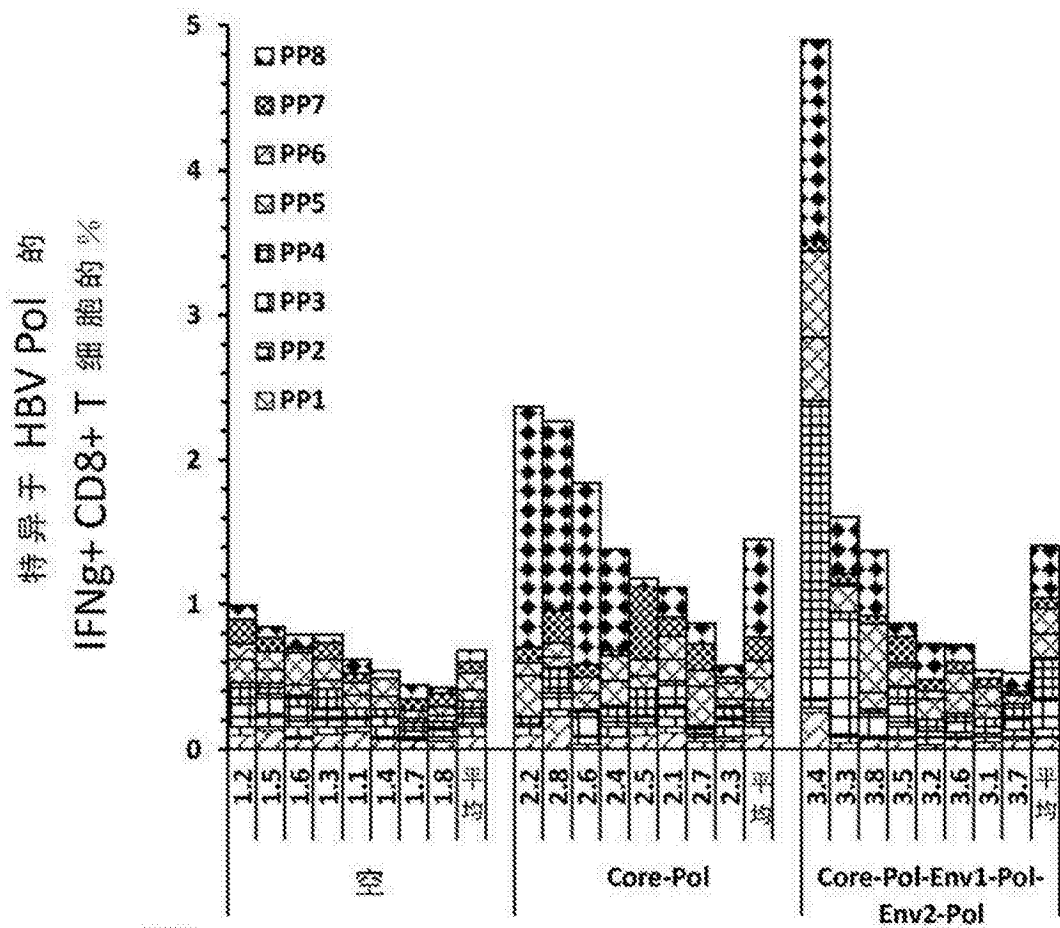


图3D

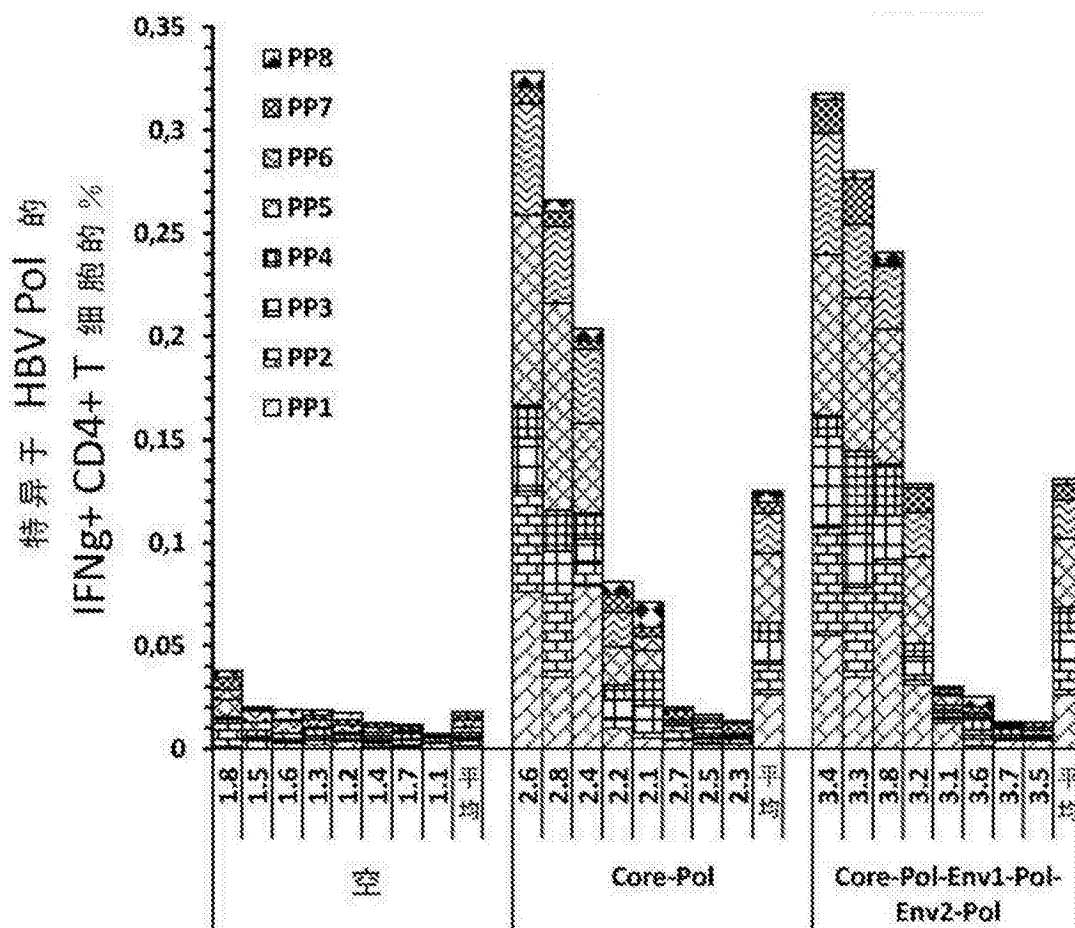


