

[19]中华人民共和国国家知识产权局

[51]Int. Cl⁷

C12N 15/86

C12N 5/10 A61K 48/00

[12] 发明专利申请公开说明书

[21] 申请号 99812304.8

[43]公开日 2001 年 11 月 21 日

[11]公开号 CN 1323350A

[22]申请日 1999.9.10 [21]申请号 99812304.8

[30]优先权

[32]1998.9.10 [33]US [31]60/099,791

[32]1999.9.9 [33]US [31]09/392,822

[86]国际申请 PCT/US99/20718 1999.9.10

[87]国际公布 WO00/15820 英 2000.3.23

[85]进入国家阶段日期 2001.4.18

[71]申请人 卡里唐公司

地址 美国加利福尼亚州

[72]发明人 俞德超

D·R·汉德森

[74]专利代理机构 中国国际贸易促进委员会专利商标事
务所

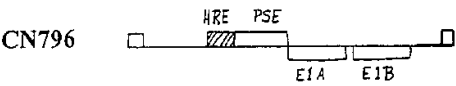
代理人 陈文平

权利要求书 3 页 说明书 39 页 附图页数 27 页

[54]发明名称 含有细胞状态特异性效应元件的腺病毒
载体和其使用方法

[57]摘要

本发明提供含有细胞状态特异性转录调节元件的腺病毒载体,其中所述转录调节元件可赋予腺病毒基因细胞状态特异性转录调节。“细胞状态”一般是指可逆的生理和/或环境状态。本发明还提供含有该载体的组合物和宿主细胞,以及使用该载体的方法。



ISSN 1008-4274

权 利 要 求 书

1. 含有处于包含了细胞状态特异性 TRE 的转录调节元件 (TRE) 控制下的腺病毒基因的腺病毒载体。
2. 权利要求 1 的腺病毒载体, 其中腺病毒基因是病毒复制所必需的。
3. 权利要求 2 的腺病毒载体, 其中腺病毒基因是早期基因。
4. 权利要求 2 的腺病毒载体, 其中腺病毒基因是晚期基因。
5. 权利要求 3 的腺病毒载体, 其中腺病毒早期基因是 E1A。
6. 权利要求 3 的腺病毒载体, 其中腺病毒早期基因是 E1B。
7. 权利要求 3 的腺病毒载体, 其中腺病毒早期基因是 E4。
8. 权利要求 1 的腺病毒载体, 其中细胞状态特异性 TRE 是人的。
9. 权利要求 1 的腺病毒载体, 其中细胞状态特异性 TRE 含有低氧效应元件 (HRE)。
10. 权利要求 9 的腺病毒载体, 其中 HRE 含有 SEQ ID NO:1。
11. 权利要求 1 的腺病毒载体, 其中细胞状态特异性 TRE 含有细胞周期特异性元件。
12. 权利要求 11 的腺病毒载体, 其中细胞周期特异性元件来自

E2F1 基因。

13. 权利要求 1 的腺病毒载体, 其中细胞状态特异性 TRE 含有热诱导元件。

14. 权利要求 1 的腺病毒载体, 还含有细胞类型特异性 TRE。

15. 权利要求 14 的腺病毒载体, 其中细胞类型特异性 TRE 是前列腺细胞特异性的。

16. 权利要求 15 的腺病毒载体, 其中前列腺细胞特异性 TRE 是 PSA-TRE。

17. 权利要求 1 的腺病毒载体, 还含有处于第二个细胞状态特异性 TRE 转录控制下的转基因。

18. 含有处于含细胞状态特异性 TRE 和细胞类型特异性 TRE 的 TRE 转录控制下的腺病毒基因的腺病毒载体。

19. 权利要求 18 的腺病毒载体, 其中腺病毒基因是早期基因。

20. 权利要求 19 的腺病毒载体, 其中腺病毒早期基因是 E1A。

21. 权利要求 20 的腺病毒载体, 其中细胞状态特异性 TRE 含有 HRE, 细胞类型特异性 TRE 是 PSA-TRE。

22. 权利要求 21 的腺病毒载体, 其中 HRE 含有 SEQ ID NO:1, PSA-TRE 含有 SEQ ID NO:3 的约第 503-约 2086 位核苷酸和 SEQ ID NO:3 的约第 5285-约 5836 位核苷酸。

23. 含有权利要求 1 的腺病毒载体的组合物。

24. 权利要求 23 的组合物，还含有药学上可接受的赋形剂。

25. 含有权利要求 1 的腺病毒载体的宿主细胞。

26. 繁殖对允许细胞状态特异性 TRE 发挥功能的细胞具有特异性的腺病毒的方法，所述方法包括将根据权利要求 1 的腺病毒与细胞结合，从而所述腺病毒获得繁殖。

27. 赋予靶细胞选择性细胞毒性的方法，所述方法包括用权利要求 1 的腺病毒载体接触允许细胞状态特异性 TRE 发挥功能的细胞，从而该载体进入细胞。

28. 抑制肿瘤生长的方法，包括向允许细胞状态特异性 TRE 发挥功能的肿瘤细胞导入权利要求 1 的腺病毒载体，其中腺病毒载体的导入导致肿瘤生长的抑制。

说 明 书

含有细胞状态特异性效应元件的腺病毒载体和其使用方法

对相关申请的引述

本申请要求 1998 年 9 月 10 日提交的美国临时专利申请 No. 60/099,791 的优先权。此在先申请的全部内容特此并入本文作为参考。

对联邦政府资助的研究下产生的发明的权利声明

(不适用)

技术领域

本发明涉及使用腺病毒载体进行的细胞转染。更具体地, 本发明涉及腺病毒载体在细胞中的细胞状态特异性复制, 而不论所述细胞的组织或细胞类型。

背景技术

尽管医药研究已取得许多进展, 在美国癌症仍是第二大致死原因。在工业化国家, 大约 5 个人中将有 1 人会死于癌症。传统的临床措施例如手术切除、放疗和化疗具有相当高的失败率, 尤其是对实体瘤而言。导致良性肿瘤的瘤形成通常可通过手术切除瘤块而完全治愈。如果肿瘤转变成恶性, 表现出向周围组织的侵袭, 则肿瘤的根除会困难得多。一旦恶性肿瘤发生转移, 根除的可能性将小的多。

除了基底细胞癌外, 仅在美国每年就有超过一百万的癌症新病例, 而且在该国每年因癌症死亡的人数超过 50 万。就全球而言, 5 种最常见的癌症是肺癌、胃癌、乳癌、结肠/直肠癌和宫颈癌, 而且每年的新病例总数超过 6 百万。

肺癌是最为顽固的实体瘤之一, 因为不能用手术治疗的病例在 60%

以上而且仅有 13% 的患者可存活 5 年。尤其是占肺癌病例总数大约一半的腺癌对化疗和放疗十分耐受。结肠直肠癌是第三种最常见癌症并且是第二大癌症致死原因。胰腺癌几乎总是致命的。因此，对于患有这些恶性肿瘤的许多患者而言目前的治疗前景不能令人满意，而且预后差。

肝细胞癌（HCC 或恶性肝细胞瘤）是世界上最为常见的癌症之一，在亚洲尤其成问题。肝细胞癌患者的治疗前景暗淡。甚至随着治疗和肝移植可行性方面的提高，也仅有少数的患者可通过切除或移植以清除肿瘤来获得治愈。对于大多数患者，目前的治疗仍不能令人满意，而且预后差。

在美国乳癌是最为常见的癌症之一，每年大约有 182,000 新病例产生而且有近 50,000 的患者死亡。在工业化国家，大约 8 名妇女中有 1 名预期会患乳癌。从 1930 年起乳腺癌的致死率一直保持不变。每年平均增加 0.2%，但在 65 岁以下的妇女中每年平均降低 0.3%。见例如 Marchant (1994) 目前对乳腺疾病的处理方法 II: 乳癌 (Contemporary Management of Breast Disease II: Breast Cancer) 北美洲临床妇产科学 (Obstetrics and Gynecology Clinics of North America) 21:555-560; 和 Colditz (1993) 癌症增刊 (Cancer Suppl.) 71:1480-1489。

尽管对乳癌的了解正在提高，但该疾病仍耐受医药干预。大多数临床主动性集中在早期诊断上，之后采取传统的干预形式尤其是手术和化疗。这些干预的成功是有限的，尤其是对肿瘤已发生转移的患者。因此，迫切需要改善目前的治疗手段以提供更为精确更为有效但伤害较小的治疗方法。

前列腺癌症是男性中发展最快的肿瘤，在美国 1995 年诊断出的新病例估计有 244,000 例，其中将出现大约 44,000 例死亡。前列腺癌症是目前男性中最频繁诊断出的癌症。前列腺癌症具有潜伏性；许多男性携带前列腺癌细胞但没有明显的疾病征兆。该病具有高发病率。（晚期）肿瘤通常向骨进行转移，并几乎总是致死的。

目前的治疗包括前列腺根除术、放射治疗、激素消融术和化疗。不幸的是，大约 80% 的患者在前列腺癌已发生向骨的转移后才获得确诊，因此限制了手术治疗的有效性。激素治疗通常随着时间的推移激素耐受性肿瘤细胞的发生而逐渐失效。尽管化疗试剂已用于前列腺癌的治疗中，但没有一个试剂表现出优于其它试剂，也没有一种药物组合似乎尤其有效。大多数前列腺癌症一般具有抗药和发展缓慢的性质，这使得标准化疗对该疾病尤其不起作用。

对于癌症治疗，主要的、事实上是压倒一切的障碍是选择性问题；即在不影响正常细胞功能的同时抑制肿瘤细胞增殖的能力。例如，在前列腺癌症的治疗中，传统肿瘤化疗的治疗比，或杀死的肿瘤细胞与杀死的正常细胞的比仅有 1.5:1。因此，需要更为有效的治疗方法和药物组合物用于治疗 and 预防瘤形成。

实体瘤经常含有血管化不良的区域，这部分是由于肿瘤细胞比构成血管的内皮细胞生长快速引起的。肿瘤细胞在此低氧情况下仍可存活，而且通常十分耐受化疗和放射治疗。在宫颈癌的最新研究中显示，肿瘤的氧状态是一个最为重要的预后因素，其重要性超过患者的年龄、绝经状态、临床阶段、体积和组织学。Hoeckel 等(1996) *Semin. Radiat. Oncol.* 6:1-8。

开发更特异的靶向形式癌症治疗方法尤其有意义，特别是对难以治疗成功的癌症而言。与导致相对非特异性和通常引起严重毒性的传统癌症治疗相反，更具有特异性的治疗的特征在于其试图选择性抑制或杀死恶性细胞而不损伤健康细胞。放射抗性和化学抗性的肿瘤尤其是一种挑战，并需要治疗这类肿瘤的方法。

对于许多这些癌症，一种可能的治疗方法是基因治疗，其中将目的基因引入恶性细胞。已开发了各种病毒载体包括腺病毒载体用作基因治疗的运载工具。在用于基因治疗的腺病毒载体的开发中，实际上仅集中在将腺病毒只用作导入目的基因的运载工具，而不将其本身作为效应物。很大程度上由于宿主的免疫应答，腺病毒复制被认为是一种不希望有的结果。在使用复制缺陷型腺病毒进行的癌症治疗中，宿

主免疫应答在两个水平上限制了重复给药的持续时间。首先，腺病毒递送载体本身的衣壳蛋白具有免疫原性。其次，病毒晚期基因常常在转导的细胞中表达，引起细胞免疫。因此，基因转移载体和转移载体的病毒基因产物的免疫原性、以及基因表达的瞬时特征均限制了细胞因子、肿瘤抑制基因、核酶、自杀基因、或将前体药物转化成活性药物的基因的重复施用效率。

已有腺病毒载体用作癌症治疗运载工具的报道。这些方法中的一些利用组织（即细胞类型）特异性转录调节元件以选择性地驱动腺病毒的复制（以及由此引起的细胞毒性）。美国专利 No. 5,698,443；也见 WO 95/11984；WO 96/17053；WO 96/34969；WO 98/35028。尽管这些方法是有效和有希望的，但仍然存在组织特异性复制可能无法胜任的其它治疗情况。

除了癌细胞外，通常还期望选择性地破坏某些不想要的细胞或组织。然而除了损伤性手术外，却缺乏可允许这种选择性细胞毒性和/或抑制的方法，尤其是非损伤性方法。

这就需要如下载体结构，该载体结构既能在抗载体的特异免疫学应答阻止进一步治疗之前以最少的施用次数基本消除所有癌细胞，又适用于特异性的集中癌症清除治疗。还需要选择性破坏、或损伤不需要细胞的能力，而无论细胞的类型和/或解剖定位。

本发明概要

本发明提供对处于给定或特定生理状态的细胞具有特异性的可复制腺病毒载体，其中所述生理状态允许或诱导处于细胞状态特异性 TRE 转录控制下的多核苷酸的表达，而且本发明提供这些载体的使用方法。在这些可复制腺病毒载体中，一或多个腺病毒基因处于细胞状态特异性转录调节元件（TRE）的转录控制之下。优选地，处于细胞状态特异性 TRE 转录控制下的腺病毒基因是腺病毒增殖所必需的基因。还可存在处于细胞状态特异性 TRE 控制下的转基因。对于本发明腺病毒载体，细胞状态特异性 TRE 在处于特定可逆生理状态的细胞中有活性，其中

所述生理状态可以是异常生理状态，即与正常生理条件下处于非分裂或调节分裂状态的相同细胞的典型或正常生理状态不同的生理状态。

由此，一方面本发明提供含有一种腺病毒基因的腺病毒载体，其中所述腺病毒基因处于细胞状态特异性 TRE 转录控制之下。在另一个实施方案中，细胞状态特异性 TRE 是人的。在另一个实施方案中，细胞状态特异性 TRE 包含细胞状态特异性启动子和增强子。在另一个实施方案中，细胞状态特异性 TRE 与细胞类型特异性 TRE 并置，并且这两个元件一起控制腺病毒基因的表达。因此，本发明提供含有包含细胞状态特异性 TRE 和细胞类型特异性 TRE 的 TRE 的腺病毒载体。

在一些实施方案中，处于细胞状态特异性 TRE 转录控制下的腺病毒基因是复制必需的腺病毒基因。在一些实施方案中，复制必需的腺病毒基因是早期基因。在另一个实施方案中，该早期基因是 E1A。在另一个实施方案中，该早期基因是 E1B。在另一个实施方案中，E1A 和 E1B 均处于细胞状态特异性 TRE 的转录控制之下。在其它实施方案中，复制必需的腺病毒基因是晚期基因。

在另一个实施方案中，细胞状态特异性 TRE 含有低氧效应元件。在另一个实施方案中，细胞状态特异性 TRE 含有 SEQ ID NO:1 的核苷酸序列。

在另一个实施方案中，细胞状态特异性 TRE 含有细胞周期特异性 TRE。该细胞周期特异性 TRE 可来源于 E2F1 的 5'侧翼区。在一个实施方案中，该细胞周期特异性 TRE 含有 SEQ ID NO:2 中所描述的核苷酸序列。

在其它实施方案中，腺病毒载体还可含有转基因，其中所述转基因处于细胞状态特异性 TRE 的转录控制之下。在一些实施方案中，该转基因是细胞毒性基因。

在其它实施方案中，腺病毒载体含有一个处于第一个细胞状态特异性 TRE 控制下的腺病毒复制必需的腺病毒基因，和第二个处于第二个细胞状态特异性 TRE 控制下的腺病毒复制必需的腺病毒基因。该第一和第二个细胞状态特异性 TRE 可相同、基本相同或不同。

在其它实施方案中，腺病毒载体含有一个处于第一个细胞状态特异性 TRE 控制下的腺病毒复制必需腺病毒基因，和一个处于第二个细胞状态特异性 TRE 控制下的转基因。该第一和第二个细胞状态特异性 TRE 可基本相同或不同。

在其它实施方案中，腺病毒载体含有一个处于细胞状态特异性 TRE 转录控制下的腺病毒基因，和第二个处于细胞类型特异性 TRE 转录控制下的腺病毒基因。在其它实施方案中，腺病毒载体含有一个处于细胞状态特异性 TRE 转录控制下的腺病毒基因，和一个处于细胞类型特异性 TRE 转录控制下的转基因。

在另一方面，本发明提供含有本文所述腺病毒载体的宿主细胞。

在另一方面，本发明提供含有本文所述腺病毒载体的药物组合物。

在另一方面，本发明提供含有本文所述腺病毒载体的试剂盒。

在另一方面，本发明提供在靶细胞（即允许或诱导细胞状态特异性 TRE 发挥功能的细胞）中赋予选择性细胞毒性的方法，该方法包括用本文所述腺病毒载体接触该细胞，从而载体进入细胞。

本发明的另一个实施方案是优先在细胞状态允许或诱导细胞状态特异性 TRE 起作用的哺乳动物细胞中复制的腺病毒。

在另一方面中，本发明提供用于增殖对细胞状态允许细胞状态特异性 TRE 发挥功能的哺乳动物细胞具有特异性的腺病毒的方法，所述方法包括将本文所述腺病毒载体与其细胞状态允许细胞状态特异性 TRE 发挥功能的哺乳动物细胞结合在一起，从而所述腺病毒获得增殖。

本发明还提供抑制肿瘤细胞，更具体地是靶肿瘤细胞（即允许或诱导细胞特异性 TRE 发挥功能的肿瘤细胞），生长的方法，该方法包括用本发明腺病毒载体接触肿瘤细胞以便该腺病毒载体进入该肿瘤细胞并表现出对该肿瘤细胞的选择性细胞毒性。

在另一方面，本发明提供用于在生物样品中鉴定其细胞状态允许细胞状态特异性 TRE 发挥功能的细胞的方法，该方法包括用本文所述腺病毒载体接触生物样品的细胞，如果有的话，并检测该腺病毒载体的复制。

附图简要说明

图 1 是腺病毒载体 CN796 的示意图，正如实施例 1 所述，该载体中 E1A 基因处于 HRE 和 PSA-TRE 的转录控制之下。

图 2 显示了大鼠烯醇酶 1 基因 5'侧翼区的 HRE 核苷酸序列(SEQ ID NO:1)。

图 3 显示了人 E2F1 基因 5'侧翼区的核苷酸序列 (SEQ ID NO:2)。星号显示转录起始位点。

图 4 描述了前列腺特异性抗原 TRE 的核苷酸序列。

图 5 描述了癌胚抗原 TRE 的核苷酸序列。

图 6 描述了人腺激肽释放酶 TRE 的核苷酸序列。

图 7 描述了粘蛋白 TRE 的核苷酸序列。

图 8 描述了大鼠 probasin TRE 的核苷酸序列。

图 9 描述了腺病毒死亡蛋白的核苷酸序列和翻译的氨基酸序列。

实施本发明的方式

我们发现并构建了含有处于细胞状态特异性转录效应元件 (TRE) 转录控制下的腺病毒基因，以至该腺病毒基因优先在其细胞状态允许该细胞状态特异性 TRE 发挥功能的细胞中转录的可复制腺病毒载体，并开发了使用这些腺病毒载体的方法。在一些优选实施方案中，本发明腺病毒载体含有至少一个处于 TRE 转录控制下的腺病毒复制必需的腺病毒基因，优选至少一个早期基因，其中所述 TRE 通过结合该细胞状态特异性 TRE 所调节转录必需的转录因子和/或辅因子而受到特异性调节。通过规定至少一个复制必需的腺病毒基因的细胞状态特异性转录，本发明提供了由于选择性复制和/或选择性转录而可用于产生特异性细胞毒效应的腺病毒载体。该载体在期望杀死靶细胞的癌症治疗中尤其有用。当期望选择性破坏和/或抑制其它非肿瘤细胞时，该载体也可用于在这些细胞中产生靶向细胞毒作用。这些载体还可用于检测在例如合适的生物（如临床）样品中是否存在其细胞状态允许细胞状

