

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12Q 1/26

C07K 14/00 C12Q 1/68

A61P 25/18 A61P 25/24

A61K 31/198



[12] 发明专利申请公开说明书

[21] 申请号 02806665.0

[43] 公开日 2005 年 1 月 19 日

[11] 公开号 CN 1568370A

[22] 申请日 2002. 1. 15 [21] 申请号 02806665.0

[30] 优先权

[32] 2001. 1. 16 [33] US [31] 60/261,883

[32] 2001. 7. 13 [33] US [31] 60/305,445

[32] 2001. 10. 22 [33] US [31] 60/345,211

[32] 2001. 11. 19 [33] US [31] 60/333,881

[86] 国际申请 PCT/IB2002/001262 2002. 1. 15

[87] 国际公布 WO2002/066672 英 2002. 8. 29

[85] 进入国家阶段日期 2003. 9. 16

[71] 申请人 根瑟特公司

地址 法国巴黎

[72] 发明人 D·科恩 I·丘马科夫

[74] 专利代理机构 上海专利商标事务所

代理人 周承泽

权利要求书 12 页 说明书 111 页 序列表 55 页
附图 6 页

[54] 发明名称 用 D - 氨基酸氧化酶和 D - 天冬氨酸氧化酶拮抗剂治疗中枢神经系统障碍

[57] 摘要

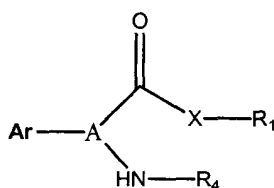
本发明揭示了 D - 氨基酸氧化酶和 D - 天冬氨酸氧化酶拮抗剂的化合物、用所述化合物治疗中枢神经系统障碍的方法，包括双相性精神障碍、精神病和精神分裂症以及含有这些拮抗剂的药学上可接受的组合物。

I S S N 1 0 0 8 - 4 2 7 4

1. 一种用于治疗中枢神经系统障碍的候选分子的鉴定方法，其特征在于：所述方法包括：
 - a) 使 D-氨基酸氧化酶或 D-天冬氨酸氧化酶多肽或其生物活性片段与测试化合物接触；
 - b) 确定所述的化合物是否 (i) 与所述的多肽结合，或者 (ii) 抑制所述的多肽的活性；以及
 - c) 如果所述的化合物与所述的多肽结合或者抑制所述的多肽的活性，就将所述的化合物施加给精神分裂症、抑郁症或双相性精神障碍动物模型，其中确定所述的化合物改善了所述的动物模型中枢神经系统障碍的典型症状，则表示所述的化合物就是治疗中枢神经系统障碍的候选分子。
2. 如权利要求 1 所述的方法，其特征在于：所述的化合物选择性与所述的多肽结合。
3. 如权利要求 1 所述的方法，其特征在于：所述的化合物选择性抑制所述的多肽的活性。
4. 如权利要求 1 至 3 中任何一项所述的方法，其特征在于：所述的化合物能够抑制一 D-氨基酸的氧化或降解，所述的 D-氨基酸选自由 D-Met、D-Pro、D-Phe、D-Tyr、D-Ile、D-Leu、D-Ala、D-Val、D-Ser、D-Arg、D-His、D-正亮氨酸、D-Trp、D-鸟氨酸、顺-4-羟基-D-脯氨酸、D-Thr、D-Trp-甲酯、N-乙酰基-D-Ala、D-Lys、D-Asp、D-Glu、D-Asn、D-Gln、D-Asp-二甲酯和 N-甲基-D-Asp 组成的组。
5. 如权利要求 4 所述的方法，其特征在于：所述的化合物能够抑制 D-丝氨酸的氧化和降解。
6. 一种治疗精神分裂症、抑郁症或双相性精神障碍的候选分子的鉴定方法，其特征在于：所述方法包括：
 - a) 使 D-氨基酸氧化酶或 D-天冬氨酸氧化酶或其生物活性片段多肽与测试化合物接触；以及
 - b) 确定所述的化合物是否选择性抑制所述的多肽的活性；其中确定所述的化合物选择性抑制所述的多肽的活性，则表示所述的化合物就是治疗精神分裂症、抑郁症或双相性精神障碍的候选分子。

7. 一种治疗精神分裂症、抑郁症或双相性精神障碍的候选分子的鉴定方法，其特征在于：所述方法包括：
 - a) 使 D-氨基酸氧化酶或 D-天冬氨酸氧化酶或其生物活性片段多肽与测试化合物接触；以及
 - b) 确定所述的化合物是否选择性与所述的多肽结合，其中确定所述的化合物选择性与所述的多肽结合，则表示所述的化合物就是治疗精神分裂症、抑郁症或双相性精神障碍的候选分子。
8. 一种编码 D-氨基酸氧化酶多肽的分离的或纯化的核酸，其特征在于：所述的核酸选自由下列核酸分子组成的组：
 - (i) 一编码多肽的核酸分子，所述的核酸分子包括一选自由 SEQ ID NO 8 到 10 组成的序列组的氨基酸序列；以及
 - (ii) 一含有核酸序列的核酸分子，所述的核酸序列选自由 SEQ ID NO 1 到 6 组成的序列组或其互补序列。
9. 一种分离的或纯化的 D-氨基酸氧化酶多肽，其特征在于：所述的 D-氨基酸氧化酶多肽含有一氨基酸序列，所述的氨基酸序列选自由 SEQ ID NO 8 到 10 组成的序列组。
10. 一种分离的或纯化的 D-氨基酸氧化酶多肽，其特征在于：所述的 D-氨基酸氧化酶多肽由含有一核酸序列的核酸分子编码，所述的核酸序列选自由 SEQ ID NO 1 到 6 组成的序列组或其互补序列。
11. 一种与精神分裂症、抑郁症或双相性精神障碍有关的关联性的检测方法，其特征在于：所述的方法包括以生物样本中 D-氨基酸氧化酶相关的双等位标记测定核苷酸或其补体的同一性。
12. 如权利要求 11 所述的方法，其特征在于：所述的生物样本取自单个受试者。
13. 如权利要求 12 所述的方法，其特征在于：所述的生物样本取自多个受试者。
14. 如权利要求 11 所述的方法，其特征在于：所述的方法还包括在所述的测定步骤之前扩增部分含有双等位标记的所述序列。
15. 如权利要求 14 所述的方法，其特征在于：通过 PCR 技术进行所述的扩增。
16. 如权利要求 11 所述的方法，其特征在于：通过杂交分析、排序分析、微排序分析或基于酶的错配检测分析而进行所述的测定。

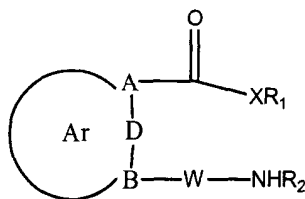
17. 一种群体中双等位标记等位基因的频率的估计方法，其特征在于：所述方法包括：
- a) 根据权利要求 11 所述的方法以所述的双等位标记使所述群体中的个体基因分型；以及
 - b) 确定所述的群体中所述的双等位标记的比例表示。
18. 一种基因型和性状之间关联性的检测方法，其特征在于：所述方法包括以下步骤：
- a) 根据权利要求 17 所述的方法，确定性状呈阳性的群体中至少一个双等位标记的频率；
 - b) 根据权利要求 17 所述的方法，确定对照群体中至少一个双等位标记的频率；以及
 - c) 确定所述的基因型与所述的性状之间是否有一种差异显著的关联性。
19. 一种个体是否有精神分裂症、抑郁症或双相性精神障碍风险的判断方法，其特征在于：所述方法包括：
- a) 根据权利要求 11 所述的方法，使至少一个双等位标记基因分型；以及
 - b) 将步骤 a) 所得的结果与发展精神分裂症、抑郁症或双相性精神障碍的风险联系起来。
20. 如权利要求 11 至 19 中任何一项所述的方法，其特征在于：所述的双等位标记选自由 SEQ ID NO 1、24、26 和 29 或其互补序列所述的标记组成的组。
21. 一种受试者对精神分裂症、抑郁症或双相性精神障碍易感因素或易感性的诊断、检测方法，其特征在于：所述方法包括：
- a) 由所述的受试者采集核酸样本；以及
 - b) 测定所述的生物样本在 D-氨基酸氧化酶相关的多态性的核苷酸或其补体的同一性。
22. 如权利要求 1-7 所述的方法，其特征在于：所述的测试化合物选自由下列化合物组成的组：
- (1) 以包括下列结构式的结构表示的化合物或其药学上可接受的盐类：