图3E

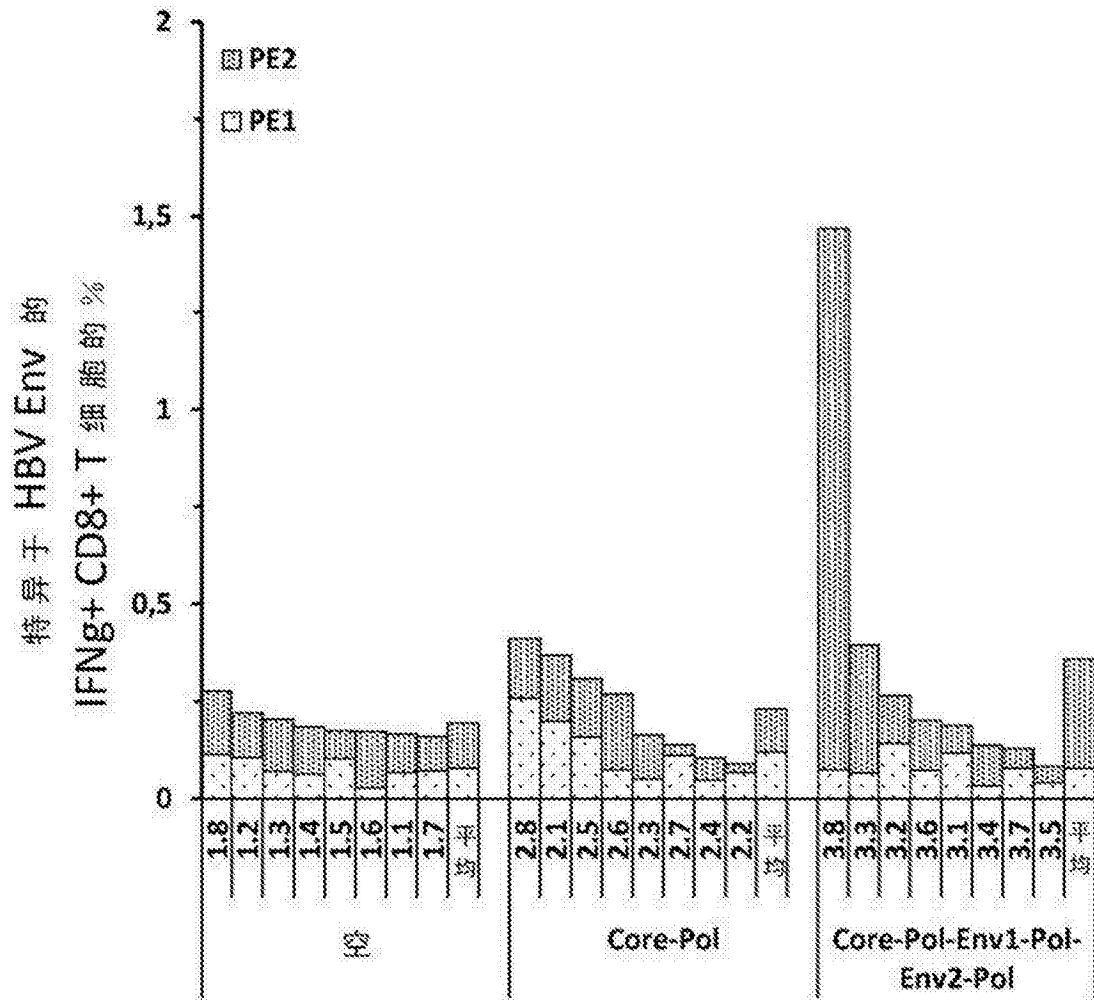


图3F

