

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12N 15/12

C12N 15/62 C12N 15/63

C07K 14/50 C07K 16/18

A61K 38/18 A61P 3/04

A01K 67/027



[12] 发明专利申请公开说明书

[21] 申请号 200410046535.9

[43] 公开日 2005 年 1 月 5 日

[11] 公开号 CN 1560249A

[22] 申请日 2000.3.9

[21] 申请号 200410046535.9

分案原申请号 00815274.8

[30] 优先权

[32] 1999.9.8 [33] US [31] PCT/US99/20594

[32] 1999.9.15 [33] US [31] PCT/US99/21090

[32] 1999.12.2 [33] US [31] PCT/US99/30999

[32] 2000.2.22 [33] US [31] PCT/US00/04414

[71] 申请人 杰南技术公司

地址 美国加利福尼亚州

[72] 发明人 蒂莫西·A·斯图尔特

伊丽莎白·汤姆林森

[74] 专利代理机构 北京市柳沈律师事务所

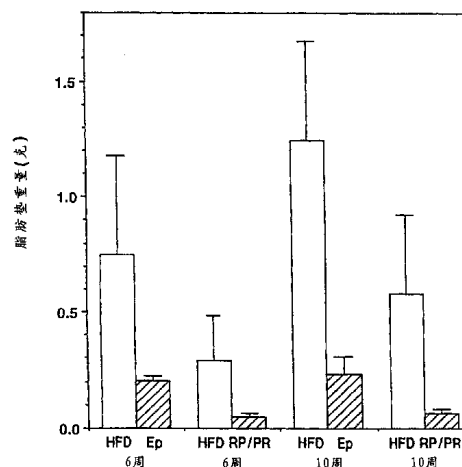
代理人 巫肖南 封新琴

权利要求书 5 页 说明书 94 页 附图 9 页

[54] 发明名称 成纤维细胞生长因子 - 19 (FGF - 19) 的核酸和多肽以及用于治疗肥胖的方法

[57] 摘要

本发明涉及成纤维细胞生长因子家族的新多肽和编码这些多肽的核酸分子。本发明也提供包含这些核酸序列、含有本发明多肽与异源多肽序列融合的嵌合多肽分子、与本发明多肽结合的抗体的载体和宿主细胞，以及涉及本发明多肽的生产方法。另外，也提供治疗肥胖的方法。



1. 分离的核酸分子, 其包括与(a) 1997年11月21日在ATCC保藏的ATCC保藏号为209480 (DNA49435-1219)的人蛋白cDNA的全长多肽编码序列的DNA分子, 或者(b) (a)所述编码序列的互补链, 具有至少约80%序列同一性的DNA。
2. 权利要求1所述分离的核酸分子, 其包括1997年11月21日在ATCC保藏的ATCC保藏号为209480 (DNA49435-1219)的人蛋白cDNA的全长多肽编码序列。
3. 编码FGF-19多肽的分离的核酸分子, 其包括能与编码图2 (SEQ ID NO: 2)中1或约23至约216位氨基酸的核酸序列的互补链杂交的DNA。
4. 权利要求3所述分离的核酸分子, 其中编码图2 (SEQ ID NO: 2)中1或约23至约216位氨基酸的核酸包括图1 (SEQ ID NO: 1)所示464或约530至约1111位的核苷酸。
5. 权利要求3所述分离的核酸分子, 其中杂交发生在严格杂交和洗涤条件下。
6. 分离的核酸分子, 其含有(a) 编码与图2 (SEQ ID NO: 2)中从1或约23至约216氨基酸残基的序列相比时, 阳性评分至少80%的多肽的DNA, 或(b) (a)所述DNA的互补链。
7. 含有至少约22个核苷酸的分离的核酸分子, 其如下产生: 在严格杂交条件下, 使测试DNA分子与(a) 编码含有图2 (SEQ ID NO: 2)中从1或约23至约216氨基酸残基序列的FGF-19多肽的DNA分子, 或(b) (a)所述DNA分子的互补链进行杂交, 并分离该测试DNA分子。
8. 权利要求7所述分离的核酸分子, 其与(a)或(b)具有至少约80%的序列同一性。
9. 分离的核酸分子, 其包含与(a)编码FGF-19多肽的DNA分子, 所述多肽包括图2 (SEQ ID NO: 2)中约1或约23至约216位氨基酸残基的序列, 或(b) (a)所述DNA分子的互补链, 具有至少约80%序列同一性的DNA。
10. 权利要求9所述分离的核酸分子, 包括图1 (SEQ ID NO: 1)中约464或约530至约1111位的核苷酸序列。
11. 权利要求9所述分离的核酸分子, 包括图1 (SEQ ID NO: 1)所示核

核苷酸序列。

12. 权利要求 9 所述分离的核酸分子，包括编码图 2 (SEQ ID NO: 2) 中约 1 或约 23 至约 216 位氨基酸残基的序列的核苷酸序列。

13. 分离的核酸分子，其包括与(a)编码由 1997 年 11 月 21 日在 ATCC 5 保藏的 ATCC 保藏号为 209480 (DNA49435-1219)的人蛋白 cDNA 所编码的相同成熟多肽的 DNA 分子，或者(b) (a)所述 DNA 分子的互补链，具有至少约 80% 序列同一性的 DNA。

14. 权利要求 13 所述分离的核酸分子，其包括编码由 1997 年 11 月 21 日在 ATCC 保藏的 ATCC 保藏号为 209480 (DNA49435-1219)的人蛋白 cDNA 10 所编码的相同成熟多肽的 DNA。

15. 含有权利要求 1 至 14 中任一核酸分子的载体。

16. 权利要求 15 所述载体，其中所述核酸分子与能被用该载体转化了的宿主细胞识别的控制序列可操作相连。

17. 以 ATCC 登记号 209480 (DNA49435-1219)保藏的核酸分子。

18. 含有权利要求 15 所述载体的宿主细胞。

19. 权利要求 18 所述宿主细胞，其中所述细胞是 CHO 细胞。

20. 权利要求 18 所述宿主细胞，其中所述细胞是大肠杆菌。

21. 权利要求 18 所述宿主细胞，其中所述细胞是酵母细胞。

22. 制备 FGF-19 多肽的方法，包括在适宜表达所述 FGF-19 多肽的条 20 件下培养权利要求 18 所述宿主细胞，并从细胞培养物中回收所述 FGF-19 多肽。

23. 分离的 FGF-19 多肽，包括与图 2 (SEQ ID NO: 2)中从 1 或约 23 至约 216 位氨基酸残基的序列至少约 80% 序列同一性的氨基酸序列。

24. 权利要求 23 所述分离的 FGF-19 多肽，其含有图 2 (SEQ ID NO: 2) 25 中从 1 或约 23 至约 216 位的氨基酸残基。

25. 分离的 FGF-19 多肽，其与 1997 年 11 月 21 日在 ATCC 以 ATCC 保藏号为 209480 (DNA49435-1219)保藏的载体中的 cDNA 插入子所编码的多肽具有至少约 80% 的序列同一性。

26. 权利要求 25 所述分离的 FGF-19 多肽，其是由 1997 年 11 月 21 日 30 在 ATCC 以 ATCC 保藏号为 209480 (DNA49435-1219)保藏的载体中的 cDNA 插入子编码的。

27. 分离的 FGF-19 多肽, 当与图 2 (SEQ ID NO: 2) 中 1 或约 23 至约 216 位的氨基酸序列比较时, 其具有至少 80% 的阳性评分。
28. 分离的 FGF-19 多肽, 其含有足以提供抗-FGF-19 抗体结合位点的图 2 (SEQ ID NO: 2) 中 1 或约 23 至约 216 位氨基酸残基的序列或其片段。
- 5 29. 分离的多肽, 其由如下制备(i)在严格条件下, 使测试 DNA 分子与 (a)编码含有图 2 (SEQ ID NO: 2) 中 1 或约 23 至约 216 位氨基酸残基的序列的 FGF-19 多肽的 DNA 分子, 或(b) (a)所述 DNA 分子的互补链进行杂交, (ii) 在适宜表达所述多肽的条件下培养含有所述测试 DNA 分子的宿主细胞, 和(iii)从细胞培养物中回收所述多肽。
- 10 30. 权利要求 29 所述分离的多肽, 其中所述测试 DNA 与(a)或(b)具有至少约 80% 的序列同一性。
31. 含有 FGF-19 多肽与异源氨基酸序列融合的嵌合分子。
32. 权利要求 31 所述嵌合分子, 其中所述异源氨基酸序列是表位标记序列。
- 15 33. 权利要求 31 所述嵌合分子, 其中所述异源氨基酸序列是免疫球蛋白的 Fc 区。
34. 与 FGF-19 多肽特异性结合的抗体。
35. 权利要求 34 所述抗体, 其中所述抗体是单克隆抗体。
36. 权利要求 34 所述抗体, 其中所述抗体是人源化抗体。
- 20 37. 权利要求 34 所述抗体, 其中所述抗体是抗体片段。
38. FGF-19 多肽的激动剂。
39. FGF-19 多肽的拮抗剂。
40. 物质组合物, 含有与可药用载体以混合物形式存在的(a) FGF-19 多肽, (b) FGF-19 多肽的激动剂, (c) FGF-19 多肽的拮抗剂, 或(d) 抗-FGF-19
- 25 抗体。
41. 能够与 FGF19 结合的生物活性剂的筛选方法, 包括:
- a) 向 FGF-19 样品中加入候选生物活性剂; 和
- b) 检测所述候选试剂与所述 FGF-19 的结合, 其中发生结合则表明是能够与 FGF-19 结合的生物活性剂。
- 30 42. 能够调节 FGF-19 活性的生物活性剂的筛选方法, 所述方法包括下述步骤:

- a) 向 FGF-19 样品中加入候选生物活性剂; 和
- (b) 检测 FGF-19 生物活性的改变, 其中发生改变则表明是能够调节 FGF-19 活性的生物活性剂。
43. 权利要求 42 所述方法, 其中所述生物活性是减少脂肪细胞对葡萄糖的摄取。
44. 权利要求 42 所述方法, 其中所述生物活性是增加 leptin 自脂肪细胞释放。
45. 鉴定 FGF-19 受体的方法, 所述方法包括: 使 FGF-19 与含有细胞膜物质的组分结合, 如果所述 FGF-19 与所述细胞膜物质上的受体形成复合物, 则确定所述受体是 FGF-19 受体。
46. 权利要求 45 所述方法, 其中 FGF-19 与所述受体结合, 并且所述方法进一步包括所述 FGF-19 与受体交联的步骤。
47. 权利要求 45 所述方法, 其中所述组分是细胞。
48. 权利要求 45 所述方法, 其中所述组分是细胞膜提取物制品。
49. 诱导 leptin 自脂肪细胞释放的方法, 所述方法包括向所述细胞施用能诱导 leptin 释放的有效量的 FGF-19。
50. 权利要求 49 所述方法, 其中所述 FGF-19 以蛋白形式施用。
51. 权利要求 49 所述方法, 其中所述 FGF-19 以核酸形式施用。
52. 诱导脂肪细胞减少葡萄糖摄取的方法, 所述方法包括向所述细胞施用能诱导葡萄糖摄取减少的有效量的 FGF-19。
53. 权利要求 52 所述方法, 其中所述 FGF-19 以蛋白形式施用。
54. 权利要求 52 所述方法, 其中所述 FGF-19 以核酸形式施用。
55. 治个体肥胖的方法, 所述方法包括向所述个体施用治疗所述肥胖有效量的含有 FGF-19 的组合物。
56. 权利要求 55 所述方法, 其中所述肥胖治疗进一步导致对与肥胖相关病症的治疗。
57. 权利要求 55 所述方法, 其中所述 FGF-19 以蛋白形式施用。
58. 权利要求 55 所述方法, 其中所述 FGF-19 以核酸形式施用。
59. 权利要求 55 所述方法, 其中所述组合物进一步包括可药用载体。
60. 权利要求 55 所述方法, 其中所述 FGF-19 与图 2 (SEQ ID NO: 2) 所示氨基酸序列具有至少约 85% 的氨基酸序列同一性。

61. 减轻个体总体重的方法，所述方法包括向所述个体施用有效量 FGF-19。

62. 权利要求 61 所述方法，其中所述 FGF-19 以蛋白形式施用。

63. 权利要求 61 所述方法，其中所述 FGF-19 以核酸形式施用。

5 64. 权利要求 61 所述方法，其中所述 FGF-19 与可药用载体一起施用。

65. 权利要求 61 所述方法，其中所述总体重减轻包括所述个体脂肪减少。

66. 权利要求 61 所述方法，其中所述 FGF-19 与图 2 (SEQ ID NO: 2) 所示氨基酸序列具有至少约 85% 的氨基酸序列同一性。

10 67. 降低个体中至少一种甘油三酯和游离脂肪酸水平的方法，所述方法包括向所述个体施用有效量的 FGF-19。

68. 权利要求 67 所述方法，其中所述 FGF-19 以蛋白形式施用。

69. 权利要求 67 所述方法，其中所述 FGF-19 以核酸形式施用。

70. 权利要求 67 所述方法，其中所述 FGF-19 与可药用载体一起施用。

15 71. 权利要求 67 所述方法，其中所述 FGF-19 与图 2 (SEQ ID NO: 2) 所示氨基酸序列具有至少约 85% 的氨基酸序列同一性。

72. 增加个体代谢率的方法，所述方法包括向所述个体施用有效量的 FGF-19。

73. 权利要求 72 所述方法，其中所述 FGF-19 以蛋白形式施用。

20 74. 权利要求 72 所述方法，其中所述 FGF-19 以核酸形式施用。

75. 权利要求 72 所述方法，其中所述 FGF-19 与可药用载体一起施用。

76. 权利要求 72 所述方法，其中所述 FGF-19 与图 2 (SEQ ID NO: 2) 所示氨基酸序列具有至少约 85% 的氨基酸序列同一性。

77. 含有包括编码 FGF-19 的转基因的基因组的啮齿动物。

成纤维细胞生长因子-19(FGF-19)的核酸
和多肽以及用于治疗肥胖的方法

5

本申请为申请日为 2000 年 3 月 9 日(PCT/US00/06471), 申请号为 00815274.8 的分案申请。

发明领域

10 本发明主要涉及鉴定和分离新的 DNA, 涉及重组产生本文中称为成纤维细胞生长因子-19(FGF-19)多肽的新多肽, 涉及使用此多肽治疗肥胖和用于制备具有治疗和药理学特性(包括与治疗肥胖有关的那些)的药物活性物质的方法、组合物和测定法。

发明背景

15 肥胖是现代社会极为普遍的慢性疾病, 不仅与社会不美观感有关, 而且与寿命缩短和大量医学问题有关, 所述大量医学问题包括不良心理发展, 生殖障碍如多囊卵巢疾病, 皮肤病学病症如感染、静脉曲张、黑棘皮症和湿疹, 运动耐力差, 糖尿病, 胰岛素抗性, 高血压, 高血胆固醇血症, 胆石病, 骨关节炎, 矫型损伤, 血栓性疾病, 癌症和冠心病。Rissanen 等, British
20 Medical Journal, 301: 835-837 (1990)。

现有的减肥疗法包括标准日常饮食和锻炼, 极低热量饮食, 行为疗法, 涉及食欲抑制剂、产热药物、食物吸收抑制剂的药物治疗, 医疗器械如颌栓结
25 术(jaw wiring), 束腰和气囊以及手术。Jung 和 Chong, Clinical Endocrinology, 35:11-20 (1991);Bray, Am.J.Clin.Nutr, 55:538S-544S(1992)。已有报道: 节制蛋白的改良型禁食方法在青少年减轻体重方面有效。Lee 等, Clin.Pediatr, 31:234-
236 (April 1992)。作为减肥疗法的热量限制引起机体贮存蛋白的分解代谢并产生负氮平衡。因此, 补充蛋白的饮食作为减少热量限制期间氮损失的措施已广受欢迎。由于这样的饮食仅产生适度的氮减少, 因此需要保持苗条身体和蛋白贮备的更有效方法。此外, 如果这类方案还导致机体脂肪损耗加速, 则能改进减肥治疗。此类治疗的各种方法包括 Weintraub 和 Bray 在 Med.
30 Clinics N. Amer., 73: 237 (1989); Bray, 在 Nutrition Reviews, 49: 33 (1991)

中讨论的那些。

考虑到肥胖在社会中的极度普遍性和前面提到的与肥胖相关的严重并发症,任何潜在地可用于减轻肥胖个体体重的治疗药物必将对他们的健康产生深刻的有益影响。本领域需要使肥胖个体全身体重减轻至其理想体重并帮助肥胖个体保持其降低后的体重水平而没有严重副作用的药物。

因此,期望提供可用于肥胖个体恢复其正常、理想体重的治疗方案。

期望进一步提供能长时期保持低体重的减肥疗法。

也期望预防肥胖,和一旦开始治疗则阻止继发于肥胖的疾病的进展或者预防其发作,所述继发病如动脉硬化和多囊卵巢疾病。

10 本发明也提供这类治疗方法及相关组合物。本发明还提供新的蛋白和核酸,以及筛选所述新蛋白和核酸的调节子的方法。本发明提供的其它方法、治疗和组合物对于本领域熟练技术人员而言将变得显而易见。

发明概述

15

已经鉴定了编码新多肽的 cDNA 克隆(本文称为 DNA49435-1219),其与成纤维细胞生长因子家族的成员 - 本申请中的"成纤维细胞生长因子-19"(FGF-19)具有某些序列相似性。

20 本发明的一个方面,提供分离的包含编码 FGF-19 多肽的核苷酸序列的核酸分子。

一方面,分离的核酸分子含有与(a)编码含图 2(SEQ ID NO: 2)所示从约 1 或约 23 至约 216 的氨基酸序列(包括两端残基在内)的 PEACH 多肽的 DNA 分子,或者(b) (a)所述 DNA 分子的互补链,具有至少约 80%、至少约 81%、至少约 82%、至少约 83%、至少约 84%、至少约 85%、至少约 86%、至少约 87%、至少约 88%、至少约 89%、至少约 90%、至少约 91%、至少约 92%、至少约 93%、至少约 94%、至少约 95%、至少约 96%、至少约 97%、至少约 98% 或至少约 99% 核酸序列同源的核苷酸序列。

另一方面,分离的核酸分子含有(a)编码含图 2(SEQ ID NO: 2)所示从约 1 或约 23 至约 216 的氨基酸序列(包括两端残基在内)的 FGF-19 多肽的核苷酸序列,或者(b) (a)所述核苷酸序列的互补链。

30 再一方面,分离的核酸分子含有与(a) 图 1 (SEQ ID NO: 1)所示由约 464

或约 530 至约 1111 的核苷酸序列(包括两端残基在内)的 DNA 分子, 或者(b) (a)所述 DNA 分子的互补链, 具有至少约 80%、至少约 81%、至少约 82%、至少约 83%、至少约 84%、至少约 85%、至少约 86%、至少约 87%、至少约 88%、至少约 89%、至少约 90%、至少约 91%、至少约 92%、至少约 93%、至少约 94%、至少约 95%、至少约 96%、至少约 97%、至少约 98% 或至少约 99% 核酸序列同源的核苷酸序列。

还一方面, 分离的核酸分子含有(a)图 1 (SEQ ID NO: 1)中从约 464 或约 530 至约 1111 位的核苷酸序列, 或者(b) (a)所述核苷酸序列的互补链。

此外, 本发明涉及分离的核酸分子, 其含有与(a) 编码由 1997 年 11 月 21 日在 ATCC 保藏的 ATCC 保藏号为 209480 (DNA49435-1219)的人蛋白 cDNA 编码的相同成熟多肽的 DNA 分子, 或者(b) (a)所述 DNA 分子的互补链, 具有至少约 80%、至少约 81%、至少约 82%、至少约 83%、至少约 84%、至少约 85%、至少约 86%、至少约 87%、至少约 88%、至少约 89%、至少约 90%、至少约 91%、至少约 92%、至少约 93%、至少约 94%、至少约 95%、至少约 96%、至少约 97%、至少约 98% 或至少约 99% 核酸序列同源的核苷酸序列。在一优选的实施方案中, 分离的核酸分子含有(a)编码与 1997 年 11 月 21 日在 ATCC 保藏的 ATCC 保藏号为 209480 (DNA49435-1219)的人蛋白 cDNA 编码的相同成熟多肽的核苷酸序列, 或者 (b) (a)所述核苷酸序列的互补链。

另一方面, 发明涉及分离的核酸分子, 其含有与(a) 1997 年 11 月 21 日在 ATCC 保藏的 ATCC 保藏号为 209480 (DNA49435-1219)的人蛋白 cDNA 的全长多肽编码序列, 或者(b) (a)所述核苷酸序列的互补链, 具有至少约 80%、至少约 81%、至少约 82%、至少约 83%、至少约 84%、至少约 85%、至少约 86%、至少约 87%、至少约 88%、至少约 89%、至少约 90%、至少约 91%、至少约 92%、至少约 93%、至少约 94%、至少约 95%、至少约 96%、至少约 97%、至少约 98% 或至少约 99% 核酸序列同源的核苷酸序列。在一优选的实施方案中, 分离的核酸分子含有(a) 1997 年 11 月 21 日在 ATCC 保藏的 ATCC 保藏号为 209480 (DNA49435-1219)的人蛋白 cDNA 的全长多肽编码序列, 或者(b) (a)所述核苷酸序列的互补链。

还一方面, 发明涉及分离的编码下面定义的活性 FGF-19 多肽的核酸分子, 其包括能与编码图 2 (SEQ ID NO: 2)中 1 或约 23 ~ 约 216 位氨基酸(包括

两个端点氨基酸残基在内)的核酸序列的互补链杂交的核苷酸序列。优选地,杂交在严格杂交条件和洗涤条件下发生。

另一方面,发明涉及分离的编码下面定义的活性 FGF-19 多肽的核酸分子,其包括与图 1(SEQ ID NO: 1)所示核苷酸约 464 或约 530 和约 1111 之间(包
5 括两个端点残基在内)核酸序列的互补链杂交的核苷酸序列。优选地,杂交在严格杂交条件和洗涤条件下发生。

再一方面,发明涉及分离的具有至少约 22 个核苷酸的核酸分子,其是在严格条件下使测试 DNA 分子与(a)编码具有图 2(SEQ ID NO: 2)中 1 或约 23 至约 216(包括两端残基在内)氨基酸残基的序列的 FGF-19 多肽的 DNA 分子,或者(b) (a)所述 DNA 分子的互补链进行杂交而产生的,而且,如果该测试 DNA 分子与(a)或(b)具有至少约 80%、至少约 81%、至少约 82%、至少
10 约 83%、至少约 84%、至少约 85%、至少约 86%、至少约 87%、至少约 88%、至少约 89%、至少约 90%、至少约 91%、至少约 92%、至少约 93%、至少约 94%、至少约 95%、至少约 96%、至少约 97%、至少约 98% 至少
15 约 99%的核酸序列同一性,则分离该测试 DNA 分子。

另一方面,发明涉及分离的核酸分子,其含有(a)编码与图 2 (SEQ ID NO: 2)约 1 或约 23 至 216 残基所示氨基酸序列(包括两端氨基酸残基在内)相比,具有至少约 80%、至少约 81%、至少约 82%、至少约 83%、至少约 84%、
20 至少约 85%、至少约 86%、至少约 87%、至少约 88%、至少约 89%、至少约 90%、至少约 91%、至少约 92%、至少约 93%、至少约 94%、至少约 95%、至少约 96%、至少约 97%、至少约 98%、至少约 99%评分阳性的多肽的核苷酸序列,或者(b) (a)所述核苷酸序列的互补链。

一个特定方面,本发明提供分离的核酸分子,包括编码没有 N-末端信号序列和/或起始蛋氨酸的 FGF-19 多肽的 DNA,或者此编码核酸分子的互
25 补链。信号肽已被试验性地确定为在图 2 (SEQ ID NO: 2)所示序列中从约氨基酸 1 位向约氨基酸 22 位延伸(包括两端残基在内)。然而,应当指出,信号肽的 C-末端边界可能不同,但是在信号肽 C-末端边界两端最可能不超过约 5 个氨基酸,如本文最初的鉴定,其中信号肽 C-末端边界可能是依据本领域用于鉴定这类氨基酸序列元件时通常使用的标准来确定的(例如, Nielsen 等,
30 Prot. Eng. 10: 1-6 (1997)和 von Heinje 等, Nucl. Acids. Res.14: 4683-4690 (1986))。而且,也应认识到,在有些情况下,从分泌的多肽上裂解的信号序

列不完全相同，从而导致一种以上的分泌类型。这些多肽和编码它们的多核苷酸是本发明考虑的内容。本申请图 2 (SEQ ID NO: 2) 所示 FGF-19 多肽的信号肽，从图 2 (SEQ ID NO: 2) 的氨基酸 1 向 X 延伸，其中 X 是图 2 (SEQ ID NO: 2) 中从 17 到 27 的任何氨基酸。因此，本发明包括 FGF-19 多肽的成熟形式，
5 其中包括那些含有图 2 (SEQ ID NO: 2) 中 X 到 216 氨基酸的多肽以及下面将要描述的它们的变体，其中 X 是图 2 (SEQ ID NO: 2) 中从 17 到 27 的任何氨基酸。也包括分离的编码这些多肽的核酸分子。

另一实施方案涉及 FGF-19 多肽序列的片段，包括用作例如杂交探针或 FGF-19 多肽中可任选地编码含抗-FGF-19 抗体结合位点之多肽的编码片段
10 的编码序列。此类核酸片段长度通常至少约 20 个核苷酸，或者至少约 30 个核苷酸，或至少约 40 个核苷酸，或至少约 50 个核苷酸或至少约 60 个核苷酸，或者至少约 70 个核苷酸，或至少约 80 个核苷酸，或至少约 90 个核苷酸，或至少约 100 个核苷酸，或者至少约 110 个核苷酸，或至少约 120 个核苷酸，或至少约 130 个核苷酸，或至少约 140 个核苷酸，或者至少约 150 个
15 核苷酸，或至少约 160 个核苷酸，或至少约 170 个核苷酸，或至少约 180 个核苷酸，或者至少约 190 个核苷酸，或至少约 200 个核苷酸，或至少约 250 个核苷酸，或至少约 300 个核苷酸，或至少约 350 个核苷酸，或至少约 400 个核苷酸，或至少约 450 个核苷酸，或至少约 500 个核苷酸，或至少约 600 个核苷酸，或至少约 700 个核苷酸，或至少约 800 个核苷酸，或至少约 900
20 个核苷酸，或至少约 1000 个核苷酸，其中本文术语“约”是指核苷酸序列有关长度加或减 10%。在一优选的实施方案中，核苷酸序列片段衍生自图 1 (SEQ ID NO: 1) 所示核苷酸序列的任何编码区。应当指出，可用常规方法测定 FGF-19 多肽-编码核苷酸序列的新片段，其通过使用众多公知序列对比程序中任何一个将 FGF-19 多肽的编码核苷酸序列与其它已知核苷酸序列对
25 比，测定出其中哪个或哪几个 FGF-19 多肽-编码核苷酸序列片段是新的。本文考虑到了所有这类 FGF-19 多肽-编码核苷酸序列，并且可以用不繁琐的实验予以测定。也考虑到由这些核苷酸分子片段编码的 FGF-19 多肽片段，优选是含有抗-FGF-19 抗体结合位点的那些 FGF-19 多肽片段。

在另一实施方案中，发明提供含有编码 FGF-19 或其变体的核苷酸序列
0 的载体。该载体可含有上述定义的任何分离的核酸分子。

也提供含有此类载体的宿主细胞。例如，宿主细胞可以是 CHO 细胞、

大肠杆菌、感染了杆状病毒的昆虫细胞或酵母。进一步提供生产 FGF-19 多肽的方法，包括在适宜表达 FGF-19 的条件下培养宿主细胞并从细胞培养物中回收 FGF-19。

在另一实施方案中，发明提供分离的由上面所定义的任一分离的核酸序列所编码的 FGF-19 多肽。

一特定方面，本发明提供分离的天然序列 FGF-19 多肽，在某些实施方案中，它包括图 2 (SEQ ID NO: 2) 所示由约 1 或约 23 到约 216 残基的氨基酸序列。

另一方案，本发明涉及分离的 FGF-19 多肽，包括与图 2 (SEQ ID NO: 2) 中约 1 或约 23 到约 216 残基所示氨基酸序列(包括两端残基在内)至少约 80%、至少约 81%、至少约 82%、至少约 83%、至少约 84%、至少约 85%、至少约 86%、至少约 87%、至少约 88%、至少约 89%、至少约 90%、至少约 91%、至少约 92%、至少约 93%、至少约 94%、至少约 95%、至少约 96%、至少约 97%、至少约 98% 或至少约 99% 同一性的氨基酸序列。

再一方面，本发明涉及分离的 FGF-19 多肽，其含有与 1997 年 11 月 21 日在 ATCC 保藏的 ATCC 保藏号为 209480 (DNA49435-1219) 的人蛋白 cDNA 编码的氨基酸序列具有至少约 80%、至少约 81%、至少约 82%、至少约 83%、至少约 84%、至少约 85%、至少约 86%、至少约 87%、至少约 88%、至少约 89%、至少约 90%、至少约 91%、至少约 92%、至少约 93%、至少约 94%、至少约 95%、至少约 96%、至少约 97%、至少约 98% 或至少约 99% 同一性的氨基酸序列。在一优选的实施方案中，分离的 FGF-19 多肽含有由 1997 年 11 月 21 日在 ATCC 保藏的 ATCC 保藏号为 209480 (DNA49435-1219) 的人蛋白 cDNA 编码的氨基酸序列。

另一方面，发明涉及分离的 FGF-19 多肽，其含有与图 2 (SEQ ID NO: 2) 中约 1 或约 23 到约 216 位残基所示氨基酸序列(包括两端残基在内)至少约 80%、至少约 81%、至少约 82%、至少约 83%、至少约 84%、至少约 85%、至少约 86%、至少约 87%、至少约 88%、至少约 89%、至少约 90%、至少约 91%、至少约 92%、至少约 93%、至少约 94%、至少约 95%、至少约 96%、至少约 97%、至少约 98% 或至少约 99% 评分阳性的氨基酸序列。

一特定方面，本发明提供分离的没有 N-末端序列和/或起始蛋氨酸的 FGF-19 多肽，该多肽由编码前述氨基酸序列的核苷酸序列编码。本文也公

开了制备该多肽的方法，方法包括：在适宜表达 FGF-19 多肽的条件下，培养含有相应编码核酸分子的载体的宿主细胞，并从培养物中回收 FGF-19 多肽。

还一方面，本发明涉及分离的 FGF-19 多肽，包括图 2 (SEQ ID NO: 2) 5 中约 1 或约 23 至约 216 残基所示氨基酸序列(包括两端残基在内)或其生物活性片段，或足以提供与抗-FGF-19 抗体结合位点的片段，其中具有生物活性或提供抗-FGF-19 抗体结合位点的 FGF-19 多肽片段的鉴定，可以用本领域公知的常规技术来完成。优选地，FGF-19 片段在性质上保留天然 FGF-19 多肽的生物活性，包括治疗肥胖的能力。

10 再一方面，本发明提供如下制备的多肽：(i)在严格条件下，使测试 DNA 分子与(a)编码具有图 2 (SEQ ID NO: 2)中约 1 或约 23 至约 216 氨基酸残基的序列(包括两端残基在内)的 FGF-19 多肽的 DNA 分子，或者(b) (a)所述 DNA 分子的互补链进行杂交，且如果该测试 DNA 分子与(a)或(b)具有至少约 80%、至少约 81%、至少约 82%、至少约 83%、至少约 84%、至少约 85%、 15 至少约 86%、至少约 87%、至少约 88%、至少约 89%、至少约 90%、至少约 91%、至少约 92%、至少约 93%、至少约 94%、至少约 95%、至少约 96%、至少约 97%、至少约 98%或至少约 99%的核酸序列同一性，则(ii)在适宜表达该多肽的条件下培养含有该测试 DNA 分子的宿主细胞，和(iii)从细胞培养物中回收该多肽。

20 在另一实施方案中，本发明提供含有与异源多肽或氨基酸序列融合的 FGF-19 多肽的嵌合分子，其中 FGF-19 多肽可含有前述任何 FGF-19 多肽或其变体或片段。此嵌合分子的一个实例是含有与免疫球蛋白表位标记序列或 Fc 区融合的 FGF-19 多肽的嵌合分子。

在另一实施方案中，发明提供下述定义的与前述 FGF-19 多肽特异结合 25 的抗体。任选地，抗体是单克隆抗体、抗体片段或单链抗体。

在另一实施方案中，本发明涉及下述定义天然 FGF-19 多肽的激动剂和拮抗剂。在一具体实施方案中，激动剂或拮抗剂是抗-FGF-19 抗体或小分子。

30 在另一实施方案中，发明涉及鉴定 FGF-19 多肽的激动剂或拮抗剂的方法，方法包括使 FGF-19 多肽与候选分子接触，并检测由所述 FGF-19 多肽介导的生物活性。优选地，FGF-19 多肽是天然 FGF-19 多肽。

在另外一个实施方案中，本发明涉及含有本文所述 FGF-19 多肽、FGF-19 多肽的激动剂或拮抗剂、或抗-FGF-19 抗体、以及载体的组合物。任选地，载体是可药用载体。

本发明另一实施方案涉及本文所述 FGF-19 多肽或其激动剂或拮抗剂或抗-FGF-19 抗体在制备用于治疗对 FGF-19 多肽、其激动剂或拮抗剂、或抗-FGF-19 抗体应答的病症的药物中的应用。

在一个实施方案中，提供了筛选能够结合 FGF-19 的生物活性剂的方法。一方面，方法包括向 FGF-19 样品中加入候选生物活性剂，并测定所述候选生物活性剂与所述 FGF-19 的结合，其中如果发生结合，则表明是能够与 FGF-19 结合的生物活性剂。

本文另外提供筛选能够调节 FGF-19 活性的生物活性剂的方法。在一个实施方案中，提供一个方法，该方法包括下述步骤：向 FGF-19 样品中加入候选生物活性剂，并检测 FGF-19 生物活性的变化，如果有变化，则表明是能够调节 FGF-19 活性的生物活性剂。在一实施方案中，FGF-19 活性是减少细胞中葡萄糖的摄取。在另一实施方案中，FGF-19 活性是增加 leptin 自细胞释放。在一优选的实施方案中，FGF-19 活性是降低葡萄糖摄取并增加 leptin 自细胞释放。优选这类细胞是脂肪细胞。在另一实施方案中，FGF-19 活性是增加脂类和碳水化合物的氧化。优选这类细胞是肝细胞或肌肉细胞。

