

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200480039355.6

[51] Int. Cl.

*C12N 15/50 (2006.01)*

*C12N 15/05 (2006.01)*

*C12N 15/63 (2006.01)*

*A61K 31/7088 (2006.01)*

*A61K 38/16 (2006.01)*

*A61K 39/215 (2006.01)*

[43] 公开日 2007 年 1 月 31 日

[11] 公开号 CN 1906302A

[51] Int. Cl. (续)

*A61P 31/14 (2006.01)*

*A61P 11/00 (2006.01)*

*A61K 48/00 (2006.01)*

[22] 申请日 2004.12.6

[21] 申请号 200480039355.6

[30] 优先权

[32] 2003.12.5 [33] US [31] 60/527,637

[86] 国际申请 PCT/CN2004/001419 2004.12.6

[87] 国际公布 WO2005/054473 英 2005.6.16

[85] 进入国家阶段日期 2006.6.29

[71] 申请人 香港大学

地址 中国香港薄扶林道

[72] 发明人 蔡美莲 李宏业 R·萨西施库马

L·M·L·潘

S·M·J·裴伟士

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 程 淼 黄可峻

权利要求书 3 页 说明书 49 页 附图 377 页

[54] 发明名称

含有 SARS - CoV 病毒核苷酸序列的基因修饰的植物和使用它进行免疫抗 SARS 的方法

[57] 摘要

本发明涉及基因修饰的植物和其后代,其构成抗严重急性呼吸综合症(Severe Acute Respiratory Syndrome, SARS)的可食用的植物衍生的粘膜疫苗和可注射的植物衍生的粘膜疫苗。本发明涉及特别转化,但不限于烟草,番茄和莴苣植物的核和/或质体以产生抗原的重组载体。在特定的实施方案中,表达核苷酸序列的质体转化载体是 pCV1, pCV6 和 pCV8,和其含有具有核苷酸(而不是氨基酸)改变以使密码子使用最适于在植物中表达的 S1 的衍生物。本发明涉及制备修饰的植物的方法,包括用于 SARS - CoV 病毒,其片段,衍生物,类似物,或变体的核表达和/或质体表达的载体转化植物。本发明还涉及免疫抗 SARS 的方法,和使用由质体和/或核载体转化的植物产生的 SARS - CoV 抗原进行抗体检测的方法。

1. 含有其核酸序列如 SEQ ID NO: 7 所示的分离的核酸，或其片段，衍生物或类似物的载体。
2. 根据权利要求 1 的载体，其中所述载体是质体转化载体。
3. 根据权利要求 2 的载体，其中所述载体是 pCV1, pCV6, pCV8, 或其含有具有至少一个核苷酸改变的 S1 的衍生物。
4. 根据权利要求 1 的载体，其中所述载体是核转化载体。
5. 根据权利要求 4 的载体，其中所述载体是 pCV2, pCV4, pCV12, 或其含有具有至少一个核苷酸改变的 S1 的衍生物。
6. 根据权利要求 1 的载体，其中所述分离的核酸编码 SARS-CoV 病毒的 S 蛋白, SARS-CoV 病毒的 M 蛋白, 或二者。
7. 根据权利要求 1 的载体，其中所述分离的核酸含有如 SEQ ID NO: 3, 5 所示的核酸序列，或其片段，衍生物或类似物。
8. 根据权利要求 1 的载体，其中所述分离的核酸编码如 SEQ ID NO: 4 或 6 所示的氨基酸序列，或其片段，衍生物或类似物。
9. 根据权利要求 1 的载体，其中所述分离的核酸在高度严谨条件下与 SEQ ID NO: 3, 5, 或 7 的核酸序列，或其互补序列杂交。
10. 含有其核酸序列如 SEQ ID NO: 7 所示的分离的核酸，或其片段，衍生物或类似物的植物细胞。
11. 根据权利要求 10 的植物细胞，其中所述分离的核酸编码 SARS-CoV 病毒的 S 蛋白, SARS-CoV 病毒的 M 蛋白, 或二者。
12. 根据权利要求 10 的植物细胞，其中所述分离的核酸含有如 SEQ ID NO: 3, 5 所示的核酸序列，或其片段，衍生物或类似物。
13. 根据权利要求 10 的植物细胞，其中所述分离的核酸编码如 SEQ ID NO: 4 或 6 所示的氨基酸序列，或其片段，衍生物或类似物。
14. 根据权利要求 10 的植物细胞，其中所述分离的核酸在高度严谨条件下与 SEQ ID NO: 3, 5, 或 7 的核酸序列，或其互补序列杂交。
15. 根据权利要求 10 的植物细胞，其中所述植物细胞分离自选自由烟草, 莴苣, 番茄, 马铃薯, 香蕉, 玉米 (corn), 稻, 谷类, 小麦, 玉米 (maize), 大麦, 苹果, 梨, 草莓, 胡萝卜, 甜菜, 薯蓣, 奇异果, 或菠菜组成的组的植物。
16. 含有根据权利要求 1 的载体的植物细胞。

17. 一种免疫抗 SARS 的方法，包括给受试者施用植物，其中所述植物含有其核酸序列如 SEQ ID NO: 7 所示的分离的核酸，或其片段，衍生物，类似物，或载体的植物。

18. 根据权利要求 17 的方法，其中所述分离的核酸编码 SARS-CoV 病毒的 S 蛋白，SARS-CoV 病毒的 M 蛋白，或二者。

19. 根据权利要求 17 的方法，其中所述分离的核酸含有如 SEQ ID NO: 3, 5 所示的核酸序列，或其片段，衍生物或类似物。

20. 根据权利要求 17 的方法，其中所述分离的核酸编码如 SEQ ID NO: 4 或 6 所示的氨基酸序列，或其片段，衍生物或类似物。

21. 根据权利要求 17 的方法，其中所述分离的核酸在高度严谨条件下与 SEQ ID NO: 3, 5, 或 7 的核酸序列，或其互补序列杂交。

22. 根据权利要求 17 的方法，其中所述给受试者施用是口服给药。

23. 根据权利要求 17 的方法，其中受试者是人。

24. 一种免疫抗 SARS 的方法，包括给受试者施用根据权利要求 1 的载体和可接受的药用载体。

25. 根据权利要求 24 的方法，其中受试者是人。

26. 含有其核酸序列如 SEQ ID NO: 7 所示的分离的核酸，或其片段，衍生物或类似物的植物。

27. 根据权利要求 26 的植物，其选自由烟草，莴苣，番茄，马铃薯，香蕉，玉米 (corn)，稻，谷类，小麦，玉米 (maize)，大麦，苹果，梨，草莓，胡萝卜，甜菜，薯蓣，奇异果，或菠菜组成的组。

28. 含有由权利要求 10 的植物细胞产生的 SARS-CoV 病毒抗原的组合物。

29. 含有由权利要求 16 的植物细胞产生的 SARS-CoV 病毒抗原的组合物。

30. 含有由权利要求 26 的植物细胞产生的 SARS-CoV 病毒抗原的组合物。

31. 一种检测样品中的对 SAR-CoV 病毒抗原的抗体的方法，包括：

(a) 使所述样品与权利要求 28 的组合物接触；并

(b) 检测与权利要求 28 的组合物结合的抗体的存在，从而检测

对 SAR-CoV 病毒抗原的抗体。

32. 根据权利要求 31 的方法，其中所述样品是生物流体。

33. 根据权利要求 32 的方法，其中所述生物流体是血液，血清，血浆，唾液，尿，粪便，痰，鼻咽吸出物，细胞或组织。

34. 一种检测样品中的对 SAR-CoV 病毒抗原的抗体的方法，包括：

(a) 使所述样品与权利要求 29 的组合物接触；并

(b) 检测与权利要求 29 的组合物结合的抗体的存在，从而检测对 SAR-CoV 病毒抗原的抗体。

35. 根据权利要求 34 的方法，其中所述样品是生物流体。

36. 根据权利要求 35 的方法，其中所述生物流体是血液，血清，血浆，唾液，尿，粪便，痰，鼻咽吸出物，细胞或组织。

37. 一种检测样品中的对 SAR-CoV 病毒抗原的抗体的方法，包括：

(a) 使所述样品与权利要求 30 的组合物接触；并

(b) 检测与权利要求 30 的组合物结合的抗体的存在，从而检测对 SAR-CoV 病毒抗原的抗体。

38. 根据权利要求 37 的方法，其中所述样品是生物流体。

39. 根据权利要求 38 的方法，其中所述生物流体是血液，血清，血浆，唾液，尿，粪便，痰，鼻咽吸出物，细胞或组织。

## 含有 SARS-CoV 病毒核苷酸序列的基因修饰的植物 和使用它进行免疫抗 SARS 的方法

本申请要求 2003 年 12 月 5 日提交的 U.S. Serial No. 60/527,637 的优先权，在此以引用的方式将其内容引入本申请。

在本申请中，某些出版物被作为参考。这些出版物以及其它的相关参考文献的完整引用可在紧邻权利要求之前找到。在此将这些出版物公开的内容引入到本申请中以更充分地描述本文中所描述的和要求保护的发明作出时的现有技术的状况。

### 发明背景

最近在中国大陆的广东省爆发了非典型性肺炎。在 2002 年 11 月到 2003 年 3 月之间，报道了 792 个病例，其中有 31 例死亡 (WHO, Severe Acute Respiratory Syndrome (SARS), Weekly Epidemiol. Rec. (2003) Vol. 78, page 86)。随着这次危机，香港医院管理局增加了对严重非典型性肺炎患者的监控。在此研究的过程中，鉴别出了许多成群的患有该病的卫生保健工作者。此外，在与那些感染者密切接触的人中有成群的肺炎事件。该疾病的严重程度和发展是不寻常的，尽管有对于已知通常与非典型性肺炎有关的细菌病原体的典型的抗生素治疗。本发明的发明人中的某些是参与对这些患者的研究的小组成员。该疾病被简称为严重急性呼吸综合症 (Severe Acute Respiratory Syndrome, "SARS")，其是由称为 SARS-CoV 病毒的新的冠状病毒引起的 (Peiris et al., 2003, Lancet Vol. 61, pages 1319-25; Fouchier et al., 2003, Nature Vol, 423, page 240)。SARS 相关的冠状病毒与以前表征的引起包括人在内的动物的呼吸和肠疾病的冠状病毒科 (*Coronaviridae*) 成员不同。

植物作为动物，包括人的营养物质，以及对于用作药物，化妆品等的物质的生产具有重要作用。通过植物核转化产生转基因植物已成功地生产出抗霍乱，诺沃克病毒 (Norwalk virus)，乙型肝炎，口蹄疫的粘膜疫苗 (Walmsley and Arntzen, 2003, Curr. Opin. Biotech. Vol, 14, pages 145-150)。在转基因植物中表达的病毒抗原有效诱导

动物中的粘膜和血清免疫反应,不管是肠胃外或是口服递送(Daniell et al., 2001, Trends Plant Sci. Vol. 6, pages 219-226)。

质体转化对重组 DNA 技术提供了巨大的好处。在植物细胞中与核基因组相比,存在高达 10,000 多个拷贝的质体基因组,这确保了质体表达的外源蛋白,或在疫苗生产中,抗原的产量的提高。而且,由于缺少花粉的传播,质体的母性遗传会导致外源基因的保持(Daniell et al., 2002, Trends Plant Sci. Vol, 7, pages 84-91)。人生长激素累积达到总可溶性蛋白的 7%,这是从核转化的烟草中得到的结果(Staub et al., 2000, Plant Journal Vol. 6, pages 547-553)的 300 倍,而人血清清蛋白累积达到总可溶性蛋白的 11.1%,这是从核转化的叶片中得到的结果(Fernandez-San Millan et al., 2003, Plant Biotech Vol. 1, pages 71-79)的 500 倍。霍乱毒素 B 亚基,首个由质体转化产生的植物衍生的疫苗,在烟草叶片中累积达到总可溶性蛋白的 4.1%(Daniell et al., 2001, J. Mol. Biol. Vol.311, pages 1001-1009)。比较而言,核转化植物中外源蛋白的产量很少超过总可溶性蛋白的 1%(Daniell et al., 2001, Trends Plant Sci. Vol. 6, pages 219-226)。

### 发明简述

本发明是基于本发明人发现表达 SARS-CoV 病毒抗原的基因修饰的植物及其后代可以用作抗 SARS 的疫苗。当基因修饰的植物或其后代被动物摄取时,或者基因修饰的植物或其后代的提取物被注入到动物,优选人中被其摄食时,产生抗 SARS-CoV 病毒抗原的抗体。抗 SARS-CoV 病毒的可食性疫苗,可作为果实(例如番茄),叶片(例如莴苣),块茎(例如马铃薯),种子(例如稻或玉米),花,茎或根递送,省去了从微生物获得重组疫苗所需的高成本的纯化程序。它进一步的优势在于容易储存,运输以及通过直接摄食给药。而且,果实,叶片,块茎,种子,花,茎或根形式的可食性疫苗可以在发展中国家培育并易于配发,省去了免疫程序的成本,也不需要冷冻和要求受过训练的卫生技术人员进行的通过注射的递送。

根据本发明,植物转化载体被工程化以提供用于引发动物的免疫反应从而预防和治疗 SARS 的 SARS-CoV 病毒抗原。所述抗原包括 SARS-CoV 病毒的蛋白质试剂或分子,其天然或人工变体,或其突变

体。含有 SARS-CoV 病毒序列的植物载体可以被工程化以提供 SARS-CoV 病毒的一个，两个，三个或更多个核苷酸序列。根据本发明，所述抗原序列可以是来自于所述 SARS-CoV 病毒，其天然或人工变体，或其突变体。

本发明提供了含有编码 SARS-CoV 病毒的多肽的片段，衍生物，类似物，或变体的 SARS-CoV 核苷酸序列的植物转化载体。在特定的实施方案中，本发明提供了转基因的和/或转质体 (transplastomic) 烟草植物。本发明提供了含有 SARS-CoV 抗原，包括 SARS-CoV 病毒的多肽或其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子的修饰的植物。这些活性包括功能活性以及免疫原性能力。本发明还提供了产生该修饰的植物的方法，包括用含有 SARS-CoV 病毒序列，其片段，衍生物，类似物，或变体的质体和/或核转化载体转化植物。在优选的实施方案中，本发明提供了表达少于 SARS-CoV 病毒的完整基因组的修饰的植物。在优选的实施方案中，该修饰的植物提供 SARS-CoV 病毒完整基因组的少于 95%，少于 90%，少于 85%，少于 80%，少于 75%，少于 70%，少于 65%，少于 60%，少于 55%，少于 50%，少于 45%，少于 40%，少于 35%，少于 30%，少于 25%，少于 20%，少于 15%，少于 10%，少于 5%，或少于 1%。在优选的实施方案中，该修饰的植物提供 SARS-CoV 病毒的至少 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 个多肽，片段，衍生物，类似物，和变体。

在另一个特定的实施方案中，本发明提供了抗 SARS 的疫苗。该疫苗含有该修饰的植物，其蛋白，或提取物。本发明还提供了免疫抗 SARS 的方法。所述方法包括给动物施用该修饰的植物以在动物中引发抗 SARS-CoV 病毒抗原的抗体的产生，所述抗原包括 SARS-CoV 病毒的多肽，或其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子。

在另一个特定的实施方案中，工程化植物转化载体以提供 SARS-CoV 病毒抗原用作血清检测试剂。本发明的一个方面涉及在样品中检测 SARS-CoV 病毒抗原的抗体的方法，其是通过将植物衍生的 SARS-CoV 病毒抗原与所述样品培育，检测与该植物衍生的 SARS-CoV 病毒抗原相结合的抗体的存在来进行的。所述样品可以是生物流体，

例如血液，血清，血浆，唾液，尿，粪便，痰，鼻咽吸出物，细胞和组织。对于抗体检测来说，这种植物衍生的抗原将会是那些由感染的细胞系产生的抗原的更便宜的替代物。

在一个实施方案中，本发明涉及含有由转化的植物细胞或植物产生的 SARS-CoV 病毒抗原的组合物。使用所述组合物检测 SARS-CoV 病毒抗原的抗体的方法也包括在内。在特定的实施方案中，方法包括使样品与含有由转化的植物细胞或植物产生的 SARS-CoV 病毒抗原的组合物接触，和检测与 SARS-CoV 病毒抗原结合的抗体。

在一个特定的实施方案中，本发明提供了含有编码 SARS-CoV 病毒的多肽的片段，衍生物，类似物，或变体的 SARS-CoV 病毒的核苷酸序列的质体转化载体。在另一个实施方案中，本发明提供了表达 SARS-CoV 病毒的多肽，其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子的质体转化载体。在特定的实施方案中，没有病毒 DNA 的载体是 pMLVHisA。在特定的实施方案中，本发明提供了通过质体转化载体 pCV1，pCV6，或 pCV8，和其含有具有核苷酸（而不是氨基酸）变化以使密码子使用最适于在植物中表达的 S1 的衍生物产生 SARS-CoV 病毒抗原的方法。

在另一个特定的实施方案中，本发明提供了含有 SARS-CoV 病毒的核苷酸序列，其片段，衍生物，类似物，或变体的核转化载体。在另一个实施方案中，本发明提供了表达 SARS-CoV 病毒的多肽，其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子的核转化载体。在特定的实施方案中，没有病毒 DNA 的载体是 pSa7。在特定的实施方案中，本发明提供了通过核转化载体 pCV2，pCV4，或 pCV12，和其含有具有核苷酸（而不是氨基酸）变化以使密码子使用最适于在植物中表达的 S1 的衍生物产生 SARS-CoV 病毒抗原的方法。

在特定的实施方案中，本发明进一步提供了其中具有含有来自于稻质体基因 *psbA* 的启动子，和终止子的载体的转质体烟草植物。包含含有 SARS-CoV 病毒的核苷酸序列的载体的植物细胞也是本发明的一个方面。含有表达 SARS-CoV 多肽，其衍生物，类似物，或变体的细胞的修饰的植物的植物部分，例如果实，叶片，块茎，种子，花，茎或根在本发明中提供。所述植物部分包括从完整植物上分离的或相



连于完整植物上的部分。在特定的实施方案中，本发明进一步使用选择标记基因 *aadA*，其具有壮观霉素抗性。在另一个实施方案中，本发明使用重组(His)<sub>5</sub>-标记蛋白的起始密码子和 *rbcL* 终止子。在优选的实施方案中，所述修饰的植物进一步含有表达蛋白酶抑制蛋白的异源核苷酸序列。

在另一个特定的实施方案中，本发明提供了用于表达 SARS-CoV 抗原的核转化载体的构建，所述抗原包括 SARS-CoV 病毒的多肽，或其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子。在特定的实施方案中，本发明提供了植物核转化载体 pCV2，pCV4 和 pCV 12。

在另一个特定的实施方案中，本发明提供了用于表达 SARS-CoV 抗原的质体转化载体的构建，所述抗原包括 SARS-CoV 病毒的多肽，或其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子。在特定的实施方案中，本发明提供了植物质体转化载体 pCV1，pCV6 和 pCV 8。

在特定的实施方案中，核转化载体用于表达一种或多种 SARS-CoV 抗原，包括 SARS-CoV 病毒的多肽，或其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子。

在特定的实施方案中，质体转化载体用于表达一种或多种 SARS-CoV 抗原，包括 SARS-CoV 病毒的多肽，或其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子。

在特定的实施方案中，质体转化载体和核转化载体用于表达一种或多种 SARS-CoV 抗原，包括 SARS-CoV 病毒的多肽，或其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子。

本发明提供了在植物中产生 SARS-CoV 病毒抗原的方法，所述抗原包括 SARS-CoV 病毒的多肽，或其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子。所述方法包括用含有与一个或多个 SARS-CoV 病毒多肽的编码序列可操作连接的启动子，和终止子的载体转化植物。包含含有一个或多个编码 SARS-CoV 病毒抗原，包括 SARS-CoV 病毒的多肽，或其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子的核酸序

列的载体的植物细胞也是本发明的一个方面。或者，所述植物细胞可以包含本发明的一个或多个载体。本发明提供了植物部分，例如所述修饰的植物的果实，叶片，块茎，种子，花，茎，根，和其它所有的解剖部分。

本发明提供了引发对 SARS-CoV 抗原，包括 SARS-CoV 病毒多肽，或其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子的免疫反应的疫苗。在一个特定的实施方案中，本发明提供了免疫抗 SARS 的方法。所述方法包括摄食可食用的转化的植物部分，所述植物部分含有编码 SARS-CoV 病毒多肽的 SARS-CoV 病毒序列的核酸分子。在另一个实施方案中，所述方法包括食用所述修饰的植物或静脉注射或摄食所述修饰的植物的提取物，该植物表达 SARS-CoV 抗原，包括 SARS-CoV 病毒多肽，或其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子。在另一个特定的实施方案中，本发明提供了免疫抗 SARS 的方法，其是给兔，山羊，牛，猪，绵羊，马，灵长类动物，灵猫，啮齿动物，浣熊，貉，狗，白鼬，鼬獾 (ferret Badger)，猫，鸟，或任何其它动物种类，包括人喂食本发明的修饰的植物或注射或摄食本发明的修饰的植物的提取物。本发明进一步提供了从免疫的动物中提取抗体。所述抗体对 SARS-CoV 抗原，包括 SARS-CoV 病毒的病毒多肽，或其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子是免疫特异的。

### 附图简述

图 1A-1W 显示了 pCV1 的限制性位点。

图 2A-2C 显示了 pCV1 (SEQ ID NO: 1) 的核酸序列。

图 3A-3B 显示了 pCV1 上特定限制酶的限制性图谱。

图 4 显示了衍生出 pCV1, pCV6 和 pCV8 的 pMLVHisA 质体转化载体。侧翼区域 *rbcL* 和 *accD* 来自于烟草质体基因组，用于在质体转化中进行同源重组。P<sub>psbA</sub> 代表用于表达插入基因的启动子。T<sub>psbA</sub> 代表终止子。P<sub>rrn</sub> 代表驱动壮观霉素抗性标记 *aadA* 表达的启动子。T<sub>rbcL</sub> 代表 *rbcL* 终止子。Met 代表重组(His)<sub>5</sub>-标记蛋白的起始密码子。(A) 显示了 pCV8 的序列 (SEQ ID NO: 2) 和限制性位点。(B) 显示了

pCV8 上特定限制酶的限制性图谱。(C) 显示了 pCV8 上特定限制酶的限制性图谱。(D) 显示了 pCV8 上特定限制酶的限制性图谱。

图 5 显示了衍生出 pCV2 和 pCV4 的 pSa7 核转化载体。RB 和 LB 代表 T (转移) -DNA 的右和左边界, 用于随机插入植物核基因组中。NOS-Pro 代表胭脂碱合酶 (NOS) 启动子。NOS-ter 代表 NOS-终止子。NPTII (Kan<sup>R</sup>) 代表具有卡那霉素抗性的新霉素磷酸转移酶。CaMV35S-Pro 代表花椰菜花叶病毒 35S 启动子。SaPIN2a cDNA 代表编码少花龙葵 (*Solanum americanum*) 蛋白酶抑制剂 IIA 的 cDNA。

图 6A-6I 显示了 SARS 病毒的完整核酸序列 (SEQ ID NO: 7)。

图 7A-7LL 显示了 pCV8 的限制性位点。

图 8A-8D 显示了 pCV8 的核酸序列 (SEQ ID NO: 2)。

图 9 显示了 pCV8 上特定限制酶的限制性图谱。

图 10 显示了 pCV8 上特定限制酶的限制性图谱。

图 11 显示了 pCV8 上特定限制酶的限制性图谱。

图 12 显示了 pCV12 核转化载体, 用于表达由 SARS-CoV S1 蛋白与绿色荧光蛋白 (GFP) 融合组成的融合蛋白。RB 和 LB 代表 T (转移) -DNA 的右和左边界, 用于随机插入植物核基因组中。NOS-ter 代表 NOS-终止子。CaMV35S-Pro 代表花椰菜花叶病毒 35S 启动子。S1 代表 SARS-CoV S1 蛋白以及 GFP, 绿色荧光蛋白。

图 13A-13B 分别显示了 M-基因的基因组 DNA 序列 (SEQ ID NO: 3), 和 M 蛋白的氨基酸序列 (SEQ ID NO: 4)。

图 14A-14B 显示了 S-基因的基因组 DNA 序列 (SEQ ID NO: 5)。

图 15 显示了 S 蛋白的氨基酸序列 (SEQ ID NO: 6)。

图 16A-16NN 显示了由 SEQ ID NO: 7 推定的三个框架的氨基酸序列。星号 (\*) 表示标志肽末端的终止密码子。

图 17A-17NN 显示了由 SEQ ID NO: 7 的互补序列推定的三个框架的氨基酸序列。星号 (\*) 表示标志肽末端的终止密码子。

图 18A-D 显示了 S1 : GFP 在农杆菌渗入的烟草叶片中的瞬时表达。代表性的烟草叶片表皮细胞在用含有表达 S1 : GFP 融合蛋白的质粒 pCV12 的根癌农杆菌 (*Agrobacterium tumefaciens*) LBA4404 (A, C) 或含有单独表达 GFP 的 pGDG 的 LBA4404 (B, D) 进行农杆菌渗入 (Yang Y. et al., In vivo analysis of plant promoters and

transcription factors by agroinfiltration of tobacco leaves. *Plant J.* 2000; 22: 543-551) 后 2 天由共焦显微镜显示。横条代表 20 $\mu$ m。

图 19A-19XX 显示了 pCV2 的限制性位点。

图 20A-20E 显示了 pCV2 的核酸序列 (SEQ ID NO: 8)。

图 21A-21NN 显示了 pCV6 的限制性位点。

图 22A-22D 显示了 pCV6 的核酸序列 (SEQ ID NO: 9)。

图 23A-23I 显示了转化后得到的植物以及对它们的分析。图 23A-23E 显示了用核转化载体 pCV2 进行农杆菌介导的对烟草和莴苣的转化后得到的再生的植物幼苗, 以及用引物 35S 和 NOS-ter 进行 PCR 然后用  $^{32}$ P-放射性标记的 S1 探针进行 Southern blot 分析对这些植物幼苗进行的分析。图 23F-23I 显示了用质体转化载体 pCV1 对烟草进行质体转化得到的再生的植物幼苗, 用 S1 引物对这些植物幼苗进行的 PCR 分析和用  $^{32}$ P-放射性标记的 S1 探针进行的 Northern blot 分析。图 23A 显示了在农杆菌介导的转化之后由烟草叶片再生的植物。图 23B 显示了 PCR 分析中使用的再生的烟草嫩芽。图 23C 显示了在农杆菌介导的转化之后由莴苣子叶再生的植物。图 23D 显示了 PCR 分析中使用的再生的莴苣嫩芽。图 23E 显示了在两个独立的烟草株系 (第 2 和 3 道), 和两个独立的莴苣株系 (第 4 和 5 道) 中存在 2.1-kb 的 S1 杂交条带 (标以箭头), 而在野生型烟草 (第 6 道) 和野生型莴苣 (第 7 道) 中缺乏此条带。另一个烟草株系检测为阴性 (第 1 道)。图 23F 和 23G 显示了微粒轰击后由烟草再生的植物幼苗。图 23H 显示了在再生的烟草 (第 2 道) 中有特定的 0.7-kb 的 PCR 条带, 而在野生型 (第 1 道) 中没有。图 23I 显示了在烟草株系 (第 2 道) 中存在 2.1-kb 的杂交 S1 mRNA 条带 (标以箭头), 而在野生型 (第 1 道) 中没有。图 23J 显示了用检测 His-标记的 S1 蛋白的 Ni-NTA 缀合物进行的表达 S1 的转质体烟草 western blot 分析。箭头指出了预期的 73-kDa 的 (His)<sub>5</sub>-S1 条带。

图 24 显示了用抗 GFP 的抗体进行的 Western blot 分析, 表明在农杆菌渗入之后烟草叶片中 S1 : GFP 的瞬时表达。从烟草叶片中提取的总蛋白 (200 $\mu$ g), 用单独表达 GFP 的质粒 pGDG (第 1 道), 或表达 S1 : GFP 融合蛋白的质粒 pCV12 渗入 (infiltrated), 在 8% 的 SDS-PAGE 凝胶上分离, 根据 Sambrook et al. (1989. *Molecular*

**Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor)** 印迹到 Hybond-C 滤膜上, 与抗 GFP 的抗体 (Clontech) 进行交叉反应。箭头指出了交叉反应的 S1 : GFP 条带 (计算大小为 99.1 kDa)。M, 分子量标记。

## 发明详述

### 术语的定义:

本文所用的术语“修饰的植物或植物部分”指植物或植物部分, 无论其与完整植物相连还是分离。它还包括通过有性或无性繁殖产生的所述修饰的植物或植物部分的后代。

本文所用的术语“变体”指天然存在的 SARS-CoV 的遗传突变体或 SARS-CoV 病毒的重组处理导致的变异, 其中每种都包括与保藏的 SARS-CoV 病毒: CCTCC-V200303 相比在基因组中的一个或多个突变, CCTCC-V200303 的序列如图 6A-6I 所示 (SEQ ID NO: 7)。术语“变体”还指给定多肽的天然存在的变异或给定多肽或蛋白质的重组处理导致的变异, 其中一个或多个氨基酸残基通过氨基酸的取代, 插入, 或缺失而被修饰。SARS-CoV 的天然变体的序列由于可能会或可能不会导致表型改变的一个或多个天然发生的突变, 包括但不限于对基因组序列的点突变, 重排, 插入, 缺失等而与 SARS-CoV 病毒的基因组序列不同。优选地, 所述变体包括相对于所述 SARS-CoV 病毒少于 25 个, 少于 20 个, 少于 15 个, 少于 10 个, 少于 5 个, 少于 4 个, 少于 3 个, 或少于 2 个氨基酸的取代, 重排, 插入, 和/或缺失。

在优选的实施方案中, 所述变体在一个或多个预测的非必需氨基酸残基 (即对于所述病毒的生物学活性, 例如感染性, 复制能力, 蛋白质合成能力, 装配能力, 和细胞毒性效果的表达不重要的氨基酸残基) 上具有保守的氨基酸取代。“保守的氨基酸取代”是指其中所述氨基酸残基被具有类似电荷的侧链的氨基酸残基取代。具有类似电荷侧链的氨基酸残基家族在本领域中已有定义。这些家族包括具有碱性侧链的 (例如赖氨酸, 精氨酸, 组氨酸), 酸性侧链的 (例如天冬氨酸, 谷氨酸), 不带电荷的极性侧链的 (例如甘氨酸, 天冬酰胺, 谷氨酰胺, 丝氨酸, 苏氨酸, 酪氨酸, 半胱氨酸), 非极性侧链的 (例如丙氨酸, 缬氨酸, 亮氨酸, 异亮氨酸, 脯氨酸, 苯丙氨酸, 甲硫氨酸

酸, 色氨酸),  $\beta$ -分支侧链的(例如苏氨酸, 缬氨酸, 异亮氨酸), 和芳香族侧链的(例如酪氨酸, 苯丙氨酸, 色氨酸, 组氨酸)氨基酸残基。在另一个实施方案中, 所述变体具有非保守性的氨基酸取代, 即氨基酸残基被不具有类似电荷侧链的氨基酸取代。

在特定的实施方案中, 可用于本发明的 SARS-CoV 序列包括那些保藏于 GenBank®的, 具有如下登记号的序列:

NC\_004718,

AY304495, AY304494, AY304493, AY304492, AY304491, AY304490, AY304489,  
AY304488, AY304487, AY304486, AY360146, AY278491, AY310120, AY278489,  
AY362699, AY362698, AY283798, AY283797, AY283796, AY283795, AY283794,  
AY268070, AY278741, AY340092, AY351680, AP006561, AP006560, AP006559,  
AP006558, AP006557, AY278554, AY348314, AY338175, AY338174, AY323977,  
AY322199, AY322198, AY322197, AH013000, AY322208, AY322207, AY322206,  
AY322205, AH012999, AY321118, AY323976, AY323975, AY323974, AY286320,  
AY290752, AY291315, AY307165, AY279354, AY278490, AY278487, AY297028,  
AY286402, AY274119, AY291451, AY271716, AY282752, AY278488, AY268049,  
AY269391,

在此将其全部引入作为参考。

本发明中使用的 SARS-CoV 病毒核苷酸序列可以是来自于突变 SARS-CoV 病毒。突变可以在所述 SARS-CoV 病毒或其变体的编码序列的全长或部分上随机引入, 例如通过饱和突变, 得到的突变体可以对其生物活性进行筛选以鉴别保留活性的突变体。在优选的实施方案中, 突变体多肽不保留野生型多肽的活性。在特定的实施方案中, 突变体多肽不保留 SARS-CoV 病毒的毒性活性。本领域已知的突变技术也可以使用, 包括但不限于定点突变, 化学突变, 体外定点突变, 使用, 例如 QuikChange® Site-Directed Mutagenesis Kit (Stratagene) 等。这种修饰的非限制性的实例包括将氨基酸替换成半胱氨酸, 以形成二硫键; 将氨基酸替换成酪氨酸, 然后对所述多肽进行化学处理以形成二酪氨酸键, 如本文中详细描述; 一个或多个氨基酸取代和/或生物或化学修饰以产生小分子(底物或抑制剂)的结合袋(binding pocket), 和/或引入侧链特异性标记(例如以表征分子相互作用或捕获蛋白质-蛋白质相互作用配体)。用于本发明的生物修饰包括烷基化,

磷酸化, 硫酸化, 氧化或还原, ADP-核糖基化, 羟基化, 糖基化, 葡糖基磷脂酰肌醇加成, 泛素化等。用于本发明的化学修饰包括, 例如改变重组病毒的电荷。当带电荷的氨基酸残基被改变为不带电荷的残基时, 正电荷或负电荷以化学方法添加到氨基酸残基上。

本文所用的术语“抗体”(“antibody”和“antibodies”)指单克隆抗体, 双特异性抗体, 多特异性抗体, 人抗体, 人源化抗体, 嵌合抗体, 骆驼化抗体, 单结构域抗体, 单链 Fvs (scFv), 单链抗体, Fab 片段, F(ab') 片段, 二硫化物连接的 Fvs (sdFv), 和抗独特型 (anti-Id) 抗体 (包括, 例如本发明的抗体的 anti-Id 抗体), 和上述任一种的表位结合片段。特别地, 抗体包括免疫球蛋白分子和免疫球蛋白分子的免疫活性片段, 即含有抗原结合位点的分子。免疫球蛋白分子可以是任何类型 (例如 IgG, IgE, IgM, IgD, IgA 和 IgY), 类 (例如 IgG1, IgG2, IgG3, IgG4, IgA1 和 IgA2) 或小类。

本文所用的术语“免疫特异地结合 SARS-CoV 病毒的多肽的抗体”指免疫特异地结合 SARS-CoV 病毒编码的多肽, 并且不非特异性结合其它多肽的抗体。免疫特异地结合 SARS-CoV 病毒的多肽的抗体与其它抗原没有交叉反应。优选地, 免疫特异地结合 SARS-CoV 病毒的多肽的抗体与其它抗原没有交叉反应。免疫特异地结合 SARS-CoV 病毒的多肽的抗体可以通过, 例如免疫检测或所属领域技术人员已知的其它技术鉴别。

本文所用的术语“表位”指在动物, 优选哺乳动物, 最优选人中具有抗原性或免疫原性活性的 SARS-CoV 病毒的片段, 多肽或蛋白质。具有免疫原性活性的表位是在动物中引起抗体反应的多肽的片段。具有抗原性活性的表位是由本领域公知的任何方法, 例如通过本文所述的免疫检测确定的抗体与之免疫特异地结合的多肽或蛋白质的片段。抗原性表位不一定具有免疫原性。

本文所用的术语“抗原性”指物质 (例如外源物体, 微生物, 药物, 抗原, 蛋白质, 肽, 多肽, 核酸, DNA, RNA 等) 在特定的生物体, 组织, 和/或细胞中引发免疫反应的能力。有时, 术语“抗原性的”与术语“免疫原性的”含义相同。

本文所用的术语“免疫原性”指物质 (例如外源物体, 微生物, 药物, 抗原, 蛋白质, 肽, 多肽, 核酸, DNA, RNA 等) 能在生物

体内激发免疫反应的性质。免疫原性部分取决于所述物质的大小，以及部分取决于宿主分子有多么不同于所述物质。高度保守的蛋白质倾向于具有相对较低的免疫原性。

本文所用的术语“在严谨条件下杂交”描述了杂交和洗涤的条件，在这种条件下互相具有至少 30%，至少 35%，至少 40%，至少 45%，至少 50%，至少 55%，至少 60%，至少 65%，至少 70%，至少 75%，至少 80%，至少 85%，至少 90%，至少 95% 相同性的核苷酸序列典型地保持互相杂交。这种杂交条件在，例如但不限于，*Current Protocols in Molecular Biology*, John Wiley & Sons, N. Y. (1989), 6.3.1-6.3.6.; *Basic Methods in Molecular Biology*, Elsevier Science Publishing Co., Inc., N. Y. (1986), pp. 75-78, and 84-87; 和 *Molecular Cloning*, Cold Spring Harbor Laboratory, N. Y. (1982), pp. 387-389 中描述，是本领域技术人员公知的。杂交和/或洗涤可以被实施的条件可以是 42-68°C，洗涤缓冲液可以含有从 0.1X 氯化钠/柠檬酸钠 (SSC)，0.5% SDS 到 6X SSC，0.5% SDS。典型地，杂交可以在 65°C (高度严谨条件)，60°C (中度严谨条件)，或 55°C (低严谨条件) 下进行过夜。滤膜可以用 0.1X SSC，0.5% SDS 在 65°C (高度严谨洗涤) 洗涤 2x15 分钟。滤膜用 0.1X SSC，0.5% SDS 在 63°C (中度严谨洗涤) 洗涤 2x15 分钟。对于低严谨洗涤，滤膜用 2X SSC，0.5% SDS 在 60°C 洗涤 2x15 分钟。严谨杂交条件的一个优选的，非限制性的实例是在大约 68°C 在 6X SSC，0.5% SDS 中杂交，然后在室温在 2X SSC，0.5% SDS 中洗涤一次或多次。严谨杂交条件的另一个优选的，非限制性的实例是在大约 45°C 在 6X SSC 中杂交，然后在大约 50-65°C 在 0.2X SSC，0.1% SDS 中洗涤一次或多次。

“分离的”或“纯化的”抗体基本不含从中获得该抗体的生物流体中的细胞材料或其它杂质蛋白。用语“基本不含细胞材料”包括抗体的制备物，其中所述抗体与用于分离所述抗体的细胞的细胞组分分离开。因此，基本不含细胞材料的抗体包括具有少于大约 30%，20%，10%，5%，2.5%，或 1% (按干重) 的杂质蛋白的抗体的制备物。在本发明的优选的实施方案中，所述抗体是分离的或是纯化的。

本文所用的术语“具有 SARS-CoV 病毒的多肽的生物活性”指与 SARS 多肽或具有如图 13B, 15, 16A-16NN, 和 17A-17NN 所示氨基



酸序列的多肽，或其片段具有一般生物学活性，类似或相同的结构域和/或具有足够的氨基酸相同性的多肽或蛋白质的特性。本发明的多肽的这种一般（Common）生物学活性包括抗原性和免疫原性。

本文所用的术语“与 SARS-CoV 病毒的多肽具有类似活性的分子”指与 SARS-CoV 病毒的多肽具有类似生物学活性和类似或相同结构域和/或与其具有足够的氨基酸相同性的多肽或蛋白质。这种生物学活性可以包括，但不限于抗原性，免疫原性，细胞毒性，激素活性，结合性质和亲和力，药理学活性，对生长增殖和分化的刺激或抑制，在细胞中诱导变化，抗病毒的，抗细菌的，抗真菌的和抗寄生虫的活性等。在优选的实施方案中，用于本发明的多肽可保留野生型 SARS-CoV 病毒的至少一种，两种，三种，四种，五种，或更多种生物学活性（例如感染性，复制能力，蛋白质合成能力，组装能力，和细胞毒性效果）。

本文所用的术语“部分”或“片段”指含有相应核酸分子上的长度至少大约

25, 30, 35, 40, 45, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1,000, 1,050, 1,100, 1,150, 1,200, 1,500, 2,000, 2,500, 3,000, 3,500, 4,000, 5,000, 6,000, 7,000, 8,000, 9,000, 10,000, 15,000, 20,000, 25,000,

或更多个连续核酸，并且具有所述核酸分子的至少一种功能特征（或者所编码的蛋白质具有由所述核酸分子编码的蛋白质的一种功能特征）的核酸分子的片段；或者含有相应蛋白质或多肽的长度至少

5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 320, 340, 360, 380, 400, 450, 500, 600, 700, 800, 900, 1,000, 2,000, 3,000, 4,000, 5,000, 6,000, 7,000, 8,000,

或更多个连续氨基酸残基，并且具有所述蛋白质或多肽的至少一个功能特征的蛋白质或多肽的片段。

本文所用的术语“类似物”在蛋白质试剂（例如蛋白质，多肽，肽，和抗体）的语境中是指与第二种蛋白质试剂（proteinaceous agent）具有类似或相同功能，但是不一定与所述第二种蛋白质试剂具有类似或相同氨基酸序列，或者与所述第二种蛋白质试剂具有类似或相同的结构的蛋白质试剂。在特定的实施方案中，抗体类似物免疫特异性地

结合与衍生出所述类似物的原始抗体相同的表位。在另一个实施方案中，抗体类似物免疫特异性地结合与衍生出所述类似物的原始抗体不同的表位。具有类似氨基酸序列的蛋白质试剂指满足至少下述之一的第二种蛋白质试剂：(a) 具有与第二种蛋白质试剂的氨基酸序列至少 30%，至少 35%，至少 40%，至少 45%，至少 50%，至少 55%，至少 60%，至少 65%，至少 70%，至少 75%，至少 80%，至少 85%，至少 90%，至少 95%，或至少 99% 相同的氨基酸序列的蛋白质试剂；(b) 由在严谨条件下与编码至少 5 个连续氨基酸残基，至少 10 个连续氨基酸残基，至少 15 个连续氨基酸残基，至少 20 个连续氨基酸残基，至少 25 个连续氨基酸残基，至少 40 个连续氨基酸残基，至少 50 个连续氨基酸残基，至少 60 个连续氨基酸残基，至少 70 个连续氨基酸残基，至少 80 个连续氨基酸残基，至少 90 个连续氨基酸残基，至少 100 个连续氨基酸残基，至少 125 个连续氨基酸残基，或至少 150 个连续氨基酸残基的第二种蛋白质试剂的核苷酸序列杂交的核苷酸序列编码的蛋白质试剂；和 (c) 由与编码第二种蛋白质试剂的核苷酸序列至少 30%，至少 35%，至少 40%，至少 45%，至少 50%，至少 55%，至少 60%，至少 65%，至少 70%，至少 75%，至少 80%，至少 85%，至少 90%，至少 95%，或至少 99% 相同的核苷酸序列编码的蛋白质试剂。与第二种蛋白质试剂具有类似结构的蛋白质试剂指与所述第二种蛋白质试剂具有类似的二级，三级或四级结构的蛋白质试剂。蛋白质试剂的结构可以通过本领域已知的方法确定，包括但不限于，肽测序，X 射线晶体学，核磁共振，圆二色性，和晶体电镜。

为了确定两个氨基酸序列或两个核酸序列的相同性百分数，所述序列出于最佳比较的目的被排列（例如可以在第一个氨基酸或核酸序列中引入缺口以与第二个氨基酸或核酸序列进行最佳排列）。然后位于相应氨基酸位置或核苷酸位置的氨基酸残基或核苷酸被比较。当第一个序列中的位置被与第二个序列中相应位置相同的氨基酸残基或核苷酸占据的时候，那么所述分子在该位置上就是相同的。两个序列之间的相同性百分数是所述序列共有的相同位置的数目的函数（即 % 相同性 = 相同的重叠位置的数目 / 位置总数目  $\times 100\%$ ）。在一个实施方案中，两个序列长度相同。

两个序列之间的相同性百分数还可以用数学算法确定。用于比较

两个序列的数学算法的优选的，非限制性的实例是 Karlin and Altschul, 1990, Proc. Natl. Acad. Sci. U.S.A. Vol, 87, pages 2264-2268 的算法，在 Karlin and Altschul, 1993, Proc. Natl. Acad. Sci U.S.A. Vol. 90, pages 5873-5877 中进行了修改。这种算法被整合到 Altschul et al., 1990, J. Mol. Biol. Vol. 215, page 403 的 NBLAST 和 XBLAST 程序中。BLAST 核苷酸搜索可以用 NBLAST 核苷酸程序参数集，例如 score=100, wordlength=12 进行，以得到与本发明的核酸分子同源的核苷酸序列。BLAST 蛋白质搜索可以用 XBLAST 程序参数集，例如 score=50, wordlength=3 进行，以得到与本发明的蛋白质分子同源的氨基酸序列。为得到比较目的的带缺口的排列，可以使用 Gapped BLAST，如 Altschul et al., 1997, Nucleic Acids Res. Vol. 25, pages 3389-3402 中所述。或者，可以使用 PSI BLAST 进行重复搜索，其发现分子之间的较远的关系 (Id.)。当使用 BLAST, Gapped BLAST, 和 PSI Blast 程序的时候，可以使用各个程序（例如 XBLAST 和 NBLAST）的默认参数（参见，例如 NCBI 网站）。用于序列比较的数学算法的另一个优选的，非限制性的实例是 Myers and Miller, 1988, CABIOS Vol.4, pages 11-17 的算法。这种算法在 ALIGN 程序（2.0 版）中使用，其是 GCG 序列比对软件包的一部分。当使用 ALIGN 程序比较氨基酸序列的时候，可以使用 PAM120 权重残基表，缺口长度罚分为 12，缺口罚分为 4。两个序列之间的相同性百分比可以用与上述类似的技术确定，允许或不允许有缺口。在计算相同性百分数的时候，典型地只有严格匹配被计算。

本文所用的，术语“类似物”在非蛋白质类似物的语境中指与第一种有机或无机分子具有类似或相同的功能，并且与第一种有机或无机分子在结构上类似的第二种有机或无机分子。

本文所用的，术语“衍生物”在蛋白质试剂（例如蛋白质，多肽，肽，和抗体）的语境中指含有通过引入氨基酸残基的取代，缺失，和/或插入而被改变的氨基酸序列的蛋白质试剂。本文所用的术语“衍生物”还指被修饰的蛋白质试剂，即通过任何类型的分子与该蛋白质试剂共价连接。例如，但不限于，抗体可以被修饰，例如通过糖基化，乙酰化，聚乙二醇化（PEGylation），磷酸化，酰胺化，用已知的保护/阻断基团衍生化，蛋白水解切割，与细胞配体或其它蛋白质连接等。

蛋白质试剂的衍生物可以用所述领域技术人员已知的技术通过化学修饰产生，包括但不限于特异性的化学切割，乙酰化，甲酰化，衣霉素的代谢合成等。进一步，蛋白质试剂的衍生物可以含有一个或多个非经典氨基酸。蛋白质试剂的衍生物具有与衍生出它的蛋白质试剂类似的或相同的功能。

本文所用的，术语“衍生物”在非蛋白质衍生物的语境中指基于第一种有机或无机分子的结构形成的第二种有机或无机分子。有机分子的衍生物包括，但不限于，例如通过加上或除去羟基，甲基，乙基，羧基或胺基基团而被修饰的分子。有机分子还可以被酯化，烷基化和/或磷酸化。

本文所用的术语“受试者”和“患者”可以互换使用。本文所用的“受试者”指动物，优选是包括非灵长类（例如兔，山羊，牛，猪，绵羊，马，麝猫，啮齿动物，浣熊，貉，狗，雪貂，鼬獾 (ferret Badger)，猫，和禽类物种）和灵长类（例如，猴如猕猴 (cynomolgus monkey) 和人）在内的哺乳动物，更优选是人。

本文所用的术语“SARS”或“SARS 相关症状”包括多种临床适应症和分类。如本文所述的，SARS 包括（1）无症状的或轻微的呼吸道疾病；（2）中度呼吸道疾病；或（3）严重呼吸道疾病。

本文所用的术语“载体 (carrier)”是用于支持或传递另一种物质例如色素，催化剂，或放射性材料的物质。

本文所用的术语“媒介物 (vehicle)”是便于药物，色素，或与之相混的其它物质的使用的物质。

本文所用的术语“赋形剂”指通常用作药物的稀释剂，媒介物，防腐剂，粘合剂，或稳定剂的情性物质，包括，但不限于蛋白质（例如血清清蛋白等），氨基酸（例如天冬氨酸，谷氨酸，赖氨酸，精氨酸，甘氨酸，组氨酸等），脂肪酸和磷脂（例如烷基磺酸盐，辛酸盐等），表面活性剂（例如 SDS，聚山梨醇酯，非离子表面活性剂等），糖类（例如蔗糖，麦芽糖，海藻糖等）和多元醇（例如甘露醇，山梨醇等）。另外参见 Remington's Pharmaceutical Sciences (by Joseph P. Remington, 18th ed., Mack Publishing Co., Easton, PA), 在此引入其全文。

本文所用的术语“可操作地结合”或“可操作地连接”指其中调

控区域（例如启动子，增强子）和要表达的核酸序列共价连接，并且其位置允许转录，和在合适的条件下，翻译的结合。

*优选的实施方案和实验细节：*

本发明提供了通过质体和/或核转化载体产生的转基因和转质体植物和它们的后代，从而产生抗 SARS-CoV 病毒抗原的粘膜疫苗。在转基因植物中表达的病毒抗原在动物中有效诱导粘膜和血清免疫反应，不管是肠胃外或是口服递送。在转基因植物中表达的病毒抗原还用作在血清学检测中进行抗体检测的试剂。本发明设法用含有来自 SARS-CoV 病毒的核苷酸序列的植物载体转化各种类型的植物，所述核苷酸序列包括但不限于编码 SARS-CoV 病毒的多肽的片段，衍生物，类似物，或变体的 SARS-CoV 病毒的核苷酸序列，或其编码与 SARS-CoV 病毒的多肽，片段，衍生物，类似物，或其变体具有类似活性的表位或蛋白质分子的核苷酸序列。

SARS-相关冠状病毒的基因组序列的分析表明它的十一个预测的开放阅读框架中的四个编码结构蛋白，包括具有 1255 个推定氨基酸的刺突糖蛋白（spike glycoprotein）（S）和具有 221 个推定氨基酸的膜糖蛋白（M）；S 和 M 可能结合形成病毒包膜。S 糖蛋白，一种 I 型膜蛋白，其功能是与宿主受体连接，具有 N-末端信号序列（S1；氨基酸 1-658）和 C-末端跨膜区，然后是胞质尾（Marra et al., 2003; Rota et al, 2003）。M 的 N-末端区域还包括不可切除的信号序列，暴露于病毒表面，而其 C-末端位于病毒膜内。虽然 S 和 M 都被认为靶定在质膜上，M 还被认为是定位于植物细胞的内质网上。在优选的实施方案中，修饰的植物或所述修饰的植物的植物部分或后代含有编码 SARS-CoV 病毒的刺突蛋白（S 或 S1），或其片段，衍生物，类似物，或变体的 SARS-CoV 病毒序列。在优选的实施方案中，修饰的植物或所述修饰的植物的植物部分或后代含有编码 SARS-CoV 病毒的刺突蛋白（S）的 N-末端区域（S1）和/或膜蛋白（M），或其片段，衍生物，类似物，或变体，或与所述 SARS-CoV 病毒的多肽具有类似活性的分子的核苷酸序列。在优选的实施方案中，所述植物或植物部分含有 SEQ ID NO: 3, 5, 或 7 的核苷酸序列，或其片段，或免疫特异性结合具有 SEQ ID NO: 3, 5, 或 7 的核苷酸序列的多肽，或其类似物，衍生

物，或片段，和/或由 SEQ ID NO: 3, 5, 或 7 的核苷酸序列编码的多肽，或其片段的核苷酸序列。本发明还涉及产生含有一个或多个质体和/或核转化载体的修饰的植物的方法，所述载体含有一个或多个编码 SARS-CoV 病毒的多肽的片段，衍生物，类似物或变体的 SARS-CoV 核苷酸序列。所有植物品种都可以用于本发明。在优选的实施方案中，所述植物是烟草，莴苣，马铃薯，番茄，香蕉，玉米 (corn)，稻，谷类，小麦，玉米 (maize)，大麦，苹果，梨，草莓，胡萝卜，甜菜，薯蓣，奇异果或菠菜。

### 转化载体的构建

本发明涉及含有 SARS-CoV 病毒的一个或多个核苷酸序列，或其片段，衍生物，类似物或变体的植物载体。在优选的实施方案中，所述植物载体是质体和核转化载体。在优选的实施方案中，所述核苷酸序列是 SARS-CoV 病毒的 M-基因和/或 S-基因。本发明还涉及含有一个或多个 SARS-CoV 病毒核苷酸序列，或其片段，衍生物，类似物或人工或天然变体的质体和核转化载体的构建。在特定的实施方案中，本发明提供了含有一个或多个核苷酸序列的植物转化载体，所述核苷酸序列是 SEQ ID NO: 7 的至少

5, 10, 15, 20, 25, 30, 35, 40, 45, 100, 150, 200, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1,000, 1,050, 1,100, 1,150, 1,200, 2,000, 3,000, 4,000, 5,000, 6,000, 7,000, 8,000, 9,000, 10,000, 11,000, 12,000, 13,000, 14,000, 15,000, 16,000, 17,000, 18,000, 19,000, 20,000, 21,000, 22,000, 23,000, 24,000, 25,000, 26,000, 27,000, 28,000, 29,000,

或更多个连续核苷酸，或其互补序列，或含有在严谨条件下与 SARS-CoV 病毒的核苷酸序列杂交的核苷酸序列的核苷酸。在特定的实施方案中，本发明提供了含有一个或多个核苷酸序列的植物转化载体，所述核苷酸序列是 SEQ ID NO: 3 的至少 5, 10, 15, 20, 25, 30, 35, 40, 45, 100, 150, 200, 300, 350, 400, 450, 500, 550, 600, 650, 或更多个连续核苷酸，或含有在严谨条件下与 SARS-CoV 病毒的核苷酸序列杂交的核苷酸序列的核苷酸。在特定的实施方案中，本发明提供了含有一个或多个核苷酸序列的植物转化载体，所述核苷酸序列是 SEQ ID NO: 5 的至少 5, 10, 15, 20, 25, 30, 35, 40, 45, 100,

150, 200, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1,000, 1,050, 1,100, 1,150, 1,200, 2,000, 3,000, 3,500 或更多个连续核苷酸, 或含有在严谨条件下与 SARS-CoV 病毒的核苷酸序列杂交的核苷酸序列的核苷酸。在特定的实施方案中, 本发明提供了含有一个或多个核苷酸序列的植物转化载体, 所述核苷酸序列编码的多肽是 SEQ ID NO: 4 的至少 5, 10, 15, 20, 25, 30, 35, 40, 45, 100, 150, 200, 或更多个连续氨基酸。在特定的实施方案中, 本发明提供了含有一个或多个核苷酸序列的植物转化载体, 所述核苷酸序列编码的多肽是 SEQ ID NO: 6 的 5, 10, 15, 20, 25, 30, 35, 40, 45, 100, 150, 200, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1,000, 1,050, 1,100, 1,150, 1,200, 或更多个连续氨基酸。在一个实施方案中, 所述植物载体含有编码 S 的氨基酸 14-1195 的核苷酸序列。在优选的实施方案中, 所述植物载体含有编码 S 的氨基酸 14-1195 上长度至少 20, 30, 40, 50, 60, 80, 100, 200, 300, 400, 500, 或更多个连续氨基酸的核苷酸序列。在特定的实施方案中, 所述多肽如图 16A-16NN 和 17A-17NN 所示, 或其片段, 衍生物或类似物。在特定的实施方案中, 本发明提供了含有一个或多个编码来自于 SARS-CoV 病毒的任何基因, 或其部分, 或变体, 片段, 类似物, 或衍生物的核苷酸序列的植物转化载体。这些基因包括, 但不限于包膜蛋白 (E 蛋白), 整体膜蛋白 (integral membrane protein) (M 蛋白), 刺突蛋白 (S 蛋白), 核衣壳蛋白 (N 蛋白), 红细胞凝集素酯酶 (HE 蛋白), 和 RNA 依赖的 RNA 聚合酶。特别有用的是那些由重组的 SARS-CoV 基因组的核酸片段编码的蛋白质物质。优选的是那些在开放阅读框架 (ORF) 内的, 特别是, 用于引发 SARS-CoV 特异性抗体或 T 细胞反应的, 不管是在体内 (例如为保护性或治疗性的目的, 或为提供诊断抗体) 或是体外 (例如通过噬菌体展示技术或另一种用于产生合成抗体的技术)。

所述病毒载体可以被工程化以提供抗原性分子, 包括 SARS-CoV 病毒的核苷酸序列, 或多肽, 包括所述病毒的重组和嵌合形式, 或所述病毒的亚基, 或其片段, 衍生物, 类似物, 或变体, 或与 SARS-CoV 病毒的多肽具有类似活性的分子。本发明进一步提供了制备 SARS-CoV 抗原的重组或嵌合形式的方法。

在另一个特定的实施方案中，本发明提供了通过单独施用疫苗制备物或 SARS-CoV 病毒抗体，或与抗病毒剂（例如金刚胺，金刚乙胺，扎那米韦（zanamivir），阿巴卡韦（abacavir），双汰芝（combivir），恩曲他滨（emtricitabine），替诺福韦（tenofovir），三协维（trizivir），恩夫韦地（enfuvirtide），更昔洛韦（gancyclovir），无环鸟苷，病毒唑，喷昔洛韦（penciclovir），奥塞米韦（oseltamivir），膦甲酸齐多夫定（AZT），去羟肌苷（didanosine, ddI），拉米夫定（lamivudine, 3TC），扎西他滨（zalcitabine, ddC），司他夫定（stavudine, d4T），奈韦拉平（nevirapine），地拉韦定（delavirdine），依非韦伦（efavirenz），fosamprenavir，安普那韦（amprenavir），阿扎那韦（atazanavir），kaletra，英地那韦（indinavir），利托那韦（ritonavir），阿糖腺苷（vidarabine），奈非那韦（nelfinavir），沙奎那韦（saquinavir），relenza，tamiflu，pleconaril，干扰素等），类固醇和皮质类固醇如强的松，可的松，氟替卡松（fluticasone）和糖皮质激素，抗生素，镇痛剂，支气管扩张剂（bronchodilators），或用于呼吸道和/或病毒感染的其它治疗组合使用，以治疗，改善，控制，或预防 SARS 的方法。在另一个特定的实施方案中，本发明的方法提供了草本植物，草本植物提取物，中药和其它药物与本发明的修饰的植物组合用于 SARS 的预防和治疗的用途。

进一步，本发明提供了含有本发明的抗病毒试剂和药学可接受的载体，媒介物或赋形剂的药物组合物。此外，本发明提供了含有脂质体包裹的植物提取物，优选纯化的，或本发明的抗病毒试剂的药物组合物。本发明还提供了含有本发明的药物组合物的试剂盒。

本发明包括编码所述 SARS-CoV 病毒的嵌合多肽的核苷酸序列的应用。在特定的实施方案中，所述嵌合多肽含有来自 SARS-CoV 病毒的两个或更多个不同株的氨基酸序列。根据本发明，本发明的修饰的植物进一步含有对于病毒基因组非天然的核苷酸序列。可用于本发明的核苷酸序列包括 SARS-CoV 病毒序列的部分，其进一步含有异源核苷酸序列。异源核苷酸序列可以来自于病毒，细菌，动物，或植物。在优选的实施方案中，所述异源核苷酸序列编码蛋白酶抑制剂。所述异源核苷酸序列使得所表达的 SARS-CoV 多肽更加稳定，并且减少了程序性细胞死亡，增加了本发明的修饰的植物的贮藏期限。在特定的



实施方案中，所述异源核苷酸序列（例如 *SaPIN2a* 或 *SaPIN2b*）使得所表达的 SARS-CoV 多肽更加稳定以使表达 SARS-CoV 多肽的修饰的植物可以在摄食，注射，或其它给药方法前被处理。*SaPIN2a* 和 *SaPIN2b*，从少花龙葵（*Solanum americanum*），属于茄科的一种草中分离的蛋白酶抑制剂 II 基因，编码使转基因植物具有昆虫抗性的丝氨酸蛋白酶抑制剂 II 蛋白（参见 2002 年 11 月 29 日提出的美国临时申请 No. 60/429,992；和 2003 年 12 月 1 日提出的美国申请 No. 10/725,829，律师案卷 No. 9661-043-999，在此引入其每一个的全文作为参考）。*SaPIN2a* 和 *SaPIN2b* 在韧皮部高表达，可能参与调节筛管元件的蛋白水解（参见，Xu et al., 2001, *Plant Mol Biol.* Vol. 47, pages 727-738；和 Xu et al., 2003, *Planta* Vol. 218, pages 623-629，在此引入其每一个的全文作为参考）。

在某些实施方案中，本发明涉及含有编码 SARS-CoV 病毒的嵌合多肽的核苷酸序列的载体和核酸分子。在特定的实施方案中，载体含有 SARS-CoV 的异源核苷酸序列。在另一个实施方案中，这种异源核苷酸序列被加入到，插入到天然或非天然序列中或取代之。根据本发明，SARS-CoV 病毒的核苷酸序列可以来自于 SARS-CoV 病毒的不同株或变体。

含有嵌合 SARS-CoV 病毒序列的植物载体特别可用于产生抗两种或多种病毒的重组疫苗（Tao et al., *J. Virol.* 72: 2955-2961；Durbin et al., 2000, *J. Virol.* 74: 6821-6831, Sliadopoulos et al., 1998, *J. Virol.* 72: 1762- 1768 (1998)；Teng et al., 2000, *J. Virol.* 74: 9317-9321)。例如，可以想象到含有 SARS-CoV 病毒核苷酸序列的植物载体可以表达 SARS-CoV 病毒变体的一个或多个肽，会保护受试者防止受到天然 SARS-CoV 和所述变体的感染。

根据本发明，所述植物载体可以被工程化以提供抗原性序列，所述抗原性序列在被受试者摄食时防止受到 SARS-CoV 和其天然变体感染。所述植物载体可以被工程化以提供 SARS-CoV 病毒的一种或多种抗原性序列。根据本发明，所述抗原性序列可以衍生自相同病毒，衍生自相同类型病毒的不同株或变体，或衍生自不同的病毒。

本发明提供了含有根据本发明的载体的宿主细胞。植物质体或核转化载体包含 SARS-CoV 病毒的核苷酸序列例如包含 SARS-CoV 基

基因组的全长，部分或片段以表达 SARS-CoV 病毒核酸。这些植物载体可以包含其它序列以产生可能含有 SARS-CoV 多肽的突变，缺失或插入的嵌合 SARS-CoV 病毒多肽。编码 SARS-CoV 病毒，SARS-CoV 病毒的多肽的片段，衍生物，类似物，或变体的核苷酸序列包括，但不限于保藏于 GenBank®的，具有如下登记号的序列：

NC\_004718, AY304495, AY304494, AY304493, AY304492, AY304491, AY304490, AY304489, AY304488, AY304487, AY304486, AY360146, AY278491, AY310120, AY278489, AY362699, AY362698, AY283798, AY283797, AY283796, AY283795, AY283794, AY268070, AY278741, AY340092, AY351680, AP006561, AP006560, AP006559, AP006558, AP006557, AY278554, AY348314, AY338175, AY338174, AY323977, AY322199, AY322198, AY322197, AH013000, AY322208, AY322207, AY322206, AY322205, AH012999, AY321118, AY323976, AY323975, AY323974, AY286320, AY290752, AY291315, AY307165, AY279354, AY278490, AY278487, AY297028, AY286402, AY274119, AY291451, AY271716, AY282752, AY278488, AY268049, AY269391,

在此全部引入作为参考。

在一个特定的实施方案中，植物细胞可以瞬时或稳定表达 SARS-CoV 病毒的一个或多个核苷酸序列。植物细胞通过转染（蛋白质或核酸载体），感染（病毒载体）或转导（病毒载体）被修饰。所述修饰的植物可以用于通过刺激体液免疫反应，细胞免疫反应，或刺激对抗原的耐受，从而调节受试者的免疫系统。本文所用的受试者是指：人，灵长类动物，兔，山羊，牛，猪，绵羊，马，灵猫，啮齿动物，浣熊，貉，狗，白鼬，鼬獾（ferret Badger），猫，鸟类，或其它非人动物。

编码 SARS-CoV 病毒抗原蛋白的核苷酸分子可以通过扩增克隆。术语“扩增”指从单个多核苷酸分子制备所述核酸的多个拷贝的过程。多核苷酸的扩增可以在体外通过本领域技术人员公知的生化过程实施。扩增试剂可以是能实现引物延伸产物的合成的任何化合物或系统，包括酶。为此目的的合适的酶包括，例如，*E. coli* DNA 聚合酶 I，*Taq* 聚合酶，*E. coli* DNA 聚合酶 I 的 Klenow 片段，T4 DNA 聚合酶，其它可获得的 DNA 聚合酶，聚合酶变体，逆转录酶，连接酶，和其它酶，包括热稳定酶（即在经受了足以引起变性的温度升高之后进行引物延伸的那些酶）。合适的酶会促使核苷酸以适当的方式组合，形

成与每个突变核苷酸链互补的引物延伸产物。通常，合成起始于每个引物的 3'末端，沿着模板链以 5'方向继续，直至合成终止，产生不同长度的分子。但是可能有扩增试剂从 5'端起始合成，沿着另一个方向继续，利用与上述相同的过程。在任何情况下，本发明的方法不限于本文所述的扩增的实施方案。

根据本发明可使用的一个体外扩增方法是美国专利 No. 4,683,202 和 4,683,195 所述的聚合酶链式反应 (PCR)。术语“聚合酶链式反应”指使用热稳定 DNA 聚合酶和两个寡核苷酸引物，来扩增 DNA 碱基序列的方法，一个引物于要扩增的序列的一端互补于 (+)-链，另一个于另一端互补于 (-)-链。因为新合成的 DNA 链随后可作为相同引物序列的另外的模板，连续轮的引物退火，链延长，和分离产生对所希望序列快速的和高度特异性的扩增。聚合酶链式反应用于检测样品中编码细胞因子的多核苷酸的存在。多种聚合酶链式方法是本领域技术人员已知的，可用于本发明的方法中。例如，DNA 可以如下所述在温度循环器 (thermocycler) 中进行 30 到 35 个循环的扩增：95°C 30 秒，52-60°C 1 分钟，72°C 1 分钟，最后延伸步骤是 72°C 5 分钟。在另一个实例中，DNA 可以在温度循环器中进行 35 个聚合酶链式反应，95°C 的变性温度 30 秒，然后是 54-58°C 的可变退火温度 1 分钟，70°C 的延伸步骤 1 分钟，以及 70°C 的最后延伸步骤。

用于扩增编码 SARS-CoV 病毒的蛋白质分子，或其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子的核苷酸序列的引物可以用任何适当的方法制备，例如传统的磷酸三酯和磷酸二酯方法或者其自动化实施方案，只要所述引物能够与感兴趣的多核苷酸杂交。在修饰的固体支持物上合成寡核苷酸的一种方法如美国专利 No. 4,458,066 中所述。引物的准确长度取决于多个因素，包括温度，缓冲液，和核苷酸组成。所述引物在存在扩增诱导试剂的条件下必须起始延伸产物的合成。

根据本发明的方法所使用的引物与要被扩增的核苷酸序列的每一条链互补。术语“互补”表示所述引物在允许聚合反应试剂发挥作用的条件下必须与各自的链杂交。换句话说，与侧翼序列互补的引物与侧翼序列杂交，并允许所述核苷酸序列扩增。优选地，被延伸的引物的 3'末端与互补侧翼链具有完全的碱基对互补性。编码所述抗原，

SARS-CoV 多肽或其片段, 衍生物, 类似物, 或变体, 或与 SARS-CoV 病毒的多肽具有类似活性的分子的核苷酸序列的引物和探针可以用已知的方法结合本发明公开的内容开发。

本领域普通技术人员知道还可用于增加靶核酸的拷贝数目的各种扩增方法。可用于本发明的多核苷酸可以在溶液中或是在与固体支持物结合以后, 通过通常用于特定核酸序列的检测的任何方法例如另一种聚合酶链式反应, 寡聚物限制 (Saiki et al., 1985, *Bio/Technology* Vol. 3, pages 1008-1012), 等位基因特异性寡核苷酸 (ASO) 探针分析 (Conner et al., 1983, *Proc. Natl. Acad. Sci. USA* Vol. 80, page 278), 寡核苷酸连接检测 (OLA) (Landegren et al., 1988, *Science* Vol. 241, page 1077), RNase 保护检测 (RPA) 等, 进一步被评估, 检测, 克隆, 测序等。DNA 分析的分子技术已有综述 (Landegren et al., 1988, *Science* Vol. 242, pages 229-237)。在 DNA 扩增后, 反应产物可以通过 Southern blot 分析检测, 不使用放射性探针。在这个过程中, 例如, 含有得自组织或受试者的多核苷酸的 DNA 的小样品被扩增, 并通过 Southern blotting 技术分析。非放射性探针或标记的使用由于扩增信号的高水平而便利。

用于扩增编码抗原, SARS-CoV 的多肽或其片段, 衍生物, 类似物, 或变体, 或与 SARS-CoV 病毒的多肽具有类似活性的分子的核苷酸序列的引物的大小是长度为至少 10, 15, 20, 25, 30 个核苷酸。特别地, 扩增 M-基因或 S-基因的引物是最优选的。优选地, G:C 比应当在 30%, 35%, 40%, 45%, 50%, 55%, 60% 以上以防止引物形成发夹结构。而且, 扩增子应当足够长以使其足以通过标准分子生物学方法检测。优选地, 扩增子的长度是至少 40, 60, 100, 200, 300, 400, 500, 600, 800, 1000, 或更多个碱基对。

可用于本发明的多核苷酸包括具有本文所示的 DNA 序列的多核苷酸, 以及额外包括编码含有相邻的和功能性抗原, SARS-CoV 多肽或其片段, 衍生物, 类似物, 或变体, 或与 SARS-CoV 病毒具有类似功能的分子的表位, 编码在高度严谨条件下与本发明所示的 DNA 序列的互补序列杂交的开放阅读框架 (ORF) 的任何核苷酸序列。作为实例而并非限制, 高度严谨杂交条件可以定义如下: 滤膜结合的 DNA 在含有 50% 去离子甲酰胺, 6X SSC, 5x Denhardt's, 1% SDS, 100  $\mu$ g/ml

变性鲑鱼精子 DNA 的溶液在 42°C 杂交过夜 (大约 4-16 小时), 在 0.1X SSC, 0.1% SDS 中在 65°C 洗涤 (Ausubel F. M. et al., eds., 1989, Current Protocols in Molecular Biology. Vol. I, Green Publishing Associates, Inc., and John Wiley & sons, Inc., New York) 并编码功能上等同的基因产物。对于寡核苷酸探针, 作为实例而并非限制, 高度严谨杂交可以指, 例如在 6XSSC/0.05% 焦磷酸钠中在 37°C (对于 14 个碱基的寡聚物), 48°C (对于 17 个碱基的寡聚物), 55°C (对于 20 个碱基的寡聚物), 和 60°C (对于 23 个碱基的寡聚物) 洗涤。

此外预期可用于本发明的多核苷酸包括在中度严谨条件下与编码抗原, SARS-CoV 的多肽或其片段, 衍生物, 类似物, 或变体, 或与 SARS-CoV 病毒的多肽具有类似活性的分子的 DNA 序列的互补序列杂交的任何核苷酸序列。作为实例而非限制, 这种中度严谨条件可以包括, 例如在 0.2X SSC/0.1% SDS 中在 42°C 洗涤 (Ausubel et al., 1989, 同上)。

此外预期可用于本发明的多核苷酸包括在低严谨条件下与编码抗原, SARS-CoV 的多肽或其片段, 衍生物, 类似物, 或变体, 或与 SARS-CoV 病毒的多肽具有类似活性的分子的 DNA 序列的互补序列杂交的任何核苷酸序列。作为实例而非限制, 使用这种低严谨条件的程序如 Shilo and Weinberg, 1981, Proc. Natl. Acad. Sci. USA Vol. 78, pages 6789-6792 中所述。变体与野生型蛋白相比在蛋白质的氨基酸序列中可含有一个或多个改变, 例如通过插入, 取代, 或缺失一个或多个氨基酸。任何改变应当不破坏所述蛋白行使其功能的能力, 但是可以增加或减少这种能力, 这取决于所述改变的特性。优选地, 氨基酸的改变是保守的。

在各个实施方案中, 抗原, SARS-CoV 病毒的多肽或片段, 变体, 类似物, 或衍生物可以表达为融合, 或嵌合蛋白产物 (含有通过肽键与异源蛋白序列 (不同蛋白质的) 连接的酶, 片段, 类似物, 或衍生物)。这种嵌合基因产物的制备可以通过将编码所希望氨基酸序列的合适的核酸序列用本领域公知的方法以正确的编码框架彼此相连, 用本领域公知的方法表达所述嵌合产物。或者, 这种嵌合产物的制备可以通过蛋白质合成技术, 例如使用肽合成仪。优选地, 所述融合蛋白中的酶的片段, 类似物, 和衍生物保留有行使所述酶的功能的能力。

对于植物核转化载体的构建, SARS-CoV 核苷酸序列, 基因, 其片段, 衍生物, 类似物或变体的表达可以通过多个调控元件中的任何元件来驱动。例如, 病毒启动子例如 CaMV 的 35S RNA 和 19S RNA 启动子 (Brisson et al., 1984, Nature Vol. 310, pages 511-514), 或者 TMV 的外壳蛋白启动子 (Takamatsu et al., 1987, EMBO J. Vol. 6, pages 307-311) 可以被使用; 或者, 植物启动子例如 RUBISCO 的小亚基 (Coruzzi et al., 1984, EMBO J. Vol. 3, pages 1671-1680; Broglie et al., 1984, Science Vol. 224, pages 838-843); 或热休克启动子, 例如, 大豆 hsp17.5-E 或 hsp17.3-B (Gurley et al., 1986, Mol. Cell. Biol. Vol. 6, pages 559-565) 可以被使用。这些构建体可以用 Ti 质粒, Ri 质粒, 植物病毒载体, 直接 DNA 转化, 生物弹/微粒轰击, 显微注射, 电穿孔等引入到植物细胞中。这些技术的综述参见, 例如, Weissbach & Weissbach, 1988, Methods for Plant Molecular Biology, Academic Press, New York, Section VIII, pp. 421-463; 和 Grierson & Corey, 1988, Plant Molecular Biology, 2d Ed., Blackie, London, Ch. 7-9。本文所用的调控元件包括但不限于诱导型和非诱导型启动子, 增强子, 操纵子和本领域已知的驱动和调控表达的其它元件。优选所述启动子能够指导在所述植物的特定组织中和/或所述植物发育的特定阶段中的表达。所述启动子对于所述植物可以是异源的或是同源的。优选所述启动子指导在果实, 例如番茄, 叶片, 例如莴苣, 植物种子的胚乳或植物的根或块茎中的表达。优选的启动子是小麦的高分子量麦谷蛋白 (HMWG) 基因。其它合适的启动子是技术人员公知的, 例如麦醇溶蛋白, 分枝酶, ADPG 焦磷酸化酶, 淀粉合酶和肌动蛋白启动子, 例如。

具有表达 SARS-CoV 抗原包括 SARS-CoV 病毒的多肽, 或其片段, 衍生物, 类似物, 或变体, 或与 SARS-CoV 病毒的多肽具有类似活性的分子的能力的转化植物可以通过用含有编码抗原, 包括 SARS-CoV 的多肽或其片段, 衍生物, 类似物, 或变体, 或与 SARS-CoV 病毒的多肽具有类似活性的分子的序列的载体转化植物细胞来构建。在一个实施方案中, 植物启动子与编码所希望的抗原的序列可操作结合。本文所用的术语“可操作结合”或“可操作连接”指其中调控区域(例如启动子, 增强子)和要表达的核酸序列共价连接, 并且其位

置允许转录，和在合适的条件下翻译的结合。在本发明的优选实施方案中，用于核转化的结合的启动子是强的并且是非组织或发育特异性的植物启动子（例如在多种或所有植物组织类型中强表达的启动子）。这种强的“组成型”启动子的实例包括，但不限于 CaMV 35S 启动子（Odell et al., 1985, Nature 313: 810-812），T-DNA 甘露氨酸合成酶启动子，以及它们的各种衍生物。在另一种优选的实施方案中，诱导型或阻遏型启动子用于在植物中表达感兴趣的 SARS-CoV 病毒，例如 Weinmann et al., 1994, The Plant Journal 5: 559-569 中所述的 *tet* 操纵子启动子；或 McNellis et al., 1998, The Plant Journal 14: 247-257 中所述的糖皮质激素诱导型启动子；或 Caddick et al., 1998, Nature Biotechnology 16: 177-180 中所述的乙醇诱导型启动子。还参见 Gatz, 1995, Methods In Cell Biology 50: 411-424，其中描述了植物的诱导型和阻遏型基因表达系统。

用于质体转化的启动子包括质体表达中的强的和组成型启动子，包括 *psbA* 启动子（*psbA* 基因编码光系统 II 32kD 蛋白质）和 16S rRNA 操纵子（*rrn*）启动子，或这些启动子的修饰，其增强了表达（Suzuki et al., 2003, Plant Cell 15: 195-205；还参见 PCT 公开 no. WO00/03012）。

在本发明的一个实施方案中，SARS-CoV 抗原定位于核表达的质外体空间。当在植物中表达的时候，通过将所述抗原与作为信号或转运体的肽一起表达为融合蛋白以使所述抗原定位于转基因植物的质外体空间，SARS-CoV 可导向质外体空间。多种信号或转运肽都可以使用，例如 Lund et al., 1992, Plant Molecular Biology 18: 47-53 中所述的 PR1b 信号序列；或 Pfitzner et al., 1987, Nucleic Acids Research 15: 4449-4465 中所述的 PR-1a, b 和 c 信号序列。含有信号或转运肽和 SARS-CoV 抗原的融合蛋白的构建可以通过将对每个组分特异性的多核苷酸彼此连接起来（例如所述多核苷酸被连接在框架中）以使当融合的多核苷酸在转基因植物中表达的时候制备了所希望的融合蛋白。技术人员应当知道如何构建用于在转基因植物的质外体空间表达 SARS-CoV 抗原，包括 SARS-CoV 多肽或其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子的多核苷酸。

在本发明的另一个实施方案中，构建具有含有与组织特异性或发

育特异性启动子，例如但不限于番茄 E8 果实特异性启动子，查耳酮合酶（CHS）启动子，patatin 启动子可操作结合的编码 SARS-CoV 抗原，SARS-CoV 多肽或其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子的序列的载体的植物是有利的。

在本发明的另一个实施方案中，用含有与修饰的或人工启动子可操作连接的编码 SARS-CoV 抗原，SARS-CoV 多肽或其片段，衍生物，类似物，或变体，或与 SARS-CoV 病毒的多肽具有类似活性的分子的序列的载体转化植物是有利的。典型地，这种启动子，通过重组不同启动子的结构元件构建，具有独特的天然启动子不具有的表达模式和/或水平。参见，例如 Salina et al., 1992, *Plant Cell* 4: 1485-1493，例如将顺式调控元件与启动子核心组合构建人工启动子。

在本发明的另一个实施方案中，SARS-CoV 抗原，SARS-CoV 多肽或其片段，衍生物，类似物，或变体，或者与 SARS-CoV 病毒的多肽具有类似活性的分子的表达可以用本领域公知的方法通过提高编码所希望蛋白或多肽的基因的拷贝数目进行工程化。

本发明提供了能指导 SARS-CoV 病毒核苷酸序列，其片段，衍生物，类似物，或变体在基因修饰的植物或其后代，包括，例如转基因核转质体植物中表达的载体。所述植物载体通过生物技术领域公知的一般重组 DNA 和克隆技术构建，参见，例如 Sambrook et al., *supra*; Ausubel et al., *supra*。这种多核苷酸构建体典型地含有编码工程化基因产物和一个或多个调控多核苷酸序列的多核苷酸序列。用于本发明的多核苷酸构建体的调控序列包括但不限于，启动子，增强子，内含子，剪接供体，剪接受体，多聚腺苷酸化序列，RNA 稳定性调节序列，或者上述任何一个的元件（例如启动子元件包括，但不限于 TATA 盒）。用于本发明的调控元件能够指导在其中希望表达编码 SARS-CoV 病毒抗原的核苷酸序列的植物种中的表达。在另一个优选的方面，所述调控元件能够指导在所感兴趣的植物种中希望表达所述工程化基因产物的细胞类型中的表达。

用于本发明的调控元件是本领域技术人员公知的，例如，已知在所感兴趣的细胞类型和植物种中表达的基因的启动子和增强子元件。用于在感兴趣的植物种中表达 SARS-CoV 病毒抗原的启动子还可以用



常规的实验方法分离，例如通过分离已知以所希望方式表达的基因的启动子区域。例如，可以用对已知在感兴趣的植物种中的感兴趣的细胞类型中表达的信使 RNA 的 5'端特异的 cDNA 探针筛选基因文库。这种 5'端 cDNA 探针优选仅是大约 100 个碱基对到大约 300 个碱基对，以使基因文库中鉴别的克隆可能包含基因的 5'端，该基因可能包含所述探针对其特异的基因的启动子区域的。所述启动子区域典型地包括转录起始位点上游的大约 1,000 到大约 2,000 个碱基对。因此，用于表达所构建的本发明的 SARS-CoV 核苷酸序列，基因，其片段，衍生物，类似物或变体的表达的启动子是已知在所感兴趣的植物种中的所感兴趣的细胞类型中表达的基因的转录起始位点的上游大约 2,000 个碱基对到下游大约 50 个碱基对，或是该多核苷酸的一部分。

为了便于正确处理 SARS-CoV 病毒抗原，可能需要包括编码这种处理所必需的肽序列的核苷酸序列。例如，被鉴别的在转基因宿主植物中发挥作用，例如，使所述抗原容易进入内质网的肽序列可能是必需的，即，信号序列。

### 植物和植物细胞的转化

植物和植物细胞核和质体可以用本领域已知的任何方法转化。在本发明的一个实施方案中，在核转化中农杆菌 (*Agrobacterium*) 被用于将本发明的载体引入。这种转化优选使用二元农杆菌 T-DNA 载体 (Bevan, 1984, *Nuc. Acid Res.* Vol. 12, pages 8711-8721) 和共培养步骤 (Horsch et al., 1985, *Science* Vol. 227, pages 1229-1231)。通常，农杆菌转化系统用于工程化双子叶植物 (Bevan et al., 1982, *Ann. Rev. Genet* Vol. 16, pages 357-384; Rogers et al., 1986, *Methods Enzymol.* Vol. 118, pages 627-641)。农杆菌转化系统也可以用于将 DNA 转化，以及转移到单子叶植物和植物细胞中 (参见 Hernalsteen et al., 1984, *EMBOJ.* Vol. 3: pages 3039-3041; Hooykaas-Van Slogteren et al., 1984, *Nature* Vol. 311, pages 763-764; Grimsley et al., 1987, *Nature* Vol. 325, pages 1677-179; Boulton et al., 1989, *Plant Mol. Biol.* Vol. 12, pages 31-40.; and Gould et al., 1991, *Plant Physiol.* Vol. 95, pages 426-434)。

在其它实施方案中，将重组核酸构建体引入到植物和植物细胞中

的各种替代方法也可以使用。这些其它方法特别用于靶是单子叶植物或植物细胞的情况。替代的基因转移和转化方法包括，但不限于，基因枪轰击 (particle gun bombardment) (基因枪 (biolistics))，通过钙<sup>2+</sup>，聚乙二醇 (PEG) -或电穿孔-介导的裸 DNA 吸收进行的原生质体转化 (参见 Paszkowski et al., 1984, EMBO J. Vol. 3, pages 2717-2722; Potrykus et al. 1985, Molec. Gen. Genet. Vol.199, pages 169-177; Fromm et al., 1985, Proc. Nat. Acad. Sci. USA Vol. 82, pages 5824-5828; 和 Shimamoto, 1989, Nature Vol. 338, pages 274-276) 和植物组织的电穿孔 (D'Halluin et al., 1992, Plant Cell Vol. 4, pages 1495-1505)。植物细胞转化的其它方法包括显微注射，碳化硅介导的 DNA 吸收 (Kaepler et al., 1990, Plant Cell Reporter Vol. 9, pages 415-418)，和微弹轰击 (microprojectile bombardment) (参见 Klein et al., 1988, Proc. Nat. Acad. Sci. USA Vol.85, pages 4305-4309; 和 GordonKamm et al., 1990, Plant Cell Vol. 2, pages 603-618)。在任何方法中，可以使用选择性标记，至少是在最开始的时候，以确定转化是否实际发生了。可用的选择性标记包括提供对抗生素，例如庆大霉素，潮霉素，卡那霉素等的抗性的酶。或者，可以使用提供通过颜色变化识别的化合物，例如 GUS，通过发光识别的化合物，例如荧光素酶的标记。

嵌合基因还可以含有基因开关机制，其决定了在什么条件下或者什么时候所述编码序列被表达。例如，这些基因开关可以是化学诱导的启动子或温度控制的启动子。

在特定的实施方案中，每个质体转化构建体 (pCV1, pCV6 或 pCV8，和其含有具有核苷酸 (而不是氨基酸) 变化以使密码子使用最适于在植物中表达的 S1 的衍生物) 通过基因枪轰击引入到烟草 (Staub and Maliga, 1994, Plant Journal Vol. 6, pages 547-553) 和番茄 (Ruf et al., 2001, Nature Biotech Vol. 19, pages 870-875) 中。在特定的实施方案中，通过核转化产生表达 S1 或 M 抗原的转基因烟草和转基因莴苣是用农杆菌介导的转化 (Horsch et al., 1985, Science Vol. 227, pages 1227-1231) 以及分别使用质粒 pCV2 [和其含有具有核苷酸 (而不是氨基酸) 变化以使密码子使用最适于在植物中表达的 S1 的衍生物] 和 pCV4 实现的。在另一个特定的实施方案中，通过核

转化产生表达 S1 或 M 抗原的转基因番茄。

根据本发明，多种植物和植物细胞系统可以用本发明的核酸构建体和上述的各种转化方法工程化以具有所希望的生理学和农学特性。在优选的实施方案中，用于工程化的靶植物和植物细胞包括，但不限于，那些单子叶和双子叶植物，例如农作物包括谷类作物（例如小麦，玉米，稻，粟，大麦），烟草，果实作物（例如番茄，苹果，梨，草莓，橙），饲料作物（例如苜蓿），根菜作物（例如胡萝卜，马铃薯，甜菜，薯蓣），叶菜作物（例如莴苣，菠菜），开花植物（例如矮牵牛花，玫瑰，菊花），针叶树和松树（例如冷杉，云杉）；用于植物修复的植物（例如重金属积聚植物）；油料作物（例如向日葵，油菜籽）；和用于实验目的的植物（例如拟南芥（*Arabidopsis*））。

### 转化植物和植物细胞的筛选

根据本发明，所希望的植物可以通过将一个或多个表达本文所述的 SARS-CoV 抗原的载体工程化到多种植物细胞类型，包括但不限于原生质体，组织培养细胞，组织和器官外植体，花粉，胚胎，以及完整植物中获得。在本发明的一个实施方案中，工程化的植物材料用下述步骤和方法选择或筛选转化体（那些掺入或整合了所引入的基因构建体的）。分离的转化体然后通过有性或无性繁殖或生长再生成植物及其后代。或者，工程化的植物材料可以在对衍生出的植物进行标记基因性状的选择或筛选之前再生成植物。从植物细胞，组织或器官，在选择或筛选标记基因之前或之后再再生植物的步骤是本领域技术人员公知的。

转化的植物细胞，愈伤组织，组织或植物可以通过选择或筛选工程化的植物材料中由转化 DNA 上存在的标记基因编码的性状被鉴别和分离。例如，选择可以通过在含有抑制量的转化基因提供了针对其的抗性的抗生素或除草剂的介质上培养工程化的植物材料来进行。进一步，转化的植物和植物细胞还可以通过筛选本发明的载体中可能存在的任何可见标记基因（例如 $\beta$ -葡糖醛酸酶，荧光素酶，B 或 C1 基因）的活性被鉴别。这种选择和筛选方法是本领域技术人员公知的。

物理和生化方法还可以用于鉴别含有本发明的基因构建体的植物或植物细胞转化体。这些方法包括但不限于：1) 用于检测和确定重

组 DNA 插入物的结构的 Southern 分析或 PCR 扩增；2) 用于检测和检查所述基因构建体的 RNA 转录物的 Northern blot, S1 RNase 保护, 引物延伸或反转录酶-PCR 扩增；3) 用于检测酶活性的酶检测, 其中该基因产物由所述基因构建体编码；4) 蛋白质凝胶电泳 (PAGE), Western blot 技术, 免疫沉淀, 或酶联免疫检测, 其中所述基因构建体产物是蛋白质。其它技术, 例如原位杂交, 酶染色, 和免疫染色, 也可以用于检测在特定植物器官和组织中所述重组构建体的存在或表达。进行所有这些检测的方法是本领域技术人员公知的。

在特定的实施方案中, 选择标记基因 *aadA*, 其赋予壮观霉素抗性, 在质体转化中由稻质体 16S rRNA 操纵子 (*rrn*) 启动子驱动。在特定的实施方案中, 选择标记基因 *nptII*, 其赋予卡那霉素抗性, 在核转化中被 NOS (胭脂碱合酶) 启动子驱动。

本发明涉及表达 SARS-CoV 病毒的一个或多个表位的转基因和转质体植物。表达 SARS-CoV 抗原的转基因或转质体植物导致在受试者摄食修饰的植物或植物部分之后, 或用本领域已知的方法给受试者施用植物提取物之后, 受试者产生抵抗 SARS-CoV 的抗体。

植物的实例是单子叶植物, 双子叶植物, 作物植物 (即为农业目的, 为动物包括人的食物生产的目的培养的任何植物种, 典型地以超过大约 10 株植物成群种植以收获完整植物或所述植物具有的植物部分, 例如果实, 花或作物的植物, 例如烟草, 谷物等), 树 (即果树, 为木材生产培养的树, 为装饰培养的树等), 任何种类的花 (即为装饰目的培养的植物, 例如在收获之后), 仙人掌。

其中 SARS-CoV 病毒抗原可以被表达的植物的更多实例包括植物界, 链形植物 (Streptophyta), 有胚植物, 维管植物, 真叶植物 (Euphyllophytes), 种子植物门, 木兰门, 百合纲 (Liliopsida), 鸭跖草亚纲 (Commelinidae), 禾本目 (Poales), 禾本科 (Poaceae), 稻属 (*Oryza*), 水稻 (*Oryza sativa*), 玉蜀黍属 (*Zea*), 玉米 (*Zea mays*), 大麦属 (*Hordeum*), 大麦 (*Hordeum vulgare*), 小麦属, 小麦 (*Triticum aestivum*), 真双子叶植物, 核心真双子叶分支 (Core eudicots), 菊亚纲 (Asteridae), 真菊分支 (Euasterids), 蔷薇亚纲 (Rosidae), II 类真蔷薇分支 (Eurosids II), 十字花目 (Brassicales), 十字花科, 阿布属 (*Arabidopsis*), 木兰纲 (Magnoliopsida), Solananae, 茄目

(Solanales), 茄科, 茄属 (*Solanum*), 和烟草属 (*Nicotiana*)。

还包括, 例如, 特别感兴趣的农作物包括茄科, 包括处理的和新鲜的市售的番茄, 胡椒和茄子; 多叶植物, 包括莴苣和菠菜; 芸苔 (*Brassicas*), 包括球花甘蓝, 球芽甘蓝, 花茎甘蓝, *cale*, 花椰菜, 红球甘蓝和白球甘蓝; 葫芦科, 包括黄瓜, 甜瓜, 西瓜, 绿皮西葫芦和南瓜; 大粒种子植物, 包括豌豆, 菜豆和甜玉米; 生根植物, 包括胡萝卜和洋葱; 无性繁殖植物, 包括浆果, 葡萄, 香蕉, 菠萝, 奇异果, 和蔷薇科果实和坚果作物; 和热带作物, 包括芒果和木瓜。

因此本发明可应用于很多种植物中, 包括, 但不限于下述属的种: 腰果 (*Anacardium*), 花生 (*Arachis*), 芦笋 (*Asparagus*), 颠茄 (*Atropa*), 燕麦 (*Avena*), 芸苔 (*Brassica*), 柑橘 (*Citrus*), 西瓜 (*Citrullus*), 辣椒 (*Capsicum*), 红花 (*Carthamus*), 椰子 (*Cocos*), 咖啡 (*Coffea*), 甜瓜 (*Cucumis*), 南瓜 (*Cucurbita*), 胡萝卜 (*Daucus*), 油棕 (*Elaeis*), 草莓 (*Fragaria*), 大豆 (*Glycine*), 棉 (*Gossypium*), 向日葵 (*Helianthus*), *Heterocallis*, 大麦 (*Hordeum*), 天仙子 (*Hyoscyamus*), 莴苣 (*Lactuca*), 亚麻 (*Linum*), 黑麦草 (*Lolium*), 羽扇豆属 (*Lupinus*), 番茄属 (*Lycopersicon*), 苹果属 (*Malus*), 木薯属 (*Manihot*), *Majorana*, 苜蓿属 (*Medicago*), 烟草属 (*Nicotiana*), 木樨榄属 (*Olea*), 稻属 (*Oryza*), 稷属 (*Panicum*), *Pannaserum*, 鳄梨属 (*Persea*), 菜豆属 (*Phaseolus*), *Pistachia*, 豌豆属 (*Pisum*), 梨属 (*Pyrus*), 李属 (*Prunus*), 萝卜属 (*Raphanus*), 蓖麻属 (*Ricinus*), 黑麦属 (*Secale*), 千里光属 (*Senecio*), 白芥 (*Sinapis*), 茄属 (*Solanum*), 高粱属 (*Sorghum*), *Theobromus*, 葫芦巴属 (*Trigonella*), 小麦属 (*Triticum*), 蚕豆属 (*Vicia*), 葡萄属 (*Vitis*), 豇豆属 (*Vigna*), 和玉蜀黍属 (*Zea*)。

### 疫苗制剂和给药

本发明还提供了含有本发明的修饰的植物的提取物的疫苗制剂, 其适合于给药以引发抗特定抗原的保护性的免疫 (体液和/或细胞介导的) 反应, 例如用于 SARS 的治疗和预防。

这种疫苗的适当的制剂包括注射剂, 作为液体溶液或悬液; 适合于在注射前配制成液体中的溶液或悬液的固体形式也可以被制备。所述制剂还可以被乳化, 或者蛋白质分子被包裹在脂质体中。活性免疫

原性成分通常与药学可接受的并且与所述活性成分相容的载体，媒介物或赋形剂混合。合适的赋形剂是，例如，水，盐水，缓冲盐水，右旋糖，甘油，乙醇，无菌等渗含水缓冲液等及其组合。此外，如果希望的话，所述疫苗制剂还可以包括微量的辅助物质例如湿润或乳化试剂，pH缓冲试剂，和/或增强疫苗有效性的佐剂。

可能有效的佐剂的实例包括但不限于：氢氧化铝，N-乙酰基-胞壁酰基-L-苏氨酸-D-异谷氨酸酰胺 (thr-MDP)，N-乙酰基-正-胞壁酰基-L-丙氨酸-D-异谷氨酸酰胺，N-乙酰基胞壁酰基-L-丙氨酸-D-异谷氨酸酰胺-L-丙氨酸-2-(1'-2'-二棕榈酰-sn-甘油基-3-羟基磷酸基氧)-乙胺。

佐剂的有效性可以通过测定抗所注射的用所述特定佐剂配制的修饰的植物的提取物的抗独特型抗体的诱导来确定。

所述组合物可以是液体溶液，悬液，乳剂，片剂，丸剂，胶囊，持续释放制剂，或粉末。口服制剂可以包括标准的载体例如药物级的甘露醇，乳糖，淀粉，硬脂酸镁，糖精钠，纤维素，碳酸镁。

在另一个实施方案中，所述组合物可以置于囊泡，特别是脂质体（参见 Langer, 1990, Science 249: 1527-1533; Treat et al., in Liposomes in the Therapy of Infectious Disease and Cancer, Lopez-Berestein and Fidler (eds.), Liss, New York, pp. 353-365 (1989); Lopez-Berestein, ibid., pp. 317-327; 一般性地参见上文）中递送。

通常，所述成分以单位剂量形式分离地或混合在一起提供，例如，作为置于密封容器例如指示活性剂数量的安瓿瓶或药囊中的干燥的冻干粉或无水浓缩物。如果所述组合物通过注射给药，可以提供无菌稀释剂的安瓿瓶以使所述成分可以在给药之前混合。

在特定的实施方案中，本发明的冻干的修饰的免疫球蛋白提供于第一个容器中；第二个容器含有由含有 50%甘油，0.25%苯酚和防腐剂（例如 0.005%亮绿）的水溶液组成的稀释剂。

本发明还提供了含有装有本发明的疫苗制剂的一种或多种成分的一个或多个容器的药物包装或试剂盒。与这些容器一起的可以由管理药物或生物产品的制造，使用或销售的政府机构所规定的形式的通知，该通知表明该机构对用于人给药的制造，使用或销售的批准。

所述组合物可以，如果希望的话，以包装或分配装置的形式提供，

其可以包含一种或多种含有所述活性成分的单位剂量形式。所述包装可以例如含有金属或塑料箔，例如泡罩包装或透皮贴片。所述包装或分配装置可以附加有给药的说明书。含有用相容的药学载体配制的本发明的提取物的组合物还可以被制备，置于合适的容器中，并标明用于治疗指定的病况。

被给予疫苗的受试者优选是哺乳动物，最优选是人，但是还可以是非人动物，包括但不限于兔，山羊，牛，猪，绵羊，马，麝猫，啮齿动物，浣熊，貉，狗，雪貂，鼬獾 (ferret Badger)，猫，和禽类。

许多方法可用于输入本发明的疫苗制剂；包括但不限于口服，脑内，皮内，肌肉内，腹膜内，静脉内，皮下，鼻内方式，以及通过划破（通过皮肤表层的划伤，例如用分叉的针）或免疫的任何其它标准方法。在特定的实施方案中，划破 (sacrification) 被使用。

在所述制剂中使用的修饰的植物的提取物的精确剂量还取决于给药的方式，和患者的状况，应根据专业医生的判断和每个患者的情况依据标准临床技术确定。有效的免疫量是足以在被给予所述疫苗制备物的宿主中产生对 SARS-CoV 病毒表位的免疫反应（即抗独特型反应）的量。有效剂量还可以从由动物模型检测系统得到的剂量-反应曲线外推。

本发明提供了抗 SARS 的疫苗接种的方法。所述方法包括给受试者施用整个修饰的植物或植物部分，提取物，以产生抗 SARS-CoV 病毒抗原的抗体。在优选的实施方案中，所述疫苗通过摄食修饰的植物或注射修饰的植物的提取物给药。在特定的实施方案中，含有用于表达 S 抗原的载体转化的修饰的植物的提取物的第一种疫苗可以在含有用于表达 M 抗原的载体转化的修饰的植物的提取物的第二种疫苗给药之前，同时，或之后给药。术语“同时”不限于预防或治疗组合物恰好在同一时间给药，还意味着本发明的组合物和其它试剂以一定顺序或在一定时间间隔内被施用给哺乳动物，以使含有多核苷酸的所述组合物可以与其它组合物一起作用以提供比它们用别的方式给药更多的益处。在不同的实施方案中，所述预防或治疗组合物相隔少于 1 小时，相隔大约 1 小时，相隔大约 1 小时到大约 2 小时，相隔大约 2 小时到大约 3 小时，相隔大约 3 小时到大约 4 小时，相隔大约 4 小时到大约 5 小时，相隔大约 5 小时到大约 6 小时，相隔大约 6 小时到

大约 7 小时，相隔大约 7 小时到大约 8 小时，相隔大约 8 小时到大约 9 小时，相隔大约 9 小时到大约 10 小时，相隔大约 10 小时到大约 11 小时，相隔大约 11 小时到大约 12 小时，相隔不超过 24 小时，或相隔不超过 48 小时给药。在优选的实施方案中，两个或多个组分在同一次患者就医期间给药。

在其它实施方案中，所述预防或治疗组合物在大约 30 分钟，相隔大约 1 小时，相隔大约 1 小时到大约 2 小时，相隔大约 2 小时到大约 3 小时，相隔大约 3 小时到大约 4 小时，相隔大约 4 小时到大约 5 小时，相隔大约 5 小时到大约 6 小时，相隔大约 6 小时到大约 7 小时，相隔大约 7 小时到大约 8 小时，相隔大约 8 小时到大约 9 小时，相隔大约 9 小时到大约 10 小时，相隔大约 10 小时到大约 11 小时，相隔大约 11 小时到大约 12 小时，相隔大约 1 到 2 天，相隔大约 2 到 4 天，相隔大约 4 到 6 天，相隔大约 1 周，相隔大约 1 到 2 周，或相隔超过 2 周给药。本领域技术人员能够通过确定所给药组合物的半衰期从而确定该时间框架。

在特定的实施方案中，本发明的预防或治疗组合物周期性地给受试者施用。周期性治疗包括施用第一种组合物一段时间，然后施用第二种组合物和/或第三种组合物一段时间，并重复这种顺序给药。周期性治疗可以减少对这些治疗的一种或多种的抗性的发生，避免或降低这些治疗之一的副作用，和/或提高治疗的功效。

在特定的实施方案中，预防或治疗组合物以少于大约 3 周的周期给药，大约每两周一次，大约每 10 天一次或大约每周一次。一个周期可以包含通过每个周期超过大约 90 分钟，每个周期超过大约 1 小时，每个周期超过大约 45 分钟的输注进行的预防或治疗组合物的给药。每个周期可以含有至少 1 周的休息，至少 2 周的休息，至少 3 周的休息。给药周期的数目是大约 1 到大约 12 个周期，更典型地是大约 2 到大约 10 个周期，更典型地是大约 2 到大约 8 个周期。

在另一个实施方案中，本发明的预防或治疗组合物以规律定量方案给药，其是通过连续的输注或没有延长的休息期的频繁给药进行。这种规律给药可以包括以没有休息期的以恒定的间隔给药。所述定量方案包括以相对低的剂量长期每天给药达一长段时间。在优选的实施方案中，使用更低剂量可以将毒副作用减至最小，并且可以去掉休息



期。在特定的实施方案中，所述预防或治疗组合物以长期低剂量或连续输注大约 24 小时到大约 2 天，到大约 1 周，到大约 2 周，到大约 3 周，到大约 1 个月，到大约 2 个月，到大约 3 个月，到大约 4 个月，到大约 5 个月，到大约 6 个月递送。这种剂量方案的时间安排可以由有技术的医生最优化。

本文提供的给药的剂量和频率包括在术语治疗有效和预防有效中。所述剂量和频率进一步典型地根据每个患者的特定因素而变化，这取决于要施用的特定治疗或预防组合物，疾病或失调的严重性和类型，给药的方式，以及患者的年龄，体重，反应，和过去的病史。合适的方案可由本领域技术人员参考这些因素并依据，例如文献中报道的和 *Physician's Desk Reference* (56<sup>th</sup> ed., 2002) 中推荐的剂量进行选择。

各种递送系统都是已知的，可用于施用本发明的治疗或预防组合物，例如包裹于脂质体中，微粒，微胶囊，能表达所述抗体或抗体片段的重组细胞，受体介导的胞吞作用(参见，例如 Wu and Wu, *J. Biol. Chem.* 262: 4429-4432 (1987))，将核酸构建为逆转录病毒或其它载体的一部分等。施用本发明的治疗或预防组合物的方法包括但不限于，肠胃外给药(例如皮内，肌肉内，腹膜内，静脉内和皮下)，硬膜外，和粘膜(例如鼻内和口服途径)。在特定的实施方案中，本发明的治疗或预防组合物进行肌肉内，静脉内，或皮下给药。所述治疗或预防组合物可以通过任何方便的方式给药，例如通过输注或弹丸注射，通过上皮或粘膜与皮肤的内层(linings)(例如口腔粘膜，直肠和肠粘膜等)的吸收，可以与其它生物活性试剂一起给药。给药可以是系统性的或是局部的。

在特定的实施方案中，可能希望对需要治疗的区域局部给予本发明的治疗或预防组合物；这可以通过，例如，但不限于局部输注，注射，或通过植入物实现，所述植入物是多孔，非多孔，或凝胶状材料的，包括膜，例如硅橡胶(sialastic)膜，或纤维。

在另一个实施方案中，所述治疗或预防组合物可以用控制释放或持续释放系统递送。在一个实施方案中，泵可用于获得控制或持续释放(参见 Langer, *supra*; Sefton, 1987, *CRC Crit. Ref, Biomed. Eng.* 14: 20; Buchwald et al., 1980, *Surgery* 88: 507; Saudek et al., 1989,

N. Engl, J. Med. 321: 574)。在另一个实施方案中，聚合材料可用于获得本发明的治疗或预防组合物的控制或持续释放（参见，例如，*Medical Applications of Controlled Release*, Langer and Wise (eds.), CRC Pres., Boca Raton, Florida (1974); *Controlled Drug Bioavailability, Drug Product Design and Performance*, Smolen and Ball (eds.), Wiley, New York (1984); Ranger and Peppas, 1983, *J. Macromol. Sci. Rev. Macromol. Chem.* 23: 61; 还参见 Levy et al., 1985, *Science* 228: 190; During et al., 1989, *Ann. Neurol.* 25: 351; Howard et al., 1989, *J. Neurosurg.* 71: 105); U. S. Patent No. 5,679,377; U.S. Patent No. 5,916,597; U.S. Patent No. 5,912,015; U. S. Patent No. 5,989,463; U. S. Patent No. 5,128,326; PCT Publication No. WO99/15154; and PCT Publication No. WO99/20253)。用于持续释放制剂的聚合物的实例包括，但不限于聚(2-甲基丙烯酸羟乙酯)，聚(甲基丙烯酸甲酯)，聚(丙烯酸)，聚(醋酸亚乙基-共聚-乙烯酯)，聚(甲基丙烯酸)，聚乙醇酸(PLG)，聚酞，聚(N-乙基吡咯烷酮)，聚(乙烯醇)，聚丙烯酰胺，聚(乙二醇)，聚交酯(PLA)，聚(丙交酯-共聚-乙醇酸)(PLGA)，和聚原酸酯(polyorthoesters)。在优选的实施方案中，用于持续释放制剂的聚合物是惰性的，不含可滤出杂质的，储存稳定的，无菌的，和生物可降解的。在另一个实施方案中，控制或持续释放系统可以置于预防或治疗的靶的附近，因此仅需要系统性剂量的一部分(参见，例如 Goodson, in *Medical Applications of Controlled Release*, supra, vol. 2, pp. 115-138 (1984))。

控制释放系统在 Langer (1990, *Science* 249: 1527-1533) 的综述中有所讨论。本领域技术人员所知的任何技术都可以用于产生含有本发明的一个或多个治疗组合物的持续释放制剂。参见，例如 U. S. Patent No. 4,526,938, PCT publication WO 91/05548, PCT publication WO96/20698, . Ning et al., 1996, "Intratumoral Radioimmunotherapy of a Human Colon Cancer Xenograft Using a Sustained-Release Gel, " *Radiotherapy & Oncology* 39: 179-189, Song et al., 1995, "Antibody Mediated Lung Targeting of Long-Circulating Emulsions, " *PDA Journal of Pharmaceutical Science & Technology* 50: 372-397, Cleek et al., 1997, "Biodegradable Polymeric Carriers for a bFGF Antibody for

Cardiovascular Application, " Pro. Int'l. Symp. Control. Rel. Bioact. Mater. 24: 853-854, and Lam et al., 1997, "Microencapsulation of Recombinant Humanized Monoclonal Antibody for Local Delivery, " Proc. Int'l. Symp. Control Rel. Bioact. Mater. 24: 759-760, 在此引入其中每一个的全文作为参考。

抗体可以在施用本发明的疫苗之后从受试者分离。产生抗 SARS 的抗体的方法包括给受试者施用有效量的修饰的植物, 和过一段时间以后从受试者的样品中分离抗体。所述样品可以是生物流体, 例如血液, 血清, 血浆, 唾液, 尿, 粪便, 痰, 鼻咽吸出物, 细胞和组织。

### 预防和治疗效用的证明

本发明提供了通过给有需要的受试者施用本发明的修饰的植物, 从而预防, 治疗, 改善, 和/或控制 SARS 的方法。用于本发明的方法的修饰的植物可以通过多种方式筛选或检测治疗或预防 SARS 的功効。

首先, 含有本发明的修饰的植物或植物提取物的疫苗制剂的免疫功効可以通过在用所述疫苗免疫后监测所检测动物的免疫反应确定。体液反应的产生可以作为一般免疫反应的指示, 其其它组成成分, 特别是细胞介导的免疫, 可能对于提供对疾病的保护是重要的。所检测动物可以包括小鼠, 兔, 黑猩猩, 最后是人受试者。本发明中制备的疫苗可以在实验中感染黑猩猩。但是, 由于黑猩猩是受保护的生物种, 对本发明的疫苗的抗体反应可以首先在许多更小的, 较便宜的动物中研究, 其目的是为了找到产生 SARS-CoV 病毒表位的最佳组合的一种或两种最好的候选的修饰的植物以用于黑猩猩的功効研究。

所述疫苗的安全性还可以通过在给所检测的受试者施用疫苗后观察或监测受试者的症状来确定。在特定的实施方案中, SARS-CoV 的全基因组用于疫苗中。在优选的实施方案中, 所述植物表达 SARS-CoV 多肽的少于 10%, 少于 20%, 少于 25%, 少于 30%, 少于 35%, 少于 40%, 少于 45%, 少于 50%, 少于 55%, 少于 60%, 少于 65%, 少于 70%, 少于 80%, 少于 90%。

所检测受试者的免疫反应可以通过各种方法进行分析, 例如用已知技术, 例如酶联免疫吸附检测 (ELISA), 免疫印迹, 放射性免疫

沉淀等检测得到的免疫血清对抗原的反应性；或在免疫的宿主中防止感染和/或减少疾病症状。

作为适当的动物检测的一个实例，本发明的疫苗组合物可以在兔中检测其诱导对 SARS 病毒抗原的免疫反应的能力。例如，可以使用雄性的无特定病原体（SPF）的年轻成年新西兰白兔。每个检测组的兔子通过任何给药方法，例如摄食修饰的植物或注射修饰的植物提取物，接受有效量的所述疫苗。对照组的兔注射含于 1 mM Tris-HCl pH 9.0 中的含有来自相同类型的未修饰的植物的相同种类的植物提取物的疫苗。每一到两周从兔中提取血液样品，用例如放射性免疫检测（Abbott Laboratories）分析血清中对所述病毒抗原特异的抗体。抗体的存在可以用 ELISA 检测。兔因为其远系繁殖的性质可能会有可变的反应，因此在小鼠中检测所述疫苗可能是有用的。

在一个实施方案中，本发明涉及从通过注射或摄食被施用了修饰的植物和/或植物提取物的受试者中收集的生物材料。所述生物材料可以检测 SARS-CoV 多肽或其片段的存在。生物材料包括，但不限于，血液，血清，血浆，唾液，尿，粪便，痰，鼻咽吸出物，细胞和组织。所述生物材料可以在施用本发明的修饰的植物和/或植物提取物之后 1 小时，2 小时，6 小时，12 小时，24 小时，2 天，3 天，或 1 周后收集。

此外，本发明的修饰的植物可以通过首先给检测受试者，其为动物或人，施用所述修饰的植物，然后分离针对所述病毒抗原的抗独特型反应中产生的抗-抗独特型抗体（即 Ab3 抗体）来检测。分离的 Ab3 然后可以检测其结合特定病毒抗原的能力（例如通过本领域已知的任何免疫检测，例如但不限于放射性免疫检测，ELISA，“sandwich”免疫检测，凝胶扩散沉淀反应，免疫扩散检测，western blots，沉淀反应，凝集检测，补体固化检测，免疫荧光检测，蛋白 A 检测，免疫电泳检测等）。

在一个方面，如果所述疫苗是抗 SARS-CoV 病毒抗原的，分离的 Ab3 治疗 SARS 的功效通过培养来自培养物或患者的 SARS-CoV 感染的细胞，使所述细胞与要检测的 Ab3 抗体接触，将接触细胞的增殖或存活与未接触所述 Ab3 抗体的细胞的增殖或存活相比较来筛选，其中接触细胞的更低水平的增殖或存活表示所述 Ab3 抗体（用 SARS-CoV

病毒抗原进行免疫引发的)在患者中有效治疗 SARS。本领域中许多标准检测可用于评价这种存活和/或生长;例如,细胞增殖可以通过测定  $^3\text{H}$ -胸腺嘧啶掺入,通过直接细胞计数,通过检测已知基因例如原癌基因(例如 *fos*, *myc*)或细胞周期标记的转录活性的改变来检测;细胞的生存力可以通过台盼蓝染色评价,分化可以依据形态变化进行视觉评价。

本发明还提供了在给有需要的受试者施用本发明的修饰的植物后得到的抗 SARS-CoV 病毒的抗体。免疫特异性结合 SARS-CoV 病毒的抗原的抗体还可以直接在体内检测。为监测本发明的抗体的效果,在施用所述疫苗之前,期间,或之后在适当的时间间隔测定所述抗原的水平。所述抗原的量的任何变化或没有变化可以被鉴别,并与对受试者的治疗效果相关联。

特别地,在 SARS 的情形中,SARS-CoV 抗原的血清水平与 SARS 的严重程度有直接关系。通常,抗原水平的降低与有效的治疗相关。所治疗的受试者中 SARS-CoV 抗原的血清水平在给所述受试者施用本发明的修饰的植物和/或植物提取物之后每 1 小时,2 小时,3 小时,6 小时,12 小时,24 小时,2 天,3 天,或 1 周被检测。

在优选的方面,所述方法可以被实施以确定不同时间点的抗原水平,并将这些值与基线水平相比较。基线水平可以是在正常的无疾病的个体中存在的标记的水平;和/或在治疗前或稳定期间的水平。这些水平然后可以和疾病进程或治疗结果相关联。

抗原的水平可以用本领域公知的任何方法测定。例如 SARS-CoV 病毒抗原可以用已知的免疫诊断方法例如 western blotting,免疫沉淀,并使用抗 SARS-CoV 抗原的任何抗体来定量。

体内对所述病毒抗原的免疫反应的强度可以用本领域已知的任何方法测定,例如但不限于延迟的超敏皮肤检测和体外细胞溶解 T-淋巴细胞活性的检测。

延迟的超敏皮肤检测在整体免疫反应和对抗原的细胞免疫的检测中具有重大价值。皮肤检测的适当技术要求所述抗原无菌储存于  $4^{\circ}\text{C}$ ,防光并在临使用前重构。25-或 27-规格的针保证了抗原的皮内而不是皮下给药。抗原皮内给药后二十四小时和 48 小时后,用标尺检测最大尺寸的红斑和硬化。对任何给定抗原或抗原组的活动减退

(hypoactivity) 通过用较高浓度的抗原检测或者, 在不确定的情况下, 通过用中间检测进行重复检测来证实。

为了检测细胞溶解性 T-淋巴细胞的活性, 从经免疫的受试者, 例如通过 Ficoll-Hypaque 梯度离心技术分离出的 T-淋巴细胞, 用具有抗 SARS 抗原的细胞在含有 10% 胎牛血清的 3 ml RPMI 介质中再次刺激。在一些实验中, 在培养介质中包含 33% 的二级混合的淋巴细胞培养物上清或 IL-2 作为 T 细胞生长因子的来源。为了测定免疫后细胞溶解性 T-淋巴细胞的初级反应, 分离的 T 细胞在有或没有具有所述抗原的细胞的条件下培养。六天后, 所述培养物用 4 小时  $^{51}\text{Cr}$ -释放检测测定其细胞毒性。如果免疫是有效的, 靶的自发的  $^{51}\text{Cr}$ -释放应当达到少于 20% 的水平 (Heike et al., J. Immunotherapy Vol. 15, pages 15-174)。

所述抗体的功效可以通过给受试者 (人受试者或者所述疾病的动物模型) 施用所述抗体, 然后监测 SARS-CoV 病毒抗原的水平或特定感染疾病的症状来检测。SARS-CoV 病毒抗原的水平可以用本领域已知的用于检测 SARS-CoV 病毒抗原的水平的方法测定, 例如病毒滴度, 在病毒的情况下, 或者细菌水平 (例如通过培养来自患者的样品) 等。病毒抗原水平的降低或 SARS 症状数目的消除, 改善或减少表示所述抗体是有效的。SARS 的症状包括, 但不限于, 温度高于  $100.4^{\circ}\text{F}$  ( $> 38^{\circ}\text{C}$ ) (根据 CDC), 寒战, 头痛, 不适, 身体疼痛, 咳嗽, 呼吸急促, 呼吸困难, 和缺氧。肺炎的射线照相证据, 呼吸窘迫综合症, 和尸检结果也用于确定受试者是否患有, 从其康复, 或没有 SARS。

### 预防和治疗效用的功效

本发明的预防和/或治疗方案的毒性和功效可以在细胞培养物或实验动物中用标准药理学方法测定, 例如测定  $\text{LD}_{50}$  (对群体的 50% 致命的剂量) 和  $\text{ED}_{50}$  (在群体的 50% 中治疗有效的剂量)。毒性和治疗效果的剂量比是治疗指数, 可以表示为  $\text{LD}_{50}/\text{ED}_{50}$  的比值。表现出大治疗指数的预防和/或治疗试剂是优选的。虽然表现出毒副作用的预防和/或治疗试剂可以使用, 但应当注意要设计将这些试剂靶定到受影响组织部位的递送系统以使对于未感染细胞的损害减至最小, 从而减少副

作用。

从细胞培养物检测和动物研究获得的数据可以用于配制一定剂量范围内的用于人的预防和/或治疗试剂。这种试剂的剂量优选在包括 ED<sub>50</sub> 并且几乎没有或没有毒性在内的循环浓度范围内。所述剂量可以在这个范围内变化，取决于所使用的剂型和所使用的给药方式。对于本发明的方法中使用的任何试剂，治疗有效剂量最初可以从细胞培养物检测来估计。剂量可以在动物模型中制定以获得包括在细胞培养物中测定的 IC<sub>50</sub>（即实现症状的半数最大抑制的待测化合物浓度）在内的循环血浆浓度范围。这种信息可以用于更精确地确定在人中可用的剂量。血浆中的水平可以，例如通过高效液相色谱检测。

### 实施例

下述实施例举例说明了含有 SARS-CoV 病毒的 M 和/或 S1 蛋白的转基因或转质体烟草植物的产生。

### 质粒构建

用于构建这些质体转化载体的骨架是衍生自质粒 pVSR326 (Reddy et al., 2002, Mol Breeding 9: 259-269; PCT international application no. PCT/EP00/12446, international publication no. WO 01/42441)。质粒 pVSR326 使用跨越 *rbcL-accD* 的烟草质体基因组序列，以便将编码β-葡糖醛酸酶 (*GUS*) 的报道基因通过同源重组靶向到叶绿体基因组中 (Reddy et al., 2002)。 *GUS* 的启动子和终止子衍生自稻质体基因 *psbA*; *psbA* 基因编码光系统 II 的 32kD 蛋白。选择标记基因 *aadA*，具有壮观霉素抗性，由稻质体 16S rRNA 操纵子 (*rrn*) 启动子驱动，*aadA* 与 *GUS* 相邻 (Reddy et al., 2002)。在 pMLVHis 载体中，除了在此区域上的独特的限制性位点，引入起始密码子和 (His)<sub>5</sub>-tag 以便于 (His)<sub>5</sub>-标记的重组蛋白的克隆和表达。

### pCV1, 用于表达 S1 抗原的质体转化载体的构建

来自 pCRII-S1 的编码去掉其起始密码子和信号肽的 S1 的 2-kb SpeI-NotI 片段被克隆到 pMLVHisA (图 4) 的 NheI 和 NotI 位点上。质粒 pCRII-S1 衍生自 pCRII (Invitrogen)，含有 S1 的 PCR 扩增片

段。在得到的称为 pCV1 (图 1) 的衍生物中, S1 的氨基酸 12-658 与十七个 pMLVHisA 衍生的残基, 包括‘ATG’起始密码子和(His)<sub>5</sub>-tag, 框内 (in-frame) 融合。质粒 pMLVHisA 设计有(His)<sub>5</sub>-tag, 使得能在 western blot 分析中使用抗(His)<sub>5</sub>-tag 的抗血清识别重组蛋白。而且, (His)<sub>5</sub>-tag 使得, 如果需要的话并且当需要的时候, 可以用 Ni-NTA Agarose (Qiagen) 亲和柱容易地纯化重组蛋白。稻质体 *psbA* 启动子驱动所述重组蛋白的表达, 终止子也是衍生自 *psbA*。烟草质体基因组上侧翼 *rbcL* 和 *accD* 序列的存在使得发生同源重组, 结果是将 S1 (和 *aadA*) 插入到靶植物的质体基因组中。

由于植物叶绿体基因组序列是高度保守的, 相同的构建体用于烟草和番茄, 二者都属于同一个家族茄科, 的质体转化。

#### pCV2, 用于表达 S1 抗原的核转化载体的构建

来自 pCRII-S1 的编码氨基酸 1-658 的 S1 的 2-kb 的 BamHI-XhoI 片段被克隆到 pSa7 (图 5), 一种植物核转化载体的 BamHI 和 Sall 位点上。在该情形中, S1 的信号肽 (氨基酸 1-13) 被保留, 因为以前的报道提示在植物核-表达的外源蛋白上存在内质网靶定信号会提高蛋白质的稳定性 (Richter et al., 2000, Nature Biotech 18: 1167-1171; Sojikul et al., 2003, Proc. Natl. Acad. Sci. 100: 2209-2214)。在得到的质粒, pCV2 (图 19) 中, S1 位于强的和组成型花椰菜花叶病毒 35S 启动子和胭脂碱合酶 (NOS) 终止子之间。该质粒在存在 *E.coli* 帮助株 HB101/pRK2013 的条件下通过三亲本杂交从大肠杆菌 (*Escherichia coli*) 引入到根癌农杆菌 (*Agrobacterium tumefaciens*) 株 LBA4404 中 (Horsch et al., 1985, Science 227: 1227-1231)。

#### pCV8, 用于表达 M 抗原的质体转化载体的构建

来自于质粒 pCRII-M 的编码 M 的 0.7-kb 的 NcoI-SacI 片段被克隆到 pMLVHisA (图 4) 的 NcoI 和 SacI 位点上生成 pCV7。质粒 pCRII-M 衍生自 pCRII (Invitrogen), 含有 M 的 PCR 扩增片段。然后, 质粒 pCV7 用 NcoI 消化, 用退火的寡聚物 ML527 (5'-CATGGCCGCGCGGGGTCTCATCATCATCATCATGG-3'; SEQ ID NO:10) 和 ML528 (5'-CATGCCATGATGATGATGATGATGATGAGAACCCCGCGCGGC-3'; SEQ ID



NO: 11) 将(His)<sub>6</sub>-tag 引入到这个位点, 产生 pCV8 (图 7)。得到的从 pCV8 表达的 M 是(His)<sub>6</sub>-标记的蛋白。

#### pCV4, 用于表达 M 抗原的核转化载体的构建

来自质粒 pCRII-M 的编码 M 的 0.7-kb 的 BamHI 片段被克隆到植物核转化载体 pSa13 的 BamHI 位点上。质粒 pSa13 衍生自质粒 pSa7 (图 5), 其是通过去除 SaPIN2a cDNA 的 0.58-kb 的 Sall 片段以后再连接而得到的。

#### pCV6,, 用于共表达 S1 和 M 抗原的质体转化载体的构建

来自于质粒 pCV3 的, 由与 psbA 终止子和 rrn 启动子融合的 M-基因, 以这个顺序构成的 1.2-kb 的部分 BamHI 片段被克隆到质粒 pCV1 的 BamHI 唯一位点上, 产生共表达 S1 和 M 抗原的质体转化载体 pCV6(图 21)。质粒 pCV3 是通过将来自 pCRII-M 的编码 M 的 0.7-kb 的 EcoRV-SacI 片段克隆到 pMLVHisA 的 StuI-SacI 位点上获得的。

#### 通过定点突变使 SARS-CoV S1 基因的密码子使用最优化

质粒 pCRII-S1 用作产生密码子使用最优化的衍生生物的突变模板。突变是通过 PCR 使用 PfuTurbo DNA 聚合酶使用 ‘QuikChange Multi site-directed mutagenesis kit’ (Stratagene), 根据制造商的说明书实施的。表 1 显示了用于定点突变的寡核苷酸。扩增程序包括在 95°C 变性 1 分钟, 然后 30 个循环的变性 (95°C 1 分钟), 退火 (55°C 1 分钟) 和延伸 (65°C 12 分钟)。扩增产物用 DpnI 处理以除去模板 DNA, 突变 DNA 用于转化 E. coli XL10-Gold 细胞。进行 DNA 测序分析以证实每个突变。引入到“最优化的” S1 中的总共 13 个核苷酸改变 (如表 1 中总结的) 然后用于质体和核转化载体中的克隆。