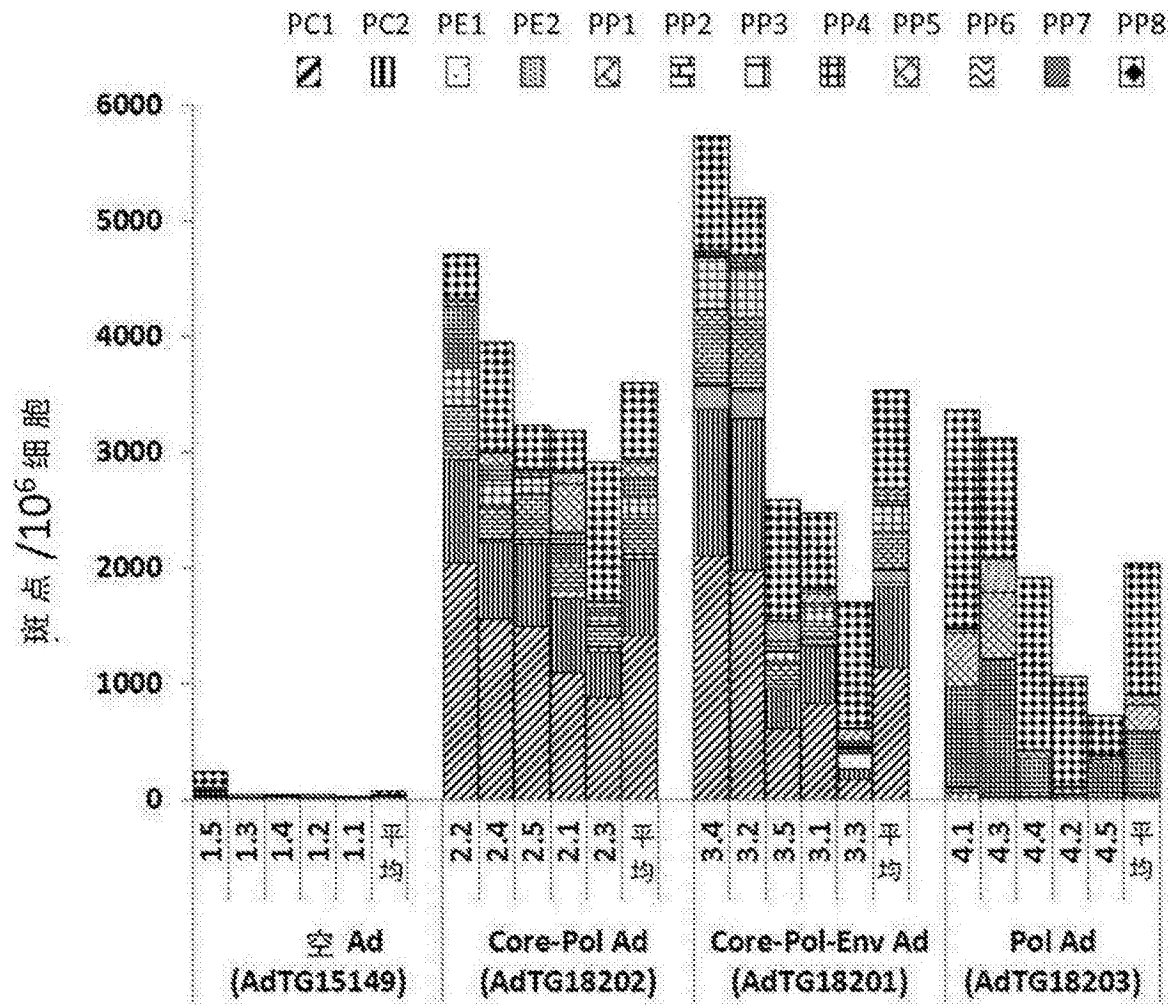


图4

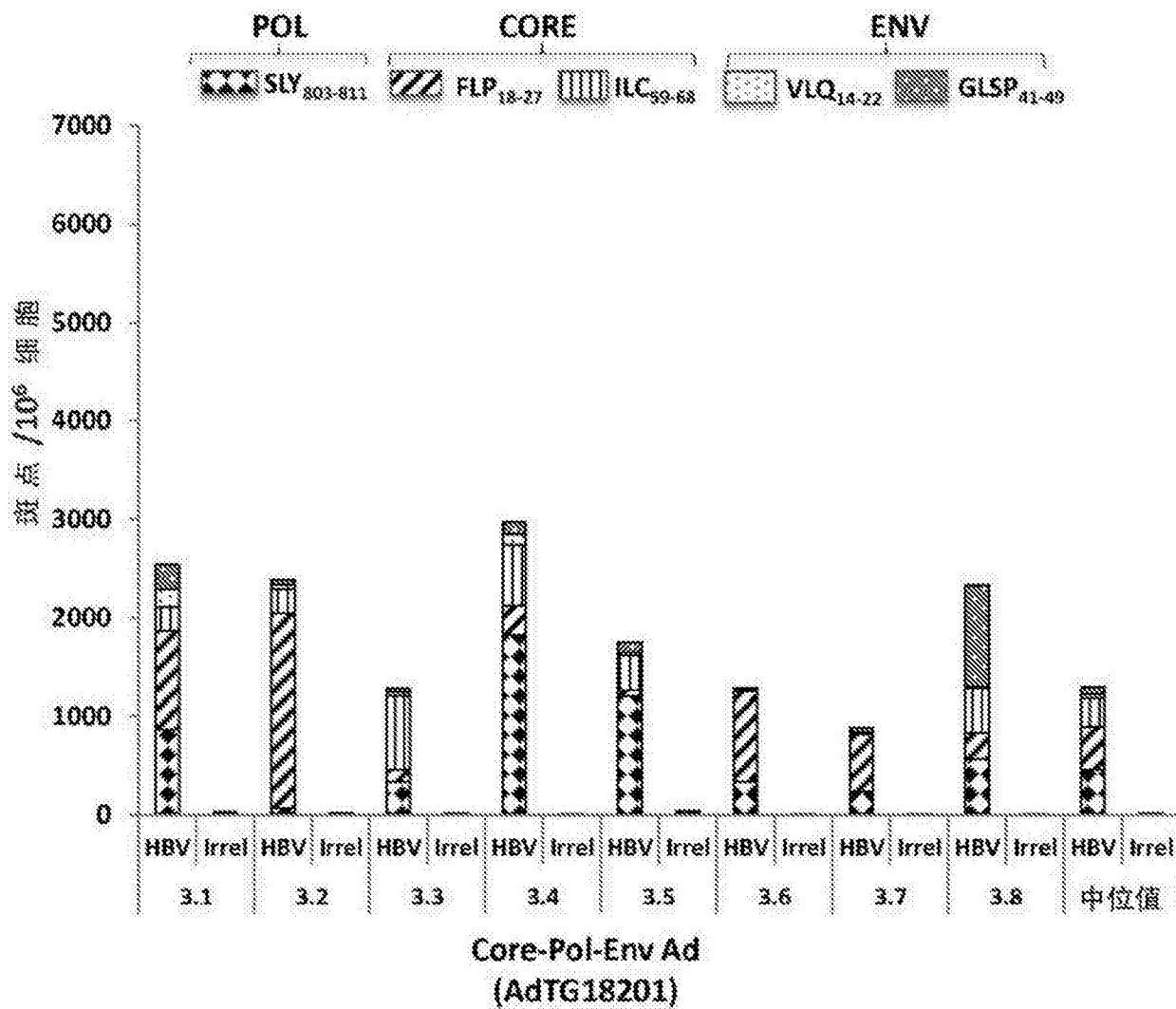