式中,

- a) A 是烷基, 例如甲基、乙基、丙基或丁基; 支链烷基, 例如异丁基、异丙基、异戊基或者环烷基, 例如环丙基、环戊基或环己基。这些基团本身可由 C_1 - C_6 烷基、卤素、羟基或氨基取代;
- b) X 是氧或氮;
- c) Ar 是芳香族单环、双环或三环稠合杂环, 其中环或者未取代或者 1 至 5 位上由氢、卤素、羟基、 $-CN$ 、 COR_2 、 $--CONR_2R_3$ 、 $--S(O)_nR_2$ 、 $--OP(O)(OR_2)OR_3$ 、 $--PO(OR_3)R_3$ 、 $--OC(O)NR_2R_3$ 、 $--COOR_2$ 、 $--CONR_2R_3$ 、 $--SO_3H$ 、 $--NR_2R_3$ 、 $--NR_2COR_3$ 、 $--NR_3COOR_3$ 、 $--SO_2NR_2R_3$ 、 $--N(R_2)SO_2R_3$ 、 $--NR_2CONR_2R_2$ 、 $--SO_2NHCOR_2$ 、 $--CONHSO_2R_2$ 、 $--SO_2NHCN$ 、 $--OR_1$ 、 C_1 - C_6 直链或支链烷基或链烯基、或者由一个或多个卤素、羟基、氨基、羧基、羧酰胺、腈、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、磷酸盐、 Ar^1 、 N_3 或其组合取代的 C_1 - C_6 支链或直链烷基或链烯基, 其中杂环含有 1-6 个选自由氧、氮、硫及其组合组成的组的杂原子;
- d) R_4 是 H、烷基、 Ar^1 、O、取代的烷基;
- e) R^1 是 $(C_1 - C_6)$ 烷基、 Ar^1 、 $(C_1 - C_4)$ 烷氧基羰基甲基、取代的烷基;
- f) R_2 和 R_3 各自独立地是氢、 C_1 - C_6 直链或支链烷基或链烯基、或者由一个或多个卤素、羟基、氨基、羧基、羧酰胺、腈、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、磷酸盐、 Ar^1 、 N_3 取代的 C_1 - C_6 支链或直链烷基或链烯基; 以及
- g) Ar^1 是单环、双环或三环、碳环或杂环, 其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1 - C_6 直链或支链烷基或链烯基、 C_1 - C_4 烷氧基、 C_1 - C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代; 其中单个环的大小是 3-7 个原子; 而且杂环含有 1-6 个选自由氧、氮、硫及其组合组成的组的杂原子;

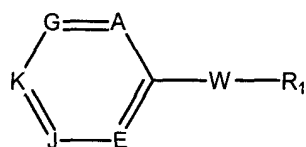
(2) 以包括下列结构式的结构表示的化合物:



式中,

- a) A 和 B 由碳或氮组成, D 可含有 0-2 个由碳或氮组成的原子;
- b) W 是 C_1-C_4 烷基, 如 $(CH_2)_n$ 、支链烷基;
- c) n 是 0-4; 此外, 当 $n = 0$ 时假定 $-NHR_2$ 与 B 共价键合;
- d) X 是氧或氮;
- e) R_2 是氢、烷基、 Ar^1 、或氧取代的烷基;
- f) R^1 是 $(C_1 - C_6)$ 烷基、 Ar^1 、 $(C_1 - C_4)$ 烷氧羰基甲基或取代的烷基;
- g) Ar 是芳香族单环、双环或三环稠合杂环, 其中环或者未取代或者 1 至 6 位上由卤素、羟基、硝基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、苯氧基、苄氧基、氨基、 C_3-C_6 环烷基或其组合取代; 其中单个环的大小是 5-6 个原子; 而且杂环含有 1-6 个选自由氧、氮、硫及其组合组成的组的杂原子; 以及
- h) Ar^1 是单环、双环或三环、碳环或杂环, 其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代; 其中单个环的大小是 3-7 个原子; 而且杂环含有 1-6 个选自由氧、氮、硫及其组合组成的组的杂原子;

(3) 以包括下列结构式的结构表示的化合物:



式中,

- a) A、G、K、J、E 是六元碳环或杂环芳族环的原子, 其中杂环含有 1-6 个选自由碳、氮及其组合组成的组的杂原子;
- b) A、G、K、J、E 可各自独立地是未取代或由氢、卤素、羟基、 $-CN$ 、 COR_2 、 $--CONR_2R_3$ 、 $--S(O)_nR_2$ 、 $--OPO(OR_2)OR_3$ 、 $--PO(OR_3)R_3$ 、 $--OC(O)NR_2R_3$ 、 $--COOR_2$ 、 $--CONR_2R_3$ 、 $--SO_3H$ 、 $--NR_2R_3$ 、 $--NR_2COR_3$ 、 $--NR_3COOR_3$ 、 $--SO_2NR_2R_3$ 、 $--N(R_2)SO_2R_3$ 、 $--NR_2CONR_2R_2$ 、 $--SO_2NHCOR_2$ 、 $--CONHSO_2R_2$ 、 $--SO_2NHCN$ 、 $--OR_1$ 、 C_1-C_6 直链或支链烷基或链烯基取代, 或者由一个或多个卤素、羟基、氨基、羧基、羧酰胺、腈、

硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、磷酸盐、 Ar^1 、 N_3 取代的 C_1 - C_6 支链或直链烷基或链烯基；

c) R_1 是 CN 、 COR_2 、 $-\text{CONR}_2\text{R}_3$ 、 $-\text{S}(\text{O})_n\text{R}_2$ 、 $-\text{OPO}(\text{OR}_2)\text{OR}_3$ 、 $-\text{PO}(\text{OR}_3)\text{R}_3$ 、 $-\text{OC}(\text{O})\text{NR}_2\text{R}_3$ 、 $-\text{COOR}_2$ 、 $-\text{CONR}_2\text{R}_3$ 、 $-\text{SO}_3\text{H}$ 、 $-\text{NR}_2\text{R}_3$ 、 $-\text{NR}_2\text{COR}_3$ 、 $-\text{NR}_3\text{COOR}_3$ 、 $-\text{SO}_2\text{NR}_2\text{R}_3$ 、 $-\text{N}(\text{R}_2)\text{SO}_2\text{R}_3$ 、 $-\text{NR}_2\text{CONR}_2\text{R}_2$ 、 $-\text{SO}_2\text{NHCOR}_2$ 、 $-\text{CONHSO}_2\text{R}_2$ 、 $-\text{SO}_2\text{NHCN}$ 、 SCN 、 COCO_2H 、 C_1 - C_6 直链或支链烷基或链烯基、或者由一个或多个卤素、羟基、氨基、羧基、羧酰胺、腈、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、磷酸盐、 Ar^1 、 N_3 取代的 C_1 - C_6 支链或直链烷基或链烯基；

d) W 是 N 、 $(\text{CH}_2)_x$ 、或 $-\text{NCH}_2$ ；

e) $x=0-4$ ；

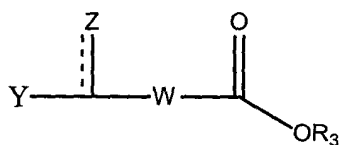
f) $n=0-2$ ；

g) R_2 和 R_3 各自独立地是氢、 C_1 - C_6 直链或支链烷基或链烯基、或者由一个或多个卤素、羟基、氨基、羧基、羧酰胺、腈、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、磷酸盐、 Ar^1 、 N_3 取代的 C_1 - C_6 支链或直链烷基或链烯基；以及

h) Ar^1 是单环、双环或三环、碳环或杂环，其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1 - C_6 直链或支链烷基或链烯基、 C_1 - C_4 烷氧基、 C_1 - C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代；其中单个环的大小是 5-6 个原子；而且杂环含有 1-6 个选自由氧、氮、硫及其组合组成的组的杂原子。

23. 如权利要求 1-7 所述的方法，其特征在于：所述的测试化合物选自由下列化合物组成的组：

(1) 以包括下列结构式的结构表示的化合物：



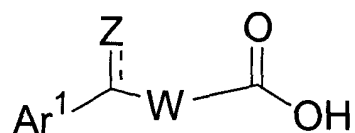
式中，

a) $\text{W} = (\text{CH}_2)_n$ ；

b) $n=0-5$ ；

c) Z 是氧或羟基；

- d) $Y=H$ 、 Ar^1 、 $R_4(CH_2)_x$ 、 $R_1S(CH_2)_x$ 、 $R_1SO(CH_2)_x$ 、 $R_1SO_2(CH_2)_x$ 、 $R_1SO_3(CH_2)_x$ 、 $HNR_1SO_2(CH_2)_x$ 、 $R_1R_2N(CH_2)_x$ 、 $R_1O(CH_2)_x$ 、 CF_3 或 OH ;
- e) $x=0-6$;
- f) R_1 、 R_2 和 R_3 各自独立地是氢、 C_1-C_6 直链或支链烷基、或者由一个或多个卤素、羟基、氨基、羧基、羧酰胺、腈、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、磷酸盐、 Ar^1 取代的 C_1-C_6 支链或直链烷基;
- g) R_4 是卤素、 CN 、 N_3 、 C_1-C_6 直链或支链烷基、或者由一个或多个卤素、羟基、硝基、烷氧基、三氟甲基、磺酸盐、膦酸盐、磷酸盐、 Ar^1 、 $--COR_1$ 、 $--COOR_1$ 、 $--CONR_1R_2$ 、 CN 、 $--NR_1$ 、 $--NR_1R_2$ 、 $--SR_1$ 、 $--SO_2NHCN$ 或者 N_3 取代的 C_1-C_6 支链或直链烷基; 以及
- h) Ar^1 是单环、双环或三环、碳环或杂环, 其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代; 其中单个环的大小是 5-6 个原子; 而且杂环含有 1-6 个选自由氧、氮、硫及其组合组成的组的杂原子;
- (2) 以包括下列结构式的结构表示的化合物:

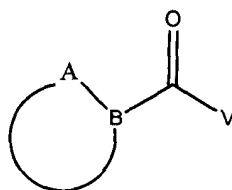


式中,

- a) Y 是 Ar^1 ;
- b) Z 是羰基或羟基;
- c) W 是 $(CH_2)_n$, 其中 ($n=0, 1$ 或 2), $R_3 = H$; 以及
- d) Ar^1 是单环、双环或三环、碳环或杂环, 其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代; 其中单个环的大小是 5-6 个原子; 而且杂环含有 1-6 个选自由氧、氮、硫及其组合组成的组的杂原子。