态特异性 TRE 发挥功能的细胞。而且，通过使用细胞状态特异性 TRE，所述腺病毒载体可选择在靶细胞中选择性地产生一或多种目的蛋白。

我们发现，本发明腺病毒载体优先在其细胞状态允许细胞状态特异性 TRE 发挥功能的细胞中复制和/或表达可操作地与细胞状态特异性 TRE 连接的腺病毒基因。与以前描述的设计以优先在特异分化细胞类型中表达的腺病毒载体不同，本发明腺病毒载体含有非细胞类型特异性的调节元件。但它们带来了细胞状态特异性复制，和/或可操作地连接的腺病毒基因和/或转基因的细胞状态特异性表达。

本发明腺病毒载体含有在表现特定的可逆和/或进行性生理（即环境或代谢）特征的细胞中起作用的细胞状态特异性 TRE。为了使细胞状态 TRE 发挥功能，靶细胞可表现出异常的生理状态，如低氧压力，或可处于异常环境条件下，如热或电离辐射。通过比较处于必要生理状态或条件下（如异常生理状态）的细胞与不表现该必要生理状态（如处于正常生理条件下）的细胞中的复制水平（即滴度），可显示出这些载体的复制优先性。因此，本发明还使用并利用了腺病毒载体被认为是不利的方面，即其复制和可能伴随的免疫原性。由于病毒复制的细胞状态特异性要求，失控感染的可能性显著降低。不希望被任何特定理论所束缚，本发明人注意到腺病毒蛋白的产生可起到激活和/或刺激如下免疫系统的作用，该免疫系统普遍和/或特异地针对产生腺病毒蛋白的靶细胞，这在患者常常出现适度或严重的免疫低下的癌症治疗中可能是一个重要的考虑因素。

本发明腺病毒载体在如下特殊治疗方案中发现尤其有用，在这些治疗方案中，治疗高度针对例如可能用其它方法不能手术治疗或不可治疗的特殊癌症。本发明的一个重要特征是，无论癌症是何种组织或细胞类型，所述载体均可用于这些治疗，而且可将它们的细胞毒性靶向确定位置。

一般技术

除非特别指出，本发明的实施可使用属于本领域技术范畴的分子

生物学（包括重组技术）、微生物学、细胞生物学、生物化学和免疫学常规技术。这些技术在文献中有充分阐述，例如“分子克隆：实验手册（Molecular Cloning: A Laboratory Manual）”，第二版（Sanbrook 等，1989）；“寡核苷酸合成（Oligonucleotide Synthesis）”（M. J. Gait 编著，1984）；“动物细胞培养（Animal Cell Culture）”（R. I. Feshney 编著，1987）；“酶学方法（Methods in Enzymology）”（Academic Press 公司）；“实验免疫学手册（Handbook of Experimental Immunology）”（D. M. Wei 和 C. C. Blackwell 编著）；“哺乳动物细胞的基因转移载体（Gene Transfer Vectors for Mammalian Cells）”（J. M. Miller 和 M. P. Calos 编著，1987）；“当代分子生物学实验方法（Current Protocols in Molecular Biology）”（F. M. Ausubel 等编著，1987）；“PCR：聚合酶链式反应（PCR: The Polymerase Chain Reaction）”（Mullis 等编著，1994）；“当代免疫学实验方法（Current Protocols in Immunology）”（J. E. Coligan 等编著，1991）。

有关腺病毒的技术，尤其参见 Felgner 和 Ringold(1989) 自然 (Nature) 337: 387-388; Berkner 和 Sharp(1983) 核酸研究 (Nucl. Acids Res.) 11: 6003-6020; Graham(1984) EMBO J. 3:2917-2922; Bett 等 (1993) 病毒学杂志 (J. Virology) 67:5911-5921; Bett 等 (1994) 美国国家科学院院刊 (Proc. Natl. Acad. Sci. USA) 91:8802-8806。

定义

本文所用的“转录效应元件”或“转录调节元件”或“TRE”是增加可操作连接的多核苷酸序列在允许该 TRE 发挥功能的宿主细胞中转录的多核苷酸序列，优选 DNA 序列。TRE 可包括增强子和/或启动子。

本文所用术语“细胞状态特异性 TRE”是使可操作连接的多核苷酸在允许细胞状态特异性 TRE 发挥功能的细胞中转录激活的 TRE，所述细胞即表现特定生理状态包括但不限于异常生理状态的细胞。因此，

“细胞状态”是指细胞可逆或进行性的给定或特定生理状态。生理状态可在内部或外部产生；例如，它可以是代谢状态（如低氧），或可由热或电离辐射产生。“细胞状态”与“细胞类型”不同，细胞类型是指细胞的分化状态，它在正常情况下是不可逆的。一般地（但不是必然），正如本文所讨论的，细胞状态具体表现为异常生理状态，其例子在下文中给出。

“正常细胞状态”或“正常生理状态”是指以正常生理状态存在并且不分裂或以调节的方式分裂的细胞状态，即处于正常生理状态的细胞。

本文中可交换使用的术语“异常细胞状态”和“异常生理状态”是指对异常生理条件起反应、或是其结果或受其影响的细胞状态。异常细胞状态既不是细胞类型特异的，也不是组织类型特异的。异常细胞状态是与处于非分裂/调节的分裂状态和正常生理情况下的相同类型细胞相比较来定义的。

“正常生理状态”是本领域技术人员熟知的。这些状态可根据细胞在机体内的位置不同而不同。例如，氧压力可随组织的不同而变化。对于为确定 TRE 是否对正常生理情况的偏离敏感而进行的体外分析，这些情况通常包括大约 20%O₂ 的氧浓度和大约 37℃ 的温度。“调节的细胞分裂”是本领域熟知的术语，是指细胞的正常有丝分裂活动。本领域技术人员明了，正常有丝分裂活动随细胞类型的不同而变化。例如，组织中的许多终末分化细胞表现出极少或没有有丝分裂活动，而造血细胞通常是有丝分裂活跃的。

本文所用“异常生理情况”或者“异常生理状态”是指偏离正常生理情况的状态，其包括但不限于以氧浓度改变为特征的生理情况，如低氧情况；偏离了生理温度的温度；触发细胞凋亡的状态；辐射，包括电离辐射和 UV 辐射；由于例如控制细胞分裂的因子如成视网膜细胞瘤蛋白（Rb）的缺乏或数量不足或失活而造成的失调细胞分裂；细胞周期时间安排的变异；病原体感染；暴露于化学物质之中；或上述情况的组合。另一个例子是可能或的确在任何细胞类型中存在的突

变，即其存在不依赖于细胞的分化状态或不与细胞的分化状态相关。

本文所用的“靶细胞”是允许或诱导细胞状态特异性 TRE 起作用以便实现可操作连接的多核苷酸的转录激活的细胞。靶细胞是显示必要生理（或环境）状态的细胞，该状态可以是异常生理状态。优选地，靶细胞是哺乳动物细胞，优选人细胞。靶细胞可以是肿瘤的，也可以不是。对于转录激活，是指靶细胞中的转录比对照细胞（即不显示必要生理状态（通常是正常生理状态）的相应细胞）中的转录增加到至少约 2 倍，优选至少约 5 倍，优选至少约 10 倍，更优选至少约 20 倍，更优选至少约 50 倍，更优选至少约 100 倍，更优选至少约 200 倍，甚至更优选至少约 400 倍至约 500 倍，甚至更优选至少约 1000 倍。正常水平通常是在激活细胞状态特异性 TRE 的情况下测试的细胞中的活性水平（如果有的话），或在显示必要生理状态的细胞中测量的缺乏细胞状态特异性 TRE 的报告结构的活性水平（如果有的话）。

细胞状态特异性 TRE 的“功能上保守的”变体是与另一个细胞状态特异性 TRE 不同，但仍保持细胞状态特异性转录活性的细胞状态特异性 TRE。细胞状态特异性 TRE 中的差异可以是由于例如一或多个单碱基突变、插入、缺失、和/或这些碱基的修饰引起的线性序列的不同造成的。该差异也可由一或多个糖和/或细胞状态特异性 TRE 的碱基之间的一或多个连接的变化引起。

“腺病毒载体”包括本发明的多核苷酸结构。本发明的多核苷酸结构可是几种形式之中的任何一种，包括但不限于 DNA、包装在腺病毒外壳中的 DNA、包装在其他病毒或病毒样形式（如单纯疱疹病毒和 AAV）中的 DNA、包在脂质体中的 DNA、与聚赖氨酸形成复合物的 DNA、与合成的聚阳离子分子形成复合物的 DNA、与运铁蛋白结合的 DNA、与化合物如 PEG 形成复合物以在免疫学上“掩蔽”该分子和/或增加半衰期的 DNA、以及与非病毒蛋白结合的 DNA。优选地，该多核苷酸是 DNA。本文所用“DNA”不仅包括碱基 A、T、C 和 G，而且包括它们的任何类似物或这些碱基的修饰形式如甲基化核苷酸、核苷酸间的修饰如不带电荷的键合和硫代酸酯、糖类似物的使用、以及修饰的和/或其它主

链结构如聚酰胺。对于本发明，腺病毒载体在靶细胞中是能够复制的。

本文中可互换使用的术语“多核苷酸”和“核酸”是指任何长度的核苷酸包括核糖核苷酸和脱氧核糖核苷酸的聚合形式。这些术语包括单链、双链或三链 DNA、基因组 DNA、cDNA、RNA、DNA-RNA 杂交体，或包含嘌呤和嘧啶碱基或其它天然的、经化学、生物化学修饰的、非天然的或衍生化核苷酸碱基的聚合物。多核苷酸的主链可包含糖和磷酸（正如 RNA 和 DNA 中典型见到的一样），或修饰或替代的糖或磷酸基团。此外，多核苷酸的主链可含有合成亚单位如氨基磷酸酯的聚合物，因此可以是寡脱氧核苷氨基磷酸酯（P-NH₂）或混合的氨基磷酸酯-磷酸二酯寡聚物。Peyrottes 等(1996)核酸研究 (Nucleic Acids Res.) 24: 1841-8; Chaturvedi 等 (1996) 核酸研究 24: 2318-23; Schultz 等 (1996) 核酸研究 24: 2966-73. 硫代磷酸酯键可用于替代磷酸二酯键。Braun 等(1988)免疫学杂志(J. Immunol.)141: 2084-9; Latimer 等 (1995) 分子免疫学 (Mol. Immunol.) 32: 1057-1064. 此外，还可通过合成互补链并在适当条件下将其退火，或用 DNA 聚合酶和适当引物从头合成互补链，从化学合成的单链多核苷酸产物获得双链多核苷酸。

下面是多核苷酸的非限制性例子：基因或基因片段、外显子、内含子、mRNA、tRNA、rRNA、核酶、cDNA、重组多核苷酸、分支多核苷酸、质粒、载体、分离的任何序列 DNA、分离的任何序列 RNA、核酸探针和引物。多核苷酸可包含修饰的核苷酸如甲基化的核苷酸和核苷酸类似物、尿嘧啶、其它糖和连接基团如氟代核糖和硫代酸酯、以及核苷酸分支。核苷酸序列可被非核苷酸成分中断。多核苷酸可在聚合后进一步修饰，如与标记成分结合。本定义中所述修饰的其它种类有加帽，用类似物替代一或多个天然存在的核苷酸，和使多核苷酸与蛋白质、金属离子、标记成分、其它多核苷酸或固体支持物结合的引入方法。优选地，该多核苷酸是 DNA。本文所用“DNA”不仅包括碱基 A、T、C 和 G，而且包括它们的任何类似物或这些碱基的修饰形式如甲基化核苷酸、核苷酸间的修饰如不带电荷的键合和硫代酸酯、糖类似物

的使用、以及修饰的和/或其它主链结构如聚酰胺。

多核苷酸或多核苷酸区与另一个序列具有某种“序列一致性”百分数（如 80%、85%、90%或 95%）是指当对齐时两个序列比较的相同碱基百分数。这种对齐（alignment）和百分同源性或序列一致性可采用本领域已知的软件程序来确定，所述程序例如“当代分子生物学实验方法”（Current Protocols in Molecular Biology）（F.M. Ausubel 等编著，1987）增补 30，7.7.18 节，表 7.7.1 中所述的程序。优选的对齐程序是 ALIGN Plus（Scientific and Educational Software, Pennsylvania），优选使用默认参数。

“处于转录控制下”是本领域熟知的术语，是指多核苷酸序列（通常是 DNA 序列）的转录取决于它可操作地连接于有助于转录起始或促进转录的元件。“可操作地连接”是指并置，其中所述元件的排列允许它们发挥功能。

“复制”和“增殖”可互换使用，是指本发明多核苷酸结构的复制或增殖能力。该术语是本领域熟知的。对于本发明，复制涉及腺病毒蛋白的产生，并且通常是指腺病毒的繁殖。复制可采用本领域标准方法和本文所述方法测量，例如爆发集落测定（burst assay）、噬菌斑测定或一步生长曲线分析。

本文所用“细胞毒性”是本领域熟知的术语，是指细胞通常的生物化学或生物学活动受到损害（即被抑制）的状态。这些活动包括但不限于代谢；细胞复制；DNA 复制；转录；翻译；分子摄取。“细胞毒性”包括细胞死亡和/或细胞裂解。指示细胞毒性的分析方法是本领域已知的，如染料排斥法、³H 胸苷摄取和噬菌斑测定。

本文所用术语“选择性细胞毒性”是指与本发明腺病毒载体对不允许细胞状态特异性 TRE 起作用的细胞（非靶细胞）赋予的细胞毒性相比，本发明的腺病毒载体对允许或诱导细胞状态特异性 TRE 起作用的细胞（靶细胞）赋予的细胞毒性。这种细胞毒性可通过例如噬菌斑测定、含有靶细胞的肿瘤的体积减小或稳定化、或肿瘤细胞特征标志或组织特异性标志如癌标志（如前列腺特异性抗原）的血清水平降低

或稳定化来测定。

对于腺病毒，“异源多核苷酸”或“异源基因”或“转基因”是野生型腺病毒中不存在的任何多核苷酸或基因。优选地，在通过腺病毒载体导入之前，转基因也不在靶细胞中表达或存在。优选的转基因的例子见下文。

对于腺病毒，“异源”启动子或增强子是指与腺病毒基因无关或不来自腺病毒基因的启动子或增强子。

对于腺病毒，“内源”启动子、增强子或 TRE 是腺病毒天然的或腺病毒来源的。

对于细胞状态特异性 TRE，“异源”启动子或增强子是指正常情况下在细胞中与细胞状态特异性 TRE 无关的或非由它来源的启动子或增强子。异源启动子或增强子的例子是白蛋白启动子或增强子和其它病毒启动子和增强子如 SV40，或细胞类型特异性 TRE 如前列腺特异性 TRE。

“细胞类型特异性 TRE”相对于其它细胞类型优先在特定细胞类型中起作用。与细胞状态相反，“细胞类型”是不可逆的细胞分化状态的反映。例如，前列腺特异抗原在前列腺细胞中表达，但基本上不在其它细胞类型如肝细胞、星形胶质细胞、心细胞、淋巴细胞等中表达。通常，细胞类型特异性 TRE 仅在一种细胞类型中有活性。当细胞类型特异性 TRE 在一种以上细胞类型中有活性时，其活性也局限于有限数目的细胞类型，即它并不是在所有细胞类型中均有活性。细胞类型特异性 TRE 可以是也可以不是肿瘤细胞特异的。

“抑制”肿瘤生长是指与未接触本文描述的腺病毒载体（即被转染）的肿瘤生长相比，减少的生长状态。肿瘤细胞生长可用本领域已知的任何方法测定，所述方法包括但不限于测量肿瘤体积、用 ^3H 胸苷掺入法测定肿瘤细胞是否增殖、或对肿瘤细胞计数。“抑制”肿瘤细胞生长意味着任何下列状态或所有下列状态：肿瘤生长减缓、延迟和停止以及肿瘤变小。

本文所用术语“瘤细胞”、“瘤形成”、“肿瘤”、“肿瘤细

胞”、“癌”和“癌细胞”（可互换使用）是指表现出相对自主生长的细胞，以至它们显示出以细胞增殖显著失控（即失调细胞分裂）为特征的异常生长表型。瘤细胞可是恶性或良性的。

“宿主细胞”包括可以或已经接受了本发明腺病毒载体的单一细胞或细胞培养物。宿主细胞包括单个宿主细胞的后代，而且由于天然、意外或有意的突变和/或改变，该后代可能不一定与初始亲本细胞（在形态学或总 DNA 互补上）完全相同。宿主细胞包括体内或体外转染或感染了本发明腺病毒载体的细胞。

“复制”或“增殖”是可互换使用的，是指本发明腺病毒载体的复制或繁殖能力。这些术语均是本领域熟知的。对于本发明，复制涉及腺病毒蛋白的产生，并且通常是指腺病毒的复制。复制可采用本领域标准方法和本文所述方法测量，例如爆发集落测定、噬菌斑测定。

“复制”和“增殖”包括直接或间接参与病毒制造过程的任何活动，包括但不限于病毒基因表达；病毒蛋白、核酸或其他成分的产生；病毒成分包装成完整病毒；和细胞裂解。

“ADP 编码序列”是编码 ADP 或其功能片段的多核苷酸。对于 ADP，就腺病毒复制而言，ADP 的“功能性片段”是指表现出细胞毒活性特别是细胞裂解活性的片段。测量细胞毒活性的方法是本领域已知的，并在本文中有描述。

“编码”ADP 多肽的多核苷酸是能被转录和/或翻译产生 ADP 多肽或其片段的多核苷酸。这种多核苷酸的反义链也被称作编码该序列。

“ADP 多肽”是含有 ADP 氨基酸序列的至少一个部分或区域（见例如 SEQ ID NO: 5），并且表现出 ADP 有关功能，特别是细胞毒性，更特别是细胞裂解功能的多肽。正如本文所讨论的，这些功能可采用本领域已知的技术测量。应当理解，由于例如保守氨基酸替代，可使用某些序列变体提供 ADP 多肽。

“描述于”SEQ ID NO 中的多核苷酸序列是指该序列具有与 SEQ ID NO 中的序列相同的连续序列。该术语包括 SEQ ID NO 的部分或区域以及 SEQ ID NO 中所包含的完整序列。

“生物学样品”包括获自个体的各种样品类型，并能用于诊断或监测分析。该定义包括血液和其它生物来源的液体样品、固体组织样品如活检标本或组织培养物或由此得到的细胞，以及它们的后代。该定义也包括获得后已被任何方式进行过操作的样品，这些操作方式如用试剂处理、溶解或富集某些成分如蛋白质或多核苷酸。术语“生物学样品”包括临床样品，也包括培养细胞、细胞上清、细胞裂解物、血清、血浆、生物液体和组织样品。

“个体”是指脊椎动物，优选哺乳动物，更优选人。哺乳动物包括但不限于畜牧动物、体育动物、啮齿类、灵长类和宠物。

“有效量”是足以实现有利或期望结果的量，所述结果可包括临床结果。有效量可通过一次或多次给药来施用。对于本发明，腺病毒载体的有效量是足以减轻、改善、稳定、逆转、减缓或延迟疾病状态发展的量。

包含细胞状态特异性 TRE 的腺病毒载体

本发明提供含有处于细胞状态特异性 TRE 转录控制下的腺病毒基因的腺病毒载体结构。优选地，该腺病毒基因可促成细胞毒性（无论直接和/或间接），更优选该基因促成或引起细胞死亡，甚至更优选该基因是腺病毒复制所必需的。可促成细胞毒性的基因的例子包括但不限于腺病毒死亡蛋白（ADP；讨论如下）。当腺病毒载体能在靶细胞即允许或诱导细胞状态特异性 TRE 发挥功能的细胞中进行选择性地（即优先地）复制繁殖时，这些细胞将优先由于腺病毒增殖而被杀死。一旦靶细胞由于选择性细胞毒和/或细胞裂解复制被破坏，腺病毒载体的复制将显著降低，因此降低了感染和不良副效应的概率。可维持体外培养物以检测混合物（如活检组织或其它适当生物样品）中是否出现（即存在）和/或重现靶细胞，如瘤细胞或其它不良细胞。为进一步保证细胞毒性，还可存在一或多个具有细胞毒效应的转基因，并使其处于选择性转录控制之下。在该实施方案中，更加可以确信靶细胞将被破坏。此外，或作为选择方案，可在腺病毒载体中包括未受到选择性