表 1. 为使密码子最优化用于 S1 的定点突变的寡核苷酸

影响的 残基	引物序列
<u>正向引物:</u>	
R18	5'-GGTAGTGACCTTGAC <u>AG</u> ATGCACCACTTTTGAT-3'; SEQ ID NO:12
T75	5'-GGGTTTCATACTATTAATCATA <u>CT</u> TTTGGCAACCCTGTCATAC-3'; SEQ ID NO:13
S113	5'-CCATGAACAACAAGTCACAG <u>TC</u> TGTGATTATTATTAACAATTCTACT-3'; SEQ ID NO:14
S169	5'-AGTACATATCTGATGCCTTT <u>CT</u> CTTGATGTTTCAGAAAAGTC-3'; SEQ ID NO:15
L209	5'-CCTATAGATGTAGTTCGTGAT <u>CT</u> CTCTTGTTTAAACACTTTG-3'; SEQ ID NO:16
T247	5'-CAAGACATTTGGGGC <u>ACT</u> TCAGCTGCAGCCTAT-3'; SEQ ID NO:17
A398	5'-GATGATGTAAGACAAATAG <u>CT</u> CCAGGACAAACTGG-3'; SEQ ID NO:18
P507	5'-TCTTTTGAAC TTTTAAATGCAC <u>CT</u> GCCACGGTTTGTGGACC-3'; SEQ ID NO:19
T509	5'-CTTTTAAATGCACCTGCC <u>ACT</u> GTTTGTGGACCAAATTATC-3'; SEQ ID NO:20
L597	5'-CTTCATCTGAAGTTGCTGTT <u>CT</u> TATCAAGATGTAACTGCAC-3'; SEQ ID NO:21
R620	5'-CAACTCACACCAGCTTGG <u>AGA</u> ATATATTCTACTGGAAACAATG-3'; SEQ ID NO:22
<u>反向引物:</u>	
R18	5'-ATCAAAAGTGGTGCA <u>TCT</u> GTCAAGGTCCTACC-3'; SEQ ID NO:23
R620	5'-CATTGTTCCAGTAGAATATAT <u>CT</u> CCAAGCTGGTGTGAGTTG-3'; SEQ ID NO:24

\*斜体的核苷酸是突变的。所影响的密码子加下划线。

## 植物转化

### 用于表达 S1 和/或 M 抗原的烟草和番茄的质体转化

每个质体转化构建体 (pCV1, pCV6 或 pCV8, 和其含有具有核苷酸 (而不是氨基酸) 变化以使密码子使用最适于在植物中表达的 S1 的衍生物) 通过基因枪轰击被引入到烟草 (Staub and Maliga, 1994, Plant Journal 6: 547-553) 和番茄 (Ruf et al., 2001, Nature Biotech

19: 870-875) 中。

### 番茄 (*Lycopersicon esculentum*) cv. UC82B 的质体转化

质体转化根据 Ruf et al., 2001 进行, 用所希望的 DNA (pCV1, pCV6 或 pCV8, 和其含有具有核苷酸 (而不是氨基酸) 变化以使密码子使用最适于在植物中表达的 S1 的衍生物) 包覆的钨颗粒轰击嫩叶。

无菌种子在添加了琼脂 (5g/l) 和蔗糖 (30 g/l) 的 Murashige and Skoog (MS) 培养基 (Murashige and Skoog, 1962, *Physiol. Plant.* 15: 473-497) 上发芽。为进行轰击, 长出的腋生分生组织的嫩叶被置于含有 MS 盐, 6-苯甲基氨基嘌呤 (1mg/l),  $\alpha$ -萘乙酸 (0.1 mg/l), 硫胺 (1mg/l), 肌型肌醇 (100 mg/l), 琼脂 (5g/l), pH 5.7 和蔗糖 (30g/l) 的 RMOP 培养基上离轴侧 (abaxial side) 面朝上。轰击用微粒递送系统, PDS 1000-He (Bio-Rad) 和其配件进行。钨颗粒 (M 17) 用所希望的 DNA 包裹, 叶片被轰击。轰击之后, 叶片被切割成小片, 置于含有壮观霉素二盐酸盐 (250 mg/l) 的 RMOP 选择培养基上。黄绿色的愈伤组织在选择培养基上经过少量循环获得含有 homoplastomic 质体的愈伤组织。从该愈伤组织将获得植物的再生。

### 烟草 (*N. tabacum*) cv. xanthi 的质体转化

质体转化根据 Svab and Maliga (Svab et al., 1993, *Proc. Natl. Acad. Acad. Sci. USA* 90: 913-917) 进行, 用所希望的 DNA (pCV1, pCV6 或 pCV8, 和其含有具有核苷酸 (而不是氨基酸) 变化以使密码子使用最适于在植物中表达的 S1 的衍生物) 包覆的钨颗粒轰击叶子。

无菌种子在添加了琼脂 (5 g/l) 和蔗糖 (30 g/l) 的 Murashige and Skoog (MS) 培养基 (Murashige and Skoog, 1962, *Physiol. Plant.* 15: 473-497) 上发芽。为进行轰击, 叶片被置于含有 MS 盐, 6-苯甲基氨基嘌呤 (1 mg/l),  $\alpha$ -萘乙酸 (0.1 mg/l), 硫胺 (1 mg/l), 肌型肌醇 (100 mg/l), 琼脂 (5 g/l), pH 5.7 和蔗糖 (30 g/l) 的 RMOP 培养基上离轴侧面朝上。轰击用微粒递送系统, PDS 1000-He (Bio-Rad) 和其配件进行。钨颗粒 (M 17) 用所希望的 DNA 包裹, 叶片被轰击。轰击之后, 叶片被切割成小片, 置于含有壮观霉素二盐酸盐 (500 mg/l)

的 RMOP 选择培养基上。用载体 pCV1 (图 23F-23G) 对烟草进行质体转化得到的再生的幼苗在选择培养基上再经过五个循环传代, 以获得含有 homoplastomic 质体的愈伤组织。显示了用 S1 引物 ML560

(5'-CAGAGAGAGTTTCCCGGATTC-3'; SEQ ID NO:25) 和 ML567 (5'-CAACCTATAGATGTAGTTCG-3'; SEQ ID NO:26)

进行的 PCR 分析 (图 23H) 和用  $^{32}\text{P}$ -放射性标记的 S1 探针进行的 northern blot 分析的结果。图 23J 显示了在检测 His-标记的 S1 蛋白时使用 Ni-NTA 缀合物对表达修饰的 S1 (具有如表 1 所示的核苷酸变化) 的转质体烟草进行的 western blot 分析。箭头指示出所期望的 73-kDa 的 (His)<sub>5</sub>-S1 条带。

### 烟草和莴苣的核转化以表达 S1 或 M 抗原

通过核转化产生表达 S1 或 M 抗原的转基因烟草和转基因莴苣是用农杆菌介导的转化 (Horsch et al., 1985, Science 227: 1227-1231) 以及分别使用质粒 pCV2 [和其含有具有核苷酸 (而不是氨基酸) 变化以使密码子使用最适于在植物中表达的 S1 的衍生物] 和 pCV4 实现的。

在农杆菌介导的转化之后, 烟草 (图 23A) 和莴苣 (图 23C) 的芽生根, 再生的芽 (图 23, B, D) 用于用分别位于 CaMV 35S 启动子和 NOS-终止子内部的引物 35S 和 NOS-ter 进行 PCR 分析。用  $^{32}\text{P}$ -标记的 S1 探针对这些 PCR 扩增的 DNA 进行的 Southern blot 分析如图 23E 所示。2.1-kb 的杂交条带证实了含有 S1 的转基因株系的存在 (图 23E, 2-5 道)。

### 烟草的核转化以表达 S1:GFP 抗原

表达由 SARS-CoV S1 蛋白与绿色荧光蛋白 (GFP) 融合而成的融合蛋白的转基因烟草的产生是通过农杆菌介导的转化 (Horsch et al. supra.) 并用质粒 pCV 12 实现的。

农杆菌渗入 (Agroinfiltration) 后两天, 选择有代表性的烟草叶片表皮细胞, 用共聚焦显微镜观察 (Yang Y. et al., *In vivo analysis of plant promoters and transcription factors by agroinfiltration of tobacco leaves*. Plant J. 2000; 22: 543-551) 含有表达 S1 : GFP 融合蛋白的质粒 pCV12 的根癌农杆菌 LBA4404 (图 18, A, C) 或含有

单独表达 GFP 的 pGDG 的 LBA4404 (图 18, B, D)。横条代表 20 $\mu$ m。

Western blot 分析用抗 GFP 的抗体进行, 显示出在农杆菌渗入之后在烟草叶片中瞬时表达 S1 : GFP。从用单独表达 GFP 的质粒 pGDG (图 23, 第 1 道) 或表达融合的 S1 : GFP 融合蛋白的质粒 pCV12 (图 23, 第 2 道) 渗入的烟草叶片中提取的总蛋白 (200 $\mu$ g) 在 8% 的 SDS-PAGE 凝胶中分离, 根据 Sambrook et al. (1989. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor) 印迹在 Hybond-C 滤膜上, 并与抗 GFP 的抗体 (Clontech) 交叉反应。结果如图 24 所示, 其中箭头指示出交叉反应的 S1 : GFP 条带 (计算大小为 99.1 kDa), M 是分子量标记。

本领域技术人员应当认识到, 或能够通过不超过常规的实验确定, 本文所述的本发明的特定实施方案的许多等同方式。这种等同方式意在包括在下述权利要求的范围内。

本说明书中提及的所有的出版物, 专利和专利申请在此以引用的方式引入到说明书中, 如同每个单独的出版物, 专利或专利申请都特别地和单独地指出以引用的方式引入本文。

本文的参考文献的引证或讨论不应当解释为承认这就是本发明的现有技术。

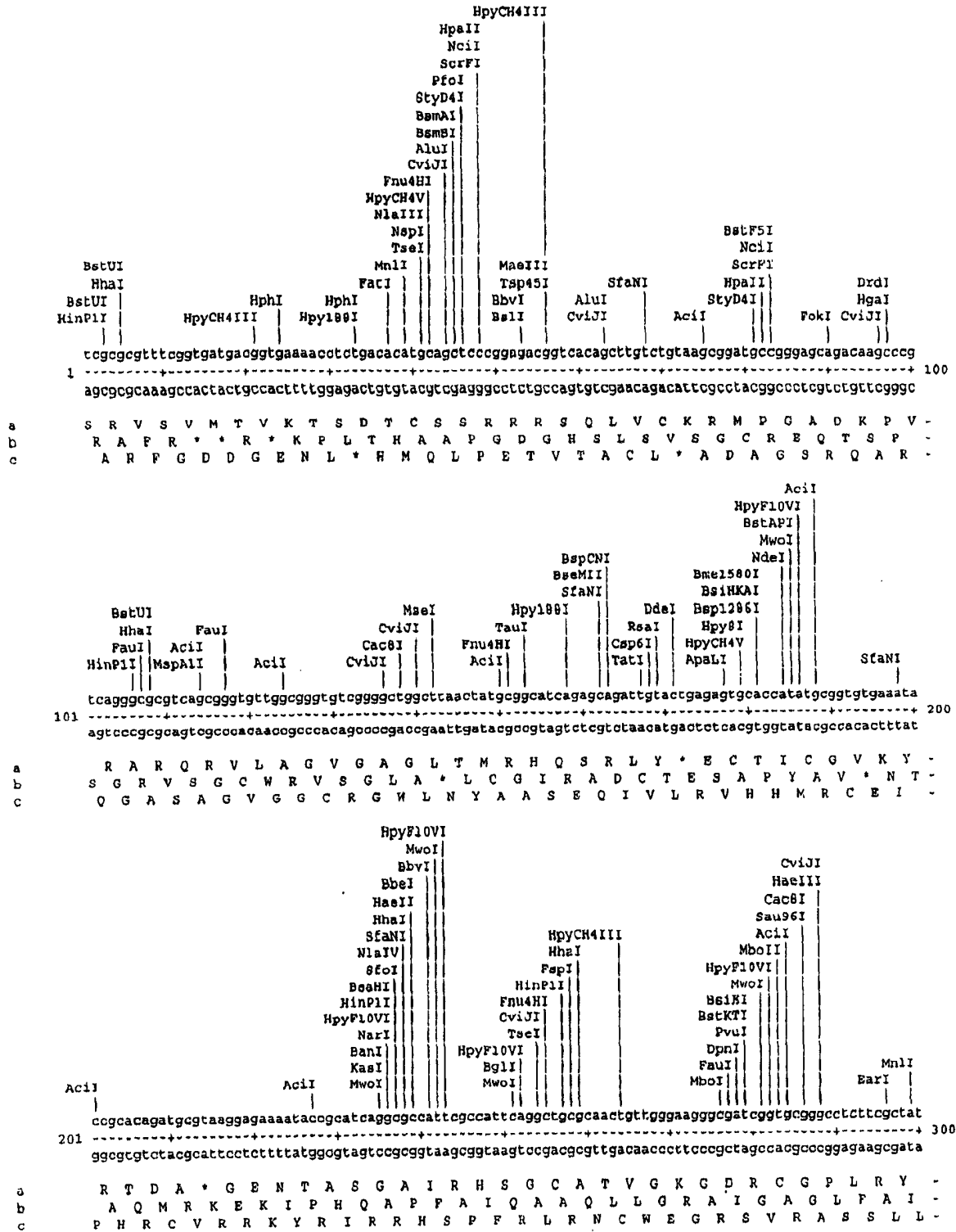


图 1-A

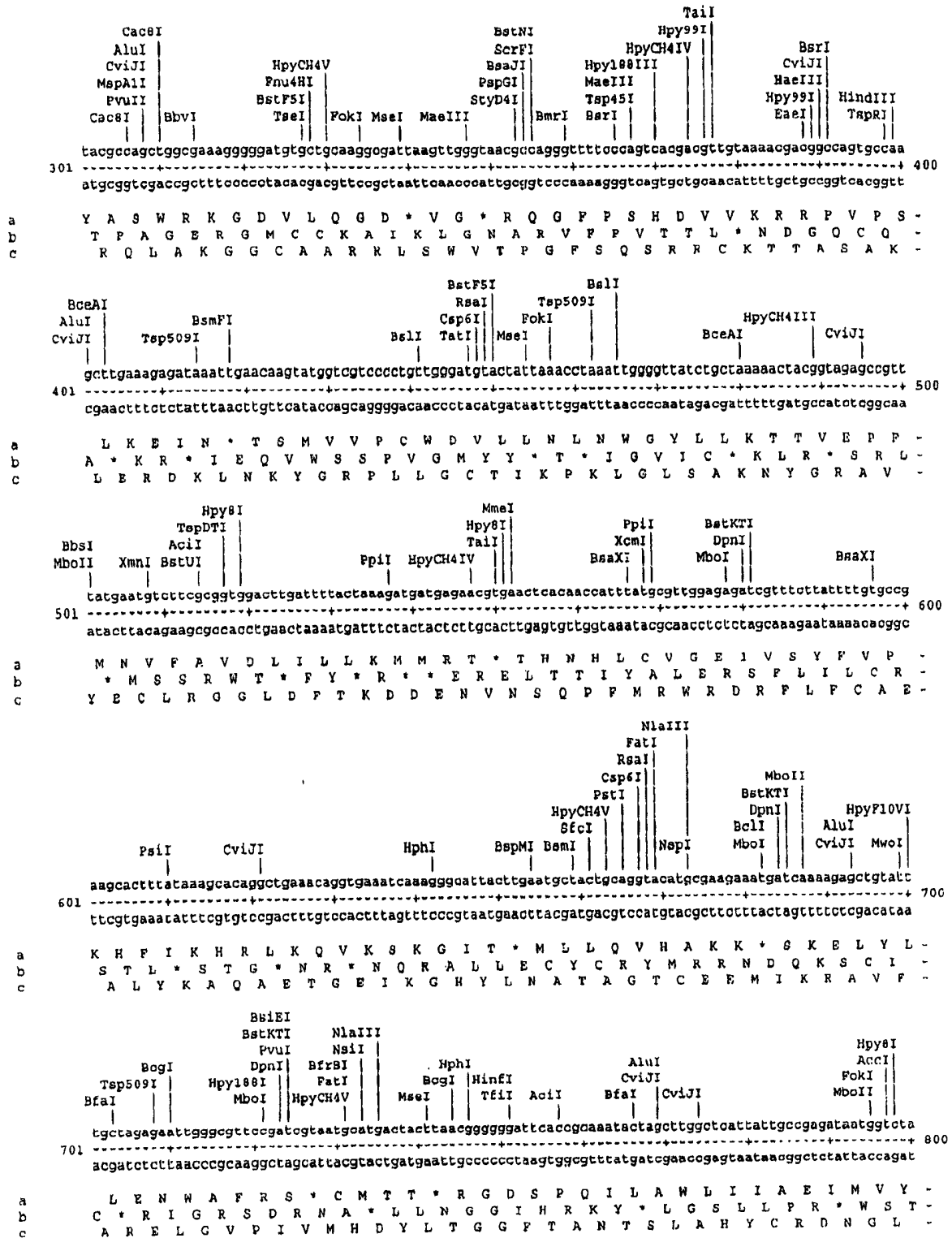


图 1-B

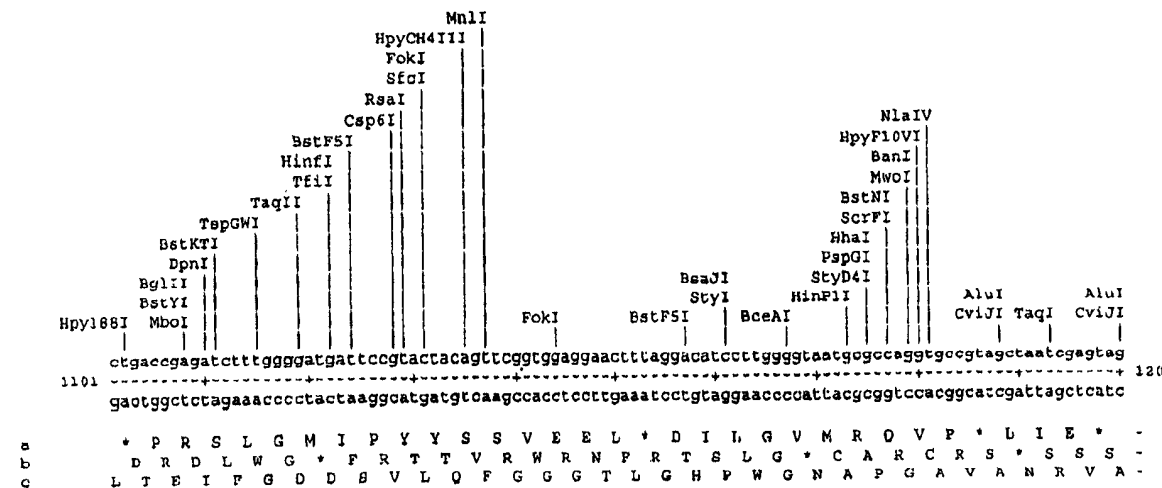
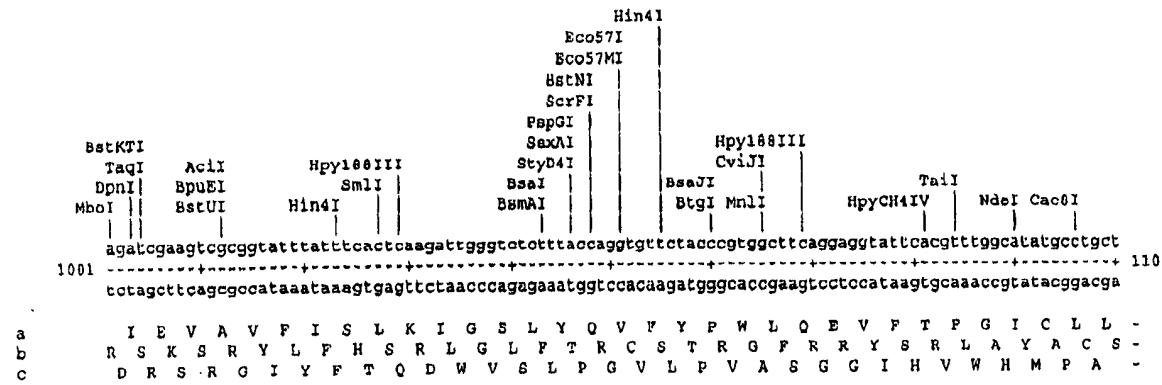
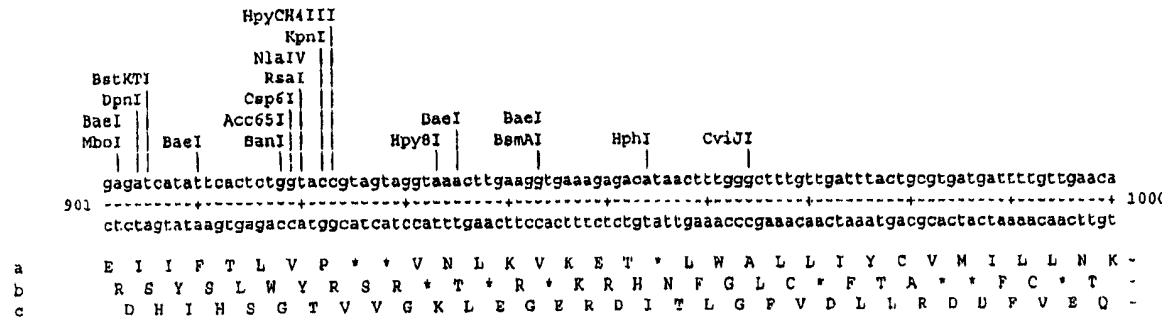
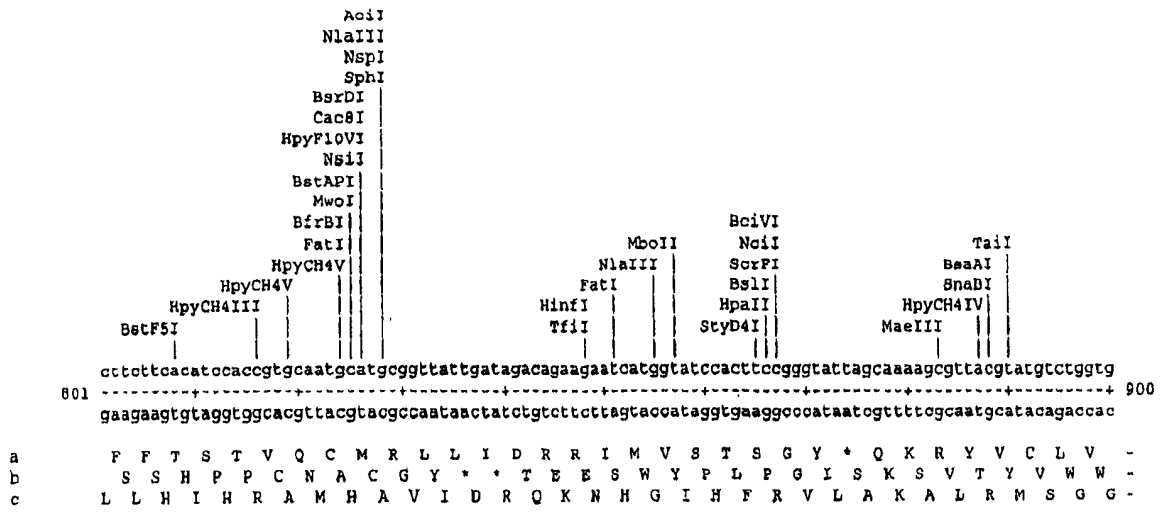


图 1-C



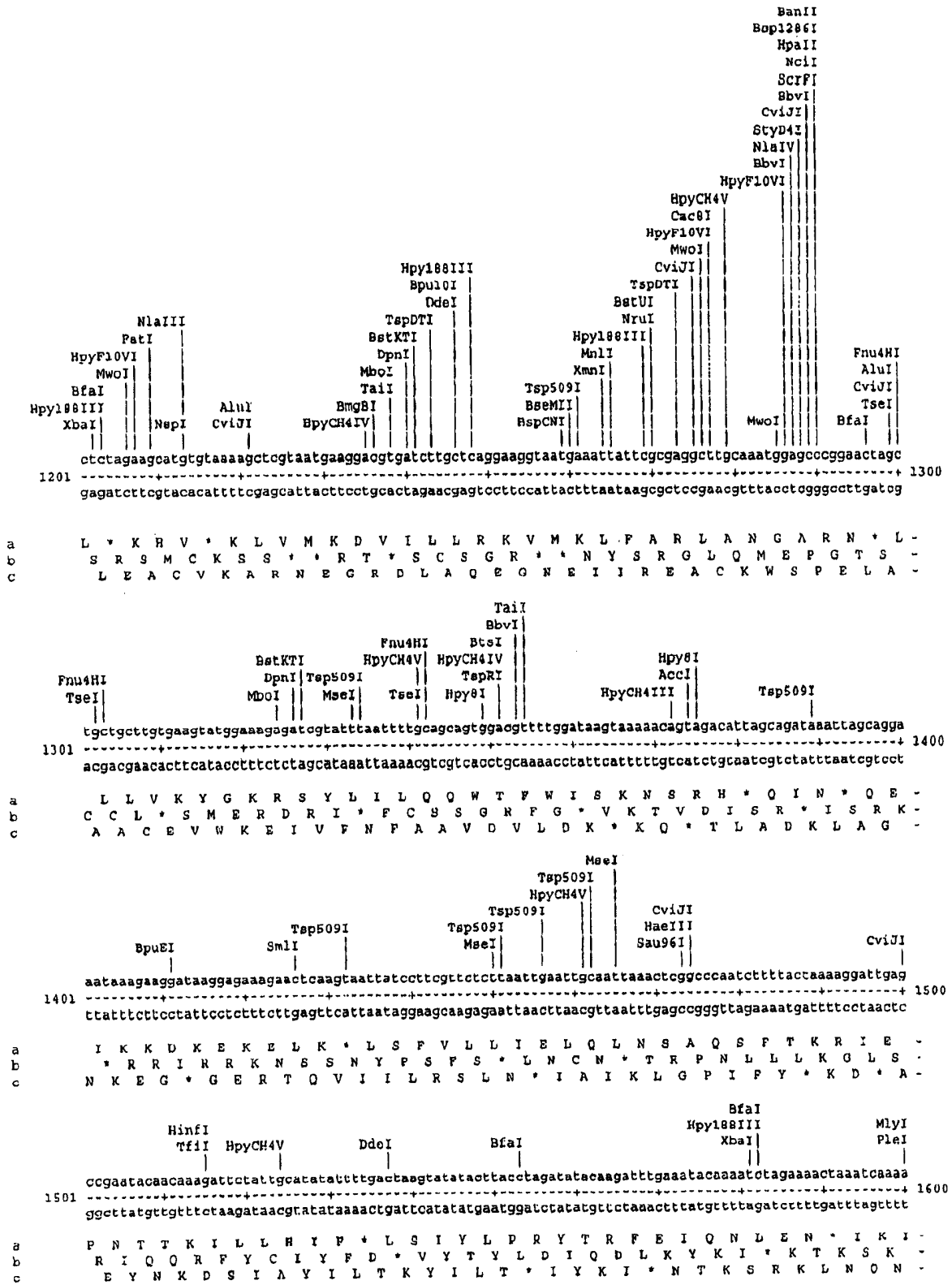


图 1-D

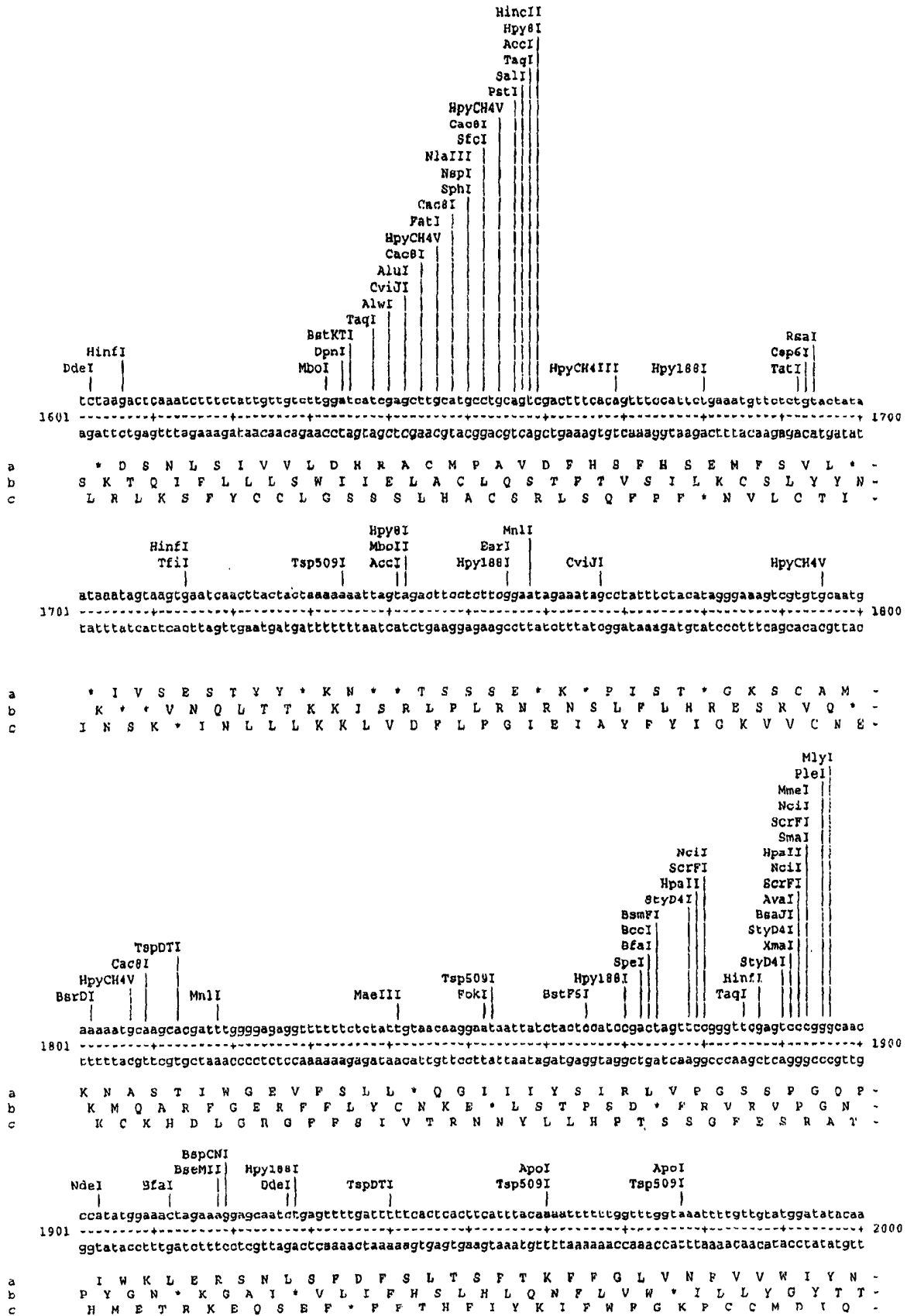


图 1-E

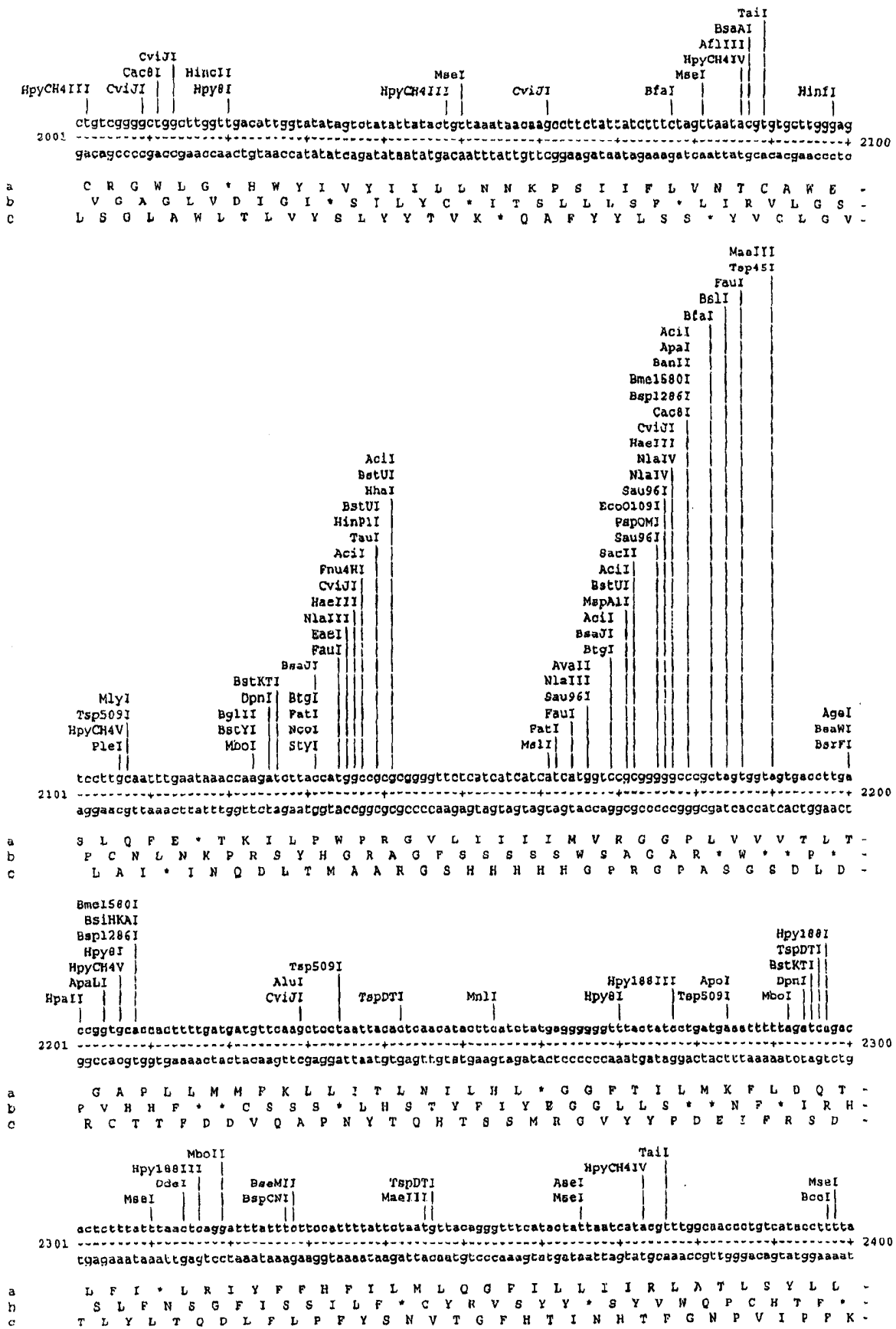


图 1-F

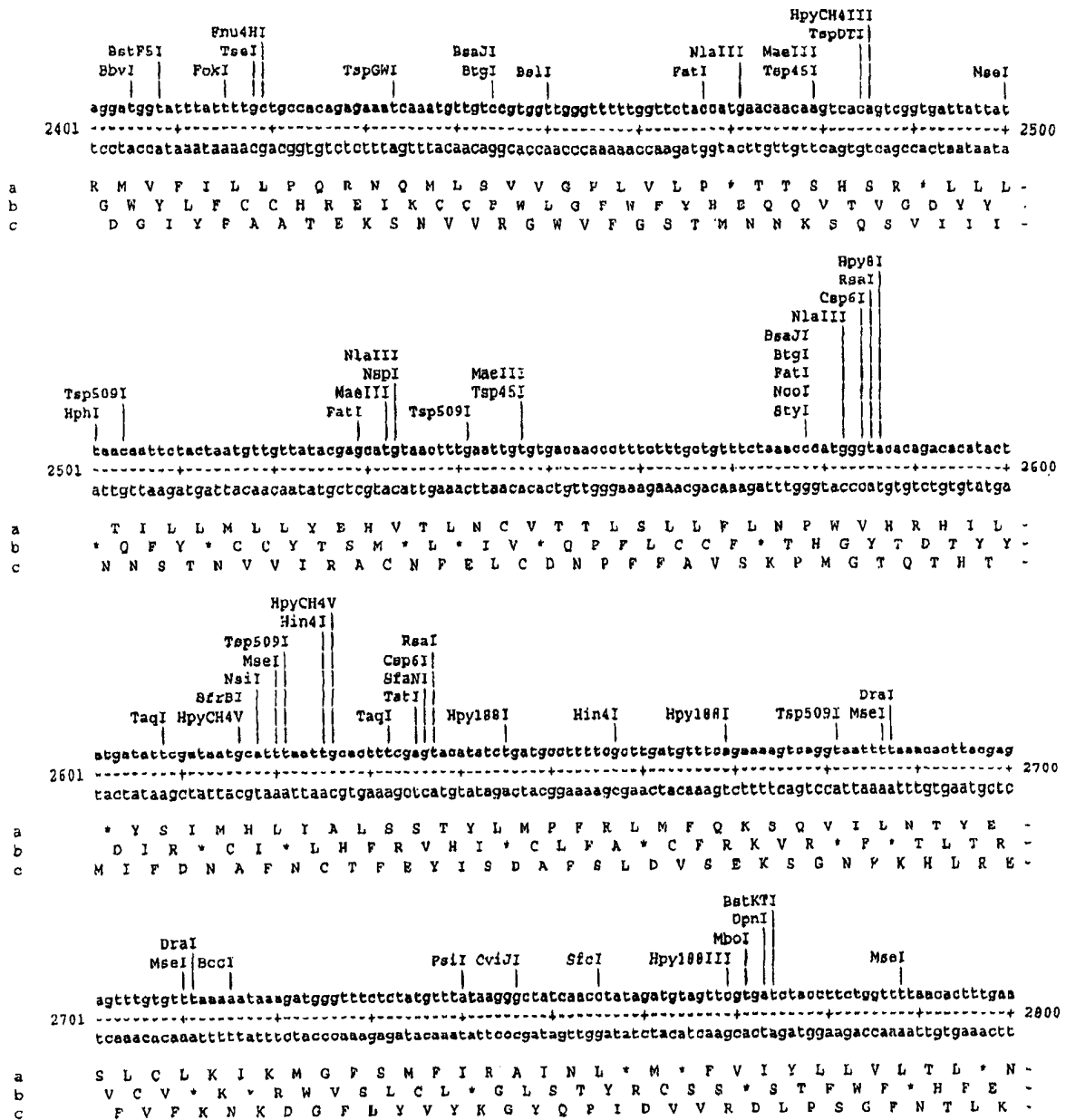


图 1-G

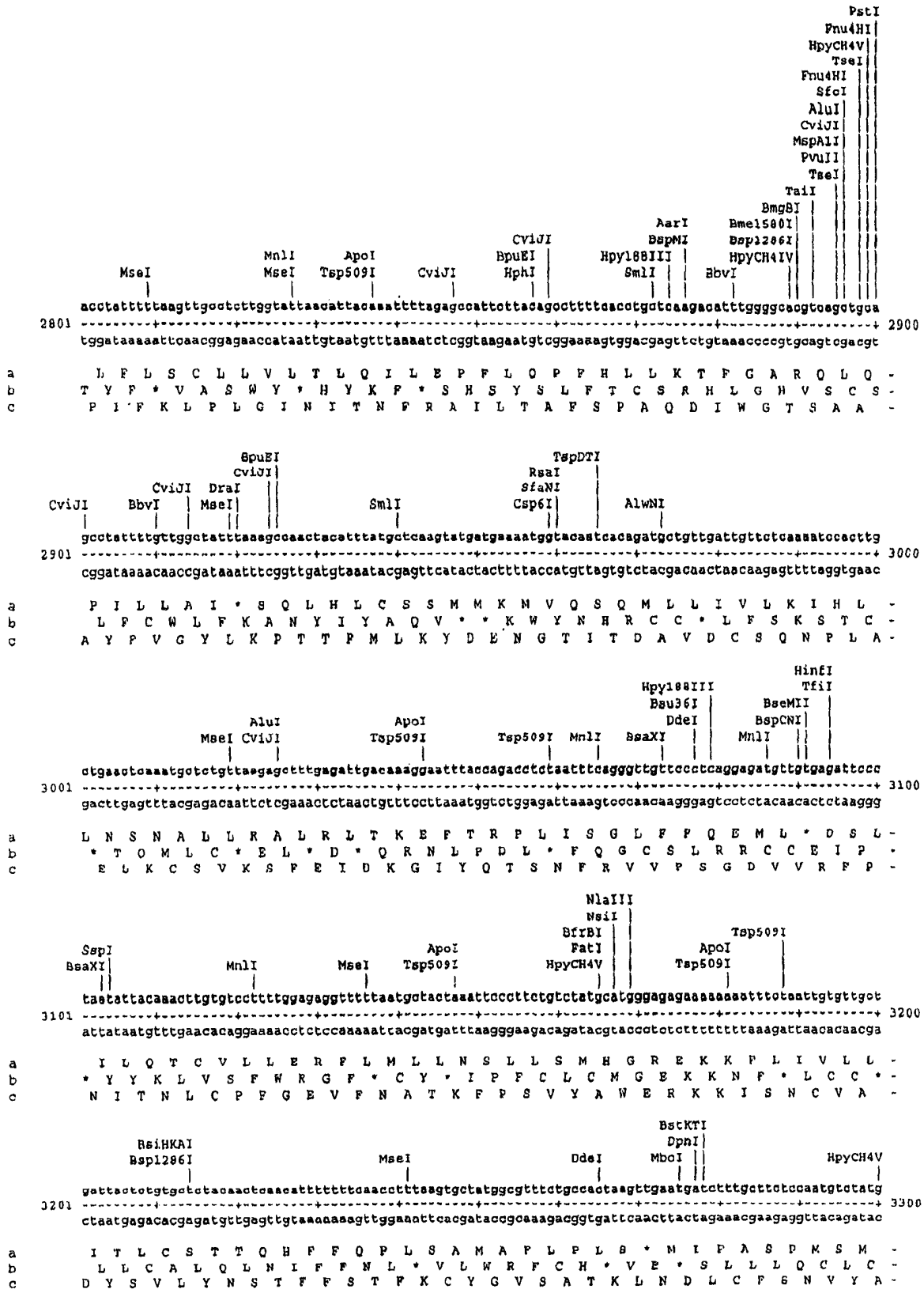


图 1-H

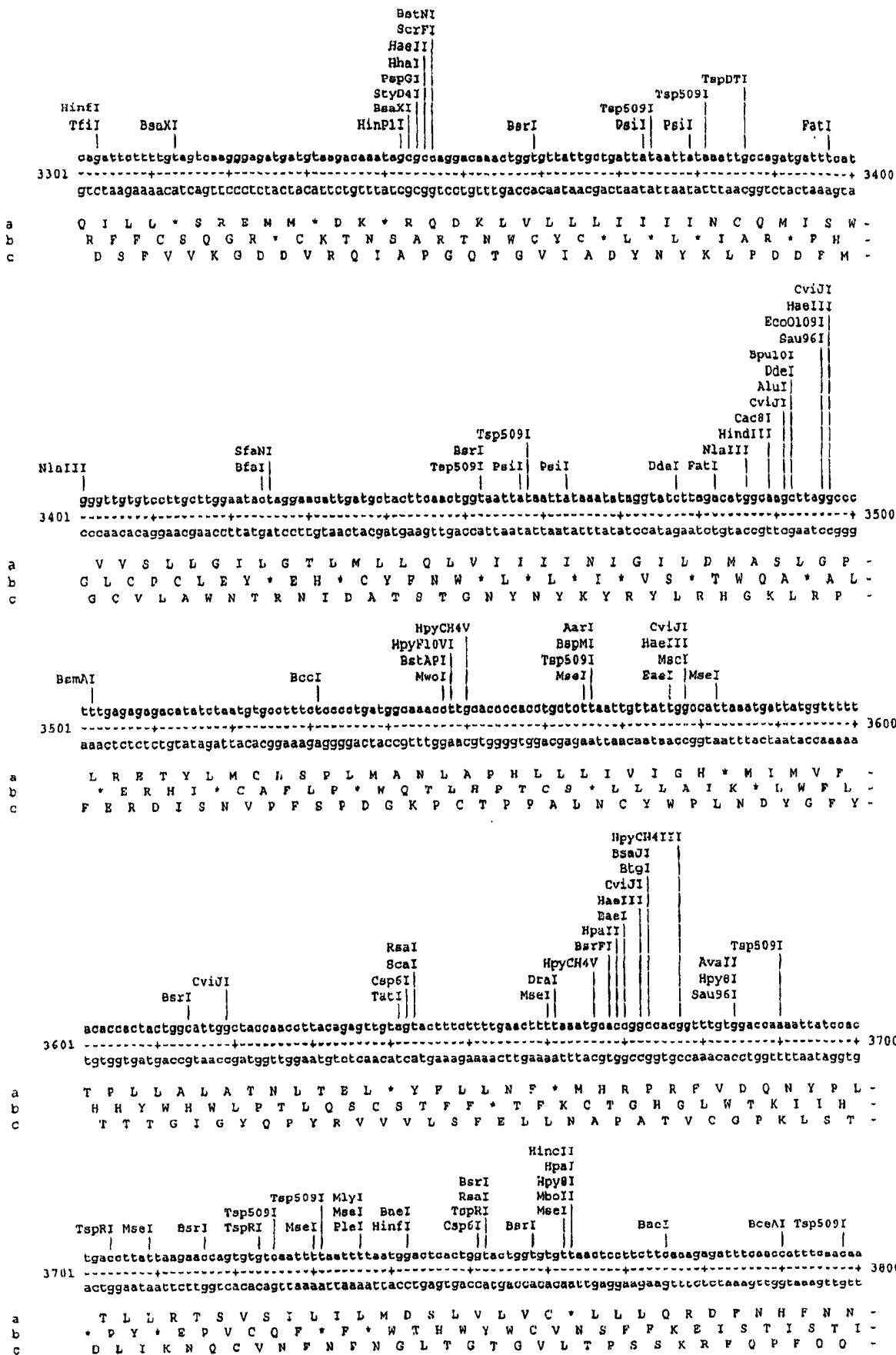


图 1-1

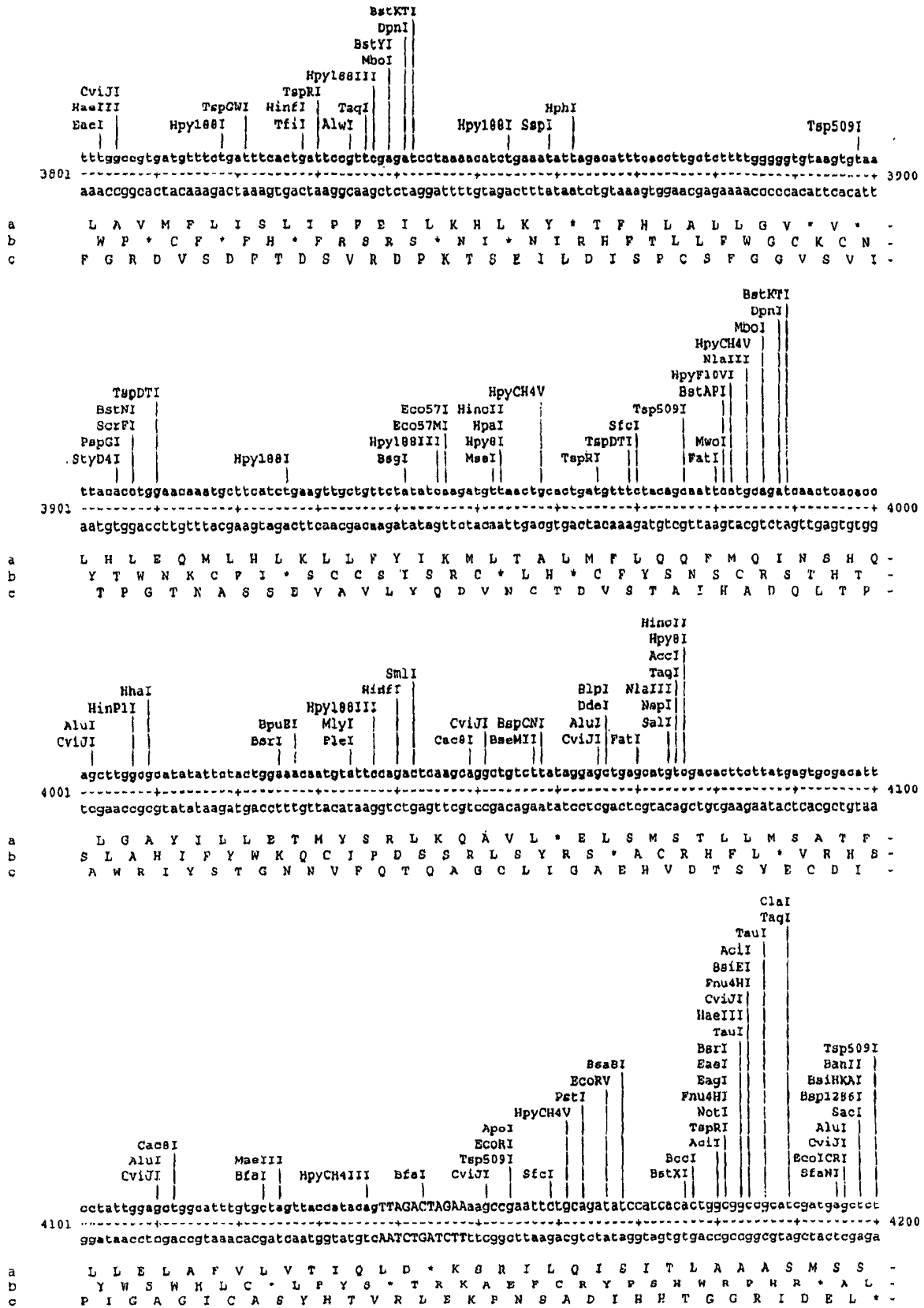


图 1-J

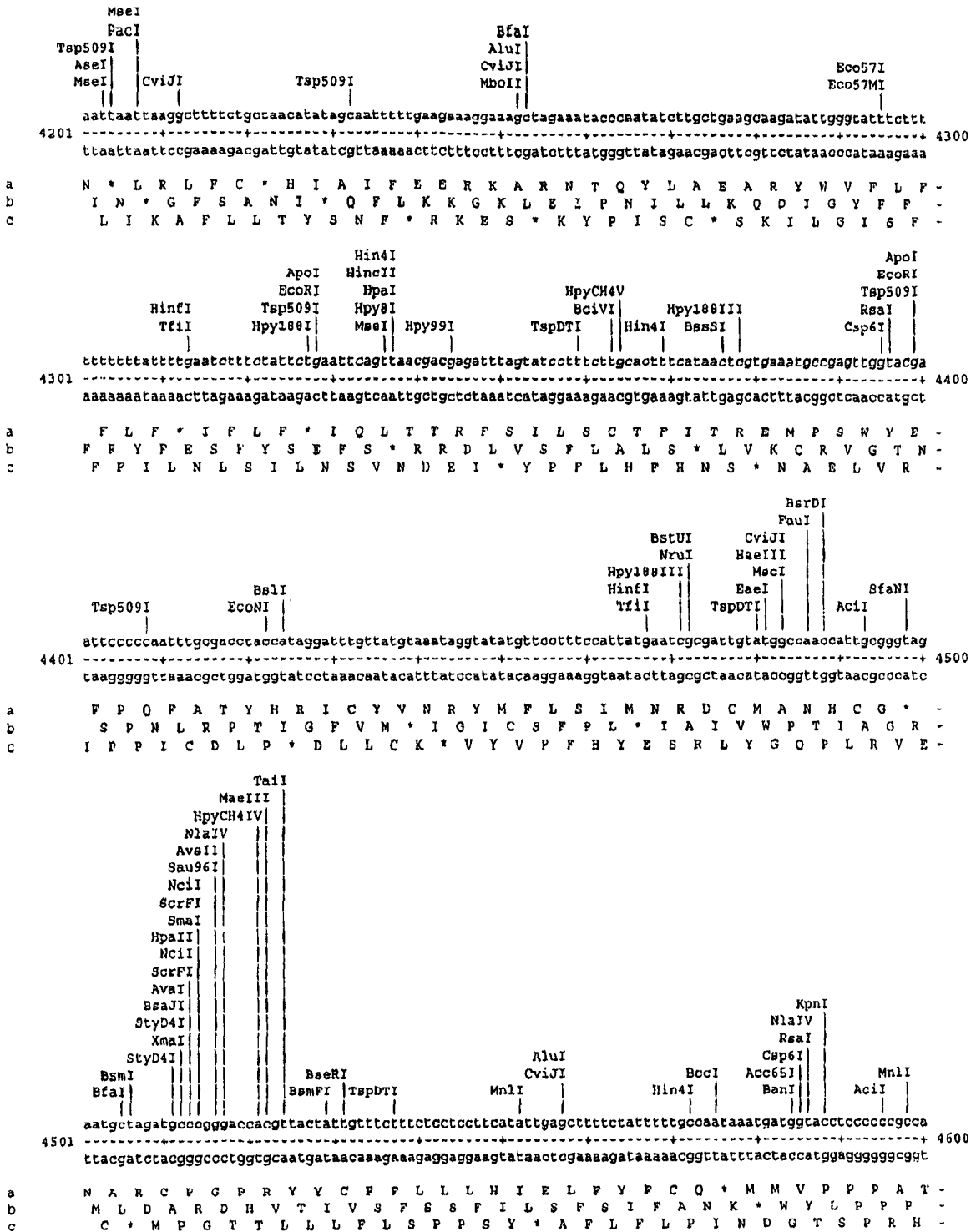


图 1-K



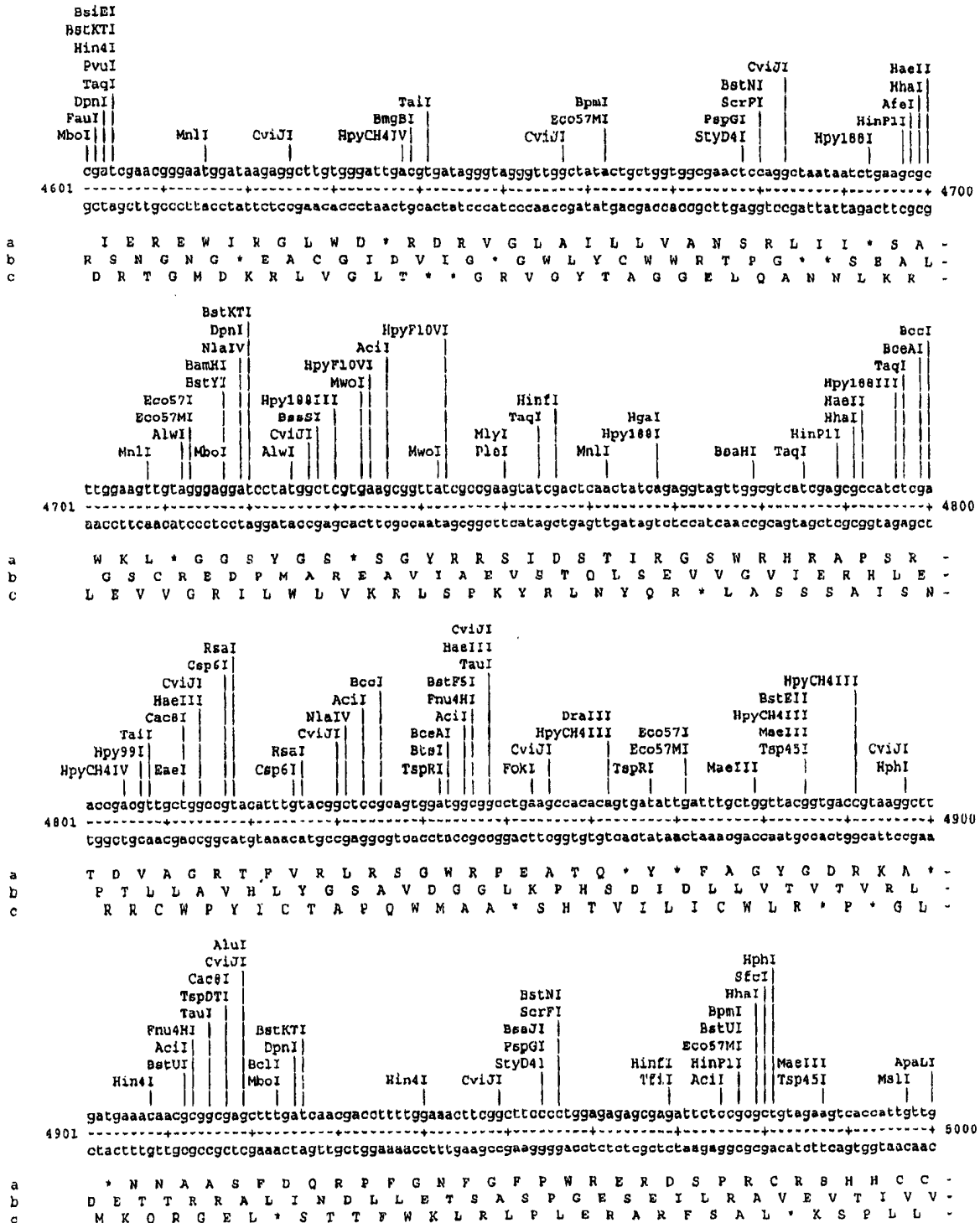


图 1-L

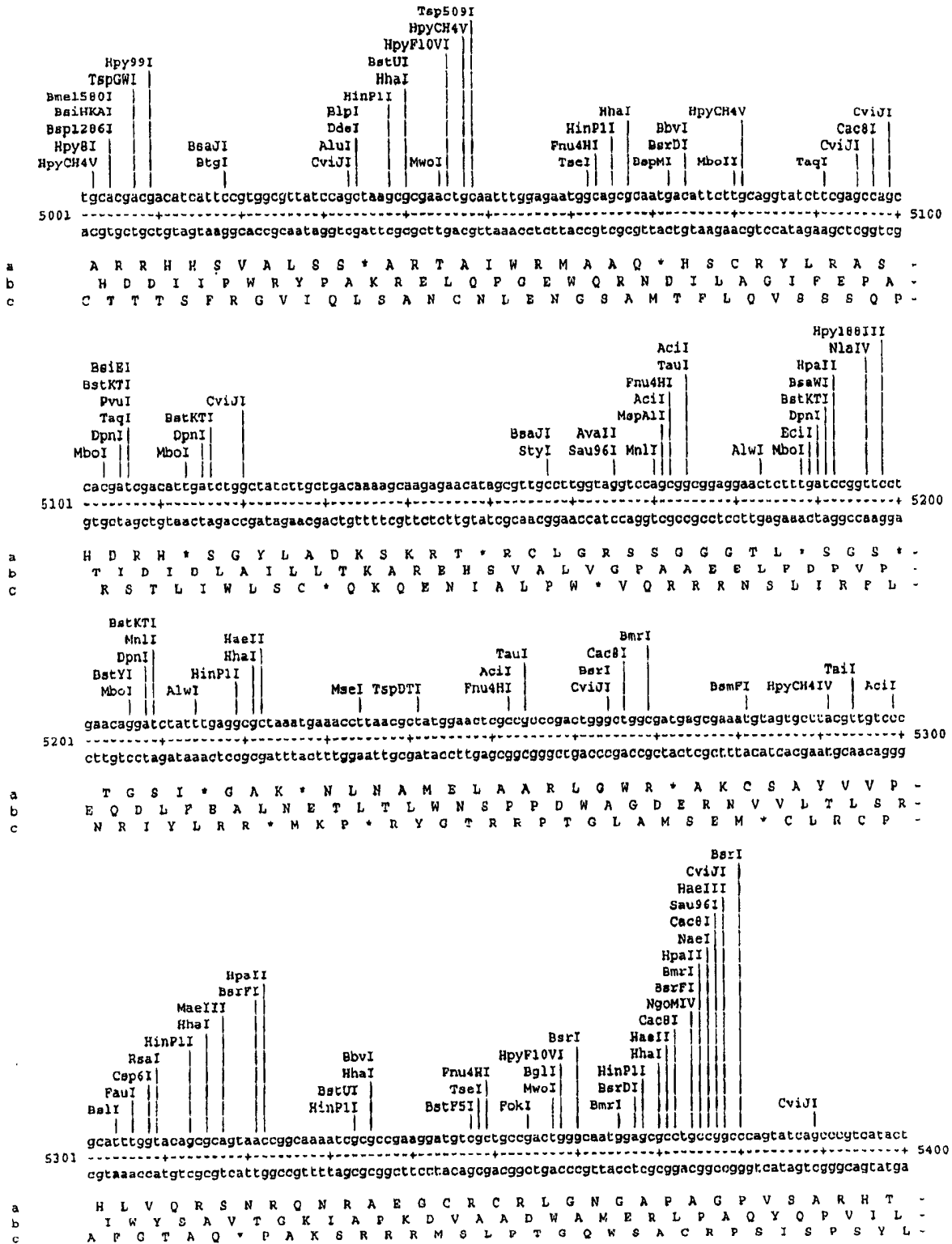


图 1-M

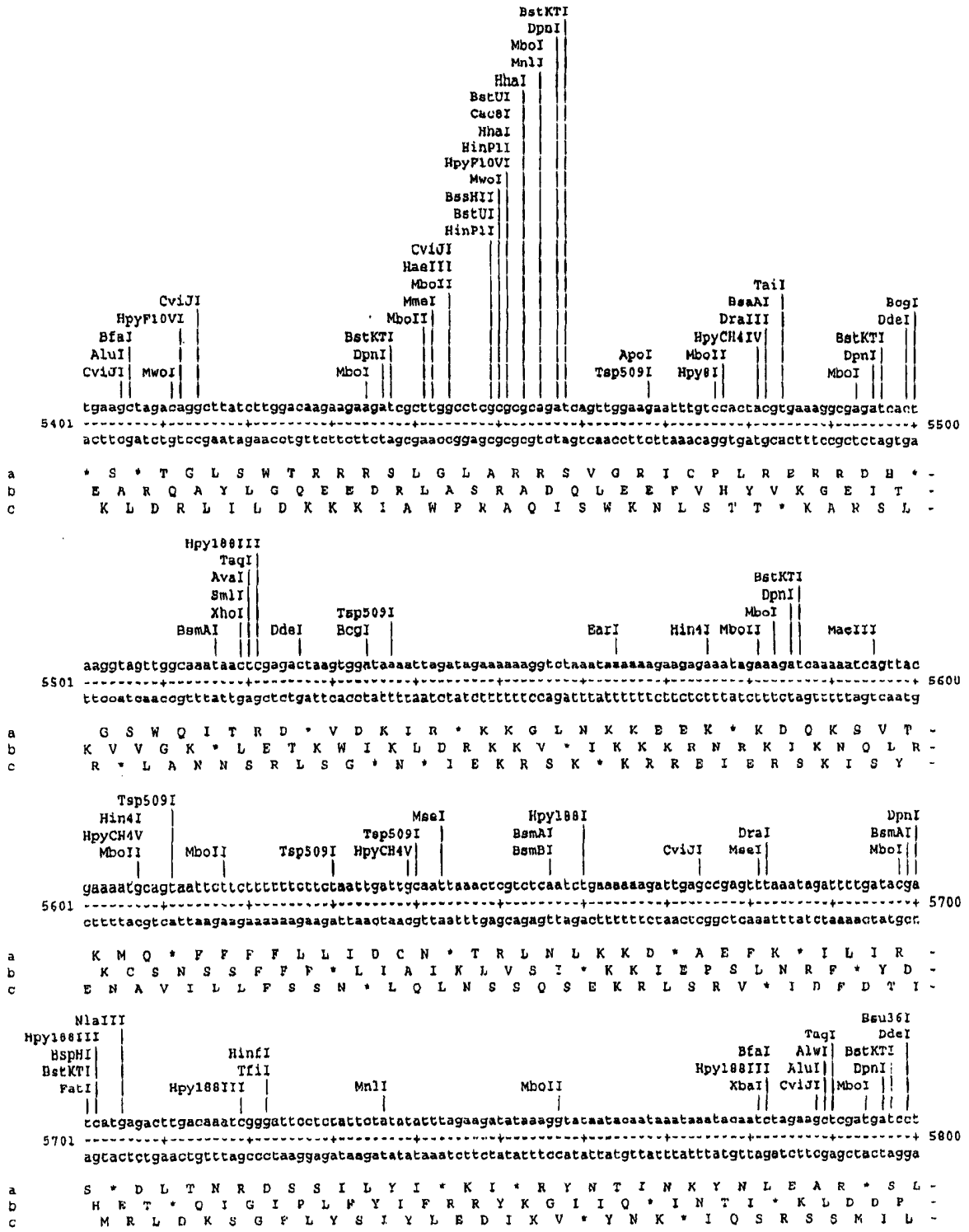


图 1-N

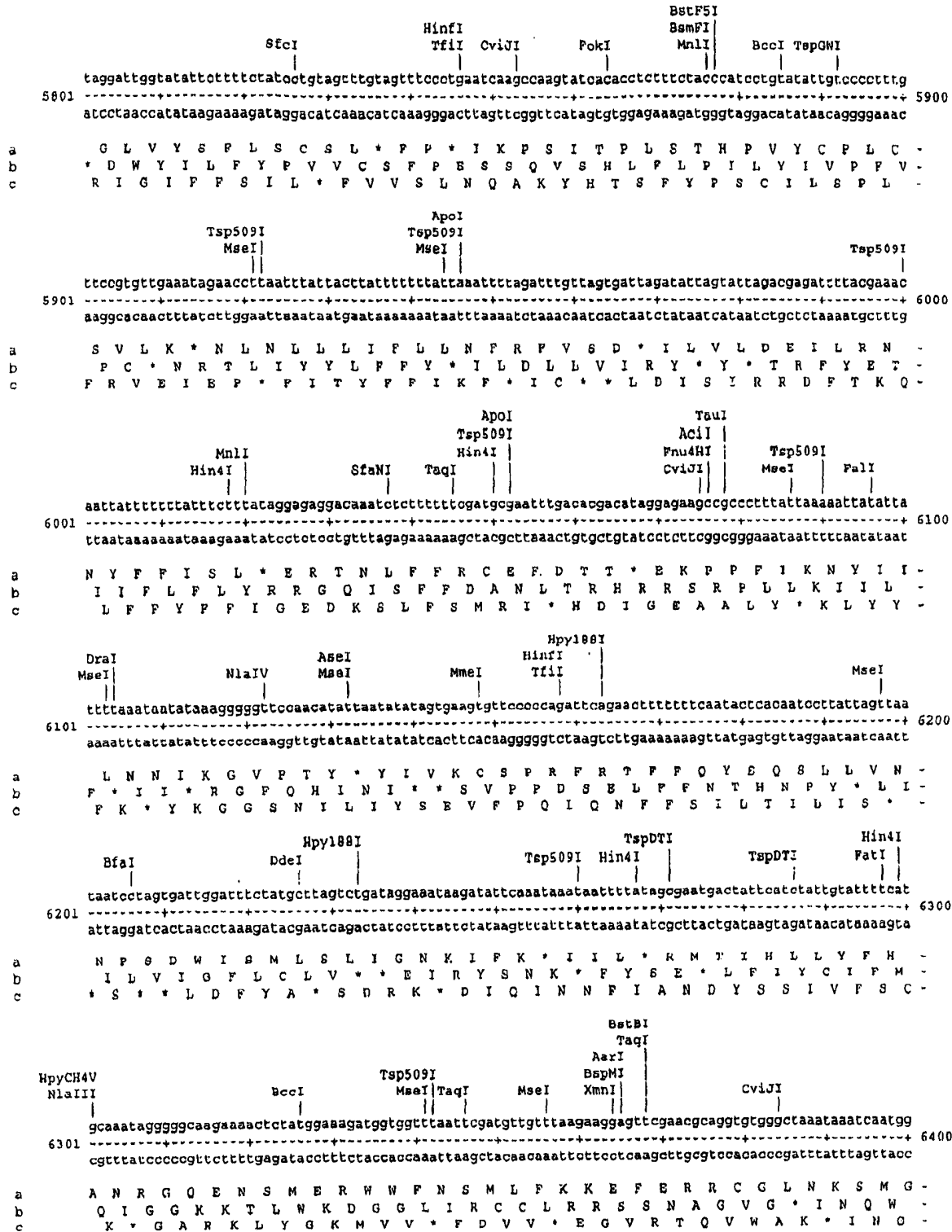


图 1-0

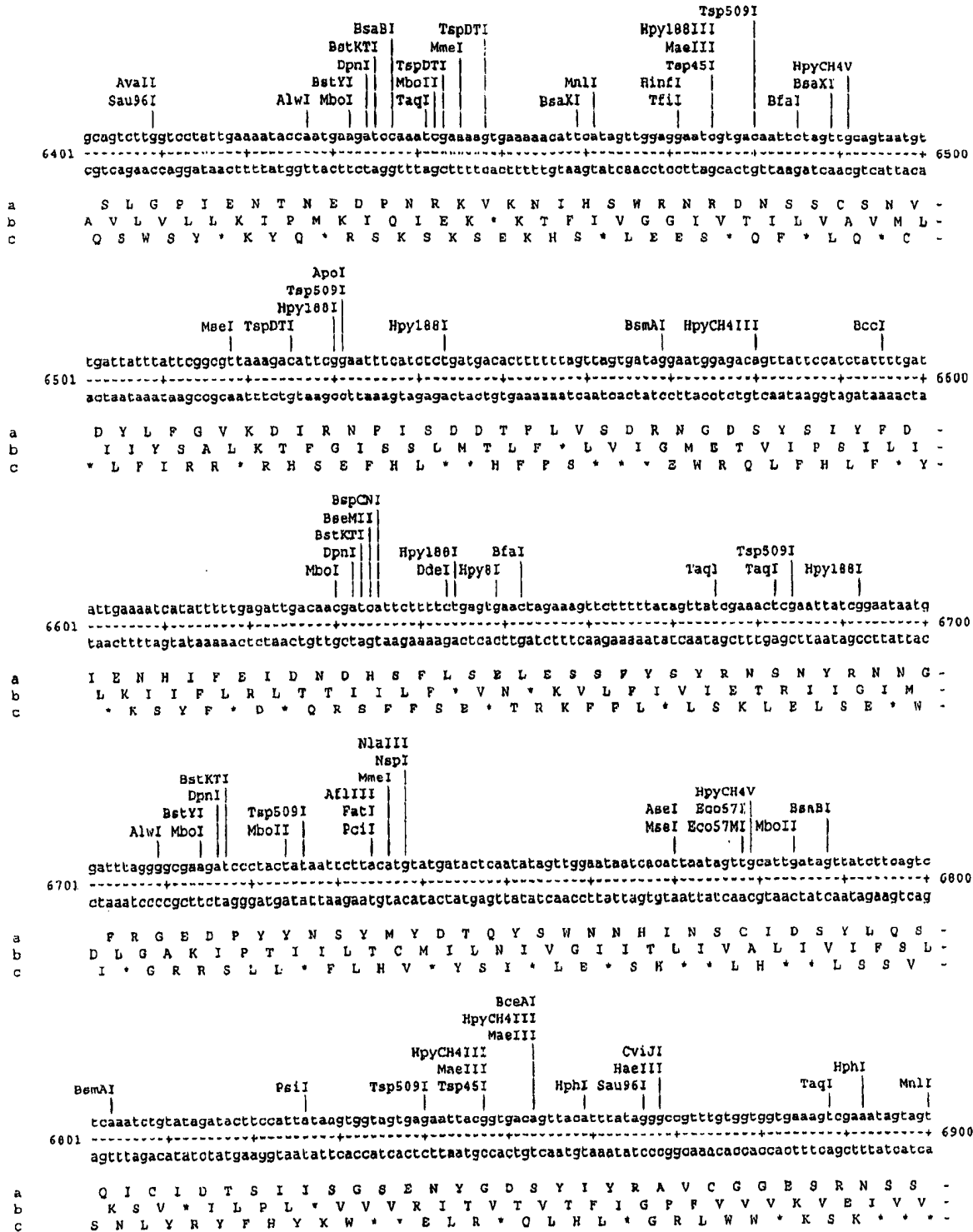


图 1-P

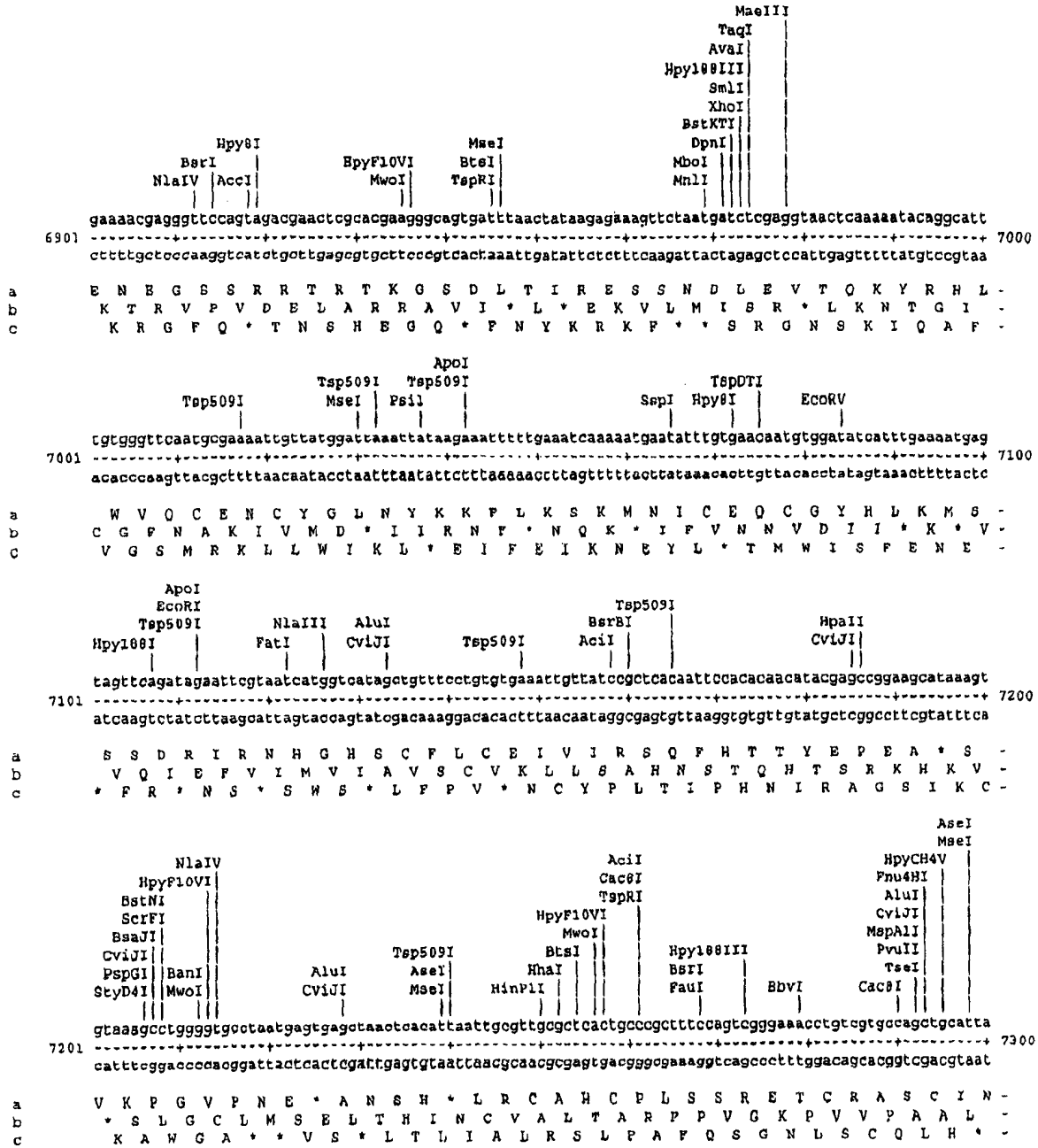


图 1-Q

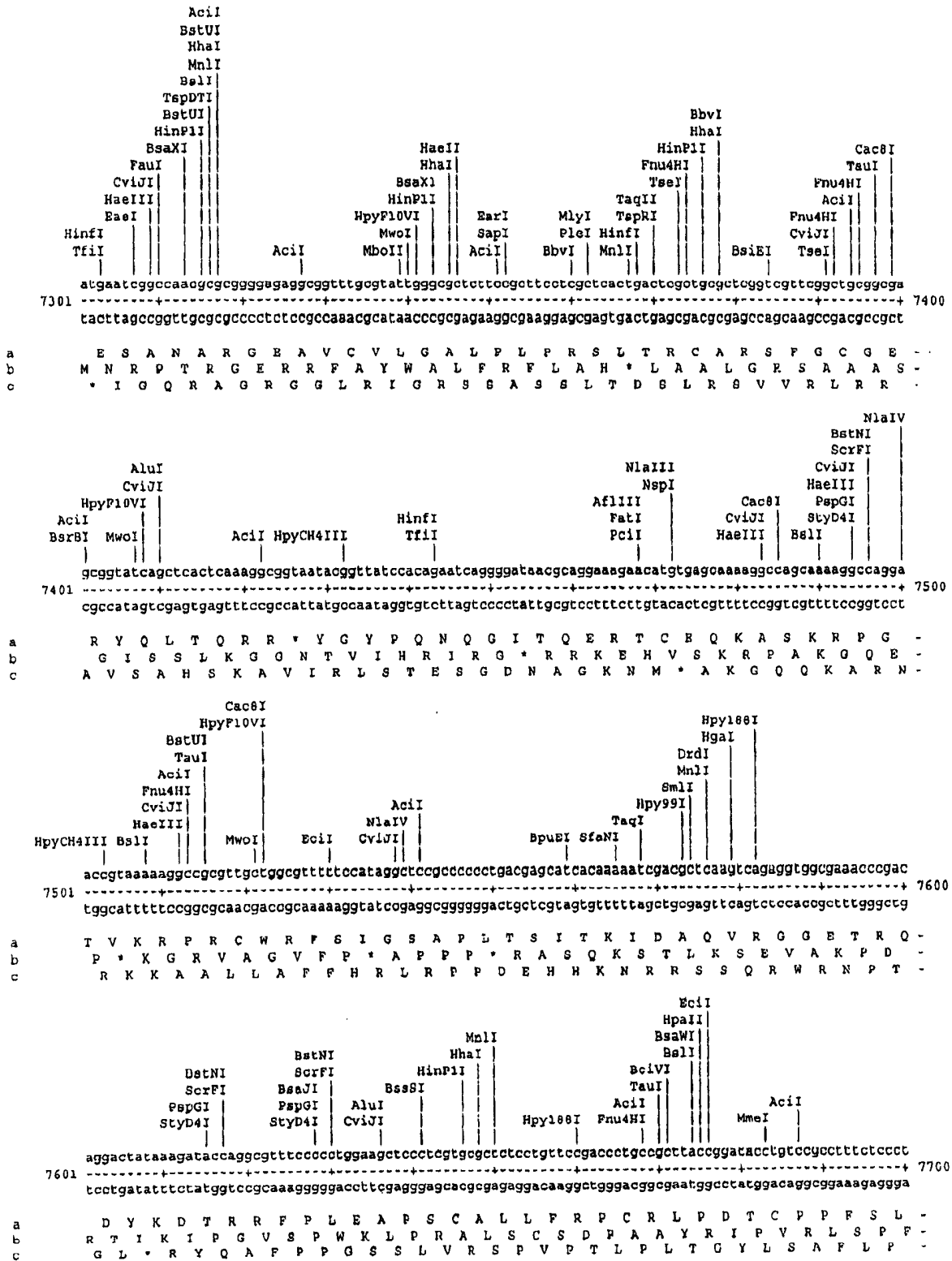


图 1-R

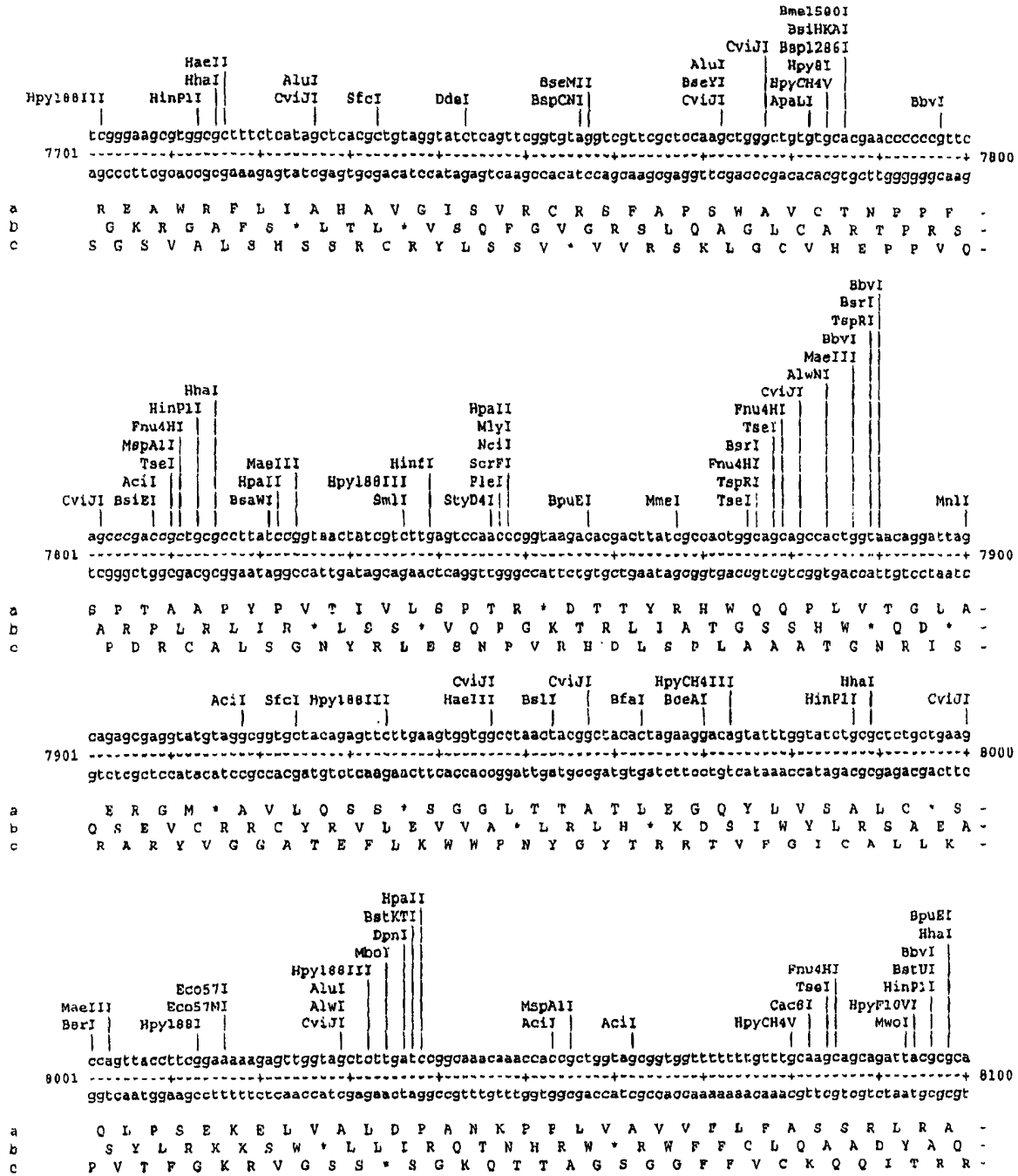


图 1-S



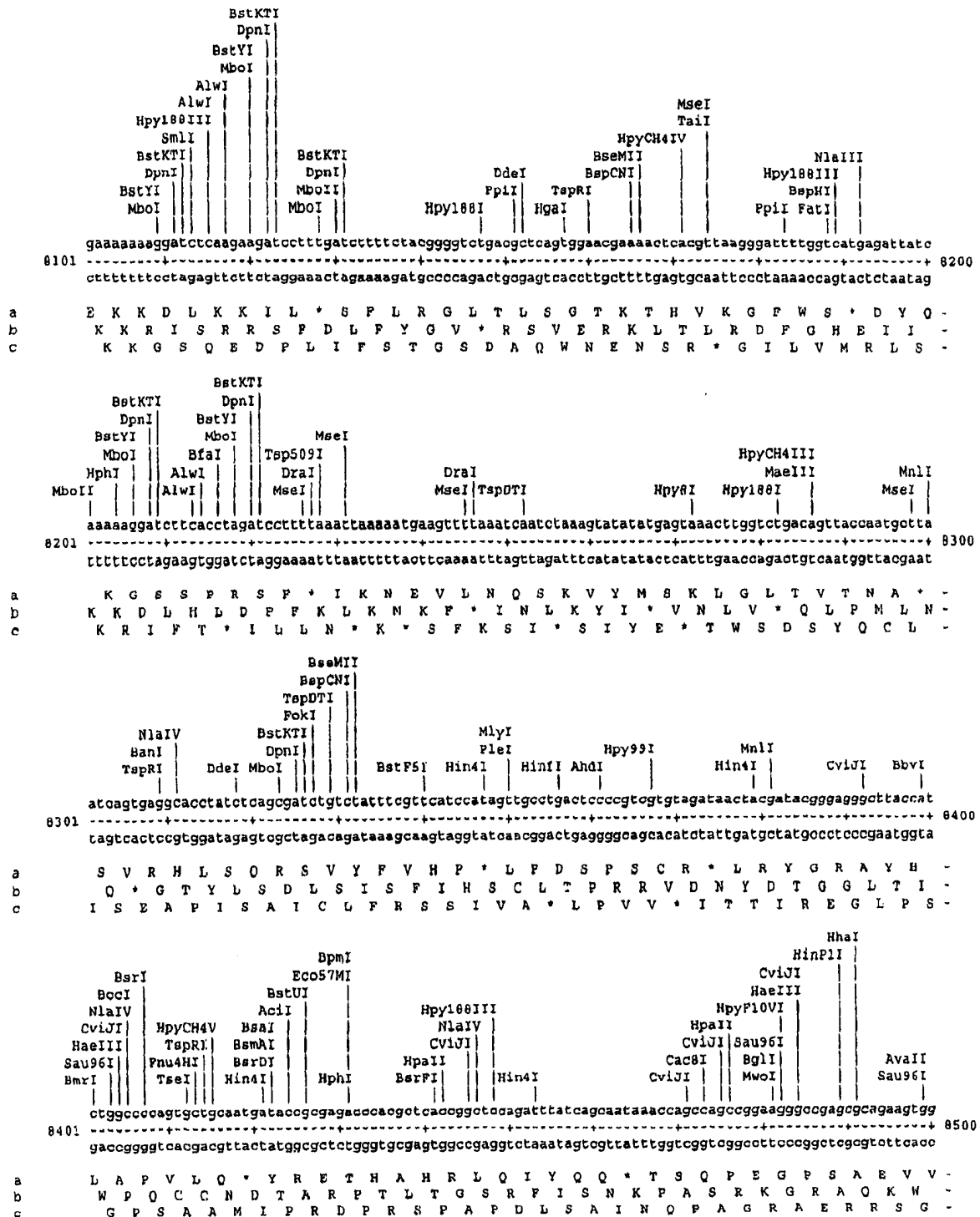


图 1-T

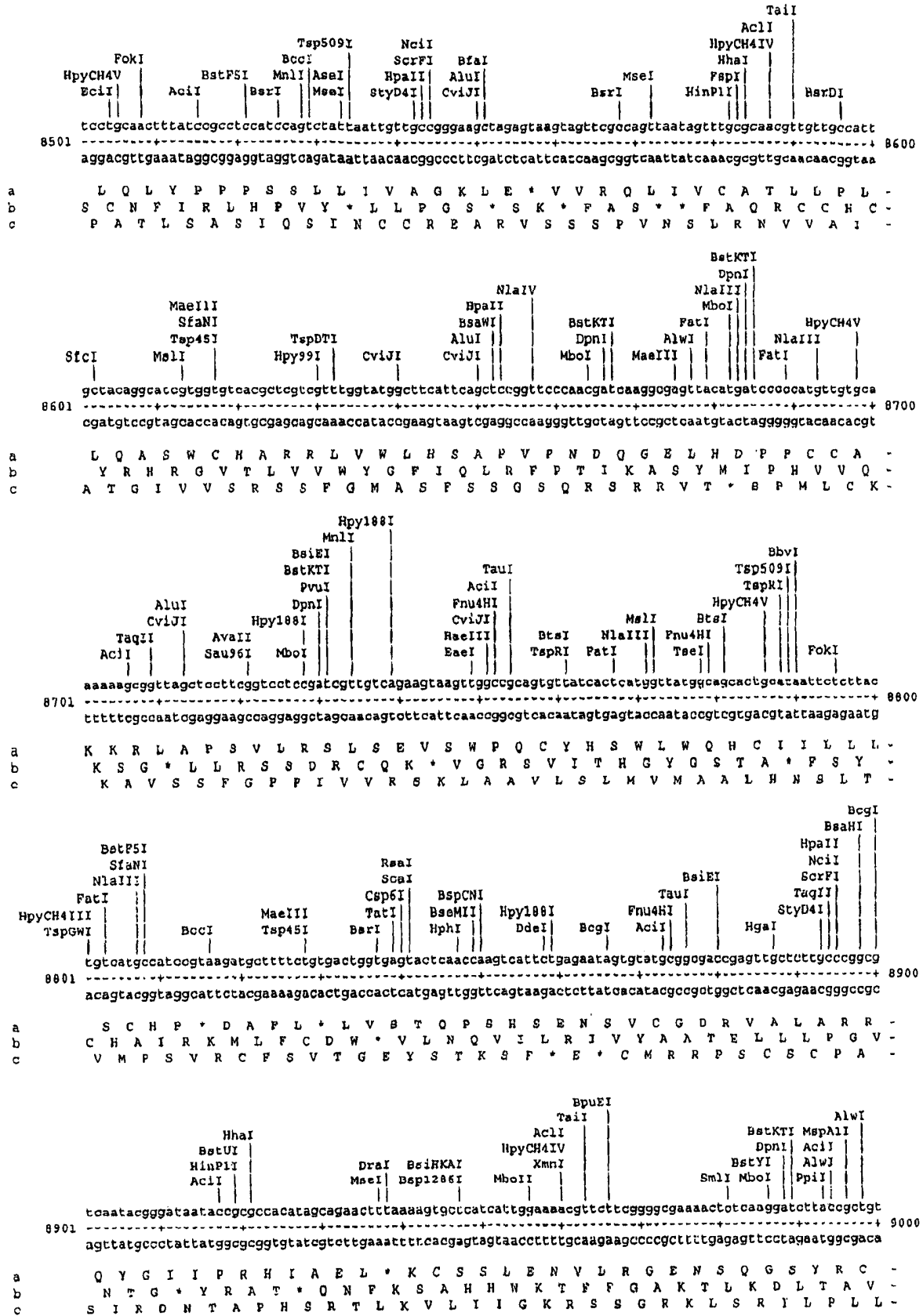


图 1-U

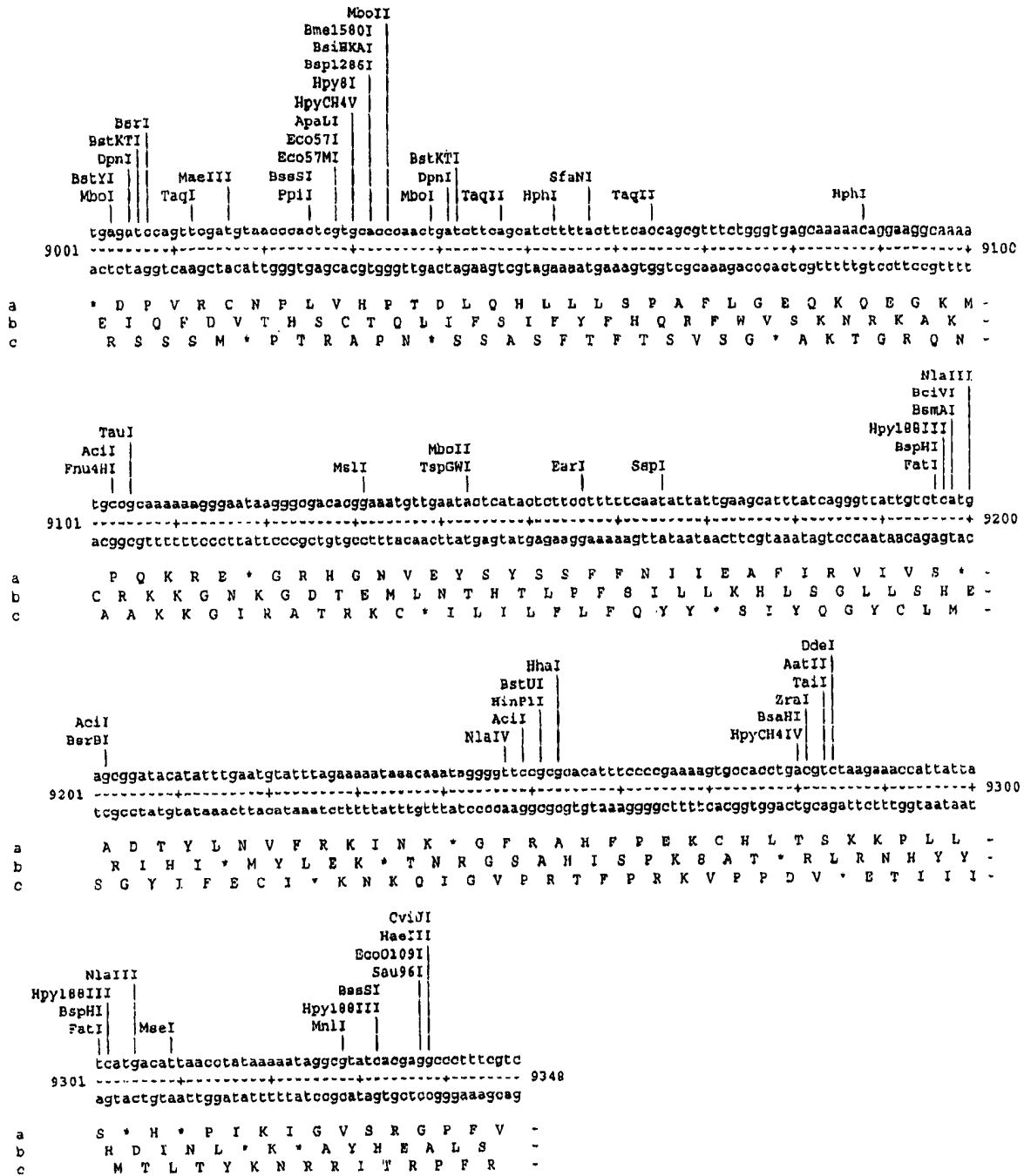


图 1-V

Enzymes that do cut and were not excluded:

AarI	AatII	AccI	Acc65I	Acii	AcII	AfeI	AflIII	AgeI	AhdI	AluI	AlwI	AlwNI
ApaI	ApaLI	ApoI	AseI	AvaI	AvaII	BaeI	BamBI	BanI	BanII	BbeI	BbsI	BbvI
BCCI	BceAI	BcgI	BciVI	BclI	Bfal	BfrBI	BglI	BgII	BlpI	Bme1580I	BmgBI	Bmri
BpmI	Bpu10I	BpuEI	BsaI	BsaAI	Bsai	BsaHI	BsaJI	BsaWI	BeaXI	BseMI	BseRI	BseYI
BsgI	BsiEI	BsiHKAI	BslI	BsmI	BsmAI	BsmBI	BsmFI	Bsp1286I	BspCNI	BspHI	BspMI	BsrI
BsrBI	BsrDI	BsrFI	BseHII	BseSI	SetAPI	SetBI	SetEI	SetF5I	BstKI	BstNI	BstUI	BstXI
BstYI	Bsu36I	StgI	BteI	Cac8I	Clai	Cep6I	CviJI	DdeI	DpnI	DraI	DraII	DrDI
EaeI	EagI	EarI	EclI	Eco57I	EcoICKI	Eco57MI	EcoNI	EcoO109I	EcoRI	EcoRV	FalI	FatI
FauI	Fnu4HI	FokI	FspI	HaeII	HaeIII	HgaI	HhaI	Hin4I	HinPI	HincII	HindIII	HinfI
HpaI	HpaII	NphI	Hpy8I	Hpy99I	Hpy188I	Hpy188III	HpyCH4III	HpyCH4IV	HpyCH4V	HpyF10VI	KasI	KpnI
MaeIII	MboI	MboII	MlyI	MmeI	MnlI	MscI	MseI	MslI	MspAI	MwoI	NaeI	NarI
NciI	NcoI	NdeI	NgoMIV	NlaIII	NlaIV	NotI	NruI	NsiI	NspI	PacI	PciI	PfoI
PleI	PpiI	PsiI	PspGI	PspOMI	PstI	PvuI	PvuII	RsaI	SacI	SacII	SalI	SapI
Sau96I	ScaI	ScrFI	SexAI	SfaNI	SfoI	SfoI	SmaI	SmlI	SnaBI	SpeI	SphI	SepI
StyI	StyD4I	TalI	TaqI	TaqII	TatI	TauI	TfiI	TseI	Tsp45I	Tsp509I	TspDTI	TspGI
TspRI	XbaI	XcmI	XhoI	XmaI	XmnI	ZraI						

Enzymes that do not cut:

AflII	AleI	AloI	AscI	AsiSI	AvrII	BbvCI	BatI	BplI	BsiWI	BspEI	BsrGI	BstZ17I
FseI	FspAI	MfeI	MluI	NheI	PflMI	PmeI	PmlI	PpuMI	PshAI	PerI	RerII	SanDI
SbfI	SfiI	SgrAI	SrfI	StuI	SwaI	TthIII						

Enzymes excluded; MinCuts: 1 MaxCuts: 100000

NONE

图 1-W

## pCV1 序列

```

tcgogcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
attcgcatt caggctcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300
tacgccagct ggcgaaagg ggatgtgctg caaggcgatt aagttgggtg acgccagggt 360
ttcccagtc acgacgttg aaaacgacgg ccagtgccaa gcttgaaaga gataaattga 420
acaagtatgg tegtccctg ttgggatgta ctattaaacc taaattgggg ttatctgcta 480
aaaactacgg tagagccgtt tatgaatgct ttcgcggtgg acttgattht actaaagatg 540
atgagaacgt gaactcaca ccatttatgc gttggagaga tcgthtctta tttgtgccc 600
aagcacttta taaagcacag gctgaaacag gtgaaatcaa agggcattac ttgattgcta 660
cgtcagttac atgcgaagaa atgatcaaaa gagctgtatt tgctagagaa ttgggcttc 720
cgatcgtaat gcatgactac ttaacggggg gattcaccgc aaatactagc ttggctcatt 780
attgccgaga taatggtcta cttcttcaca tccaccgtgc aatgcatgcg gttattgata 840
gacagaagaa tcatggtatc cacttcgggg tattagcaaa agcgttacgt atgtctggtg 900
gagatcatat tcaactcgtt accgtagtag gtaaactga aggtgaaaga gacataact 960
tgggctttgt tgatttactg cgtgatgatt ttgtgaaaca agatcgaagt cgcggtatth 1020
atttcaactca agattgggtc tcttaccag gtgtctacc cgtggttca ggaggtatth 1080
acgtttggca tatgcctgct ctgaccgaga tcttgggga tgattccgta ctacagttcg 1140
gtggaggaac tttaggacat cttggggta atgcgccagg tgccgtagct aatcagtag 1200
ctctagaagc atgtgtaaaa gctcgtaatg aaggacgtga tcttctcag gaaggtaatg 1260
aaattattcg cgaggcttgc aaatggagcc cggaaactagc tgctgcttgt gaagtatgga 1320
aagagatcgt atttaatttt gcagcagtg acgttttggg taagtaaaaa cagtagacat 1380
tagcagataa attagcagga aataaagaag gataaggaga aagaactcaa gtaattatcc 1440
ttcgttctct taattgaatt gcaattaaac tcggcccaat ctttactaa aaggattgag 1500
cgaatacaa caaagattct attgcatata tttgactaa gtatatactt acctagatat 1560
acaagatttg aaatacaaaa tctagaaaac taaatcaaaa tctaagactc aaatcttct 1620
attgttgtct tggatcatcg agcttgcctg cctgcagtcg acttccacag tttccattct 1680
gaaatgttct ctgtactata ataatagta agtgaatcaa ctactacta aaaaaattag 1740
tagacttctt cttcggaaata gaaatagcct atttctacat agggaaagtc gtgtgcaatg 1800
aaaaatgcaa gcacgatttg gggagaggtt tttctctat tgtaacaagg aataattatc 1860
tactccatcc gactagttcc gggctcgagt cccgggcaac ccatatggaa actagaaagg 1920
agcaatctga gttttgattt tcaactcact tcatttaca aatthtttg tttggtaat 1980
tttgttgtat ggatatacaa ctgtcggggc tggttgggtt gacattggtg tatagtctat 2040
attatactgt taaatacaa gocttctatt atcttctag ttaatacgtg tgctggggag 2100
tccttgcaat ttgaataaac caagatctta ccatggcgc gcgggttct catcatc 2160
atcatggtcc gcgggggccc actagtgtga gtgacctga ccggtgcacc actttgtg 2220
atgttcaagc tctaattac actcaacata ctctctat gaggggggtt tactatctg 2280
atgaaattht tagatcagac actcttatt taactcagga thtattctt ccatthatt 2340
ctaagtthc aggtthtcat actattaatc atagtttg caacctgtc atacctth 2400
aggatggtat thattttgct gccacagaga aatcaaatgt tgcctggtt tgggttht 2460
gttctaccat gaacaacaag tcacagtcg tgattattat taacaattct actaatgt 2520
ttatacagac atgtaactth gaattgtgtg acaaccctt cttgctgtt tctaacc 2580
tgggtacaca gacacatact atgatattc ataatgcatt taattgcact ttcgagta 2640
tatctgatgc cthtctgct gatgttctg aaaagtcagg taathttaa cacttacag 2700
agthtgtgt taaaaataaa gatgggttct tctatgtth taagggtat caacctag 2760
atgtagtthc tgatctacct tctggttht acacttgaa acctattht aagttgctc 2820
ttggtattaa cattacaaat thtagagcca thctacagc cthtcaact gctcaagac 2880
ttggggcac gtcagctgca gectattht ttggtatth aaagccaact acatthtgc 2940
tcaagtatga tgaaaatggt acaatcacag atgctgtg ttgttctca aatccactg 3000
ctgaactcaa atgctctgtt aagacttht agattgaca aggaatthc cagacctct 3060
atthcagggt tgttccctca ggagatgttg tgagattccc taatattaca aactgtgtc 3120
ctthtggaga ggtthttht gctactaaat tccctctgt ctatgcatg gagagaaaa 3180
aaatthctaa thgtgtgtg gattactctg tgcctacaa ctcaactth tthtcaact 3240
ttaagtgtc tggcgttht gccactaagt tgaatgatct thgttctcc aatgtctat 3300
cagattctth tgtagtcaag ggagatgat taagacaaat agcgcagga caactggtg 3360

```

图 2-A

```

ttattgctga ttataattat aaattgccag atgatttcat gggttgtgtc cttgcttggg 3420
atactaggaa cattgatgct acttcaactg gtaattataa ttataaatat aggtatctta 3480
gacatggcaa gcttaggcc tttgagagag acatatctaa tgtgccttcc tcccctgatg 3540
gcaaaccttg caccacacct gctcttaatt gttattggee attaaatgat tatggttttt 3600
acaccactac tggcattggc taccaacctt acagagtgt agtactttct tttgaacttt 3660
taaatgcacc ggccacgggt tgtggaccaa aattatccac tgaccttatt aagaaccagt 3720
gtgtcaattt taattttaat ggactcactg gtactgggtg gttaaactcct tcttcaaaga 3780
gatttcaacc atttcaacaa tttggccgtg atgtttctga tttcactgat tccgttcgag 3840
atcctaaaac atctgaaata ttagacattt caccttgctc ttttgggggt gtaagtgtaa 3900
ttacacctgg aacaaatgct tcatctgaag ttgctgttct atatcaagat gttaaactgca 3960
ctgatgtttc tacagcaatt catgcagatc aactcacacc agcttggcgc atatatctta 4020
ctggaaacaa tgtattccag actcaagcag gctgtcttat aggagctgag catgtcgaca 4080
cttcttatga gtgcgacatt cctattggag ctggcatttg tgctattacc atacagttag 4140
actagaaaag ggcgcaattc gcagatatcc atcacactgg cggccgcac gatgagctct 4200
aattaattaa ggcctttctg ctaacatata gcaatttttg aagaaaggaa agctagaaa 4260
acccaatatac ttgctgaagc aagatattgg gtatttcttt tttttttatt ttgaatcttt 4320
ctattctgaa ttcagttaac gacgagattt agtatccttt cttgcacttt cataactcgt 4380
gaaatgccga gttggtacga attccccaa tttgcgacct accataggat ttgtlatgta 4440
aatagggtata tgttctttc cattatgaat cgcgattgta tggccaacca ttgogggtag 4500
aatgctagat gccgggacc acgttactat tgtttctttc tctctcttca tattgagctt 4560
ttctattttt gccataaat gatgggtacct cccccgcca cgategaacg ggaatggata 4620
agaggcttgt gggattgacg tgatagggtg gggttggcta tactgctggg ggcgaactcc 4680
aggctaataa tctgaagcgc ttggaagttg tagggaggat cctatggctc gtgaagcggg 4740
tatcgccgaa gtatcgactc aactatcaga ggtagttggc gtcacgcagc gccatctcga 4800
accgacgttg ctggccgtac atttgtacgg ctccgcagtg gatggcggcc tgaagccaca 4860
cagtgatatt gatttgetgg ttacggtgac cgtaaggctt gatgaaacaa cgcggcgagc 4920
tttgatcaac gaccttttg aaacttcggc ttcccctgga gagagcgaga ttctccgcgc 4980
tgtagaagtc accattgttg tgcaogacga catcattccg tggcgttacc cagctaaccg 5040
cgaactgcaa ttggagaat ggcagcgcaa tyacattctt gcaggtatct tcgagccagc 5100
cacgatcgac attgatctgg ctatcttget gcacaaaagca agagaacata gcgttgctt 5160
ggtagggtcca gggcgaggg aactcttga tccggttctt gaacaggatc tatttgaggc 5220
gctaaatgaa acctaacgc tatggaactc gccgcccagc tgggctggcg atgagcggaa 5280
tgtagtgtct acgttgtccc gcatttggtg cagcgcagta accggcaaaa tcgcccggaa 5340
ggatgtcgct gccgactggg caatggagcg cctgccggcc cagtatcagc ccgtcact 5400
tgaagctaga caggcttacc ttggacaaga agaagatcgc ttggcctcgc gcgcagatca 5460
gttgaagaa ttgtccact acgtgaaagg cagatcact aaggtagttg gcaataaact 5520
cgagactaag tggataaaat tagatagaaa aaaggctctaa ataaaaaga agagaaatag 5580
aaagatcaaa aatcagttac gaaaatgcag taattcttct ttttcttct aattgattgc 5640
aattaaactc gtctcaatct gaaaaagat tgagccgagt ttaaatagat tttgatacga 5700
tcatgagact tgacaaatcg ggattctct attctatata tttagaagat ataaaggat 5760
aatacaataa ataaatacaa tctagaagct cagatgctt taggattggt atattctttt 5820
ctatctgtgta gtttgtagtt tccctgaatc aagccaagta tcacacctct tctaccat 5880
cctgtatatt gtcccccttg tccgtgttg aatagaacc ttaatttatt acttattttt 5940
ttattaaatt ttagatttgt tagtgattag atattagtat tagacgagat tttacgaaac 6000
aattattttt ttatttcttt ataggagagg acaaatctct ttttctgatg cgaatttgac 6060
acgacatagg agaagccgcc ctttattaaa aattatatta ttttaataa tataaagggg 6120
gttccaacat attaatatat agtgaagtgt tccccagat tcagaacttt ttttcaatac 6180
tcacaatcct tattagttaa taatcctagt gattggattt ctatgcttag tctgatagga 6240
aataagatat tcaataaat aattttatag cgaatgacta ttcactatt gtattttcat 6300
gcaaataggg ggcaagaaaa ctctatggaa agatgggtgg ttaattcgat gttgtttaag 6360
aaggagtctg aacgcaggtg tgggctaaat aatcaatgg gcagtcttg tcctattgaa 6420
aataccaatg aagatccaaa tcgaaaagtg aaaaacattc atagtggag gaactgtgac 6480
aattctagtt gcagtaatgt tgattattta ttcggcgtta aagacattcg gaatttcatc 6540
tctgatgaca ctttttagt tagtgatagg aatggagaca gttattcoat ctattttgat 6600
attgaaaatc atatttttga gattgacaac gatcattctt tcttgagtga actagaaagt 6660
tctttttata gttatogaaa ctogaattat cggaaataatg gatttagggg cgaagatccc 6720
tactataatt cttacatgta tgatactcaa tatagtggg ataatcacat taatagttgc 6780
attgatagtt atcttcagtc tcaaatctgt atagatactt ccattataag tggtagtgag 6840
aattacggtg acagttacat ttatagggcc gtttgtggtg gtgaaagtgc aaatagtagt 6900

```

图 2-B

```

gaaaacgagg gttccagtag acgaactcgc acgaaggcca gtgatttaac tataagagaa 6960
agttcctaag atctcgaggt aactcaaaaa tacaggcatt tgtgggttca atgcgaaaat 7020
tgttatggat taaattataa gaaatTTTTg aaatcaaaaa tgaatatttg tgaacaatgt 7080
ggatatcatt tgaaaatgag tagttcagat agaattcgta atcatggcca tagctgtttc 7140
ctgtgtgaaa ttgttatccg ctcaaatc cacacaacat acgagccgga agcataaagt 7200
gtaaagcctg ggggtgcctaa tgagttagct aactcacatt aattgcgttg cgtcactgc 7260
ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg 7320
ggagaggcgg tttgcgtatt gggcgctctt ccgcttccct gctcactgac tcgctgcgct 7380
cggtcggtcg gctgcggcga gcggtatcag ctactcaaa ggcggtaata cggttatcca 7440
cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga 7500
accgtaaaaa ggcgcggtg ctggcgtttt tccataggct ccgccccct gacgagcctc 7560
acaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa agataaccag 7620
cgtttcccc tggaagctcc ctctgctgct ctccgttcc gaccctgccc cttaccggat 7680
acctgtccgc ctttctccct tcgggaagcg tggcgctttc tcatagctca cgtgttaggt 7740
atctcagttc ggtgtaggtc gttcgtcca agctgggctg tgtgcaagaa cccccgctc 7800
agcccgaccg ctgcccctta tccggtaact atcgtcttga gtccaaccg gtaagacacg 7860
acttatcgcc actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg 7920
gtgctacaga gttcttgaag tgggtggccta actacggcta cactagaagg acagtatttg 7980
gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg 8040
gcaaacaaac caccgctggg agcgggtggt tttttgtttg caagcagcag attacgcgca 8100
gaaaaaaagg atctcaagaa gatcctttga tctttctac ggggtctgac gctcagtggg 8160
acgaaaactc acgtaaggg attttgggtc tgagattatc aaaaaggatc ttcacctaga 8220
tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt 8280
ctgacagtta ccaatgctta atcagtgagg cacctatctc agcagctgt ctatttctgt 8340
catccatagt tgctgactc ccgctcgtgt agataactac gatacgggag ggcttaccat 8400
ctggccccag tgctgcaatg ataccgag acccacgctc accggctcca gatttatcag 8460
caataaacca gccagccgga agggccgagc gcagaagtgg tctgcaact ttatccgct 8520
ccatocagtc tattaattgt tgcggggaag ctagagtaag tagttcgcca gttaatagtt 8580
tgcgcaacgt tgttgccatt gctadaggca tctgtgtgct acgctcgtcg tttggtatgg 8640
cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatcccc atgtgtgca 8700
aaaaagcggg tagctccttc ggtcctccga tctgtgctag aagtaagttg gccgcaggtg 8760
tatcactcat ggttatggca gcaactgcata attctcttac tgtoatgcca tccgtaagat 8820
gcttttctgt gactggtgag tactcaacca agtcattctg agaatagttg atgcccgcac 8880
cgagttgctc ttgcccggcg tcaatacggg ataataccgc gccacatagc agaactttaa 8940
aagtgtcat cattggaaaa cgttottcgg ggcgaaaact ctcaaggatc ttaccgctgt 9000
tgagatccag ttgatgtaa cccactcgtg cacccaactg atcttcagca tcttttactt 9060
tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa 9120
ggcgacacg gaaatggtga atactcatac tcttctttt tcaatattat tgaagcattt 9180
atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa aataaacaaa 9240
taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa accattatta 9300
tcatgacatt aacctataaa aatagggcgt tccagaggcc ctttctgc 9348

```

图 2-C

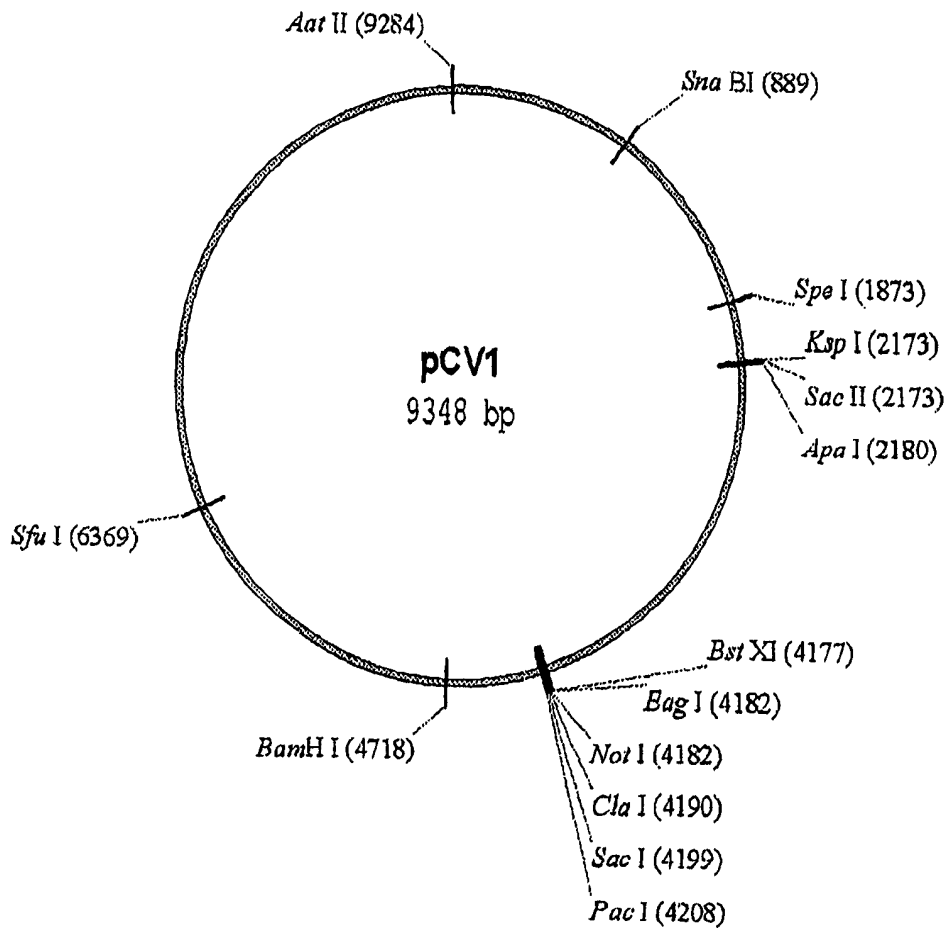


图 3-A



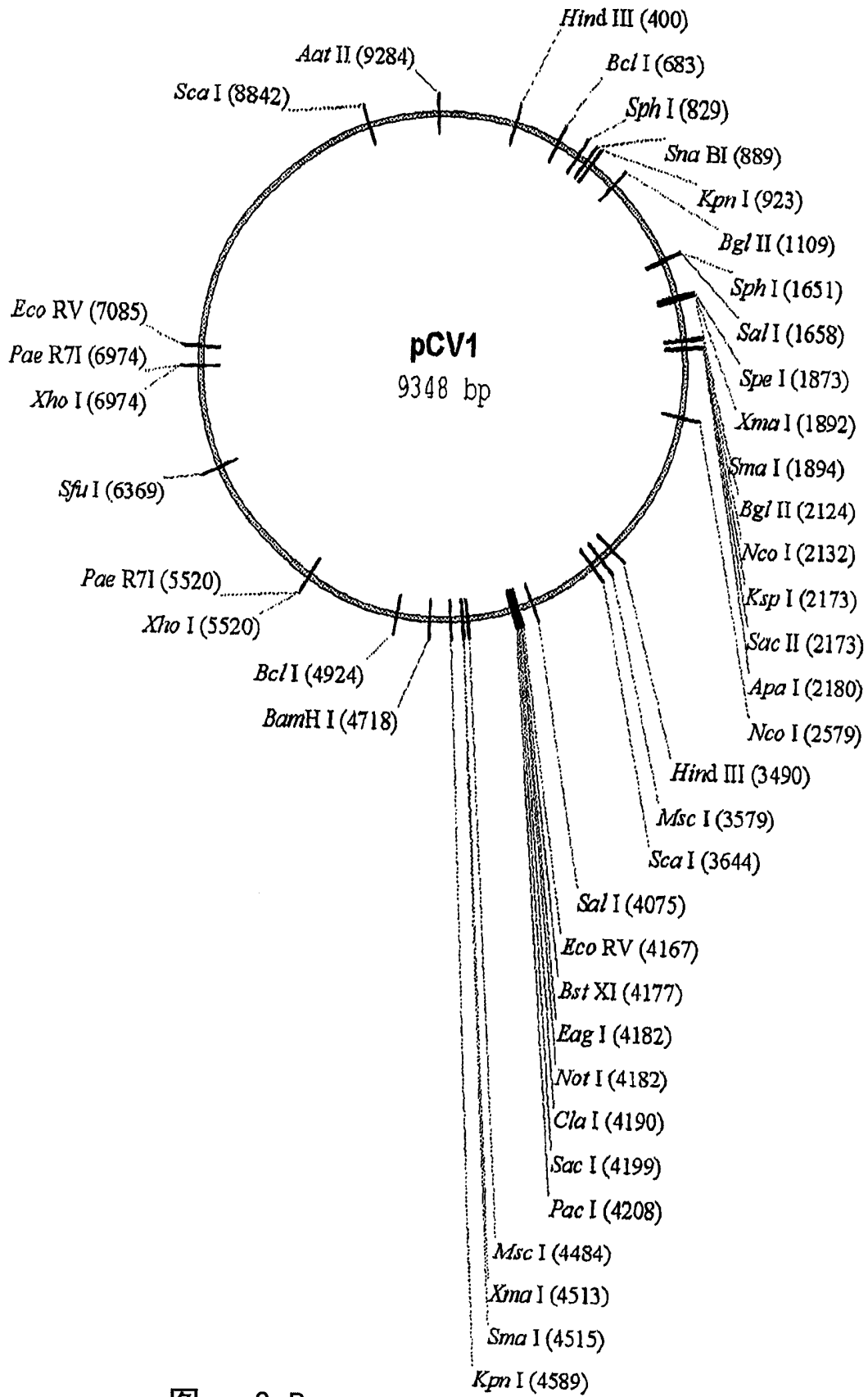


图 3-B

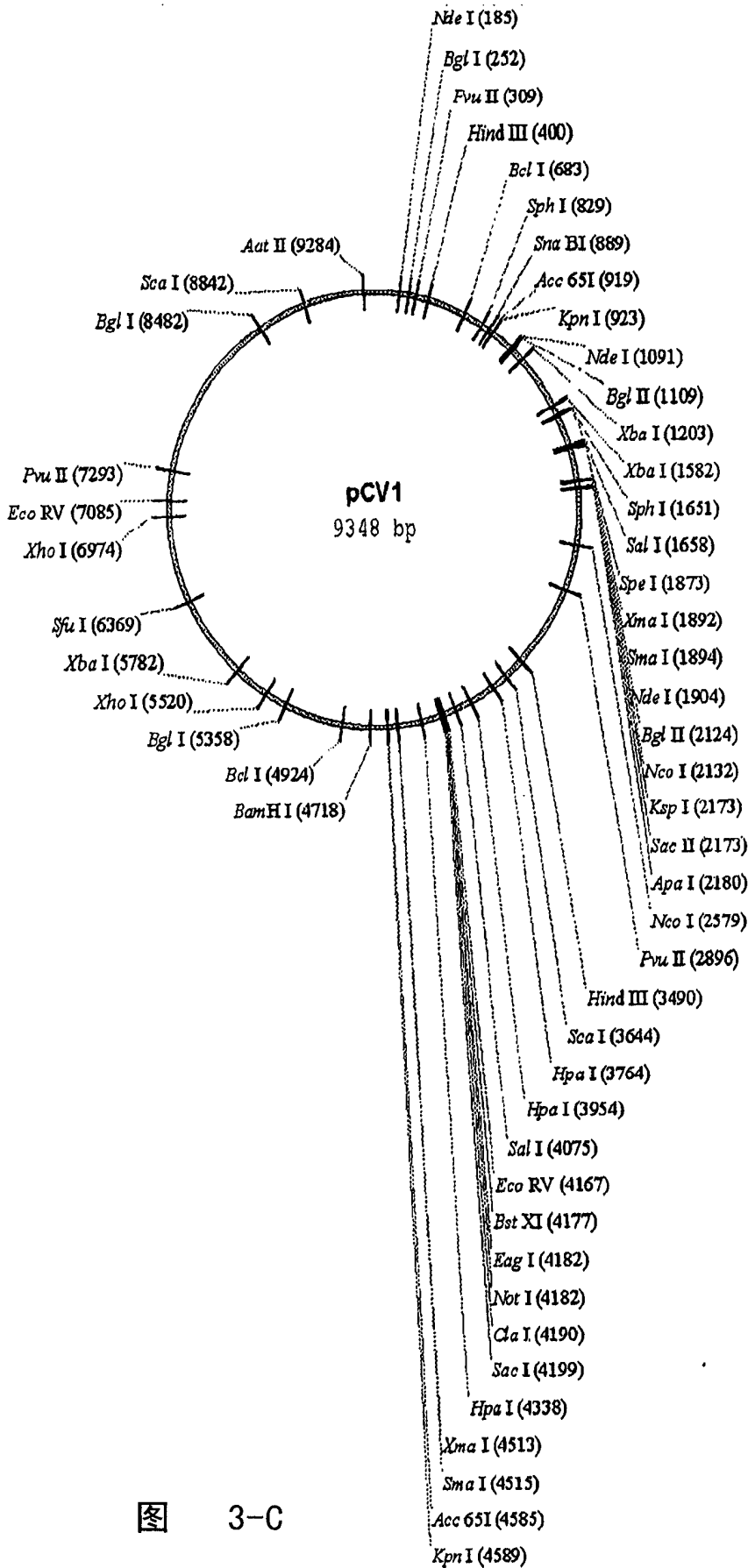


图 3-C

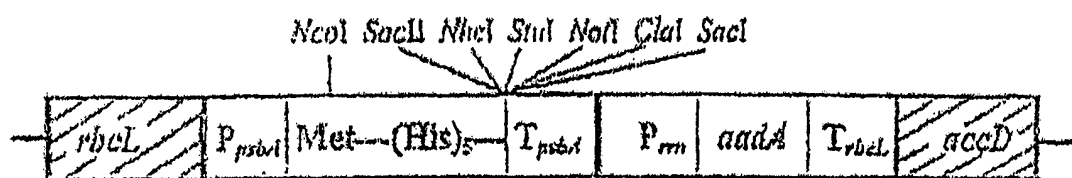


图 4

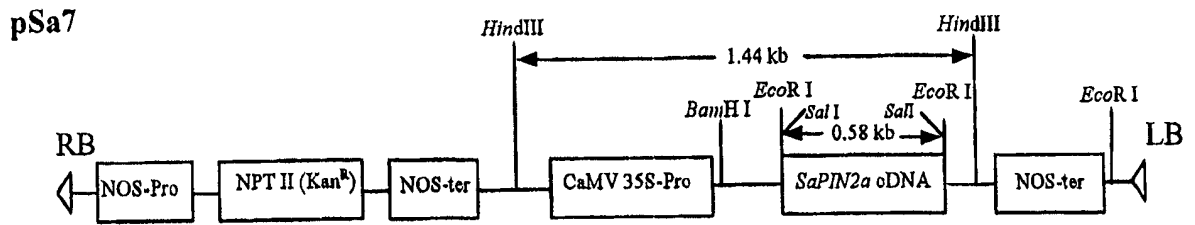


图 5

1	ATATTAGGTT	TTTACCTACC	CAGGAAAAGC	CAACCAACCT	CGATCTCTTG	TAGATCTGTT
61	CTCTAAACGA	ACTTTAAAAAT	CTGTGTAGCT	GTCGCTCGGC	TGCATGCCTA	GTGCACCTAC
121	GCAGTATAAA	CAATAATAAA	TTTTACTGTC	GTTGACAAGA	AACGAGTAAC	TCGTCCCTCT
181	TCTGCAGACT	GCTTACGGTT	TCGTCCGTGT	TGCAGTCGAT	CATCAGCATA	CCTAGGTTTC
241	GTCCGGGTGT	GACCGAAAGG	TAAGATGGAG	AGCCTTGTTT	TTGGTGTCAA	CGAGAAAACA
301	CACGTCCAAC	TCAGTTTGCC	TGTCCTTCAG	GTTAGAGACG	TGCTAGTGCG	TGGCTTCGGG
361	GACTCTGTGG	AAGAGGCCCT	ATCGGAGGCA	CGTGAACACC	TCAAAAATGG	CACTTGTGGT
421	CTAGTAGAGC	TGGAAAAGG	CGTACTGCC	CAGCTTGAAC	AGCCCTATGT	GTTCATTAAA
481	CGTTCTGATG	CCTTAAGCAC	CAATCACGGC	CACAAGGTCG	TTGAGCTGGT	TGCAGAAATG
541	GACGGCATT	AGTACGGTCG	TAGCGGTATA	ACACTGGGAG	TACTCGTGCC	ACATGTGGGC
601	GAAACCCCAA	TTGCATACCG	CAATGTTCTT	CTTCGTAAGA	ACGGTAATAA	GGGAGCCGGT
661	GGTCATAGCT	ATGGCATCGA	TCTAAAGTCT	TATGACTTAG	GTGACGAGCT	TGGCACTGAT
721	CCCATTTGAG	ATTATGAACA	AAACTGGAAC	ACTAAGCATG	GCAGTGGTGC	ACTCCGTGAA
781	CTCACTCGTG	AGCTCAATGG	AGGTGCAGTC	ACTCGCTATG	TCGACAACAA	TTTCTGTGGC
841	CCAGATGGGT	ACCCTCTTGA	TTGCATCAAA	GATTTTCTCG	CACGCGCGGG	CAAGTCAATG
901	TGCACTCTTT	CCGAACAACT	TGATTACATC	GAGTCGAAGA	GAGGTGTCTA	CTGCTGCCGT
961	GACCATGAGC	AAATTAAGAG	TGCCAAGAAA	TTTGACACTT	TCAAAGGGGA	ATGCCCAAAG
1021	ACACCCCTCG	AAATTAAGAG	AAAAGTCAAA	GTCATTCAAC	CACGTGTTGA	AAAGAAAAG
1081	TTTGTGTTTC	CTCTTAACTC	AAAAGTCAAA	GTCATTCAAC	CACGTGTTGA	AAAGAAAAG
1141	ACTGAGGGTT	TCATGGGGCG	TATACGCTCT	GTGTACCCTG	TTGCATCTCC	ACAGGAGTGT
1201	AACAATATGC	ACTTGTCTAC	CTTGATGAAA	TGTAATCATT	GCGATGAAGT	TTCATGGCAG
1261	ACGTGCGACT	TTCTGAAAGC	CACTTGTGAA	CATTGTGGCA	CTGAAAATTT	AGTTATTGAA
1321	GGACCTACTA	CATGTGGGTA	CCTACCTACT	AATGCTGTAG	TGAAAATGCC	ATGTCTGCC
1381	TGTCAAGACC	CAGAGATTGG	ACCTGAGCAT	AGTGTTCAG	ATTATCACAA	CCACTCAAAC
1441	ATTGAAACTC	GACTCCGCAA	GGGAGGTAGG	ACTAGATGTT	TTGGAGGCTG	TGTGTTTGCC
1501	TATGTTGGCT	GCTATAATAA	GCGTGCCTAC	TGGGTTCCCTC	GTGCTAGTGC	TGATATTGGC
1561	TCAGGCCATA	CTGGCATTAC	TGGTGACAAT	GTGGAGACCT	TGAATGAGGA	TCTCCTTGAG
1621	ATACTGAGTC	GTGAACGTGT	TAACATTAAC	ATTGTTGGCG	ATTTTCATTT	GAATGAAGAG
1681	GTTGCCATCA	TTTTGGCATT	TTTCTCTGCT	TCTACAAGTG	CCTTTATTGA	CACTATAAAG
1741	AGTCTTGATT	ACAAGTCTTT	CAAAACCATT	GTTGAGTCCT	GCGGTAACCTA	TAAAGTTACC
1801	AAGGGAAAAG	CCGTAAAAGG	TGCTTGGAAC	ATTGGACAAC	AGAGATCAGT	TTTAACACCA
1861	CTGTGTGGTT	TTCCCTCACA	GGCTGCTGGT	GTTATCAGAT	CAATTTTTGC	GCGCACACTT
1921	GATGCAGCAA	ACCACTCAAT	TCCTGATTTG	CAAAGAGCAG	CTGTCACCAT	ACTTGATGGT
1981	ATTTCTGAAC	AGTCATTACG	TCTTGTGAC	GCCATGGTTT	ATACTTCAGA	CCTGCTCACC
2041	AACAGTGTCA	TTATTATGGC	ATATGTAAC	GGTGGTCTTG	TACAACAGAC	TTCTCAGTGG
2101	TTGTCTAATC	TTTTGGGCAC	TACTGTTGAA	AAACTCAGGC	CTATCTTTGA	ATGGATTGAG
2161	GCGAAACTTA	GTGCAGGAGT	TGAATTTCTC	AAGGATGCTT	GGGAGATTCT	CAAATTTCTC
2221	ATTACAGGTG	TTTTTGACAT	CGTCAAGGGT	CAAATACAGG	TTGCTTCAGA	TAACATCAAG
2281	GATTGTGTAA	AATGCTTCAT	TGATGTTGTT	AACAAGGCAC	TCGAAAATGTG	CATTGATCAA
2341	GTCACATATC	CTGGCGCAA	GTTGCGATCA	CTCAACTTAG	GTGAAGTCTT	CATCGCTCAA
2401	AGCAAGGGAC	TTTACCCTCA	GTGTATACGT	GGCAAGGAGC	AGCTGCAACT	ACTCATGCCT
2461	CTTAAGGCAC	CAAAAGAAGT	AACCTTCTT	GAAGGTGATT	CACATGACAC	AGTACTTACC
2521	TCTGAGGAGG	TTGTTCTCAA	GAACGGTGAA	CTCGAAGCAC	TCGAGACGCC	CCTTGATAGC
2581	TTCACAAATG	GAGCTATCGT	CGGCACACCA	GTCTGTGTAA	ATGGCCTCAT	GCTCTTAGAG
2641	ATTAAGGACA	AAGAACAATA	CTGCGCATTG	TCTCCTGGTT	TACTGGCTAC	AAACAATGTC
2701	TTTCGCTTAA	AAGGGGGTGC	ACCAATTAAA	GGTGTAACTT	TTGGAGAAGA	TACTGTTTGG
2761	GAAGTTCAAG	GTTACAAGAA	TGTGAGAATC	ACATTTGAGC	TTGATGAACG	TGTTGACAAA
2821	GTGCTTAATG	AAAAGTGCTC	TGTCTACACT	GTTGAATCCG	GTACCGAAGT	TACTGAGTTT
2881	GCATGTGTTG	TAGCAGAGGC	TGTTGTGAAG	ACTTTACAAC	CAGTTTCTGA	TCTCCTTACC
2941	AACATGGGTA	TTGATCTTGA	TGAGTGGAGT	GTAGCTACAT	TCTACTTATT	TGATGATGCT
3001	GGTGAAGAAA	ACTTTTCATC	ACGTATGTAT	TGTTCCTTTT	ACCCTCCAGA	TGAGGAAGAA
3061	GAGGACGATG	CAGAGTGTGA	GGAAGAAGAA	ATTGATGAAA	CCTGTGAACA	TGAGTACGGT
3121	ACAGAGGATG	ATTATCAAGG	TCTCCCTCTG	GAATTTGGTG	CCTCAGCTGA	AACAGTTCGA
3181	GTTGAGGATG	AAGAAGAGGA	AGACTGGCTG	GATGATACTA	CTGAGCAATC	AGAGATTGAG
3241	CCAGAACCAG	AACCTACACC	TGAAGAACCA	GTTAATCAGT	TTACTGGTTA	TTTAAAACCT
3301	ACTGACAATG	TTGCCATTAA	ATGTGTTGAC	ATCGTTAAGG	AGGCACAAAG	TGCTAATCCT