24. 如权利要求 1-7 所述的方法, 其特征在于: 所述的测试化合物以包括下列

结构式的结构表示:

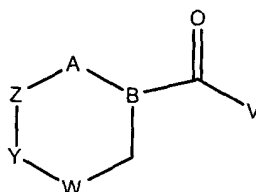


式中,

- a) A 和 B 一起形成一个 5-8 元饱和或部分不饱和杂环, 其含有至少一个处于化学稳定氧化状态的附加杂原子 O、S、SO、SO₂、NH、或 NR¹;
- b) V 是 O、OR₁、NR₂、NR₁R₂、CHR₁R₂、CH₂R₃、CHR₃R₄ 或 CH₂N₃;
- c) R₁ 和 R₂ 各自独立地是氢、C₁-C₆ 直链或支链烷基、或者由一个或多个卤素、羟基、氨基、羧基、羧酰胺、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、或 Ar¹ 取代的 C₁-C₆ 支链或直链烷基;
- d) R₃ 和 R₄ 或者是卤素、C₁-C₆ 直链或支链烷基、或者由一个或多个羟基、氨基、羧基、羧酰胺、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、Ar¹、--OC(O)R₁、--COOR₁、--CONR₁R₂、CN, NR₁、NR₁R₂、SR₁、SO₂NHCN 或 N₃ 取代的 C₁-C₆ 支链或直链烷基; 以及
- e) Ar¹ 是单环、双环或三环、碳环或杂环, 其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、C₁-C₆ 直链或支链烷基或链烯基、C₁-C₄ 烷氧基、C₁-C₄ 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代; 其中单个环的大小是 5-6 个原子; 而且杂环含有 1-6 个选自由氧、氮、硫及其组合组成的组的杂原子。

25. 如权利要求 24 所述的方法, 其特征在于: 所述的化合物是脘硫醚酮亚胺和环硫堍。

26. 如权利要求 1-7 所述的方法, 其特征在于: 所述的测试化合物以包括下列结构式的结构表示:



式中,

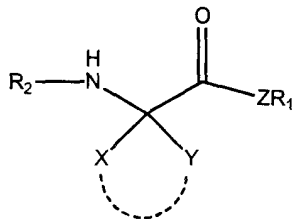
- a) W-Y-Z-A-B 包括六元饱和或部分饱和碳环或杂环, 其中杂环含有选自由氧、氮、硫及其任意组合组成的组的杂原子;

- b) B 是 C、CH 或 N;
- c) A、W、Y、Z 各自独立地是 CH_2 、 CHR_3 、 CR_3R_4 、O、S、 SO 、 SO_2 、NH、 NR_1 、 NR_1R_2 或 $\text{C}=\text{O}$;
- d) V 是 O、 OR_1 、 NR_2 、 NR_1R_2 、 CHR_1R_2 、 CH_2R_3 、 CHR_3R_3 或 CH_2N_3 ;
- e) R_1 和 R_2 各自独立地是氢、 C_1 - C_6 直链或支链烷基、或者由一个或多个卤素、羟基、氨基、羧基、羧酰胺、腈、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、磷酸盐或 Ar^1 取代的 C_1 - C_6 支链或直链烷基;
- f) R_3 和 R_4 各自独立地是卤素、 $-\text{OC}(\text{O})\text{R}_1$ 、 $-\text{COOR}_1$ 、 $-\text{CONR}_1\text{R}_2$ 、CN、 $-\text{NR}_1$ 、 $-\text{NR}_1\text{R}_2$ 、 $-\text{SR}_1$ 、 $-\text{SO}_2\text{NHCN}$ 、 N_3 、 C_1 - C_6 直链或支链烷基、或者由一个或多个卤素、羟基、硝基、烷氧基、三氟甲基、磺酸盐、膦酸盐、 Ar^1 、 $-\text{OC}(\text{O})\text{R}_1$ 、 $-\text{COOR}_1$ 、 $-\text{CONR}_1\text{R}_2$ 、CN、 $-\text{NR}_1$ 、 $-\text{NR}_1\text{R}_2$ 、 $-\text{SR}_1$ 、 $-\text{SO}_2\text{NHCN}$ 或 N_3 取代的 C_1 - C_6 支链或直链烷基; 以及
- g) Ar^1 是单环、双环或三环、碳环或杂环, 其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1 - C_6 直链或支链烷基或链烯基、 C_1 - C_4 烷氧基、 C_1 - C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代; 其中单个环的大小是 5-6 个原子; 而且杂环含有 1-6 个选自由氧、氮、硫及其任意组合组成的组的杂原子。

27. 如权利要求 26 所述的方法, 其特征在于: 所述的化合物是选自由氨基乙基半胱氨酸酮亚胺 (2H-1, 4-噻嗪-5, 6-二氢-3-羧酸)、硫代吗啉-2-羧酸、羊毛硫氨酸酮亚胺和 1, 4-硫代吗啉-3, 5-二羧酸组成的组。

28. 如权利要求 1-7 所述的方法, 其特征在于: 所述的测试化合物选自由下列化合物组成的组:

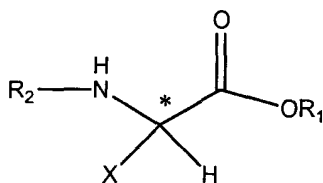
(1) 以包括下列结构式的结构表示的化合物:



式中,

a) Z 是 O 或 NH;

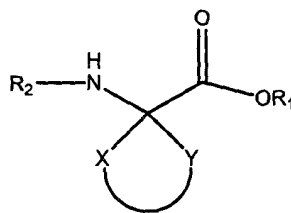
- b) R^1 是 (C_1-C_6) 烷基、 Ar^1 、或者 (C_1-C_4) 烷氧基羰基甲基；
- c) X、Y 各自独立地是氢、 Ar^1 、 (C_1-C_6) 烷基(可由杂原子断开或取代，如氮、磷、氧、硫或硅，杂原子本身可由 (C_1-C_3) 烷基取代一次或多次)、 (C_2-C_6) 链烯基、 (C_1-C_6) 卤代烷基、或卤素。当 X 和 Y 各自都是碳时，它们可共价结合形成一种饱和或部分不饱和独立地由碳、氮、氧和硫组成的 3-8 元碳环化合物，此外，环的原子本身可以是未取代或由卤素、羟基、羧基、硝基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、苯氧基、苄氧基、氨基、取代的烷基、 Ar^1 或其组合取代；
- d) R_2 是氢、烷基、 Ar^1 或氧取代的烷基；以及
- e) Ar^1 是单环、双环或三环、碳环或杂环，其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代；其中单个环的大小是 3-7 个原子；而且杂环含有 1-6 个选自由氧、氮、硫及其任意组合组成的组的杂原子。
- (2) 以包括下列结构式的结构表示的化合物：



式中，

- a) * = 不对称中心，及
- b) R^1 = (C_1-C_6) 烷基、 Ar^1 、 (C_1-C_4) 烷氧基羰基甲基及
- c) X 是氢、 (C_1-C_6) 烷基(可由杂原子断开或取代，如氮、磷、氧、硫或硅，杂原子本身可由 (C_1-C_3) 烷基取代一次或多次)、 (C_2-C_6) 链烯基、 (C_1-C_6) 卤代烷基、卤素或 Ar^1 ；
- d) R_2 是氢、烷基、 Ar^1 或氧取代的烷基；
- e) Ar^1 是单环、双环或三环、碳环或杂环，其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代；其中单个环的大小是 3-7 个原子；而且杂环含有 1-6 个选自由氧、氮、硫及其任意组合组成的组的杂原子。

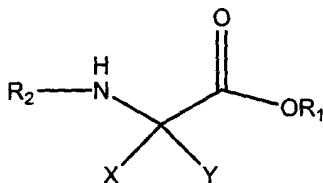
(3) 以包括下列结构式的结构表示的化合物:



式中,

- a) X 和 Y 各自是碳;
- b) X 和 Y 以 3-8 个碳原子的饱和或部分饱和环连接, 这样一个环本身的 1 至 5 位上可由卤素、羟基、羧基、氨基、硝基、氰基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、或取代的烷基取代;
- c) R^1 是 (C_1-C_6) 烷基、 Ar^1 或 (C_1-C_4) 烷氧羰基甲基;
- d) R_2 是氢、烷基、 Ar^1 或氧取代的烷基; 以及
- e) Ar^1 是单环、双环或三环、碳环或杂环, 其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代; 其中单个环的大小是 3-7 个原子; 而且杂环含有 1-6 个选自由氧、氮、硫及其任意组合组成的组的杂原子。

(4) 以包括下列结构式的结构表示的化合物:

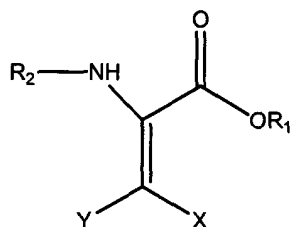


式中,

- a) X、Y 各自独立地是氢、 Ar^1 、 (C_1-C_6) 烷基(可由杂原子断开或取代, 如氮、磷、氧、硫或硅, 杂原子本身可由 (C_1-C_3) 烷基取代一次或多次)、 (C_2-C_6) 链烯基、 (C_1-C_6) 卤代烷基、或卤素如萘基或苯基;
- b) R_2 是氢、烷基、 Ar^1 或氧取代的烷基; 以及
- c) Ar^1 是单环、双环或三环、碳环或杂环, 其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、苯氧基、苄氧基、氨基、