转录控制或处于选择性转录控制之下的促成细胞毒性和/或细胞死亡的腺病毒基因（如 ADP）。

本发明腺病毒载体中使用的细胞状态特异性 TRE 可获自任何物种，优选哺乳动物。一些响应细胞状态或与细胞状态相关而表达的基因已有过描述。可使用任何这些细胞状态相关基因来产生细胞状态特异性 TRE。

细胞状态的一个例子是细胞周期。其表达与细胞周期有关的典型基因是 E2F-1，一种广泛表达的生长调节基因，该基因在 S 期显示出最高转录活性。Johnson 等（1994）*Genes Dev.* 8: 1514-1425。RB 蛋白以及 RB 家族的其它成员，与 E2F-1 形成特异性复合物，从而抑制了它激活转录的能力。因此，对 E2F-1 起反应的启动子受到 RB 的下调。许多肿瘤细胞破坏了 RB 的功能，这可导致应答 E2F-1 的启动子的去抑制，反过来又导致细胞分裂的失调。

因此，在一个实施方案中，本发明提供其中腺病毒基因（优选复制必需基因）处于细胞状态特异性 TRE 的转录控制之下的腺病毒载体，其中该细胞状态特异性 TRE 包含细胞周期激活的或细胞周期特异的 TRE。在一个实施方案中，细胞周期激活的 TRE 是 E2F1 TRE。在一个实施方案中，该 TRE 包含图 3 和 SEQ ID NO:2 中描述的序列。

另一组被细胞状态调节的基因是其表达响应低氧状态而增加的基因。Bunn 和 Poyton（1996）*Physiol. Rev.* 76:839-885；Dachs 和 Stratford（1996）*Br. J. Cancer* 74:5126-5132；Guillemin 和 Krasnow（1997）*细胞 (Cell)* 89:9-12。许多肿瘤的血供应不足，部分原因是肿瘤细胞典型地比构成血管的内皮细胞生长快，从而导致肿瘤中出现低氧区。Folkman（1989）*J. Natl. Cancer Inst.* 82:4-6；和 Kallinowski（1996）*The Cancer J.* 9:37-40。低氧反应的一个重要介导者是转录复合物 HIF-1 或低氧诱导因子 1，其可与数个基因的调节区中的低氧效应元件（HRE）相互作用，这些基因包括血管内皮生长因子和几个编码糖酵解酶包括烯醇酶 1 的基因。小鼠 HRE 序列已被鉴定和进行了特性分析。Firth 等（1994）美国国家科学院院刊

91:6496-6500。一个来自大鼠烯醇酶 1 启动子的 HRE 描述于 Jiang 等 (1997) 癌症研究 (Cancer Res.) 57:5328-5335。一个来自大鼠烯醇化酶 1 启动子的 HRE 描述于图 2 和 SEQ ID NO:1 中。

因此, 在一个实施方案中, 腺病毒载体包含处于含 HRE 的细胞状态特异性 TRE 转录控制下的腺病毒基因, 优选复制必需的腺病毒基因。在一个实施方案中, 该细胞状态特异性 TRE 含有图 2 和 SEQ ID NO:1 中描述的 HRE。

其它细胞状态特异性 TRE 包括热诱导 (即热休克) 启动子、和应答辐射暴露包括电离辐射和 UV 辐射的启动子。例如, 早期生长反应 1 (Egr-1) 基因的启动子区域含有可被电离辐射诱导的元件。Hallahan 等 (1995) Nat. Med. 1:786-791; 和 Tsai-Morris 等 (1989) 核酸研究 16:8835-8846。包括热诱导元件的热诱导启动子已有描述。见例如 Welsh (1990) “生物学和医学中的应激蛋白 (Stress Proteins in Biology and Medicine)”, Morimoto, Tisseres, 和 Georgopoulos 编, Cold Spring Harbor Laboratory Press; 和 Perisic 等 (1989) 细胞 (Cell) 59:797-806。因此, 在一些实施方案中, 细胞状态特异性 TRE 含有电离辐射效应元件。在一个实施方案中, 该 TRE 含有 Egr-1 基因的 5'侧翼序列。在其它实施方案中, 细胞状态特异性 TRE 含有热休克效应元件, 或热诱导元件。

细胞状态特异性 TRE 还可含有多聚体。例如, 一个 HRE 可含有至少两个、至少三个、至少四个或至少五个串联低氧效应元件。这些多聚体还可含有异源启动子和/或增强子序列。

细胞状态特异性 TRE 可以有也可以没有沉默子。沉默子 (即负调控元件) 的存在可辅助关闭非许可细胞 (即处于正常状态的细胞) 中的转录 (以及由此关闭复制)。因此, 沉默子的存在可通过更有效地阻止非靶细胞中的腺病毒载体复制, 赋予更强的细胞状态特异性复制。另一方面, 缺少沉默子可有助于靶细胞中复制的实现, 因此由于靶细胞中更为有效的复制产生了增强的细胞状态特异性复制。

在其它实施方案中, 腺病毒载体含有一个处于第一个细胞状态特

异性 TRE 控制下的腺病毒复制必需的腺病毒基因，和第二个处于第二个细胞状态特异性 TRE 控制下的腺病毒复制必需的腺病毒基因。该第一个和第二个细胞状态特异性 TRE 可相同也可不同，并且可基本一致也可不一致。“基本一致”是指这两个 TRE 间的序列一致性的必要程度。这些 TRE 间的序列一致程度达到至少约 80%、优选至少约 85%、更优选至少约 90%、甚至更优选至少约 95%、尤其更优选至少约 98%、以及最优选 100%。序列一致性可通过使用即本领域已知的序列对齐程序进行序列比较来确定，这些程序如描述于“当代分子生物学实验手册”（Current Protocols in Molecular Biology）（F.M. Ausubel 等编，1987）增刊 30，7.7.18 节，表 7.7.1 中的程序。优选的对齐程序是 ALIGN Plus（科学和教育软件 Scientific and Educational Software, Pennsylvania），优选使用默认参数。此外，严谨条件下的杂交也可显示序列一致性程度。严谨条件是本领域已知的；严谨条件的一个例子是 80℃（或更高温度）和 6X SSC（或更低浓度的 SSC）。（为增加严谨性）其它杂交条件和参数是：解育温度为 25℃、37℃、50℃和 68℃；缓冲液浓度为 10XSSC、6XSSC、1XSSC、0.1XSSC（其中 1XSSC 是 0.15M NaCl 和 15mM 柠檬酸缓冲液）以及使用其它缓冲体系的它们的等价物；甲酰胺浓度为 0%、25%、50%和 75%；解育时间从约 24 小时到约 5 分钟；1、2 或更多次洗涤步骤；洗涤解育时间为 1、2 或 15 分钟；洗涤液为 6XSSC、1XSSC、0.1XSSC 或去离子水。

第一个和第二个细胞状态特异性 TRE 一致或基本一致的腺病毒结构，尤其是其中这些 TRE 控制着早期基因（如 E1A 和 E1B）转录的腺病毒结构可能表现出不稳定性，这种不稳定性在某些情况下可能是有利的，例如当自主的“自体破坏”性质可关闭病毒时，由此即可控制繁殖程度。因此，在一些实施方案中，与相互之间具有较少一致性的两个 TRE（即具有更多的序列差异或不同）比较，第一个和第二个细胞状态特异性 TRE，或第一个和第二个 TRE（其中之一是细胞状态特异性 TRE）具有充分一致性以引起不稳定性。优选的实施方案中这两个 TRE 分别控制着 E1A 和 E1B。“不稳定性”是指当病毒在细胞中复制

时没有保持腺病毒载体结构的完整性，该不稳定性可采用本领域的标准方法例如 Southern 分析来测定。在其它实施方案中，第一个和第二个 TRE 充分不同并/或位于该载体中以便该载体稳定（即可保持该腺病毒载体结构的完整性）。

在其它实施方案中，腺病毒载体含有一个处于第一个细胞状态特异性 TRE 控制下的腺病毒复制必需的腺病毒基因，和一个处于第二个细胞状态特异性 TRE 控制下的转基因。第一个和第二个细胞状态特异性 TRE 相互之间可基本一致也可不一致。

在一些实施方案中，细胞状态特异性 TRE 可与另一个 TRE 如不同的细胞状态特异性 TRE 或细胞类型特异性 TRE 并置。“并置”是指一个细胞状态特异性 TRE 和第二个 TRE 控制着相同基因的转录，或至少两者与同一个基因贴近。对于这些实施方案，该细胞状态特异性 TRE 和第二个 TRE 可采取许多构型，包括但不限于(a)相互邻近（即相邻）；(b)均位于被转录控制基因的 5'端（即它们之间可有间插序列）；(c)一个 TRE 位于基因的 5'端，而另一个 TRE 位于基因的 3'端。例如，实施例 1 描述和图 1 显示的，细胞类型特异性 TRE 可与细胞状态特异性 TRE 并置以控制可操作地连接的腺病毒基因的转录。可使用这些复合 TRE 以赋予可操作地连接的多核苷酸同时细胞状态特异性和细胞类型特异性地表达，并由此产生显著增加的特异性和/或效力。以下给出了细胞类型特异性 TRE 的例子。此外，“复合”TRE 可用于赋予不同，但可能是协同的，细胞状态特异性。例如，一个复合 TRE 可赋予对低氧和热休克的特异性。

实施例 1 描述了在 E1A 上游有一个复合 TRE 的腺病毒结构，该复合 TRE 由一个 HRE 和一个前列腺特异性 TRE，PAS-TRE，（由与第-541 至+12 位 PSA 启动子序列融合的第-5322 至-3738 位增强子序列组成；见美国专利 5,871,726；5,648,478）构成。因此，在一些实施方案中，本发明提供含有处于 TRE 转录控制下的复制必需腺病毒基因的腺病毒载体，其中所述腺病毒基因优选早期基因，优选 E1A 或 E1B，所述 TRE 含有 HRE（优选含有 SEQ ID NO:1 描述的 67 个碱基片段或由其构成）

和前列腺细胞特异性 TRE, 该前列腺细胞特异性 TRE 优选含有一个 PSA 增强子 (优选第-5322 至-3738 位序列; 或 SEQ ID NO:3 的约第 503 位至约 2086 位 (图 4 的第约 503 至约 2086 位碱基)) 和一个启动子, 优选含有一个 PSA 增强子和一个 PSA 启动子 (SEQ ID NO:3 的第约 5285 至约 5836 位)。

正如本领域技术人员易于理解的, 细胞状态特异性 TRE 是多核苷酸序列, 而且就这一点而言其可不受多种序列改变的影响而表现功能。核苷酸替换、添加和缺失的方法是本领域已知的, 并且易于进行的功能性分析 (例如 CAT 或荧光素酶报告基因分析) 允许普通技术人员确定序列变异体是否表现出必要的细胞状态特异性转录功能。因此, 本发明还包括本文公开的核酸序列的功能保持的变体, 这些变体包括核苷酸替换、添加和/或缺失。尽管不希望被单一的理论所束缚, 但本发明者注意到可能某些修饰会导致所获表达水平的改变, 包括表达水平的增强。对某些更具有侵袭性的癌症形式, 或是当 (例如由于个体的无免疫应答状态) 需要更为快速和/或攻击性的方式杀伤细胞时, 可能尤其期望实现所获表达水平的改变, 优选表达水平的增加。

作为如何测定细胞状态特异性 TRE 活性的一个例子, 可采用本领域已知的方法例如化学合成、定点诱变、PCR 和/或重组方法, 产生多核苷酸序列或一组这样的序列。将待测定的序列插入含有合适报告基因的载体中, 所述报告基因包括但不限于氯霉素乙酰转移酶 (CAT)、 β -半乳糖苷酶 (由 lacZ 基因编码)、荧光素酶 (由 luc 基因编码)、绿色荧光蛋白、碱性磷酸酶、和辣根过氧化物酶。这些载体和分析均易于获得, 尤其可从商业途径获得。采用本领域已知的转染方法, 如磷酸钙沉淀、电穿孔、脂质体 (脂转染)、和 DEAE-葡聚糖, 将由此构建的质粒转染至合适的宿主细胞中, 以检测推测的细胞状态特异性 TRE 所控制的报告基因的表达。合适的宿主细胞包括任何细胞类型, 包括但不限于 Hep3B, Hep G2, HuH7, HuH1/C12, LNCaP, HBL-100, Chang 肝细胞, MCF-7, HLF, HLE, 3T3, HUVEC, 和 HeLa 细胞。把待测 TRE-报告基因结构转染的宿主细胞置于导致细胞状态改变的条

件下（例如导致异常生理学状态的条件）。未经这些条件处理的相同细胞，即处于正常生理学条件并因此处于正常生理学状态的细胞充当对照。通过使用标准分析测量报告基因的表达水平获得结果。处于特定状态的细胞和对照细胞之间的表达比较结果反映转录激活的有或无。“转录激活”已在上文中进行了定义。

各种细胞状态特异性 TRE 之间的比较可通过例如测定和比较处于正常和异常生理状态下的单个细胞系中的表达水平来进行。这些比较也可通过测定和比较处于可逆环境条件（例如热）的单个细胞系与未经这些条件处理的细胞中的表达水平来进行。应该理解，细胞状态特异性 TRE 的绝对转录活性取决于几个因素，例如靶细胞的性质，细胞状态特异性 TRE 的递送方式和形式，以及被选择性转录激活的编码序列。为消除所用各种质粒的大小影响，活性可表示为每摩尔转染质粒的相对活性。此外，转录水平（即 mRNA）可采用标准 Northern 分析和杂交技术测量。转染水平（即转染效率）通过共转染含有处于不同 TRE，例如巨细胞病毒 (CMV) 的立即早期启动子，控制下的编码不同报告基因的质粒来测量。该分析还可显示负调控区域即沉默子。

对于如何测定低氧诱导的例子，可使用例如描述于 Jiang 等 (1997) 癌症研究 (Cancer Research) 57:5328-5335 或 Dachs 等 (1997) Nature Med. 3:515-520 中的分析方法。例如，将含有推测的 HRE 或其多个串联拷贝以及一个最小启动子元件的结构导入宿主细胞，其中所述结构可操作地连接并控制可检测蛋白或可用于给出可检测信号的蛋白的编码多核苷酸的转录。然后，将该宿主细胞置于常氧条件下（例如 20% O_2 ），以及不同程度的低氧条件下例如 5%、2%、1%、0.1% 或更少的 O_2 。然后，测量可操作地连接的多核苷酸（报告基因）的表达产物。

另外，可评价推测的细胞状态特异性 TRE 给表现出必要生理状态如热或电离辐射的细胞赋予腺病毒复制优先性的能力。对于该分析，用含有可操作地与推测的细胞状态特异性 TRE 连接的复制必需腺病毒基因的结构转染表现出必要生理状态的细胞。病毒在这些细胞中的复制与该结构在例如处于正常生理情况下的细胞（即没有表现出必要生

理状态的细胞)中引起的病毒复制相比较。

可以使用各种血清型的腺病毒,例如 Ad2、Ad5、Ad12 和 Ad40。为了进行阐述,本文采用血清型 Ad5 作为示例。

当细胞状态特异性 TRE 与繁殖必需的腺病毒基因一起使用时,复制能力优先在表现细胞状态的靶细胞中实现。优选地,该基因是早期基因,例如 E1A、E1B、E2 或 E4。(E3 对于病毒复制是非必需的。)更优选,处于细胞状态 TRE 控制下的早期基因是 E1A 和/或 E1B。可将不止一个早期基因置于一个细胞状态特异性 TRE 的控制之下。实施例 1 提供了这种结构的更为详细的描述。

E1A 基因在病毒感染之后(0-2 小时)立即表达,早于任何其它的病毒基因。E1A 蛋白起正向转录调控反式作用因子的作用,对于其它早期病毒基因 E1B、E2、E3、E4 和启动子近端的主要晚期基因的表达是必需的。尽管如此命名,但由主要晚期启动子驱动的启动子近端的基因在 Ad5 感染之后的早期表达。Flint(1982) *Biochem. Biophys. Acta* 651:175-208; Flint (1986) 病毒研究进展(*Advances Virus Research*) 31:169-228; Grand (1987) 生物化学杂志(*Biochem. J.*) 241:25-38。当缺乏功能性 E1A 基因时,病毒感染无法进行,因为无法产生病毒 DNA 复制必需的基因产物。Nevins (1989) *Adv. Virus Res.* 31:35-81。病毒基因组中 Ad5 E1A 的转录起始位点位于第 498 位, E1A 蛋白的 ATG 起始位点位于第 560 位。

E1B 蛋白以反式作用方式起作用,是晚期 mRNA 从细胞核向细胞质运输所必需的。E1B 表达缺陷导致不良的晚期病毒蛋白表达,而且致使不能关闭宿主细胞蛋白的合成。E1B 启动子已被认为可能是 Ad40 和 Ad5 宿主范围差异的决定元件:临床上,Ad40 是肠道病毒,而 Ad5 引起急性结膜炎。Bailey、Mackay 等(1993)病毒学(*Virology*) 193:631; Bailey 等(1994)病毒学 202:695-706)。Ad5 的 E1B 启动子由一个单一的 Sp1 高亲和性识别位点和一个 TATA 盒构成。

腺病毒的 E2 区编码腺病毒基因组复制相关的蛋白,包括 72kDa DNA 结合蛋白、80kD 前体末端蛋白和病毒 DNA 聚合酶。Ad5 的 E2 区从两个

启动子向右转录，这两个启动子称为 E2 早期和 E2 晚期启动子，分别位于 76.0 和 72.0 图距单位。E2 晚期启动子在感染后期阶段瞬时激活，并且不依赖于 E1A 反式激活蛋白，而 E2 早期启动子在病毒复制早期阶段是关键。

E2 晚期启动子与相对链编码的基因的编码序列相重叠，因此不利于遗传操作。然而，E2 早期启动子仅与对侧链上编码 33kD 蛋白的序列有几个碱基对的重叠。值得注意的是，SpeI 限制性位点（Ad5 的第 27082 位）是上面提及的 33kD 蛋白的终止密码子的部分，其方便地把主要 E2 早期转录起始位点以及 TATA 结合蛋白位点与上游转录因子 E2F 和 ATF 的结合位点分隔开。因此，具有 SpeI 末端的细胞状态 TRE 在+链中 SpeI 位点的插入将破坏 Ad5 的内源性 E2 早期启动子，而且将允许 E2 转录物的细胞状态限制性表达。

E4 基因有许多转录产物。E4 区编码两个多肽，这两个多肽负责刺激病毒基因组 DNA 的复制和晚期基因的表达。开放阅读框 (ORF) 3 和 6 的蛋白产物均能通过结合 E1B 的 55kD 蛋白以及 E2F 与 DP-1 形成的异二聚体发挥这些功能。ORF6 蛋白需要与 E1B 的 55kD 蛋白相互作用才具有活性，而 ORF3 蛋白却不需要。在缺少 ORF3 和 ORF6 的功能性蛋白时，噬菌斑的产生效率低于野生型病毒的 10^{-6} 。为进一步将病毒复制限制于表现必要生理情况或状态的细胞中，可缺失 E4 的 ORF1-3，使病毒 DNA 复制和晚期基因合成依赖于 E4 的 ORF6 蛋白。通过该突变与其中 E1B 区受到细胞状态特异性 TRE 调节的序列结合，可获得 E1B 功能和 E4 功能均依赖于驱动 E1B 的细胞状态特异性 TRE 的病毒。

与本发明相关的主要晚期基因是编码腺病毒病毒体蛋白的基因 L1、L2、L3、L4 和 L5。所有这些基因（典型地编码结构蛋白）可能是腺病毒复制所必需的。晚期基因均处于主要晚期启动子 (MLP) 的控制之下，该启动子位于 Ad5 的 +5986 至 +6048 位。