图 6-A

3361	ATGGTGATTG	TAAATGCTGC	TACATACAC	CTGAAACATG	GTGGTGGTGT	AGCAGGTGCA
3421	CTCAACAAGG	CAACCAATGG	TGCCATGCAA	AAGGAGAGTG	ATGATTACAT	TAAGCTAAAT
3481	GGCCCTCTTA	CAGTAGGAGG	GTCTTGTTTG	CTTTCTGGAC	ATAATCTTGC	TAAGAAGTGT
3541	CTGCATGTTG	TTGGACCTAA	CCTAAATGCA	GGTGAGGACA	TCCAGCTTCT	TAAGGCAGCA
3601	TATGAAAATT	TCAATTCACA	GGACATCTTA	CTTGACCACAT	TGTTGTCCAGC	AGGCATATTT
3661	GGTGCTAAAC	CACTTCAGTC	TTTACAAGTG	TGCGTGCAGA	CGGTTCCGTAC	ACAGGTTTAT
3721	ATTGCACTCA	ATGACAAAGC	TCFTTATGAG	CAGGTTGTCA	TGGATTATCT	TGATAACCTG
3781	AAGCCTAGAG	TGGAAGCACC	TAAACAAGAG	GAGCCACCAA	ACACAGAAGA	TTCCAAAAC
3841	GAGGAGAAAT	CTGTCGTACA	GAAGCCTGTC	GATGTGAAGC	CAAAAATTAA	GGCCTGCATT
3901	GATTAGGTTA	CCACAACACT	GGAAGAAACT	AAGTTTCTTA	CCAATAAGTT	ACTCTGTTT
3961	GCTGATATCA	ATGGTAAGCT	TTACCATGAT	TCTCAGAACA	TGCTTAGAGG	TGAAGATATG
4021	TCTTCCCTTG	AGAAGGATGC	ACCTTACATG	GTAGGTGATG	TTATCACTAG	TGGTGATATC
4081	ACTTGTGTTG	TAATACCCTC	CAAAAAGGCT	GGTGGCACTA	CTGAGATGCT	CTCAAGAGCT
4141	TTGAAGAAAG	TGCCAGTTGA	TGAGTATATA	ACCACGTACC	CTGGACAAGG	ATGTGCTGGT
4201	TATACACTTG	AGGAAGCTAA	GACTGCTCTT	AAGAAATGCA	AATCTGCATT	TTATGTACTA
4261	CCTTCAGAAG	CACCTAATGC	TAAGGAAGAG	ATTCTAGGAA	CTGTATCCTG	GAATTTGAGA
4321	GAAATGCTTG	CTCATGCTGA	AGAGACAAGA	AAATTAATGC	CTATATGCAT	GGATGTTAGA
4381	GCCATAATGG	CAACCATCCA	ACGTAAGTAT	AAAGGAATTA	AAATTCAAGA	GGGCATCGTT
4441	GACTATGGTG	TCCGATCTT	CTTTTATACT	AGTAAAGAGC	CTGTAGCTTC	TATTATACG
4501	AAGCTGAACT	CTCTAAATGA	GCCGCTTGTC	ACAATGCCAA	TTGGTTATGT	GACACATGGT
4561	TTTAACTTTG	AAGAGGCTGC	GCGCTGTATG	CGTTCTCTTA	AAGCTCCTGC	CGTAGTGTCA
4621	GTATCATCAC	CAGATGCTGT	TACTACATAT	AATGGATACC	TCACTTCGTC	ATCAAGACA
4681	TCTGAGGAGC	ACTTTGTAGA	AACAGTTTCT	TTGGCTGGCT	CTTACAGAGA	TTGGTCCTAT
4741	TCAGGACAGC	GTACAGAGTT	AGGTGTTGAA	TTTCTTAAGC	GTGGTGACAA	AATTTGTATC
4801	CACACTCTGG	AGAGCCCCTG	CGAGTTTCAT	CTTGACGGTG	AGGTTCTTTC	ACTTGACAAA
4861	CTAAAGAGTC	TCTTATCCCT	GCGGGAGGTT	AAGACTATAA	AAGTGTTCAC	AACTGTGGAC
4921	AACACTAATC	TCCACACACA	GCTTGTGGAT	ATGTCTATGA	CATATGGACA	GCAGTTTGGT
4981	CCAACATACT	TGGATGGTGC	TGATGTTACA	AAAATTAAC	CTCATGTAAA	TCATGAGGGT
5041	AAGACTTTCT	TTGTACTACC	TAGTGATGAC	ACACTACGTA	GTGAAGCTTT	CGAGTACTAC
5101	CATACTCTTG	ATGAGAGTTT	TCTTGGTAGG	TACATGTCTG	CTTTAAACCA	CACAAAGAAA
5161	TGGAATTTTC	CTCAAGTTGG	TGTTTTAACT	TCAATTAAT	GGGCTGATAA	CAATTTGTTAT
5221	TTGTCTAGTG	TTTTATFAGC	ACTTCAACAG	CTTGAAGTCA	AATTCATGC	ACCAGCACTT
5281	CAAGAGGCTT	ATTATAGAGC	CCGTGCTGGT	GATGCTGCTA	ACTTTTGTGC	ACTCATCTC
5341	GCTTACAGTA	ATAAACTGT	TGCGGAGCTT	GGTGATGTCA	GAGAAACTAT	GACCCATCTT
5401	CTACAGCATG	CTAATTTGGA	ATCTGCAAAG	CGAGTTCTTA	ATGTGGTGTG	TAAACATTTG
5461	GGTCAGAAAA	CTACTACCTT	AACGGGTGTA	GAAGCTGTGA	TGTATATGG	TACTCTATCT
5521	TATGATAATC	TTAAGACAGG	TGTTTTCCATT	CCATGTGTGT	GTGGTCGTGA	TGCTACACAA
5581	TATCTAGTAC	AACAAGAGTC	TTCTTTTGT	ATGATGTCTG	CACCACCTGC	TGAGTATAAA
5641	TTACAGCAAG	GTACATTCTT	ATGTGCGAAT	GAGTACACTG	GTAACATATCA	GTGTGGTCAT
5701	TACACTCATA	TAACTGCTAA	GGAGACCCTC	TATCGTATTG	ACGGAGCTCA	CCTTACAAAG
5761	ATGTCAGAGT	ACAAAGGACC	AGTGACTGAT	GTTTTCTACA	AGGAAACATC	TTACTACTACA
5821	ACCATCAAGC	CTGTGTCGTA	TAAACTCGAT	GGAGTTACTT	ACACAGAGAT	TGAACCAAAA
5881	TTGGATGGGT	ATTATAAAAA	GGATAATGCT	TACTATACAG	AGCAGCCTAT	AGACCTTGTA
5941	CCAACCAAC	CATTACCAA	TGCGAGTTTT	GATAATTTCA	AACTCACATG	TTCTAACACA
6001	AAATTTGCTG	ATGATTTAAA	TCAAATGACA	GGCTTCACAA	AGCCAGCTTC	ACGAGAGCTA
6061	TCTGTACAT	TCTTCCCAGA	CTTGAATGGC	GATGTAGTGG	CTATTGACTA	TAGACACTAT
6121	TCAGCGAGTT	TCAAGAAAGG	TGCTAAATTA	CTGCATAAGC	CAATTGTTTG	GCACATTAAC
6181	CAGGCTACAA	CCAAGACAAC	GTTCAAACCA	AACACTTGGT	GTTTACGTTG	TCTTTGGAGT
6241	ACAAAGCCAG	TAGTACTTTC	AAATTCATTT	GAAGTTCTGG	CAGTAGAAGA	CACACAAGGA
6301	ATGGACAATC	TTGCTTGTGA	AAGTCAACAA	CCCACCTCTG	AAGAAGTAGT	GGAAAATCCT
6361	ACCATACAGA	AGGAAGTCAT	AGAGTGTGAC	GTGAAAACATA	CCGAAGTTGT	AGGCAATGTC
6421	ATACTTAAAC	CATCAGATGA	AGGTGTTAAA	GTAACACAAG	AGTTAGGTCA	TGAGGATCTT
6481	ATGGCTGCTT	ATGTGGAATA	CACAAGCATT	ACCATTAAGA	AACCTAATGA	GCTTCTACTA
6541	GCCTTAGGTT	TAAAAACAAT	TGCCACTCAT	GGTATTGCTG	CAATTAATAG	TGTTCCCTGG
6601	AGTAAAATTT	TGGCTTATGT	CAACCATTTC	TTAGGACAAG	CAGCAATTAC	AACATCAAAT
6661	TGCGCTAAGA	GATTAGCACA	ACGTGTGTTT	AACAATTATA	TGCCCTTATG	GTTTACATTA
6721	TTGTTCCAAT	TGTGTACTTT	TACTAAAAGT	ACCAATTCTA	GAATTAGAGC	TTCACTACCT
6781	ACAACATATTG	CTAAAAATAG	TGTTAAGAGT	GTTGCTAAAT	TATGTTTGGG	TGCCGGCATT

图 6-B

6841	AATTATGTGA	AGTCACCCAA	ATTTTCTAAA	TTGTTCCACAA	TCGCTATGTG	GCTATTGTGTG
6901	TTAAGTATTT	GCTTAGGTPC	TCTAATCTGT	GTAAGTCTGT	CTTTTGGTGT	ACTCTTATCT
6961	AATTTTGGTG	CTCCTTCTTA	TTGTAATGGC	GTTAGAGAAT	TGTATCTTAA	TTCGTCTAAC
7021	GTTACTACTA	TGGATTCTG	TGAAGGTTCT	TTTCCTTGCA	GCATTTGTTT	AAGTGGATTA
7081	GACTCCCTTG	ATTCTTATCC	AGCTCTTGAA	ACCATTCAGG	TGACGATTTT	ATCGTACAAG
7141	CTAGACTTGA	CAATTTTAGG	TCTGGCCGCT	GAGTGGGTTT	TGGCATATAT	GTTGTTTACA
7201	AAATTCCTTT	ATTTATTAGG	TCTTTCAGCT	ATAATGCAGG	TGTTCTTTGG	CTATTTTGCT
7261	AGTCATTTCA	TCAGCAATTC	TTGGCTCATG	TGGTTTATCA	TTAGTATTGT	ACAAATGGCA
7321	CCCGTTTCTG	CAATGGTTAG	GATGTACATC	TTCTTTGCTT	CTTCTACTA	CATATGGAAG
7381	AGCTATGTTT	ATATCATGGA	TGGTTGCACC	TCTTCGACTT	GCATGATGTG	CTATAAGCGC
7441	AATCGTGCCA	CACGCGTTGA	GTGTACAAC	ATTGTTAATG	GCATGAAGAG	ATCTTCTTAT
7501	GTCTATGCAA	ATGGAGGCCG	TGGCTTCTGC	AAGACTCACA	ATGGAATTG	TCTCAATTGT
7561	GACACATTTT	GCACGGTAG	TACATTCATT	AGTGATGAAG	TTGCTCGTGA	TTTGTCACTC
7621	CAGTTTAAAA	GACCAATCAA	CCCTACTGAC	CAGTCATCGT	ATATTGTTGA	TAGTGTGGCT
7681	GTGAAAAATG	GCGCGCTTCA	CCTCTACTTT	GACAAGGCTG	GTCAAAAGAC	CTATGAGAGA
7741	CATCCGCTCT	CCCATTTTGT	CAATTTAGAC	AATTTGAGAG	CTAACACAC	TAAAGGTTCA
7801	CTGCCTATTA	ATGTCATAGT	TTTTGATGGC	AAGTCCAAAT	GCGACGAGTC	TGCTTCTAAG
7861	TCTGCTTCTG	TGTACTACAG	TCAGCTGATG	TGCCAACCTA	TTCTGTTGCT	TGACCAAGCT
7921	CTTGATCAA	ACGTTGGAGA	TAGTACTGAA	GTTTCCGTTA	AGATGTTTGA	TGCTTATGTC
7981	GACACCTTTT	CAGCAACTTT	TAGTGTCTCT	ATGGAAAAAC	TTAAGGCAT	TGTTTGTACA
8041	GCTCACAGCG	AGTTAGCAAA	GGGTGTAGCT	TTAGATGGTG	TCCTTCTAC	ATTCGTGTCA
8101	GCTGCCCGAC	AAGGTGTTGT	TGATACCGAT	GTTGACACAA	AGGATGTTAT	TGAATGTCTC
8161	AAACTTTCAC	ATCACTCTGA	CTTAGAAGTG	ACAGGTGACA	GTTGTAACAA	TTTCATGCTC
8221	ACCTATAATA	AGGTTGAAAA	CATGACGCC	AGAGATCTTG	GCGCAATGAT	TGACTGTAAT
8281	GCAAGGCATA	TCAATGCCCA	AGTAGCAAAA	AGTCACAATG	TTTCACTCAT	CTGGAATGTA
8341	AAAGACTACA	TGTCTTTATC	TGAACAGCTG	CGTAAACAAA	TTCTGACTGC	TGCCAAGAAG
8401	AACAACATAC	CTTTTACACT	AACTTGTGCT	ACAAC TAGAC	AGGTTGTCAA	TGTCATAACT
8461	ACTAAAATCT	CACTCAAGGG	TGGTAAGATT	GTTAGTACTT	GTTTTAAACT	TATGCTTAAG
8521	GCCACATTAT	TGTGCGTTCT	TGCTGCATTG	GTTTGTATA	TCGTTATGCC	AGTACATACA
8581	TTGTCAATCC	ATGATGGTTA	CACAAATGAA	ATCATTGGTT	ACAAAGCCAT	TCAGGATGGT
8641	GTCACTCGTG	ACATCATTTT	TACTGATGAT	TGTTTTGCAA	ATAAACATGC	TGGTTTTGAC
8701	GCATGGTTTA	GCCAGCGTGG	TGGTTCATA	AAAAATGACA	AAAGCTGCC	TGTTGCTGAG
8761	GCTATCATT	CAAGAGAGAT	TGGTTTCATA	GTGCCTGGCT	TACCGGTAC	TGTGCTGAGA
8821	GCAATCAATG	GTGACTTCTT	GCATTTCTTA	OCTCGTGT	TTAGTGTCTG	TGGCAACATT
8881	GCTACACAC	CTTCCAAACT	CATTGAGTAT	AGTGATTTTG	CTACCTCTGC	TTGCGTCTT
8941	GCTGCTGAGT	GTACAATTTT	TAAGGATGCT	ATGGGCAAAC	CTGTGCCATA	TTGTTATGAC
9001	ACTAATTTGC	TAGAGGGTTC	TATTTCTTAT	AGTGAGCTTC	GTCCAGACAC	TCGTTATGTG
9061	CTTATGGATG	GTTCCATCAT	ACAGTTTCTT	AACACTTACC	TGGAGGGTTC	TGTTAGAGTA
9121	GTAAACAATT	TTGATGCTGA	GTAAGTGTGA	CATGGTACAT	GCGAAAGGTC	AGAAGTAGGT
9181	ATTTGCCTAT	CTACCAGTGG	TAGATGGGTT	CTTAATAATG	AGCATTACAG	AGCTCTATCA
9241	GGAGTTTCT	GTGGTGTGGA	TGCGATGAAT	CTCATAGCTA	ACATCTTAC	TCCTCTGTG
9301	CAACCTGTGG	GTGCTTTAGA	TGTGTCTGCT	TCAGTAGTGG	CTGGTGGTAT	TATTGCCATA
9361	TTGGTGACTT	GTGCTGCCTA	CTACTTTATG	AAATTCAGAC	GTGTTTTTGG	TAGTACAAC
9421	CATGTTGTTG	CTGCTAATGC	ACTTTTGT	TTGATGTCTT	TCACTATACT	CTGTCTGGTA
9481	CCAGCTTACA	GCTTCTGCC	GGGAGTCTAC	TCAGTCTTTT	ACTTGTACTT	GACATTCTAT
9541	TTACCAATG	ATGTTTCATT	CTTGGCTCAC	CTTCAATGGT	TTGCCATGTT	TTCTCTTATT
9601	GTGCCTTTTT	GGATAACAGC	AATCTATGTA	TTCTGTATTT	CTCTGAAGCA	CTGCCATTGG
9661	TTCTTTAACA	ACTATCTTAG	GAAAAGAGTC	ATGTTTAAATG	GAGTTACATT	TAGTACCTTC
9721	GAGGAGGCTG	CTTTGTGTAC	CTTTTGTCTC	AACAAGGAAA	TGTACCTAAA	ATTGCGTAGC
9781	GAGACTGT	TGCCACTTAC	ACAGTATAAC	AGGTATCTTG	CTCTATATAA	CAAGTACAAG
9841	TATTTAGTG	GAGCCTTAGA	TACTACCAGC	TATCGTGAAG	CAGCTTGCTG	CCACTTAGCA
9901	AAGGCTCTAA	ATGACTTTAG	CAACTCAGGT	GCTGATGTTT	TCTACCAACC	ACCACAGACA
9961	TCAATCACTT	CTGCTGTTCT	GCAGAGTGGT	TTAGGAAAA	TGGCATTTCC	GTCAGGCAAA
10021	GTTGAAGGGT	GCATGGTACA	AGTAACTGT	GGAAC TACA	CTCTTAATGG	ATTGTGGTTG
10081	GATGACACAG	TATACTGTCC	AAGACATGTC	ATTTGCACAG	CAGAAGACAT	GCTTAATCCT
10141	AACTATGAAG	ATCTGCTCAT	TGCGAAATCC	AACCATAGCT	TTCTTGTTC	GGCTGGCAAT
10201	GTTCAACTTC	GTGTTATTGG	CCATTCTATG	CAAAATTGTC	TGCTTAGGCT	TAAAGTTGAT
10261	ACTTCTAACC	CTAAGACACC	CAAGTATAAA	TTTGTCCGTA	TCCAACCTGG	TCAAACATTT

图 6-C

10321	TCAGTTCTAG	CATGCTACAA	TGGTTCACCA	TCTGGTGTTC	ATCAGTGTGC	CATGAGACCT
10381	AATCATACCA	TTAAAGGTTT	TTTCCTTAAT	GGATCATGTG	GTAGTGTGG	TTTTAACATT
10441	GATTATGATT	GCGTGTCTTT	CTGCTATATG	CATCATATGG	AGCTTCCAAC	AGGAGTACAC
10501	GCTGGTACTG	ACTTAGAAGG	TAAATTCAT	GGTCCATTTG	TTGACAGACA	AACTGCACAG
10561	GCTGCAGGTA	CAGACACAAC	CATAACATTA	AATGTTTTGG	CATGGCTGTA	TGCTGCTGTT
10621	ATCAATGGTG	ATAGGTGGTT	TCTTAATAGA	TTCAACACTA	CTTTGAATGA	CTTTAACCTT
10681	GTGGCAATGA	AGTACAACATA	TGAACCTTTG	ACACAAGATC	ATGTTGACAT	ATTGGGACCT
10741	CTTTCTGCTC	AAACAGGAAT	TGCCGCTTIA	GATATGTGTG	CTGCTTTGAA	AGAGCTGCTG
10801	CAGAATGGTA	TGAATGGTCC	TACTATCCTT	GGTAGCACTA	TTTTAGAAGA	TGAGTTTACA
10861	CCATTTGATG	TTGTTAGACA	ATGCTCTGGT	GTTACCTTCC	AAGGTAAGTT	CAAGAAAATT
10921	GTTAAGGGCA	CTCATCATTG	GATGCTTTTA	ACTTTCCTGA	CATCACTATT	GATCTTGTTT
10981	CAAAGTACAC	AGTGGTCACT	GTTTTCTTTT	GTTTACGAGA	ATGCTTTCCT	GCCATTTACT
11041	CTTGGTATTA	TGGCAATFGC	TGCACTGTCT	ATGCTGCTTG	TTAAGCATAA	GCACGCATTC
11101	TTGTGCTTGT	TTCTGTTACC	TTCTCTTGCA	ACAGTTGCTT	ACTTTAATAT	GGTCTACATG
11161	CCTGCTAGCT	GGGTGATGCG	TATCATGACA	TGGCTTGAAT	TGGCTGACAC	TAGCTTGTCT
11221	GGTTATAGGC	TTAAGGATTG	TGTTATGTAT	GCTTCAGCTT	TAGTTTTGCT	TATTCTCATG
11281	ACAGCTCGCA	CTGTTTATGA	TGATGCTGCT	AGACGTGTTT	GGACACTGAT	GAATGTCATT
11341	ACACTTGTTT	ACAAAGTCTA	CTATGTAAT	GCTTTAGATC	AAGCTATTTT	CATGTGGGCC
11401	TTAGTTATTT	CTGTAACCTC	TAACATTTCT	GGTGTCTGTTA	CGACTATCAT	GTTTTTAGCT
11461	AGAGCTATAG	TGTTTGTGTG	TGTTGAGTAT	TACCCATTGT	TATTTATTAC	TGGCAACACC
11521	TTACAGTSTA	TCATGCTTGT	TTATTTGTTT	TTAGGCTATT	GTGCTGCTG	CTACTTTGGC
11581	CTTTCTGTGT	TACTCAACCG	TTACTTCAGG	CTTACTCTTG	GTGTTTATGA	CTACTTGGTC
11641	TCTACACAAG	AATTTAGGTA	TATGAACCTC	CAGGGGCTTT	TGCCCTCCTAA	GAGTAGTATT
11701	GATGCTTTCA	AGCTTAACAT	TAAGTTGTTG	GGTATTGGAG	GTAACCATG	TATCAAGGTT
11761	GCTACTGTAC	AGTCTAAAAT	GTCTGACGTA	AAGTGCACAT	CTGTGGTACT	CTCTCGGTT
11821	CTTCAACAAC	TTAGAGTAGA	GTCATCTTCT	AAATTTGTGG	CACAATGTGT	ACAACCTCC
11881	AATGATATTC	TTCTTGCAAA	AGACACAAC	GAAGCTTTTC	AGAAGATGGT	TTCTCTTTTG
11941	TCTGTTTTGC	TATCCATGCA	GGGTGCTGTA	GACATTAATA	GGTTGTGCGA	GGAAATGCTC
12001	GATAACCGTG	CTACTCTTCA	GGCTATTGCT	TCAGAATTTA	GTTCTTTACC	ATCATATGCC
12061	GCTTATGCCA	CTGCCCAGGA	GGCCTATGAG	CAGGCTGTAG	CTAATGGTGA	TTCTGAAGTC
12121	GTTCTCAAAA	AGTTAAAGAA	ATCTTTGAAT	GTGGCTAAAT	CTGAGTTTGA	CCGTGATGCT
12181	GCCATGCAAC	GCAAGTTGGA	AAAGATGGCA	GATCAGGCTA	TGACCCAAAT	GTACAAACAG
12241	GCAAGATCTG	AGGACAAGAG	GGCAAAAGTA	ACTAGTGCTA	TGCCAAACAAT	GCTCTTCACT
12301	ATGCTTAGGA	AGCTTGATAA	TGATGCACCT	AACAACATTA	TCAACAATGC	GCGTGTGGT
12361	TGTGTTCCAC	TCAACATCAT	ACCATTGACT	ACAGCAGCCA	AACCTCATGG	TGTTGTCCCT
12421	GATTATGGTA	CCTACAAGAA	CACTTGTGAT	GGTAACACCT	TTACATATGC	ATCTGCACCT
12481	TGGGAAATCC	AGCAAGTTGT	TGATGCCGAT	AGCAAGATTG	TTCAACTTAG	TGAAATTAAC
12541	ATGGACAATT	CACCAAATTT	GGCTTGGCCT	CTTATTTGTTA	CAGCTTAAG	AGCCAACTCA
12601	GCTGTTAAAC	TACAGAATAA	TGAACCTAGT	CCAGTAGCAC	TACGACAGAT	GTCTGTGCG
12661	GCTGGTACCA	CACAAACAGC	TTGTACTGAT	GACAAATGCAC	TTGCCTACTA	TAACAATTCG
12721	AAGGGAGGTA	GGTTTGTGCT	GGCATTACTA	TCAGACCACC	AAGATCTCAA	ATGGGCTAGA
12781	TTCCCTAAGA	GTGATGGTAC	AGGTACAATT	TACACAGAAC	TGGAACCACC	TTGTAGGTTT
12641	GTTACAGACA	CACCAAAGG	GCCTAAAGTG	AAATACTTGT	ACTTCATCAA	AGGCTTAAAC
12901	AACCTAAATA	GAGGTATGGT	GCTGGGCAGT	TTAGCTGCTA	CAGTACGTCT	TCAGGCTGGA
12961	AATGCTACAG	AAGTACCTGC	CAATTCAACT	GTGCTTTCCT	TCGTGCTTT	TGCAGTAGAC
13021	CCTGCTAAAG	CATATAAGGA	TTACCTAGCA	AGTGGAGGAC	AACCAATCAC	CAACTGTGTG
13081	AAGATGTTGT	GTACACACAC	TGGTACAGGA	CAGGCAATTA	CTGTAACACT	AGAAGCTAAC
13141	ATGGACCAAG	AGTCCTTTGG	TGGTGCCTCA	TGTTGTCTGT	ATGTTAGATG	CCACATTGAC
13201	CATCCAAATC	CTAAAGGATT	CTGTGACTTG	AAAGGTAAGT	ACGTCCAAAT	ACCTACCACT
13261	TGTGCTAATG	ACCCAGTGGG	TTTTACACTT	AGAAACACAG	TCGTACCCGT	CTGCGGAATG
13321	TGGAAAGGTT	ATGGCTGTAG	TTGTGACCAA	CTCCGCGAAC	CCTTGATGCA	GTCTGCGGAT
13381	GCATCAACGT	TTTTAAACGG	GTTTGCGGTG	TAAGTGCAGC	CCGTCTTACA	CCGTGCGGCA
13441	CAGGCACTAG	TACTGATGTC	GTCTACAGGG	CTTTTGATAT	TTACAACGAA	AAAAGTGCTG
13501	GTTTTGCAAA	GTTTCTAAAA	ACTAATTGCT	GTGCTTCCA	GGAGAAGGAT	GAGGAAGGCA
13561	ATTATATTAGA	CTCTTACTTT	GTAGTTAAGA	GGCATACTAT	GTCTAACTAC	CAACATGAAG
13621	AGACTATTTA	TAACTTGGTT	AAAGATTGTC	CAGCGGTTGC	TGTCCATGAC	TTTTTCAAGT
13681	TTAGAGTAGA	TGGTGACATG	GTACCACATA	TATCACGTCA	GCGTCTAACC	AAATACACAA
13741	TGGCTGATTT	AGTCTATGCT	CTACGTCATT	TTGATGAGGG	TAATTTGTGAT	ACATTAAGG

图 6-D



13801	AAATACTCGT	CACATACAAT	TGCTGTGATG	ATGATTATTT	CAATAAGAAG	GATTGGTATG
13861	ACTTCGTAGA	GAATCCTGAC	ATCTTACGCG	TATATGCTAA	CTTAGGTGAG	CGTGTACGCC
13921	AATCATTATT	AAAGACTGTA	CAATTCGTGG	ATGCTATGCG	TGATGCAGGC	ATTGTAGGCG
13981	TACTGACATT	AGATAATCAG	GATCTTAATG	GGAAGTGGTA	CGATTTGCGT	GATTTTCGTAC
14041	AAGTAGGACC	AGGCTGCGGA	GTTCCATTG	TGGATTGATA	TTACTCATTG	CTGATGCCCA
14101	TCCTCACTTT	GACTAGGGCA	TTGGCTGCTG	AGTCCCATAT	GGATGCTGAT	CTCGCAAAAC
14161	CACCTATTAA	GTGGGATTTG	CTGAAATATG	ATTTTACGGA	AGAGAGACTT	TGTCTCTTCG
14221	ACCGTTATTT	TAAATATTGG	GACCAGACAT	ACCATCCCAA	TTGTATTAAC	TGTTTGGATG
14281	ATAGGTGTAT	CCTTCATTGT	GCAAACTTTA	ATGTGTATT	TTCTACTGTG	TTTCCACCTA
14341	CAAGTTTGG	ACCACTAGTA	AGAAAAATAT	TTGTAGATGG	TGTTCCTTTT	GTTGTTTCAA
14401	CTGGATACCA	TTTTCGTGAG	TTAGGAGTCG	TACATAATCA	GGATGTAAAC	TTACATAGCT
14461	CGCGTCTCAG	TTTCAAGGAA	CTTTTAGTGT	ATGCTGCTGA	TCCAGCTATG	CATGCAGCTT
14521	CTGGCAATTT	ATTGCTAGAT	AAACGCCTA	CATGCTTTTC	AGTAGCTGCA	CTAACAAACA
14581	ATGTTGCTTT	TCAAACCTGTC	AAACCCGGTA	ATTTTAATAA	AGACTTTTAT	GACTTTGCTG
14641	TGTCTAAAG	TTTCTTTAAG	GAAGGAAGTT	CTGTTGAACT	AAAACACTTC	TTCTTTGCTC
14701	AGGATGGCAA	CGCTGCTATC	AGTGATTATG	ACTATTATCG	TTATAACTCTG	CCAACAATGT
14761	GTGATATCAG	ACAACTCCTA	TTTCGTAGTTG	AAGTTGTTGA	TAAATACTTT	GATTGTTACG
14821	ATGGTGGCTG	TATTAATGCC	AACCAAGTAA	TCGTTAACAA	TCTGGATAAA	TCAGCTGGTT
14861	TCCCATTTAA	TAAATGGGGT	AAGGCTAGAC	TTTATTATGA	CTCAATGAGT	TATGAGGATC
14941	AAGATGCAC	TTTCGCGTAT	ACTAAGCGTA	ATGTCATCCC	TACTATAACT	CAAATGAATC
15001	TTAAGTATGC	CATTAGTGCA	AAGAATAGAG	CTCGCACCGT	AGCTGGTGTC	TCTATCTGTA
15061	GTACTATGAC	AAATAGACAG	TTTCATCAGA	AATTATTGAA	GTCAATAGCC	GCCACTAGAG
15121	GAGCTACTGT	GGTAATTGGA	ACAAGCAAGT	TTTACGGTGG	CTGGCATAAT	ATGTTAAAAA
15181	CTGTTACAG	TGATGTAGAA	ACTCCACACC	TTATGGGTGG	GGATTATCCA	AAATGTGACA
15241	GAGCCATGCC	TAACATGCTT	AGGATAATGG	CCTCTCTTGT	TCTTGCTCGC	AAACATAACA
15301	CTTGCTGTAA	CTTATCACAC	CGTTTCTACA	GGTTAGCTAA	CGAGTGTGCG	CAAGTATTAA
15361	GTGAGATGGT	CATGTGTGGC	GGCTCACTAT	ATGTTAAACC	AGGTGGAACA	TCATCCGGTG
15421	ATGCTACAAC	TGCTTATGCT	AATAGTGTCT	TTAACATTTG	TCAAGCTGTT	ACAGCCAATG
15481	TAAATGCAC	TCTTTCAACT	GATGGTAATA	AGATAGCTGA	CAAGTATGTC	CGCAACTTAC
15541	AACACAGGCT	CTATGAGTGT	CTCTATAGAA	ATAGGGATGT	TGATCATGAA	TTTCGTGGATG
15601	AGTTTACGC	TTACCTGCGT	AAACATTTCT	CCATGATGAT	TCTTTCGAT	GATGCCGTTG
15661	TGTGCTATAA	CAGTAACTAT	GCGGCTCAAG	GTTTAGTAGC	TAGCATTAAAG	AACTTTAAGG
15721	CAGTCTTTA	TTATCAAAAT	AATGTGTTCA	TGTCTGAGGC	AAAATGTTGG	ACTGAGACTG
15781	ACCTTACTAA	AGGACCTCAC	GAATTTTGCT	CACAGCATA	AATGCTAGTT	AAACAAGGAG
15841	ATGATTACGT	GTACCTGCCT	TACCCAGATC	CATCAAGAAT	ATTAGGCGCA	GGCTGTTTTG
15901	TCGATGATAT	TGTCAAAACA	GATGGTACAC	TTATGATTGA	AAGGTTGCGT	TCACTGGCTA
15961	TTGATGCTTA	CCCACTTACA	AAACATCCTA	ATCAGGAGTA	TGCTGATGTC	TTTCACTTGT
16021	ATTTACAATA	CATTAGAAAAG	TTACATGATG	AGCTTACTGG	CCACATGTTG	GACATGTATT
16081	CGTAATGCT	AACTAATGAT	AACACCTCAC	GGTACTGGGA	ACCTGAGTTT	TATGAGGCTA
16141	TGTACACACC	ACATACAGTC	TTGCAGGCTG	TAGGTGCTTG	TGTATTGTGC	AATTCACAGA
16201	CTTCACTTCG	TTGCGGTGCC	TGTATTAGGA	GACCATTCCCT	ATGTTGCAAG	TGCTGCTATG
16261	ACCATGTCTAT	TTCAACATCA	CACAAAATTAG	TGTTGTCTGT	TAATCCCTAT	GTTTGCATATG
16321	CCCCAGTTG	TGATGTCACT	GATGTGACAC	AACTGTATCT	AGGAGGTATG	AGCTATTATT
16381	GCAAGTCACA	TAAGCCTCCC	ATTAGTTTTT	CATTATGTGC	TAATGGTCAG	GTTTTTGGTT
16441	TATACAAAAA	CACATGTGTA	GGCAGTGACA	ATGTCACTGA	CTTCAATGCG	ATAGCAACAT
16501	GTGATTGGAC	TAATGCTGGC	GATTACATAC	TTGCCAACAC	TTGTAATGAG	AGACTCAAGC
16561	TTTTCGCAGC	AGAAACGCTC	AAAGCCACTG	AGGAAACATT	TAAGCTGTCA	TATGGTATTG
16621	CCACTGTACG	CGAAGTACTC	TCTGACAGAG	AATTGCATCT	TTTATGGGAG	GTTGGAAAAC
16661	CTAGACCACC	ATTGAACAGA	AACTATGTCT	TTACTGGTTA	CCGTGTAAC	AAAAATAGTA
16741	AAGTACAGAT	TGGAGAGTAC	ACCTTTGAAA	AAGGTGACTA	TGGTGTATGCT	GTTGTGTACA
16801	GAGGTACTAC	GACATACAAG	TTGAATGTTG	GTGATTACTT	TGTGTTGACA	TCTCACACTG
16861	TAAATGCCACT	TAGTGCACCT	ACTTAGTGC	CACAAGAGCA	CTATGTGAGA	ATTACTGGCT
16921	TGTACCCAAC	ACTCAACATC	TCAGATGAGT	TTTCTAGCAA	TGTTGCAAA	TATCAAAAGG
16981	TCGGCATGCA	AAAGTACTCT	ACACTCCAAG	GACCACCTGG	TACTGGTAA	AGTCATTTTG
17041	CCATCGGACT	TGCTCTCTAT	TACCCATCTG	CTCGCATAGT	GTATACGGCA	TGCTCTCATG
17101	CAGCTGTTGA	TGCCCTATGT	GAAAAGGCAT	TAAAATATTT	GCCCATAGAT	AAATGTAGTA
17161	GAATCATACC	TGCGCTGCG	CGCGTAGAGT	GTTTTGATAA	ATTCAAAGTG	AATTCACAC
17221	TAGAACAGTA	TGTTTTCTGC	ACTGTAAATG	CATTGCCAGA	AACAACTGCT	GACATTGTAG

图 6-E

17281	TCTTTGATGA	AATCTCTATG	GCTACTAATT	ATGACTTGAG	TGTTGTCAAT	GCTAGACTTC
17341	GTGAAAACA	CTACGTCTAT	ATTGGCGATC	CTGCTCAATT	ACCAGCCCCC	CGCACATTGC
17401	TGACTAAAGG	CACACTAGAA	CCAGAATATT	TTAATTCAGT	GTGCAGACTT	ATGAAAACAA
17461	TAGGTCCAGA	CATGTTCCCT	GGAACTTGTC	GCCGTGTGCC	TGCTGAAATT	GTTGACACTG
17521	TGAGTGCTTT	AGTTTATGAC	AATAAGCTAA	AAGCACACAA	GGATAAGTCA	GCTCAATGCT
17561	TCAAAATGTT	CTACAAAGGT	GTTATTACAC	ATGATGTTTC	ATCTGCAATC	AACAGACCTC
17641	AAATAGGCGT	TGTAAGAGAA	TTTCTTACAC	GCAATCCTGC	TTGGAGAAAA	GCTGTTTTTA
17701	TCTCACCTTA	TAATTCACAG	AACGCTGTAG	CTTCAAAAAA	CTTAGGATTG	CCTACGCAGA
17761	CTGTTGATTC	ATCACAGGGT	TCTGAATATG	ACTATGTCAT	ATTCACACAA	ACTACTGAAA
17821	CAGCACACTC	TTGTAATGTC	AACCGCTTCA	ATGTGGCTAT	CACAAGGGCA	AAAATTGGCA
17881	TTTTGTGCAT	AATGTCTGAT	AGAGATCTTT	ATGACAAACT	GCAATTTACA	AGCTAGAAA
17941	TACCACGTCG	CAATGTGGCT	ACATTACAAG	CAGAAAATGT	AACTGGACTT	TTTAAGGACT
18001	GTAGTAAGAT	CATTACTGGT	CTTCATCCTA	CACAGGCACC	TACACACCTC	AGCGTTGATA
18061	TAAAAATCAA	GACTGAAGGA	TTATGTSTTG	ACATACCAGG	CATACCAAAG	GACATGACCT
18121	ACCGTAGACT	CATCTCTATG	ATGGGTTTCA	AAATGAATTA	CCAAGTCAAT	GGTTACCCTA
18181	ATATGTTTAT	CACCCGCGAA	GAAGCTATTC	GTCACGTTTC	TGCGTGGATT	GGCTTTGATG
18241	TAGAGGGCTG	TCATGCAACT	AGAGATGCTG	TGGTACTAA	CCTACCTCTC	CAGCTAGGAT
18301	TTTCTACAGG	TGTTAACTTA	GTAGCTGTAC	CGACTGGTTA	TGTTGACACT	GAAAATAACA
18361	CAGAATTCAC	CAGAGTTAAT	GCAAAACCTC	CACCAGGTGA	CCAGTTAAA	CATCTTATAC
18421	CACTCATGTA	TAAAGGCTTG	CCCTGGAATG	TAGTGCCTAT	TAAGATAGTA	CAAATGCTCA
18481	GTGATACACT	GAAAGGATTG	TCAGACAGAG	TCGTGTTCTG	CCTTTGGGCG	CATGGCTTTG
18541	AGCTTACATC	AATGAAGTAC	TTTGTCAAGA	TTGGACCTGA	AAGAACGTGT	TGTCTGTGTG
18601	ACAAAACGTGC	AACTTGCTTT	TCTACTTCAT	CAGATACTTA	TGCCCTGCTG	AATCATTCTG
18651	TGGGTTTTGA	CTATGTCTAT	AACCCATTTA	TGATTGATGT	TCAGCAGTGG	GGCTTTACGG
18721	GTAACCTTCA	GAGTAACCAT	GACCAACATT	GCCAGGTACA	TGGAATGCA	CATGTGGCTA
18781	GTTGTGATGC	TATCATGACT	AGATGTTTAG	CAGTCCATGA	GTGCTTGTG	AAGCGCGTTG
18841	ATTGGTCTGT	TGAATACCCT	ATTATAGGAG	ATGAACTGAG	GGTTAATTCT	GCTTGCAGAA
18901	AAGTACAACA	CATGGTTGTG	AAGTCTGCAT	TGCTTGCTGA	TAAGTTTCCA	GTTCTTCATG
18961	ACATTGGAAA	TCCAAAGGCT	ATCAAGTGTG	TGCCCTCAGG	TGAAGTAGAA	TGGAAGTTCT
19021	ACGATGCTCA	GCCATGTAGT	GACAAAGCTT	ACAAAATAGA	GGAACTCTTC	TATCTTTATG
19081	CTACACATCA	CGATAAATTC	ACTGATGGTG	TTTGTGTTGT	TTGGAAATGT	AACGTTGATC
19141	GTTACCCAGC	CAATGCAATT	GTGTGTAGGT	TTGACACAAG	AGTCTTGTCA	AACTTGAAC
19201	TACCAGGCTG	TGATGGTGGT	AGTTTGTATG	TGAATAAGCA	TGCATTCCAC	ACTCCAGCTT
19261	TCGATAAAAG	TGCATTTACT	AATTTAAAGC	AATTGCCTTT	CTTTTACTAT	TCTGATAGTC
19321	CTTGTGAGTC	TCATGGCAAA	CAAGTAGTGT	CGGATATTGA	TTATGTTCCA	CTCAAATCTG
19381	CTACGTGTAT	TACACGATGC	AATTTAGGTG	GTGCTGTTTG	CAGACACCAT	GCAAATGAGT
19441	ACCACAGTA	CTTGGATGCA	TATAATATGA	TGATTTCTGC	TGGATTTAGC	CTATGGATTT
19501	ACAAACAATT	TGATACTTAT	AACCTGTGGA	ATACATTTAC	CAGGTTACAG	AGTTTAGAAA
19561	ATGTGGCTTA	TAATGTTGTT	AATAAAGGAC	ACTTTGATGG	ACACGCCGGC	GAAGCACCTG
19621	TTTCCATCAT	TAATAATGCT	GTTTACACAA	AGGTAGATGG	TATTGATGTG	GAGATCTTTG
19681	AAAATAAGAC	AACACTTCCT	GTTAATGTTG	CATTTGAGCT	TTGGGCTAAG	CGTAACATTA
19741	AACCACTGCC	AGAGATTAAG	ATACTCAATA	ATTTGGGTGT	TGATATCGCT	GCTAATACTG
19801	TAATCTGGGA	CTACAAAAGA	GAAGCCCCAG	CACATGTATC	TACAATAGGT	GTCTGCACAA
19861	TGACTGACAT	TGCCAAGAAA	CCTACTGAGA	GTGCTTGTTC	TTCACTTACT	GTCTTGTTTG
19921	ATGGTAGAGT	GGAAGGACAG	GTAGACCTTT	TTAGAAAACGC	CCGTAATGGT	GTTTAAATAA
19931	CAGAAGTTC	AGTCAAAGGT	CTAACACCTT	CAAAGGGACC	AGCACAGCT	AGCGTCAATG
20041	GAGTCACATT	AATTGGAGAA	TCAGTAAAAA	CACAGTTTAA	CTACTTTAAG	AAAGTAGACG
20101	GCATTATTC	ACAGTTGCCT	GAAACCTACT	TTACTCAGAG	CAGAGACTTA	GAGGATTTTA
20161	AGCCCAGATC	ACAAATGGAA	ACTGACTTTC	TCGAGCTCGC	TATGGATGAA	TTCATACAGC
20221	GATATAAGCT	CGAGGGCTAT	GCCTTCGAAC	ACATCGTTTA	TGGAGATTTC	AGTCATGGAC
20281	AACFTGGCGG	TCTTCATTTA	ATGATAGGCT	TAGCCAAGCG	CTCACAAGAT	TCACCACTTA
20341	AATTAGAGGA	TTTTATCCCT	ATGGACAGCA	CAGTGAAAAA	TTACTTCATA	ACAGATGCCG
20401	AAACAGGTTT	ATCAAATGT	GTGTGTTCTG	TGATTGATCT	TTACTTGTAT	GACTTTGTCT
20461	AGATAATAAA	GTCACAAGAT	TTGTCAAGTA	TTTCAAAAGT	GGTCAAGGTT	ACAATTGACT
20521	ATGCTGAAAT	TTCATTCATG	CTTTGGTGTA	AGGATGGACA	TGTTGAAACC	TTCTACCCAA
20581	AACTACAAGC	AAGTCAAGCG	TGGCAACCAG	GTGTGCGAT	GCCTAACTTG	TACAAGATGC
20641	AAAGATPGCT	TCTTGAAGAG	TGTGACCTTC	AGAATTATGG	TGAAAATGCT	GTATATCCAA
20701	AAGGAATAAT	GATGAATGTC	GCAAAGTATA	CTCAACTGTG	TCAACTTTA	AATACACTTA

图 6-F

20761	CTTTAGCTGT	ACCCTACAAC	ATGAGAGTTA	TTCAC TTTGG	TGCTGGCTCT	GATAAAGGAG
20821	TTGCACCAGG	TACAGCTGTG	CTCAGACAAT	GGTGGCCAAC	TGGCACACTA	CTTCTCGATT
20881	CAGATCTTAA	TGACTTCGTC	TCCGACGCAG	ATTCTACTTT	AATTGGAGAC	TGTGCAACAG
20941	TACATACGGC	TAATAAATGG	GACCTTATTA	TTAGCGATAT	GTATGACCCT	AGGACCAAAC
21001	ATGTGACAAA	AGAGAATGAC	TCTAAAGAAG	GGTTTTTCAC	TTATCTGTGT	GGATTATAAA
21061	AGCAAAAAC	AGCCCTGGGT	GGTTCATAG	CTGTAAAGAT	AACAGAGCAT	TCTTGGAAATG
21121	CTGACCTTTA	CAAGCTTATG	GGCCATTTCT	CATGGTGGAC	AGCTTTTGT	ACAAATGTAA
21181	ATGCATCATC	ATCGGAAGCA	TTTTTAATTG	GGGCTAACTA	TCTTGGCAAG	CCGAAGGAAC
21241	AAATTGATGG	CTATACCATG	CATGCTAACT	ACATTTTCTG	GAGGAACACA	AATCCTATCC
21301	AGTTGTCTTC	CTATTCACTC	TTTGACATGA	GCAAAATTTCC	TCTTAAATTA	AGAGGAACTG
21361	CTGTAATGTC	TCTTAAGGAG	AATCAAATCA	ATGATATGAT	TTATTCTCTT	CTGGAAAAAG
21421	TAGGCTTAT	CATTAGAGAA	AACAACAGAG	TTGTGGTTTC	AAGTGATATT	CTTGTTAACA
21481	ACTAAACGAA	CATGTTTATT	TTCTTATTAT	TTCTTACTCT	CACTAGTGGT	AGTGACCTTG
21541	ACCGGTGCAC	CACTTTTGAT	GATGTTCAAG	CTCCTAATTA	CACTCAACAT	ACTTCATCTA
21601	TGAGGGGGGT	TTACTATCCT	GATGAAATTT	TTAGATCAGA	CACTCTTAT	TTAACTCAGG
21661	ATTTATTTCT	TCCATTTTAT	TCTAATGTTA	CAGGGTTTCA	TACTATTAAT	CATACCTTTG
21721	GCAACCCCTG	CATACCTTTT	AAGGATGGTA	TTTATTTTGC	TGCCACAGAG	AAATCAAATG
21781	TTGTCCGTGG	TTGGGTTTTT	GGTTCACCA	TGAACAACAA	GTACAGTCCG	GTGATATTA
21841	TTAAACAATTC	TACTAATGTT	GTTATACGAG	CATGTAACCT	TGAATTGTGT	GACAACCTTT
21901	TCTTTGCTGT	TTCTAAACCC	ATGGSTACAC	AGACACATAC	TATGATATTC	GATAATGCAT
21951	TTAATTGCAC	TTTCGAGTAC	ATATCTGATG	CCTTTTCGCT	TGATGTTTCA	GAAAAGTCAG
22021	GTAATTTTAA	ACACTTACGA	GAGTTTGTGT	TAAAAATAA	AGATGGGTTT	CTCTATGTTT
22081	ATAAGGGCTA	TCAACCTATA	GATGTAGTTC	GTGATCTACC	TTCTGGTTTT	AACACTTTGA
22141	AACCTATTTT	TAAGTTGCCT	CTTGGTATTA	ACATTACAAA	TTTTAGAGCC	ATTCTTACAG
22201	CCTTTTCACC	TGCTCAAGAC	ATTTGGGSCA	CGTCAGCTGC	AGCCTATTTT	GTTGGCTATT
22261	TAAAGCCAAC	TACATTTATG	CTCAAGTATG	ATGAAAATGG	TACAATCACA	GATGCTGTTG
22321	ATTGTTCTCA	AAATCCACTT	GCTGAACTCA	AATGCTCTGT	TAAGAGCTTT	GAGATTGACA
22381	AAGGAATTTA	CCAGACCTCT	AATTTCAGGG	TTGTTCCCTC	AGGAGATGTT	GTGAGATTCC
22441	CTAATATTAC	AAACTTGTGT	CCTTTTGGAG	AGGTTTTTAA	TOCTACTAAA	TTCCCTTCTG
22501	TCTATGCATG	GGAGAGAAAA	AAAATTTCTA	ATTGTGTTGC	TGATTACTCT	GTGCTCTACA
22561	ACTCAACATT	TTTTTCAACC	TTTAAAGTGT	ATGGCGTTTC	TGCCACTAAG	TTGAATGATC
22621	TTTGCTTCTC	CAATGTCTAT	GCAGATTCTT	TTGTAGTCAA	GGGAGATGAT	GTAAGACAAA
22681	TAGCCGCTAG	ACAAACTGGT	GTTATTGCTG	ATTATAATTA	TAAATTGCCA	TAAATTCTCA
22741	TTGGTTGTGT	CCTTGCTTGG	AATACTAGGA	ACATTTGATC	TACTTCAACT	GGTAATTATA
22801	ATTATAAATA	TAGGTATCTT	AGACATGGCA	AGCTTAGGCC	CTTTGAGAGA	GACATATCTA
22861	ATGTGCCTTT	CTCCCTGAT	GGCAAACCTT	GCACCCACC	TGCTCTTAAT	TGTTATTGGC
22921	CATTAAATGA	TTATGTTTTT	TACACCACTA	CTGGCATTGG	CTACCAACCT	TACAGAGTTG
22981	TAGTACTTTC	TTTTGAACCT	TTAAATGCAC	CGGCCACGGT	TTGTGGACCA	AAATTTATCC
23041	CTGACCTTAT	TAAGAACCAG	TGTGTCAATT	TTAATTTTAA	TGGACTCACT	GGTACTGGTG
23101	TGTTAACTCC	TTCTTCAAAG	AGATTTCAAC	CATTTCAACA	ATTTGGCCGT	GATGTTTCTG
23161	ATTTCACTGA	TTCCGTTTGA	GATCCTAAAA	CATCTGAAAT	ATTAGACATT	TCACCTTGCT
23221	CTTTTGGGGG	TGTAAGTGT	ATTACACCTG	GAACAATATG	TTCATCTGAA	GTTGCTGTTT
23281	TATATCAAGA	TGTTAACTGC	ACTGATGTTT	CTACAGCAAT	TCATGCAGAT	CAACTCACAC
23341	CAGCTTGGCG	CATATATTCT	ACTGGAACA	ATGTATTCCA	GACTCAAGCA	GGCTGTCTTA
23401	TAGGAGCTGA	GCATGTCGAC	ACTTCTTATG	AGTGGACAT	TCCTATTGGA	GCTGGCATT
23461	GTGCTAGTTA	CCATACAGTT	TCTTTATTAC	GTAGTACTAG	CCAAAAATCT	ATTGTGGCTT
23521	ATACTATGTC	TTTAGGTGCT	GATAGTTCAA	TTGCTTACTC	TAATAACACC	ATTGCTATAC
23581	CTACTAACTT	TTCAATTAGC	ATTACTACAG	AAGTAATGCC	TGTTTCTATG	GCTAAAACCT
23641	CCGTAGATTG	TAATATGTAC	ATCTGCGGAG	ATTCTACTGA	ATGTGCTAAT	TTGCTTCTCC
23701	AATATGGTAG	CTTTTGCACA	CAACTAAATC	GTGCACTCTC	AGGTATTGCT	GCTGAACAGG
23761	ATCGCAACAC	ACGTGAAGTG	TTCCGCTCAAG	TCAAACAAAT	GTACAAAACC	CCAACCTTGA
23821	AATATTTTGG	TGTTTTTAAT	TTTTTCAAAA	TATTACCTGA	CCCTCTAAG	CCAACCTAAGA
23881	GGTCTTTTAT	TGAGGACTTG	CTCTTTAATA	AGGTGACACT	CGCTGATGCT	GGCTTCATGA
23941	AGCAATATGG	CGAATGCCTA	GGTGATATTA	ATGCTAGAGA	TCTCATTTGT	GCGCAGAAGT
23001	TCAATGGACT	TACAGTGTG	CCACCTCTGC	TCACTGATGA	TATGATTGCT	GCCTACACTG
24061	CTGCTCTAGT	TAGTGGTACT	GCCACTGCTG	GATGGACATT	TGGTGTGCTG	GCTGCTCTTC
24121	AAATACCTTT	TGCTATGCAA	ATGGCATATA	GGTTCATATG	CATTTGGAGTT	ACCCAAATG
24181	TTCTCTATGA	GAACCAAAAA	CAAAATCGCCA	ACCAATTTAA	CAAGGCGATT	AGTCAAATTC

图

6-G

24241	AAGAATCACT	TACAACAACA	TCAACTGCAT	TGGGCAAGCT	GCAAGACGTT	GTTAACCAGA
24301	ATGCTCAAGC	ATTAAACACA	CTTGTAAAC	AACCTAGCTC	TAATTTGGT	GCAATTTCAA
24361	GTGTGCTAAA	TGATATCCTT	TCGCGACTTG	ATAAAGTCGA	GGCGGAGGTA	CAAATTTGACA
24421	GGTTAATTAC	AGGCAGACTT	CAAAGCCTTC	AAACCTATGT	AACACAACAA	CTAATCAGGG
24481	CTGCTGAAAT	CAGGGCTTCT	GCTAATCTTG	CTGCTACTAA	AATGTCTGAG	TGTGTTCTTG
24541	GACAATCAA	AAGAGTTGAC	TTTTGTGGAA	AGGGCTACCA	CCTTATGTCC	TTCCACAAG
24601	CAGCCCCGA	TGGTGTGTC	TCCTACATG	TCAGTATGT	GCCATCCCAG	GAGAGGAAT
24661	TCACCACAGC	GCCAGCAATT	TGTCATGAAG	GCAAAGCATA	CTTCCCTCGT	GAAGGTGTTT
24721	TTGTGTTTAA	TGGCACTTCT	TGGTTTATTA	CACAGAGGAA	CTTCTTTTCT	CCACAAATAA
24781	TTACTACAGA	CAATACATTT	GTCTCAGGAA	ATTGTGATGT	CGTTATTGGC	ATCATTAAACA
24841	ACACAGTTTA	TGATCCTCTG	CAACCTGAGC	TTGACTCATT	CAAAGAAGAG	CTGGACAAGT
24901	ACTTCAAAA	TCATACATCA	CCAGATGTTG	ATCTTGGCGA	CATTTCAGGC	ATTAACGCTT
24961	CTGTCTGCAA	CATTCAAAAA	GAAATTGACC	GCCTCAATGA	GGTCGCTAAA	AATTTAAATG
25021	AATCACTCAT	TGACCTTCAA	GAATTGGGAA	AATATGAGCA	ATATATTAAA	TGGCCTTGGT
25081	ATGTTTGGCT	CGGCTTCATT	GCTGGACTAA	TTGCCATCGT	CATGTTTACA	ATCTTGGCTT
25141	GTTGCATGAC	TAGTTGTGTC	AGTTGCCTCA	AGGGTGCATG	CTCTTGTGGT	TCTTGTGCA
25201	AGTTTGATGA	GGATGACTCT	GAGCCAGTTC	TCAAGGGTGT	CAAATTACAT	TACACATAAA
25261	CGAACTTATG	GATTTGTTTA	TGAGATTTTT	TACTCTTGA	TCAATTACTG	CACAGCCAGT
25321	AAAAATTGAC	AATGCTTCTC	CTGCAAGTAC	TGTTTATGCT	ACAGCAACGA	TACCGTACA
25381	AGCCTCACTC	CCTTTCGGAT	GGCTTGTAT	TGGCGTTGCA	TTTCTTGTG	TTTTTCAGAG
25441	CGCTACCAA	ATAATTGCCG	TCAATAAAG	ATGGCAGCTA	GCCCTTTATA	AGGGCTTCCA
25501	GTTCAATTGC	AATTTACTGC	TGCTATTTGT	TACCATCTAT	TCACATCTTT	TGCTTGTCCG
25561	TCACGGTAAG	GAGCGCAAT	TTTTGTACCT	CTATGCTTGT	ATATATTTTC	TACAATGCAT
25621	CAACGCATGT	AGAATTATTA	TGAGATGTTG	GCTTTGTTGG	AAGTGCRAAT	CCAAGAACCC
25681	ATTACTTTAT	GATGCCAACT	ACTTTGTTTG	CTGGCACACA	CATAACTATG	ACTACTGTAT
25741	ACCATATAAC	AGTGTACACG	ATACAATTGT	CGTTACTGAA	GGI GACGGCA	TTTCAACACC
25801	AAAACTCAAA	GAAGACTACC	AAATTGGTGG	TTATTCTGAG	GATAGGCCT	CAGGTGTTAA
25861	AGACTATGTC	GTTGTACATG	GCTATTTTAC	CGAAGTTTAC	TACCAGCTTG	AGTCTACACA
25921	AATTACTACA	GACACTGGTA	TTGAAAATGC	TACATTTCTC	ATCTTTAACA	AGCTTGTAAA
25981	AGACCACCG	AATGTGCAAA	TACACACAAT	CGACGGCTCT	TCAGGAGTTG	CTAATCCAGC
26041	AATGGATCCA	ATTTATGATG	AGCCGACGAC	GACTACTAGC	GTGCCTTTGT	AAGCACAGA
26101	AAGTGAGTAC	GAACCTATGT	ACTCATTCGT	TTCCGAAAGAA	ACAGGTACGT	TAATAGTTAA
26161	TAGCCTACTT	CTTTTCTTG	CTTTCTGTTG	ATTTCTGCTA	GTCACTTAG	CCATCCTTAC
26221	TGCGCTTCGA	TTGTGTGCGT	ACTGCTCAA	TATTGTAAAC	GTGAGTTTAG	TAAAACCAAC
26281	GGTTTACGTC	TACTCGCGTG	TTAAAAATCT	GAACCTTCT	GAAGGAGTTC	CTGATCTTCT
26341	GGCTAAACG	AACTAACTAT	TATTATTATT	CTGTTTGGAA	CTTTAACATT	GCTTATCATG
26401	GCAGACAACG	GTACTATTAC	CGTTGAGGAG	CTTAAACAAC	TCCTTGAACA	ATGGAACTTA
26461	GTAATAGGTT	TCCTATTCTT	AGCCTGGATT	ATGTTACTAC	AATTTGCCTA	TTCTAATCCG
26521	AACAGTTTTT	TGTACATAAT	AAAGCTTGT	TTCCCTTGGC	TCTTGTGGCC	AGTAACACTT
26581	GCTTGTTTTG	TGCTTGTGCT	TGTCTACAGA	ATTAATTGGG	TGACTGGCGG	GATTGCGATT
26641	GCAATGGCTT	GTATTGTAGG	CTGTATGTTG	CTTAGCTACT	TCGTTGCTTC	CTCAGGCTG
26701	TTTGTCTCGTA	CCCCTCAAT	GTGGTCATTC	AACCAGAAA	CAAACATTCT	TCTCAATGTC
26761	CCTCTCCGGG	GGACAATTGT	GACCAGACCG	CTCATGGAAA	GTGAACTTGT	CATTGGTGTCT
26821	GTGATCATT	GTGGTCACTT	GCGAATGGCC	GGACACTCC	TAGGGCGCTG	TGACATTAAG
26881	GACCTGCCAA	AAGAGATCAC	TGTGGCTACA	TCACGAACGC	TTTCTTATTA	CAAATTAGGA
26941	GCGTCGCAGC	GTGTAGGCAC	TGATTCAGGT	TTTGTCTCAT	ACAACCGCTA	CCGTATTGGA
27001	AACTATAAAT	TAAATACAGA	CCACGCCGGT	AGCAACGACA	ATATTGCCTT	GCTAGTACAG
27061	TAAGTGACAA	CAGATGTTTC	ATCTTGTGTA	CTTCCAGGTT	ACAATAGCAG	AGATATTGAT
27121	TATCATTTATG	AGGACTTTCA	GGATTGCTAT	TTGGAATCTT	GACGTTATAA	TAAGTTCAAT
27181	AGTGAGACAA	TTATTTAAGC	CTCTAACTAA	GAAGAATTAT	TCGGAGTTAG	ATGATGAAGA
27241	ACCTATGGAG	TTAGATTATC	CATAAAACGA	ACATGAAAAT	TATTCTCTTC	CTGACATTGA
27301	TTGTATTTAC	ATCTTGCAG	CTATATCACT	ATCAGGAGTG	TGTTAGAGGT	ACGACTGTAC
27361	TACTAAAAGA	ACCTTGCCCA	TCAGGAACAT	ACGAGGGCAA	TTCAACATTT	CACCCCTCTG
27421	CTGACAATAA	ATTTGCACATA	ACTTGCACATA	GCACACACTT	TGCTTTTGGT	TGTGCTGACG
27481	TACTTTCGACA	TACCTATCAG	CTGCGTGCAA	GATCAGTTTC	ACCAAACTT	TTTATCAGAC
27541	AAGAGGAGGT	TCAACAAGAG	CTTACTCTGC	CACTTTTTCT	CATGTTTGGT	GCTCTAGTAT
27601	TTTTAATACT	TTGCTTCACC	ATTAAGAGAA	AGACAGAATG	AATGAGCTCA	CTTTAATTGA
27661	CTTCTATTTG	TGCTTTTATG	CCTTCTGCT	ATTCCTTGT	TTAATAATGC	TTATATATTT

图 6-H

27721	TTGGTTTTCA	CTCGAAATCC	AGGATCTAGA	AGAACCTTGT	ACCAAAGTCT	AAACGAACAT
27781	GAAACTTCTC	ATTGTTTTGA	CTTGATTTTC	TCTATGCAGT	TGCATATGCA	CTGTAGTACA
27B41	GCGCTGTGCA	TCTAATAAAC	CTCATGTGCT	TGAAGATCCT	TGTAAGGTAC	AACACTAGGG
27901	GTAATACTTA	TAGCACTGCT	TGGCTTTGTG	CTCTAGGAAA	GGTTTTACCT	TTTCATAGAT
27961	GGCACACTAT	GGTTCAAACA	TGCACACCTA	ATGTTACTAT	CAACTGTCAA	GATCCAGCTG
28021	GTGGTGCGCT	TATAGCTAGG	TGTTGOTACC	TTCATGAAGG	TCACCAAAC	GCTGCATTTA
26081	GAGACGTACT	TGTTGTTTTA	AATAAACGAA	CAAATTAATA	TGTCTGATAA	TGGACCCCAA
28141	TCAAACCAAC	GTAGTGCCCC	CCGCATTACA	TTTGGTGGAC	CCACAGATTC	AACTGACAA
28201	AACCAGAAATG	GAGGACGCAA	TGGGGCAAGG	CCAAAACAGC	GCCGACCCCA	AGGTTTACCC
28261	AATAATACTG	CGTCTTGGTT	CACAGCTCTC	ACTCAGCATG	GCAAGGAGGA	ACTTAGATTC
28321	CCTCGAGGCC	AGGGCGTTCC	AATCAACACC	AATAGTGGTC	CAGATGACCA	AATTGGCTAC
28381	TACCGAAGAG	CTACCCGACG	AGTTCGTGGT	GGTGACGGCA	AAATGAAAGA	GCTCAGCCCC
28441	AGATGGTACT	TCTATTACCT	AGGAACTGGC	CCAGAAGCTT	CACTTCCCTA	CGGCGCTAAC
28501	AAAGAAGGCA	TCGTATGGGT	TGCAACTGAG	GGAGCCTTGA	ATACACCCAA	AGACCACATT
28561	GGCACCCGCA	ATCCTAATAA	CAATGCTGCC	ACCGTGCTAC	AACTTCTCTA	AGGAACAACA
28621	TTGCCAAAAG	GCTTCTACGC	AGAGGGAAGC	AGAGGCGGCA	GTCAAGCCTC	TTCTCGCTCC
28681	TCATCACGTA	GTCGCGGTAA	TTCAAGAAAT	TCAACTCCTG	GCAGCAGTAG	GGGAAATTCT
28741	CCTGCTCGAA	TGGCTAGCGG	AGGTGGTGAA	ACTGCCCTCG	CGCTATTGCT	GCTAGACAGA
28601	TTGAACCAGC	TTGAGAGCAA	AGTTTCTGGT	AAAGGCCAAC	AACAACAAGG	CCAAACTGTC
28861	ACTAAGAAAT	CTGCTGCTGA	GGCATCTAAA	AAGCCTCGCC	AAAAACGTAC	TGCCACAAAA
28921	CAGTACAACG	TCACTCAAGC	ATTTGGGAGA	CGTGGTCCAG	AACAAACCCA	AGGAAATTTT
28981	GGGGACCAAG	ACCTAATCAG	ACAAGGAAC	GATTACAAC	ATTGGCCGCA	AATTGCACAA
29041	TTTGCTCCAA	GTGCCTCTGC	ATTCCTTTGGA	ATGTCACGCA	TTGGCATGGA	AGTCACACCT
29101	TCGGGAACAT	GGCTGACTTA	TCATGGAGCC	ATTAAATTGG	ATGACAAAGA	TCCACAATTC
29161	AAAGACAACG	TCATACTGCT	GAACAAGCAC	ATTGACGCAT	ACAAAACATT	CCCACCAACA
29221	GAGCCTAAAA	AGGACAAAAA	GAAAAAGACT	GATGAAGCTC	AGCCTTTGCC	GCAGAGACAA
29281	AAGAAGCAGC	CCACTGTGAC	TCTTCTTCTT	GCGGCTGACA	TGGATGATTT	CTCCAGACAA
29341	CTTCAAATTT	CCATGAGTGG	AGCTTCTGCT	GATTCAACTC	AGGCATAAAC	ACTCATGATG
29401	ACCACACAAG	GCAGATGGGC	TATGTAAACG	TTTTCGCAAT	TCCGTTTACG	ATACATAGTC
29461	TACTCTTGTG	CAGAATGAAT	TCTCGTAACT	AAACAGCACA	AGTAGGTTTA	GTTAACTTTA
29521	ATCTCACATA	GCAATCTTTA	ATCAATGTGT	AACATTAGGG	AGGACTTGAA	AGAGCCACCA
29581	CATTTTCATC	GAGGCCACGC	GGAGTACGAT	CGAGGGTACA	GTGAATAATG	CTAGGGAGAG
29641	CTGCCTATAT	GGAAGAGCCC	TAATGTGTAA	AATTAATTTT	AGTAGTGCTA	TCCCCATGTC
29701	ATTTTAATAG	CTTCTTAGGA	GAATGACAAA	AAAAAAAAAA	AA	

图 6-1



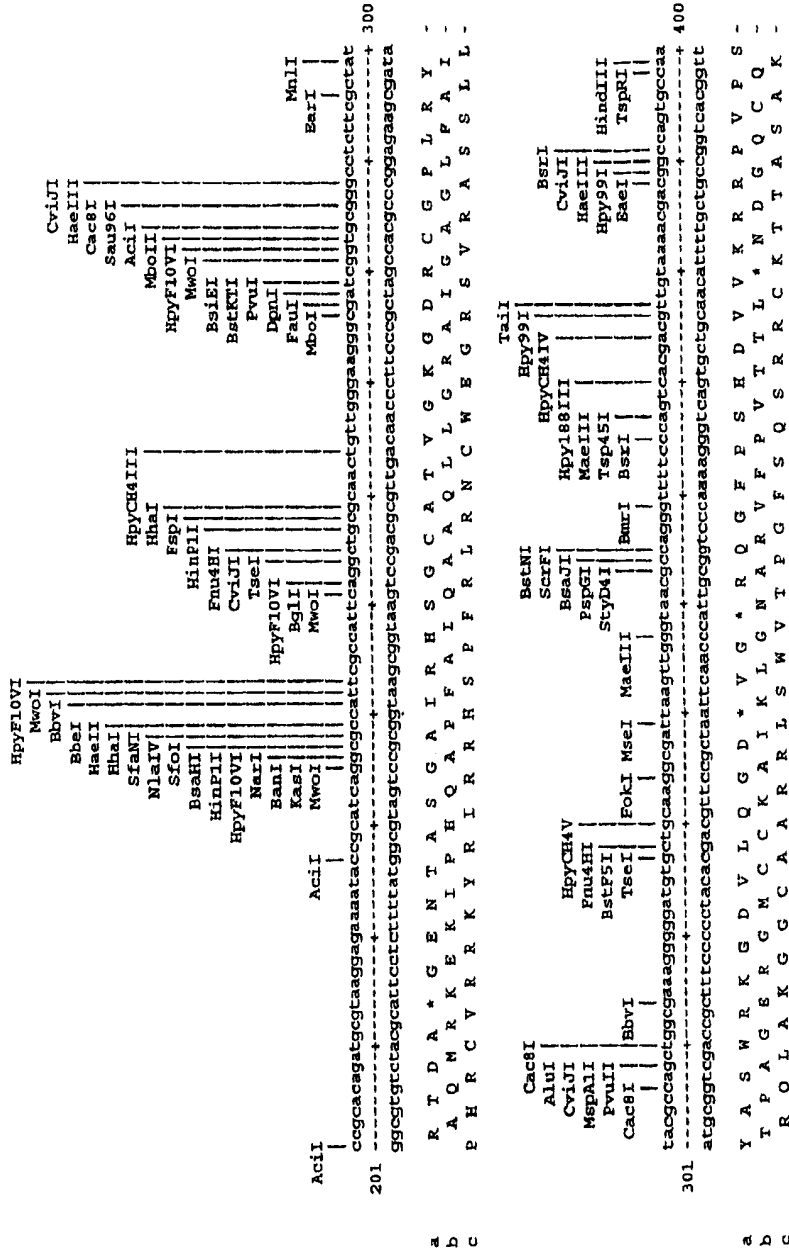


图 7-B





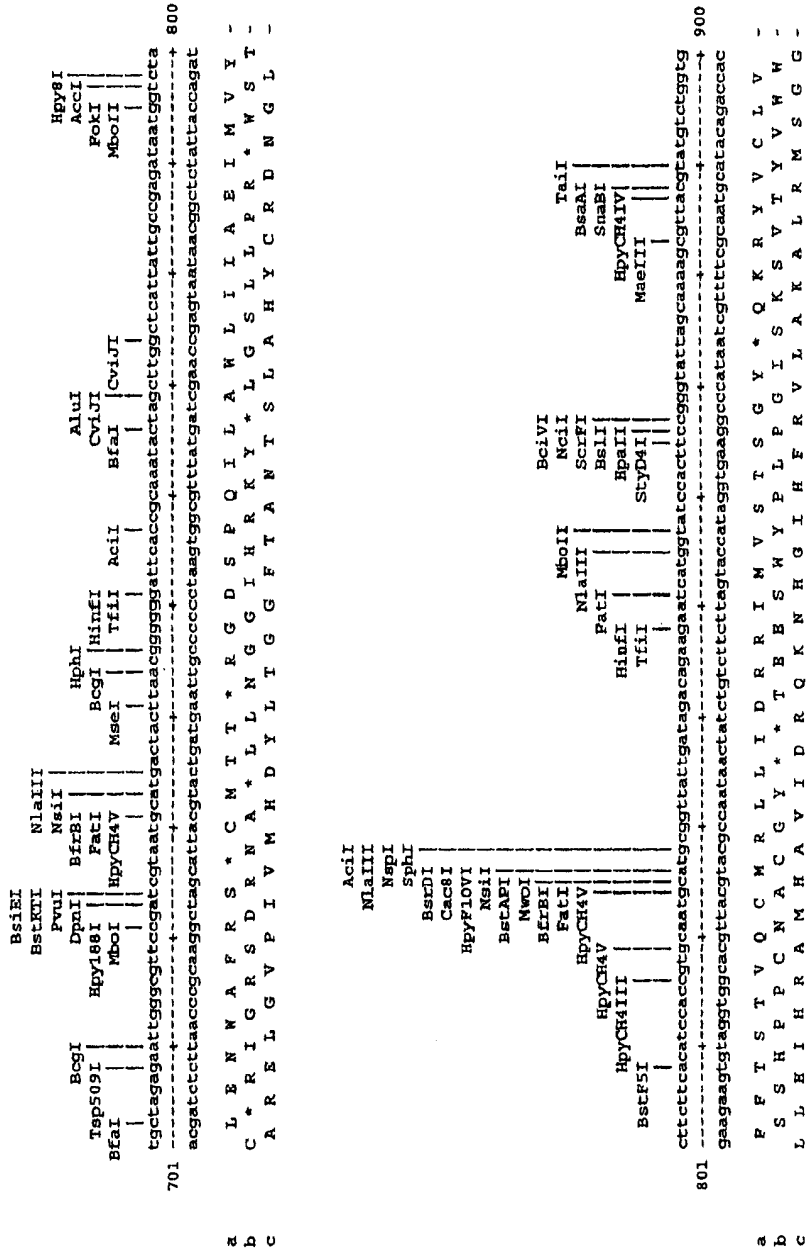


图 7-D

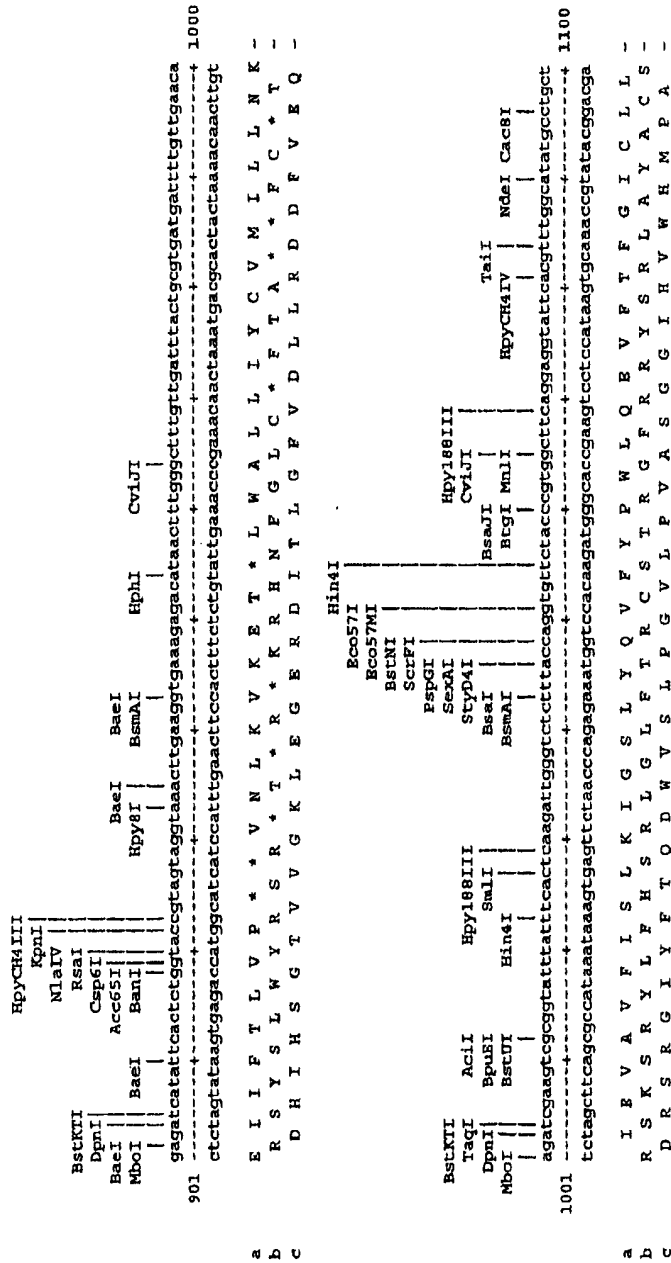


图 7-E

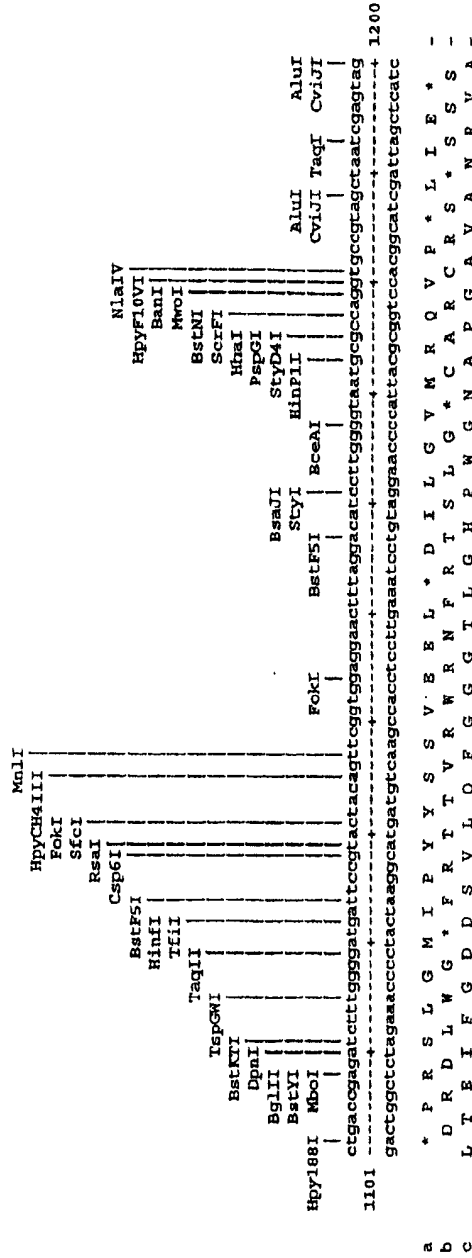


图 7-F

```

      BanII
      Bsp1286I
      HpaII
      NciII
      SbfI
      BbvI
      CviJI
      SfyDI
      NlaIV
      BbvI
      HpyF10VI
      HpyCH4V
      Cac8I
      HpyF10VI
      MwoI
      CviJI
      TspDTI
      BstOI
      NruI
      Hpy188III
      MnlI
      XmnI
      Tsp509I
      BseMI
      BspCMI
      MwoI
      Hpy188III
      Hpa10I
      DdeI
      TspDTI
      BstKTI
      DpnI
      MboI
      TaiI
      BngBI
      HpyCH4IV
      NspI
      AluI
      CviJI
      Hpy188III
      XbaI
      ctctagaagcatgtgtaaaagctcgtatgaaggacgtctctgactagagcttcctctcactagacacgctctccattctcctccactttaataagcgtcccgacgtttaccctcggccttgatcg
1201 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1300

```

a L \* K H V \* K L V M K D V I L L R K V M K L F A R L A N G A R N \* L -  
 b S R S M C K K S + \* R T + S C S G R + \* N Y S R G L Q M E P G T S - -  
 c L E A C V K A R N E G R D L A Q E G N E I I R E A C K W S P E L A -

图 7-G

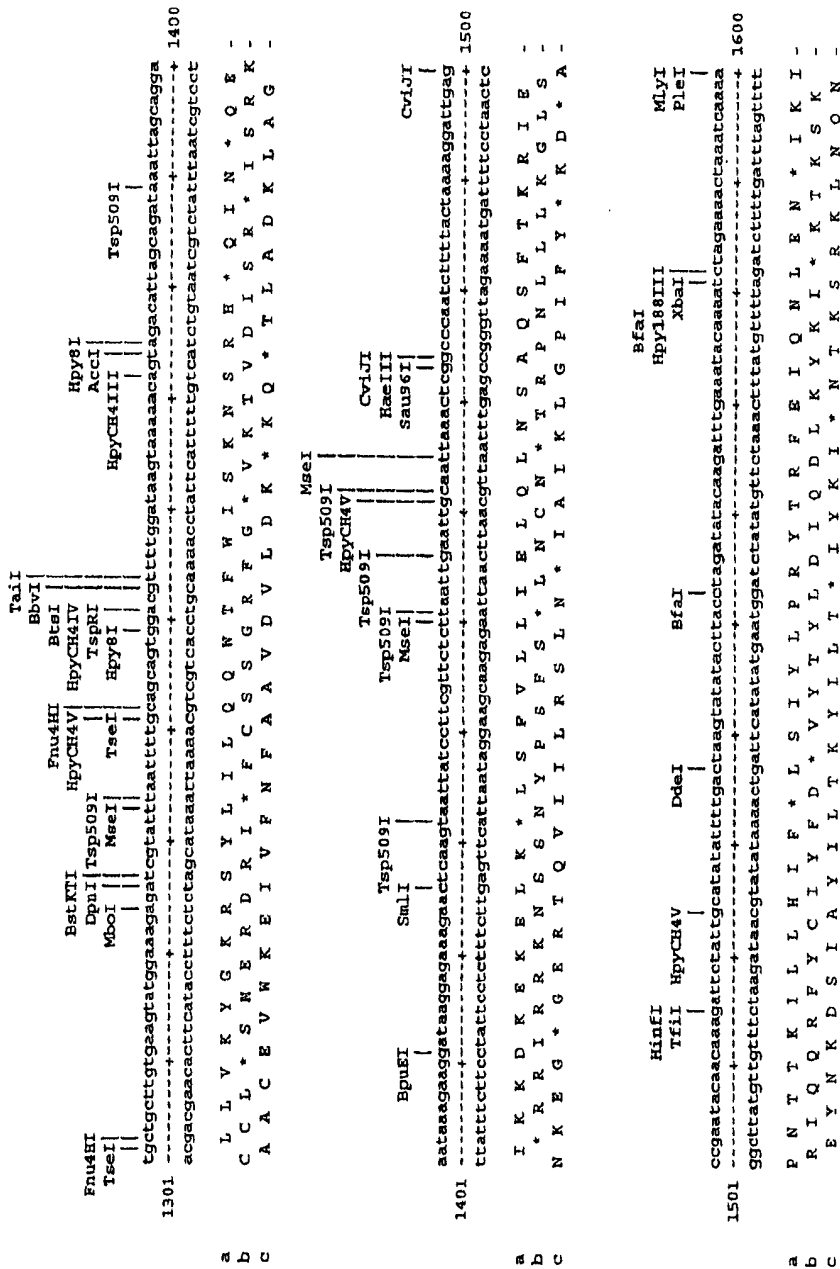


图 7-G

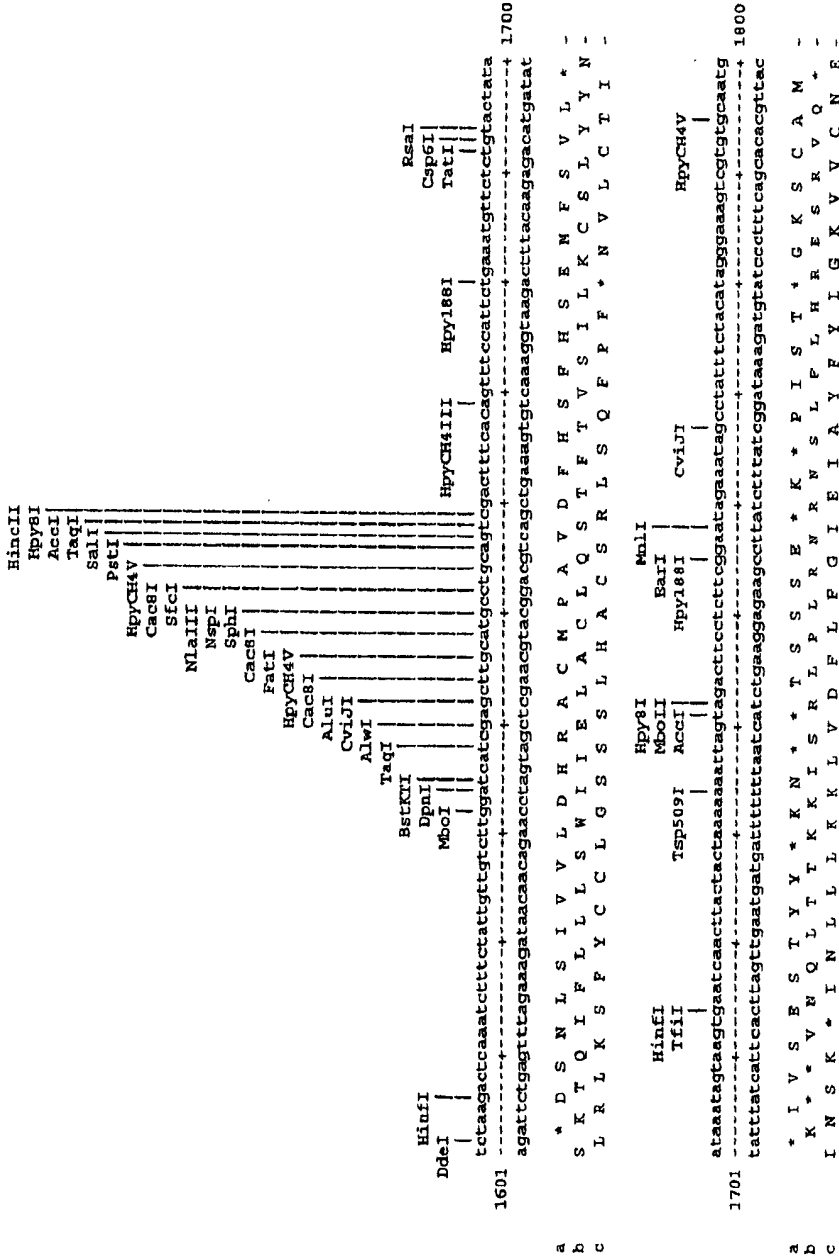


图 7-H

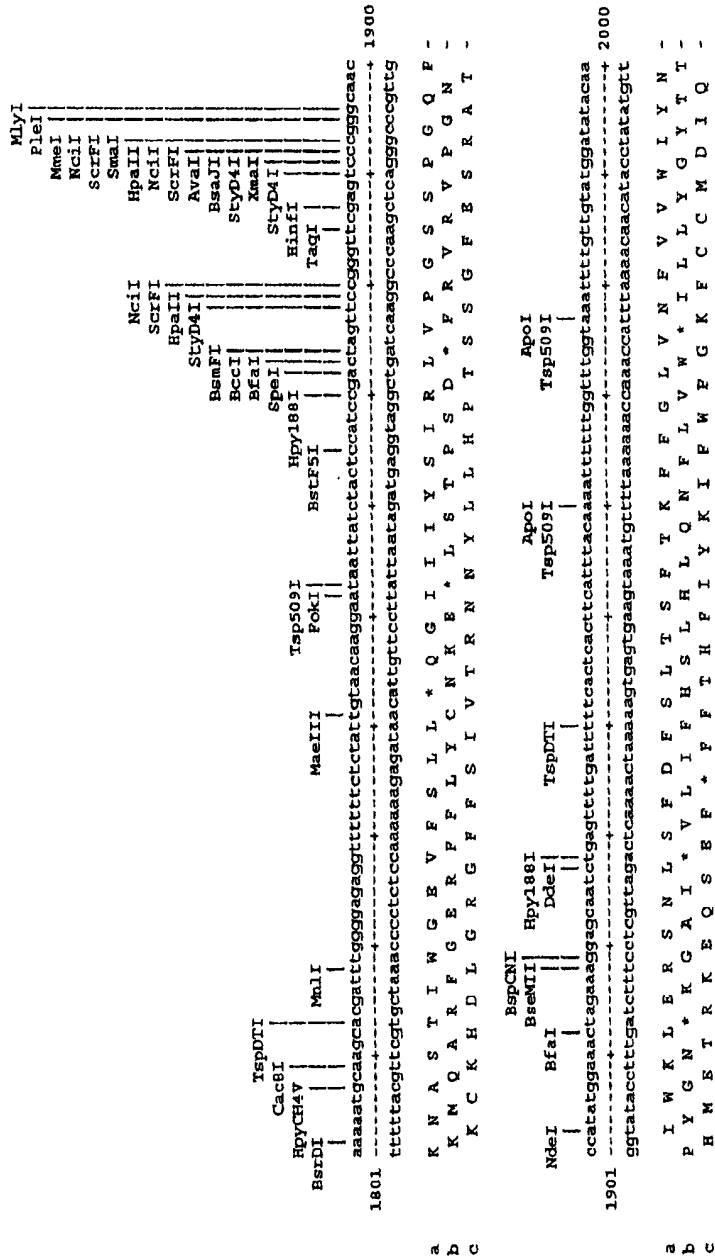


图 7-1

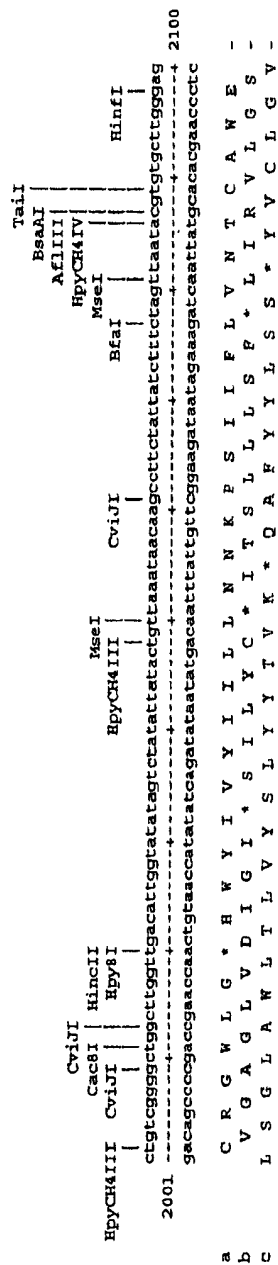


图 7-J



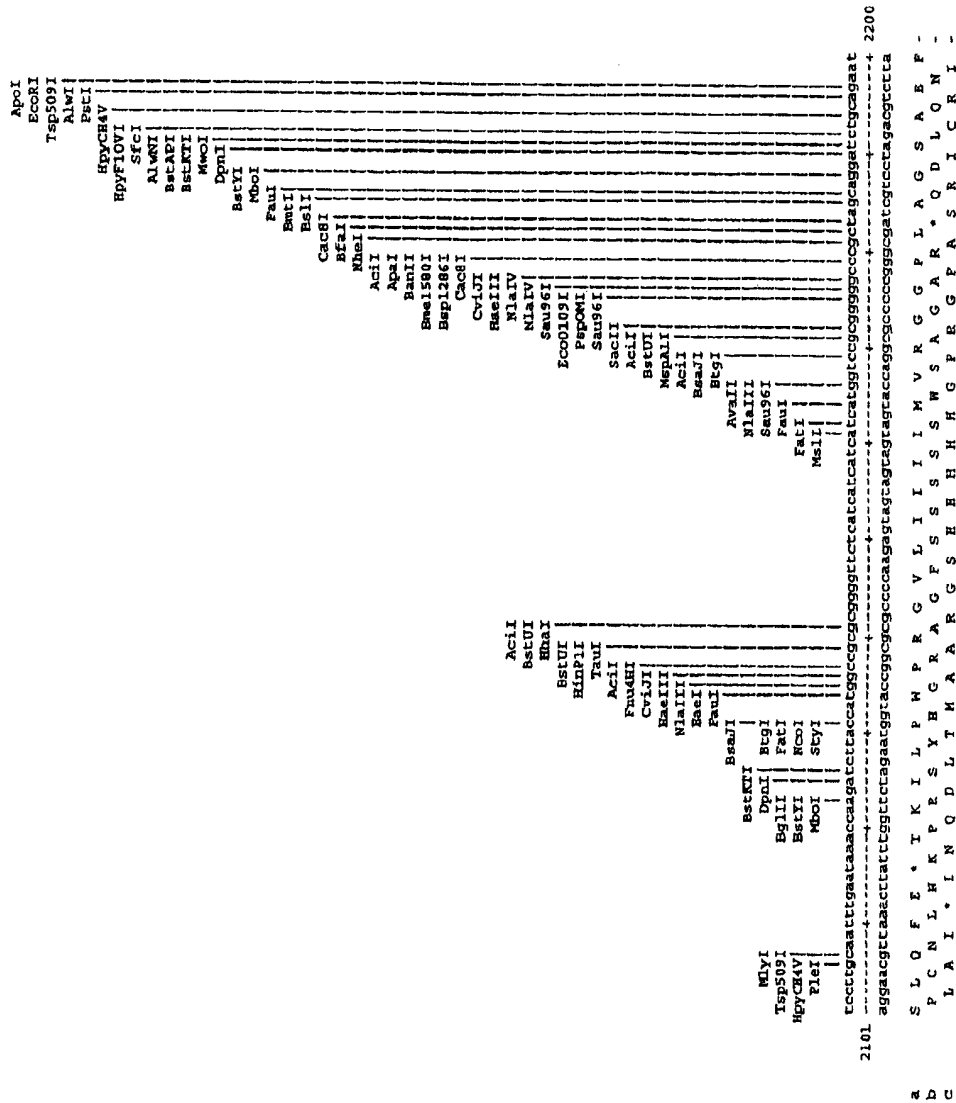


图 7-K

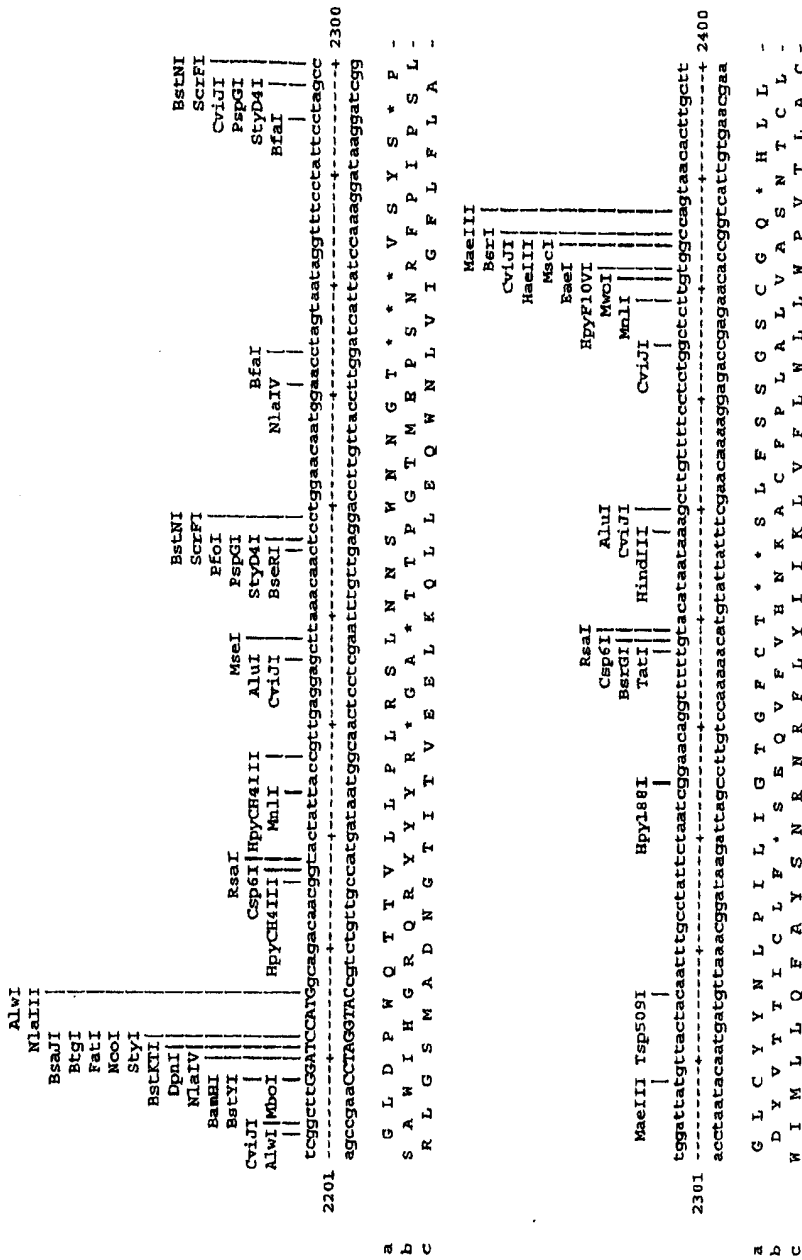


图 7-L

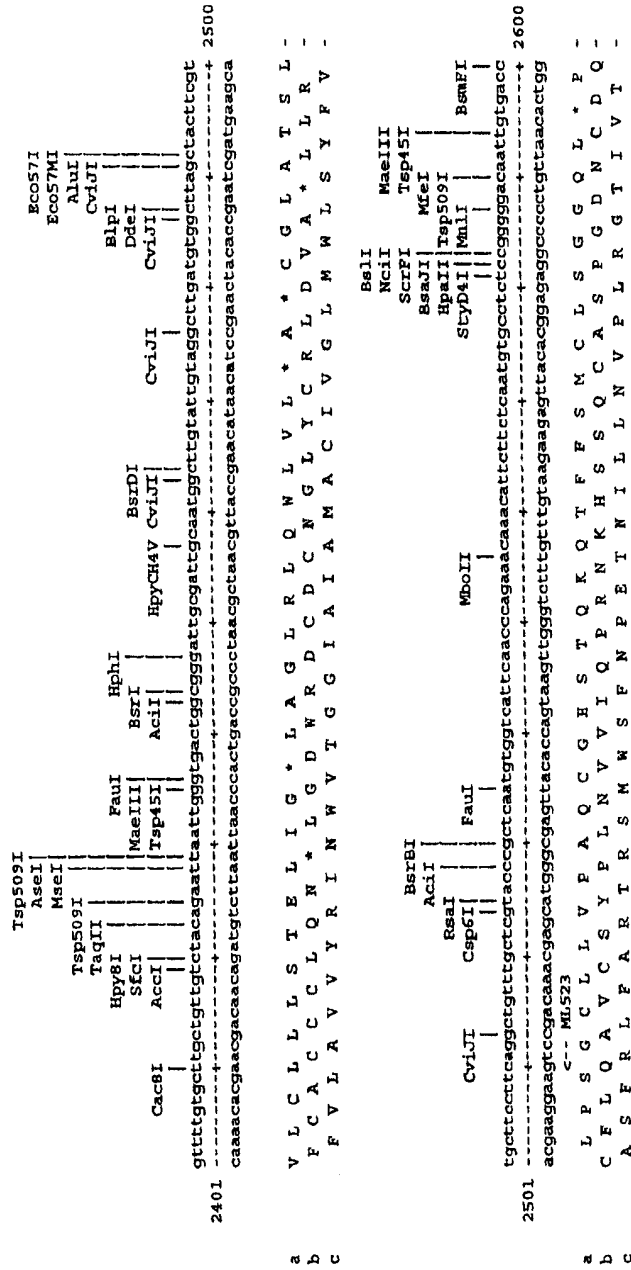


图 7-M

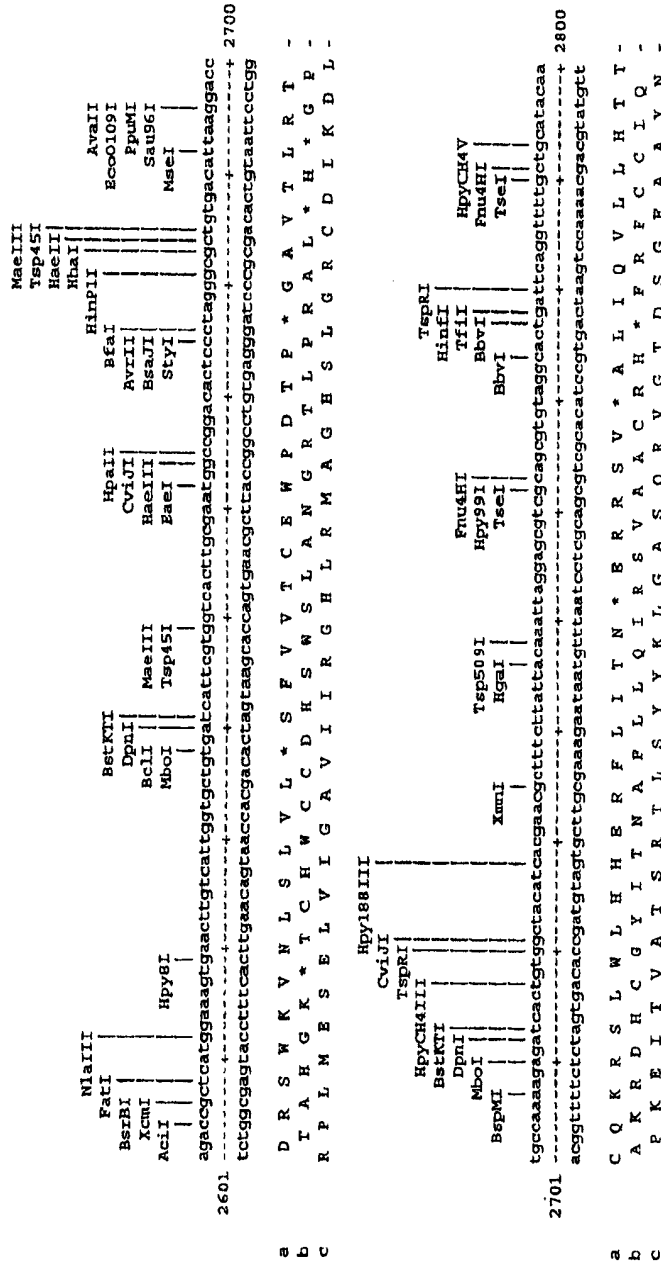


图 7-N



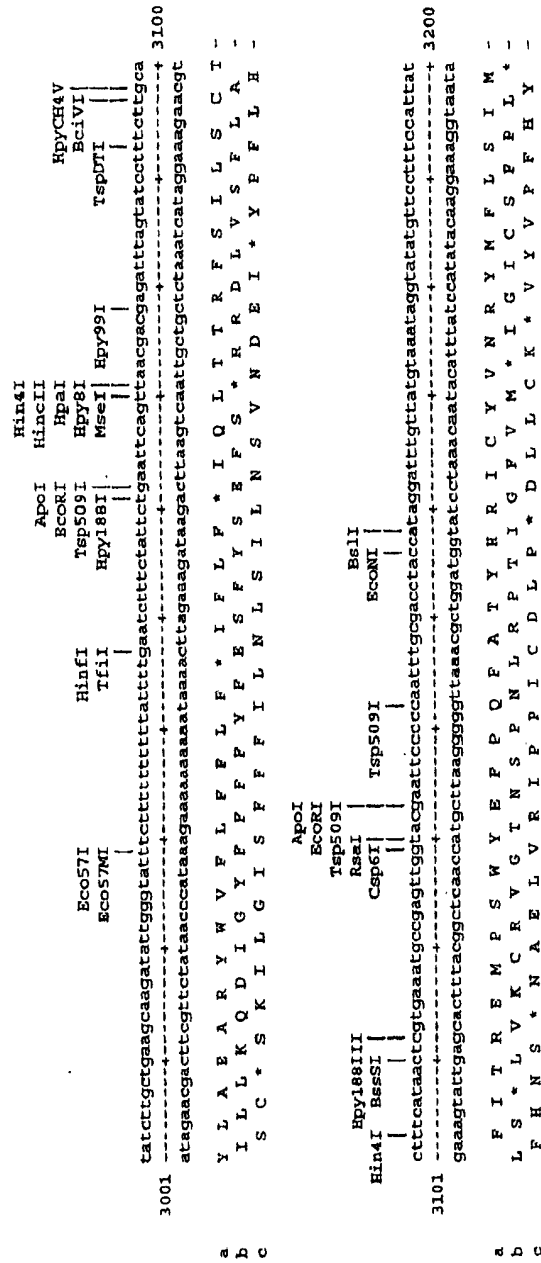


图 7-P

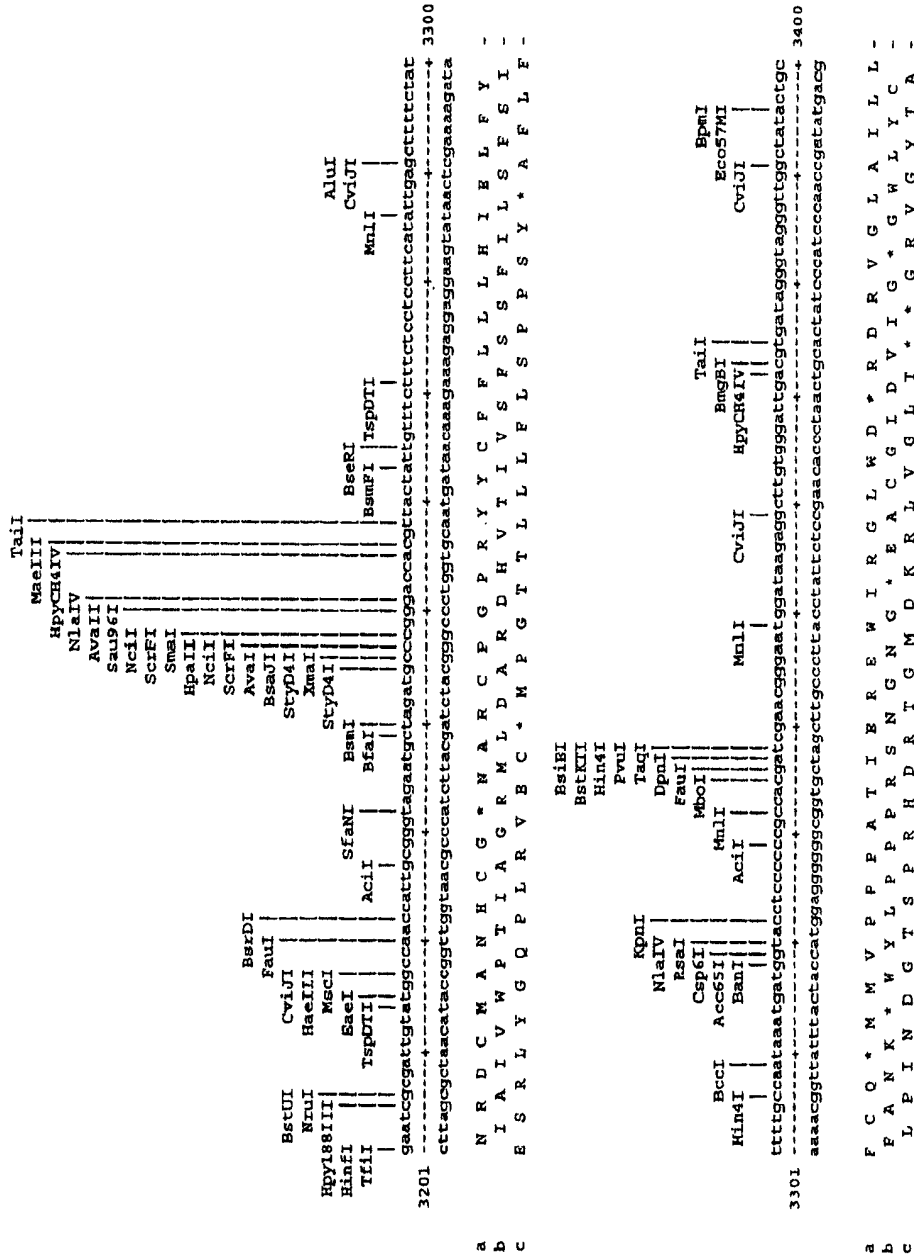


图 7-0





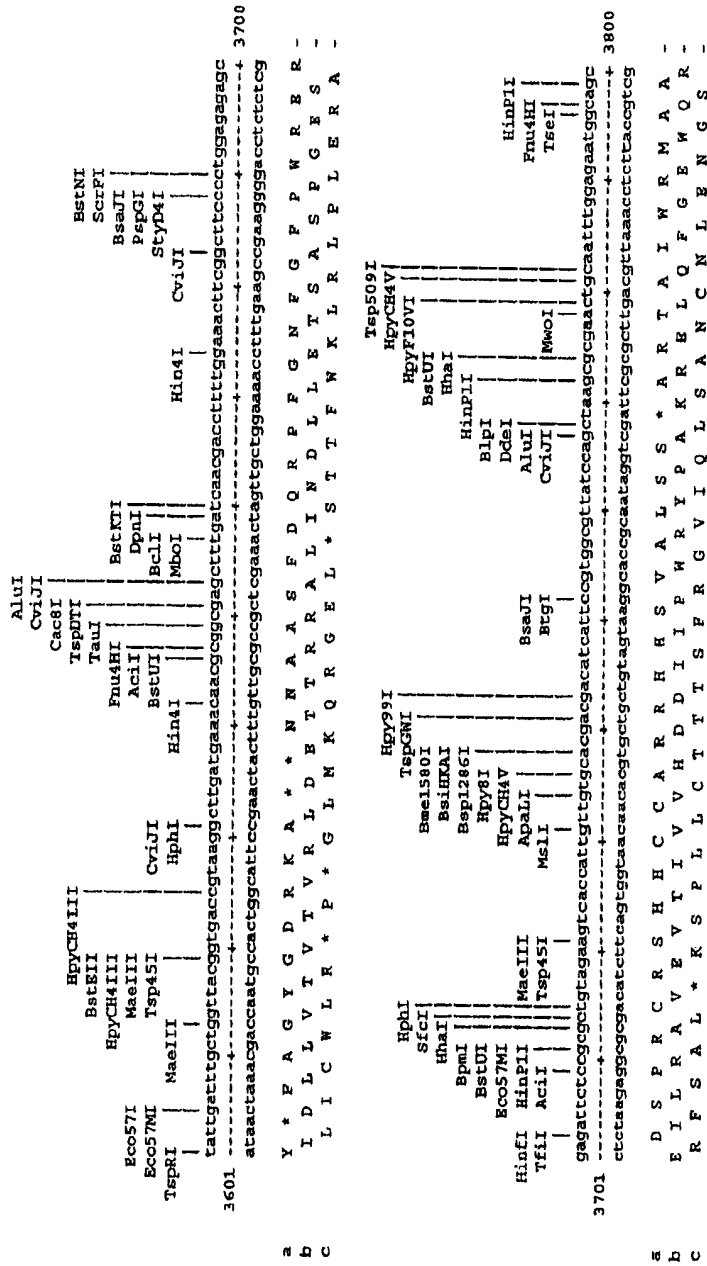


图 7-S

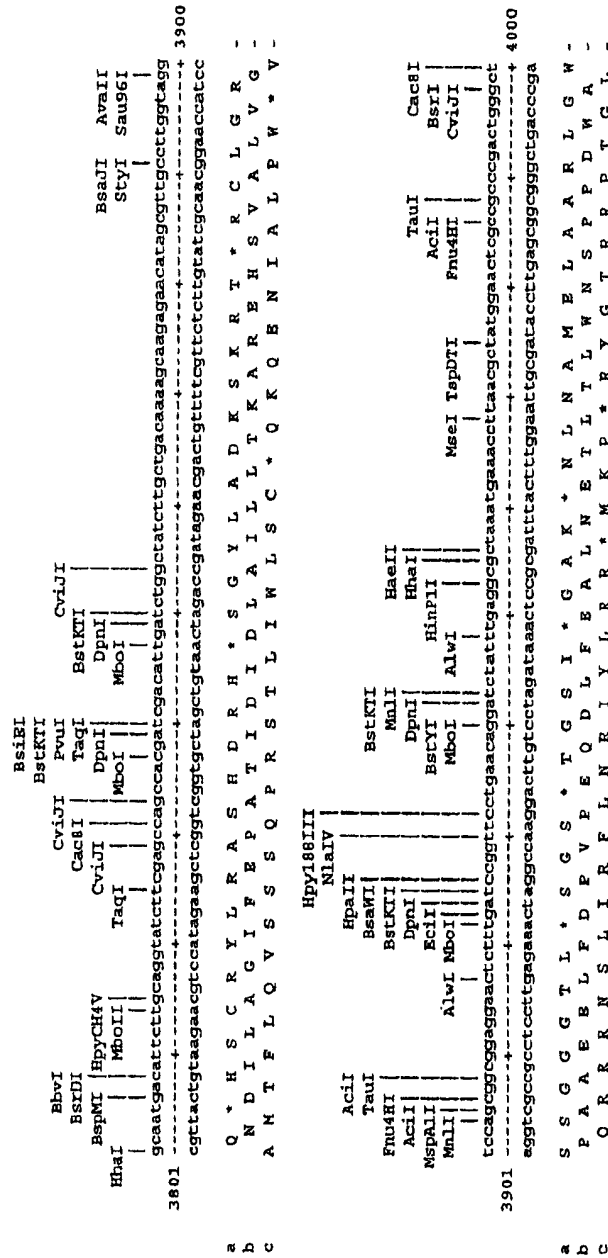


图 7-T

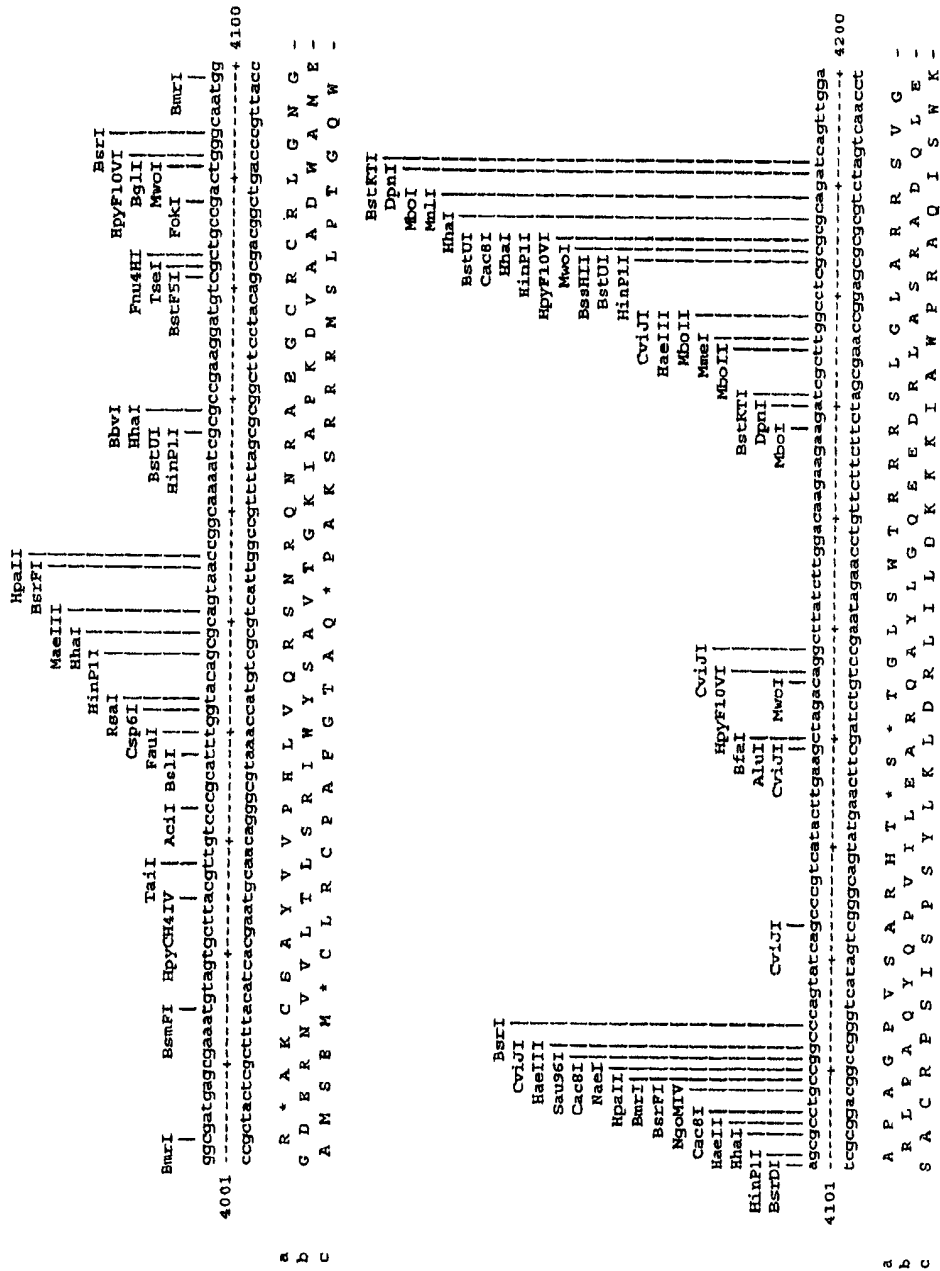


图 7-U

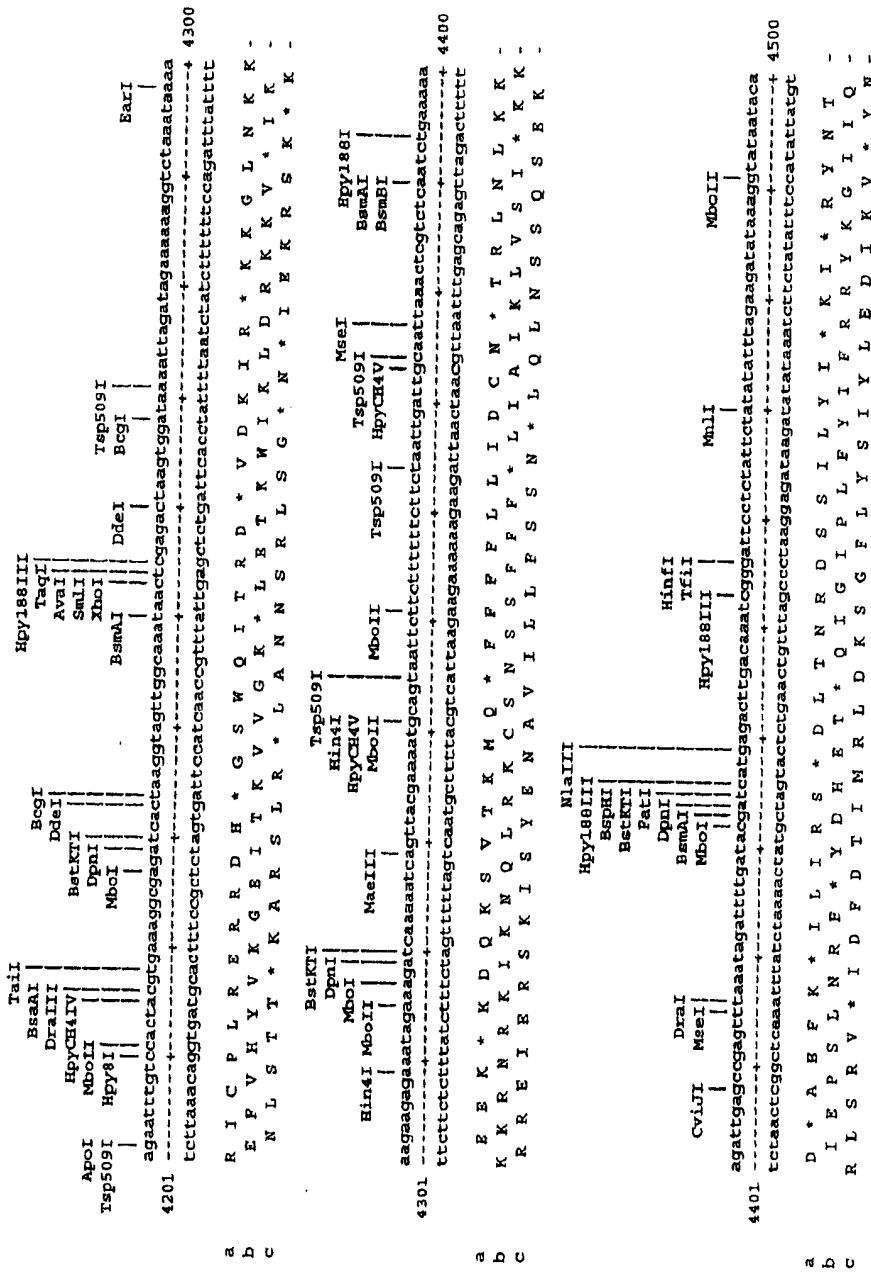


图 7-V



```

          TauI
          AclI |
          Fnu4HI |
          CviJI | |
          Tsp509I |
          MseI | |
          FseI | |
          DraI |
          MseI | |
          NlaIV |
          AseI |
          MseI |
          MmeI |
          Bpy188I |
          HinfI |
          TfilI |
tagggaagccgccctttatcaaaattatatttttaaatataaaaggggtcccaacattatataatagtagaagttccccagattccagaa
4801 atcctctcgcggaataatTTTTTaaataaataattattttcccccaaggtgtataattatataatcacttccaaaggggtcctaagtctt 4900
a * E K P P F I K N Y I I L N N I K G V P T Y * Y I V X C S P R F R I -
b R R S R P L L K I I L F * I I * R G F Q H I N I * * S V P P D S E -
c G B A A L Y * K L Y Y F K * Y K G G S N I L I Y S E V F P Q I Q N -

          Epy188I
          MseI |
          BfaI |
          DdeI |
          Tsp509I |
cttttttccaaactcacaatcccttattacttaataatcctagtgattgattttctcttagcttagtaggaataaagatatccaaataaataatttt
4901 gaaaaagttatgagtttaggaataatcaattattaggtacactaacctaaagatacgaactcctttattcttataagttttattataaaa
a F F Q Y S O S L L V N N P S D W I S M L S L I G N K Y F K * I I L -
b L F F N T H N P Y * L I L V I G F L C L V * * E I R Y S N K * F Y -
c F F S I L T I L I S * * S * * L D F Y A * S D R K * D I Q I N N F -

          RpyCH4V
          NlaIII |
          TspDTI |
          Hin4I |
          TspDTI |
          FseI | |
          MseI |
          Tsp509I |
          MseI |
          TaqI |
          MseI |
          XmiI | |
          AarI |
          BepMI |
          XmiI | |
atagcgaatgactattcattctatttctcatgcaaataggggcaagaanaactctatggaagatgggttataatcgatgttggtaaagaaggag
5001 tatcgttactgataagtagatacaaaaagtaagttatcccccttttttgagatacctttccaccacaataaagctacaacaattcttctctc
a * R M T I H L L Y F H A N R G O E N S M E R W F N S M L F K K E -
b S E * L P I Y C I F M Q I G G K K T L W K D G G L Y R C C L R R S -
c I A N D Y S S I V F S C K * G A R K L Y G K M V * F D V V * E G V -