在另一实施方案中，发明提供鉴定 FGF-19 受体的方法。在一优选实施方案中，该方法包括使 FGF-19 与含有细胞膜物质的组分结合，如果所述 FGF-19 与所述细胞膜物质上的受体形成复合物，则鉴定所述受体就是 FGF-19 受体。在一实施方案中，方法包括使所述 FGF-19 与受体交联的步骤。细胞膜取自完整细胞或者细胞膜提取物制品(extract preparation)。

本发明的另一方面，提供诱导 leptin 自细胞(优选脂肪细胞)释放的方法，在一实施方案中，该方法包括向细胞施用能有效诱导 leptin 释放的量的 FGF-19。

在本文提供的方法中，FGF-19 可以以表达 FGF-19 的核酸形式或以蛋白形式来施用。如下面将详述的那样，可通过输注或持续释放制剂而施用 FGF-19。优选地，FGF-19 与可药用载体一起施用于个体。

也提供诱导细胞(优选脂肪细胞)的葡萄糖摄取减少的方法。在一实施方案中，该方法包括向细胞施用能有效诱导葡萄糖摄取减少的量的 FGF-19。

本发明另一方面提供治疗个体肥胖的方法。在一实施方案中，该方法包括向个体施用含有有效治疗肥胖的量的 FGF-19 的组合物。也可以以这种方式治疗与肥胖有关的病症如心血管疾病。

本文也提供降低个体总体重的方法，包括向所述个体施用有效量的 FGF-19。在一优选的实施方案中，减少个体肥胖(脂肪)。

此外，本文提供降低个体的至少一种甘油三酯和游离脂肪酸水平的方法，包括向所述个体施用有效量的 FGF-19。本文也提供增加个体代谢率的方法，包括向所述个体施用有效量 FGF-19。

本文还提供在各种条件和状态下检测 FGF-19 及调节物的效应的动物模型。在一实施方案中，提供含有编码 FGF-19 的转基因的基因组的动物，优选是啮齿类动物。

附图简述

图 1 显示含有编码天然序列 FGF-19 的核苷酸序列(核苷酸 464-1111)cDNA 的核苷酸序列(SEQ ID NO: 1)，其中该核苷酸序列(SEQ ID NO: 1)是本文称为"DNA49435-1219"的克隆。也以粗字体和下划线标出各自的起始和终止密码子的位置。

图 2 显示衍生自 SEQ ID NO: 1 中编码序列的天然序列 FGF-19 多肽的氨基酸序列(SEQ ID NO: 2)。也显示各种其它重要多肽结构域的近似位置。

图 3A 和 3B 显示证明 MLC-FGF-19 转基因小鼠体重较其未转基因的同窝鼠体重轻(图 3A)及具有较低的循环 leptin 水平的直方图(图 3B)。图 3A 显示 6 周龄 FGF-19 转基因鼠(实心棒)和未转基因(野生型)同窝鼠(点彩棒)在自由进食期间(最左侧)、禁食 6 小时和 24 小时、和结束 24 小时禁食后 24 小时(最右侧)的重量。图 3B 显示图 3A 中所示相同组小鼠在 leptin 测定中的血清(垂直棒)。

图 4A-4D 是表明 FGF-19 转基因鼠增加食物摄取和尿量但仍具有正常血细胞比容的直方图。监测一组小鼠自由进食期间和结束 24 小时禁食之后 24 小时的食物摄取(图 4A)、水摄取(图 4B)、排尿量(图 4C)和血细胞比容(图 4D)情况，其中每组中 FGF-19 转基因鼠的结果用实心黑棒表示，野生型鼠的结果用点彩棒表示。

图 5 是表示 FGF-19 转基因鼠增加氧消耗率的直方图。显示了 FGF-19 转基因鼠(实心黑棒)和野生型(点彩棒)在日间和夜间、禁食 24 小时和结束 24 小时禁食后 24 小时的氧消耗情况。

图 6A 和 6B 是表明与野生型小鼠(点彩棒)相比, FGF-19 转基因鼠(实心黑棒)降低甘油三酯(图 6A)和游离脂肪酸(图 6B)的直方图。

图 7A 和 7B 是表明与输注不含 FGF-19 的安慰剂(点彩棒)相比, 给未转基因鼠输注 FGF-19(实心黑棒)导致食物摄取增加(图 7A)和氧消耗增加(图 7B)的直方图, 其中"n"代表黑夜, 而"d"代表白天。

图 8A 和 8B 是表明 FGF-19 增加 leptin 自脂肪细胞释放(图 8A)和减少脂肪细胞摄取葡萄糖(图 8B)的直方图。

图 9 是显示 FGF-19 转基因鼠(阴影棒)或野生型(实心黑棒)分别高脂饮食(HFD)一段时间后脂肪垫重量的直方图, 其中沿水平方向自左开始, 分别是 6 周时附睾(HFD Ep)的和腹膜后与肾周(HFD RP/PR)的结果, 接着是 10 周时附睾的结果, 然后是腹膜后和肾周的结果。

图 10 是显示 FGF-19 转基因鼠(阴影棒)或野生型鼠(实心黑棒)随着时间的葡萄糖耐量情况的直方图(均高脂饮食 10 周)。

优选实施方案详述

I. 定义

本文所用术语"FGF-19 多肽"、"FGF-19 蛋白"和"FGF-19", 包括天然序列 FGF-19 和 FGF-19 多肽变体(将进一步定义)。FGF-19 多肽可以从各种来源(如人组织或其它来源)分离得到的, 也可以用重组和/或合成方法制得。

"天然序列 FGF-19"包括具有与得自天然 FGF-19 相同的氨基酸序列的多肽。此天然序列 FGF-19 可以从自然界分离得到的, 或者用重组和/或合成方法制得的。术语"天然序列 FGF-19"特别包括 FGF-19 的天然截短型或分泌型(例如, 细胞外结构域序列), 天然变体型(例如, 可替换式剪切型)和天然等位变体。在本发明一实施方案中, 天然序列 FGF-19 是包括图 2 (SEQ ID NO: 2)中 1 至 216 氨基酸的成熟或全长天然序列 FGF-19。并且, 尽管图 2 (SEQ ID NO: 2)公开的 FGF-19 多肽显示由本文定义为位置 1 的蛋氨酸残基开始, 但是图 2 (SEQ ID NO: 2)中氨基酸 1 位上游或下游的另一蛋氨酸残基作为

FGF-19 多肽的起始氨基酸残基是可以想象到的和可能的。

"FGF-19 变体多肽"指如下定义的与下列氨基酸序列具有至少约 80% 氨基酸序列同一性的活性 FGF-19 多肽: (a)图 2 (SEQ ID NO: 2)所示 FGF-19 多肽的 1 或约 23 至 216 残基的氨基酸序列, (b)图 2 (SEQ ID NO: 2)所示 FGF-19 多肽的 X 至 216 残基的氨基酸序列, 其中 X 是图 2 (SEQ ID NO: 2)中 17 至 27 的任何氨基酸残基, 或者(c)衍生自图 2 (SEQ ID NO: 2)所示氨基酸序列的另一特定片段。此类 FGF-19 变体多肽包括, 例如, 在图 2 (SEQ ID NO: 2)所示序列的 N-和/或 C-末端以及在一个或多个内部结构域内增加或删减一个或多个氨基酸残基的 FGF-19 多肽。通常, FGF-19 变体多肽与(a)图 2 (SEQ ID NO: 2)所示 FGF-19 多肽的 1 或约 23 至 216 残基的氨基酸序列, (b)图 2 (SEQ ID NO: 2)所示 FGF-19 多肽的 X 至 216 残基的氨基酸序列, 其中 X 是图 2 (SEQ ID NO: 2)中 17 至 27 的任何氨基酸残基, 或者(c)衍生自图 2 (SEQ ID NO: 2)所示氨基酸序列的另一特定片段, 具有至少约 80%、至少约 81%、至少约 82%、至少约 83%、至少约 84%、至少约 85%、至少约 86%、至少约 87%、至少约 88%、至少约 89%、至少约 90%、至少约 91%、至少约 92%、至少约 93%、至少约 94%、至少约 95%、至少约 96%、至少约 97%、至少约 98% 或至少约 99% 的氨基酸序列同一性。FGF-19 变体多肽不包括天然 FGF-19 多肽序列。一般地, FGF-19 变体多肽的长度是至少约 10 个氨基酸, 或至少约 20 个氨基酸, 或者至少约 30 个氨基酸, 或至少约 40 个氨基酸, 或至少约 50 个氨基酸, 或至少约 60 个氨基酸, 或者至少约 70 个氨基酸, 或至少约 80 个氨基酸, 或至少约 90 个氨基酸, 或至少约 100 个氨基酸, 或者至少约 150 个氨基酸, 或至少约 200 个氨基酸, 或至少约 300 或更多个氨基酸。

本文所说有关 FGF-19 多肽序列的"氨基酸序列同一性百分比(%)"是指候选序列的氨基酸残基, 在进行序列对比, 并在必要时导入空隙以获取最大百分比的序列同一性, 而不将任何保守取代视为序列同一性时与 FGF-19 序列的氨基酸残基相同的百分数。可使用本领域各种方法进行序列对比以便测定氨基酸序列同一性百分比, 例如, 使用公众可得到的计算机软件如 BLAST、BLAST-2、ALIGN、ALIGN-2 或 Megalign (DNASTAR)软件。本领域技术人员可以决定测量对比的适宜参数, 包括针对所比较的序列全长获得最大对比所需的任何算法。然而, 为此目的, 氨基酸序列同一性 % 值是使用

下述序列对比计算机程序 ALIGN-2 而获得的, 其中 ALIGN-2 程序的全部源代码见表 1。ALIGN-2 序列对比计算机程序的作者是 Genentech, Inc., 表 1 所示源代码和用户资料一起已经提交地处 Washington D. C., 20559 的美国版权局, 其美国版权注册登记号为 TXU510087。公众通过 Genentech, Inc., South
5 San Francisco, California 可以得到 ALIGN-2 程序, 或者可从表 1 提供的源代码进行编制。ALIGN2 程序应当为在 UNIX 操作系统, 优选在数码 UNIX V4.0D 使用而进行编制。ALIGN-2 程序设定了所有序列对比参数并且不变。

为了本发明目的, 给定氨基酸序列 A 相对于给定氨基酸序列 B 的氨基酸序列同一性%(或者说: 给定氨基酸序列 A 具有或含有给定氨基酸序列 B 相同氨基酸序列的%)如下计算:
10

X/Y 比值乘以 100

其中 X 是用序列对比程序 ALIGN-2 比较 A 和 B 的氨基酸残基后计算出的相同氨基酸数, 其中 Y 是 B 的氨基酸残基总数。可以理解, 当氨基酸序列 A 与氨基酸序列 B 的长度不相等时, A 相对于 B 的氨基酸序列同一性%
15 将不等于 B 相对于 A 的氨基酸序列同一性%。作为计算氨基酸序列同一性的一个实例, 表 2 和 3 说明如何计算指定为"对照蛋白"的氨基酸序列与指定为"PRO"的氨基酸序列的序列同一性%。

除非另外特别声明, 否则本文所用所有氨基酸序列同一性%值是按照上述使用 ALIGN-2 序列对比计算机程序获得的。然而, 氨基酸序列同一性%
20 也可使用序列对比程序 NCBI-BLAST2 (Altschul 等, Nucleic acids Res. 25: 3389-3402 (1997))进行测定。NCBI-BLAST2 序列对比程序可从 <http://www.ncbi.nlm.nih.gov> 下载, 或从 National Institute of Health, Bethesda, MD 得到。NCBI-BLAST2 使用数种检索参数, 其中所有这些参数设定为默认值, 包括例如, 非掩蔽=yes, 链(strand)=全部(all), 预期的出现=10, 最低复杂长度=
!5 15/5, 多-程 e-值 = 0.01, 多-程常数= 25, 最终缺口化对比的陡降(dropoff)= 25, 和评分矩阵= BLOSUM62。

当使用 NCBI-BLAST2 进行氨基酸序列比较时, 给定氨基酸序列 A 针对给定氨基酸序列 B 的氨基酸序列同一性%(或者说, 给定氨基酸序列 A 具有或含有与给定氨基酸序列 B 相同的氨基酸序列的%)如下计算:

0

X/Y 比值乘以 100

其中 X 是用序列对比程序 NCBI-BLAST2 比较 A 和 B 氨基酸残基后计

算出的相同氨基酸数量，其中 Y 是 B 的氨基酸残基总数量。可以理解，当氨基酸序列 A 与氨基酸序列 B 的长度不相等时，A 针对 B 的氨基酸序列同一性 % 将不等于 B 针对 A 的氨基酸序列同一性 %。

"FGF-19 变体多核苷酸"或"FGF-19 变体核酸序列"是指编码下述定义的
5 活性 FGF-19 多肽的的核苷酸分子，它与(a)编码图 2 (SEQ ID NO: 2)所示 FGF-19 多肽的 1 或约 23 至 216 残基的氨基酸序列的核苷酸序列，(b) 编码图 2 (SEQ ID NO: 2)所示 FGF-19 多肽的 X 至 216 残基的氨基酸序列的核苷酸序列，其中 X 是图 2 (SEQ ID NO: 2)中 17 至 27 的任何氨基酸残基，或者
10 (c) 编码衍生自图 2(SEQ ID NO: 2)所示氨基酸序列的另一特定片段的核苷酸序列，具有至少约 80% 的氨基酸序列同一性。通常，FGF-19 变体多核苷酸与(a) 编码图 2 (SEQ ID NO: 2)所示 FGF-19 多肽的 1 或约 23 至 216 残基的氨基酸序列的核苷酸序列，(b) 编码图 2 (SEQ ID NO: 2)所示 FGF-19 多肽的 X 至 216 残基的氨基酸序列的核苷酸序列，其中 X 是图 2 (SEQ ID NO: 2)中 17 至 27 的任何氨基酸残基，或者(c) 编码衍生自图 2 (SEQ ID NO: 2)所示
15 氨基酸序列的另一特定片段的核苷酸序列，具有至少约 80%、至少约 81%、至少约 82%、至少约 83%、至少约 84%、至少约 85%、至少约 86%、至少约 87%、至少约 88%、至少约 89%、至少约 90%、至少约 91%、至少约 92%、至少约 93%、至少约 94%、至少约 95%、至少约 96%、至少约 97%、至少约 98% 或至少约 99% 的核酸序列同一性。FGF-19 多核苷酸变体
20 不包括天然 FGF-19 核苷酸序列。

一般地，FGF-19 变体多核苷酸的长度是至少约 30 个核苷酸，或至少约 60 个核苷酸，或至少约 90 个核苷酸，或至少约 120 个核苷酸，或者至少约 150 个核苷酸，或至少约 180 个核苷酸，或至少约 210 个核苷酸，或至少约 240 个核苷酸，或者至少约 270 个核苷酸，或至少约 300 个核苷酸，或至少
:5 约 450 个核苷酸，或至少 600 个核苷酸，或至少 900 或更多个核苷酸。

本文所述有关编码 FGF-19 多肽的核酸序列的"核酸序列同一性百分比 (%)"是指候选序列中与编码 FGF-19 多肽的核酸序列相比时，具有相同核苷酸的百分数，这是在对比序列，和如果必要则引入空隙，以达到最大序列同一性百分比的前提下获得的。为测定核苷酸序列同一性百分比而进行的序列
0 对比可使用本领域各种方法，例如，使用公众可得到的计算机软件如 BLAST、BLAST-2、ALIGN、ALIGN-2 或 Megalign(DNASTAR)软件。本领

域技术人员可以决定用于测量对比的适宜参数,包括针对所比较序列的全长获得最大对比所需的任何算法。然而,为此目的,核苷酸序列同一性%值是如下述使用序列对比计算机程序 ALIGN-2 而获得的,其中 ALIGN-2 程序的全部源代码见表 1。ALIGN-2 序列对比计算机程序的作者是 Genentech, Inc., 5 表 1 所示源代码和用户资料一起已经提交地处 Washington D. C., 20559 的美国版权局,其美国版权注册登记号为 TXU510087。公众通过 Genentech, Inc., South San Francisco, California 可以得到 ALIGN-2 程序,或者从表 1 提供的源代码进行编制。ALIGN2 程序应当为在 UNIX 操作系统,优选在数码 UNIX V4.0D 使用而进行编制。ALIGN-2 程序设定了所有序列对比参数并且不变。

10 为了本发明目的,给定核酸序列 C 相对于给定核酸序列 D 的核酸序列同一性%(或者说:给定核酸序列 C 具有或含有与给定核酸序列 D 相同的核酸序列的%)如下计算:

W/Z 比值乘以 100

其中 W 是用序列对比程序 ALIGN-2 比较 C 和 D 后计算出的相同核苷酸的数量,和其中 Z 是 D 的核苷酸总数量。可以理解,当核酸序列 C 与核酸序列 D 的长度不相等时,C 针对 D 的核酸序列同一性%将不等于 D 针对 C 的核酸序列同一性%。作为计算核酸序列同一性%的一个实例,表 4 和 5 说明如何计算指定为"对照 DNA"的核酸序列与指定为"PRO-DNA"的核酸序列的同一性%。

20 除非另外特别声明,否则本文所用所有核酸序列同一性%值是按照上述使用 ALIGN-2 序列对比计算机程序获得的。然而,核酸序列同一性%也可使用序列对比程序 NCBI-BLAST2 (Altschul 等, Nucleic Acids Res. 25: 3389-3402 (1997))进行测定。NCBI-BLAST2 序列对比程序可从 <http://www.ncbi.nlm.nih.gov> 下载,或从 National Institute of Health, Bethesda, MD 得到。

25 NCBI-BLAST2 使用数种检索参数,其中所有这些参数设定为默认值,包括例如,非掩蔽=yes,链=全部,预期的出现=10,最低复杂长度=15/5,多-程 e-值=0.01,多-程常数=25,最终缺口化对比的陡降=25,和评分矩阵=BLOSUM62。

30 当使用 NCBI-BLAST2 进行序列比较时,给定核酸序列 C 与给定核酸序列 D 的核酸序列同一性%(或者说,给定核酸序列 C 具有或含有与给定核酸序列 D 相同的核酸序列的%)如下计算:

W/Z 比值乘以 100

其中 W 是用序列对比程序 NCBI-BLAST2 比较 C 和 D 后计算出的相同核苷酸数量, 和其中 Z 是 D 的核苷酸总数量。可以理解, 当核酸序列 C 与核酸序列 D 的长度不相等时, C 针对 D 的核酸序列同一性%将不等于 D 针对 C 的核酸序列同一性%。

在另一实施方案中, FGF-19 变体多核苷酸是编码活性 FGF-19 多肽的核酸分子, 它能够与编码图 2 (SEQ ID NO: 2) 所示全长 FGF-19 多肽的核苷酸序列杂交 (优选地在严格杂交和洗涤条件下杂交)。FGF-19 变体多肽可以是由 FGF-19 变体多核苷酸编码。

10 在如上述进行氨基酸序列同一性比较时使用的术语"阳性", 不仅包括所比较的序列中相同的氨基酸残基, 也包括其中具有相似特性的氨基酸残基。针对目的氨基酸残基评分为阳性值的氨基酸残基是那些与目的氨基酸残基相同或是目的氨基酸残基的首选取代的氨基酸残基。(见下表 6 中定义)。

15 为了本发明目的, 给定氨基酸序列 A 相对于给定氨基酸序列 B 的氨基酸序列阳性值% (或者说: 给定氨基酸序列 A 具有或含有与给定氨基酸序列 B 相同的氨基酸序列的阳性%) 如下计算:

X/Y 比值乘以 100

20 其中 X 是用上述序列对比程序 ALIGN-2 比较 A 和 B 后得出的如上定义的评分阳性氨基酸数量, 其中 Y 是 B 的氨基酸残基总数量。可以理解, 当氨基酸序列 A 与氨基酸序列 B 的长度不相等时, A 针对 B 的阳性%将不等于 B 针对 A 的阳性%。

25 当用"分离的"来描述本文的各种多肽时, 是指已从天然环境的组成中鉴定和分离和/或回收的多肽。优选地, 分离的多肽与其天然结合的所有组分没有任何结合。多肽的天然环境中的污染成分是那些将干扰使用多肽进行诊断和治疗的物质, 可包括酶、激素和其它蛋白性或非-蛋白性溶质。在优选的实施方案中, 多肽被纯化为: (1) 使用转杯式蛋白测序仪, 达到足以获得 N-末端或内部氨基酸序列至少 15 个残基的程度, 或者(2) 达到在非-还原或还原条件下经 SDS-PAGE 电泳和考马斯蓝或优选银染证实为同质性的程度。分离的多肽包括在重组细胞内的原位多肽, 因为 FGF-19 自然环境中的至少一个组分是不存在的。然而, 通常用至少一个纯化步骤来制备分离的多肽。

30 编码 FGF-19 多肽的"分离的"核酸分子是从 FGF-19 编码核酸的天然来

源中通常与其结合的至少一种杂质核酸分子中鉴定并分离的核酸分子。优选地，所述分离的核酸与其天然结合的所有组分没有任何结合。一分离的编码 FGF-19 的核酸分子不同于其天然形式或组合(setting)。因此分离的编码 FGF-19 的核酸分子区别于其在天然细胞中的存在形式。但是，分离的编码 FGF-19 多肽的核酸分子，包括通常表达 FGF-19 的细胞中所含的 FGF-19-编码核酸分子，例如，所述核酸分子与天然细胞所在的染色体位置不同。

术语“控制序列”指在特定宿主生物中使可操作连接的编码序列表达所需的 DNA 序列。适于原核生物的调控序列可包括如启动子，任选操纵子序列，和核糖体结合位点。已知真核生物细胞利用启动子，聚腺苷酸化信号和增强子。

当一种核酸与另一种核酸序列处于功能相关的位置时，该核酸与该另一种核酸“可操作相连”。例如如果一种多肽表达为参与该多肽的分泌的前体蛋白，则将前序列或分泌前导序列的 DNA 与该多肽的 DNA 可操作相连；启动子或增强子当能影响编码序列的转录时可与该编码序列可操作相连；或核糖体结合位点当其位置能促进翻译时可与编码序列可操作相连。通常“可操作相连”指被连接的 DNA 序列是相邻的，而且，如果是分泌前导序列，则相邻且为阅读状态。然而，增强子不必是相邻的。连接可通过在方便的限制性位点相连而实现。如果不存在这样的位点，可根据常规实践使用合成的寡核苷酸衔接子或接头。

本文所用术语“抗体”是指最广义上的抗体，并特别包括例如单链抗-FGF-19 单克隆抗体(包括激动剂、拮抗剂和中和抗体)、具有多表位特异性的抗-FGF-19 抗体组合物、单链抗-FGF-19 抗体和抗-FGF-19 抗体的片段(见下述)。本文中术语“单克隆抗体”，是指来自基本上同质的抗体群的抗体，即除了可能少量存在的天然突变以外，该抗体群中的各个抗体均相同。

杂交反应的“严格度”由本领域技术人员很容易地确定，通常根据经验在考虑探针长度、洗涤温度和盐浓度的基础上计算得出。一般而言，长的探针需要较高的温度以便正确退火，而短的探针需要较低温度。杂交通常取决于变性的 DNA 在低于解链温度的环境中存在互补链时重退火能力。探针和杂交序列之间所需同源程度越高，则可使用的相对温度越高。结果是，相对温度较高将使反应条件趋于更加严格，而较低温度则使反应的严谨度较小。有关杂交反应的严格度的其它细节和解释，见 Ausubel 等，Current Protocols in

Molecular Biology, Wiley Interscience Publishers, (1995).

本文的"严格条件"或"高度严格条件",可以定义为:(1) 使用低离子强度和高温洗涤,例如, 50°C 0.015M 氯化钠/0.0015M 柠檬酸钠/0.1% 十二烷基硫酸钠;(2) 在杂交过程中 42°C 使用变性剂如甲酰胺,譬如,含 0.1% 牛血清白蛋白的 50% (v/v) 甲酰胺/0.1% Ficoll/0.1% 聚乙烯吡咯烷酮/50mM 磷酸钠缓冲液 pH 6.5 及 750 mM 氯化钠, 75 mM 柠檬酸钠;或者(3) 42°C 使用 50% 甲酰胺、5 x SSC (0.75M NaCl, 0.075 M 柠檬酸钠)、50 mM 磷酸钠(pH 6.8)、0.1% 焦磷酸钠、5 x Denhardt's 溶液、超声处理的鲑精 DNA(50 µg/ml)、0.1% SDS 和 10% 硫酸葡聚糖,于 42°C 在 0.2 x SSC(氯化钠/柠檬酸钠)中和于 55°C 在 50% 甲酰胺中洗涤,接着于 55°C 用 0.1 x 含 EDTA 的 SSC 进行高度严格洗涤。

"中等严格条件"可以按照 Sambrook 等, Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989 的定义,包括使用严格度低于前述的洗涤溶液和杂交条件(例如,温度,离子强度和 SDS%)。中等严格条件的一个实例是,于 37°C 在如下组成的溶液中过夜温育: 20% 甲酰胺, 5 x SSC (150 mM NaCl, 15 mM 柠檬酸三钠), 50 mM 磷酸钠(pH 7.6), 5 x Denhardt's 溶液, 10% 硫酸葡聚糖和 20 mg/ml 变性剪切的鲑精 DNA,接着于 37-50°C 在 1 x SSC 中洗涤杂交膜。本领域熟练技术人员懂得如何必要地调整温度、离子强度等条件以适应诸如探针长度等因素。

本文使用的术语"表位标记",是指含有与"标记多肽"融合了的 FGF-19 多肽的嵌合多肽。该标记多肽具有足够多的残基以提供针对所制备抗体的表位,而且要足够短以便它不干预与其融合的多肽的活性。标记多肽也优选确实是独一无二的,这样抗体与其它表位基本上不发生交叉反应。通常,适宜的标记多肽至少有 6 个氨基酸残基,一般为约 8~50 个氨基酸残基(优选为 10~20 个氨基酸残基)。

在本文中术语"免疫粘附素"定义为抗体样分子,其将异源"粘附素"蛋白(例如受体、配体或酶)的结合结构域与免疫球蛋白恒定区组合。结构上,免疫粘附素包括具有所需结合特异性之粘附素氨基酸残基与免疫球蛋白恒定区序列的融合,所述粘附素氨基酸不同于抗体的抗原识别和结合位点(抗原结合位点)(即为"异源性")。免疫粘附分子的粘附素部分通常是包括受体或配体的至少一个结合位点的邻接氨基酸序列。免疫粘附素中的免疫球蛋白

恒定区序列, 可以得自任何免疫球蛋白, 如 IgG-1、IgG-2、IgG-3 或 IgG-4 亚型、IgA (包括 IgA-1 和 IgA-2)、IgE、IgD 或 IgM。

本文中"活性"是指保留完整或天然 FGF-19 的生物学和/或免疫学活性的 FGF-19 形式, 其中"生物学"活性指完整或天然 FGF-19 引起的生物学功能(抑制或兴奋), 不是指诱导产生针对完整或天然 FGF-19 所拥有的表位的抗体的能力; 而"免疫学"活性则是指诱导产生针对完整或天然 FGF-19 所拥有的表位的抗体的能力。优选的生物活性包括下述任何一个或多个活性: 增加个体的代谢(或代谢速率), 降低个体体重, 减轻个体肥胖, 减少葡萄糖摄取进入脂肪细胞, 增加 leptin 自脂肪细胞释放, 降低个体的甘油三酯水平和减少个体的游离脂肪酸。应当懂得, FGF-19 的部分活性是由 FGF-19 直接诱导的, 而有些活性是间接诱导的, 但是, 每种活性均是 FGF-19 存在的结果, 如果没有 FGF-19, 将不会产生这些结果。

本文中术语"拮抗剂", 使用其广义含义, 包括部分或完全阻断、抑制或中和本文所述天然 FGF-19 多肽生物活性的任何分子。类似地, 术语"激动剂"也使用其广义含义, 包括模拟本文所述天然 FGF-19 多肽生物活性的任何分子。适宜的激动剂或拮抗剂分子特别包括: 激动剂或拮抗剂抗体或抗体片段、天然 FGF-19 多肽的片段或氨基酸序列变体、肽、小的有机分子等。鉴定 FGF-19 多肽的激动剂或拮抗剂的方法可以是, 使 FGF-19 多肽与待测激动剂或拮抗剂分子接触, 并测定正常情况下与 FGF-19 多肽有关的一个或多个生物活性的可测得变化。

"治疗"是指治疗性治疗和预防性措施, 其目的是阻止或延缓(减轻)所针对的病理性病症。那些需要治疗的个体包括, 已经患有病症或有患病倾向或者为了预防患上病症的个体。

"慢性"给药, 指与快速给药模式相反, 将试剂以连续给药模式施用, 以便在长时间内保持初始治疗效果(活性)。
"间断"给药是指治疗不是没有间歇地连续进行, 而是以周期性为特征。

意欲治疗的"哺乳动物"是指归为哺乳动物的任何动物, 包括人、家畜和农场动物, 以及动物园里的动物、参与运动项目的动物或宠物如狗、马、猫、牛等。优选所述哺乳动物为人类。

"个体"是指任何主体, 优选是哺乳动物, 更优选是人。

"肥胖"指哺乳动物的体重指数(BMI)至少为 25.9 的状况, 该指数按照体

重(kg)除以高度(米)的平方来计算。通常,具有正常体重的人的 BMI 为 19.9~ 小于 25.9。本文所述肥胖可以是任何起因,无论是遗传还是环境因素所致。

导致肥胖或说是肥胖原因的病症的实例包括: 过度进食和食欲过盛、多囊卵巢疾病、颅咽管瘤、Prader Willi 综合征、弗勒赫利希综合征(Frohlich's syndrome)、II 型糖尿病、GH-缺陷型个体、身材矮小常态变异、特纳综合征(Turner's syndrome), 和其它表现为代谢活动减低或静态能量消耗占不含脂肪的物质总量(fat-free mass)的比例下降的病理性病症, 例如, 患急性淋巴细胞性白血病的儿童。

"与肥胖有关的病症"是指是造成肥胖或使肥胖加重的病症, 例如但不限于皮肤病如感染、静脉曲张、黑棘皮症、以及湿疹, 运用耐力差, 糖尿病, 胰岛素抵抗, 高血压, 血胆固醇过多症, 胆石病, 骨关节炎, 整形术损伤, 血栓性疾病, 癌症和冠心病(或心血管病), 特别是那些与个体高甘油三酯和游离脂肪酸有关的心血管病症。

与一个或多个其它治疗剂"联合"给药包括同时给药和以任何顺序的联合给药。

本文所用"载体"包括在所用剂量和浓度下对细胞或哺乳动物无毒性的可药用载体、赋形剂、稳定剂。可药用载体常常是含水的 pH 缓冲的溶液。可药用载体的实例包括缓冲剂如磷酸盐、柠檬酸盐和其它有机酸; 抗氧化剂包括抗坏血酸; 低分子量(小于 10 个残基)多肽; 蛋白, 如白蛋白、明胶或免疫球蛋白; 亲水聚合物如聚乙烯吡咯烷酮; 氨基酸如甘氨酸、谷氨酰胺、天冬酰胺、精氨酸或赖氨酸; 单糖, 二糖和其它碳水化合物, 包括葡萄糖、甘露糖、或糊精; 螯合剂如 EDTA; 糖醇如甘露醇或山梨醇; 成盐离子如钠; 和/或非离子表面活性剂如 TWEENTM、聚乙二醇(PEG)和 PLURONICTM。

"抗体片段"包含完整抗体的一部分, 优选是完整抗体的抗原结合或可变区。抗体片段的实例包括 Fab、Fab'、F(ab')₂ 和 Fv 片段; 二价抗体; 线形抗体(Zapata 等, Protein Eng. 8 (10): 1057-1062 [1995]); 单链抗体分子; 和由抗体片段形成的多特异抗体。

木瓜蛋白酶消化抗体可产生两个相同的各带有单个抗原结合位点的抗原结合片段(称为 "Fab" 片段)和残余的 "Fc" 片段, Fc 段的名称反应了其易于结晶的能力。经胃蛋白酶处理可产生具有两个抗原结合位点并仍然能交联抗原的 F(ab')₂ 片段。

“Fv”是含有完整的识别和结合位点的最小抗体片段。此区由一个重链可变区与一个轻链可变区紧密地非共价连接形成的二聚体组成。在这个构象中每个可变区的三个 CDR 相互作用，在 V_H - V_L 二聚体表面限定一个抗原结合位点。这六个 CDR 共同赋予抗体以抗原结合特异性。然而，即使是单个可变区(或 F_v 的一半仅含有三个抗原特异性 CDR)也具有识别和结合抗原的能力，尽管与完整的结合位点相比其亲和力较低。

Fab 段还包括轻链恒定区和重链的第一个恒定区(CH1)。Fab'与 Fab 的差别在于 Fab'在重链 CH1 的羧基末端多出几个残基，包括抗体铰链区的一个或多个半胱氨酸。Fab'-SH 在本文中是指恒定区半胱氨酸残基中至少有一个游离巯基的 Fab'。F(ab')₂ 抗体片段在最初产生为 Fab'片段对，在它们之间具有铰链区半胱氨酸。抗体片段的其它化学偶联是众所周知的。

脊椎动物任何物种的抗体(免疫球蛋白)的“轻链”，可依据其恒定区氨基酸序列而归为完全不同的两型(称为 κ 和 λ)中的一型。

根据其重链恒定区的氨基酸序列，可将免疫球蛋白分为不同类。主要有 5 类免疫球蛋白：IgA、IgD、IgE、IgG 和 IgM，其中一些还可进一步分成“亚类”(同种型)，例如 IgG1、IgG2、IgG3、IgG4、IgA1 和 IgA2。

单链 Fv”或“scFv”抗体片段，包括抗体的 V_H 和 V_L 结构域，这些结构域存在于单个多肽链上。优选 Fv 多肽在 V_H 和 V_L 结构域之间还包含一个多肽接头，它能使 sFv 形成抗原结合所需的结构。关于 scFv 的综述见 Pluckthun 在《单克隆抗体的药理学》，第 113 卷，Rosenburg 和 Moore 编，Springer-Verlag, New York, 第 269-315 页(1994)。

术语“二价抗体”是指具有两个抗原结合位点的小分子抗体片段，这些片段在一条多肽链(V_H - V_L)上含有相连的一个重链可变区(V_H)和一个轻链可变区(V_L)。利用一种非常短的接头，其使得同一条链上的两个结构域无法配对，不得与另一条链上的互补结构域配对，从而形成两个抗原结合位点。在 EP 404,097; WO93/11161; 和 Hollinger 等，美国国家科学院院刊，90: 6444-6448(1993)中有对二价抗体的更详细描述。