或其组合取代；其中单个环的大小是 3-7 个原子；而且杂环含有 1-6 个选自由氧、氮、硫及其任意组合组成的组的杂原子。

29. 如权利要求 1-7 所述的方法，其特征在于：所述的测试化合物以包括下列结构式的结构表示：



式中，

- a) R^1 是 (C_1-C_6) 烷基、 Ar^1 或 (C_1-C_4) 烷氧羰基甲基；
 - b) R_2 是氢、烷基、 Ar^1 或氧取代的烷基；
 - c) Y 是氢、 Ar^1 、 (C_1-C_6) 烷基(可由杂原子断开或取代，如氮、磷、氧、硫或硅，杂原子本身可由 (C_1-C_3) 烷基取代一次或多次)、 (C_2-C_6) 链烯基、 (C_1-C_6) 卤代烷基、或卤素；以及
 - d) X 是烷基或苯基。
30. 一种能够抑制 D-氨基酸转化为相应的 α -酮酸的化合物在制备治疗精神分裂症或双相性精神障碍的药物中的应用，其特征在于：所述的化合物选自由苯甲酸盐、氨基乙基半胱氨酸-酮亚胺、氨基乙基半胱氨酸(硫赖氨酸)、半胱胺、泛酰巯基乙胺、胱硫醚和 S-腺苷甲硫氨酸及其衍生物组成的组。
31. 一种能够抑制 D-氨基酸转化为相应的 α -酮酸的化合物在制备治疗精神分裂症或双相性精神障碍的药物中的应用，其特征在于：所述的化合物选自由权利要求 22 至 30 所述的化合物组成的组。
32. 如权利要求 30 或 31 所述的应用，其特征在于：所述的化合物能够抑制 D-丝氨酸的氧化或降解。
33. 如权利要求 32 所述的应用，其特征在于：所述的化合物可抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶多肽的活性。

用 D-氨基酸氧化酶和 D-天冬氨酸氧化酶拮抗剂治疗中枢神经系统障碍

技术领域

本发明提供用于治疗与中枢神经系统障碍相关的诸如精神分裂症、抑郁症及其他情绪失常的疾病的化合物的鉴定装置、个体对所述疾病罹病倾向的判断装置、以及所述疾病的疾病诊断和预后装置。更加具体地说，本发明涉及用 D-氨基酸氧化酶(DAO)和 D-天冬氨酸氧化酶(DDO)拮抗剂治疗所述障碍的装置。

发明背景

技术装备方面的进步使基础和临床研究人员能够对健康和患病中的脑及神经系统功能进行越来越深入的研究。就与中枢神经系统(CNS)障碍包括精神障碍和神经退化性疾病有关的神经化学和基因机理而论，人们业已提出了许多神经生物学上及药理学上的假设。然而，中枢神经系统障碍很复杂，病原学上知之甚少以及症状重迭，性状不显，故而难以判断。所以，进一步的治疗方案和药物研制的努力会更加复杂，并集中在多基因病因上，而且需要一些新测试去把患病人群分开，从而为患有中枢神经系统障碍的病人提供更准确的诊断和预后信息。

中枢神经系统障碍的神经学基础

神经递质起全身的信号递质的作用。所以影响神经传递的疾病可导致严重后果。例如，三十多年来解释许多精神障碍如抑郁症的生物基础的主导理论一直是单胺假说。这一理论提出抑郁症部分是由于三种主要生物胺即多巴胺、去甲肾上腺素和/或 5-羟色胺中一种有缺陷。

除了单胺假说之外，很多争论都趋向重视考虑脑的整体功能，不再只考虑单个神经系统。在本文中，中枢胺能系统包括第二和第三信使系统的双重特异性作用的重要性已显现出来。

中枢神经系统障碍的内分泌基础

显然，主要单胺系统，即多巴胺、去甲肾上腺素和 5-羟色胺不能完全解释许多中枢神经系统障碍的病理生理学。具体地说，很明显，中枢神经系统障碍可能有一内分泌组分；下丘脑-垂体-肾上腺轴(HPA)在中枢神经系统障碍的病理生理学中

起着重要作用，包括促肾上腺皮质激素释放因子和糖皮质激素的作用。

在下丘脑-垂体-肾上腺轴中，下丘脑位于调节激素分泌的级系顶部。它制造和释放作用于脑底的垂体的肽(小氨基酸链)，刺激或抑制各种激素将垂体释放到血液中。这些激素包括生长激素、促甲状腺激素和促肾上腺皮质激素，控制靶腺释放出其它激素。除了在神经系统外起作用外，随垂体激素而释放的激素也反馈回到垂体和下丘脑。它们在那里输出抑制信号，以限制生物合成过量的激素。

中枢神经系统障碍

神经递质和内分泌失调与运动障碍(例如，帕金森氏病、杭廷顿舞蹈病、运动神经元病等等)、情绪失常(例如单相性抑郁症、双相性精神障碍、焦虑等等)以及涉及智力的疾病(例如阿耳茨海默氏病、雷维氏体痴呆症、精神分裂症等)相牵连。另外，这些系统也与许多其它疾病，例如昏迷、头部创伤、脑梗塞、癫痫、酒精中毒和代谢源的智力迟钝状态有连带关系，尤其见于小童。

中枢神经系统障碍可包括很多疾病，而且，相当多为遗传因子的疾病。中枢神经系统障碍的例子包括神经退化性疾病、精神病(psychotic disorders)、情绪失常、自闭症、药物依赖和酒精中毒、弱智及其它精神疾病(psychiatric diseases)，包括认知能力(cognitive)、焦虑、饮食、冲动控制(impulse-control)及人格障碍(personality disorders)。疾病的定义可采用精神病的诊断和统计手册第四版(Diagnosis and Statistical Manual of Mental Disorders, DSM-IV-R)的分类。

即使只考虑中枢神经系统障碍的小亚型(subset)，由于对精神病、精神分裂症和双相性精神障碍缺乏足够的治疗和对其疾病的分子基础又缺乏认识，很显然，需要在治疗发明上有新目标及改进的治疗方法。对于精神分裂症和双相性精神障碍，所有用于治疗精神分裂症的已知分子都有副作用，而且只对疾病症状起作用。故目前迫切需要有一些没有副作用、并针对涉及精神分裂症和双相性精神障碍因果机制的目标的新分子。所以，有助发现和表征这些目标的工具是必要的，而且十分有用。

精神分裂症和双相性精神障碍的家族聚集、从双胞胎和收养研究所得的证据以及全球发病率没有变化，都表明虽然环境危险因子在一定程度上是必要、充分或互相作用的原因，但精神分裂症和双相性精神障碍主要还是遗传条件。例如，普通人中有 1%会患上精神分裂症，但若一个人的祖父母有精神分裂症，则他患病的危险增大至约 3%；若父母一方有精神分裂症，则他患病的危险增大至约 10%；若父母

双方都有精神分裂症，则他患病的危险增大至约 40%。

染色体 13q31-q33 上的精神分裂症易感性基因的鉴定

鉴定涉及特定中枢神经系统障碍如精神分裂症等具体性状(trait)的基因主要有两种策略，目前均用于基因图谱，它们是：连锁分析(linkage analysis)和关联性研究(association studies)。连锁分析需要研究其成员有多人患病的家族，这种分析现在用于检测单个或寡基因固有性状。与之相反，关联性研究检查与阴性性状(T-)对照组相比无亲戚关系的性状(T+)个体的标记等位基因频率，这种研究通常用于检测多基因遗传。

遗传连锁或“连锁性”是基于在减数分裂过程中通过交换对染色体上两个相邻序列中的那一个含有最少重组的作分析。要进行遗传连锁，先使诸如微卫星标记等染色体标记准确定位于基因组上。遗传连锁分析根据系统树、疾病的传播和标记的传递计算靶基因与所用染色体标记重组的概率。因此，如果预定标记的特定等位基因与疾病一起传播的频率比带有该等位基因(重组水平在0和0.5之间)的机会要多，就可以推断在标记附近可找到所述靶基因。用这种技术可以确定一些表示家族性癌症的遗传倾向性的基因。为了能够用遗传连锁进行研究，患有遗传型疾病的家族成员必须符合“信息性”标准：每一代要有数个成员患病(可获得其体质上的DNA)，而且该家族最好有众多兄弟姐妹。

以前的连锁研究结果支持染色体 13 可能在 13q32 上隐藏一个精神分裂症易感性基因座的假设(Blouin JL 等人, *Nature Genetics*, 20: 70-73 页, 1998 年; Lin MW 等人, *Hum. Genet.*, 99(3): 417-410 页, 1997 年)。这些发现表明通过进行连锁研究可以确定染色体 13q32 位点上存在精神分裂症基因座。连锁研究也已成功用于绘制清楚显示孟德尔遗传模式以及具有高外显率的简单遗传性状,但这种方法存在许多缺点。首先,连锁分析受制于选择适合每种研究性状的遗传模型的可靠性。而且,用连锁分析所得到的分辨率也有限,需要进行补充研究来细化最初通过这种方法鉴定的常见 20Mb 区。此外,当应用于复杂的基因性状时,例如由于多基因和/或环境因子综合作用而产生的性状时,连锁分析也被证实无能为力。在这样的情况下,需要庞大的人力和财力才能招募到足够数目的患病家族成员,这是用连锁分析对这些情形所需要的。最后,连锁分析不能用于研究无法获得大量信息的家族的性状。

新的精神分裂症基因: g34872 (sbgl)