```

图 7-X

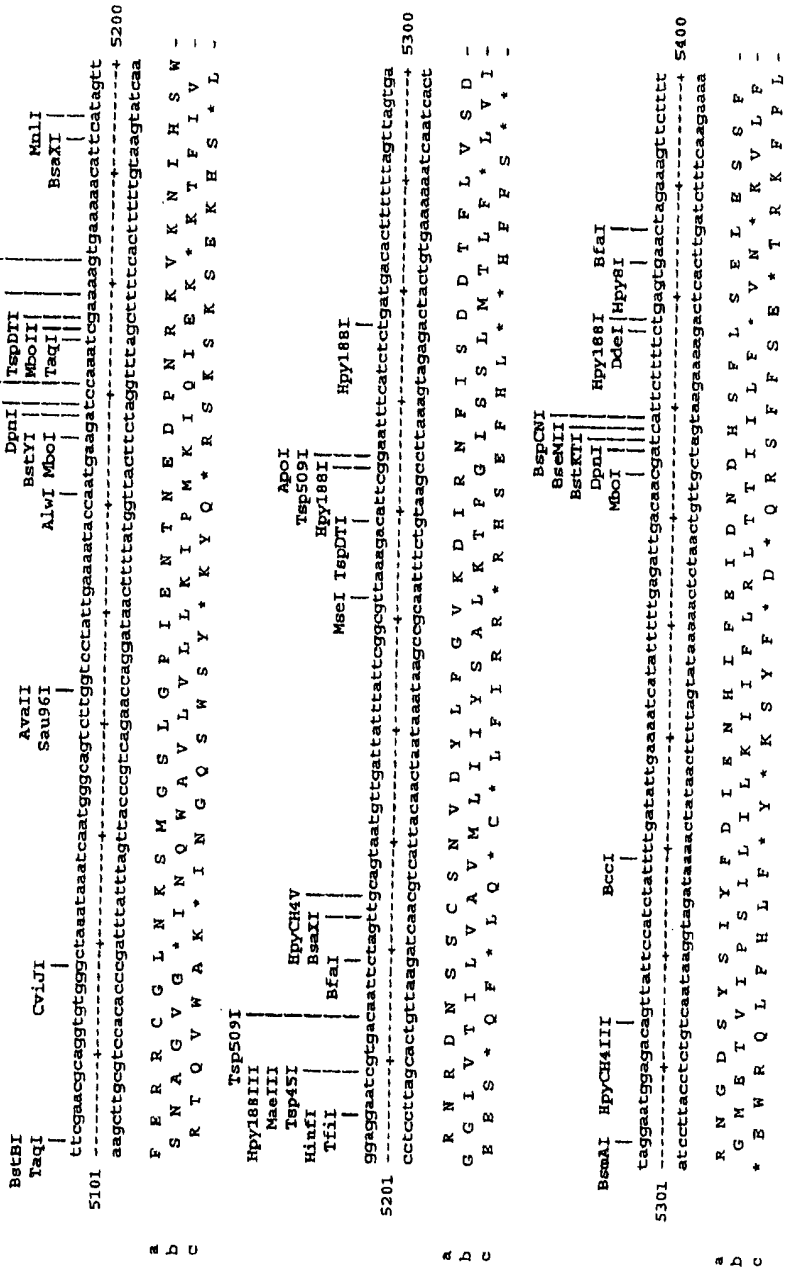


图 7-Y

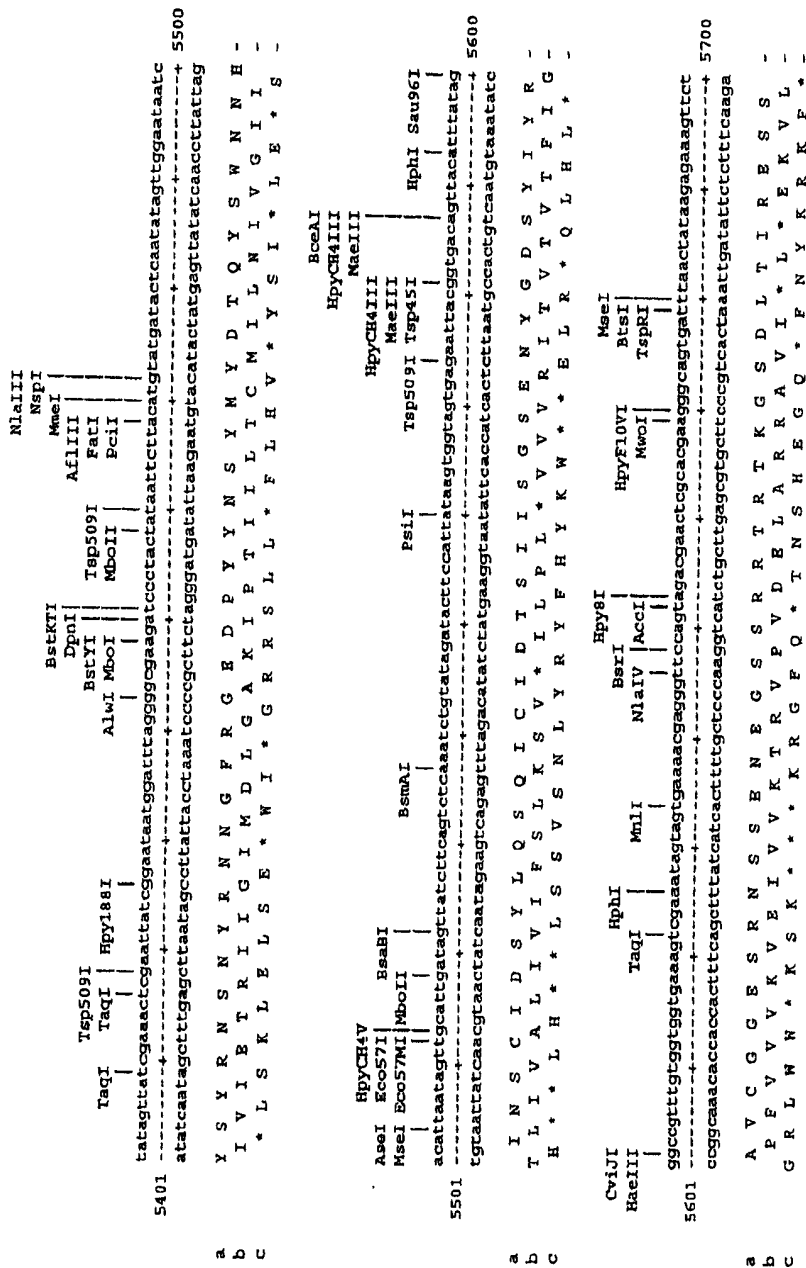


图 7-Z



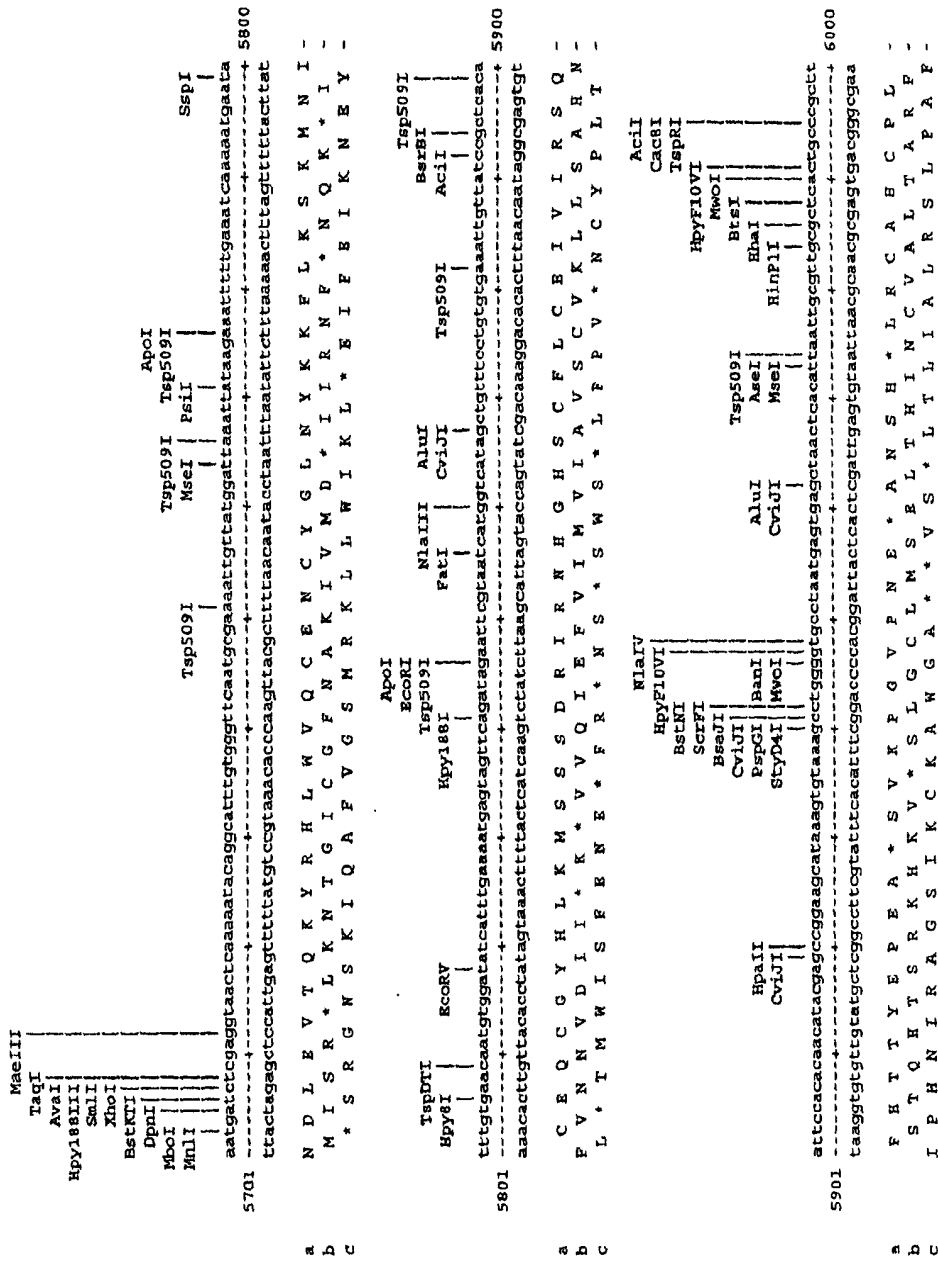


图 7-AA



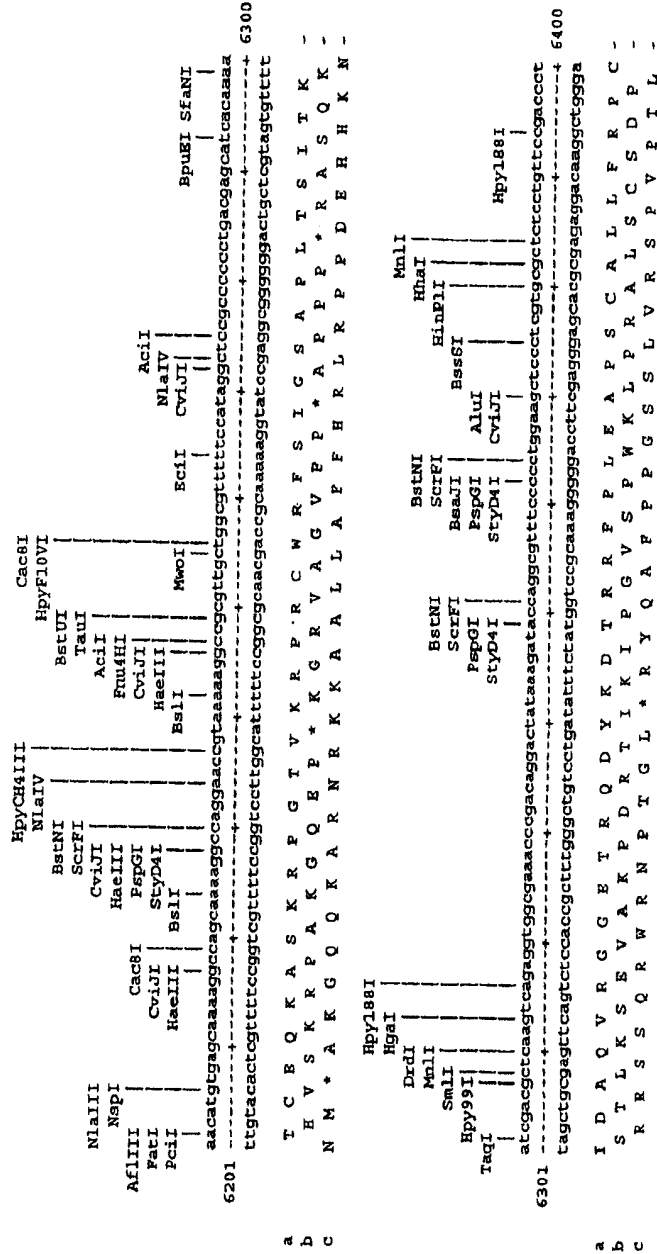


图 7-CC

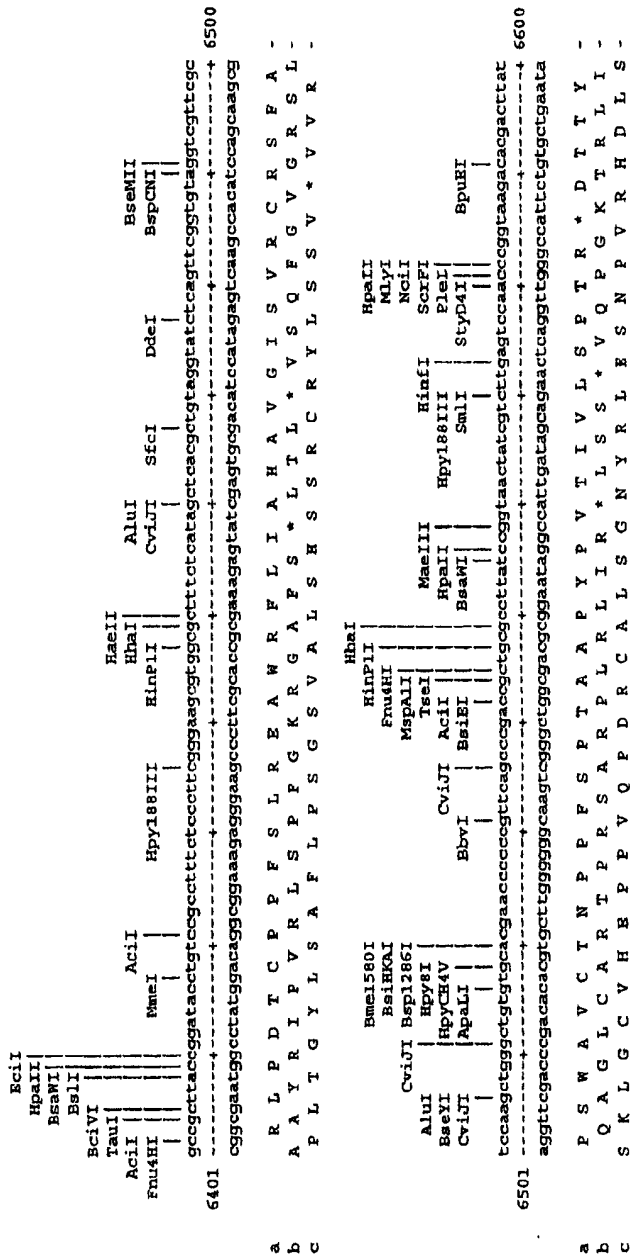


图 7-DD

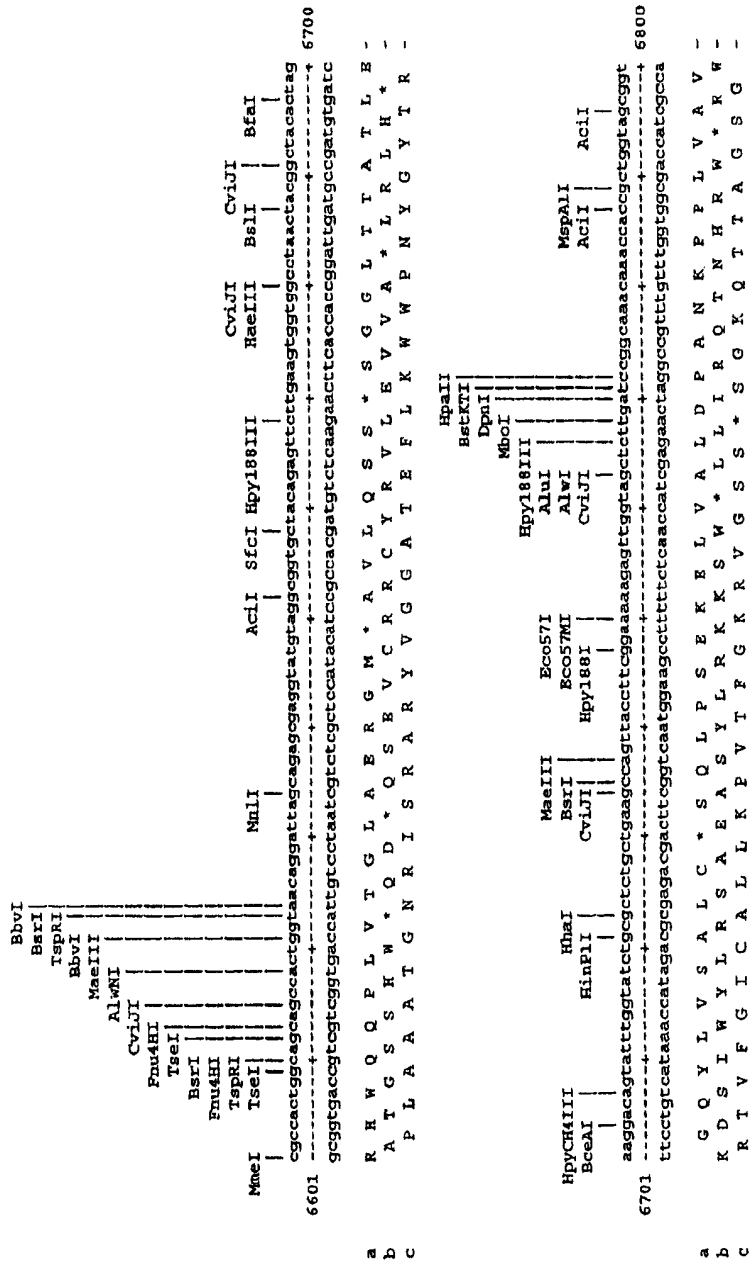


图 7-EE

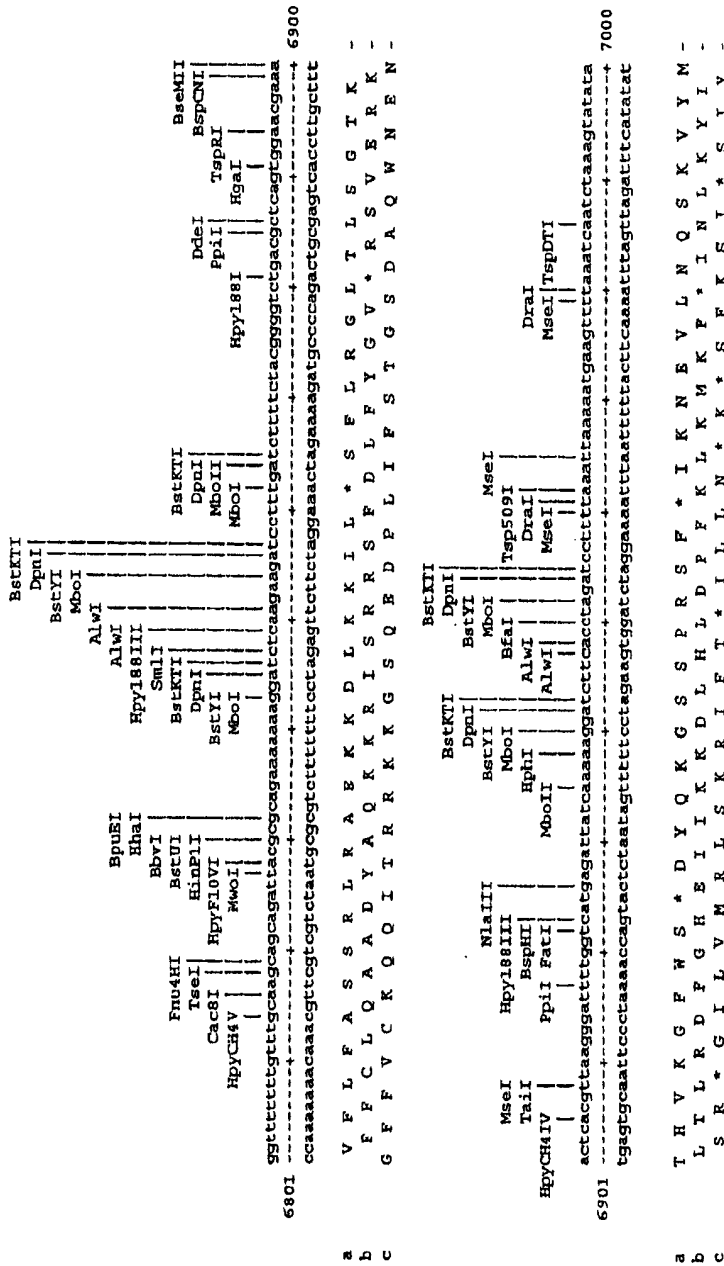


图 7-FF

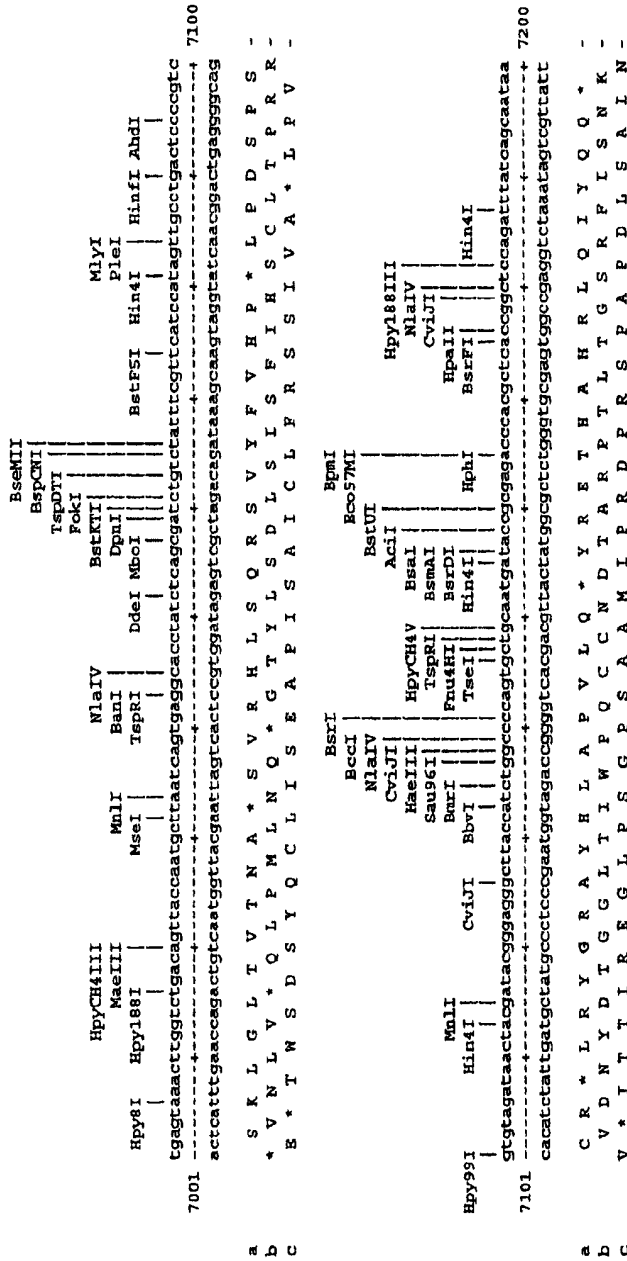


图 7-7GG

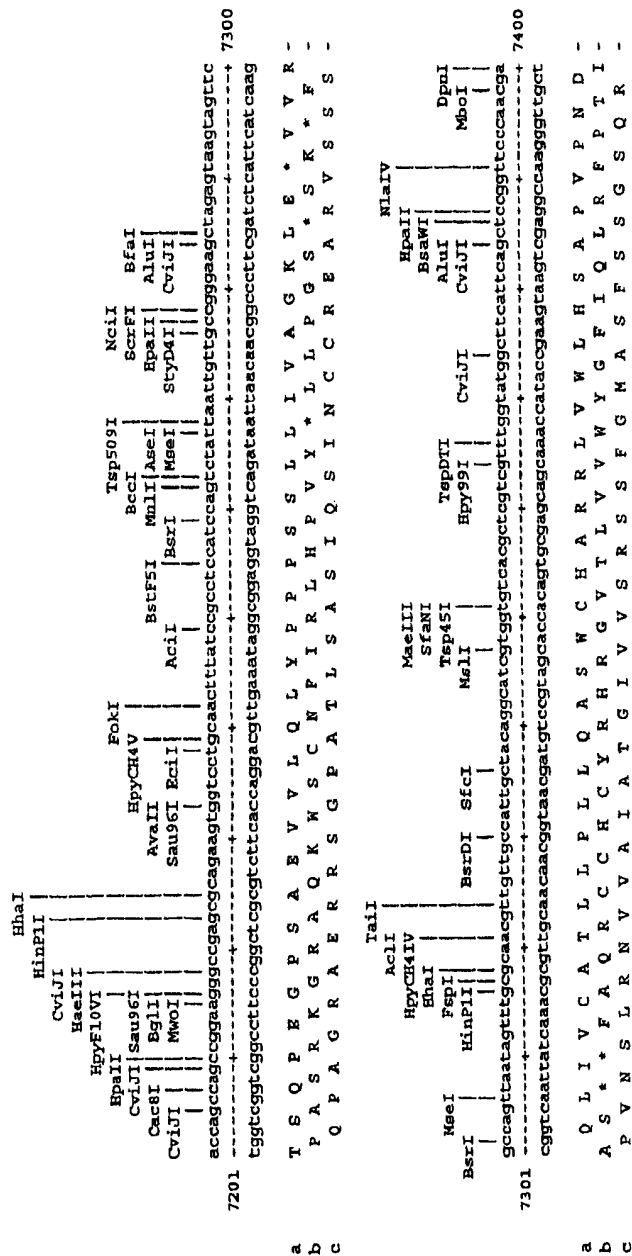


图 7-HH



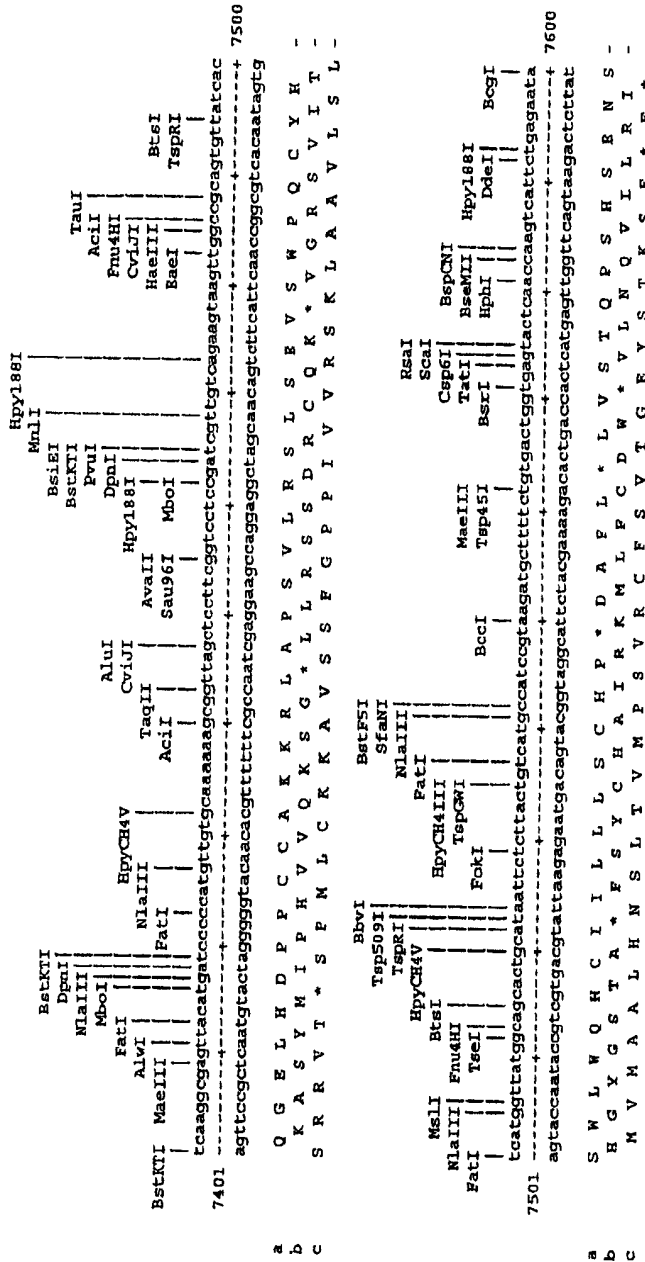


图 7-11



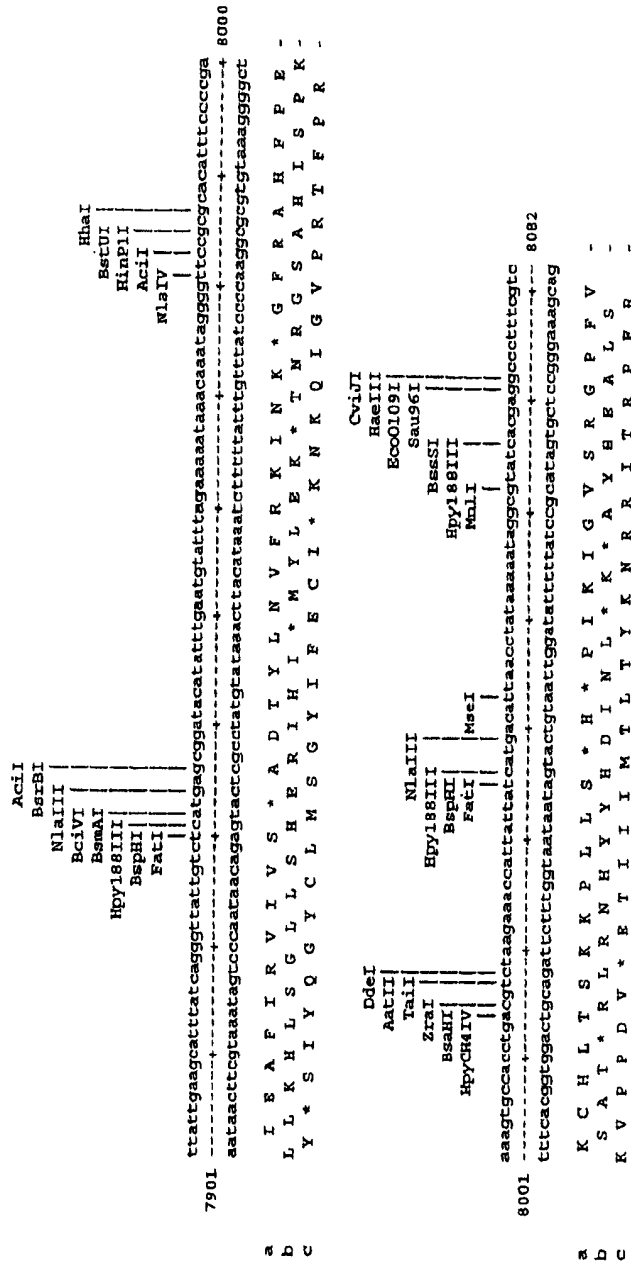


图 7-KK

切割的不被排除在外的酶:

AarI	AatII	AccI	Acc65I	AccII	AclI	AfeI	AflIII	AhdI	AluI	AlwI	AlwNI	ApaI
ApalI	ApoI	AseI	AvaI	AvaII	AvrII	BaeI	BamHI	BanI	BanII	BbeI	BbsI	BbvI
BccI	BceNI	BcgI	BclI	BclII	BfaI	BfrBI	BglI	BglII	BipI	BmeI580I	BngBI	BnrI
BdtI	BglI	BpuI	BpuEI	BsaI	BsaAI	BsaBI	BsaHI	BsaJI	BsaWI	BsaXI	BsaXII	BseRI
BseRI	BsiEI	BslI	BslII	BsmI	BsmAI	BsmBI	BsmFI	BspI286I	BspCNI	BspHI	BspNI	BsrI
BstBI	BstDI	BstPI	BstRGI	BstSII	BstSI	BstAPI	BstBI	BstEII	BstFSI	BstKFI	BstNI	BstUI
BstXI	BstYI	Bsu36I	BtgI	BtsI	Cac8I	Cep6I	CviJI	DdeI	DpnI	DraI	DraIII	DrdI
EaeI	EagI	EakI	EclI	Eco57I	EcoICRI	Eco57MI	EcoNI	EcoO109I	EcoRI	EcoRV	FalI	FatI
FauI	Fnu4HI	FokI	FspI	HaeII	HaeIII	HgaI	HhaI	Hin4I	HinPI	HincII	HindIII	HinfI
HpaI	HpaII	HphI	Hpy8I	Hpy9I	Hpy188I	Hpy188III	HpyCH4III	HpyCH4IV	HpyCH4V	HpyF10VI	KasI	KpnI
MaeIII	MboI	MboII	MfeI	MlyI	MaeI	NlaI	MscI	MseI	MslI	MspALI	MwoI	NaeI
NarI	NciI	NcoI	NdeI	NgoMIV	NheI	NlaIII	NlaIV	NruI	NsiI	NspI	PacI	PciI
PfoI	PleI	PpiI	FpumI	FsiI	FspGI	FspOML	FstI	FvuI	PvuII	RsaI	SacI	SacII
Sali	SapI	Sau96I	ScaI	SceFI	SerAI	SfaNI	SfcI	SmaI	SmaII	SmlI	SnaBI	SpeI
SphI	SepI	StyI	StyD4I	TaiI	TaqI	TaqII	TatI	TauI	TfiI	TseI	Tsp45I	Tsp509I
TspDTI	TspGWI	TspRI	XbaI	XcuI	XhoI	XnaI	XnuI	ZraI				

不切割的酶:

AflII	AgeI	AlcI	AscI	AsiSI	BbvCI	BglI	BsiNI	BspEI	BstZ17I	Clal
EseI	FspAI	MluI	NotI	PmeI	PmlI	PshAI	RsrII	SauBI	SbfI	SfiI
SgrAI	SrfI	StuI	SwaI	TthIII						

被排除在外的酶: MinCuts: 1 MaxCuts: 100000

无

图 7-LL

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc toagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatagcgg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
attcgccatt caggctgctc aactgttggg aaggcgatc ggtgcgggcc tcttcgctat 300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360
tttcccagtc acgacgttgt aaaacgacgg ccagtgccaa gcttgaaga gataaattga 420
acaagtatgg tcgtcccctg ttgggatgta ctattaacc taaattgggg ttatctgcta 480
aaaactacgg tagagocgtt tatgaatgtc ttcgcggtgg acttgatttt actaaagatg 540
atgagaacgt gaactcacia ccatttatgc gttggagaga tcgtttctta ttttgcgcg 600
aagcacttta taaagcacag gctgaaacag gtgaaatcaa agggcattac ttgaaatgcta 660
ctgcaggtac atgcgaagaa atgatcaaaa gagctgtatt tgctagagaa ttgggcggtc 720
cgatcgtaat gcatgactac ttaacggggg gattcacccg aaatactagc ttggctcatt 780
attgcccaga taatggtcta cttcttcaca tccacogtgc aatgcatgcg gttattgata 840
gacagaagaa tcatggtatc cacttccggg tattagcaaa agcgttacgt atgtctgggtg 900
gagatcatat tcaactctggt acogtagtag gtaaacttga aggtgaaaga gacataactt 960
tgggctttgt tgatttactg cgtgatgatt ttgtgaaaca agatcgaagt cgcggtattt 1020
atttcactca agattgggtc tctttaccag gtgttctacc cgtggcttca ggaggatttc 1080
acgtttggca tatgcctgct ctgaccgaga tctttgggga tgattccgta ctacagttcg 1140
gtggaggaac tttaggacat ccttggggta atgcgccagg tgccgtagct aatcgagtag 1200
ctctagaagc atgtgtaaaa gctcgtaatg aaggacgtga tcttgctcag gaaggtaatg 1260
aaattattcg cgaggcttgc aaatggagcc cggaaactagc tgctgcttgt gaagtatgga 1320
aagagatcgt atttaatttt gcagcagtg acgttttggg taagtataaaa cagttagacat 1380
tagcagataa attagcagga aataaagaag gataaggaga aagaactcaa gtaattatcc 1440
ttcgttctct taattgaatt gcaattaac tcggcccaat cttttactaa aaggattgag 1500
ccgaatacaa caaagattct attgcatata ttttgactaa gtatatactt acctagatat 1560
acaagatttg aaatacaaaa tctagaaaac taaatcaaaa tctaagactc aaatctttct 1620
attgttgtct tggatcatcg agcttgcagc cctgcagtcg actttcacag tttccattct 1680
gaaatgttct ctgtaactata ataaatagta agtgaatcaa cttactacta aaaaaattag 1740
tagacttctt cttcgggaata gaaatagcct atttctacat agggaaagtc gtgtgcaatg 1800
aaaaatgcaa gcacgatttg gggagagggt ttttctctat tgtaacaagg aataattatc 1860
tactccatcc gactagttcc gggttcgagt cccgggcaac ccatatggaa actagaaagg 1920
agcaatctga gttttgattt ttcactcact tcatttacaa aattttttgg tttggtaaat 1980
tttgttgtat ggatatacaa ctgtcggggc tggcttgggt gacattggta tatagtctat 2040
attatactgt taaataacaa gccttctatt atctttctag ttaatacgtg tgcttgggag 2100
tccttgcaat ttgaataaac caagatctta ccatggccgc gcgggggtct catcatcatc 2160
atcatcatgg catggcagac aacggtacta ttaccgttga ggagcttaaa caactcctgg 2220
aacaatggaa cctagtaata ggttctctat tccatagcctg gattatgtta ctacaatttg 2280
cctattctaa tcggaacagg tttttgtaca taataaagct tgttttctct tggctcttct 2340

```

图 8-A

```

ggccagtaac acttgcttgt tttgtgcttg ctggtgtcta cagaattaat tgggtgactg 2400
gcgggattgc gattgcaatg gcttgtattg taggcttgat gtggetttagc tacttctgtg 2460
cttccttcag gctgtttgct cgtaccgct caatgtggtc attcaaccca gaaacaaca 2520
ttcttctcaa tgtgcctctc cgggggacaa ttgtgaccag accgctcatg gaaagtgaac 2580
ttgtcattgg tgctgtgatc attcgtggtc acttgcgat ggccggacac tccctagggc 2640
gctgtgacat taaggacctg ccaaaagaga tcactgtggc tacatcacga acgcttctt 2700
attacaatt aggagcgtcg cagcgtgtag gcactgattc aggttttctg gcatacaacc 2760
gctaccgat tggaaactat aaattaata cagaccacgc cggtagcaac gacaatattg 2820
ctttgctagt acagtaagaa ttcaagccga attccagcac actggcggcc gttactagt 2880
gatccgagct ctaattaatt aaggcttttc tgctaacata tagcaatttt tgaagaaagg 2940
aaagctagaa ataccaata tcttgctgaa gcaagatatt gggtatcttct tttttttta 3000
ttttgaaatc ttctattctg aattcagtta acgacgagat ttagtatcct ttcttgcact 3060
ttcataactc gtgaaatgcc gagttggtac gaattcccc aatttgcgac ctaccatagg 3120
atctgttatg taaataggta tatgttcctt tccattatga atcgcgattg tatggccaac 3180
cattgcccgtt agaatgctag atgcccggga ccacgttact attgtttctt tctcctcctt 3240
catattgagc ttttctattt ttgccaaata atgatggtac ctccccccgc cacgatcgaa 3300
cgggaaatgga taagaggctt gtgggattga cgtgataggg tagggttggc tatactgctg 3360
gtggcgaaact ccaggctaata aatctgaagc gcttgggaagt tgtaggaggg atcctatggc 3420
tcgtgaagcg gttatcgccg aagtatcgac tcaactatca gaggtagtgg gcgtcatcga 3480
gcgccatctc gaaccgacgt tgctggccgt acatttgtac ggctccgcag tggatggcgg 3540
cctgaagcca cacagtgata ttgatttctt ggttacggtg accgtaaggc ttgatgaaac 3600
aacgcggcga gctttgatca acgacctttt ggaaacttcg gcttccccctg gagagagcga 3660
gattctccgc gctgtagaag tcaccattgt tgtgcaogac gacatcattc cgtggcgctta 3720
tccagctaag cgcgaactgc aatttggaga atggcagcgc aatgacattc ttgcaggtat 3780
cttcgagcca gccacgatcg acattgatct ggctatcttg ctgacaaaag caagagaaca 3840
tagcgttgc tggtaggtc cagcggcgga ggaactctt gatccggttc ctgaacagga 3900
tctatttgag gcgctaataa aaacctaac gctatggaac tcgcccctcg actgggctgg 3960
cgatgagcga aatgtagtgc ttacgttctc ccgcttttgg tacagcgcag taaccggcaa 4020
aatcgcgcgc aaggatgtcg ctgccgactg ggcaatggag cgcctgcgg cccagtatca 4080
gcccgtcata ottgaagcta gacaggctta tcttggacaa gaagaagatc gcttggcctc 4140
ggcgcagat cagttggaag aatttgtcca ctacgtgaaa ggcgagatca ctaaggtagt 4200
tggcaataaa ctcgagacta agtggataaa attagataga aaaaaggctt aaataaaaaa 4260
gaagagaaat agaaagatca aaaaatcagtt acgaaaatgc agtaattctt cttttttctt 4320
ctaattgatt gcaattaaac tcgtctcaat ctgaaaaaag attgagccga gtttaaatag 4380
atcttgatac gatcatgaga cttgacaaat cgggattcct ctattctata tatttagaag 4440
atataaagg atataacaat aaataaatac aatctagaag ctcgatgatc cttaggattg 4500
gtatattctt ttctatctg tagttttagt ttccctgaa tcaagccaag tatcacacct 4560
ctttctacc atcctgtata ttgtccctt tgttccgtg tgaatataga ccttaattta 4620

```

图 8-B

```

ttacttattt ttttattaaa ttttagattt gttagtattt agatattagt attagacgag 4680
atthttacgaa acaattattt ttttatttct ttataggaga ggacaaatct cttttttcga 4740
tgcgaatttg acacgacata ggagaagccg ccotthatta aaaattatat tatttttaa 4800
aatataaagg gggttccaac atattaatat atagtgaagt gttccccag attcagaact 4860
ttttttcaat actcacaatc cttattagtt aataatccta gtgattggat ttctatgctt 4920
agtctgatag gaaataagat attcaataa ataattttat agcgaatgac tattcatcta 4980
ttgtatthtc atgcaaatag ggggcaagaa aactotatgg aaagatggtg gtttaattcg 5040
atgttgthta agaaggagtt cgaacgcagg tgtgggctaa ataaatcaat gggcagctt 5100
ggctctattg aaaaatccaa tgaagatcca aatcgaaaag tgaaaaacat tcatagttgg 5160
aggaatcgtg acaattctag ttgcagtaat gttgattatt tattcggcgt taaagacatt 5220
cggaatttca tctctgatga cactthttta gttagtata ggaatggaga cagtatttcc 5280
atctatthtg atattgaaaa tcatatthtt gagattgaca acgatcattc ttttctgagt 5340
gaactagaaa gttctthttta tagttatcga aactcgaatt atcggaataa tggatttagg 5400
ggcgaagatc cctactataa ttcttactg tatgatactc aatatagttg gaataatcac 5460
attaatagtt gcattgatag ttatcttcag tctcaaatct gtabagatac ttccattata 5520
agtggtagtg agaattacgg tgacagttac atthtatagg ccgtttgtgg tgggtaagt 5580
cgaaatagta gtgaaaacga gggttccagt agacgaactc gcacgaaggg cagtgattta 5640
actataagag aaagttctaa tgatctcgag gtaactcaaa aatacaggca tttgtgggtt 5700
caatgcaaaa atthgtatgg atthaaattat aagaaatttt tgaatcaaa aatgaatatt 5760
tgtgaacaat gtggatatca tttgaaaatg agtagttcag atagaattcg taatcatggt 5820
catagctgth tctctgtgta aattgthtct cgtccacaat tccacacaac atacgagccg 5880
gaagcataaa gtgtaaaagcc tggggtgctt aatgagtgag ctaactcaca ttaattgctt 5940
tgcgctcact gcccgtthtc cagtccggaa acctgtcgtg ccagctgcat taatgaatcg 6000
gccaacgcgc ggggagaggc ggthtgctga ttggcgctc ttccgcttcc tccgctcactg 6060
actcgtcgc ctcggtcgtt cggctgcggc gagcggatc agctcactca aaggcggtaa 6120
taccgthtct cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 6180
aaaaggccag gaaccgtaaa aaggccgcgt tgotggcgtt tttccatagg ctccgcccc 6240
ctgacgagca tcacaaaaat cgcgctcaa gtcagaggtg gcaaaaaccg acaggactat 6300
aaagatacca ggcgtthtcc cctggaagct ccctcgtgcg ctctcctgth ccgaccctgc 6360
cgcttaccgg atacctgtcc gcctthtctc ctccgggaa cgtggcgtt tctcatagct 6420
cacgctgtag gtatctcagt tccgtgtagg tccgtcgtc caagctgggc tgtgtgcagc 6480
aaccocccgt tcagcccgac cgtcgcgct tatccggtaa ctatcgtctt gactccaacc 6540
cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga 6600
ggtatgtagg cgggtctaca gagthcttga agtggtggcc taactacggc tactactagaa 6660
ggacagttat tggatctctc gctctgctga agccagttac ctccgaaaa agagttggta 6720
gctcttgatc cggcaaaaa accaccgctg gtagcgggtg tttttttgth tgcaagcagc 6780
agattaccgc cagaaaaaaa ggatctcang aagatcctth gatctthtct acggggtctg 6840
acgctcagtg gaaccgaaaa tcacgttaag ggattthgth catgagatta tcaaaaagga 6900

```

图 8-C

```

tcttcaccta gatcctttha aattaanaaat gaagttttaa atcaatctaa agtatatatg 6960
agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct 7020
gtctatttcg ttcattccata gttgcctgac tccccgtcgt gtagataact acgatacggg 7080
agggcctacc atctggcccc agtgctgcaa tgatacgcg agaocccagc tcaccggctc 7140
cagatattatc agcaataaac cagccagccc gaagggccga gcgcagaagt ggtcctgcaa 7200
ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc 7260
cagttaatag tttgcgcaac gttggtgcca ttgctacagg catcgtggtg tcacgctcgt 7320
cgtttggtat ggcttcattc agctccggtt cccaacgata aaggcgagtt acatgatccc 7380
ccatggttg gcaaaaaagcg gttagctcct tcggtcctcc gatcgttgc agaagtaagt 7440
tggccgcagt gttatcactc atgggtatgg cagcactgca taattctctt actgtcatgc 7500
catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt 7560
gtatgcggcg accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata 7620
gcagaacttt aaaagtgtc atcattggaa aacggtcttc ggggcgaaa ctctcaagga 7680
tcttaccgct gttgagatcc agttcgatgt aaccactcg tgcaccaac tgatcttcag 7740
catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa 7800
aaaaggggat aagggcgaca cggaaatggt gaatactcat actcttcctt tttcaatatt 7860
attgaagcat ttatcaggtt tattgtctca tgagcggata catatttgaa tgtattttaga 7920
aaaataaaca aataggggtt ccgcgcacat tccccgaaa agtgcacct gacgtctaag 7980
aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg ccctttcgtc 8040

```

图 8-D



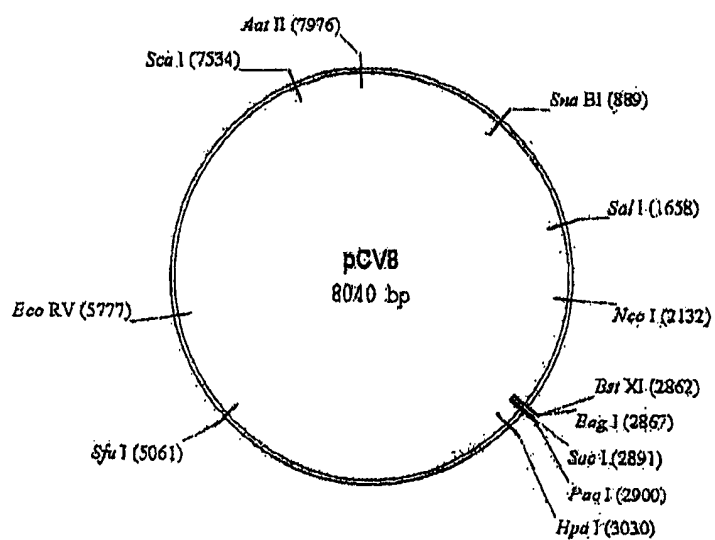


图 9

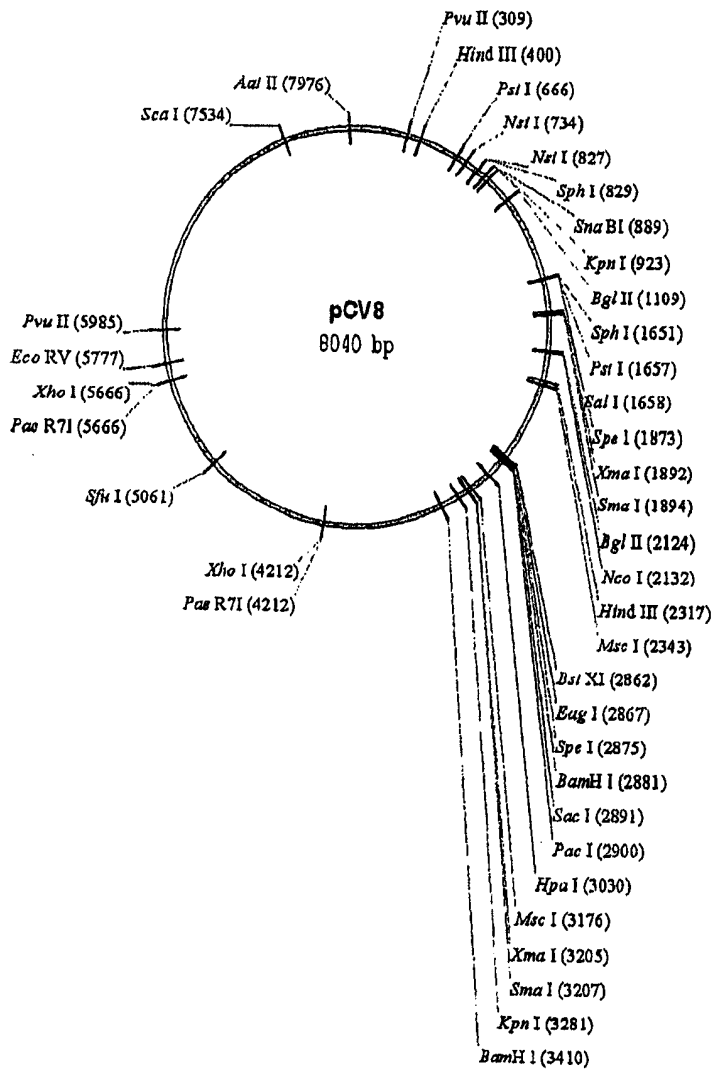


图 10

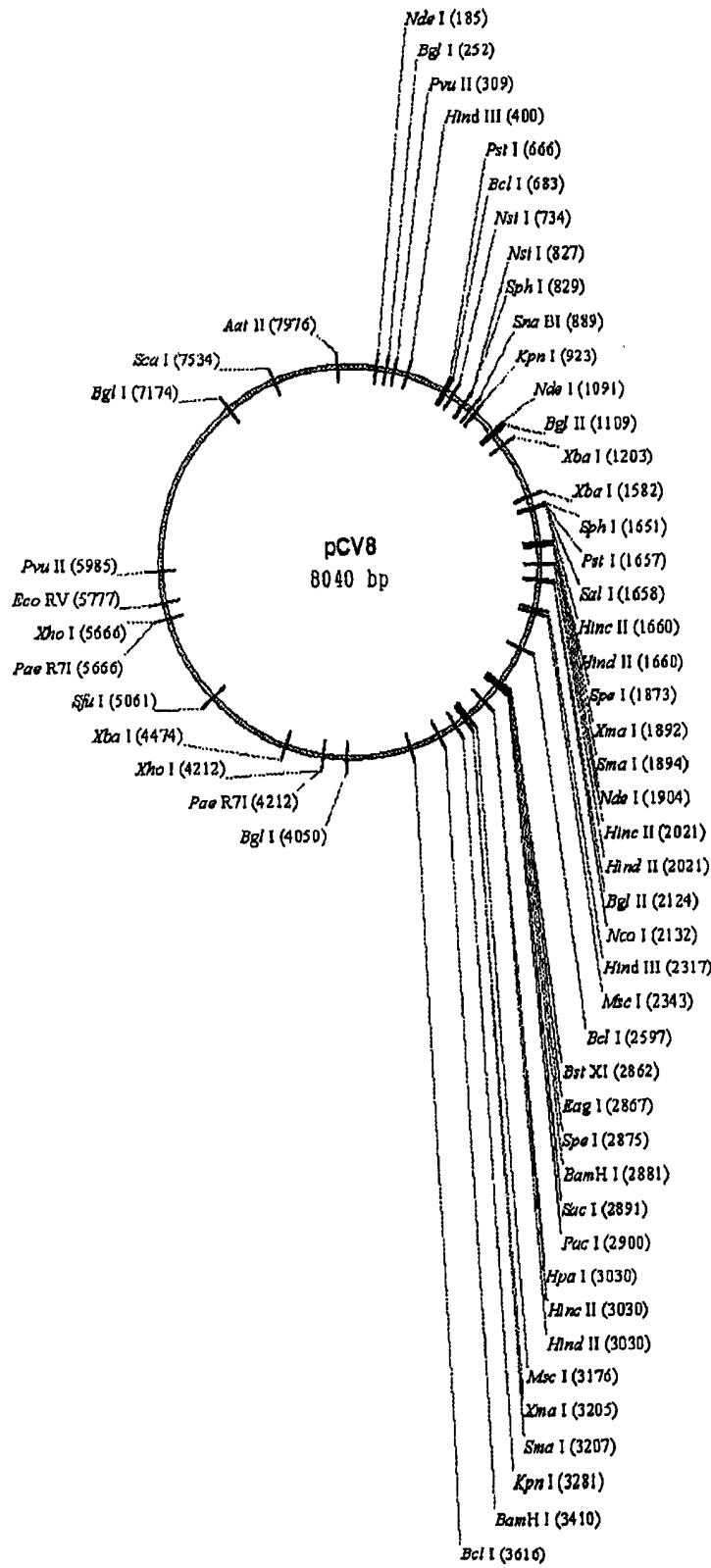


图 11

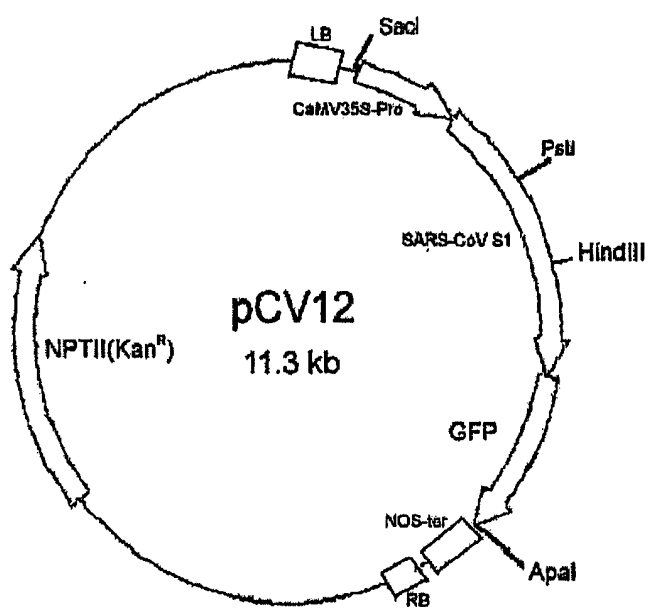


图 12

```

atggcagaca acggtactat taccgttgag gagcttaaac aactcctgga acaatggaac 60
ctagtaatag gtttcctatt cctagcctgg attatgttac tacaatttgc ctattctaatt 120
cggaacaggt ttttgatcat aataaagctt gtttcctctt ggcctttgtg gccagtaaca 180
cttgcttgtt ttgtgcttgc tgcgtctac agaattaatt gggtagctgg cgggattgcg 240
attgcaatgg cttgtattgt aggcttgatg tggetttagct acttogttgc ttccttcagg 300
ctgtttgctc gtacccgctc aatgtggca ttcaaccag aaacaaacat tcttctcaat 360
gtgcctctcc gggggacaat tgtgaccaga ccgctcatgg aaagtgaact tgtcattggt 420
gctgtgatca ttctgggtca cttgcgaatg gccggacact ccctagggcg ctgtgacatt 480
aaggacctgc caaaagagat cactgtggct acatcacgaa cgctttctta ttacaaatta 540
ggagcgtcgc agcgtgtagg cactgattca ggttttgctg catacaaccg ctaccgtatt 600
ggaaactata aattaaatac agaccacgcc ggtagcaacg acaatattgc tttgctagta 660
cagtaa

```

图 13A

```

1 MADNGTITVE ELKQLLEQWN LVIGFLFLAW IMLLQFAYSN RNRFLYIYIKL VFLWLLWPVT
61 LACFVLAAVY RINWVTGGIA IAMACIVGLM WLSYFVASFR LFARTSMWS FNPETNILLN
121 VPLRGTIVTR PLMESELVIG AVIIRGHLRM AGHSLGRCDI KDLPEITVA TSRTLSYYKL
181 GASQRVGTDS GFAAYNRYRI GNYKLNTHA GSNDNIALLV Q

```

图 13B

```

atgtttattt tcttattatt tcttactctc actagtggta gtgacctga ccggtgcacc 60
acttttgatg atgttcaagc tcctaattac actcaacata cttcatctat gaggggggtt 120
tactatcctg atgaaatttt tagatcagac actctttatt taactcagga tttattttctt 180
ccattttatt ctaatggttac agggtttcat actattaatc atacgttttg caaccctgtc 240
atacctttta aggatgggat ttattttgct gccacagaga aatcaaatgt tgtccgtggg 300
tgggtttttg gttctaccat gaacaacaag tcacagtegg tgattattat taacaattct 360
actaatgttg ttatacgagc atgtaacttt gaattgtgtg acaacccttt ctttgcgtgt 420
tctaaaccca tgggtacaca gacacatact atgatattcg ataatgcatt taattgcact 480
ttcgagtaca tatctgatgc cttttcgctt gatgtttcag aaaagtcagg taattttaa 540
cacttacgag agtttggtgt taaaaataaa gatgggtttc tctatgttta taagggtat 600
caacctatag atgtagtctg tgatctacct tctggtttta acactttgaa acctattttt 660
aagttgcttc ttggtattaa cattacaaat tttagagcca ttcttacagc cttttcacct 720
gctcaagaca tttggggcac gtcagctgca gcctattttg ttggctattt aaagccaact 780
acatttatgc tcaagtatga tgaaaatggt acaatcacag atgctgttga ttgttctcaa 840
aatccacttg ctgaactcaa atgctctgtt aagagctttg agattgcaa aggaatttac 900
cagacctcta atttcagggt tgttccctca ggagatgttg tgagattccc taattattca 960
aacttggtgc cttttggaga ggtttttaat gctactaaat tcccttctgt ctatgcattg 1020
gagagaaaaa aaatttctaa ttgtggtgct gattactctg tgctctacaa ctcaacattt 1080
ttttcaacct ttaagtgtca tggcgtttct gccactaagt tgaatgatct ttgcttctcc 1140
aatgtctatg cagattcttt tgtagtcaag ggagatgatg taagacaaat agcgcaggga 1200
caaactgggtg ttattgctga ttataattat aaattgccag atgatttcat gggttggttc 1260
cttgcctgga atactaggaa cattgatgct acttcaactg gtaattataa ttataaatat 1320
aggtatctta gacatggcaa gcttagggcc tttgagagag acatatactaa tgtgcctttc 1380
tccctgatg gcaaaccttg caccaccact gctcttaatt gttattggcc attaaatgat 1440
tatggttttt acaccactac tggcattggc taccacactt acagagttgt agtactttct 1500
tttgaacttt taaatgcacc ggccaoggtt tgtggaccda aattatccac tgaccttatt 1560
aagaaccagt gtgtcaattt taattttaat ggactcactg gtaactgggt gtttaactcct 1620
tcttcaaaga gatttcaacc atttcaacaa tttggcogtg atgtttctga tttcaotgat 1680
tccgttcgag atcctaaaac atctgaaata ttagacattt caccttgccg ttttgggggt 1740
gtaagtgtaa ttacacctgg aacaaatgct tcatctgaag ttgctgttct atatcaagat 1800
gttaactgca ctgatgttct tacagcaatt catgcagatc aactcacacc agcttggcgc 1860
atataattca ctggaaacaa tgtattccag actcaagcag gctgtcttat aggagctgag 1920
catgtcgaca cttcttatga gtgcgacatt cctattggag ctggcatttg tgctagttac 1980
catacagttt ctttattacg tagtactagc caaaaatcta ttgtggctta tactatgtct 2040
ttagggtgctg atagttcaat tgcttactct aataacacca ttgctatacc tactaacttt 2100
tcaattagca ttactacaga agtaatgcct gtttctatgg ctaaaacctc cgtagattgt 2160
aatagtaca tctgcggaga ttctactgaa tgtgctaatt tgcttctcca atatggtagc 2220
ttttgcacac aactaaatcg tgcactctca ggtattgctg ctgaacagga tgcgaacaca 2280

```

图 14-A

```

cgtgaagtgt tcgctcaagt caaacaaatg tacaaaaccc caactttgaa atattttgg 2340
ggttttaatt ttccacaaat attacctgac cctctaaagc caactaagag gtcttttatt 2400
gaggacttgc tctttaataa ggtgacactc gctgatgctg gcttcatgaa gcaatatggc 2460
gaatgcctag gtgatattaa tgctagagat ctcatctgtg cgcagaagtt caatggactt 2520
acagtgttgc cacctctgct cactgatgat atgattgctg cctacactgc tgctctagtt 2580
agtgttactg ccactgctgg atggacattt ggtgctggcg ctgctcttca aatacctttt 2640
gctatgcaaa tggcatatag gttcaatggc attggagtta cccaaaatgt tctctatgag 2700
aaccaaaaac aantcgccaa ccaatttaac aaggcgatta gtcaaattca agaatacctt 2760
acaacaacat caactgcatt gggcaagctg caagacgttg ttaaccagaa tgctcaagca 2820
ttaaacacac ttgttaaaca acttagctct aattttggtg caatttcaag tgtgctaaat 2880
gatacctttt cgcgacttga taaagtcgag gcggagggtac aaattgacag gttaattaca 2940
ggcagacttc aaagccttca aacctatgta acacaacaac taatcagggc tgctgaaatc 3000
agggcttctg ctaatcttgc tgctactaaa atgtctgagt gtgttcttgg acaatcaaaa 3060
agagttagct tttgtgaaa gggctaccac cttatgtcct tcccacaagc agccccgcac 3120
gggtgtgtct tcctacatgt cacgtatgtg ccatcccagg agaggaaact caccacagcg 3180
ccagcaatth gtcataagag caaagcatac ttccctcgtg aagggtgttt tgtgtttaat 3240
ggcacttctt ggtttattac acagaggaac ttcttttctc cacaaataat tactacagac 3300
aatacatttg tctcaggaaa ttgtgatgtc gttattggca tcattaacaa cacagtttat 3360
gatcctctgc aacctgagct tgactcattc aaagaagagc tggacaagta cttcaaaaat 3420
catacatcac cagatgttga tcttggcgac atttcaggca ttaacgcttc tgtcgtcaac 3480
attcaaaaag aaattgaccg cctcaatgag gtcgctaaaa atttaaatga atcactcatt 3540
gaccttcaag aattgggaaa atatgagcaa tatattaaat ggccttggta tgtttggctc 3600
ggcttcattg ctggactaat tgccatcgtc atggttacia tcttgctttg ttgcatgact 3660
agttgttgca gttgcctcaa gggcgcctgc tcttgtggtt cttgctgcaa gtttgatgag 3720
gatgactctg agccagttct caagggtgtc aaattacatt acacataa 3768

```

图 14-B

1 MFIFLLFLTL TSGSDLRCT TFDDVQAPNY TQHTSSMRGV YYPDEIFRSD TLYLTQDLFL  
 61 PFYSNVTGFH TINHTFGNPV IPFKDGIYFA ATEKSNVVRG WVFGSTMNKN SQSVIIINNS  
 121 TNVVIRACNF ELCDNPFPAV SKPMGTQHT MIFDNAFNCT FEYISDAFSL DVSEKSGNFK  
 181 HLREFVFKNK DGFLYVYKGY QPIDVVRDLP SGFNTLKPIF KLPLGINITN FRAILTAFSP  
 241 AQDIWGTSAA AYFVGYLKPT TFMLKYDENG TITDAVDCSQ NPLAELKCSV KSFEIDKGIY  
 301 QTSNFRVVPV GDVVRFPNIT NLCPFGEVFN ATKFPVYAW ERKKISNCVA DYSVLYNSTF  
 361 FSTFKCYGVS ATKLNDLCFS NVYADSFVVK GDDVRQIAPG QTGVIADYNY KLPDDPFMGCV  
 421 LWNTRNIDA TSTGNVNYKY RYLRHGKLRP FERDISNVPF SPDGKPCPTP ALNCYWPLND  
 481 YGFYTTTGIG YQPYRVVLS FELLNAPATV CGPKLSTDLI KNQCVNFNFN GLTGTGVLTP  
 541 SSKRFQPFQ FGRDVSDFTD SVRDPKTSEI LDISPCSFSG VSVITPGTNA SSEVAVLYQD  
 601 VNCTDVSTAI HADQLTPAWR IYSTGNVVFQ TQAGCLIGAE HVDTSYECDI PIGAGICASY  
 661 HTVSLLRSTS QKSIVAYTMS LGADSSIAYS NNTIAIPTNF SISITTEVMP VSMAKTSVDC  
 721 NMYICGDSTE CANLLQYGS FCTQLNRALS GIAAEQDRNT REVFAQVKQM YKTPTLKYFG  
 781 GENFSQILPD PLKPTKRSFI EDLLFNKVTI ADAGFMKQYG ECLGDINARD LICAQKFNGL  
 841 TVLPPLTDD MIAAYTAALV SGTATAGWTF GAGAALQIPF AMQMAYRFNG IGVTQNVLYE  
 901 NQKQIANQFN KAISQIQESL TTTSTALGKL QDVVNQNAQA LNTLVKQLSS NFGAISSVLN  
 961 DILSRDLKVE AEVQIDRLIT GRLQSLQTYV TQQLIRAAEI RASANLAATK MSECVLGQSK  
 1021 RVDFCGKGYH LMSFPQAAPH GVVFLHVTYV PSQERNFTTA PAICHEGKAY FPREGVVFVN  
 1081 GTSWFITQRN FFSFQIITD NTFVSGNCDV VIGIINNTVY DPLQPELDSF KEELDKYFKN  
 1141 HTSPDVLGD ISGINASVVN IQKEIDRLNE VAKNLNESLI DLQELGKYEQ YIKWPVYVWL  
 1201 GFIAGLIAIV MVTILLCCMT SCCSCLKGAC SCGSCCKFDE DDSEPVLKGV KLHYT

图 15



```

1 - ATATTAGGTTTTTACCTACCCAGGAAAAGCCAACCAACCTCGATCTCTTGTAGATCTGTT - 60
- I L G F Y L P R K S Q P T S I S C R S V
- Y * V F T Y P G K A N Q P R S L V D L F
- I R F L P T Q E K P T N L D L L * I C S
61 - CTCTAAACGAACTTTAAAATCTGTGTAGCTGTGCTCGGCTGCATGCCTAGTGACCTAC - 120
- L * T N F K I C V A V A R L H A * C T Y
- S K R T L K S V * L S L G C M P S A P T
- L N E L * N L C S C R S A A C L V H L R
121 - GCAGTATAACAATAATAAATTTACTGTGCGTTGACAAGAAACGAGTAACTCGTCCCTCT - 180
- A V * T I I N F T V V D K K R V T R P S
- Q Y K Q * * I L L S L T R N E * L V P L
- S I N N N K F Y C C R * Q E T S N S S L F
181 - TCTGCAGACTGCTTACGGTTTCGTCCGTGTTGACGTCATCAGCATACCTAGGTTTC - 240
- S A D C L R F R P C C S R S S A Y L G F
- L Q T A Y G F V R V A V D H Q H T * V S
- C R L L T V S S V L Q S I I S I P R F R
241 - GTCCGGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGTCTTGGTGTCAACGAGAAAACA - 300
- V R V * P K G K M E S L V L G V N E K T
- S G C D R K V R W R A L F L V S T R K H
- P G V T E R * D G E P C S W C Q R E N T
301 - CACGTCCAACCTAGTTTGCCTTCCCTCAGGTTAGAGACGTGCTAGTGCCTGGCTTCGGG - 360
- H V Q L S L P V L Q V R D V L V R G F G
- T S N S V C L S F R L E T C * C V A S G
- R P T Q F A C P S G * R R A S A W L R G
361 - GACTCTGTGGAGAGGCCCTATCGGAGGCACGTGAACACCTCAAAAATGGCACTTGTGGT - 420
- D S V E E A L S E A R E H L K N G T C G
- T L W K R P Y R R H V N T S K M A L V V
- L C G R G P I G G T * T P Q K W H L W S
421 - CTAGTAGAGCTGGAAAAGGCGTACTGCCCCAGCTTGAACAGCCCTATGTGTTTCATAAA - 480
- L V E L E K G V L P Q L E Q P Y V F I K
- * * S W K K A Y C P S L N S P M C S L N
- S R A G K R R T A P A * T A L C V H * T
481 - CGTTCGTAGTCCCTAAGCACCAATCACGGCCACAAGGTCGTTGAGCTGGTTGCAGAAATG - 540
- R S D A L S T N H G H K V V E L V A E M
- V L M P * A P I T A T R S L S W L Q K W
- F * C L K H Q S R P Q G R * A G C R N G
541 - GACGGCATTAGTACGGTTCGTAGCGGTATAACACTGGGAGTACTCGTGCCACATGTGGGC - 600
- D G I Q Y G R S G I T L G V L V P H V G
- T A F S T V V A V * H W E Y S C H M W A
- R H S V R S * R Y N T G S T R A T C G R
601 - GAAACCCCAATTGCATACCGCAATGTTCTTCTTCGTAAGAACGGTAATAAGGGAGCCGGT - 660
- E T P I A Y R N V L L R K N G N K G A G
- K P Q L H T A M F F V R T V I R E P V
- N P N C I P Q C S S S * E R * * G S R W
661 - GGTCTAGCTATGGCATCGATCTAAAGTCTTATGACTTAGGTTGACGAGCTTGGCACTGAT - 720
- G H S Y G I D L K S Y D L G D E L G T D
- V I A M A S I * S L M T * V T S L A L I
- S * L W H R S K V L * L R * R A W H * S

```

图 16-A

```

721 - CCCATTGAAGATTATGAACAAAACCTGGAACACTAAGCATGGCAGTGGTGCCTCCGTGAA - 780
- P I E D Y E Q N W N T K H G S G A L R E
- P L K I M N K T G T L S M A V V H S V N
- H * R L * T K L E H * A W Q W C T P * T
781 - CTCACTCGTGAGCTCAATGGAGGTGCAGTCACTCGCTATGTCGACAAACAATTTCTGTGGC - 840
- L T R E L N G G A V T R Y V D N N F C G
- S L V S S M E V Q S L A M S T T I S V A
- H S * A Q W R C S H S L C R Q Q F L W P
841 - CCAGATGGGTACCCTCTTGATTGCATCAAAGATTTTCTCGCACGCGGGCAAGTCAATG - 900
- P D G Y P L D C I K D F L A R A G K S M
- Q M G T L L I A S K I F S H A R A S Q C
- R W V P S * L H Q R F S R T R G Q V N V
901 - TGCCTCTTCCGAACAACCTTGATTACATCGAGTGAAGAGAGGTGTCTACTGCTGCCGT - 960
- C T L S E Q L D Y I E S K R G V Y C C R
- A L F P N N L I T S S R R E V S T A A V
- H S F R T T * L H R V E E R C L L L P *
961 - GACCATGAGCATGAAATTGCCTGGTTCCTGAGCGCTCTGATAAGAGCTACGAGCACCAG -
1020
- D H E H E I A W F T E R S D K S Y E H Q
- T M S M K L P G S L S A L I R A T S T R
- P * A * N C L V H * A L * * E L R A P D
1021 - ACACCCTTCGAAATTAAGAGTGCCAAGAAATTTGACACTTTCAAAGGGGAATGCCCAAG -
1080
- T P F E I K S A K K F D T F K G E C P K
- H P S K L R V P R N L T L S K G N A Q S
- T L R N * E C Q E I * H F Q R G M P K V
1081 - TTTGTGTTTCTCTTAACCTCAAAGTCAAAGTCATTCAACCAGTGTGAAAAGAAAAG -
1140
- F V F P L N S K V K V I Q P R V E K K K
- L C F L L T Q K S K S F N H V L K R K R
- C V S S * L K S Q S H S T T C * K E K D
1141 - ACTGAGGGTTTCATGGGGCGTATACGCTCTGTGTACCCTGTTGCATCTCCACAGGAGTGT -
1200
- T E G F M G R I R S V Y P V A S P Q E C
- L R V S W G V Y A L C T L L H L H R S V
- * G F H G A Y T L C V P C C I S T G V *
1201 - AACAAATATGCACCTTGCTACCTTGATGAAATGTAATCATTGCGATGAAGTTTCATGGCAG -
1260
- N N M H L S T L M K C N H C D E V S W Q
- T I C T C L P * * N V I I A M K F H G R
- Q Y A L V Y L D E M * S L R * S F M A D
1261 - ACGTGCGACTTTCTGAAAGCCACTTGTAACATTGTGGCACTGAAAATTTAGTTATTGAA -
1320
- T C D F L K A T C E H C G T E N L V I E
- R A T F * K P L V N I V A L K I * L L K
- V R L S E S H L * T L W H * K F S Y * R
1321 - GGACCTACTACATGTGGGTACCTACCTACTAATGCTGTAGTGAATGCCATGTCCTGCC -
1380
- G P T T C G Y L P T N A V V K M P C P A
- D L L H V G T Y L L M L * * K C H V L P
- T Y Y M W V P T Y * C C S E N A M S C L
1381 - TGTCAAGACCCAGAGATTGGACCTGAGCATAGTGTTCAGATTATCACACCCTCAAC -
1440
- C Q D P E I G P E H S V A D Y H N H S N
- V K T Q R L D L S I V L Q I I T T T Q T
- S R P R D W T * A * C C R L S Q P L K H
1441 - ATTGAAACTCGACTCCGCAAGGGAGGTAGACTAGATGTTTGGAGGCTGTGTGTTGCC -
1500

```

图 16-B

- I E T R L R K G G R T R C F G G C V F A  
- L K L D S A R E V G L D V L E A V C L P  
- \* N S T P Q G R \* D \* M F W R L C V C L

```

1501 - TATGTTGGCTGCTATAATAAGCGTGCCTACTGGGTTCTCGTGCTAGTGCTGATATTGGC -
1560 - Y V G C Y N K R A Y W V P R A S A D I G
- M L A A I I S V P T G F L V L V L I L A
- C W L L * * A C L L G S S C * C * Y W L
1561 - TCAGGCCATACTGGCATTACTGGTGACAAATGTGGAGACCTTGAATGAGGATCTCCTTGAG -
1620 - S G H T G I T G D N V E T L N E D L L E
- Q A I L A L L V T M W R P * M R I S L R
- R P Y W H Y W * Q C G D L E * G S P * D
1621 - ATACTGAGTCGTGAACGTGTTAACATTAACATTGTTGGCGATTTTCATTGATGAAGAG -
1680 - I L S R E R V N I N I V G D F H L N E E
- Y * V V N V L T L T L L A I F I * M K R
- T E S * T C * H * H C W R F S F E * R G
1681 - GTTGCCATCATTTTGGCACTTTCTCTGCTTCTACAAGTGCCTTTATTGACACTATAAAG -
1740 - V A I I L A S F S A S T S A F I D T I K
- L P S F W H L S L L L Q V P L L T L * R
- C H H F G I F L C F Y K C L Y * H Y K E
1741 - AGTCTTGATTACAAGTCTTTCAAACCATTTGTTGAGTCCTGCGGTAACATAAAGTTACC -
1800 - S L D Y K S F K T I V E S C G N Y K V T
- V L I T S L S K P L L S P A V T I K L P
- S * L Q V F Q N H C * V L R * L * S Y Q
1801 - AAGGGAAGCCCGTAAAGGTGCTTGGAACATTGGACAACAGAGATCAGTTTAAACACCA -
1860 - K G K P V K G A W N I G Q Q R S V L T P
- R E S P * K V L G T L D N R D Q F * H H
- G K A R K R C L E H W T T E I S F N T T
1861 - CTGTGTGGTTTTCCCTCACAGGCTGCTGGTGTATCAGATCAATTTTTCGCGCACACTT -
1920 - L C G F P S Q A A G V I R S I F A R T L
- C V V F P H R L L V L S D Q F L R A H L
- V W F S L T G C W C Y Q I N F C A H T *
1921 - GATGCAGCAAACCACTCAATTCCTGATTTGCAAAGAGCAGCTGTCACCATACTTGATGGT -
1980 - D A A N H S I P D L Q R A A V T I L D G
- M Q Q T T Q F L I C K E Q L S P Y L M V
- C S K P L N S * F A K S S C H H T * W Y
1981 - ATTTCTGACAGTCATTACGCTTTGTCGACGCCATGGTTTATACTTCAGACCTGCTCACC -
2040 - I S E Q S L R L V D A M V Y T S D L L T
- F L N S H Y V L S T P W F I L Q T C S P
- F * T V I T S C R R H G L Y F R P A H Q
2041 - AACAGTGTCAATATATGGCATATGTAACCTGGTGGTCTGTACAACAGACTTCTCAGTGG -
2100 - N S V I I M A Y V T G G L V Q Q T S Q W
- T V S L L W H M * L V V L Y N R L L S G
- Q C H Y Y G I C N W W S C T T D F S V V
2101 - TTGTCTAATCTTTTGGGCACTACTGTTGAAAACTCAGGCCTATCTTTGAATGGATTGAG -
2160 - L S N L L G T T V E K L R P I F E W I E
- C L I F W A L L L K N S G L S L N G L R
- V * S F G H Y C * K T Q A Y L * M D * G
2161 - GCGAAACTTAGTGCAGGAGTTGAATTTCTCAAGGATGCTGGGAGATTCTCAAATTTCTC -
2220 - A K L S A G V E F L K D A W E I L K F L

```

图 16-C

---

- R N L V Q E L N F S R M L G R F S N F S  
- E T \* C R S \* I S Q G C L G D S Q I S H  
2221 - ATTACAGGTGTTTTGACATCGTCAAGGGTCAAATACAGGTTGCTTCAGATAACATCAAG -  
2280  
- I T G V F D I V K G Q I Q V A S D N I K  
- L Q V F L T S S R V K Y R L L Q I T S R  
- Y R C F \* H R Q G S N T G C F R \* H Q G

```

2281 - GATTGTGTAATAATGCTTCATTGATGTTCTTAACAAGGCACTCGAAATGTGCATTGATCAA -
2340
- D C V K C F I D V V N K A L E M C I D Q
- I V * N A S L M L L T R H S K C A L I K
- L C K M L H * C C * Q G T R N V H * S S
2341 - GTCACATATCGCTGGCGCAAAGTTGCGATCACTCAACTTAGGTGAAGTCTTCATCGCTCAA -
2400
- V T I A G A K L R S L N L G E V F I A Q
- S L S L A Q S C D H S T * V K S S S L K
- H Y R W R K V A I T Q L R * S L H R S K
2401 - AGCAAGGGACTTTACCGTCAGTGATACGTGGCAAGGAGCAGCTGCAACTACTCATGCCT -
2460
- S K G L Y R Q C I R G K E Q L Q L L M P
- A R D F T V S V Y V A R S S C N Y S C L
- Q G T L P S V Y T W Q G A A A T T H A S
2461 - CTTAAGGCACCAAAGAAGTAACCTTTCTTGAAGGTGATTCACATGACACAGTACTTACC -
2520
- L K A P K E V T F L E G D S H D T V L T
- L R H Q K K * P F L K V I H M T Q Y L P
- * G T K R S N L S * R * F T * H S T Y L
2521 - TCTGAGGAGGTTGTTCTCAAGAACGGTGAACCTCGAAGCACTCGAGACGCCCGTTGATAGC -
2580
- S E E V V L K N G E L E A L E T P V D S
- L R R L F S R T V N S K H S R R P L I A
- * G G C S Q E R * T R S T R D A R * * L
2581 - TTCACAAATGGAGCTATCGTCGGCACACCACTCTGTGTAATGGCCTCATGCTCTTAGAG -
2640
- F T N G A I V G T P V C V N G L M L L E
- S Q M E L S S A H Q S V * M A S C S * R
- H K W S Y R R H T S L C K W P H A L R D
2641 - ATTAAAGGACAAAGAACAATACTGCGCATTGTCTCCTGGTTTACTGGCTACAAACAATGTC -
2700
- I K D K E Q Y C A L S P G L L A T N N V
- L R T K N N T A H C L L V Y W L Q T M S
- * G Q R T I L R I V S W F T G Y K Q C L
2701 - TTTGCTTAAAAGGGGGTGCACCAATTAAGGTGTAACTTTGGAGAAGATACTGTTTGG -
2760
- F R L K G G A P I K G V T F G E D T V W
- F A * K G V H Q L K V * P L E K I L F G
- S L K R G C T N * R C N L W R R Y C L G
2761 - GAAGTTCAAGGTTACAAGAATGTGAGAATCACATTTGAGCTTGATGAACGTTTGACAAA -
2820
- E V Q G Y K N V R I T F E L D E R V D K
- K F K V T R M * E S H L S L M N V L T K
- S S R L Q E C E N H I * A * * T C * Q S
2821 - GTGCTTAATGAAAAGTGCTCTGTCTACACTGTTGAATCCGGTACCGAAGTTACTGAGTTT -
2880
- V L N E K C S V Y T V E S G T E V T E F
- C L M K S A L S T L L N P V P K L L S L
- A * * K V L C L H C * I R Y R S Y * V C
2881 - GCATGTGTTGTAGCAGAGGCTGTTGTGAAGACTTTACAACCAGTTTCTGATCTCCTTACC -
2940
- A C V V A E A V V K T L Q P V S D L L T
- H V L * Q R L L * R L Y N Q F L I S L P
- M C C S R G C C E D F T T S F * S P Y Q
2941 - AACATGGGTATTGATCTTGATGAGTGGAGTGTAGCTACATTCTACTTATTTGATGATGCT -
3000
- N M G I D L D E W S V A T F Y L F D D A
- T W V L I L M S G V * L H S T Y L M M L

```

图 16-D

---

- H G Y \* S \* \* V E C S Y I L L I \* \* C W  
3001 - GGTGAAGAAAACCTTTTCATCACGTATGTATTGTCCTTTACCCTCCAGATGAGGAAGAA -  
3060  
- G E E N F S S R M Y C S F Y P P D E E E  
- V K K T F H H V C I V P F T L Q M R K K  
- \* R K L F I T Y V L F L L P S R \* G R R

```

3061 - GAGGACGATGCAGAGTGTGAGGAAGAAGAAATTGATGAAACCTGTGAACATGAGTACGGT -
3120 - E D D A E C E E E E I D E T C E H E Y G
- R T M Q S V R K K K L M K P V N M S T V
- G R C R V * G R R N * * N L * T * V R Y
3121 - ACAGAGGATGATTATCAAGGTCTCCCTCTGGAATTTGGTGCCTCAGCTGAAACAGTTTCA -
3180 - T E D D Y Q G L P L E F G A S A E T V R
- Q R M I I K V S L W N L V P Q L K Q F E
- R G * L S R S P S G I W C L S * N S S S
3181 - GTTGAGGAAGAAGAAGAGGAAAGACTGGCTGGATGATACTACTGAGCAATCAGAGATTGAG -
3240 - V E E E E E E D W L D D T T E Q S E I E
- L R K K K R K T G W M I L L S N Q R L S
- * G R R R G R L A G * Y Y * A I R D * A
3241 - CCAGAACCAGAACCTACACCTGAAGAACCAGTTAATCAGTTTACTGGTTATTTAAACTT -
3300 - P E P E P T P E E P V N Q F T G Y L K L
- Q N Q N L H L K N Q L I S L L V I * N L
- R T R T Y T * R T S * S V Y W L F K T Y
3301 - ACTGACAATGTTGCCATTAATGTGTTGACATCGTTAAGGAGGCACAAAGTGCTAATCCT -
3360 - T D N V A I K C V D I V K E A Q S A N P
- L T M L P L N V L T S L R R H K V L I L
- * Q C C H * M C * H R * G G T K C * S Y
3361 - ATGGTGATTGTAATGCTGCTAACATACACCTGAAACATGGTGGTGGTGTAGCAGGTGCA -
3420 - M V I V N A A N I H L K H G G G V A G A
- W * L * M L L T Y T * N M V V V * Q V H
- G D C K C C * H T P E T W W W C S R C T
3421 - CTCAACAAGGCAACCAATGGTGCCATGCAAAGGAGAGTGATGATTACATTAAGCTAAAT -
3480 - L N K A T N G A M Q K E S D D Y I K L N
- S T R Q P M V P C K R R V M I T L S * M
- Q Q G N Q W C H A K G E * * L H * A K W
3481 - GGCCTCTTACAGTAGGAGGTCTTGTGCTTCTGGACATAATCTTGCTAAGAAGTGT -
3540 - G P L T V G G S C L L S G H N L A K K C
- A L L Q * E G L V C F L D I I L L R S V
- P S Y S R R V L F A F W T * S C * E V S
3541 - CTGCATGTTGTTGGACCTAACCTAAATGCAGGTGAGGACATCCAGCTTCTTAAGGCAGCA -
3600 - L H V V G P N L N A G E D I Q L L K A A
- C M L L D L T * M Q V R T S S F L R Q H
- A C C W T * P K C R * G H P A S * G S I
3601 - TATGAAAATTTCAATTCACAGGACATCTTACTTGACCATTGTTGTCAGCAGGCATATTT -
3660 - Y E N F N S Q D I L L A P L L S A G I F
- M K I S I H R T S Y L H H C C Q Q A Y L
- * K F Q F T G H L T C T I V V S R H I W
3661 - GGTGCTAAACCACCTTCAGTCTTTACAAGTGTGCGTGCAGACGGTTTCGTACACAGGTTTAT -
3720 - G A K P L Q S L Q V C V Q T V R T Q V Y
- V L N H F S L Y K C A C R R F V H R F I
- C * T T S V F T S V R A D G S Y T G L Y
3721 - ATTGCAGTCAATGACAAAGCTCTTTATGAGCAGGTTGTCATGGATTATCTTGATAACCTG -
3780 - I A V N D K A L Y E Q V V M D Y L D N L

```

图 16E



---

```
- L Q S M T K L F M S R L S W I I L I T *  
- C S Q * Q S S L * A G C H G L S * * P E  
3781 - AAGCCTAGAGTGAAGCACCTAAACAAGAGGAGCCACCAACACAGAAGATTCCTAAAAC -  
3840  
- K P R V E A P K Q E E P P N T E D S K T  
- S L E W K H L N K R S H Q T Q K I P K L  
- A * S G S T * T R G A T K H R R F Q N *
```

```

3841 - GAGGAGAAATCTGTCGTACAGAAGCCTGTCGATGTGAAGCCAAAAATTAAGCCTGCATT -
3900
- E E K S V V Q K P V D V K P K I K A C I
- R R N L S Y R S L S M * S Q K L R P A L
- G E I C R T E A C R C E A K N * G L H *
3901 - GATGAGGTTACCACAACACTGGAAGAACTAAGTTTCTTACCAATAAGTTACTCTTGTTT -
3960
- D E V T T T L E E T K F L T N K L L L F
- M R L P Q H W K K L S F L P I S Y S C L
- * G Y H N T G R N * V S Y Q * V T L V C
3961 - GCTGATATCAATGGTAAGCTTTACCATGATTCTCAGAACATGCTTAGAGGTGAAGATATG -
4020
- A D I N G K L Y H D S Q N M L R G E D M
- L I S M V S F T M I L R T C L E V K I C
- * Y Q W * A L P * F S E H A * R * R Y V
4021 - TCTTCTTGAGAAGGATGCACCTTACATGGTAGGTGATGTTATCACTAGTGGTGATATC -
4080
- S F L E K D A P Y M V G D V I T S G D I
- L S L R R M H L T W * V M L S L V V I S
- F P * E G C T L H G R * C Y H * W * Y H
4081 - ACTTGTGTGTAATACCCTCCAAAAGGCTGGTGGCACTACTGAGATGCTCTCAAGAGCT -
4140
- T C V V I P S K K A G G T T E M L S R A
- L V L * Y P P K R L V A L L R C S Q E L
- L C C N T L Q K G W W H Y * D A L K S F
4141 - TTGAAGAAAGTGCCAGTTGATGAGTATATAACCACGTACCCTGGACAAGGATGTGCTGGT -
4200
- L K K V P V D E Y I T T Y P G Q G C A G
- * R K C Q L M S I * P R T L D K D V L V
- E E S A S * * V Y N H V P W T R M C W L
4201 - TATACACTGAGGAAGCTAAGACTGCTCTTAAGAAATGCAAATCTGCATTTTATGTACTA -
4260
- Y T L E E A K T A L K K C K S A F Y V L
- I H L R K L R L L L R N A N L H F M Y Y
- Y T * G S * D C S * E M Q I C I L C T T
4261 - CCTCAGAAGCACCTAATGCTAAGGAAGAGATTCTAGGAAGTGTATCCTGGAATTTGAGA -
4320
- P S E A P N A K E E I L G T V S W N L R
- L Q K H L M L R K R F * E L Y P G I * E
- F R S T * C * G R D S R N C I L E F E R
4321 - GAAATGCTTGCTCATGCTGAAGAGACAAGAAAATTAATGCCTATATGCATGGATGTTAGA -
4380
- E M L A H A E E T R K L M P I C M D V R
- K C L L M L K R Q E N * C L Y A W M L E
- N A C S C * R D K K I N A Y M H G C * S
4381 - GCCATAATGGCAACCATCCAACGTAAGTATAAAGGAATTAATAATCAAGAGGGCATCGTT -
4440
- A I M A T I Q R K Y K G I K I Q E G I V
- P * W Q P S N V S I K E L K F K R A S L
- H N G N H P T * V * R N * N S R G H R *
4441 - GACTATGGTCCGATTCTTCTTTTATACTAGTAAAGAGCCTGTAGCTTCTATTATTACG -
4500
- D Y G V R F F F Y T S K E P V A S I I T
- T M V S D S S F I L V K S L * L L L L R
- L W C P I L L L Y * * R A C S F Y Y Y E
4501 - AAGCTGAACTCTCTAAATGAGCCGCTGTGCACAATGCCAATTGGTTATGTGACACATGGT -
4560

```

图 16-F

- K L N S L N E P L V T M P I G Y V T H G  
- S \* T L \* M S R L S Q C Q L V M \* H M V  
- A E L S K \* A A C H N A N W L C D T W F

```

4561 - TTTAATCTTGAAGAGGCTGCGCGCTGTATGCGTTCTCTTAAAGCTCCTGCCGTAGTGTCA -
4620
- F N L E E A A R C M R S L K A P A V V S
- L I L K R L R A V C V L L K L L P * C Q
- * S * R G C A L Y A F S * S S C R S V S
4621 - GTATCATCACCAGATGCTGTACTACATATAATGGATACCTCACTTCGTCATCAAAGACA -
4680
- V S S P D A V T T Y N G Y L T S S S K T
- Y H H Q M L L L H I M D T S L R H Q R H
- I I T R C C Y Y I * W I P H F V I K D I
4681 - TCTGAGGAGCACTTTGTAGAAACAGTTTCTTTGGCTGGCTCTTACAGAGATTGGTCCTAT -
4740
- S E E H F V E T V S L A G S Y R D W S Y
- L R S T L * K Q F L W L A L T E I G P I
- * G A L C R N S F F G W L L Q R L V L F
4741 - TCAGGACAGCGTACAGAGTTAGGTGTTGAATTTCTTAAAGCGGTGACAAAATTGTGTAC -
4800
- S G Q R T E L G V E F L K R G D K I V Y
- Q D S V Q S * V L N F L S V V T K L C T
- R T A Y R V R C * I S * A W * Q N C V P
4801 - CACACTCTGGAGAGCCCCGTCGAGTTTCATCTTGACGGTGAGTTCTTTCACTTGACAAA -
4860
- H T L E S P V E F H L D G E V L S L D K
- T L W R A P S S F I L T V R F F H L T N
- H S G E P R R V S S * R * G S F T * Q T
4861 - CTAAGAGTCTCTTATCCCTGCGGGAGGTTAAGACTATAAAAAGTTTCACTCAACTGTGGAC -
4920
- L K S L L S L R E V K T I K V F T T V D
- * R V S Y P C G R L R L * K C S Q L W T
- K E S L I P A G G * D Y K S V H N C G Q
4921 - AACACTAATCTCCACACACAGCTTGTGGATATGCTATGACATATGGACAGCAGTTTGGT -
4980
- N T N L H T Q L V D M S M T Y G Q Q F G
- T L I S T H S L W I C L * H M D S S L V
- H * S P H T A C G Y V Y D I W T A V W S
4981 - CCAACATACTGGATGGTGCTGATGTTACAAAATTAAACCTCATGTAATCATGAGGGT -
5040
- P T Y L D G A D V T K I K P H V N H E G
- Q H T W M V L M L Q K L N L M * I M R V
- N I L G W C * C Y K N * T S C K S * G *
5041 - AAGACTTCTTTGTACTACCTAGTGATGACACACTACGTAGTGAAGCTTTTCGAGTACTAC -
5100
- K T F F V L P S D D T L R S E A F E Y Y
- R L S L Y Y L V M T H Y V V K L S S T T
- D F L C T T * * * H T T * * S F R V L P
5101 - CATACTCTTGATGAGAGTTTCTTTGGTAGGTACATGTCTGCTTTAAACCACAAAAGAAA -
5160
- H T L D E S F L G R Y M S A L N H T K K
- I L L M R V F L V G T C L L * T T Q R N
- Y S * * E F S W * V H V C F K P H K E M
5161 - TGGAAATTTCTCAAGTTGGTGGTTAACTTCAATTAATGGGCTGATAACAATTGTTAT -
5220
- W K F P Q V G G L T S I K W A D N N C Y
- G N F L K L V V * L Q L N G L I T I V I
- E I S S S W W F N F N * M G * * Q L L F
5221 - TTGTCTAGTGTTTTATTAGCACTTCAACAGCTTGAAGTCAAATTCATGCACCAGCACTT -
5280
- L S S V L L A L Q Q L E V K F N A P A L

```

图 16-G

- -  
- C L V F Y \* H F N S L K S N S M H Q H F  
- V \* C F I S T S T A \* S Q I Q C T S T S  
5281 - CAAGAGGCTTATTATAGAGCCCGTGCTGGTGATGCTGCTAACTTTTGTGCACTCATACTC  
5340  
- Q E A Y Y R A R A G D A A N F C A L I L  
- K R L I I E P V L V M L L T F V H S Y S  
- R G L L \* S P C W \* C C \* L L C T H T R

```

5341 - GCTTACAGTAATAAACTGTTGGCGAGCTTGGTGATGTCAGAGAACTATGACCCATCTT -
5400 - A Y S N K T V G E L G D V R E T M T H L
- L T V I K L L A S L V M S E K L * P I F
- L Q * * N C W R A W * C Q R N Y D P S S
5401 - CTACAGCATGCTAATTGGAATCTGCAAGCGAGTTCCTAATGTTGGTGTGTAACATTGT -
5460 - L Q H A N L E S A K R V L N V V C K H C
- Y S M L I W N L Q S E F L M W C V N I V
- T A C * F G I C K A S S * C G V * T L W
5461 - GGTCAGAAACTACTACCTTAACGGGTGTAGAAGCTGTGATGTATATGGTACTCTATCT -
5520 - G Q K T T T L T G V E A V M Y M G T L S
- V R K L L P * R V * K L * C I W V L Y L
- S E N Y Y L N G C R S C D V Y G Y S I L
5521 - TATGATAATCTTAAGACAGGTGTTCCATTCCATGTGTGTGGTGTGCTACACAA -
5580 - Y D N L K T G V S I P C V C G R D A T Q
- M I I L R Q V F P F H V C V V V M L H N
- * * S * D R C F H S M C V W S * C Y T I
5581 - TATCTAGTACAACAAGAGTCTTCTTTTGTATGATGCTGCACCACCTGCTGAGTATAAA -
5640 - Y L V Q Q E S S F V M M S A P P A E Y K
- I * Y N K S L L L L * C L H H L L S I N
- S S T T R V F F C Y D V C T T C * V * I
5641 - TTACAGCAAGGTACATTCTTATGTGCGAATGAGTACACTGGTAACTATCAGTGTGGTCAT -
5700 - L Q Q G T F L C A N E Y T G N Y Q C G H
- Y S K V H S Y V R M S T L V T I S V V I
- T A R Y I L M C E * V H W * L S V W S L
5701 - TACACTCATATAACTGCTAAGGAGACCCCTCTATCGTATTGACGGAGCTCACCTTACAAAG -
5760 - Y T H I T A K E T L Y R I D G A H L T K
- T L I * L L R R P S I V L T E L T L Q R
- H S Y N C * G D P L S Y * R S S P Y K D
5761 - ATGTCAGAGTACAAAGGACCAGTGACTGATGTTTCTACAAGGAAACATCTTACACTACA -
5820 - M S E Y K G P V T D V F Y K E T S Y T T
- C Q S T K D Q * L M F S T R K H L T L Q
- V R V Q R T S D * C F L Q G N I L H Y N
5821 - ACCATCAAGCCTGTGTCGTATAAACTCGATGGAGTTACTTACACAGAGATTGAACCAAAA -
5880 - T I K P V S Y K L D G V T Y T E I E P K
- P S S L C R I N S M E L L T Q R L N Q N
- H Q A C V V * T R W S Y L H R D * T K I
5881 - TTGGATGGGTATTATAAAAAGGATAATGCTTACTATACAGAGCAGCCTATAGACCTTGTA -
5940 - L D G Y Y K K D N A Y Y T E Q P I D L V
- W M G I I K R I M L T I Q S S L * T L Y
- G W V L * K G * C L L Y R A A Y R P C T
5941 - CCAACTCAACCATTACCAAATGCGAGTTTGTGATAATTTCAAACCTCACATGTTCTACACA -
6000 - P T Q P L P N A S F D N F K L T C S N T
- Q L N H Y Q M R V L I I S N S H V L T Q
- N S T I T K C E F * * F Q T H M F * H K
6001 - AAATTTGCTGATGATTTAAATCAAATGACAGGCTTCACAAGCCAGCTTCAGGAGAGCTA -
6060 - K F A D D L N Q M T G F T K P A S R E L

```

图 16-H

- N L L M I \* I K \* Q A S Q S Q L H E S Y  
- I C \* \* F K S N D R L H K A S F T R A I  
6061 - T C T G T C A C A T T C T T C C C A G A C T T G A A T G G C G A T G T A G T G G C T A T T G A C T A T A G A C A C T A T  
6120  
- S V T F F P D L N G D V V A I D Y R H Y  
- L S H S S Q T \* M A M \* W L L T I D T I  
- C H I L P R L E W R C S G Y \* L \* T L F

```

6121 - TCAGCGAGTTTCAAGAAAGGTGCTAAATTACTGCATAAGCCAATTGTTTGGCACATTAAC -
6180
- S A S F K K G A K L L H K P I V W H I N
- Q R V S R K V L N Y C I S Q L F G T L T
- S E F Q E R C * I T A * A N C L A H * P
6181 - CAGGCTACAACCAAGACAACGTTCAACCAACACTTGGFGTTTACGTTGTCTTTGGAGT -
6240
- Q A T T K T T F K P N T W C L R C L W S
- R L Q P R Q R S N Q T L G V Y V V F G V
- G Y N Q D N V Q T K H L V F T L S L E Y
6241 - ACAARGCCAGTAGATACTTCAAATTCATTTGAAGTTCTGGCAGTAGAAGACACACAAGGA -
6300
- T K P V D T S N S F E V L A V E D T Q G
- Q S Q * I L Q I H L K F W Q * K T H K E
- K A S R Y F K F I * S S G S R R H T R N
6301 - ATGGACAATCTTGCTTGTGAAAGTCAACAACCCACCTCTGAAGAAGTAGTGGAAAATCCT -
6360
- M D N L A C E S Q Q P T S E E V V E N P
- W T I L L V K V N N P P L K K * W K I L
- G Q S C L * K S T T H L * R S S G K S Y
6361 - ACCATACAGAAGGAAGTCATAGAGTGTGACCTGAAAACCTACCGAAGTTGTAGGCAATGTC -
6420
- T I Q K E V I E C D V K T T E V V G N V
- P Y R R K S * S V T * K L P K L * A M S
- H T E G S H R V * R E N Y R S C R Q C H
6421 - ATACTTAAACCATCAGATGAAGGTGTTAAAGTAACACAAGAGTTAGGTCATGAGGATCTT -
6480
- I L K P S D E G V K V T Q E L G H E D L
- Y L N H Q M K V L K * H K S * V M R I L
- T * T I R * R C * S N T R V R S * G S Y
6481 - ATGGCTGCTTATGTGGAAAACACAAGCATTACCATTAAGAAACCTAATGAGCTTTCACATA -
6540
- M A A Y V E N T S I T I K K P N E L S L
- W L L M W K T Q A L P L R N L M S F H *
- G C L C G K H K H Y H * E T * * A F T S
6541 - GCCTTAGGTTTAAAACAATTGCCACTCATGGTATTGCTGCAATTAATAGTGTTCCTTGG -
6600
- A L G L K T I A T H G I A A I N S V P W
- P * V * K Q L P L M V L L Q L I V F L G
- L R F K N N C H S W Y C C N * * C S L E
6601 - AGTAAAATTTTGGCTTATGTCAAACCATTCTTAGGACAAGCAGCAATTACAACATCAAAT -
6660
- S K I L A Y V K P F L G Q A A I T T S N
- V K F W L M S N H S * D K Q Q L Q H Q I
- * N F G L C Q T I L R T S S N Y N I K L
6661 - TGGCCTAAGAGATTAGCACACGTTGTTTAAACAATTATATGCCTTATGTGTTACATTA -
6720
- C A K R L A Q R V F N N Y M P Y V F T L
- A L R D * H N V C L T I I C L M C L H Y
- R * E I S T T C V * Q L Y A L C V Y I I
6721 - TTGTTCCAATTGTGTACTTTTACTAAAAGTACCAATCTAGAATTAGAGCTTCACTACCT -
6780
- L F Q L C T F T K S T N S R I R A S L P
- C S N C V L L L K V P I L E L E L H Y L
- V P I V Y F Y * K Y Q F * N * S F T T Y
6781 - ACAACTATTGCTAAAATAGTGTTAGAGTGTGCTAAATTATGTTTGGATGCCGGCATT -
6840
- T T I A K N S V K S V A K L C L D A G I

```

图 16-1



---

- Q L L L K I V L R V L L N Y V W M P A L  
- N Y C \* K \* C \* E C C \* I M F G C R H \*  
6841 - AATTATGTGAAGTCACCCAAATTTTCTAAATTGTTCAACAATCGCTATGTTGGCTATTGTTG -  
6900  
- N Y V K S P K F S K L F T I A M W L L L  
- I M \* S H P N F L N C S Q S L C G Y C C  
- L C E V T Q I F \* I V H N R Y V A I V V

```

6901 - TTAAGTATTTGCTTAGGTTCTCTAATCTGTGTAAGTCTGCTTTTGGTGTACTCTTATCT -
6960
- L S I C L G S L I C V T A A F G V L L S
- * V F A * V L * S V * L L L L V Y S Y L
- K Y L L R F S N L C N C C F W C T L I *
6961 - AATTTTGGTGCTCCTTCTTATTGTAATGGCGTTAGAGAATTGTATCTTAATTCGTCTAAC -
7020
- N F G A P S Y C N G V R E L Y L N S S N
- I L V L L L I V M A L E N C I L I R L T
- F W C S F L L * W R * R I V S * F V * R
7021 - GTACTACTATGGATTTCTGTGAAGGTTCTTTTCTTGCAGCATTGTTTAAAGTGGATTA -
7080
- V T T M D F C E G S F P C S I C L S G L
- L L L W I S V K V L F L A A F V * V D *
- Y Y Y G F L * R F F S L Q H L F K W I R
7081 - GACTCCCTTGATTCTTATCCAGCTCTTGAACCATTGAGGTGACGATTTTCATCGTACAAG -
7140
- D S L D S Y P A L E T I Q V T I S S Y K
- T P L I L I Q L L K P F R * R F H R T S
- L P * F L S S S * N H S G D D F I V Q A
7141 - CTAGACTTGACAATTTTAGTCTGGCCGCTGAGTGGGTTTGGCATATATGTTGTTCCACA -
7200
- L D L T I L G L A A E W V L A Y M L F T
- * T * Q F * V W P L S G F W H I C C S Q
- R L D N F R S G R * V G F G I Y V V H K
7201 - AAATTCCTTTATTTATTAGTCTTTTTCAGCTATAATGCAGGTGTTCTTTGGCTATTTTGTCT -
7260
- K F F Y L L G L S A I M Q V F F G Y F A
- N S F I Y * V F Q L * C R C S L A I L L
- I L L F I R S F S Y N A G V L W L F C *
7261 - AGTCATTTTCATCAGCAATTTCTGGCTCATGTGGTTTATCATTAGTATTGTACAAATGGCA -
7320
- S H F I S N S W L M W F I I S I V Q M A
- V I S S A I L G S C G L S L V L Y K W H
- S F H Q Q F L A H V V Y H * Y C T N G T
7321 - CCCGTTTCTGCAATGGTTAGGATGTACATCTTCTTTGCTTCTTTCTACTACATATGGAAG -
7380
- P V S A M V R M Y I F F A S F Y Y I W K
- P F L Q W L G C T S S L L L S T Y G R
- R F C N G * D V H L L C F F L L H M E E
7381 - AGCTATGTTTCATATCATGGATGGTTGCACCTCTTCGACTTGCATGATGTGCTATAAGCGC -
7440
- S Y V H I M D G C T S S T C M M C Y K R
- A M F I S W M V A P L R L A * C A I S A
- L C S Y H G W L H L F D L H D V L * A Q
7441 - AATCGTGCCACACGCGTTGAGTGTACAACCTATTGTTAATGGCATGAAGAGATCTTTCTAT -
7500
- N R A T R V E C T T I V N G M K R S F Y
- I V P H A L S V Q L L L M A * R D L S M
- S C H T R * Y Y N Y C * W H E E I F L C
7501 - GTCTATGCAAATGGAGGCCGTTGCTTCTGCAAGACTCACAAATGGAATTGTCTCAATTGT -
7560
- V Y A N G G R G F C K T H N W N C L N C
- S M Q M E A V A S A R L T I G I V S I V
- L C K W R P W L L Q D S Q L E L S Q L *
7561 - GACACATTTTGCAGTGGTAGTACATTCATTAGTGATGAAGTTGCTCGTGATTTGTCACTC -
7620
- D T F C T G S T F I S D E V A R D L S L

```

图 16-J

---

- T H F A L V V H S L V M K L L V I C H S  
- H I L H W \* Y I H \* \* \* S C S \* F V T P  
7621 - CAGTTTAAAAGACCAATCAACCCTACTGACCAGTCATCGTATATTGTTGATAGTGTGCT -  
7680  
- Q F K R P I N P T D Q S S Y I V D S V A  
- S L K D Q S T L L T S H R I L L I V L L  
- V \* K T N Q P Y \* P V I V Y C \* \* C C C

```

7681 - GTGAAAAATGGCGCGCTTCACCTCTACTTTGACAAGGCTGGTCAAAAGACCTATGAGAGA -
7740
- V K N G A L H L Y F D K A G Q K T Y E R
- * K M A R F T S T L T R L V K R P M R D
- E K W R A S P L L * Q G W S K D L * E T
7741 - CATCCGCTCTCCCATTTTGTCAATTTAGACAATTTGAGAGCTAACAACTAAAGGTTCA -
7800
- H P L S H F V N L D N L R A N N T K G S
- I R S P I L S I * T I * E L T T L K V H
- S A L P F C Q F R Q F E S * Q H * R F T
7801 - CTGCCTATTAATGTCATAGTTTTTGATGGCAAGTCCAAATGCGACGAGTCTGCTTCTAAG -
7860
- L P I N V I V F D G K S K C D E S A S K
- C L L M S * F L M A S P N A T S L L L S
- A Y * C H S F * W Q V Q M R R V C F * V
7861 - TCTGCTTCTGTGTACTACAGTCAGCTGATGTGCCAACCTATTCTGTTGCTTGACCAAGCT -
7920
- S A S V Y Y S Q L M C Q P I L L L D Q A
- L L L C T T V S * C A N L F C C L T K L
- C F C V L Q S A D V P T Y S V A * P S S
7921 - CTTGTATCAAACGTTGGAGATAGTACTGAAGTTCCGTTAAGATGTTTGATGCTTATGTC -
7980
- L V S N V G D S T E V S V K M F D A Y V
- L Y Q T L E I V L K F P L R C L M L M S
- C I K R W R * Y * S F R * D V * C L C R
7981 - GACACCTTTTCAGCAACTTTTAGTTCCTATGAAAACTTAAGGCACCTGTTGCTACA -
8040
- D T F S A T F S V P M E K L K A L V A T
- T P F Q Q L L V F L W K N L R H L L L Q
- H L F S N F * C S Y G K T * G T C C Y S
8041 - GCTCACAGCGAGTTAGCAAAGGTTAGCTTTAGATGGTGTCTTCTACATTCGTTCA -
8100
- A H S E L A K G V A L D G V L S T F V S
- L T A S * Q R V * L * M V S F L H S C Q
- S Q R V S K G C S F R W C P F Y I R V S
8101 - GCTGCCGACAAGGTGTGTTGATACCGATGTTGACACAAAGGATGTTATGAATGTCTC -
8160
- A A R Q G V V D T D V D T K D V I E C L
- L P D K V L L I P M L T Q R M L L N V S
- C P T R C C * Y R C * H K G C Y * M S Q
8161 - AAATTTACATCACTCTGACTTAGAAGTGACAGGTGACAGTTGTAACAATTTATGCTC -
8220
- K L S H H S D L E V T G D S C N N F M L
- N F H I T L T * K * Q V T V V T I S C S
- T F T S L * L R S D R * Q L * Q F H A H
8221 - ACCTATAATAAGGTTGAAAACATGACGCCAGAGATCTTGGCGCATGTATTGACTGTAAT -
8280
- T Y N K V E N M T P R D L G A C I D C N
- P I I R L K T * R P E I L A H V L T V M
- L * * G * K H D A Q R S W R M Y * L * C
8281 - GCAAGGCATATCAATGCCCAAGTAGCAAAAAGTCACAATGTTTCACTCATCTGGAATGTA -
8340
- A R H I N A Q V A K S H N V S L I W N V
- Q G I S M P K * Q K V T M F H S S G M *
- K A Y Q C P S S K K S Q C F T H L E C K
8341 - AAAGACTACATGTCTTTATCTGAACAGCTGCCTAAACAAATTCGTACTGCTGCCAAGAAG -
8400
- K D Y M S L S E Q L R K Q I R T A A K K

```

图 16-K

---

- K T T C L Y L N S C V N K F V L L P R R  
- R L H V F I \* T A A \* T N S Y C C Q E E  
8401 - AACACATACCTTTTACACTTGTGCTACAACCTAGACAGGTTGTCAATGTCATAACT -  
8460  
- N N I P F T L T C A T T R Q V V N V I T  
- T T Y L L H \* L V L Q L D R L S M S \* L  
- Q H T F Y T N L C Y N \* T G C Q C H N Y

```

8461 - ACTAAAATCTCACTCAAGGGTGGTAAGATTGTTAGTACTTGTMTTAACTTATGCTTAAG -
8520 - T K I S L K G G K I V S T C F K L M L K
- L K S H S R V V R L L V L V L N L C L R
- * N L T Q G W * D C * Y L F * T Y A * G
8521 - GCCACATTATTGTGCGTTCTTGTGTCATTGGTTTGTATATCGTTATGCCAGTACATACA -
8580 - A T L L C V L A A L V C Y I V M P V H T
- P H Y C A F L L H W F V I S L C Q Y I H
- H I I V R S C C I G L L Y R Y A S T Y I
8581 - TTGTCAATCCATGATGGTTACACAAATGAAATCATTGGTTACAAAGCCATTCCAGGATGGT -
8640 - L S I H D G Y T N E I I G Y K A I Q D G
- C Q S M M V T Q M K S L V T K P F R M V
- V N P * W L H K * N H W L Q S H S G W C
8641 - GTCACTCGTGACATCATTCTACTGATGATTGTTTGGCAAATAAACATGCTGGTTTGGAC -
8700 - V T R D I I S T D D C F A N K H A G F D
- S L V T S F L L M I V L Q I N M L V L T
- H S * H H F Y * * L F C K * T C W F * R
8701 - GCATGTTTAGCCAGCGTGGTGGTTCATACAAAATGACAAAAGCTGCCCTGTAGTAGCT -
8760 - A W F S Q R G G S Y K N D K S C P V V A
- H G L A S V V V H T K M T K A A L * * L
- M V * P A W W F I Q K * Q K L P C S S C
8761 - GCTATCATTACAGAGAGATTGGTTTCATAGTGCTGGCTTACCGGGTACTGTGCTGAGA -
8820 - A I I T R E I G F I V P G L P G T V L R
- L S L Q E R L V S * C L A Y R V L C * E
- Y H Y K R D W F H S A W L T G Y C A E S
8821 - GCAATCAATGGTGACTTCTTGCATTTTCTACCTCGTGTMTTGTAGTGTGGCAACATT -
8880 - A I N G D F L H F L P R V F S A V G N I
- Q S M V T S C I F Y L V F L V L L A T F
- N Q W * L L A F S T S C F * C C W Q H L
8881 - TGCTACACACCTTCCAAACTCATTGAGTATAGTGATTTTGCTACCTCTGCTTGCCTTCTT -
8940 - C Y T P S K L I E Y S D F A T S A C V L
- A T H L P N S L S I V I L L P L L A F L
- L H T F Q T H * V * * F C Y L C L R S C
8941 - GCTGCTGAGTGTACAATTTTAAGGATGCTATGGGCAAACCTGTGCCATATGTTATGAC -
9000 - A A E C T I F K D A M G K P V P Y C Y D
- L L S V Q F L R M L W A N L C H I V M T
- C * V Y N F * G C Y G Q T C A I L L * H
9001 - ACTAATTTGCTAGAGGGTCTATTTCTTATAGTGAGCTTCGTCCAGACACTCGTTATGTG -
9060 - T N L L E G S I S Y S E L R P D T R Y V
- L I C * R V L F L I V S F V Q T L V M C
- * F A R G F Y F L * * A S S R H S L C A
9061 - CTTATGGATGGTTCATCATACAGTTTCTAACACTTACCTGGAGGGTCTGTAGAGTA -
9120 - L M D G S I I Q F P N T Y L E G S V R V
- L W M V P S Y S F L T L T W R V L L E *
- Y G W F H H T V S * H L P G G F C * S S
9121 - GTAACAACCTTTTGTAGTGTGACTGTAGACATGGTACATGCGAAAGGTCAGAAGTAGGT -
9180 - V T T F D A E Y C R H G T C E R S E V G
- * Q L L M L S T V D M V H A K G Q K * V

```

图 16-L

---

- N N F \* C \* V L \* T W Y M R K V R S R Y  
9181 - A T T G C C T A T C T A C C A G T G G T A G A T G G G T T C T T A A T A A T G A G C A T T A C A G A G C T C T A T C A -  
9240  
- I C L S T S G R W V L N N E H Y R A L S  
- F A Y L P V V D G F L I M S I T E L Y Q  
- L P I Y Q W \* M G S \* \* \* A L Q S S I R

```

9241 - GGAGTTTCTGTGGTGTGATGCGATGAATCTCATAGCTAACATCTTTACTCCTCTTGTG -
9300
- G V F C G V D A M N L I A N I F T P L V
- E F S V V L M R * I S * L T S L L L L C
- S F L W C * C D E S H S * H L Y S S C A
9301 - CAACCTGTGGGTGCTTTAGATGTGTCTGCTTCAGTAGTGGCTGGTATTATTGCCATA -
9360
- Q P V G A L D V S A S V V A G G I I A I
- N L W V L * M C L L Q * W L V V L L P Y
- T C G C F R C V C F S S G W W Y Y C H I
9361 - TTGGTGACTTGTGCTGCCTACTACTTTATGAAATTCAGACGTGTTTTGGTGAGTACAAC -
9420
- L V T C A A Y Y F M K F R R V F G E Y N
- W * L V L P T T L * N S D V F L V S T T
- G D L C C L L Y E I Q T C F W * V Q P
9421 - CATGTTGTTGCTGCTAATGCACCTTTGTTTTGATGTCTTCACTATACTCTGTCTGGTA -
9480
- H V V A A N A L L F L M S F T I L C L V
- M L L L L M H F C F * C L S L Y S V W Y
- C C C C * C T F V F D V F H Y T L S G T
9481 - CCAGCTTACAGCTTCTGCCGGAGTCTACTCAGTCTTTTACTTGTACTTGACATTCTAT -
9540
- P A Y S F L P G V Y S V F Y L Y L T F Y
- Q L T A F C R E S T Q S F T C T * H S I
- S L Q L S A G S L L S L L L V L D I L F
9541 - TTCACCAATGATGTTTTATTCTTGGCTCACCTTCAATGGTTGCCATGTTTTCTCCTATT -
9600
- F T N D V S F L A H L Q W F A M F S P I
- S P M M F H S W L T F N G L P C F L L L
- H Q * C F I L G S P S M V C H V F S Y C
9601 - GTGCCCTTTTGGATAACAGCAATCTATGATTCTGTATTCTCTGAAGCACTGCCATTGG -
9660
- V P F W I T A I Y V F C I S L K H C H W
- C L F G * Q Q S M Y S V F L * S T A I G
- A F L D N S N L C I L Y F S E A L P L V
9661 - TTCTTTAACAACCTATCTTAGGAAAAGAGTCATGTTAATGGAGTTACATTTAGTACCTTC -
9720
- F F N N Y L R K R V M F N G V T F S T F
- S L T T I L G K E S C L M E L H L V P S
- L * Q L S * E K S H V * W S Y I * Y L R
9721 - GAGGAGGCTGCTTTGTGTACCTTTTGTCTCAACAAGCAAAATGTACCTAAAATGCGTAGC -
9780
- E E A A L C T F L L N K E M Y L K L R S
- R R L L C V P F C S T R K C T * N C V A
- G G C F V Y L F A Q Q G N V P K I A * R
9781 - GAGACACTGTTGCCACTTACACAGTATAACAGGTATCTTGTCTATATAACAAGTACAAG -
9840
- E T L L P L T Q Y N R Y L A L Y N K Y K
- R H C C H L H S I T G I L L Y I T S T S
- D T V A T Y T V * Q V S C S I * Q V Q V
9841 - TATTTAGTGGAGCCTTAGATACTACCAGCTATCGTGAAGCAGCTTGTGCCACTTAGCA -
9900
- Y F S G A L D T T S Y R E A A C C H L A
- I S V E P * I L P A I V K Q L A A T * Q
- F Q W S L R Y Y Q L S * S S L L P L S K
9901 - AAGGCTCTAAATGACTTTAGCAACTCAGGTGCTGATGTTCTCTACCAACCACCACAGACA -

```

图 16-M



9960  
 - K A L N D F S N S G A D V L Y Q P P Q T  
 - R L \* M T L A T Q V L M F S T N H H R H  
 - G S K \* L \* Q L R C \* C S L P T T T D I

图 16-N

9961 - TCAATCACTTCTGCTGTTCTGCAGAGTGGTTTTAGGAAAATGGCATTCCCCTCAGGCAAA -  
 10020  
 - S I T S A V L Q S G F R K M A F P S G K  
 - Q S L L L F C R V V L G K W H S R Q A K  
 - N H F C C S A E W F \* E N G I P V R Q S  
 10021 - GTTGAAGGGTGCATGGTACAAGTAACCTGTGGAACACAACTCTTAATGGATTGTGGTTG -  
 10080  
 - V E G C M V Q V T C G T T T L N G L W L  
 - L K G A W Y K \* P V E L Q L L M D C G W  
 - \* R V H G T S N L W N Y N S \* W I V V G  
 10081 - GATGACACAGTATACTGTCCAAGACATGTCATTTGCACAGCAGAAGACATGCTTAATCCT -  
 10140  
 - D D T V Y C P R H V I C T A E D M L N P  
 - M T Q Y T V Q D M S F A Q Q K T C L I L  
 - \* H S I L S K T C H L H S R R H A \* S \*  
 10141 - AACTATGAAGATCTGCTCATTGCGAAATCCAACCATAGCTTTCTTGTTCAGGCTGGCAAT -  
 10200  
 - N Y E D L L I R K S N H S F L V Q A G N  
 - T M K I C S F A N P T I A F L F R L A M  
 - L \* R S A H S Q I Q P \* L S C S G W Q C  
 10201 - GTTCAACTTCGTGTTATTGGCCATTCTATGCAAAATGTCTGCTTAGGCTTAAAGTTGAT -  
 10260  
 - V Q L R V I G H S M Q N C L L R L K V D  
 - F N F V L L A I L C K I V C L G L K L I  
 - S T S C Y W P F Y A K L S A \* A \* S \* Y  
 10261 - ACTTCTAACCCCTAAGACACCCCAAGTATAAATTTGTCCGTATCCAACCTGGTCAACATTT -  
 10320  
 - T S N P K T P K Y K F V R I Q P G Q T F  
 - L L T L R H P S I N L S V S N L V K H F  
 - F \* P \* D T Q V \* I C P Y P T W S N I F  
 10321 - TCAGTTCTAGCATGCTACAATGGTTCCACCATCTGGTGTATTATCAGTGTGCCATGAGACCT -  
 10380  
 - S V L A C Y N G S P S G V Y Q C A M R P  
 - Q P \* H A T M V H H L V F I S V P \* D L  
 - S S S M L Q W F T I W C L S V C H E T \*  
 10381 - AATCATACCATTAAAGGTTCTTTCCTTAATGGATCATGTGGTAGTGTGGTTTTAACATT -  
 10440  
 - N H T I K G S F L N G S C G S V G F N I  
 - I I P L K V L S L M D H V V V L V L T L  
 - S Y H \* R F F P \* W I M W \* C W F \* H \*  
 10441 - GATTATGATTGCGTGTCTTCTGCTATATGCATCATATGGAGCTTCCAACAGGAGTACAC -  
 10500  
 - D Y D C V S F C Y M H H M E L P T G V H  
 - I M I A C L S A I C I I W S F Q Q E Y T  
 - L \* L R V F L L Y A S Y G A S N R S T R