图5

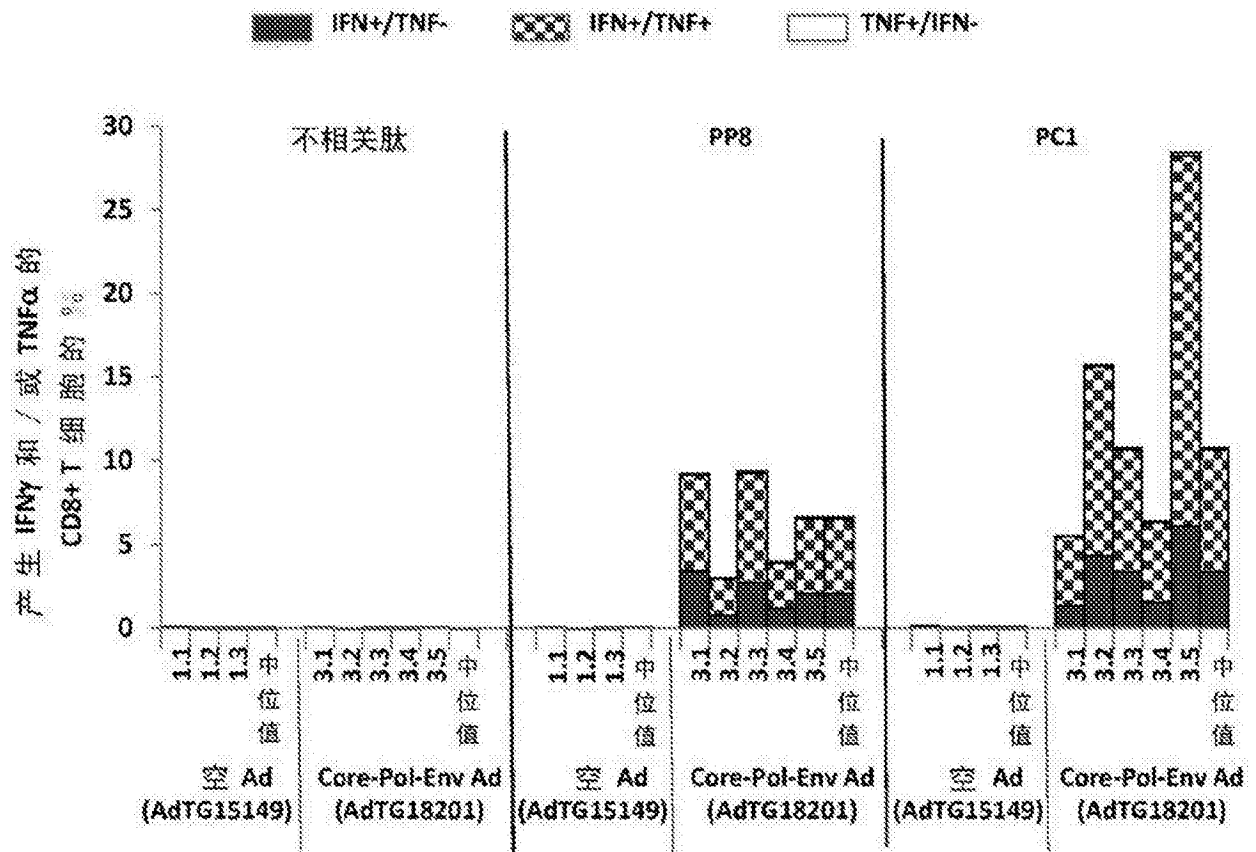


图6

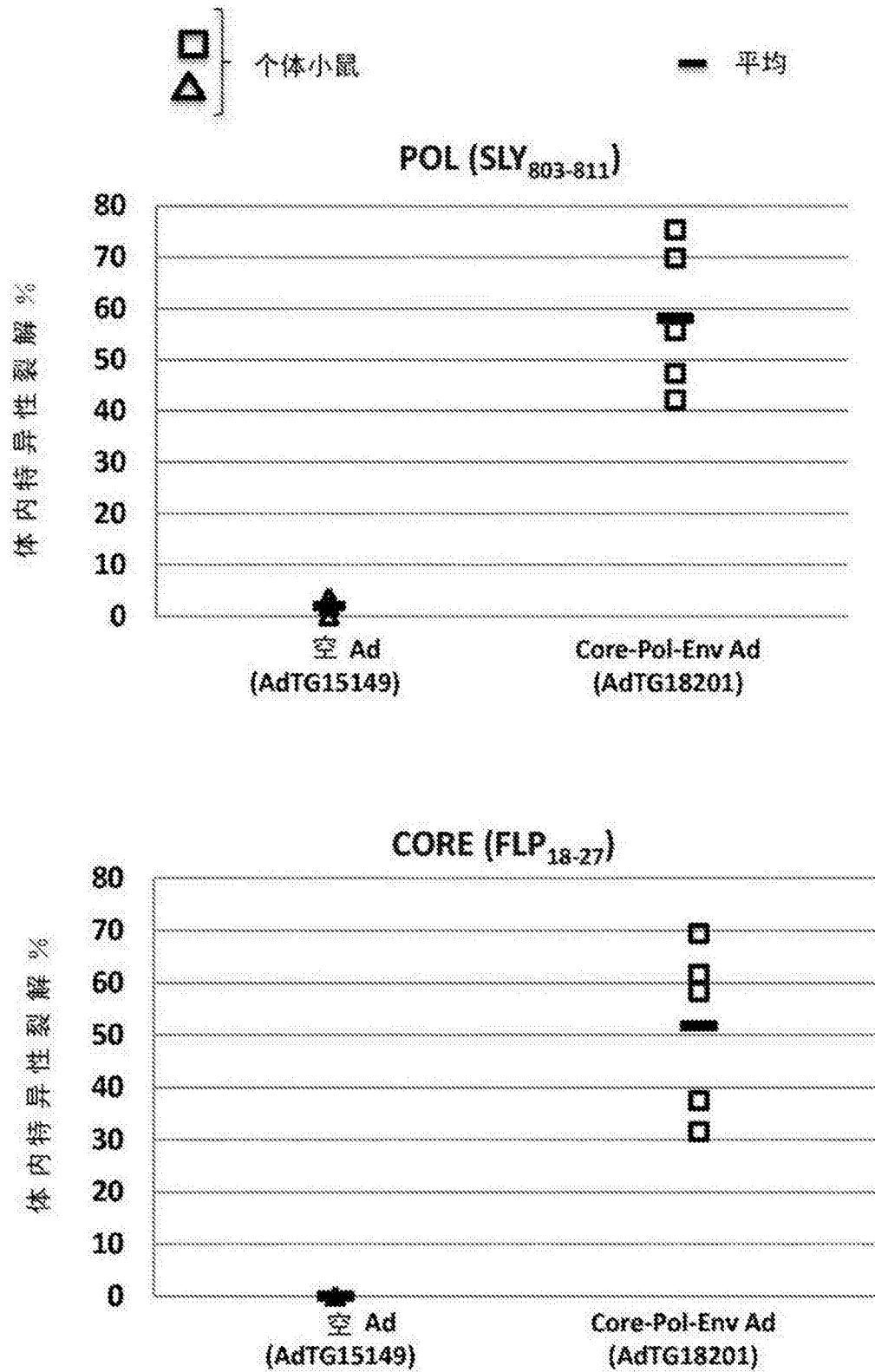