最近, 采用进行关联性研究的另一种方法, 而不是采用连锁研究来鉴定一种与新型精神分裂症和双相情感障碍有关的基因, 其被称为 g34872 基因, 位于染色体 13q31-q33 基因座上。这种变换方法包括在感兴趣的区内产生双等位标记(主要是单核苷酸多态性), 鉴定与精神分裂症连锁失衡的标记, 以及在没有亲戚关系的精神分裂症和双相情感障碍病人及对照组人群中进行关联性研究。

简而言之, 用位于染色体 13q31-q33 区的 27 个公共序列标记位点(STSs)构建一覆盖候选基因组区的细菌人工染色体(BAC)重迭群(contig), 以便筛选一个 7 种基因组相当的专有 BAC 文库。从这些材料产生新的序列标记位点, 构建该区的密集型物理图谱。总之, 275 个 STSs 可以鉴定 255 个 BAC, 这些 BAC 全部依大小排列, 并通过原位染色体杂交技术作图以作验证。以一些位于人染色体 13q31-q33 区的 BAC 插入片段的亚克隆插入端的局部排序产生新的双等位标记。在分析的第一阶段, 在精神分裂症病例和对照组中分析染色体 13q31-q33 候选基因座之间 9 个不同 BAC 上第一组 34 个双等位标记, 从而鉴定精神分裂症相关亚区(subregion)。第一次分析后, 如上所述, 产生另外一些双等位标记, 以绘制高度密集型靶区图谱。鉴定最小一组有 35 个 BAC, 使其完全排序, 结果出现几个重迭群, 包括一个 900kb 以上包含靶区序列的重迭群。

这些双等位标记用在关联性研究中, 为了使含有候选精神分裂症基因 g34872 的感兴趣的特定亚区更加细化。在对不同人群进行的一些研究中双等位标记被基因型分型, 以确认与亚区的关联性。关联性研究首先在法裔加拿大人群中精神分裂症病人和对照组的两种不同筛选样本, 其分别包括 139 病例和 141 个对照受试者以及 215 病例和 241 个对照受试者, 以及在阿根廷人群中双相情感障碍病例和对照受试者中进行。用这些人进行的一些研究后所得的结果显示大约 150kb 的基因组区与精神分裂症密切相关。这种关联在用美国精神分裂症人群的病例和对照受试者进行的独立研究以及其它阿根廷人和法裔加拿大人群的样本中再次得到证实。

已发现, 与精神分裂症相关的约 150kb 基因组区含有候选基因 g34872。除了对 g34872 基因的内含子-外显子结构进行特性分析外, 还鉴定了包括组织特异性 mRNA 剪接变异型在内的一系列 mRNA 剪接变异型, 而且证明了 mRNA 的存在。其后, 一肽片段自 g34872 多肽产物中衍生, 其氨基酸序列示于 SEQ ID No 5 中, 这表明当它们被注入小鼠腹膜内, 移位动作频率下降, 而刻板性动作(stereotypy)增加。于 2000 年 3 月 30 日提交、题为“与精神分裂症有关的基因、蛋白质及双等位标记”的共同待批美国专利申请号 09/539, 333 与共同待批国际专利申请号 PCT/IB00/00435 都对 g34872 基因的鉴定作了进一步的讨论, 以上两件申请均纳入

其整体内容作为参考。

与蛋白质相互作用的 g34872 及精神分裂症

目前迫切需要鉴定涉及精神分裂症和双相情感障碍的基因。同时也有必要鉴定一些包含在 g34872 途径的基因以及一些其产物在功能上与 g34872 基因产物相互作用的基因。这些基因在治疗精神分裂症或双相情感障碍中提供一些新的介入点,而且可以进一步研究和表征 g34872 基因和有关生物途径。认识这些基因和与精神分裂症有关的生物途径使研究人员能理解精神分裂症和双相情感障碍的病原,从而研究出针对这些病因的药物和治疗方案。此外,也急切需要寻找检测精神分裂症和双相情感障碍易感性以及预防或跟踪这些疾病发展的新方法。诊断工具也证明十分有用。事实上,越早鉴定受试者有演变精神分裂症的危险,就可以施以早期和/或预防性治疗。而且,准确评估药物的最终效能以及病人对该药物的最终耐量可使临床医生提高精神分裂症和双相情感障碍治疗方案的效益/风险比。

因此,本发明涉及任何一种为与 g34872 多肽相互作用的蛋白质编码的基因,本文称之 g34872 结合配偶体。本发明人通过酵母双杂交技术克隆了多种 g34872 结合配偶体。本发明人证实了所述 g34872 结合配偶体的组包含 D-氨基酸氧化酶。认识 g34872 结合配偶体有助开发治疗中枢神经系统疾病的药物,其由选自由 g34872、D-氨基酸氧化酶和任何一种其它 g34872 结合配偶体组成的组的基因介导。此外,认识 g34872 结合配偶体还可以提供 g34872、g34872 结合配偶体、g34872 结合配偶体复合物或 g34872 与其结合配偶体相互作用的检测装置。

与蛋白质相互作用的 g34872 及精神分裂症: D-氨基酸氧化酶

最早有记载的其中一种酶是 D-氨基酸氧化酶(DAO),它是在二十世纪三十年代发现的第二种黄素蛋白。D-氨基酸氧化酶将 D-氨基酸转化为相应的 α -酮酸。这是通过使 D-氨基酸催化脱氢生成亚氨基对应物和还原黄素产物复合物来完成。然后用双氧(再)氧化还原的黄素,生成 FADox 和双氧水,而亚氨基酸自动水解成酮酸和 NH_4^+ 。虽然大部分生物体和哺乳类组织都有 D-氨基酸氧化酶,但它在脊椎动物中的生理作用还不清楚。D-氨基酸氧化酶可以氧化 D-Met、D-Pro、D-Phe、D-Tyr、D-Ile、D-Leu、D-Ala 和 D-Val。D-Ser、D-Arg、D-His、D-正亮氨酸和 D-Trp 的氧化速度很低。D-鸟氨酸、顺-4-羟基-D-脯氨酸、D-Thr、D-Trp-甲酯、N-乙酰基-D-Ala 和 D-Lys 的氧化速度极低。D-Asp、D-Glu 及其衍生物、Gly 和所有 L-氨基酸不氧化(或者检测不到它们的氧化速度)。D-天冬氨酸氧化酶只能氧化 D-Asp、D-Glu 及

其下列衍生物：D-Asn、D-Gln、D-Asp-二甲酯和N-甲基-D-Asp。

中枢神经系统障碍是一类神经障碍。中枢神经系统障碍可以由药物诱发；可以因遗传罹病倾向性、感染或创伤而引起；或者由未知的病原引起。中枢神经系统障碍包括神经精神障碍、神经病和精神病(mental illnesses)；包括神经退化性疾病、行为障碍、认知障碍和认知情感障碍(cognitive affective disorders)。一些中枢神经系统障碍的临床表现是由于中枢神经系统机能异常(即因释放不适当水平的神经递质、神经递质受体的不适当性质和/或神经递质与神经递质受体之间不适当的相互作用)而引起的。另一些中枢神经系统障碍可由胆碱不正常分泌、多巴胺能不正常分泌、肾上腺素能不正常分泌和/或5-羟色胺能不正常分泌所造成。较常见的中枢神经系统障碍包括早衰性痴呆(早期发作的阿耳茨海默氏病)、老年性痴呆(阿耳茨海默氏型痴呆)、帕金森氏症包括帕金森氏病、杭廷顿舞蹈病、迟发性运动障碍、运动过度、躁狂症、注意力缺陷障碍、焦虑、诵读困难、精神分裂症、精神病、双相性精神障碍、抑郁症和图雷特氏综合征(Tourette's syndrome)。

神经递质和内分泌失调与运动障碍(例如，帕金森氏病、杭廷顿舞蹈病、运动神经元病等等)、情绪失常(例如单相性抑郁症、双相性精神障碍、焦虑等等)以及涉及智力的疾病(例如阿耳茨海默氏病、雷维氏体痴呆症、精神分裂症等)相牵连。另外，这些系统也与许多其它疾病，例如昏迷、头部创伤、脑梗塞、癫痫、酒精中毒和代谢源的智力迟钝状态有连带关系，尤其见于小童。

精神分裂症

在发达国家中，大约有1%成年人在其一生的某一点会出现精神分裂症。估计全球有四千五百万人患有精神分裂症，其中在发达国家超过三千三百万。此外，精神分裂症占去全部精神卫生费用的四分之一，并且精神医院的病床有三分之一是精神分裂症病人。大部分精神分裂症患者没有工作能力。社会为精神分裂症所耗的费用十分庞大。例如，美国治疗精神分裂症的直接费用估计接近国民生产总值的0.5%。精神分裂症患者的标准化死亡率估计比一般人高2到4倍，他们的预期寿命较一般人短20%。

精神分裂症患者最常见的死因是自杀(10%患者自杀)，风险比一般人高20倍。精神分裂症患者由于心脏病与呼吸和消化系统疾病而死亡的数目也不断上升。

精神分裂症包括一组具有“阳性”或“阴性”症状的精神病。阳性症状由幻觉、妄想和思想障碍组成；阴性症状包括情绪低落、缺乏意志和减少活动。

近年已经识别了许多生化异常情况，结果发展了一些基于神经递质的假说；

最为人熟知的一种假说是“多巴胺假说”，其中一种形式宣称 D2 受体水平上中脑边缘的多巴胺通路过度活动。然而，研究人员发现在多巴胺能系统的各种受体与精神分裂症之间的关联性并不一致。