```

10501 - GCTGGTACTGACTTAGAAGGTAATCTATGGTCCATTTGTTGACAGACAACTGCACAG -
10560
- A G T D L E G K F Y G P F V D R Q T A Q
- L V L T * K V N S M V H L L T D K L H R
- W Y * L R R * I L W S I C * Q T N C T G
10561 - GCTGCAGGTACAGACACAACCATAACATTAATGTTTGGCATGGCTGTATGCTGCTGT -
10620
- A A G T D T T I T L N V L A W L Y A A V
- L Q V Q T Q P * H * M F W H G C M L L L
- C R Y R H N H N I K C F G M A V C C C Y
10621 - ATCAATGGTGATAGGTGGTTCTTAATAGATTCACCACTACTTTGAATGACTTTAACCTT -
10680
- I N G D R W F L N R F T T T L N D F N L
- S M V I G G F L I D S P L L * M T L T L
- Q W * * V V S * * I H H Y F E * L * P C

```

图 16-0

```

10681 - GTGGCAATGAAGTACAACCTATGAACCTTTGACACAAGATCATGTTGACATATTGGGACCT -
10740
- V A M K Y N Y E P L T Q D H V D I L G P
- W Q * S T T M N L * H K I M L T Y W D L
- G N E V Q L * T F D T R S C * H I G T S
10741 - CTTTCTGCTCAAACAGGAATGGCGTCTTAGATATGTGTGCTGCTTTGAAAGAGCTGCTG -
10800
- L S A Q T G I A V L D M C A A L K E L L
- F L L K Q E L P S * I C V L L * K S C C
- F C S N R N C R L R Y V C C F E R A A A
10801 - CAGAATGGTATGAATGGTCTACTATCCTTGGTAGCACTATTTAGAGATGAGTTTACA -
10860
- Q N G M N G R T I L G S T I L E D E F T
- R M V * M V V L S L V A L F * K M S L H
- E W Y E W S Y Y P W * H Y F R R * V Y T
10861 - CCATTTGATGTTGTTAGACAATGCTCTGGTGTACCTTCCAAGGTAAGTTCAAGAAAATF -
10920
- P F D V V R Q C S G V T F Q G K F K K I
- H L M L L D N A L V L P S K V S S R K L
- I * C C * T M L W C Y L P R * V Q E N C
10921 - GTTAAGGGCACTCATCATTTGGATGCTTTTAACTTTCTTGACATCACTATTGATTCTTGT -
10980
- V K G T H H W M L L T F L T S L L I L V
- L R A L I I G C F * L S * H H Y * F L F
- * G H S S L D A F N F L D I T I D S C S
10981 - CAAAGTACACAGTGGTCACTGTTTTCTTTGTTTACGAGAATGCTTTCTTGCATTTACT -
11040
- Q S T Q W S L F F F V Y E N A F L P F T
- K V H S G H C F S L F T R M L S C H L L
- K Y T V V T V F L C L R E C F L A I Y S
11041 - CTTGGTATTATGGCAATTGCTGCATGCTATGCTGCTTGTAAAGCATAAGCACGCATTC -

```

```

11100 - L G I M A I A A C A M L L V K H K H A F
      - L V L W Q L L H V L C C L L S I S T H S
      - W Y Y G N C C M C Y A A C * A * A R I L
11101 - TTGTGCTTGTTCCTGTTACCTTCTCTTGCAACAGTTGCTTACITTAATATGGTCTACATG -
11160 - L C L F L L P S L A T V A Y F N M V Y M
      - C A C F C Y L L L Q Q L L T L I W S T C
      - V L V S V T F S C N S C L L * Y G L H A
11161 - CCTGCTAGCTGGGTGATGCGTATCATGACATGGCTTGAATTGGCTGACACTAGCTTGTCT -
11220 - P A S W V M R I M T W L E L A D T S L S
      - L L A G * C V S * H G L N W L T L A C L
      - C * L G D A Y H D M A * I G * H * L V W
11221 - GGTATAGGCTTAAGGATTGTGTTATGTATGCTTCAGCTTTAGTTTTGCTTATTCTCATG -
11280 - G Y R L K D C V M Y A S A L V L L I L M
      - V I G L R I V L C M L Q L * F C L F S *
      - L * A * G L C Y V C F S F S F A Y S H D
11281 - ACAGCTCGCACTGTTTATGATGATGCTGCTAGACGTGTTTGGACACTGATGAATGTCATT -
11340 - T A R T V Y D D A A R R V W T L M N V I
      - Q L A L F M M M L L D V F G H * * M S L
      - S S H C L * * C C * T C L D T D E C H Y
11341 - ACACTTGTTTACAAAGTCTACTATGGTAATGCTTTAGATCAAGCTATTTCCATGTGGGCC -
11400 - T L V Y K V Y Y G N A L D Q A I S M W A
      - H L F T K S T M V M L * I K L F P C G P
      - T C L Q S L L W * C F R S S Y F H V G L

```

图 16-P

```

11401 - TTAGTTATTTCTGTRACCTCTAACTATTCCTGGTGTTCGTTACGACTATCATGTTTPTAGCT -
11460 - L V I S V T S N Y S G V V T T I M F L A
      - * L F L * P L T I L V S L R L S C F * L
      - S Y F C N L * L F W C R Y D Y H V F S *
11461 - AGAGCTATAGTGTTCGTTGTTGAGTATTACCCATTGTTATTATTACTGGCAACACC -
11520 - R A I V F V C V E Y Y P L L F I T G N T
      - E L * C L C V L S I T H C Y L L L A T P
      - S Y S V C V C * V L P I V I Y Y W Q H L
11521 - TTACAGTGTATCATGCTTGTATTATTGTTTCTTAGGCTATTGTTGCTGCTACTTTGGC -
11580 - L Q C I M L V Y C F L G Y C C C C Y F G
      - Y S V S C L F I V S * A I V A A A T L A
      - T V Y H A C L L F L R L L L L L L L W P
11581 - CTTTTCTGTTTACTCAACCGTTACTTCAGGCTTACTCTTGGTGTATGACTACTTGGTC -
11640 - L F C L L N R Y F R L T L G V Y D Y L V
      - F S V Y S T V T S G L L L V F M T T W S
      - F L F T Q P L L Q A Y S W C L * L L G L

```

```

11641 - TCTACACAAGAATTTAGGTATATGAACTCCAGGGGCTTTTGCCTCCTAAGAGTAGTATT -
11700
- S T Q E F R Y M N S Q G L L P P K S S I
- L H K N L G I * T P R G F C L L R V V L
- Y T R I * V Y E L P G A F A S * E * Y *
11701 - GATGCTTTCAGCTTAACATTAAGTTGTTGGGTATTGGAGGTAACCATGTATCAAGGTT -
11760
- D A F K L N I K L L G I G G K P C I K V
- M L S S L T L S C W V L E V N H V S R L
- C F Q A * H * V V G Y W R * T M Y Q G C
11761 - GCTACTGTACAGTCTAAAATGTCTGACGTAAGTGCACATCTGTGGTACTGCTCTCGGTT -
11820
- A T V Q S K M S D V K C T S V V L L S V
- L L Y S L K C L T * S A H L W Y C S R F
- Y C T V * N V * R K V H I C G T A L G S
11821 - CTTCAACAACCTAGAGTAGAGTCATCTTCTAAATGTGGGCACAATGTGTACAACCTCCAC -
11880
- L Q Q L R V E S S S K L W A Q C V Q L H
- F N N L E * S H L L N C G H N V Y N S T
- S T T * S R V I F * I V G T M C T T P Q
11881 - AATGATATTCTTCTTGCAAAAGACACAACCTGAAGCTTTTCGAGAAGATGGTTTCTCTTTTG -
11940
- N D I L L A K D T T E A F E K M V S L L
- M I F F L Q K T Q L K L S R R W F L F C
- * Y S S C K R H N * S F R E D G F S F V
11941 - TCTGTTTGGCTATCCATGCAGGGTCTGTAGACATTAATAGGTTGTGCGAGGAAATGCTC -
12000
- S V L L S M Q G A V D I N R L C E E M L
- L F C Y P C R V L * T L I G C A R K C S
- C F A I H A G C C R H * * V V R G N A R
12001 - GATAACCGTGCTACTCTTCAGGCTATTGCTTCAGAATTTAGTTCTTTACCATCATATGCC -
12060
- D N R A T L Q A I A S E F S S L P S Y A
- I T V L L F R L L L Q N L V L Y H H M P
- * P C Y S S G Y C F R I * F F T I I C R
12061 - GCTTATGCCACTGCCAGGAGGCTATGAGCAGGCTGTAGCTAATGGTGATTCTGAAGTC -
12120
- A Y A T A Q E A Y E Q A V A N G D S E V
- L M P L P R R P M S R L * L M V I L K S
- L C H C P G G L * A G C S * W * F * S R

```

 16-Q

```

12121 - GTTCTCAAAAGTTAAAGAAATCTTTGAATGTGGCTAAATCTGAGTTTGACCGTGATGCT -
12180
- V L K K L K K S L N V A K S E F D R D A
- F S K S * R N L * M W L N L S L T V M L
- S Q K V K E I F E C G * I * V * P * C C
12181 - GCCATGCAACGCAAGTTGGAARAGATGCCAGATCAGGCTATGACCCAAATGTACAAACAG -
12240
- A M Q R K L E K M A D Q A M T Q M Y K Q
- P C N A S W K R W Q I R L * P K C T N R

```

```

- H A T Q V G K D G R S G Y D P N V Q T G
12241 - GCAAGATCTGAGGACAAGAGGGCAAAAAGTAACTAGTCTATGCAAACAATGCTCTTCACT -
12300
- A R S E D K R A K V T S A M Q T M L F T
- Q D L R T R G Q K * L V L C K Q C S S L
- K I * G Q E G K S N * C Y A N N A L H Y
12301 - ATGCTTAGGAAGCTTGATAATGATGCACTTAACAACATTATCAACAATGCGCGTGATGGT -
12360
- M L R K L D N D A L N N I I N N A R D G
- C L G S L I M M H L T T L S T M R V M V
- A * E A * * * C T * Q H Y Q Q C A * W L
12361 - TGTGTCCACTCAACATCATACCATTGACTACAGCAGCCAAACTCATGGTTGTTGTCCCT -
12420
- C V P L N I I P L T T A A K L M V V V P
- V F H S T S Y H * L Q Q P N S W L L S L
- C S T Q H H T I D Y S S Q T H G C C P *
12421 - GATTATGGTACCTACAAGAACACTTGTGATGGTAACACCTTTACATATGCATCTGCACTC -
12480
- D Y G T Y K N T C D G N T F T Y A S A L
- I M V P T R T L V M V T P L H M H L H S
- L W Y L Q E H L * W * H L Y I C I C T L
12481 - TGGGAATCCAGCAAAGTTGTTGATGCGGATAGCAAGATTGTTCAACTTAGTGAATTAAC -
12540
- W E I Q Q V V D A D S K I V Q L S E I N
- G K S S K L L M R I A R L F N L V K L T
- G N P A S C * C G * Q D C S T * * N * H
12541 - ATGGACAATTCACCAAATTTGGCTTGGCCTCTATTGTTACAGCTCTAAGAGCCAACCTCA -
12600
- M D N S P N L A W P L I V T A L R A N S
- W T I H Q I W L G L L L L Q L * E P T Q
- G Q F T K F G L A S Y C Y S S K S Q L S
12601 - GCTGTAAACTACAGAATAATGAACTGAGTCCAGTAGCACTACGACAGATGTCCTGTGCG -
12660
- A V K L Q N N E L S P V A L R Q M S C A
- L L N Y R I M N * V Q * H Y D R C P V R
- C * T T E * * T E S S S T T T D V L C G
12661 - GCTGGTACCACACAAACAGCTTGTACTGATGACAATGCACTTGCCTACTATAACAATTCG -
12720
- A G T T Q T A C T D D N A L A Y Y N N S
- L V P H K Q L V L M T M H L P T I T I R
- W Y H T N S L Y * * Q C T C L L * Q F E
12721 - AAGGGAGGTAGGTTTGTGCTGGCATTACTATCAGACCACCAAGATCTCAAATGGGCTAGA -
12780
- K G G R F V L A L L S D H Q D L K W A R
- R E V G L C W H Y Y Q T T K I S N G L D
- G R * V C A G I T I R P P R S Q M G * I
12781 - TTCCTAAGAGTGATGGTACAGGTACAATTTACAGAACTGGAACCACCTTGTAGGTTT -
12840
- F P K S D G T G T I Y T E L E P P C R F
- S L R V M V Q V Q F T Q N W N H L V G L
- P * E * W Y R Y N L H R T G T T L * V C
12841 - GTTACAGACACCAAAGGGCCTAAAGTGAATACTTGTACTTCATCAAAGGCTTAAAC -
12900
- V T D T P K G P K V K Y L Y F I K G L N
- L Q T H Q K G L K * N T C T S S K A * T
- Y R H T K R A * S E I L V L H Q R L K Q

```

## 图 16-R

```

12901 - AACCTAAATAGAGGTATGGTGCTGGGCAGTTTAGCTGCTACAGTACGTCTTCAGGCTGGA -
12960
- N L N R G M V L G S L A A T V R L Q A G
- T * I E V W C W A V * L L Q Y V F R L E
- P K * R Y G A G Q F S C Y S T S S G W K
12961 - AATGCTACAGAAGTACCTGCCAATCAACTGTGCTTTCCTTCTGTGCTTTTGCCAGTAGAC -
13020
- N A T E V P A N S T V L S F C A F A V D
- M L Q K Y L P I Q L C F P S V L L Q * T
- C Y R S T C Q F N C A F L L C F C S R P
13021 - CCTGCTAAAGCATATAAGGATTACCTAGCAAGTGGAGGACACCAATCACCAACTGTGTG -
13080
- P A K A Y K D Y L A S G G Q P I T N C V
- L L K H I R I T * Q V E D N Q S P T V *
- C * S I * G L P S K W R T T N H Q L C E
13081 - AAGATGTTGTGTACACACTGGTACAGGACAGGCAATTACTGTAACACCAGAAGCTAAC -
13140
- K M L C T H T G T G Q A I T V T P E A N
- R C C V H T L V Q D R Q L L * H Q K L T
- D V V Y T H W Y R T G N Y C N T R S * H
13141 - ATGGACCAAGAGTCCCTTGGTGGTGCTTCATGTTGTCTGTATTGTAGATGCCACATTGAC -
13200
- M D Q E S F G G A S C C L Y C R C H I D
- W T K S P L V V L H V V C I V D A T L T
- G P R V L W W C F M L S V L * M P H * P
13201 - CATCCAATCCTAAAGGATTCTGTGACTTGAAGGTAAGTACGTCCAATACCTACCCT -
13260
- H P N P K G F C D L K G K Y V Q I P T T
- I Q I L K D S V T * K V S T S K Y L P L
- S K S * R I L * L E R * V R P N T Y H L
13261 - TGTGCTAATGACCCAGTGGGTTTTACACTTAGAAACACAGTCTGTACCCTCTGCCGAATG -
13320
- C A N D P V G F T L R N T V C T V C G M
- V L M T Q W V L H L E T Q S V P S A E C
- C * * P S G F Y T * K H S L Y R L R N V
13321 - TGGAAAGGTTATGGCTGTAGTTGTGACCAACTCCGCGAACCTTGATGCAGTCTGCGGAT -
13380
- W K G Y G C S C D Q L R E P L M Q S A D
- G K V M A V V V T N S A N P * C S L R M
- E R L W L * L * P T P R T L D A V C G C
13381 - GCATCAACGTTTTTAAACGGGTTTGCGGTGAAGTGCAGCCCGTCTTACACCGTGCGGCA -
13440
- A S T F L N G F A V * V Q P V L H R A A
- H Q R F * T G L R C K C S P S Y T V R H
- I N V F K R V C G V S A A R L T P C G T
13441 - CAGGCACTAGTACTGATGTCGTCTACAGGGCTTTTGATATTTACAACGAAAAAGTGCTG -
13500
- Q A L V L M S S T G L L I F T T K K V L
- R H * Y * C R L Q G F * Y L Q R K K C W
- G T S T D V V Y R A F D I Y N E K S A G
13501 - GTTTTGCAAAGTTCCTAAAACTAATTGCTGTGCGTTCAGGAGAAGGATGAGGAAGGCA -
13560
- V L Q S S * K L I A V A S R R R R M R K A
- F C K V P K N * L L S L P G E G * G R Q
- F A K F L K T N C C R F Q E K D E E G N
13561 - ATTTATTAGACTCTTACTTTGTAGTTAAGAGGCATACTATGTCCTAACTACCAACATGAAG -
13620

```

---

- I Y \* T L T L \* L R G I L C L T T N M K  
- F I R L L L C S \* E A Y Y V \* L P T \* R  
- L L D S Y F V V K R H T M S N Y Q H E E  
13621 - AGACTATTTATAACTTGGTTAAAGATTGTCAGCGGTTGCTGTCCATGACTTTTTCAAGT -  
13680  
- R L F I T W L K I V Q R L L S M T F S S  
- D Y L \* L G \* R L S S G C C P \* L F Q V  
- T I Y N L V K D C P A V A V H D F F K F

图 16-S

```

13681 - TTAGAGTAGATGGTGACATGGTACCACATATATCACGTCAGCGTCTAACTAAATACACAA -
13740
- L E * M V T W Y H I Y H V S V * L N T Q
- * S R W * H G T T Y I T S A S N * I H N
- R V D G D M V P H I S R Q R L T K Y T M
13741 - TGGCTGATTTAGTCTATGCTCTACGTCATTTTGATGAGGGTAATTGTGATACATTAAG -
13800
- W L I * S M L Y V I L M R V I V I H * K
- G * F S L C S T S F * * G * L * Y I K R
- A D L V Y A L R H F D E G N C D T L K E
13801 - AAATACTCGTCACATACAATTGCTGTGATGATGATTATTTCAATAAGAAGGATTGGTATG -
13860
- K Y S S H T I A V M M I I S I R R I G M
- N T R H I Q L L * * * L F Q * E G L V *
- I L V T Y N C C D D D Y F N K K D W Y D
13861 - ACTTCGTAGAGAATCCTGACATCTTACGCGTATATGCTAACTTAGGTGAGCGGTACGCC -
13920
- T S * R I L T S Y A Y M L T * V S V Y A
- L R R E S * H L T R I C * L R * A C T P
- F V E N P D I L R V Y A N L G E R V R Q
13921 - AATCATTATTAAGACTGTACAATTCTGCGATGCTATGCGTGATGCAGGCATTGTAGGCC -
13980
- N H Y * R L Y N S A M L C V M Q A L * A
- I I I K D C T I L R C Y A * C R H C R R
- S L L K T V Q F C D A M R D A G I V G V
13981 - TACTCACATTAGATAATCAGGATCTTAATGGGAAGTACGATTTCGGTGATTTCTGAC -
14040
- Y * H * I I R I L M G T G T I S V I S Y
- T D I R * S G S * W E L V R F R * F R T
- L T L D N Q D L N G N W Y D F G D F V Q
14041 - AAGTAGCACCAGGCTCGGAGTTCCTATTGTGGATTATATTACTCATTGCTGATGCCCA -
14100
- K * H Q A A E F L L W I H I T H C * C P
- S S T R L R S S Y C G F I L L I A D A H
- V A P G C G V P I V D S Y Y S L L M P I
14101 - TCCTCACTTTGACTAGGGCATTGGCTGCTGAGTCCCATATGGATGCTGATCTGCAAAC -
14160
- S S L * L G H W L L S P I W M L I S Q N
- P H F D * G I G C * V P Y G C * S R K T
- L T L T R A L A E S H M D A D L A K P
14161 - CACTTATTAAGTGGGATTGCTGAAATATGATTTTACGGAAGAGAGACTTTGTCTCTCG -
14220
- H L L S G I C * N M I L R K R D F V S S
- T Y * V G F A E I * F Y G R E T L S L R
- L I K W D L L K Y D F T E E R L C L F D
14221 - ACCGTTATTTAAATATGGGACCAGACATACCATCCCAATTGTATTAAGTGTGGATG -
14280
- T V I L N I G T R H T I P I V L T V W M
- P L F * I L G P D I P S Q L Y * L F G *
- R Y F K Y W D Q T Y H P N C I N C L D D
14281 - ATAGGTGTATCCTTCATTGTGCAAACCTTAATGTGTTATTTCTACTGTGTTCCACCTA -
14340
- I G V S F I V Q T L M C Y F L L C F H L
- * V Y P S L C K L * C V I F Y C V S T Y
- R C I L H C A N F N V L F S T V F P P T
14341 - CAAGTTTTGGACCACTAGTAAGAAAATATTTGTAGATGGTTCCTTTTGTGTTTCAA -
14400

```



---

- Q V L D H \* \* E K Y L \* M V F L L L F Q  
- K F W T T S K K N I C R W C S F C C F N  
- S F G P L V R K I F V D G V P F V V S T  
14401 - CTGGATACCATTTTCGTGAGTTAGGAGTCGTACATAATCAGGATGTAAACTTACATAGCT -  
14460  
- L D T I F V S \* E S Y I I R M \* T Y I A  
- W I P P S \* V R S R T \* S G C K L T \* L  
- G Y H F R E L G V V H N Q D V N L H S S

## 图 16-T

```

14461 - CGCGTCTCAGTTTCAAGGAACCTTTAGTGTATGCTGCTGATCCAGCTATGCATGCAGCTT -
14520
- R V S V S R N F * C M L L I Q L C M Q L
- A S Q F Q G T F S V C C * S S Y A C S F
- R L S F K E L L V Y A A D P A M H A A S
14521 - CTGGCAATTTATTGCTAGATAAACGCACTACATGCTTTTCAGTAGCTGCACTAACAAACA -
14580
- L A I Y C * I N A L H A F Q * L H * Q T
- W Q F I A R * T H Y M L F S S C T N K Q
- G N L L L D K R T T C F S V A A L T N N
14581 - ATGTTGCTTTTCAAACCTGCAACCCGGTAATTTAATAAAGACTTTTATGACTTTTGCTG -
14640
- M L L F K L S N P V I L I K T F M T L L
- C C F S N C Q T R * F * * R L L * L C C
- V A F Q T V K P G N F N K D F Y D F A V
14641 - TGTCTAAAGGTTTCTTTAAGGAAGGAAGTTCTGTTGACTAAAACACTTCTTCTTTGCTC -
14700
- C L K V S L R K E V L L N * N T S S L L
- V * R F L * G R K F C * T K T L L L C S
- S K G F F K E G S S V E L K H F F F A Q
14701 - AGGATGGCAACGCTGCTATCAGTGATTATGACTATTATCGTTATAATCTGCCAACAAATGT -
14760
- R M A T L L S V I M T I I V I I C Q Q C
- G W Q R C Y Q * L * L L S L * S A N N V
- D G N A A I S D Y D Y Y R Y N L P T M C
14761 - GTGATATCAGACAACTCCTATTGCTAGTTGAAGTTGTTGATAAATACTTTGATTGTTACG -
14820
- V I S D N S Y S * L K L L I N T L I V T
- * Y Q T T P I R S * S C * * I L * L L R
- D I R Q L L F V V E V V D K Y F D C Y D
14821 - ATGGTGGCTGTATTAATGCCAACCAAGTAATCGTTAACAACTGCGATAAATCAGCTGGTT -
14880
- M V A V L M P T K * S L T I W I N Q L V
- W W L Y * C Q P S N R * Q S G * I S W F
- G G C I N A N Q V I V N N L D K S A G F
14881 - TCCCATTAATAAATGGGTAAGGCTAGACTTTATATGACTCAATGAGTTATGAGGATC -
14940
- S H L I N G V R L D F I M T Q * V M R I
- P I * * M G * G * T L L * L N E L * G S
- P F N K W G K A R L Y Y D S M S Y E D Q
14941 - AAGATGCACTTTTCGCGTATACTAAGCGTAATGTCATCCCTACTATAACTCAAATGAATC -
15000
- K M H F S R I L S V M S S L L * L K * I
- R C T F R V Y * A * C H P Y Y N S N E S
- D A L F A Y T K R N V I P T I T Q M N L
15001 - TTAAGTATGCCATTAGTGCAAAGAATAGAGCTCGCACCGTAGCTGGTGTCTCTATCTGTA -
15060
- L S M P L V Q R I E L A P * L V S L S V
- * V C H * C K E * S S H R S W C L Y L *
- K Y A I S A K N R A R T V A G V S I C S
15061 - GTACTATGACAAATAGACAGTTTTCATCAGAAATTATTGAAGTCAATAGCCGCCACTAGAG -
15120
- V L * Q I D S F I R N Y * S Q * P P L E
- Y Y D K * T V S S E I I E V N S R H * R
- T M T N R Q F H Q K L L K S I A A T R G
15121 - GAGCTACTGTGGTAATTGGAACAAGCAAGTTTACGGTGGCTGGCATAATATGTTAAAAA -

```

15180

- E L L W \* L E Q A S F T V A G I I C \* K  
- S Y C G N W N K Q V L R W L A \* Y V K N  
- A T V V I G T S K F Y G G W H N M L K T

## 图 16-U

```

15181 - CTGTTTACAGTGATGTAGAAACTCCACACCTTATGGGTTGGGATTATCCAAATGTGACA -
15240
- L F T V M * K L H T L W V G I I Q N V T
- C L Q * C R N S T P Y G L G L S K M * Q
- V Y S D V E T P H L M G W D Y P K C D R
15241 - GAGCCATGCCTAACATGCTTAGGATAATGGCCTCTCTTGTCTTGCTCGCAAACATAACA -
15300
- E P C L T C L G * W P L L F L L A N I T
- S H A * H A * D N G L S C S C S Q T * H
- A M P N M L R I M A S L V L A R K H N T
15301 - CTTGCTGTAACCTTATCACACCGTTTCTACAGTTAGCTAACGAGTGTGCGCAAGTATTAA -
15360
- L A V T Y H T V S T G * L T S V R K Y *
- L L * L I T P F L Q V S * R V C A S I K
- C C N L S H R F Y R L A N E C A Q V L S
15361 - GTGAGATGGTCATGTGTGGCGGCTCACTATATGTTAAACCAGGTGGAACATCATCCGGTG -
15420
- V R W S C V A A H Y M L N Q V E H H P V
- * D G H V W R L T I C * T R W N I I R *
- E M V M C G G S L Y V K P G G T S S G D
15421 - ATGCTACAACCTGCTTATGCTAATAGTGTCTTTAACATTGTGCAAGCTGTACAGCCAATG -
15480
- M L Q L L M L I V S L T F V K L L Q P M
- C Y N C L C * * C L * H L S S C Y S Q C
- A T T A Y A N S V F N I C Q A V T A N V
15481 - TAAATGCACTTCTTTCAACTGATGGTAATAAGATAGCTGACAAGTATGTCGGCAATCTAC -
15540
- * M H F F Q L M V I R * L T S M S A I Y
- K C T S F N * W * * D S * Q V C P Q S T
- N A L L S T D G N K I A D K Y V R N L Q
15541 - AACACAGGCTCTATGAGTGTCTCTATAGAAATAGGGATGTTGATCATGAARTCGTGGATG -
15600
- N T G S M S V S I E I G M L I M N S W M
- T Q A L * V S L * K * G C * S * I R G *
- H R L Y E C L Y R N R D V D H E F V D E
15601 - AGTTTTACGCTTACCTGCGTAAACATTTCTCCATGATGATTCTTTCTGATGATGCCGTTG -
15660
- S F T L T C V N I S P * * F F L M M P L
- V L R L P A * T F L H D D S F * * C R C
- F Y A Y L R K H F S M M I L S D D A V V
15661 - TGTGCTATAACAGTAACATATGCGGCTCAAGGTTTAGTAGCTAGCATTAGAAGCTTTAAGG -
15720
- C A I T V T M R L K V * * L A L R T L R
- V L * Q * L C G S R F S S * H * E L * G
- C Y N S N Y A A Q G L V A S I K N F K A
15721 - CAGTTCTTATTATCAAATAATGTGTTTCATGTCTGAGGCAAATGTTGGACTGAGACTG -
15780
- Q F F I I K I M C S C L R Q N V G L R L
- S S L L S K * C V H V * G K M L D * D *
- V L Y Y Q N N V F M S E A K C W T E T D
15781 - ACCTTACTAAAGGACCTCACGAATTTGCTCACAGCATACAATGCTAGTTAAACAAGGAG -
15840
- T L L K D L T N F A H S I Q C * L N K E
- P Y * R T S R I L L T A Y N A S * T R R
- L T K G P H E F C S Q H T M L V K Q G D
15841 - ATGATTACGTGTACCTGCCTTACCCAGATCCATCAAGAATATTAGGCCAGGCTGTTTTG -
15900

```

---

```
- M I T C T C L T Q I H Q E Y * A Q A V L  
- * L R V P A L P R S I K N I R R R L F C  
- D Y V Y L P Y P D P S R I L G A G C F V  
15901 - TCGATGATATTGTCAAACAGATGGTACACTTATGATTGAAAGGTTTCGTCACTGGCTA -  
15960  
- S M I L S K Q M V H L * L K G S C H W L  
- R * Y C Q N R W Y T Y D * K V R V T G Y  
- D D I V K T D G T L M I E R F V S L A I
```

## 图 16-V

```

15961 - TTGATGCTTACCCACTTACAAAACATCCTAATCAGGAGTATGCTGATGTCTTTCACTTGT -
16020
- L M L T H L Q N I L I R S M L M S F T C
- * C L P T Y K T S * S G V C * C L S L V
- D A Y P L T K H P N Q E Y A D V F H L Y
16021 - ATTTACAATACATTAGAAAAGTTACATGATGAGCTTACTGGCCACATGTTGGACATGTATT -
16080
- I Y N T L E S Y M M S L L A T C W T C I
- F T I H * K V T * * A Y W P H V G H V F
- L Q Y I R K L H D E L T G H M L D M Y S
16081 - CCGTAATGCTAACTAATGATAACACCTCACGGTACTGGGAACCTGAGTTTTATGAGGCTA -
16140
- P * C * L M I T P H G T G N L S F M R L
- R N A N * * * H L T V L G T * V L * G Y
- V M L T N D N T S R Y W E P E F Y E A M
16141 - TGTACACACCACATACAGTCTTGCAGGCTGTAGGTGCTTGTGTATTGTGCAATTCACAGA -
16200
- C T H H I Q S C R L * V L V Y C A I H R
- V H T T Y S L A G C R C L C I V Q F T D
- Y T P H T V L Q A V G A C V L C N S Q T
16201 - CTTCACTTCGTTGCGGTGCCTGTATTAGGAGACCATTCTATGTTGCAAGTCTGCTATG -
16260
- L H F V A V P V L G D H S Y V A S A A M
- F T S L R C L Y * E T I P M L Q V L L *
- S L R C G A C I R R P F L C C K C C Y D
16261 - ACCATGTCAATTTCAACATCACACAATTAGTGTGTCTGTTAATCCCTATGTTTGCAATG -
16320
- T M S F Q H H T N * C C L L I P M F A M
- P C H F N I T Q I S V V C * S L C L Q C
- H V I S T S H K L V L S V N P Y V C N A
16321 - CCCAGGTTGTGATGTCACTGATGTGACACAACCTGTATCTAGGAGGTATGAGCTATTATT -
16380
- P Q V V M S L M * H N C I * E V * A I I
- P R L * C H * C D T T V S R R Y E L L L
- P G C D V T D V T Q L Y L G G M S Y Y C
16381 - GCAAGTCACATAAGCCTCCCATAGTTTCCATTATGTGCTAATGGTCAGGTTTTTGGTT -
16440
- A S H I S L P L V F H Y V L M V R F L V
- Q V T * A S H * F S I M C * W S G F W F
- K S H K P P I S F P L C A N G Q V F G L
16441 - TATACAAAAACACATGTGTAGGCAGTGACAATGTCACTGACTTCAATGCCATAGCAACAT -
16500
- Y T K T H V * A V T M S L T S M R * Q H
- I Q K H M C R Q * Q C H * L Q C D S N M
- Y K N T C V G S D N V T D F N A I A T C
16501 - GTGATTGGACTAATGCTGGCGATTACATACTTGCCAACACTTGTACTGAGAGACTCAAGC -
16560
- V I G L M L A I T Y L P T L V L R D S S
- * L D * C W R L H T C Q H L Y * E T Q A
- D W T N A G D Y I L A N T C T E R L K L
16561 - TTTTCGCAGCAGAAACGCTCAAAGCCACTGAGGAAACATTTAAGCTGTATATGGTATTG -
16620
- F S Q Q K R S K P L R K H L S C H M V L
- F R S R N A Q S H * G N I * A V I W Y C
- F A A E T L K A T E E T F K L S Y G I A
16621 - CCACTGTACGCGAAGTACTCTGTGACAGAGAATTGCATCTTTCATGGGAGGTTGGAAAAC -

```

16680

- P L Y A K Y S L T E N C I F H G R L E N  
- H C T R S T L \* Q R I A S F M G G W K T  
- T V R E V L S D R E L H L S W E V G K P

## 图 16-W

```

16681 - CTAGACCACCATTGAACAGAACTATGTCTTTACTGGTTACCGTGTAAGTAAATAGTA -
16740
- L D H H * T E T M S L L V T V * L K I V
- * T T I E Q K L C L Y W L P C N * K * *
- R P P L N R N Y V F T G Y R V T K N S K
16741 - AAGTACAGATTGGAGAGTACACCTTTGAAAAAGGTGACTATGGTGATGCTGTTGTGTACA -
16800
- K Y R L E S T P L K K V T M V M L L C T
- S T D W R V H L * K R * L W * C C C V Q
- V Q I G E Y T F E K G D Y G D A V V Y R
16801 - GAGGTACTACGACATACAAGTTGAATGTTGGTGATTACTTTTGTGTTGACATCTCACACTG -
16860
- E V L R H T S * M L V I T L C * H L T L
- R Y Y D I Q V E C W * L L C V D I S H C
- G T T T Y K L N V G D Y F V L T S H T V
16861 - TAATGCCACTTAGTGCACCTACTCTAGTGCCACAAGAGCACTATGTGAGAATTACTGGCT -
16920
- * C H L V H L L * C H K S T M * E L L A
- N A T * C T Y S S A T R A L C E N Y W L
- M P L S A P T L V P Q E H Y V R I T G L
16921 - TGTACCCAACTCAACATCTCAGATGAGTTTTCTAGCAATGTGCAAATTATCAAAGG -
16980
- C T Q H S T S Q M S F L A M L Q I I K R
- V P N T Q H L R * V F * Q C C K L S K G
- Y P T L N I S D E F S S N V A N Y Q K V
16981 - TCGGCATGCAAAAGTACTCTACACTCCAAGGACCACCTGGTACTGGTAAGAGTCATTTTG -
17040
- S A C K S T L H S K D H L V L V R V I L
- R H A K V L Y T P R T T W Y W * E S F C
- G M Q K Y S T L Q G P P G T G K S H F A
17041 - CCATCGGACTTGCTCTCTATTACCCATCTGCTCGCATAGTGATACGGCATGCTCTCATG -
17100
- P S D L L S I T H L L A * C I R H A L M
- H R T C S L L P I C S H S V Y G M L S C
- I G L A L Y Y P S A R I V Y T A C S H A
17101 - CAGCTGTTGATGCCCTATGTGAAAAGGCATTAATAATTTGCCCATAGATAAATGTAGTA -
17160
- Q L L M P Y V K R H * N I C P * I N V V
- S C * C P M * K G I K I F A H R * M * *
- A V D A L C E K A L K Y L P I D K C S R
17161 - GAATCATACCTGCGCGTGCGCGTAGAGTGTGATAAATCAAAGTGAATCAACAC -
17220
- E S Y L R V R A * S V L I N S K * I Q H
- N H T C A C A R R V F * * I Q S E F N T
- I I P A R A R V E C F D K F K V N S T L
17221 - TAGAACAGTATGTTTTCTGCACTGTAATGCATTGCCAGAAACAAGTCTGACATTGTAG -
17280
- * N S M F S A L * M H C Q K Q L L T L *
- R T V C F L H C K C I A R N N C * H C S
- E Q Y V F C T V N A L P E T T A D I V V
17281 - TCTTTGATGAAATCTCTATGGCTACTAATTATGACTTGAGTGTGTCAAATGCTAGACTTC -
17340
- S L M K S L W L L I M T * V L S M L D F
- L * * N L Y G Y * L * L E C C Q C * T S
- F D E I S M A T N Y D L S V V N A R L R

```



17341 - GTGCAAACACTACGTCTATATTGGCGATCCTGCTCAATTACCAGCCCCCGCACATTGC -  
17400  
- V Q N T T S I L A I L L N Y Q P P A H C  
- C K T L R L Y W R S C S I T S P P H I A  
- A K H Y V Y I G D P A Q L P A P R T L L  
17401 - TGACTAAAGGCACACTAGAACCAATATTTTAATTCAGTGTGCAGACTTATGAAAACAA -  
17460  
- \* L K A H \* N Q N I L I Q C A D L \* K Q  
- D \* R H T R T R I F \* F S V Q T Y E N N  
- T K G T L E P E Y F N S V C R L M K T I

## 图 16-X

```

17461 - TAGGTCCAGACATGTTCCCTTGGAACTTGTGCGCCGTTGTCTGCTGAAATGTTGACACTG -
17520
- * V Q T C S L E L V A V V L L K L L T L
- R S R H V P W N L S P L S C * N C * H C
- G P D M F L G T C R R C P A E I V D T V
17521 - TGAGTGCTTTAGTTTATGACAATAAGCTAAAAGCACACAAGGATAAGTCAGCTCAATGCT -
17580
- * V L * F M T I S * K H T R I S Q L N A
- E C F S L * Q * A K S T Q G * V S S M L
- S A L V Y D N K L K A H K D K S A Q C F
17581 - TCAAAATGTTCTACAAAGGTGTTATTACACATGATGTTTCATCTGCAATCAACAGACCTC -
17640
- S K C S T K V L L H M M F H L Q S T D L
- Q N V L Q R C Y Y T * C F I C N Q Q T S
- K M F Y K G V I T H D V S S A I N R P Q
17641 - AAATAGCGCTTGTAGAGAATTTCTTACACGCAATCTGCTTGGAGAAAAGCTGTTTAA -
17700
- K * A L * E N F L H A I L L G E K L F L
- N R R C K R I S Y T Q S C L E K S C F Y
- I G V V R E F L T R N P A W R K A V F I
17701 - TCTCACCTTATAATTCACAGAACGCTGTAGCTTCAAAAATCTTAGGATTGCCTACGCAGA -
17760
- S H L I I H R T L * L Q K S * D C L R R
- L T L * F T E R C S F K N L R I A Y A D
- S P Y N S Q N A V A S K I L G L P T Q T
17761 - CTGTTGATTCAACAGGGTCTGAATATGACTATGTCATATTCACACAACTACTGAAA -
17820
- L L I H H R V L N M T M S Y S H K L L K
- C * F I T G F * I * L C H I H T N Y * N
- V D S S Q G S E Y D Y V I F T Q T T E T
17821 - CAGCACACTCTTGAATGTCAACCGCTTCAATGTGGCTATCACAGGGCAAAAATTGGCA -
17880
- Q H T L V M S T A S M W L S Q G Q K L A
- S T L L * C Q P L Q C G Y H K G K N W H
- A H S C N V N R F N V A I T R A K I G I
17881 - TTTTGTGCATAATGTCTGATAGAGATCTTTATGACAACTGCAATTTACAAGTCTAGAAA -
17940
- F C A * C L I E I F M T N C N L Q V * K
- F V H N V * * R S L * Q T A I Y K S R N
- L C I M S D R D L Y D K L Q F T S L E I
17941 - TACCACCTCGCAATGTGGCTACATTACAAGCAGAAAATGTAAGTGGACTTTTAAAGGACT -
18000
- Y H V A M W L H Y K Q K M * L D F L R T
- T T S Q C G Y I T S R K C N W T F * G L
- P R R N V A T L Q A E N V T G L F K D C
18001 - GTAGTAGATCATTACTGGTCTTCATCCTACACAGGCACCTACACACCTCAGCGTTGATA -
18060
- V V R S L L V F I L H R H L H T S A L I
- * * D H Y W S S S Y T G T Y T P Q R * Y
- S K I I T G L H P T Q A P T H L S V D I
18061 - TAAAATCAAGACTGAAGGATTATGTGTTGACATACCAGGCATACCAAAGGACATGACCT -
18120
- * N S R L K D Y V L T Y Q A Y Q R T * P
- K I Q D * R I M C * H T R H T K G H D L
- K F K T E G L C V D I P G I P K D M T Y
18121 - ACCGTAGACTCATCTCTATGATGGGTTTCAAAATGAATTACCAAGTCAATGGTTACCTA -

```

18180

- T V D S S L \* W V S K \* I T K S M V T L  
- P \* T H L Y D G F Q N E L P S Q W L P \*  
- R R L I S M M G F K M N Y Q V N G Y P N

## 图 16- Y

```

18181 -
ATATGTTTATCACCCGCGAAGAAGCTATTTCGTCACGTTTCGTGCGTGGATTGGCTTTGATG - 18240
- I C L S P A K K L F V T F V R G L A L M
- Y V Y H P R R S Y S S R S C V D W L * C
- M F I T R E E A I R H V R A W I G F D V
18241 - TAGAGGGCTGTCATGCAACTAGAGATGCTGTGGGTACTAACCTACCTCTCCAGCTAGGAT -
18300
- * R A V M Q L E M L W V L T Y L S S * D
- R G L S C N * R C C G Y * P T S P A R I
- E G C H A T R D A V G T N L P L Q L G F
18301 - TTTCTACAGGTGTTAACTTAGTAGCTGTACCGACTGGTTATGTTGACACTGAAAATAACA -
18360
- F L Q V L T * * L Y R L V M L T L K I T
- F Y R C * L S S C T D W L C * H * K * H
- S T G V N L V A V P T G Y V D T E N N T
18361 - CAGAATTCACCAGAGTTAATGCAAACCTCCACCAGGTGACCAGTTTAAACATCTTATAC -
18420
- Q N S P E L M Q N L H Q V T S L N I L Y
- R I H Q S * C K T S T R * P V * T S Y T
- E F T R V N A K P P P G D Q F K H L I P
18421 - CACTCATGTATAAAGGCTTGCCCTGGAATGTAGTGCCTATTAAGATAGTACAAATGCTCA -
18480
- H S C I K A C P G M * C V L R * Y K C S
- T H V * R L A L E C S A Y * D S T N A Q
- L M Y K G L P W N V V R I K I V Q M L S
18481 - GTGATACACTGAAAGGATTGTCAGACAGAGTTCGTGTTTCGTCTTTGGGCGCATGGCTTTG -
18540
- V I H * K D C Q T E S C S S F G R M A L
- * Y T E R I V R Q S R V R P L G A W L *
- D T L K G L S D R V V F V L W A H G F E
18541 - AGCTTACATCAATGAAGTACTTTGTCAAGATTGGACCTGAAAGAACGTGTTGTCTGTGTG -
18600
- S L H Q * S T L S R L D L K E R V V C V
- A Y I N E V L C Q D W T * K N V L S V *
- L T S M K Y F V K I G P E R T C C L C D
18601 - ACAAACGTGCAACTTGCTTTTCTACTTCATCAGATACTTATGCCTGCTGGAATCATTCTG -
18660
- T N V Q L A F L L H Q I L M P A G I I L
- Q T C N L L F Y F I R Y L C L L E S F C
- K R A T C F S T S S D T Y A C W N H S V
18661 - TGGGTTTTGACTATGTCTATAACCCATTTATGATTGATGTTTCAGCAGTGGGGCTTTACGG -
18720
- W V L T M S I T H L * L M F S S G A L R
- G F * L C L * P I Y D * C S A V G L Y G
- G F D Y V Y N P F M I D V Q Q W G F T G
18721 - GTAACCTTCAGAGTAACCATGACCAACATTGCCAGGTACATGGAATGCACATGTGGCTA -
18780
- V T F R V T M T N I A R Y M E M H M W L
- * P S E * P * P T L P G T W K C T C G *
- N L Q S N H D Q H C Q V H G N A H V A S
18781 - GTTGTGATGCTATCATGACTAGATGTTTAGCAGTCCATGAGTGCTTTGTTAAGCGCGTTG -
18840
- V V M L S * L D V * Q S M S A L L S A L
- L * C Y H D * M F S S P * V L C * A R *
- C D A I M T R C L A V H E C F V K R V D
18841 - ATTGGTCTGTTGAATACCCATTATAGGAGATGAACTGAGGGTTAATTCTGCTGCAGAA -

```

18900  
- I G L L N T L L \* E M N \* G L I L L A E  
- L V C \* I P Y Y R R \* T E G \* F C L Q K  
- W S V E Y P I I G D E L R V N S A C R K  
18901 - AAGTACAACACATGGTTGTGAAGTCTGCATGCTTGCTGATAAGTTCCAGTTCTTCATG -  
18960  
- K Y N T W L \* S L H C L L I S F Q F F M  
- S T T H G C E V C I A C \* \* V S S S S \*  
- V Q H M V V K S A L L A D K F P V L H D

## 图 16-Z

```

18961 - ACATTGGAAATCCAAAGGCTATCAAGTGTGTGCCTCAGGCTGAAGTAGAATGGAAGTTCT -
19020
- T L E I Q R L S S V C L R L K * N G S S
- H W K S K G Y Q V C A S G * S R M E V L
- I G N P K A I K C V P Q A E V E W K F Y
19021 - ACGATGCTCAGCCATGTAGTGACAAAGCTTACAAAATAGAGGAACCTTCTATTCCTTATG -
19080
- T M L S H V V T K L T K * R N S S I L M
- R C S A M * * Q S L Q N R G T L L F L C
- D A Q P C S D K A Y K I E E L F Y S Y A
19081 - CTACACATCAGATAAAATCACTGATGGTGTGTTGTTTGGAAATGTAACGTTGATC -
19140
- L H I T I N S L M V F V C P G I V T L I
- Y T S R * I H * W C L F V L E L * R * S
- T H H D K F T D G V C L F W N C N V D R
19141 - GTTACCCAGCCAATGCAATTGTGTGTAGGTTTGACACAAGAGTCTTGTCAAACCTGAACT -
19200
- V T Q P M Q L C V G L T Q E S C Q T * T
- L P S Q C N C V * V * H K S L V K L E L
- Y P A N A I V C R F D T R V L S N L N L
19201 - TACCAGGCTGTGATGGTGGTAGTTTGTATGTGAATAAGCATGCATTCCACACTCCAGCTT -
19260
- Y Q A V M V V V C M * I S M H S T L Q L
- T R L * W W * F V C E * A C I P H S S F
- P G C D G G S L Y V N K H A F H T P A F
19261 - TCGATAAAGTGCAATTTACTAATTTAAAGCAATTGCCTTTCTTTACTATTCTGATAGTC -
19320
- S I K V H L L I * S N C L S F T I L I V
- R * K C I Y * F K A I A F L L L F * * S
- D K S A F T N L K Q L P F F Y Y S D S P
19321 - CTTGTGAGTCTCATGGCAAACAAGTAGTGTGCGGATATTGATTATGTTCCACTCAAATCTG -
19380
- L V S L M A N K * C R I L I M F H S N L
- L * V S W Q T S S V G Y * L C S T Q I C
- C E S H G K Q V V S D I D Y V P L K S A
19381 - CTACGTGTATTACAGGATGCAATTTAGGTGGTGTCTTTGCAGACACCATGCAAATGAGT -
19440
- L R V L H D A I * V V L F A D T M Q M S
- Y V Y Y T M Q F R W C C L Q T P C K * V
- T C I T R C N L G G A V C R H H A N E Y
19441 - ACCGACGTACTTGGATGCATATAATATGATGATTTCTGCTGGATTTAGCCTATGGATTT -
19500
- T D S T W M H I I * * F L L D L A Y G F
- P T V L G C I * Y D D F C W I * P M D L
- R Q Y L D A Y N M M I S A G F S L W I Y
19501 - ACAAACAATTTGATACTTATAACCTGTGGAATACATTTACCAGGTTACAGAGTTTAGAAA -
19560
- T N N L I L I T C G I H L P G Y R V * K
- Q T I * Y L * P V E Y I Y Q V T E F R K
- K Q F D T Y N L W N T F T R L Q S L E N
19561 - ATGTGGCTTATAATGTTGTTAATAAAGGACACTTTGATGGACACGCCGCGAAGCACCTG -
19620
- M W L I M L L I K D T L M D T P A K H L
- C G L * C C * * R T L * W T R R R S T C
- V A Y N V V N K G H F D G H A G E A P V
19621 - TTTCCATCATTAATAATGCTGTTTACCAAAGGTAGATGGTATTGATGFGAGATCTTTG -

```

19680  
 - F P S L I M L F T Q R \* M V L M W R S L  
 - F H H \* \* C C L H K G R W Y \* C G D L \*  
 - S I I N N A V Y T K V D G I D V E I F E

图 16-AA

19681 - AAAATAAGACAACACTTCTGTTAATGTTGCATTTGAGCTTTGGGCTAAGCGTAACATTA -  
 19740  
 - K I R Q H F L L M L H L S F G L S V T L  
 - K \* D N T S C \* C C I \* A L G \* A \* H \*  
 - N K T T L P V N V A F E L W A K R N I K  
 19741 - AACCAGTGCCAGAGATTAAGATACTCAATAATTGGGTGTTGATATCGCTGCTAATACTG -  
 19800  
 - N Q C Q R L R Y S I I W V L I S L L I L  
 - T S A R D \* D T Q \* F G C \* Y R C \* Y C  
 - P V P E I K I L N N L G V D I A A N T V  
 19801 - TAATCTGGGACTACAAAAGAGAAGCCCCAGCACATGTATCTACAATAGGTGTCTGCACAA -  
 19860  
 - \* S G T T K E K P Q H M Y L Q \* V S A Q  
 - N L G L Q K R S P S T C I Y N R C L H N  
 - I W D Y K R E A P A H V S T I G V C T M  
 19861 - TGACTGACATTGCCAAGAAACCTACTGAGAGTGCTTGTCTTCACTTACTGTCTTGTGTTG -  
 19920  
 - \* L T L P R N L L R V L V L H L L S C L  
 - D \* H C Q E T Y \* E C L F F T Y C L V \*  
 - T D I A K K P T E S A C S S L T V L P D  
 19921 - ATGGTAGAGTGAAGGACAGGTAGACCTTTTTAGAAACGCCCGTAATGGTGTTTTAATAA -  
 19980  
 - M V E W K D R \* T F L E T P V M V F \* \*  
 - W \* S G R T G R P P \* K R P \* W C F N N  
 - G R V E G Q V D L F R N A R N G V L I T  
 19981 - CAGAAGGTTCAAGGTTCAACACCTTCAAAGGGACCAGCACAAGCTAGCGTCAATG -  
 20040  
 - Q K V Q S K V \* H L Q R D Q H K L A S M  
 - R R F S Q R S N T F K G T S T S \* R Q W  
 - E G S V K G L T P S K G P A Q A S V N G  
 20041 - GAGTCACATTAATTGGAGAATCAGTAAAAACACAGTTTAACTACTTTAAGAAAGTAGACG -  
 20100  
 - E S H \* L E N Q \* K H S L T T L R K \* T  
 - S H I N W R I S K N T V \* L L \* E S R R  
 - V T L I G E S V K T Q F N Y F K K V D G  
 20101 - GCATTATTCAACAGTTGCCTGAAACCTACTTTACTCAGAGCAGAGACTTAGAGGATTTTA -  
 20160  
 - A L F N S C L K P T L L R A E T \* R I L  
 - H Y S T V A \* N L L Y S E Q R L R G F \*  
 - I I Q Q L P E T Y F T Q S R D L E D F K  
 20161 - AGCCAGATCACAAATGGAACACTGACTTCTCGAGCTCGCTATGGATGAATTCATACAGC -  
 20220  
 - S P D H K W K L T F S S S L W M N S Y S  
 - A Q I T N G N \* L S R A R Y G \* I H T A  
 - P R S Q M E T D F L E L A M D E F I Q R  
 20221 - GATATAAGCTCGAGGGCTATGCCTTCGAACACATCGTTTATGGAGATTTCACTCATGGAC -

20280  
- D I S S R A M P S N T S F M E I S V M D  
- I \* A R G L C L R T H R L W R F Q S W T  
- Y K L E G Y A F E H I V Y G D F S H G Q  
20281 - AACTTGGCGGTCTTCATTTAATGATAGGCTTAGCCAAGCGCTCACAAAGATTCACCACTTA -  
20340  
- N L A V F I \* \* \* A \* P S A H K I H H L  
- T W R S S F N D R L S Q A L T R F T T \*  
- L G G L H L M I G L A K R S Q D S P L K  
20341 - AATTAGAGGATTTTATCCCTATGGACAGCACAGTGAAAAATTACTTCATAACAGATGGGC -  
20400  
- N \* R I L S L W T A Q \* K I T S \* Q M R  
- I R G F Y P Y G Q H S E K L L H N R C A  
- L E D F I P M D S T V K N Y F I T D A Q  
20401 - AAACAGGTTTCATCAAAATGTGTGTCTCTGTGATTGATCTTTTACTTGATGACTTTGTGC -  
20460  
- K Q V H Q N V C V L \* L I F Y L M T L S  
- N R F I K M C V F C D \* S F T \* \* L C R  
- T G S S K C V C S V I D L L L D D F V E



## 图 16-BB

20461 - AGATAATAAAGTCACAAGATTGTGTCAGTGATTTCAAAGTGGTCAAGGTTACAATTGACT -  
 20520  
 - R \* \* S H K I C Q \* F Q K W S R L Q L T  
 - D N K V T R F V S D F K S G Q G Y N \* L  
 - I I K S Q D L S V I S K V V K V T I D Y  
 20521 - ATGCTGAAATTTCAATTCATGCTTTGGTGTAAAGGATGGACATGTTGAAACCTTCTACCCAA -  
 20580  
 - M L K F H S C F G V R M D M L K P S T Q  
 - C \* N F I H A L V \* G W T C \* N L L P K  
 - A E I S F M L W C K D G H V E T F Y P K  
 20581 - AACTACAAGCAAGTCAAGCGTGGCAACCAGGTGTTGCGATGCCCTAACTTGTACAAGATGC -  
 20640  
 - N Y K Q V K R G N Q V L R C L T C T R C  
 - T T S K S S V A T R C C D A \* L V Q D A  
 - L Q A S Q A W Q P G V A M P N L Y K M Q  
 20641 - AAAGAATGCTTCTTGAAGTGTGACCTTCAGAATTATGGTGAAGTGTGTTATACCAA -  
 20700  
 - K E C F L K S V T F R I M V K M L L Y Q  
 - K N A S \* K V \* P S E L W \* K C C Y T K  
 - R M L L E K C D L Q N Y G E N A V I P K  
 20701 - AAGGAATAATGATGAATGTGCGCAAAGTATACTCAACTGTGCAATACTTAAATACACTTA -  
 20760  
 - K E \* \* \* M S Q S I L N C V N T \* I H L  
 - R N N D E C R K V Y S T V S I L K N T Y  
 - G I M M N V A K Y T Q L C Q Y L N Y T T  
 20761 - CTTTAGCTGTACCCTACAACATGAGAGTTATTCACCTTTGGTGTGGCTCTGATAAAGGAG -  
 20820  
 - L \* L Y P T T \* E L F T L V L A L I K E  
 - F S C T L Q H E S Y S L W C W L \* \* R S  
 - L A V P Y N M R V I H F G A G S D K G V  
 20821 - TTGCACCAGGTACAGCTGTGCTCAGACAATGGTTCGCAACTGGCACACTACTTGTGCGATT -  
 20880  
 - L H Q V Q L C S D N G C Q L A H Y L S I  
 - C T R Y S C A Q T M V A N W H T T C R F  
 - A P G T A V L R Q W L P T G T L L V D S  
 20881 - CAGATCTTAATGACTTCGTCTCCGACGAGATTCTACTTTAATGGAGACTGTGCAACAG -  
 20940  
 - Q I L M T S S P T Q I L L \* L E T V Q Q  
 - R S \* \* L R L R R R F Y F N W R L C N S  
 - D L N D F V S D A D S T L I G D C A T V  
 20941 - TACATACGGCTAATAAATGGGACCTTATTATTAGCGATATGTATGACCCTAGGACCAAAC -  
 21000  
 - Y I R L I N G T L L L A I C M T L G P N  
 - T Y G \* \* M G P Y Y \* R Y V \* P \* D Q T  
 - H T A N K W D L I I S D M Y D P R T K H  
 21001 - ATGTGACAAAAGAGAATGACTCTAAAGAAGGGTTTTCACTTATCTGTGTGGATTATAA -  
 21060  
 - M \* Q K R M T L K K G F S L I C V D L \*  
 - C D K R E \* L \* R R V F H L S V W I Y K  
 - V T K E N D S K E G F F T Y L C G F I K  
 21061 - AGCAAAACTAGCCCTGGGTGGTTCATAGCTGTAAAGATAACAGAGCATCTTGGGAATG -  
 21120  
 - S K N \* P W V V L \* L \* R \* Q S I L G M  
 - A K T S P G W F Y S C K D N R A F L E C  
 - Q K L A L G G S I A V K I T E H S W N A  
 21121 - CTGACCTTTACAAGCTTATGGGCCATTTCTCATGGTGGACAGCTTTTGTACAAATGTAA -

21180

- L T F T S L W A I S H G G Q L L L Q M \*  
- \* P L Q A Y G P F L M V D S F C Y K C K  
- D L Y K L M G H F S W W T A F V T N V N

## 图 16-CC

```

21181 - ATGCATCATCATCGGAAGCATTTTTAATTGGGGCTAACTATCTTGGCAAGCCGAAGGAAC -
21240 - M H H H R K H F * L G L T I L A S R R N
- C I I I G S I F N W G * L S W Q A E G T
- A S S S E A F L I G A N Y L G K P K E Q
21241 - AAATTGATGGCTATACCATGCATGCTAACTACATTTTCTGGGGAACACAAATCCTATCC -
21300 - K L M A I P C M L T T F S G G T Q I L S
- N * W L Y H A C * L H F L E E H K S Y P
- I D G Y T M H A N Y I F W R N T N P I Q
21301 - AGTTGTCTTCTTACTCTTTGACATGAGCAAATTCCTCTTAAATTAAGAGGAAGT -
21360 - S C L P I H S L T * A N F L L N * E E L
- V V F L P T L * H E Q I S S * I K R N C
- L S S Y S L F D M S K F P L K L R G T A
21361 - CTGTAATGTCTTTAAGGAGAATCAAATCAATGATATGATTTATTCTTCTTGGAAAAAG -
21420 - L * C L L R R I K S M I * F I L F W K K
- C N V S * G E S N Q * Y D L F S S G K R
- V M S L K E N Q I N D M I Y S L L E K G
21421 - GTAGGCTTATCATTAGAGAAAACAACAGAGTTGTGGTTTCAAGTGATATCTTGTTAACA -
21480 - V G L S L E K T T E L W F Q V I F L L T
- * A Y H * R K Q Q S C G F K * Y S C * Q
- R L I I R E N N R V V V S S D I L V N N
21481 - ACTAAACGAACATGTTTATTCTTATTATTCTTACTCTCACTAGTGGTAGTGACCTTG -
21540 - T K R T C L F S Y Y F L L S L V V V T L
- L N E H V Y F L I I S Y S H * W * * P *
- * T N M F I F L L F L T L T S G S D L D
21541 - ACCGGTGCACCACTTTTGATGATGTTCAAGCTCCTAATTACACTCAACATACTTCATCTA -
21600 - T G A P L L M M F K L L I T L N I L H L
- P V H H F * * C S S S * L H S T Y F I Y
- R C T T F D D V Q A P N Y T Q H T S S M
21601 - TGAGGGGGTTTACTATCCTGATGAAATTTTGTAGATCAGACACTCTTTATTTAACTCAGG -
21660 - * G G F T I L M K F L D Q T L F I * L R
- E G G L L S * * N F * I R H S L F N S G
- R G V Y Y P D E I F R S D T L Y L T Q D
21661 - ATTTATTTCTTCCATTTTATTCTAATGTTACAGGGTTTCATACTATTAATCATACTTTG -
21720 - I Y F F H F I L M L Q G F I L L I I R L
- F I S S I L F * C Y R V S Y Y * S Y V W
- L F L P F Y S N V T G F H T I N H T F G
21721 - GCAACCCTGTCATACCTTTAAGGATGGTATTTATTTTGCTGCCACAGAGAAATCAAATG -
21780 - A T L S Y L L R M V F I L L P Q R N Q M
- Q P C H T F * G W Y L F C C H R E I K C
- N P V I P F K D G I Y F A A T E K S N V
21781 - TTGTCCGTGGTTGGGTTTGGTTCTACCATGAACAACAAGTCACAGTCGGTGATTATTA -
21840 - L S V V G F L V L P * T T S H S R * L L
- C P W L G F W F Y H E Q Q V T V G D Y Y
- V R G W V F G S T M N N K S Q S V I I I

```

21841 - TTAACAATTCTACTAATGTTGTTATACGAGCATGTAAC TTGAATTGTGTGACAACCCTT -  
21900 .  
- L T I L L M L L Y E H V T L N C V T T L  
- \* Q F Y \* C C Y T S M \* L \* I V \* Q P F  
- N N S T N V V I R A C N F E L C D N P F

```

21901 - TCTTTGCTGTTTCTAAACCCATGGGTACACAGACACATACTATGATATTCGATAATGCAT -
21960 - S L L F L N P W V H R H I L * Y S I M H
- L C C F * T H G Y T D T Y Y D I R * C I
- F A V S K P M G T Q T H T M I F D N A F
21961 - TTAATTCACCTTTCGAGTACATATCTGATGCCTTTTCGCTTGATGTTTCAGAAAAGTCAG -
22020 - L I A L S S T Y L M P F R L M F Q K S Q
- * L H F R V H I * C L F A * C F R K V R
- N C T F E Y I S D A F S L D V S E K S G
22021 - GTAATTTTAAACACTTACGAGAGTTTGTGTTTAAAAATAAGATGGGTTTCTCTATGTTT -
22080 - V I L N T Y E S L C L K I K M G F S M F
- * F * T L T R V C V * K * R W V S L C L
- N F K H L R E F V F K N K D G F L Y V Y
22081 - ATAAGGGCTATCAACCTATAGATGTAGTTCGCTGATCTACCTTCTGGTTTAAACACTTGA -
22140 - I R A I N L * M * F V I Y L L V L T L *
- * G L S T Y R C S S * S T F W F * H F E
- K G Y Q P I D V V R D L P S G F N T L K
22141 - AACCTATTTTAAAGTTGCCTCTGGTATTAACATTACAAATTTTAGAGCCATTCTTACAG -
22200 - N L F L S C L L V L T L Q I L E P F L Q
- T Y F * V A S W Y * H Y K F * S H S Y S
- P I F K L P L G I N I T N F R A I L T A
22201 - CCTTTCACCTGCTCAAGACATTTGGGGCACGTCAGCTGCAGCCTATTTTGTGGCTATT -
22260 - P F H L L K T F G A R Q L Q P I L L A I
- L F T C S R H L G H V S C S L F C W L F
- F S P A Q D I W G T S A A A Y F V G Y L
22261 - TAAAGCCAACACTATTTATGCTCAAGTATGATGAAAATGGTACAATCACAGATGCTGTTG -
22320 - * S Q L H L C S S M M K M V Q S Q M L L
- K A N Y I Y A Q V * * K W Y N H R C C *
- K P T T F M L K Y D E N G T I T D A V D
22321 - ATTGTTCTCAAATCCACTTGTGGAACCTCAAATGCTCTGTTAAGAGCTTTGAGATTGACA -
22380 - I V L K I H L L N S N A L L R A L R L T
- L F S K S T C * T Q M L C * E L * D * Q
- C S Q N P L A E L K C S V K S F E I D K
22381 - AAGGAATTTACCAGACCTCTAATTTACAGGTTGTTCCCTCAGGAGATGTTGTGAGATTCC -
22440 - K E F T R P L I S G L F P Q E M L * D S
- R N L P D L * F Q G C S L R R C C E I P
- G I Y Q T S N F R V V P S G D V V R F P
22441 - CTAATATTACAACTTGTGTCCTTTTGGAGAGGTTTTTAATGCTACTAAATCCCTTCTG -
22500 - L I L Q T C V L L E R F L M L L N S L L
- * Y Y K L V S F W R G F * C Y * I P F C
- N I T N L C P F G E V F N A T K F P S V
22501 - TCTATGCATGGGAGAGAAAAAATTTCTAATTTGTTGCTGATTACTCTGTGCTTACA -
22560 - S M H G R E K K F L I V L L I T L C S T
- L C M G E K K N F * L C C * L L C A L Q
- Y A W E R K K I S N C V A D Y S V L Y N

```

图 16-DD

---

22561 - ACTCAACATTTTTTCAACCTTTAAGTGCTATGGCGTTTCTGCCACTAAGTTGAATGATC -  
22620  
- T Q H F F Q P L S A M A F L P L S \* M I  
- L N I F F N L \* V L W R F C H \* V E \* S  
- S T F F S T F K C Y G V S A T K L N D L

```

22621 - TTTGCTTCTCCAATGTCTATGCAGATTCTTTTGTAGTCAAGGGAGATGATGTAAGACAAA -
22680 - F A S P M S M Q I L L * S R E M M * D K
- L L L Q C L C R F F C S Q G R * C K T N
- C F S N V Y A D S F V V K G D D V R Q I
22681 - TAGCGCCAGGACAACTGGTGTATTGCTGATTATAATTATAAATTGCCAGATGATTCA -
22740 - * R Q D K L V L L L I I I I N C Q M I S
- S A R T N W C Y C * L * L * I A R * F H
- A P G Q T G V I A D Y N Y K L P D D F M
22741 - TGGGTGTGTCCTTGCTTGAATACTAGGAACATTGATGCTACTTCAACTGGTAATTATA -
22800 - W V V S L L G I L G T L M L L Q L V I I
- G L C P C L E Y * E H * C Y F N W * L *
- G C V L A W N T R N I D A T S T G N Y N
22801 - ATTATAAATATAGGTATCTTAGACATGGCAAGCTTAGGCCCTTTGAGAGAGACATATCTA -
22860 - I I N I G I L D M A S L G P L R E T Y L
- L * I * V S * T W Q A * A L * E R H I *
- Y K Y R Y L R H G K L R P F E R D I S N
22861 - ATGTGCCTTTCTCCCTGATGGCAAACCTTGACCCACCTGCTCTTAATTGTTATTGGC -
22920 - M C L S P L M A N L A P H L L L I V I G
- C A F L P * W Q T L H P T C S * L L L A
- V P F S P D G K P C T P P A L N C Y W P
22921 - CATTAAATGATTATGGTTTTTACACCACTACTGGCATTGGCTACCAACCTTACAGAGTTG -
22980 - H * M I M V F T P L L A L A T N L T E L
- I K * L W F L H H Y W H W L P T L Q S C
- L N D Y G F Y T T T G I G Y Q P Y R V V
22981 - TAGTACTTTCTTTGAACTTTTAAATGCACCGCCACGGTTTGTGGACAAAATTATCCA -
23040 - * Y F L L N F * M H R P R F V D Q N Y P
- S T F F * T F K C T G H G L W T K I I H
- V L S F E L L N A P A T V C G P K L S T
23041 - CTGACCTTATTAAGAACCAGTGTGTCAATTTTAAATTTTAAATGGACTACTGGTACTGGTG -
23100 - L T L L R T S V S I L I L M D S L V L V
- * P Y * E P V C Q F * F * W T H W Y W C
- D L I K N Q C V N F N F N G L T G T G V
23101 - TGTTAACTCCTTCTTCAAAGAGATTTCAACCATTTCAACAATTTGGCCGTGATGTTTCTG -
23160 - C * L L L Q R D F N H F N N L A V M F L
- V N S F F K E I S T I S T I W P * C F *
- L T P S S K R F Q P F Q Q F G R D V S D
23161 - ATTTCACTGATTCGGTTCGAGATCCTAAAACATCTGAAATATTAGACATTTACCTTGCT -
23220 - I S L I P F E I L K H L K Y * T F H L A
- F H * F R S R S * N I * N I R H F T L L
- F T D S V R D P K T S E I L D I S P C S
23221 - CTTTGGGGGTGTAAGTGAATTACACCTGGAACAAATGCTTCATCTGAAGTTGCTGTTG -
23280 - L L G V * V * L H L E Q M L H L K L L F
- F W G C K C N Y T W N K C F I * S C C S
- F G G V S V I T P G T N A S S E V A V L
23281 - TATATCAAGATGTTAACTGCACTGATGTTTCTACAGCAATTCATGCAGATCAACTCACAC -

```

图 16-EE

23340

- Y I K M L T A L M F L Q Q F M Q I N S H  
- I S R C \* L H \* C F Y S N S C R S T H T  
- Y Q D V N C T D V S T A I H A D Q L T P



```

23341 - CAGCTTGGCGCATATATTCTACTGGAAACAATGTATTCCAGACTCAAGCAGGCTGTCTTA -
23400
- Q L G A Y I L L E T M Y S R L K Q A V L
- S L A H I F Y W K Q C I P D S S R L S Y
- A W R I Y S T G N N V F Q T Q A G C L I
23401 - TAGGAGCTGAGCATGTGACACTTCTTATGAGTGGCAGATTCCCTATFGAGCTGGCATT -
23460
- * E L S M S T L L M S A T F L L E L A F
- R S * A C R H F L * V R H S Y W S W H L
- G A E H V D T S Y E C D I P I G A G I C
23461 - GTGCTAGTTACCATACAGTTTCTTTATTACGTAGTACTAGCCAAAATCTATTGTGGCTT -
23520
- V L V T I Q F L Y Y V V L A K N L L W L
- C * L P Y S F F I T * Y * P K I Y C G L
- A S Y H T V S L L R S T S Q K S I V A Y
23521 - ATACTATGTCTTTAGGTGCTGATAGTTCAATTGCTTACTCTAATAACACCATTGCTATAC -
23580
- I L C L * V L I V Q L L T L I T P L L Y
- Y Y V F R C * * F N C L L * * H H C Y T
- T M S L G A D S S I A Y S N N T I A I P
23581 - CTACTACTTTTCAATTAGCATTACTACAGAAGTAATGCCTGTTTCTATGGCTAAAACCT -
23640
- L L T F Q L A L L Q K * C L F L W L K P
- Y * L F N * H Y Y R S N A C F Y G * N L
- T N F S I S I T T E V M P V S M A K T S
23641 - CCGTAGATTGTAATATGTACATCTGCGGAGATTCTACTGAATGTGCTAATTTGCTCTCC -
23700
- P * I V I C T S A E I L L N V L I C F S
- R R L * Y V H L R R F Y * M C * F A S P
- V D C N M Y I C G D S T E C A N L L L Q
23701 - AATATGGTAGCTTTTGCACACAACTAAATCGTGCCTCTCAGGTATTGCTGCTGAACAGG -
23760
- N M V A F A H N * I V H S Q V L L L N R
- I W * L L H T T K S C T L R Y C C * T G
- Y G S F C T Q L N R A L S G I A A E Q D
23761 - ATCGCAACACAGTGAAGTGTGCTCAAGTCAAACAAATGTACAAAACCCCACTTTGA -
23820
- I A T H V K C S L K S N K C T K P Q L *
- S Q H T * S V R S S Q T N V Q N P N F E
- R N T R E V F A Q V K Q M Y K T P T L K
23821 - AATATTTTGGTGGTTTAAATTTTTCACAAATATTACCTGACCCTCTAAAGCCAACCTAAGA -
23880
- N I L V V L I F H K Y Y L T L * S Q L R
- I F W W F * F F T N I T * P S K A N * E
- Y F G G F N F S Q I L P D P L K P T K R
23881 - GGTCTTTTATTGAGGACTTGCTCTTTAATAAGGTGACACTCGCTGATGCTGGCTTCATGA -
23940
- G L L L R T C S L I R * H S L M L A S *
- V F Y * G L A L * * G D T R * C W L H E
- S F I E D L L F N K V T L A D A G F M K
23941 - AGCAATATGGCGAATGCCTAGGTGATATTAATGCTAGAGATCTCATTGTGCGCAGAAGT -
24000
- S N M A N A * V I L M L E I S F V R R S
- A I W R M P R * Y * C * R S H L C A E V
- Q Y G E C L G D I N A R D L I C A Q K F
24001 - TCAATGGACTTACAGTGTGGCCACCTGCTCACTGATGATATGATTGCTGCCTACACTG -

```

图 16-FF

24060

- S M D L Q C C H L C S L M I \* L L P T L  
- Q W T Y S V A T S A H \* \* Y D C C L H C  
- N G L T V L P P L L T D D M I A A Y T A

```

24061 - CTGCTCTAGTTAGTGGTACTGCCACTGCTGGATGGACATTTGGTGCTGGCGCTGCTCTTC -
24120 - L L * L V V L P L L D G H L V L A L L F
- C S S * W Y C H C W M D I W C W R C S S
- A L V S G T A T A G W T F G A G A A L Q
24121 - AAATACCTTTTGCTATGCAAATGGCATATAGGTTCAATGGCATTGGAGTTACCCAAAATG -
24180 - K Y L L L C K W H I G S M A L E L P K M
- N T F C Y A N G I * V Q W H W S Y P K C
- I P F A M Q M A Y R F N G I G V T Q N V
24181 - TTCTCTATGAGAACC AAAACAAATCGCCACCAATTTAACAAGGCGATTAGTCAAATTC -
24240 - F S M R T K N K S P T N L T R R L V K F
- S L * E P K T N R Q P I * Q G D * S N S
- L Y E N Q K Q I A N Q F N K A I S Q I Q
24241 - AAGAATCACTTACAACAACATCAACTGCATTGGGCAAGCTGCAAGACGTTGTTAACCAGA -
24300 - K N H L Q Q H Q L H W A S C K T L L T R
- R I T Y N N I N C I G Q A A R R C * P E
- E S L T T T S T A L G K L Q D V V N Q N
24301 - ATGCTCAAGCATAAACACACTTGTAAACAACCTTAGCTCTAATTTTGGTGCAATTTCAA -
24360 - M L K H * T H L L N N L A L I L V Q F Q
- C S S I K H T C * T T * L * F W C N F K
- A Q A L N T L V K Q L S S N F G A I S S
24361 - GTGTGCTAAATGATATCCTTTCCGCACTTGATAAAGTCGAGGCGGAGGTACAAATTGACA -
24420 - V C * M I S F R D L I K S R R R Y K L T
- C A K * Y P F A T * * S R G G G T N * Q
- V L N D I L S R L D K V E A E V Q I D R
24421 - GGTAAATTACAGGCAGACTTCAAAGCCTTCAAACCTATGTAACACAACAACCTAATCAGGG -
24480 - G * L Q A D F K A F K P M * H N N * S G
- V N Y R Q T S K P S N L C N T T T N Q G
- L I T G R L Q S L Q T Y V T Q Q L I R A
24481 - CTGCTGAAATCAGGGCTTCTGCTAATCTTGCTGCTACTAAAATGTCTGAGTGTGTTCTTG -
24540 - L L K S G L L L I L L L L K C L S V F L
- C * N Q G F C * S C C Y * N V * V C S W
- A E I R A S A N L A A T K M S E C V L G
24541 - GACAATCAAAAAGAGTTGACTTTTGTGAAAGGCTACCACCTTATGTCCTTCCCACAAG -
24600 - D N Q K E L T F V E R A T T L C P S H K
- T I K K S * L L W K G L P P Y V L P T S
- Q S K R V D F C G K G Y H L M S F P Q A
24601 - CAGCCCGCATGGTGTGCTTCTTCTACATGTCAGTATGTGCCATCCAGGAGGAACT -
24660 - Q P R M V L S S Y M S R M C H P R R G T
- S P A W C C L P T C H V C A I P G B E L
- A P H G V V F L H V T Y V P S Q E R N F
24661 - TCACCACAGCGCCAGCAATTTGTCATGAAGGCAAGCATACTTCCCTCGTGAAGGTGTTT -
24720 - S P Q R Q Q F V M K A K H T S L V K V F
- H H S A S N L S * R Q S I L P S * R C F
- T T A P A I C H E G K A Y F P R E G V F
24721 - TTGTGTTAATGGCACTTCTTGGTTTATTACACAGAGGAACTTCTTTTCTCCACAAATAA -
24780

```

图 16-GG

---

- L C L M A L L G L L H R G T S F L H K \*  
- C V \* W H F L V Y Y T E E L L F S T N N  
- V F N G T S W F I T Q R N F F S P Q I I  
24781 - TTACTACAGACAATACATTTGTCTCAGGAAATTGTGATGTCGTTATTGGCATCATTACA -  
24840  
- L L Q T I H L S Q E I V M S L L A S L T  
- Y Y R Q Y I C L R K L \* C R Y W H H \* Q  
- T T D N T F V S G N C D V V I G I I N N

```

24841 - ACACAGTTTATGATCCTCTGCAACCTGAGCTTGACTCATTCAAAGAAGAGCTGGACAAGT -
24900 - T Q F M I L C N L S L T H S K K S W T S
- H S L * S S A T * A * L I Q R R A G Q V
- T V Y D P L Q P E L D S F K E E L D K Y
24901 - ACTTCAAAAATCATAATCACCAGATGTTGATCTTGGCGACATTTTCAGGCATTAACGCTT -
24960 - T S K I I H H Q M L I L A T F Q A L T L
- L Q K S Y I T R C * S W R H F R H * R F
- F K N H T S P D V D L G D I S G I N A S
24961 - CTGTCGTCAACATTCAAAAAGAAATTGACCCGCTCAATGAGGTCGCTAAAAATTTAAATG -
25020 - L S S T F K K K L T A S M R S L K I * M
- C R Q H S K R N * P P Q * G R * K F K *
- V V N I Q K E I D R L N E V A K N L N E
25021 - AATCACTCATTGACCTTCAAGAATTGGGAAAATATGAGCAATATATTAATGGCCTTGTT -
25080 - N H S L T F K N W E N M S N I L N G L G
- I T H * P S R I G K I * A I Y * M A L V
- S L I D L Q E L G K Y E Q Y I K W P W Y
25081 - ATGTTTGGCTCGGCTTTCATTGCTGGACTAATTGCCATCGTCATGGTTACAATCTTGCTTT -
25140 - M F G S A S L L D * L P S S W L Q S C F
- C L A R L H C W T N C H R H G Y N L A L
- V W L G F I A G L I A I V M V T I L L C
25141 - GTTGCATGACTAGTTGTTGCAGTTGCCCTCAAGGGTGCATGCTCTTTGTGGTTCTTGCTGCA -
25200 - V A * L V V A V A S R V H A L V V L A A
- L H D * L L Q L P Q G C M L L W F L L Q
- C M T S C C S C L K G A C S C G S C C K
25201 - AGTTTGATGAGGATGACTCTGAGCCAGTTCTCAAGGGTGTCAAATTACATTACACATAAA -
25260 - S L M R M T L S Q F S R V S N Y I T H K
- V * * G * L * A S S Q G C Q I T L H I N
- F D E D D S E P V L K G V K L H Y T * T
25261 - CGAACTTATGGATTTGTTTATGAGATTTTTTACTCTTGGATCAATTAAGTGCACAGCCAGT -
25320 - R T Y G F V Y E I F Y S W I N Y C T A S
- E L M D L F M R F F T L G S I T A Q P V
- N L W I C L * D F L L L D Q L L H S Q *
25321 - AAAAAATTGACAATGCTTCTCCTGCAAGTACTGTTTCATGCTACAGCAACGATACCGCTACA -
25380 - K N * Q C F S C K Y C S C Y S N D T A T
- K I D N A S P A S T V H A T A T I P L Q
- K L T M L L L Q V L F M L Q Q R Y R Y K
25381 - AGCCTCACTCCCTTTCGGATGGCTTGTATTGGCGTTGCATTTCTTGCTGTTTTTCAGAG -
25440 - S L T P F R M A C Y W R C I S C C F S E
- A S L P F G W L V I G V A F L A V F Q S
- P H S L S D G L L L A L H F L L F F R A
25441 - CGCTACCAAAATAATTGCGCTCAATAAAGATGGCAGCTAGCCCTTATAAGGGCTTCCA -
25500 - R Y Q N N C A Q * K M A A S P L * G L P
- A T K I I A L N K R W Q L A L Y K G F Q
- L P K * L R S I K D G S * P F I R A S S
25501 - GTTCATTTGCAATTTACTGCTGCTATTTGTACCATCTATTACATCTTTTGCTTGCTGCG -

```

图 16-HH

25560

- V H L Q F T A A I C Y H L F T S F A C R  
- F I C N L L L L F V T I Y S H L L L V A  
- S F A I Y C C Y L L P S I H I F C L S L

25561 - TGCAGGTAAGGAGGCGCAATTTTTGTACCTCTATGCCTTGATATATTTTCTACAATGCAT -  
25620  
- C R \* G G A I F V P L C L D I F S T M H  
- A G K E A Q F L Y L Y A L I Y F L Q C I  
- Q V R R R N F C T S M P \* Y I F Y N A S  
25621 - CAACGCATGTAGAATTATTATGAGATGTTGGCTTTGTTGGAAGTGCAAATCCAAGACCC -  
25680  
- Q R M \* N Y Y E M L A L L E V Q I Q E P  
- N A C R I I M R C W L C W K C K S K N P  
- T H V E L L \* D V G F V G S A N P R T H  
25681 - ATTACTTTATGATGCCAACTACTTTGTTTGCTGGCACACACATAACTATGACTACTGTAT -  
25740  
- I T L \* C Q L L C L L A H T \* L \* L L Y  
- L L Y D A N Y F V C W H T H N Y D Y C I  
- Y F M M P T T L F A G T H I T M T T V Y  
25741 - ACCATATAACAGTGTCCAGATACAATTGTCGTTACTGAAGGTGACGGCATTTCACACCC -  
25800  
- T I \* Q C H R Y N C R Y \* R \* R H F N T  
- P Y N S V T D T I V V T E G D G I S T P  
- H I T V S Q I Q L S L L K V T A F Q H Q  
25801 - AAAACTCAAAGAAGACTACCAAATGGTGGTTATTCTGAGGATAGGCACTCAGGTGTAA -  
25860  
- K T Q R R L P N W W L F \* G \* A L R C \*  
- K L K E D Y Q I G G Y S E D R H S G V K  
- N S K K T T K L V V I L R I G T Q V L K  
25861 - AGACTATGTCGTTGTACATGGCTATTTACCGAAGTTTACTACCAGCTTGAGTCTACACA -  
25920  
- R L C R C T W L F H R S L L P A \* V Y T  
- D Y V V V H G Y F T E V Y Y Q L E S T Q  
- T M S L Y M A I S P K F T T S L S L H K  
25921 - AATTACTACAGACTGGTATTGAAAATGTACATTCTTCATCTTAAACAAGCTTGTTAA -  
25980  
- N Y Y R H W Y \* K C Y I L H L \* Q A C \*  
- I T T D T G I E N A T F F I F N K L V K  
- L L Q T L V L K M L H S S S L T S L L K  
25981 - AGACCCACCGAATGTGCAAATACACAAATCGACGGCTCTTCAGGAGTTGCTAATCCAGC -  
26040  
- R P T E C A N T H N R R L F R S C \* S S  
- D P P N V Q I H T I D G S S G V A N P A  
- T H R M C K Y T Q S T A L Q E L L I Q Q  
26041 - AATGGATCCAATTTATGATGAGCCGACGACTACTAGCGTGCCTTTGTAAGCACAAGA -  
26100  
- N G S N L \* \* A D D D Y \* R A F V S T R  
- M D P I Y D E P T T T T S V P L \* A Q E  
- W I Q F M M S R R R L L A C L C K H K K  
26101 - AAGTGAGTACGAACCTTATGTACTCATTGTTTCGGAAGAAACAGGTACGTTAATAGTTAA -  
26160  
- K \* V R T Y V L I R F G R N R Y V N S \*  
- S E Y E L M Y S F V S E E T G T L I V N  
- V S T N L C T H S F R K K Q V R \* \* L I  
26161 - TAGCGTACTTCTTTTCTGCTTTTCGTTGATTCTTGCTAGTCACACTAGCCATCCTTAC -  
26220  
- \* R T S F S C F R G I L A S H T S H P Y  
- S V L L F L A F V V F L L V T L A I L T  
- A Y F F F L L S W Y S C \* S H \* P S L L  
26221 - TGCGCTTCGATTGTGTGCGTACTGCTGCAATATGTTAACGTGAGTTTAGTAAAACCAAC -  
26280

图 16-11

```
- C A S I V C V L L Q Y C * R E F S K T N  
- A L R L C A Y C C N I V N V S L V K P T  
- R F D C V R T A A I L L T * V * * N Q R  
26281 - G G T T A C G T C T A C T C G C G T G T T A A A A A T C T G A A C T C T T C T G A A G G A G T T C C T G A T C T T C T -  
26340  
- G L R L L A C * K S E L F * R S S * S S  
- V Y V Y S R V K N L N S S E G V P D L L  
- F T S T R V L K I * T L L K E F L I F W
```



```

26341 - GGTCTAAACGAACTAACTATTATTATTATTCTGTTTGGAACTTTAACATTGCTTATCATG -
26400
- G L N E L T I I I I L F G T L T L L I M
- V * T N * L L L L F C L E L * H C L S W
- S K R T N Y Y Y Y S V W N F N I A Y H G
26401 - GCAGACAACGGTACTATTACCCTTGAGGAGCTTAAACAACCTCCTGGAAACAATGGAACCTA -
26460
- A D N G T I T V E E L K Q L L E Q W N L
- Q T T V L L P L R S L N N S W N N G T *
- R Q R Y Y Y R * G A * T T P G T M E P S
26461 - GTAATAGGTTTCCTATTCCCTAGCCTGGATTATGTTACTACAATTTGCCTATTCTAATCGG -
26520
- V I G F L F L A W I M L L Q F A Y S N R
- * * V S Y S * P G L C Y Y N L P I L I G
- N R F P I P S L D Y V T T I C L F * S E
26521 - AACAGGTTTTGTACATAATAAAGCTTGTTCCTCTGGCTCTGTGGCCAGTAACACTT -
26580
- N R F L Y I I K L V F L W L L W P V T L
- T G F C T * * S L F S S G S C G Q * H L
- Q V F V H N K A C F P L A L V A S N T C
26581 - GCTTGTTCCTGCTGCTGTTGCTACAGAATTAATTGGGTGACTGGCGGGATTGCGATT -
26640
- A C F V L A V V Y R I N W V T G G I A I
- L V L C L L L S T E L I G * L A G L R L
- L F C A C C C L Q N * L G D W R D C D C
26641 - GCAATGGCTTGTATTGTAGGCTTGATGTGGCTTAGCTACTTCGTTGCTTCCTTACGGCTG -
26700
- A M A C I V G L M W L S Y F V A S F R L
- Q W L V L * A * C G L A T S L L P S G C
- N G L Y C R L D V A * L L R C F L Q A V
26701 - TTTGCTCGTACCCGCTCAATGTGGTCATTCAACCCAGAAACAACATTCTTCTCAATGTG -
26760
- F A R T R S M W S F N P E T N I L L N V
- L L V P A Q C G H S T Q K Q T F F S M C
- C S Y P L N V V I Q P R N K H S S Q C A
26761 - CCTCTCCGGGGACAATTGTGACCAGACCGCTCATGGAAAGTGAACCTTGTCAATGGTGCT -
26820
- P L R G T I V T R P L M E S E L V I G A
- L S G G Q L * P D R S W K V N L S L V L
- S P G D N C D Q T A H G K * T C H W C C
26821 - GTGATCATTGCTGGTCACTTGCGAATGGCCGGACACTCCCTAGGGCGCTGTGACATTAAG -
26880
- V I I R G H L R M A G H S L G R C D I K
- * S F V V T C E W P D T P * G A V T L R
- D H S W S L A N G R T L P R A L * H * G
26881 - GACCTGCCAAAAGAGATCACTGTGGCTACATCAGAACGCTTTTCTATTACAAATTAGGA -
26940
- D L P K E I T V A T S R T L S Y Y K L G
- T C Q K R S L W L H H E R F L I T N * E
- P A K R D H C G Y I T N A F L L Q I R S
26941 - GCGTCGCAGCGTGTAGGCACTGATTCAGGTTTTGCTGCATACAACCGCTACCGTATTGGA -
27000
- A S Q R V G T D S G F A A Y N R Y R I G
- R R S V * A L I Q V L L H T T A T V L E
- V A A C R H * F R F C C I Q P L P Y W K
27001 - AACTATAAATTAATACAGACCACGCCGGTAGCAACGACAATATTGCTTTGCTAGTACAG -
27060

```

图 16-JJ

- N Y K L N T D H A G S N D N I A L L V Q  
- T I N \* I Q T T P V A T T I L L C \* Y S  
- L \* I K Y R P R R \* Q R Q Y C F A S T V

```

27061 - TAAGTGACAACAGATGTTTCATCTTGTTGACTTCCAGGTTACAATAGCAGAGATATTGAT -
27120
- * V T T D V S S C * L P G Y N S R D I D
- K * Q Q M F H L V D F Q V T I A E I L I
- S D N R C F I L L T S R L Q * Q R Y * L
27121 - TATCATTATGAGGACTTTCAGGATGCTATTTGGAATCTTGACGTTATAATAAGTTCAAT -
27180
- Y H Y E D F Q D C Y L E S * R Y N K F N
- I I M R T F R I A I W N L D V I I S S I
- S L * G L S G L L F G I L T L * * V Q *
27181 - AGTGAGACAATTATTTAAGCCTCTAACTAAGAAGATTATTCGGAGTTAGATGATGAAGA -
27240
- S E T I I * A S N * E E L F G V R * * R
- V R Q L F K P L T K K N Y S E L D D E E
- * D N Y L S L * L R R I I R S * M M K N
27241 - ACCTATGGAGTTAGATTATCCATAAAACGAACATGAAAATTATTCTCTTCCGACATTGA -
27300
- T Y G V R L S I K R T * K L F S S * H *
- P M E L D Y P * N E H E N Y S L P D I D
- L W S * I I H K T N M K I I L F L T L I
27301 - TTGTATTACATCTTGCGAGCTATATCACTATCAGGAGTGTGTTAGAGGTACGACTGTAC -
27360
- L Y L H L A S Y I T I R S V L E V R L Y
- C I Y I L R A I S L S G V C * R Y D C T
- V F T S C E L Y H Y Q E C V R G T T V L
27361 - TACTAAAGAACCTTGCCCATCAGGAACATACGAGGGCAATTCACCATTTACCCTCTTG -
27420
- Y * K N L A H Q E H T R A I H H F T L L
- T K R T L P I R N I R G Q F T I S P S C
- L K E P C P S G T Y E G N S P F H P L A
27421 - CTGACAATAAATTTGCACTAACTTGCACTAGCACACTTTGCTTTGCTTGTGCTGACG -
27480
- L T I N L H * L A L A H T L L L L V L T
- * Q * I C T N L H * H T L C F C L C * R
- D N K F A L T C T S T H F A F A C A D G
27481 - GTACTCGACATACCTATCAGCTGCGTGCAAGATCAGTTTCACCAAACCTTTTCATCAGAC -
27540
- V L D I P I S C V Q D Q F H Q N F S S D
- Y S T Y L S A A C K I S F T K T F H Q T
- T R H T Y Q L R A R S V S P K L F I R Q
27541 - AAGAGGAGGTTCAACAAGAGCTCTACTCGCCACTTTTTCTCATTGTTGCTGCTCTAGTAT -
27600
- K R R F N K S S T R H F F S L L L L * Y
- R G G S T R A L L A T F S H C C C S S I
- E E V Q Q E L Y S P L F L I V A A L V F
27601 - TTTTAATACTTTGCTTCCACCATTAAGAGAAAGACAGAATGAATGAGCTCACTTTAATTGA -
27660
- F * Y F A S P L R E R Q N E * A H F N *
- F N T L L H H * E K D R M N E L T L I D
- L I L C F T I K R K T E * M S S L * L T
27661 - CTTCTATTTGTGCTTTTTAGCCTTTCGCTATTCCTTGTTTAATAATGCTTATTATATT -
27720
- L L F V L F S L S A I P C F N N A Y Y I
- F Y L C F L A F L L F L V L I M L I I F
- S I C A F * P P C Y S L F * * C L L Y F

```

图 16-KK

---

27721 - TTGGTTTCACTCGAAATCCAGGATCTAGAAGAACCTTGTACCAAAGTCTAAACGAACAT -  
27780  
- L V F T R N P G S R R T L Y Q S L N E H  
- W F S L E I Q D L E E P C T K V \* T N M  
- G F H S K S R I \* K N L V P K S K R T \*

```

27781 - GAAACTTCTCATTGTTTTGACTTGTATTTCTCTATGCAGTTGCATATGCACTGTAGTACA -
27840 - E T S H C F D L Y F S M Q L H M H C S T
- K L L I V L T C I S L C S C I C T V V Q
- N F S L F * L V F L Y A V A Y A L * Y S
27841 - GCGCTGTGCATCTAATAAACCTCATGTGCTTGAAGATCCTTGTAAAGGTACAACACTAGGG -
27900 - A L C I * * T S C A * R S L * G T T L G
- R C A S N K P H V L E D P C K V Q H * G
- A V H L I N L M C L K I L V R Y N T R G
27901 - GTAATACTTATAGCACTGCTTGGCTTGTGCTCTAGGAAAGGTTTTACCTTTTCATAGAT -
27960 - V I L I A L L G F V L * E R F Y L F I D
- * Y L * H C L A L C S R K G F T F S * M
- N T Y S T A W L C A L G K V L P F H R W
27961 - GGCACACTATGGTTCAAACATGCACACCTAATGTTACTATCAACTGTCAAGATCCAGCTG -
28020 - G T L W F K H A H L M L L S T V K I Q L
- A H Y G S N M H T * C Y Y Q L S R S S W
- H T M V Q T C T P N V T I N C Q D P A G
28021 - GTGGTGCCTTATAGCTAGGTGTTGGTACCTTCATGAAGGTCAACAACTGCTGCATTTA -
28080 - V V R L * L G V G T F M K V T K L L H L
- W C A Y S * V L V P S * R S P N C C I *
- G A L I A R C W Y L H E G H Q T A A F R
28081 - GAGACGTACTTGTGTTTTAAATAAACGAACAAATAAAATGTCTGATAATGGACCCCAA -
28140 - E T Y L L F * I N E Q I K M S D N G P Q
- R R T C C F K * T N K L K C L I M D P N
- D V L V V L N K R T N * N V * * W T P I
28141 - TCAAACCAACGTAGTGCCCCCGCATTACATTTGGTGGACCCACAGATTCAACTGACAAT -
28200 - S N Q R S A P R I T F G G P T D S T D N
- Q T N V V P P A L H L V D P Q I Q L T I
- K P T * C P P H Y I W W T H R F N * Q *
28201 - AACCAGAATGGAGGACGCAATGGGGCAAGGCCAAAACAGCGCCGACCCCAAGGTTTACCC -
28260 - N Q N G G R N G A R P K Q R R P Q G L P
- T R M E D A M G Q G Q N S A D P K V Y P
- P E W R T Q W G K A K T A P T P R F T Q
28261 - AATAATACTGCGTCTTGGTTCACAGCTCTCACTCAGCATGGCAAGGAGGAAGCTTAGATTC -
28320 - N N T A S W F T A L T Q H G K E E L R F
- I I L R L G S Q L S L S M A R R N L D S
- * Y C V L V H S S H S A W Q G G T * I P
28321 - CCTCGAGGCCAGGGCGTTCCAATCAACACCAATAGTGGTCCAGATGACCAATGGCTAC -
28380 - P R G Q G V P I N T N S G P D D Q I G Y
- L E A R A F Q S T P I V V Q M T K L A T
- S R P G R S N Q H Q * W S R * P N W L L
28381 - TACCGAAGAGCTACCCGACGAGTTCGTGGTGGTACGGCAAATGAAAGAGCTCAGCCCC -
28440 - Y R R A T R R V R G G D G K M K E L S P
- T E E L P D E F V V V T A K * K S S A P
- P K S Y P T S S W W * R Q N E R A Q P Q

```

图 16-LL

---

28441 - AGATGGTACTTCTATTACCTAGGAACTGGCCAGAGCTTCACTTCCCTACGGCGCTAAC -  
28500  
- R W Y F Y Y L G T G P E A S L P Y G A N  
- D G T S I T \* E L A Q K L H F P T A L T  
- M V L L L P R N W P R S F T S L R R \* Q

```

28501 - AAAGAAGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATT -
28560
- K E G I V W V A T E G A L N T P K D H I
- K K A S Y G L Q L R E P * I H P K T T L
- R R H R M G C N * G S L E Y T Q R P H W
28561 - GGCACCCGCAATCCTAATAACAATGCTGCCACCGTCTACAACCTCCTCAAGGAACAACA -
28620
- G T R N P N N N A A T V L Q L P Q G T T
- A P A I L I T M L P P C Y N F L K E Q H
- H P Q S * * Q C C H R A T T S S R N N I
28621 - TTGCCAAAAGGCTTCTACGCAGAGGGAAGCAGAGGCGGCAAGCCTCTTCTCGCTCC -
28680
- L P K G F Y A B G S R G G S Q A S S R S
- C Q K A S T Q R E A E A A V K P L L A P
- A K R L L R R G K Q R R Q S S L F S L L
28681 - TCATCACGTAGTCGCGTAATTCAAGAAATCAACTCCTGGCAGCAGTAGGGAAATTC -
28740
- S S R S R G N S R N S T P G S S R G N S
- H H V V A V I Q E I Q L L A A V G E I L
- I T * S R * F K K F N S W Q Q * G K F S
28741 - CCTGCTCGAATGGCTAGCGGAGGTGGTGAACCTGCCCTCGCGCTATTGCTGCTAGACAGA -
28800
- P A R M A S G G G E T A L A L L L L D R
- L L E W L A E V V K L P S R Y C C * T D
- C S N G * R R W * N C P R A I A A R Q I
28801 - TTGAACCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAAACAAGGCCAAACTGTC -
28860
- L N Q L E S K V S G K G Q Q Q Q G Q T V
- * T S L R A K F L V K A N N N K A K L S
- E P A * E Q S F W * R P T T T R P N C H
28861 - ACTAAGAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCAAAACGTAAGTCCACAAA -
28920
- T K K S A A E A S K K P R Q K R T A T K
- L R N L L L R H L K S L A K N V L P Q N
- * E I C C * G I * K A S P K T Y C H K T
28921 - CAGTACAACGTCACCTCAAGCATTGGGAGACGTGGTCCAGAACAACCAAGGAAATTC -
28980
- Q Y N V T Q A F G R R G P E Q T Q G N F
- S T T S L K H L G D V V Q N K P K E I S
- V Q R H S S I W E T W S R T N P R K F R
28981 - GGGGACCAAGACCTAATCAGACAAGGAACTGATTACAACATTGGCCGCAAATGCACAA -
29040
- G D Q D L I R Q G T D Y K H W P Q I A Q
- G T K T * S D K E L I T N I G R K L H N
- G P R P N Q T R N * L Q T L A A N C T I
29041 - TTTGCTCCAAGTGCCTCTGCATTCTTGGAAATGTCACGCATTGGCATGGAAGTCACACCT -
29100
- F A P S A S A F F G M S R I G M E V T P
- L L Q V P L H S L E C H A L A W K S H L
- C S K C L C I L W N V T H W H G S H T F
29101 - TCGGGAACATGGCTGACTTATCATGGAGCCATTAAATGGATGACAAAGATCCACAATTC -
29160
- S G T W L T Y H G A I K L D D K D P Q F
- R E H G * L I M E P L N W M T K I H N S
- G N M A D L S W S H * I G * Q R S T I Q
29161 - AAAGACACGTCATACTGTGAACAAGCACATTGACGCATACAAAACATTCCCACCAACA -
29220

```

图 16-MM

---

- K D N V I L L N K H I D A Y K T F P P T  
- K T T S Y C \* T S T L T H T K H S H Q Q  
- R Q R H T A E Q A H \* R I Q N I P T N R  
29221 - GAGCCTAAAAAGGACAAAAAGAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAA -  
29280  
- E P K K D K K K K T D E A Q P L P Q R Q  
- S L K R T K R K R L M K L S L C R R D K  
- A \* K G Q K E K D \* \* S S A F A A E T K



```

29281 - AAGAAGCAGCCCCTGTGACTCTTCTTCTGCGGCTGACATGGATGATTTCTCCAGACAA -
29340 -
- K K Q P T V T L L P A A D M D D F S R Q
- R S S P L * L F F L R L T W M I S P D N
- E A A H C D S S S C G * H G * F L Q T T
29341 - CTTCAAATTCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATG -
29400 -
- L Q N S M S G A S A D S T Q A * T L M M
- F K I P * V E L L L I Q L R H K H S * *
- S K F H E W S F C * F N S G I N T H D D
29401 - ACCACACAAGGCAGATGGGCTATGTAAACGTTTTTCGCAATTCGGTTTACGATACATAGTC -
29460 -
- T T Q G R W A M * T F S Q F R L R Y I V
- P H K A D G L C K R F R N S V Y D T * S
- H T R Q M G Y V N V F A I P F T I H S L
29461 - TACTCTGTGCAGAATGAATTCTCGTAACTAACAGCACAAGTAGGTTTACTTAACCTTA -
29520 -
- Y S C A E * I L V T K Q H K * V * L T L
- T L V Q N E F S * L N S T S R F S * L *
- L L C R M N S R N * T A Q V G L V N F N
29521 - ATCTCACATAGCAATCTTTAATCAATGTGTAAACATTAGGGAGGACTTGAAAGAGCCACCA -
29580 -
- I S H S N L * S M C N I R E D L K E P P
- S H I A I P N Q C V T L G R T * K S H H
- L T * Q S L I N V * H * G G L E R A T T
29581 - CATTTCATCGAGGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAG -
29640 -
- H F H R G H A E Y D R G Y S E * C * G E
- I F I E A T R S T I E G T V N N A R E S
- F S S R P R G V R S R V Q * I M L G R A
29641 - CTGCCATATGGAAGAGCCCTAATGTGTAAATTAATTTTAGTAGTGCTATCCCCATGTG -
29700 -
- L P I W K S P N V * N * F * * C Y P H V
- C L Y G R A L M C K I N F S S A I P M *
- A Y M E E P * C V K L I L V V L S P C D
29701 - ATTTTAATAGCTTCTTAGGAGAATGACAAAAAAAAAAAAAAAAA - 29742
- I L I A S * E N D K K K K K X
- F * * L L R R M T K K K K X
- F N S F L G E * Q K K K K X

```

图 16-NN

```

1 - TTTTTTTTTTTTTTGTTCATCTCCTAAGAAGCTATTAATAACACATGGGGATAGCACTA - 60
- F F F F F V I L L R S Y * N H M G I A L
- F F F F L S F S * E A I K I T W G * H Y
- F F F F C H S P K K L L K S H G D S T T
61 - CTAAAATTAATTTTACACATTAGGGCTCTCCATATAGGCAGCTCTCCCTAGCATTATTC - 120
- L K L I L H I R A L P Y R Q L S L A L F
- * N * F Y T L G L F H I G S S P * H Y S
- K I N F T H * G S S I * A A L P S I I H
121 - ACTGTACCCTCGATCGTACTCCGCGTGGCCTCGATGAAAATGGTGGCTCTTTCAAGTC - 180
- T V P S I V L R V A S M K M W W L F Q V
- L Y P R S Y S A W P R * K C G G S F K S
- C T L D R T P R G L D E N V V A L S S P
181 - CTCCTAATGTTACACATTGATTAAGATTGCTATGTGAGATTAAGTTAACTAAACCTA - 240
- L P N V T H * L K I A M * D * S * L N L
- S L M L H I D * R L L C E I K V N * T Y
- P * C Y T L I K D C Y V R L K L T K P T
241 - CTTGTGCTGTTTAGTTACGAGAATTCATTCTGCACAAGAGTAGACTATGTATCGTAAACG - 300
- L V L F S Y E N S F C T R V D Y V S * T
- L C C L V T R I H S A Q E * T M Y R K R
- C A V * L R E F I L H K S R L C I V N G
301 - GAATTGCGAAAACGTTTACATAGCCCATCTGCCTTGTGTGGTCATCATGAGTGTATATGC - 360
- E L R K R L H S P S A L C G H H E C L C
- N C E N V Y I A H L P C V V I M S V Y A
- I A K T F T * P I C L V W S S * V F M P
361 - CTGAGTTGAATCAGCAGAAGCTCCACTCATGAATTTGAAGTTGTCTGGAGAATCATC - 420
- L S * I S R S S T H G I L K L S G E I I
- * V E S A E A P L M E F * S C L E K S S
- E L N Q Q K L H S W N F E V V W R N H P
421 - CATGTCAGCCGCAGGAGAAGAGTACAGTGGGCTGCTCTTTTGTCTCTGCGGCAAAGG - 480
- H V S R R R K K S H S G L L L L S L R Q R
- M S A A G R R V T V G C F F C L C G K G
- C Q P Q E E E S Q W A A S F V S A A K A
481 - CTGAGCTTCATCAGTCTTTTTCTTTTTGTCTTTTTAGGCTCTGTTGGTGGGAATGTTTT - 540
- L S F I S L F L F V L F R L C W W E C F
- * A S S V F F F L S F L G S V G G N V L
- E L H Q S F S F C P F * A L L V G M F C
541 - GTATGCGTCAATGTGCTTGTTCAGCAGTATGACGTTGTCTTTGAATTGTGGATCTTTGTC - 600
- V C V N V L V Q Q Y D V V F E L W I F V
- Y A S M C L F S S M T L S L N C G S L S
- M R Q C A C S A V * R C L * I V D L C H
601 - ATCCAATTTAATGGCTCCATGATAAGTCAGCCATGTTCCCGAAGGTGTGACTTCCATGCC - 660
- I Q F N G S M I S Q P C S R R C D F H A
- S N L M A P * * V S H V P E G V T S M P
- P I * W L H D K S A M F P K V * L P C Q
661 - AATGCGTGACATTCCAAAGAATGCAGAGGCCTTGGAGCAAATGTGCAATTTGCGGCCA - 720
- N A * H S K E C R G T W S K L C N L R P
- M R D I P K N A E A L G A N C A I C G Q
- C V T F Q R M Q R H L E Q I V Q F A A N

```

图 17-A

图 17-B

```

721 - ATGTTTGAATCAGTTCCTTGCTGATTAGGTCTTGGTCCCCGAAATTCCTTGGGTTTG - 780
- M F V I S S L S D * V L V P E I S L G L
- C L * S V P C L I R S W S P K F P W V C
- V C N Q F L V * L G L G P R N F L G F V
781 - TTCTGGACCACGTCTCCCAAATGCTTGAGTGACGTTGTACTGTTTGTGGCAGTACGTTT - 840
- F W T T S P K C L S D V V L F C G S T F
- S G P R L P N A * V T L Y C F V A V R F
- L D H V S Q M L E * R C T V L W Q Y V F
841 - TTGGCGAGGCTTTTGTAGATGCCTCAGCAGCAGATTCTTAGTGACAGTTGGCCCTGTTG - 900
- L A R L F R C L S S R F L S D S L A L L
- W R G F L D A S A A D F L V T V W P C C
- G E A F * M P Q Q Q I S * * Q F G L V V
901 - TTGTTGGCCTTACCAGAACTTTGCTCTCAAGCTGGTCAATCTGTCTAGCAGCAATAG - 960
- L L A F T R N F A L K L V Q S V * Q Q *
- C W P L P E T L L S S W F N L S S S N S
- V G L Y Q K L C S Q A G S I C L A A I A
961 - CGCGAGGGCAGTTTACCACCTCCGCTAGCCATTGAGCAGGAGAATTTCCCCTACTGCT -
1020
- R E G S F T T S A S H S S R R I S P T A
- A R A V S P P P L A I R A G E F P L L L
- R G Q F H H L R * P F E Q E N F P Y C C
1021 - GCCAGGAGTTGAATTTCTTGAATTACCGCGACTACGTGATGAGGAGCGAGAAGAGGCTTG -
1080
- A R S * I S * I T A T T * * G A R R G L
- P G V E F L E L P R L R D E E R E E A *
- Q E L N F L N Y R D Y V M R S E K R L D
1081 - ACTGCCGCTCTGCTTCCCTCTGCGTAGAAGCCTTTTGGCAATGTTGTTCTTGAGGAAG -
1140
- T A A S A S L C V E A F W Q C C S L R K
- L P P L L P S A * K P F G N V V P * G S
- C R L C F P L R R S L L A M L F L E V
1141 - TTGTAGCACGGTGGCAGCATTGTTATTAGGATTGGGGTGCCAAATGTTGGTCTTTGGGTGT -
1200
- L * H G G S I V I R I A G A N V V F G C
- C S T V A A L L L G L R V P M W S L G V
- V A R W Q H C Y * D C G C Q C G L W V Y
1201 - ATTCAGGCTCCCTCAGTTGCAACCCATACGATGCCTTCTTGTAGCGCGTAGGGAAG -
1260
- I Q G S L S C N P Y D A F F V S A V G K
- F K A P S V A T H T M P S L L A P * G S
- S R L P Q L Q P I R C L L C * R R R E V
1261 - TGAAGCTTCTGGGCCAGTTCCTAGGTAATAGAAGTACCATCTGGGGCTGAGCTCTTCAT -
1320
- * S P W A S S * V I E V P S G A E L F H
- E A S G P V P R * * K Y H L G L S S F I
- K L L G Q F L G N R S T I W G * A L S F
1321 - TTTGCCGTACCACCACGAACCTCGTCCGGTAGCTCTTCCGGTAGCCRAATTTGGTCATC -
1380
- F A V T T T N S S G S S S V V A N L V I
- L P S P P R T R R V A L R * * P I W S S
- C R H H H E L V G * L F G S S Q F G H L
1381 - TGGACCACTATGGTGTGATTGGAACGCCCTGGCCTCGAGGGAATCTAAGTTCCCTCTT -
1440
- W T T I G V D W N A L A S R E S K F L L
- G P L L V L I G T P W P R G N L S S S L
- D H Y W C * L E R P G L E G I * V P P C
1441 - GCCATGCTGAGTGAGAGCTGTGAACCAAGACGCAGTATTATTGGGTAAACCTTGGGGTGC -
1500

```

- A M L S E S C E P R R S I I G \* T L G S  
 - P C \* V R A V N Q D A V L L G K P W G R  
 - H A E \* E L \* T K T Q Y Y W V N L G V G

图 17-C

1501 - GCGCTGTTTTGGCCTTGCCCCATTGCGTCCTCCATTCTGGTTATTGTTCAGTTGAATCTGT -  
 1560  
 - A L F W P C P I A S S I L V I V S \* I C  
 - R C F G L A P L R P P F W L L S V E S V  
 - A V L A L P H C V L H S G Y C Q L N L W  
 1561 - GGGTCCACCAATGTAATGCGGGGGGCACTACGTTGGTTTGATTGGGGTCCATTATCAGA -  
 1620  
 - G S T K C N A G G . T T L V \* L G S I I R  
 - G P P N V M R G A L R W F D W G P L S D  
 - V H Q M \* C G G H Y V G L I G V H Y Q T  
 1621 - CATTTAATTTGTTTCGTTTATTAAACAACAAGTACGTCTCTAAATGCAGCAGTTTGGT -  
 1680  
 - H F N L F V Y L K Q Q V R L \* M Q Q F G  
 - I L I C S F I \* N N K Y V S K C S S L V  
 - F \* F V R L F K T T S T S L N A A V W \*  
 1681 - GACCTTCATGAAGGTACCAACACCTAGCTATAAGCGCACCACCAGCTGGATCTTGACAGT -  
 1740  
 - D L H E G T N T \* L \* A H H Q L D L D S  
 - T F M K V P T P S Y K R T T S W I L T V  
 - P S \* R Y Q H L A I S A P P A G S \* Q L  
 1741 - TGATAGTAACATTAGGTGTGCATGTTGAACCATAGTGTGCCATCTATGAAAAGGTAAAA -  
 1800  
 - \* \* \* H \* V C M F E P \* C A I Y E K V K  
 - D S N I R C A C L N H S V P S M K R \* N  
 - I V T L G V H V \* T I V C H L \* K G K T  
 1801 - CCTTTCCTAGAGCACAAAGCCAAGCAGTGCCTATAAGTATTACCCCTAGTGTGTACCTTA -  
 1860  
 - P F L E H K A K Q C Y K Y Y P \* C C T L  
 - L S \* S T K P S S A I S I T P S V V P Y  
 - F P R A Q S Q A V L \* V L P L V L Y L T  
 1861 - CAAGGATCTTCAAGCACATGAGGTTTATTAGATGCACAGCGCTGTACTACAGTGCATATG -  
 1920  
 - Q G S S S T \* G L L D A Q R C T T V H M  
 - K D L Q A H E V Y \* M H S A V L Q C I C  
 - R I F K H M R F I R C T A L Y Y S A Y A  
 1921 - CAACTGCATAGAGAAATACAAGTCAAACAATGAGAAGTTTCATGTTTCGTTTAGACTTTG -  
 1980  
 - Q L H R E I Q V K T M R S F M F V \* T L  
 - N C I E K Y K S K Q \* E V S C S F R L W  
 - T A \* R N T S Q N N E K F H V R L D F G  
 1981 - GTACAAGGTTCTTCTAGATCCTGGATTTTCGAGTGA AAAACCAAATATAATAAGCATTATT -  
 2040  
 - V Q G S S R S W I S S E N Q N I I S I I  
 - Y K V L L D P G F R V K T K I \* \* A L L  
 - T R F F \* I L D F E \* K P K Y N K H Y \*  
 2041 - AAAACAAGGAATAGCAGAAAGGCTAAAAGCACAAATAGAAGTCAATTAAGTGAGCTCA -  
 2100  
 - K T R N S R K A K K H K \* K S I K V S S

```

- K Q G I A E R L K S T N R S Q L K * A H
- N K E * Q K G * K A Q I E V N * S E L I
2101 - TTCATTCTGTCTTTCTCTTAATGGTGAAGCAAAGTATTAATAAATACTAGAGCAGCAACAA -
2160
- F I L S F S * W * S K V L K I L E Q Q Q
- S F C L S L N G E A K Y * K Y * S S N N
- H S V F L L M V K Q S I K N T R A A T M
2161 - TGAGAAAAGTGGCGAGTAGAGCTCTTGTGAACCTCCTCTTGTCTGATGAAAAGTTTG -
2220
- * E K V A S R A L V E P P L V * * K V L
- E K K W R V E L L L N L L L S D E K F W
- R K S G E * S S C * T S S C L M K S F G

```

图 17-D

```

2221 - GTGAAACTGATCTTGCACGCAGCTGATAGGTATGTCGAGTACCGTCAGCACAAAGCAAAG -
2280
- V K L I L H A A D R Y V E Y R Q H K Q K
- * N * S C T Q L I G M S S T V S T S K S
- E T D L A R S * * V C R V P S A Q A K A
2281 - CAAAGTGTGTGCTAGTGCAGTTAGTGCAAATTTATTGTCAGCAAGAGGGTGAAATGGTG -
2340
- Q S V C * C K L V Q I Y C Q Q E G E M V
- K V C A S A S * C K F I V S K R V K W *
- K C V L V Q V S A N L L S A R G * N G E
2341 - AATTGCCCTCGTATGTTCTGATGGGCAAGTTCCTTTTAGTAGTACAGTCTGACCTCTAA -
2400
- N C P R M F L M G K V L L V V Q S Y L *
- I A L V C S * W A R F F * * Y S R T S N
- L P S Y V P D G Q G S F S S T V V P L T
2401 - CACACTCCTGATAGTGATATAGCTCGCAAGATGTAATACAATCAATGTCAGGAAGAGAA -
2460
- H T P D S D I A R K M * I Q S M S G R E
- T L L I V I * L A R C K Y N Q C Q E E N
- H S * * * Y S S Q D V N T I N V R K R I
2461 - TAATTTTCATGTTGTTTATGGATAATCTAACTCCATAGGTTCTTCATCATCTAACTCC -
2520
- * F S C S F Y G * S N S I G S S S S S N S
- N F H V R F M D N L T P * V L H H L T P
- I F M F V L W I I * L H R F F I I * L R
2521 - GAATAATCTTCTTAGTTAGAGGCTTAAATAATTGTTCTCACTATTGAACTTATATAAG -
2580
- E * F F L V R G L N N C L T I E L I I T
- N N S S * L E A * I I V S L L N L L * R
- I I L L S * R L K * L S H Y * T Y Y N V
2581 - TCAAGATCCAAATAGCAATCCTGAAAGTCTCATAATGATAATCAATATCTCTGCTATT -
2640
- S R F Q I A I L K V L I M I I N I S A I
- Q D S K * Q S * K S S * * * S I S L L L
- K I P N S N P E S P H N D N Q Y L C Y C
2641 - GTAACCTGGAAGTCAACAAGATCAACATCTGTTGTCACTTACTGTACTAGCAAAGCAAT -
2700

```

```

- V T W K S T R * N I C C H L L Y * Q S N
- * P G S Q Q D E T S V V T Y C T S K A I
- N L E V N K M K H L L S L T V L A K Q Y
2701 - ATTGTCGTTGCTACCGCGTGGTCTGTATTTAATTTATAGTTTCCAATACGGTAGCGGTT -
2760
- I V V A T G V V C I * F I V S N T V A V
- L S L L P A W S V F N L * F P I R * R L
- C R C Y R R G L Y L I Y S F Q Y G S G C
2761 - GTATGCAGCAAAACCTGAATCAGTGCCTACACGCTGCGACGCTCCTAATTTGTAATAAGA -
2820
- V C S K T * I S A Y T L R R S * F V I R
- Y A A K P E S V P T R C D A P N L * * E
- M Q Q N L N Q C L H A A T L L I C N K K
2821 - AAGCGTTCGTGATGTAGCCACAGTGATCTCTTTTGGCAGGTCCTTAATGTCACAGCGCCC -
2880
- K R S * C S H S D L F W Q V L N V T A P
- S V R D V A T V I S F G R S L M S Q R P
- A F V M * P Q * S L L A G P * C H S A L
2881 - TAGGGAGTGTCCGGCCATTCCGCAAGTGACCACGAATGATCACAGCACCAATGACAAGTTC -
2940
- * G V S G H S Q V T T N D H S T N D K F
- R E C P A I R K * P R M I T A P M T S S
- G S V R P P F A S D H E * S Q H Q * Q V H
2941 - ACTTTCCATGAGCGGTCGTTGTCACAATTGTCCCCGGAGAGGCACATTGAGAAGAATGTT -
3000
- T F H E R S G H N C P P E R H I E K N V
- L S M S G L V T I V P R R G T L R R M F
- F P * A V W S Q L S P G E A H * E E C L

```

```

3001 - TGTTCCTGGGTTGAATGACCACATTGAGCGGGTACGAGCAAACAGCCTGAAGGAAGCAAC -
3060
- C F W V E * P H * A G T S K Q P E G S N
- V S G L N D H I E R V R A N S L K E A T
- F L G * M T T L S G Y E Q T A * R K Q R
3061 - GAAGTAGCTAAGCCACATCAAGCCTACAATACAAGCCATTGCAATCGCAATCCCGCCAGT -
3120
- E V A K P H Q A Y N T S H C N R N P A S
- K * L S H I K P T I Q A I A I A I P P V
- S S * A T S S L Q Y K P L Q S Q S R Q S
3121 - CACCCAATTAATCTGTAGACAACAGCAAGCACAAAACAAGCAAGTGTACTGGCCACAA -
3180
- H P I N S V D N S K H K T S K C Y W P Q
- T Q L I L * T T A S T K Q A S V T G H K
- P N * F C R Q Q Q A Q N K Q V L L A T R
3181 - GAGCCAGAGGAAAAACAAGCTTTATTTATGTACAAAAACCTGTTCCGATTAGAATAGGCCAA -
3240
- E P E E N K L Y Y V Q K P V P I R I G K
- S Q R K T S F I M Y K N L F R L E * A N
- A R G K Q A L L C T K T C S D * N R Q I
3241 - TTGTAGTAACATAATCCAGGCTAGGAATAGGAAACCTATTACTAGGTTCCATTGTTCCAG -
3300
- L * * H N P G * E * E T Y Y * V P L F Q
- C S N I I Q A R N R K P I T R F H C S R
- V V T * S R L G I G N L L L G S I V P G
3301 - GAGTGTGTTAAGCTCCTCAACGGTAATAGTACCGTTGTCTGCCATGATAAGCAATGTTAA -
3360
- E L F K L L N G N S T V V C H D K Q C *
- S C L S S S T V I V P L S A M I S N V K
- V V * A P Q R * * Y R C L P * * A M L K
3361 - AGTTCCAAACAGATAATAATAATAGTTAGTTCGTTAGACCAGAAGATCAGGAACCTCT -
3420
- S S K Q N N N N S * F V * T R R S G T P
- V P N R I I I I V S S F R P E D Q E L L
- F Q T E * * * * L V R L D Q K I R N S F
3421 - TCAGAAGAGTTCAGATTTTTAACACGCGAGTAGACGTAAACCGTTGGTTTTACTAAACTC -
3480
- S E E F R F L T R E * T * T V G F T K L
- Q K S S D F * H A S R R K P L V L L N S
- R R V Q I F N T R V D V N R W F Y * T H
3481 - ACGTTAACAAATATTGCAGCAGTACGCACACAATCGAAGCGCAGTAAGGATGGCTAGTGTG -
3540
- T L T I L Q Q Y A H N R S A V R M A S V
- R * Q Y C S S T H T I E A Q * G W L V *
- V N N I A A V R T Q S K R S K D G * C D
3541 - ACTAGCAAGAATACCACGAAGCAAGAAAAAGAAGTACGCTATTAATTAACGTACCT -
3600
- T S K N T T K A R K R S T L L T I N V P
- L A R I P R K Q E K E V R Y * L L T Y L
- * Q E Y H E S K K K K Y A I N Y * R T C
3601 - GTTTCCTCCGAAACGAATGAGTACATAAGTTCGTACTCACTTCTTGTGCTTACAAAGGC -
3660
- V S S E T N E Y I S S Y S L S C A Y K G
- F L P K R M S T * V R T H F L V L T K A
- F F R N E * V H K F V L T F L C L Q R H
3661 - ACGCTAGTAGTCGTCGGCTCATCAAAATTGGATCCATTGCTGGATTAGCAACTCCT -

```

图 17-E

3720

- T L V V V V G S S \* I G S I A G L A T P  
- R \* \* S S S A H H K L D P L L D \* Q L L  
- A S S R R R L I I N W I H C W I S N S \*



```

3721 - GAAGAGCCGTCGATTGTGTATTGACATTTCGGTGGGTCTTTAACAAGCTTGTAAAG -
3780
- E E P S I V C I C T F G G S L T S L L K
- K S R R L C V F A H S V G L * Q A C * R
- R A V D C V Y L H I R W V F N K L V K D
3781 - ATGAAGAATGTAGCATTTCATACCAGTGTCTGTAGTAATTTGTGTAGACTCAAGCTGG -
3840
- M K N V A F S I P V S V V I C V D S S W
- * R M * H F Q Y Q C L * * F V * T Q A G
- E E C S I F N T S V C S N L C R L K L V
3841 - TAGTAAACTTCGGTGAAATAGCCATGTACAACGACATAGTCTTTAACACCTGAGTGCCTA -
3900
- * * T S V K * P C T T T * S L T P E C L
- S K L R * N S H V Q R H S L * H L S A Y
- V N F G E I A M Y N D I V F N T * V P I
3901 - TCCTCAGAATAACCACCAATTTGGTAGTCTTCTTGAGTTTGGTGTGAAATGCCGTCA -
3960
- S S E * P P I W * S S L S F G V E M P S
- P Q N N H Q F G S L L * V L V L K C R H
- L R I T T N L V V F F E F W C * N A V T
3961 - CCTTCAGTAACGACAATTTGTATCTGTGACACTGTTATATGGTATACAGTAGTCATAGTTA -
4020
- P S V T T I V S V T L L Y G I Q * S * L
- L Q * R Q L Y L * H C Y M V Y S S H S Y
- F S N D N C I C D T V I W Y T V V I V M
4021 - TGTTGTGCCAGCAAACAAGTAGTTGGCATCATAAAGTAATGGGTCTTGGATTGCAC -
4080
- C V C Q Q T K * L A S * S N G F L D L H
- V C A S K Q S S W H H K V M G S W I C T
- C V P A N K V V G I I K * W V L G F A L
4081 - TTCCAACAAGCCAACATCTCATAATAATTTACATGCGTTGATGCATTGTAGAAAATAT -
4140
- F Q Q S Q H L I I I L H A L M H C R K Y
- S N K A N I S * * F Y M R * C I V E N I
- P T K P T S H N N S T C V D A L * K I Y
4141 - ATCAAGGCATAGAGGTACAAAATTCGCGCTCCTTACCTGCAGCGACAAGCAAAGATGT -
4200
- I K A * R Y K N C A S L P A A T S K R C
- S R H R G T K I A P P Y L Q R Q A K D V
- Q G I E V Q K L R L L T C S D K Q K M *
4201 - GAATAGATGGTAACAATAGCAGCAGTAAATGCAATGAACTGGAAGCCCTTATAAAG -
4260
- E * M V T N S S S K L Q M N W K P L * R
- N R W * Q I A A V N C K * T G S P Y K G
- I D G N K * Q Q * I A N E L E A L I K G
4261 - GCTAGCTGCCATCTTTATTGAGCGCAATTATTTGGTAGCGCTCTGAAAAACAGCAAGA -
4320
- A S C H L L L S A I I L V A L * K T A R
- L A A I F Y * A Q L F W * R S E K Q Q E
- * L P S F I E R N Y F G S A L K N S K K
4321 - AATGCAACGCCAATAACAAGCCATCCGAAAGGGAGTGAGGCTTGTAGCGGTATCGTTGCT -
4380
- N A T P I T S H P K G S E A C S G I V A
- M Q R Q * Q A I R K G V R L V A V S L L
- C N A N N K P S E R E * G L * R Y R C C
4381 - GTAGCATGAACAGTACTTGCAGGAGAAGCATTGTCAATTTTACTGGCTGTGCAGTAAT -

```

图 17-F

4440 - V A \* T V L A G E A L S I F T G C A V I  
- \* H E Q Y L Q E K H C Q F L L A V Q \* L  
- S M N S T C R R S I V N F Y W L C S N \*  
4441 - GATCCAAGAGTAAAAATCTCATAAACAAATCCATAAGITTCGTTTATGTGTATGTAATT -  
4500 - D P R V K N L I N K S I S S F M C N V I  
- I Q E \* K I S \* T N P \* V R L C V M \* F  
- S K S K K S H K Q I H K F V Y V \* C N L