图7a

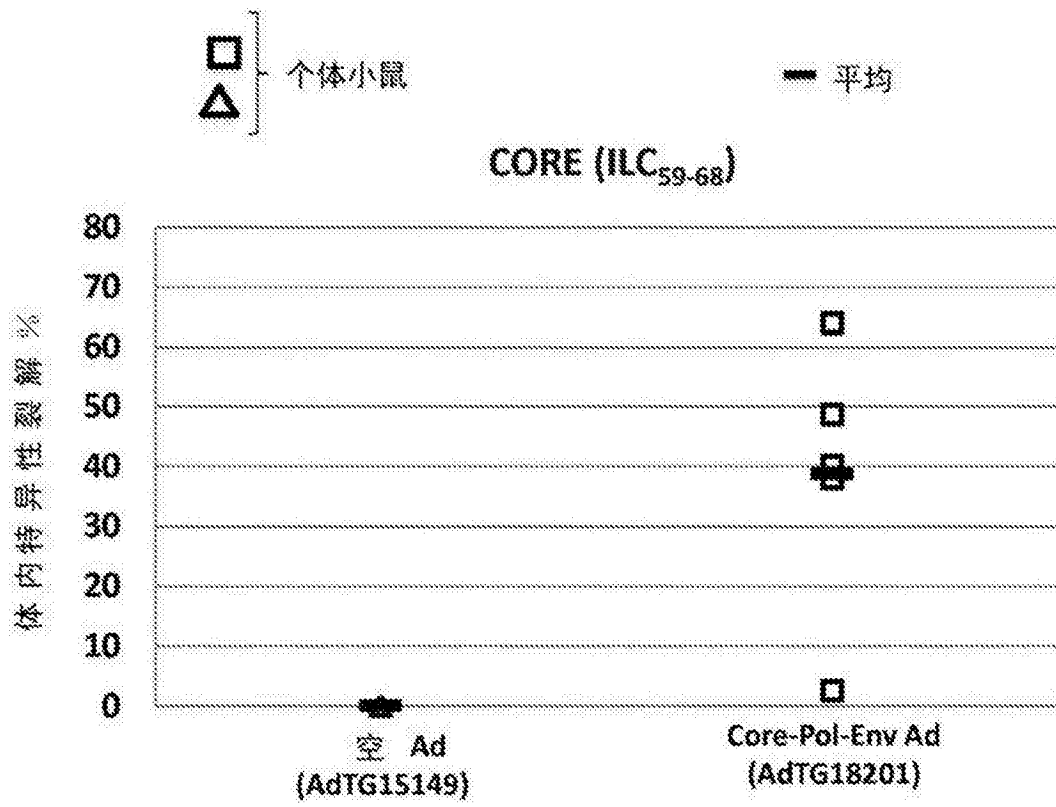


图7b