“分离的”抗体是已从其天然环境的组成中鉴定、分离和/或回收的抗体。其天然环境的污染成分是干扰抗体的诊断或治疗应用的物质，可包括酶、激素和其它蛋白性或非蛋白性溶质。在优选的实施方案中，该抗体的纯度应达到：(1)经 Lowery 法确定的抗体重量的 95% 以上，最优选所述重量的 99%

以上, (2)足以获得用旋转杯状(spinning cup)序列分析仪所测至少 15 个残基的 N 端或内部氨基酸序列, (3)通过还原或非还原条件下的 SDS-PAGE 以及考马斯亮蓝染色、优选银染色所证实的同质性。分离的抗体包括重组细胞内的原位抗体, 因为该抗体的自然环境中的至少一种组分已不存在。一般情况下分离的抗体可通过至少一个纯化步骤来制备。

本文中"标记"一词, 是指直接或间接地与抗体偶联从而产生"标记"抗体的可检测的化合物或组分。标记可以是其本身可被测得(例如放射性同位素标记或荧光标记), 或者如果是酶标记时, 可催化底物化合物或组分发生化学变化, 这种变化是可检测的。

10 "固相"是指本发明抗体可粘附的非水性基质。本文中固相的实例包括部分或全部由玻璃(如控制孔径的玻璃)、多糖(如琼脂糖)、聚丙烯酰胺、聚苯乙烯、聚乙烯醇和硅氧烷制成的那些固相。在某些实例中, 根据上下文的内容, 固相可包括测试板的孔; 在其它例子中, 它可以是纯化柱(如亲和层析柱)。该术语也包括美国专利 4275149 中公开的分散颗粒的不连续固相。

15 "脂质体"是由能向哺乳动物有效运送药物(如本文公开的抗 FGF-19 多肽和/或其抗体)的各类脂质、磷脂和/或表面活性剂组成的小分子囊泡。脂质体的组分通常排列为双层形式, 与生物膜的脂质排列相似。

本文中"小分子"定义为分子量低于约 500 道尔顿的分子。

表 1

(此处插入原文第 34 - 50 页的计算机命令符)

Table 1

```

/*
 *
 * C-C increased from 12 to 15
 * Z is average of EQ
 * B is average of ND
 * match with stop is _M; stop-stop = 0; J (joker) match = 0
 */
#define _M      -8      /* value of a match with a stop */

int
  _day[26][26] = {
/* A */ { 2, 0, -2, 0, 0, -4, 1, -1, -1, 0, -1, -2, -1, 0, _M, 1, 0, -2, 1, 1, 0, 0, -6, 0, -3, 0},
/* B */ { 0, 3, -4, 3, 2, -5, 0, 1, -2, 0, 0, -3, -2, 2, _M, -1, 1, 0, 0, 0, 0, -2, -5, 0, -3, 1},
/* C */ {-2, -4, 15, -5, -5, -4, -3, -3, -2, 0, -5, -6, -5, -4, _M, -3, -5, -4, 0, -2, 0, -2, -8, 0, 0, -5},
/* D */ { 0, 3, -5, 4, 3, -6, 1, 1, -2, 0, 0, -4, -3, 2, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 2},
/* E */ { 0, 2, -5, 3, 4, -5, 0, 1, -2, 0, 0, -3, -2, 1, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 3},
/* F */ {-4, -5, -4, -6, -5, 9, -5, -2, 1, 0, -5, 2, 0, -4, _M, -5, -5, -4, -3, -3, 0, -1, 0, 0, 7, -5},
/* G */ { 1, 0, -3, 1, 0, -5, 5, -2, -3, 0, -2, -4, -3, 0, _M, -1, -1, -3, 1, 0, 0, -1, -7, 0, -5, 0},
/* H */ {-1, 1, -3, 1, 1, -2, -2, 6, -2, 0, 0, -2, -2, 2, _M, 0, 3, 2, -1, -1, 0, -2, -3, 0, 0, 2},
/* I */ {-1, -2, -2, -2, -2, 1, -3, -2, 5, 0, -2, 2, 2, -2, _M, -2, -2, -2, -1, 0, 0, 4, -5, 0, -1, -2},
/* J */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */ {-1, 0, -5, 0, 0, -5, -2, 0, -2, 0, 5, -3, 0, 1, _M, -1, 1, 3, 0, 0, 0, -2, -3, 0, -4, 0},
/* L */ {-2, -3, -6, -4, -3, 2, -4, -2, 2, 0, -3, 6, 4, -3, _M, -3, -2, -3, -3, -1, 0, 2, -2, 0, -1, -2},
/* M */ {-1, -2, -5, -3, -2, 0, -3, -2, 2, 0, 0, 4, 6, -2, _M, -2, -1, 0, -2, -1, 0, 2, -4, 0, -2, -1},
/* N */ { 0, 2, -4, 2, 1, -4, 0, 2, -2, 0, 1, -3, -2, 2, _M, -1, 1, 0, 1, 0, 0, -2, -4, 0, -2, 1},
/* O */ { _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, 0, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M},
/* P */ { 1, -1, -3, -1, -1, -5, -1, 0, -2, 0, -1, -3, -2, -1, _M, 6, 0, 0, 1, 0, 0, -1, -6, 0, -5, 0},
/* Q */ { 0, 1, -5, 2, 2, -5, -1, 3, -2, 0, 1, -2, -1, 1, _M, 0, 4, 1, -1, -1, 0, -2, -5, 0, -4, 3},
/* R */ {-2, 0, -4, -1, -1, -4, -3, 2, -2, 0, 3, -3, 0, 0, _M, 0, 1, 6, 0, -1, 0, -2, 2, 0, -4, 0},
/* S */ { 1, 0, 0, 0, 0, -3, 1, -1, -1, 0, 0, -3, -2, 1, _M, 1, -1, 0, 2, 1, 0, -1, -2, 0, -3, 0},
/* T */ { 1, 0, -2, 0, 0, -3, 0, -1, 0, 0, 0, -1, -1, 0, _M, 0, -1, -1, 1, 3, 0, 0, -5, 0, -3, 0},
/* U */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */ { 0, -2, -2, -2, -2, -1, -1, -2, 4, 0, -2, 2, 2, -2, _M, -1, -2, -2, -1, 0, 0, 4, -6, 0, -2, -2},
/* W */ {-6, -5, -8, -7, -7, 0, -7, -3, -5, 0, -3, -2, -4, -4, _M, -6, -5, 2, -2, -5, 0, -6, 17, 0, 0, -6},
/* X */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */ {-3, -3, 0, -4, -4, 7, -5, 0, -1, 0, -4, -1, -2, -2, _M, -5, -4, -4, -3, -3, 0, -2, 0, 0, 10, -4},
/* Z */ { 0, 1, -5, 2, 3, -5, 0, 2, -2, 0, 0, -2, -1, 1, _M, 0, 3, 0, 0, 0, 0, -2, -6, 0, -4, 4}
};

```

Table 1 (cont')

```

/*
*/
#include <stdio.h>
#include <ctype.h>
#define MAXJMP 16 /* max jumps in a diag */
#define MAXGAP 24 /* don't continue to penalize gaps larger than this */
#define JMPS 1024 /* max jmps in an path */
#define MX 4 /* save if there's at least MX-1 bases since last jmp */
#define DMAT 3 /* value of matching bases */
#define DMIS 0 /* penalty for mismatched bases */
#define DINS0 8 /* penalty for a gap */
#define DINS1 1 /* penalty per base */
#define PINS0 8 /* penalty for a gap */
#define PINS1 4 /* penalty per residue */
struct jmp {
    short n[MAXJMP]; /* size of jmp (neg for dely) */
    unsigned short x[MAXJMP]; /* base no. of jmp in seq x */
}; /* limits seq to 2^16 -1 */
struct diag {
    int score; /* score at last jmp */
    long offset; /* offset of prev block */
    short ijmp; /* current jmp index */
    struct jmp jp; /* list of jmps */
};
struct path {
    int spc; /* number of leading spaces */
    short n[JMP]; /* size of jmp (gap) */
    int x[JMP]; /* loc of jmp (last elem before gap) */
};
char *ofile; /* output file name */
char *name[2]; /* seq names: getseqs() */
char *prog; /* prog name for err msgs */
char *seq[2]; /* seqs: getseqs() */
int dmax; /* best diag: nw() */
int dmax0; /* final diag */
int dna; /* set if dna: main() */
int endgaps; /* set if penalizing end gaps */
int gapx, gapy; /* total gaps in seqs */
int len0, len1; /* seq lens */
int ngapx, ngapy; /* total size of gaps */
int smax; /* max score: nw() */
int *xbm; /* bitmap for matching */
long offset; /* current offset in jmp file */
struct diag dx; /* holds diagonals */
struct path pp[2]; /* holds path for seqs */
char *calloc(), *malloc(), *index(), *strcpy();
char *getseq(), *g_calloc();

```

Table 1 (cont')

```

/* Needleman-Wunsch alignment program
*
* usage: progs file1 file2
* where file1 and file2 are two dna or two protein sequences.
* The sequences can be in upper- or lower-case and may contain ambiguity
* Any lines beginning with ';', '>' or '<' are ignored
* Max file length is 65535 (limited by unsigned short x in the jmp struct)
* A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
* Output is in the file "align.out"
*
* The program may create a tmp file in /tmp to hold info about traceback.
* Original version developed under BSD 4.3 on a vax 8650
*/
#include "nw.h"
#include "day.h"
static  _dbval[26] = {
    1, 14, 2, 13, 0, 0, 4, 11, 0, 0, 12, 0, 3, 15, 0, 0, 0, 5, 6, 8, 8, 7, 9, 0, 10, 0
};
static  _pbval[26] = {
    1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
    128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
    1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
    1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};
main(ac, av)
int      ac;
char     *av[];
{
    prog = av[0];
    if (ac != 3) {
        fprintf(stderr, "usage: %s file1 file2\n", prog);
        fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
        fprintf(stderr, "The sequences can be in upper- or lower-case\n");
        fprintf(stderr, "Any lines beginning with ';', '>' or '<' are ignored\n");
        fprintf(stderr, "Output is in the file \"align.out\"\n");
        exit(1);
    }
    namex[0] = av[1];
    namex[1] = av[2];
    seqx[0] = getseq(namex[0], &len0);
    seqx[1] = getseq(namex[1], &len1);
    xbm = (dna)? _dbval : _pbval;
    endgaps = 0; /* 1 to penalize endgaps */
    ofile = "align.out"; /* output file */
    nw(); /* fill in the matrix, get the possible jmps */
    readjmps(); /* get the actual jmps */
    print(); /* print stats, alignment */
    cleanup(0); /* unlink any tmp files */
}
main

```

Table 1 (cont')

```

/* do the alignment, return best score: main()
* dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
* pro: PAM 250 values
* When scores are equal, we prefer mismatches to any gap, prefer
* a new gap to extending an ongoing gap, and prefer a gap in seqx
* to a gap in seq y.
*/
nw()
{
    char          *px, *py;          /* seqs and ptrs */
    int           *ndely, *dely;     /* keep track of dely */
    int           ndelx, delx;       /* keep track of delx */
    int           *tmp;              /* for swapping row0, row1 */
    int           mis;               /* score for each type */
    int           ins0, ins1;        /* insertion penalties */
    register      id;                /* diagonal index */
    register      ij;                /* jmp index */
    register      *col0, *col1;      /* score for curr, last row */
    register      xx, yy;            /* index into seqs */
    dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));
    ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
    dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
    col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
    col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
    ins0 = (dna)? DINS0 : PINS0;
    ins1 = (dna)? DINS1 : PINS1;
    smax = -10000;
    if (endgaps) {
        for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
            col0[yy] = dely[yy] = col0[yy-1] - ins1;
            ndely[yy] = yy;
        }
        col0[0] = 0; /* Waterman Bull Math Biol 84 */
    }
    else
        for (yy = 1; yy <= len1; yy++)
            dely[yy] = -ins0;
    /* fill in match matrix
    */
    for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
        */
        if (endgaps) {
            if (xx == 1)
                col1[0] = delx = -(ins0+ins1);
            else
                col1[0] = delx = col0[0] - ins1;
            ndelx = xx;
        }
        else {
            col1[0] = 0;
            delx = -ins0;
            ndelx = 0;
        }
    }
}

```

Table 1 (cont')

...DW

```

for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
  mis = col0[yy-1];
  if (dna)
    mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
  else
    mis += _day[*px-'A'][*py-'A'];

  /* update penalty for del in x seq;
   * favor new del over ongong del
   * ignore MAXGAP if weighting endgaps
   */
  if (endgaps || ndely[yy] < MAXGAP) {
    if (col0[yy] - ins0 >= dely[yy]) {
      dely[yy] = col0[yy] - (ins0+ins1);
      ndely[yy] = 1;
    } else {
      dely[yy] -= ins1;
      ndely[yy]++;
    }
  } else {
    if (col0[yy] - (ins0+ins1) >= dely[yy]) {
      dely[yy] = col0[yy] - (ins0+ins1);
      ndely[yy] = 1;
    } else
      ndely[yy]++;
  }

  /* update penalty for del in y seq;
   * favor new del over ongong del
   */
  if (endgaps || ndelx < MAXGAP) {
    if (col1[yy-1] - ins0 >= delx) {
      delx = col1[yy-1] - (ins0+ins1);
      ndelx = 1;
    } else {
      delx -= ins1;
      ndelx++;
    }
  } else {
    if (col1[yy-1] - (ins0+ins1) >= delx) {
      delx = col1[yy-1] - (ins0+ins1);
      ndelx = 1;
    } else
      ndelx++;
  }

  /* pick the maximum score; we're favoring
   * mis over any del and delx over dely
   */

```

Table 1 (cont')

```

id = xx - yy + len1 - 1;
if (mis >= delx && mis >= dely[yy])
    col1[yy] = mis;
else if (delx >= dely[yy]) {
    col1[yy] = delx;
    ij = dx[id].ijmp;
    if (dx[id].jp.n[0] && (!dna || (ndelx >= MAXJMP
    && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
        dx[id].ijmp++;
        if (++ij >= MAXJMP) {
            writejumps(id);
            ij = dx[id].ijmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
        }
    }
    dx[id].jp.n[ij] = ndelx;
    dx[id].jp.x[ij] = xx;
    dx[id].score = delx;
}
else {
    col1[yy] = dely[yy];
    ij = dx[id].ijmp;
if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
    && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
        dx[id].ijmp++;
        if (++ij >= MAXJMP) {
            writejumps(id);
            ij = dx[id].ijmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
        }
    }
    dx[id].jp.n[ij] = -ndely[yy];
    dx[id].jp.x[ij] = xx;
    dx[id].score = dely[yy];
}
if (xx == len0 && yy < len1) {
    /* last col
    */
    if (endgaps)
        col1[yy] -= ins0+ins1*(len1-yy);
    if (col1[yy] > smax) {
        smax = col1[yy];
        dmax = id;
    }
}
}
if (endgaps && xx < len0)
    col1[yy-1] -= ins0+ins1*(len0-xx);
if (col1[yy-1] > smax) {
    smax = col1[yy-1];
    dmax = id;
}
}
tmp = col0; col0 = col1; col1 = tmp;
}
(void) free((char *)ndely);
(void) free((char *)dely);
(void) free((char *)col0);
(void) free((char *)col1);
}

```

Table 1 (cont')

```

/*
 *
 * print() -- only routine visible outside this module
 *
 * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[]: print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
 * stars() - -put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
 */

#include "nw.h"

#define SPC 3
#define P_LINE 256 /* maximum output line */
#define P_SPC 3 /* space between name or num and seq */

extern _day[26][26];
int olen; /* set output line length */
FILE *fx; /* output file */

print()
{
    int lx, ly, firstgap, lastgap; /* overlap */

    if ((fx = fopen(ofile, "w")) == 0) {
        fprintf(stderr, "%s: can't write %s\n", prog, ofile);
        cleanup(1);
    }
    fprintf(fx, "< first sequence: %s (length = %d)\n", namex[0], len0);
    fprintf(fx, "< second sequence: %s (length = %d)\n", namex[1], len1);
    olen = 60;
    lx = len0;
    ly = len1;
    firstgap = lastgap = 0;
    if (dmax < len1 - 1) { /* leading gap in x */
        pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
    else if (dmax > len1 - 1) { /* leading gap in y */
        pp[1].spc = firstgap = dmax - (len1 - 1);
        lx -= pp[1].spc;
    }
    if (dmax0 < len0 - 1) { /* trailing gap in x */
        lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
    }
    else if (dmax0 > len0 - 1) { /* trailing gap in y */
        lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
    getmat(lx, ly, firstgap, lastgap);
    pr_align();
}

```

Table 1 (cont')

```

/*
 * trace back the best path, count matches
 */
static
getmat(lx, ly, firstgap, lastgap)                                getmat
    int    lx, ly;                                /* "core" (minus endgaps) */
    int    firstgap, lastgap;                       /* leading trailing overlap */
{
    int    nm, i0, i1, siz0, siz1;
    char    outx[32];
    double    pct;
    register    n0, n1;
    register char    *p0, *p1;

    /* get total matches, score
     */
    i0 = i1 = siz0 = siz1 = 0;
    p0 = seqx[0] + pp[1].spc;
    p1 = seqx[1] + pp[0].spc;
    n0 = pp[1].spc + 1;
    n1 = pp[0].spc + 1;

    nm = 0;
    while ( *p0 && *p1 ) {
        if (siz0) {
            p1++;
            n1++;
            siz0--;
        }
        else if (siz1) {
            p0++;
            n0++;
            siz1--;
        }
        else {
            if (xbrm[*p0-'A']&xbrm[*p1-'A'])
                nm++;
            if (n0++ == pp[0].x[i0])
                siz0 = pp[0].n[i0++];
            if (n1++ == pp[1].x[i1])
                siz1 = pp[1].n[i1++];
            p0++;
            p1++;
        }
    }

    /* pct homology:
     * if penalizing endgaps, base is the shorter seq
     * else, knock off overhangs and take shorter core
     */
    if (endgaps)
        lx = (len0 < len1)? len0 : len1;
    else
        lx = (lx < ly)? lx : ly;
    pct = 100.*(double)nm/(double)lx;
    fprintf(fx, "\n");
    fprintf(fx, "< %d match%s in an overlap of %d: %.2f percent similarity\n",
        nm, (nm == 1)? "" : "es", lx, pct);
}

```

Table 1 (cont')

```

fprintf(fx, "< gaps in first sequence: %d", gapx);
if (gapx) {
    (void) sprintf(outx, " (%d %s%s)",
        ngapx, (dna)? "base":"residue", (ngapx == 1)? "" : "s");
    fprintf(fx, "%s", outx);

fprintf(fx, ", gaps in second sequence: %d", gapy);
if (gapy) {
    (void) sprintf(outx, " (%d %s%s)",
        ngapy, (dna)? "base":"residue", (ngapy == 1)? "" : "s");
    fprintf(fx, "%s", outx);
}
if (dna)
    fprintf(fx,
        "\n< score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
        smax, DMAT, DMIS, DINS0, DINS1);
else
    fprintf(fx,
        "\n< score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
        smax, PINS0, PINS1);
if (endgaps)
    fprintf(fx,
        "< endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
        firstgap, (dna)? "base" : "residue", (firstgap == 1)? "" : "s",
        lastgap, (dna)? "base" : "residue", (lastgap == 1)? "" : "s");
else
    fprintf(fx, "< endgaps not penalized\n");
}

static      nm;          /* matches in core -- for checking */
static      lmax;        /* lengths of stripped file names */
static      ij[2];       /* jmp index for a path */
static      nc[2];       /* number at start of current line */
static      ni[2];       /* current elem number -- for gapping */
static      siz[2];
static char *ps[2];      /* ptr to current element */
static char *po[2];      /* ptr to next output char slot */
static char out[2][P_LINE]; /* output line */
static char star[P_LINE]; /* set by stars() */

/*
 * print alignment of described in struct path pp[]
 */
static
pr_align()
{
    int      nn;          /* char count */
    int      more;
    register i;

    for (i = 0, lmax = 0; i < 2; i++) {
        nn = stripname(name[i]);
        if (nn > lmax)
            lmax = nn;

        nc[i] = 1;
        ni[i] = 1;
        siz[i] = ij[i] = 0;
        ps[i] = seqx[i];
        po[i] = out[i];
    }
}

```

pr_align

Table 1 (cont')

```

for (nn = nm = 0, more = 1; more; ) {
    for (i = more = 0; i < 2; i++) {
        /*
         * do we have more of this sequence?
         */
        if (!*ps[i])
            continue;

        more++;

        if (pp[i].spc) { /* leading space */
            *po[i]++ = ' ';
            pp[i].spc--;
        }
        else if (siz[i]) { /* in a gap */
            *po[i]++ = '-';
            siz[i]--;
        }
        else { /* we're putting a seq element
                */
            *po[i] = *ps[i];
            if (islower(*ps[i]))
                *ps[i] = toupper(*ps[i]);
            po[i]++;
            ps[i]++;

            /*
             * are we at next gap for this seq?
             */
            if (ni[i] == pp[i].x[ij[i]]) {
                /*
                 * we need to merge all gaps
                 * at this location
                 */
                siz[i] = pp[i].n[ij[i]++];
                while (ni[i] == pp[i].x[ij[i]])
                    siz[i] += pp[i].n[ij[i]++];
            }
            ni[i]++;
        }
    }
    if (++nn == olen || !more && nn) {
        dumpblock();
        for (i = 0; i < 2; i++)
            po[i] = out[i];
        nn = 0;
    }
}

/*
 * dump a block of lines, including numbers, stars: pr_align()
 */
static
dumpblock()
{
    register i;

    for (i = 0; i < 2; i++)
        *po[i]-- = '\0';
}

```

...pr_align

dumpblock

Table 1 (cont')

```

                                ...dumpblock
(void) putc('\n', fx);
for (i = 0; i < 2; i++) {
    if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
        if (i == 0)
            nums(i);
        if (i == 0 && *out[1])
            stars();
        putline(i);
        if (i == 0 && *out[1])
            fprintf(fx, star);
        if (i == 1)
            nums(i);
    }
}
}

/*
 * put out a number line: dumpblock()
 */
static
nums(ix)                                nums
{
    int    ix;        /* index in out[] holding seq line */
    char    nline[P_LINE];
    register    i, j;
    register char    *pn, *px, *py;

    for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
        *pn = ' ';
    for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
        if (*py == ' ' || *py == '-')
            *pn = ' ';
        else {
            if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                j = (i < 0)? -i : i;
                for (px = pn; j; j /= 10, px--)
                    *px = j%10 + '0';
                if (i < 0)
                    *px = '-';
            }
            else
                *pn = ' ';
            i++;
        }
    }
    *pn = '\0';
    nc[ix] = i;
    for (pn = nline; *pn; pn++)
        (void) putc(*pn, fx);
    (void) putc('\n', fx);
}

/*
 * put out a line (name, [num], seq, [num]): dumpblock()
 */
static
putline(ix)                                putline
{
    int    ix;

```

Table 1 (cont')

```

...putline

int          i;
register char *px;

for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
    (void) putc(*px, fx);
for (; i < lmax+P_SPC; i++)
    (void) putc(' ', fx);

/* these count from 1:
 * ni[] is current element (from 1)
 * nc[] is number at start of current line
 */
for (px = out[ix]; *px; px++)
    (void) putc(*px&0x7F, fx);
(void) putc('\n', fx);
}

/*
 * put a line of stars (seqs always in out[0], out[1]): dumpblock()
 */
static
stars()
{
    int          i;
    register char *p0, *p1, cx, *px;

    if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
        !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
        return;
    px = star;
    for (i = lmax+P_SPC; i; i--)
        *px++ = ' ';

    for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
        if (isalpha(*p0) && isalpha(*p1)) {
            if (xbm[*p0-'A']&xbm[*p1-'A']) {
                cx = '*';
                nm++;
            }
            else if (!dna && _day[*p0-'A'][*p1-'A'] > 0)
                cx = '.';
            else
                cx = ' ';
        }
        else
            cx = ' ';
        *px++ = cx;
    }
    *px++ = '\n';
    *px = '\0';
}

```

Table 1 (cont')

```
/*
 * strip path or prefix from pn, return len: pr_align()
 */
static
stripname(pn)
    char *pn; /* file name (may be path) */
{
    register char *px, *py;

    py = 0;
    for (px = pn; *px; px++)
        if (*px == '/')
            py = px + 1;
    if (py)
        (void) strcpy(pn, py);
    return(strlen(pn));
}
```

Table 1 (cont')

```

/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * g_malloc() -- calloc() with error checkin
 * readjumps() -- get the good jumps, from tmp file if necessary
 * writejumps() -- write a filled array of jumps to a tmp file: nw()
 */
#include "nw.h"
#include <sys/file.h>

char    *jname = "/tmp/homgXXXXXX";          /* tmp file for jumps */
FILE    *fj;

int      cleanup();                          /* cleanup tmp file */
long     lseek();

/*
 * remove any tmp file if we blow
 */
cleanup(i)                                  cleanup
{
    int    i;
    if (fj)
        (void) unlink(jname);
    exit(i);
}

/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
char    *
getseq(file, len)                           getseq
{
    char    *file;    /* file name */
    int     *len;     /* seq len */

    char    line[1024], *pseq;
    register char *px, *py;
    int     natgc, tlen;
    FILE    *fp;

    if ((fp = fopen(file, "r")) == 0) {
        fprintf(stderr, "%s: can't read %s\n", prog, file);
        exit(1);
    }
    tlen = natgc = 0;
    while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++)
            if (isupper(*px) || islower(*px))
                tlen++;
    }
    if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr, "%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
    }
    pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
}

```

Table 1 (cont')

```

py = pseq + 4;
*len = tlen;
rewind(fp);

while (fgets(line, 1024, fp)) {
    if (*line == ';' || *line == '<' || *line == '>')
        continue;
    for (px = line; *px != '\n'; px++) {
        if (isupper(*px))
            *py++ = *px;
        else if (islower(*px))
            *py++ = toupper(*px);
        if (index("ATGCU", *(py-1)))
            natgc++;
    }
}
*py++ = '\0';
*py = '\0';
(void) fclose(fp);
dna = natgc > (tlen/3);
return(pseq+4);
}

char *
g_alloc(msg, nx, sz)
char *msg;          /* program, calling routine */
int nx, sz;         /* number and size of elements */
{
    char *px, *calloc();

    if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
        if (*msg) {
            fprintf(stderr, "%s: g_alloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
            exit(1);
        }
    }
    return(px);
}

/*
 * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()
 */
readjmps()
{
    int fd = -1;
    int siz, i0, i1;
    register i, j, xx;

    if (fj) {
        (void) fclose(fj);
        if ((fd = open(jname, O_RDONLY, 0)) < 0) {
            fprintf(stderr, "%s: can't open() %s\n", prog, jname);
            cleanup(1);
        }
    }
    for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; ; i++) {
        while (1) {
            for (j = dx[dmax].ijmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                ;
        }
    }
}

```

...getseq

g_alloc

readjmps

Table 1 (cont?)

...readjumps

```

    if (j < 0 && dx[dmax].offset && fj) {
        (void) lseek(fd, dx[dmax].offset, 0);
        (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
        (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
        dx[dmax].ijmp = MAXJMP-1;
    }
    else
        break;
}
if (i >= JMPS) {
    fprintf(stderr, "%s: too many gaps in alignment\n", prog);
    cleanup(1);
}
if (j >= 0) {
    siz = dx[dmax].jp.n[j];
    xx = dx[dmax].jp.x[j];
    dmax += siz;
    if (siz < 0) { /* gap in second seq */
        pp[1].n[i1] = -siz;
        xx += siz;
        /* id = xx - yy + len1 - 1
        */
        pp[1].x[i1] = xx - dmax + len1 - 1;
        gapy++;
        ngapy -= siz;
/* ignore MAXGAP when doing endgaps */
        siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
        i1++;
    }
    else if (siz > 0) { /* gap in first seq */
        pp[0].n[i0] = siz;
        pp[0].x[i0] = xx;
        gapx++;
        ngapx += siz;
/* ignore MAXGAP when doing endgaps */
        siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
        i0++;
    }
}
else
    break;
}

/* reverse the order of jumps
*/
for (j = 0, i0--, j < i0; j++, i0--) {
    i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
    i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
}
for (j = 0, i1--, j < i1; j++, i1--) {
    i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
    i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
}
if (fd >= 0)
    (void) close(fd);
if (fj) {
    (void) unlink(jname);
    fj = 0;
    offset = 0;
}
}

```

Table 1 (cont')

```

/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
writejmps(ix)                                writejmps
{
    int ix;
    char *mktemp();

    if (!fj) {
        if (mktemp(jname) < 0) {
            fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
            cleanup(1);
        }
        if ((fj = fopen(jname, "w")) == 0) {
            fprintf(stderr, "%s: can't write %s\n", prog, jname);
            exit(1);
        }
    }
    (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
    (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
}

```

表 2

PRO XXXXXXXXXXXXXXXX (长度= 15 个氨基酸)

对照蛋白 XXXXXYYYYYYY (长度= 12 个氨基酸)

- 氨基酸序列同一性%=(两个多肽序列之间用 ALIGN-2 测定的相同匹配氨基酸残基的数量)除以(PRO 多肽中氨基酸残基的总数)= 5 除以 15 = 33.3 %

表 3

PRO XXXXXXXXXXX (长度= 10 个氨基酸)

对照蛋白 XXXXXYYYYYYYZZYZ (长度=15 个氨基酸)

- 氨基酸序列同一性%=(两个多肽序列之间用 ALIGN-2 测定的相同匹配氨基酸残基的数量)除以(PRO 多肽中氨基酸残基的总数)= 5 除以 10 = 50 %

表 4

PRO-DNA NNNNNNNNNNNNN (长度= 14 个核苷酸)

- 15 对照 DNA NNNNNLLLLLLLLL (长度=16 个核苷酸)

核酸序列同一性%=(两个核酸序列之间用 ALIGN-2 测定的相同匹配核苷酸的数量)除以(PRO-DNA 核酸序列中核苷酸的总数)= 6 除以 14 = 42.9 %

表 5

- 20 PRO-DNA NNNNNNNNNNN (长度= 12 个核苷酸)

对照 DNA NNNLLLV (长度= 9 个核苷酸)

核酸序列同一性%=(两个核酸序列之间用 ALIGN-2 测定的相同匹配核苷酸的数量)除以(PRO-DNA 核酸序列中核苷酸的总数)= 4 除以 12 = 33.3 %

- 25 II. 本发明的组合物和方法

A. 全长 FGF-19 多肽

- 30 本发明提供编码本申请称为 FGF-19 的多肽的新鉴定和分离的核苷酸序列(或者也称为 UNQ334)。更具体而言, 如下面的实施例进一步详细公开的那样, 已经鉴定和分离出编码 FGF-19 多肽的 cDNA。应注意到在不同表达循环中产生的蛋白可能具有不同的 PRO 数, 但是, 任何特定 DNA 和其编码的蛋白的 UNQ 数是独一无二的, 而且将不会变化。然而, 为了简明起见,

在本说明书中将 DNA49435-1219 编码的蛋白以及前述定义的 FGF-19(有时也称 PR0533)的所有其它天然同系物和变体均称为"FGF-19", 而不论其来源或制备模式如何。

5 如下面的实施例所公开的那样, 本文的 DNA49435-1219 cDNA 克隆已经在 ATCC 保藏。该克隆的实际核苷酸序列, 可由本领域熟练技术人员使用本领域常规方法通过对保藏克隆进行序列分析而很容易地测得。使用本领域常规技术可以预测该核苷酸序列编码的氨基酸序列。关于本文所述 FGF-19 多肽和其编码核酸, 申请人已经确定此时在阅读框上什么是与现有序列信息最一致的。

10 使用上述 ALIGN-2 序列对比计算机程序, 业已发现: 全长天然 FGF-19 序列(见图 2 和 SEQ ID NO: 2)与 AF007268_1 具有一定的氨基酸序列同一性。因此, 目前据信本申请公开的 FGF-19 多肽是成纤维细胞生长因子蛋白家族的新成员, 而且可能具有该蛋白家族通常具有的一种或多种生物学和/或免疫学活性或者特性。

15 B. FGF-19 变体

除了本文描述的全长天然序列的 FGF-19 多肽外, 还可以制备 FGF-19 变体。通过在 FGF-19DNA 中引入适当的核苷酸改变和/或合成所需 FGF-19 多肽, 可制备 FGF-19 变体。本领域技术人员懂得, 氨基酸变化将会改变 FGF-19 的翻译后过程, 例如改变糖基化位点的数量和位置或者改变膜锚着

20 特性。

使用例如在美国专利 5364934 号中阐述的保守和非-保守突变的任一技术和指导路线, 可以在天然全长序列 FGF-19 或者在本文所述 FGF-19 的各种结构域中制造变化。变化可以是取代、缺失或插入编码 FGF-19 的一个或多个密码子, 它们导致相对于天然序列 FGF-19 的 FGF-19 氨基酸序列变化。

25 任选地, 变化是 FGF-19 一个或多个结构域中的至少一个氨基酸被任何其它氨基酸取代。将 FGF-19 序列与已知同源蛋白分子进行比较, 并使高度同源区氨基酸序列数量变化最小, 可以建立决定哪个氨基酸残基可以被插入、取代或缺失而对所需活性无负面影响的指导原则。氨基酸取代可以是一个氨基酸被另一个具有类似结构和/或类似化学特性的氨基酸代替的结果, 如丝氨酸

30 置换亮氨酸, 即保守氨基酸置换。插入或缺失可任选地在约 1~5 个氨基酸的范围内。通过系统地在氨基酸序列中进行插入、缺失或取代并检测所得

变体显示全长或成熟天然序列的活性的情况，可以确定能允许的变化。

本文提供 FGF-19 多肽片段。该类片段可以是，例如，与全长天然蛋白相比，在 N-末端或 C-末端截短的或者是缺乏内部残基的片段。某些片段缺乏那些对于 FGF-19 多肽预期生物活性非必要的氨基酸残基。

- 5 FGF-19 片段可以按照大量常规技术中的任何一个来制备。可以化学合成所需肽片段。另一方法涉及通过酶消化产生 FGF-19 片段，例如，用已知在特定氨基酸残基所限定的位点裂解蛋白的酶处理蛋白，或用适宜的限制性内切酶消化 DNA 并分离所需片段。另一适宜的技术包括分离并用聚合酶链反应(PCR)扩增编码所需多肽片段的 DNA 片段。将限定 DNA 片段所需末端
- 10 的寡核苷酸用于 PCR 引物的 5'和 3'端。优选地，FGF-19 多肽片段具有图 2 (SEQ ID NO: 2)所示天然 FGF-19 多肽的至少一种生物学和/或免疫学活性。

在特定实施方案中，目标保守取代冠以优选取代的名称示于下表 6 中。如果这类取代导致生物活性变化，则可引入起更多实质性改变并筛选产物，所述改变在表 6 中命名为示范性取代，或者如下文有关氨基酸种类的详述。

15

表 6

原始残基	示范性取代	优选取代
Ala (A)	val; leu; ile	val
Arg (R)	lys; gln; asn	lys
20 Asn (N)	gln; his; lys; arg	gln
Asp (D)	glu	glu
Cys (C)	ser	ser
Gln (Q)	asn	asn
Glu (E)	asp	asp
25 Gly (G)	pro; ala	ala
His (H)	asn; gln; lys; arg	arg
Ile (I)	leu; val; met; ala; phe; norleucine	leu
Leu (L)	norleucine; ile; val; met; ala; phe	ile
30 Lys (K)	arg; gln; asn	arg

	Met (M)	leu; phe; ile	leu
	Phe (F)	leu; val; ile; ala; tyr	leu
	Pro (P)	ala	ala
	Ser (S)	thr	thr
5	Thr (T)	ser	ser
	Trp (W)	tyr; phe	tyr
	Tyr (Y)	trp; phe; thr; ser	phe
	Val (V)	ile; leu; met; phe; ala; norleucine	leu

10

对 FGF-19 多肽的生物学特性的实质性修改可通过选择性取代来完成, 所述取代的效应在维持(a)取代区多肽骨架的结构, 例如片层结构或螺旋构象, (b)该分子靶位点的电荷或疏水性, (c)侧链的大小这几方面有显著差异。天然残基根据共有的侧链特性可分为:

- 15 (1) 疏水性: 正亮氨酸, met, ala, val, leu, ile;
 (2) 中性亲水性: cys, ser, thr;
 (3) 酸性: asp, glu;
 (4) 碱性: asn, gln, his, lys, arg;
 (5) 影响链取向的残基: gly, pro;和
 20 (6) 芳香族: trp, tyr, phe.