除了凭借表现必要生理状态（例如，异常生理状态如低氧情况）的细胞中的优先复制能力赋予选择性细胞毒性和/或细胞裂解活性外，本发明的腺病毒载体可进一步包括处于细胞状态特异性 TRE 控制下的

异源基因（转基因）。通过这种方式，可将各种遗传性能引入靶细胞尤其是癌细胞。例如，在某些情况中，由于如癌性靶细胞的相对顽固性质或不寻常的侵袭性，可能期望增强细胞毒活性的程度和/或比率。这可通过偶联细胞状态特异的复制性细胞毒活性与例如 HSV-tk 和/或胞嘧啶脱羧酶（cd）的细胞特异性表达来实现，其中胞嘧啶脱羧酶使细胞能够将 5-氟胞嘧啶（5-FC）代谢成化疗试剂 5-氟尿嘧啶（5-FU）。这些类型转基因的使用还可赋予附加效应。

可通过腺病毒载体引入的其它有价值的转基因包括编码细胞毒性蛋白如白喉毒素、蓖麻毒素或相思豆毒素 A 链的基因（Palmiter 等（1987）细胞（Cell）50:435；Maxwell 等（1987）分子细胞生物学（Mol. Cell Biol.）7:1576；Behringer 等（1988）Genes Dev. 2:453；Messing 等（1992）Neuron 8:507；Piatak 等（1988）生物化学杂志（J. Biol. Chem.）263:4937；Lamb 等（1985）Eur. J. Biochem. 148:265；Frankel 等（1989）分子细胞生物学 9:415），编码能触发细胞凋亡的因子的基因，编码反义转录物或核酶的序列（除了其它功能外，可针对编码增殖必需蛋白如结构蛋白或转录因子的 mRNA）；病原体在细胞内增殖的病毒或其它病原性蛋白；编码核酸酶（如 RNase A）或蛋白酶（如 amsin、木瓜蛋白酶、蛋白酶 K、羧肽酶等）的工程性细胞质变体的基因，或编码 Fas 的基因，等等。其它有意义的基因包括细胞因子、抗原、跨膜蛋白等等，例如 IL-1、-2、-12、GM-CSF、G-CSF、M-CSF、IFN- α 、- β 、- γ 、TNF- α 、- β 、TGF- α 、- β 、NGF 等。正效应物基因可在更早的时期，在因复制导致的细胞毒活性出现之前使用。

在一个实施方案中，腺病毒载体中维持有 E3 区编码的腺病毒死亡蛋白（ADP）。处于主要晚期启动子（MLP）控制之下的 ADP 基因可能编码在促使宿主细胞裂解中起重要作用的蛋白（ADP）。Tollefson 等（1996）病毒学杂志（J. Virol.）70(4):2296；Tollefson 等（1992）病毒学杂志 66(6):3633。因此，含有 ADP 基因的腺病毒载体可使该腺病毒更有效力，使更为有效的治疗和/或更低的剂量要求成为可能。

因此，本发明提供进一步包括 ADP 编码多核苷酸序列的本文所述

病毒载体。编码 ADP 的 DNA 序列和 ADP 的氨基酸序列描述于图 9。简单地说，可采用本领域已知的技术如 PCR 优选地从 Ad2 获得 ADP 编码序列（因为对该链中的 ADP 已有更为完善的特性分析）。优选地，进一步获得 Y 前导序列（对于晚期基因的正确表达重要的序列），并将其与 ADP 编码序列连接。然后，可将 ADP 编码序列（有或无 Y 前导序列）引入腺病毒基因组，例如 E3 区中（在该区域 ADP 编码序列将由 MLP 驱动）。也可将 ADP 编码序列插入腺病毒基因组的其它位置，例如 E4 区。或者，可将 ADP 编码序列可操作地与（有或无增强子的）异源启动子连接，这些异源启动子包括但不限于另一种病毒启动子、细胞状态特异 TRE 如低氧效应元件、或细胞类型特异性 TRE 如来源于癌胚抗原 (CEA)、粘蛋白、和大鼠 probasin 基因的细胞类型特异性 TRE。

进一步包含细胞类型特异性元件的本发明腺病毒载体

除了借助表现必要生理状态的细胞中的优先复制能力和/或编码细胞毒因子的基因的优先转录，赋予选择性细胞毒和/或细胞裂解活性外，本发明的腺病毒载体还可以包括处于细胞类型特异性 TRE 控制下的腺病毒基因和/或异源基因（转基因）。通过这种方式，细胞毒活性被进一步限制在特定细胞类型中。

例如，优先在前列腺细胞中起作用的 TRE 包括但不限于，来源于前列腺特异抗原基因的 TRE (PSA-TRE)（美国专利 No. 5,648,478），来源于腺激肽释放酶 1 基因的 TRE（来源于人基因的，hKLK2-TRE），和来源于 probasin 基因的 TRE (PB-TRE)（国际专利申请 No. PCT/US98/04132）。所有这三个基因均优先在前列腺细胞中表达，而且其表达是可由雄激素诱导的。一般，响应雄激素诱导的基因表达需要雄激素受体 (AR) 的存在。

PSA 仅由正常、增生、和恶性前列腺上皮细胞合成；因此，它的组织特异性表达使它成为一个极好的良性前列腺增生 (BPH) 和前列腺癌 (CaP) 生物标记。PSA 的正常血清水平通常低于 5ng/ml，水平的升高提示出现了 BPH 或 CaP. Lundwall 等(1987) FEBS Lett. 214:317;

Lundwall (1989) *Biochem. Biophys. Res. Comm.* 161:1151; 和 Riegmann 等(1991) *Molec. Endocrin.* 5:1921.

用于提供尤其在前列腺细胞中依赖于雄激素的细胞特异性的 PSA 基因区域涉及大约 6.0 千碱基。Schuur 等 (1996) 生物化学杂志 271:7043-7051。在人中大约 1.5kb 的增强子区域位于相对 PSA 基因转录起始位点的核苷酸第-5322 位至-3738 位。PSA 启动子由相对转录起始位点的核苷酸约第-540 位至+12 位的序列组成。这两个遗传元件的并置产生具有完整功能的最小前列腺特异性增强子/启动子(PSE)TRE。只要保证必要的功能性,也可在本发明中使用 PSA 基因大约 6.0kb 区的其它部分。在实施例 1 中描述了腺病毒载体 CN796, 该病毒载体含有含 HRE 和 PSA-TRE 的复合 TRE, 其中 PSA-TRE 含有与第-541 至+12 位的 PSA 启动子融合的第-5322 至-3738 位的 PSA 增强子。该 PSA-TRE 来自腺病毒载体 CN706。Rodriguez 等(1997) 癌症研究(Cancer Research) 57:2559-2563。因此, 在一个实施方案中, 腺病毒载体含有处于复合 TRE 转录控制之下的腺病毒 E1A 基因, 其中所述的复合 TRE 由细胞状态特异性 TRE(HRE)和细胞类型特异性 TRE(PSA-TRE)构成。

本发明中使用的 PSE 和 PSA TRE 均来自图 4 (SEQ ID NO:3) 所描述的序列。增强子元件为图 4 (SEQ ID NO:3) 的大约第 503 位至大约 2086 位核苷酸。启动子为图 4 (SEQ ID NO:3) 的大约第 5285 位至大约 5836 位核苷酸。因此, 在一些实施方案中, 复合 TRE 包含含有 SEQ ID NO:1 的 HRE 和含有 SEQ ID NO:3 的大约第 503 位至大约 2086 位核苷酸的 PSA-TRE。在其它实施方案中, 复合 TRE 包含含有 SEQ ID NO:1 的 HRE 和含有 SEQ ID NO:3 的大约第 5285 位至大约 5836 位核苷酸以及 SEQ ID NO:3 的大约第 503 位至大约 2086 位核苷酸的 PSA-TRE。正如上面所描述的, 这些复合(HRE/PSA) TRE 可操作地连接于复制必需的腺病毒基因, 尤其是早期基因如 E1A 或 E1B。实施例 1 描述了此结构。

在本发明中, 含有细胞状态特异性 TRE 和细胞类型特异性 TRE 的可复制腺病毒载体可使用优先在特定肿瘤细胞中起作用的细胞类型特

异性 TRE。肿瘤细胞特异性 TRE 和相应的潜在靶细胞的非限制性例子包括来自如下基因的 TRE: 甲种胎儿球蛋白 (AFP) (肝癌), 粘蛋白样糖蛋白 DF3 (MUC1) (乳癌), 癌胚抗原 (CEA) (结肠直肠、胃、胰腺、乳腺、和肺癌), 纤溶酶原激活物尿激酶 (uPA) 和其受体基因 (乳腺、结肠和肝癌), HER-2/neu(c-erbB2/neu) (乳腺、卵巢、胃、和肺癌)。

其它细胞类型特异性 TRE 可来自例如下述基因 (括号内是 TRE 特异性地在其中起作用的细胞类型): 血管内皮生长因子受体 (内皮), 白蛋白 (肝), 因子 VII (肝), 脂肪酸合成酶 (肝), von Willebrand 因子 (大脑内皮), α 肌动蛋白和肌球蛋白重链 (均在平滑肌内), 合成酶 I (小肠), Na-K-Cl 转运体 (肾)。其它细胞类型特异性 TRE 是本领域已知的。

AFP 是一种癌胚蛋白, 其表达主要限制在内胚层来源的发育组织 (卵黄囊、胚肝、和肠) 中, 尽管其表达水平在不同的组织和发育阶段有极大的变化。AFP 具有临床意义, 因为 AFP 的血清浓度在大多数肝细胞瘤患者中升高, 在晚期疾病患者中可发现高水平的 AFP。患者中的 AFP 血清水平似乎受肝细胞癌中的 AFP 表达调节, 而不是周围正常肝中的。因此, AFP 基因似乎被调节为肝肿瘤细胞特异性表达。

AFP 基因来源的细胞类型特异性 TRE 已获得鉴定。例如, 人 AFP 特异性增强子活性的克隆和性质分析描述在 Watanabe 等 (1987) 生物化学杂志 262:4812-4818。完整的 5'AFP 侧翼区 (含有启动子、推测的沉默子, 和增强子元件) 包含在转录起始位点上游的大约 5kb 序列中。

人的 AFP 增强子区位于相对于 AFP 基因转录起始位点的大约第 -3954 至大约 -3335 位核苷酸之间。人 AFP 启动子包括从大约第 -174 位至大约第 +29 位核苷酸的区域。这两个遗传元件的并置产生完整功能的 AFP-TRE。Ido 等 (1995) 描述了一个 HCC 特异的 259bp 启动子片段 (核苷酸第 -230 位至 +29 位) (癌症研究 55:3105-3109)。AFP 增强子含有两个区, 表示为 A 和 B, 位于相对起始转录位点的核苷酸第 -3954

至-3335位。启动子区含有典型的TATA和CAAT盒。优选地，AFP-TRE含有至少一个增强子区。更优选，AFP-TRE含有两个增强子区。

对于含有AFP-TRE的腺病毒载体，适合的靶细胞可以是任何允许AFP-TRE发挥功能的细胞类型。优选表达或产生AFP的细胞，包括但不限于表达AFP的肿瘤细胞。这类细胞的例子有肝细胞癌细胞、生殖腺和其它生殖细胞肿瘤（尤其是内胚窦瘤）、脑瘤细胞、卵巢肿瘤细胞、胰腺腺泡细胞癌(Kawamoto等(1992) *Hepatogastroenterology* 39:282-286)、原发性胆囊肿瘤(Katsuragi等(1989) *Rinsko Hoshasen* 34:371-374)、子宫内膜腺癌细胞(Koyama等(1996) *Jpn. J. Cancer Res.* 87:612-617)、和任何前述肿瘤的转移瘤（转移瘤可在肺、肾上腺、骨髓和/或脾中发生）。在某些情况中，从一些胰腺和胃癌向肝转移的疾病产生AFP。尤其优选肝细胞癌细胞和它的任何转移瘤。AFP的产生可采用本领域标准分析如RIA、ELISA或Western印迹杂交（免疫分析方法）确定AFP蛋白产生水平，或Northern印迹杂交确定AFP mRNA的产生水平来测量。或者，可通过细胞激活AFP-TRE转录（即允许AFP-TRE发挥功能）的能力对这些细胞进行鉴定和/或特性分析。

尿激酶纤溶酶原激活物(uPA)蛋白和它的细胞表面受体尿激酶纤溶酶原激活物受体(uPAR)在许多最为频繁发生的瘤形成中表达，似乎是癌症转移中的重要蛋白。这两个蛋白均参与了乳腺、结肠、前列腺、肝、肾、肺和卵巢癌。调节uPA和uPAR转录的转录调节元件已有深入研究。Riccio等(1985)核酸研究(*Nucleic Acids Res.*) 13:2759-2771; Cannio等(1991)核酸研究 19:2303-2308。

CEA是一个180,000道尔顿的肿瘤相关抗原糖蛋白，存在于胃肠道的内皮来源瘤形成例如结肠直肠、胃和胰腺癌，以及其它腺癌例如乳腺和肺癌的上面。CEA具有临床意义，因为在极大多数CEA阳性肿瘤患者中可检测到循环CEA。在肺癌中，全部病例的大约50%有循环CEA，在腺癌中通常检测到高浓度的CEA（大于20ng/ml）。大约50%的胃癌患者是CEA血清学阳性的。

已显示CEA基因的5'上游侧翼序列赋予细胞特异性活性。CEA启

动子区位于该基因 5'侧翼区中, 为翻译起始位点上游的大约头 424 个核苷酸, 该区域在产生 CEA 的细胞中比在不产生 CEA 的 HeLa 细胞中提供更高的启动子活性, 这显示该区域可赋予细胞特异性活性。Schrewe 等(1990)分子细胞生物学 10:2738-2748。此外, 还发现了细胞特异性增强子区(WO/95/14100)。完整的 5'CEA 侧翼区(含有启动子、推测的沉默子、和增强子元件)似乎包含在转录起始位点上游的大约 14.5kb 中。Richards 等(1995); WO 95/14100。Richards 等(1995)对 CEA 基因 5'侧翼区的进一步分析显示有两个上游区, -13.6 至 -10.7kb 或 -6.1 至 -4.0kb, 当与多聚启动子连接时可在产生 CEA 的 LoVo 和 SW1463 细胞中引起报告基因结构的高水平选择性表达。Richards 等(1995)还将该启动子区定位在相对转录起始位点的核苷酸第-90 位至+69 位, 其中核苷酸第-41 位至-18 位区是表达必需的。WO95/14100 描述了一系列赋予细胞特异性活性的 5'侧翼 CEA 片段, 例如核苷酸约第-299 至+69 位; 核苷酸约第-90 至+69 位; 核苷酸约第-14,500 至-10,600 位; 核苷酸约第-13,600 至-10,600 位; 核苷酸约第-6100 至-3800 位。此外, 细胞特异转录活性可通过描述于 SEQ ID NO:6 的核苷酸第-402 至+69 位 CEA 片段赋予可操作地连接的基因。本发明使用的任何 CEA-TRE 均来自哺乳动物细胞, 包括但不限于人细胞。因此, 只要可在腺病毒载体中表现出必要的期望功能, 任何 CEA-TRE 均可在本发明中使用。文献中已描述过 CEA 序列的克隆和特性分析, 因此可用于本发明的实施并且无需在本文中作详细描述。

MUC1 基因的蛋白产物(称为粘蛋白或 MUC1 蛋白; episialin; 多形上皮细胞粘蛋白或 PEM; EMA; DF3 抗原; NPGP; PAS-O; 或 CA15.3 抗原)正常主要表达在衬在腺体或胃、胰腺、肺、气管、肾、子宫、唾液腺和乳腺的导管中的上皮细胞的顶表面。Zotter 等(1988)Cancer Rev. 11-12:55-101; 和 Girling 等(1989)Int. J. Cancer 43:1072-1076。然而, 在 75%-90%的人乳癌中粘蛋白过表达。Kufe 等(1984)Hybridoma 3:223-232。综述见 Hilkens (1988) Cancer Res. 11-12:25-54; 和 Taylor-Papadimitriou 等(1990) J. Nucl. Med.

Allied Sci. 34:144-150. 粘蛋白的表达与乳腺肿瘤的分化程度相关。Lundy 等(1985) Breast Cancer Res. Treat. 5:269-276. 该过表达似乎受控于转录水平上。

人乳腺癌细胞 MCF-7 和 ZR-75-1 中的 MUC1 基因过表达似乎在转录水平上受到调节。Kufe 等(1984); Kovarik (1993) 生物化学杂志 268:9917-9926; 和 Abe 等(1993) J. Cell. Physiol. 143:226-231. MUC1 基因的调节序列已被克隆, 包括转录起始位点上游的大约 0.9kb, 其含有似乎参与细胞特异性转录的 TRE. Abe 等(1993) 美国科学院院刊(Proc. Natl. Acad. Sci. USA) 90:282-286; Kovarik 等(1993); 和 Kovarik 等(1996) 生物化学杂志 271:18140-18147.

本发明中使用的所有 MUC1-TRE 均来自哺乳动物细胞, 包括但不限于人细胞。优选地, 该 MUC1-TRE 是人的。在一个实施方案中, MUC1-TRE 可含有 MUC1 基因的完整 0.9kb 5'侧翼序列。在其它实施方案中, MUC1-TRE 含有如下序列(相对于 MUC1 基因的转录起始位点): 核苷酸约第-725 至+31 位, 核苷酸约第-743 至+33 位, 核苷酸约第-750 至+33 位, 核苷酸约第-598 至+485 位(可操作地与启动子连接)。

c-erbB2/neu 基因(HER-2/neu 或 HER)是转化性基因, 编码 185kd 表皮生长因子受体相关的跨膜糖蛋白。人的 c-erbB2/neu 基因在胚胎发育期间表达, 然而在成年人中, 在许多正常组织的上皮中(通过免疫组织化学)仅可微弱地检测到该蛋白。c-erbB2/neu 基因的扩增和/或过表达已和许多人类癌症联系在一起, 这些癌症包括乳腺、卵巢、子宫、前列腺、胃和肺癌。c-erbB2/neu 蛋白过表达的临床后果在乳腺和卵巢癌中已进行了深入研究。在 20-40%的乳腺导管内癌和 30%的卵巢癌中有 c-erbB2/neu 蛋白的过表达, 这与这两种疾病亚型的预后差相关。人、大鼠和小鼠的 c-erbB2/neu TRE 已获得鉴定, 显示出赋予 c-erbB2/neu 表达细胞特异性活性。Tal 等(1987)分子细胞生物学 7:2597-2601; Hudson 等(1990)生物化学杂志 265:4389-4393; Grooteclaes 等(1994) Cancer Res. 54:4193-4199; Ishii 等(1987)美国国家科学院院刊 84:4374-4378; Scott 等(1994)生物化学杂志

269: 19848-19858.