双相性精神障碍

双相性精神障碍是较常见的精神障碍，发病率约占人口的 1.3%，据报导，精神病临床所见中大约一半的情绪失常者后果严重，甚至可能致残。业已发现，性别不同，双相性精神障碍也不同，取决于精神障碍的类型；例如，发现男性和女性同样可患上双相性精神障碍 I，但据说，双相性精神障碍 II 多见于女性。双相性精神障碍的发病年龄一般在十多岁，通常在患者二十岁出头的时候被诊断出。双相性精神障碍也发生于老年人，一般是由神经障碍或其它健康状况所造成。除了对患者的社会发展有严重影响外，据报导，约 15% 双相性精神障碍患者自杀成功。

双相性精神障碍以亢奋和经常性抑郁阶段为特点；亢奋阶段指躁狂或轻度躁狂，抑郁阶段可能交替发作或各个阶段混合发作，严重程度不同，持续时间也各异。由于双相性精神障碍有不同形式，表现为不同症状，对双相性精神障碍亚型的定义和整体概念扩展为包括以前被认为患有不同精神障碍的病人，都导致双相性精神障碍的分类成为一个广泛研究的主题。双相性精神障碍通常有一些共同的临床迹象、症状、治疗和一般精神病的神经生物特征，所以要作出准确诊断对精神病医生来说是一个颇大的挑战。另外，因为双相性精神障碍和各种情绪失常与精神障碍的病程差别很大，所以越早对疾病进行表征，制订长期控制疾病的方案变得至关重要。

社会对双相性精神障碍的负担十分庞大。与该疾病相关的躁狂症使工作能力下降，并导致精神病，最后的结果往往需要住院治疗。无论在直接还是间接费用上，这种疾病对病人家属和亲戚都造成极大负担，社会上对该疾病的歧视有时甚至延续几代。这样的歧视通常会造成孤立和忽视。此外，越早发病，被中断的教育与社会发展的效应越严重。

双相性精神障碍的 DSM-IV 分类根据躁狂症或轻度躁狂症的程度和持续时间分为四种精神障碍，以及根据明显的健康状况或其治疗或者药物滥用分为两种精神障碍。躁狂症的症状是情绪高涨或暴躁、注意力分散、行为冲动、过度活跃、夸大、过份乐观、脑筋转动很快以及说话急速。四种双相性精神障碍的征状是躁狂症有特定的程度和持续时间，DSM-IV 包括：

- 双相性精神障碍 I 型，包括躁狂症至少持续一星期的病人；
- 双相性精神障碍 II 型，包括轻度躁狂症至少持续四天的病人，特点是亢奋症状比躁狂症轻，病人从来未有过躁狂症，以前有过重度抑郁症；

- 其它未注明(NOS)之双相性精神障碍, 包括有双相性精神障碍 II 的症状但亢奋期不足 4 天的病人, 或者有轻度躁狂症但从来未有过重度抑郁症的病人;
- 循环性精神病, 包括有各种不符合轻度躁狂症或重度抑郁症标准的躁狂和抑郁症状, 但这些症状持续超过两年, 中间无症状的时间不超过两个月的病人。

按 DSM-IV 分类的其余两种双相性精神障碍是症状明显或由各种健康状况及其治疗所造成的精神障碍, 以及涉及或与药物滥用有关的精神障碍。医学上可引起双相性精神障碍的精神障碍通常包括内分泌失调和脑血管受损, 引起双相性精神障碍的医学疗法已知包括糖皮质激素和滥用刺激物。与使用或滥用药物相关的精神障碍称为“药物诱导的具有躁狂或混合特征的情绪失常”。

诊断双相性精神障碍富于挑战性。一个特别困难之处在于某些病人同时有躁狂症和焦虑症或抑郁症的混合状态, 但因为至少一个星期不是每天都全部符合躁狂症和重度抑郁所规定的标准, 所以不在 DSM-IV 分类之内。其它的困难还包括基于持续期对病人按 DSM-IV 组分类, 因为病人通常反复处于亢奋期和抑郁期, 但程度又不同。具体地说, 据报导, 在引起“快速循环”病程中用抗抑郁药可使病程变坏。使诊断变得更加困难的是以下一个事实: 双相性病人, 特别是处于称为第III阶段的躁狂症, 与双相性精神障碍病人具有相同的混乱思想和行为。此外, 精神病医生必须区分焦虑型抑郁症与混合型躁狂症; 重度抑郁症病人一般会焦躁, 导致有类似双极性症状。另一个复杂因素是双极性病人滥用药物特别是酒精的比例非常高。而酗酒病人中躁狂症的比率较低, 众所周知, 药物滥用者可表现为亢奋症状。所以, 这些困难会将双极性病人诊断为滥用药物。

抑郁症

抑郁症在医学上是一种严重疾病, 全球有三亿四千万人患有抑郁症。与正常的情感体验悲伤、沮丧、心情欠佳不同, 临床上的抑郁症是持续和长期的, 能严重影响个人的生活能力。所以, 抑郁症是世界上丧失能力的主因, 估计仅美国每年为此耗费五百三十亿美元。

抑郁症的症状包括情绪低落、对活动失去兴趣和悲观、食欲或体重改变、失眠或嗜睡、精神性运动激动或阻抑、疲倦或精神不振、无价值感受或过份自责、焦虑、精神难以集中或难作决定、以及重复想死亡或自杀。如果病人有五种或以上这些症状, 并且两星期内几乎每天日常生活与处事能力都受到影响, 就可以诊断为单

极性重度抑郁症(或重度抑郁情感障碍)。抑郁症一般出现在青春期的末期或成人期的早期；不过，最近有证据显示，三十年前出生的人可能在很早期就出现抑郁症。

世界卫生组织预测，到2020年抑郁症对发展中国家的人民而言是最严重的疾病，到那时抑郁症将是死亡和丧失能力的第二大诱因。除了它引起难以忍受的痛苦外，重度抑郁症最大的危险是自杀。估计25%的病人在患有重度抑郁症5年内试图自杀。另外，抑郁症也引发心脏病发作、中风、糖尿病和癌症等经常性和严重并发症。根据一项持续13年的近期研究，有重度抑郁症的个体心脏病发作可能是无此病史的人的4倍。

高达50%有中枢神经系统障碍如帕金森氏病、阿耳茨海默氏病的病人的一个特征是抑郁症。

在抑郁症病人的脑脊液中发现多巴胺代谢物HVA的水平很低。另外，多巴胺激动剂对抑郁症具有治疗反应。

目前的抗抑郁药用于针对多种抑郁症状，提高胺能突触中神经递质的浓度。不同的药理机制将抗抑郁药分为七种不同类型。二种典型机制是三环抗抑郁药(TCAs)和单胺氧化酶抑制剂(MAOIs)。最常用的药物是选择性5-羟色胺再吸收抑制剂(SSRIs)。其它三类抗抑郁药，如SSRIs提高5-羟色胺能的神经传递，但它们也有其它作用，即双重抑制5-羟色胺和去甲肾上腺素再吸收；抑制5-羟色胺-2拮抗性/再吸收； α_2 拮抗性加上5-羟色胺-2和-3拮抗性。选择性去甲肾上腺素和多巴胺再吸收抑制剂限定了一类新的抗抑郁药，它们对5-羟色胺系统无直接作用。

对于中枢神经系统障碍，如精神分裂症、双相性精神障碍、抑郁症和其它情绪失常，所有用于治疗已知分子都有副作用，而且只对疾病症状起作用。故迫切需要有一些没有相关副作用或减少副作用、并针对涉及这些中枢神经系统障碍病因机制的目标的新分子。有必要提供一种有用的方法，给容易患上或已患上中枢神经系统障碍的病人施加D-氨基酸氧化酶拮抗剂化合物来预防和治疗这样一些障碍。此外，也有必要提供一种有用的方法，给容易患上或已患上中枢神经系统障碍的病人施加D-天冬氨酸氧化酶拮抗剂化合物来预防和治疗这样一些障碍。

对于中枢神经系统障碍，如帕金森氏病、阿耳茨海默氏病和其它神经退化性疾病，目前可用于治疗的药物组合物十分有限，用于治疗已知分子都有副作用，而且只对疾病症状起作用。故迫切需要有一些没有相关副作用或减少副作用、并针对涉及这些中枢神经系统障碍病因机制的目标的新分子。有必要提供一种有用方法，给容易患上或已患上中枢神经系统障碍的病人施加D-氨基酸氧化酶活化剂化合物来预防和治疗这样一些障碍。此外，也有必要提供一种有用方法，给容易患上

或已患上中枢神经系统障碍的病人施加 D-天冬氨酸氧化酶活化剂化合物来预防和治疗这样一些障碍。

本发明的药物组合物用于预防和治疗这些中枢神经系统障碍。

治疗

由于目前还不能治愈中枢神经系统障碍，例如精神分裂症、双相性精神障碍、抑郁症和其它情绪失常，所以治疗的目的是使症状减轻，如果可能的话，减到缓解的程度。因为症状具有相似性，因此精神分裂症、抑郁症和双相性精神障碍往往用同一些药物治疗。两种疾病经常用抗精神病药和安定药治疗。