非保守取代将限定上述某一类的成员被另一类取代。也可将取代的残基引入到保守取代位点, 或者更优选地引入其余(非-保守)位点。

可以使用本领域已知方法如寡核苷酸-介导的(定点)诱变、丙氨酸扫描和 PCR 诱变技术来产生变异。可以对克隆的 DNA 进行定点诱变[Carter 等, Nucl. Acids Res., 13: 4331 (1986); Zoller 等, Nucl. Acids Res., 10: 6487 (1987)]、盒式诱变[Wells 等, Gene, 34: 315 (1985)]、限制选择诱变[Wells 等, Philos.转. R. Soc. London SerA, 317: 415 (1986)]或其它已知技术以产生 FGF-19 变体 DNA。

扫描氨基酸分析也可沿邻接序列确定一个或多个氨基酸。优选的扫描氨基酸是相对小的中性氨基酸。此类氨基酸包括丙氨酸, 甘氨酸、丝氨酸和半胱氨酸。通常, 丙氨酸是此组中优选的扫描氨基酸, 因为它的 β -碳上无侧链并且不太可能改变变体的主链构象[Cunningham 和 Wells, Science, 244:

1081-1085 (1989)]. 优选丙氨酸的另一原因是因为它是最常用氨基酸。而且, 它常常既出现在掩蔽位置也出现在暴露位置[Creighton, *The Proteins*, (W. H. Freeman & Co., N. Y.); Chothia, *J. Mol. Biol.*, 150: 1 (1976)]. 如果丙氨酸取代不能产生足够量的变体, 则可使用 isoteric 氨基酸。

5 C. FGF-19 的修饰

本发明包括 FGF-19 的共价修饰。一种共价修饰包括使 FGF-19 多肽的靶氨基酸残基与能够与 FGF-19 的选定侧链或 N-或 C-末端残基发生反应的有机衍生化剂进行反应。具有双官能剂的衍生化是非常有用的, 例如, FGF-19 与纯化抗-FGF-19 抗体方法中所用水溶性载体基质或表面交联, 反之亦然。

10 通常使用的交联剂包括例如, 1,1-双(二偶氮乙酰基)-2-苯乙烷, 戊二醛, N-羟琥珀酰亚胺酯如与 4-叠氮基水杨酸的酯, 高双官能亚胺基酯, 包括二琥珀酰亚胺酯如 3,3'-连二硫酸双(琥珀酰亚胺基丙酸酯)、双官能马来酰亚胺酯如双-N-马来酰亚胺-1,8-辛酯和化学试剂如甲基-3-[(对-叠氮苯基)二硫]丙炔亚胺酯。

15 其它修饰包括谷氨酰胺和天冬酰胺残基脱酰胺分别成为相应的谷氨酰和门冬氨酰残基, 脯氨酸和赖氨酸羟基化, 丝氨酰或苏氨酰残基的羟基磷酸化, 赖氨酸、精氨酸和组氨酸侧链的 α -氨基甲基化[T. E. Creighton, *Proteins: Structure and Molecular Properties*, W. H. Freeman & Co., San Francisco, pp. 79-86 (1983)], 胺的 N-末端乙酰化, 和任何 C-末端羧基的酰胺化。

20 本发明还包括 FGF-19 多肽的另一类共价修饰, 所述共价修饰包括改变多肽的天然糖基化模式。"改变天然糖基化模式", 在这里的目的是为了删除天然序列 FGF-19 中一个或多个碳水化合物部分(或者通过除去发生糖基化的位点或者通过化学和/或酶手段而删除糖基化), 和/或在天然序列 FGF-19 中加入一个或多个本不存在的糖基化位点。此外, 该短语包括天然蛋白糖基化的性质变化, 包括不同碳水化合物部分在性质和比例上的变化。

25 通过改变氨基酸序列可以实现在 FGF-19 多肽中加入糖基化位点。此变更可以例如, 通过向天然序列 FGF-19(为了 O-连接的糖基化位点)中加入、或者取代一个或多个丝氨酸或苏氨酸残基而完成。任选地, 可以通过 DNA 水平的变化, 特别是通过使编码 FGF-19 多肽的 DNA 中预先选定的碱基发生突变从而产生将会翻译成期望氨基酸的密码子, 来改变 FGF-19 氨基酸序
30 列。

在 FGF-19 多肽上增加碳水化合物部分数量的另一方法是，将配糖体化学性或酶性地与多肽偶联。现有技术有此类方法的描述，例如 1987 年 9 月 11 日公布的 WO 87/05330，以及 Aplin 和 Wriston, CRC Crit. Rev. Biochem. 第 259-306 页(1981)。

- 5 用化学或酶方法，或者通过突变性取代编码作为糖基化靶点的氨基酸残基的密码子，来实现去除 FGF-19 多肽中的碳水化合物部分。化学性去糖基化技术为本领域所公知，并描述在例如下述文献中：Hakimuddin 等, Arch. Biochem. Biophys., 259: 52 (1987)和 Edge 等, Anal. Biochem., 118:131 (1981)。使用 Thotakura 等在 Meth. Enzymol., 138: 350 (1987)中描述的各种内-和外-糖苷酶，可以完成多肽中碳水化合物部分的酶裂解。

FGF-19 的另一类共价修饰，包括以美国专利 4640835、4496689、4301144、4670417、4791192 或 4179337 中所述方式，将 FGF-19 多肽连与各种非蛋白类聚合物之一，例如聚乙二醇(PEG)、聚丙烯二醇或聚氧化亚烷基。

- 15 本发明 FGF-19 也可被修饰成嵌合分子，包括将 FGF-19 与另一异源多肽或氨基酸序列融合。

在一实施方案中，嵌合分子含有 FGF-19 与标记多肽的融合体，其中的标记多肽具有抗-标记抗体可选择性结合的表位。表位标记一般位于 FGF-19 的氨基-或羧基-末端。使用抗标记多肽的抗体，可以检测到此类表位-标记型 FGF-19 的存在。另外，表位标记的提供，使得利用抗-标记抗体或另一类与表位标记结合的亲和基质来亲和纯化 FGF-19 变得非常容易。各种标记多肽及其各自抗体是本领域众所周知的。实例包括聚-组氨酸(poly-his)或聚-组氨酸-甘氨酸(poly-his-gly)标记; flu HA 标记多肽及其抗体 12CA5 [Field 等, Mol. Cell. Biol., 8: 2159-2165 (1988)]; c-myc 标记和它的 8F9、3C7、6E10、G4、
20 B7 及 9E10 抗体[Evan 等, Molecular and Cellular Biology, 5: 3610-3616 (1985)]; 和单纯疱疹病毒糖蛋白 D (gD)标记及其抗体[Paborsky 等, Protein Engineering, 3 (6): 547-553 (1990)]。

其它标记多肽包括 Flag-肽[Hopp 等, BioTechnology, 6: 12041210 (1988)]; KT3 表位肽[Martin 等, Science, 255: 192-194 (1992)]; α -微管蛋白表位肽
30 [Skinner 等, J. Biol. Chem., 266: 15163-15166 (1991)];和 T7 基因 10 蛋白肽标记[Lutz-Freyermuth 等, Proc. Natl. Acad. Sci. USA, 87: 6393-6397 (1990)]。

在另一实施方案中，嵌合分子可包括 FGF-19 与免疫球蛋白或免疫球蛋白特定区域的融合。对于二价形式的嵌合分子(也指"免疫粘附素")，融合可以是与 IgG 分子 Fc 区的融合。Ig 融合优选包括 FGF-19 多肽的可溶形式(跨膜结构域缺失或失活)取代 Ig 分子内至少一个可变区。在一特别优选的实施方案中，免疫球蛋白融合包括 IgG1 分子的铰链区、CH2 和 CH3 区，或铰链区、CH1、CH2 和 CH3 区。免疫球蛋白融合体的制备也参见 1995 年 6 月 27 日发布的美国专利 5428130。

D. 制备 FGF-19

下面的描述主要涉及通过培养转化或转染了含 FGF-19 核酸的载体的细胞来制备 FGF-19。因此，当然考虑到本领域熟知的供选择方法可用于制备 FGF-19。例如，通过利用固相合成技术的直接合成肽的方法来制备 FGF-19 序列或其一部分[参见例如，Stewart 等，Solid-Phase Peptide Synthesis, W. H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85: 2149-2154 (1963)]。使用手工或自动操作技术，进行蛋白体外合成。例如，按照制造商的指示，使用实用生物系统肽合成仪(Foster City, CA)，可以完成自动合成。可以分别化学合成 FGF-19 的各部分，然后利用化学或酶学方法制得全长 FGF-19。

1. 编码 FGF-19 的 DNA 的分离

可以从 cDNA 文库获得编码 FGF-19 的 DNA，所述 cDNA 文库是从据信具有 FGF-19 mRNA 并且以可测得水平表达 FGF-19 的组织制备的。相应地，如实施例中所描述的那样，从人组织制备的 cDNA 文库中可以很方便地得到人的 FGF-19 DNA。也可以从基因组文库或者通过已知合成方法(例如自动核酸合成)获得 FGF-19-编码基因。

使用为识别目标基因或其编码的蛋白而设计的探针(如 FGF-19 抗体或至少约 20-80 个碱基的寡核苷酸)，可以筛选文库。使用标准操作，如 Sambrook 等在 Molecular Cloning: A Laboratory Manual(New York: Cold Spring Harbor Laboratory, 1989 出版)中所描述的方法，用选择的探针进行 cDNA 或基因组文库的筛选。分离 FGF-19 编码基因的另一个可选用方法是使用 PCR 方法[Sambrook 等，出处同上；Dieffenbach 等，PCR Primer: A Laboratory Manual(Cold Spring Harbor Laboratory, 1995 出版)]。

下述实施例描述 cDNA 文库筛选技术。选作探针的寡核苷酸序列应当

具有足够的长度并且足够清晰,以便使假阳性出现机率最小。寡核苷酸优选是标记的,这样通过与正在筛选的文库中的 DNA 杂交而可被检测到。标记方法为本领域所公知,包括使用放射性标记物如 ^{32}P -标记的 ATP、生物素或酶标记。包括中等严格和高度严格的杂交条件,见 Sambrook 等出处同上。

5 可将筛选文库方法中鉴定的序列与其它已知的保藏序列或公共数据库如 GenBank 或其它私有序列数据库中的序列进行比较。使用本领域已知的和本文描述的方法,可以测定分子中限定区域或全长序列的序列同一性(无论在氨基酸水平还是在核苷酸水平)。

使用本文首次公开的推定氨基酸序列和如果必要,使用 Sambrook 等, 10 出处同上描述的常规引物延伸程序,以检测没有被逆转录为 cDNA 的 mRNA 的前体和加工中间体,可以从选定的 cDNA 或基因组文库获得具有蛋白编码序列的核酸。

2. 宿主细胞的选择和转化

用生产本文所述 FGF-19 的表达或克隆载体转染或转化宿主细胞,并在 15 常规营养培养基中培养宿主细胞,对所述培养基进行改良使之适宜于诱导增强子、选择性转化体或扩增编码所需序列的基因。无需繁琐的试验,本领域熟练技术人员就可以选择培养条件如培养基、温度、pH 等。通常,使细胞培养物产量最大化的原则、方案和实用技术见于 Mammalian Cell Biotechnology: a Practical Approach, M. Butler 编著(IRL 出版,1991)和 20 Sambrook 等出处同上。

转染真核细胞和转化原核细胞的方法为本领域熟练技术人员所已知,例如, CaCl_2 、 CaPO_4 、脂质体-介导的方法和电穿孔法。根据所用的宿主细胞,使用适合于该宿主细胞的标准技术进行转化。如 Sambrook 等出处同上所述的使用氯化钙的钙处理,或电穿孔法一般用于原核生物。如 Shaw 等在 Gene, 25 23: 315(1983)和 1989 年 6 月 29 日公布的 WO 89/05859 中描述的那样,根癌农杆菌的感染用于转化某些植物细胞。对于没有细胞壁的哺乳动物来说,可以使用 Graham 和 van derEb 在 Virology, 52: 456-457(1978)中所述的磷酸钙沉淀法。哺乳动物细胞宿主系统转化的总的情况在美国专利 4399216 中已有描述。根据 Van Solingen 等, J. Bact., 130: 946(1977)和 Hsiao 等, Proc. Natl. Acad. 30 Sci.(USA), 76: 3829(1979)所述方法,通常可以完成向酵母的转化。然而,也可使用将 DNA 引入细胞的其它方法,例如通过核的显微注射、电穿孔、细

菌原生质体与完整细胞的融合，或者聚阳离子如 polybrene、聚鸟氨酸。转化哺乳动物细胞的各种方法见 Keown 等, *Methods in Enzymology*, 185: 527-537(1990)和 Mansour 等, *Nature*, 336: 348-352(1988)。

克隆或表达本文所述载体中 DNA 的适宜的宿主细胞包括原核生物、酵母或高等真核细胞。适宜的原核生物包括但不限于革兰氏阴性或革兰氏阳性细菌等真细菌，譬如肠杆菌科(*Enterobacteriaceae*)如大肠杆菌。各种大肠杆菌株均是公众可以得到的，如大肠杆菌 K12 株 MM294(ATCC 31446); 大肠杆菌 X1776(ATCC 31537); 大肠杆菌株 W3110(ATCC 27325)和 K5 772(ATCC 53635)。其它适宜的原核宿主细胞包括肠杆菌科如埃希氏杆菌属，例如，大肠杆菌, 肠杆菌属, 欧文菌属, 克雷白菌属, 变形菌杆属, 沙门菌属(如鼠伤寒沙门菌), 沙雷菌属(如粘质沙雷菌)和志贺菌属等，以及芽孢杆菌属如枯草芽孢杆菌和地衣芽孢杆菌(例如 1989 年 4 月 12 日出版的 DD 266710 中所述地衣芽孢杆菌 41P)等，假单胞菌属的铜绿菌假单胞菌，及链霉菌。这些实例是用于说明而并非限制于此。W3110 株是一个特别优选的宿主或母宿主，因为它 5 是重组 DNA 产物发酵的常用宿主株。优选地，宿主细胞分泌少量蛋白水解酶。例如，修饰 W3110 株以使编码宿主内源性蛋白的基因发生遗传突变，此类宿主的实例包括大肠杆菌 W3110 株 1 A2，该株具有完整基因型 *tonA* ;大肠杆菌 W3110 株 9E4，该株具有完整基因型 *tonA ptr3* ; 大肠杆菌 W3110 株 27C7(ATCC 55244)，该株具有完整基因型 *tonA ptr3 phoA* 10 *E15(argF-lac)169 degP ompT kanr*; 大肠杆菌 W3110 株 37D6，该株具有完整基因型 *tonA ptr3 phoA E15(argF-lac)169 degP ompT rbs7 ilvG kant*; 大肠杆菌 W3110 株 40B4，它是 37D6 发生非-卡那霉素耐药 *degP* 缺失突变的株;和具有 1990 年 8 月 7 日发布的美国专利 4946783 中披露的突变型周质蛋白酶的大肠杆菌株。或者，克隆的体外方法，例如 PCR 或其它核酸聚合酶反应， 15 是适宜的。

除了原核生物，真核微生物如丝状真菌或酵母也是适于使编码 FGF-19 的载体克隆或表达的宿主。酿酒酵母，或常用的面包酵母，在低等真核宿主微生物中最为常用。其它真核微生物包括：粟酒裂殖酵母 (*Schizosaccharomyces pombe*)(Beach 和 Nurse, *Nature*, 290: 140 [1981]; 1985 年 30 5 月 2 日公布的 EP 139383); 克鲁维酵母属(*Kluyveromyces*)宿主(美国专利 4943529; Fleer 等, *Bio/Technology*, 9: 968-975(1991))，例如乳酸克鲁维酵母

(*K.lactis*)(MW98-8C, CBS683, CBS4574; Louvencourt 等, J. Bacteriol., 154(2): 737-742 [1983])、脆壁克鲁维酵母(*K.fragilis*)(ATCC 12424)、保加利亚克鲁维酵母(*K.bulgaricus*)(ATCC 16045)、威克曼氏克鲁维酵母(*K. wickeramii*)(ATCC 24178)、*K.waltii*(ATCC 56500)、果蝇克鲁维酵母(*K. drosophilorum*)(ATCC 36906; Van den Berg 等, Bio/Technology, 8:135(1990))、耐热克鲁维酵母(*K. thermotolerans*)和克鲁维氏酵母属(*K. marxianus*)等; *yarrowia*(EP 402226); 巴斯德毕赤酵母(*pichia pastoris*)(EP 183070; Sreekrishna 等, J. Basic Microbiol., 28: 265-278[1988]); 念珠菌属; *Trichoderma reesia*(EP 244234); 粗糙链孢霉(Case 等, Proc. Natl. Acad. Sci. USA, 76: 5259-5263 [1979]); 许旺氏酵母属(*schwanniomyces*)如西方许旺氏酵母(*schwanniomyces occidentalis*)(1990年10月31日公布的 EP 394538)等; 和丝状真菌, 例如链孢霉属、青霉属、*Tolypocladium*(1991年1月10日公布的 WO 91/00357)以及曲霉属宿主如构巢曲霉(Ballance 等, Biochem. Biophys. Res. Commun., 112: 284289 [1983]; Tilburn 等, Gene, 26: 205-221 [1983]; Yelton 等, Proc. Natl. Acad. Sci. USA, 81: 1470-1474 [1984])和黑曲霉等(Kelly 和 Hynes, EMBO J., 4: 475-479 [1985])。在本文中嗜甲基(Methylotropic)酵母是适宜的, 包括但不限于选自下述属的能够在甲醇中生长的酵母: 汉逊氏酵母属(*Hansenula*), 念珠菌属(*Candida*), 克勒克酵母属(*Kloeckera*), 毕赤酵母属(*Pichia*), 酵母属, 球拟酵母属和红酵母属。此类酵母的例证性特定酵母属的清单见 C. Anthony, The Biochemistry of Methylotrophs, 269(1982)。

用于表达糖基化 FGF-19 的适合宿主细胞来自多细胞生物。无脊椎动物细胞的实例包括昆虫细胞(例如果蝇属 S2 和草地夜蛾 Sf9)和植物细胞。有用的哺乳动物宿主细胞系包括中国仓鼠卵巢(CHO)细胞和 COS 细胞。更特异的实例包括转化了 SV40(COS-7, ATCC CRL 1651)的猴肾 CV1 细胞系; 人胚肾细胞系(293 细胞或亚克隆培养成在悬浮培养液中生长的 293 细胞, Graham 等, J. Gen Virol. 36: 59(1977)); 中国仓鼠卵巢细胞/-DHFR(CHO, Urlaub 和 Chasin, Proc. Natl. Acad. Sci. USA, 77:4216(1980)); 鼠赛尔托利细胞(TM4, Mather, Biol. Reprod., 23: 243-251(1980)); 人肺细胞(W138, ATCC CCL 75); 人肝脏细胞(Hep G2, HB 8065); 和鼠乳腺癌细胞(MMT 060562, ATCC CCL51)。选择合适宿主细胞是本领域常识。

3. 复制型载体的选择和应用

可将编码 FGF-19 的核酸(例如 cDNA 或基因组 DNA)插入复制型载体中以进行克隆(DNA 扩增)或表达。各种载体是公众可以获得的。载体可以是例如质粒、粘粒、病毒颗粒或噬菌体的形式。使用各种方法可将合适的核酸序列插入载体。通常,利用本领域已知技术将 DNA 插入到合适的限制性核酸内切酶位点。载体组成部分一般包括但不限于:一个或多个信号序列、复制

5 起点、一个或多个标记基因、增强子元件、启动子和转录终止序列。使用本领域熟练技术人员熟知的标准连接技术,构建包括这些组成部分中一个或多个元件的适宜载体。

FGF-19 不仅可直接重组制备,而且还可制备成与异源多肽的融合多肽,所述异源多肽为信号序列或在成熟蛋白或多肽的 N 末端上具有特异性裂解

10 位点的其它多肽。通常,信号序列是载体的一个组分或者是被插入载体的 FGF-19-编码 DNA 的一部分。信号序列可以是选自例如碱性磷酸酶、青霉素酶、lpp 或热稳定肠毒素 II 前导区的原核信号序列。对于酵母分泌,可以是例如酵母转化酶前导区、 α 因子前导区(包括糖酵母属和克鲁维酵母属的 α

15 因子前导区,后者在美国专利 5010182 中有述),或酸性磷酸酶前导区、白色念珠菌葡萄糖淀粉酶前导区(1990 年 4 月 4 日公布的 EP 362179),或者 1990 年 11 月 15 日公布的 WO90/13646 中所述信号序列。在哺乳动物细胞表达时,可使用哺乳动物信号序列以直接分泌蛋白,如得自相同物种或相关物种分泌型多肽的信号序列以及病毒分泌前导区。

20 表达载体和克隆载体均含有使载体在一个或多个选定宿主细胞中复制的核酸序列。在多种细菌、酵母和病毒中,这样的序列众所周知。来自质粒 pBR322 的复制起点适宜于大多数革兰氏阴性细菌,2 μ 质粒复制起始点适于酵母,各种病毒复制起始点(SV40, 多瘤病毒,腺病毒, VSV 或 BPV)适宜于哺乳动物细胞中的克隆载体。

25 表达载体和克隆载体通常包含筛选基因,也称为可筛选标记。典型的筛选基因编码蛋白,该蛋白(a)提供对抗生素或其它毒素,如氨基青霉素、新霉素、氨基喋呤或四环素等的抗性,(b)弥补营养缺陷,或(c)提供从复合培养基中不能得到的关键营养物质,例如编码芽孢杆菌属 D-丙氨酸消旋酶的基因。

30 适于哺乳动物细胞的筛选标记实例是那些使能接纳 FGF-19 核酸的细胞得到鉴定的标记,例如 DHFR 或胸苷激酶。当使用野生型 DHFR 时,适当

的宿主细胞是按照 Urlaub 等在 *Proc. Natl. Acad. Sci. USA* 77:4216(1980)中描述的方法制备和繁殖的 DHFR 活性缺陷型中国仓鼠卵巢(CHO)细胞系。适用于酵母的合适筛选基因是存在于酵母质粒 YRp7 中的 *trp1* 基因[Stinchcomb 等, *Nature*, 282: 39(1979); Kingsman 等, *Gene*, 7:141(1979); Tschemper 等, *Gene*, 10:157(1980)]。5 *trp1* 基因为不能在色氨酸中生长的酵母突变株(例如 ATCC 44076 或 PEP4-1)提供了筛选标记[Jones, *Genetics*, 85: 12(1977)]。

表达和克隆载体通常含有可操作地连接于 FGF-19-编码核酸序列的启动子, 以指导 mRNA 合成。被各种潜在宿主细胞识别的启动子是众所周知的。适用于原核宿主的启动子, 包括 β -内酰胺酶和乳糖启动子系统[Chang 10 等, *Nature*, 275: 615(1978); Goeddel 等, *Nature*, 281: 544(1979)], 碱性磷酸酶, 色氨酸(*trp*)启动子系统[Goeddel, *Nucleic acids Res.*, 8: 4057(1980); EP 36776], 和杂化启动子如 *tac* 启动子[deBoer 等, *Proc. Natl. Acad. Sci. USA*, 80: 21-25(1983)]。适用于细菌系统的启动子, 也含有可操作地连接于编码 FGF-19 的 DNA 的 Shine-Dalgarno(S. D.)序列。

15 适用于酵母宿主的启动序列的实例包括 3-磷酸甘油酸激酶[Hitzeman 等, *J. Biol. Chem.*, 255: 2073(1980)]或其它糖酵解酶 [Hess 等, *J. Adv. Enzyme Reg.*, 7: 149(1968); Holl 和 *Biochemistry*, 17: 4900(1978)]的启动子, 所述其它糖酵解酶如烯醇化酶, 甘油醛-3-磷酸脱氢酶, 己糖激酶, 丙酮酸脱羧酶, 磷酸果糖激酶, 葡萄糖-6 磷酸异构酶, 3-磷酸甘油变位酶, 丙酮酸激酶, 磷酸 20 丙糖异构酶, 磷酸葡萄糖异构酶和葡萄糖激酶。

其它的酵母启动子, 即那些具有由生长条件控制转录的优点的诱导型启动子, 是下述基因的启动子区, 即乙醇脱氢酶 2、异细胞色素 C、酸性磷酸酶、与氮代谢相关的降解酶、金属硫蛋白、甘油醛-3-磷酸脱氢酶和负责麦芽糖和半乳糖利用的酶。在 EP 73657 中进一步描述了用于酵母表达系统的适当载体和启动子。 25

在哺乳动物宿主细胞中, 由载体转录 FGF-19 可受下述启动子调控, 所述启动子来自病毒基因组如多瘤病毒、鸡痘病毒(1989 年 7 月 5 日公布的 UK 2211504)、腺病毒(如腺病毒 2)、牛乳头瘤病毒、禽肉瘤病毒、巨细胞病毒、逆转录病毒、乙型肝炎病毒和猴病毒 40(SV40)的启动子, 或者来自异源哺乳 30 动物的启动子, 如肌动蛋白启动子或免疫球蛋白启动子等, 和来自热休克启动子, 前提是这些启动子与宿主细胞系统相容。

在载体中插入增强子序列，可以增加编码 FGF-19 的 DNA 在高等真核生物中的转录。增强子是作用于启动子以增加转录的 DNA 的顺式作用元件，一般约 10~300 bp。目前已经知道了很多哺乳动物基因(珠蛋白，弹性蛋白酶、白蛋白、 α 胎蛋白和胰岛素)的增强子序列。然而，人们通常使用真核
5 细胞病毒的增强子。实例包括在其复制起始点晚期侧的 SV40 增强子(bp 100-270)，巨细胞病毒早期启动子增强子，在其复制起始点晚期侧的多形瘤增强子，和腺病毒增强子。所述增强子可以剪接入 FGF-19 编码序列的 5' 或 3' 端，但优选位于启动子的 5' 端。

用于真核宿主细胞(酵母、真菌、昆虫、植物、动物、人或来自其它多
10 细胞生物的有核细胞)的表达载体，还包括对转录终止和稳定 mRNA 结构所必须的序列。这些序列通常来自真核或病毒 DNA 或 cDNA 的 5'(偶尔为 3')非翻译区。这些区域包含转录为编码 FGF-19 的 mRNA 的非翻译区中聚腺苷酸化片段的核苷酸片段。

在重组脊椎动物细胞培养中适应 FGF-19 合成的其它适宜方法、载体和
15 宿主细胞，见 Gething 等, *Nature*, 293: 620-625(1981); Mantei 等, *Nature*, 281: 40-46(1979); EP 117060;和 EP 117058 中的描述。

4. 检测基因扩增/表达

使用基于本文提供的序列的合适标记的探针，利用常规技术如 Southern
印迹、测定 mRNA 转录量的 Northern 印迹[Thomas, *Proc. Natl. Acad. Sci. USA*,
20 77: 5201-5205(1980)]、点印迹(DNA 分析)或原位杂交，可以直接在样品中测定基因扩增和/或表达。或者，使用能够识别特异双链体的抗体，所述双链体包括 DNA 双链体、RNA 双链体和 DNA-RNA 杂化双链体或者 DNA-蛋白双链体。依次标记抗体，对双链体与表面哪个部位结合进行分析，从而基于双链体在表面的形成，探测到与双链体结合的抗体。

25 或者，为直接定量测定表达的基因产物，用免疫学方法测定基因表达，所述方法如细胞或组织切片的免疫组化染色和细胞培养物或体液分析。适用于免疫组化染色和/或样品液分析的抗体，可以是单克隆抗体或多克隆抗体，并且可以在任何哺乳动物中制备抗体。可以方便地制备抗天然序列 FGF-19 多肽的抗体，或抗基于本文提供的 DNA 序列的合成肽的抗体，或者抗与
30 FGF-19DNA 融合并编码特异抗体表位的外源序列的抗体。

5. 多肽纯化

可从培养基或宿主细胞裂解物中回收 FGF-19。如果 FGF-19 与膜结合，那么使用适宜去垢剂溶液(例如 Triton-X 100)或通过酶裂解使之自细胞膜释放。使用各种物理或化学手段，如冻融循环、超声、机械破裂或细胞溶解剂，使表达 FGF-19 所用的细胞破裂。

- 5 可以期望从重组细胞蛋白或多肽中纯化 FGF-19。下述程序是例证性适宜的纯化步骤：在离子-交换柱上分级分离；乙醇沉淀；反相 HPLC；在硅石或阳离子-交换树脂如 DEAE 上层析；聚焦层析；SDS-PAGE；硫酸铵沉淀；使用例如交联葡聚糖(Sephadex)G-75 进行凝胶过滤；过蛋白 A 琼脂糖柱以除去污染物如 IgG；和结合表位-标记形式 FGF-19 的金属螯合剂柱。可以使用蛋白
- 10 纯化的各种方法，这些方法是本领域熟知的，并在例如 Deutscher, Methods in Enzymology, 182(1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York(1982)中作了描述。纯化步骤的选择，取决于例如，生产方法的性质和所产生的特定 FGF-19。

E.FGF-19 的应用

- 15 编码 FGF-19 的核苷酸序列(或其互补链)在分子生物学领域具有多种用途，包括用作杂交探针，在染色体和基因图谱和在制备反义 RNA 和 DNA 中的用途。FGF19 核酸对于通过本文所述重组技术来制备 FGF-19 多肽，也是非常有用的。

- 全长天然序列 FGF-19 基因(SEQ ID NO: 1)或其部分，可用作 cDNA 文库的杂交探针以分离出全长 FGF-19 cDNA，或分离出与图 1(SEQ ID NO: 1)中公开的 FGF-19 序列具有期望序列同一性的其它 cDNA(例如，那些编码天然 FGF-19 变体或来自其它种属的 FGF-19 的 cDNA)。任选地，探针长度是约 20~约 50 个碱基。杂交探针可源自于 SEQ ID NO: 1 核苷酸序列的至少部分
- 20 新区，其中所述新区是那些不需要十分繁琐实验就可确定的区域，或从包括
- 25 启动子、增强子元件和内含子的天然序列 FGF19 基因组序列中确定的区域。举例来说，筛选方法包括：使用已知 DNA 序列合成约 40 个碱基的选定探针，来分离 FGF-19 基因的编码区。杂交探针可标记上各种标志，包括放射性核苷酸如 ^{32}P 或 ^{35}S 或酶标记，所述酶标记如通过亲和素/生物素偶联系统将碱性磷酸酶与探针耦合。具有与本发明 FGF-19 基因互补序列的标记探针，可
- 30 用于筛选 cDNA 文库、基因组 DNA 或 mRNA，以测定探针与哪一文库发生杂交。下面的实施例将进一步详细描述杂交技术。

使用本文披露的方法，本申请公开的 EST 序列可类似地用作探针。

FGF-19 核酸的其它有用片段，包括含有能与靶 FGF-19mRNA(正义)或 FGF-19DNA(反义)序列结合的单链核酸序列(RNA 或 DNA)的反义或正义寡核苷酸。根据本发明，反义或正义寡核苷酸包括 FGF-19DNA 编码区的片段。

- 5 该片段一般含有至少约 14 个核苷酸，优选地由约 14~30 个核苷酸。基于编码一给定蛋白的 cDNA 序列得到反义或正义寡核苷酸的能力，在例如 Stein 和 Cohen(Cancer Res. 48: 2659, 1988)及 van der Krol 等(BioTechniques 6: 958, 1988)的文章中有描述。

- 反义或正义寡核苷酸与靶核酸序列的结合，导致通过一种或多种方式阻
10 断靶序列转录或翻译的双螺旋形成，所述阻断方式包括增加双螺旋降解、转录或翻译的成熟前终止或者其它方式。因此，反义寡核苷酸可用于阻断 FGF-19 蛋白的表达。反义或正义寡核苷酸进一步包括具有修饰过的糖磷酸二酯主链的寡核苷酸(或其它糖链接，如 WO 91/06629 中描述的那些)，其中
15 此类糖链接是耐内源性核酸酶的。具有抗性糖链接的寡核苷酸在体内是稳定的(即，能够抗酶降解)，同时保留序列与靶核苷酸序列结合特异性。

- 正义或反义寡核苷酸的其它实例包括那些与有机分子共价键合的寡核苷酸，如 WO 90/10048 中描述的那些寡核苷酸，和其它增加寡核苷酸与靶核酸序列亲和力的分子，如聚-(L-赖氨酸)。另外，嵌入剂(如椭圆玫瑰树碱)和烷基化剂或金属络合物可挂接在正义或反义寡核苷酸上，以修饰反义或正义
20 寡核苷酸对靶核苷酸序列的结合特异性。