提供以上列出的细胞类型特异性 TRE 作为在本发明中起作用的 TRE 的非限制性例子。其它细胞特异性 TRE 是本领域已知的，同样，对可疑 TRE 进行鉴定和检测其细胞特异性的方法也是本领域已知的。

本发明腺病毒载体的使用

所述腺病毒载体可以以多种形式使用，包括但不限于裸多核苷酸（通常是 DNA）结构；与利于多核苷酸进入细胞的试剂如阳离子脂质体或其它阳离子化合物如多聚赖氨酸形成复合物的多核苷酸结构；包装在感染性腺病毒颗粒中（该颗粒可使该腺病毒载体具有更强的免疫原性）；包装在其它颗粒病毒形式如 HSV 或 AAV 中；与增强或减弱免疫应答的试剂（如 PEG）形成复合物；与利于体内转染的试剂如 DOTMATM、DOTAPTM、和多胺形成复合物。因此，本发明还提供可优先在产生细胞状态的细胞中复制的腺病毒载体。“优先复制”是指腺病毒载体在表现必要生理状态的细胞中比在不表现该状态的细胞中复制更有效。优选地，该腺病毒复制高出至少大约 2 倍，优选至少大约 5 倍，更优选至少大约 10 倍，进而更优选至少大约 50 倍，甚至更优选至少大约 100 倍，进而更优选至少大约 400 倍-500 倍，进而更优选至少大约 1000 倍，最优选至少大约 1×10^6 倍。最优选地，该腺病毒载体仅在表现必要生理状态的细胞中复制（即，在不表现必要生理状态的细胞中不复制或以非常低的水平复制）。

如果将腺病毒载体包装在腺病毒中，还可对腺病毒本身进行选择以进一步增强靶向性。例如，腺病毒尾丝通过辅助定向运动介导与细胞受体的最初接触。见例如 Amberg 等(1997)Virol. 227:239-244。如果特定亚属的腺病毒血清型表现出对靶细胞类型的定向运动和/或对非靶细胞类型的亲和性降低，则可使用该亚属（或多个亚属）以进一步增加细胞毒性和/或细胞裂解的细胞特异性。

可采用多种途径向靶细胞递送腺病毒载体，这些途径包括但不限于脂质体、本领域已知的一般转染方法（例如磷酸钙沉淀或电穿孔）、

直接注射、和静脉内灌注。递送方式将大部分取决于特定的腺病毒载体（包括它的形式）以及靶细胞的类型和定位（即是体外还是体内细胞）。

如果以包装腺病毒的形式使用，腺病毒载体可置于合适的生理可接受载体中按每剂量大约 10^4 至大约 10^{14} 的量施用。感染复数一般在大约 0.001-100 的范围内。如果以多核苷酸结构（即未包装成病毒）的形式施用，可施用大约 0.01ug-约 1000ug 的腺病毒载体。根据预期的使用和宿主的免疫应答能力，可一次或多次施用该腺病毒载体，也可以以多个同时注射的方式施用。如果不希望有免疫应答，可通过使用各种免疫抑制剂降低免疫应答，以便允许重复施用而不出现强烈的免疫应答。如果包装为其它病毒形式如 HSV，施用剂量根据该特定病毒的标准知识（这些知识可容易地从例如出版文献中获得）确定，也可根据经验确定。

含有本发明腺病毒载体的宿主细胞

本发明还提供含有（即转化了）本文所描述的腺病毒载体的宿主细胞。原核和真核宿主细胞均可使用，只要载体中存在宿主中维持所必需的序列，例如合适的复制原点。为了方便，还提供了可筛选标记。原核宿主细胞包括细菌细胞，例如大肠杆菌和分支杆菌。真核宿主细胞有酵母、昆虫、鸟、植物和哺乳动物。宿主系统是本领域已知的，无需在本文中作详细描述。

本发明的组合物

本发明还提供含有本文所描述的腺病毒载体的组合物，包括药物组合物。这些组合物（尤其是药物组合物）对体内施用是有用的，例如在测量个体中的转导程度和/或细胞杀死效率时。在药物可接受赋形剂中含有本发明腺病毒载体（一般为有效量的腺病毒载体）的药物组合物适于按单位剂量形式给个体系统施用，所述剂量形式包括无菌注射用药物溶液或悬浮液、无菌非注射用药物溶液或口服溶液或悬浮液，

水包油或油包水乳化液等。用于注射用和非注射用药物递送的制剂是本领域已知的，陈述于 Remington 制药科学 (Remington's Pharmaceutical Sciences) 第 19 版，Mack 出版社 (1995)。药物组合物还包括冻干和/或重构形式的本发明腺病毒载体 (包括包装成病毒如腺病毒的载体)。

可使用其它组合物，它们对于本文所描述的检测方法是有益的。对于这些组合物，腺病毒载体通常悬浮在合适的溶剂或溶液如缓冲系统中。这些溶剂系统是本领域熟知的。

本发明试剂盒

本发明还包括含有本发明腺病毒载体的试剂盒。这些试剂盒可用于诊断和/或监测目的，优选检测。使用这些试剂盒的程序可由临床实验室、实验性实验室、医师或私人个体操作。本发明的试剂盒可检测适当生物学样品如活检标本中是否存在产生细胞状态的细胞。

本发明试剂盒包含存在于合适包装中的本文所述腺病毒载体。该试剂盒可选择提供其它在该程序中有用的成分，包括但不限于缓冲液、生色试剂、标签、反应表面、检测工具、对照样品、说明书和解释性资料。

本发明腺病毒载体的制备

本发明腺病毒载体可采用本领域标准的重组技术制备。一般，细胞状态特异性 TRE 插在目标腺病毒基因的 5' 端，目标腺病毒基因优选一或多个早期基因 (尽管也可使用晚期基因)。细胞状态特异性 TRE 可采用寡核苷酸合成 (如果序列已知) 或重组方法 (例如 PCR 和/或限制性酶) 制备。存在于天然腺病毒 DNA 序列中的或通过例如寡核苷酸介导的诱变和 PCR 方法引入的便利限制性位点，为细胞状态特异性 TRE 提供了插入位点。因此，可采用标准重组方法如 PCR 将用于退火 (即插入) 细胞状态特异性 TRE 的便利限制性位点加在细胞状态特异性 TRE 的 5' 和 3' 末端。

用于制备本发明腺病毒载体的多核苷酸可采用本领域标准方法获得，例如化学合成、重组方法、和/或从生物资源获得需要的序列。

腺病毒载体可通过使用两个质粒方便地制备，一个质粒提供腺病毒的左手区，另一个质粒提供右手区，为了进行同源重组这两个质粒共同享有至少大约 500 核苷酸的中间区。通过这种方式，每个质粒均可随意地独立操作，之后共转染入一个感受态宿主中，其中所述感受态宿主为腺病毒的增殖提供合适的互补基因，或用于从细胞状态特异性 TRE 进行转录起始的合适转录因子。一般可采用合适的转导方法例如阳离子脂质体将质粒导入合适的宿主细胞中如 293 细胞。或者，可通过体外连接腺病毒基因组的左右手部分构建含有所有腺病毒基因组复制必需部分的重组腺病毒衍生物。Berkner 等 (1983) 核酸研究 11:6003-6020; Bridge 等 (1989) 病毒学杂志 63:631-638。

为方便起见，可采用提供腺病毒必需部分的质粒。质粒 pXC.1 (McKinnon (1982) 基因 (Gene) 19:33-42) 含有 Ad5 的野生型左手端。pBHG10 (Bett 等 (1994) 美国科学院院刊 91:8802-8806; Microbix Biosystems Inc., Toronto) 提供 E3 缺失的 Ad5 右手端。E3 缺失为在病毒中插入未缺失内源性增强子/启动子的 3kb 细胞状态 TRE 提供了空间。Bett 等 (1994)。E3 基因位于 E4 的相对链上 (r 链)。pBHG11 提供甚至更大的 E3 缺失 (又缺失了 0.3kb)。Bett 等 (1994)。

对于早期基因的操作，病毒基因组中 Ad5 E1A 的转录起始位点在第 498 位，E1A 蛋白的 ATG 起始位点在第 560 位。该区域可用于插入细胞状态特异性 TRE。限制性位点可通过使用聚合酶链式反应 (PCR) 引入，所用引物可限制于 Ad5 基因组，或可包含携带 Ad5 基因组 DNA 的质粒的部分。例如，使用 pBR322 时，引物可使用 pBR322 主链中的 EcoRI 位点以及 Ad5 1339 位的 XbaI 位点。通过进行两步 PCR，在该区域中部重叠的引物引入一个 30 序列改变，导致产生一个单一的限制性位点，这样即可在该位点插入异源 TRE。

也可使用相同的策略插入异源 TRE 以调节 E1B。Ad5 的 E1B 启动子由一个单一的 Sp1 高亲和性识别位点和一个 TATA 盒构成。该区域从

1636 延伸至 1701。通过在该区域插入异源 TRE，即可实现 E1B 基因的靶细胞特异性转录。通过使用经调节 E1A 的异源 TRE 修饰的左手区作为模板引入异源 TRE 以调节 E1B，所获腺病毒载体将依赖于细胞状态特异性转录因子来表达 E1A 和 E1B。

同样，细胞状态特异性 TRE 可插在 E2 基因的上游，以使其表达具有细胞状态特异性。E2 早期基因位于 Ad5 图谱的 27050-27150 之间，由一个主要和一个次要转录起始位点(后者负责大约 5% 的 E2 转录物)，两个非典型的 TATA 盒，两个 E2F 转录因子结合位点和一个 AFP 转录因子结合位点构成。E2 启动子结构的详细综述见 Swaminathan 等 *Curr. Topics in Micro. and Imm.* (1995) 199(第 3 部分):177-194。

对于 E4，必须使用腺病毒基因组的右手部分。E4 的转录起始位点主要位于 35609 位，TATA 盒位于 35638 位，ORF1 的第一个 ATG/CTG 位于 35532 位。Virtanen 等(1984)病毒学杂志 51:822-831。使用任何以上策略用于其它基因，均可将细胞状态特异性 TRE 引入转录起始位点的上游。为了在 E4 区构建突变体，在可反式提供 E4 蛋白以弥补这些蛋白合成缺陷的 W162 细胞中，进行共转染和同源重组(Weinberg 等(1983)美国科学院院刊 80:5383-5386)。或者，可通过体外连接制备这些结构。

使用本发明腺病毒载体的方法

本发明腺病毒载体可用于非常多的目的，这些载体将随期望或预期的结果而不同。因此，本发明包括使用上述腺病毒载体的方法。

在一个实施方案中，提供在靶细胞(即表现出允许细胞状态特异性 TRE 起作用的必要生理状态的细胞)中赋予选择性细胞毒性的方法，所述靶细胞通常但并不必然在个体(优选人)中，该方法包括用本文所述腺病毒载体接触细胞以便该腺病毒进入细胞。细胞毒性可采用本领域标准分析方法例如染料排斥、³H 胸苷掺入、和/或裂解方法测量。

在另一个实施方案中，提供增殖对允许细胞状态特异性 TRE 发挥功能的哺乳动物细胞具有特异性的腺病毒的方法。这些方法要求腺病

毒载体与哺乳动物细胞的结合，藉此所述腺病毒获得增殖。

本发明还提供抑制肿瘤细胞生长的方法，所述肿瘤细胞通常但并不必然在个体（优选人）中，该方法包括用本发明腺病毒载体接触肿瘤细胞，以便该腺病毒载体进入肿瘤细胞并表现出对肿瘤细胞的选择性细胞毒性。肿瘤细胞的生长可通过本领域已知的任何方法进行评价，这些方法包括但不限于测量肿瘤的体积、采用³H胸苷掺入法确定肿瘤细胞是否增生、或对肿瘤细胞进行计数。

本发明还包括在生物学样品中检测靶细胞（即允许或诱导细胞状态特异性 TRE 发挥功能的细胞）的方法。这些方法对于监测个体（即哺乳动物）的临床和/或生理状态是尤其有用的，无论个体处于实验环境还是临床环境。对于这些方法，用腺病毒载体接触生物学样品的细胞，并检测腺病毒载体的复制。合适的生物学样品是指其中可能或怀疑存在表现必要生理（和/或环境）状态如异常生理状态的细胞（如处于低氧情况中并表现出低氧情况中细胞的特征性表型例如 HIF-1 的表达的细胞）的样品。一般，在哺乳动物中，合适的临床样品是怀疑其中存在表现必要生理状态的癌细胞的样品，例如处于低氧情况下的实体瘤中的细胞。这些细胞可通过例如穿刺活检或其它手术程序获得。可对待接触的细胞进行处理如选择性富集和/或增溶作用以改善分析条件。在这些方法中，靶细胞可采用本领域标准的体外检测腺病毒增殖的方法来检测。这些标准方法的例子包括但不限于爆发集落测定（测量病毒的产量）和噬菌斑测定（测量每个细胞的感染颗粒数）。还可通过测量特异腺病毒 DNA 的复制检测增殖，该方法也是标准分析方法。

提供以下实施例旨在进行说明而不是限制本发明。

实施例

实施例 1

含有处于低氧效应元件和 PSA-TRE 转录控制下的 E1A 的腺病毒载体
一般技术

人胚胎肾细胞系 293 有效表达 Ad5 的 E1A 和 E1B 基因，并表现出

高的腺病毒 DNA 转染效率。为产生重组腺病毒，用一个左端 Ad5 质粒和一个右端 Ad5 质粒共转染 293 细胞。同源重组产生具有在 293 细胞中复制所必需的遗传元件的腺病毒，293 细胞可反式提供 E1A 和 E1B 蛋白以弥补这些蛋白的合成缺陷。

采用阳离子脂质体如 Lipofectin(DOTMA:DOPE™, Life Technologies) 向 293 细胞共转染待组合的质粒。方法如下：通过合并这两个质粒，然后将质粒 DNA 溶液（500u1 无血清或其它添加物的最小必需培养基（MEM）中每种质粒 10ug）与 200u1 相同缓冲液中的 4 倍摩尔数过量脂质体混合。然后将 DNA-脂混合物置于细胞之上，37℃ 于 5% CO₂ 中孵育 16 小时。孵育后将培养基改变为含有 10% 胎牛血清的 MEM，再次将细胞于 5% CO₂ 中 37℃ 孵育 10 天，其中更换两次培养基。在此结束之后将细胞和培养基转移至试管中，冻融三次，将适当稀释的裂解物用于感染 293 细胞以检测噬菌斑形式的单个病毒。

收获的噬菌斑是经两次纯化的噬菌斑，并通过 PCR 和偶尔通过 DNA 测序分析病毒是否存在期望的序列。为了进一步的实验，通过氯化铯梯度离心大规模纯化该病毒。

E1A 处于细胞状态特异性 TRE 转录控制下的腺病毒载体

制备含有低氧效应元件（HRE）的腺病毒载体。CN796 是其中 E1A 处于 HRE 和 PSA-TRE 构成的复合 TRE 转录控制下的腺病毒载体，该载体通过 CN515 和 pBHG10 的共转染制备。CN515 通过在 CN65 的 BglIII 位点插入一个 HRE eno1 的 67 碱基对片段（Jiang 等（1997）癌症研究 57:5328-5335）（SEQ ID NO:1; 图 2）构建。CN65 是含有人 PSA 基因的增强子和启动子的质粒，-5322 位至-3738 位增强子融合于-542 位至+12 位 PSA 启动子。该 PSA-TRE 正是质粒 CN706 中包含的 PSA-TRE。Rodriguez 等(1997)Cancer Res. 57:2559-2563。

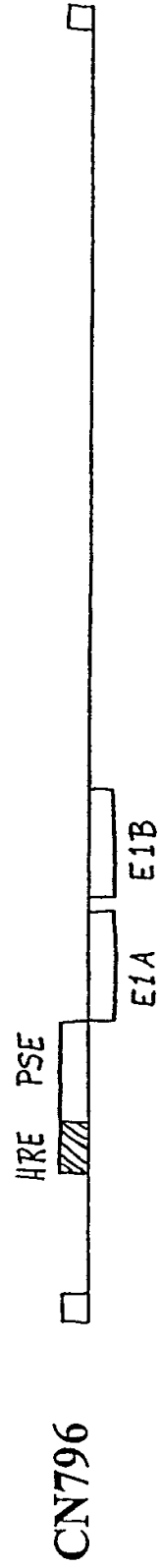
体外病毒的生长

在噬菌斑测定中分析重组腺病毒的生长选择性，在该测定中单个

感染颗粒通过多轮感染和复制产生一个可见的噬菌斑。将病毒原液稀释成相同的 pfu/ml，然后用以感染单层细胞 1 小时。然后将接种物除去，并用含有培养基的半固体琼脂覆盖该细胞，于 37℃ 孵育 10 天。然后计数单层中的噬菌斑并计算各种细胞上的感染性病毒滴度。数据对 293 细胞上的 CN702（野生型）的滴度标准化。

说明书附图

图 1



CN796

图 2

ccccgagg cagtgcac gaggctcagg gcgtgcgt gattcgacgagagaccccg gggtgcag gccgga

图 4A

aagcttctag ttttcttttc ccggtgacat cgtggaaagc actagcauct ctaagcaatg 60
 atctgtgaca atattcacag tgtaatgcc a tccagggaac tcaactgagc cttgatgtcc 120
 agagatTTTT gtgtTTTTTT ctgagactga gtctcgctct gtgccaggct ggagtgcagt 180
 ggtgcaacct tggctcactg caagctccgc ctctggggtt cagccattc tctgcctca 240
 gcctcctgag tagctgggac tacaggcacc cgcaccacg cctggctaatt tttttgtat 300
 ttttagtaga gatgggggtt cactgtgtta gccaggatgg tctcagtctc ctgacctcgt 360
 gatctgcca ccttggcctc ccaaagtgtt gggatgacag gcgtgagcca ccgcgcctgg 420
 ccgatatcca gagatTTTTT ggggggctcc atcacacaga catgttgac: gtcttcatgg 480
 ttgactttta gtatccagcc cctctagaaa tctagctgat atagtgtggc tcaaaacctt 540
 cagcacaaat cacaccgtta gactatctgg tgtggcccaa accttcaggt gaacaaaggg 600
 actctaattc ggcaggatat tccaaagcat tagagatgac ctcttgcaaa gaaaaagaaa 660
 tggaaaagaa aaagaaagaa aggaaaaaaa aaaaaaaaa gagatgacct ctcaggctct 720
 gaggggaaac gcctgaggtc tttgagcaag gtcagtcctc tgttgacag tctccctcac 780
 agggtcattg tgacgatcaa atgtgttcac gtgtatgagg caccagcaca tgcctggctc 840
 tggggagtgc cgtgtaagtg tatgcttgca ctgctgaatg cttgggatgt gtcagggatt 900
 atcttcagca cttacagatg ctcatctcat cctcacagca tcaactatggg atgggtatta 960
 ctggcctcat ttgatggaga aagtggctgt ggctcagaaa ggggggacca ctgaccagg 1020
 gacactctgg atgtgggga ctccagagac catgaccact caccaactgc agagaaatta 1080
 attgtggcct gatgtccctg tcttgagag ggtggagggtg gaccttact aacctctac 1140
 cttgaccctc tcttttaggg ctctttctga cctccaccat ggtactagga cccattgta 1200
 ttctgtacc tcttgactct atgaccccca ctgccactg catccagctg ggtcccctcc 1260
 tatctctatt ccagctggc cagtgcagtc tcaagtccca cctgtttgtc agtaactctg 1320
 aaggggctga cttttactg acttgcaaac aaataagcta actttccaga gttttgtgaa 1380
 tgctggcaga gtccatgaga ctctgagtc agaggcaaag gcttttactg ctacagctt 1440
 agcagacagc atgaggttca tgttcacatt agtacacctt gccccccca aatcttgtag 1500
 ggtgaccaga gcagtctagg tggatgctgt gcagaagggg tttgtgccac tggtgagaaa 1560
 cctgagatta ggaatcctca atcttatact gggacaactt gcaaactgc tcagcctttg 1620
 tctctgatga agatattatc ttcatgatct tggattgaaa acagacctac tctggaggaa 1680