例如，对于精神分裂症，抗精神病药最常用和最有效。通常开给精神分裂症的抗精神病药有四大类。第一类安定药例如是氯丙嗪(chlorpromazine)，由于减轻了阳性(精神病)症状以及防止复发，对精神分裂症病人的治疗取得重大进步。服用氯丙嗪的病人已经能够离开精神病院，在社区或他们的家中生活。但这些药物远未为理想。约 20%到 30%的病人对它们完全无反应，其他人最终旧病复发。这些药物被称为安定药是因为它们对神经产生严重副作用，包括手臂和腿僵硬和颤抖、肌肉痉挛、身体移动反常和静坐不能(行走和坐立不安)。这些副作用令人心烦以致很多病人都拒绝服用这些药物。此外，安定药也不能改善精神分裂症所谓的阴性症状，副作用甚至会加重这些症状。所以，尽管安定药有着明显的有益效果，但是一些短期有良好反应的病人的生活和处事能力在最后甚至会恶化。

标准安定药的这些公知缺陷刺激人们寻找新治疗，并导致发现一类新的药物，称为非典型安定药。第一种非典型安定药是氯扎平(clozapine)，对大约三分之一对标准安定药无反应的病人有效。它似乎可减轻阴性和阳性症状，或对阴性症状的加重程度至少较标准安定药轻。而且，它对整体的生活和处事能力具有有益的效果，并且可降低精神分裂症病人自杀的危险。它也不会象标准安定药那样产生令人心烦的神经症状，或者提高激素催乳素的血液水平，催乳素过量可使女性月经异常和不孕，使男性性无能或乳房增大。许多不能忍受标准安定药的病人都能够服用氯扎平。然而，氯扎平也有严重限制。它最初从市场上收回是因为它可引起粒细胞缺乏症，这是一种有可能致命的不能产生白血细胞的疾病。粒细胞缺乏症乃是一种需要小心监控并且进行周期性验血的预兆。氯扎平还可能引起疾病发作或其它不良副作用(例如，渴睡、血压下降、流涎、尿床和体重增加)。所以，它通常只用于对其它药物无反应的病人。

研究人员已研制成第三类抗精神症药，它具有氯扎平的功效但没有其缺点。

其中一种是利培酮(risperidone)。早期研究表明,它对阳性症状的效果像标准安定药一样,对阴性症状的效果可能更好。它对神经产生的副作用比氯扎平多,但较标准安定药少。然而,它会提高催乳素水平。现在利培酮可开药方给更多的精神病患者,越来越多医生似乎给对标准药物无反应的病人优先考虑采用利培酮而不是氯扎平,因为他们认为它较安全。另一种新药是奥扎平(Olanzapine),它对阳性症状的效果至少与标准安定药一样有效,对阴性症状的效果则更好。正常临床剂量时它的神经副作用很少,它也不会显著提高催乳素水平。虽然它没有氯扎平大部分副作用,但一些服用奥扎平的病人可能变得镇静或晕眩、引发口腔干燥、或体重增加。在一些罕见的情况下,肝脏功能测试会有瞬时异常。

精神分裂症的结果研究通常基于医院的治疗研究,可能不能代表精神分裂症患者人群。极端的结果是 20%的患者似乎发作过一次精神病后完全康复,而 14-19%的病人会发展为慢性持续性精神病,不会完全康复。一般来说,五年的临床结果似乎遵循以下三种情况的规律:约 35%的病病人的结果很差;36%的病病人的结果良好,而其余的病病人的结果为中间情况。精神分裂症预后似乎在五年后不会变差。

不论何种原因,不断有证据显示精神分裂症不在发病初期进行长期治疗可能对结果产生负面影响。不过,药物的选用通常会延迟病人第一次发病的时间。病人可能没有意识到他们已经病了,或者他们可能害怕寻求协助;家人有时希望问题简单消失或者无法说服病人去治疗;当潜在的副作用使诊断不确定时,医生可能犹豫开出抗精神药处方。事实上,在疾病最早表现时,很难将精神分裂症与双极性躁狂-抑郁精神障碍、重度抑郁症、与药物滥用有关的精神障碍、以及与压力有关的精神障碍区分开来。因为这些疾病的最佳治疗是不同的,所以这些精神障碍的长期预后也使治疗的开始时间不同。

对于中枢神经系统障碍,如精神分裂症、双相性精神障碍、抑郁症和其它情绪失常,用于治疗已知分子都有副作用,而且只对疾病症状起作用。故迫切需要有一些没有副作用、并针对涉及中枢神经系统障碍因果机制的目标的新分子。所以,有助发现和表征这些目标的工具是必要的,而且十分有用。

精神分裂症和双相性精神障碍的家族聚集、从双胞胎和收养研究所得的证据以及全球发病没有变化,都表明虽然环境危险因子在一定程度上是必要、充分或相互作用的原因,但精神分裂症和双相性精神障碍主要还是遗传条件。例如,普通人中有 1%会患上精神分裂症,但若一个人的祖父母有精神分裂症,则他患病的危险增大至约 3%;若父母一方有精神分裂症,则他患病的危险增大至约 10%;若父母双方都有精神分裂症,则他患病的危险增大至约 40%。

所以，目前迫切需要鉴定涉及这些中枢神经系统障碍中包含的基因。认识这些基因可使研究人员能理解精神分裂症、抑郁症、双相性精神障碍以及其它情绪失常的病原，从而可能研究出针对这些疾病病因的药物和治疗方案，而不仅仅针对它们的症状。

此外，也很需要有检测这些中枢神经系统障碍如精神分裂症和双相性精神障碍、抑郁症易感性以及预防或跟踪这些疾病发展的新方法。诊断工具也证明十分有用。事实上，越早识别受试者有演变精神分裂症的危险，就可以施以早期和/或预防性治疗。而且，准确评估药物的最终效能以及病人对该药物的最终耐量可使临床医生提高中枢神经系统障碍治疗方案的效益/风险比，例如精神分裂症、抑郁症、双相性精神障碍以及其它情绪失常的效益/风险比。

发明内容

本发明基于新的多态性，包括位于人染色体 13q31-q33 基因座上双等位标记的鉴定、位于人染色体 13q31-q33 基因座上与精神分裂症相关的新基因的鉴定和特性分析、以及位于人染色体 13q31-q33 基因座上双等位标记等位基因与一组人受试者已确定和表征的疾病之间的基因相关性。美国专利申请号 09/539, 333 与国际专利申请号 PCT/IB00/00435 提交了新的多态性和与精神分裂症相关的基因，它们的公开资料纳入其整体内容作为参考。

可按照本发明的方法治疗的中枢神经系统障碍包括早衰性痴呆(早期发作的阿耳茨海默氏病)、老年性痴呆(阿耳茨海默氏型痴呆)、帕金森氏症包括帕金森氏病、杭廷顿舞蹈病、迟发性运动障碍、运动过度、躁狂症、注意力缺陷活动过度障碍(ADHD)、注意力缺陷障碍(ADD)、焦虑症、诵读困难、精神病、精神分裂症、双相性精神障碍、重度抑郁症、躁狂症、轻度躁狂症、抑郁症、孤独症、滥用药物、过度攻击(excessive aggression)、抽搐症和图雷特氏综合征。本发明优选的中枢神经系统障碍包括精神分裂症、抑郁症和双相性精神障碍。精神分裂症和精神分裂症样精神障碍的较优选例子包括：精神分裂症(紧张型)、精神分裂症(无组织型)、精神分裂症(妄想型)、精神分裂症(混合型)、精神分裂症(残余型)、精神分裂症样精神障碍、短期反应性精神病、情感分裂性精神障碍、诱发性精神病、精神分裂型人格障碍、精神分裂样人格障碍、妄想型人格障碍和妄想(类偏执狂)性精神障碍。

本发明涉及给容易患上中枢神经系统障碍的受试者提供这样一种精神障碍的治疗方法、给患有中枢神经系统障碍的受试者提供治疗方法。具体地说，该方法包括给患者施加一定量的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶拮抗剂或抑制剂化合物，一定程度上有效制止或改善中枢神经系统障碍的进展、制止或改善中枢神经系

统障碍的症状以及制止或改善中枢神经系统障碍的復發。

本发明涉及给容易患上中枢神经系统障碍的受试者提供这样一种精神障碍的预防方法、给患有中枢神经系统障碍的受试者提供治疗方法。具体地说，该方法包括给患者施加一定量的 D-氨基酸氧化酶和 D-天冬氨酸氧化酶拮抗剂化合物，一定程度上有效预防中枢神经系统障碍的进展(即提供保护作用)、预防中枢神经系统障碍的症状和预防中枢神经系统障碍的复发。

本发明还涉及 D-氨基酸氧化酶的基因组序列、D-氨基酸氧化酶基因中发现的新型外显子、D-氨基酸氧化酶基因中发现的新型多态双等位标记(SNPs)、容易患上中枢神经系统障碍的人的检测方法，拮抗、抑制、或降低 D-氨基酸氧化酶活性的新方法，增加、激动或促进 D-氨基酸氧化酶活性的新方法以及影响 D-氨基酸氧化酶活性的新型组合物。本发明还涉及核酸分子，其含有编码 g34872(sbg1)蛋白质的新型人基因的基因组序列、由该核酸分子编码的蛋白质及其抗体，共同待批的美国专利申请号 09/539, 333 与国际专利申请号 PCT/IB00/00435 对此作了叙述，它们的公开资料纳入其整体内容作为参考。本发明也涉及编码 g34872、D-氨基酸氧化酶和 D-天冬氨酸氧化酶蛋白质、及其变异型的 cDNA 序列。与 g34872、D-氨基酸氧化酶和 D-天冬氨酸氧化酶基因组或 cDNA 序列特异性杂交的寡核苷酸探针或引物，及使用所述引物和探针的 DNA 扩增和检测方法也是本发明的一部分内容。