- 使用任何基因转移方法可将反义或正义寡核苷酸引入到含有靶核酸序列的细胞中，所述方法包括例如，CaP04-介导的 DNA 转染、电穿孔、或使用基因转移载体如 Epstein-Barr 病毒。在一优选的方法中，将反义或正义寡核苷酸插入到一适宜的逆转录病毒载体中，在体内或体外使含有靶核酸序列
25 的细胞与重组逆转录病毒载体接触。适宜的逆转录病毒载体包括但不限于，得自小鼠逆转录病毒 M-MuLV、N2(得自 M-MuLV 的逆转录病毒)，或称为 DCT5A、DCT5B 和 DCT5C(见 WO 90/13641)的双倍复制病毒。

- 也可以通过使正义或反义寡核苷酸与配体结合分子形成共轭物的形式，而将正义或反义寡核苷酸引入到含有靶核苷酸序列的细胞中，见 WO
30 91/04753 的描述。适宜的配体结合分子包括但不限于：细胞表面受体、生长因子、其它细胞因子或能与细胞表面受体结合的其它配体。优选地，配体结

合分子的共轭基本上不干扰配体结合分子与其相应分子或受体结合的能力，或阻断正义或反义寡核苷酸或其共轭型进入细胞。

或者，按照 WO 90/10448 描述的方法，通过形成寡核苷酸-脂复合物，而将正义或反义寡核苷酸引入到含有靶核酸序列的细胞中。优选地，正义或反义寡核苷酸-脂复合物是使用内源性脂酶而与细胞分离的复合物。

也可在 PCR 技术中使用探针，以产生用于鉴定最接近 FGF-19 编码序列的一组序列。

编码 FGF-19 的核苷酸序列还可用于构建在描绘编码 FGF-19 的基因图谱和对患有遗传病个体进行基因分析中使用的杂交探针。使用已知技术，如原位杂交、抗已知染色体标记物的连锁分析和与文库的杂交筛选技术，可将本文提供的核苷酸序列整入到染色体和染色体的特定区域。

FGF-19 的编码序列编码能与另一蛋白结合的蛋白时(例如，当 FGF-19 是受体时)，FGF-19 可用在鉴定参与结合反应的其它蛋白或分子的分析中。使用此类方法，可以鉴定受体/配体结合反应的抑制剂。还可将参与此结合反应的蛋白用于筛选肽或结合反应的小分子抑制剂或激动剂。而且，受体 FGF-19 可用于分离相关配体。筛选分析可用于设计寻找模拟天然 FGF-19 生物活性或 FGF19 受体的主要化合物。此类筛选分析包括使用化学用数据库进行高-通过量筛选分析，使之特别适宜于鉴定小分子候选药物。小分子包括合成的有机或无机化合物。可使用各种形式的分析法，包括蛋白-蛋白结合分析，生化筛选分析，免疫测定和以细胞为基础的分析，这些分析法均是本领域公知的。

编码 FGF-19 或其修饰形式的核酸，也可用于生产转基因动物或"剔除"动物，所得动物反过来对于开发和筛选适合的治疗剂是非常有用的。转基因动物(例如小鼠或大鼠)是含有转基因细胞的动物，所述转基因被引入到动物或孕期祖代动物中，例如胚胎期。转基因是整合入细胞基因组的 DNA，由所述细胞发展为转基因动物。在一实施方案中，使用已经建立的技术和用于生产含有表达编码 FGF-19DNA 的细胞的转基因动物的基因组序列，编码 FGF-19 的 cDNA 可用于克隆编码 FGF-19 的基因组 DNA。生产转基因动物的方法，特别是诸如小鼠或大鼠这样的动物，已经成为本领域常规技术，并在例如美国专利 4736866 和 4870009 中有描述。通常，带有组织-特异增强子的 FGF-19 转基因被整合入特定细胞中。包含在胚胎期被整合入动物干细

胞一个拷贝编码 FGF-19 转基因的转基因动物，可用于检测增加编码 FGF-19 DNA 表达的效果。此类动物可用作检测认为有可能对与过度表达有关的病理病症具有保护作用的药剂的试验动物。根据本发明的一个方面，与未经治疗的携带转基因的动物相比，用药剂治疗的动物，其病理病症的发生率降低，这表明药剂对于病理病症具有潜在的治疗干预作用。

或者，非-人源性 FGF-19 用于构建 FGF19 的"剔除"动物，该动物具有缺陷的或改变的 FGF-19 基因编码，这是将内源性 FGF-19 基因编码和改变的编码 FGF-19 基因组 DNA 之间的同源重组物引入动物胚胎干细胞的结果。例如，按照已经建立的技术，可将编码 FGF-19 的 cDNA 用于克隆编码 FGF-19 的基因组 DNA。编码 FGF-19 基因组 DNA 的一部分，可被删除或用其它基因替换，如被用于监控整合的选择性标志的基因编码替换。通常，将数千碱基对未被改变的侧面 DNA(5'或 3'末端)引入到载体中[参见例如 Thomas 和 Capecchi, Cell, 51: 503(1987)对同源重组载体的描述]。将载体引入胚胎干细胞系(例如通过电穿孔)，并选出引入了已经与内源性 DNA 同源重组的 DNA 的细胞[参见例如，Li 等, Cell, 69: 915(1992)]。接着，将选出的细胞注入动物(例如小鼠或大鼠)的胚囊中，以形成 to form 聚集嵌合体[参见例如, Bradley, in Teratocarcinomas 和 Embryonic Stem Cells : A Practical Approach, E. J. Robertson 编著(IRL, Oxford, 1987), 113-152 页]。将嵌合胚胎植入适宜的假妊娠雌性养动物体内，胚胎生长到足月即成为"剔除"动物。使用标准技术，鉴定出在其干细胞中携带有同源重组 DNA 的子代动物，用这些子代动物繁殖出在所有动物细胞中均含有同源重组 DNA 的动物。剔除动物的特征在于，例如，抵御某些病理病症的能力和由于动物缺乏 FGF-19 多肽而发生病理病症。

也可将编码 FGF-19 多肽的核酸用于基因治疗。在基因治疗应用中，将基因引入细胞以便达到在体内合成治疗有效的基因产物，例如置换缺陷基因。"基因治疗"既包括经单次治疗就可获得永久效果的常规基因治疗，也包括基因治疗剂的给药，后者涉及一次或重复多次施用治疗有效量的 DNA 或 mRNA。反义 RNA 和 DNA 可用作阻断体内特定基因表达的治疗剂。业已证明，短链反义寡核苷酸可整合入细胞中，在所述细胞中，短链反义寡核苷酸起抑制剂的作用，而不论由于细胞膜的限制性摄取所导致其低细胞内浓度如何。(Zamecnik 等, Proc. Natl. Acad. Sci. USA 83: 4143-4146 [1986])。可修饰寡

核苷酸以增强其摄取, 例如永不带电荷的基团取代带阴性电荷的磷酸二酯。

有各种可利用的技术来把核酸引入到活细胞中。技术依是否将核酸转移进体外培养细胞或意欲的宿主体内细胞中而异。将核酸转移入哺乳动物体外细胞的适宜技术包括使用脂质体、电穿孔、微注射、细胞融合、DEAE-葡聚糖、磷酸钙沉淀法等。目前流行的体内基因转移技术包括用病毒(通常为逆转录病毒)载体转染和用病毒包衣蛋白-脂质体介导的转染(Dzau 等, Trends in Biotechnology 11, 205-210 [1993])。有些情况下, 希望提供带有针对靶细胞的物质的核酸源, 所述物质如对细胞表面的膜蛋白具有特异性的抗体, 或靶细胞上受体的配体等。使用脂质体时, 将与参与胞吞的细胞表面膜蛋白结合的蛋白用作靶向和/或促进摄取的物质, 例如对特定型细胞具有向性的病毒壳体蛋白或其片段、在循环中发生内化的蛋白的抗体、靶向于细胞内定位(localization)并增加细胞内半衰期的蛋白。有关受体-介导的胞吞的技术在例如 Wu 等, J. Biol. Chem. 262, 4429-4432(1987);和 Wagner 等, Proc. Natl. Acad. Sci. USA 87, 3410-3414(1990)中有描述。关于基因标志和基因治疗方案的综述, 请参见 Anderson 等, Science 256, 808-813(1992)。

本文所述 FGF-19 多肽也可用作电泳中蛋白的分子量标志。

本文所述编码 FGF-19 多肽或其片段的核酸分子在染色体鉴定中非常有用。在此方面, 由于基于实际序列数据只有相对少数几个可利用的染色体标志剂, 因而一直存在着对于鉴定新染色体标志的需求。

20 本发明 FGF-19 核酸分子可用作染色体标志。

本发明 FGF-19 多肽和核酸分子也可用于组织分型, 其中本发明 FGF-19 多肽在不同组织间有差别地表达。FGF-19 核酸分子还可用于制备 PCR、Northern 分析、Southern 分析和 Western 分析用的探针。

25 本文所述 FGF-19 多肽和其调节剂也可用作治疗剂。按照已知制备有益药物组合物的方法, 可制得本发明 FGF-19 多肽和其调节剂, 所制得的 FGF-19 产品与可药用载体成呈混合物结合。为了治疗制剂贮存方便, 将具有所需纯度的活性组分与任选的可药用载体、赋形剂或稳定剂混合(Remington's Pharmaceutical Sciences, 第 16 版, Osol, A. 编著(1980)), 制备成冷冻干燥制剂或水溶液形式。可接受的载体、赋形剂或稳定, 在所用剂量和浓度范围对接受者没有毒性, 其包括缓冲剂如磷酸盐、柠檬酸盐和其它有机酸; 抗氧化剂包括抗坏血酸; 低分子量(小于约 10 个残基)多肽; 蛋白, 如血清白蛋白、明

胶或免疫球蛋白;亲水性聚合物如聚乙烯吡咯烷酮, 氨基酸如甘氨酸、谷氨酰胺、天冬酰胺、精氨酸或赖氨酸;单糖、二糖和其它碳水化合物如葡萄糖甘露糖或糊精;螯合剂如EDTA;糖醇如甘露醇或山梨醇;成盐离子如钠;和/或非离子表面活性剂如TWEENTM、PLURONICSTM或PEG。

- 5 用于体内给药的制剂必须是无菌的。这可以在冷冻干燥和重新配制之前或之后通过除菌滤膜过滤而轻易实现。

本文的治疗组合物一般盛放在具有无菌存取口的容器中,例如该容器是静脉注射袋或带有能通过皮下注射针刺穿的塞子的小瓶。

- 10 给药途径与已知方法一致,例如注射或静脉输注、腹膜内、颅内、肌内、眼内、动脉内或损伤内给药,局部用药或使用持续释放系统给药。

- 本发明药物组合物的剂量和所需药物浓度将依预想的特定用途而异。决定适宜剂量或给药途径是普通内科医师完全能够很好胜任的事情。动物试验为决定治疗人类患者所用剂量提供可靠指导。可按照 Yacobi 等编著的 Toxicokinetics and New Drug Development(Pergamon Press, New York 1989) 一书中, Mordenti, J.和 Chappell 在 W."The use of interspecies scaling in toxicokinetics"第 42-96 页所给出的方法换算种属间有效剂量的比例。

- 15 体内施用 FGF-19 多肽或其激动剂或拮抗剂时,取决于给药途径,每日标准剂量范围为约 10 ng/kg 至高达 100 mg/kg 哺乳动物体重或更多,优选地约 1 μ g/kg/日 ~ 10mg/kg/日。文献提供了有关特定剂量的指导和释放方法,参见例如美国专利 4657760; 5206344;或 5225212。对于不同治疗化合物和不同病症而言,可以预期不同制剂能够奏效,例如,那些施用于器官或组织的制剂使得以不同于施用于其它器官或组织的方式进行释放成为必需。

- 20 当期望 FGF-19 多肽或调节剂在具有适宜于治疗需要施用 FGF-19 多肽或调节剂的任何疾病或病症的释放特性的制剂中持续释放给药时,考虑使用微囊包封。持续释放用的重组蛋白微囊包封已被成功地用于人生长激素 (rhGH)、干扰素-(rhIFN-)、白介素-2 和 MN rgpl20 的持续释放。Johnson 等, Nat. Med., 2: 795-799(1996); Yasuda, Biomed. Ther., 27: 1221-1223(1993); Hora 等, Bio/Technology 8: 755-758(1990); 在 Powell 和 Newman 编著的 Vaccine Design: The Subunit and Adjuvant Approach(Plenum Press: New York, 1995)一书中第 25 439-462 页, Cleland, "Design and Production of Single Immunization Vaccines Using Polylactide Polyglycolide Microsphere Systems"; WO 97/03692, WO 30

96/40072, WO 96/07399;和美国专利 5654010。

由于聚 lactic-coglycolic acid(PLGA)聚合物具有生物相容性很宽范围的生物降解特性, 所以使用该聚合物来制备蛋白的持续释放制剂。PLGA 的降解产物, 即乳酸和羟基乙酸, 在体内被快速清除。而且, 可以根据该聚合物的分子量 and 构成来调节其降解特性, 可以是数月 to 数年不等。在 M. Chasin 和 R. Langer(编著), Biodegradable Polymers as Drug Delivery Systems(Marcel Dekker: New York, 1990)一书中第 1-41 页, Lewis, "Controlled release of active agents from lactide/glycolide polymer"。

本文提供的含有 FGF-19 治疗剂和组合物, 可具有多种用途。这些用途包括治疗肥胖个体或与肥胖有关的病症。一方面, 将有效量 FGF-19 施用于需要 FGF-19 的个体, 从而治疗病症。优选地, 病症是需要治疗的至少一种下述病症: 代谢增加、体重下降、机体脂肪减少、甘油三酯减少、游离脂肪酸减少、葡萄糖自脂肪细胞释放增加和/或 leptin 自脂肪细胞释放增加。使用标准方法, 可以测定其中的每一项参数, 例如, 通过测定耗氧量来测定代谢速率, 使用秤测量体重, 以及通过测量尺寸来测定脂肪情况。

另外, 使用标准方法, 可以测定甘油三酯、游离脂肪酸、葡萄糖和 leptin 的含量。在下面的特定实施例中, 例证性给出了每一项参数的测定情况。

FGF-19 和含 FGF-19 组合物, 优选在体内使用。但是, 如下面所述的那样, 也可在体外给药, 例如在下述筛选 FGF-19 调节剂的方法中使用。此外, 应当懂得, 也可通过使用动物模型和采自患者的样品来鉴定 FGF-19 的调节剂。

发明包括筛选化合物的方法, 以鉴定那些模拟或增强 FGF-19 多肽(激动剂)或者阻止或抑制 FGF-19 多肽(拮抗剂)效果的化合物。在本文中, 把激动剂和拮抗剂称作调节剂。拮抗剂候选药物筛选试验, 目的是鉴定与由本文确定的基因编码的 FGF-19 多肽结合或复合的化合物, 或者相反, 鉴定那些干扰基因编码的多肽与其它细胞蛋白相互作用的化合物。此类筛选试验包括使用高通量化学数据库筛选分析, 使之特别适于鉴定小分子候选药物。

试验可按照任何方式进行, 包括蛋白-蛋白结合分析、生化筛选分析、免疫测定和基于细胞的分析, 这些分析方法在本领域已被很好描述。

拮抗剂试验共同的方面是, 它们使候选药物与由本文确定的核酸编码的 FGF-19 多肽接触足够时间以使得这两种组分相互作用。

在结合试验中，相互作用是结合，形成的复合物可被分离或在反应混合物中被检测。一特定实施方案中，通过共价或非-共价地附着，而把由本文确定的基因编码的 FGF-19 多肽或候选药物固定在固相上，例如，固定在微量滴定盘上。非-共价附着，一般通过用 FGF-19 多肽溶液包封固体表面并进行干燥而实现。或者，固定抗体，例如，固定的对 FGF-19 多肽具有特异性的单克隆抗体，可用于使 FGF-19 多肽固定到固体表面。试验这样进行：向固定组分中加入非固定组分，所述非固定组分可以标记上可检测的标记物，譬如，包衣表面含有使 FGF-19 多肽固定的组分。当反应完成时，除去未反应的组分(例如，通过洗涤方式)，并检测固定在固体表面的复合物。当原来非固定组分带有可检测标记物时，检测到固定在表面的标记物，就表明复合物形成。如果原来非-固定组分没有携带标记物，则可以检测到复合，例如，通过使用与固定复合物特异结合的抗体。

如果候选化合物与本文确定的基因编码的特定 FGF-19 多肽相互作用但是又不与之结合，那么，使用已知的检测蛋白-蛋白相互作用的方法来分析与那个多肽的相互作用。此类分析方法包括惯常的方法，例如，交联、共-免疫沉淀和通过梯度或层析柱的共-纯化。此外，按照 Fields 与其同事(Fields 和 Song, *Nature*(340: 245-246(1989)); Chien 等, *Proc. Natl. Acad. Sci. USA*, 88: 9578-9582(1991))和 Chevray 和 Nathans, *Proc. Natl. Acad. Sci. USA*, 89: 5789-5793(1991)公开的方法，使用酵母-基遗传系统可以检测蛋白-蛋白相互作用。许多转录激活剂，如酵母 GAL4，它由两个完全分散的分子结构域构成，其一作为 DNA-结合区，另一个作为激活转录的区。前述出版物中描述的酵母表达系统(一般指"二因子杂合系统")利用此特性，并使用两个杂交蛋白，其一的靶蛋白与 GAL4 的 DNA-结合区融合，另一个的候选激活蛋白与激活区融合。在 GAL4-激活的增强子控制下 GAL1-lacZ 报告基因的表达，依赖于通过蛋白-蛋白相互作用重建的 GAL4 活性。用对-半乳糖苷酶的显色底物检测含有相互作用多肽的菌丛。从 Clontech 可以商购利用二因子杂合技术鉴定两个特异蛋白间蛋白-蛋白相互作用的全套试剂盒(MATCHMAKER)。该系统也可被扩展到描绘参与特异蛋白相互作用的蛋白结构域，以及对于这些相互作用起关键作用的精确的氨基酸残基。

干预本文鉴定的基因编码的 FGF-19 多肽与其它细胞内或细胞外组分相互作用的化合物可如下测定：通常在控制条件下，制备含有产物基因和细胞

内或细胞外组分产物的反应混合物，使两种产物相互作用和结合一定时间。为了测试候选化合物抑制结合的能力，在含有和不含有受试化合物的条件下进行反应。另外，可把安慰剂加入到第三反应混合物中，作为阳性对照。按照上述方法，监测混合物中受试化合物与细胞内或细胞外组分之间的结合(复合物形成)。在对照反应中有复合物形成，而在含有受试化合物的反应混合物中没有复合物形成，表明受试复合物干预受试化合物与其配伍化合物之间的相互作用。

为分析拮抗剂，可将 FGF-19 多肽与被筛选化合物一起加入到细胞中，一定活性和能力的化合物抑制共存的 FGF-19 多肽的有关活性，则表明化合物是 FGF-19 多肽的拮抗剂。或者，在适当条件下使 FGF-19 多肽和潜在拮抗剂与结合在膜上的 FGF-19 多肽受体或重组受体结合，通过竞争性抑制实验来检测拮抗剂。可标记 FGF-19 多肽，如放射性标记，这样，结合在受体上的 FGF-19 多肽分子数量，就可用于测定潜在拮抗剂的效力。使用本领域已知的众多方法，例如，配体 panning 和 FACS 排序，可以鉴定编码受体的基因。Coligan 等, *Current Protocols in Immun.*, 1(2): Chapter 5(1991)。优选使用表达克隆，其中多聚腺苷酸化 RNA 是从对 FGF-19 多肽易感细胞中制备的，由该 RNA 建立的 cDNA 文库被分成数个库，用于转染 COS 细胞或其它对 FGF-19 多肽不易感的细胞。使玻璃载波片上生长的转染细胞与标记的 FGF-19 多肽接触。可用各种方法标记 FGF-19 多肽，包括碘化或位点特异蛋白激酶的识别位点的包埋。固定和温育后，载玻片接受放射自显影分析。鉴定阳性库，制备亚-库，并使用交互作用的亚-库进行再转染，重新进行筛选过程，直至得到编码推定受体的单个克隆。

作为受体鉴定的另一途径，可使标记的 FGF-19 多肽与细胞膜或表达受体分子的提取制品光亲和连接。用 PAGE 除去交联材料，暴露于 X-线胶片。可切除含受体的标记复合物，分解为肽片段，接受蛋白微测序。从微测序得到的氨基酸序列，用于设计一套筛选 cDNA 文库的简并寡核苷酸探针，以鉴定编码推定受体的基因。

在拮抗剂另一分析方法中，在候选化合物存在的条件下，哺乳动物细胞或表达受体的膜制品与标记的 FGF-19 多肽温育。然后测定化合物增强或阻断相互作用的能力。

潜在拮抗剂的更特异实例包括与免疫球蛋白和 FGF-19 多肽的融合物结

合的寡核苷酸，特别是抗体包括但不限于多克隆和单克隆抗体及抗体片段、单链抗体、抗-独特型抗体，和嵌合抗体或人源化抗体或抗体片段，以及人抗体和抗体片段。或者，潜在拮抗剂是密切相关的蛋白，例如，FGF-19 多肽的变种，后者识别受体但不引起效应，从而竞争性抑制 FGF-19 多肽的作用。

5 在本文的一实施方案中，做了竞争性结合实验，FGF 受体 4 或 FGF-19 抗体用作竞争剂。

10 另一潜在 FGF-19 多肽拮抗剂是使用反义技术制备的反义 RNA 或 DNA 构建体，其中例如，反义 RNA 或 DNA 分子通过与靶 mRNA 杂交和阻止蛋白翻译，起着直接阻断 mRNA 翻译的作用。反义技术可用于控制通过三链螺旋结构或反义 DNA 或 RNA 的基因表达，两种方法均建立在多聚核苷酸与 DNA 或 RNA 结合的基础上。例如，多聚核苷酸序列的 5'编码部分，该部分编码本文所述成熟 FGF-19 多肽，被用于设计约 10~40 个碱基对长度的反义 RNA 寡核苷酸。DNA 寡核苷酸被设计成参与转录基因区的互补链(三链螺旋参见 Lee 等, Nucl. Acids Res., 6: 3073(1979); Cooney 等, Science, 241: 15 456(1988); Dervan 等, Science, 251: 1360(1991)), 由此阻止转录和 FGF-19 多肽产生。反义 RNA 寡核苷酸在体内与 mRNA 杂交，阻断 mRNA 分子翻译成 FGF-19 多肽(反义，见 Okano, Neurochem., 56: 560(1991); 寡脱氧核苷酸作为基因表达的反义抑制剂(CRC Press: Boca Raton, FL, 1988)。也可把上述寡核苷酸释放入细胞中，这样反义 RNA 或 DNA 可在体内表达，而抑制 20 FGF-19 多肽的产生。当使用反义 DNA 时，优选衍生自翻译起始位点寡脱氧核苷酸，例如，靶基因核苷酸序列的约 -10 和+10 位。

25 潜在的拮抗剂包括与活性位点结合的小分子、受体结合位点、或生长因子或 FGF-19 多肽其它相关结合位点，从而阻断 FGF-19 多肽的正常生物活性。

小分子的实例包括但不限于，小肽或类-肽分子，优选溶解的肽和合成的非-肽有机或无机化合物。

30 核酶是能够催化 RNA 特异性裂解的具有酶性质的 RNA 分子。核酶通过序列-特异性杂交于靶 RNA 互补链，接着通过核酸内裂解而发挥作用。使用已知技术，可以鉴定位于潜在 RNA 靶内的特异核酶裂解位点。进一步的细节请参见例如 Rossi, Current Biology, 4: 469-471(1994),和 PCT 出版物 WO

97/33551(1997年9月18日公布)。

用于抑制转录的三链螺旋构型的核酸分子,应当是单链并由脱氧核苷酸构成。这些寡核苷酸碱基组成是设计好的,这样它通过 Hoogsteen 碱基对原则促进三链螺旋的形成,这一般需要双螺旋中一条链上的嘌呤或嘧啶相当大的延伸。更详细情况请参见例如, PCT 出版物 WO 97/33551, 出处同上。

用上述讨论的筛选试验的任何一个或多个分析方法,和/或使用本领域技术人员熟知的任何其它筛选技术,可以鉴定这些小分子。

值得欣慰的是,本文提供的分析方法可用于筛选多种候选生物活性剂。本文使用的术语"候选生物活性剂"、"候选剂"或"候选药物"或语法上的等价物,是指为了筛选能够直接或间接改变细胞活性表型或 FGF-19 序列表达(既包括核酸序列也包括蛋白序列)的生物活性剂,而被测试的任何分子,例如蛋白、寡肽、小的有机分子、多糖、多核苷酸、嘌呤类等。

尽管代表性的候选剂是有机分子,但是候选剂可以包括众多化学种类,优选分子量大于 100 而小于约 2500 道尔顿(d)的小有机化合物。本文可进一步限定分子量为 50 d~2000 d 的小分子。在另一实施方案中,小分子的分子量小于 1500、或小于 1200、或小于 1000、或小于 750、或小于 500 d。在一实施方案中,本文使用的小分子分子量为约 100~200 d。候选剂包括与蛋白相互作用所必需的结构上的官能团,特别是氢键,和通常包括至少一个氨基、羰基、羟基或羧基,优选至少两个化学官能团。候选剂常常包含被一个或多个上述官能团取代的环碳或杂环结构和/或芳香或多芳香结构。候选剂也见于生物分子,包括肽、糖、脂肪酸、类固醇、嘌呤、嘧啶、衍生物、结构类似物或它们的组合。特别优选肽。

候选剂得自广泛的各种来源,包括合成或天然化合物库。例如,有大量方法可用于随机和定向合成各种有机化合物和生物分子,方法包括随机化寡核苷酸的表达。或者,细菌、真菌、植物和动物提取物形式的天然化合物库是可以利用或很方便制造的。另外,通过常规的化学、物理和生化学方法,可以很容易地修饰天然或合成制造的库和化合物。可对已知的药剂进行定向或随机化学修饰,如酰化、烷化、酯化、amidification,以产生结构类似物。

在一优选的实施方案中,候选生物活性剂是蛋白。这里"蛋白"是指至少 2 个共价连接的氨基酸,它包括蛋白、多肽、寡肽和肽。蛋白由天然氨基酸和肽键构成,或由合成的类肽结构(peptidomimetic structures)构成。因此,本

文所述"氨基酸"或"肽残基",既指天然氨基酸,也指合成的氨基酸,例如,为了本发明的目的,高-苯丙氨酸、瓜氨酸和正亮氨酸被认为是氨基酸。"氨基酸"也包括亚基残基如脯氨酸和羟脯氨酸。侧链可以是(R)或(S)构型。在一优选的实施方案中,氨基酸是(S)或 L-构型。如果使用非天然侧链,则可以
5 使用非-氨基酸取代基,例如为预防或阻止体内降解。

在一优选的实施方案中,候选生物活性剂是天然蛋白或天然存在蛋白的片段。因此,例如可以使用含有蛋白的细胞提取物,或者随机或定向消化的蛋白性细胞提取物。以这种方式,可以制备用于本发明筛选方法的原核生物和真核生物蛋白库。此实施方案中,特别优选细菌、真菌、病毒和哺乳动物
10 蛋白库,以后者为优选,特别优选人蛋白。

在一优选的实施方案中,候选生物活性剂是约 5~约 30 个氨基酸的肽,优选约 5~约 20 个氨基酸,特别优选约 7~约 15 个氨基酸。肽可能是如上略述的天然存在蛋白的消化产物、随机肽或"边缘"随机肽。这里的"随机肽"或语法上的等价物是指每一个核酸和肽基本上分别由随机核苷酸和氨基酸
15 构成。既然这些随机肽(或核酸,见下述)通常是化学合成的,那么它们可在任何位置整合入任意核苷酸或氨基酸。可设计合成路线以产生随机蛋白或核酸,使得沿着序列长度形成所有或大多数可能的结合,由此形成随机候选活性蛋白剂库。

在一实施方案中,文库是完全随机的,在任何位置均没有序列优先或恒定。在一优选的实施方案中,文库是偏倚性的,也即序列内的某些位置或者
20 保持恒定,或者只选自有限数目的可能性。譬如,在一优选的实施方案中,核苷酸或氨基酸残基在限定种类内是随机的,例如,是疏水氨基酸、亲水残基、空间位阻性偏倚(小的或大的)残基,目的是为了产生核酸结合域,为产生交联用的半胱氨酸,为了 SH-3 区的脯氨酸,为了磷酸化位点的丝氨酸、
25 苏氨酸、酪氨酸或组氨酸等,或为了与嘌呤结合等。

在一优选的实施方案中,候选生物活性剂是核酸。本文所述"核酸"或"寡核苷酸"或语法上的等价物是指至少 2 个核苷酸共价连接。本发明核酸通常含有磷酸二酯键,尽管在某些情况下,如下面提到的那样,包括具有可变主链的核酸类似物,后者含有例如磷酰胺(Beaucage 等, Tetrahedron 49(10):
30 1925(1993)和其中的引用文献; Letsinger, J. Org. Chem. 35: 3800(1970); Sprinzl 等, Eur. J. Biochem. 81: 579(1977); Letsinger 等, Nucl. Acids Res. 14:

3487(1986); Sawai 等, Chem. Lett. 805(1984), Letsinger 等, J. Am. Chem. Soc. 110: 4470(1988);和 Pauwels 等, *Chemica Scripta* 26: 141 9(1986)), 硫代磷酸酯 (Mag 等, *Nucleic acids Res.* 19: 1437(1991);和美国专利 5644048), 二硫代磷酸酯 (Briu 等, J. Am. Chem. Soc. 111: 2321(1989), 氧-甲基磷酰胺酯 (phosphoramidite) 连接 (参见 Eckstein, *Oligonucleotide and Analogues: A Practical Approach*, Oxford University Press), 和肽核酸主链及连接 (见 Egholm, J. Am. Chem. Soc. 114: 1895(1992); Meier 等, *Chem. Int. Ed. Engl.* 31: 1008(1992); Nielsen, *Nature*, 365: 566(1993); Carlsson 等, *Nature* 380: 207(1996), 所有这些文献均引入参考)。其它类似核酸包括那些带正电主链的核酸 (Denpcy 等, *Proc. Natl. Acad. Sci. USA* 92: 6097(1995); 非-离子主链 (美国专利 5386023, 5637684, 5602240, 5216141 和 4469863; Kiedrowshi 等, *Angew. Chem. Intl. Ed. English* 30: 423(1991); Letsinger 等, J. Am. Chem. Soc. 110: 4470(1988); Letsinger 等, *Nucleoside & glycoside* 13: 1597(1994); Chapters 2 和 3, *ASC Symposium Series 580, "Carbohydrate Modifications in antisense Research"*, Ed. Y. S. Sanghui 和 P. Dan Cook; Mesmaeker 等, *Bioorganic & Medicinal Chem. Lett.* 4: 395(1994); Jeffs 等, *J. Biomolecular NMR* 34: 17(1994); *Tetrahedron Lett.* 37: 743(1996)) 和非-核糖主链, 包括在美国专利 5235033、5034506 以及 Chapters 6 和 7, *ASC Symposium Series 580, "Carbohydrate Modifications in antisense Research"*, Ed. Y. S. Sanghui P. Dan Cook 中描述的那些。含有一个或多个碳环糖的核酸也包括在核酸的定义中 (参见 Jenkins 等, *Chem. Soc. Rev.*(1995)ppl69-176)。数种核酸类似物在 Rawls, *C & E News* June 2, 1997 page 35 中作了描述。所有这些文献均以其全部在此引入参考。对核糖-磷酸酯主链进行的这些修饰, 是为了便于添加附加部分如标记物, 或者为了增加分子在生理环境中的稳定性和半衰期。此外, 可以制备天然核酸与类似物的混合物。或者, 制备不同核酸类似物的混合物, 和天然存在核酸与类似物的混合物。按指明的那样, 核酸可以是单链或双链, 或者既含有双链也含有单链序列部分。核酸可以是 DNA (既可是基因组也可是 cDNA)、RNA 或杂交体, 其中核酸含有脱氧核糖和核糖核苷酸的任何组合, 和碱基任意组合, 包括尿嘧啶、腺嘌呤、胸腺嘧啶、胞嘧啶、鸟嘌呤、次黄苷、xathanine hypoxathanine、异胞嘧啶、异鸟嘌呤等。

如上面对蛋白所作的总括性描述那样, 核酸候选生物活性剂可以是天然

存在核酸、随机核酸或"偏倚的"随机核酸。例如,如上面对蛋白所作描述那样,可使用原核或真核基因组的消化产物。

在一优选的实施方案中,候选生物活性剂是有机化学分子,文献中记载了可利用的各种有机化学分子。

5 如上所述,在一优选的实施方案中,可对单个基因和基因产物(蛋白)进行筛选。在一优选的实施方案中,按照下面实施例中所描述的,已经鉴定出与特定组织差异地表达的基因有关的基因或蛋白,因而所述病症与这些组织相关。因此,在一实施方案中,把筛选设计成首先找到与 FGF-19 结合的候选剂,然后将寻找到的这些试剂用于评价调节 FGF-19 活性的候选剂活性的
10 试验中。所以,正如将被本领域技术人员领会的那样,可以进行大量不同的试验。

也可对调节 FGF-19 活性的试剂进行筛选。在一优选的实施方案中,能够调节 FGF-19 生物活性剂的筛选方法,包括将候选生物活性剂加入到 FGF-19 的样品中,并测定 FGF-19 生物活性的改变。"FGF-19 活性的调节",
15 包括活性增加、活性降低或者活性类型的变化。因此,在该实施方案中,候选剂应当既与 FGF-19 结合(尽管这可能是不必要的)又影响 FGF-19 的所述生物学或生化学活性。方法既包括上述体外筛选方法,也包括体内改变 FGF-19 存在、表达、分布、活性或含量的细胞的筛选方法。

因此,在该实施方案中,方法包括将样品与候选生物活性剂结合,并评价对 FGF-19 活性的影响。"FGF-19 蛋白活性"或语法上的等价物,在这里是指 FGF-19 蛋白的至少一种上述生物活性。
20

在一优选的实施方案中,FGF-19 蛋白的活性增加;在另一优选实施方案中,FGF-19 蛋白的活性下降。因而,在某些实施方案中优选作为拮抗剂的生物活性剂,在其它实施方案中优选作为激动剂的生物活性剂。