图 4B

catattgtat cgattgtcct tgacagtaaa caaatctgtt gtaagagaca ttatctttat 1740
 tatctaggac agtaagcaag cctggatctg agagagatat catcttgcaa ggatgcctgc 1800
 ttacaaaca tccttgaaac aacaatccag aaaaaaaaaag gtgttgctgt ctttgctcag 1860
 aagacacaca gatacgtgac agaaccatgg agaattgcct cccaacgctg ttcagccaga 1920
 gccttccacc cttgtctgca ggacagtctc aacgttccac cattaaatac ttcttctatc 1980
 acatcctgct tctttatgcc taaccaaggt tctaggctcc gatcgactgt gtctggcagc 2040
 actccactgc caaaccaga ataaggcagc gctcaggatc ccgaaggggc atggctgggg 2100
 atcagaactt ctgggtttga gtgaggagt ggtccaccct cttgaatttc aaaggaggaa 2160
 gaggctggat gtgaaggtag tgggggaggg aaagtgtcag ttccgaactc ttaggtcaat 2220
 gagggaggag actggttaagg tcccagctcc cgaggtagtg atgtgggaat ggcctaagaa 2280
 tctcatatcc tcaggaagaa ggtgctggaa tcctgagggg tagagttctg ggtatatttg 2340
 tggcttaagg ctctttggcc cctgaaggca gaggctggaa ccattaggtc cagggtttgg 2400
 ggtgatagta atgggatctc ttgattcctc aagagtctga ggatcgaggg ttgcccattc 2460
 ttccatcttg ccacctaate cttactccac ttgagggtat caccagccct tctagctcca 2520
 tgaaggctcc ctgggcaagc acaatctgag catgaaagat gccccagagg ccttgggtgt 2580
 catccactca tcatccagca tcacactctg aggggtgtggc cagcaccatg acgtcatggt 2640
 gctgtgacta tccctgcagc gtgcctctcc agccacctgc caaccgtaga gctgcccatac 2700
 ctctctgtgt gggagtggcc tgcattgtgc caggctgagg cctagtgtca gacagggagc 2760
 ctggaatcat agggatccag gactcaaaag tgctagagaa tggccatattg tcaccatcca 2820
 tgaaatctca agggcttctg ggtggagggc acagggacct gaacttatgg tttcccaagt 2880
 ctattgctct cccaagttag tctcccagat acgaggcact gtgccagcat cagccttatc 2940
 tccaccacat cttgtaaaag gactaccag ggcctgatg aacaccatgg tgtgtacagg 3000
 agtagggggt ggaggcacgg actcctgtga ggtcacagcc aaggagcat catcatgggt 3060
 ggggaggagg caatggacag gcttgagaac ggggatgtgg ttgtatttgg ttttctttgg 3120
 ttagataaag tgctgggtat aggattgaga gtggagtatg aagaccagtt aggatggagg 3180
 atcagattgg agttgggtta gataaagtgc tgggtatagg attgagagtg gagtatgaag 3240
 accagttagg atggaggatc agattggagt tgggttagag atggggtaaa attgtgctcc 3300
 ggatgagttt gggattgaca ctgtggaggt ggtttgggat ggcattggctt tgggatggaa 3360

图 4C

atagatttgt tttgatgttg gctcagacat ccttggggat tgaactgggg atgaagctgg 3420
 gtttgatttt ggaggtagaa gacgtggaag tagctgtcag atttgacagt ggccatgagt 3480
 tttgtttgat ggggaatcaa acaatggggg aagacataag gggtggcttg ttaggttaag 3540
 ttgcgttggg ttgatggggg cgggctgtg tataatgcag ttggattggg ttgtattaaa 3600
 ttgggttggg tcaggttttg gttgaggatg agttgaggat atgcttgggg acaccggatc 3660
 catgaggttc tcaactggagt ggagacaaac ttcctttcca ggatgaatcc agggaagcct 3720
 taattcacgt gtaggggagg tcaggccact ggctaagtat atccttcac tccagctcta 3780
 agatgggtctt aaattgtgat tatctatata cacttctgtc tccctcactg tgcttggagt 3840
 ttacctgac actcaactag aaacagggga agattttatc aaattctttt tttttttttt 3900
 ttttttttga gacagagtct cactctgttg ccagggctgg agtgcagtgg cgcagtctcg 3960
 gctcactgca acctctgcct ccaggttca agtgattctc ctgcctcagc ctccctgagtt 4020
 gctgggatta caggcatgca gcacatgcc cagctaattt ttgtattttt agtagagatg 4080
 gggtttcacc aatgtttgcc aggtggcct cgaactcctg acctgggtgat ccacctgcct 4140
 cagcctccca aagtgtctggg attacaggcg tcagccaccg cgcccagcca cttttgtcaa 4200
 attcttgaga cacagctcgg gctggatcaa gtgagctact ctggttttat tgaacagctg 4260
 aaataaccaa ctttttggaa attgatgaaa tcttacggag ttaacagtgg aggtaccagg 4320
 gctcttaaga gttcccgatt ctctcttgag actacaaatt gtgattttgc atgccacctt 4380
 aatctttttt tttttttttt taaatcgagg ttccagtctc attctatttc ccaggctgga 4440
 gttcaatagc gtgatcacag ctcaactgtg ccttgaaact ctggccttaa gagattctcc 4500
 tgcttcgggc tcccaatagc taagactaca gtagtccacc accatatcca gataattttt 4560
 aaattttttg gggggccggg cacagtggct cacgcctgta atcccaacac catgggaggc 4620
 tgagatgggt ggatcacgag gtcaggagtt tgagaccagc ctgaccaaca tggtgaaact 4680
 ctgtctctac taaaaaaaaa aaaaatagaa aaattagccg ggcgtgggtg cacacggcac 4740
 ctgtaatccc agctactgag gaggtgagg caggagaatc actgaacctc agaaggcaga 4800
 gggtgcaatg agccgagatt gcgcactgc actccagcct gggtgacaga gtgagactct 4860
 gtctcaaaaa aaaaaaattt tttttttttt ttgtagaga tggatcttgc tttgtttctc 4920
 tggttggcct tgaactcctg gcttcaagtg atcctcctac cttggcctcg gaaagtgttg 4980
 ggattacagg cgtgagccac catgactgac ctgtcgtaa tcttgaggta cataaacctg 5040
 gctcctaaag gctaaaggct aaatatttgt tggagaaggg gcattggatt ttgcatgagg 5100

图 4D

```

atgattctga cctgggaggg caggtcagca ggcattctctg ttgcacagat agagtgtaca 5160
ggtctggaga acaaggagtg gggggttatt ggaattccac attgtttgct gcacgttgga 5220
ttttgaaatg ctagggaact ttgggagact catatctctg ggctagagga tctgtggacc 5280
acaagatctt tttatgatga cagtagcaat gtatctgtgg agctggattc tgggttgga 5340
gtgcaaggaa aagaatgtac taaatgccaa gacatctatt tcaggagcat gaggaataaa 5400
agttctagtt tctggtctca gagtgggtgca gggatcaggg agtctcaca tctcctgagt 5460
gctggtgtct tagggcacac tgggtcttgg agtgcaaagg atctaggcac gtgaggcttt 5520
gtatgaagaa tcggggatcg taccacccc ctgtttctgt ttcattctgg gcatgtctcc 5580
tctgcctttg tcccctagat gaagtctcca tgagctacaa gggcctggtg catccagggt 5640
gatctagtaa ttgcagaaca gcaagtgcta gctctccctc ccctccaca gctctgggtg 5700
tgaggagggg ttgtccagcc tccagcagca tggggagggc cttggtcagc ctctgggtgc 5760
cagcagggca ggggcggagt cctggggaat gaaggtttta tagggctcct gggggaggct 5820
ccccagcccc aagctt 5836

```

图 5A

aagcttttta gtgctttaga cagtgaactg gtctgtctaa cccaagtga cttgggtcca	60
tactcagccc cagaagtga ggtgaagct ggtggagcc aaaccaggca agcctaccct	120
cagggtccc agtggcctga gaaccattgg acccaggacc cattacttct agggtaagga	180
aggtacaaac accagatcca accatggtct ggggggacag ctgtcaaag cctaaaaata	240
tacctgggag aggagcaggc aaactatcac tgccccaggt tctctgaaca gaaacagagg	300
ggcaacccaa agtccaaatc caggtgagca ggtgcaccaa atgccagag atatgacgag	360
gcaagaagtg aaggaaccac ccctgcatca aatgttttgc atgggaagga gaagggggtt	420
gctcatgttc ccaatccagg agaatgcatt tgggatctgc cttcttctca ctcttggtt	480
agcaagacta agcaaccagg actctggatt tggggaaaga cgtttatttg tggaggccag	540
tgatgacaat cccacgaggg cctaggtgaa gagggcagga aggtctgaga cactggggac	600
tgagtgaata ccacacccat gatctgcacc acccatggat gctccttcat tgctcacctt	660
tctgttgata tcagatggcc ccattttctg taccttcaca gaaggacaca ggctagggtc	720
tgtgcatggc cttcatcccc ggggcatgt gaggacagca ggtgggaaag atcatgggtc	780
ctctgggtc ctgcagggcc agaacattca tcaccatac tgacctcta gatgggaatg	840
gcttccctgg ggtgggcca acggggcctg ggcaggggag aaaggacgtc aggggacagg	900
gaggaagggt catcgagacc cagcctggaa ggttcttgtc tctgaccatc caggatttac	960
ttccctgcat ctaccttgg tcattttccc tcagcaatga ccagctctgc ttctgatct	1020
cagcctcca ccctggacac agcaccacag tccctggccc ggctgcatcc acccaatacc	1080
ctgataaccc aggaccatt acttctaggg taaggagggt ccaggagaca gaagctgagg	1140
aaaggctctga agaagtcaca tctgtcctgg ccagagggga aaaaccatca gatgtgaac	1200
caggagaatg ttgaccacag aaagggaccg aggaccaag aaaggagtca gaccaccagg	1260
gtttgcctga gaggaaggat caaggccccg agggaaagca gggctggctg catgtgcagg	1320
acactgggtg ggcataatgt tcttagattc tccctgaatt cagtgtccct gccatggcca	1380
gactctctac tcaggcctgg acatgtctga ataggacaat ggccttgtcc tctctcccca	1440
ccatttgcca agagacataa aggacattcc aggacatgcc ttcctgggag gtccagggtc	1500
tctgtctcac acctcaggga ctgtagttac tgcatcagcc atggtaggtg ctgatctcac	1560
ccagcctgtc caggcccttc cactctccac tttgtgacca tgtccaggac caccctcag	1620
atcctgagcc tgcaaatacc cccttgcctg gtgggtggat tcagtaaaca gtgagctcct	1680

图 5B

atccagcccc	cagagccacc	tctgtcacct	tctgtctggg	catcatccca	ccttcacaag	1740
cactaaagag	catggggaga	cctggctagc	tgggtttctg	catcacaaag	aaaataatcc	1800
cccaggttcg	gattcccagg	gctctgtatg	tggagctgac	agacctgagg	ccaggagata	1860
gcagaggtca	gccctaggga	gggtgggtca	tccaccagg	ggacaggggt	gcaccagcct	1920
tgctactgaa	agggcctccc	caggacagcg	ccatcagccc	tgcttgagag	ctttgctaaa	1980
cagcagtcag	aggaggccat	ggcagtggct	gagctcctgc	tccagggccc	aacagaccag	2040
accaacagca	caatgcagtc	cttccccaac	gtcacaggtc	accaaaggga	aactgagggtg	2100
ctacctaac	ttagagccat	caggggagat	aacagcccaa	tttcccaaac	aggccagttt	2160
caatcccatg	acaatgacct	ctctgctctc	attcttccca	aaataggacg	ctgattctcc	2220
cccacccatg	atttctccct	tgtcccgga	gccttttctg	ccccctatga	tctgggcaact	2280
cctgacacac	acctcctctc	tggtgacata	tcagggtccc	tcactgtcaa	gcagtccaga	2340
aaggacagaa	ccttgacag	cgcccatctc	agcttcaccc	ttctccttc	acagggttca	2400
gggcaaagaa	taaatggcag	aggccagtga	gccagagat	ggtgacaggc	agtgaccacg	2460
gggcagatgc	ctggagcagg	agctggcggg	gccacaggga	gaaggtgatg	caggaaggga	2520
aaccagaaa	tgggcaggaa	aggaggacac	aggctctgtg	gggctgcagc	ccagggttgg	2580
actatgagtg	tgaagccatc	tcagcaagta	aggccaggtc	ccatgaacaa	gagtgggagc	2640
acgtggcttc	ctgctctgta	tatggggtgg	gggattccat	gccccataga	accagatggc	2700
cggggttcag	atggagaagg	agcaggacag	gggatcccca	ggataggagg	acccagtggt	2760
ccccaccag	gcagggtgact	gatgaatggg	catgcagggt	cctcctgggc	tgggctctcc	2820
ctttgtccct	caggattcct	tgaaggaaca	tccggaagcc	gaccacatct	acctggtggg	2880
ttctggggag	tccatgtaaa	gccaggagct	tgtgttgcta	ggaggggtca	tggcatgtgc	2940
tgggggcacc	aaagagagaa	acctgagggc	aggcaggacc	tggcttgagg	aggcatggga	3000
gccagatgg	ggagatggat	gtcaggaaa	gctgccccat	cagggagggt	gatagcaatg	3060
gggggtctgt	gggagtgggc	acgtgggatt	ccctgggctc	tgccaagttc	cctcccatag	3120
tcacaacctg	gggacactgc	ccatgaagg	gcgcctttgc	ccagccagat	gctgctgggt	3180
ctgcccattc	actaccctct	ctgctccagc	cactctgggt	ctttctccag	atgcctgga	3240
cagccctggc	ctgggcctgt	cccctgagag	gtgttgagg	aaactgagtc	tctggggaca	3300
ctctcatcag	agtctgaaag	gcacatcagg	aaacatccct	ggtctccagg	actaggcaat	3360

图 5C

gaggaaaggg cccagctcc tccctttgcc actgagaggg tgcaccctgg gtggccacag	3420
tgactttctgc gtctgtccca gtcaccctga aaccacaaca aaaccccagc cccagaccct	3480
gcaggtacaa tacatgtggg gacagtctgt acccagggga agccagttct ctcttcctag	3540
gagaccgggc ctgagggtg tgcccggggc aggcgggggc agcacgtgcc tgtccttgag	3600
aactcgggac ctttaagggtc tctgctctgt gaggcacagc aaggatcctt ctgtccagag	3660
atgaaagcag ctctgcccc tctctgacc tcttcctct tcccaaatct caaccaacaa	3720
ataggtgttt caaatctcat catcaaact tcatccatcc acatgagaaa gcttaaaacc	3780
caatggattg acaacatcaa gagttggaac aagtggacat ggagatgtta cttgtggaaa	3840
tttagatgtg ttcagctatc gggcaggaga atctgtgtca aattccagca tggttcagaa	3900
gaatcaaaaa gtgtcacagt ccaaagtgc aacagtgcag gggataaaac tgtggtgcat	3960
tcaaactgag ggatattttg gaacatgaga aaggaaggga ttgctgctgc acagaacatg	4020
gatgatctca cacatagagt tgaaagaaag gagtcaatcg cagaatagaa aatgatcact	4080
aattccacct ctataaagtt tccaagagga aaaccaatt ctgctgctag agatcagaat	4140
ggaggtgacc tgtgccttgc aatggctgtg agggtcacgg gagtgtcact tagtgcaggc	4200
aatgtgccgt atcttaatct gggcagggtt ttcatgagca cataggaatg cagacattac	4260
tgctgtgttc attttacttc accggaaaag aagaataaaa tcagccgggc gcggtggctc	4320
acgcctgtaa tcccagcact ttagaaggct gaggtgggca gattacttga ggtcaggagt	4380
tcaagaccac cctggccaat atggtgaaac cccggctcta ctaaaaatac aaaaattagc	4440
tgggcatggt ggtgcgcgcc tgtaatccca gctactcggg aggctgaggc tggacaattg	4500
cttggaccca ggaagcagag gttgcagtga gccaaagtgt tgccactgca ctccagcttg	4560
ggcaacagag ccagactctg taaaaaaaaa aaaaaaaaaa aaaaaagaa agaaagaaaa	4620
agaaaagaaa gtataaaatc tctttgggtt aacaaaaaaaa gatccacaaa acaaacacca	4680
gctcttatca aacttacaca actctgccag agaacaggaa acacaaatac tcattaactc	4740
acttttgtgg caataaaacc ttcattgtcaa aaggagacca ggacacaatg .aggaagtaaa	4800
actgcaggcc ctacttgggt gcagagaggg aaaatccaca aataaaacat taccagaagg	4860
agctaagatt tactgcattg agttcattcc ccaggatatgc aaggtgattt taacacctga	4920
aatcaatca ttgcctttac tacatagaca gattagctag aaaaaatta caactagcag	4980
aacagaagca atttggcctt cctaaaattc cacatcatat catcatgatg gagacagtgc	5040
agacgccaat gacaataaaa agagggacct ccgtcacccg gtaaacaatgt ccacacagct	5100

图 5D

ccagcaagca cccgtcttcc cagtgaatca ctgtaacctc ccctttaatc agccccaggc	5160
aaggctgcct gcgatggcca cacaggctcc aaccctgagg cctcaacctc ccgcagaggc	5220
tctcctttgg ccaccccatg gggagagcat gaggacaggg cagagccctc tgatgccac	5280
acatggcagg agctgacgcc agagccatgg gggctggaga gcagagctgc tggggtcaga	5340
gcttcctgag gacaccagg cctaagggaa ggcagctccc tggatggggg caaccaggct	5400
ccgggctcca acctcagagc ccgcatggga ggagccagca ctctaggcct ttcctagggt	5460
gactctgagg ggaccctgac acgacaggat cgctgaatgc acccgagatg aaggggccac	5520
cacgggaccc tgctctctg gcagatcagg agagagtggg acaccatgcc agggcccat	5580
ggcatggctg cgactgaccc aggccactcc cctgcatgca tcagcctcgg taagtcacat	5640
gaccaagccc aggaccaatg tggaaggaa gaaacagcat cccttttagt gatggaaccc	5700
aaggctcagtg caaagagagg ccatgagcag ttaggaaggg tggccaacc tacagcacia	5760
accatcatct atcataagta gaagccctgc tccatgaccc ctgcatttaa ataaacgttt	5820
gttaaagtga tcaaattccc tcaccatgag agctcacctg tgtgtaggcc catcacacac	5880
acaaacacac acacacacac acacacacac acacacacac acagggaaag tgcaggatcc	5940
tggacagcac caggcaggct tcacaggcag agcaaacagc gtgaatgacc catgcagtgc	6000
cctgggcccc atcagctcag agaccctgtg agggctgaga tggggctagg caggggagag	6060
acttagagag ggtggggcct ccagggaggg ggctgcaggg agctgggtac tggcctccag	6120
ggagggggct gcagggagct ggggtactgc ctccaggag ggggctgcag ggagctgggt	6180
actgccctcc agggaggggg ctgcaggag ctgggtactg ccctccaggg agggggctgc	6240
agggagctgg gtactgccct ccagggaggc aggagactg ttccaacag agagcacatc	6300
ttcctgcagc agctgcacag acacaggagc ccccatgact gccctgggcc aggggtgtga	6360
ttccaaattt cgtgccccat tgggtgggac ggaggtgac cgtgacatcc aaggggcatc	6420
tgtgattcca aacttaaact actgtgcta caaaatagga aataacccta ctttttctac	6480
tatctcaaat tccctaagca caagctagca ccctttaaat caggaagttc agtcaactct	6540
ggggtcctcc catgccccca gtctgacttg cagggtgcaca ggggtggctga catctgtcct	6600
tgctcctcct cttggctcaa ctgccgcccc tcctgggggt gactgatggg caggacaagg	6660
gacctagag ctggcccat gattgacagg aaggcaggac ttggcctcca ttctgaagac	6720
taggggtgtc aagagagctg ggcattccac agagctgcac aagatgacgc ggacagaggg	6780