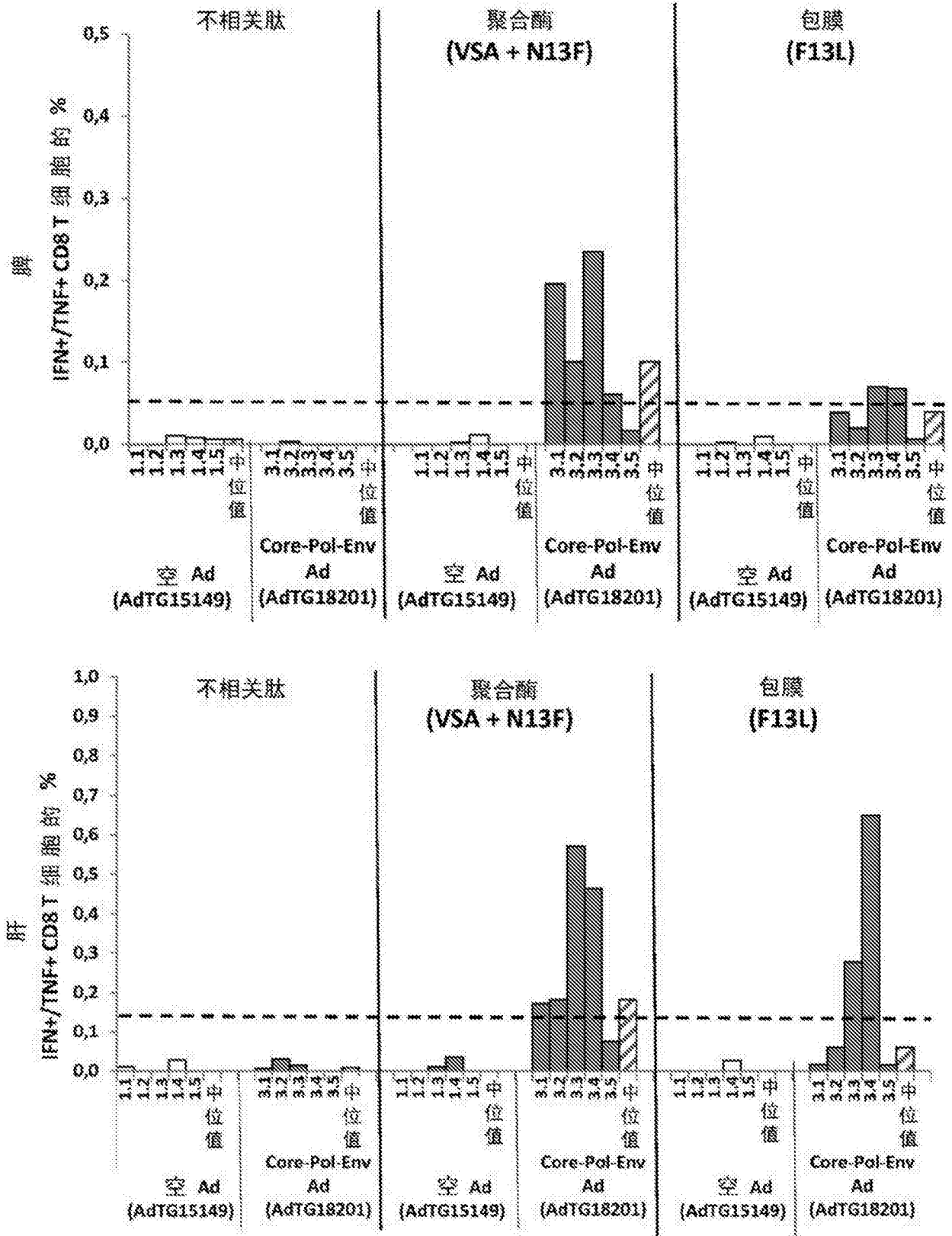


图8

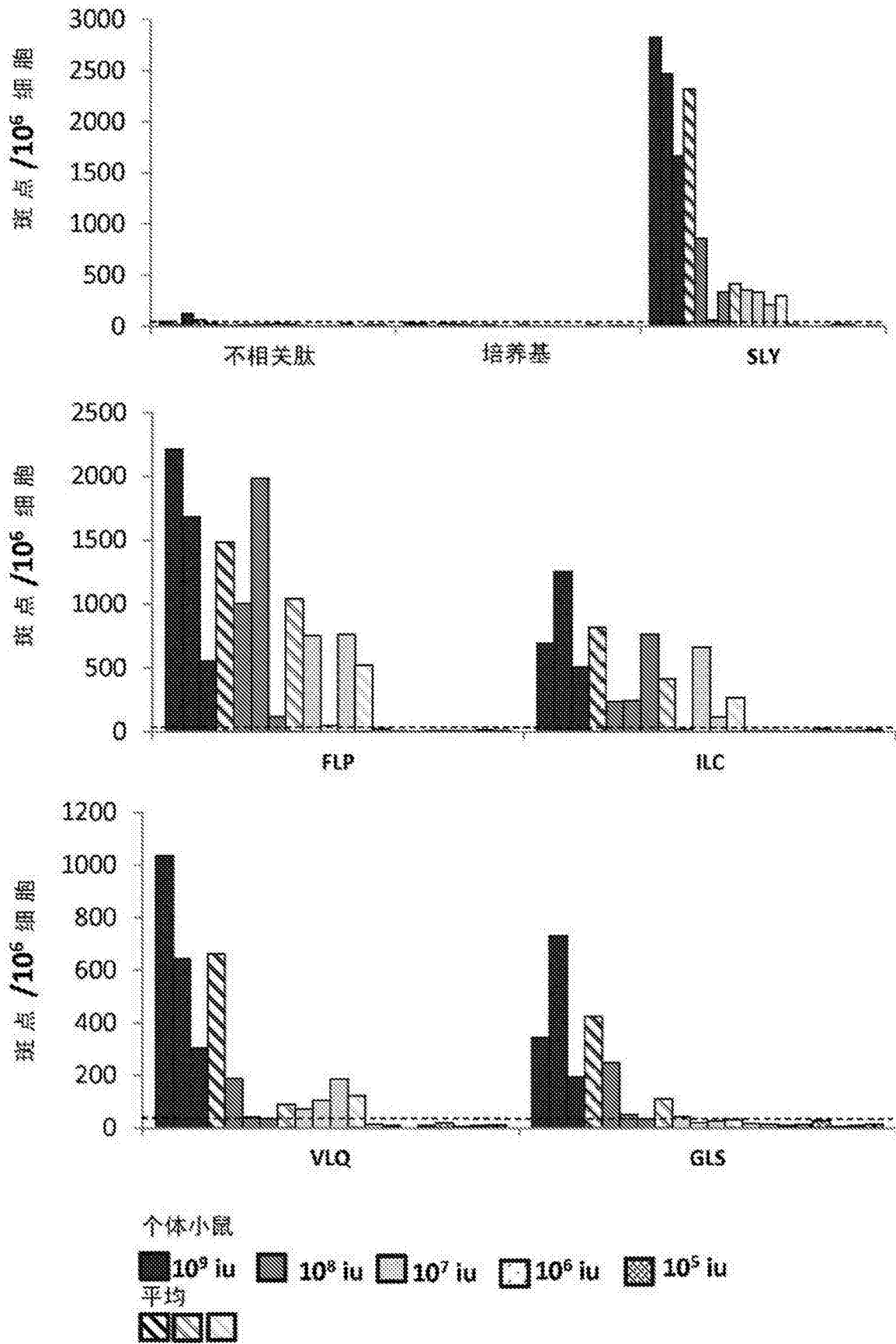