25 本发明的一个方面,药物筛选试验中使用含 FGF-19 序列的细胞,检测候选药物对 FGF-19 的影响。细胞类型包括正常细胞、肿瘤细胞和脂肪细胞。

评定 FGF-19 活性的方法为本领域所已知,并且在下述实施例中作例证性描述,所述活性如在葡萄糖摄取、leptin 释放、代谢、甘油三酯和游离脂肪酸水平、体重和机体脂肪方面的变化。

30 在一优选的实施方案中,方法包括将上述候选生物活性剂加入到含有 FGF-19 的细胞中。优选的细胞类型包括几乎任何细胞。细胞含有编码 FGF-19

蛋白的核酸, 优选重组物。在一优选的实施方案中, 候选剂库在多细胞中进行检测。

一方面, 试验在存在或缺乏生理信息物的情况下或者在暴露于生理信息物之前或之后进行评定, 所述生理信息物例如激素、抗体、肽、抗原、细胞因子、生长因子、动作电位, 药理学试剂包括化疗剂、放疗剂、致癌剂或其它 cells(即细胞-细胞接触)。在另一实例中, 在细胞周期进程的不同期进行测定。

本文提供的 FGF-19 序列也可在诊断方法中使用。

FGF-19 的过度表达表明不正常代谢率, 而其低表达则表明肥胖倾向。而且, 可以分析取自患者的样品中 FGF-19 是否变异或失去功能。通常, 此类方法包括将取自患者的样品以及 FGF-19 的表达与对照进行比较。

F. 抗-FGF-19 抗体

本发明进一步提供抗-FGF-19 抗体。示范性抗体包括多克隆抗体、单克隆抗体、人源化抗体、双特异抗体和异源偶联抗体。

1. 多克隆抗体

抗-FGF-19 抗体包括多克隆抗体。制备多克隆抗体的方法为本领域熟练技术人员所已知。多克隆抗体可在哺乳动物中产生, 例如, 通过一次或多次注射致免疫剂, 如果需要, 可加入佐剂。通常, 对哺乳动物皮下或腹膜内多次注射致免疫剂和/或佐剂。致免疫剂可包括 FGF-19 多肽或其融合蛋白。将致免疫剂与针对所免疫的哺乳动物具有免疫原性的蛋白进行偶联是有效的。所述致免疫蛋白的实例包括但不限于: 匙孔血蓝蛋白、血清白蛋白、牛甲状腺球蛋白和大豆胰蛋白酶抑制剂。可以使用的佐剂的实例包括弗氏完全佐剂和 MPL-TDM 佐剂(单磷脂酰脂质 A, 合成的海藻糖双白喉菌酸酯(dicorynomycolate))。无需繁琐的试验, 本领域技术人员就可选择致免疫方案。

2. 单克隆抗体

或者, 抗-FGF-19 抗体是单克隆抗体。可以使用杂交瘤方法制备单克隆抗体, 如 Kohler 和 Milstein 在 Nature, 256:495(1975)中描述的方法。在杂交瘤方法中, 通常用致免疫剂免疫小鼠、仓鼠或其它适宜宿主动物, 以引发产生或能够产生将与致免疫剂特异性结合的抗体的淋巴细胞。或者, 在体外致敏淋巴细胞。

致免疫剂通常包括 FGF-19 多肽或其融合蛋白。一般而言, 如果需要来

源于人的细胞, 则使用外周血淋巴细胞("PBLs"), 如果需要非人哺乳动物来源的细胞, 那么就使用脾细胞或淋巴结细胞。然后, 使用适宜的融合试剂如聚乙二醇, 将淋巴细胞与无限增殖细胞系融合, 从而形成杂交瘤细胞[Goding, 单克隆抗体: Principles 和 Practice, Academic Press,(1986)59-103 页]。无限增殖细胞系通常是转化的哺乳动物细胞, 特别是啮齿类、牛和人类起源的骨髓瘤细胞。通常使用大鼠或小鼠骨髓瘤细胞系。可在适宜培养基中培养杂交瘤细胞, 所述培养基优选含有一种或多种抑制未融合无限增殖细胞生长或生存的物质。例如, 母代细胞缺乏酶次黄嘌呤鸟嘌呤磷酸核糖转移酶(HGPRT 或 HPRT), 杂交瘤的培养基通常包括次黄嘌呤、氨基蝶呤和胸腺嘧啶核苷("HAT 培养基"), 这些物质抑制 HGPRT-缺陷细胞生长。

优选的无限增殖细胞系, 是那些有效融合、支持筛选出的抗体-生成细胞稳定地高水平表达抗体, 并且对培养基如 HAT 敏感的细胞系。更优选的无限增殖细胞系是鼠骨髓瘤系, 后者可得自例如 Salk Institute Cell Distribution Center, San Diego, California 和美国模式培养物保藏所, Manassas, Virginia。人骨髓瘤和鼠-人杂交骨髓瘤细胞系也已用于产生人单克隆抗体 [Kozbor, J.Immunol., 133: 3001(1984); Brodeur 等, Monoclonal antibody Production Techniques and Applications, Marcel Dekker, Inc., New York,(1987)pp. 51-63]。

接下来, 可以检定培养杂交瘤细胞的培养基中是否存在直接抗 FGF-19 的单克隆抗体。优选地, 用免疫沉淀或体外结合试验如放免分析(RIA)或酶联免疫吸附分析(ELISA), 来测定杂交瘤细胞产生的单克隆抗体的结合特异性。这些技术和试验分析方法是本领域所已知的。使用例如 Scatchard analysis of Munson 和 Pollard, Anal. Biochem., 107: 220(1980)公开的方法, 可以测定单克隆抗体的结合亲和力。

鉴定出所需的杂交瘤细胞之后, 经有限稀释法将克隆进行亚克隆处理并采用标准方法使之生长[Goding, 出处同上]。用于此目的适宜的培养基包括例如, Dulbecco's 改进的 Eagle's 培养基和 RPMI-1640 培养基。或者, 杂交瘤细胞可以在哺乳动物体内以腹水形式生长。

使用常规免疫球蛋白纯化方法, 例如蛋白 A-琼脂糖凝胶、羟基磷灰石层析法、凝胶电泳、渗析或亲和层析, 可以从培养基或腹水中分离或纯化由亚克隆分泌的单克隆抗体。

也可使用重组 DNA 方法生产单克隆抗体,例如在美国专利 4816567 中描述的方法。使用常规技术(例如,使用能够与编码鼠抗体重链和轻链的基因特异结合的寡核苷酸探针),很容易地分离编码本发明单克隆抗体的 DNA 并进行序列测定。本发明的杂交瘤细胞作为此 DNA 的优选来源。一旦分离出该 DNA,就可将其植入到表达载体中,然后用后者转染宿主细胞,所述宿主细胞如猴 COS 细胞、中国仓鼠卵巢(CHO)细胞或不产生免疫球蛋白的骨髓瘤细胞,以便在重组宿主细胞中合成单克隆抗体。也可对 DNA 进行修饰,例如人重链和轻链恒定区的编码序列被相应的鼠序列置换[美国专利 4816567; Morrison 等,出处同上],或者将非-免疫球蛋白多肽的全部或部分编码序列共价连接到免疫球蛋白的编码序列上。此类非-免疫球蛋白多肽可取代本发明抗体的恒定区,或可取代本发明抗体一个抗原结合位点的可变区,以产生嵌合二价抗体。

抗体可以是单价抗体。制备单价抗体的方法为本领域所已知。例如,有一个方法涉及免疫球蛋白轻链和修饰后重链的重组表达。通常在 Fc 区的任何位点将重链截短,从而防止重链交联。或者,用其它氨基酸残基代替相关的半胱氨酸残基或者半胱氨酸残基缺失以防止交联。

体外方法也适于制备单价抗体。使用本领域常规技术可以完成抗体消化而产生抗体片段,特别是 Fab 片段。

3. 人抗体和人源化抗体

本发明抗-FGF-19 抗体进一步包括人源化抗体或人抗体。非-人(例如,鼠)的人源化抗体是嵌合免疫球蛋白、免疫球蛋白链或其片段(如 Fv, Fab, Fab', F(ab')₂ 或抗体的其他抗原结合亚序列),它们包含非人免疫球蛋白的最小序列。人源化抗体包括人免疫球蛋白(受者抗体)中受者互补决定区(CDR)残基被具有所需特异性、亲和力和性能的小鼠、大鼠、家兔等非人源物种抗体(供体抗体)的 CDR 残基所取代。在一些实例中,人免疫球蛋白的 Fv 框架区残基由相应的非人类残基所取代。人源化抗体也可包含受者抗体或供体 CDR 或框架序列中均不存在的残基。通常,人源化抗体基本上包括至少一个、通常两个可变区的全部,其中 CDR 的全部或基本上全部对应于非人免疫球蛋白的相应部分,而 FR 序列的全部或基本上全部是人免疫球蛋白序列。人源化抗体还任选包括免疫球蛋白恒定区(Fc),通常为人免疫球蛋白的恒定区的至少一部分。详见 Jones 等,自然,321: 522-525(1986); Riechmann 等,

自然 332: 323-329(1988); 和 Presta, *Curr. Op. Struct. Biol.* 2: 593-596(1992)。

人源化非-人抗体的方法是本领域技术人员所已知的。通常, 人源化抗体中已导入一或多个源自非人类的氨基酸残基。这些非人类氨基酸残基常称为“引进的”残基, 它们通常来自“引进的”可变区。人源化过程基本如

5 Winter 及其同事(Jones 等, 自然, 321:522-525(1986); Riechmann 等, 自然, 332:323-327(1988); Verhoeyen 等, 科学, 239:1534-1536(1988))所述, 用一个或多个 CDR 序列取代人类抗体的相应序列来进行。因此, 这样的“人源化”抗体是嵌合抗体(美国专利 4816567), 其中完整人类可变区的很少一部分被非人类物种的相应序列取代。实践中, 人源化抗体通常是人的抗体, 其中一些 CDR 残基且可能有部分 FR 残基被啮齿类抗体中类似位点的残基取代。

使用本领域已知的各种技术可以制备人抗体, 所述方法包括噬菌体展示文库[Hoogenboom 和 Winter, *J. Mol. Biol.*, 227: 381(1991); Marks 等, *J. Mol. Biol.*, 222: 581(1991)]。也可使用 Cole 等和 Boerner 等描述的技术制备人单克隆抗体[Cole 等, *Monoclonal Antibody and Cancer Therapy*, Alan R. Liss,(1985) 15 第 7 页, 和 Boerner 等, *J. Immunol.*, 147 1: 86-95(1991)]。类似地, 通过把人免疫球蛋白的基因座插入转基因动物例如鼠来生产人抗体, 所述宿主动物的内源性免疫球蛋白基因已被部分或全部灭活。攻击后, 观察到人抗体产生, 这在各个方面均与在人体所见的极为相似, 包括基因重排、装配和抗体所有组成成分。有关进展在例如下列专利和科学出版物中作了描述: 美国专利 20 5545807; 5545806; 5569825; 5625126; 5633425; 5661016; Marks 等, *Bio/Technology* 10, 779-783(1992); Lonberg 等, *Nature* 368 856-859(1994); Morrison, *Nature* 368, 812-13(1994); Fishwild 等, *Nature Biotechnology* 14, 845-51(1996); Neuberger, *Nature Biotechnology* 14, 826(1996); Lonberg 和 Huszar, *Intern. Rev. Immunol.* 13 65-93(1995)。

25 4. 双特异抗体

双特异抗体是单克隆抗体, 优选具有针对至少两种不同抗原的结合特异性的人抗体或人源化抗体。在本发明中, 一种结合特异性针对 FGF-19, 另一种针对其它抗原, 优选针对细胞表面蛋白或受体或受体亚单位。

制备双特异抗体的方法为本领域所已知。双特异性抗体的传统制备方法, 是基于两种免疫球蛋白重链-轻链对的共表达, 其中这两条链具有不同 30 特异性(Millstein 等, *Nature*, 305: 537-539(1983))。由于免疫球蛋白重链轻

链随机分配, 这些杂交瘤(细胞杂交瘤(quadroma))可能产生 10 种不同抗体分子的混合物, 其中只有一种具有正确的双特异性结构。对所述正确分子的纯化(通常通过亲和层析步骤来进行)非常复杂, 且产量很低。类似的方法见 WO93/08829 和 Traunecker 等, EMBO J, 10: 3655-3659(1991)。

- 5 可将具有所需结合特异性(抗体-抗原结合位点)的抗体可变区与免疫球蛋白恒定区序列融合。该融合优选与包含铰链区的至少一部分、CH2 及 CH3 区的免疫球蛋白重链恒定区融合。优选使含有轻链结合所需位点的第一重链恒定区(CH1)出现在至少在一种融合中。可将编码免疫球蛋白重链融合体, 以及必要时, 编码免疫球蛋白轻链的 DNA 插入不同表达载体, 共转染至适当宿主生物。产生双特异抗体的进一步细节内容, 请参见例如 Suresh 等, Methods in Enzymology, 121: 210(1986)。

10 根据 WO 96/27011 所述的另一种方法, 可改造一对抗体分子之间的界面, 使得从重组细胞培养中获得的异源二聚体的百分比最大。优选的界面包括抗体恒定区 CH3 结构域的至少一部分。在该方法中, 源于第一抗体分子界面上的一条或多条氨基酸小侧链被较大侧链(如酪氨酸或色氨酸)取代。与
15 所述大侧链大小相同或相近的互补“沟”可通过将氨基酸大侧链用小侧链(如丙氨酸或苏氨酸)取代而在第二抗体分子的界面上形成。这使得异二聚体的产量比不想要的终产物如同型二聚体的高。

20 双特异抗体可以制备全长抗体或抗体片段(如 $F(ab')_2$ 双特异抗体)。从抗体片段制备双特异性抗体的技术已有文献。例如, 双特异性抗体可利用化学连接制备。Brennan 等, 科学 229: 81(1985)中描述了将完整抗体经蛋白水解制备片段的方法。这些片段在二巯基复合剂亚砷酸钠存在时被还原, 从而稳定相邻的巯基, 并阻止分子间二硫键的形成。生成的 Fab' 片段被转化为硫硝基苯甲酸盐(TNB)衍生物。其中一种 Fab'-TNB 衍生物经巯基乙胺还原成
25 Fab'-硫醇, 再与等分子量的其它 Fab'-TNB 衍生物混合形成双特异性抗体。如此产生的双特异性抗体可作为酶的选择性固相化中所用的试剂。

30 近期的进展促进了 Fab'-SH 片段从大肠杆菌的直接回收, 该片段可经化学偶联形成双特异性抗体。Shalaby 等, 实验医学杂志, 175:217-225(1992)中描述了完全人源化双特异性抗体 $F(ab')_2$ 分子的产生。每一 Fab' 片段分别从大肠杆菌中分泌出来, 体外直接化学偶联形成双特异性抗体。如此制备的双特异性抗体能与过度表达 ErbB2 受体的细胞和正常人 T 细胞结合, 还能引

发人类细胞毒淋巴细胞对人乳腺肿瘤细胞的裂解活性。

直接从重组细胞培养中制备并分离双特异性抗体片段的多种技术也已有描述。例如，可用亮氨酸拉链制备双特异性抗体。Kostelny 等，免疫学杂志，148(5):1547-1553(1992)。将来自 Fos 和 Jun 蛋白的亮氨酸拉链肽与两种不同抗体的 Fab' 部分通过基因融合而连接。使抗体的同型二聚体在铰链区被还原成单体，然后被再氧化形成抗体的异二聚体。该方法也可用于制备抗体同型二聚体。由 Hollinger 等，美国国家科学院学报，90:6444-6448(1993)) 描述的“二价抗体”技术提供了另一种制备双特异性抗体片段的方法。所述片段中含有重链可变区(V_H)，其通过接头与轻链可变区(V_L)相连，该接头非常短，使得同一链的两个结构域之间无法配对。因此，同一片段上的 V_H 和 V_L 结构域被迫与另一片段上的互补 V_L 和 V_H 结构域配对，从而形成两个抗原结合位点。此外还报道了另一种用单链 Fv(sFv)二聚体来制备双特异性抗体的策略。见 Gruber 等，免疫学杂志，152: 5368(1994)。还考虑了二价以上的抗体。如可制备三特异性抗体。Tutt 等，免疫学杂志，147: 60(1991)。

例证性双特异性抗体可与本文指定的 FGF-19 多肽上的两种不同表位结合。或者，可将抗-FGF-19 多肽臂与白细胞上引发分子的臂结合，从而强化针对表达特定 FGF-19 多肽细胞的细胞防御机制，所述引发分子如 T 细胞受体分子(CD2、CD3、CD28 或 B7)，或 IgG Fc 受体(Fc γ R)如 Fc γ R I (CD64)、Fc γ R II (CD32)和 Fc γ R III(CD16)。双特异性抗体还可用于将细胞毒制剂定位至表达特定 FGF-19 多肽的细胞上。这些抗体具有 FGF-19-结合臂和结合细胞毒制剂或放射性同位素半抗原的臂，如 EOTUBE、DPTA、DOTA 或 TETA。另一有关的双特异抗体结合 FGF-19 多肽并且进一步结合组织因子(TF)。

5. 异源偶联抗体

异源偶联抗体也包括在本发明范围之内。异源偶联的抗体由两个共价连接的抗体组成。有观点认为，这类抗体可用于将免疫细胞导向不想要的细胞(美国专利 4676980)，也可用于治疗 HIV 感染(WO91/00360, WO92/200373, EP03089)。可考虑使用蛋白化学合成中已知方法，包括涉及使用交联剂，在体外制备异源偶联抗体。举例来说，使用二硫化物交换反应或通过形成硫醚键，可以构建免疫毒素。用于此目的适当试剂的实例，包括免疫硫醇盐和甲基-4-巯基，巯基 butyrimidate 以及例如在美国专利 4676980 中描述的那些。

6. 效应物功能的工程改造

在效应物功能方面可能希望改进本发明的抗体，以增强在治疗(例如治疗癌症)时的效果。例如，可在Fc区域中引入半胱氨酸残基，从而在该区域中形成链间二硫键。这样制得的同型二聚体可能有改进的内化能力和/或提高的补体介导细胞杀伤能力及依赖于抗体的细胞毒性(ADCC)。参见 Caron 等, J. Exp Med. 176:1191-1195(1992) 和 Shopes, B. J. Immunol. 148:2918-2922(1992)。也可如 Wolff 等, Cancer Research 53:2560-2565(1993) 描述的那样，用异二聚功能性交联剂来制备抗肿瘤活性增强的同型二聚抗体。或者，可将抗体经工程化改造成具有双Fc区，从而可能增强补体裂解和 ADCC 能力。见 Stevenson 等, Anti-Cancer Drug Design 3:219-230(1989)。

7. 免疫偶联物

本发明还涉及免疫偶联物，其中含有与细胞毒药物偶联的抗体，所述细胞毒药物如化疗剂、毒素(例如小分子毒素或细菌、真菌、植物或动物来源的酶活性毒素，包括其片段和/或变体)或放射性同位素(即放射偶联物)。

适用于制备这类免疫偶联物的化疗剂在上文已作描述。可以应用的酶活性毒素及其片段包括：白喉毒素 A 链、白喉毒素的非结合活性片段、外毒素 A 链(来自铜绿假单胞菌)、蓖麻毒蛋白 A 链、相思豆毒蛋白 A 链、蒴莲根毒素 A 链、 α -帚曲毒素、油桐(*Aleutites fordii*)蛋白、石竹素蛋白、美洲商陆(*Phytolaca Americana*)蛋白(PAPI, PAPII, PAP-S)、苦瓜(*Momordica charantia*)抑制因子、麻疯树毒蛋白、巴豆毒蛋白、肥皂草(*Saponaire officinalis*)抑制剂，白树毒素，米托菌素(mitogellin)、局限曲菌素、酚霉素、依诺霉素和单端孢菌毒素(tricothecenes)。多种放射性同位素可用于制备放射性偶联的抗体，实例包括 Bi²¹²、I¹³¹、In¹³¹、Y⁹⁰ 和 Re¹⁸⁶。

抗体与细胞毒制剂的偶联物可通过多种双功能蛋白偶联剂来连接，所述双功能蛋白偶联剂如：N-琥珀酰亚氨基-3-(2-吡啶基二巯基)丙酸酯(SPDP)，琥珀酰亚氨基-4-(N-马来酰亚氨基甲基)环己烷-1-羧酸酯，亚氨基硫烷(iminothiolane)(IT)，亚氨酸酯的双功能衍生物(如亚氨基己二酸二甲酯盐酸盐)，活性酯类(如二琥珀酰亚胺基辛二酸酯)，醛类(如戊二醛(glutareldehyde))，双-叠氮化合物(如双(对-叠氮基苯甲酰基)己二胺)，双-重氮衍生物(如双-(对-重氮苯甲酰基)-乙二胺)，二异氰酸酯(如亚甲代苯基 2,6-二异氰酸酯)，和双-活性氟化合物(如 1,5-二氟-2,4-二硝基苯)。例如，蓖麻毒

蛋白免疫毒素可如 Vitetta 等, 科学 238: 1098(1987)所述制备。C¹⁴ 标记的 1-异硫氰酸苯甲基-3-甲基二亚乙基三氯五乙酸酯(MX-DTPA)是将放射性核苷酸偶联至抗体的偶联剂之一。见 WO94/11026。

5 在另一实施方案中, 抗体可与肿瘤预靶向中应用的“受体”(如亲和素)偶联, 将该抗体-受体偶联物施用于患者, 之后用清除剂除去循环中未结合的偶联物, 再给予已偶联了细胞毒制剂(如放射性核苷酸)的“配体”(如抗生物素蛋白)。

8. 免疫脂质体

10 本文公开的抗体还可制成免疫脂质体。含抗体的脂质体可通过本领域已知方法制备, 如 Epstein 等, 美国国家科学院学报 82: 3688(1985); Hwang 等, 美国国家科学院学报 77: 4030(1980); 美国专利 4485045 和 4544545 及 1997 年 10 月 23 日公开的 WO97/38731。在美国专利 5013566 中公开了循环时间已增加了的脂质体。

15 特别有用的脂质体可利用包含磷脂酰胆碱、胆固醇和 PEG 衍生的磷脂酰乙醇胺(PEG-PE)的脂质组合物经反相蒸发法而制得。脂质体挤压通过限定孔径的滤膜, 获得具有所需直径的脂质体。本发明抗体的 Fab' 片段可如 Martin 等, 生物学化学杂志, 257: 286-288(1982)所述的那样, 经二硫化物交换反应与脂质体偶联。可任选在所述脂质体中包含一种化疗剂。见 Gabizon 等, J. National Cancer Inst, 81(19)1484(1989)。

20 9. 抗体的药学组合物

与本文鉴定的 FGF-19 多肽特异性结合的抗体, 以及用前述公开的筛选试验鉴定的其它分子, 可以药学组合物的形式施用来治疗各种病症。

25 如果 FGF-19 多肽是在细胞内并且使用完整抗体作为抑制剂, 则优选内化抗体。然而, 也可用脂质转染物或脂质体将抗体或抗体片段递送至细胞内。使用抗体片段时, 优选与靶蛋白结合结构域特异结合的最小抑制剂片段。例如, 基于抗体可变区序列, 可把肽分子设计为保留与靶蛋白序列结合的能力。此类肽可以化学合成和/或用重组 DNA 技术产生。参见例如 Marasco 等, Proc. Natl. Acad. Sci. USA, 90: 7889-7893(1993)。这里所述制剂也包含一种以上对于所治特定适应征而言必要的活性化合物, 优选具有彼此之间无不利影响的
30 互补活性的那些化合物。或者或另外, 组合物可含有增强其功能的试剂, 诸如细胞毒素剂、细胞因子、化疗剂或生长抑制剂。这些分子适宜以对所治疗

目的而言有效的量联合。

也可将活性组分包封在按照例如凝聚技术或界面聚合技术制备的微胶囊中，例如，分别在胶态药物递送系统(例如脂质体，白蛋白微球体，微乳剂，纳米颗粒和纳米胶囊)或粗滴乳化剂中的羟甲基纤维素或明胶微胶囊和聚-(甲基丙烯酸甲酯)微胶囊。这些技术公开在 Remington's Pharmaceutical Sciences, 出处同上中。

用于体内给药的制剂必须是无菌的。这可以通过除菌滤膜过滤而轻易实现。

也可制备控释制剂。控释制剂的适当实例包括含有抗体的固态疏水聚合物的半通透性基质，所述基质为具有一定形状的制品，如膜或微胶囊。控释制剂实例包括聚酯、水凝胶(如聚(2-羟基乙基-异丁烯酸酯)或聚(乙烯醇)，聚交酯(美国专利 3,773,919)，L-谷氨酸与 γ 乙基-L-谷氨酸酯的共聚物，不可降解的乙烯乙酸乙酯，可降解的乳酸-羟基乙酸共聚物如 LUPRON DEPOT™(由乳酸-羟基乙酸共聚物和酸亮丙瑞林(leuprolide)组成的可注射的微球体)，以及聚 D-(-)-3-羟丁酸。尽管聚合物如乙烯-乙酸乙烯酯和乳酸-羟基乙酸能持续释放分子 100 天以上，但是某些水凝胶释放蛋白的时间却较短。当胶囊化的抗体长时间停留在体内时，它们会由于暴露在 37°C 水分下而变性或凝聚，从而导致生物活性损失，且免疫原性可能会改变。可以根据涉及的机理来设计稳定化的合理策略。例如，如果发现凝聚的机理是通过硫代二硫键互换而形成了分子间 S-S 键，则可通过修饰巯基残基、自酸性溶液中冻干、控制湿度程度、采用合适的添加剂和开发特殊的聚合物基质组合物来达到稳定化。

G. 应用抗-FGF-19 抗体

本发明抗-FGF-19 抗体具有多种用途。例如，抗-FGF-19 抗体可用于 FGF-19 的诊断分析，譬如，检测 FGF-19 在特定细胞、组织或血清的表达。可使用本领域已知的各种诊断分析试验，如竞争性结合试验、直接和间接夹心试验，及免疫沉淀试验。Zola, 单克隆抗体：技术指南，147-158 页(CRC Press, Inc. 1987)。还可将所述抗体标记上可检测部分。可检测部分应当能够直接或间接产生可测得信号。例如，可测得部分是放射性同位素如 ^3H 、 ^{14}C 、 ^{32}P 、 ^{35}S 或 ^{125}I ，荧光或化学发光化合物如异硫氰酸荧光素、罗丹明或萤光素，或者酶如碱性磷酸酶、 β -半乳糖苷酶或辣根过氧化物酶。本领域已知的将抗体与可检测部分偶联的任何方法均可使用，包括在 Hunter 等, Nature, 144:

945(1962); David 等, Biochemistry, 13: 1014(1974); Pain 等, J. Immunol. Meth., 40: 219(1981);和 Nygren, J. Histochem.和 Cytochem., 30: 407(1982)中描述的方法。

5 抗-FGF-19 抗体也适用于从重组细胞培养物或天然来源中亲和纯化 FGF-19。在该过程中, 使用本领域已知方法将抗 FGF-19 抗体固定在适当载体如 Sephadex 树脂或滤纸上。然后, 将固定的抗体用含有欲纯化的 FGF-19 的样品包被, 接下来用能够基本上将样品中与固定载体结合的 FGF-19 之外的所有物质去除的适宜溶剂洗涤载体。最后, 用能从抗体中释出 FGF-19 的另一适宜溶剂洗涤载体。

10 下述实施例只是为了描述的目的而提供, 并非旨在以任何方式限定本发明的保护范围。

本说明书中引证的所有专利和参考文献, 以其全部在此引入作为参考。

实施例

15 除非另外声明, 否则实施例中所提到的可商购试剂, 按照制造商的说明使用。在下述实施例和整个说明书中用 ATCC 登记号定义的细胞的来源是在美国典型培养物保藏中心, Manassas, VA 进行保藏。

实施例 1

分离编码人 FGF-19 的 cDNA 克隆

20 EST 序列登记号为 AF007268 的鼠成纤维细胞生长因子(FGF-15)用于检索各种公用 EST 数据库(如 GenBank, Dayhoff 等)。使用计算机程序 BLAST 或 BLAST2[Altschul 等, Methods In Enzymology, 266:460-480(1996)]比较 EST 序列 6 框架翻译与 ECD 蛋白序列来进行检索。检索结果击中 GenBank EST AA220994, 后者已被鉴定为 STRATAGENE NT2 神经元前体 937230。AA220994 的序列在这里也称为 DNA47412。

25 基于 DNA47412 序列, 合成寡核苷酸: 1)以使用 PCR 方法识别含有目标序列的 cDNA 文库, 和 2)用作分离 FGF-19 全长编码序列的克隆的探针。正向和反向 PCR 引物一般含有 20~30 个核苷酸, 并且常被设计成产生长度约 100-1000 bp 的 PCR 产物。探针序列长度通常为 40-55 bp。有些情况下, 当共有序列大于约 1-1.5kbp 时, 则合成另外的寡核苷酸。为了筛选数个全长克
30 隆的文库, 按照 Ausubel 等在 Current Protocols in Molecular Biology, 出处同上中描述的方法, 采用 PCR 扩增法用 PCR 引物对从文库中筛选 DNA。然后,

使用探针寡核苷酸和引物对之一，用阳性文库分离编码目标基因的克隆。

合成 PCR 引物(正向和反向):

正向 PCR 引物 5'-ATCCGCCAGATGGCTACAATGTGTA-3'(SEQ ID NO: 3), 和

5 反向 PCR 引物 5'-CCAGTCCGGTGACAAGCCCAA-3'(SEQ ID NO: 4)。

另外，由具有下述核苷酸序列的 DNA47412 序列构建合成性寡核苷酸杂交探针:

杂交探针

10 5'-GCCTCCCGGTCTCCCTGAGCAGTGCCAAACAGCGGCAGTGTA-3'
(SEQ ID NO: 5)。

用于构建 cDNA 文库的 RNA，是从人胎儿视网膜组织分离出来的。用于分离 cDNA 克隆的 cDNA 文库，可使用商购试剂采用标准方法进行构建，所述试剂如购自 Invitrogen, San Diego, CA。用含有 NotI 位点的寡 dT 引导
15 cDNA，将其钝端与 SalI 半激酶化衔接子连接，用 NotI 裂解，用凝胶电泳分出适当大小，并按规定方向克隆入适合的克隆载体中独一无二的 XhoI 和 NotI 位点(如 pRKB 或 pRKD; pRK5B 是不含 SfiI 位点的 pRKSD 前体;见 Holmes 等, Science, 253: 1278-1280(1991))处。

对上述分离的克隆进行 DNA 测序，得到全长 FGF-19 多肽的全长 DNA
20 序列(本文定为 DNA49435-1219 [图 1, SEQ ID NO: 1])和 FGF-19 多肽的衍生的蛋白序列。

上述鉴定的全长克隆含有单个开放阅读框，其在核苷酸 464-466 处有明显的翻译起始位点，并在核苷酸 1112-1114 处有终止信号(图 1, SEQ ID NO: 1)。该预测的多肽前体长 216 个氨基酸，理论分子量约 24003 道尔顿，估计
25 pI 约 6.99。示于图 2(SEQ ID NO: 2)的全长 FGF-19 序列的分析证明，图 2 所示序列中存在多个重要的多肽结构域，其中那些重要的多肽结构域的位置与上面所述是接近的。染色体绘图证明，FGF-19-编码核酸对应于人染色体 11q13.1, q13.1 带。克隆 DNA49435-1219 已经于 1997 年 11 月 21 日在 ATCC 保藏，ATCC 保藏号为 209480。

30 Dayhoff 数据库(version 35.45 SwissProt 35)分析，利用了对图 2(SEQ ID NO: 2)所示全长序列的 ALIGN-2 序列对比分析，其证明 FGF-19 氨基酸序列

和下述 Dayhoff 序列之间具有序列同一性: AF007268_1、S54407、P_W52596、FGF2_XENLA、P_W53793、AB002097_1、P_R27966、HSU67918_1、S23595 和 PR_70824。

实施例 2

5 FGF-19 作为杂交探针的应用

下述方法描述编码 FGF-19 的核苷酸序列作为杂交探针的应用。

包括全长或成熟 FGF-19 的编码序列的 DNA, 用作筛选人组织 cDNA 文库或人组织基因组文库中同源 DNA(如编码 FGF-19 的天然变体的那些)的探针。

10 在高度严格条件下进行杂交, 并洗涤含有任何一个文库 DNA 的膜。得自 FGF-19 的放射标记探针与滤膜的杂交, 是于 42°C 在下述溶液中进行的: 50% 甲酰胺, 5x SSC, 0.1% SDS, 0.1% 焦磷酸钠, 50 mM 磷酸钠, pH 6.8, 2x Denhardt's 溶液, 和 10% 葡聚糖硫酸酯, 杂交 20 小时。滤膜洗涤则是于 42°C 在 0.1x SSC 和 0.1% SDS 的水溶液中进行。

15 然后, 使用本领域已知标准技术, 可以鉴定那些相对于编码全长天然序列 FGF-19 的 DNA 具有预期序列同一性的 DNA。

实施例 3

FGF-19 在大肠杆菌中的表达

20 该实施例描述通过在大肠杆菌中的重组表达制备未糖基化形式的 FGF-19。

使用选定的 PCR 引物, 对编码 FGF-19 的 DNA 序列进行初步扩增。引物应当含有对应于所选表达载体中限制性酶切位点的限制性内切酶位点。可使用各种表达载体。适宜载体的一个实例是 pBR322(得自大肠杆菌; 见 Bolivar 等, Gene, 2: 95(1977)), 它含有氨苄青霉素和四环素抗药性基因。该载体用限制性内切酶消化并脱磷酸化。然后把 PCR 扩增的序列连接到该载体中。载体优选包括下列序列: 编码抗生素抗性基因的序列、trp 启动子、多聚组氨酸前导序列(包括第一个 6 STII 密码子、多聚组氨酸序列和肠激酶裂解位点)、FGF-19 编码区、 λ 转录终止子和 argU 基因。

30 接着, 使用 Sambrook 等, 出处同上描述的方法, 用连接混合物转化所选大肠杆菌株。转化体通过其在 LB 平皿上的生长能力来鉴定, 然后选出抗生素抗性菌落。分离质粒 DNA, 并用限制性分析和 DNA 测序确认。

选出的克隆可在液体培养基(如添加了抗生素的 LB 肉汤)中生长过夜。随后, 过夜培养物用于接种更大规模的培养。然后细胞生长至所需光密度, 在此期间表达启动子被打开。