图 5E

tgacacaggg	ctcagggctt	cagacgggtc	gggaggctca	gctgagagtt	cagggacaga	6840
cctgaggagc	ctcagtggga	aaagaagcac	tgaagtggga	agttctggaa	tgttctggac	6900
aagcctgagt	gctctaagga	aatgctccca	ccccgatgta	gcctgcagca	ctggacggtc	6960
tgtgtacctc	cccgtgccc	atcctctcac	agccccgcc	tctagggaca	caactcctgc	7020
cctaacatgc	atctttcctg	tctcattcca	cacaaaaggg	cctctggggg	ccctgttctg	7080
cattgcaagg	agtggaggtc	acgttcccac	agaccaccca	gcaacagggg	cctatggagg	7140
tgcggtcagg	aggatcacac	gtcccccat	gccagggga	ctgactctgg	gggtgatgga	7200
ttggcctgga	ggccactgg	cccctctgtc	cctgagggga	atctgcaccc	tggaggctgc	7260
cacatccctc	ctgattcttt	cagctgaggg	cccttcttga	aatcccaggg	aggactcaac	7320
ccccactggg	aaaggcccag	tgtggacggg	tccacagcag	cccagctaag	gcccttggac	7380
acagatcctg	agtgagagaa	cctttaggga	cacagggtga	cggccatgtc	cccagtcccc	7440
acacagagca	ggggcatctg	gaccctgagt	gtgtagctcc	cgcgactgaa	cccagccctt	7500
ccccaatgac	gtgaccctg	gggtggctcc	aggtctccag	tccatgccac	caaaatctcc	7560
agattgaggg	tcctcccttg	agtcctgat	gcctgtccag	gagctgcccc	ctgagcaaat	7620
ctagagtgca	gagggctggg	attgtggcag	taaaagcagc	cacatttgct	tcaggaagga	7680
aagggaggac	atgagctcca	ggaaggcgga	tggcgctctc	tagtgggcgc	ctcctgttaa	7740
tgagcaaaaa	ggggccagga	gagttgagag	atcagggctg	gccttggact	aaggctcaga	7800
tggagaggac	tgaggtgcaa	agagggggct	gaagtagggg	agtggtcggg	agagatggga	7860
ggagcaggta	aggggaagcc	ccagggaggg	cgggggaggg	tacagcagag	ctctccactc	7920
ctcagcattg	acatttgggg	tggctcgtgct	agtgggggtc	tgtaagttgt	aggggtgttca	7980
gcaccatctg	gggactctac	ccactaaatg	ccagcaggac	tccctcccca	agctctaaca	8040
accaacaatg	tctccagact	ttccaaatgt	cccctggaga	gcaaaattgc	ttctggcaga	8100
atcactgac	tacgtcagtc	tctaaaagtg	actcatcagc	gaaatccttc	acctcttggg	8160
agaagaatca	caagtgtgag	aggggtagaa	actgcagact	tcaaaatctt	tccaaaagag	8220
ttttacttaa	tcagcagttt	gatgtcccag	gagaagatac	atttagagtg	tttagagttg	8280
atgccacatg	gctgcctgta	cctcacagca	ggagcagagt	gggttttcca	agggcctgta	8340
accacaactg	gaatgacact	cactgggtta	cattacaaag	tggaatgtgg	ggaattctgt	8400
agactttggg	aagggaaatg	tatgacgtga	gccacagcc	taaggcagtg	gacagtccac	8460
tttgaggctc	tcaccatcta	ggagacatct	cagccatgaa	catagccaca	tctgtcatta	8520

图 5F

gaaaacatgt tttattaaga ggaaaaatct aggctagaag tgctttatgc tcttttttct	8580
ctttatgttc aaattcatat acttttagat cattccttaa agaagaatct atccccctaa	8640
gtaaatgtta tcaactgactg gatagtgttg gtgtctcact cccaaccctt gtgtggtgac	8700
agtgccttgc ttccccagcc ctggggccctc tctgattcct gagagctttg ggtgctcctt	8760
cattaggagg aagagaggaa ggggtgtttt aatattctca ccattcaccc atccacctct	8820
tagacactgg gaagaatcag ttgcccactc ttggatttga tctctgaatt aatgacctct	8880
atttctgtcc ctgtccatt tcaacaatgt gacaggccta agaggtgcct tctccatgtg	8940
atttttgagg agaaggttct caagataagt tttctcacac ctctttgaat tacctccacc	9000
tgtgtcccca tcaccattac cagcagcatt tggacccttt ttctgttagt cagatgcttt	9060
ccacctcttg aggggtgtata ctgtatgctc tctacacagg aatatgcaga ggaaatagaa	9120
aaagggaaat cgcattacta ttcagagaga agaagacctt tatgtgaatg aatgagagtc	9180
taaaatccta agagagccca tataaaatta ttaccagtgc taaaactaca aaagttacac	9240
taacagtaaa ctagaataat aaaacatgca tcacagttgc tggtaaagct aaatcagata	9300
tttttttctt agaaaaagca ttccatgtgt gttgcagtga tgacaggagt gcccttcagt	9360
caatatgctg cctgtaattt ttgttccctg gcagaatgta ttgtcttttc tccctttaa	9420
tcttaaatgc aaaactaaag gcagctcctg ggccccctcc ccaaagtcag ctgcctgcaa	9480
ccagccccac gaagagcaga ggcctgagct tccctggtca aaataggggg ctagggagct	9540
taaccttgct cgataaagct gtgttcccag aatgtcgtc ctgttcccag gggcaccagc	9600
ctggagggtg gtgagcctca ctggtggcct gatgcttacc ttgtgccctc acaccagtgg	9660
tcaactggaac cttgaacact tggctgtcgc ccggatctgc agatgtcaag aacttctgga	9720
agtcaaatta ctgccactt ctccagggca gatacctgtg aacatccaaa accatgccac	9780
agaaccctgc ctgggggtcta caacacatat ggactgtgag caccaagtcc agccctgaat	9840
ctgtgaccac ctgccaagat gcccctaact gggatccacc aatcactgca catggcaggc	9900
agcgaggctt ggaggtgctt cgccacaagg cagcccctaat ttgctgggag tttcttggca	9960
cctggtagtg gtgaggagcc ttgggaccct caggattact ccccttaagc atagtgggga	10020
cccttctgca tccccagcag gtgccccgct cttcagagcc tctctctctg aggtttaccc	10080
agacccctgc accaatgaga ccatgctgaa gcctcagaga gagagatgga gctttgacca	10140
ggagccgctc ttccttgagg gccagggcag ggaaagcagg aggcagcacc aggagtggga	10200

图 5G

acaccagtgt ctaagcccct gatgagaaca ggggtggtctc tcccatatgc ccataccagg	10260
cctgtgaaca gaatcctcct tetgcagtga caatgtctga gaggacgaca tgtttccag	10320
cctaacgtgc agccatgccc atctaccac tgcctactgc aggacagcac caaccaggga	10380
gctgggaagc tgggagaaga catggaatac ccatggcttc tcaccttcct ccagtcagtc	10440
gggcaccatt tatgcctagg acaccacct gccggcccca ggctcttaag agttaggtca	10500
cctaggtgcc tctgggaggc cgaggcagga gaattgcttg aaccggggag gcagaggttg	10560
cagtgaagcc agatcacacc actgcactcc agcctgggtg acagaatgag actctgtctc	10620
aaaaaaaaag agaaagatag catcagtggc taccaagggc taggggcagg ggaaggtgga	10680
gagttaatga ttaatagtat gaagtttcta tgtgagatga tgaaaatgtt ctggaaaaaa	10740
aaatatagtg gtgaggatgt agaataatt gaataaatt aacggcattt aattgtacac	10800
ttaacatgat taatgtggca tattttatct tatgtatttg actacatcca agaaacactg	10860
ggagagggaa agcccaccat gtaaaataca cccaccctaa tcagatagtc ctcatgtac	10920
ccaggtacag gccctcatg acctgcacag gaataactaa ggatttaagg acatgaggct	10980
tcccagccaa ctgcaggtgc acaacataaa tgtatctgca aacagactga gagtaaagct	11040
gggggcacaa acctcagcac tgccaggaca cacacccttc tcgtggattc tgactttatc	11100
tgaccgggcc cactgtccag atcttgttgt gggattggga caaggaggt cataaagcct	11160
gtccccaggg cactctgtgt gagcacagga gacctccca cccccacc gttaggtctc	11220
cacacataga tctgaccatt aggcattgtg aggaggactc tagcgcgggc tcagggatca	11280
caccagagaa tcaggtacag agaggaagac ggggctcgag gagctgatgg atgacacaga	11340
gcaggggtcc tgcagtccac aggtccagct caccctggtg taggtgcccc atccccctga	11400
tccaggcatc cctgacacag ctccctcccg gagcctctc ccaggtgaca catcagggtc	11460
cctcactcaa gctgtccaga gagggcagca ccttgagacag cgcccacccc acttcactct	11520
tcctccctca cagggtcgag ggctcagggc tcaagtctca gaacaaatgg cagaggccag	11580
tgagcccaga gatggtgaca gggcaatgat ccaggggcag ctgcctgaaa cgggagcagg	11640
tgaagccaca gatgggagaa gatggttcag gaagaaaaat ccaggaatgg gcaggagagg	11700
agaggaggac acaggctctg tggggctgca gcccaggatg ggactaagtg tgaagacatc	11760
tcagcaggtg aggccaggtc ccatgaacag agaagcagct cccacctccc ctgatgcacg	11820
gacacacaga gtgtgtggtg ctgtgcccc agagtcgggc tctcctgttc tgggtcccg	11880
ggagtgaaga gtgaggttga cttgtccctg ctctctctg ctacccaac attcaccttc	11940

图 5H

tcctcatgcc cctctctctc aaatatgatt tggatctatg tccccgcca aatctcatgt	12000
caaattgtaa accccaatgt tggaggtggg gccttgtag aagtgattgg ataatgcggg	12060
tggattttct gctttgatgc tgtttctgtg atagagatct cacatgatct ggttgtttaa	12120
aagtgtgtag cacctctccc ctctctctct ctctctctta ctcatgctct gccatgtaag	12180
acgttcctgt tcccccttca ccgtccagaa tgattgtaag tttcttgagg cctccccagg	12240
agcagaagcc actatgcttc ctgtacaact gcagaatgat gagcgaatta aacctctttt	12300
ctttataaat taccagctct caggtatttc tttatagcaa tgcgaggaca gactaataca	12360
atcttctact ccagatccc cgcacacgct tagccccaga catcactgcc cctgggagca	12420
tgcacagcgc agcctcctgc cgacaaaagc aaagtcacaa aaggtgacaa aaatctgcat	12480
ttggggacat ctgattgtga aagagggagg acagtacact tgtagccaca gagactgggg	12540
ctcaccgagc tgaaacctgg tagcactttg gcataacatg tgcatgacct gtgttcaatg	12600
tctagagatc agtggtgagt aaaacagcct ggtctggggc cgctgctgtc ccacttccc	12660
tctgtccac cagagggcgg cagagttcct ccaccctgg agcctcccca ggggctgctg	12720
acctccctca gccgggcca cagcccagca gggccaccc tcaccgggt cacctcggcc	12780
cacgtcctcc tcgccctccg agtcctcac acggactctg tcagctcctc cctgcagcct	12840
atcggccgcc cacctgaggc ttgtcgccg ccacttgag gcctgtcggc tgccctctgc	12900
aggcagctcc tgtcccctac acccctcct tccccgggt cagctgaaag ggcgtctccc	12960
aggcagctc cctgtgatct ccaggacagc tcagtctctc acaggctccg acgcccccta	13020
tgctgtcacc tcacagccct gtcattacca ttaactcctc agtcccatga agttcactga	13080
gcgcctgtct cccggttaca ggaaaactct gtgacagga ccacgtctgt cctgctctct	13140
gtggaatccc agggcccagc ccagtgcctg acacggaaca gatgctccat aaatactgg	13200
taaatgtgtg ggagatctct aaaaagaagc atatcacctc cgtgtggccc ccagcagtca	13260
gagtctgttc catgtggaca caggggact ggcaccagca tgggaggagg ccagcaagt	13320
cccgcggctg cccaggaat gaggcctcaa ccccagagc ttcagaaggg aggacagagg	13380
cctgcaggga atagatcctc cggcctgacc ctgcagccta atccagagtt cagggtcagc	13440
tcacaccacg tcgacctgg tcagcatccc tagggcagtt ccagacaagg ccggaggtct	13500
cctcttgccc tccagggggg gacattgcac acagacatca ctcaggaaac ggattcccct	13560
ggacaggaac ctggctttgc taaggaagtg gaggtggagc ctggtttcca tcccttgctc	13620

图 5I

caacagaccc	ttctgatctc	tcccacatac	ctgctctgtt	cctttctggg	tcctatgagg	13680
accctgttct	gccaggggtc	cctgtgcaac	tccagactcc	ctcctgggtac	cacatgggg	13740
aaggtgggg	gatcacagga	cagtcagcct	cgcagagaca	gagaccaccc	aggactgtca	13800
gggagaacat	ggacaggccc	tgagccgcag	ctcagccaac	agacacggag	agggagggtc	13860
cccctggagc	cttccccaa	gacagcagag	cccagagtca	cccacctccc	tccaccacag	13920
tcctctcttt	ccaggacaca	caagacacct	ccccctccac	atgcaggatc	tggggactcc	13980
tgagacctct	gggcctgggt	ctccatccct	gggtcagtgg	cgggggttgg	ggtactggag	14040
acagagggct	ggtcctctcc	cagccaccac	ccagtgagcc	tttttctagc	cccagagcc	14100
acctctgtca	ccttctgtt	gggcatcatc	ccaccttccc	agagccctgg	agagcatggg	14160
gagacccggg	acctgtctgg	gtttctctgt	cacaaaggaa	aataatcccc	ctgggtgtgac	14220
agaccaagg	acagaacaca	gcagaggtca	gcactgggga	agacaggttg	tcctcccagg	14280
ggatgggggt	ccatccacct	tgccgaaaag	atttgtctga	ggaactgaaa	atagaaggga	14340
aaaaagagga	gggacaaaag	aggcagaaat	gagaggggag	gggacagagg	acacctgaat	14400
aaagaccaca	cccatgaccc	acgtgatgct	gagaagtact	cctgccctag	gaagagactc	14460
agggcagagg	gaggaaggac	agcagaccag	acagtcacag	cagccttgac	aaaacgttcc	14520
tggaactcaa	gctcttctcc	acagaggagg	acagagcaga	cagcagagac	catggagtct	14580
ccctcgcccc	ctccccacag	atggtgcatc	ccctggcaga	ggctcctgct	cacaggtgaa	14640
gggaggacaa	cctgggagag	ggtgggagga	gggagctggg	gtctcctggg	taggacaggg	14700
ctgtgagacg	gacagagggc	tcctgttgga	gcctgaatag	ggaagaggac	atcagagagg	14760
gacaggagtc	acaccagaaa	aatcaaattg	aactggaatt	ggaaaggggc	aggaaaacct	14820
caagagttct	attttcttag	ttaattgtca	ctggccacta	cgttttttaa	aatcataata	14880
actgcatcag	atgacacttt	aaataaaaac	ataaccaggg	catgaaacac	tgtcctcatc	14940
cgcctaccgc	ggacattgga	aaataagccc	caggctgtgg	agggccctgg	gaaccctcat	15000
gaactcatcc	acaggaatct	gcagcctgtc	ccaggcactg	gggtgcaacc	aagatc	15056

图 6A

gaattcagaa ataggggaag gttgaggaag gacactgaac tcaaagggga tacagtgatt 60
 ggtttatttg tcttctcttc acaacattgg tgctggagga attcccaccc tgaggttatg 120
 aagatgtctg aacacccaac acatagcact ggagatatga gctcgacaag agtttctcag 180
 ccacagagat tcacagccta gggcaggagg acactgtacg ccaggcagaa tgacatggga 240
 attgcgctca cgattggcct gaagaagcaa ggactgtggg aggtgggctt tgtagtaaca 300
 agagggcagg gtgaactctg attcccatgg gggaatgtga tggctctgtt acaaattttt 360
 caagctggca ggaataaaaa cccattacgg tgaggacctg tggagggcgg ctgccccaac 420
 tgataaagga aatagccagg tgggggcctt tcccattgta ggggggacat atctggcaat 480
 agaagccttt gagacccttt aggggtacaag tactgaggca gcaaataaaa tgaaatctta 540
 tttttcaact ttatactgca tgggtgtgaa gatataattg tttctgtaca gggggtgagg 600
 gaaaggaggg gaggaggaaa gttcctgcag gtctggtttg gtcttgtgat ccaggggggc 660
 ttggaactat ttaaattaaa ttaaattaaa acaagcgact gttttaaatt aaattaaatt 720
 aaattaaatt ttactttatt ttatcttaag ttctgggcta catgtgcagg acgtgcagct 780
 ttgttacata ggtaaacgtg tgccatgggtg gtttgcgtga cctatcaacc catcacctag 840
 gtattaagcc cagcatgcat tagctgtttt tcctgacgct ctccctctcc ctgactccca 900
 caacaggccc cagtgtgtgt tgttcccctc cctgtgtcca tgtgttctca ttgttcagct 960
 cccacttata agtgagaaca tgtggtgttt ggttttctgt ttctgtgtta gtttgctgag 1020
 gataatggct tccacctcca tccatgttcc tgcaaaggac gtgatcttat tcttttttat 1080
 ggttgcatag aaattgtttt tacaaatcca attgatattg tatttaatta caagttaatc 1140
 taattagcat actagaagag attacagaag atattaggta cattgaatga ggaaatatat 1200
 aaaataggac gaaggtgaaa tattaggtag gaaaagtata atagttgaaa gaagtaaaaa 1260
 aaaatatgca tgagtagcag aatgtaaaag aggtgaagaa cgtaatagtg actttttaga 1320
 ccagattgaa ggacagagac agaaaaattt taaggaattg ctaaaccatg tgagtgttag 1380
 aagtacagtc aataacatta aagcctcagg aggagaaaag aataggaaaag gaggaaatat 1440
 gtgaataaat agtagagaca tgtttgatgg attttaaaat atttgaaaga cctcacatca 1500
 aaggattcat accgtgccat tgaagaggaa gatggaaaag ccaagaagcc agatgaaagt 1560
 tagaaatatt attggcaaag cttaaattgt aaaagtccta gagagaaaag atggcagaaa 1620
 tattggcggg aaagaatgca gaacctagaa tataaattca tccaacagt ttggtagtgt 1680

图 6B

gcagctgtag ccttttctag ataatacaact attgtcatac atcgcttaag cgagtgtaaa 1740
 atggtctcct cactttatctt atttatatat ttatttagtt ttgagatgga gcctcgctct 1800
 gtctcctagg ctggagtgca atagtgcgat accactcact gcaacctctg cctcctctgt 1860
 tcaagtgatt ttcttacctc agcctcccgga gtagctggga ttacagggtgc gtgccaccac 1920
 acccggctaa tttttgtatt tttttagag acggggtttt gccatgttg ccaggctgg 1980
 cttgaactcc tgacatcagg tgatccacct gccttggcct cctaaagtgc tgggattaca 2040
 ggcatgagcc accgtgccca accactttat ttatttttta tttttatctt taaatttcag 2100
 cttctatttg aaatacaggg ggcacatata taggattgtt acatgggtat attgaactca 2160
 ggtagtgatc atactacca acaggtaggt tttcaacca ctccccctct tttctcccc 2220
 attctagtag tgtgcagtgt ctattgttct catgtttatg tctatgtgtg ctccagggtt 2280
 agctcccacc tgtaagttag aacgtgtggt atttgatttt ctgtccctgt gtttaattcac 2340
 ttaggattat ggcttcacgc tccattcata ttgctgtaaa ggatatgatt catttttcat 2400
 ggcatgcag tattccatat tgcgtataga tcacattttc tttctttttt ttttttgaga 2460
 cggagtcttg ctttgcctgc taggctggag tgcagtagca cgatctcggc tcaactgcaag 2520
 cttcacctcc ggggttcacg tcattcttct gtctcagctt cccaagtagc tgggactaca 2580
 ggcgcccgc accacgtccg gctaattttt ttgtgtgttt ttagtagaga tgggggtttc 2640
 actgtgttag ccaggatggg cttgatctcc tgacctgtg gtccacctgc ctgggtctcc 2700
 caaagtgtg ggattacagg ggtgagccac tgcgcccggc ccatatatac cacattttct 2760
 ttaaccaatc caccattgat gggcaactag gtagattcca tggattccac agttttgcta 2820
 ttgtgtgcag tgtggcagta gacatatgaa tgaatgtgtc tttttggtat aatgatttgc 2880
 attccttttg gtatacagtc attaatagga gtgctgggtt gaacgggtgc tctgtttaaa 2940
 attcctttgag aattttccaa actggttgcc atagagagca aactaattta catttccacg 3000
 aacagtatat aagcattccc ttttctccac agctttgtca tcatggtttt ttttttctt 3060
 tattttaaaa aagaatatgt tgttgttttc ccagggtaca tgtgcaggat gtgcagggtt 3120
 gttacatagg tagtaaactg gagccatggt ggtttgctgc acctgtcaac ccattacctg 3180
 ggtatgaagc cctgcctgca ttagctcttt tccctaatgc tctcactact gccccacct 3240
 caccctgaca gggcaaacag acaacctaca gaatgggagg aaatttttgc aatctattca 3300
 tctgacaaaag gtcaagaata tccagaatct acaaggaact taagcaaatt ttacttttt 3360