图9

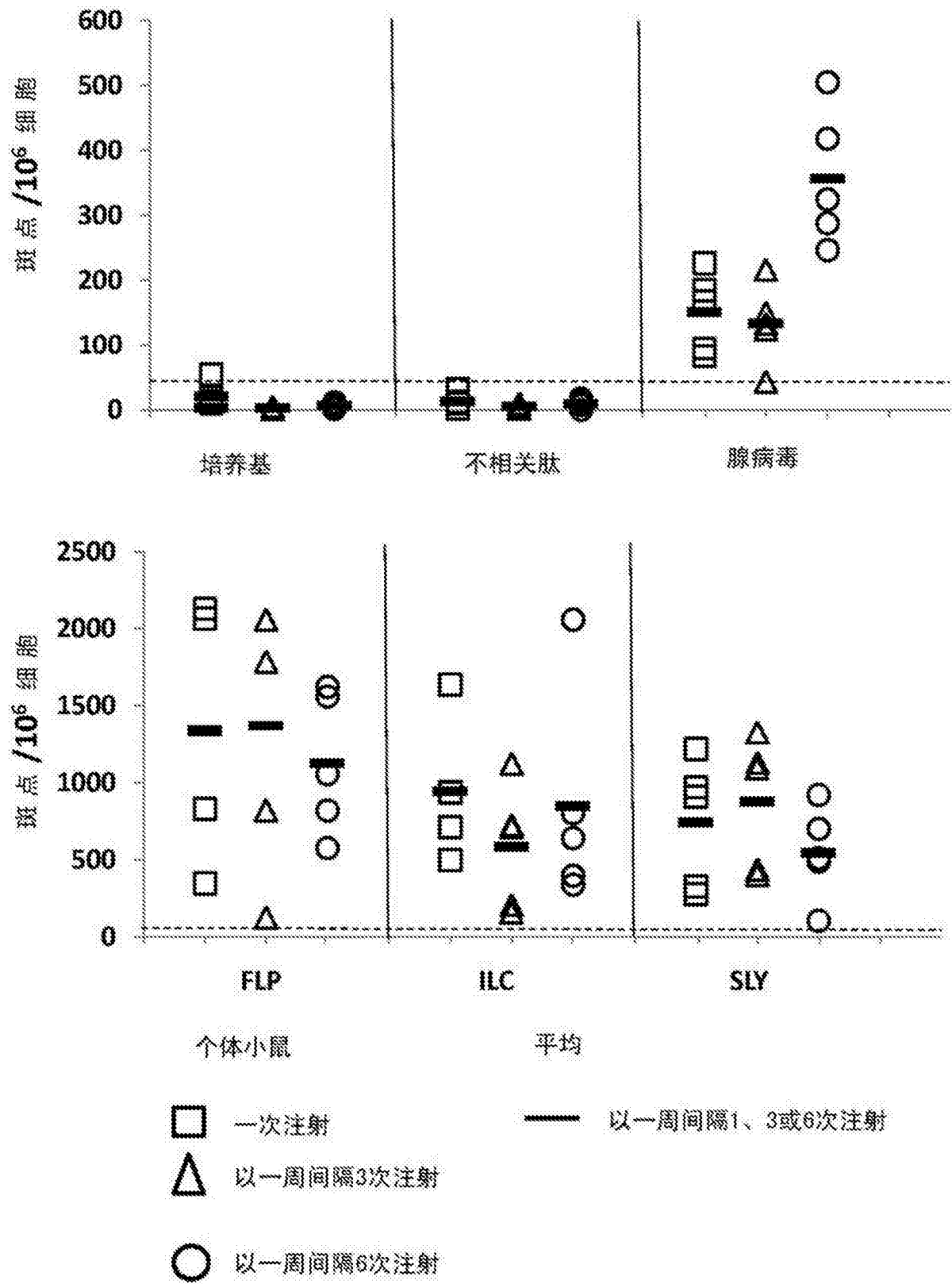


图10