5 培养细胞数小时后, 离心收获细胞。使用本领域已知的各种试剂可以溶解离心所得细胞小团, 然后, 使用金属螯合柱在允许紧密结合蛋白的条件下纯化已溶解的 FGF-19 蛋白。

经下述操作, FGF-19 可在大肠杆菌以多聚组氨酸标记形式表达。使用选定的 PCR 引物, 初步扩增编码 FGF-19 的 DNA。引物含有对应于所选表达载体中限制性酶切位点的限制性酶切位点和其它对于提供有效、可靠翻译起始、在金属螯合柱上快速纯化、并用肠激酶进行蛋白水解去除有用的序列。10 接下来, 将 PCR 扩增的多聚组氨酸标记的序列连接到表达载体中, 后者用于转化基于菌株 52 的大肠杆菌宿主 (W3110 fuhA(tonA)lon galE rpoHts(htpRts)clpP(lacIq)。转化体首先于 30°C 在含有 50 mg/ml 羧苄青霉素的 LB 中振荡培养, 直至 O. D.600 达到 3-5。接着, 把培养物用 CRAP 培养基(在 15 500 mL 水中混合 3.57 g(NH₄)₂SO₄, 0.71 g 柠檬酸钠•2H₂O, 1.07 g KCl, 5.36 g Difco 酵母提取物, 5.36 g Sheffield hycase SF, 以及 110 mM MPOS、pH 7.3、0.55% (w/v)葡萄糖和 7 mM MgSO₄)稀释 50-100 倍, 于 30°C 振荡培养约 20-30 小时。取样用 SDS-PAGE 分析证实表达, 把大批量培养物离心成细胞团。冷冻细胞团直至纯化和重折叠(refolding)。

20 在 10 倍体积(w/v)的 7M 胍、20 mM Tris、pH 8 缓冲剂中, 重新悬浮得自 0.5~1 L 发酵物(6-10 g 小团)的大肠杆菌糊状物。加入固体亚硫酸钠和连四硫酸钠使终浓度分别为 0.1M 和 0.02 M, 在 4°C 搅拌溶液过夜。该步骤产生变性蛋白, 后者的所有半胱氨酸残基被 sulfitolization 封闭。溶液在 Beckman 超速离心机(Ultracentrifuge)上 40000 rpm 离心 30 分钟。上清液用 3-5 倍体积 25 的金属螯合柱缓冲液(6M 胍, 20 mM Tris, pH 7.4)稀释, 并通过 0.22 微米滤器过滤以使上清液澄清。将澄清的提取液上样至已用金属螯合柱缓冲液平衡的 5 ml Qiagen Ni-NTA 金属螯合柱上。用另一含有 50 mM 咪唑(Calbiochem, Utrol grade)的 pH 7.4 缓冲液洗柱。用含 250 mM 咪唑的缓冲液洗脱蛋白。汇集含有所需蛋白的级分并于 4°C 贮存。利用基于蛋白氨基酸序列计算出的消光(extinction)系数, 通过 280 nm 处的吸收度估算蛋白浓度。 30

将样品用新鲜配制的重折叠缓冲液缓慢稀释, 从而使蛋白重折叠, 所述

重折叠缓冲液的构成是：pH 8.6 20 mM Tris、0.3 M NaCl、2.5 M 尿素、5 mM 半胱氨酸、20 mM 甘氨酸和 1 mM EDTA。选择重折叠液的体积以使蛋白终浓度为 50~100 微克/ml。于 4°C 轻轻搅拌重折叠液 12-36 小时。加入 TFA 至终浓度为 0.4% (pH 约为 3) 来淬灭重折叠液。进一步纯化蛋白质之前，溶液
5 通过 0.22 微米滤器过滤，加入乙腈至终浓度为 2-10%。在 Poros R1/H 反相柱上对重折叠的蛋白进行层析分析，使用的流动相缓冲液为 0.1% TFA，用 10~80% 乙腈梯度洗脱。有 A280 处吸收的等份试样在 SDS 聚丙烯酰胺凝胶上分析，收集含有均匀重折叠的蛋白的级分。通常，大多数正确重折叠的蛋白在最低乙腈浓度时洗脱下来，这是由于这些类蛋白是最紧密的，它们的疏水内部被掩蔽因而避免与反相树脂相互作用。聚集型蛋白通常在高乙腈浓度时洗脱下来。除了能将所需形式的蛋白与错误折叠的蛋白区分开来，该反相步骤还可从样品中去除内毒素。

收集含有所需折叠的 FGF-19 多肽的级分，用氮气流直接轻吹溶液而去除乙腈。用 0.14 M 氯化钠和 4% 甘露醇经透析或通过用配制缓冲液平衡的
15 G25 Superfine(Pharmacia)树脂进行凝胶过滤，而将蛋白配制成 pH 6.8 的 20 mM HEPES 溶液，并进行过滤除菌。

实施例 4

FGF-19 在哺乳动物细胞中的表达

该实施例描述通过在哺乳动物细胞中的重组表达制备潜在糖基化形式
20 的 FGF-19。

载体 pRK5(见 EP 307247, 公开日为 1989 年 3 月 15 日)用作表达载体。任选地，使用诸如 Sambrook 等，出处同上中描述的连接方法，将 FGF-19DNA 连接至具有选定的限制酶的 pRK5 中，所述限制酶允许插入 FGF-19DNA。由此产生的载体称为 pRK5-FGF-19。

25 在一实施方案中，所选宿主细胞是 293 细胞。人 293 细胞(ATCC CCL 1573)在组织培养板上的培养基中生长至融合，所述培养基如添加了胎牛血清和任选的营养成分和/或抗生素的 DMEM。约 10 μ g pRK5-FGF-19DNA 与约 1 μ g 编码 VA RNA 基因[Thimmappaya 等, Cell, 31: 543(1982)]的 DNA 混合，溶于 500 μ l 1 mM Tris-HCl、0.1 mM EDTA、0.227 M CaCl₂ 中。向该混
30 合物中滴加 500 μ l 50 mM HEPES(pH 7.35)、280 mM NaCl、1.5 mM NaPO₄，25°C 10 分钟形成沉淀。悬浮该沉淀物，加入 293 细胞，37°C 静置约 4 小时。

吸出培养基,用30秒的时间加入20%甘油的PBS溶液2ml。然后用不含血清培养基洗涤293细胞,加入新鲜培养基,保温细胞约5天。

转染后约24小时,弃除培养基,代之以培养基(单独)或含有 $200\mu\text{Ci/ml}$ ^{35}S -半胱氨酸和 $200\mu\text{Ci/ml}$ ^{35}S -蛋氨酸的培养基。保温12小时后,收集条件
5 培养液,在旋转滤器上浓缩,上样至15%SDS凝胶上。电泳后使凝胶干燥,并于胶片上曝光一段时间,从而揭示FGF-19多肽的存在。可进一步温育含有转染型细胞的培养物(在无血清培养基中),在所选择的生物鉴定法中检测培养物。

在另一技术中,使用Sompanyrac等在Proc. Natl. Acad. Sci., 12:
10 7575(1981)中描述的硫酸葡聚糖法,可将FGF-19瞬时引入293细胞。293细胞在旋转烧瓶中生长至最大密度,加入 $700\mu\text{g}$ pRK5-FGF-19DNA。采自旋转烧瓶的细胞通过离心进行第一次浓缩,用PBS洗涤。在细胞小团上将DNA-葡聚糖沉淀物温育4小时。细胞用20%甘油处理90秒,用组织培养基洗涤,再次装入含有组织培养基, $5\mu\text{g/ml}$ 牛胰岛素和 $0.1\mu\text{g/ml}$ 牛运铁蛋白的旋转
15 烧瓶中。约4天后,离心该调整的培养液,过滤去除细胞和碎片。然后,含有表达的FGF-19的样品可用任何所选择的方法,如透析和/或柱层析进行浓缩和纯化。

在另一实施方案中,FGF-19可在CHO细胞中表达。使用已知试剂如 CaPO_4 或DEAE葡聚糖,可将pRK5-FGF19转染入CHO细胞。如上所述,
20 温育细胞培养物,培养基用培养基(单独)或含有放射标记物如 ^{35}S -蛋氨酸的培养基替换。测定了FGF-19多肽的存在后,可用无血清培养基替换该培养基。优选地,培养物温育约6天后,收获调整的培养液。接着,用任何选择的方法浓缩并纯化该含有表达的FGF-19的培养液。

也可在宿主CHO细胞中表达表位-标记的FGF-19。FGF-19可以从pRK5
25 载体中亚克隆出来。可对该亚克隆插入物进行PCR,使之与选择的表位标记如多聚组氨酸-标记融合在杆状病毒表达载体上的同一读码框中。接着,可将多聚组氨酸标记的FGF-19插入物亚克隆入含有选择标志(如用于选择稳定克隆的DHFR)的SV40驱动型载体中。最后,用该SV40驱动型载体转染CHO细胞(如上所述)。可按上述进行标记,以证实表达。然后,用任何选择
30 的方法,如用 Ni^{2+} -螯合剂亲和层析法,可以浓缩并纯化含有已表达的多聚组氨酸标记的FGF-19的培养基。

还可通过瞬时表达方法在 CHO 和/或 COS 细胞中表达 FGF-19, 或通过另一稳定表达方法在 CHO 细胞中表达 FGF-19。

在 CHO 细胞中的稳定表达可使用如下方法。将蛋白表达为 IgG 构建体(免疫粘附素), 其中编码每种蛋白的可溶形式(例如细胞外结构域)的序列与 IgG1 恒定区序列融合, 所述 IgG1 恒定区包括铰链区、CH2 区和 CH2 区和/或为多聚组氨酸标记的形式。

PCR 扩增后, 使用 Ausubel 等在 Current Protocols of Molecular Biology. Unit 3.16, John Wiley and Sons(1997)中所述标准技术, 将各个 DNA 亚克隆入 CHO 表达载体中。构建 CHO 表达载体, 使其具有目标 DNA 的 5'和 3'相容性限制酶切位点, 以便于 cDNA 的往返穿梭。在 CHO 细胞中使用载体的表达在 Lucas 等, Nucl. Acids Res. 24: 9(1774-1779)(1996)中有描述, 并且使用 SV40 早期启动子/增强子来驱动目标 cDNA 和二氢叶酸还原酶(DHFR)的表达。DHFR 表达允许选择转染后维持稳定的质粒。

使用商购转染试剂 Superfect®(Quiagen)、Dospere® 或 Fugene®(Boehringer Mannheim), 将 12 微克所需质粒 DNA 引入约 1 千万个 CHO 细胞。按 Lucas 等, 出处同上所述方法培养细胞。在每一安瓿中冷冻约 3×10^7 个细胞以备如下所述进一步生长和增殖。

把含有质粒 DNA 的安瓿放入水浴中解冻并涡旋混合。用移液管将内容物移入盛有 10 mL 培养基的离心管中, 1000 rpm 离心 5 分钟。吸出上清液, 用 10 mL 选择性培养基(0.2 μ m 过滤的 PS20 和 5% 0.2 μ m 渗滤的胎牛血清)悬浮细胞。然后将细胞等分至含有 90 mL 选择性培养基的 100 mL 旋转器中。1-2 天后, 将细胞转移至盛有 150 mL 选择性生长培养基的 250 mL 旋转器中, 并于 37°C 温育。再过 2-3 天, 在 250 mL、500 mL 和 2000 mL 旋转器中接种 3×10^5 细胞/mL。离心后用新鲜培养基替换细胞培养基, 在生产培养基中重悬浮。尽管可以使用任何适宜的 CHO 培养基, 但是实际使用的是 1992 年 1 月 16 日颁布的美国专利 5122469 中所述生产培养基。3L 生产旋转器中接种 1.2×10^6 细胞/mL。在 0 天, 测定细胞数 pH ic。第 1 天, 从旋转器取样, 开始喷雾过滤的空气。第 2 天, 从旋转器中取样, 温度调至 33°C, 加 500 g/L 葡萄糖 30 mL 和 10% 止泡剂 0.6 mL(例如 35% 聚二甲基硅氧烷乳液, Dow Corning 365 Medical Grade Emulsion)。整个生产过程中, 必要时调整 pH 使之保持在 7.2 左右。10 天后, 或直至活力降至 70% 以下时, 离心收获细胞

培养物并通过 0.22 μm 滤膜过滤。滤出液贮于 4 $^{\circ}\text{C}$ 或立即上柱纯化。

对于多聚组氨酸标记的构建体, 使用 Ni-NTA 柱(Qiagen)纯化蛋白。纯化前, 向调整的培养液中加入咪唑至浓度为 5 mM。于 4 $^{\circ}\text{C}$, 将该调整的培养液以 4-5 ml/min 速度泵到 6 ml Ni-NTA 柱上, 后者已用含 0.3 M NaCl 和 5 mM 咪唑的缓冲液 20 mM Hepes、pH 7.4 平衡。装柱后, 用另外的平衡缓冲液洗柱, 用含有 0.25 M 咪唑的平衡缓冲液洗脱蛋白。随后, 用 25 ml G25 Superfine(Pharmacia)柱使高度纯化的蛋白脱盐至含有 10 mM Hepes、0.14 M NaCl 和 4% 甘露醇 pH 6.8 的贮藏缓冲液中, 并贮于 -80 $^{\circ}\text{C}$ 。

按照下述步骤, 从调整的培养液中纯化免疫粘附素(含 Fc)构建体。将调整的培养液泵入 5 ml 蛋白 A 柱(Pharmacia), 后者已经用 pH 6.8 的 20 mM 磷酸钠缓冲液平衡。装柱后, 在用 pH 3.5 100 mM 柠檬酸洗脱前, 先用平衡液彻底洗柱。通过将 1 ml 级分集中到含有 275 μL pH 9 1M Tris 缓冲液的试管中, 而将洗脱的蛋白立即中和。接着, 按照前面针对 poly-his 标记的蛋白的描述将高度纯化的蛋白脱盐至贮存缓冲液中。用 SDS 聚丙烯酰胺凝胶电泳和用 Edman 降解技术进行 N 末端氨基酸测序, 来测定同质性。

实施例 5

FGF-19 在酵母中的表达

下述方法方法描述 FGF-19 在酵母中的重组表达。

首先, 构建酵母表达载体, 以使 FGF-19 受 ADH2/GAPDH 启动子的驱动而在细胞内产生或分泌。将编码 FGF-19 的 DNA 和启动子插入所选质粒的适宜限制性酶切位点, 以引导 FGF-19 的细胞内表达。为了分泌 FGF-19, 可将编码 FGF-19 的 DNA 克隆入所选质粒中, 与 DNA 一起插入的还有编码 ADH2/GAPDH 启动子的 DNA、天然 FGF-19 信号肽或哺乳动物其它信号肽, 或者例如酵母 α -因子或转化酶分泌信号/前导序列以及连接序列(如果需要)以使 FGF-19 表达。

然后, 可用上述表达质粒转化酵母细胞如酵母株 AB 110, 并在选择的发酵培养基中培养。通过用 10% 三氯醋酸沉淀和用 SDS-PAGE 分离, 接着用考马斯蓝染色凝胶, 可以检测转化酵母的上清液。

随后, 经离心而从发酵培养基中除去酵母细胞, 接着使用选择的柱式过滤器浓缩培养基, 由此而分离和纯化重组 FGF-19。使用选择的柱层析树脂, 可以进一步纯化含 FGF-19 的浓缩物。

实施例 6

FGF-19 在杆状病毒-感染的昆虫细胞中的表达

下述方法描述重组 FGF-19 在杆状病毒-感染的昆虫细胞中的表达。

FGF-19 的编码序列融合在杆状病毒表达载体所含的表位标记的上游。

- 5 这类表位标记包括多聚组氨酸标记和免疫球蛋白标记(如 IgG 的 Fc 区)。可使用各种质粒, 包括商购质粒如 pVL1393(Novagen)。简言之, 使用与 5'和 3'区互补的引物经 PCR 扩增 FGF-19 的编码序列或 FGF-19 编码序列的所需部分如编码跨膜蛋白胞外结构域的序列或如果蛋白是胞外蛋白, 编码成熟蛋白的序列。5'引物可添加侧翼(选择的)限制性酶切位点。然后用那些选择的
- 10 限制性内切酶消化产物并亚克隆入表达载体中。

重组杆状病毒是这样产生的: 使用脂质转染试剂(购自 GIBCO-BRL), 将上述质粒和 BaculoGold™ 病毒 DNA(Pharming)共-转染入草地夜蛾 (*Spodoptera frugiperda*)("Sf9")细胞(ATCC CRL 1711)中。28℃温育 4-5 天后, 收获释放出的病毒, 用于进一步扩增。病毒性感染和蛋白表达按照 O'Reilley

15 等在 *Baculovirus Express Vectors: A Laboratory Manual*, Oxford: Oxford University Press(1994)中的描述进行。

接下来, 可纯化表达的多聚组氨酸标记的 FGF-19, 例如, 通过如下 Ni²⁺-螯合亲和层析法进行。重组病毒-感染的 Sf9 细胞的提取物的制备, 按照 Rupert 等在 *Nature*, 362: 175-179(1993)中所述方法进行。简要讲, 洗涤 Sf9

20 细胞, 在超声缓冲液(25 mL HEPES, pH 7.9; 12.5 mM MgCl₂; 0.1 mM EDTA; 10%甘油; 0.1%NP-40; 0.4 M KCl)中重悬浮, 冰浴中超声 2 次, 各 20 秒钟。离心清除超声后的产物, 上清液用填充缓冲液(50 mM 磷酸盐, 300 mM NaCl, 10%甘油, pH 7.8)50 倍稀释, 经 0.45 μm 滤器过滤。制备 Ni²⁺-NTA 琼脂糖柱(购自 Qiagen), 床体积 5 mL, 用 25 mL 水洗涤, 用 25 mL 填充缓冲液平

25 衡。以每分钟 0.5 mL 的速度将已过滤的细胞提取液上柱, 用填充缓冲液洗柱至 A₂₈₀基线, 在此点开始收集级分。接着, 用第二洗涤缓冲液(50 mM 磷酸盐; 300 mM NaCl, 10%甘油, pH 6.0)洗柱, 此过程洗脱出非特异结合的蛋白。重又达到 A₂₈₀基线后, 用第二洗涤缓冲液中 0~500 mM 的咪唑梯度洗柱。收集 1 mL 级分, 用 SDS-PAGE 和银染色或用碱性磷酸酶(Qiagen)偶

30 联型 Ni²⁺-NTA 进行 Western 印迹, 进行分析。汇总含洗脱的组氨酸¹⁰-标记的 FGF-19 的级分, 针对填充缓冲液透析。

或者,使用已知层析技术,包括例如蛋白 A 或蛋白 G 柱层析法,对 IgG 标记(或 Fc 标记)的 FGF-19 进行纯化。

实施例 7

制备结合 FGF-19 的抗体

5 该实施例描述能够与 FGF-19 特异结合的单克隆抗体的制备。

生产单克隆抗体的技术为本领域技术人员所已知,并在例如 Goding, 出出处同上中有描述。可以使用的免疫原包括纯化的 FGF-19、含有 FGF-19 的融合蛋白,和在细胞表面表达重组 FGF-19 的细胞。无需繁琐实验,本领域熟练技术人员就可对免疫原作出选择。

10 小鼠如 Balb/c 经皮下或腹腔内注射 1-100 微克在完全 Freund's 佐剂中乳化的 FGF-19 免疫原来免疫。或者,免疫原在 MPL-TDM 佐剂(Ribi Immunochemical Research, Hamilton, MT)中乳化,注射入动物的后足垫。10~12 天后,再次用在所选佐剂中乳化的免疫原对免疫后的鼠进行加强免疫。此后数周,还可增加免疫注射而对小鼠加强免疫。经后-眼眶放血方式定期
15 采集小鼠血清样品,进行 ELISA 分析,以检测抗-FGF-19 的抗体。

已经检测到适宜的抗体滴度后,对抗体“阳性”动物最后一次静脉注射 FGF-19。3 或 4 天后,处死小鼠,取其脾细胞。然后将脾细胞与所选鼠骨髓瘤细胞系(如 P3 X 63 AgU.1, 得自 ATCC CRL 1597 号)融合(使用 35% 聚乙二醇)。将融合产生的杂交瘤细胞接种到 96 孔组织培养板上,所述培养板上含
20 有 HAT(次黄嘌呤、氨基蝶呤和胸腺嘧啶核苷)培养基以抑制非-融合细胞、骨髓杂合体 and 脾细胞杂合体的增殖。

根据 ELISA 中的抗 FGG-19 活性来筛选杂交瘤细胞。筛选能分泌所需抗 FGF1-19 单克隆抗体的“阳性”杂交瘤细胞的测定,是本领域常识。

25 可给同系 Balb/c 小鼠腹腔注射阳性杂交瘤细胞,以产生含有抗-FGF-19 单克隆抗体的腹水。或者,使杂交瘤细胞在组织培养烧瓶或滚瓶中生长。使用硫酸铵沉淀,继之用凝胶排阻层析法,可以完成对腹水中产生的单克隆抗体的纯化。或者,使用基于抗体与蛋白 A 或蛋白 G 的结合的亲和层析法。

实施例 8

使用特异抗体纯化 FGF-19 多肽

30 可使用蛋白纯化领域的各种标准技术,进行天然或重组 FGF-19 多肽的纯化。譬如,使用目的 FGF-19 多肽的特异性抗体的免疫亲和层析法可纯化

原-FGF-19多肽、成熟 FGF-19多肽或前-FGF-19多肽。总体上讲，免疫亲和柱是通过将抗-FGF-19多肽的抗体与活化的层析树脂共价偶联而构建的。

使用硫酸铵沉淀或者通过在固相化蛋白 A(Pharmacia LKB Biotechnology, Piscataway, N. J.)上的纯化，可以从免疫血清制备多克隆免疫球蛋白。同样地，用硫酸铵沉淀或在固相化蛋白 A 上的层析，可从小鼠腹水中制备单克隆抗体。将部分纯化的免疫球蛋白共价附着于层析树脂如 CnBr-活化的 SEPHAROSETM(Pharmacia LKB Biotechnology)。抗体与树脂共价偶联，从而封闭树脂，按照制造商的指示洗脱所产生的树脂。

通过制备含有可溶性 FGF-19多肽的细胞级分，可在 FGF-19多肽的纯化中使用此免疫亲和柱。该制备通过溶解经加入去污剂的差异离心或使用本领域公知方法获得的完整细胞或亚细胞级分而实现。或者，含有信号序列的可溶性 FGF-19多肽可以以有效量分泌到细胞生长所在的培养基中。

使含有可溶性 FGF-19多肽的制剂流经免疫亲和柱，在使得优先吸收 FGF-19多肽的条件下(例如，存在去污剂时的高离子强度缓冲液)洗柱。然后，在使抗体/FGF-19多肽的结合分裂的条件下(例如，低 pH 缓冲液如 pH 约 2-3，或高浓度离液剂如尿素或硫氰酸盐离子)洗脱，并收集 FGF-19多肽。

实施例 9

药物筛选

通过在多种药物筛选技术中使用 FGF-19多肽或其结合片段，本发明对于筛选化合物特别有用。这类检测所用 FGF-19多肽或片段可以是溶液中的游离形式，也可以是固定在固体载体上、产生于细胞表面或位于细胞内。药物筛选方法之一是利用已被表达 FGF-19多肽或片段的重组核酸稳定转化的真核或原核宿主细胞。在竞争性结合试验中筛选抗此类转化细胞的药物。活细胞或是固定形式的此类细胞，无论为有活性的或已固定的，均可用于标准结合试验。可以测定例如 FGF-19多肽或其片段与被测试剂之间的复合物的形成。或者，可以检验由被测试剂造成的 FGF-19多肽与其靶细胞或靶受体之间复合物形成的减少。

因此，本发明提供筛选药物或任何其它可影响 FGF-19多肽-相关疾病或病症的试剂的方法。这些方法包括将这类试剂与 FGF-19多肽或其片段接触，并使用本领域熟知方法分析(i)该试剂与 FGF-19多肽或片段之间复合物的存在，或(ii)FGF-19多肽或片段与细胞之间的复合物的存在。在此类竞争性结

合试验中, FGF-19 多肽或片段通常被标记。适宜温育之后, 游离 FGF-19 多肽或片段与结合形式 FGF-19 多肽或片段分开, 游离或未形成复合物的标记物的量, 可以衡量特定试剂与 FGF-19 多肽的结合能力或者对 FGF19 多肽/细胞复合物的阻碍能力。

- 5 用于药物筛选的另一技术提供高效率筛选具有针对多肽的适宜结合亲和力的化合物, 详细描述见 1984 年 9 月 13 日公开的 WO 84/03564 中。简要讲, 在固体基底(如塑料钉或其它表面)上合成大量不同的小肽检测化合物。如用于 FGF-19 多肽的那样, 将肽检测化合物与 FGF-19 多肽反应并洗涤。用本领域熟知方法检测结合的 FGF-19 多肽。也可将纯化的 FGF-19 多肽直接包被在前述药物筛选技术所用的板上。此外, 非-中和抗体可用于捕获所述肽并将所述肽固定在固体载体上。

- 10 本发明也考虑了使用竞争性药物筛选试验, 在所述试验中, 能够与 FGF-19 多肽特异性结合的中和抗体与被测化合物竞争地同 FGF-19 多肽或其片段结合。以这种方式, 抗体可用于检测与 FGF-19 多肽具有一个或多个相同表位的任何肽的存在。

15 实施例 10

合理的药物设计

- 合理药物设计的目标是生产目标生物活性多肽(即 FGF-19 多肽)的结构类似物或与多肽相互作用的小分子例如激动剂、拮抗剂或抑制剂。其中任何实例可用于形成药物, 所述药物是更具活性或更稳定形式的 FGF-19 多肽, 或者药物增强或干扰 FGF-19 多肽体内功能(参见 Hodgson, Bio/Technology, 9: 19-21(1991))。

- 25 在一种方法中, 用 x-射线晶体照相术、计算机模型设计或者最通常这二者的结合, 来测定 FGF-19 多肽或 FGF-19 多肽-抑制剂复合物的三维结构。为描述结构和测定分子的活性位点, 必须确定 FGF-19 多肽的形状和电荷。不常用的方法是, 通过基于同源蛋白结构的模型, 来获取关于 FGF-19 多肽结构的有用信息。两种情况均使用相关的结构信息来设计类似的 FGF-19 多肽样分子或用于鉴定有效抑制剂。合理药物设计的有用实例包括 Braxton 和 Wells, Biochemistry, 31: 7796-7801(1992)所证实的具有改进活性或稳定性的分子, 或者 Athauda 等, J. Biochem., 113: 742-746(1993)所述作为天然肽的抑制剂、激动剂或拮抗剂的分子。

还有一种可能是分离靶特异抗体，通过如上所述的功能分析进行筛选，然后解决其晶体结构。大体而言，这种途径产生药物母核(pharmacore)，可基于该母核设计药物。通过产生针对功能性、药理活性抗体的抗-独特型抗体(抗-独特型)，可完全绕开蛋白晶体构象。作为镜象的镜象，预计抗-独特型抗体的结合位点是原始受体的类似物。因此，抗-独特型抗体可用于从化学或生物学产生的肽库中鉴定和分离肽。随后，分离的肽可用作药物母核。

本发明可提供足够量的 FGF-19 多肽来完成诸如 X-线晶体照相术此类的分析研究。另外，本文提供的 FGF-19 多肽氨基酸序列的信息，将对那些使用计算机模型技术代替 x-射线晶体照相术或除使用 x-射线晶体照相术之外还使用计算机模型技术的人员提供指导。

实施例 11

FGF-19 转基因鼠的体重、Leptin 水平、摄食量、排尿量、耗氧量以及甘油三酯和游离脂肪酸水平的研究

如上所述，FGF-19 最近被确定为与成纤维细胞生长因子有关的分泌蛋白生长因子家族中的一员。本文中 FGF-19 的特征是与 FGF 受体 4 相互作用但不表现为促分裂原。为进一步研究该蛋白的功能，生产了表达人 FGF-19 的转基因鼠。

更具体而言，将编码人 FGF-19 的 cDNA 克隆入含有肌球蛋白轻链启动子的质粒中。该启动子足以使该转基因发生肌肉特异性转录。FGF-19cDNA 的 5'端也包括剪接受体和供体以增加表达水平，FGF-19cDNA 的 3'端包括剪接受体和供体以及多聚腺苷酸附加信号，以增加转录水平和提供转录终止位点。

使用适宜的限制性内切酶，从细菌载体序列中释放出包括 MLC 启动子、5'剪接受体和供体、FGF-19cDNA，3'剪接受体和供体以及转录终止位点的 DNA(转基因)，随后在琼脂糖凝胶上通过分级分离而纯化。纯化的 DNA 注射入受精鼠卵的一个原核中，产生转基因鼠，并按文献所述方法进行鉴定 (Genetic Modification of Animals; Tim Stewart; In Exploring Genetic Mechanisms 第 565-598 页; 1997 年 M Singer 和 P Berg 编著; University Science Books; Sausalito, Calif)。小鼠长至 6 周龄后如下测量水摄取、食物消耗、排尿量和血细胞比容。相同小鼠长至 8 周龄后测量 leptin、甘油三酯和游离脂肪酸。

由下述结果讨论所见，这些鼠的耗氧率表明其食物摄取和代谢率增加。尽管摄食增加，但是这些鼠的体重较其它未转基因同窝鼠的体重显著减少。体重降低似乎是肥胖症减轻的结果，因为 leptin 与人和啮齿动物的脂肪组织的量密切相关，而 leptin 在转基因鼠中是减少的。为更进一步支持这一点，

5 从鼻子到臀部的长度测量评定表明，转基因鼠呈常态线性增长。至于体温、身长(骨长)和血液学检测值，这些动物均正常。与增加摄食一致的是，转基因鼠的排尿量增加。由于这些鼠并没有表现出饮用更多水，而且经正常血细胞比容测定证实也没有表现脱水，那么增加的排尿量可能是源自对增加的摄食的代谢。由于 FGF-19 减少肥胖而不改变肌肉质量或长骨形成，所以表明

10 FGF-19 是肥胖和相关病症治疗中的有效治疗剂。

更具体地，在不同禁食和摄食条件下，在各种时间点称量 MLC-FGF-19 转基因鼠的体重。特别是，6 周龄时对各组雌性 FGF-19 转基因鼠和其非-转基因同窝鼠，在随意进食、禁食 6 小时和 24 小时以及结束 24 小时禁食后 24 小时进行称重。如图 3A 所示，所有条件下，FGF-19 转基因鼠(实心棒)

15 体重低于其野生型、未转基因同窝鼠(点彩棒)。

图 3B 显示对图 3A 所示同组鼠的血清中 leptin 的分析。FGF-19 转基因小鼠中降低的 leptin 与其由于肥胖症减轻而致的体重减轻(图 3A)是一致的。

监测一组 6 周龄转基因鼠的摄食(图 4A)、摄水(图 4B)、排尿(图 4C)和血细胞比容(图 4D)情况。如图所见，FGF-19 转基因鼠(实心棒)较野生型同窝鼠消耗更多食物，但摄水量并不增加。尽管水消耗量没有变化，但是转基因鼠产尿量增加(图 4C)。尽管产尿量增加，但是通过正常血细胞比容可以证明转基因鼠并未出现脱水(图 4D)。

20

可以用代谢速率增加来解释体重下降(图 3)伴有食物消耗增加(图 4)。通过测定耗氧量来测定代谢速率。如图 5 所示，在禁食 24 小时和在结束 24 小时禁食后 24 小时的明、暗光周期中，FGF-19 转基因鼠的代谢速率均见增加。

25

肥胖和甘油三酯与游离脂肪酸的升高是心血管疾病的危险因素。由于 FGF-19 减少心血管疾病危险因素中的一个因素(肥胖(图 3))，因此，对于 FGF-19 是否也降低其它危险因素也进行了研究。如图 6 所见，FGF-19 转基因鼠中甘油三酯和游离脂肪酸(FFA)水平也下降。

30 实施例 12

FGF-19 输注导致食物摄取增加和耗氧量增加

为了证实在 FGF-19 转基因鼠所见到的效果是由 FGF-19 蛋白引起，通过渗透压驱动型植入泵给药方式，给各组非-转基因 FvB 鼠输入重组 FGF-19(1 mg/kg/日，静脉)。如图 7A-B 所示，与单纯输入载液的鼠相比，施用重组人 FGF-19 引起摄食增加。而且，耗氧量测定显示 FGF-19 输入还导致代谢速率增加。

实施例 13

FGF-19 降低葡萄糖摄取和增加 Leptin 自脂肪细胞释放

为进一步研究 FGF-19 改变代谢的机制，将重组人 FGF-19 加入到大鼠脂肪细胞原代培养中，并检测葡萄糖摄取和 leptin 自这些细胞释放的情况。如图 8A、B 所示，FGF-19 增加原代大鼠脂肪细胞中 leptin 的释放，并减少葡萄糖摄入。

实施例 14

高脂饮食 FGF-19 转基因鼠的葡萄糖耐量和脂肪垫重量的研究

通常，高脂饮食的鼠(和人)体重将会增加并变得肥胖，而且葡萄糖耐量下降或者患上糖尿病。为了检验接受 FGF-19 是否对肥胖和葡萄糖耐量有影响，基本上按照 Rebuffe-Scrive 等在 Metabolism 1993 年第 42 卷第 11 期第 405-1409 页和 Surwit 等在 Metabolism 1995 年第 44 卷第 5 期第 645-651 页中描述的方法，只是将钠含量改变为符合正常饮食标准(饮食来自 Research Diets Inc. Catalog no. D12330N)，使一群转基因小鼠和其未转基因的同窝鼠(鼠龄和性别相当)进食高脂饮食。

正常鼠粮或高脂饮食喂养 10 周后，小鼠(雌性转基因鼠和其未转基因同窝鼠)接受葡萄糖耐量测试。这样每只鼠腹腔内注射 1.0 mg 葡萄糖/kg 体重，注射后，按照一定时间间隔测定血液中葡萄糖浓度。图 10 中的曲线显示鼠体内葡萄糖水平，证明 8/9 的进食高脂饮食的雌性未转基因鼠被确定为糖尿病(2 小时葡萄糖水平大于 200 mg/dl;(World Book of Diabetes in Practice. Vo] 3; Ed Krall, L. P.; Elsevier))，而 0/5 的进食相同饮食的转基因鼠患糖尿病。

高脂饮食喂食雄性鼠 6 或 10 周后处死动物，通过测量特定脂肪积存的重量来测定肥胖程度。如图 9 所示，进食高脂饮食的转基因鼠较未转基因同窝鼠具有显著少的脂肪。