图 6C

aataatagcc actctgactg gcgtgaaatg gtatctcatt gtggttttca tttgaatttc 3420
tctgatgac agtgacgatg agcatttttt catatttggt ggctgcttgt acgtcttttg 3480
agaagtgtct cttcatgcct tttggccact ttaatgggat tattttttgc ttttagttt 3540
aagttcctta tagattctgg atattagact tcttattgga tgcatagttt gtgaatactc 3600
tcttccattc tgtaggttgt ctgtttactc tattgatggc ttcttttgct gtgccgaagc 3660
atcttagttt aattagaaac cacctgccaa tttttgtttt tgttgcaatt gcttttgggg 3720
acttagtcat aaactctttg ccaaggtctg ggtcaagaag agtatttcct aggttttctt 3780
ctagaatttt gaaagtctga atgtaaacad ttgcattttt aatgcatctt gagttagttt 3840
ttgtatatgt gaaaggtcta ctctcatttt ctttccctct ttctttcttt ctttcttttc 3900
tttctttctt tcttcttttc tttctttctt tcttcttttc tttctttttg tccttctttc 3960
tttctttctt tcttctttct tcttcttttc tttttttttt ttgatggagt attgctctgt 4020
tgcccaggct gcagtgcagc ggcacgatct cggctcactg caacctctgc ctctggggtt 4080
caactgattc tcctgcatca gccttccaag tagctgggat tataggcgcc cgccaccacg 4140
ccogactaat ttttgtattt ttagtagaga cggggttggt ccatgttggc caggctgggt 4200
tgaaactcct gacctcaaac gatctgcctg ccttggcctc ccaaagtgtt gggattacag 4260
gtgtgagcca ctgtgccag ccaagaatgt cattttctaa gaggtccaag aacctcaaga 4320
tattttggga ccttgagaag agaggaattc atacaggtat tacaagcaca gcctaattggc 4380
aaatctttgg catggcttgg cttcaagact ttaggtcttt aaaagtcgaa tccaaaaatt 4440
tttataaaag ctccagctaa gctaccttaa aaggggcctg tatggctgat cactcttctt 4500
gctatacttt acacaaataa acaggccaaa tataatgagg ccaaaattta ttttgcaaat 4560
aaattggtcc tgctatgatt tactcttggg aagaacaggg aaaatagaga aaaatttaga 4620
ttgcatctga cctttttttc tgaattttta tatgtgccta caatttgagc taaatcctga 4680
attattttct gggtgcaaaa actctctaaa gaagaacttg gttttcattg tcttcgtgac 4740
acatttatct ggctctttac tagaacagct ttcttgtttt tgggtgttcta gcttgtgtgc 4800
cttacagttc tactcttcaa attattgtta tgtgtatctc atagttttcc ttcttttgag 4860
aaaactgaag ccatggtatt ctgaggacta gagatgactc aacagagctg gtgaatctcc 4920
tcatatgcaa tccactgggc tcgatctgct tcaaattgct gatgcactgc tgctaaagct 4980
atacatttta aacctcact aaaggatcag ggaccatcat ggaagaggag gaaacatgaa 5040
attgtaagag ccagattcgg ggggtagagt gtggagggtca gagcaactcc accttgaata 5100

图 6D

agaaggtaaa gcaacctatc ctgaaagcta acctgccatg gtggcttctg attaacctct 5160
 gttctaggaa gactgacagt ttgggtctgt gtcattgccc aaatctcatg ttaaattgta 5220
 atccccagtg ttccggaggtg ggacttggtg gtaggtgatt cggtcatggg agtagatttt 5280
 cttctttgtg gtgttacagt gatagtgagt gagttctcgt gagatctggt catttaaaag 5340
 tgtgtggccc ctcccctccc tctcttggtc ctctactgc catgtaagat acctgctcct 5400
 gctttgcctt ctaccataag taaaagcccc ctgaggcctc ccagaagca gatgccacca 5460
 tgcttcctgt acagcctgca gaaccatcag ccaattaaac ctcttttctg tataaattac 5520
 cagtcttgag tatctcttta cagcagtgtg agaacggact aatacaaggg tctccaaaat 5580
 tccaagttta tgtattcttt ctgccaaat agcaggatt taccataaat cctgtcctta 5640
 ggtcaaacaa ccttgatggc atcgtacttc aattgtctta cacattcctt ctgaatgact 5700
 cctcccctat ggcatataag ccctgggtct tgggggataa tggcagaggg gtccaccatc 5760
 ttgtctggct gccacctgag acacggacat ggcttctggt ggtaagtctc tattaaatgt 5820
 ttctttctaa gaaactggat ttgtcagctt gtttctttgg cctctcagct tcctcagact 5880
 ttggggtagg ttgcacaacc ctgcccacca cgaaacaaat gtttaatatg ataaatatgg 5940
 atagatataa tccacataaa taaaagctct tggagggccc tcaataattg ttaagagtgt 6000
 aaatgtgtcc aaagatggaa aatgtttgag aactactgtc ccagagattt tcctgagttc 6060
 tagagtgtgg gaatatagaa cctggagctt ggcttcttca gcctagaatc aggagtatgg 6120
 ggctgaagtc tgaagcttgg cttcagcagt ttggggttgg cttccggagc acatatttga 6180
 catgttgca ctgtgatttg gggtttggtt tttgctctga atcctaattgt ctgtccttga 6240
 ggcattctaga atctgaaatc tgtggtcaga attctattat cttgagtagg acatctccag 6300
 tcctggttct gccttctagg gctggagtct gtagtcagt acccggtctg gcatttcaac 6360
 ttcatatata gtgggctatc ttttggtcca tgtttcaacc aaacaaccga ataaaccatt 6420
 agaacctttc cccacttccc tagctgcaat gttaaacccta ggatttctgt ttaatagggt 6480
 catatgaata atttcagcct gatccaactt tacattcctt ctaccgttat tctacacca 6540
 ccttaaaaat gcattcccaa tatattcctt ggattctacc tatatatggt aatcctggct 6600
 ttgccagttt ctagtgcatt aacataacct atttacattc ttttacttta aagtggaaat 6660
 aagagtcctt ctgcagagtt caggagttct caagatggcc cttacttctg acatcaattg 6720
 agatttcaag ggagtcgcca agatcatcct caggttcagt gattgctggt agccctcata 6780

图 6E

taactcaatg aaagctgtta tgctcatggc tatggtttat tacagcaaaa gaatagagat 6840
 gaaaatctag caaggaaga gttgcatggg gcaaagacaa ggagagctcc aagtgcagag 6900
 attcctgttg ttttctccca gtggtgtcat ggaaagcagt atcttctcca tacaatgatg 6960
 tgtgataata ttcagtgtat tgccaatcag ggaactcaac tgagccttga ttatattgga 7020
 gcttggttgc acagacatgt cgaccacctt catggctgaa cttagtact tagccctcc 7080
 agacgtctac agctgatagg ctgtaacca acattgtcac cataaatcac attgttagac 7140
 tatccagtgt ggcccaagct cccgtgtaaa cacaggcact ctaaacaggc aggatatttc 7200
 aaaagcttag agatgacctc ccaggagctg aatgcaaaga cctggcctct ttgggcaagg 7260
 agaatccttt accgcacact ctcttcaca gggttattgt gaggatcaaa tgtggtcatg 7320
 tgtgtgagac accagcacat gtctggctgt ggagagtgc ttctatgtgt gctaacattg 7380
 ctgagtgtca agaaagtatt aggcattggt ttcagcactc acagatgtc atctaactct 7440
 cacaacatgg ctacaggggtg ggcactacta gcctcatttg acagaggaaa ggactgtgga 7500
 taagaagggg gtgaccaata ggtcagagtc attctggatg caaggggctc cagaggacca 7560
 tgattagaca ttgtctgcag agaaattatg gctggatgtc tctgccccgg aaaggggat 7620
 gcactttcct tgaccccta tctcagatct tgactttgag gttatctcag acttctcta 7680
 tgataccagg agcccatcat aatctctctg tgcctctcc ccttctcag tcttactgcc 7740
 cactcttccc agctccatct ccagctggcc aggtgtagcc acagtaccta actctttgca 7800
 gagaactata aatgtgtatc ctacagggga gaaaaaaaaa aagaactctg aaagagctga 7860
 cattttaccg acttgcaaac acataagcta acctgccagt tttgtgctgg tagaactcat 7920
 gagactcctg ggtcagaggc aaaagatttt attaccaca gctaaggagg cagcatgaac 7980
 tttgtgttca catttggtca ctttgcccc caattcatat gggatgatca gagcagttca 8040
 ggtggatgga cacaggggtt tgtggcaaag gtgagcaacc taggcttaga aatcctcaat 8100
 cttataagaa ggtactagca aacttgcca gtctttgtat ctgacggaga tattatcttt 8160
 ataattgggt tgaaagcaga cctactctgg aggaacatat tgtatttatt gtctgaaca 8220
 gtaaacaat ctgctgtaaa atagacgtta actttattat ctaaggcagt aagcaaacct 8280
 agatctgaag gcgataccat cttgcaaggc tatctgctgt acaaatatgc ttgaaaagat 8340
 ggtccagaaa agaaaacggt attattgcct ttgtcagaa gacacacaga aacataagag 8400
 aaccatggaa aattgtctcc caacactgtt ccccagagc cttccactct tgtctgcagg 8460
 acagtcttaa catcccatca ttagtgtgtc taccacatct ggcttcaccg tgcctaacca 8520

图 6F

agattttctag gtccagttcc ccaccatgtt tggcagtgcc ccactgccaa cccagaata 8580
 agggagtgc cagaattccg aggggacatg ggtggggatc agaacttctg ggcttgagtg 8640
 cagagggggc ccatactcct tggttccgaa ggaggaagag gctggagggtg aatgtccttg 8700
 gaggggagga atgtgggttc tgaactctta aatccccaag ggaggagact ggtaagggtcc 8760
 cagcttccga ggtactgacg tgggaatggc ctgagagggtc taagaatccc gtatcctcgg 8820
 gaaggagggg ctgaaattgt gaggggttga gttgcagggg tttgttagct tgagactcct 8880
 tgggtgggtcc ctgggaagca aggactggaa ccattgggtc cagggttttg tgtgaaggta 8940
 atgggatctc ctgattctca aagggtcaga ggactgagag ttgcccatgc tttgatcttt 9000
 ccatctactc cttactccac ttgagggtaa tcacctactc ttctagtctc acaagagtgc 9060
 gcctgcgcga gtataatctg cacatgtgcc atgtcccgag gcctggggca tcatccactc 9120
 atcattcagc atctgcgcta tgcgggcgag gccgggcgcca tgacgtcatg tagctgcgac 9180
 tatccctgca gcgcgcctct cccgtcacgt cccaaccatg gagctgtgga cgtgcgtccc 9240
 ctggtggatg tggcctgcgt ggtgccaggc cggggcctgg tgtccgataa agatcctaga 9300
 accacaggaa accaggactg aaagggtgta gagaatggcc atatgtcgct gtccatgaaa 9360
 tctcaaggac ttctgggttg agggcacagg agcctgaact tacgggtttg cccaggtcca 9420
 ctgtcctccc aagtgagtct ccagatacg aggcactgtg ccagcatcag cttcatctgt 9480
 accacatctt gtaacaggga ctaccagga ccctgatgaa caccatgggtg tgtgcaggaa 9540
 gagggggtga aggcattggac tcctgtgttg tcagagccca gagggggcca tgacgggttg 9600
 ggaggaggct gtggactggc tcgagaagtg ggatgtggtt gtgtttgatt tcctttggcc 9660
 agataaagtg ctggatatag cattgaaaac ggagtatgaa gaccagttag aatggaggggt 9720
 caggttggag ttgagttaca gatggggtaa aattctgctt cggatgagtt tggggattgg 9780
 caatctaaag gtggtttggg atggcatggc tttgggatgg aaataggttt gtttttatgt 9840
 tggctgggaa ggggtgtggg attgaattgg ggatgaagta ggtttagttt tggagataga 9900
 atacatggag ctggctattg catgcgagga tgtgcattag tttggtttga tctttaaata 9960
 aaggaggcta ttagggttgt cttgaattag attaagttgt gttgggttga tgggttgggc 10020
 ttgtgggtga tgtggttga ttgggctgtg ttaaattggt ttgggtcagg ttttggttga 10080
 ggttatcatg gggatgagga tatgcttggg acatggattc aggtggttct cattcaagct 10140
 gaggcaaatt tcctttcaga cggtcattcc agggaacgag tggttgtgtg ggggaaatca 10200

图 6G

ggccactggc tgtgaatata cctctatcct ggtcttgaat tgtgattata tatgtccatt 10260
 ctgtctcctt cactgtactt ggaattgata tggtcattca gctggaaatg ggggaagatt 10320
 ttgtcaaatt cttgagacac agctgggtct ggatcagcgt aagccttcct tctggtttta 10380
 ttgaacagat gaaatcacat ttttttttc aaaatcacag aaatcttata gagttaacag 10440
 tggactctta taataagagt taacaccagg actcttattc ttgattcttt tctgagacac 10500
 caaaatgaga tttctcaatg ccaccctaatt tctttttttt tttttttttt tttttgagac 10560
 acagtctggg tcttttgctc tgtcactcag gctggagcgc agtgggtgta tcatagctca 10620
 ctgaaccctt gacctcctgg acttaaggga tctcctgct tcagcctcct gagtagatgg 10680
 ggctacaggt gcttgccacc acacctggct aattaaattt tttttttttt tttgtagaga 10740
 aagggtctca ctttggtgcc ctggctgata ttgaacttct gacttcaagt gattcttcag 10800
 ccttggaact ccaaagcact gggattgctg gcatgagcca ctcaccgtgc ctggcttgca 10860
 gcttaatctt ggagtgtata aacctggctc ctgatagcta gacatttcag tgagaaggag 10920
 gcattggatt ttgcatgagg acaattctga cctaggaggg caggtcaaca ggaatccccg 10980
 ctgtacctgt acgtgtaca ggcattgaga atgaggagtg aggaggccgt accggaaccc 11040
 catattgttt agtggacatt ggattttgaa ataataggga acttgggtctg ggagagtcatt 11100
 atttctggat tggacaatat gtggtatcac aaggttttat gatgaggagg aaatgtatgt 11160
 ggggaaccat tttctgagtg tggaagtga agaatcagag agtagctgaa tgccaacgct 11220
 tctatttcag gaacatggta agttggagggt ccagctctcg ggctcagacg ggtataggga 11280
 ccaggaagtc tcacaatccg atcattctga tatttcaggg catattaggt ttggggtgca 11340
 aaggaagtac ttgggactta ggcacatgag actttgtatt gaaaatcaat gattggggct 11400
 ggccgtgggt ctcacgcctg taatctcatc actttgggag accgaagtgg gaggatggct 11460
 tgatctcaag agttggacac cagcctaggc aacatggcca gacctctct ctacaaaaaa 11520
 attaaaaatt agctggatgt ggtggtgcat gcttggtgct tcagctatcc tggaggctga 11580
 gacaggagaa tcggttgagt ctgggagttc aaggctacag ggagctgca tcacgccgct 11640
 gactccagc ctgggaacaa gagtgagact gtctcagaat ttttttaaaa aagaatcagt 11700
 gatcatcca acccctgttg ctgttcattc tgagcctgcc ttctctggct ttgttccta 11760
 gatcacatct ccatgatcca taggcctgac ccaatctgac ctcacaccgt gggaaatgct 11820
 ccagactgat ctagtatgtg tggaacagca agtgctggct ctcctcccc ttccacagct 11880
 ctgggtgtgg gagggggttg tccagcctcc agcagcatgg ggagggcctt ggtcagcatc 11940

图 6H

taggtgccaa cagggcaagg gcgggggcct ggagaatgaa ggctttatag ggctcctcag 12000
ggaggccccc cagccccaaa ctgcaccacc tggccgtgga caccggt 12047

图 7

```

cgagcggccc ctcagcttcg gcgccagcc ccgcaaggct cccggtgacc actagagggc 60
gggaggagct cctggccagt ggtggagagt ggcaaggaag gaccctaggg ttcacggag 120
cccaggttta ctcccttaag tggaaatttc tccccact cctccttggc tttctccaag 180
gaggggaaccc aggctgctgg aaagtccggc tggggcgggg actgtgggtt caggggagaa 240
cggggtgtgg aacgggacag ggagcggtta gaagggtggg gctattccgg gaagtgggtg 300
ggggagggag cccaaaacta gcacctagtc cactcattat ccagccctct tatttctcgg 360
ccgctctgct tcagtggacc cggggagggc ggggaagtgg agtgggagac ctaggggtgg 420
gcttcccgac cttgctgtac aggacctga cctagctggc tttgttccc atccccacgt 480
tagttgttgc cctgaggcta aaactagagc ccaggggccc caagtccag actgcccctc 540
ccccctcccc cggagccagg gagtggttgg tgaaaggggg aggccagctg gagaacaaac 600
gggtagtcag ggggttgagc gattagagcc cttgtaccct acccaggaat ggttggggag 660
gaggaggaag aggtaggagg taggggaggg ggcgggggtt tgtcacctgt cacctgctcg 720
ctgtgcctag ggcgggcggg cggggagtgg ggggaccggt ataaagcggg aggcgcctgt 780
gcccgctcca cctctcaagc agccagcgcc tgcctgaatc tgttctgccc cctccccacc 840
catttcacca ccaccatg
858

```

图 8

```

aagcttccac aagtgcattt agcctctcca gtattgctga tgaatccaca gttcagggtc 60
aatggcggtc aaaacttgat caaaaatgac cagactttat attcttacac caacatctat 120
ctgattggag gaatggataa tagtcatcat gtttaaacad ctaccattcc agttaagaaa 180
atatgatagc atcttggtct tagtcttttt cttaataggg acataaagcc cacaaataaa 240
aatatgcctg aagaatggga caggcattgg gcattgtcca tgcctagtaa agtactccaa 300
gaacctattt gtatactaga tgacacaatg tcaatgtctg tgtacaactg ccaactggga 360
tgcaagacac tgcccatgcc aatcatcctg aaaagcagct ataaaaagca ggaagctact 420
ctgcaccttg tcagtgaggt ccagatacct acag 454

```

图 9

g atg acc ggc tca acc atc gcg ccc aca acg gac tat cgc aac acc act	49
Met Thr Gly Ser Thr Ile Ala Pro Thr Thr Asp Tyr Arg Asn Thr Thr	
1 5 10 15	
gct acc gga cta aca tct gcc cta aat tta ccc caa gtt cat gcc ttt	97
Ala Thr Gly Leu Thr Ser Ala Leu Asn Leu Pro Gln Val His Ala Phe	
20 25 30	
gtc aat gac tgg gcg agc ttg gac atg tgg tgg ttt tcc ata gcg ctt	145
Val Asn Asp Trp Ala Ser Leu Asp Met Trp Trp Phe Ser Ile Ala Leu	
35 40 45	
atg ttt gtt tgc ctt att att atg tgg ctt att tgt tgc cta aag cgc	193
Met Phe Val Cys Leu Ile Ile Met Trp Leu Ile Cys Cys Leu Lys Arg	
50 55 60	
aga cgc gcc aga ccc ccc atc tat agg cct atc att gtg ctc aac cca	241
Arg Arg Ala Arg Pro Pro Ile Tyr Arg Pro Ile Ile Val Leu Asn Pro	
65 70 75 80	
cac aat gaa aaa att cat aga ttg gac ggt ctg aaa cca tgt tct ctt	289
His Asn Glu Lys Ile His Arg Leu Asp Gly Leu Lys Pro Cys Ser Leu	
85 90 95	
ctt tta cag tat gat taa	307
Leu Leu Gln Tyr Asp	
100	