```

4501 - TGACACCCTTGAGAACTGGCTCAGAGTCATCCTCATCAAACCTTGACGCAAGAACCACAAG -
4560
- * H P * E L A Q S H P H Q T C S K N H K
- D T L E N W L R V I L I K L A A R T T R
- T P L R T G S E S S S S N L Q Q E P Q E
4561 - AGCATGCACCCCTTGAGGCAACTGCAACAACTAGTCATGCAACAAAGCAAGATTGTAACCA -
4620
- S M H P * G N C N N * S C N K A R L * P
- A C T L E A T A T T S H A T K Q D C N H
- H A P L R Q L Q Q L V M Q Q S K I V T M
4621 - TGACGATGGCAATTAGTCCAGCAATGAAGCCGAGCCAAACATACCAAGGCCATTTAATAT -
4680
- * R W Q L V Q Q * S R A K H T K A I * Y
- D D G N * S S N B A E P N I P R P F N I
- T M A I S P A M K P S Q T Y Q G H L I Y
4681 - ATTGCTCATATTTTCCCAATTCTTGAAGGTCAATGAGTGATTCAATTTAAATTTTGTAGCGA -
4740
- I A H I F P I L E G Q * V I H L N F * R
- L L I F S Q F L K V N E * F I * I F S D
- C S Y F P N S * R S M S D S F K F L A T
4741 - CCTCATGAGGCGGTCAATTTCTTTTGAATGTTGACGACAGAAGCGTTAATGCCTGAAA -
4800
- P H * G G Q F L F E C * R Q K R * C L K
- L I E A V N F F L N V D D R S V N A * N
- S L R R S I S F * M L T T E A L M P E M
4801 - TGTCGCCAAGATCAACATCTGGTGATGTATGATTTTGAAGTACTGTCCAGCTCTTCTT -
4860
- C R Q D Q H L V M Y D F * S T C P A L L
- V A K I N I W * C M I F E V L V Q L F F
- S P R S T S G D V * F L K Y L S S S S L
4861 - TGAATGAGTCAAGCTCAGGTTGCAGAGGATCATAAACTGTGTTGTTAATGATGCCAATAA -
4920
- * M S Q A Q V A E D H K L C C * * C Q *
- E * V K L R L Q R I I N C V V N D A N N
- N E S S S G C R G S * T V L L M M P I T
4921 - CGACATCACAATTCCTGAGACAAATGTATGTCTGTAGTAATTATTTGTGGAGAAAGA -
4980
- R H H N F L R Q M Y C L * * L F V E K R
- D I T I S * D K C I V C S N Y L W R K E
- T S Q F P E T N V L S V V I I C G E K K
4981 - AGTTCCTCTGTGTAATAAACAAGAAGTGCCATTAACACAAAACACCTTCACGAGGGA -
5040
- S S S V * * T K K C H * T Q K H L H E G
- V P L C N K P R S A I K H K N T F T R E
- F L C V I N Q E V P L N T K T P S R G K
5041 - AGTATGCTTTGCCTTCATGACAAATGCTGGCGCTGTGGTGAAGTTCCTCTCCTGGGATG -
5100
- S M L C L H D K L L A L W * S S S P G M
- V C F A F M T N C W R C G E V P L L G W
- Y A L P S * Q I A G A V V K F L S W D G
5101 - GCACATACGTGACATGTAGGAAGACAACACCATGCGGGGCTGCTTGTGGGAGGACATAA -
5160
- A H T * H V G R Q H H A G L L V G R T *
- H I R D M * E D N T M R G C L W E G H K
- T Y V T C R K T T P C G A A C G K D I R
5161 - GGTGGTAGCCCTTCCACAAAAGTCAACTCTTTTGTATTGTCCAAGAACACACTCAGACA -

```

图 17-G

5220

- G G S P F H K S Q L F L I V Q E H T Q T  
- V V A L S T K V N S F \* L S K N T L R H  
- W \* P F P Q K S T L F D C P R T H S D I

```

5221 - TTTTAGTAGCAGCAAGATTAGCAGAAGCCCTGATTTTCAGCAGCCCTGATTAGTTGTTGTG -
5280
- F * * Q Q D * Q K P * F Q Q P * L V V V
- F S S S K I S R S P D F S S P D * L L C
- L V A A R L A B A L I S A A L I S C C V
5281 - TTACATAGGTTTGAAGGCTTTGAAGTCTGCCTGTAATTAACCTGTCAATTTGTACCTCG -
5340
- L H R F E G F E V C L * L T C Q F V P P
- Y I G L K A L K S A C N * P V N L Y L R
- T * V * R L * S L P V I N L S I C T S A
5341 - CCTCGACTTTATCAAGTCGCGAAAGGATATCATTTAGCACACTTGAATTCACCAAAAT -
5400
- P R L Y Q V A K G Y H L A H L K L H Q N
- L D F I K S R K D I I * H T * N C T K I
- S T L S S R E R I S F S T L E I A P K L
5401 - TAGAGCTAAGTTGTTTAAACAGTGTGTTTAAATGCTTGAGCATTCTGGTTAACACGCTCT -
5460
- * S * V V * Q V C L M L E H S G * Q R L
- R A K L F N K C V * C L S I L V N N V L
- E L S C L T S V F N A * A F W L T T S C
5461 - GCAGCTTGCCCAATGCAGTTGATGTTGTTGTAAGTGATTCTTGAATTTGACTAATCGCCT -
5520
- A A C P M Q L M L L * V I L E F D * S P
- Q L A Q C S * C C C K * F L N L T N R L
- S L P N A V D V V S D S * I * L I A L
5521 - TGTTAAATGTTGGCGATTTGTTTTGGTTCTCATAGAGAACATTTTGGGTAACCCAA -
5580
- C * I G W R F V F G S H R E H F G * L Q
- V K L V G D L F L V L I E N I L G N S N
- L N W L A I C F W F S * R T F W V T P M
5581 - TGCCATTGAACCTATATGCCATTTGCATAGCAAAGGTATTTGAAGAGCAGCGCCAGCAC -
5640
- C H * T Y M P F A * Q K V F E E Q R Q H
- A I E P I C H L H S K R Y L K S S A S T
- P L N L Y A I C I A K G I * R A A P A P
5641 - CAAATGTCCATCCAGCAGTGGCAGTACCCTAAGTACAGCAGCAGTGTAGGCAGCAATCA -
5700
- Q M S I Q Q W Q Y H * L E Q Q C R Q Q S
- K C P S S S G S T T N * S S S V G S N H
- N V H P A V A V P L T R A A V * A A I I
5701 - TATCATCAGTGAGCAGAGGTGGCAACACTGTAAGTCCATTGAACTTCTGCGCACAAATGA -
5760
- Y H Q * A E V A T L * V H * T S A H K *
- I I S E Q R W Q H C K S I E L L R T N E
- S S V S R G G N T V S P L N F C A Q M R
5761 - GATCTTAGCATTAAATATCACCTAGGCATTGCGCATATTGCTTCATGAAGCCAGCATCAG -
5820
- D L * H * Y H L G I R H I A S * S Q H Q
- I S S I N I T * A F A I L L H E A S I S
- S L A L I S P R H S P Y C F M K P A S A
5821 - CGAGTGTACCTTATTAAGAGCAAGTCCTCAATAAAGACCTTTAGTTGGCTTTAGAG -
5880
- R V S P Y * R A S P Q * K T S * L A L E
- E C H L I K E Q V L N K R P L S W L * R
- S V T L L K S K S S I K D L L V G F R G
5881 - GGTCAAGTAATTTGTGAAAAATTAACCACCAAAATATTTCAAAGTTGGGGTTTGT -

```

图 17-H

5940  
- G Q V I F V K N \* N H Q N I S K L G F C  
- V R \* Y L \* K I K T T K I F Q S W G F V  
- S G N I C E K L K P P K Y F K V G V L Y  
5941 - ACATTTGTTTGACTTGAGCGAACACTTCACGTGTGTTGCGATCCTGTTTCAGCAGCAATAC -  
6000  
- T F V \* L E R T L H V C C D P V Q Q Q Y  
- H L F D L S E H F T C V A I L F S S N T  
- I C L T \* A N T S R V L R S C S A A I P

```

6001 - CTGAGAGTGCACGATTTAGTTGTGTGCAAAGCTACCATATTGGAGAAGCAAATTAGCAC -
6060
- L R V H D L V V C K S Y H I G E A N * H
- * E C T I * L C A K A T I L E K Q I S T
- E S A R F S C V Q K L P Y W R S K L A H
6061 - ATTCAGTAGAATCTCCGCAGATGTACATATTACAATCTACGGAGGTTTTAGCCATAGAAA -
6120
- I Q * N L R R C T Y Y N L R R F * P * K
- F S R I S A D V H I T I Y G G F S H R N
- S V E S P Q M Y I L Q S T E V L A I E T
6121 - CAGGCATTACTTCTGTAGTAATGCTAATTGAAAAGTTAGTAGGTATAGCAATGGTGTAT -
6180
- Q A L L L * * C * L K S * * V * Q W C Y
- R H Y F C S N A N * K V S R Y S N G V I
- G I T S V V M L I E K L V G I A M V L L
6181 - TAGAGTAAGCAATTGAACTATCAGCACCTAAAGACATAGTATAAGCCACAATAGATTTTT -
6240
- * S K Q L N Y Q H L K T * Y K P Q * I F
- R V S N * T I S T * R H S I S H N R F L
- E * A I E L S A P K D I V * A T I D F W
6241 - GGCTAGTACTACGTAATAAAGAACTGTATGGTAACTAGCACAAATGCCAGCTCCAATAG -
6300
- G * Y Y V I K K L Y G N * H K C Q L Q *
- A S T T * * R N C M V T S T N A S S N R
- L V L R N K E T V W * L A Q M P A P I G
6301 - GAATGTCGCACTCATAAGAAGTGTGACATGCTCAGCTCCTATAAGACAGCCTGCTTGAG -
6360
- E C R T H K K C R H A Q L L * D S L L E
- N V A L I R S V D M L S S Y K T A C L S
- M S H S * E V S T C S A P I R Q P A * V
6361 - TCTGGAATACATTGTTTCCAGTAGAATATATGCGCCAAGCTGGTGTGAGTTGATCTGCAT -
6420
- S G I H C F Q * N I C A K L V * V D L H
- L E Y I V S S R I Y A P S W C E L I C M
- W N T L F P V E Y M R Q A G V S * S A *
6421 - GAATTGCTGTAGAAACATCAGTGCAGTTAACATCTTGATATAGAACAGCAACTTCAGATG -
6480
- E L L * K H Q C S * H L D I E Q Q L Q M
- N C A R N I S A V N I L I * N S N F R *
- I A V E T S V Q L T S * Y R T A T S D E
6481 - AAGCATTGTTCCAGGTGTAATTACACTTACACCCCAAAGAGCAAGGTGAAATGTCTA -
6540
- K H L F Q V * L H L H P Q K S K V K C L
- S I C S R C N Y T Y T P K R A R * N V *
- A F V P G V I T L T P P K E Q G E M S N
6541 - ATATTTAGATGTTTTAGGATCTCGAACGGAATCAGTCAAATCAGAAACATCACGGCCAA -
6600
- I F Q M F * D L E R N Q * N Q K H H G Q
- Y F R C F R I S N G I S E I R N I T A K
- I S D V L G S R T E S V K S E T S R P N
6601 - ATTGTTGAAATGGTTGAAATCTCTTTGAGAAGGAGTTAACACACCAGTACCAGTGAGTC -
6660
- I V E M V E I S L K K E L T H Q Y Q * V
- L L K W L K S L * R R S * H T S T S E S
- C * N G * N L F E E G V N T P V P V S P
6661 - CATTAAATTAATAATGACACACTGGTTCTTAATAAGGTCAGTGGATAATTTGGTCCAC -

```

图 17-1

6720

- H \* N \* N \* H T G S \* \* G Q W I I L V H  
- I K I K I D T L V L N K V S G \* F W S T  
- L K L K L T H W F L I R S V D N F G P Q



```

6721 - AAACCGTGGCCGGTGCATTTAAAAGTTCAAAGAAAGTACTACAACCTCTGTAAGGTTGGT -
6780
- K P W P V H L K V Q K K V L Q L C K V G
- N R G R C I * K F K R K Y Y N S V R L V
- T V A G A F K S S K E S T T T L * G W *
6781 - AGCCAATGCCAGTAGTGGTGTAAAAACCATAATCATTTAATGGCCAATRAACAATTAAGAG -
6840
- S Q C Q * W C K N H N H L M A N N N * E
- A N A S S G V K T I I I * W P I T I K S
- P M P V V V * K P * S F N G Q * Q L R A
6841 - CAGGTGGGGTGCAAGGTTTGCCATCAGGGGAGAAAGGCACATTAGATATGTCTCTCTCAA -
6900
- Q V G C K V C H Q G R K A H * I C L S Q
- R W G A R F A I R G E R H I R Y V S L K
- G G V Q G L P S G E K G T L D M S L S K
6901 - AGGGCCTAAGCTTGCCATGTCTAAGATACCTATATTTATAATTATAATTACCAGTTGAAG -
6960
- R A * A C H V * D T Y I Y N Y N Y Q L K
- G P K L A M S K I P I F I I I T S * S
- G L S L P C L R Y L Y L * L * L P V E V
6961 - TAGCATCAATGTTCTAGTATTCCAAGCAAGGACACAACCCATGAAATCATCTGGCAATT -
7020
- * H Q C S * Y S K Q G H N P * N H L A I
- S I N V P S I P S K D T T H E I I W Q F
- A S M F L V F Q A R T Q P M K S S G N L
7021 - TATAATTATAATCAGCAATAACACCAGTTTGTCTGGCGCTATTTGTCTTACATCATCTC -
7080
- Y N Y N Q Q * H Q F V L A L F V L H H L
- I I I I S N N T S L S W R Y L S Y I I S
- * L * S A I T P V C P G A I C L T S S P
7081 - CCTTGACTACAAAAGAAATCTGCATAGACATTGGAGAAGCAAAGATCATTCAACTTAGTGG -
7140
- P * L Q K N L H R H W R S K D H S T * W
- L D Y K R I C I D I G E A K I I Q L S G
- L T T K E S A * T L E K Q R S F N L V A
7141 - CAGAAACGCCATAGCACTTAAAGGTTGAAAAAATGTTGAGTTGTAGAGCACAGAGTAAT -
7200
- Q K R H S T * R L K K M L S C R A Q S N
- R N A I A L K G * K K C * V V E H R V I
- E T P * H L K V E K N V E L * S T E * S
7201 - CAGCAACACAATTAGAAATTTTTTTCTCTCCCATGCATAGACAGAAGGGAATTTAGTAG -
7260
- Q Q H N * K F F F S P M H R Q K G I * *
- S N T I R N F F S L P C I D R R E F S S
- A T Q L E I F F L S H A * T E G N L V A
7261 - CATTAAAAACCTCTCCAAAAGGACACAAGTTTGTAAATATTAGGGAATCTCACAACATCTC -
7320
- H * K P L Q K D T S L * Y * G I S Q H L
- I K N L S K R T Q V C N I R E S H N I S
- L K T S P K G H K F V I L G N L T T S P
7321 - CTGAGGGAACAACCCCTGAAATTAGAGGCTGGTAAATTCCTTTGTCAATCTCAAAGCTCT -
7380
- L R E Q P * N * R S G K F L C Q S Q S S
- * G N N P E I R G L V N S F V N L K A L
- E G T T L K L E V W * I P L S I S K L L
7381 - TAACAGAGCATTGAGTTCAGCAAGTGGATTTGAGACAATCAACAGCATCTGTGATTG -
7440
- * Q S I * V Q Q V D F E N N Q Q H L * L

```

图 17-J

- N R A F E F S K W I L R T I N S I C D C  
- T E H L S S A S G F \* E Q S T A S V I V  
7441 - T A C C A T T T T C A T C A T A C T T G A G C A T A A T G T A G T T G G C T T T A A A T A G C C A A C A A A T A G G -  
7500  
- Y H F H H T \* A \* M \* L A L N S Q Q N R  
- T I F I I L E H K C S W L \* I A N K I G  
- P F S S Y L S I N V V G F K \* P T K \* A

```

7501 - CTGCAGCTGACGTGCCCCAAATGTCTTGAGCAGGTGAAAAGGCTGTAAGAATGGCTCTAA -
7560
- L Q L T C P K C L E Q V K R L * E W L *
- C S * R A P N V L S R * K G C K N G S K
- A A D V P Q M S * A G E K A V R M A L K
7561 - AATTTGTAATGTTAATACCAAGAGGCCAACTTAAAAATAGGTTTCAAAGTGTTAAAAACCAG -
7620
- N L * C * Y Q E A T * K * V S K C * N Q
- I C N V N T K R Q L K N R F Q S V K T R
- F V M L I P R G N L K I G F K V L K P E
7621 - AAGGTAGATCACGAACTACATCTATAGGTTGATAGCCCTTATAACATAGAGAAACCCAT -
7680
- K V D H E L H L * V D S P Y K H R E T H
- R * I T N Y I Y R L I A L I N I E K P I
- G R S R T T S I G * * P L * T * R N P S
7681 - CTTTATTTTTAAACACAAACTCTCGTAAGTGTTTAAAAATTACCTGACTTTTCTGAAACAT -
7740
- L Y F * T Q T L V S V * N Y L T F L K H
- F I F K H K L S * V F K I T * L F * N I
- L F L N T N S R K C L K L P D F S E T S
7741 - CAAGCGAAAAGGCATCAGATATGTACTCGAAAGTGCAATTAATGCATTATCGAATATCA -
7800
- Q A K R H Q I C T R K C N * M H Y R I S
- K R K G I R Y V L E S A I K C I I E Y H
- S E K A S D M Y S K V Q L N A L S N I I
7801 - TAGTATGTGTCTGTGTACCCATGGGTTTAGAAACAGCAAAGAAAGGGTTGTCACACAATT -
7860
- * Y V S V Y P W V * K Q Q R K G C H T I
- S M C L C T H G F R N S K E R V V T Q F
- V C V C V P M G L E T A K K G L S H N S
7861 - CAAAGTTACATGCTCGTATAACAACATTAGTAGAATTGTTAATAATAATCACCGACTGTG -
7920
- Q S Y M L V * Q H * * N C * * * S P T V
- K V T C S Y N N I S R I V N N N H R L *
- K L H A R I T T L V E L L I I I T D C D
7921 - ACTTGTGTTCATGGTAGAACCAAAAACCAACCACGGACAACATTTGATTTCTCTGTGG -
7980
- T C C S W * N Q K P N H G Q H L I S L W
- L V V H G R T K N P T T D N I * F L C G
- L L F M V E P K T Q P R T T F D F S V A
7981 - CAGCAAATAAATACCATCCTTAAAAGGTATGACAGGGTTGCCAAACGTATGATTAATAG -
8040
- Q Q N K Y H P * K V * Q G C Q T Y D * *
- S K I N T I L K R Y D R V A K R M I N S
- A K * I P S L K G M T G L P N V * L I V
8041 - TATGAAACCCTGTAACATTAGAATAAAATGGAAGAAATAAATCCTGAGTTAAATAAAGAG -
8100
- Y E T L * H * N K M E E I N P E L N K E
- M K P C N I R I K W K K * I L S * I K S
- * N P V T L E * N G R N K S * V K * R V
8101 - TGTCTGATCTAAAATTTTCATCAGGATAGTAACCCCTCATAGATGAAGTATGTTGAG -
8160
- C L I * K F H Q D S K P P S * M K Y V E
- V * S K N F I R I V N P P H R * S M L S
- S D L K I S S G * * T P L I D E V C * V
8161 - TGTAATTAGGAGCTTGAACATCATCAAAGTGGTGCACCGGTCAAGGTCACTACCACTAG -

```

图 17-K

8220

- C N \* E L E H H Q K W C T G Q G H Y H \*  
- V I R S L N I I K S G A P V K V T T S  
- \* L G A \* T S S K V V H R S R S L P L V

```

8221 - TGAGAGTAAGAAATAATAAGAAAATAAACATGTTTCGTTTAGTTGTTAACAAGAATATCAC -
8280
- * E * E I I R K * T C S F S C * Q E Y H
- E S K K * * E N K H V R L V V N K N I T
- R V R N N K K I N M F V * L L T R I S L
8281 - TTGAAACCACAACCTCTGTTGTTTCTCTAATGATAAGCCTACCTTTTTCCAGAAGAGAAT -
8340
- L K P Q L C C F L * * * A Y L F P E E N
- * N H N S V V F S N D K P T F F Q K R I
- E T T T L L F S L M I S L P F S R R E *
8341 - AAATCATATCATTGATTGATTCCTCTAAGAGACATTACAGCAGTTCCTCTTAATTAA -
8400
- K S Y H * F D S P * E T L Q Q F L L I *
- N H I I D L I L L K R H Y S S S S * F K
- I I S L I * F S L R D I T A V P L N L R
8401 - GAGGAAATTTGTCATGTCAAAGAGTGAATAGGAAGACAACCTGGATAGGATTTGTGTTC -
8460
- E E I C S C Q R V N R K T T G * D L C S
- R K F A H V K E * I G R Q L D R I C V P
- N L L M S K S E * E D N W I G F V F L
8461 - TCCAGAAAATGTAGTTAGCATGCATGGTATAGCCATCAATTTGTTCTTCGGCTTGCCAA -
8520
- S R K C S * H A W Y S H Q F V P S A C Q
- P E N V V S M H G I A I N L F L R L A K
- Q K M * L A C M V * P S I C S F G L P R
8521 - GATAGTTAGCCCCAATTAATAAATGCTTCCGATGATGATGCATTTACATTTGTAACAAAAG -
8580
- D S * P Q L K M L P M M M H L H L * Q K
- I V S P N * K C F R * * C I Y I C N K S
- * L A P I K N A S D D A F T F V T K A
8581 - CTGTCCACCATGAGAAATGGCCCATAGCTTGTAAAGGTCAGCATTCCAAGAATGCTCTG -
8640
- L S T M R N G P * A C K G Q H S K N A L
- C P P * E M A H K L V K V S I P R M L C
- V H H E K W P I S L * R S A F Q E C S V
8641 - TTATCTTTACAGCTATAGAACCACCCAGGGCTAGTTTTGCTTTATAAATCCACACAGAT -
8700
- L S L Q L * N H P G L V F A L * I H T D
- Y L Y S Y R T T O G * F L L Y K S T Q I
- I F T A I E P P R A S F C F I N P H R *
8701 - AAGTGAAAAACCCTTCTTTAGAGTCATTCTCTTTTGTACATGTTTGGTCCTAGGGTCAT -
8760
- K * K T L L * S H S L L S H V W S * G H
- S E K P F F R V I L F C H M F G P R V I
- V K N P S L E S F S F V T C L V L G S Y
8761 - ACATATCGCTAATAATAAGGTCCCATTTATTAGCCGTATGTACTGTTGCACAGTCTCCAA -
8820
- T Y R * * * G P I Y * P Y V L L H S L Q
- H I A N N K V P F I S R M Y C C T V S N
- I S L I I R S H L L A V C T V A Q S P I
8821 - TTAAAGTAGAATCTGCGTCCGAGACGAAGTCATTAAGATCTGAATCGACAAGTAGTGTGC -
8880
- L K * N L R R R R S H * D L N R Q V V C
- * S R I C V G D E V I K I * I D K * C A
- K V E S A S E T K S L R S E S T S S V P
8881 - CAGTTGGCAACCATTTGTCCTGAGCACAGCTGTACCTGGTGCAACTCCTTTATCAGAGCCAG -
8940

```

图 17-L

---

- Q L A T I V \* A Q L Y L V Q L L Y Q S Q  
- S W Q P L S E H S C T W C N S F I R A S  
- V G N H C L S T A V P G A T P L S E P A  
8941 - CACCAAAGTGAATAACTCTCATGTTGTAGGGTACAGCTAAAGTAAGTGTATTTAAGTATT -  
9000  
- H Q S E \* L S C C R V Q L K \* V Y L S I  
- T K V N N S H V V G Y S \* S K C I \* V L  
- P K \* I T L M L \* G T A K V S V F K Y \*

9001 - GACACAGTTGAGTATACTTTGCGACATTCATCATTATTCCTTTTGGTATAACAGCATT -  
9060  
- D T V E Y T L R H S S L F L L V \* Q H F  
- T Q L S I L C D I H H Y S F W Y N S I F  
- H S \* V Y F A T F I I I P F G I T A F S  
9061 - CACCATAATTCTGAAGGTCACACTTTTCAAGAAGCATTCTTTCATCTTGACAAAGTTAG -  
9120  
- H H N S E G H T F Q E A F F A S C T S \*  
- T I I L K V T L F K K H S L H L V Q V R  
- P \* F \* R S H F S R S I L C I L Y K L G  
9121 - GCATCGCAACACCTGGTTGCCACGCTTGACTTGCTTGTAGTTTTGGGTAGAAGGTTTCAA -  
9180  
- A S Q H L V A T L D L L V V L G R R F Q  
- H R N T W L P R L T C L \* F W V E G F N  
- I A T P G C H A \* L A C S F G \* K V S T  
9181 - CATGTCCATCCTTACACCAAAGCATGAATGAAATTTTCAGCATAGTCAATTGTAACCTTGA -  
9240  
- H V H P Y T K A \* M K F Q H S Q L \* P \*  
- M S I L T P K H E \* N F S I V N C N L D  
- C P S L H Q S M N E I S A \* S I V T L T  
9241 - CCACCTTTGAAATCACTGACAAATCTTGTGACTTATTATCTCGACAAAGTCATCAAGTA -  
9300  
- P L L K S L T N L V T L L S R Q S H Q V  
- H F \* N H \* Q I L \* L Y Y L D K V I K \*  
- T F E I T D K S C D F I I S T K S S K  
9301 - AAAGATCAATCACAGAACACACACTTTTGATGAACCTGTTTGCATCTGTTATGAAGT -  
9360  
- K D Q S Q N T H I L M N L F A H L L \* S  
- R I N H R T H T F \* \* T C L R I C Y E V  
- R S I T E H T H F D E P V C A S V M K \*  
9361 - AATTTTCACTGTGCTGTCCATAGGGATAAAATCCTCTAATTTAAGTGGTGAATCTTGTG -  
9420  
- N F S L C C P \* G \* N P L I \* V V N L V  
- I F H C A V H R D K I L \* F K W \* I L \*  
- F F T V L S I G I K S S N L S G E S C E  
9421 - AGCGCTTGGCTAAGCCTATCATTAATGAAGACCGCCAAGTTGTCCATGACTGAAATCTC -  
9480  
- S A W L S L S L N E D R Q V V H D \* N L  
- A L G \* A Y H \* M K T A K L S M T E I S  
- R L A K P I I K \* R P P S C P \* L K S P  
9481 - CATAAACGATGTGTTTCAAGGCATAGCCCTCGAGCTTATATCGCTGTATGAATTCATCCA -  
9540  
- H K R C V R R H S P R A Y I A V \* I H P  
- I N D V F E G I A L E L I S L Y E F I H  
- \* T M C S K A \* P S S L Y R C M N S S I  
9541 - TAGCGAGCTCGAGAAAGTCAGTTTCCATTTGTGATCTGGGCTTAAATCCTCTAAGTCTC -  
9600  
- \* R A R E S Q F P F V I W A \* N P L S L  
- S E L E K V S F H L \* S G L K I L \* V S  
- A S S R K S V S I C D L G L K S S K S L  
9601 - TGCTCTGAGTAAAGTAGGTTTCAGGCAACTGTTGAATAATGCCGTCTACTTTCCTAAAGT -  
9660  
- C S E \* S R F Q A T V E \* C R L L S \* S  
- A L S K V G F R Q L L N N A V Y F L K V  
- L \* V K \* V S G N C \* I M P S T F L K \*  
9661 - AGTTAACTGTGTTTTACTGATTTCCCAATTAATGTGACTCCATTGACGCTAGCTTGTG -  
9720

图 17-M

---

- S \* T V F L L I L Q L M \* L H \* R \* L V  
- V K L C F Y \* F S N \* C D S I D A S L C  
- L N C V F T D S P I N V T P L T L A C A  
9721 - CTGGTCCTTTGAAGGTGTTAGACCTTTGACTGAACTTCTGTTATTAACACCATTAC -  
9780  
- L V P L K V L D L \* L N L L L L K H H Y  
- W S L \* R C \* T F D \* T F C Y \* N T I T  
- G P F E G V R P L T E P S V I K T P L R



```

9781 - GGGCGTTTCTAAAAAGGTCTACCTGTCCTTCCACTCTACCATCAAACAAGACAGTAAGTG -
9840
- G R F * K G L P V L P L Y H Q T R Q * V
- G V S K K V Y L S F H S T I K Q D S K *
- A F L K R S T C P S T L P S N K T V S E
9841 - AAGAACAAGCACTCTCAGTAGGTTTCTTGGCAATGTCAGTCATTGTGCAGACACCTATTG -
9900
- K N K H S Q * V S W Q C Q S L C R H L L
- R T S T L S R F L G N V S H C A D T Y C
- E Q A L S V G F L A M S V I V Q T P I V
9901 - TAGATACATGTGCTGGGGCTTCTCTTTTGTAGTCCCAGATTACAGTATTAGCAGCGATAT -
9960
- * I H V L G L L F C S P R L Q Y * Q R Y
- R Y M C W G F S F V V P D Y S I S S D I
- D T C A G A S L L * S Q I T V L A A I S
9961 - CAACACCCAAATTATTGAGTATCTTAATCTCTGGCACTGGTTAATGTTACGCTTAGCCC -
10020
- Q H P N Y * V S * S L A L V * C Y A * P
- N T Q I I E Y L N L W H W F N V T L S P
- T P K L L S I L I S G T G L M L R L A Q
10021 - AAAGCTCAAATGCAACATTAACAGGAAGTGTGCTTATTTTCAAAGATCTCCACATCAA -
10080
- K A Q M Q H * Q E V L S Y F Q R S P H Q
- K L K C N I N R K C C L I F K D L H I N
- S S N A T L T G S V V L F S K I S T S I
10081 - TACCATCTACCTTTGTGTAACAGCATTATTAATGATGGAAACAGGTGCTTCGCCGGCGT -
10140
- Y H L P L C K Q H Y * * W K Q V L R R R
- T I Y L C V N S I I N D G N R C F A G V
- P S T F V * T A L L M M E T G A S P A C
10141 - GTCCATCAAAGTGCCTTTATTACAACATTATAAGCCACATTTTCTAAACTCTGTAACC -
10200
- V H Q S V L Y * Q H Y K P H F L N S V T
- S I K V S F I N N I I S H I F * T L * P
- P S K C P L L T T L * A T F S K L C N L
10201 - TGGTAAATGTATTCCACAGGTTATAAGTATCAAATGTTTGTAAATCCATAGGCTAATC -
10260
- W * M Y S T G Y K Y Q I V C K S I G * I
- G K C I P Q V I S I K L F V N P * A K S
- V N V F H R L * V S N C L * I H R L N P
10261 - CAGCAGAAATCATCATATTATATGCATCCAAGTACTGTCGGTACTCATTGTCATGGTGC -
10320
- Q Q K S S Y Y M H P S T V G T H L H G V
- S R N H H I I C I Q V L S V L I C M V S
- A E I I I L Y A S K Y C R Y S F A W C L
10321 - TGCAAACAGCACCACTAAATTGCATCGTGAATACAGTAGCAGATTGAGTGGACAT -
10380
- C K Q H H L N C I V * Y T * Q I * V E H
- A N S T T * I A S C N T R S R F E W N I
- Q T A P P K L H R V I H V A D L S G T *
10381 - AATCAATATCCGACACTACTTGTGGCCATGAGACTCACAAGGACTATCAGAATAGTAAA -
10440
- N Q Y P T L L V C H E T H K D Y Q N S K
- I N I R H Y L F A M R L T R T I R I V K
- S I S D T T C L P * D S Q G L S E * * K
10441 - AGAAAGGCAATTGCTTTAAATTAGTAAATGCACTTTATCGAAAGCTGGAGTGTGGAATG -
10500

```

图 17-N

---

```
- R K A I A L N * * M H F Y R K L E C G M
- E R Q L L * I S K C T F I E S W S V E C
- K G N C F K L V N A L L S K A G V W N A
10501 - CATGCTTATTCACATACAAACTACCACCATCACAGCCTGGTAAGTTCAAGTTGACAAGA -
10560
- H A Y S H T N Y H H H S L V S S S L T R
- M L I H I Q T T T I T A W * V Q V * Q D
- C L F T Y K L P P S Q P G K F K F D K T
```

```

10561 - CTCTTGTGTCAAACCTACACACAATTGCATTGGCTGGGTAACGATCAACGTTACAATTCC -
10620
- L L C Q T Y T Q L H W L G N D Q R Y N S
- S C V K P T H N C I G W V T I N V T I P
- L V S N L H T I A L A G * R S T L Q F Q
10621 - AAAACAAACAAACACCCATCAGTGAATTTATCGTGATGTGTAGCATAAGAATAGAAGAGTT -
10680
- K T N K H H Q * I Y R D V * H K N R R V
- K Q T N T I S E F I V M C S I R I E E F
- N K Q T P S V N L S * C V A * E * K S S
10681 - CCTCTATTTTGTAGCTTTGTCACTACATGGCTGAGCATCGTAGAACTTCCATTCTACTT -
10740
- P L F C K L C H Y M A E H R R T S I L L
- L Y F V S F V T T W L S I V E L P F Y F
- S I L * A L S L H G * A S * N F H S T S
10741 - CAGCCTGAGGCACACACTTGATAGCCTTTGGATTTCCAATGTCATGAAGAAGTGGAACT -
10800
- Q P E A H T * * P L D F Q C H E E L E T
- S L R H T L D S L W I S N V M K N W K L
- A * G T H L I A F G F P M S * R T G N L
10801 - TATCAGCAAGCAATGCAGACTTCACAACCATGTGTTGTACTTTTCTGCAAGCAGAATTAA -
10860
- Y Q Q A M Q T S Q P C V V L F C K Q N *
- I S K Q C R L H N H V L Y F S A S R I N
- S A S N A D F T T M C C T F L Q A E L T
10861 - CCCTCAGTTTCATCTCCTATAATAGGGTATTCAACAGACCAATCAACGCGCTTAACAAAGC -
10920
- P S V H L L * * G I Q Q T N Q R A * Q S
- P Q F I S Y N R V F N R P I N A L N K A
- L S S S P I I G Y S T D Q S T R L T K H
10921 - ACTCATGGACTGCTAAACATCTAGTCATGATAGCATCAACTAGCCACATGTGCATTTC -
10980
- T H G L L N I * S * * H H N * P H V H F
- L M D C * T S S H D S I T T S H M C I S
- S W T A K H L V M I A S Q L A T C A F P
10981 - CATGTACCTGGCAATGTTGGTCATGGTTACTCTGAAGGTTACCCGTAAGCCCCACTGCT -
11040
- H V P G N V G H G Y S E G Y P * S P T A
- M Y L A M L V M V T L K V T R K A P L L
- C T W Q C W S W L L * R L P V K P H C *
11041 - GAACATCAATCATAAATGGGTTATAGACATAGTCAAAACCCACAGAATGATTCCAGCAGG -
11100
- E H Q S * M G Y R H S Q N P Q N D S S R
- N I N H K W V I D I V K T H R M I P A G
- T S I I N G L * T * S K P T E * F Q Q A
11101 - CATAAGTATCTGATGAAGTAGAAAAGCAAGTTGCACGTTTGTACACAGACAACACGTTT -
11160
- H K Y L M K * K S K L H V C H T D N T F
- I S I * * S R K A S C T F V T Q T T R S
- * V S D E V E K Q V A R L S H R Q H V L
11161 - TTTCAAGTCCAATCTTGACAAAGTACTTCATTGATGTAAGCTCAAAGCCATGCGCCCAA -
11220
- F Q V Q S * Q S T S L M * A Q S H A P K
- F R S N L D K V L H * C K L K A M R P K
- S G P I L T K Y F I D V S S K P C A Q R
11221 - GGACGAACACGACTCTGTCTGACAATCCTTTTCAGTGTATCACTGAGCATTGTACTATCT -

```

图 17-0

11280

- G R T R L C L T I L S V Y H \* A F V L S  
- D E H D S V \* Q S F Q C I T E H L Y Y L  
- T N T T L S D N P F S V S L S I C T I L

```

11281 - TAATACGGCACTACATTCAGGGCAAGCCTTTATACATGAGTGGTATAAGATGTTTAACT -
11340
- * Y A L H S R A S L Y T * V V * D V * T
- N T H Y I P G Q A P I H E W Y K M F K L
- I R T T F Q G K P L Y M S G I R C L N W
11341 - GGTACCTGGTGGAGGTTTTGCATTAACCTCTGGTGAATTCGTGTTATTTTCAGTGTCAA -
11400
- G H L V E V L H * L W * I L C Y F Q C Q
- V T W W R F C I N S G E F C V I F S V N
- S P G G G F A L T L V N S V L F S V S T
11401 - CATAACCGTCCGTACAGCTACTAAGTTAACACCTGTAGAAAATCCTAGCTGGAGAGGTA -
11460
- H N Q S V Q L L S * H L * K I L A G E V
- I T S R Y S Y * V N T C R K S * L E R *
- * P V G T A T K L T P V E N P S W R G R
11461 - GGTAGTACCCACAGCATCTCTAGTTGCATGACAGCCCTCTACATCAAAGCCAATCCACG -
11520
- G * Y P Q H L * L H D S P L H Q S Q S T
- V S T H S I S S C M T A L Y I K A N P R
- L V P T A S L V A * Q P S T S K P I H A
11521 - CACGAACGTGACGAATAGCTTCTTCGGGGTGATAAACATATTAGGGTAACCATTGACTT -
11580
- H E R D E * L L R G * * T Y * G N H * L
- T N V T N S F F A G D K H I R V T I D L
- R T * R I A S S R V I N I L G * P L T W
11581 - GGTAAATTCATTTGAAACCCATCATAGAGATGAGTCTACGGTAGGTCATGTCCTTTGGTA -
11640
- G N S F * N P S * R * V Y G R S C P L V
- V I H F E T H H R D E S T V G H V L W Y
- * F I L K P I I E M S L R * V M S F G M
11641 - TGCCTGGTATGTCAACACATAATCCTTCAGTCTTGAATTTTATATCAACGCTGAGGTGTG -
11700
- C L V C Q H I I L Q S * I L Y Q R * G V
- A W Y V N T * S F S L E F Y I N A E V C
- P G M S T H N P S V L N F I S T L R C V
11701 - TAGGTGCCTGTGTAGGATGAAGACCAGTAATGATCTTACTACAGTCTTAAAGTCCAG -
11760
- * V P V * D E D Q * * S Y Y S P * K V Q
- R C L C R M K T S N D L T T V L K K S S
- G A C V G * R P V M I L L Q S L K S P V
11761 - TTACATTTCTGCTTGTAAATGTAGCCACATTGCGACGTGGTATTTCTAGACTTGTAAT -
11820
- L H F L L V M * P H C D V V F L D L * I
- Y I F C L * C S H I A T W Y F * T C K L
- T F S A C N V A T L R R G I S R L V N C
11821 - GCAGTTTGTATAAAGATCTCTATCAGACATATATGCACAAAATGCCAATTTTGCCTTG -
11880
- A V C H K D L Y Q T L C T K C Q F L P L
- Q F V I K I S I R H Y A Q N A N F C P C
- S L S * R S L S D I M H K M P I F A L V
11881 - TGATAGCCACATTGAAGCGGTTGACATTACAAGAGTGTGCTGTTTCAGTAGTTTGTGTGA -
11940
- * * P H * S G * H Y K S V L F O * F V *
- D S H I E A V D I T R V C C F S S L C E
- I A T L K R L T L Q E C A V S V V C V N
11941 - ATATGACATAGTCATATTCAGAACCCTGTGATGAATCAACAGTCTGCGTAGGCAATCCTA -

```

图 17-P

12000

- I \* H S H I Q N P V M N Q Q S A \* A I L  
- Y D I V I F R T L \* \* I N S L R R Q S \*  
- M T \* S Y S E P C D E S T V C V G N P K

```

12001 - AGATTTTTGAAGCTACAGCGTTCTGTGAATTATAAGGTGAGATAAAAAACAGCTTTTCTCC -
12060
- R F L K L Q R S V N Y K V R * K Q L F S
- D F * S Y S V L * I I R * D K N S F S P
- I F E A T A F C E L * G E I K T A F L Q
12061 - AAGCAGGATTGCGTGTAGAAATTCTCTTACAACGCCTATTTGAGGTCTGTTGATTGCAG -
12120
- K Q D C V * E I L L Q R L F E V C * L Q
- S R I A C K K F S Y N A Y L R S V D C R
- A G L R V R N S L T P I * G L L I A D
12121 - ATGAAACATCATGTGTAATAACACCTTTGTAGAACATTTTGAAGCATTGAGCTGACTTAT -
12180
- M K H H V * * H L C R T F * S I E L T Y
- * N I M C N N T F V E H F E A L S * L I
- E T S C V I T P L * N I L K H * A D L S
12181 - CCTTGTGTGCTTTTAGCTTATGTGCATAAACTAAAGCACTCACAGTGTCAACAATTTGAG -
12240
- P C V L L A Y C H K L K H S Q C Q Q F Q
- L V C F * L I V I N * S T H S V N N F S
- L C A F S L L S * T K A L T V S T I S A
12241 - CAGGACAACGGCGACAAGTTCCAAGGAACATGTCTGGACCTATGTTTTTCATAAGTCTGC -
12300
- Q D N G D K F Q G T C L D L L F S * V C
- R T T A T S S K E H V W T Y C F H K S A
- G Q R R Q V P R N M S G P I V F I S L H
12301 - ACACTGAATTAAAATATTCTGGTTCTAGTGTGCCTTTAGTCAGCAATGTGCGGGGGCTG -
12360
- T L N * N I L V L V C L * S A M C G G L
- H * I K I F W F * C A F S Q Q C A G G W
- T E L K Y S G S S V P L V S N V R G A G
12361 - GTAATTGAGCAGGATCGCCAATATAGACGTAGTGTTCACGAAAGTCTAGCATGACAA -
12420
- V I E Q D R Q Y R R S V L H E V * H * Q
- * L S R I A N I D V V F C T K S S I D N
- N * A G S P I * T * C F A R S L A L T T
12421 - CACTCAAGTCATAATTAGTAGCCATAGAGATTTTCATCAAAGACTACAATGTCAGCAGTTG -
12480
- H S S H N * * P * R F H Q R L Q C Q Q L
- T Q V I I S S H R D F I K D Y N V S S C
- L K S * L V A I E I S S K T T M S A V V
12481 - TTTCTGGCAATGCATTTACAGTGCAGAAAACATACTGTTCTAGTGTGAATTCACCTTTGA -
12540
- F L A M H L Q C R K H T V L V L N S L *
- F W Q C I Y S A E N I L F * C * I H F E
- S G N A F T V Q K T Y C S S V E F T L N
12541 - ATTTATCAAAACACTCTACGGCGCAGCGCAGGTATGATTCTACTACATTTATCTATGG -
12600
- I Y Q N T L R A H A Q V * F Y Y I Y L W
- F I K T L Y A R T R R Y D S T T F I Y G
- L S K H S T R A R A G M I L L H L S M G
12601 - GCAAATATTTAATGCCTTTTCACATAGGGCATCAACAGCTGCATGAGAGCATGCCGTAT -
12660
- A N I L M P F H I G H Q Q L H E S M P Y
- Q I F * C L F T * G I N S C M R A C R I
- K Y F N A F S H R A S T A A * E H A V Y
12661 - ACACATGCGAGCAGATGGGTAATAGAGAGCAAGTCCGATGGCAAATGACTCTTACCAG -

```

图 17-0

12720

- T L C E Q M G N R E . Q V R W Q N D S Y Q  
- H Y A S R W V I E S K S D G K M T L T S  
- T M R A D G \* \* R A S P M A K \* L L P V



```

12721 - TACCAGGTGGTCCTTGGAGTGTAGACTCTTTTGCATGCCGACCTTTTGATAAATTGCAA -
12780
- Y Q V V L G V * S T F A C R P F D N L Q
- T R W S L E C R V L L H A D L L I I C N
- P G G P W S V E Y F C M P T F * * F A T
12781 - CATTGCTAGAAAACATCATCTGAGATGTTGAGTGTGGGTACAAGCCAGTAATTCACAT -
12840
- H C * K T H L R C * V L G T S Q * F S H
- I A R K L I * D V E C W V Q A S N S H I
- L L E N S S E M L S V G Y K P V I L T *
12841 - AGTGCTCTTGTGGCACTAGAGTAGGTGCACTAAGTGGCATTACAGTGTGAGATGTCAACA -
12900
- S A L V A L E * V H * V A L Q C E M S T
- V L L W H * S R C T K W H Y S V R C Q H
- C S C G T R V G A L S G I T V * D V N T
12901 - CAAAGTAATCACCAACATTCAACTTGTATGTCGTAGTACCTCTGTACACAACAGCATCAC -
12960
- Q S N H Q H S T C M S * Y L C T Q Q H H
- K V I T N I Q L V C R S T S V H N S I T
- K * S P T F N L Y V V V P L Y T T A S P
12961 - CATAGTCACCTTTTCAAAGGTGACTCTCCAATCTGTACTTTACTATTTTGTAGTTACAC -
13020
- H S H L F Q R C T L Q S V L Y Y F * L H
- I V T F F K G V L S N L Y F T I F S Y T
- * S P F S K V Y S P I C T L L F L V T R
13021 - GGTAAACCAGTAAAGACATAGTTTCTGTTCAATGGTGGTCTAGGTTTTCCAACCTCCCATG -
13080
- G N Q * R H S F C S M V V * V F Q P P M
- V T S K D I V S V Q W W S R F S N L P *
- * P V K T * F L F N G G L G F P T S H E
13081 - AAAGATGCAATTCTCTGTACAGAGTACTTCGCGTACAGTGGCAATACCATATGACAGCT -
13140
- K D A I L C Q R V L R V Q W Q Y H M T A
- K M Q F S V R E Y F A Y S G N T I * Q L
- R C N S L S E S T S R T V A I P Y D S L
13141 - TAAATGTTTCTCAGTGGCTTTGAGCGTTTCTGCTGCGAAAAGCTTGAGTCTCTCAGTAC -
13200
- * M F P Q W L * A F L L R K A * V S Q Y
- K C F L S G F E R F C C E K L E S L S T
- N V S S V A L S V S A A K S L S L S V Q
13201 - AAGTGTGGCAAGTATGTAATCGCCAGCATTAGTCCAATCACATGTTGCTATCGCATTGA -
13260
- K C W Q V C N R Q H * S N H M L L S H *
- S V G K Y V I A S I S P I T C C Y R I E
- V L A S M * S P A L V Q S H V A I A L K
13261 - AGTCAGTGACATTGTCACTGCCTACACATGTGTTTTGTATAAACCAAAAACCTGACCAT -
13320
- S Q * H C H C L H M C F C I N Q K P D H
- V S D I V T A Y T C V F V * T K N L T I
- S V T L S L P T H V F L Y K P K T * P L
13321 - TAGCACATAATGGAAAACATGAGGGGCTTATGTGACTTGCAATAATAGCTCATACTC -
13380
- * H I M E N * W E A Y V T C N N S S Y L
- S T * W K T N G R L M * L A I I A H T S
- A H N G K L M G G L C D L Q * * L I P P
13381 - CTAGATACAGTTGTGTACATCAGTGACATCAACCTGGGGCATTGCAACATAGGGAT -

```

图 17-R

13440

- L D T V V S H Q \* H H N L G H C K H R D  
- \* I Q L C H I S D I T T W G I A N I G I  
- R Y S C V T S V T S Q P G A L Q T \* G L

```

13441 - TAACAGACAACACTAATTTGTGTGATGTTGAAATGACATGGTCATAGCAGCACTTGCAAC -
13500
- * Q T T L I C V M L K * H G H S S T C N
- N R Q H * F V * C * N D M V I A A L A T
- T D N T N L C D V E M T W S * Q H L Q H
13501 - ATAGGAATGGTCTCCTAATACAGGCACCGCAACGAAGTGAAGTCTGTGAATTGCACAATA -
13560
- I G M V S * Y R H R N E V K S V N C T I
- * E W S P N T G T A T K * S L * I A Q Y
- R N G L L I Q A P Q R S E V C E L H N T
13561 - CACAAGCACCTACAGCCTGCAAGACTGTATGTGGTGTGTACATAGCCTCATAAAAACCTCAG -
13620
- H K H L Q P A R L Y V V C T * P H K T Q
- T S T Y S L Q D C M W C V H S L I K L R
- Q A P T A C K T V C G V Y I A S * N S G
13621 - GTTCCCGTACCGTGAGGTGTTTATCATTAGTTAGCATTACGGAATACATGTCCAACATGT -
13680
- V P S T V R C Y H * L A L R N T C P T C
- F P V P * G V I I S * H Y G I H V Q H V
- S Q Y R E V L S L V S I T E Y M S N M W
13681 - GGCCAGTAAGCTCATCATGTAACCTTCTAATGTATTGTAATAACAAGTGAAGACATCAG -
13740
- G Q * A H H V T F * C I V N T S E R H Q
- A S K L I M * L S N V L * I Q V K D I S
- P V S S S C N F L M Y C K Y K * K T S A
13741 - CATACTCCTGATTAGGATGTTTTGTAAGTGGGTAAGCATCAATAGCCAGTGACACGAACC -
13800
- H T P D * D V L * V G K H Q * P V T R T
- I L L I R M F C K W V S I N S Q * H E P
- Y S * L G C F V S G * A S I A S D T N L
13801 - TTTCAATCATAAGTGTACCATCTGTTTTGACAATATCATCGACAAAACAGCCTGCGCCTA -
13860
- F Q S * V Y H L F * Q Y H R Q N S L R L
- F N H K C T I C F D N I I D K T A C A *
- S I I S V P S V L T I S S T K Q P A P N
13861 - ATATTCCTGATGGATCTGGGTAAGGCAGGTACACGTAATCATCTCCTTGTFTAACAGCA -
13920
- I F L M D L G K A G T R N H L L V * L A
- Y S * W I W V R Q V H V I I S L F N * H
- I L D G S G * G R Y T * S S P C L T S I
13921 - TTGTATGCTGTGAGCAAATTCGTGAGGTCTTTAGTAAGGTCAGTCTCAGTCCAACATT -
13980
- L Y A V S K I R E V L * * G Q S Q S N I
- C M L * A K F V R S F S K V S L S P T F
- V C C E Q N S * G P L V R S V S V Q H F
13981 - TTGCCTCAGACATGAACACATTATTTGATAATAAAGAACTGCCTTAAAGTCTTAAATGC -
14040
- L P Q T * T H Y F D N K E L P * S S * C
- C L R H E H I I L I I K N C L K V L N A
- A S D M N T L F * * * R T A L K F L M L
14041 - TAGCTACTAAACCTTGAGCCGCATAGTTACTGTTATAGCACACAACGGCATCATCAGAAA -
14100
- * L L N L E P H S Y C Y S T Q R H H Q K
- S Y * T L S R I V T V I A H N G I I R K
- A T K P * A A * L L L * H T T A S S E R
14101 - GAATCATCATGGAGAAATGTTTACGCAGGTAAGCGTAAAACCTCATCCACGAATTCATGAT -
14160

```

图 17-S

---

- E S S W R N V Y A G K R K T H P R I H D  
- N H H G E M F T Q V S V K L I H E F M I  
- I I M E K C L R R \* A \* N S S T N S \* S  
14161 - CAACATCCCTATTTCTATAGAGACACTCATAGAGCCTGTGTTGTAGATTGCGGACATACT -  
14220  
- Q H P Y F Y R D T H R A C V V D C G H T  
- N I P I S I E T L I E P V L \* I A D I L  
- T S L F L \* R H S \* S L C C R L R T Y L

```

14221 - TGTCAGCTATCTTATTACCATCAGTTGAAAGAAGTGCATTTACATTGGCTGTAACAGCTT -
14280
- C Q L S Y Y H Q L K E V H L H W L * Q L
- V S Y L I T I S * K K C I Y I G C N S L
- S A I L L P S V E R S A F T L A V T A *
14281 - GACAAATGTTAAAGACACTATTAGCATAAGCAGTTGTAGCATCACCGGATGATGTTCCAC -
14340
- D K C * R H Y * H K Q L * H H R M M F H
- T N V K D T I S I S S C S I T G * C S T
- Q M L K T L L A * A V V A S P D D V P P
14341 - CTGGTTTAAACATATAGTGAGCCGCCACACATGACCATCTCACTTAATACTTGGGCACACT -
14400
- L V * H I V S R H T * P S H L I L A H T
- W F N I * * A A T H D H L T * Y L R T L
- G L T Y S E P P H M T I S L N T C A H S
14401 - CGTTAGCTAACCTGTAGAAACGGTGTGATAAGTTACAGCAAGTGTATGTTTGGCAGCAA -
14460
- R * L T C R N G V I S Y S K C Y V C E Q
- V S * P V E T V * * V T A S V M F A S K
- L A N L * K R C D K L Q Q V L C L R A R
14461 - GAACAAGAGAGGCCATTATCCTAAGCATGTTAGGCATGGCTCTGTCACATTTTGGATAAT -
14520
- E Q E R P L S * A C * A W L C H I L D N
- N K R G H Y P K H V R H G S V T F W I I
- T R E A I I L S M L G M A L S H F G * S
14521 - CCCAACCCATAAGGTGTGGAGTTTCTACATCACTGTAAACAGTTTTTAAACATATTATGCC -
14580
- P N P * G V E F L H H C K Q F L T Y Y A
- P T H K V W S F Y I T V N S F * H I M P
- Q P I R C G V S T S L * T V F N I L C Q
14581 - AGCCACCGTAAAACCTTGCTTGTTCCAATTACCACAGTAGCTCCTCTAGTGGCGCTATTG -
14640
- S H R K T C L F Q L P Q * L L * W R L L
- A T V K L A C S N Y H S S S S S G G Y *
- L P * N L L V P I T T V A P L V A A I D
14641 - ACTTCAATAATTTCTGATGAAACTGTCTATTGTTCATAGTACTACAGATAGAGACACCAG -
14700
- T S I I S D E T V Y L S * Y V R * R H Q
- L Q * F L M K L S I C H S T T D R D T S
- F N N F * * N C L F V I V L Q I E T P A
14701 - CTACGGTCCGAGCTTATTCTTTGCACTAATGGCATACTTAAGATTCAATTTGAGTTATAG -
14760
- L R C E L Y S L H * W H T * D S F E L *
- Y G A S S I L C T N G I L K I H L S Y S
- T V R A L P F A L M A Y L R F I * V I V
14761 - TAGGGATGACATTACGCTTAGTATACGCGAAAAGTGCATCTTGATCCTCATAACTCATTG -
14820
- * G * H Y A * Y T R K V H L D P H N S L
- R D D I T L S I R E K C I L I L I T H *
- G M T L R L V Y A K S A S * S S * L I E
14821 - AGTCATAATAAAGTCTAGCCTTACCCATTATTAATGCGGAAACCAGCTGATTATCCA -
14880
- S H N K V * P Y P I Y * M G N Q L I Y P
- V I I K S S L T P F I K W E T S * F I Q
- S * * S L A L P H L L N G K P A D L S R
14881 - GATTGTTAACGATTACTTGGTTGGCATTAAATACAGCCACCATCGTAACAATCAAAGTATT -

```

图 17-T

14940 .  
- D C \* R L L G W H \* Y S H H R N N Q S I  
- I V N D Y L V G I N T A T I V T I K V F  
- L L T I T W L A L I Q P P S \* Q S K Y L

```

14941 - TATCAACAACCTTCAACTACGAATAGGAGTTGTCTGATATCACACATTGTTGGCAGATTAT -
15000 - Y Q Q L Q L R I G V V * Y H T L L A D Y
- I N N F N Y E * E L S D I T H C W Q I I
- S T T S T T N R S C L I S H I V G R L *
15001 - AACGATAATAGTCATAATCACTGATAGCAGCGTTGCCATCCTGAGCAAAGAAGAAGTGT -
15060 - N D N S H N H * * Q R C H P E Q R R S V
- T I I V I I T D S S V A I L S K E E V F
- R * * S * S L I A A L P S * A K K K C F
15061 - TTAGTTCAACAGAAC'TCCTTCTTAAGAAACCTTTAGACACAGCAAAGTCATAAAAGT -
15120 - L V Q Q N F L P * R N L * T Q Q S H K S
- * F N R T S F L K E T F R H S K V I K V
- S S T E L P S L K K P L D T A K S * K S
15121 - CTTTATTAATAATACCGGTTTGACAGTTTGAAAAGCAACATGTTTGTAGTGCAGCTA -
15180 - L Y * N Y R V * Q F E K Q H C L L V Q L
- F I K I T G F D S L K S N I V C * C S Y
- L L K L P G L T V * K A T L F V S A A T
15181 - CTGAAAAGCATGTAGTGCCTTATCTAGCAATAAATTGCCAGAAGCTGCATGCATAGCTG -
15240 - L K S M * C V Y L A I N C Q K L H A * L
- * K A C S A F I * Q * I A R S C M H S W
- E K H V V R L S S N K L P E A A C I A G
15241 - GATCAGCAGCATACACTAAAAGTTCCTGAAACTGAGCGGAGCTATGTAAGTTTACAT -
15300 - D Q Q H T L K V P * N * D A S Y V S L H
- I S S I H * K F L E T E T R A M * V Y I
- S A A Y T K S S L K L R R E L C K F T S
15301 - CCTGATTATGTACGACTCCTAACTCAGAAAATGGTATCCAGTTGAAACAACAAAAGGAA -
15360 - P D Y V R L L T H E N G I Q L K Q Q K E
- L I M Y D S * L T K M V S S * N N K R N
- * L C T T P N S R K W Y P V E T T K G T
15361 - CACCATCTACAAATATTTTCTTACTAGTGGTCCAAAACCTGTAGGTGGAACACAGTAG -
15420 - H H L Q I F F L L V V Q N L * V E T Q *
- T I Y K Y F S Y * W S K T C R W K H S R
- P S T N I F L T S G P K L V G G N T V E
15421 - AAAATAACACATTAAGTTTGACAAATGAAGGATACACCTATCATCCAACAGTTAATAC -
15480 - K I T H * S L H N E G Y T Y H P N S * Y
- K * H I K V C T M K D T P I I Q T V N T
- N N T L K F A Q * R I H L S S K Q L I Q
15481 - AATTGGGATGGTATGTCTGGTCCCAATATTTAAAATAACGGTCAAGAGACAAAGTCTCT -
15540 - N W D G M S G P N I * N N G R R D K V S
- I G M V C L V P I F K I T V E E T K S L
- L G W Y V W S Q Y L K * R S K R Q S L S
15541 - CTTCCGTAATAATCATATTCAGCAAATCCCACTTAATAAGTGGTTTTCGAGATCAGCAT -
15600 - L P * N H I S A N P T * * V V L R D Q H
- F R K I I F Q I P L N K W F C E I S I
- S V K S Y F S K S H L I S G F A R S A S
15601 - CCATATGGGACTCAGCAGCCAATGCCCTAGTCAAAGTGAGGATGGGCATCAGCAATGAGT -

```

图 17-U

15660

- P Y G T Q Q P M P \* S K \* G W A S A M S  
- H M G L S S Q C P S Q S E D G H Q Q \* V  
- I W D S A A N A L V K V R M G I S N E \*



```

15661 - AATATGAATCCACAATAGGAATCCGCAGCCTGGTGCTACTTGTACGAAATCACCGAAAT -
15720
- N M N P Q * E L R S L V L L V R N H R N
- I * I H N R N S A A W C Y L Y E I T E I
- Y E S T I G T P Q P G A T C T K S P K S
15721 - CGTACCAGTTCCCATTAAGATCCTGATTATCTAATGTCAGTACGCCTACAATGCCTGCAT -
15780
- R T S S H * D P D Y L M S V R L Q C L H
- V P V P I K I L I I * C Q Y A Y N A C I
- Y Q F P L R S * L S N V S T P T M P A S
15781 - CACGCATAGCATCGCAGAATTGTACAGTCTTTAATAATGATTGGCGTACACGCTCACCTA -
15840
- H A * H R R I V Q S L I M I G V H A H L
- T H S I A E L Y S L * * * L A Y T L T *
- R I A S Q N C T V F N N D W R T R S P K
15841 - AGTTAGCATATACGCGTAAGATGTCAGGATTCTCTACGAAGTCATACCAATCCTTCTTAT -
15900
- S * H I R V R C Q D S L R S H T N P S Y
- V S I Y A * D V R I L Y E V I P I L L I
- L A Y T R K M S G F S T K S Y Q S F L L
15901 - TGAATAATCATCATCACAGCAATTGTATGTGACGAGTATTTCTTTAATGTATCAAAAT -
15960
- * N N H H H S N C M * R V F L L M Y H N
- E I I I I T A I V C D E Y F F * C I T I
- K * S S S Q Q L Y V T S I S F N V S Q L
15961 - TACCCTCATCAAAATGACGTAGAGCATAGACTAAATCAGCCATTGTGTATTAGTTAGAC -
16020
- Y P H Q N D V E H R L N Q P L C I * L D
- T L I K M T * S I D * I S H C V F S * T
- P S S K * R R A * T K S A I V Y L V R R
16021 - GCTGACGTGATATATGTGGTACCATGTCACCATCTACTCTAAACTTGAAAAAGTCATGGA -
16080
- A D V I Y V V P C H H L L * T * K S H G
- L T * Y M W Y H V T I Y S K L E K V M D
- * R D I C G T M S P S T L N L K K S W T
16081 - CAGCAACCGCTGGACAATCTTTAACCAGTTATAAATAGTCTCTTCATGTTGGTAGTTAG -
16140
- Q Q P L D N L * P S Y K * S L H V G S *
- S N R W T I F N Q V I N S L F M L V V R
- A T A G Q S L T K L * I V S S C W * L D
16141 - ACATAGTATGCCTCTTAACTACAAAGTAAGAGTCTAATAAATGCTTCCTCCTCATCCTTCT -
16200
- T * Y A S * L Q S K S L I N C L P H P S
- H S M P L N Y K V R V * * I A F L I L L
- I V C L L T T K * E S N K L P S S S F S
16201 - CCTGGAAGCGACAGCAATTAGTTTTTAGGAACTTGCAAAACCAGCACTTTTTTCGTTGT -
16260
- P G S D S N * F L G T L Q N Q H F F R C
- L E A T A I S F * E L C K T S T F F V V
- W K R Q Q L V F R N F A K P A L F S L *
16261 - AAATATCAAAAGCCCTGTAGACGACATCAGTACTAGTGCCTGTGCCGCACGGTGTAAAGAC -
16320
- K Y Q K P C R R H Q Y * C L C R T V * D
- N I K S P V D D I S T S A C A A R C K T
- I S K A L * T T S V L V P V P H G V R R
16321 - GGGCTGCACTTACACCGCAAACCCGTTTAAAAACGTTGATGCATCCGCAGACTGCATCAA -

```

图 17-V

16380

- G L H L H R K P V \* K R \* C I R R L H Q  
- G C T Y T A N P F K N V D A S A D C I K  
- A A L T P Q T R L K T L M H P Q T A S R

```

16381 - GGGTTCGCGGAGTTGGTCACAACCTACAGCCATAACCTTCCACATTCCGCAGACGGTACA -
16440
- G F A E L V T T T A I T F P H S A D G T
- G S R S W S Q L Q P * P F H I P Q T V Q
- V R G V G H N Y S H N L S T F R R R Y R
16441 - GACTGTGTTTCTAAGTGTA AACCCACTGGGTCATTAGCACAGTGGTAGGTATTGGAC -
16500
- D C V S K C K T H W V I S T S G R Y L D
- T V F L S V K P T G S L A Q V V G I W T
- L C F * V * N P L G H * H K W * V F G R
16501 - GTACTTACCTTTCAAGTCACAGAATCCTTTAGGATTTGGATGGTCAATGTGGCATCTACA -
16560
- V L T F Q V T E S F R I W M V N V A S T
- Y L P F K S Q N P L G F G W S M W H L Q
- T Y L S S H R I L * D L D G Q C G I Y N
16561 - ATACAGACAACATGAAGCACCACCAAGGACTCTTGGTCCATGTTAGCTTCTGGTGTAC -
16620
- I Q T T * S T T K G L L V H V S F W C Y
- Y R Q H E A P P K D S W S M L A S G V T
- T D N M K H H Q R T L G P C * L L V L Q
16621 - AGTAATTGCCTGTCCGTACCAGTGTGTGTACACAACATCTTCACACAGTTGGTGATTGG -
16680
- S N C L S C T S V C T Q H L H T V G D W
- V I A C P V P V C V H N I F T Q L V I G
- * L P V L Y Q C V Y T T S S H S W * L V
16681 - TTGTCCCTCCACTTGCTAGGTAATCCTTATATGCTTTAGCAGGGTCTACTGCAAAAGCACA -
16740
- L S S T C * V I L I C F S R V Y C K S T
- C P P L A R * S L Y A L A G S T A K A Q
- V L H L L G N P Y M L * Q G L L Q K H R
16741 - GAAGGAAAGCACAGTTGAATTGGCAGGTACTTCTGTAGCATTTCAGCCTGAAGACGTAC -
16800
- E G K H S * I G R Y P C S I S S L K T Y
- K E S T V E L A G T S V A F P A * R R R T
- R K A Q L N W Q V L L * H F Q P E D V L
16801 - TGTAGCAGCTAAACTGCCAGCACCATCTATTTAGGTTGTTAAGCCTTTGATGAA -
16860
- C S S * T A Q H H T S I * V V * A F D E
- V A A K L P S T I P L F R L F K P L M K
- * Q L N C P A P Y L Y L G C L S L * * S
16861 - GTACAAGTATTTCACTTTAGGCCCTTTGGTGTGTCTGTAAACAACTACAAGGTGGTTC -
16920
- V Q V F H F R P F W C V C N K P T R W F
- Y K Y F T L G P F G V S V T N L Q G G S
- T S I S L * A L L V C L * Q T Y K V V P
16921 - CAGTCTGTGTAAATTGTACCTGTACCATCACTCTTAGGGAATCTAGCCATTGAGATC -
16980
- Q F C V N C T C T I T L R E S S P F E I
- S S V * I V P V P S L L G N L A H L R S
- V L C K L Y L Y H H S * G I * P I * D L
16981 - TTGGTGGTCTGATAGTAATGCCAGCACAAACCTACCTCCCTTCGAATTGTTATAGTAGGC -
17040
- L V V * * * C Q H K P T S L R I V I V G
- W W S D S N A S T N L P P F E L L * * A
- G G L I V M P A Q T Y L P S N C Y S R Q
17041 - AAGTGCATTGTCATCAGTACAAGCTGTTTGTGTGGTACCAGCCGCACAGGACATCTGTCC -

```

图 17-W

17100

- K C I V I S T S C L C G T S R T G H L S  
- S A L S S V Q A V C V V P A A Q D I C R  
- V H C H Q Y K L F V W Y Q P H R T S V V

```

17101 - TAGTGCTACTGGACTCAGTTCATTATTCTGTAGTTTAAACAGCTGAGTTGGCTCTTAGAGC -
17160
-.* C Y W T Q F I I L * F N S * V G S * S
- S A T G L S S L F C S L T A E L A L R A
- V L L D S V H Y S V V * Q L S W L L E L
17161 - TGTAACAATAAGAGGCCAAGCCAAATTTGGTGAATTGTCCATGTTAATTCACCTAAGTTG -
17220
- C N N K R P S Q I W * I V H V N F T K L
- V T I R G Q A K F G E L S M L I S L S *
- * Q * E A K P N L V N C P C * F H * V E
17221 - AACAACTCTTGCTATCCGCATCAACAACCTTGCTGGATTTCCAGAGTGCAGATGCATATGT -
17280
- N N L A I R I N N L L D F P E C R C I C
- T I L L S A S T T C W I S Q S A D A Y V
- Q S C Y P H Q Q L A G F P R V Q M H M *
17281 - AAAGGTGTACCATCACAAGTGTCTTGTAGGTACCATAATCAGGGACAACAACCATGAG -
17340
- K G V T I T S V L V G T I I R D N N H E
- K V L P S Q V F L * V P * S G T T T M S
- R C Y H H K C S C R Y H N Q G Q Q P * V
17341 - TTTGGCTGCTAGTCAATGGTATGATGTTGAGTGGAAACACAACCATCACGGCATTGTT -
17400
- F G C C S Q W Y D V E W N T T I T R I V
- L A A V V N G M M L S G T Q P S R A L L
- W L L * S M V * C * V E H N H H A H C *
17401 - GATAATGTTGTTAAGTGCATCATTATCAAGCTTCCTAAGCATAGTGAAGAGCATTGTTT -
17460
- D N V V K C I I I K L P K H S E E H C L
- I M L L S A S L S S F L S I V K S I V C
- * C C * V H H Y Q A S * A * * R A L F A
17461 - CATAGCACTAGTTACTTTTGCCTCTTGTCCTCAGATCTTGCCTGTTTGTACATTTGGGT -
17520
- H S T S Y F C P L V L R S C L F V H L G
- I A L V T F A L L S S D L A C L Y I W V
- * H * L L L P S C P Q I L P V C T F G S
17521 - CATAGCCTGATCTGCCATCTTTTCCAACCTTGGTTCATGGCAGCATCACGGTCAAACCTC -
17580
- H S L I C H L F Q L A L H G S I T V K L
- I A * S A I F S N L R C M A A S R S N S
- * P D L P S F P T C V A W Q H H G Q T Q
17581 - AGATTTAGCCACATTCAAAGATTTCTTTAACTTTTGGAGAAGACTTCAGAATCACCATT -
17640
- R F S H I Q R F L * L F E N D F R I T I
- D L A T F K D F F N F L R T T S E S P L
- I * P H S K I S L T F * E R L Q N H H *
17641 - AGCTACAGCCTGCTCATAGGCCTCCTGGGCAGTGGCATAAGCGGCATATGATGGTAAAGA -
17700
- S Y S L L I G L L G S G I S G I * W * R
- A T A C S * A S W A V A * A A Y D G K E
- L Q P A H R P P G Q W H K R H M M V K N
17701 - ACTAAATCTGAAGCAATAGCCTGAAGAGTAGCACGGTTATCGAGCATTTCCTCGCACAA -
17760
- T K F * S N S L K S S T V I E H F L A Q
- L N S E A I A * R V A R L S S I S S H N
- * I L K Q * P E E * H G Y R A P P R T T
17761 - CCTATTAATGTCTACAGCACCCCTGCATGGATAGCAAACAGACAAAAGAGAAACCATCTT -
17820

```

图 17-X

---

- P I N V Y S T L H G \* Q N R Q K R N H L  
- L L M S T A P C M D S K T D K R E T I F  
- Y \* C L Q H P A W I A K Q T K E K P S S  
17821 - CTCGAAAGCTTCAGTTGTGTCTTTTGCAAGAAGAATATCATTGGGAGTTGTACACATTG -  
17880  
- L E S F S C V F C K K N I I V E L Y T L  
- S K A S V V S F A R R I S L W S C T H C  
- R K L Q L C L L Q E E Y H C G V V H I V

```

17881 - TGCCACAAATTTAGAAGATGACTCTACTCTAAGTTGTTGAAGAACCAGAGCAGTACCAC -
17940
- C P Q F R R * L Y S K L L K N R E Q Y H
- A H N L E D D S T L S C * R T E S S T T
- P T I * K M T L L * V V E E P R A V P Q
17941 - AGATGTGCACCTTTACGTACACATTTAGACTGTACAGTAGCAACCTTGATACATGGTTT -
18000
- R C A L Y V R H F R L Y S S N L D T W F
- D V H F T S D I L D C T V A T L I H G L
- M C T L R Q T F * T V Q * Q P * Y M V Y
18001 - ACCTCCAATACCCAACAACCTAATGTTAAGCTTGAAGCATCAATACTACTCTTAGGAGG -
18060
- T S N T Q Q L N V K L E S I N T T L R R
- P P I P N N L M L S L K A S I L L L G G
- L Q Y P T T * C * A * K H Q Y Y S * E A
18061 - CAAAAGCCCCTGGGAGTTTATATACCTAAATCTTGTGTAGAGACCAAGTAGTCATAAAC -
18120
- Q K P L G V H I P K F L C R D Q V V I N
- K S P W E F I Y L N S C V E T K * S * T
- K A P G S S Y T * I L V * R P S S H K H
18121 - ACCAAGAGTAAGCCTGAAGTAACGGTTGAGTAAACAGAAAAGGCCAAAGTAGCAGCAGCA -
18180
- T K S K P E V T V E * T E K A K V A A A
- P R V S L K * R L S K Q K R P K * Q Q Q
- Q E * A * S N G * V N R K G Q S S S S N
18181 - ACAATAGCCTAAGAAACAATAAACAGCATGATACACTGTAAGGTGTTGCCAGTAATAAA -
18240
- T I A * E T I N K H D T L * G V A S N K
- Q * P K K Q * T S M I H C K V L P V I N
- N S L R N N K Q A * Y T V R C C Q * * I
18241 - TAACAATGGGTAATACTCAACACACAAACACTATAGCTCTAGCTAAAACATGATAGT -
18300
- * Q W V I L N T H K H Y S S S * K H D S
- N N G * Y S T H T N T I A L A K N M I V
- T M G N T Q H T Q T L * L * L K T * * S
18301 - CGTAACGACACCAGAATAGTTAGAGTTACAGAAATAACTAAGGCCACATGGAATAGC -
18360
- R N D T R I V R G Y R N N * G P H G N S
- V T T P E * L E V T E I T K A H M E I A
- * R H Q N S * R L Q K * L R P T W K * L
18361 - TTGATCTAAAGCATTACCATAGTAGACTTTGTAAACAAGTGAATGACATTATCAGTGT -
18420
- L I * S I T I V D F V N K C N D I H Q C
- * S K A L P * * T L * T S V M T F I S V
- D L K H Y H S R L C K Q V * * H S S V S
18421 - CCAAACACGCTTAGCAGCATATCATAAACAGTGCAGCTGTGATGAGAATAAGCAAAC -
18480
- P N T S S S I I I N S A S C H E N K Q N
- Q T R L A A S S * T V R A V M R I S K T
- K H V * Q H H H K Q C E L S * E * A K L
18481 - TAAAGCTGAAGCATAACACAATCCTTAAGCCTATAACCAGACAAGCTAGTGTGAGC -
18540
- * S * S I H N T I L K P I T R Q A S V S
- K A E A Y I T Q S L S L * P D K L V S A
- K L K H T * H N P * A Y N Q T S * C Q P
18541 - CAATCAAGCCATGTGATAGCAGCATACCCAGCTAGCAGGCATGTAGACCATATAAA -
18600

```

图 17-Y

---

- Q F K P C H D T H H P A S R H V D H I K  
- N S S H V M I R I T Q L A G M \* T I L K  
- I Q A M S \* Y A S P S \* Q A C R P Y \* S  
18601 - GTAAGCAACTGTTGCAAGAGAAGGTAACAGAAACAAGCACAAAGAATGCGTGCTTATGCTT -  
18660  
- V S N C C K R R \* Q K Q A Q E C V L M L  
- \* A T V A R E G N R N K H K N A C L C L  
- K Q L L Q E K V T E T S T R M R A Y A \*



```

18661 - AACAAAGCAGCATAGCACATGCAGCAATTGCCATAATACCAAGAGTAAATGGCAAGAAAGC -
18720
- N K Q H S T C S N C H N T K S K W Q E S
- T S S I A H A A I A I I P R V N G K K A
- Q A A * H M Q Q L P * Y Q E * M A R K H
18721 - ATTCTCGTAAACRAAGAAAACAGTGACCCTGTGTACTTTGAACAAGAATCAATAGTGA -
18780
- I L V N K E K Q * P L C T L N K N Q * *
- F S * T K K N S D H C V L * T R I N S D
- S R K Q R K T V T T V Y F E Q E S I V M
18781 - TGTCAGAAAGTTAAAAGCATCCAATGATGAGTGCCTTAAACAATTTCTTGAACCTTACC -
18840
- C Q E S * K H P M M S A L N N F L E L T
- V K K V K S I Q * * V P L T I F L N L P
- S R K L K A S N D E C P * Q F S * T Y L
18841 - TTGGAAGGTAACACCAGAGCATTGTCTAACACATCAAATGGTGTAACCTCATCTTCTAA -
18900
- L E G N T R A L S N N I K W C K L I F *
- W K V T P E H C L T T S N G V N S S S K
- G R * H Q S I V * Q H Q M V * T H L L K
18901 - AATAGTGCTACCAAGGATAGTACGACCATTATACCATTCTGCAGCAGCTCTTTCAAAGC -
18960
- N S A T K D S T T I H T I L Q Q L F Q S
- I V L P R I V R P F I P F C S S S F K A
- * C Y Q G * Y D H S Y H S A A A L S K Q
18961 - AGCACATATCTAAGACGGCAATCTCTGTTGAGCAGAAAGAGTCCCAATATGTCAAC -
19020
- S T H I * D G N S C L S R K R S Q Y V N
- A H I S K T A I P V * A E R G P N M S T
- H T Y L R R Q F L F E Q K E V P I C Q H
19021 - ATGATCTGTGTCAAAGGTTATAGTTGTACTTCATTGCCACAAGGTAAAGTCATTCAA -
19080
- M I L C Q R F I V V L H C H K V K V I Q
- * S C V K G S * L Y F I A T R L K S F K
- D L V S K V H S C T S L P Q G * S H S K
19081 - AGTAGTGGTGAATCTATTAAGAAACCCTATCACCATTGATAACAGCAGCATACAGCCA -
19140
- S S G E S I K K P P I T I D N S S I Q P
- V V V N L L R N H L S P L I T A A Y S H
- * W * I Y * E T T Y H H * * Q Q H T A M
19141 - TGCCAAAACATTTAATGTTATGGTTGTGTCTGTACCTGCAGCCTGTGCAGTTTGTCTGTC -
19200
- C Q N I * C Y G C V C T C S L C S L S V
- A K T F N V M V V S V P A A C A V C L S
- P K H L M L W L C L Y L Q P V Q F V C Q
19201 - AACAAATGGACCATAGAATTTACCTTCTAAGTCAGTACCAGCGTGTACTCTCTGTTGGAAG -
19260
- N K W T I E F T F * V S T S V Y S C W K
- T N G P * N L P S K S V P A C T P V G S
- Q M D H R I Y L L S Q Y Q R V L L L E A
19261 - CTCCATATGATGCATATAGCAGAAAGACACGCAATCATAATCAATGTTAAAACCAACT -
19320
- L H M M H I A E R H A I I I N V K T N T
- S I * C I * Q K D T Q S * S M L K P T L
- P Y D A Y S R K T R N H N Q C * N Q H Y
19321 - ACCACATGATCCATTAAGGAAAGAACCTTTAATGGTATGATTAGGTCTCATGGCACACTG -

```

图 17-Z

19380

- T T \* S I K E R T F N G M I R S H G T L  
- P H D P L R K E P L M V \* L G L M A H \*  
- H M I H \* G K N L \* W Y D \* V S W H T D

```

19381 - ATAACACCAGATGGTGAACCATTTGTAGCATGCTAGAACTGAAAATGTTTGACCAGGTTG -
19440
- I N T R W * T I V A C * N * K C L T R L
- * T P D G E P L * H A R T E N V * P G W
- K H Q M V N H C S M L E L K M F D Q V G
19441 - GATACGGACAAATTTATACTTGGGTGTCTTAGGGTTAGAAGTATCAACTTTAAGCCTAAG -
19500
- D T D K F I L G C L R V R S I N F K P K
- I R T N L Y L G V L G L E V S T L S L S
- Y G Q I Y T W V S * G * K Y Q L * A * A
19501 - CAGACAATTTGCATAGAATGGCCAATAACACGAAGTTGAACATTGCCAGCCTGAACAAG -
19560
- Q T I L H R M A N N T K L N I A S L N K
- R Q F C I E W P I T R S * T L P A * T R
- D N F A * N G Q * H E V E H C Q P E R E
19561 - AAAGCTATGGTTGGATTGCGAATGAGCAGATCTTCATAGTTAGGATTAAGCATGTCTTC -
19620
- K A M V G F A N E Q I F I V R I K H V F
- K L W L D L R M S R S S * L G L S M S S
- S Y G W I C E * A D L H S * D * A C L L
19621 - TGCTGTGCAATGACATGTCTTGGACAGTATACTGTGTTCATCCAACCACAATCCATTAAG -
19680
- C C A N D M S W T V Y C V I Q P Q S I K
- A V Q M T C L G Q Y T V S S N H N P L R
- L C K * H V L D S I L C H P T T I H * E
19681 - AGTTGTAGTTCCACAGTTACTTGTACCATGCACCCTTCAACTTTGCCTGACGGGAATGC -
19740
- S C S S T G Y L Y H A P F N F A * R E C
- V V V P Q V T C T M H P S T L P D G N A
- L * F H R L L V P C T L Q L C L T G M P
19741 - CATTTTCCTAAAACCCTCTGCAGAACAGCAGAAGTGATTGATGTCTGTGGTGGTTGGTA -
19800
- H F P K T T L Q N S R S D * C L W W L V
- I F L K P L C R T A E V I D V C G G W *
- F S * N H S A E Q Q K * L M S V G V G R
19801 - GAGAACATCAGCACCTGAGTTGCTAAAGTCATTTAGAGCCTTTGCTAAGTGGCAGCAAGC -
19860
- E N I S T * V A K V I * S L C * V A A S
- R T S A P E L L K S F R A F A K W Q Q A
- E H Q H L S C * S H L E P L L S G S K L
19861 - TGCTTCAGATAGCTGGTAGTATCTAAGGCTCCACTGAAATACTTGTACTTGTATATAG -
19920
- C F T I A G S I * G S T E I L V L V I *
- A S R * L V V S K A P L K Y L Y L L Y R
- L H D S W * Y L R L H * N T C T C Y I E
19921 - AGCAAGATACCTGTTATACTGTGTAAGTGGCAACAGTGTCTCGCTACGCAATTTTAGGTA -
19980
- S K I P V I L C K W Q Q C L A T Q F * V
- A R Y L L Y C V S G N S V S L R N F R Y
- Q D T C Y T V * V A T V S R Y A I L G T
19981 - CATTTCTTGTGAGCAAAAAGGTACACAAAGCAGCCTCCTCGAAGGTACTAAATGTAAC -
20040
- H F L V E Q K G T Q S S L L E G T K C N
- I S L L S K K V H K A A S S K V L N V T
- F P C * A K R Y T K Q P P R R Y * M * L
20041 - TCCATTAAACATGACTCTTTTCTAAGATAGTTGTTAAAGAACCAATGGCAGTGCTTCAG -

```

图 17-AA

20100

- S I K H D S F P K I V V K E P M A V L Q  
- P L N M T L F L R \* L L K N Q W Q C F R  
- H \* T \* L F S \* D S C \* R T N G S A S E

```

20101 - AGAAATACAGAATACATAGATTGCTGTTATCCAAAAGGCACAATAGGAGAAAACATGGC -
20160
- R N T E Y I D C C Y P K R H N R R K H G
- E I Q N T * I A V I Q K G T I G E N M A
- K Y R I H R L L L S K K A Q * E K T W Q
20161 - AAACCATGGAAGGTGAGCCAAGAATGAAACATCATTTGGTGAATAGAATGTCAAGTACAA -
20220
- K P L K V S Q E * N I I G E I E C Q V Q
- N H * R * A K N E T S L V K * N V K Y K
- T I E G E P R M K H H W * N R M S S T S
20221 - GTAAAAGACTGAGTAGACTCCCGGCAGAAAGCTGTAAGCTGGTACCAGACAGAGTATAGT -
20280
- V K D * V D S R Q K A V S W Y Q T E Y S
- * K T E * T P G R K L * A G T R Q S I V
- K R L S R L P A E S C K L V P D R V * *
20281 - GAAAGACATCAAAAACAAAAGTGCATTAGCAGCAACAACATGGTTGTACTACCAAAAAC -
20340
- E R H Q K Q K C I S S N N M V V L T K N
- K D I K N K S A L A A T T W L Y S P K T
- K T S K T K V H * Q Q Q H G C T H Q K H
20341 - ACGTCTGAATTTTCATAAAGTAGTAGGCAGCACAAGTCACCAATATGGCAATAATACCACC -
20400
- T S E F H K V V G S T S H Q Y G N N T T
- R L N F I K * * A A Q V T N M A I I P P
- V * I S * S S R Q H K S P I W Q * Y H Q
20401 - AGCCACTACTGAGCAGACACATCTAAAGCACCCACAGTTGCACAAAGAGGAGTAAAGAT -
20460
- S H Y * S R H I * S T H R L H K R S K D
- A T T E A D T S K A P T G C T R G V K M
- P L L K Q T H L K H P Q V A Q E E * R C
20461 - GTTAGCTATGAGATTCATCGCATCAACACCACAGAAAACCTCCTGATAGAGCTCTGTAATG -
20520
- V S Y E I H R I N T T E N S * * S S V M
- L A M R F I A S T P Q K T P D R A L * C
- * L * D S S H Q H H R K L L I E L C N A
20521 - CTCATTATTAAGAACCCATCTACCCTGAGTAGATAGGCAATACCTACTTCTGACCTTTC -
20580
- L I I K N P S T T G R * A N T Y F * P F
- S L L R T H L P L V D R Q I P T S D L S
- H Y * E P I Y H W * I G K Y L L L T F R
20581 - GCATGTACCATGTCTACAGTACTCAGCATCAAAGTTGTTACTACTCTAACAGAACCCTC -
20640
- A C T M S T V L S I K S C Y Y S N R T L
- H V P C L Q Y S A S K V V T T L T E P S
- M Y H V Y S T Q H Q K L L L L * Q N P P
20641 - CAGGTAAGTGTTAGGAAACTGTATGATGGAACCATCCATAAGCACATAACGAGTGTCTGG -
20700
- Q V S V R K L Y D G T I H K H I T S V W
- R * V L G N C M M E P S I S T * R V S G
- G K C * E T V * W N H P * A H N E C L D
20701 - ACGAAGCTCACTATAAGAAATAGAACCCTCTAGCAAATTAGTGTACATAACAATATGGCAC -
20760
- T K L T I R N R T L * Q I S V I T I W H
- R S S L * E I E P S S K L V S * Q Y G T
- E A H Y K K * N P L A N * C H N N M A Q

```

图 17-BB

---

20761 - AGGTTTGCCCATAGCATCCTTAAAAATTGTACACTCAGCAGCAAGAACGCAAGCAGAGGT -  
20820  
- R F A H S I L K N C T L S S K N A S R G  
- G L P I A S L K I V H S A A R T Q A E V  
- V C P \* H P \* K L Y T Q Q Q E R K Q R \*

```

20821 - AGCAAAATCACTATACTCAATGAGTTTGGGAAGGTGTGTAGCAAATGTTGCCAACAGCACT -
20880 - S K I T I L N E F G R C V A N V A N S T
- A K S L Y S M S L E G V * Q M L P T A L
- Q N H Y T Q * V W K V C S K C C Q Q H *
20881 - AAAAACACGAGGTAGAAAATGCAAGAAGTCACCATTGATTGCTCTCAGCACAGTACCCGG -
20940 - K N T R * K M Q E V T I D C S Q H S T R
- K T R G R K C K K S P L I A L S T V P G
- K H E V E N A R S H H * L L S A Q Y P V
20941 - TAAGCCAGGCACTATGAAACCAATCTCTCTTGTAAATGATAGCAGCTACTACAGGGCAGCT -
21000 - * A R H Y E T N L S C N D S S Y Y R A A
- K P G T M K P I S L V M I A A T T G Q L
- S Q A L * N Q S L L * * * Q L L Q G S F
21001 - TTTGTCATTTTTGTATGAACCACCAGCTGGCTAAACCATGCGTCAAACCAAGCATGTTT -
21060 - F V I F V * T T T L A K P C V K T S M F
- L S F L Y E P P R W L N H A S K P A C L
- C H F C M N H H A G * T M R Q N Q H V Y
21061 - ATTTGCAAAACAATCATCAGTAGAAATGATGTCACGAGTGACACCATCCTGAATGGCTTT -
21120 - I C K T I I S R N D V T S D T I L N G F
- F A K Q S S V E M M S R V T P S * M A L
- L Q N N H Q * K * C H E * H H P E W L C
21121 - GTAACCAATGATTTTCATTTGTGTAACCATCATGGATTGACAATGTATGTACTGGCATAAC -
21180 - V T N D F I C V T I M D * Q C M Y W H N
- * P M I S F V * P S W I D N V C T G I T
- N Q * F H L C N H H G L T M Y V L A * R
21181 - GATATAACAAACCAATGCAGCAAGAACGCACAATAATGTGGCCTTAAGCATAAGTTTAAA -
21240 - D I T N Q C S K N A Q * C G L K H K F K
- I * Q T N A A R T H N N V A L S I S L K
- Y N K P M Q Q E R T I M W P * A * V * N
21241 - ACAAGTACTAACAATCTTACCACCCTTGAGTGAGATTTTAGTAGTTATGACATTGACAAAC -
21300 - T S T N N L T T L E * D F S S Y D I D N
- Q V L T I L P P L S E I L V V M T L T T
- K Y * Q S Y H P * V R F * * L * H * Q P
21301 - CTGTCTAGTTGTAGCACAAAGTTAGTGTAAGGATGTTGTTCTTCTTGGCAGCAGTACG -
21360 - L S S C S T S * C K R Y V V L L G S S T
- C L V V A Q V S V K G M L F F L A A V R
- V * L * H K L V * K V C C S S W Q Q Y E
21361 - AATTTGTTTACGCAGCTGTTTCAGATAAAGACATGTAGTCTTTTACATTCCAGATGAGTGA -
21420 - N L F T Q L F R * R H V V F Y I P D E *
- I C L R S C S D K D M * S F T F Q M S E
- F V Y A A V Q I K T C S L L H S R * V K
21421 - AACATTGTGACTTTTTGCTACTTGGGCATTGATATGCCTTGATTACAGTCAATACATGC -
21480 - N I V T F C Y L G I D M P C I T V N T C
- T L * L F A T W A L I C L A L Q S I H A
- H C D F L L L G H * Y A L H Y S Q Y M R
21481 - GCCAAGATCTCTGGGCGTCATGTTTTCAACCTTATTATAGGTGAGCATGAAATTGTTACA -

```

图 17-CC

21540

- A K I S G R H V F N L I I G E H E I V T  
- P R S L G V M F S T L L \* V S M K L L Q  
- Q D L W A S C F Q P Y Y R \* A \* N C Y N



```

21541 - ACTGTCACCTGTCACCTTCTAAGTCAGAGTGATGTGAAAAGTTTGAGACATTCAATAACATC -
21600
- T V T C H F * V R V M * K F E T F N N I
- L S P V T S K S E * C E S L R H S I T S
- C H L S L L S Q S D V K V * D I Q * H P
21601 - CTTTGTGTCAACATCGGTATCAACAACACCTTGTGCGGCAGCTGACACGAATGTAGAAAG -
21660
- L C V N I G I N N T L S G S * H E C R K
- F V S T S V S T T P C R A A D T N V E R
- L C Q H R Y Q Q H L V G Q L T R M * K G
21661 - GACACCATCTAAAGCTACACCCTTTGCTAACTCGCTGTGAGCTGTAGCAACAAGTGCCTT -
21720
- D T I * S Y T L C * L A V S C S N K C L
- T P S K A T P F A N S L * A V A T S A L
- H H L K L H P L L T R C E L * Q Q V P *
21721 - AAGTTTTTCCATAGGAACACTAAAAGTTGCTGAAAAGGTGTGACATAAGCATCAAACAT -
21780
- K F F H R N T K S C * K G V D I S I K H
- S F S I G T L K V A E K V S T * A S N I
- V F P * E H * K L L K R C R H K H Q T S
21781 - CTTAACGGAAACTTCAGTACTATCTCCAACGTTTGATACAAGAGCTTGGTCAAGCAACAG -
21840
- L N G N F S T I S N V * Y K S L V K Q Q
- L T E T S V L S P T F D T R A W S S N R
- * R K L Q Y Y L Q R L I Q E L G Q A T E
21841 - AATAGGTTGGCACATCAGCTGACTGTAGTACACAGAAGCAGACTTAGAAGCAGACTCGTC -
21900
- N R L A H Q L T V V H R S R L R S R L V
- I G W H I S * L * Y T E A D L E A D S S
- * V G T S A D C S T Q K Q T * K Q T R R
21901 - GCATTTGGACTTGCCATCAAAAACCTATGACATTAATAGGCAGTGAACCTTTAGTGTGT -
21960
- A F G L A I K N Y D I N R Q * T F S V V
- H L D L P S K T M T L I G S E P L V L L
- I W T C H Q K L * H * * A V N L * C C *
21961 - AGCTCTCAAATGTCTAAATTGACAAAATGGGAGAGCGGATGTCTCTCATAGGTCTTTT -
22020
- S S Q I V * I D K M G E R M S L I G L L
- A L K L S K L T K W E S G C L S * V F *
- L S N C L N * Q N G R A D V S H R S F D
22021 - ACCAGCCTTGTCAAAGTAGAGGTGAAGCGGCCATTTTTACAGCAACACTATCAACAAT -
22080
- T S L V K V E V K R A I F H S N T I N N
- P A L S K * R * S A P F F T A T L S T I
- Q P C Q S R G E A R H F S Q Q H Y Q Q Y
22081 - ATACGATGACTGGTCAGTAGGGTTGATTGGTCTTTAAACTGGAGTGACAAATCAGGAGC -
22140
- I R * L V S R V D W S F K L E * Q I T S
- Y D D W S V G L I G L L N W S D K S R A
- T M T G Q * G * L V F * T G V T N H E Q
22141 - AACTTCATCACTAATGAATGTACTACCAGTGCAAATGTGTACAATTGAGACAATTCCA -
22200
- N F I T N E C T T S A K C V T I E T I P
- T S S L M N V L P V Q N V S Q L R Q F Q
- L H H * * M Y Y Q C K M C H N * D N S N
22201 - ATTGTGAGTCTTGCAGAAGCCACGGCCTCCATTTGCATAGACATAGAAAGATCTCTTCAT -
22260

```

图 17-DD

---

- I V S L A E A T A S I C I D I E R S L H  
- L \* V L Q K P R P P F A \* T \* K D L F M  
- C E S C R S H G L H L H R H R K I S S C  
22261 - GCCATTAACAATAGTTGTACTCAACGCGTGTGGCAGATTGCGCTTATAGCACATCAT -  
22320  
- A I N N S C T L N A C G T I A L I A H H  
- P L T I V V H S T R V A R L R L \* H I M  
- H \* Q \* L Y T Q R V W H D C A Y S T S C

```

22321 - GCAAGTCGAAGAGGTGCAACCATCCATGATATGAACATAGCTCTTCCATATGTAGTAGAA -
22380
- A S R R G A T I H D M N I A L P Y V V E
- Q V E E V Q P S M I * T * L F H M * * K
- K S K R C N H P * Y E H S S S I C S R K
22381 - AGAAGCAAAGAAGATGTACATCCTAACCATGCGAAACGGGTGCCATTTGTACAATACT -
22440
- R S K E D V H P N H C R N G C H L Y N T
- E A K K M Y I L T I A E T G A I C T I L
- K Q R R C T S * P L Q K R V P F V Q Y *
22441 - AATGATAAACACATGAGCCAAGAATTGCTGATGAAATGACTAGCAAATAGCCAAAGAA -
22500
- N D K P H E P R I A D E M T S K I A K E
- M I N H M S Q E L L M K * L A K * P K N
- * * T T * A K N C * * N D * Q N S Q R T
22501 - CACCTGCATTATAGCTGAAAGACCTAATAAATAAAGAATTTTGTGAACAACATATATGC -
22560
- H L H Y S * K T * * I K E F C E Q H I C
- T C I I A E R P N K * K N F V N N I Y A
- P A L * L K D L I N K R I L * T T Y M P
22561 - CAAAACCCACTCAGCGGCCAGACCTAAAATGTCAAGTCTAGCTTGTACGATGAAATCGT -
22620
- Q N P L S G Q T * N C Q V * L V R * N R
- K T H S A A R P K I V K S S L Y D E I V
- K P T Q R P D L K L S S L A C T M K S S
22621 - CACCTGAATGGTTTCAAGAGCTGGATAAGAATCAAGGGAGTCTAATCCACTTAAACAAT -
22680
- H L N G F K S W I R I K G V * S T * T N
- T * M V S R A G * E S R E S N P L K Q M
- P E W F Q E L D K N Q G S L I H L N K C
22681 - GCTGCAAGGAAAAGAACCTTACAGAAATCCATAGTAGTAACGTTAGACGAATTAAGATA -
22740
- A A R K R T F T E I H S S N V R R I K I
- L Q G K E P S Q K S I V V T L D E L R Y
- C K E K N L H R N P * * * R * T N * D T
22741 - CAATCTCTAACGCCATTACAATAAGAAGGAGCACCAAATTAGATAAGAGTACACCAA -
22800
- Q F S N A I T I R R S T K I R * E Y T K
- N S L T P L Q * E G A P K L D K S T P K
- I L * R H Y N K K E H Q N * I R V H Q K
22801 - AGCAGCAGTTACACAGATTAGAGAACCTAAGCAAATACTTAAACAATAGCCACATAGC -
22860
- S S S Y T D * R T * A N T * Q Q * P H S
- A A V T Q I R E P K Q I L N N N S H I A
- Q Q L H R L E N L S K Y L T T I A T * R
22861 - GATGTGAACAATTTAGAAAATTTGGGTGACTTACATAATTAATGCCGGCATCCAAACA -
22920
- D C E Q F R K F G * L H I I N A G I Q T
- I V N N L E N L G D F T * L M P A S K H
- L * T I * K I W V T S H N * C R H P N I
22921 - TAATTTAGCAACACTCTTAACACTATTTTTAGCAATAGTTGTAGGTAGTGAAGCTCTAAT -
22980
- * F S N T L N T I F S N S C R * * S S N
- N L A T L L T L F L A I V V G S E A L I
- I * Q H S * H Y F * Q * L * V V K L * F
22981 - TCTAGAATGGTACTTTTAGTAAAAGTACACAATTGGAACAATAATGTAAACACATAAGG -

```

图 17-EE

23040

- S R I G T F S K S T Q L E Q \* C K H I R  
- L E L V L L V K V H N W N N N V N T \* G  
- \* N W Y F \* \* K Y T I G T I M \* T H K A

```

23041 - CATATAATTGTTAAACACACAGTTGTGCTAATCTCTTAGCGCAATTTGATGTTGTAATTGC -
23100 - H I I V K H T L C * S L S A I * C C N C
- I * L L N T R C A N L L A Q F D V V I A
- Y N C * T H V V L I S * R N L M L * L L
23101 - TGCTTGTCCCTAAGAATGGTTTGACATAAGCCAAAATTTACTCCAAGGAACACTATTAAT -
23160 - C L S * E W F D I S Q N F T P R N T I N
- A C P K N G L T * A K I L L Q G T L L I
- L V L R M V * H K P K F Y S K E H Y * L
23161 - TGCAGCAATACCATGAGTGGCAATGTTTTTAAACCTAAGGCTAGTGAAGCTCATTAGG -
23220 - C S N T M S G N C F * T * G * * K L I R
- A A I P * V A I V F K P K A S E S S L G
- Q Q Y H E W Q L F L N L R L V K A H * V
23221 - TTTCTTAATGGTAATGCTTGTGTTTTCCACATAAGCAGCCATAAGATCCTCATGACCTAA -
23280 - F L N G N A C V F H I S S H K I L M T *
- F L M V M L V F S T * A A I R S S * P N
- S * W * C L C F P H K Q P * D P H D L T
23281 - CTCTTGTGTACTTTAACACCTTCATCTGATGGTTAAGTATGACATTGCCTACAACCTC -
23340 - L L C Y F N T F I * W F K Y D I A Y N F
- S C V T L T P S S D G L S M T L P T T S
- L V L L * H L H L M V * V * H C L Q L R
23341 - GGTAGTTTTACGTCACACTCTATGACTTCCTTCTGTATGGTAGGATTTTCCACTACTTC -
23400 - G S F H V T L Y D F L L Y G R I F H Y F
- V V F T S H S M T S F C M V G F S T T S
- * F S R H T L * L P S V W * D F P L L L
23401 - TTCAGAGGTGGGTGTTGACTTTTACAAGCAAGATTGTCCATTCCTTGTGTCTTCTAC -
23460 - F R G G L L T F T S K I V H S L C V F Y
- S E V G C * L S Q A R L S I P C V S S T
- Q R W V V D F H K Q D C P F L V C L L L
23461 - TGCCAGAACTTCAAATGAATTTGAAGTATCTACTGGCTTTGTACTCCAAAGACAACGTAA -
23520 - C Q N F K * I * S I Y W L C T P K T T *
- A R T S N E F E V S T G F V L Q R Q R K
- P E L Q M N L K Y L L A L Y S K D N V N
23521 - ACACCAAGTGTGTTGGTTGAACGTTGTCTTGGTTGTAGCCTGGTTAATGTGCCAAACAA -
23580 - T P S V W F E R C L G C S L V N V P N N
- H Q V F G L N V V L V V A W L M C Q T I
- T K C L V * T L S W L * P G * C A K Q L
23581 - TGGCTTATGCAGTAATTTAGCACCTTCTTGAACCTCGCTGAATAGTGTCTATAGTCAAT -
23640 - W L M Q * F S T F L E T R * I V S I V N
- G L C S N L A P F L K L A E * C L * S I
- A Y A V I * H L S * N S L N S V Y S Q *
23641 - AGCCACTACATCGCCATTCAAGTCTGGGAAGAATGTGACAGATAGCTCTCGTGAAGCTGG -
23700 - S H Y I A I Q V W E E C D R * L S * S W
- A T T S P F K S G K N V T D S S R E A G
- P L H R H S S L G R M * Q I A L V K L A
23701 - CTTTGTGAAGCCTGTCAATTTGATTTAAATCATCAGCAAATTTGTGTTAGAACAATGTGAG -
23760

```

图 17-FF

---

- L C E A C H L I \* I I S K F C V R T C E  
- F V K P V I \* F K S S A N F V L E H V S  
- L \* S L S F D L N H Q Q I L C \* N M \* V  
23761 - TTTGAAATTATCAAACCTCGCATTGGTAATGGTTGAGTTGGTACAAGGTCTATAGGCTG -  
23820  
- F E I I K T R I W \* W L S W Y K V Y R L  
- L K L S K L A F G N G \* V G T R S I G C  
- \* N Y Q N S H L V M V E L V Q G L \* A A

```

23821 - CTCTGTATAGTAAGCATTATCCTTTTATAATACCCATCCAATTTGGTTCAATCTCTGT -
23880
- L C I V S I I L F I I P I Q F W F N L C
- S V * * A L S F L * Y P S N F G S I S V
- L Y S K H Y P F Y N T H P I L V Q S L C
23881 - GTAAGTAACTCCATCGAGTTTATACGACACAGGCTTGATGGTTGTAGTGTAAAGATGTTTC -
23940
- V S N S I E F I R H R L D G C S V R C F
- * V T P S S L Y D T G L M V V V * D V S
- K * L H R V Y T T Q A * W L * C K M F P
23941 - CTTGTAGAAAACATCAGTCACTGGTCCTTGTACTCTGACATCTTTGTAAGGTGAGCTCC -
24000
- L V E N I S H W S F V L * H L C K V S S
- L * K T S V T G P L Y S D I F V R * A P
- C R K H Q S L V L C T L T S L * G E L R
24001 - GTCAATACGATAGAGGGTCTCCTTAGCAGTTATATGAGTGTAAAGACCACACTGATAGTT -
24060
- V N T I E G L L S S Y M S V M T T L I V
- S I R * R V S L A V I * V * * P H * * L
- Q Y D R G S P * Q L Y E C N D H T D S Y
24061 - ACCAGTGTACTCATTGCGACATAAGAATGTACCTTGCTGTAATTTATACTCAGCAGGTGG -
24120
- T S V L I R T * E C T L L * F I L S R W
- P V Y S F A H K N V P C C N L Y S A G G
- Q C T H S H I R M Y L A V I Y T Q Q V V
24121 - TGCAGACATCATAACAAAAGAAGACTCTTGTGTACTAGATATTGTGTAGCATCAGCACC -
24180
- C R H H N K R R L L L Y * I L C S I T T
- A D I I T K E D S C C T R Y C V A S R P
- Q T S * Q K K T L V V L D I V * H H D H
24181 - ACACACACATGGAATGGAACACCTGTCTTAAGATTATCATAAGATAGAGTACCCATATA -
24240
- T H T W N G N T C L K I I I R * S T H I
- H T H G M E T P V L R L S * D R V P I Y
- T H M E W K H L S * D Y H K I E Y P Y T
24241 - CATCACAGCTTCTACACCCGTTAAGGTAGTAGTTTCTGACCACAATGTTTACACACCAC -
24300
- H H S F Y T R * G S S F L T T M F T H H
- I T A S T P V K V V V F * P Q C L H T T
- S Q L L H P L R * * F S D H N V Y T P H
24301 - ATTAAGAACTCGCTTTGCAGATTCCAAATTAGCATGCTGTAGAAGATGGGTCATAGTTTC -
24360
- I K N S L C R F Q I S M L * K M G H S F
- L R T R F A D S K L A C C R R W V I V S
- * E L A L Q I P N * H A V E D G S * F L
24361 - TCTGACATCACCAGCTCGCCAACAGTTTATTACTGTAAGCGAGTATGAGTGCACAAAA -
24420
- S D I T K L A N S F I T V S E Y E C T K
- L T S P S S P T V L L L * A S M S A Q K
- * H H Q A R Q Q F Y Y C K R V * V H K S
24421 - GTTAGCAGCATCACCAGCAGGGCTCTATAATAAGCCTCTTGAAGTGCTGGTGCATTGAA -
24480
- V S S I T S T G S I I S L L K C W C I E
- L A A S P A R A L * * A S * S A G A L N
- * Q H H Q H G L Y N K P L E V L V H * I
24481 - TTTGACTTCAAGCTGTTGAAGTGCTAATAAAACACTAGACAAATAACAATGTTATCAGC -

```

图 17-GG

24540

- F D F K L L K C \* \* N T R Q I T I V I S  
- L T S S C \* S A N K T L D K \* Q L L S A  
- \* L Q A V E V L I K H \* T N N N C Y Q P



```

24541 - CCATTTAATTGAAGTTAAACCACCAACTTGAGGAAATTTCCATTTCTTTGTGGTTTAA -
24600
- P F N * S * T T N L R K F P F L C V V *
- H L I E V K P P T * G N F H F F V W F K
- I * L K L N H Q L E E I S I S L C G L K
24601 - AGCAGACATGTACCTACCAAGAAAACCTCATCAAGAGTATGGTAGTACTCGAAAGCTTC -
24660
- S R H V P T K K T L I K S M V V L E S F
- A D M Y L P R K L S S R V W * Y S K A S
- Q T C T Y Q E N S H Q E Y G S T R K L H
24661 - ACTACGTAGTGTGTCATCACTAGGTAGTACAAAGAAAGCTTACCCTCATGATTTACATG -
24720
- T T * C V I T R * Y K E S L T L M I Y M
- L R S V S S L G S T K K V L P S * F T *
- Y V V C H H * V V Q R K S Y P H D L H E
24721 - AGGTTTAAATTTTGTAAACATCAGCACCATCCAAGTATGTTGGACCAAACCTGCTGTCCATA -
24780
- R F N F C N I S T I Q V C W T K L L S I
- G L I F V T S A P S K Y V G P N C C P Y
- V * F L * H Q H H P S M L D Q T A V H M
24781 - TGTCATAGACATATCCACAAGCTGTGTGTGGAGATTAGTGTTCACAGTTGTGAACAC -
24840
- C H R H I H K L C V E I S V V H S C E H
- V I D I S T S C V W R L V L S T V V N T
- S * T Y P Q A V C G D * C C P Q L * T L
24841 - TTTTATAGTCTTAACCTCCCAGGGATAAGAGACTCTTAGTTGTCAAGTGAAGAAC -
24900
- F Y S L N L P Q G * E T L * F V K * K N
- F I V L T S R R D K R L F S L S S E R T
- L * S * P P A G G I R D S L V C Q V K E P
24901 - CTCACCGTCAAGATGAAACTCGACGGGGCTCTCCAGAGTGTGGTACACAATTTTGTACC -
24960
- L T V K M K L D G A L Q S V V H N F V T
- S P S R * N S T G L S R V W Y T I L S P
- H R Q D E T R R G S P E C G T Q F C H H
24961 - ACGCTTAAGAAATTCACACCTAACTCTGTACGCTGTCTGAATAGGACCAATCTCTGTA -
25020
- T L K K F N T * L C T L S * I G P I S V
- R L R N S T P N S V R C P E * D Q S L *
- A * E I Q H L T L Y A V L N R T N L C K
25021 - AGAGCCAGCCAAAGAAACTGTTTCTACAAAGTGCTCCTCAGATGTCTTTGATGACGAAGT -
25080
- R A S Q R N C F Y K V L L R C L * * R S
- E P A K E T V S T K C S S D V F D D E V
- S Q P K K L F L Q S A P Q M S L M T K *
25081 - GAGGTATCCATTATATGTAGTAACAGCATCTGGTGATGATACTGACACTACGGCAGGAGC -
25140
- E V S I I C S N S I W * * Y * H Y G R S
- R Y P L Y V V T A S G D D T D T T A G A
- G I H Y M * * Q H L V M I L T L R Q E L
25141 - TTTAAGAGAACGCATACAGCGCGAGCCTCTTCAAGATTAAAACCATGTGCACATAACC -
25200
- F K R T H T A R S L F K I K T M C H I T
- L R E R I Q R A A S S R L K P C V T * P
- * E N A Y S A Q P L Q D * N H V S H N Q

```

图 17-HH

---

25201 - AATTGGCATTGTGACAAGCGGCTCATTAGAGAGTTCAGCTTCGTAATAATAGAAGCTAC -  
25260  
- N W H C D K R L I \* R V Q L R N N R S Y  
- I G I V T S G S F R E F S F V I I E A T  
- L A L \* Q A A H L E S S A S \* \* \* K L Q

```

25261 - AGGCTCTTTACTAGTATAAAAAGAAGAATCGGACACCATAGTCAACGATGCCCTCTTGAAT -
25320
- R L F T S I K E E S D T I V N D A L L N
- G S L L V * K K N R T P * S T M P S * I
- A L Y * Y K R R I G H H S Q R C P L E F
25321 - TTTAATTCCTTTATACTTACGTTGGATGGTTGCCATTATGGCTCTAACATCCATGCATAT -
25380
- F N S F I L T L D G C H Y G S N I H A Y
- L I P L Y L R W M V A I M A L T S M H I
- * F L Y T Y V G W L P L W L * H P C I *
25381 - AGGCATTAATTTCTTGTCTCTTCAGCATGAGCAAGCATTTCTCTCAAATCCAGGATAC -
25440
- R H * F S C L P S M S K H F S Q I P G Y
- G I N F L V S S A * A S I S L K P Q D T
- A L I F L S L Q H E Q A F L S N S R I Q
25441 - AGTTCCTAGAATCTCTTCCTTAGCATTAGGTGCTTCTGAAGGTAGTACATAAAATGCAGA -
25500
- S S * N L F L S I R C F * R * Y I K C R
- V P R I S S L A L G A S E G S T * N A D
- F L E S L P * H * V L L K V V H K M Q I
25501 - TTTGCATTTCTTAAGAGCAGTCTTAGCTTCTCAAGTGTATAACCAGCACATCCTTGTC -
25560
- F A F L K S S L S F L K C I T S T S L S
- L H F L R A V L A S S S V * P A H P C P
- C I S * E Q S * L P Q V Y N Q H I L V Q
25561 - AGGGTACGTGGTTATATACTCATCAACTGGCATTCTTCAAAGCTCTTGAGAGCATCTC -
25620
- R V R G Y I L I N W H F L Q S S * E H L
- G Y V V I Y S S T G T F F K A L E S I S
- G T W L Y T H Q L A L S S K L L R A S Q
25621 - AGTAGTGCCACCAGCCTTTTGGAGGGTATTACAACACAAGTGATATCACCCTAGTGAT -
25680
- S S A T S L F G G Y Y N T S D I T T S D
- V V P P A F L E G I T T Q V I S P L V I
- * C H Q P F W R V L Q H K * Y H H * *
25681 - AACATCACCTACCATGTAAGGTGCATCCTTCTCAAGGAAAGACATATCTTACCTCTAAG -
25740
- N I T Y H V R C I L L K E R H I F T S K
- T S P T M * G A S F S R K D I S S P L S
- H H L P C K V H P S Q G K T Y L H L * A
25741 - CATGTTCTGAGAATCATGGTAAAGCTTACCATTGATATCAGCAAACAAGAGTAACTTATT -
25800
- H V L R I M V K L T I D I S K Q E * L I
- M F * E S W * S L P L I S A N K S N L L
- C S E N H G K A Y H * Y Q Q T R V T Y W
25801 - GGTAAGAACTTAGTTTCTCCAGTGTGGTAACTCATCAATGCAGGCCTTAATTTT -
25860
- G K K L S F F Q C C G N L I N A G L N F
- V R N L V S S S V V V T S S M Q A L I F
- * E T * F L P V L W * P H Q C R P * F L
25861 - TGGCTTACATCGACAGGCTTCTGTACGACAGATTCTCCTCAGTFTTGGAACTCTTCTGT -
25920
- W L H I D R L L Y D R F L L S F G I F C
- G F T S T G F C T T D F S S V L E S S V
- A S H R Q A S V R Q I S P Q F W N L C
25921 - GTTGGTGGCTCCTTGTTTAGGTGCTTCCACTCTAGGCTTCAGGTATCAAGATAATC -
25980

```

图 17-11

---

```
- V W W L L L F R C F H S R L Q V I K I I
- F G G S S C L G A S T L G F R L S R * S
- L V A P L V * V L P L * A S G Y Q D N P
25981 - CATGACAACCTGCTCATAAAGAGCTTTGTCATTGACTGCAATATAAACCTGTGTACGAAC -
26040
- H D N L L I K S F V I D C N I N L C T N
- M T T C S * R A L S L T A I * T C V R T
- * Q P A H K E L C H * L Q Y K P V Y E P
```

```

26041 - CGTCTGCACGCACACTTGTAAAGACTGAAGTGGTTTAGCACCAAATATGCCTGCTGACAA -
26100
- R L H A H L * R L K W F S T K Y A C * Q
- V C T H T C K D * S G L A P N M P A D N
- S A R T L V K T E V V * H Q I C L L T T
26101 - CAATGGTGCAAGTAAGATGTCTGTGAATTGAAATTTTCATATGCTGCCTTAAGAAGCTG -
26160
- Q W C K * D V L * I E I F I C C L K K L
- N G A S K M S C E L K F S Y A A L R S W
- M V Q V R C P V N * N F H M L P * E A G
26161 - GATGTCCTCACCTGCATTTAGGTTAGGTCCAACAACATGCAGACACTTCTTAGCAAGATT -
26220
- D V L T C I * V R S N N M Q T L L S K I
- M S S P A F R L G P T T C R H F L A R L
- C P H L H L G * V Q Q H A D T S * Q D Y
26221 - ATGTCCAGAAAGCAAACAAGACCCTCTACTGTAAGAGGGCCATTAGCTTAATGTAATC -
26280
- M S R K Q T R P S Y C K R A I * L N V I
- C P E S K Q D P P T V R G P F S L M * S
- V Q K A N K T L L L * E G H L A * C N H
26281 - ATCACTCTCCTTTTGCATGGCACCATTGGTTGCCTTGTGAGTGCACCTGCTACACCACC -
26340
- I T L L L H G T I G C L V E C T C Y T T
- S L S F C M A P L V A L L S A P A T P P
- H S P F A W H H W L P C * V H L L H H H
26341 - ACCATGTTTCAGGTGTATGTAGCAGCATTTACAATCACCATAGGATTAGCACATTTGTGC -
26400
- T M F Q V Y V S S I Y N H H R I S T L C
- P C F R C M L A A F T I T I G L A L C A
- H V S G V C * Q H L Q S P * D * H F V P
26401 - CTCCTAACGATGTCAACACATTTAATGGCAACATTGTCAGTAAGTTTAAATAACCACT -
26460
- L L N D V N T F N G N I V S K F * I T S
- S L T M S T H L M A T L S V S F K * P V
- P * R C Q H I * W Q H C Q * V L N N Q *
26461 - AAACGTAACTGGTTCTTCAGGTGTAGGTTCTGGTTCTGGCTCAATCTCTGATTGCTC -
26520
- K L I N W F F R C R F W F W L N L * L L
- N * L T G S S G V G S G S G S I S D C S
- T D * L V L Q V * V L V L A Q S L I A Q
26521 - AGTAGTATCATCCAGCCAGTCTTCTTCTTCTTCTCAACTCGAACTGTTTCAGCTGA -
26580
- S S I I Q P V F L F F F L N S N C F S *
- V V S S S Q S S S S S S S T R T V S A E
- * Y H P A S S L P L L L P Q L E L F Q L R
26581 - GGCACCAAATCCAGAGGGAGACCTTGATAATCATCCTCTGTACCGTACTCATGTTTACA -
26640
- G T K F Q R E T L I I I L C T V L M F T
- A P N S R G R P * * S S S V P Y S C S Q
- H Q I P E G D L D N H P L Y R T H V H R
26641 - GGTTTCATCAATTTCTTCTTCTCCTCACACTCTGCATCGTCTTCTTCTTCTCATCTGGAGG -
26700
- G F I N F F F L T L C I V L F F L I W R
- V S S I S S S S H S A S S S S S S S G G
- F H Q F L L P H T L H R P L L P H L E G

```

图 17-JJ

26701 - GTAAAAGGAACAATACATACGTGATGAAAAGTTTTCTTCACCAGCATCATCAAATAAGTA -  
26760  
- V K G T I H T \* \* K V F F T S I I K \* V  
- \* K E Q Y I R D E K F S S P A S S N K \*  
- K R N N T Y V M K S F L H Q H H Q I S R

```

26761 - GAATGTAGCTACACTCCACTCATCAAGATCAATACCCATGTTGGTAAGGAGATCAGAAAC -
26820
- E C S Y T P L I K I N T H V G K E I R N
- N V A T L H S S R S I P M L V R R S E T
- M * L H S T H Q D Q Y P C W * G D Q K L
26821 - TGGTTGTAAGTCTTCACAACAGCCTCTGCTACAACACATGCAAACCTCAGTAACCTCGGT -
26880
- W L * S L H N S L C Y N T C K L S N F G
- G C K V F T T A S A T T H A N S V T S V
- V V K S S Q Q P L L Q H M Q T Q * L R Y
26881 - ACCGGATTCAACAGTGTAGACAGAGCACTTTTCATTAAGCACTTTGTCAACACGTTTCATC -
26940
- T G F N S V D R A L F I K H F V N T F I
- P D S T V * T E H F S L S T L S T R S S
- R I Q Q C R Q S T F H * A L C Q H V H Q
26941 - AAGCTCAAATGTGATTCTCACATTCTTGTAACCTTGAACCTCCCAAACAGTATCTTCTCC -
27000
- K L K C D S H I L V T L N F P N S I F S
- S S N V I L T F L * P * T S Q T V S S P
- A Q M * F S H S C N L E L P K Q Y L L Q
27001 - AAAGGTTACACCTTTAATTGGTGCACCCCTTTAAGCGAAAGACATTGTTTGTAGCCAG -
27060
- K G Y T F N W C T P F * A K D I V C S Q
- K V T P L I G A P P F K R K T L F V A S
- R L H L * L V H P L L S E R H C L * P V
27061 - TAAACCAGGAGACAATGCGCAGTATTGTTCTTTGTCCCTTAATCTCTAAGAGCATGAGGCC -
27120
- * T R R Q C A V L F F V L N L * E H E A
- K P G D N A Q Y C S L S L I S K S M R P
- N Q E T M R S I V L C P * S L R A * G H
27121 - ATTTACACAGACTGGTGTGCCGACGATAGCTCCATTTGTGAAGCTATCAACGGCGTCTC -
27180
- I Y T D W C A D D S S I C E A I N G R L
- F T Q T G V P T I A P F V K L S T G V S
- L H R L V C R R * L H L * S Y Q R A S R
27181 - GAGTGCCTTCGAGTTCACCGTTCCTTGAGAACACCTCCTCAGAGGTAAGTACTGTGTGCATG -
27240
- E C F E F T V L E N N L L R G K Y C V M
- S A S S S P F L R T T S S E V S T V S C
- V L R V H R S * E Q P P Q R * V L C H V
27241 - TGAATCACCTTCAAGAAAGGTTACTTCTTTTGGTGCCTTAAGAGGCATGAGTAGTTGCAG -
27300
- * I T F K K G Y F F W C L K R H E * L Q
- E S P S R K V T S F G A L R G M S S C S
- N H L Q E R L L L L V P * E A * V V A A
27301 - CTGCTCCTTGCCACGTATACACTGACGGTAAAGTCCCTTGCTTTGAGCGATGAAGACTTC -
27360
- L L L A T Y T L T V K S L A L S D E D F
- C S L P R I H * R * S P L L * A M K T S
- A P C H V Y T D G K V P C F E R * R L H
27361 - ACCTAAGTTGAGTGATCGCAACTTTGCGCCAGCGATAGTACTTGATCAATGCACATTTC -
27420
- T * V E * S Q L C A S D S D L I N A H F
- P K L S D R N F A P A I V T * S M H I S
- L S * V I A T L R Q R * * L D Q C T F R
27421 - GAGTGCCTTGTTAACAACATCAATGAAGCATTTTACACAATCCTTGATGTTATCTGAAGC -
27480

```



图 17-KK

- E C L V N N I N E A F Y T I L D V I \* S  
- S A L L T T S M K H F T Q S L M L S E A  
- V P C \* Q H Q \* S I L H N P \* C Y L K Q  
27481 - AACCTGTATTGACCCTTGACGATGTCAAAAACACCTGTAATGAGAAATTGAGAAATCTC -  
27540  
- N L Y L T L D D V K N T C N E K F E N L  
- T C I \* P L T M S K T P V M R N L R I S  
- P V F D P \* R C Q K H L \* \* E I \* E S P