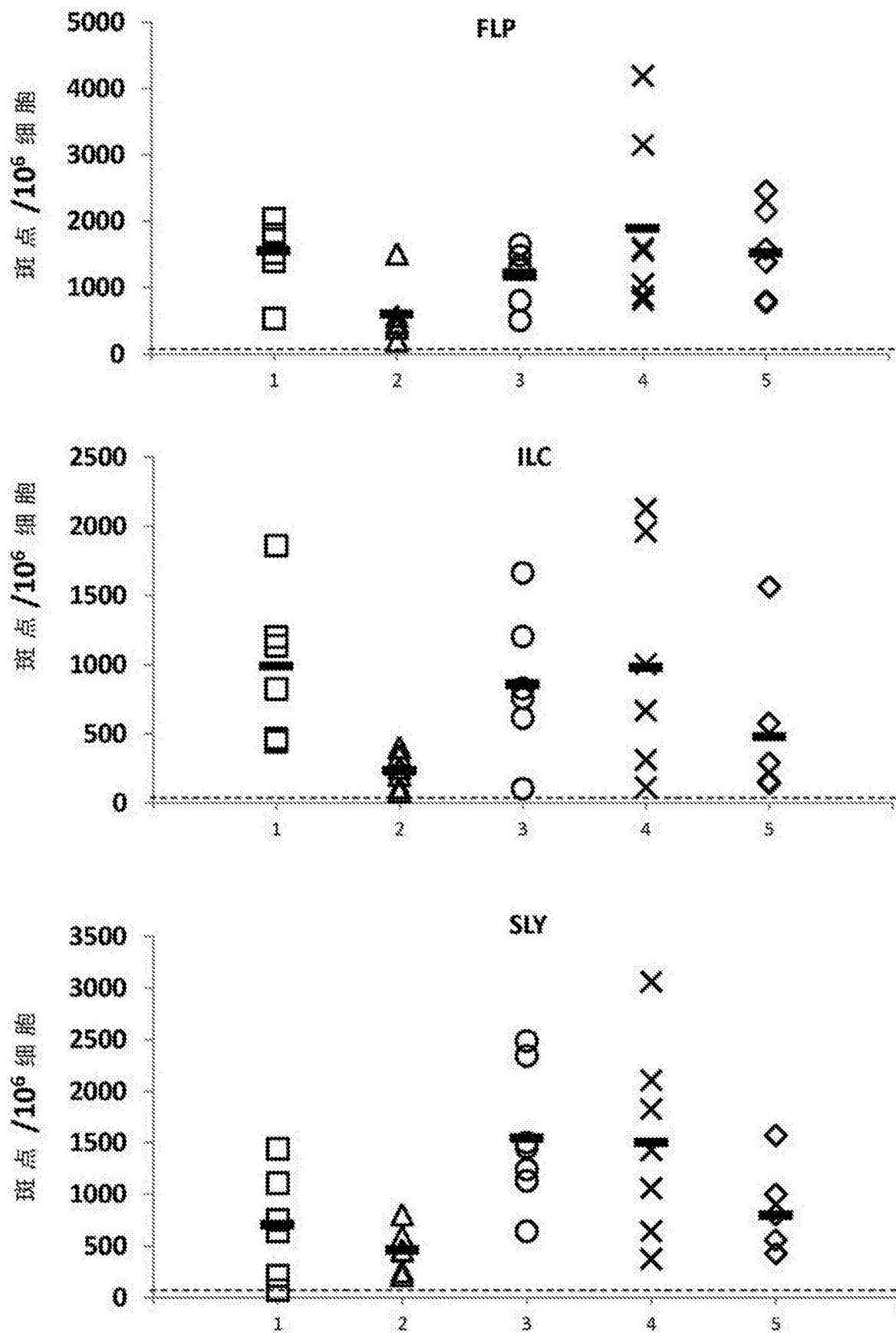


图11

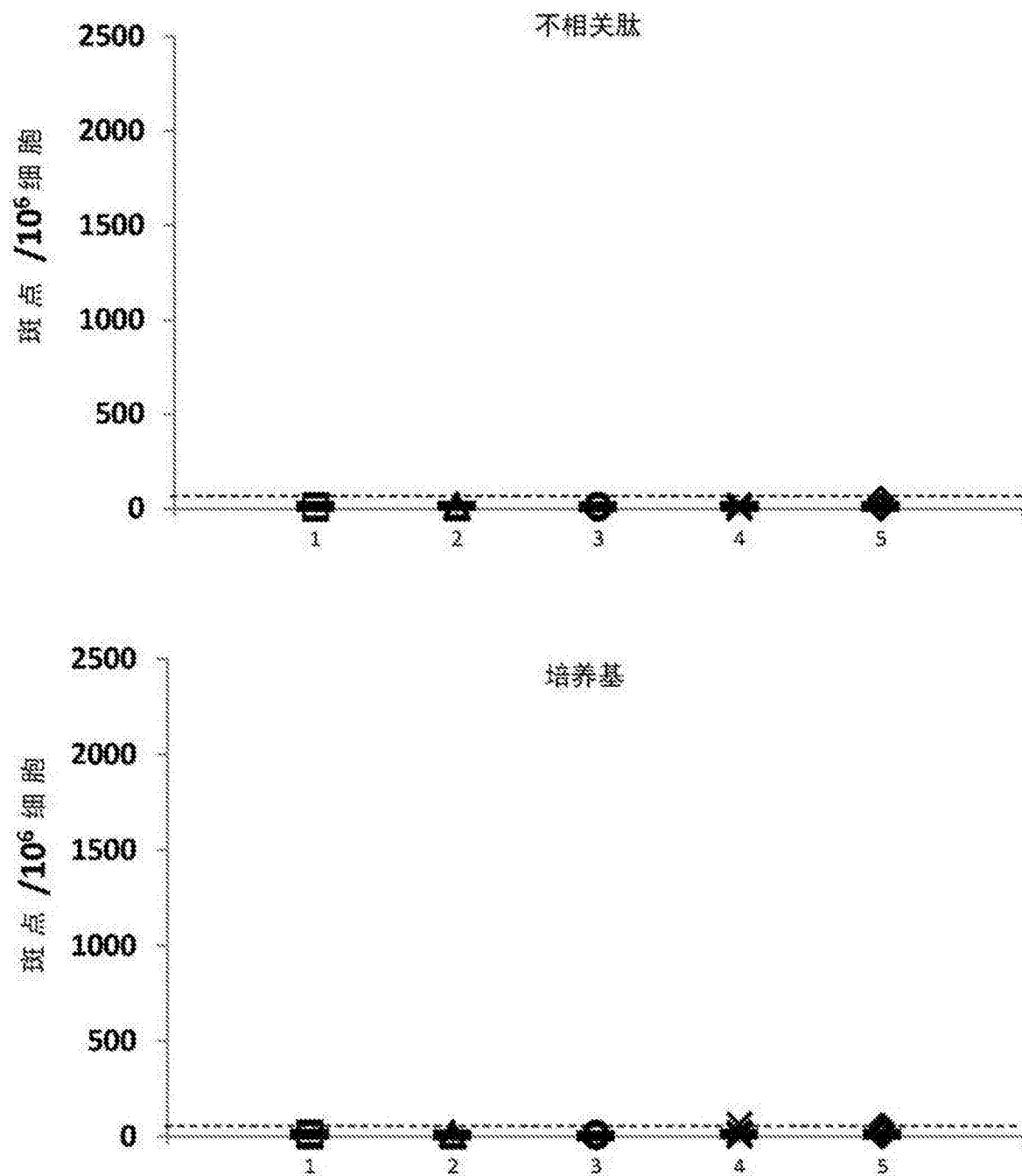


图12