材料保藏

下述材料已经在美国典型培养物保藏中心(the American Type Culture

Collection, 10801 University Blvd., Manassas, VA 20110-2209, USA)(ATCC)保藏:

<u>材料</u>	<u>ATCC 保藏号</u>	<u>保藏日</u>
DNA49435-1219	209480	1997.11.21

5 该保藏是按照国际承认的用于专利程序的微生物保藏的布达佩斯条约及其章程(布达佩斯条约)进行的。这确保自保藏日起 30 年的时间内保藏培养物保持存活。这使得通过 ATCC 可以获得在布达佩斯条约期限内的保藏物，并服从 Genentech, Inc.与 ATCC 之间的协议，后者确保公众基于颁布的相关美国专利或者基于对任何美国或外国专利申请者开放，而可以永久性和无限

10 制利用保藏培养物的传代物，无论哪一个先来，根据美国专利商标委员会 35 USC 122 和所依据的委员会规则(包括 37 CFR 1.14, 特别参见 886 OG 638)规定授予的权利，确保个体得到传代培养物。

本申请的委托人已经同意，如果保藏培养物材料在适宜条件下培养过程中死亡或丢失或遭到破坏，应当在接到通知后迅速用另一相同材料替换保藏

15 材料。按照专利法的规定，保藏材料的可获得性，不得解释为是对违背任何政府当局所赋予权利而实施发明的许可。

前面撰写的说明书已经足以能够使本领域技术人员实施本发明。本发明不受保藏构建体范围的限制，因为保藏方案旨在作为本发明一个特定方面的说明，具有等同功能的任何构建体也包括在本发明范围之内。本文的保藏材

20 料既不构成准许本书面说明书所包含的不足以使本发明任何方面(包括其最佳模式)实施的内容的加入，也不解释为将权利要求的保护范围限定为特定的例证说明。事实上，除了本文所示和所述之外，对前面所描述的本发明进行各种变改，对于本领域技术人员而言是显而易见的，并落入本发明的保护范围。

序列表

- 5 <110> 杰南技术公司 (Genentech, Inc.)
Stewart, Timothy A.
Tomlinson, Elizabeth
- <120> 成纤维细胞生长因子-19 (FGF-19) 的核酸和多肽以及用于治疗肥胖的方法
- 10 <130> P1219P1PCT
- <140> PCT/US00/06471
<141> 2000-03-09
- 15 <150> PCT/US99/20594
<151> 1999-09-08
- <150> PCT/US99/21090
<151> 1999-09-15
- 20 <150> PCT/US99/30999
<151> 1999-12-20
- <150> PCT/US00/04414
<151> 2000-02-22
- 25 <160> 5
- <210> 1
<211> 2137
30 <212> DNA
<213> 人 (Homo Sapien)
- <400> 1
35 gctcccagcc aagaacctcg gggccgctgc gcggtgggga ggagttcccc 50
gaaaccggc cgctaagcga ggcctcctcc tcccgcagat ccgaacggcc 100
tgggcggggt caccocggct gggacaagaa gccgccgcct gcctgcccgg 150
40 gcccggggag ggggctgggg ctggggccgg aggcggggtg tgagtgggtg 200
tgtgcggggg gcggaggctt gatgcaatcc cgataagaaa tgctcgggtg 250
tcttgggcac ctaccogtgg ggcccgtaag gcgctactat ataaggctgc 300
45 cggcccggag ccgccgcgcc gtcagagcag gagcgcctgcg tccaggatct 350
agggccacga ccatcccaac ccggcactca cagccccgca gcgcatcccg 400
50 gtcgccgccc agcctcccgc acccccatcg ccggagctgc gccgagagcc 450

ccagggaggt gccatgcgga gcgggtgtgt ggtgggccac gtatggatcc 500
 tggccggcct ctggctggcc gtggccgggc gccccctcgc cttctcggac 550
 5 ggggggcccc acgtgcacta cggctggggc gaccccatcc gcctgcggca 600
 cctgtacacc tccggcccc acgggctctc cagctgcttc ctgcgcatcc 650
 10 gtcccgacgg cgtcgtggac tgcgcgctgg gccagagcgc gcacagtttg 700
 ctggagatca aggcagtcgc tctgcggacc gtggccatca agggcgtgca 750
 cagcgtgcgg tacctctgca tgggcgccga cggcaagatg caggggctgc 800
 15 ttcagtactc ggaggaagac tgtgctttcg aggaggagat cggcccagat 850
 ggctacaatg tgtaccgatc cgagaagcac cgctcccgg tctccctgag 900
 20 cagtgc meta cagcggcagc tgtaca meta cagaggcttt cttccactct 950
 ctcatctcct gccatgctg cccatggtcc cagaggagcc tgaggacctc 1000
 aggggccact tggaaatcga catgtttctc tggccccctgg agaccgacag 1050
 25 catggacca tttgggcttg tcaccggact ggaggccgtg aggagtccca 1100
 gctttgagaa gtaactgaga ccatgcccgg gcctcttcac tgctgccagg 1150
 30 ggctgtggta cctgcagcgt gggggacgtg cttctacaag aacagtcctg 1200
 agtccacgtt ctgtttagct ttagga meta acatctagaa gttgtacata 1250
 ttcagagttt tccattggca gtgccagttt ctagccaata gacttgtctg 1300
 35 atcataacat tgtaagcctg tagcttgccc agctgctgcc tgggccccca 1350
 ttctgctccc tcgaggttgc tggacaagct gctgcaactgt ctcagttctg 1400
 40 cttgaatacc tccatcgatg gggaactcac ttcctttgga aaaattctta 1450
 tgtcaagctg aaattctcta atttttctc atcaattccc caggagcagc 1500
 cagaagacag gcagtagttt taatttcagg aacaggtgat ccaactctgta 1550
 45 aaacagcagg taaatttcac tcaaccccat gtgggaattg atctatatct 1600
 ctacttccag ggaccatttg cccttccaa atccctccag gccagaactg 1650
 50 actggagcag gcatggcca ccaggcttca ggagtagggg aagcctggag 1700
 ccccactcca gccctgggac aacttgagaa ttccccctga ggccagttct 1750

gtcattgatg ctgtcctgag aataacttgc tgtcccgggtg tcacctgctt 1800
 5 ccactccca gccaccagc cctctgcca cctcacatgc ctccccatgg 1850
 attggggcct cccaggcccc ccaccttatg tcaacctgca cttcttgttc 1900
 aaaaatcagg aaaagaaaag attgaagac cccaagtctt gtcaataact 1950
 10 tgctgtgtgg aagcagcggg ggaagaccta gaaccctttc cccagcactt 2000
 ggttttccaa catgatattt atgagtaatt tattttgata tgtacatctc 2050
 ttattttctt acattattta tgccccaaa ttatatttat gtatgtaagt 2100
 15 gaggtttggt ttgtatatta aatggagtt tgtttgt 2137
 <210> 2
 <211> 216
 20 <212> PRT
 <213> 人(Homo Sapien)
 <400> 2
 25 Met Arg Ser Gly Cys Val Val Val His Val Trp Ile Leu Ala Gly
 1 5 10 15
 Leu Trp Leu Ala Val Ala Gly Arg Pro Leu Ala Phe Ser Asp Ala
 20 25 30
 30 Gly Pro His Val His Tyr Gly Trp Gly Asp Pro Ile Arg Leu Arg
 35 40 45
 His Leu Tyr Thr Ser Gly Pro His Gly Leu Ser Ser Cys Phe Leu
 50 55 60
 35 Arg Ile Arg Ala Asp Gly Val Val Asp Cys Ala Arg Gly Gln Ser
 65 70 75
 40 Ala His Ser Leu Leu Glu Ile Lys Ala Val Ala Leu Arg Thr Val
 80 85 90
 Ala Ile Lys Gly Val His Ser Val Arg Tyr Leu Cys Met Gly Ala
 95 100 105
 45 Asp Gly Lys Met Gln Gly Leu Leu Gln Tyr Ser Glu Glu Asp Cys
 110 115 120
 Ala Phe Glu Glu Glu Ile Arg Pro Asp Gly Tyr Asn Val Tyr Arg
 125 130 135
 50 Ser Glu Lys His Arg Leu Pro Val Ser Leu Ser Ser Ala Lys Gln

		140		145		150
		Arg Gln Leu Tyr Lys Asn Arg Gly Phe		Leu Pro Leu Ser His Phe		
		155		160		165
5		Leu Pro Met Leu Pro Met Val Pro Glu		Glu Pro Glu Asp Leu Arg		
		170		175		180
		Gly His Leu Glu Ser Asp Met Phe Ser		Ser Pro Leu Glu Thr Asp		
10		185		190		195
		Ser Met Asp Pro Phe Gly Leu Val Thr		Gly Leu Glu Ala Val Arg		
		200		205		210
15		Ser Pro Ser Phe Glu Lys				
		215				
		<210> 3				
		<211> 26				
20		<212> DNA				
		<213> 人工序列				
		<220>				
		<223> 人工合成的寡核苷酸探针				
25		<400> 3				
		atccgcccag atggctacaa tgtgta				26
		<210> 4				
30		<211> 22				
		<212> DNA				
		<213> 人工序列				
		<220>				
35		<223> 人工合成的寡核苷酸探针				
		<400> 4				
		ccagtccggt gacaagccca aa				22
40		<210> 5				
		<211> 42				
		<212> DNA				
		<213> 人工序列				
45		<220>				
		<223> 人工合成的寡核苷酸探针				
		<400> 5				
		gcctcccgt ctccctgagc agtgccaaac agcggcagtg ta				42

GCTCCCAGCCAAGAACCTCGGGGCCGCTGCGCGGTGGGGAGGAGTTCCCCGAAACCCGGCCG
CTAAGCGAGGCCTCCTCCTCCCAGATCCGAACGGCCTGGGCGGGGTCACCCCGGCTGGGA
CAAGAAGCCGCCGCTGCCTGCCCGGGCCGGGGAGGGGGCTGGGGCTGGGGCCGGAGGCGG
GGTGTGAGTGGGTGTGTGCGGGGGCGGAGGCTTGATGCAATCCCGATAAGAAATGCTCGGG
TGTCTTGGGCACCTACCCGTGGGGCCGTAAGGCGCTACTATATAAGGCTGCCGGCCCGGAG
CCGCCGCGCGTCAGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCCACGACCATCCCAACCC
GGCACTCACAGCCCCGAGCGCATCCCGGTGCGCCCCAGCCTCCCGCACCCCATCGCCGG
AGCTGCGCCGAGAGCCCCAGGGAGGTGCCATGCGGAGCGGGTGTGTGGTGGTCCACGTATGG
ATCCTGGCCGGCCTCTGGCTGGCCGTGGCCGGGCGCCCCCTCGCCTTCTCGGACGCGGGGCC
CCACGTGCACTACGGCTGGGGCGACCCCATCCGCCTGCGGCACCTGTACACCTCCGGCCCC
ACGGGCTCTCAGCTGCTTCTGCGCATCCGTGCGGACGGCGTCTGGACTGCGCGCGGGGC
CAGAGCGCGCACAGTTTGTGAGATCAAGGCAGTCGCTCTGCGGACCGTGGCCATCAAGGG
CGTGCACAGCGTGGGTACCTCTGCATGGGCGCCGACGGCAAGATGCAGGGGCTGCTTCAGT
ACTCGGAGGAAGACTGTGCTTTCGAGGAGGAGATCCGCCAGATGGCTACAATGTGTACCGA
TCCGAGAAGCACCGCCTCCCGGTCTCCCTGAGCAGTGCCAAACAGCGGCAGCTGTACAAGAA
CAGAGGCTTCTTCCACTCTCTCATTCCCTGCCATGCTGCCATGGTCCCAGAGGAGCCTG
AGGACCTCAGGGGCCACTTGAATCTGACATGTTCTCTTCGCCCTGGAGACCGACAGCATG
GACCCATTTGGGCTTGTACCCGACTGGAGGCCGTGAGGAGTCCAGCTTTGAGAAGTAACT
GAGACCATGCCCGGGCCTCTTCACTGCTGCCAGGGGCTGTGGTACCTGCAGCGTGGGGGACG
TGCTTCTACAAGAACAGTCCTGAGTCCACGTTCTGTTTAGCTTTAGGAAGAAACATCTAGAA
GTTGTACATATTCAGAGTTTTCCATTGGCAGTGCCAGTTTCTAGCCAATAGACTTGTCTGAT
CATAACATTGTAAGCCTGTAGCTTGCCAGCTGCTGCCTGGGCCCCCATCTGCTCCCTCGA
GGTTGCTGGACAAGCTGCTGCACTGTCTCAGTCTGCTTGAATACCTCCATCGATGGGGAAC
TCACTTCTTTGGAAAAATTCTTATGTCAAGCTGAAATTCTCTAATTTTTTCTCATCACTTC
CCCAGGAGCAGCCAGAAGACAGGCAGTAGTTTTAATTTAGGAACAGGTGATCCACTCTGTA
AAACAGCAGGTAATTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGG
ACCATTTGCCCTTCCCAAATCCCTCCAGGCCAGAACTGACTGGAGCAGGCATGGCCCACCAG
GTTTCCAGGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACTTGAGAATTC
CTGAGGCCAGTTCTGTATGGATGCTGTCTGAGAATAACTTGCTGTCCCGGTGTACCTGC
TTCCATCTCCAGCCCACCAGCCCTCTGCCACCTCACATGCCTCCCCATGGATTGGGGCCT
CCCAGGCCCCCACCTTATGTCAACCTGCACTTCTTGTTCAAAAATCAGGAAAAGAAAAGAT
TTGAAGACCCCAAGTCTTGTCAATAACTTGCTGTGTGGAAGCAGCGGGGGAAGACCTAGAAC
CCTTTCCCAGCACTTGGTTTTCCAACATGATATTTATGAGTAATTTATTTTGTATGTACA
TCTCTTATTTTCTTACATTATTTATGCCCCCAAATTATATTTATGTATGTAAGTGAGTTTG
TTTTGTATATTAATAATGGAGTTTGT

图 1

MRSGCVVHVWILAGLWLAVAGRPLAFSDAGPHVHYGWDPPIRLRLHLYTSGPHGLSSCFLRI
RADGVVDCARGQSAHSLLEIKAVLRVIAIKGVHVRVYLCMGADGKMQGLLOYSEEDCAFEE
EIRPDGYNVYRSEKHRLPVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPEEPEDLRGHLES
MFSSPLETDSMDPFGLVTGLEAVRSPSEK

信号肽:
氨基酸 1-22

N-肉豆蔻酰化位点:
氨基酸 15-21, 54-60, 66-72, 201-207

原核膜脂蛋白脂质附着位点:
氨基酸 48-59

HBMF/FGF 结构域:
氨基酸 80-131

图 2

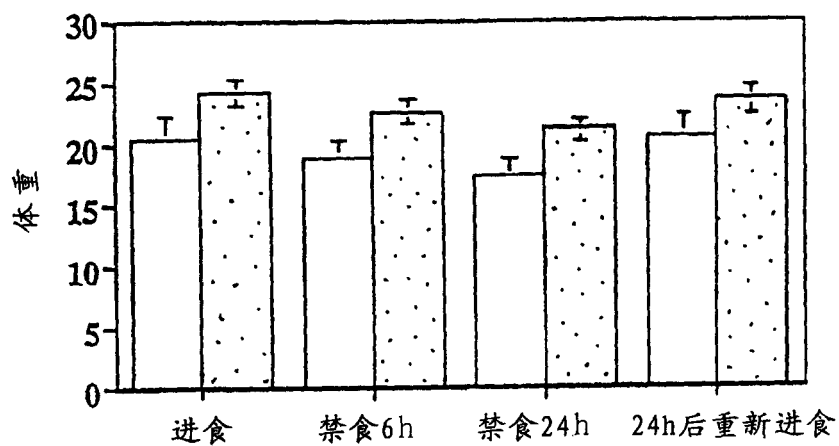


图 3A

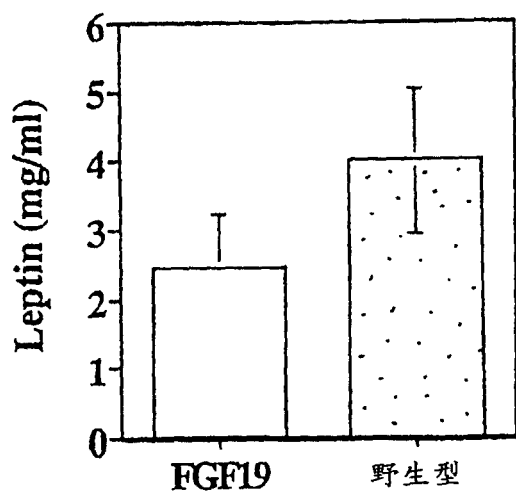


图 3B

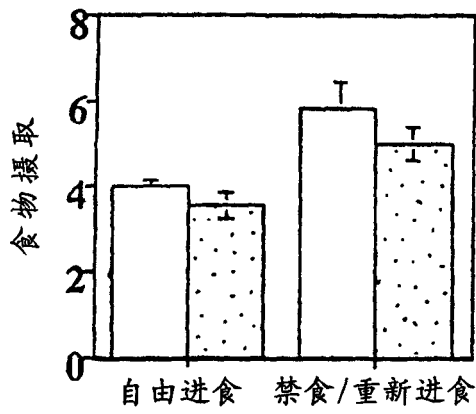


图 4A

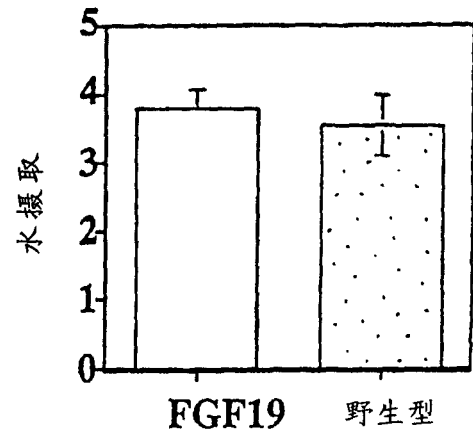


图 4B

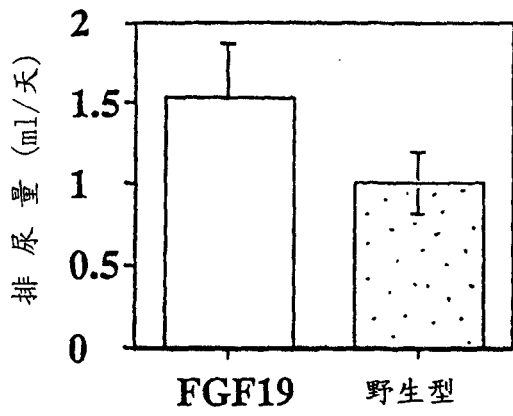


图 4C

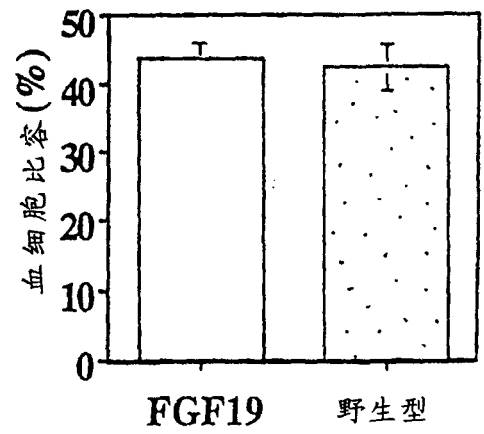


图 4D

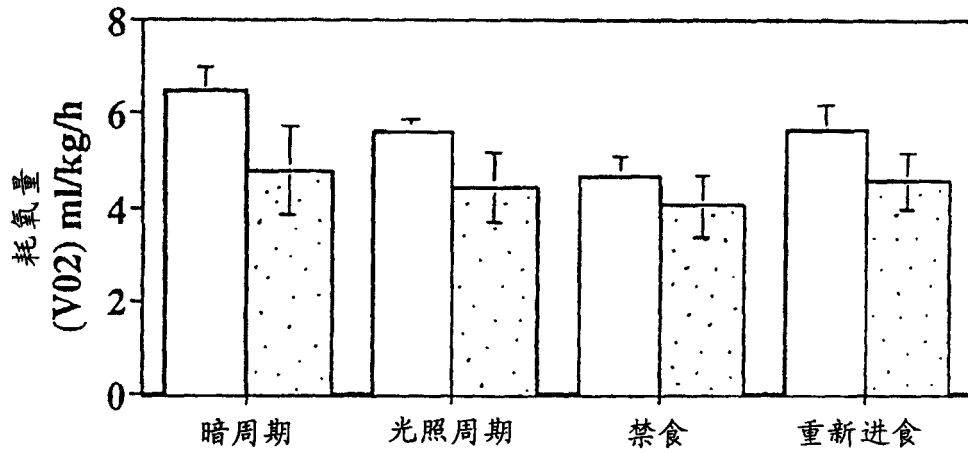


图 5

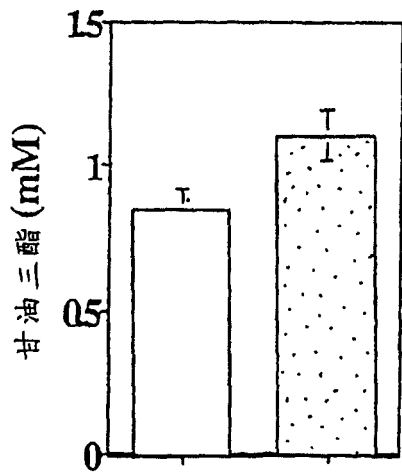


图 6A

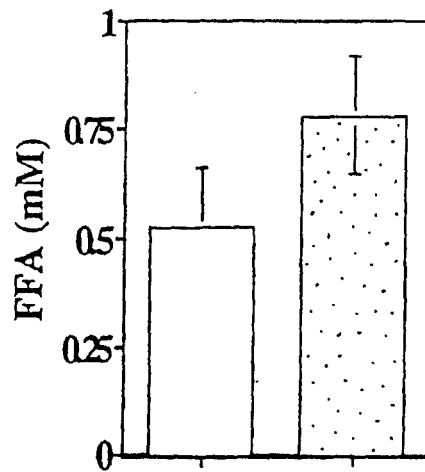


图 6B

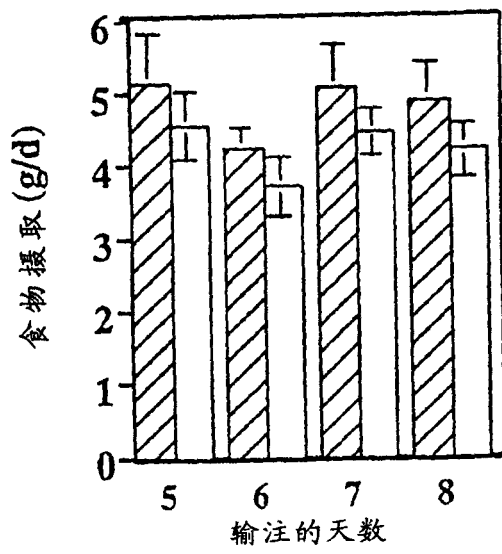


图 7A

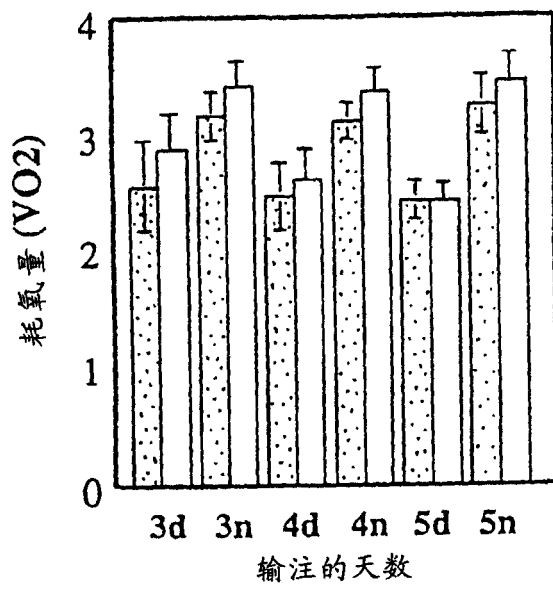


图 7B

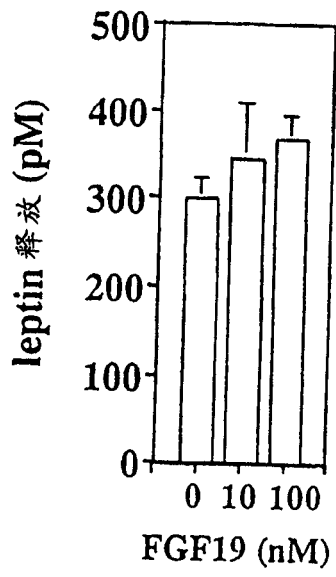


图 8A

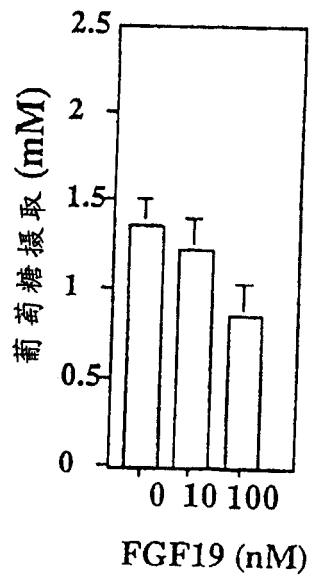


图 8B

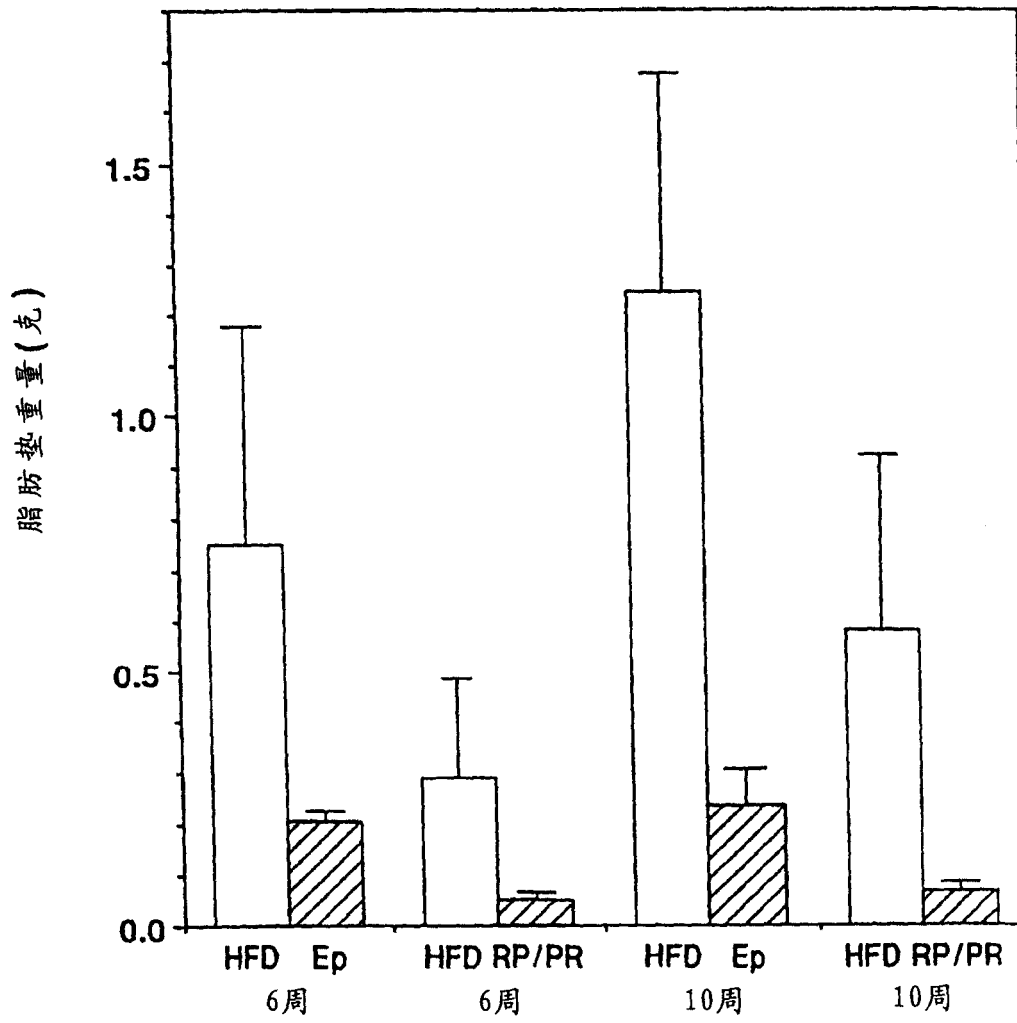


图 9

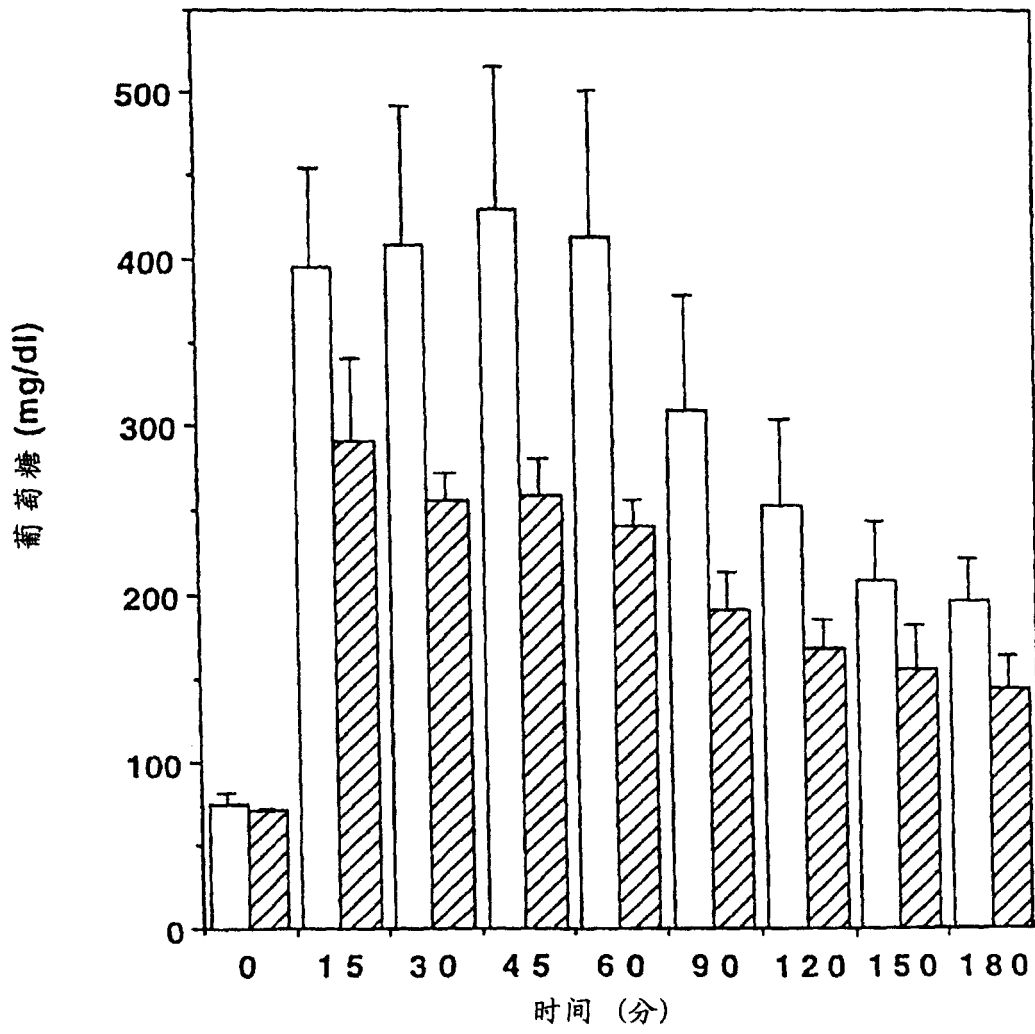


图 10

专利名称(译)	成纤维细胞生长因子 - 19 (FGF - 19) 的核酸和多肽以及用于治疗肥胖的方法		
公开(公告)号	CN1560249A	公开(公告)日	2005-01-05
申请号	CN200410046535.9	申请日	2000-03-09
[标]申请(专利权)人(译)	健泰科生物技术公司		
申请(专利权)人(译)	杰南技术公司		
当前申请(专利权)人(译)	杰南技术公司		
[标]发明人	蒂莫西·阿斯图尔特 伊丽莎白·汤姆林森		
发明人	蒂莫西·A·斯图尔特 伊丽莎白·汤姆林森		
IPC分类号	A01K67/027 A61K31/70 A61K38/18 A61K38/22 A61K48/00 A61P1/16 A61P3/00 A61P3/04 A61P3/06 A61P3/10 A61P7/02 A61P9/00 A61P9/10 A61P9/12 A61P15/00 A61P17/00 A61P19/00 A61P19/02 A61P31/00 A61P35/00 A61P43/00 C07K14/47 C07K14/50 C07K16/18 C07K16/46 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12N15/12 C12N15/62 C12N15/63 C12P21/02 C12P21/08 C12Q1/02 C12Q1/68 G01N33/15 G01N33/50 G01N33/53		
CPC分类号	C12N15/09 G01N33/53 A61K48/00 C07K14/47 C07K16/46 C12Q1/02 C12P21/02 C07K16/18 A61K38/22 C12N15/62 C12N15/63 C12Q1/68 C07K14/50 A61K38/18 C12N5/10 A01K67/027 G01N33/15 G01N33/50 A61K31/70 A01K67/0275 A01K2217/052 A01K2227/105 A01K2267/01 A61K38/00 A61P1/16 A61P3/00 A61P3/04 A61P3/06 A61P3/10 A61P7/02 A61P9/00 A61P9/10 A61P9/12 A61P15/00 A61P17/00 A61P19/00 A61P19/02 A61P31/00 A61P35/00 A61P43/00 C12N2710/14143 G01N33/74 G01N2333/50		
优先权	PCT/US1999/020594 1999-09-08 WO PCT/US1999/021090 1999-09-15 WO PCT/US1999/030999 1999-12-20 WO PCT/US2000/004414 2000-02-22 WO		
外部链接	Espacenet SIPO		

摘要(译)

本发明涉及成纤维细胞生长因子家族的新多肽和编码这些多肽的核酸分子。本发明也提供包含这些核酸序列、含有本发明多肽与异源多肽序列融合的嵌合多肽分子、与本发明多肽结合的抗体的载体和宿主细胞，以及涉及本发明多肽的生产方法。另外，也提供治疗肥胖的方法。