```

27541 - CCAAGCATCCTTGAGAAATCAACTCCTGCACTAAGTTTCGCCTCAATCCATTCAAAGAT -
27600
- P S I L E K F N S C T K F R L N P F K D
- Q A S L R N S T P A L S F A S I H S K I
- K H P * E I Q L L H * V S P Q S I Q R *
27601 - AGGCCTGAGTTTTTCAACAGTAGTGCCCAAAGATTAGACAACCACTGAGAAGTCTGTG -
27660
- R P E F F N S S A Q K I R Q P L R S L L
- G L S F S T V V P K R L D N H * E V C C
- A * V F Q Q * C P K D * T T T E K S V V
27661 - TACAAGACCACCAGTTACATATGCCATAATAATGACACTGTTGGTGAGCAGGTCTGAAGT -
27720
- Y K T T S Y I C H N N D T V G E Q V * S
- T R P P V T Y A I I M T L L V S R S E V
- Q D H Q L H M P * * * H C W * A G L K Y
27721 - ATAAACCATGGCGTCGACAAGACGTAATGACTGTTTCAGAAATACCATCAAGTATGGTGAC -
27780
- I N H G V D K T * * L F R N T I K Y G D
- * T M A S T R R N D C S E I P S S M V T
- K P W R R Q D V M T V Q K Y H Q V W * Q
27781 - AGCTGCTCTTGCAAATCAGGAATTGAGTGGTTTGCATCAAGTGTGCGCGCAAAAAT -
27840
- S C S L Q I R N * V V C C I K C A R K N
- A A L C K S G I E W F A A S S V R A K I
- L L F A N Q E L S G L L H Q V C A Q K L
27841 - TGATCTGATAACACCAGCAGCCTGTGAGGGAAAACCACACAGTGGTGTAAACTGATCT -
27900
- * S D N T S S L * G K T T Q W C * N * S
- D L I T P A A C E G K P H S G V K T D L
- I * * H Q Q P V R E N H T V V L K L I S
27901 - CTGTTGTCCAATGTTCCAAGCACCTTTTACGGGCTTCCCCTGGTAACTTTATAGTTACC -
27960
- L L S N V P S T F Y G L S L G N F I V T
- C C P M F Q A P F T G F P L V T L * L P
- V V Q C S K H L L R A F P W * L Y S Y R
27961 - GCAGGACTCAACAATGGTTTGAAGACTTGTATCAAGACTCTTTATAGTGTCAATAAA -
28020
- A G L N N G F E R L V I K T L Y S V N K
- Q D S T M V L K D L * S R L F I V S I K
- R T Q Q W F * K T C N Q D S L * C Q * R
28021 - GGCACCTGTAGAAGCAGAGAAAGATGCCAAAATGATGGCAACCTCTTCATTCAAATGAAA -
28080
- G T C R S R E R C Q N D G N L F I Q M K
- A L V E A E K D A K M M A T S S F K * K
- H L * K Q R K M P K * W Q P L H S N E N
28081 - ATCGCCAACAATGTTAATGTTAACGTTTACGACTCAGTATCTCAAGGAGATCCTCATT -
28140
- I A N N V N V N T F T T Q Y L K E I L I
- S P T M L M L T R S R L S I S R R S S F
- R Q Q C * C * H V H D S V S Q G D P H S
28141 - CAAGTCTCCACATTGTCACCAGTAATGCCAGTATGGCCTGAGCCAATATCAGCACTAGC -
28200
- Q G L H I V T S N A S M A * A N I S T S
- K V S T L S P V M P V W P E P I S A L A
- R S P H C H Q * C Q Y G L S Q Y Q H * H
28201 - ACGAGGAACCCAGTAGGCACGCTTATATAGCAGCCAACATAGGCAACACACAGCCTCC -

```

图 17-LL

28260

- T R N P V G T L I I A A N I G K H T A S  
- R G T Q \* A R L L \* Q P T \* A N T Q P P  
- E E P S R H A Y Y S S Q H R Q T H S L Q

28261 - AAAACATCTAGTCCTACCTCCCTTGGCGAGTCGAGTTTCAATGTTTGAGTGGTTGTGATA -  
28320 - K T S S P T S L A E S S F N V \* V V V I  
- K H L V L P P L R S R V S M F E W L \* \*  
- N I \* S Y L P C G V E F Q C L S G C D N  
28321 - ATCTGCAACACTATGCTCAGGTCCAATCTCTGGGTCTTGACAGGCAGGACATGGCATTTT -  
28380 - I C N T M L R S N L W V L T G R T W H F  
- S A T L C S G P I S G S \* Q A G H G I F  
- L Q H Y A Q V Q S L G L D R Q D M A F S  
28381 - CACTACAGCATTAGTAGGTAGGTACCCACATGTAGTAGGTCTTCAATAACTAAATTTTC -  
28440 - H Y S I S R \* V P T C S R S F N N \* I F  
- T T A L V G R Y P H V V G P S I T K P S  
- L Q H \* \* V G T H M \* \* V L Q \* L N F Q  
28441 - AGTGCCACAATGTTCAAGAAGTGGCTTTCAGAAAGTCGCACGTCTGCCATGAAACTTCATC -  
28500 - S A T M F T S G F Q K V A R L P \* N F I  
- V P Q C S Q V A F R K S H V C H E T S S  
- C H N V H K W L S E S R T S A M K L H R  
28501 - GCAATGATTACATTTTCATCAAGGTAGACAAGTGCATATTGTTACTCCTGTGGAGATGC -  
28560 - A M I T F H Q G R Q V H I V T L L W R C  
- Q \* L H F I K V D K C I L L H S C G D A  
- N D Y I S S R \* T S A Y C Y T P V E M Q  
28561 - AACAGGGTACACAGAGCGTATACGCCCCATGAAACCCTCAGTCTTTTTCTTTTCAACAGC -  
28620 - N R V H R A Y T P H E T L S L F L F N T  
- T G Y T E R I R P M K P S V F F F S T R  
- Q G T Q S V Y A P \* N P Q S F S F Q H V  
28621 - TGGTGAATGACTTTGACTTTTGAGTTAAGAGGAAACACAAACTTTGGGCATTCCCTTT -  
28680 - W L N D F D F \* V K R K H K L W A F P F  
- G \* M T L T F E L R G N T N F G H S P L  
- V E \* L \* L L S \* E E T Q T L G I P L \*  
28681 - GAAAGTGTCAAATTTCTTGGCACTCTTAATTTGAAAGGGTGTCTGGTCTCGTAGCTCTT -  
28740 - E S V K F L G T L N F E G C L V L V A L  
- K V S N F L A L L I S K G V W C S \* L L  
- K C Q I S W H S \* F R R V S G A R S S Y  
28741 - ATCAGAGCGCTCAGTGAACCAGGCAATTTTCATGCTCATGGTCACGGCAGCAGTAGACACC -  
28800 - I R A L S E P G N F M L M V T A A V D T  
- S E R S V N Q A I S C S W S R Q Q \* T P  
- Q S A Q \* T R Q F H A H G H G S S R H L  
28801 - TCTCTCGACTCGATGTAATCAAGTTGTTTCGAAAGAGTGCACATTGACTTGCCCGCGG -  
28860 - S L R L D V I K L F G K S A H \* L A R A  
- L F D S M \* S S C S E R V H I D L P A R  
- S T R C N Q V V R K E C T L T C P R V  
28861 - TGCGAGAAATCTTTGATGCAATCAAGAGGGTACCCATCTGGGCCACAGAAATTGTTGTC -  
28920 - C E K I F D A I K R V P I W A T E I V V  
- A R K S L M Q S R G Y P S G P Q K L L S  
- R E N L \* C N Q E G T H L G H R N C C R  
28921 - GACATAGCGAGTACTGCACCTCCATTGAGCTCAGAGTGAGTTCACGGAGTGCACCACT -

图 17-MM

28980 .  
- D I A S D C T S I E L T S E F T E C T T  
- T \* R V T A P P L S S R V S S R S A P L  
- H S E \* L H L H \* A H E \* V H G V H H C  
28981 - GCCATGCTTAGTGTCCAGTTTTGTTTCATAATCTTCAATGGGATCAGTGCCAAGCTCGTC -  
29040  
- A M L S V P V L F I I F N G I S A K L V  
- P C L V F Q F C S \* S S M G S V P S S S  
- H A \* C S S F V H N L Q W D Q C Q A R H

```

29041 - ACCTAAGTCATAAGACTTTAGATCGATGCCATAGCTATGACCACCGGCTCCCTTATTACC -
29100
- T * V I R L * I D A I A M T T G S L I T
- P K S * D F R S M P * L * P P A P L L P
- L S H K T L D R C H S Y D H R L P Y Y R
29101 - GTTCTTACGAAGAAGAACATTGCGGTATGCAATTGGGGTTTCGCCACATGTGGCACCAG -
29160
- V L T K K N I A V C N W G F A H M W H E
- F L R R R T L R Y A I G V S P T C G T S
- S Y E E E H C G M Q L G F R P H V A R V
29161 - TACTCCCAGTGTATACCGCTACGACCGTACTGAATGCCGTCATTTCTGCAACCAGCTC -
29220
- Y S Q C Y T A T T V L N A V H F C N Q L
- T P S V I P L R P Y * M P S I S A T S S
- L P V L Y R Y D R T E C R P F L Q P A Q
29221 - AACGACCTTGTGGCCGTGATTGGTGCTTAAGGCATCAGAACGTTTAAATGAACACATAGGG -
29280
- N D L V A V I G A * G I R T F N E H I G
- T T L W P * L V L K A S E R L M N T * G
- R P C G R D W C L R H Q N V * * T H R A
29281 - CTGTTCAAGCTGGGGCAGTACGCCCTTTTTCCAGCTCTACTAGACCACAAGTGCCATTTTT -
29340
- L F K L G Q Y A F F Q L Y * T T S A I F
- C S S W G S T P F S S S T R P Q V P F L
- V Q A G A V R L F P A L L D H K C H F *
29341 - GAGGTGTTACCGTCCGCTCCGATAGGGCCTCTTCCACAGAGTCCCCGAAGCCAGCCACTAG -
29400
- E V F T C L R * G L F H R V P E A T H *
- R C S R A S D R A S S T E S P K P R T S
- G V H V P P I G P L P Q S P R S H A L A
29401 - CACGTCTTAACCTGAAGGACAGGCAAACCTGAGTTGGACGTGTGTTTTCTCGTTGACACC -
29460
- H V S N L K D R Q T E L D V C F L V D T
- T S L T * R T G K L S W T C V F S L T P
- R L * P E G Q A N * V G R V F S R * H Q
29461 - AAGAACAAGGCTCTCCATCTTACCTTTTCGGTCACACCCGGACGAAACCTAGGTATGCTGA -
29520
- K N K A L H L T F R S H P D E T * V C *
- R T R L S I L P F G H T R T K P R Y A D
- E Q G S P S Y L S V T P G R N L G M L M
29521 - TGATCGACTGCAACACGGACGAAACCGTAAGCAGTCTGCAGAAGAGGGACGAGTTACTCG -
29580
- * S T A T R T K P * A V C R R G T S Y S
- D R L Q H G R N R K Q S A E E G R V T R
- I D C N T D E T V S S L Q K R D E L L V
29581 - TTTCTTGTCAACGACAGTAAATTTATTATTGTTTATACTGCGTAGGTGCACTAGGCATG -
29640
- F L V N D S K I Y Y C L Y C V G A L G M
- F L S T T V K F I I V Y T A * V H * A C
- S C Q R Q * N L L L F I L R R C T R H A
29641 - CAGCCGAGCGACGCTACACAGATTTTAAAGTTTCGTTTAGAGAACAGATCTACAAGAGAT -
29700
- Q P S D S Y T D F K V R L E N R S T R D
- S R A T A T Q I L K F V * R T D L Q E I
- A E R Q L H R F * S S F R E Q I Y K R S

```

图 17-NN

29701 - CGAGGTTGGTTGGCTTTTCCTGGGTAGGTAAAAACCTAATAT - 29742  
- R G W L A F P G \* V K T \* Y X  
- E V G W L F L G R \* K P N X  
- R L V G F S W V G K N L I X

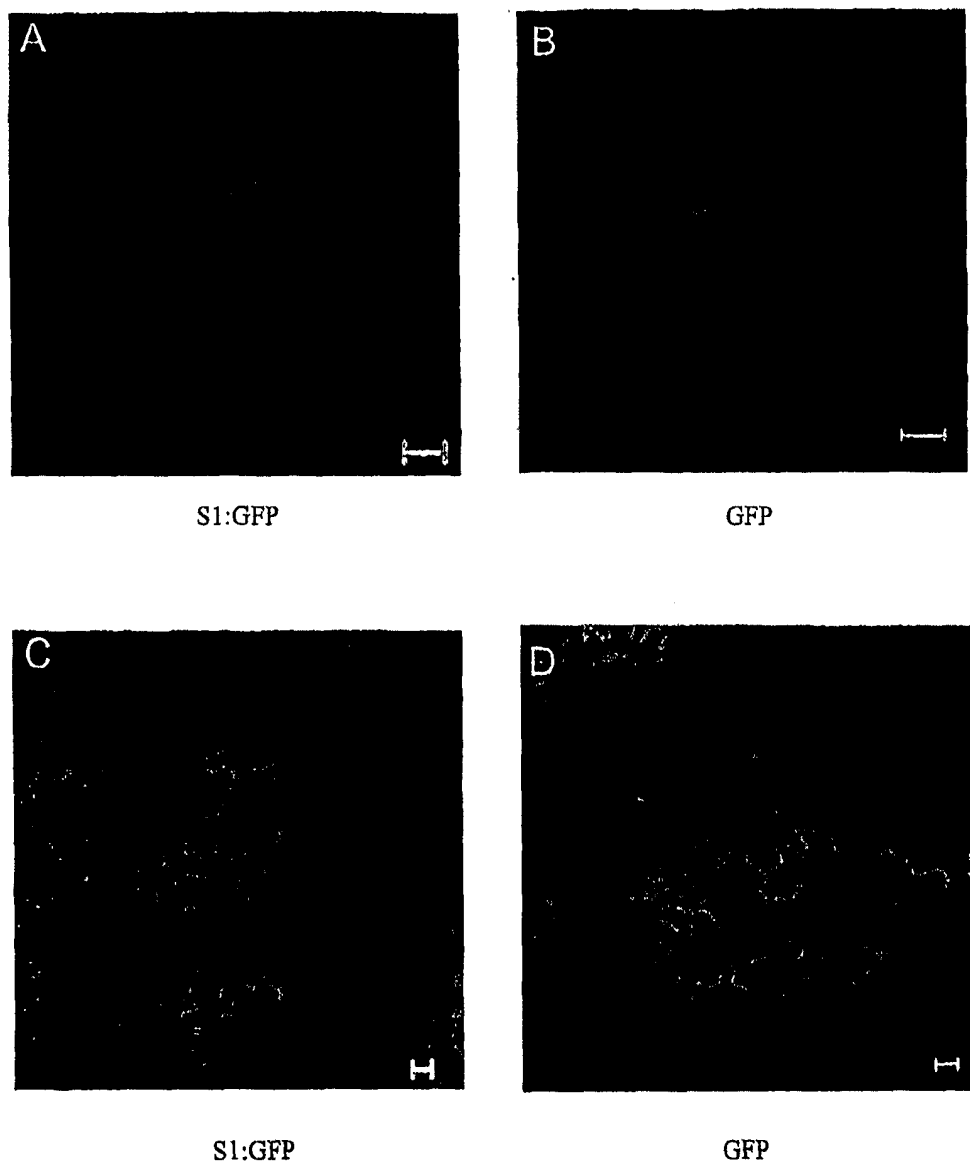


图 18

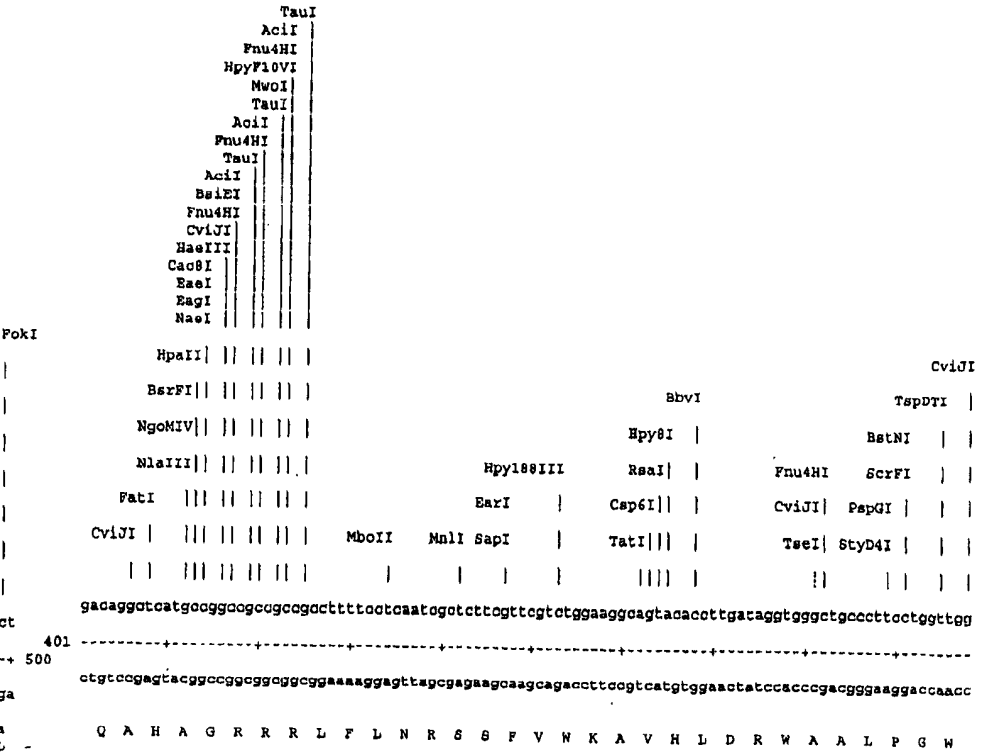
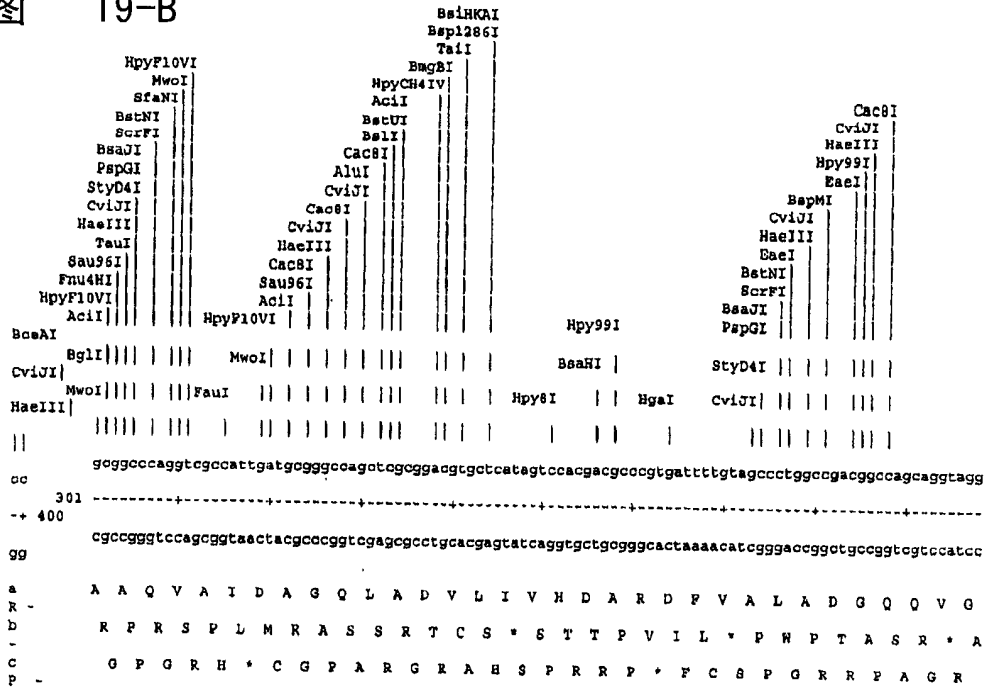




```

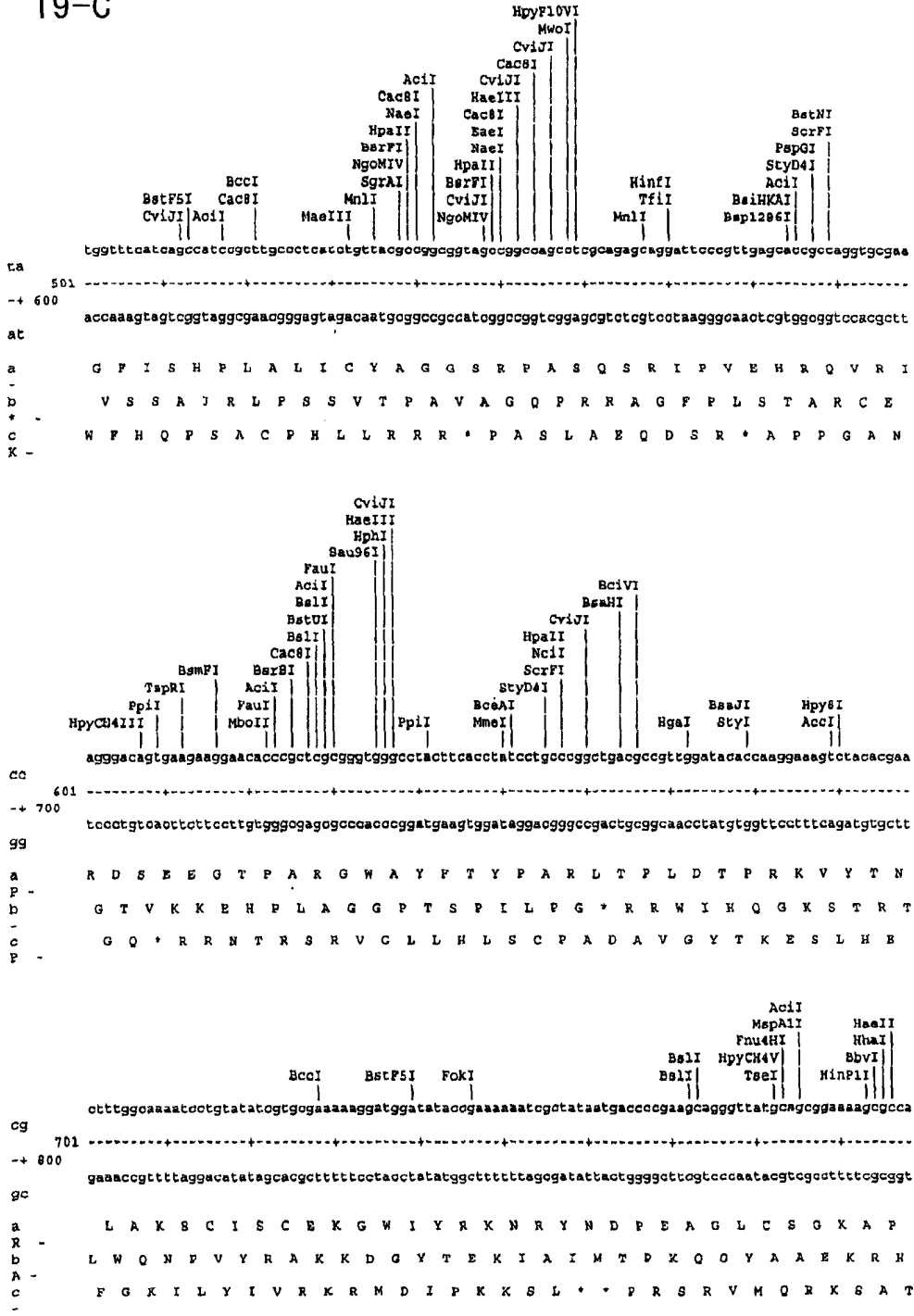
      BstKTI              FauI  | |                TauI
      DpnI|  BpyF10VI              HhaI  | |                AclI  ||
      EarI||      BglI|  BstKTI          NotI  | |                Pvu4HI  ||
      MboI|||      MwoI|  HphI              HinP1I | |                CviJI  ||
      TagI|||      MnlI ||  DpnI| AlwI          AclI  | |                CviJI  HaeIII  ||
MboII  ||| |  CacSI | |  MboI | BceAI  SfaNI | | MwoI |  Hpy99I  |  HphI  CacSI  || |
      | || | |  || | |  || |  | || | | | |  |  |  |  || | |
ag gcttctcttcgatcttcgccagcaggcgaggatcgtggcatcaccgaaccgcgccgtgcgcgggtcgtcggtgagccagagtttcagcagggcggccc
201 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
-+ 300 ogaagagaagotagaaggggtcgtcccgotootagoaccgtagtgggtggcgggcaagccagccagccactoggtotcaagtcgctccggcggg
tc
a   L L F D L R Q Q G E D R G I T E F R R A R V V G E P E F Q Q A A Q
b   F S S I F A S R A R I V A S P M R A V R G S S V S Q S F S R P P
R - A S L R S S P A G R G S W H H H T A P C A G R R * A R V S A G R P
c
g -
```

图 19-B



b  
L-  
c  
-  
D R L M P A A A A P S S I A L R S S G R Q Y T L I G G L P F L V G  
T G S C R P P P P P Q S L F V R L E G S T P \* \* V G C P S W L A

图 19-C

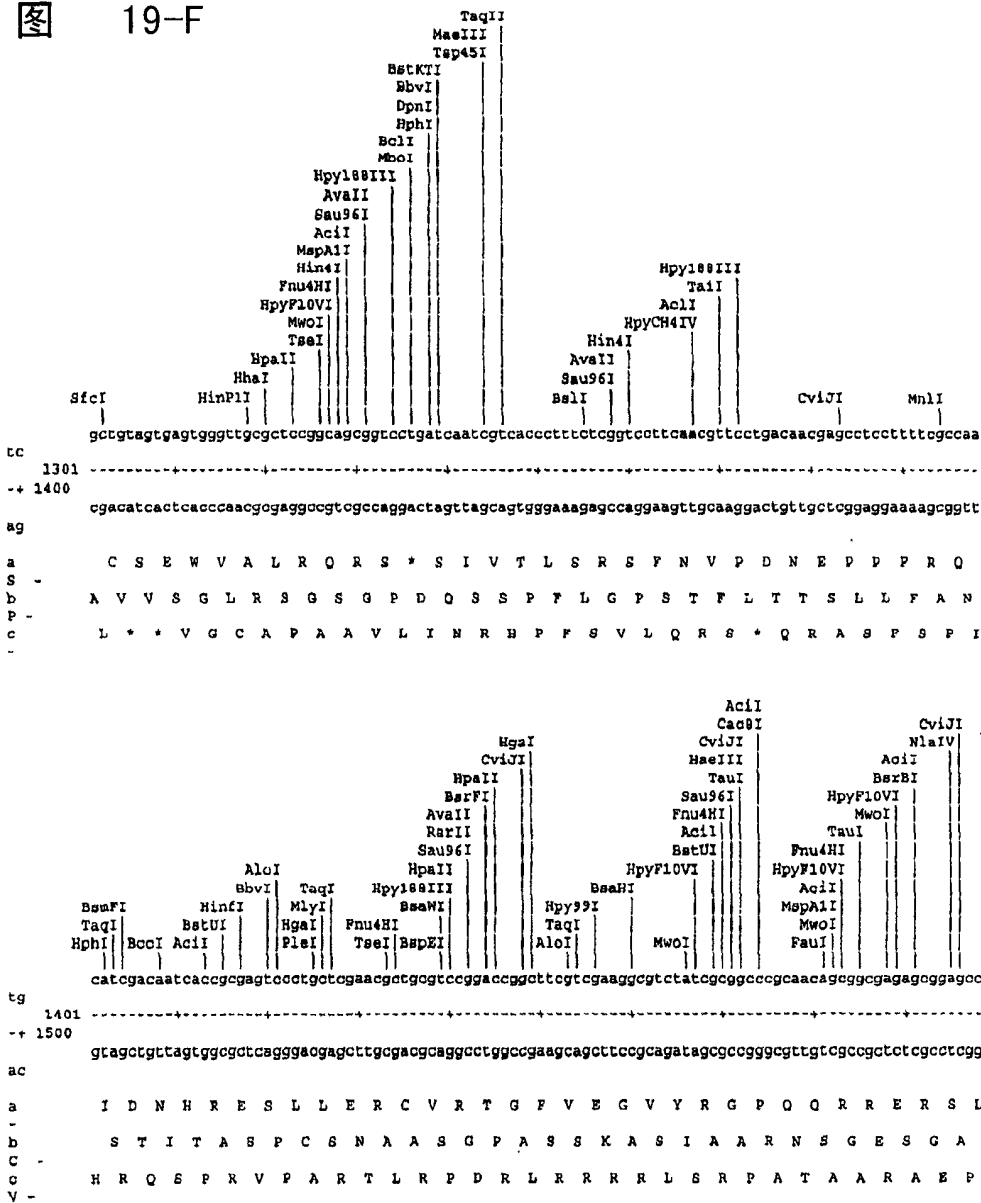








19-F









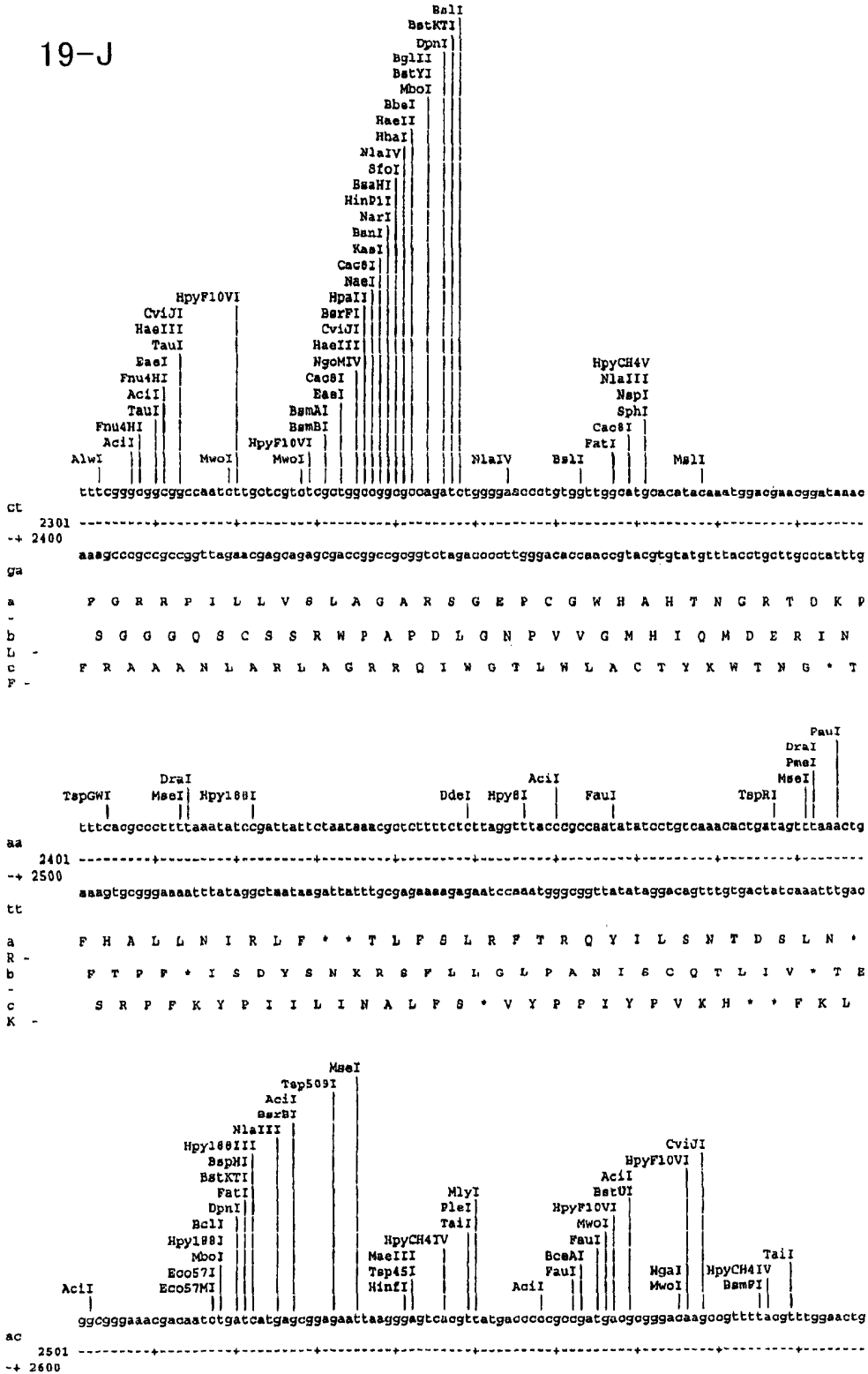






b  
v -  
c  
-  
F T R T T S P P T M A F C W R C M R W C N L P A H L C W A R C R I  
S L E R R K K P L W H S A G A V C V G A I C L R T C A G R A V G S

图 19-J



```
cgcccttgcctgtagactagtagctcctcttaattccctcagtgcaatactggggggcctactgcccctgttggcaaaatgaaaacctgac
tg
A   R E T T I + S + A E N + G S H V M T P A D D A G Q A V L R L E L
T   G G K R Q S D H E R R I X G V T L * P P P M T R D K P F Y V W N +
b   A G N D N L I M S G E L R E S R Y D P R R + R Q T S R F T F G T D
O   -
C   -
```





C I Q I I C T G S G S F R M Y E Q D G L H A G S P A A W V E R L F G  
.



```

aaggagctggctgctattggggaagtgcggggcaggatctcctgtcatctcaccttgcctcctgagagaagtatccatcctggtgatgcaatgc
gg      .
          3101 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
          +- 3200
          |
          | ttcctgaccgacgataaccogettcaaggccccctcctagaggacagtagagtgggaacgaggacggctcttctcataggtagtagtaccgactacggttacg
          |
          | cc
          |
          | a      G T G C Y W A K C R G R I S C H L T L L L P R K Y P S W L M Q C
          | G      -
          | b      K G L A A I G R S A G A C S P V I S P C S C R E S I H K G * C N A
          | A      -
          | c      R D W L L L G E V F G Q D L L S S H L A P A E K V S I M A D A M R
          |
          | -
    
```

图 19-M

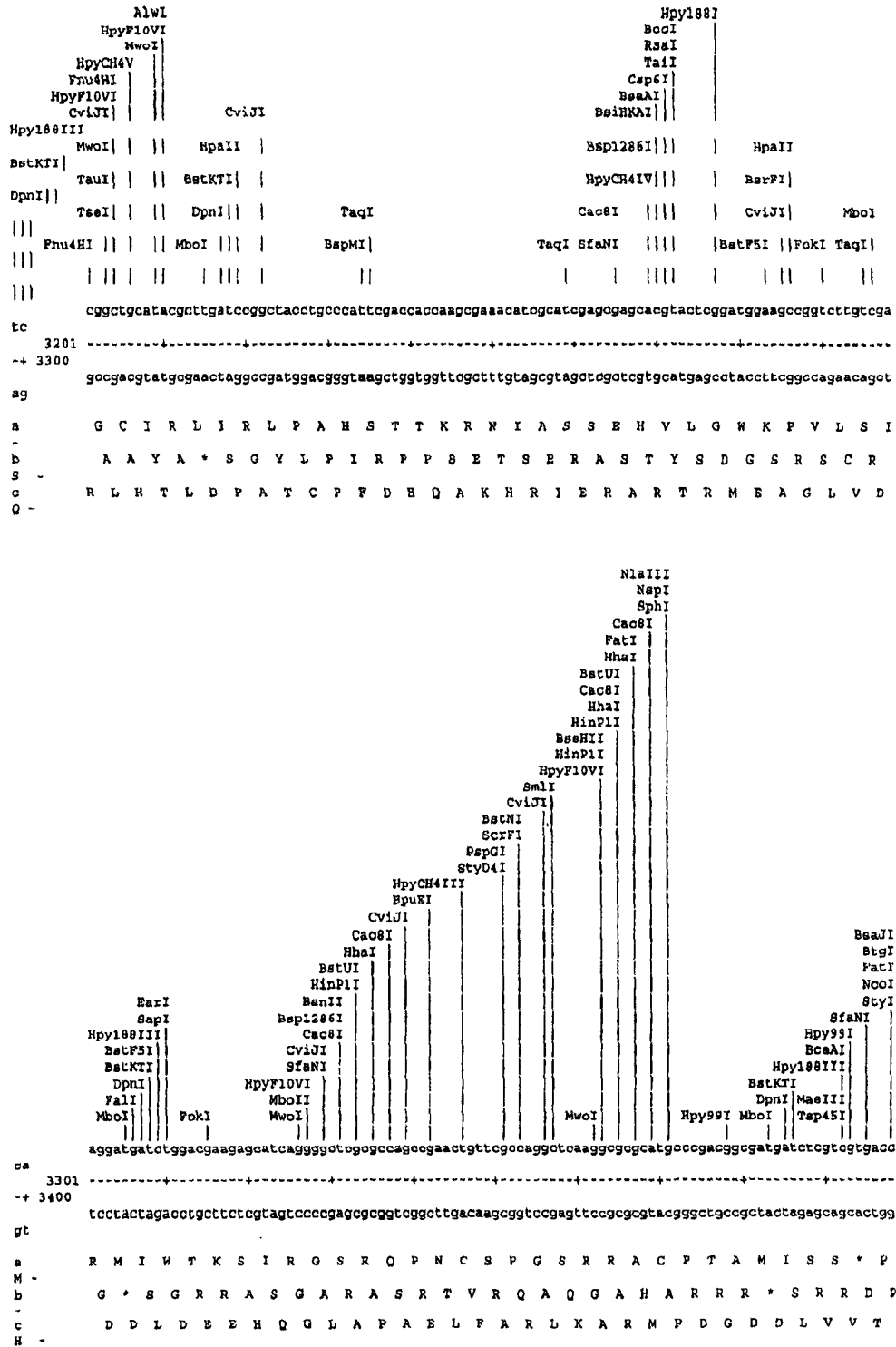


图 19-N

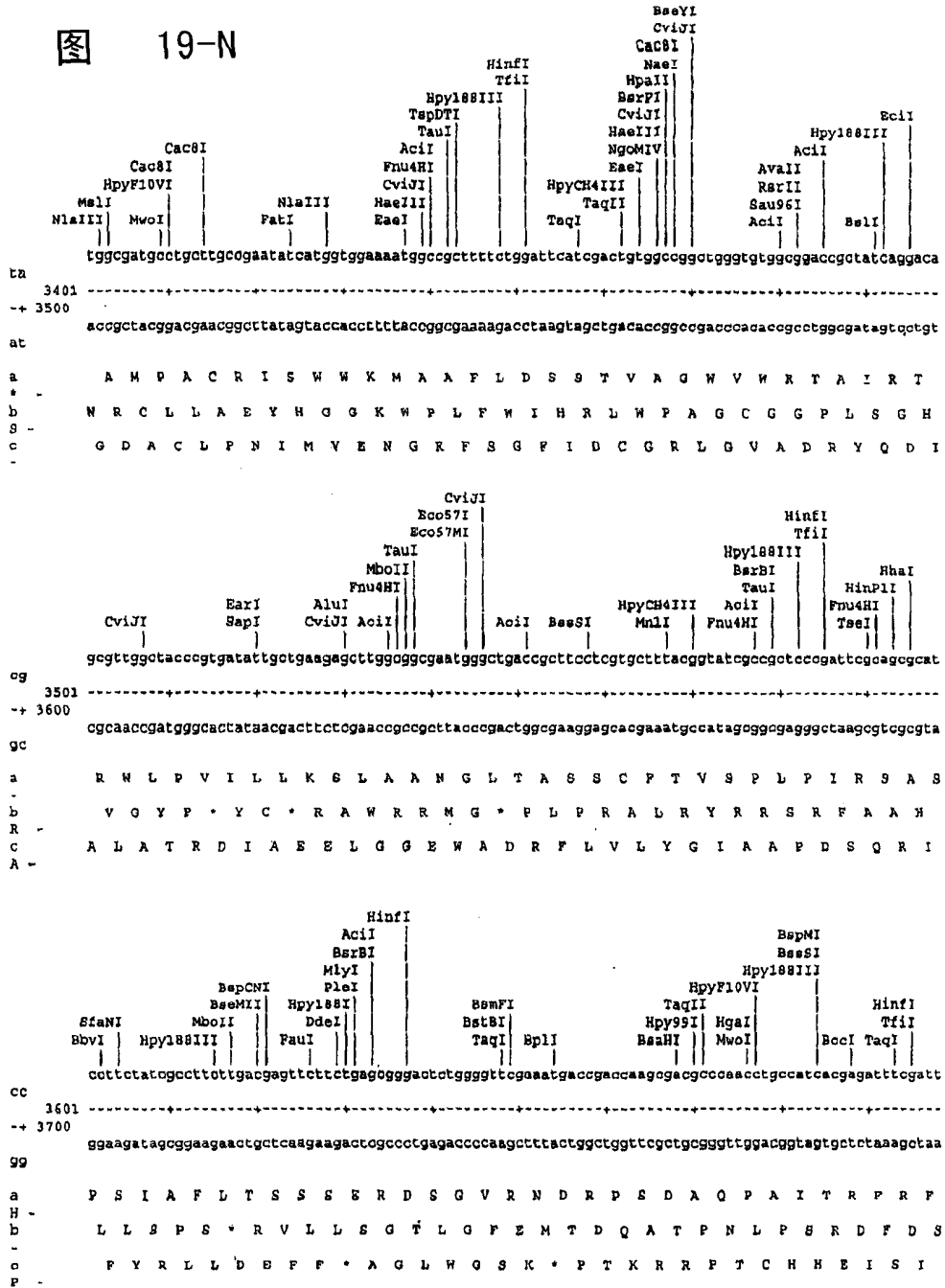
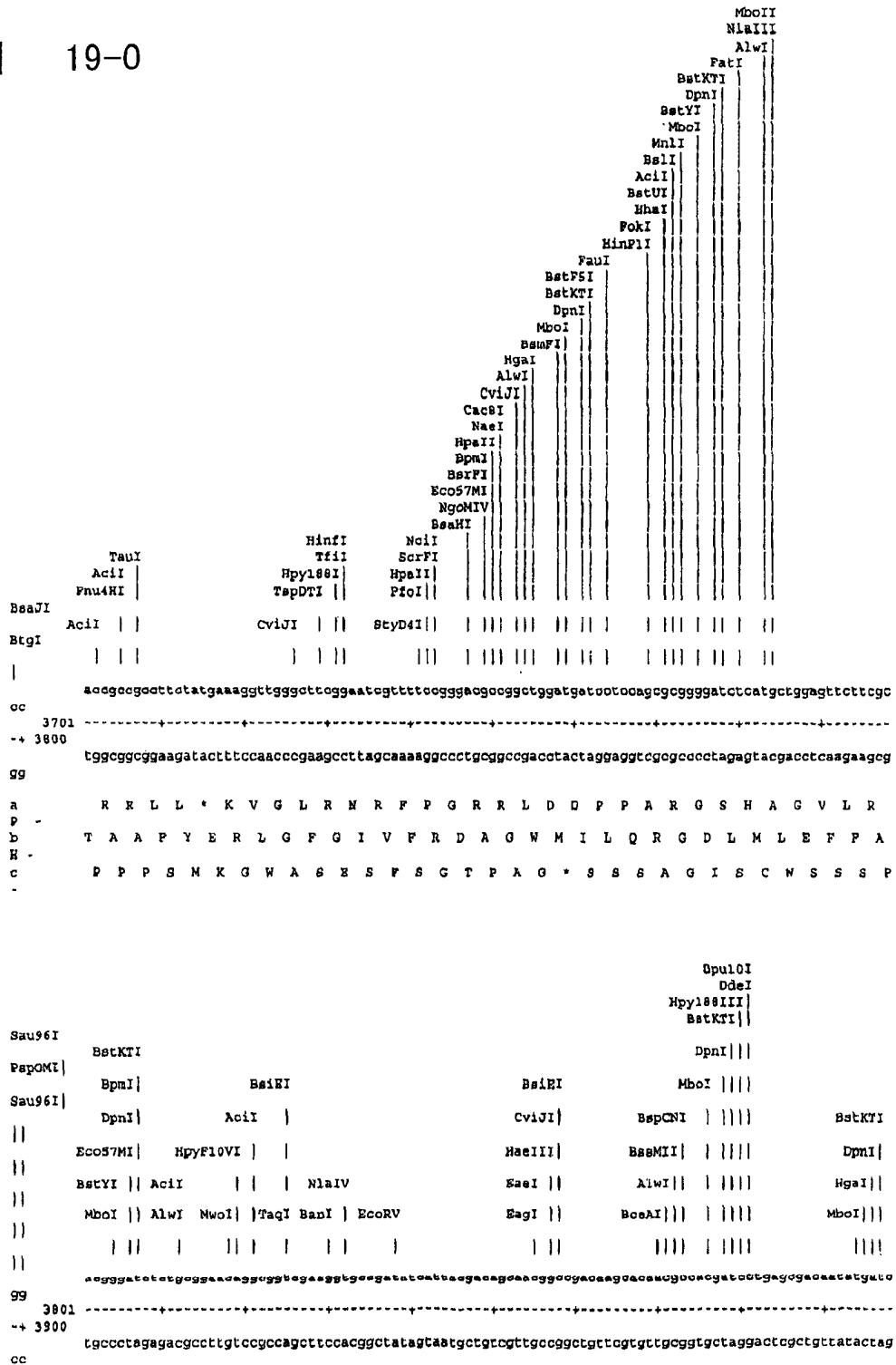


图 19-0



a R D L C G T G G R R C R Y H Y D S N G R Q A Q R H D P E R Q Y D R  
-  
b G I S A E Q A V H G A D I I T T A T A D K H M A T I L S D N M I  
G -  
c T G S L R M R R S K V P I S L R Q Q R P T S T T P R E \* A T I \* S  
G -

图 19-P

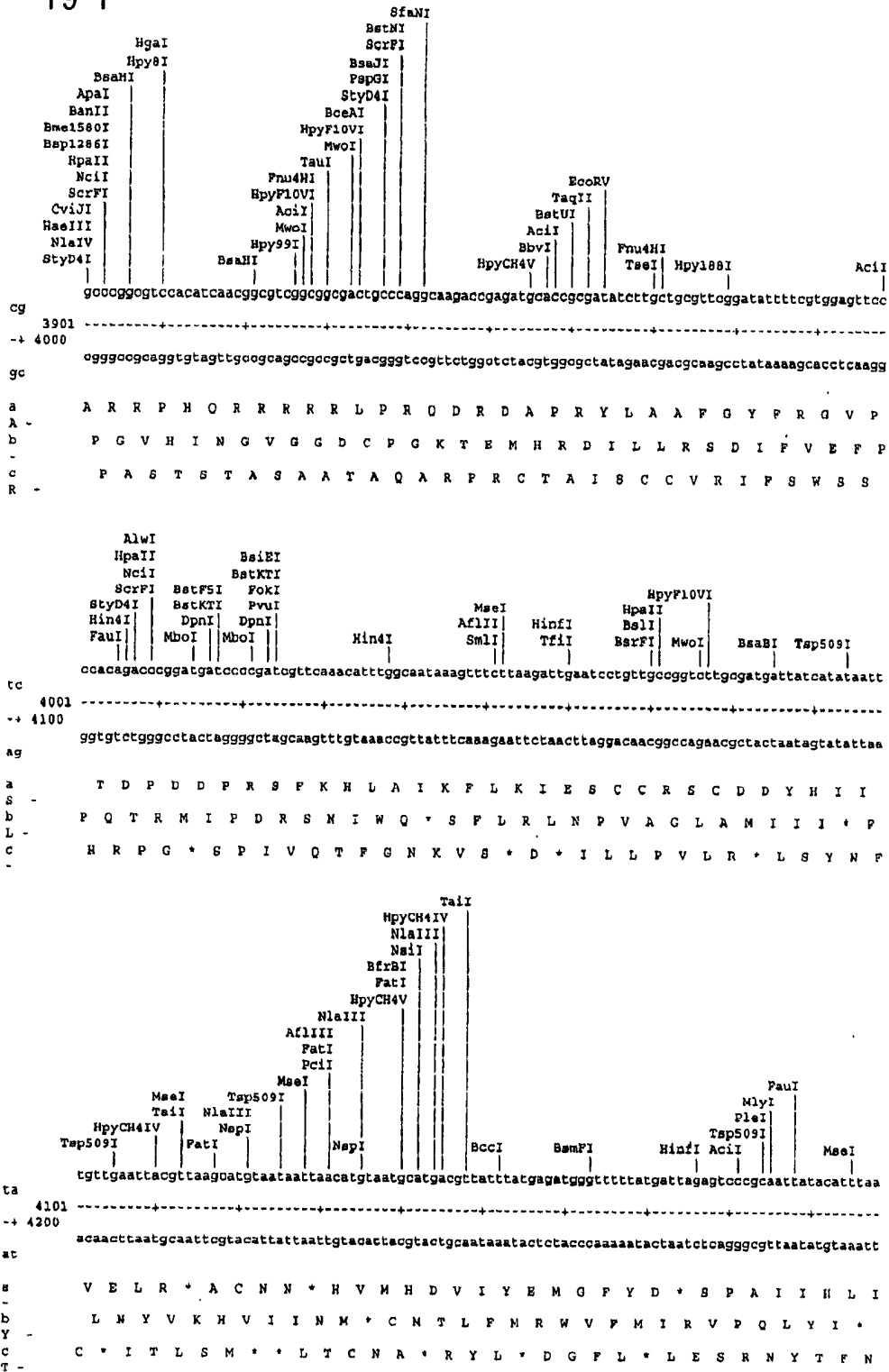
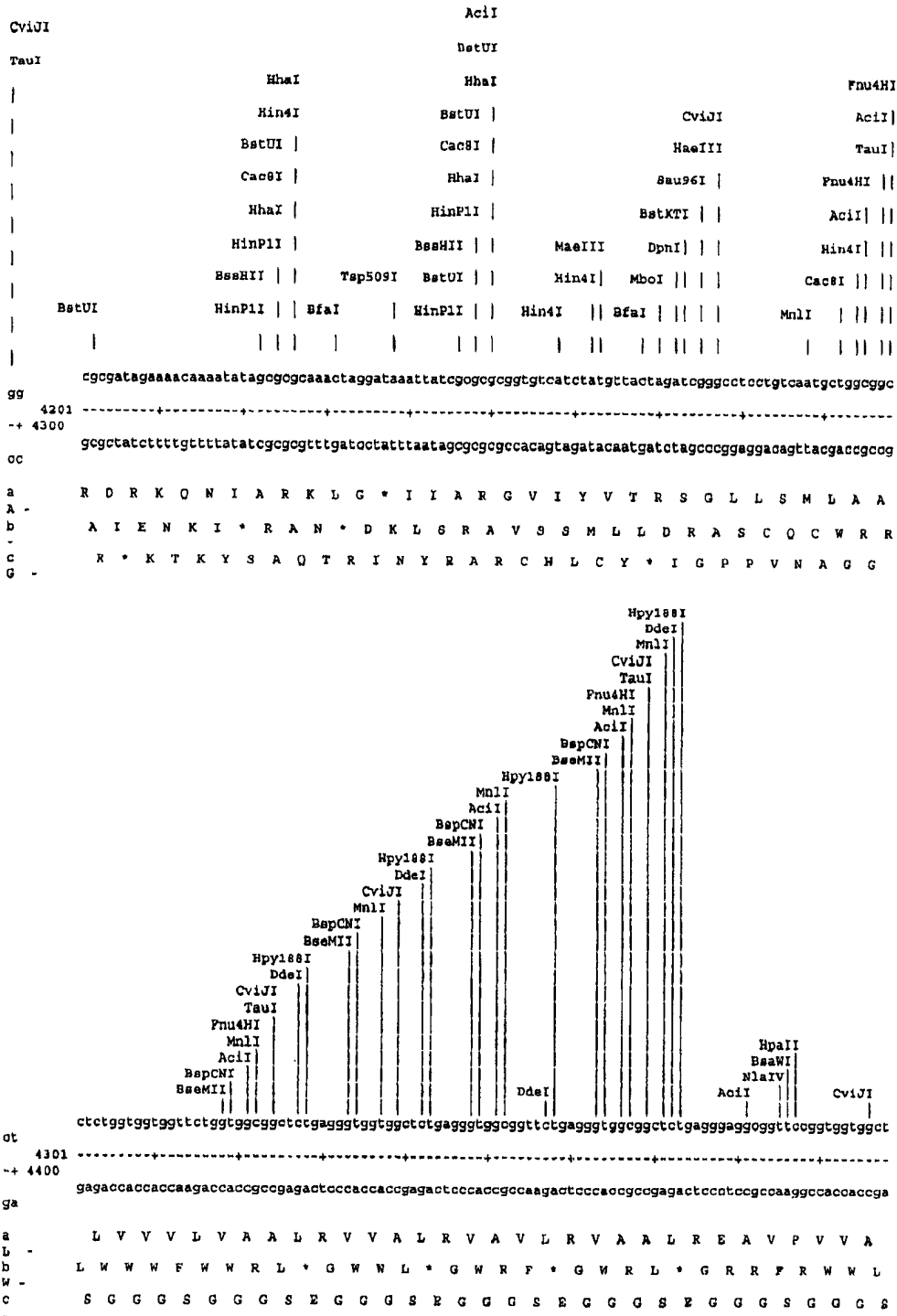




图 19- Q



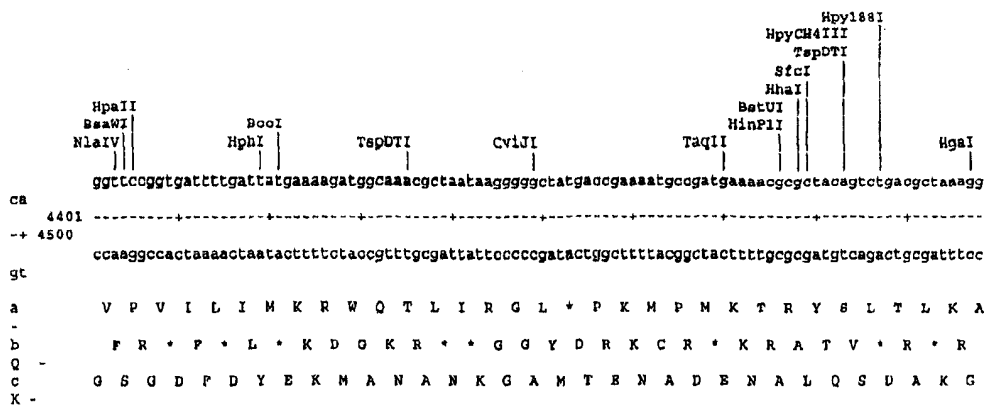
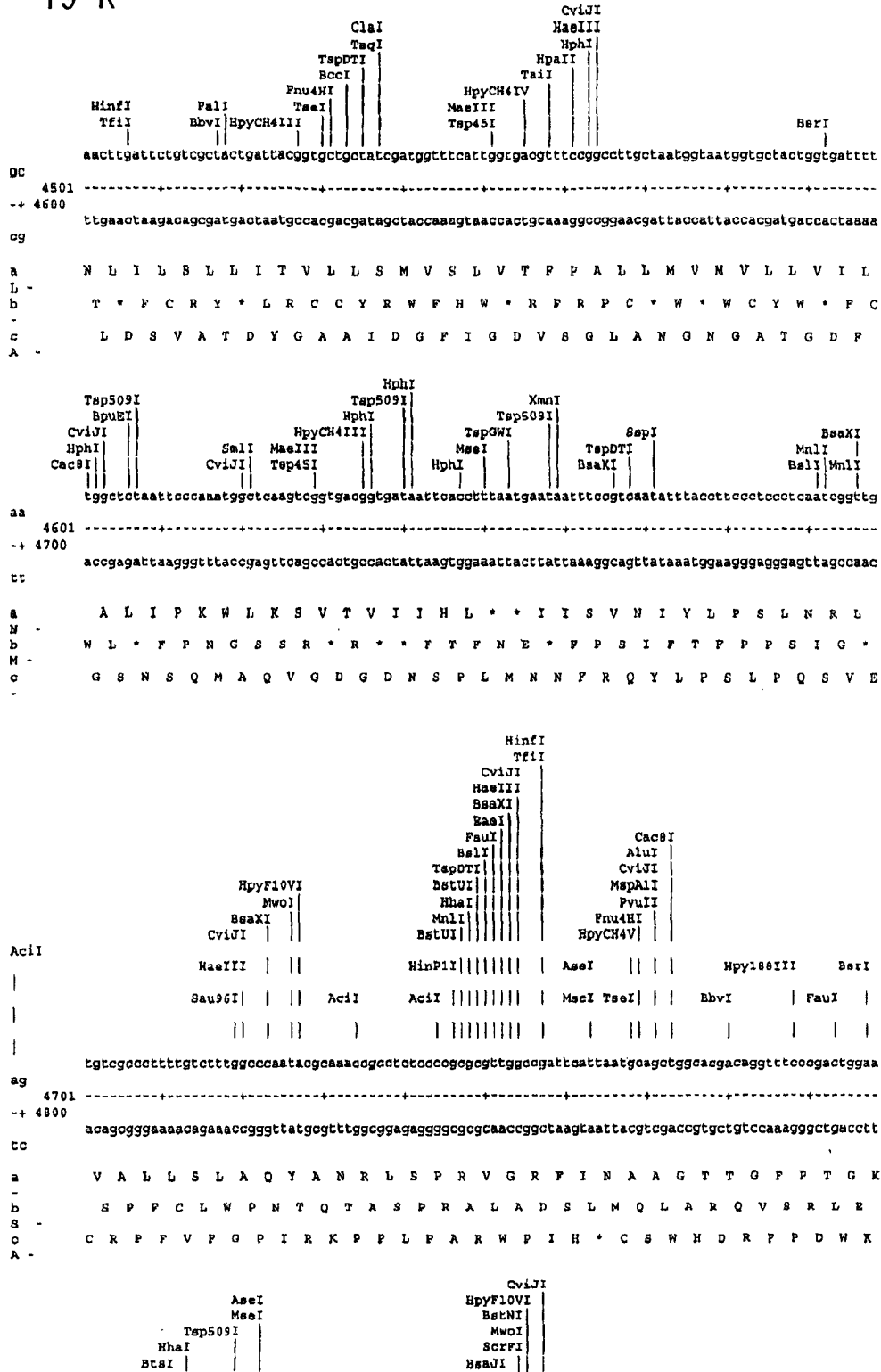
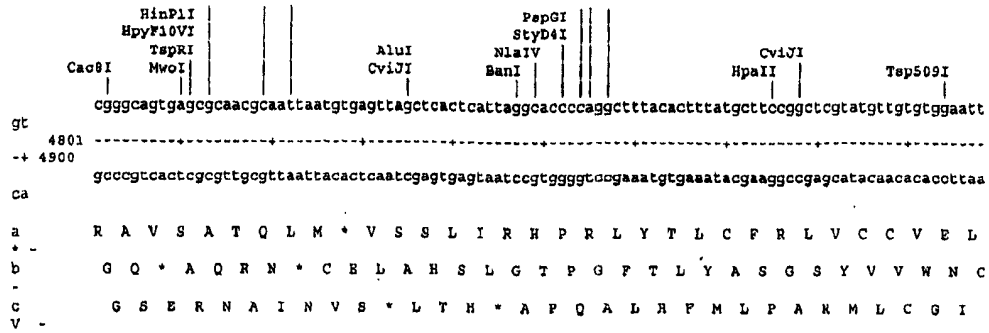


图 19-R







C  
F - L K M Q S K D S G L T A S R T Q R X I Y F S R S E V L P Q Y O R

图 19-T

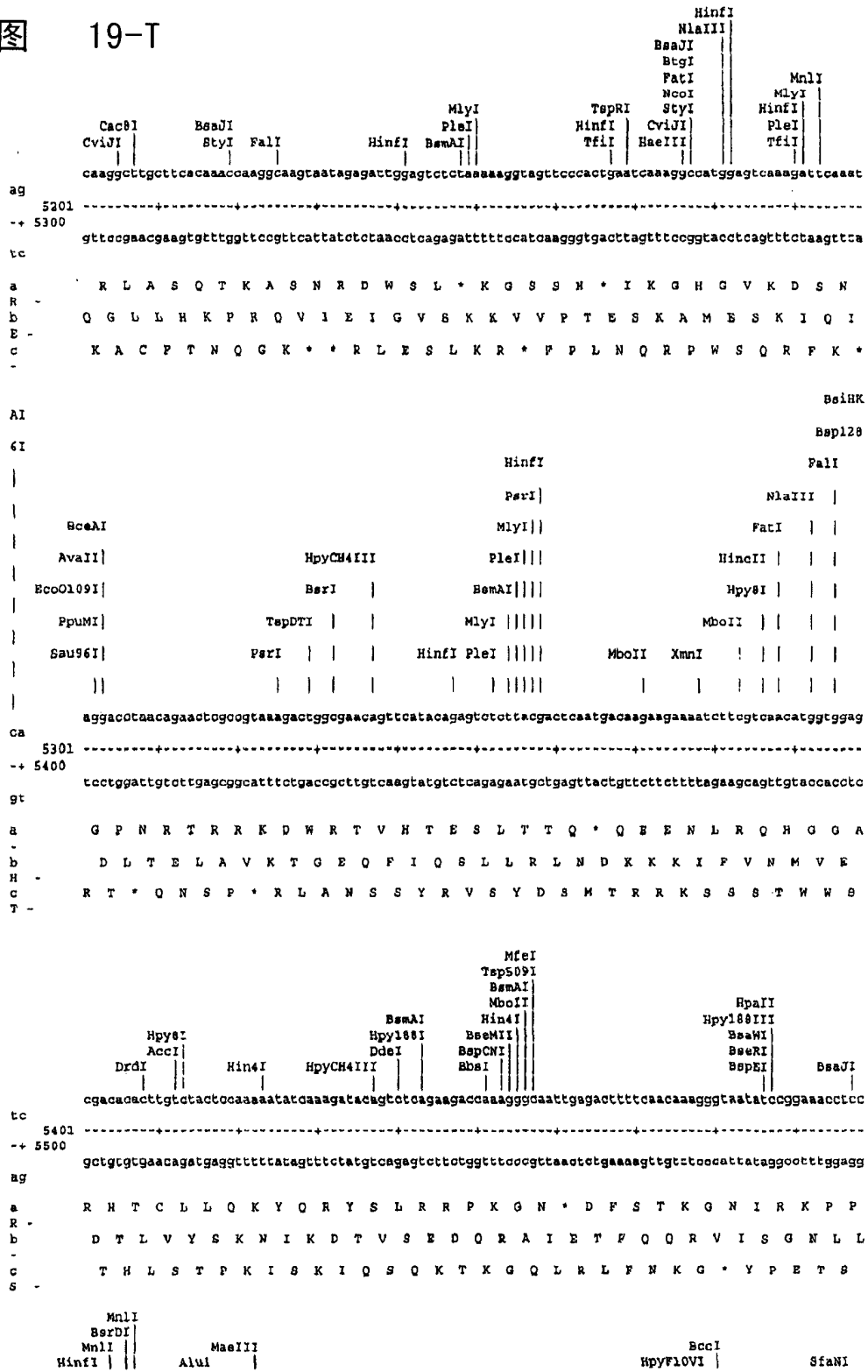
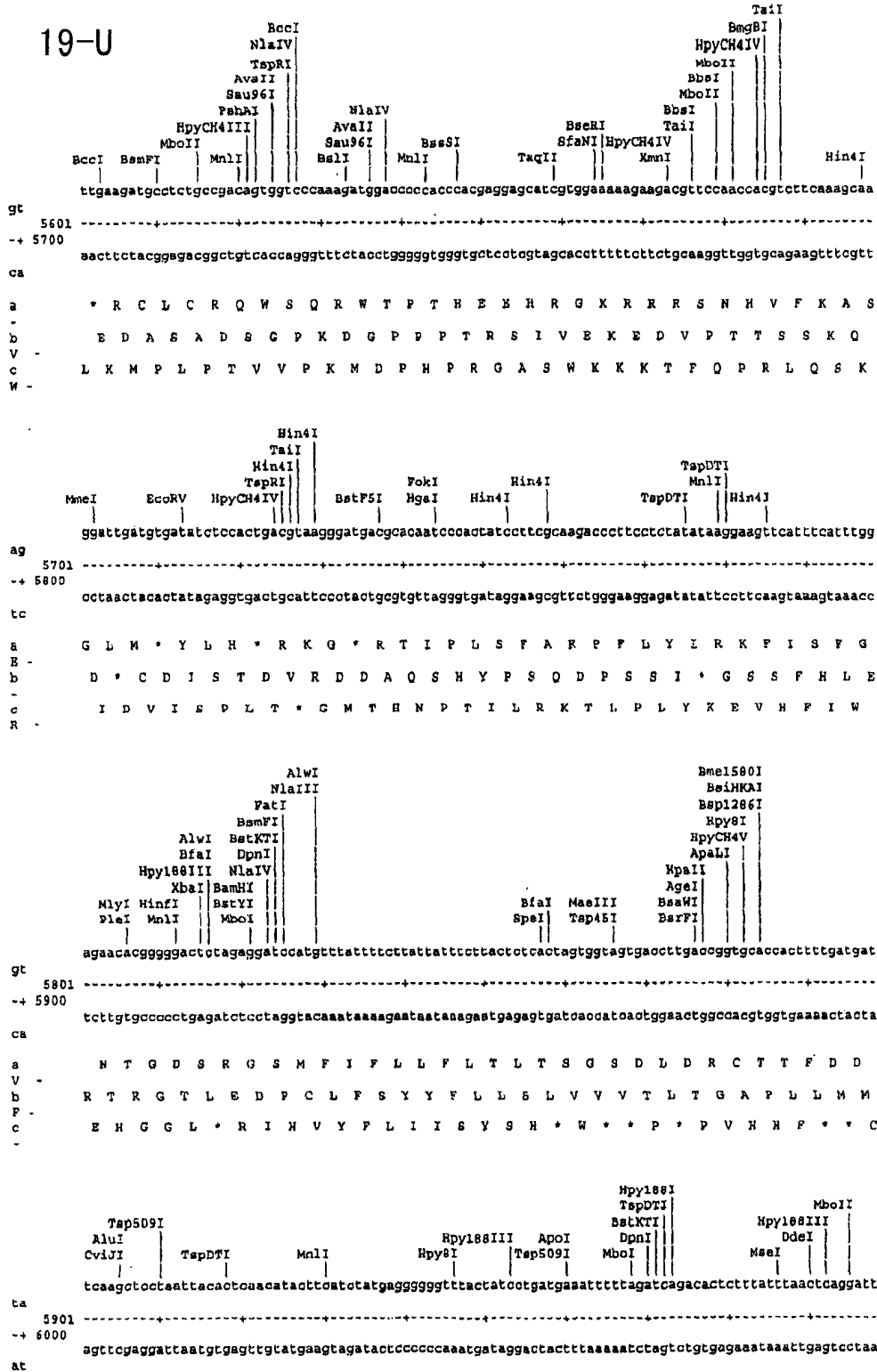






图 19-U



---

a        Q A P N Y T Q H T S S M R G V Y Y P D E I F R S D T L Y L T Q D L  
-  
b        K L L I T L N I L H L \* G G F T I L M K F L D Q T L F I \* L R I  
Y -  
c        S S S \* L H S T Y F I Y E G G L L S \* \* N F \* I R H S L F N S G F  
I -

图 19-V

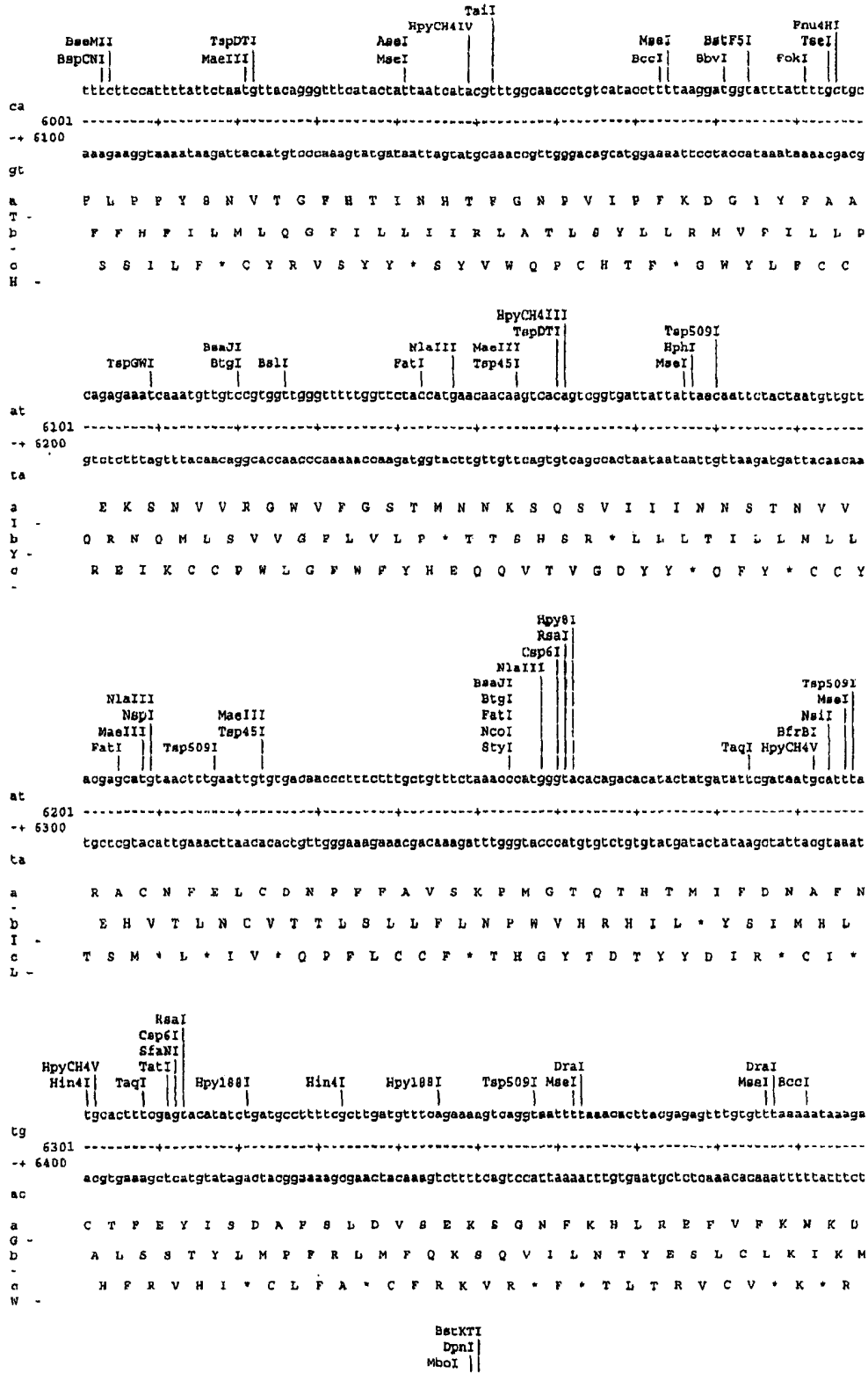




图 19-W

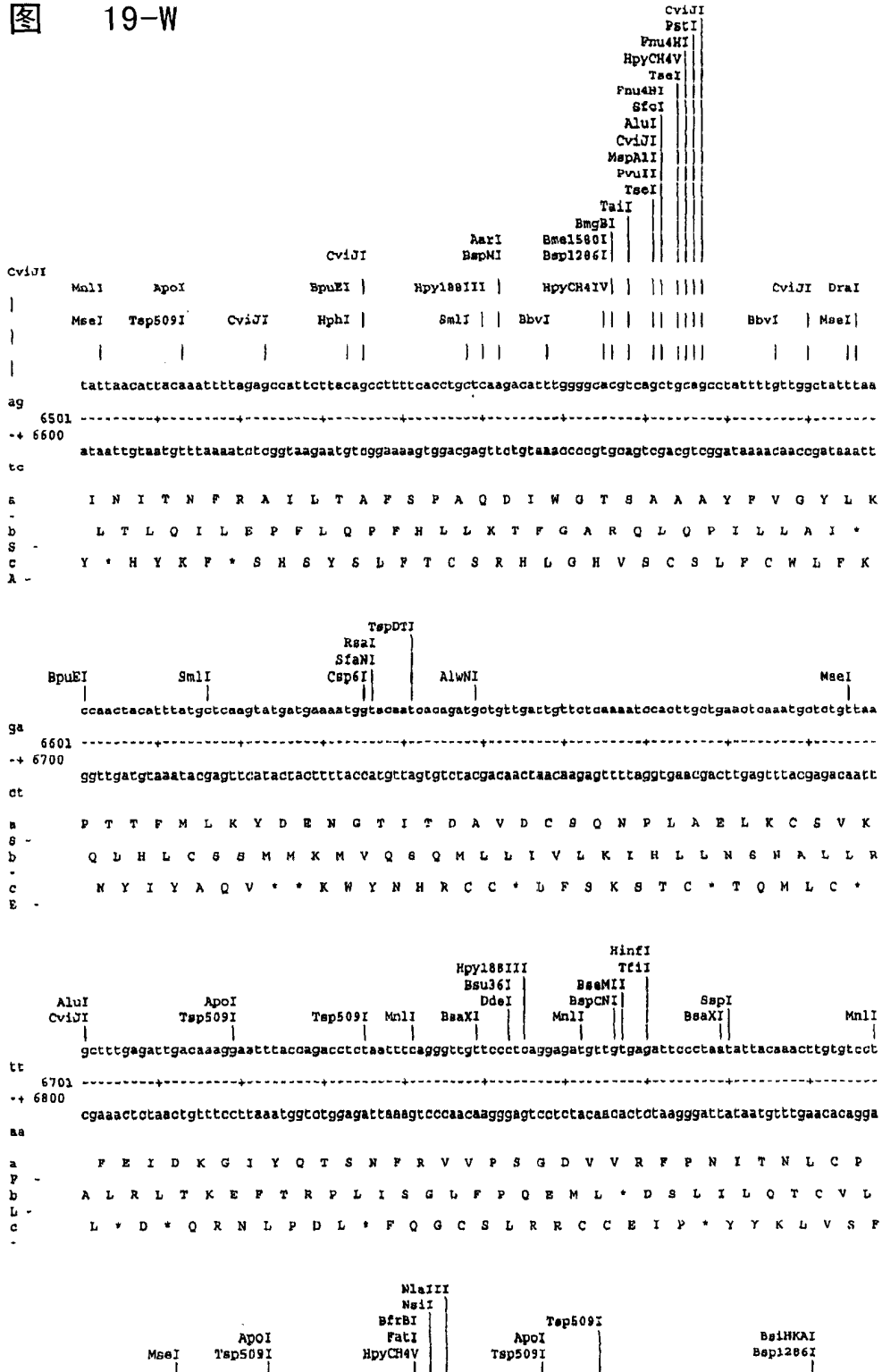
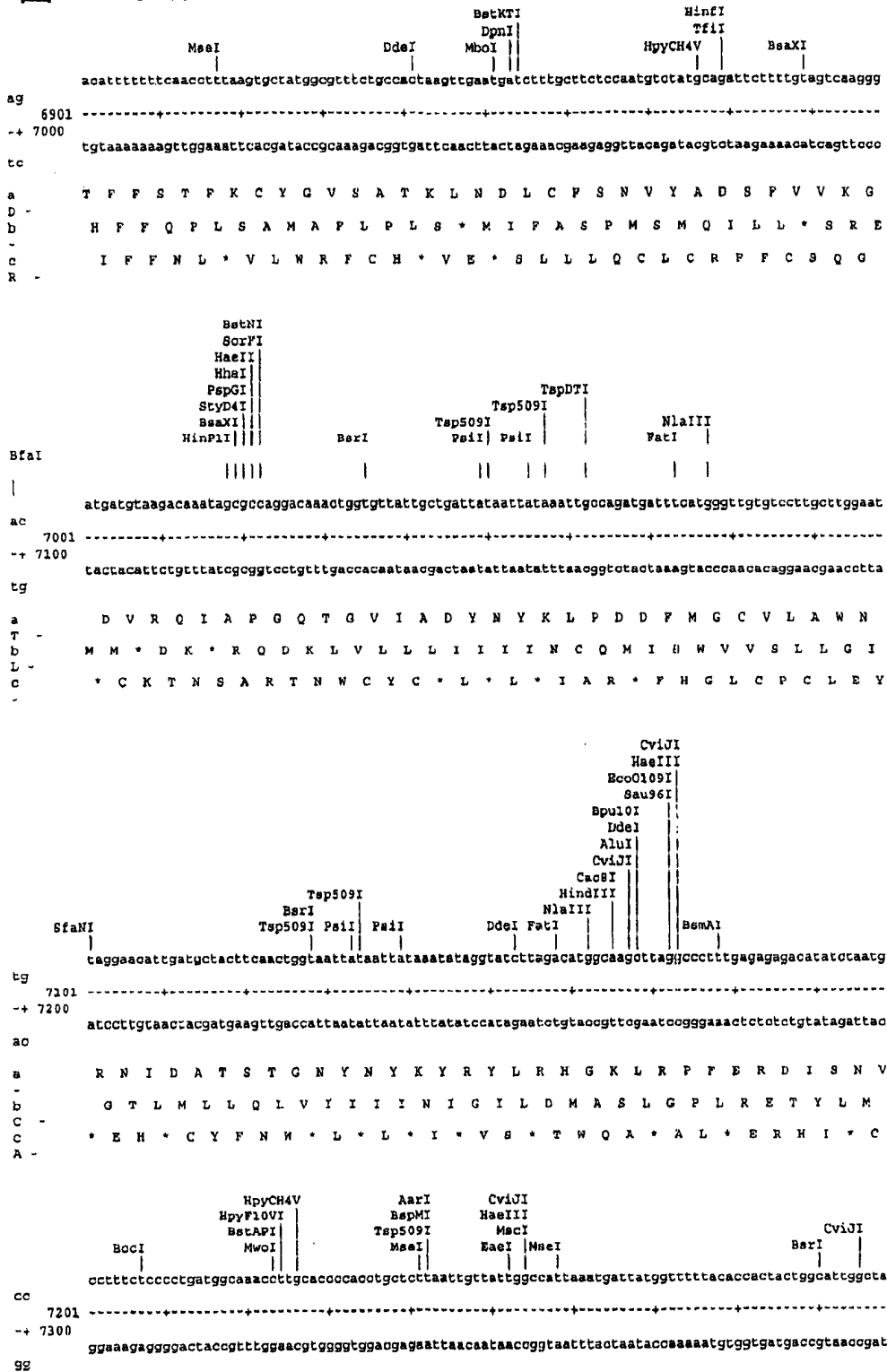




图 19-X



a P F S P D G K E C T P P A L N C Y W P L N D Y G P Y T T G I G Y  
0 -  
b L S P L M A N L A P H L L L I V I G H \* M I M V F T P L L A L A T  
-  
c P L P \* W Q T L H P T C S \* L L L A I K \* L W P L H H Y W H W L  
P -

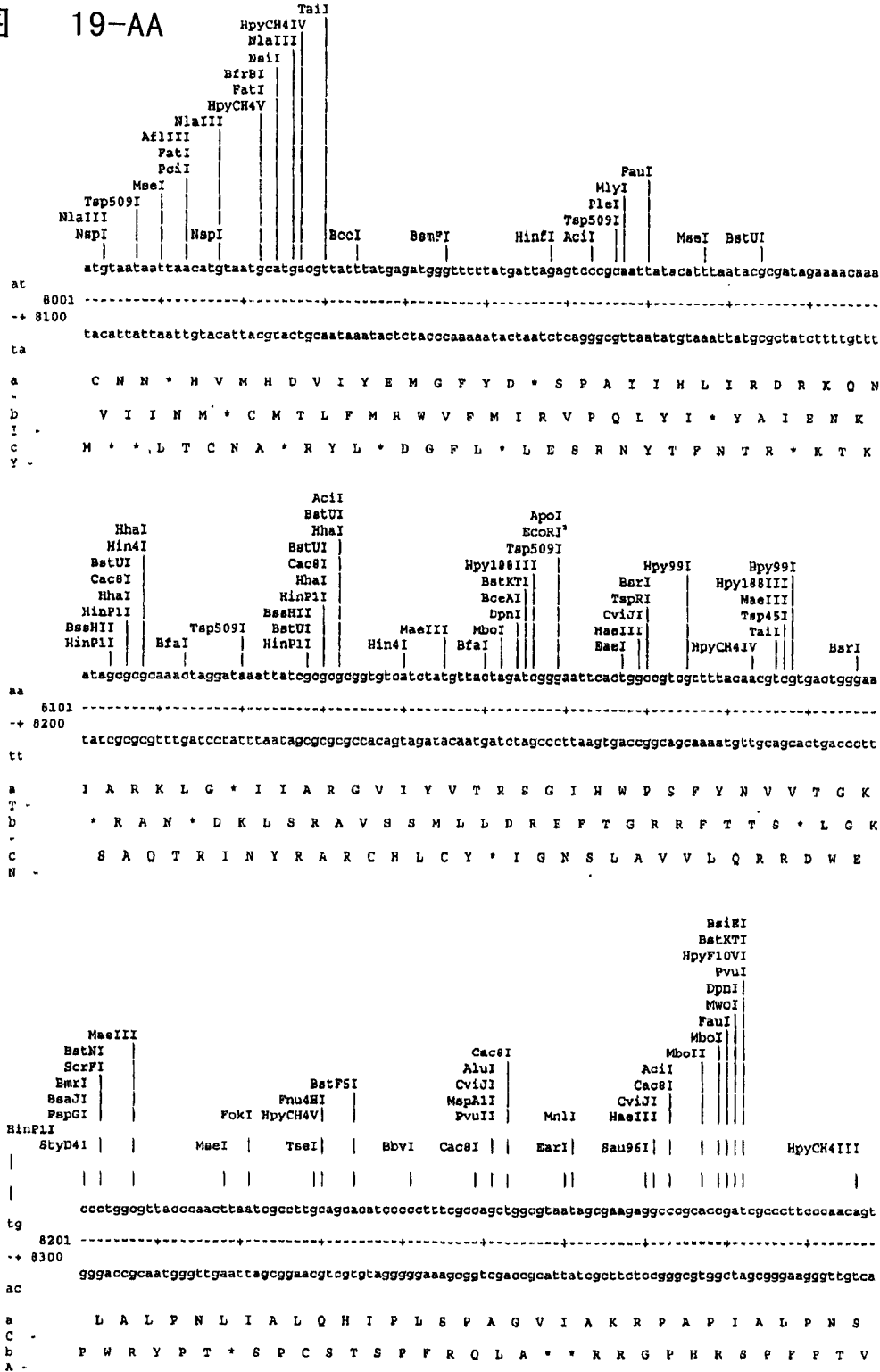




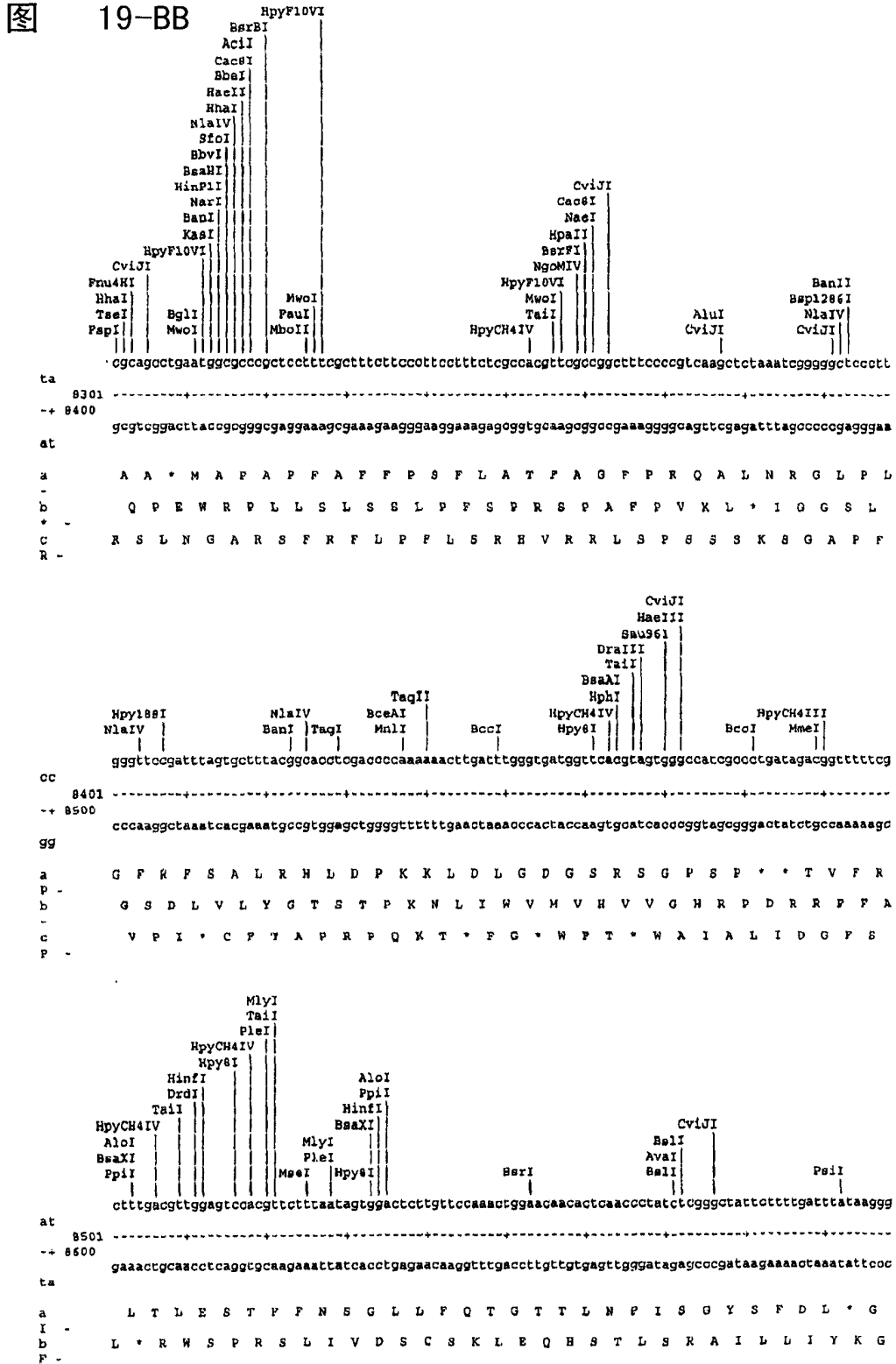




图 19-AA



C P G V T Q L N K L A A H P P A S W R N S E E A R T D R P S Q Q L



c F D V G V H V L \* \* W T L V P N W N N T Q P Y L G L P F \* F I R D  
-





```
      cggogtoagcgggagagcgcgttgttaaggcggcagactttgctcatggtaccgatgctattcggaagaaggcchactaagotgcccgggttgaaacag
ga      -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
      8901 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
      -+ 9000 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
      ct      gcgcgagtcgcacctctcggaaacattcggcogtctgaaacgagtagatggctacgataagccttcttgccggttgattcgacggccocaaactttgtgc
a      G V S G R A V V R R Q T L L M L P M L P G R T A T X L P G L K H G
b      A S A G E P L * G G R L C S C Y R C Y S E E R Q L S C R V * N T
d      R R Q R E S R C K A A D F A H V T D A I R K N G N * A A G P E T R
c
m
```

图 19-DD

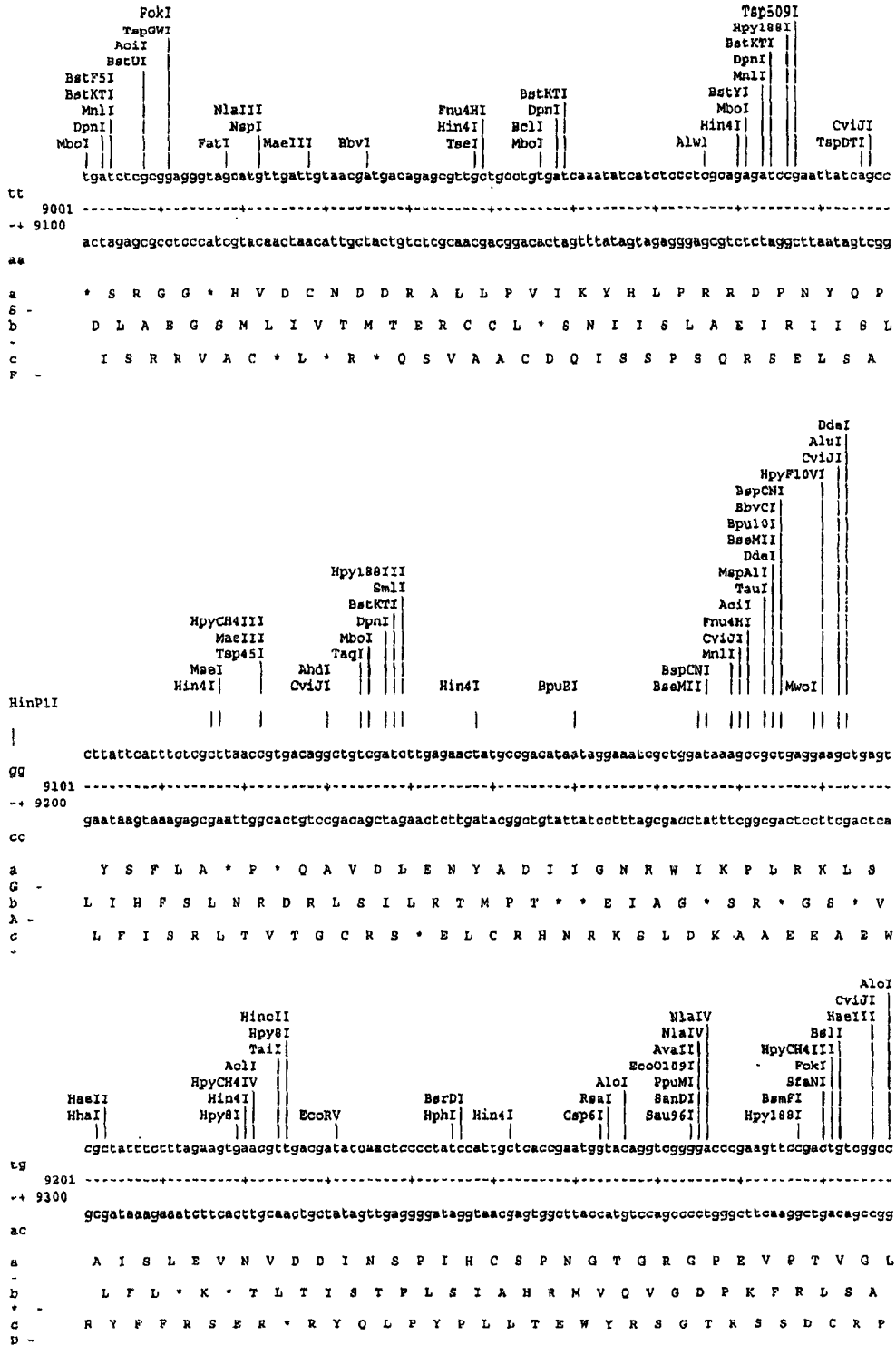
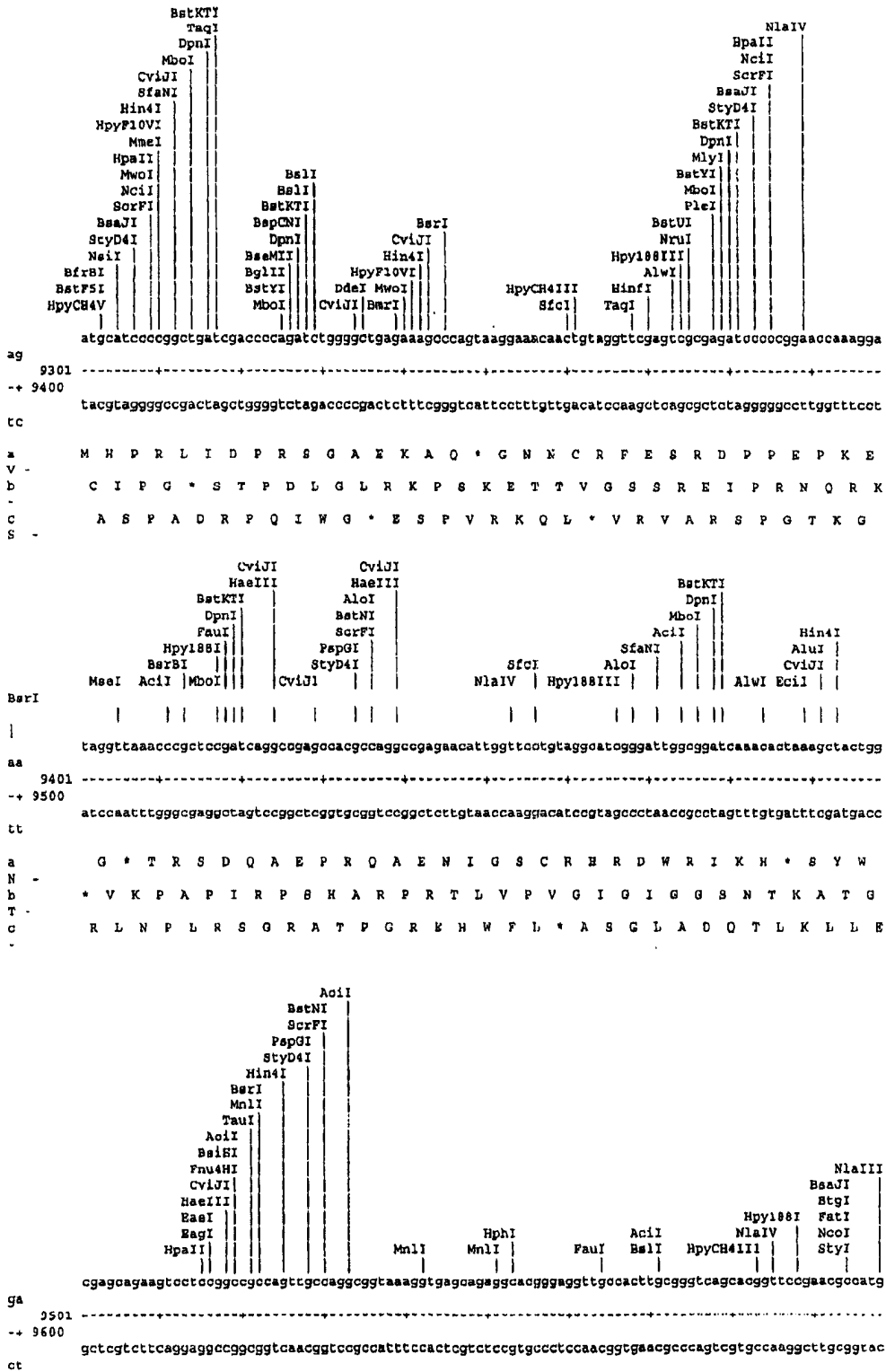
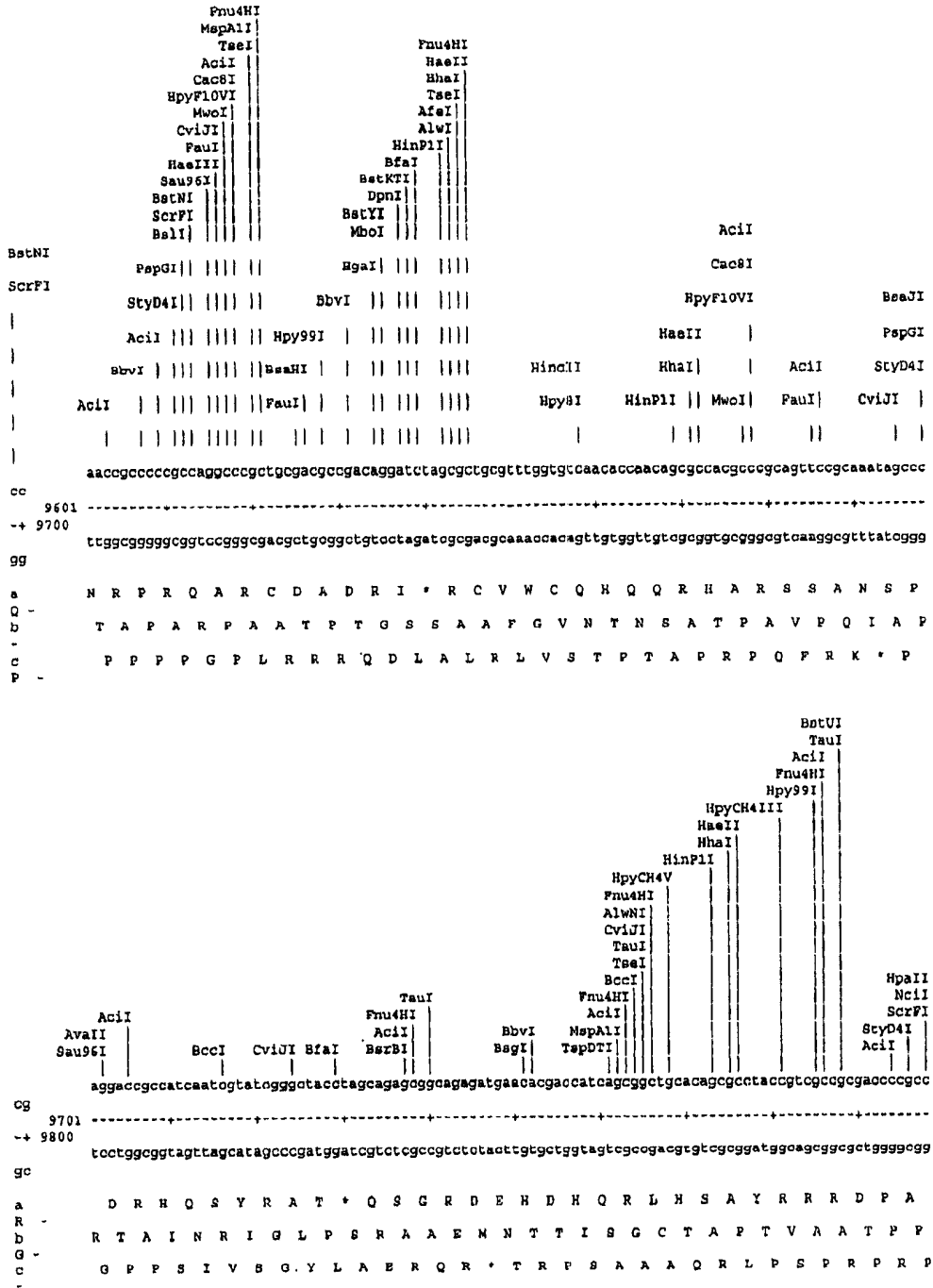


图 19-EE



a        E Q K S S G R Q L P G G K G E Q R H G R L P L A G Q H G S E R H G  
-  
b        S R S P P A A S C Q A V K V S R G T G G C H L R V S T V P N A M  
E -  
c        R A E V L R P P V A R R \* R \* A E A R E V A T C G S A R F R T P W  
K -

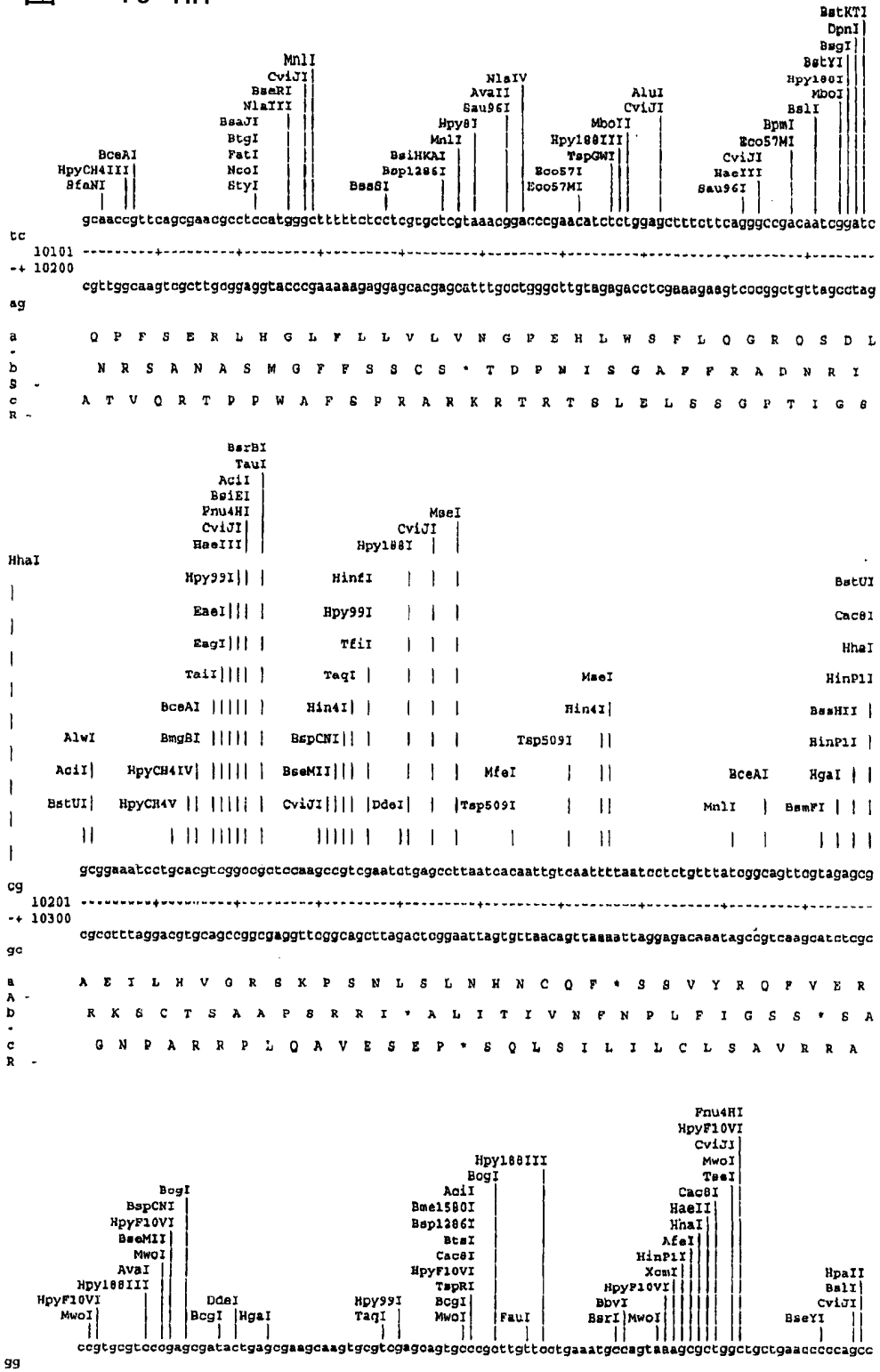
图 19-FF





a           M R L V S V L G A V L L F E D R Q P N D L A V D V G A E C H G I  
s -  
b           R C A L \* A S L G P S S C L K T D S P M I S P S M \* A P N A T A S  
r -  
c           D A F C E R F W G R P F V \* R P T A Q \* S R R R C R R R M P R H L  
-

图 19-HH

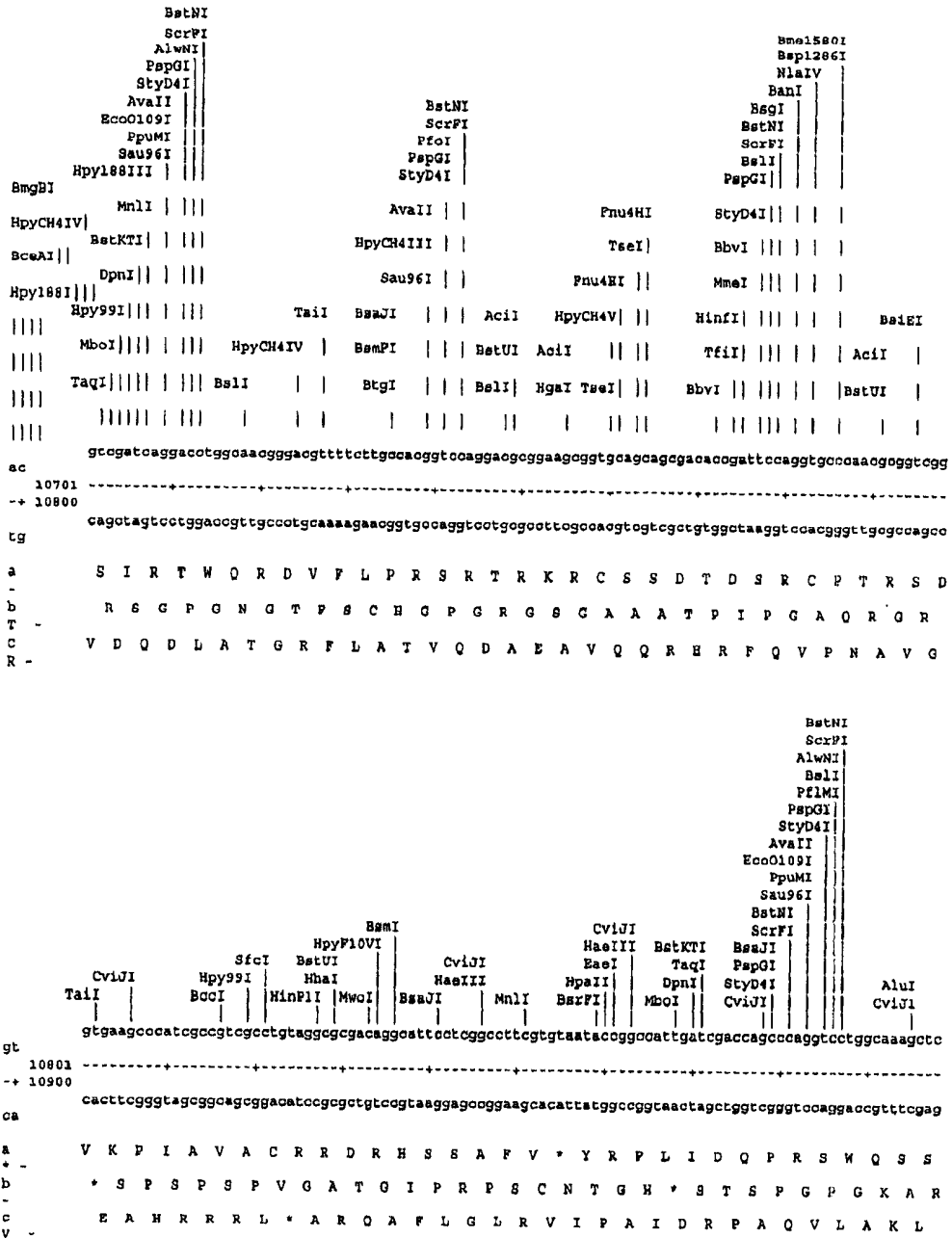




```
10301 -----+-----+-----+-----+-----+-----+-----+-----+-----  
--> 10400  
cc     ggcaagcaggggtcgcctatgactcgcttcggttcacgcagotcgtcaegggggaacaaggactttaagggtcatttcgcgaccgaacgacttgggggtcgg  
a      V R P E R Y * A K Q V R R A V P A C S * N A S K A L A A E P P A  
g -  
b      P C V P S D T E R S K C V E Q C P L V P E M P V K R W L L N P Q P  
e -  
c      R A S R A I L S E A S A S S S A R L P L K C Q * S A G C * T P S R  
-
```



图 19-JJ





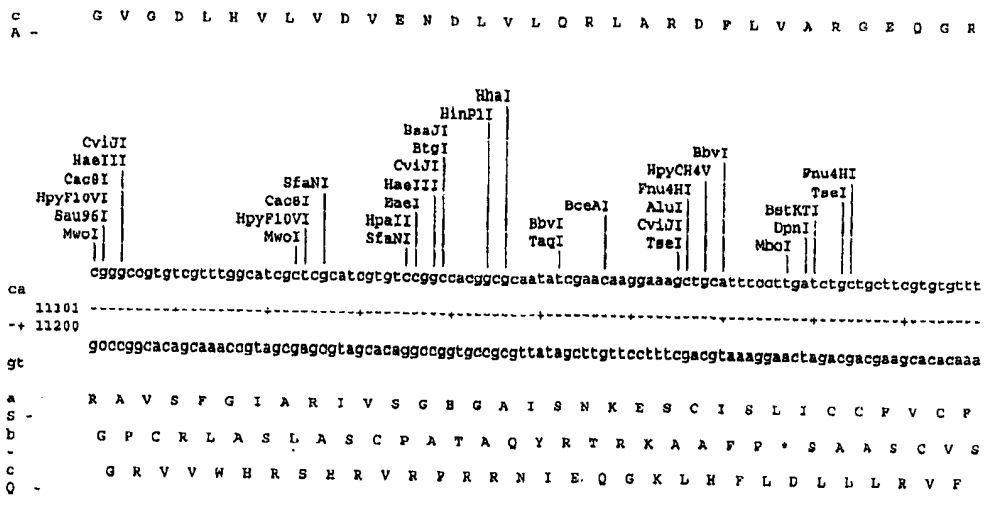


图 19-LL

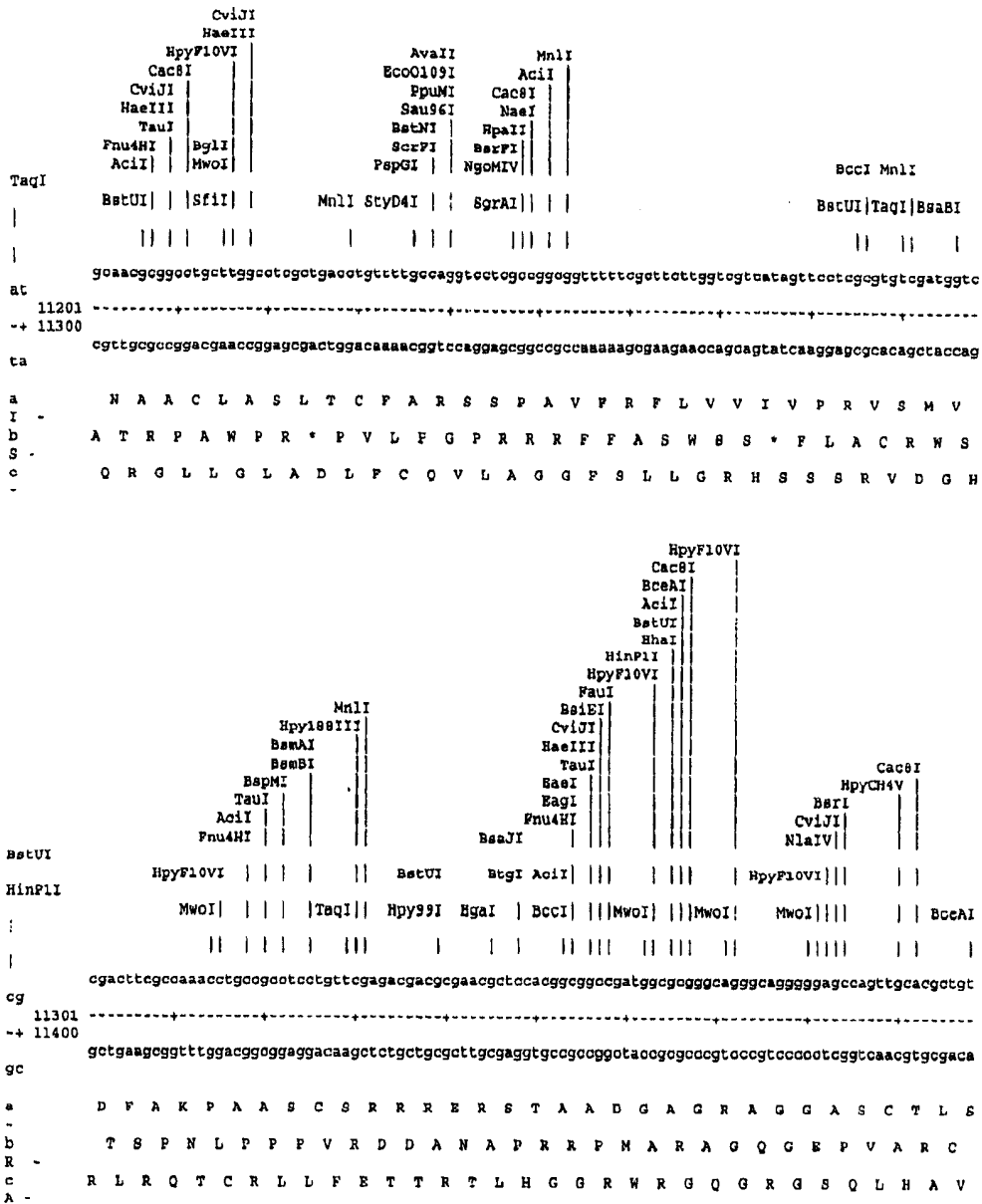
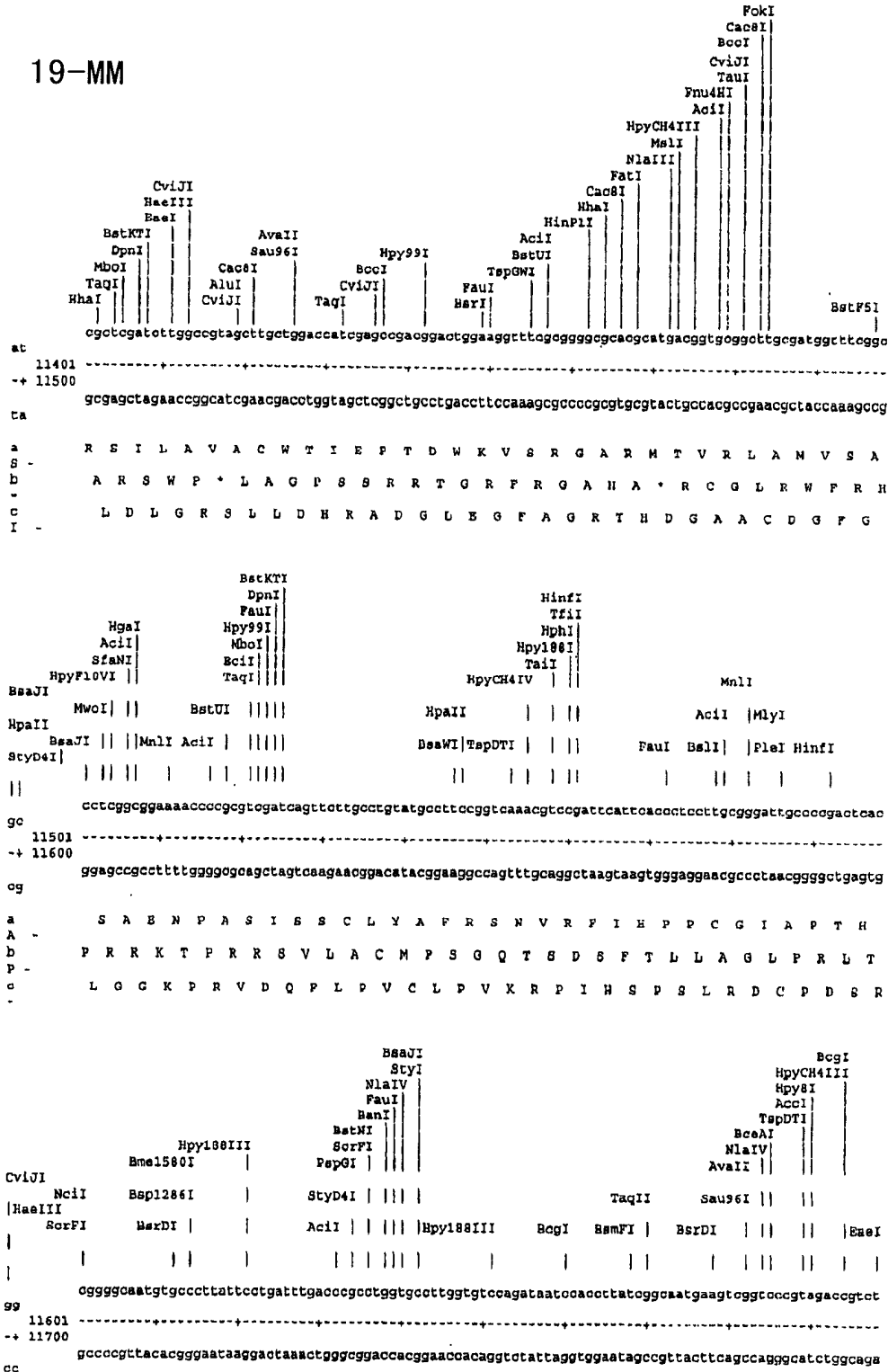


图 19-MM

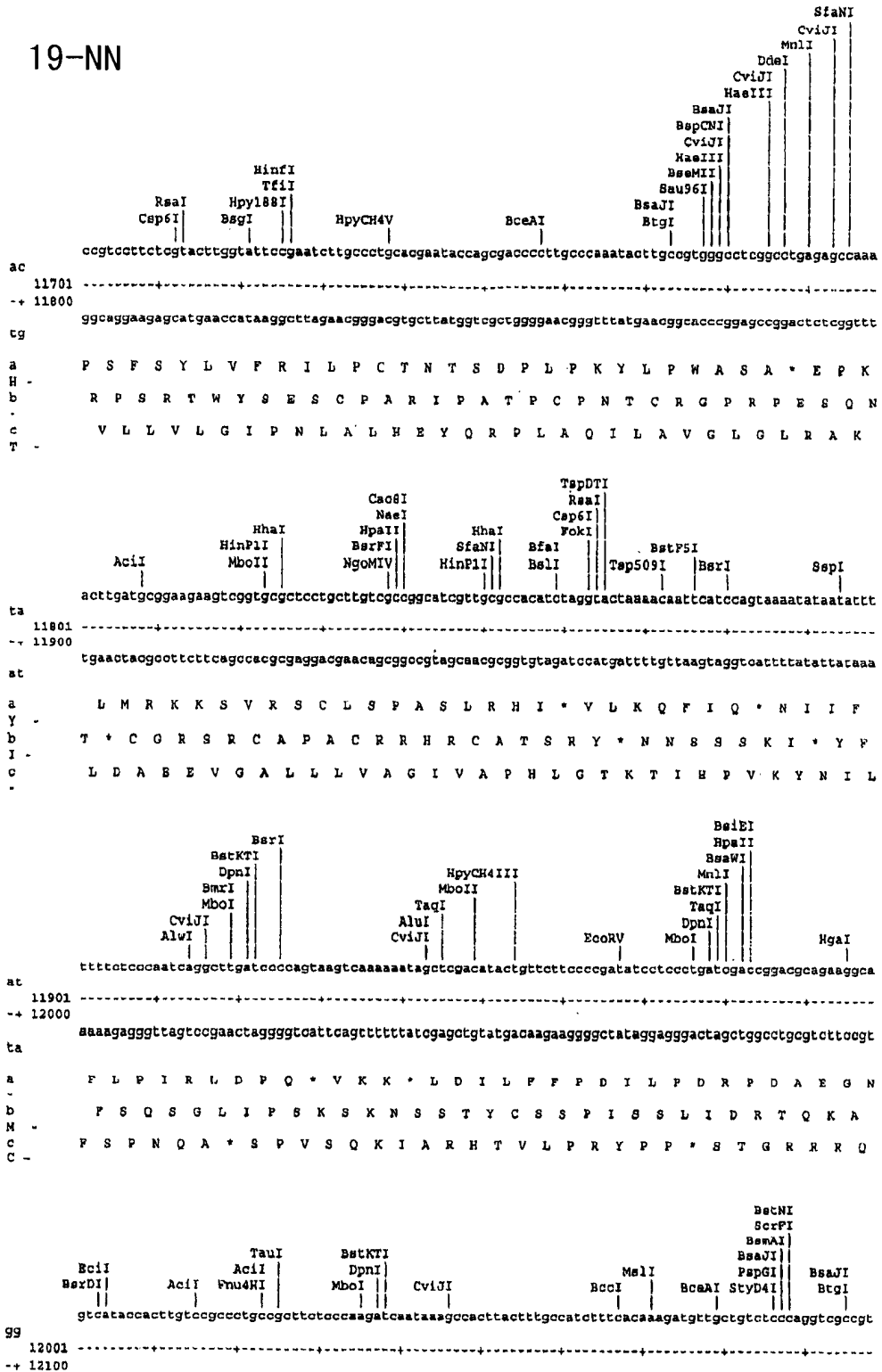


---

a        G A M C P Y S \* F D P P G A L V S R \* S T L S A M K E V P \* T V W  
-  
b        G Q C A L I P D L T R L V P W C P D N P P Y R Q \* S R S R R P S  
G -  
c        R G N V P L P L I \* P A W C L G V Q I I H L I G N E V G P V D R L  
A -



图 19-NN

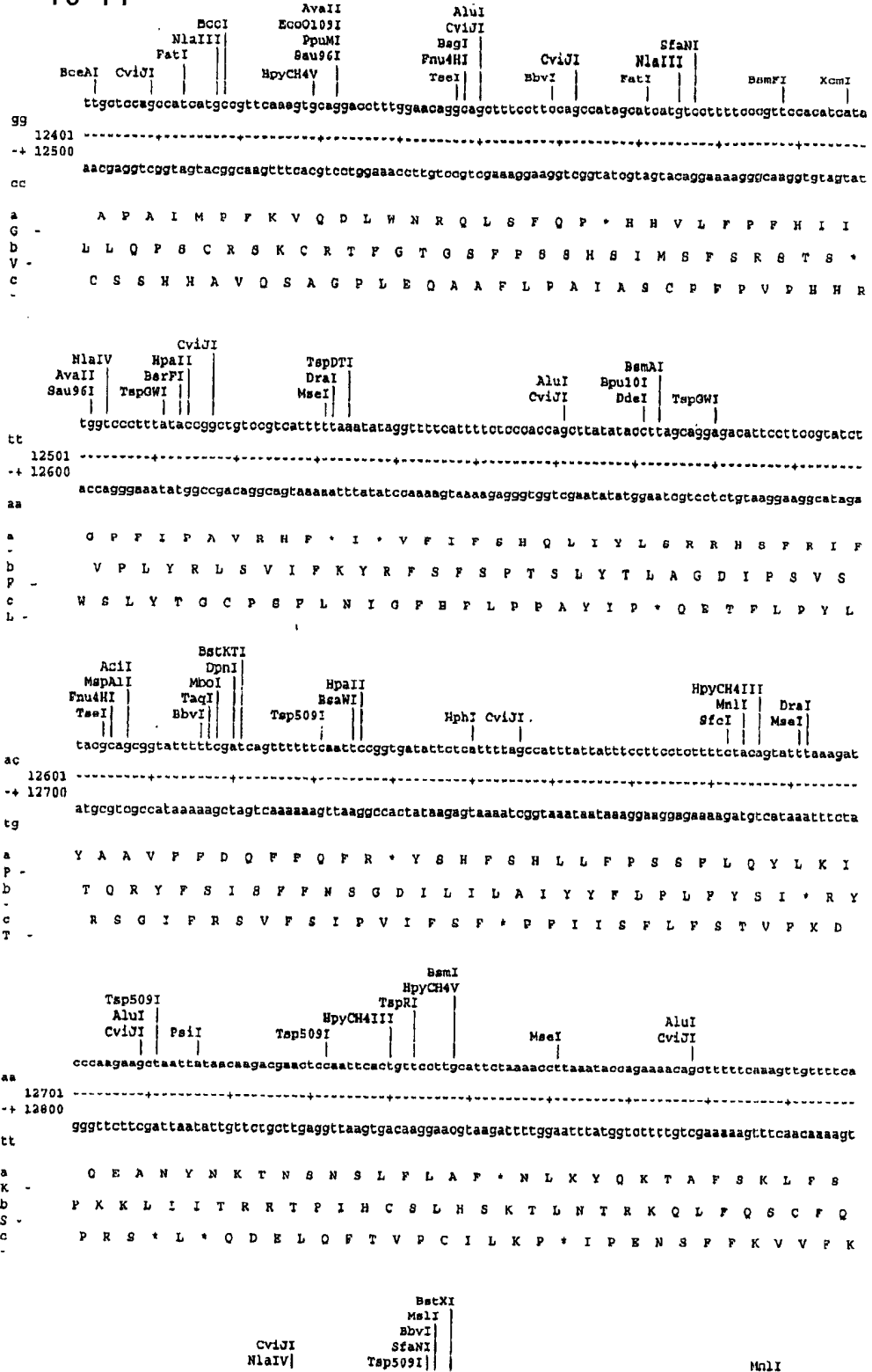


```
cc      cagtatgggtaacaggcgggacggcgaagagggttctagttatttcggatgaatgaaacggtagaagtgttctacaacgacagagggtccagoggca
a      V I P L V R F A A S P K I N K A T Y F A I F H K D V A V S Q V A V
G -
b      S Y H L S A L P L L P R S I K P L T L P S F T X M L L S P R S P W
-
c      H T T C F P C R F S Q D Q * S H L L C H L S Q R C C C L P G R R
G -
```





19-PP



```

                Hpy99I  | | |
                TaqI   | | |
                TspGWI  | | |
                TaqII   | | | |
                Fnu4HI  | | |
                TseI    | | |
                HphI    | |
                BccI    |
                Eco57MI |
gt  gttggcgtataaacatagtatcgacgggagccgattttgaaaccacaattatgggtgatgctgocaaacttactgatttagtgtatgatgggtgttttggag
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
12801
-+ 12900
ca  caaccgcataattgtatcatagctgcctcgggctaaaactttgggttaataccactacgacgggtgaatgactaaatcacatactaccacaaaaactc
a   L A Y N I V S T E P I L K P Q L W V M L P T Y * F S V * W C F * G
-   N R I T * Y R R S R F * N H N Y G * C C Q L T D L V Y D G V F E
b   V G V * H S I D G A D F E T T I M G D A A N L L I * C M M V F L R
V -
c -
C -

```



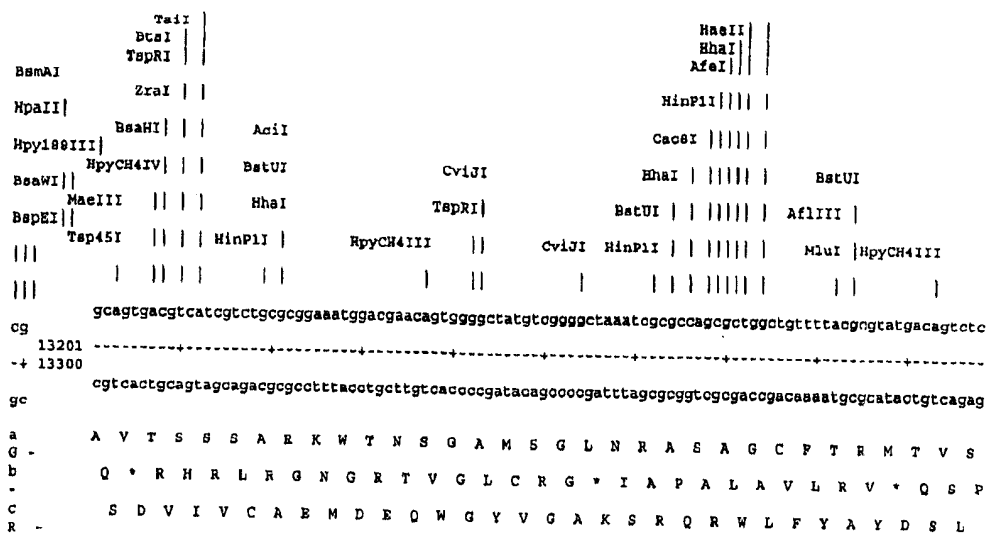
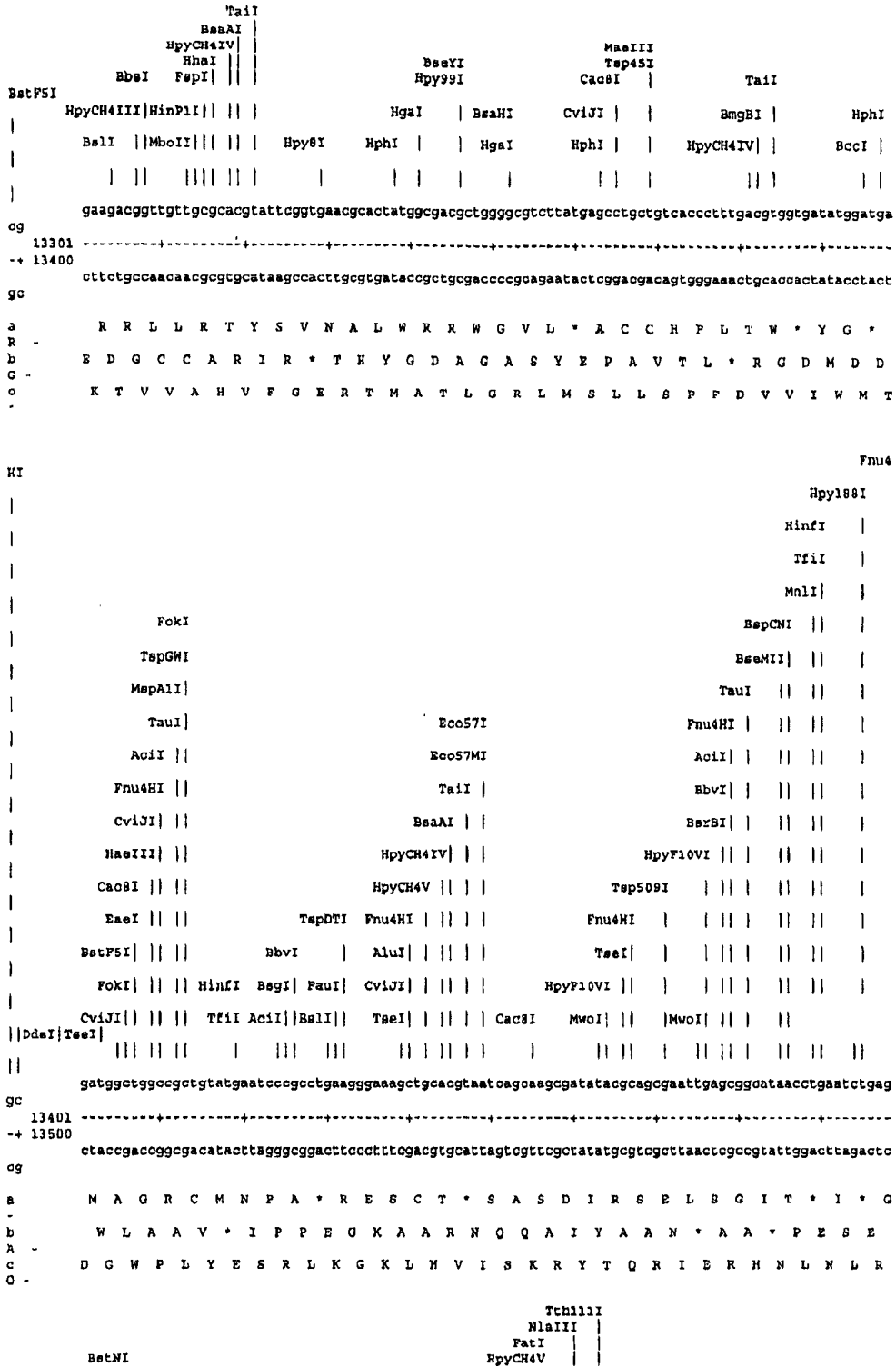


图 19-RR





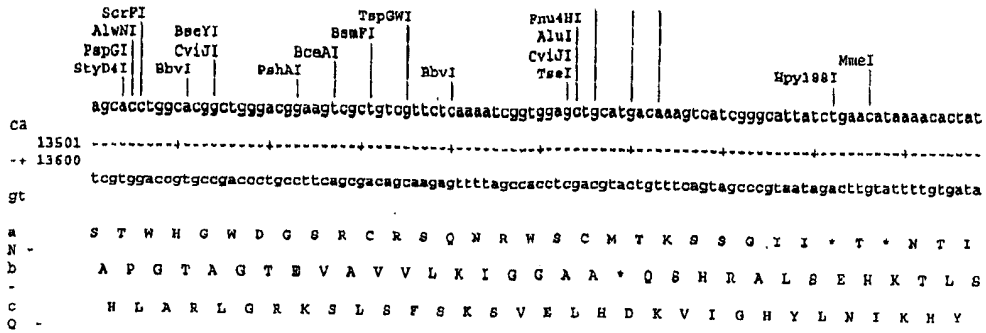
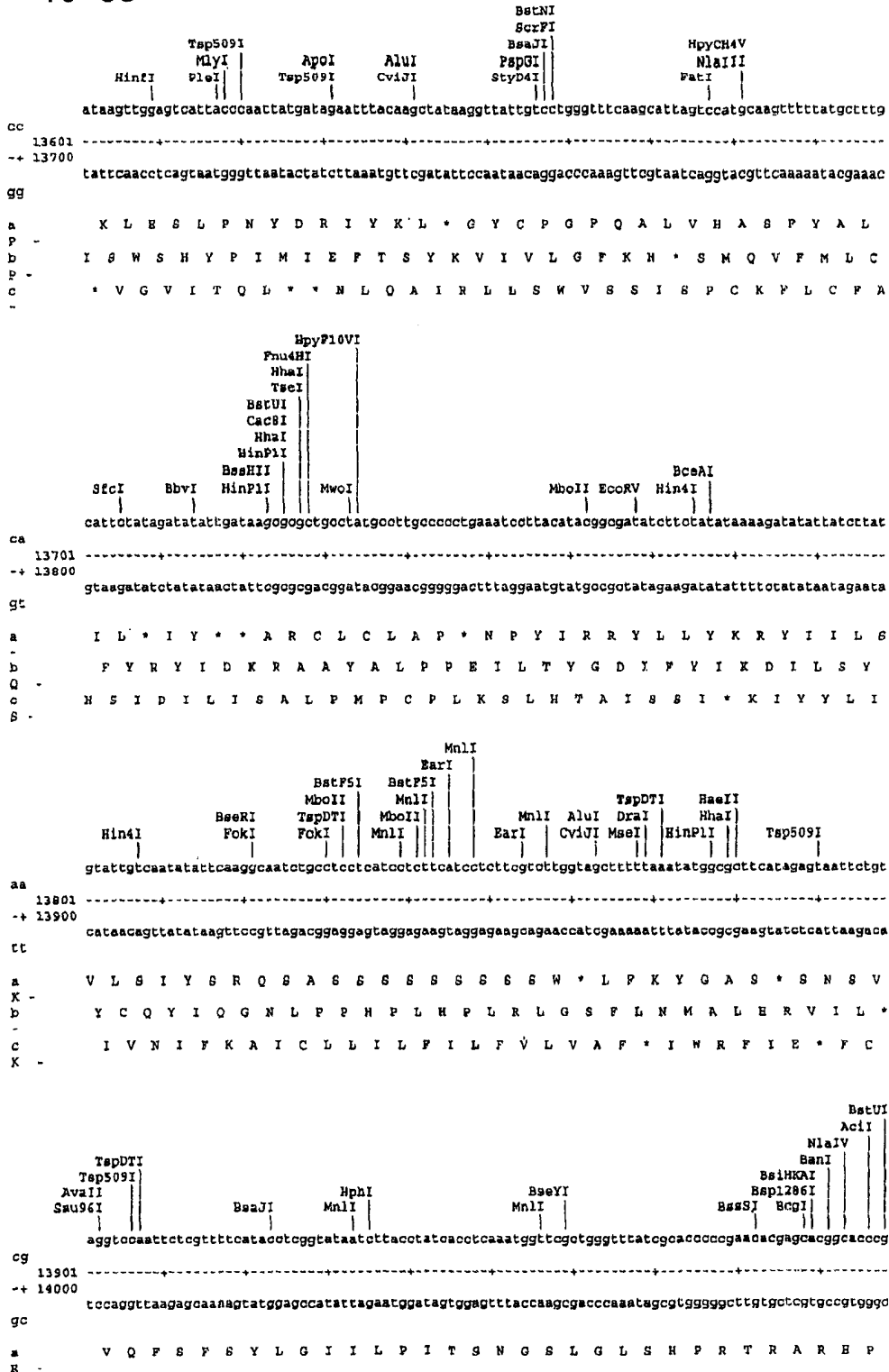


图 19-SS

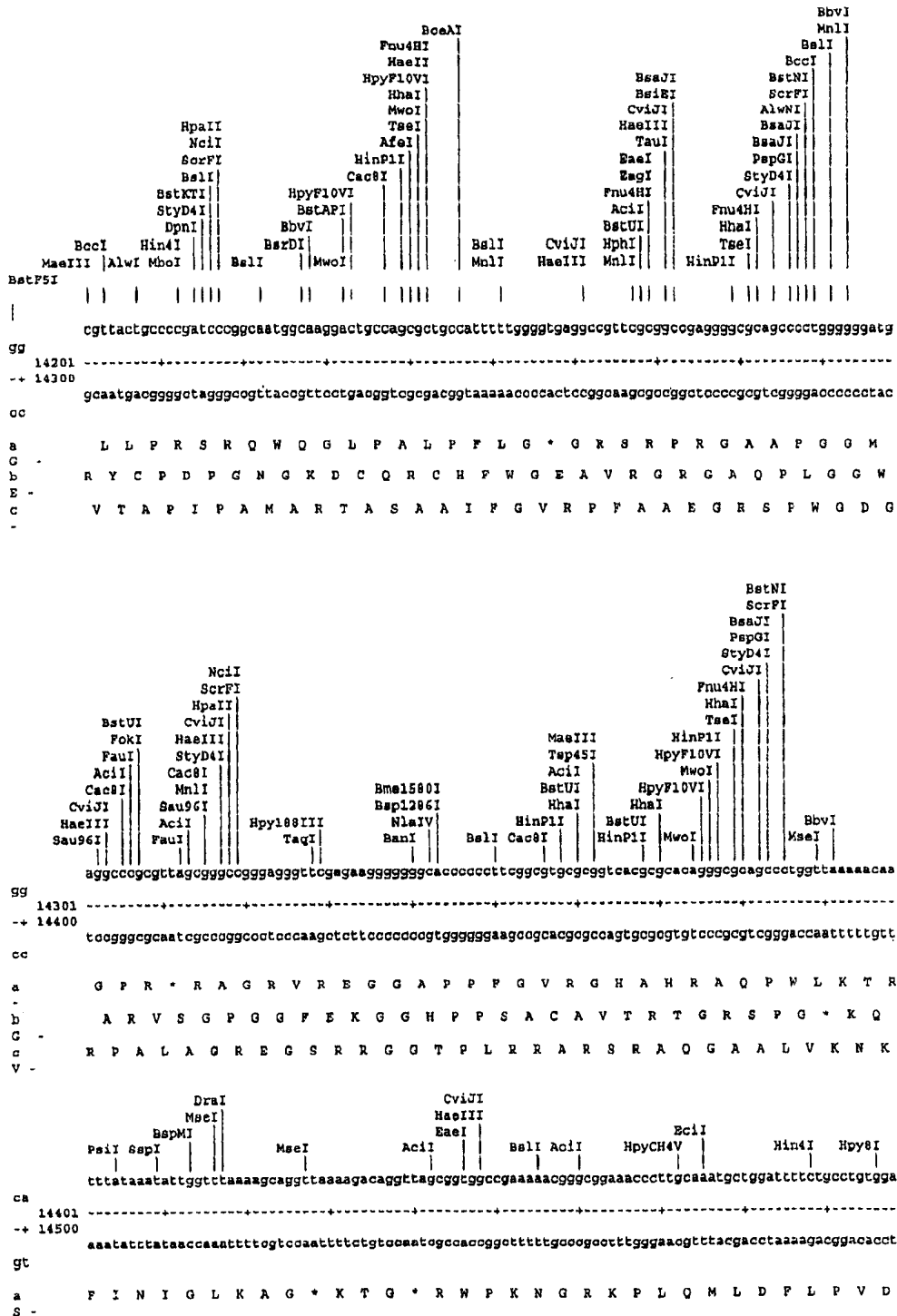


D R S N S R F H T S V \* S Y L S P Q M V R W V Y R T P E H E H G T R  
D - G P I L V F I P R Y N L T Y H L K W F A G P I A P P N T S T A P A  
C -  
-



a - R R P H C P A P G R \* M S M P A P A A R Q C F R A S R S G \* S P I  
b - A A L T A R H L V A E C R C Q H L R H V N A S G R R A R A D R P S  
c - P P S L P G T W S L N V D A S T C G T S M L P G V A L G L I A H  
d -

图 19-UU



b        L \* I L V \* K Q V X R Q V S G G R K T G G N P C K C W I F C L W T  
-  
c        Y K Y W P X S R L K D R L A V A E K R A E T L A N A G F S A C G  
Q -





```
tc gaggtgcagcggcggcgtttgctcggacggggagtagacagttgogggcggccactcagccggggagttcacagttgcaggggggagtagacag
a L H V A G R N R A C P S S V N A A P G E S A P Q V S T S A P H L S
v - S T S P A E I E P A P H L S T D R R V S R P L K C Q R P P L I C Q
b - P R R R P K S S L P L I C Q R R A G * V G P S S V N V R P S S V
c -
s -
```

图 19-WW

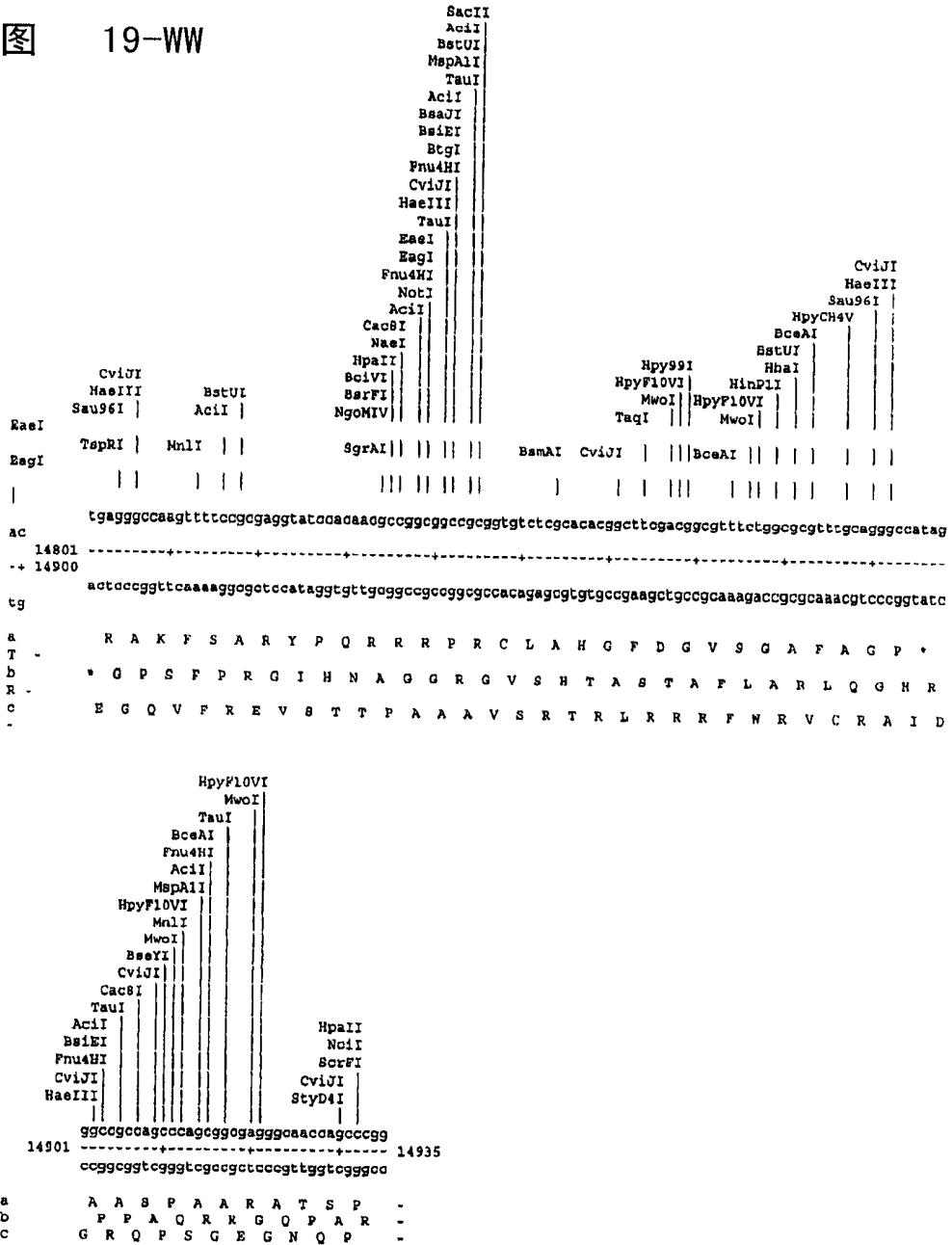


图 19-XX

切割的不被排除在外的酶:

AarI	AatII	AccI	AccI	AccI	AfeI	AflII	AflIII	AgeI	AhdI	AleI
AloI	AluI									
AlwI	AlwNI	ApaI	ApaLI	APOI	AsaI	AvaI	AvaII	BaeI	BamHI	BanI
BanII	BbsI	BbvI	BbvCI	BccI	BceAI	BcgI	BciVI	BclI	BfaI	BfrBI
BglII	BlpI	BmgBI	BmzI	BmtI	BplI	BpmI	BpuI	BpuEI	BsaI	BsaAI
BmeI500I	BsaJI									BsaBI
BsaWI	BsaXI	BseMII	BseRI	BseYI	BsgI	BsIEI	BsIHKAI	BslI	BsmI	BsmAI
BsmBI	BsmFI									
BspI286I	BspCNI	BspEI	BspHI	BspMI	BsrI	BsrBI	BsrDI	BsrFI	BssHII	BssSI
BstAPI	BstBI	BstKTI	BstNI	BstUI	BstXI	BstYI	Bsu36I	BtgI	BtsI	Cac8I
BstEII	BstP5I									
Clal	Csp6I	DpnI	DraI	DraIII	DrdI	EaeI	EagI	EarI	EclI	Eco57I
CviJI	DdeI									
Eco57MI	EcoNI	EcoRV	FalI	FatI	FauI	Fnu4HI	FokI	FspI	FspAI	HaeII
EcoO109I	EcoRI									
HaeIII	HgaI	HinPI	HincII	HindIII	HinfI	HpaI	HpaII	HphI	Hpy8I	Hpy99I
HhaI	Hin4I									
Hpy188I	Hpy188III	HpyCH4V	HpyF10VI	KasI	MaeIII	MboI	MboII	MfeI	MluI	MlyI
HpyCH4III	HpyCH4IV									
MmeI	MnlI	MaeI	MaeI	MaeI	MaeI	MaeI	MaeI	MaeI	MaeI	MaeI
MacI	MeeI	MaiI	MspAI	MwoI	NaeI	NarI	NciI	NcoI	NdeI	NgoMIV
NheI	NlaIII	NotI	NruI	NsiI	NspI	PciI	PfIMI	PfoI	PleI	PmeI
PpuMI	PshAI	PepOMI	ParI	PstI	PvuI	PvuII	RsaI	RsrII	SacII	SaiI
PaiI	PspGI									
SanDI	SapI	ScarI	ScrFI	SexAI	SfaNI	SfcI	SfiI	SfoI	SgrAI	SmlI
Sau56I	SbfI									
SpeI	SphI	StyD4I	TaiI	TaqI	TaqII	TatI	TauI	TfiI	TecI	Tsp45I
SspI	StyI									
Tsp509I	TspDTI									
Tsp6WI	TspRI	TthIII	XbaI	XcmI	XhoI	XmnI	ZraI			

不切割的酶:

Acc65I	AscI	AsiSI	AvrII	BsiWI	BsrGI	Bst217I	EcoICRI	PaeI	KpnI	PacI
PmlI	SacI									
SmaI	SnaBI	SrfI	StuI	SwaI	XmaI					

被排除在外的酶: MinCuts: 1 MaxCuts: 100000  
无

图 20-A

```

tgagcgtcgc aaaggcgctc ggtcttgccct tgctcgtcgg tgatgtactt caccagctcc 60
gCGaagtGcG tcttcttgat ggagcgcAtg gggAogtGct tggcaatcaC gCGcaCccC 120
CGGccgtttt agcggctaaa aaagtcAtgg ctctgcccctc gggCGgacca cGcccatcat 180
gaccttgcca agctcgctcct gcttctcttc gatcttcgcc agcaggggcga ggatcgTggc 240
atcaccgaac cgcgcgctgc gCGggctcgtc ggtgagccag agtttcagca ggcCGccCag 300
gCGggccCag tCGccattga tgcgggccag ctCGcggaCg tgctcatagt cCacgacgC 360
cgtgattttg tagccctggc cGacggccag caggtaggcc gacaggctca tgcCGgccc 420
CGccgccttt tCctcaatcg ctcttcgttc gtctggaagg cagtacacct tgataggtgg 480
gtcgcccttc ctggttggtc tggtttcAtc agccatccgc ttgcccctcat ctgttacgC 540
ggCGgtagcc ggcCagcctc gCagagcagg attcccgTg agcaccgcca ggtgCGaata 600
aggGacagtg aagaaggaac acccgctcGc gggTgggCct acttCacctA tCctgcccGg 660
ctgacgcccgt tggatacacc aaggaaagtc tacacgaacc ctTtgGcaaa atCctgtata 720
tcgtgCGaaa aaggatggat ataccgaaaa aatcGctata atgacccCGa agcagggtta 780
tGcagcGgaa aagcGccacg ctTccCGaag ggagaaaggc ggacaggTat cCGtaagcG 840
gCagggTcgG aacaggagag cGcagagggg agcttccagG gggaaacgC 900
atagtCctgt cgggtttcGc cacctctgac ttgagcgTcG attttTgtga tGctcgtcag 960
ggggcgGag cctatggaaa aacGccagca acGcgGcct tttacggTtC ctggcctttt 1020
gctggccttt tGctcacatg ttctttctcG cgttatcccc tGattctgtg gataaocGta 1080
ttacgcctt tGagtGagct gataccgctc gCcgagccG aacgaccgag cGcagcgagT 1140
cagtGagcGa ggaagcGgaa gagcGccaga aggcGccag agaggccGag cGcgccGctg 1200
aggctTggac gctaggGcag ggcatgaaaa agcccGtagc gggctgctac gggcGctctga 1260
cGcggtGgaa agggggaggg gatgtTgtct acatggctct gctgtagtga gTgggtTgCg 1320
ctcGcgagc ggtcctgAtc aatcGtcacc ctTtctcGgt ccttcaacgt tCctgacaac 1380
gagcctcctt ttCGccaatc catGgacaat caccGcgagT cCctgctcga acGctgCgtc 1440
cGgaccgGct tCgtcgaagg cgtctatcGc ggcCGcaac agcggcGaga gCGgagcctg 1500
ttcaacggTg cGcGcgGct cGcGcgcatc gctgTcgccG gcctgctcct caagcagGc 1560
cccAACagTg aagtagctga ttgtcatcag cGcattgaCg gCgtccCcg cCGaaaaac 1620
cGcctcGcag aggaagcGaa gctgCgctc ggcCGtttcc atctgCGgtg cGcccGgtcG 1680
cgtgCcgGca tggatgCgCg cGccatcGcG gtaggCGagc agcGcctgC tgaagctgCg 1740
ggcattccCG atcagaaatg agcGccagTc gTcgtcGgct ctCGgcaccG aatcGctatg 1800
attctcGcc agcatggctt cGgCagTgc gTcGagcagc gcccGcttgt tCctgaagTg 1860
ccagTaaagc gCcggtgct gaaccccaaa cGcttccGcc agtttGcgtg tCgtcagac 1920
gtctacGcG acctcGttca acaggtccag ggcGgcaCg atcactgtat tCGgtgca 1980
ctTgtcAtg ctTgacactt tatcactgat aaacataata tGtccaccaaa ctTatcagTg 2040
ataaagaatc cGcGcgttca atcGgaccag cGgagGctgG tccGgaggc agacGtGaaa 2100
cccAACatac cCctgatcgt aattctgagc actgtcGcG tCgacGctgt cGgcatcGgC 2160
ctgattatg cGgtgctgC gggcctcctg cGcGatctgG ttcactcGaa cGacgtcacc 2220
gcccactatg gCattctgct gGcGctgtat gCgtTggtGc aatttGcctg cGcaccTgtg 2280
ctggGcGcG tGtCGgAtcG ttTcGgGcG cGgccaatct tGctcgtctc gctggcGgC 2340
gCagatctG gggAACcctg tggTtggcat gCacatacaa atgGacGaac gGataaacct 2400
tttcaCgccc ttttaaatat cCgattatc taataaacGc tcttttctct taggttacc 2460
cGccaatata tCctgtcaaa cactgatagT ttaaacTgaa gGcGggaac gacaatctga 2520
tcatgagcGg agaattaaGg gagtcacgtt atgacccCcg cCgatgacGc gggacaagc 2580
gttttacgtt tggaaactgac agaaccGcaa cgttgaagga gCcactcagc cGcgggtttc 2640
tggagtttaa tGagctaagc acatacgtca gaaaccattA ttGcGcgttc aaaagtcGc 2700
taaggtcact atcagctagc aaatatttct tGtcaaaaat gctccactga cgttccataa 2760
attcCctcG gTatccaatt agagtctcat attcactctc aatccaaata atctgCaccG 2820
gatctggatc gTttcGcatg attgaacaag atggattgca cGcaggttct cGcgccGctt 2880
gggtgGagag gctattcGgC tatgactggg cacaacagac aatcGgctGc tctgagcCg 2940
cCgtgttCcg gctgtcagcG cagggGcGcc cGgttctttt tGtcaagacc gacctgtcCg 3000
gtgcccTgaa tgaactgCag gacGagGcag cGcggtatc gTggtgGc acgacggGc 3060

```

图 20-B

```

ttccttgccg agctgtgctc gacgttgcca ctgaagcggg aagggactgg ctgctattgg 3120
gogaagtgcc ggggcaggat ctctgtcat ctcacctgce tcctgccgag aaagtatcca 3180
tcatggctga tgcaatgccg cgggtgcata cgcttgatcc ggctacctgc ccattcgacc 3240
accaagcgaa acatcgcatc gagcggagcac gtactcggat ggaagccggc cttgtcgatc 3300
aggatgatct ggacgaagag catcaggggc tcgcgccagc ggaactgttc gccaggctca 3360
aggcgcgatc gcccgacggc gatgatctcg tcgtgacca tggcgatgcc tgcttgccga 3420
atatcatggt ggaaaatggc cgcttttctg gattcatcga ctgtggccgg ctgggtgtgg 3480
cggaccgcta tcaggacata gcgttggeta cccgtgatat tgctgaagag cttggcggcg 3540
aatgggctga ccgcttcctc gtgctttacg gtatcgccgc tcccgatctc cagcgcctcg 3600
ccttctatcg ccttcttgac gagttcttct gagcgggact ctggggttcg aaatgaccga 3660
ccaagcgacg cccaacctgc catcacgaga tttcgattcc accgccgcct tctatgaaag 3720
gttgggcttc ggaatcgttt tccgggacgc cggctggatg atcctccagc gccggggatct 3780
catgctggag ttcttcgccc acgggatctc tcgggaacag gcggtcgaag gtgccgatat 3840
cattacgaca gcaacggccg acaagcacia gccacgatc ctgagcgaca atatgatcgg 3900
gcccggcctc cacatcaacg gcgtcggcgg cgactgccca ggcaagaccg agatgcaccg 3960
cgatatcttg ctgcgttcgg atattttcgt ggagttcccg ccaagaccgc ggatgatccc 4020
cgatcgttca aacatttggc aataaagttt cttagattg aatcctgttg ccggtccttg 4080
gatgattatc atataatttc tgttgaatta cgttaagcat gtaataatta acatgtaatg 4140
catgacgtta tttatgagat gggtttttat gattagatc ccgcaattat acatttaata 4200
cgcgatagaa aacaaaatat agcgcgcaaa ctaggataaa ttatcgcgcg cgggtgctatc 4260
tatgttacta gatcgggcct cctgtcaatg ctggcggcgg ctctgggtgtt ggttctggtg 4320
gcggctctga ggggtggtggc tctgagggtg gcggttctga ggggtggcggc tctgagggag 4380
gcggttccgg tggtggctct ggttccggtg attttgatta tgaaaagatg gcaaacgcta 4440
ataagggggc tatgaccgaa aatgccgatg aaaaacgcct acagtctgac gctaaaggca 4500
aacttgatc tgtogetact gattacggtg ctgtatcga tggtttcatt ggtgacgttt 4560
ccggccttgc taatggtaat ggtgctactg gtgattttgc tggctcta at tcccaaatgg 4620
ctcaagctcg tgaoggtgat aattcaccct taatgaataa tttccgtcaa tatttaact 4680
ccctccctca atcggttgaa tgcgcgccct ttgtctttgg cccaatacgc aaaccgcctc 4740
tccccgcgcg ttggccgatt cattaatgca gctggcaoga caggtttccc gactggaaag 4800
cgggcagtg ggcgaaacgca attaatgtga gttagctcac tcattaggca ccccaggctt 4860
tacactttat gcttccggct cgtatgttgt gtggaattgt gagcggataa caatttcaca 4920
caggaacag ctatgacct gattaogcca agcttgoatg cctgcaggtc cccagattag 4980
ccttttcaat ttcagaaaaga atgctaacc acagatgggt agagaggctt acgcagcag 5040
tctcatcaag acgatctacc cgagcaataa tctccaggaa atcaaatacc ttccaagaa 5100
ggttaaagat gcagtcaaaa gattcaggac taactgcac aagaacacag agaaagatat 5160
atltctcaag atcagaagta ctattccagt atggacgat caaggottgc ttcaaaaacc 5220
aaggcaagta atagagattg gagtctctaa aaaggtagtt cccactgaat caaaggccat 5280
ggagtcaaa attcaaatag aggacctaac agaactcgc gtaaaactg gcgaacagtt 5340
catacagagt ctcttacgac tcaatgacaa gaagaaaatc ttcgtcaaca tgggtggagca 5400
cgacacactt gtctactoca aaaaatcaaa agatacagtc tcagaagacc aaagggcaat 5460
tgagactttt caacaaagg taatatccgg aaacctcctc ggattccatt gccagctat 5520
ctgtcacttt atttgaaga tagtgaaaa ggaaggtggc tcctacaaat gccatcattg 5580
cgataaagg aaggccatcg ttgaaagatgc ctctgccgac agtgggtcca aagatggacc 5640
ccccccacg aggagcatcg tggaaaaaga agacgttcca accacgtctt caaagcaagt 5700
ggattgatgt gatatctcca ctgacgtaag ggatgacgca caatcccact atccttcgca 5760
agacccttcc tctatataag gaagtccatt tcatltggag agaacacggg ggactctaga 5820
ggatccatgt ttattttctt attatttctt actctcacta gtggtagtga ccttgaccgg 5880
tgcaccactt ttgatgatgt tcaagctcct aattacactc aacatacttc atctatgagg 5940
ggggtttact atcctgatga aatltttaga tcagacactc tttatttaac tcaggattta 6000
tttcttccat tttatcttaa tgttacaggg tttcactacta ttaatcatac gtttggcaac 6060
cctgtcatac cttttaagga tggtatttat tttgtgcca cagagaaatc aaatgttgc 6120
cgtgggtggg tttttggttc taccatgaac aacaagtcac agtcgggtgat tattattaac 6180
aattctacta atgttgttat acgagcatgt aactttgaat tgtgtgacaa ccctttcttt 6240

```

图 20-C

```

gctgtttcta aacctatggg tacacagaca catactatga tattegataa tgcatttaat 6300
tgcactttcg agtacatatac tgatgccttt tcgcttgatg tttcagaaaa gtcaggtaat 6360
tttaaacact tacgagagtt tgtgtttaaa aataaagatg ggtttctcta tgtttataag 6420
ggctatcaac ctatagatgt agttcgtgat ctacettctg gtttbaacac tttgaaacct 6480
atthttaagt tgcctcttgg tattaacatt acaaatttta gagccattct tacagccttt 6540
tcacctgctc aagacatttg gggcacgtca gctgcagcct atthttgttg ctatttaaag 6600
ccaactacat ttatgctcaa gtatgatgaa aatggtacaa tcacagatgc tgttgattgt 6660
tctcaaaatc cacttgctga actcaaatgc tctgttaaga gctttgagat tgacaaagga 6720
atthaccaga octctaattt cagggttgtt ccctcaggag atgttgtgag atthccctaat 6780
attcaaaact tgtgtccttt tggagaggtt thtaatgcta ctaaattccc thctgtctat 6840
gcatggggaga gaaaaaaat thctaattgt gttgtctgatt actctgtgct ctacaactca 6900
acatthtttt caacctthaa gtgctatggc gthttctgcc ctaagttgaa tgatctttgc 6960
thctccaatg tctatgcaga thctthtgta gtcaagggag atgatgtaag acaaatagcg 7020
ccaggacaaa ctggtgttat tgctgattat aattataaat tgccagatga thtcatgggt 7080
tgtgtccttg cttggaatac taggaacatt gatgctactt caactggtaa thataattat 7140
aaatataggt atcttagaca tggcaagctt aggccttttg agagagacat atctaatgtg 7200
cctthctccc ctgatggcaa accttgcacc ccacctgctc thaatgtta thggccatta 7260
aatgattatg gththtacac cactactggc atthggctacc aaccttacag atthtgtaga 7320
ctthctthtg aactthtaaa tgcaccggcc acggtttgtg gacaaaaatt atthctgac 7380
ctthattaaga accagtgtgt caatthtaat thtaatggac tcactggtac tgggtgtgta 7440
actccttctt caaagagatt tcaaacattt caacaatttg gccctgatgt thctgatttc 7500
actgattccg thcgagatcc taaaacatct gaaatattag acatthcacc thgtctthtt 7560
gggggtgtaa gtgtaattac acctggaaca aatgcttcat ctgaagttgc thgtctat 7620
caagatgtta actgcactga tgtthctaca gcaattcatg cagatcaact cacaccagct 7680
tggegcatac attctactgg aaacaatgta thccagactc aagcaggctg thctatagga 7740
gctgagcatg tgcacacttc thtatgagtgc gacatthcta thggagctgg catthtgct 7800
attaccatac agttagacta gaaaagggcg aatthctgac atathcatca cactggggcg 7860
cgctcgacct cgaggggggg ccggttacag ctogaatttc ccgatcgtt caaacattg 7920
gcaataaagt thcttaagat tgaatctgt tgccggtctt gcgatgatta thcatataatt 7980
tctgttgaat tacgttaagc atgtaataat taacatgtaa tgcagacgt ththtatgag 8040
atgggtthtt atgattagag thcccgcaatt atacatthaa thacggatag aaaaacaaat 8100
atagcgcgca aactaggata aattatcgcg cgcggtgtca thctatgttac tagatcggga 8160
atthcactggc cgtcgtthta caacgtcgtg actgggaaaa ccctggcgtt acccaacta 8220
atcgccttgc agcacatccc cctthcggca gctggcgtaa tagcgaagag gcccgaccg 8280
atcgccttcc caaacagttg cgcagcctga atggcgcccg thctcttctc thtctctccc 8340
thctthctcg ccacgttctc cggctthccc cgtcaagctc thaatcgggg gctcccttha 8400
gggttccgat thagtgtctt acggcacctc gacccaaaa aacttgattt ggggtgatgt 8460
thcactagtg ggcatcgcc ctgatagacg gththctgcc cthtgacgtt ggagthccag 8520
thctthtaata gtggactctt gthccaaact ggaacaacac thcaacctat ctcggtctat 8580
thctthtgatt tataagggat thtgccgatt thggaaccac catcaaacag gattthctc 8640
tgctggggca aaccagcgtg gaccgcttgc tgcaactctc thcagggccag gcggtgagg 8700
gcaatcagct gthgcccgtc thcactggtg aaagaaaaac caccacagta catthaaac 8760
gtccgcaatg thgtattaag thgtctaagc gtcaatthgt thacaccaca atathctctg 8820
ccaccagcca gccaacagct ccccgaccgg cagctcggca caaatcacc actcgatac 8880
ggcagcccat cagthcggga cggcgtcagc gggagagccg thgttaaggcg gcagactthg 8940
ctcatgttac cgatgctatt cgggaagaac gcaactaagc thcggggtt gaaacacgga 9000
tgatctcgcg gaggttagca thgttgattgt aacgatgaca gagcgttgct gcctgtgatc 9060
aaatatcacc thcctcgcag agatccgaat thctagcctt cthattcatt thctcgtta 9120
ccgtgacagg ctgtcgatct thgagaactat gccgacataa thggaatcg ctggataaag 9180
ccgtgagga agctgagtg cgctatthct thagaagtg acgttgacga thctaacctc 9240
cctatccatt gctcacggaa thggtacaggt cggggacccg aagthctcagc thctcggctg 9300
atgcatcccc ggctgatcga ccccgatct ggggtgaga aagccagta aggaacaac 9360

```

## 图 20-D

```

tgtaggttcg agtcgcgaga tccccggaa ccaaaggaag taggttaaac ccgctccgat 9420
caggccgagc cacgccaggc cgagaacatt ggttcctgta ggcatcggga ttggcggatc 9480
aaacactaaa gctactggaa cgagcagaag tcctccggcc gccagttgcc aggcggtaaa 9540
ggtgagcaga ggcacgggag gttgccactt gcgggtcagc acggttccga acgocatgga 9600
aaccgcccc gccaggcccg ctgcgacgcc gacaggatct agcgctgctg ttggtgtcaa 9660
caccaacagc gccacgcccg cagttecga aatagcccc aggacggcca tcaatcgtat 9720
cgggctacct agcagagcgg cagagatgaa cacgaccatc agcggctgca cagcgcctac 9780
cgtcgcccg agccccccc gcaggcggta gaccgaaata aacaacaagc tcagaatag 9840
cgaaatatta agtgcgcca ggatgaagat gcgcaccac cagattccc ttggaatctg 9900
tcggacgatc atcacgagca ataaaccgc cggcaacgcc cgcagcagca tacoggcgac 9960
ccctcgccct cgtgttcgg gctccacgaa aaogccggac agatgcccct tgtgagcgtc 10020
cttggggccg tcctcctgtt tgaagaccga cagoccaatg atctcggcgt cgatgtaggc 10080
gccgaatgcc acggcatctc gcaaccgttc agcgaacgcc tccatggcct ttttctctc 10140
gtgctcgtaa acggaccoga acatctctgg agctttcttc agggccgaca atcggatctc 10200
goggaaaatc tgcacgtogg ccgctccaag ccgtcgaatc tgagcettaa tcacaattgt 10260
caattttaat cctctgttta tcggcagttc gtagagcgcg ccgtgctgcc cgagcgatac 10320
tgagcgaagc aagtgcgtcg agcagtgccc gcttgttctt gaaatgccag taagcgcgtg 10380
gctgctgaac cccagccgg aactgacccc acaaggccct agcgtttgca atgcaccagg 10440
tcacattga cccaggcgtg ttccaccagg ccgctgcctc gcaactctc cagggtctcg 10500
ccgacctgct cgcgccactt cttcacgccc gtggaatccg atccgcacat gaggcgggaag 10560
gtttccagct tgagcgggta cggctcccgg tgcgagctga aatagtcgaa catccgtcgg 10620
gcgctcggcg acagcttgg gtaottctcc catatgaatt tcgtgtagtg gtcgccagca 10680
aacagcagca cagattctcc gtcgatcagg acctggcaac gggacgtttt cttgccacgg 10740
tcaggacgc ggaagcggtg cagcagcgac accgattcca ggtgcccaac gcggtcggac 10800
gtgaagccca tcgcccgtcg ctgtaggcgc gacaggcatt cctcggcctt cgtgtaatac 10860
cggccattga tcgaccagcc caggctcctgg caaagctcgt agaacgtgaa ggtgatcggc 10920
tcgcccagatg ggggtgcgctt cgcgtactcc aacacctgct gccacaccag ttcgtcatcg 10980
tcggcccgc gctcgacgcc ggtgtaggtg atcttcacgt ccttgttgac gtgaaaaatg 11040
acctgtttt gcagcgcctc gcgcccggatt ttcttgttgc gcgtggtgaa cagggcagag 11100
cgggcccgtg cgtttggcat cgctcgatc gtgtcccggc acggcgaat atcgacaacg 11160
gaaagctgca tttccttgat ctgtgcttc gtgtgtttca gcaacggcgc ctgcttgccc 11220
tcgctgacct gttttgccag gtcctcgccc gcggtttttc gcttcttggg cgtcatagtt 11280
cctcgcgtgt cgatggatc cgaactcgc aaacctgccc cctcctgttc gagacgagc 11340
gaacgctcca cggcggccga tggcgcgggc agggcagggg gagccagttg cacgctgtcg 11400
cgctcgatct tggccgtagc ttgctggacc atcgagccga cggactggaa ggtttcgcgg 11460
ggcgcacgca tgacgggtgc gcttgcgatg gtttcggcat cctcggcggg aaaccccgcg 11520
tcgatcagtt ottgcctgta tgocctccgg tcaaacgtcc gattcattca cctccttgc 11580
gggatggccc cgactcacgc cggggcaatg tgcccttatt cctgatttga cccgcctggg 11640
gccttggtgt ccagataatc caccttatcg gcaatgaagt cggctccgta gaccgtctgg 11700
ccgctcctct cgtacttggg attccgaatc ttgccctgca cgaataccag cgacccttg 11760
cccaataact tgccgtgggc ctggcctgga gagccaaaac acttgatgcg gaagaagtcg 11820
gtgcgctcct gcttgcgcc ggcategttg cgocacatct aggtactaaa acaattcctc 11880
cagtaaaata taatatttta tttctctcca atcaggcttg atccccagta agtcaaaaa 11940
tagctcgaca tactgttctt ccccgatata ctccctgata gaccggacgc agaaggcaat 12000
gtcataccac ttgtcgcgcc tgccgcttct ccaagatca ataaagccac ttactttgcc 12060
atctttcaca aagatgttgc tgtctcccag gtcgcccgtg gaaaagacaa gtcctctctc 12120
gggcttttcc gcttttaaaa aatcatacag ctgcgcggga tctttaaatg gagtgccttc 12180
ttcccagttt tcgcaatcca categgccag atcgttatc agtaagtaat ccaattcggc 12240
taagcggctg tctaagctat tcgtataagg acaatccgat atgtcgatgg agtgaaagag 12300
cctgatgcac tcgcataca gctcgataat cttttcaggg ctttgttcat cttcatactc 12360
ttccgagcaa aggacgccat cggcctcact catgagcaga ttgctccagc catcatgccc 12420

```

图 20-E

```

ttcaaagtgc aggacctttg gaacaggcag ctttccttcc agccatagca tcatgtcctt 12480
ttcccgttcc acatcatagg tggteccctt ataccggctg tocgtcattt ttaaataatag 12540
gttttcattt tctcccacca gcttatatac cttagcagga gacattcctt ccgatatctt 12600
tacgcagcgg tatttttoga tcagtttttt caattccggg gatattctca ttttagccat 12660
ttattatttc cttcctcttt tctacagtat ttaaagatac cccaagaagc taattataac 12720
aagaogaaact ccaattcact gttccttgca ttctaaaacc ttaaatacca gaaaacagct 12780
ttttcaaagt tgttttcaaa gttggcgtat aacatagtat cgcagggagcc gattttgaaa 12840
ccacaattat ggggtgatgct gccaaacttac tgatttagtg batgatgggtg tttttgaggt 12900
gctccagtgg cttctgtgtc tatcagctgt ccctcctggt cagctactga cgggggtgggt 12960
cgtaacggca aaagcacggc cggacatcag cgctatctct gctctcactg ccgtaaaaca 13020
tggcaactgc agttcactta caccgcttct caaccgggta cgcaccagaa aatcatlgat 13080
atggccatga atggcgttgg atgcccggca acagcccgca ttatggcgtg tggcctcaac 13140
acgattttac gtcacttaaa aaactcaggc cgcagtcggg aacctcggc atacagccgg 13200
gcagtgacgt catcgtctgc gcggaatgg acgaacagtg gggctatgtc ggggctaaat 13260
cgcgcagcgg ctggcgtgtt tacgegtatg acagctcctg gaagacgggt gttgcccagc 13320
tattcgggtg acgcactatg cgcagctcgg ggcgtcttat gagcctgctg tcaccctttg 13380
acgtgggtgat atggatgacg gatggctggc cgctgtatga atcccgcctg aagggaaagc 13440
tgcacgtaat cagcaagcga tatacgcagc gaattgagcg gcataacctg aatctgaggc 13500
agcacctggc acggctggga cggaaagtcg tgtogttctc aaaatcgggt gagotgcatg 13560
acaaagtcac cgggcattat ctgaacataa aacactatca ataagttgga gtoattacc 13620
aattatgata gaatttaca gctataaggt tattgtcctg ggtttcaagc attagtcctt 13680
gcaagttttt atgctttgce cattctatag atatatatgat aagcgcgctg cctatgctt 13740
gccccctgaa atccttadat acggcgatat cttctatata aaagatata tctcttatca 13800
gtattgtcaa tatattcaag gcaatctgcc tctcatcct cttcatcctc ttogtcttgg 13860
tagcttttta aatatggcgc ttcataagat aattctgtaa aggtccaatt ctogttttca 13920
tacctcggta taatcttacc tatcacctca aatgggtcgc tgggtttatc gcacccccga 13980
acacgagcac ggcacccgcg accactatgc caagaatgcc caaggtaaaa attgcccggc 14040
ccgccatgaa gtcctgtaat gccccgacgg ccgaagtga gggcaggccg ccacccaggc 14100
cgcgcctc actgcccggc acctggtcgc tgaatgtoga tgccagcacc tgccggocgt 14160
caatgcttcc gggcgtcggc ctccggctga tcgcccaccc cgttactgccc ccgatcccgg 14220
caatggcaag gactgccagc gctgccattt ttgggggtgag gccgttcgcg gccgaggggg 14280
gcagcccctg ggggatggg aggcccgcgt tagcgggccc ggagggttcg agaagggggg 14340
gcacccccct tcggcgtgcg cggctcacgcg cacagggcgc agccctgggt aaaaaaagg 14400
tttataaata ttggtttaaa agcaggttaa aagacaggtt agcgggtggc gaaaaacggg 14460
cggaaacctc tgcaaagct ggattttctg cctgtggaca gccctcaaa tgtcaatagg 14520
tgcccccctc atctgtcagc actctgcccc tcaagtgtca aggatcgcgc cctcatctg 14580
tcagtagtcg cgcctctcaa gtgtcaatac cgcagggcac ttatcccag gcttgtccac 14640
atcatctgtg ggaaactcgc gtaaaatcag gcgttttcgc cgatttgca ggctggccag 14700
ctccacgtcg ccggccgaaa tcgagcctgc cctcatctg tcaacgcgc gccgggtgag 14760
tcggccctc aagtgtcaac gtcggccctc catctgtcag tgagggccaa gttttccgg 14820
aggatccac aacgcggcg gccgcggtgt ctgcacacg gcttcgacgg cgtttctggc 14880
gcgtttgcag ggccatagac ggccgcagc ccagcggcga gggcaaccag cccgg 14935

```



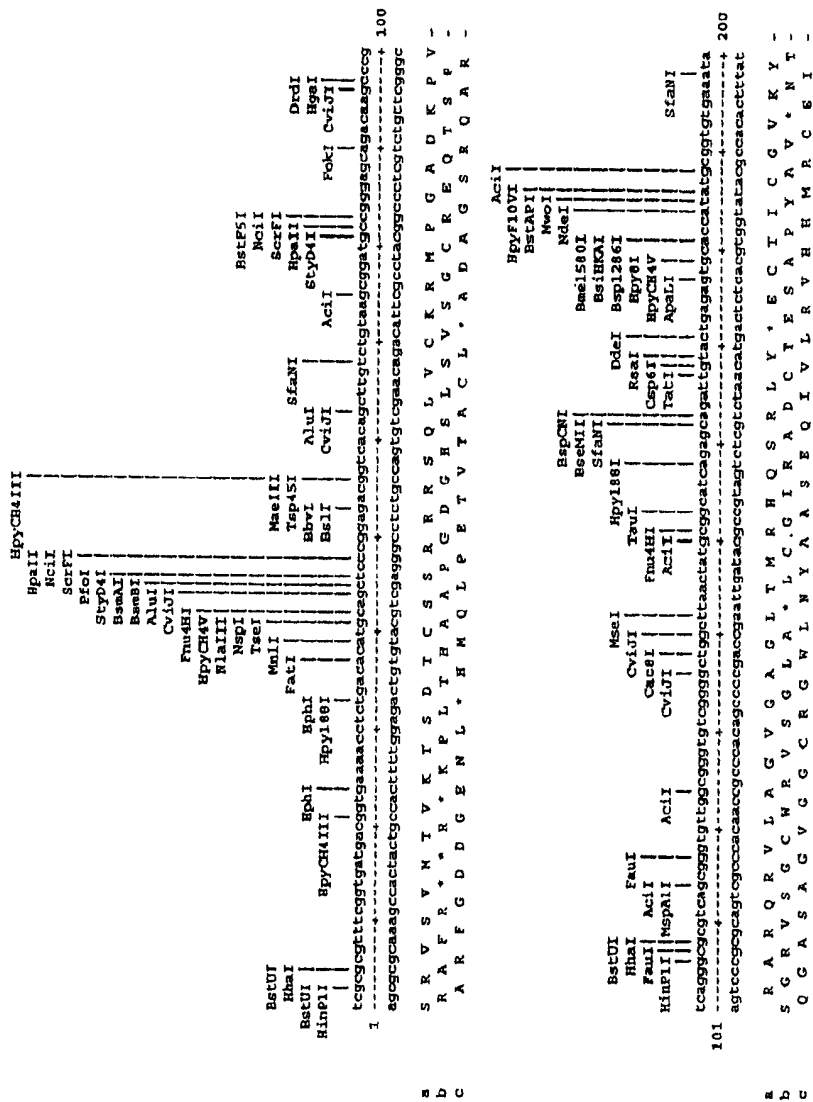


图 21-A

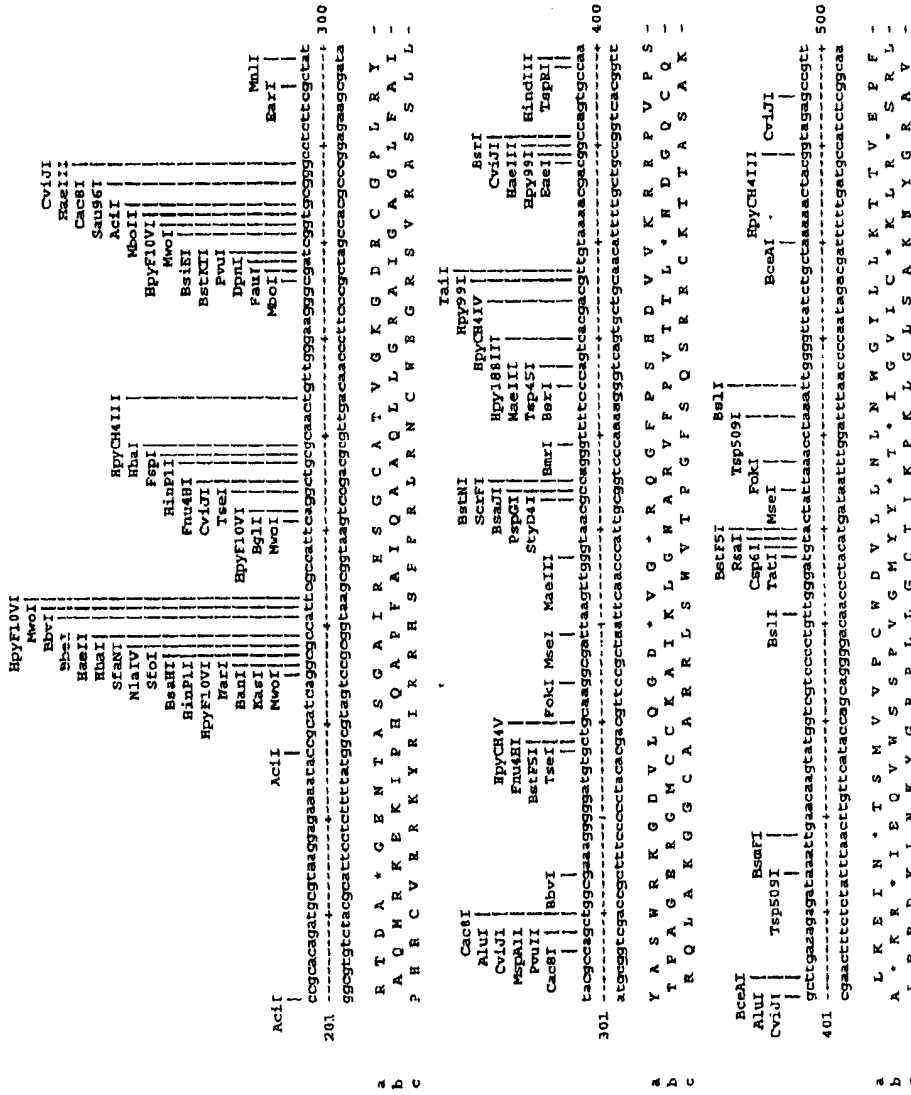


图 21-B



```

AcII
NlaIII
NspI
SphI
BsrDI
Cac8I
RpyFIOVI
NsiI
BclAPI
BsrBI
MwoI
FatI
HpyCRAV
HpyCR4II
HpyCR4III
BstFSI
cttcttcacatccacgcgcaacgcgatgggtcattgtagacagaagaatcctggtaccactcccggtattagcaaaagcgttagctatgtctggtg
801  Saagaagttagtggcagctacgtaagcacaatactctgctcttagtacataggtgaaggoccatataatcgcttctcgcaatgcataagaccac
F P T S T V Q C M R L L I D K R I M V S T S G Y * Q K R Y V C L V -
S S H P P C M A C G V * T E E S M Y P L P G I S K S V T Y V W W -
L L E I E R A M H A V I D R Q K N H G I H F R V L A K A L R N S G C -
HpyCR4III
SpmI
NlaIV
RsaI
Csp6I
BaeI
MboI
BaeI
Hpy8I
BsmAI
HphI
CviJI
gagstcatcttcaactcggtaaccgtagtggtaaaccttgaagaggaacaacttgggtctgtgattctactcgcgfgatgttttggtaaca
901  ctctagataagtggaccattggcatccatttgaactcaccatttctctgtatggaaaccggaaacaactaaatgacgcactactaaacaacttgt
E I F T L V P * V N L K V K E T * L W A L L I Y C V M I L L N K -
R S Y S L M Y R S R * T * R * K R H N F G L C * F T A * F C * T -
D H I H S G T V Y G K L E G G E R D I T L G F V D L L R D D F V E Q -

```

图 21-D

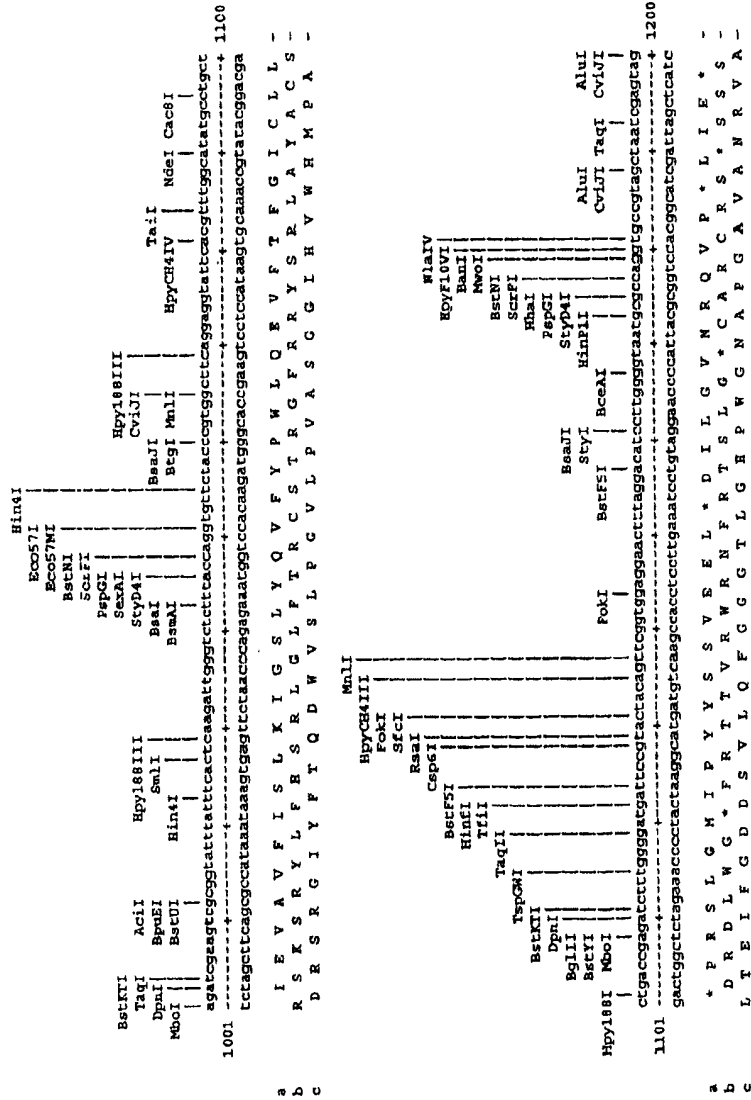


图 21-E

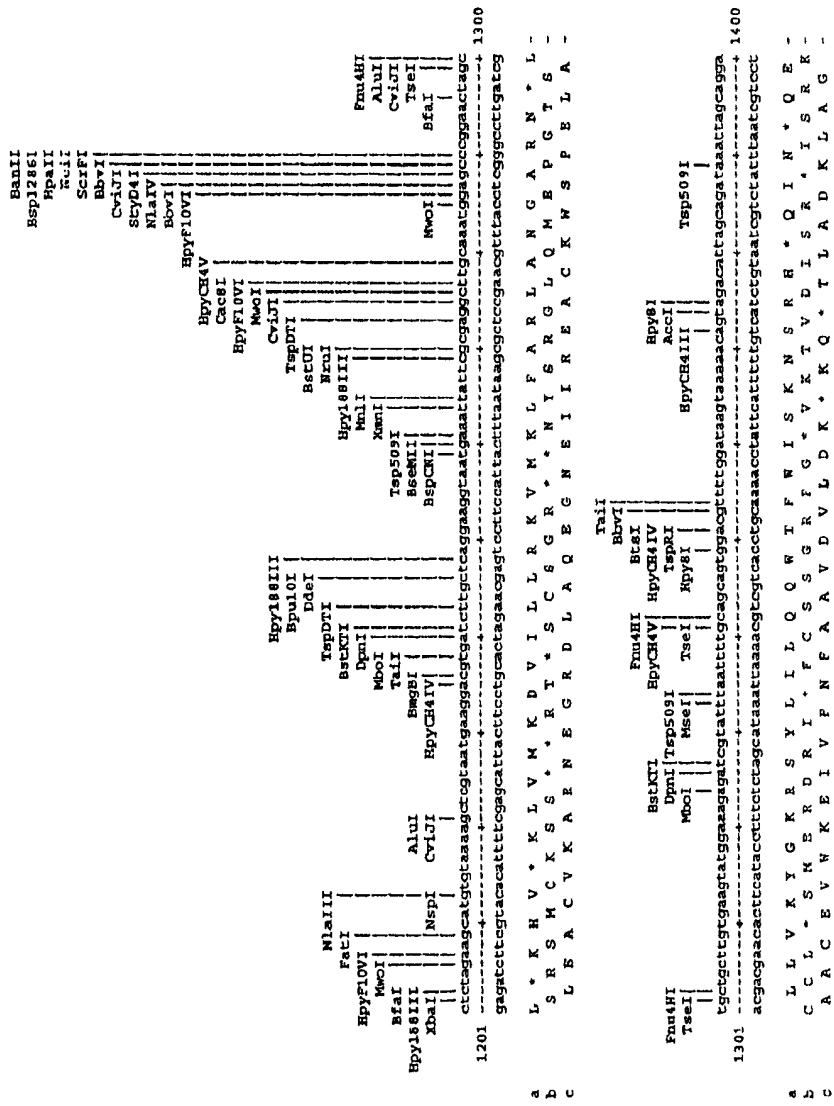


图 21-F

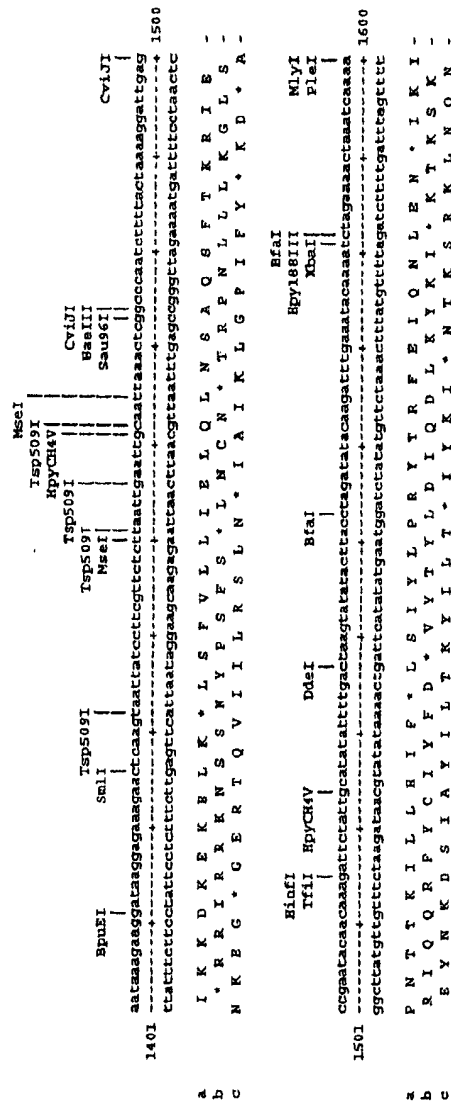


图 21-G







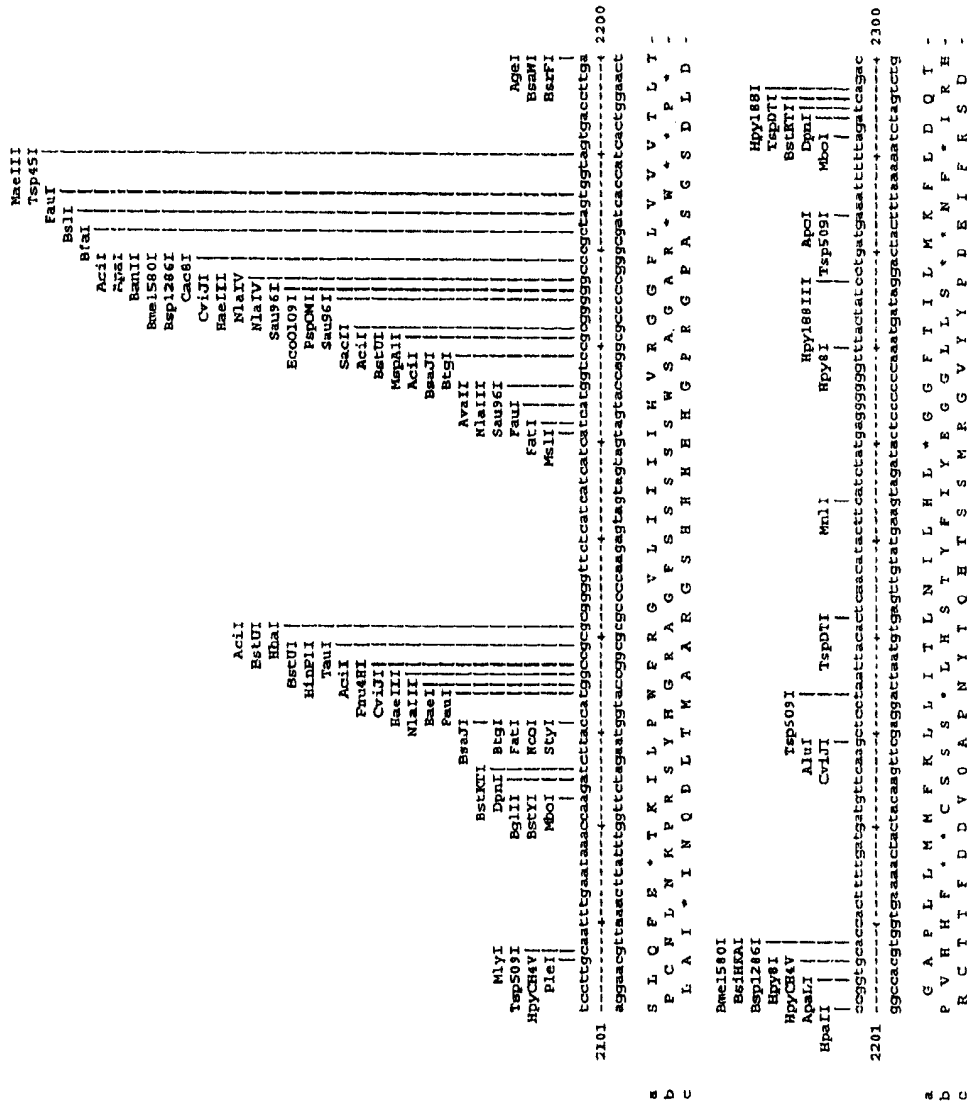


图 21-J



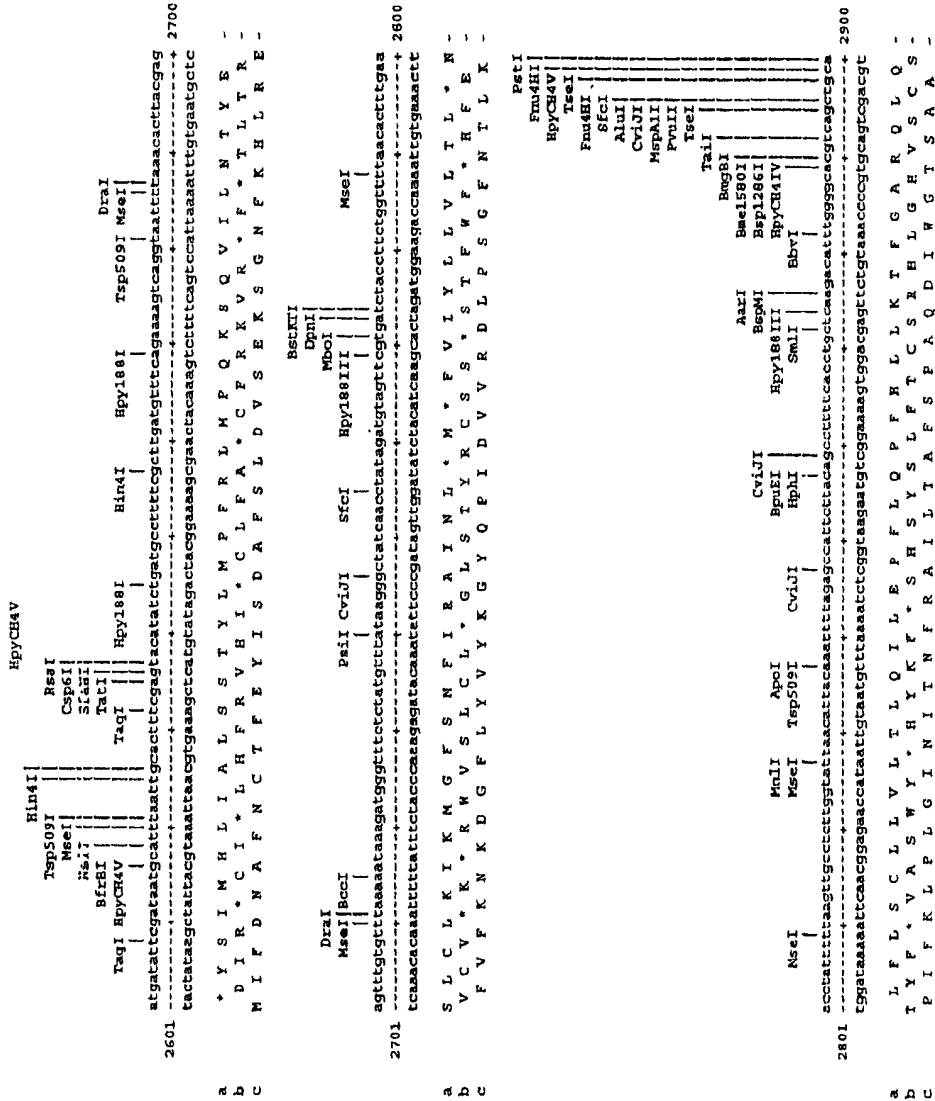


图 21-L

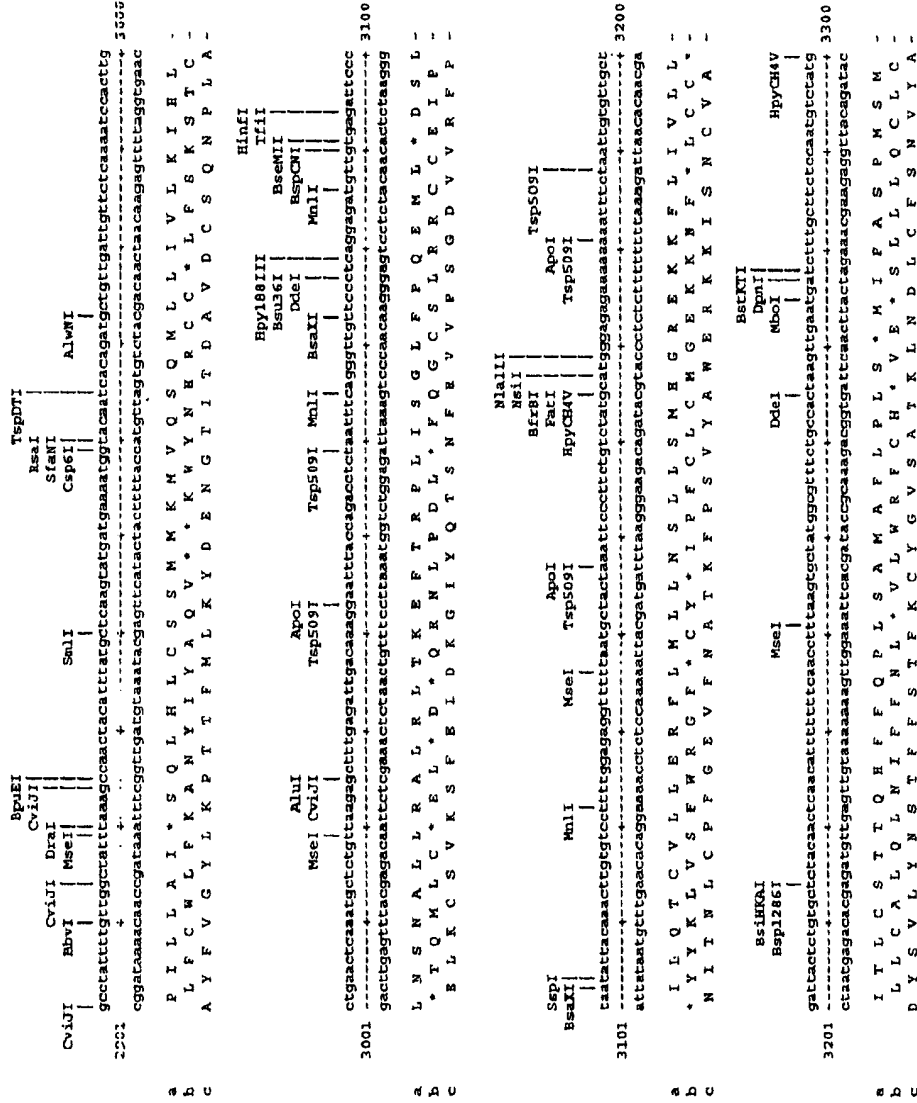


图 21-M

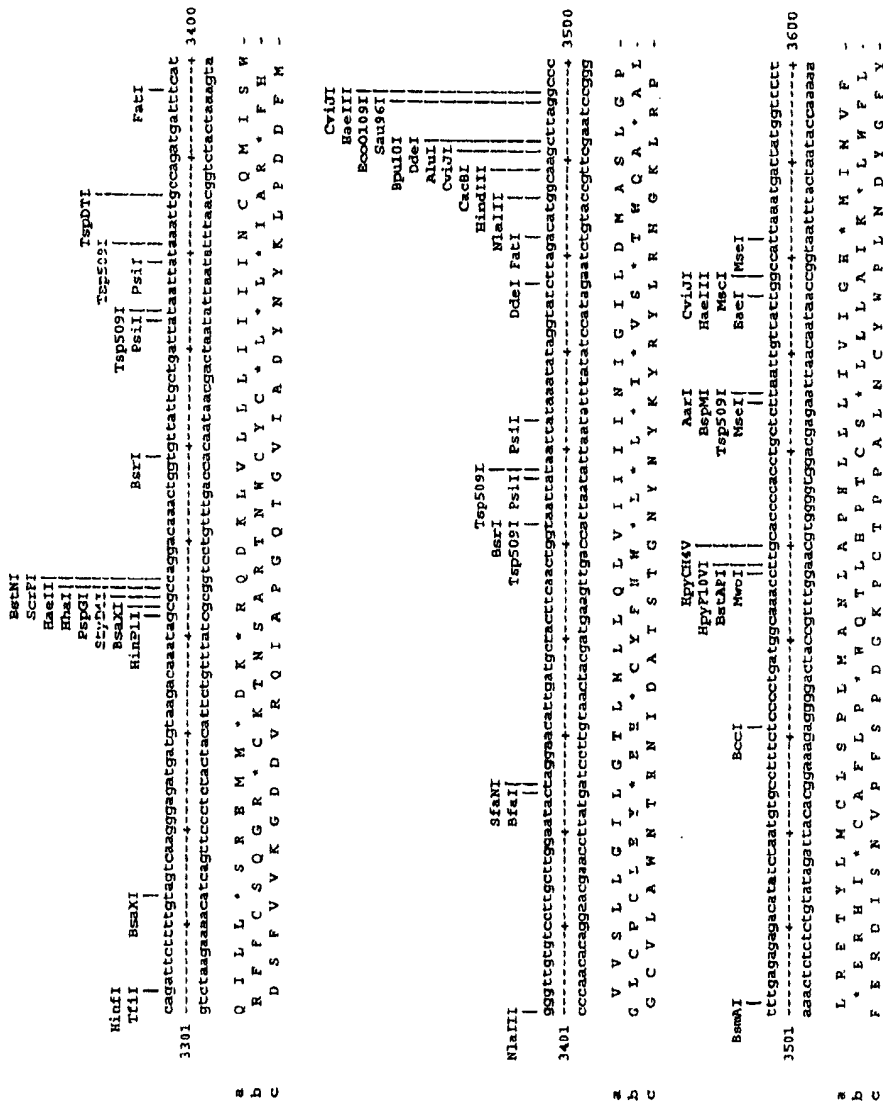


图 21-N









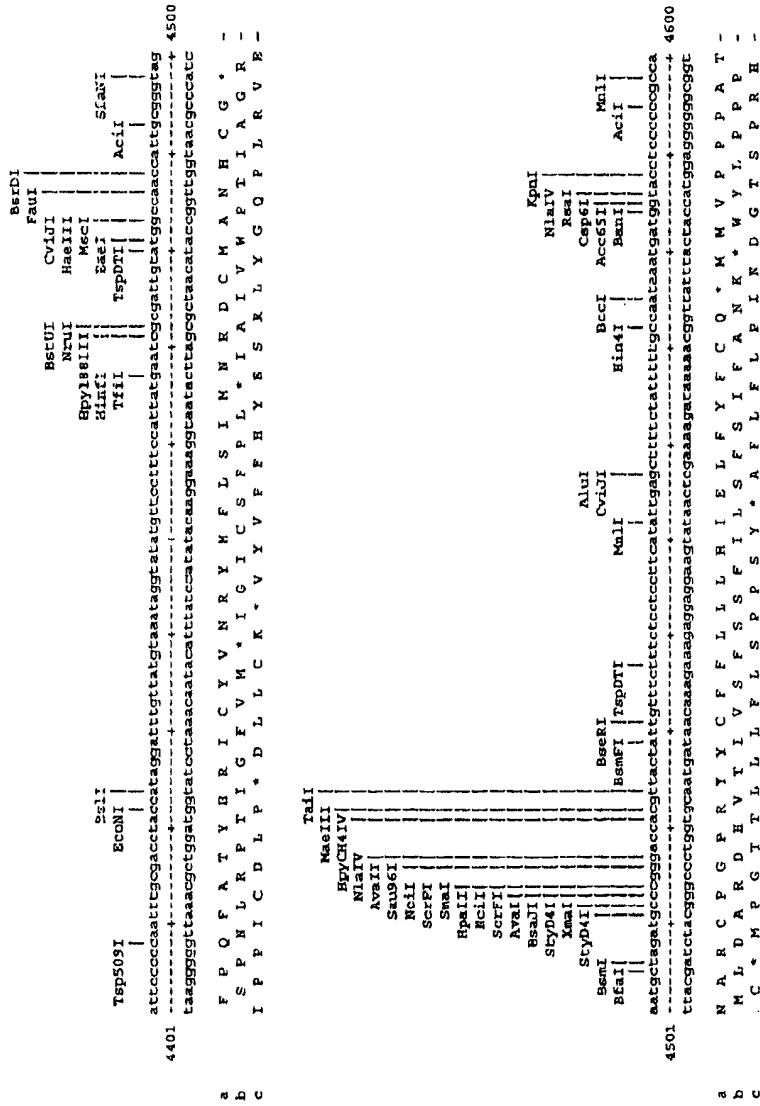


图 21-R

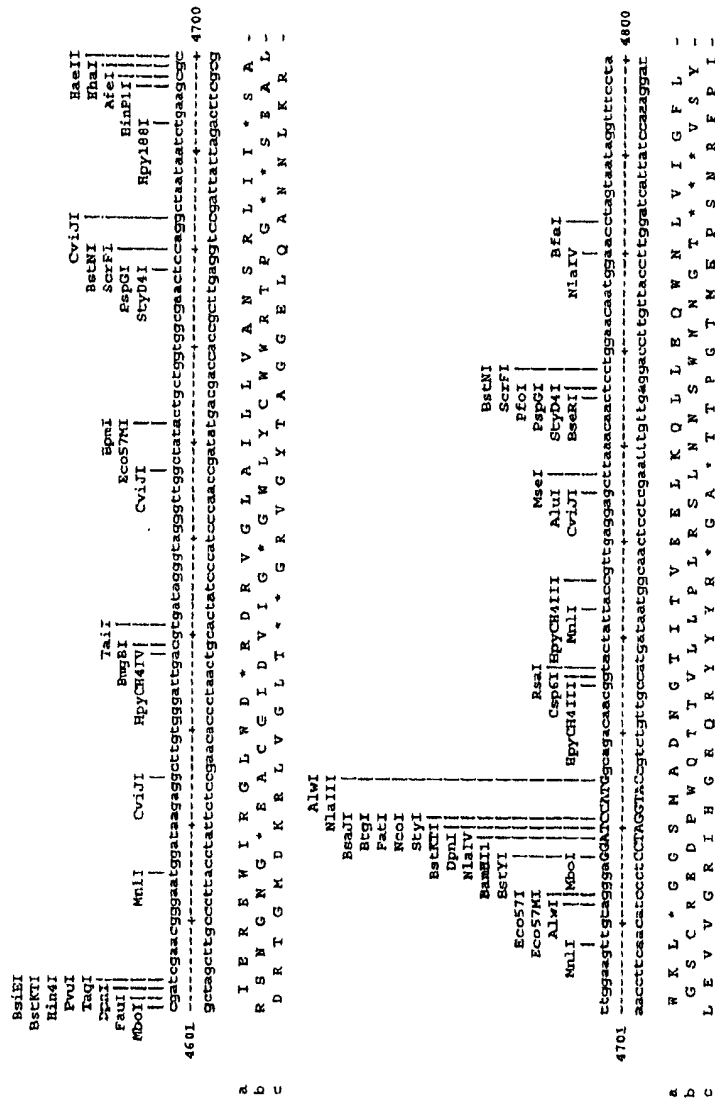


图 21-S

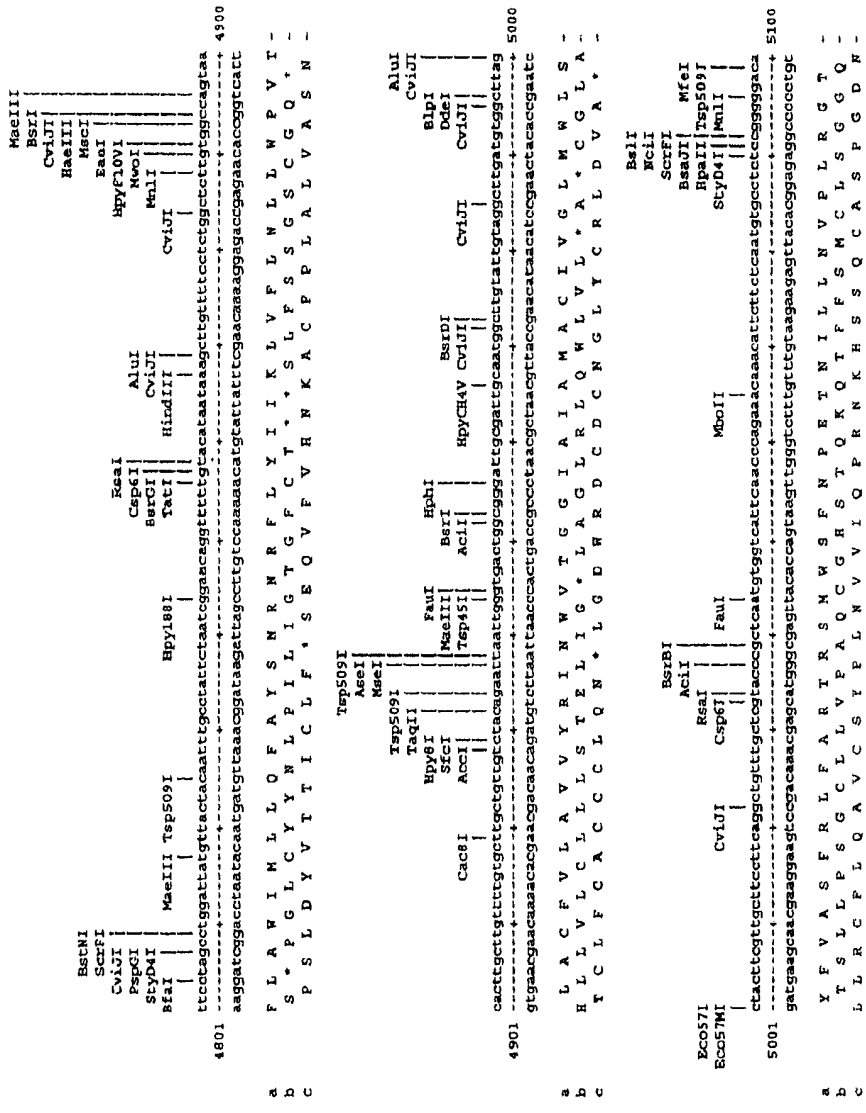


图 21-T











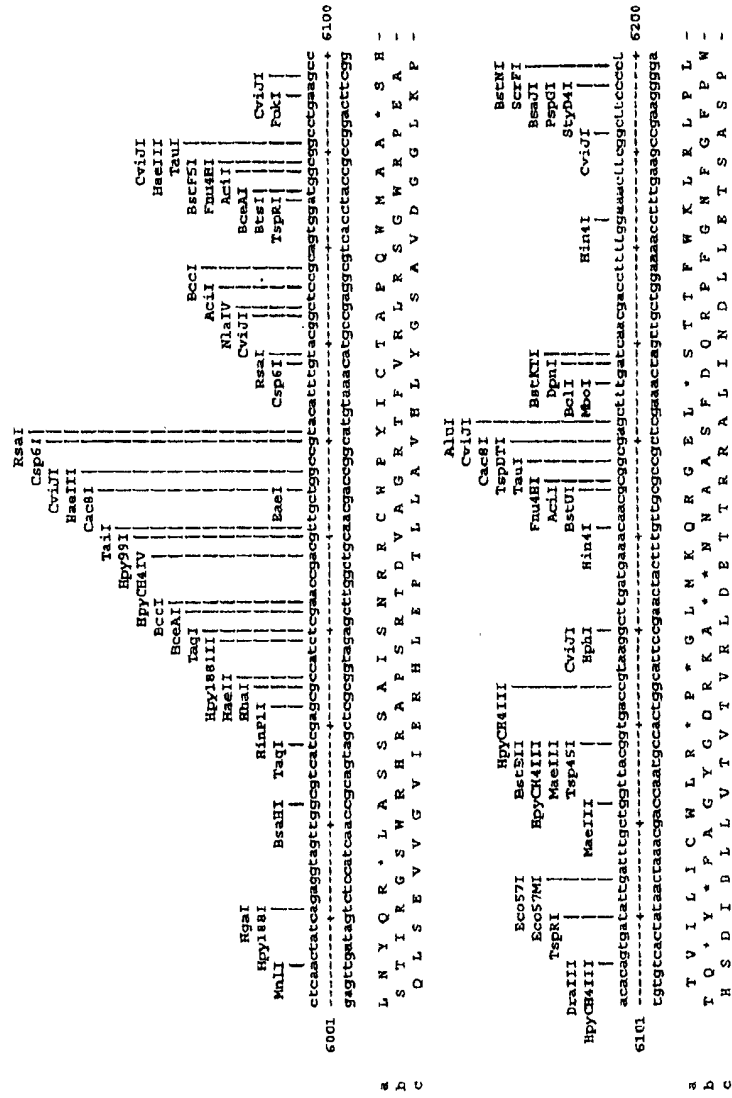


图 21-Y

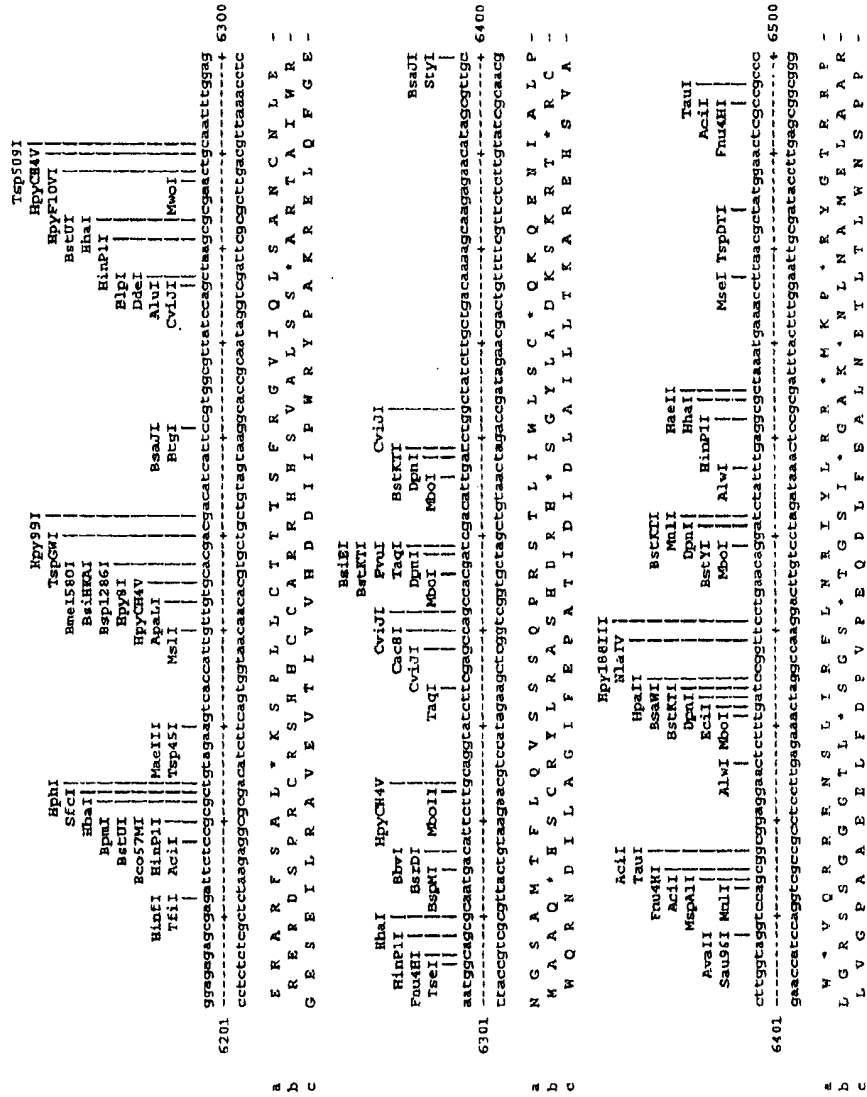


图 21-Z





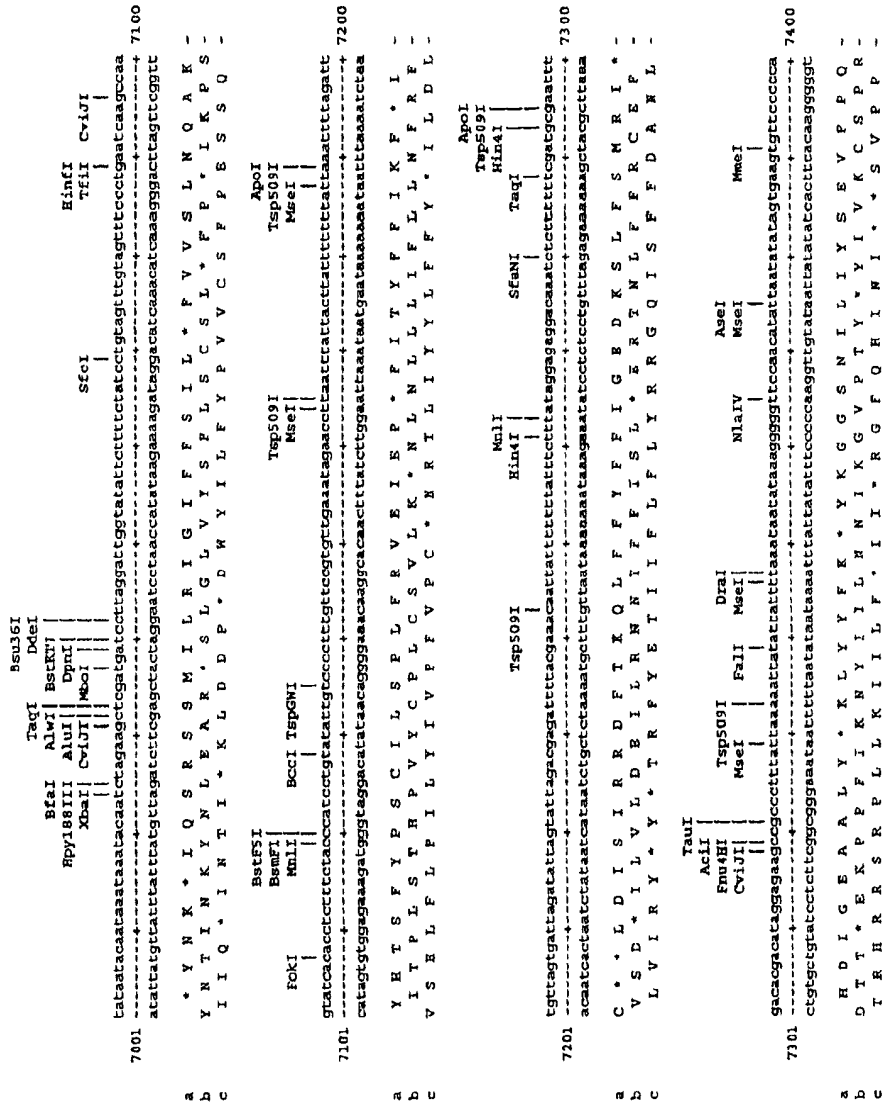


图 21-CC

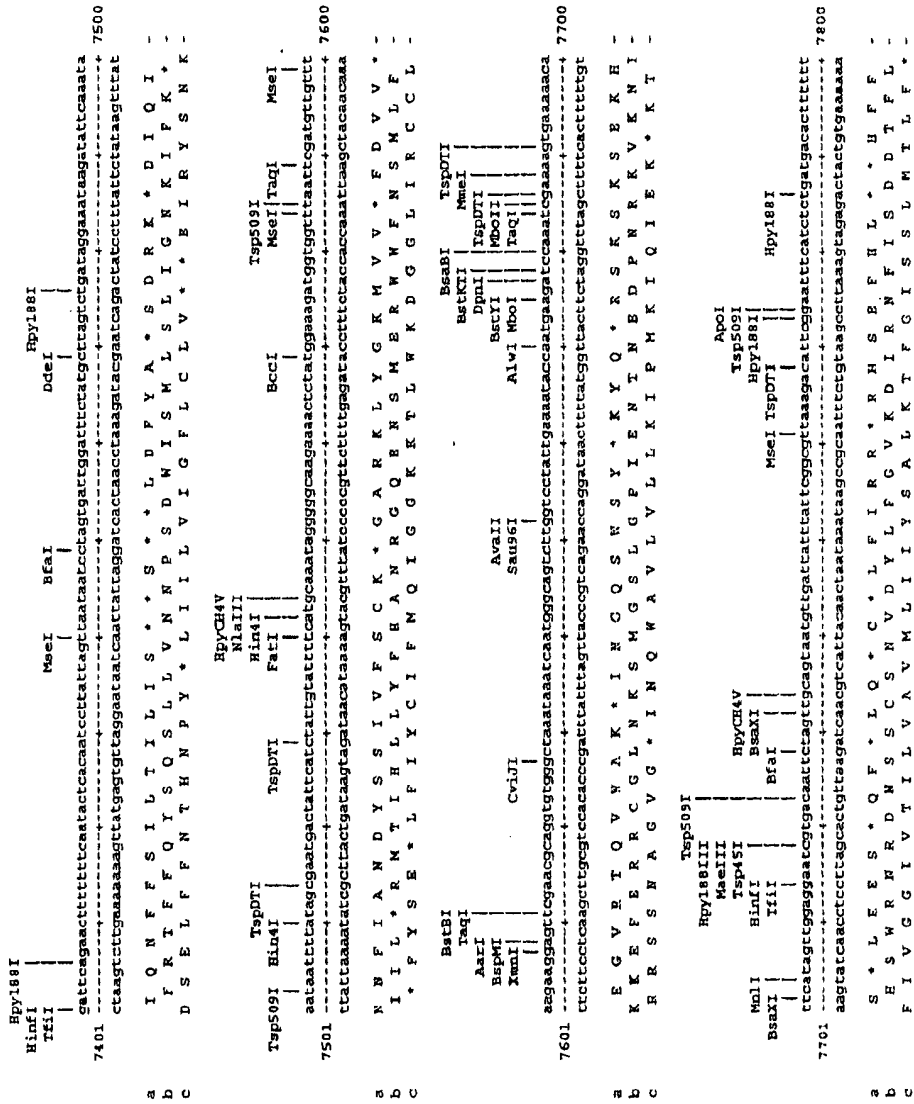


图 21-DD

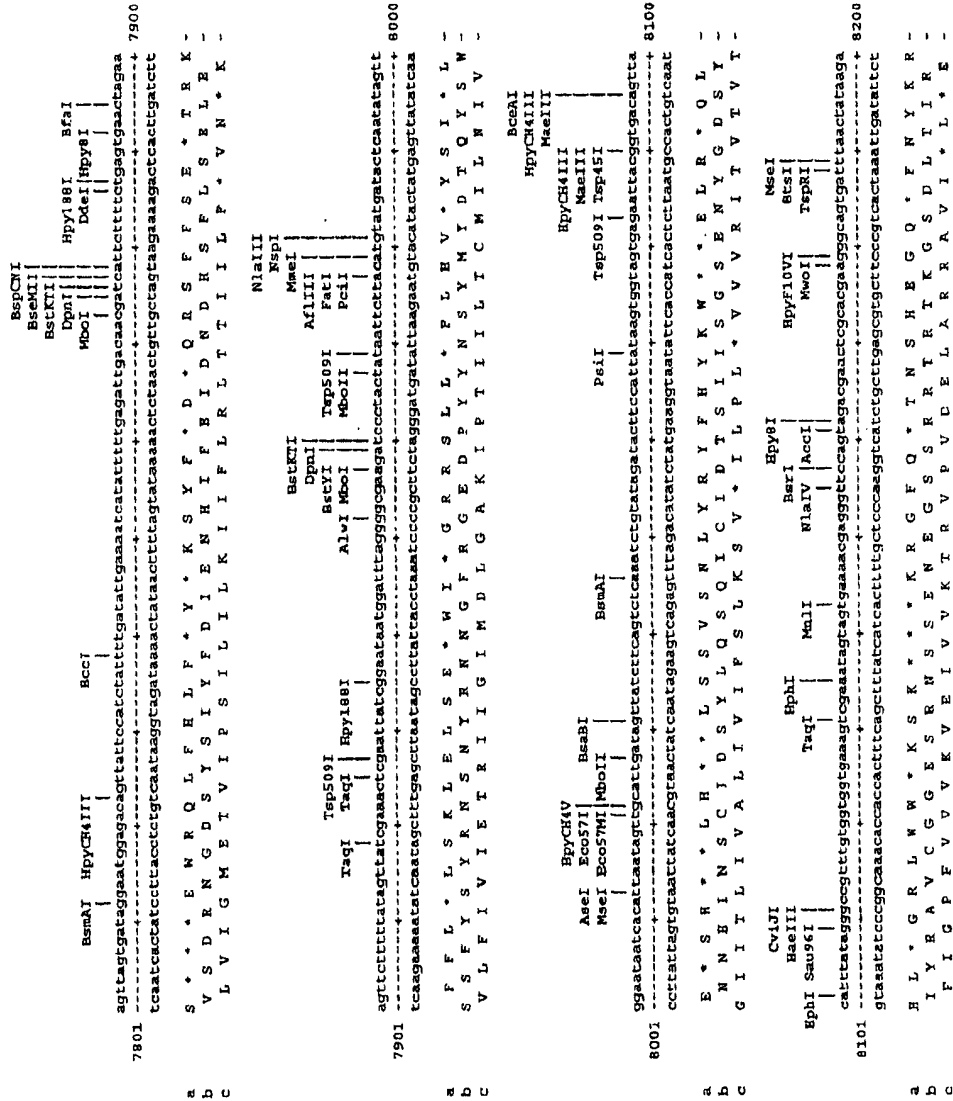


图 21-EE

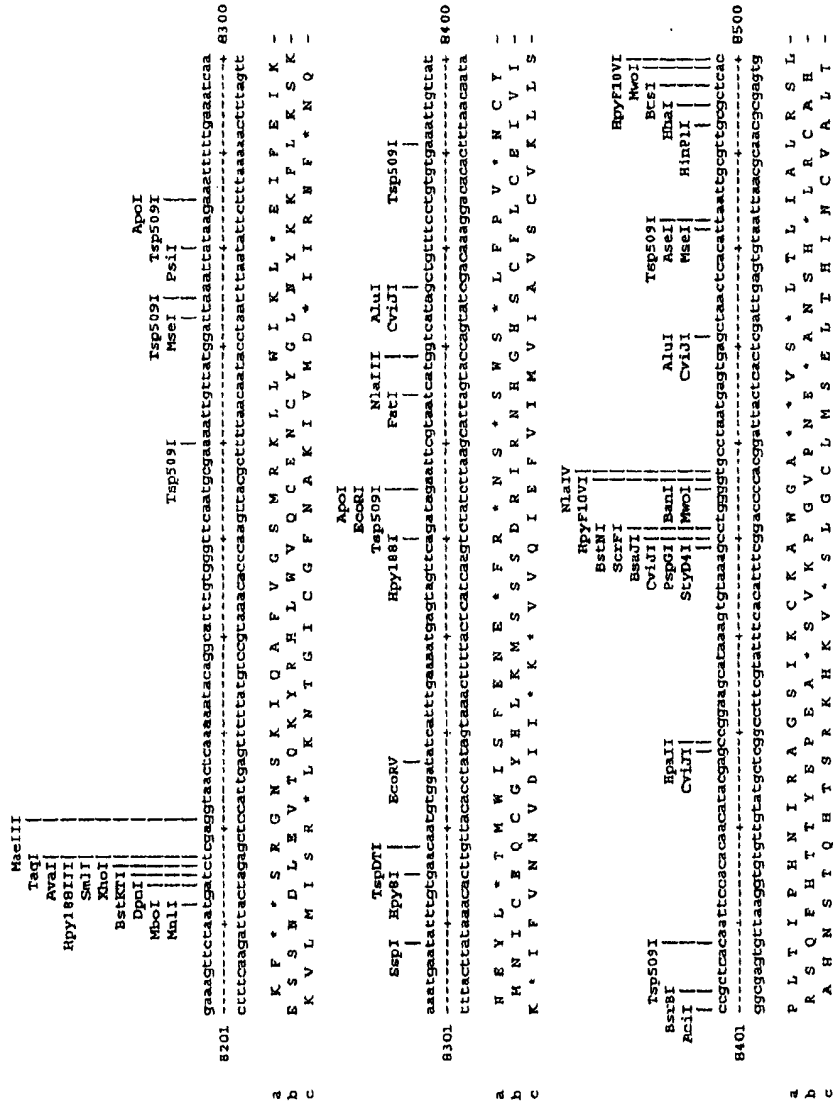


图 21-FF





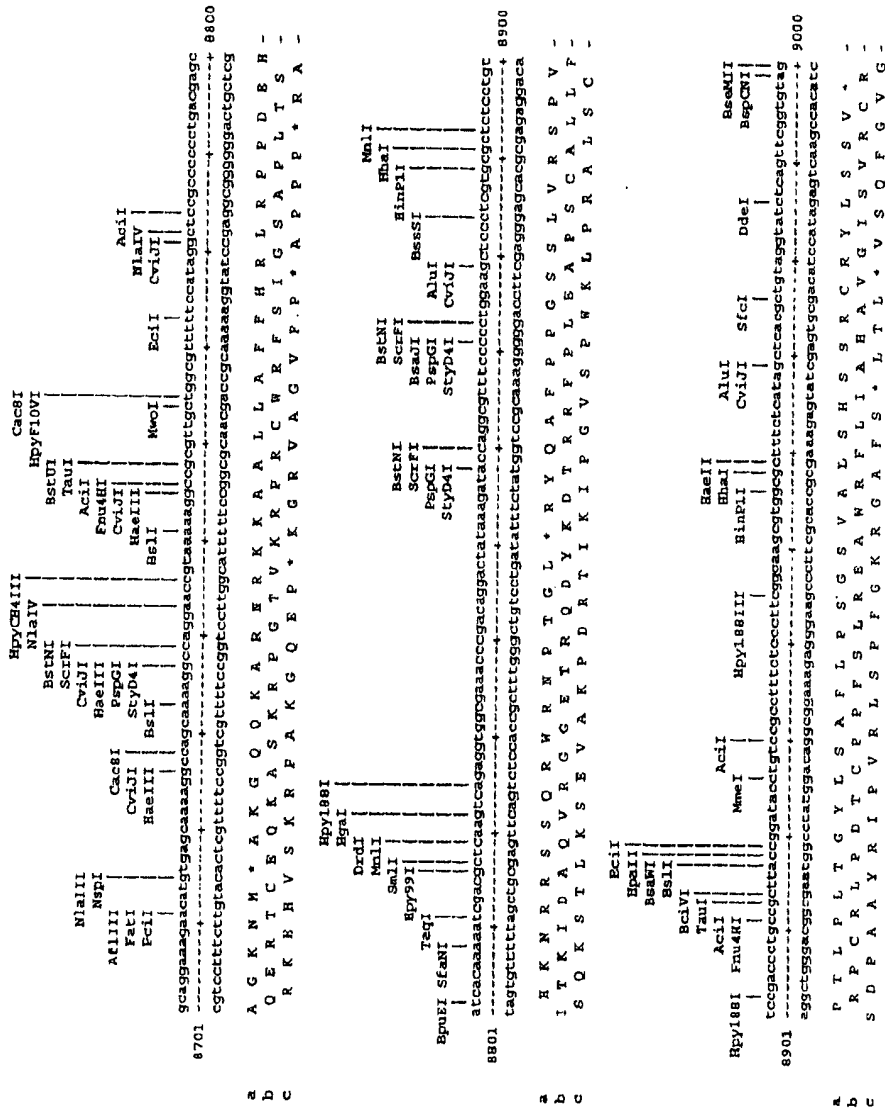


图 21-HH









```

10201      aacggtcttcggggcgaaaactctcaaggattctacgctggttgatccagttcgatgaaaccactcggcaccacacgattctcagcatcttba
      tttcgaagaagccccgctttfagagttccagatggcgaactctagcacaagctacattgggtgagcagctgggtgactagaagctgtagaaaat 10300
a   K R S S G R K L S R I L P L R L R S S M * P T R A P N * S S A S F T -
b   N V L R G E N S Q G S Y R C * D P V R C N P L V H P T D L Q H L L -
c   T F F G A E T L K D L T A V E I Q F D V T H S C T Q L I F S I F Y -

      TaqII      HpaI      EcoRI      MboII
      |          |          |          |
ctttccaccgctttctgggtgacaaaabaggaagcacaatgcccgaataaagggaatgaagggcagcagcggaaatgtgaaactcacaactcttctt 10301
      gaaagtggctcgaagaaccactcgttttbtctctcgttttaccggcgtttttcccttattcccgctgctctttacaactctcagcagagaagga 10400
a   F T S V S G * A K I G R Q N R A K K G I R A T R K C * I L I L F L -
b   L S P A F L G E Q K Q E G K M P Q K R E * G R H G N V E Y S Y S S F -
c   F R Q R F W V S K N R K A K C R K K G N K G D T E M L N T R T L P -

      AclI      BstEII
      |          |
tttccaatattgagcattatcagggattgtctcactgagggatcattttgaaatgatttgaanaataaacaaggggtccggcaca 10401
      aaaagttataaacttggtaaatgcccmaataacagagtaactcgtctataactcacaataatctttttttttatcccccaagcgctgtg
a   F Q Y Y * S I Y Q G Y C L M S G Y I F E C I * K N K Q I G V P R T -
b   F N I I E A F I R V I V S * A D T Y L N V F R K I N K * G F R A H -
c   F S I L L K H L S G L L S H S R I H I * M Y L E K * T N R G S A H I -

```

图 21-MM





pCV6 序列

```

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatategc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
attegccatt caggctcgcg aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggg 360
tttcccagtc acgacgttgt aaaacgacgg ccagtgccaa gcttgaaga gataaattga 420
acaagtatgg tcgtcccctg ttgggatgta ctattaacc taaattgggg tctctgcta 480
aaaactacgg tagagccggt tatgaatgtc ttgcgggtgg acttgatttt actaaagatg 540
atgagaacgt gaactcacia ccatttatgc gttggagaga tcgtttotta ttttgtccg 600
aagcacttta taaagcacag gctgaaacag gtgaaatcaa agggcattac ttgaatgcta 660
ctgcaggtag atgccaagaa atgatcaaaa gagctgtatt tgctagagaa ttgggcgttc 720
cgatcgtaat gcatgactac ttaacggggg gattcacccg aaatactagc ttggctcatt 780
attgccgaga taatggtcta cttcttcaca tccacogtgc aatgcatgcg gttattgata 840
gacagaagaa tcatggtatc cacttcgggg tattagcaaa agcgttacgt atgtctggtg 900
gagatcatat tcactctggt accgtagtag gtaaacttga aggtgaaga gacataactt 960
ttggccttgt tgatttactg cgtgatgatt ttggtgaaca agatcgaagt cgcggtat 1020
atctcactca agattgggtc tctttaccag gtgttctacc cgtggcttca ggaggtat 1080
acgtttggca tatgcctgct ctgaccgaga tctttgggga tgattccgta ctacagttcg 1140
gtggaggaac tttaggacat ccttggggta atgcgccagg tgccgtaget aatcgagttag 1200
ctctagaagc atgtgtaaaa gctcgtaatg aaggacgtga tcttgctcag gaaggtaatg 1260
aaattattcg ctaggcttgc aaatggagcc cggaaactagc tgetgcttgt gaagtatgga 1320
aagagatcgt atttaatttt gcagcagtgg acgttttggg taagtaaaaa cagtacatc 1380
tagcagataa attagcagga aataaagaag gataaggaga aagaactcaa gtaattatcc 1440
ttcgttctct taattgaatt gcaattaaac tcggcccaat cttttactaa aaggattgag 1500
ccgaatacaa caaagattct attgcatata ttttgactaa gtatatactt acctagatat 1560
acaagatttg aaatacaaaa tctagaaaaa taaatcaaaa tctaagactc aaatctttct 1620
attgttgtct tggatcatcg agcttgcctg cctgcagtcg actttcacag tttccattct 1680
gaaatgttct ctgtactata ataatatgta agtgaatcaa cttactacta aaaaaattag 1740
tagacttctt cttcgggaata gaaatgctct atttctacat agggaaagtc gtgtgcaatg 1800
aaaaatgcaa gcacgatttg gggagagggt ttttctctat tgtaacaagg aataattatc 1860
tactccatcc gactagttcc gggttcgagt ccggggcaac ccatatggaa actagaaagg 1920
agcaatctga gttttgattt ttcactcact tcatttacia aattttttgg tttggtaaat 1980
tttgttgat ggatatacaa ctgtcggggc tggcttggtt gacattggta tatagtctat 2040
attatactgt taaataacaa gccttctatt atctttctag ttaatacgtg tgcttgggag 2100
tccttgcaat ttgaataaac caagatctta ccatggccgc ggggggttct catcatcctc 2160
atcatggtcc gggggggccc gctagtggta gtgacctga cgggtgcacc acttttgatg 2220
atgttcaagc tcctaattac actcaacata cttcatctat ggggggggtt tactatctcg 2280
atgaaatttt tagatcagac actctttatt taactcagga tttatttctt ccattttatt 2340
ctaattgtac agggtttcat actattaatc atacgtttgg caaccctgct atacccttta 2400
aggatggtat ttattttgct gccacagaga aatcaaatgt tgtccgtggt tgggtttttg 2460
gttctacat gaacaacaag tcacagtcgg tgattattat taacaattct actaatgttg 2520
ttatacgagc atgtaacttt gaattgtgtg acaacccttt ctttgctggt tctaaacca 2580
tgggtacaca gacacatact atgatattcg ataatgcatt taattgcaat ttcagagtaca 2640
tatctgatgc cttttcgott gatgtttcag aaaagtcagg taattttaaa cacttacgag 2700
agtttgtgtt taaaaataaa gatgggttct tctatgttta taaggctat caacctatag 2760
atgtagttcg tgatctacct tctggtttta acactttgaa acctattttt aagttgctc 2820
ttggtattaa cattacaaat tttagagcca ttcttacagc cttttcaact gctcaagaca 2880

```

图 22-A

```

tttggggcac gtcagctgca gccatatttg ttggctattt aaagccaact acatttatgc 2940
tcaagtatga tgaaaatggt acaatcacag atgctgttga ttgttctcaa aatccacttg 3000
ctgaactcaa atgctctgtt aagagctttg agattgacaa aggaatttac cagacctcta 3060
atctcagggg tgttcctca ggagatgttg tyagattccc taatattaca aacttggtgc 3120
cttttgagga ggtttttaa gctactaaat tcccttctgt ctatgcatgg gagagaaaa 3180
aaatctctaa ttgtgttgc gattactctg tgctctacaa ctcaacatct tttcaacct 3240
ttaagtgcga tggcgtttct gccactaagt tgaatgatct ttgottctcc aatgtctatg 3300
cagattcttt tgtagtcaag ggagatgatg taagacaaat agcgcagga caaactggtg 3360
ttattgctga ttataattat aaattgccag atgatttcat gggttgtgtc cttgcttggg 3420
atactaggaa cattgatgct acttcaactg gtaattataa ttataaatat aggtatctta 3480
gacatggcaa gcttaggccc ttgagagag acatatctaa tggccttcc tccctgatg 3540
gcaaaccttg caccacact gctcttaatt gttattggcc attaaatgat tatggttttt 3600
acaccactac tggcattggc taccacactt acagagttgt agtactttct ttgaaacttt 3660
taaattgcacc ggccacggtt tgtggaccaa aattatccac tgaccttatt aagaaccagt 3720
gtgtcaattt taattttaa ggactcactg gtaactggtt gttactcct tcttcaaaga 3780
gatttcaacc atttcaacaa ttggccgtg atgtttctga ttactgat tccgttcgag 3840
atcctaaaac atctgaaata ttagacattt caccttgctc ttttgggggt gtaagtgtaa 3900
ttacacctgg aocaaatgct tcatctgaag ttgctgttct atatcaagat gtaactgca 3960
ctgatgttcc tacagcaatt catgcagatc aactcacacc agcttggcgc atatttcta 4020
ctggaacaaa tgtattccag actcaagcag gctgtcttat aggagctgag catgtcgaca 4080
cttcttatga gtgcgacatt cctattggag ctggcatttg tgctattacc atacagttag 4140
actagaaaag ggcgaattct gcagatattc atcacactgg cggccgcctc gatgagctct 4200
aattaattaa ggcttttctg ctaacatata gcaatttttg aagaaaggaa agctagaaat 4260
acccaatata ttgctgaagc aagatattgg gtatttcttt ttttttatt ttgaatcttt 4320
ctattctgaa ttoagttaac gacgagattt agtatccttt cttgcacttt cataactcgt 4380
gaaatgcoga gttggtaoga attccccaa ttggcgaact aoccataggat ttgttatgta 4440
aataggtata tgttccttcc cattatgaat cgcgattgta tggccaacca ttgcgggtag 4500
aatgctagat gcccgggacc acgttactat tgtttcttcc tccctcctca tattgagctt 4560
ttctattttt gccaaataat gatggtaoat cccccgcga cgcgcgaacg ggaatggata 4620
agaggcttgt gggattgacg tgatagggtg gggttggcta tactgctggg ggcgaactcc 4680
aggctaataa tctgaagcgc ttggaagtgt tagggaggat ccatggcaga caacgggtact 4740
attaccgttg aggagcttaa caaactcctg gaacaatgga acctagtat aggtttccta 4800
ttcctagcct ggattatggt actacaattt gccatttcta atcggaacag gtttttctac 4860
ataataaagc ttgtttctct ctggctcttg tggccagtaa cacttgcttg ttttctgctt 4920
gctgttctct acagaattaa ttgggtgact ggcgggattg cgattgcaat ggcttctatt 4980
gtaggcttga tgtggcttag ctacttctgt gcttcttca ggctgtttgc tctgaccgcg 5040
tcaatgtggg cattcaaccc agaaacaaac attcttctca atgtgcctct cggggggaca 5100
attgtgacca gaccgctcat ggaaagtga cttgtcattg gtgctgtgat cattctgggt 5160
cacttgcgaa tggccggaca ctccctaggg cgcgttgaca ttaaggacct gccaaaagag 5220
atcactgtgg ctacatcacg aacgcttctc tattacaaat taggagogtc gcagcgtgta 5280
ggcactgatt caggttttgc tgcatacaac cgctaccgta ttggaaacta taaattaaat 5340
acagaccacg ccggtagcaa cgacaatatt gctttgctag tacagtaaga attcaagccg 5400
aattccagca cactggcggc cgtaactagt ggatccgagc tctaattaat taaggctttt 5460
ctgctaacat atagcaattt ttgaagaaag gaaagctaga aatacccaat atcttcttga 5520
agcaagatat tgggtatttc tttttttttt attttgaatc ttcttattct gaattcagtt 5580
aacgacgaga tttagtatcc tttcttgca tttcataact cgtgaaatgc cgagttggta 5640
cgaattcccc caatttgoga cctaccatag gatttgttat gtaaataggg atatgttctt 5700
ttocattatg aatcgcgatt gtatggocaa ccattgcggg tagaatgcta gatgccggg 5760
accacgttac tattgtttct ttotcctct tcatattgag cttttctatt tttgccaata 5820
aatgatggta cctcccccg ccacgatcga acgggaatgg ataagaggct tgtgggattg 5880
acgtgatagg gtagggttgg ctatactgct ggtggcgaac tccaggctaa taactggaag 5940
cgcttggag ttgtagggag gatcctatgg ctctgtaagc ggttatcgcc gaagtatcga 6000
ctcaactatc agaggtagtt ggcgctatcg agcgcctatc cgaaccgacg ttgctggccg 6060

```

图 22-B

```

tacatttga cggctcogca gtggatggcg gcctgaagcc acacagtgat attgatttgc 6120
tggttacggg gaccgtaagg cttgatgaaa caacogggcg agctttgatc aacgaccttt 6180
tggaaacttc ggcttcccc ggagagagcg agattctccg cgctgtagaa gtcaccattg 6240
ttgtgcacga cgacatcatt ccgtggcggt atccagctaa gcgcgaactg caatttggag 6300
aatggcagcg caatgacatt cttgcaggta tcttcgagcc agccacgatc gacattgatc 6360
tggctatctt gctgacaaaa gcaagagaac atagcgttgc cttggtaggt ccagcggcgg 6420
aggaacctct tgatccgggt cctgaacagg atctatttga ggccgctaaat gaaaccttaa 6480
cgctatggaa ctgcgcccgc gactgggctg gcgatgagcg aaatgtagtg cttacgttgt 6540
cccgcatttg gtacagcgca gtaaccggca aaatcgcgcc gaaggatgtc gctgccgact 6600
gggcaatgga gcgcctgccc gcccgatc agcccgctcat acttgaagct agacaggcct 6660
atcttggaca agaagaagat cgcttggcct cgccgcgaga tcagttggaa gaatttgtcc 6720
actacgtgaa aggcgagatc actaaggtag ttggcaata actcgagact aagtggataa 6780
aattagatag aaaaaaggtc taataaaaa agaagagaaa tagaaagatc aaaaatcagt 6840
tacgaaaaatg cagtaattct tctttttct tctaattgat tgcaattaaa ctcgctcaa 6900
tctgaaaaaa gattgagccg agtttaata gattttgata cgatcatgag acttgacaaa 6960
tcgggattcc tctattctat atatttagaa gatataaagg tataatacaa taaataaata 7020
caatctagaa gctcgatgat ccttaggatt ggtatattct tttctatcct gtagtttga 7080
gtttccctga atcaagccaa gtatcacacc tctttctacc catcctgtat attgtccct 7140
ttgttccgtg ttgaaataga accttaattt attactatt tttttattaa attttagatt 7200
tgttagtgat tagataatag tattagaoga gattttacga aacaattatt tttttatttc 7260
tttataggag aggacaaatc tctttttctg atgcgaattt gacacgacat aggagaagcc 7320
gccctttatt aaaaattata ttattttaa taatataaag ggggttccaa catattaata 7380
tatagtgaag tgttccccca gattcagaac ttttttcaa tactcaaat ccttatagtt 7440
taataatcct agtgattgga tttctatgct tagtctgata gaaataaga tattcaata 7500
aataatttta tagcgaatga ctattcatct attgtatttt catgcaata gggggcaaga 7560
aaactctatg gaaagatggt ggtttaatc gatgttgtt aagaaggagt tcgaacgcag 7620
gtgtgggcta aataaatcaa tgggcagtct tggctctatt gaaatacca atgaagatcc 7680
aaatgaaaa gtgaaaaaca ttcatagttg gaggaatcgt gacaattcta gttgcagtaa 7740
tgttgattat ttattcggcg ttaaagacat tcggaatttc atctctgatg acactttttt 7800
agttagtgat aggaatggag acagttattc catctatttt gatattgaaa atcatatttt 7860
tgagattgac aacgatcatt cttttctgag tgaactagaa agttcttttt atagttatcg 7920
aaactcgaat tatcggaaat atggatttag gggcgaagat cctactata attcttcat 7980
gtatgatact caatatagtt ggaataatca cattaatagt tgcattgata gttatcttca 8040
gbtcaaatc tgtatagata cttccattat aagtggtagt gagaattacg gtgacagtta 8100
catttatagg gccgtttgtg gtggtgaaag tcgaaatagt agtgaaaacg agggttccag 8160
tagacgaact cgcacgaagg gcagtgattt aactataaga gaaagtcta atgatctca 8220
ggtaactcaa aaatacagcg atttgtgggt tcaatgcgaa aattgttatg gattaaatta 8280
taagaaatth ttgaaatcaa aaatgaatat ttgtgaacaa tgtggatc atttgaaat 8340
gagtagttca gatagaatc gtaatcatgg toatagctgt ttctgtgtg aaattgttat 8400
ccgctcacia tccacacaaa catacagacc ggaagcataa agtgtaaagc ctggggtgcc 8460
taatgagtga gctaactcac attaattgcg ttgcgctcac tgcccgttt ccagtcggga 8520
aacctgtcgt gccagctgca ttaatgaatc ggccaaagcg cggggagagg cggtttgcgt 8580
attgggogct cttccgtctc ctgcctcact gactcgctgc gctcggtcgt tgggctgccc 8640
cgagcgggat cagctcactc aaagcgggta atacggttat ccacagaatc aggggataac 8700
gcaggaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggcccg 8760
ttgctggcgt tttccatag gctccgccc cctgacgagc atcaaaaaa tcgacgctca 8820
agtcagaggt ggcaaaacc gacaggacta taaagatacc aggcgtttcc cctggaagc 8880
tcctcgtgct gctctcctgt tccgacctg ccgcttaccg gatacctgct cgccttctc 8940
ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag 9000
gtcgttcgct ccaagctggg ctgtgtgcac gaacccccg ttcagcccga ccgctgcgc 9060
ttatccggta actatcgtct tgagtcaaac ccggtaagac acgacttacc gccactggca 9120
gcagccactg gtaacaggat tagcagagcg aggtatgtag cgggtgtac agagttctt 9180
aagtgtggc ctaactacgg ctacactaga aggacagtat ttggtatctg cgctctgctg 9240

```

图 22-C

```

aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaaa aaccaccgct 9300
ggtagcggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa 9360
gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgtaa 9420
gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaa 9480
tgaagtttta aatcaatcta aagtatata gagtaacctt ggtctgacag ttaccaatgc 9540
ttaatcagtg aggcacctat ctcagcgatc tgtctatttc gttcatccat agttgcctga 9600
ctccccgtcg tgtagataac tacgatacgg gagggcttac catctggccc cagtgcctga 9660
atgataccgc gagaccacg ctccaccggt ccagatttat cagcaataaa ccagccagcc 9720
ggaagggccc agcgcagaag tggctctgca actttatccg cctccatcca gtctattaat 9780
tggtgcggg aagctagagt aagtagtctg ccagttaata gtttgcgcaa cgttgttgc 9840
attgctacag gcatcgtggt gtcacgctcg tcgtttgta tggcttcatt cagctccggt 9900
tcccaacgat caaggcgagt tacatgatcc cccatggtgt gcaaaaaagc ggtagctcc 9960
ttcggctctc cgatcgttgt cagaagtaag ttggccgtag tgttatact catggttatg 10020
gcagcactgc ataattctct tactgtcatg ccatccgtaa gatgctttc tgtgactggt 10080
gagtactcaa ccaagtcatt ctgagaatag tgtatgcggc gaccgagttg ctcttgccc 10140
gcgtcaatac gggataatac cgcgccacat agcagaactt taaaagtgct catcattgga 10200
aaacgttctt cggggcgaaa actctcaagg atcttacgcg tgttgagatc cagttcgatg 10260
taaccactc gtgcacccaa ctgatctca gcactttta ctttcaocag cgtttctggg 10320
tgagcaaaaa caggaaggca aaatgccga aaaaaggaa taagggcgac acggaaatgt 10380
tgaatactca tactcttctt ttttcaatat tattgaagca tttatcaggg ttattgtctc 10440
atgagcggat acatatttga atgtatttag aaaaataaac aaatagggt tccgcgcaca 10500
tttccccgaa aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat 10560
aaaaataggc gtatcacgag gccctttcgt c 10591

```

图 22-D

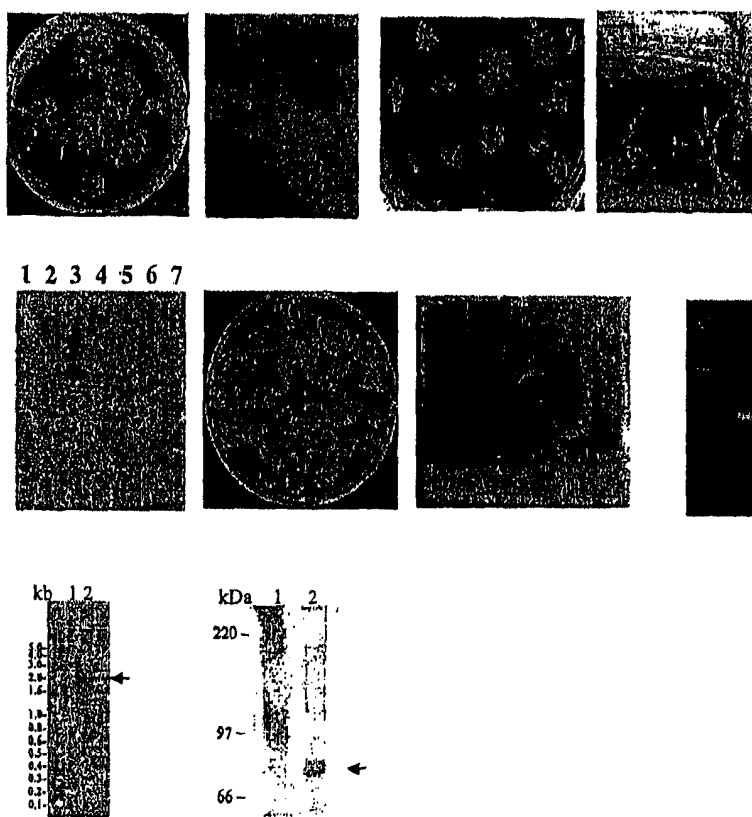


图 23

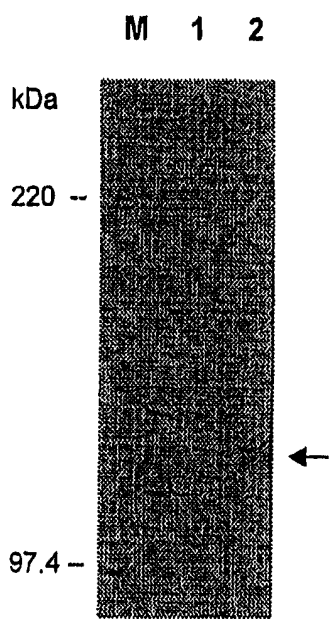


图 24