本发明的另一个目的包括含有上述任何一段核酸序列的重组载体，特别是含有 g34872、D-氨基酸氧化酶和 D-天冬氨酸氧化酶调控序列或编码 g34872、D-氨基酸氧化酶和 D-天冬氨酸氧化酶蛋白质的序列的重组载体，以及含有所述核酸序列或重组载体的细胞宿主和转基因非人的动物。

本发明还涉及 g34872、D-氨基酸氧化酶和 D-天冬氨酸氧化酶基因的双等位标记及其应用。本发明包括用于本发明双等位标记基因分型(genotyping)的探针和引物。

本发明的一实施例包括附着于固相支持物上的任何一种本发明的多核苷酸。多核苷酸可以含有本说明书所揭示的序列；所述多核苷酸可任选地包括任何本说明书所叙述的任何一种多核苷酸，可任选地由或基本上由本说明书所叙述的任何一种多核苷酸组成；可任选地在杂交分析、排序分析、微排序分析或基于酶的错配检测分析中进行所述的确定；所述多核苷酸可任选地附着于固相支持物、阵列或选通式阵列(addressable array)；所述多核苷酸可任选地带上标记。

最后，本发明的目的在于基于 g34872、D-氨基酸氧化酶或 D-天冬氨酸氧化酶核苷酸和多核苷酸在疾病中的作用，提供药物筛选分析和治疗精神分裂症、双相性

精神障碍或相关中枢神经系统障碍的药物的筛选方法。本发明的一个目的在于基于 g34872、D-氨基酸氧化酶或 D-天冬氨酸氧化酶在疾病中的作用，提出精神分裂症动物模型，包括小鼠、灵长类动物、非人灵长类动物双相性精神障碍或相关中枢神经系统障碍。本发明也涉及抑制 g34872、D-氨基酸氧化酶或 D-天冬氨酸氧化酶表达的物质或分子的筛选方法，以及应用这些筛选方法来筛选与 g34872、D-氨基酸氧化酶或 D-天冬氨酸氧化酶多肽相互作用，或者调节 g34872、D-氨基酸氧化酶或 D-天冬氨酸氧化酶多肽的活性的物质或分子。

如上所述，本发明的某些方面基于精神分裂症和双相性精神障碍与 g34872 基因和 D-氨基酸氧化酶基因的双等位标记等位基因之间的基因关联性的鉴定。本发明提供一些合适的方法来进一步建立 g34872 和 D-氨基酸氧化酶基因座的双等位标记等位基因与由于施加作用于包括精神分裂症、抑郁症或双相性精神障碍在内的中枢神经系统障碍或症状的药物，或者精神分裂症或双相性精神障碍症状的药物，包括氯丙嗪、氯扎平、利培酮、奥扎平、舍吲哚(sertindole)、喹的平(quetiapine)和齐拉西酮(ziprasidone)而产生的副作用或有益效果之间的其它基因关联性。

本发明提供一些合适的方法来进一步建立特定性状(trait)的 D-氨基酸氧化酶与 g34872 的双等位标记等位基因之间的其它基因关联性。本发明提供一些人对精神分裂症或双相性精神障碍遗传易感性的分子检测方法和产物。它们可用于疾病的诊断、分期、预后和监控，所述方法还可包括在治疗方法之内。本发明也提供有效设计和评估合适的治疗方案，包括药物的最佳使用方法、潜在的候选新药物的筛选的个人化策略。

本发明的一优选实施例包括对有需要的中枢神经系统障碍病人实施的治疗方法，该方法包括给所述病人施加一有效量的含有 D-氨基酸氧化酶拮抗剂或抑制剂或者 D-天冬氨酸氧化酶抑制剂或拮抗剂的组合物或化合物。

较优选是精神病的治疗方法，该方法包括给精神病病人施加一有效治疗量的含有 D-氨基酸氧化酶拮抗剂或抑制剂或者 D-天冬氨酸氧化酶抑制剂或拮抗剂的组合物或化合物。

较优选是精神分裂症的治疗方法，该方法包括给精神分裂症病人施加一有效治疗量的含有 D-氨基酸氧化酶拮抗剂或抑制剂或者 D-天冬氨酸氧化酶抑制剂或拮抗剂的组合物或化合物。

较优选是双相性精神障碍的治疗方法，该方法包括给双相性精神障碍病人施加一有效治疗量的含有 D-氨基酸氧化酶拮抗剂或抑制剂或者 D-天冬氨酸氧化酶抑制剂或拮抗剂的组合物或化合物。

本发明一优选实施例包括对有需要的中枢神经系统障碍病人实施的治疗方法，该方法包括给所述病人施加一有效量的含有 D-氨基酸氧化酶拮抗剂或抑制剂或者 D-天冬氨酸氧化酶拮抗剂或抑制剂的组合物或化合物。

较优选是精神病的治疗方法，该方法包括给精神病病人施加一有效治疗量的含有 D-氨基酸氧化酶拮抗剂或抑制剂或者 D-天冬氨酸氧化酶拮抗剂或抑制剂的组合物或化合物。

较优选是精神分裂症的治疗方法，该方法包括给精神分裂症病人施加一有效治疗量的含有 D-氨基酸氧化酶拮抗剂或抑制剂或者 D-天冬氨酸氧化酶拮抗剂或抑制剂的组合物或化合物。

较优选是双相性精神障碍的治疗方法，该方法包括给双相性精神障碍病人施加一有效治疗量的含有 D-氨基酸氧化酶拮抗剂或抑制剂或者 D-天冬氨酸氧化酶抑制剂或拮抗剂的组合物或化合物。

本发明一优选实施例包括对有需要的中枢神经系统障碍病人实施的治疗方法，该方法包括给所述病人施加一有效量的含有 g34872 拮抗剂或抑制剂的组合物或化合物。

较优选是精神病的治疗方法，该方法包括给精神病病人施加一有效治疗量的含有 g34872 拮抗剂或抑制剂的组合物或化合物。

较优选是精神分裂症的治疗方法，该方法包括给精神分裂症病人施加一有效治疗量的含有 g34872 拮抗剂或抑制剂的组合物或化合物。

较优选是双相性精神障碍的治疗方法，该方法包括给双相性精神障碍病人施加一有效治疗量的含有 g34872 拮抗剂或抑制剂的组合物或化合物。

本发明一优选实施例包括对有需要的中枢神经系统障碍病人实施的治疗方法，该方法包括给所述病人施加一有效量的含有 D-氨基酸氧化酶拮抗剂或抑制剂或者 D-天冬氨酸氧化酶拮抗剂或抑制剂的组合物或化合物，并结合 g34872 拮抗剂或抑制剂组合物或化合物。

较优选是精神病的治疗方法，该方法包括给精神病病人施加一治疗有效量的含有 D-氨基酸氧化酶拮抗剂或抑制剂或者 D-天冬氨酸氧化酶拮抗剂或抑制剂的组合物或化合物，并结合 g34872 拮抗剂或抑制剂组合物或化合物。

较优选是精神分裂症的治疗方法，该方法包括给精神分裂症病人施加一有效治疗量的含有 D-氨基酸氧化酶拮抗剂或抑制剂或者 D-天冬氨酸氧化酶拮抗剂或抑制剂的组合物或化合物，并结合 g34872 拮抗剂或抑制剂组合物或化合物。

较优选是双相性精神障碍的治疗方法，该方法包括给双相性精神障碍病人施

加一有效治疗量的含有 D-氨基酸氧化酶拮抗剂或抑制剂或者 D-天冬氨酸氧化酶拮抗剂或抑制剂的组合物或化合物,并结合 g34872 拮抗剂或抑制剂组合物或化合物。

本发明一优选实施例包括对有需要的中枢神经系统障碍病人实施的治疗方法,该方法包括给所述病人施加一有效量的结合有 D-氨基酸氧化酶拮抗剂或抑制剂、D-天冬氨酸氧化酶拮抗剂或抑制剂和 g34872 拮抗剂或抑制剂组合物或化合物的组合物或化合物。

较优选是精神病的治疗方法,该方法包括给精神病病人施加一治疗有效量的结合有 D-氨基酸氧化酶拮抗剂或抑制剂、D-天冬氨酸氧化酶拮抗剂或抑制剂和 g34872 拮抗剂或抑制剂组合物或化合物的组合物或化合物。

较优选是精神分裂症的治疗方法,该方法包括给精神分裂症病人施加一有效治疗量的结合有 D-氨基酸氧化酶拮抗剂或抑制剂、D-天冬氨酸氧化酶拮抗剂或抑制剂和 g34872 拮抗剂或抑制剂组合物或化合物的组合物或化合物。

较优选是双相性精神障碍的治疗方法,该方法包括给双相性精神障碍病人施加一有效治疗量的结合有 D-氨基酸氧化酶拮抗剂或抑制剂、D-天冬氨酸氧化酶拮抗剂或抑制剂和 g34872 拮抗剂或抑制剂组合物或化合物的组合物或化合物。

本发明一优选实施例包括对有需要的中枢神经系统障碍病人实施的治疗方法,该方法包括给所述病人施加一有效量的含有至少以下一种物质的组合物或化合物: D-氨基酸氧化酶拮抗剂或抑制剂、D-天冬氨酸氧化酶拮抗剂或抑制剂,或者 g34872 拮抗剂或抑制剂组合物或化合物。

较优选是精神病的治疗方法,该方法包括给精神病病人施加一治疗有效量的含有至少以下一种物质的组合物或化合物: D-氨基酸氧化酶拮抗剂或抑制剂、D-天冬氨酸氧化酶拮抗剂或抑制剂,或者 g34872 拮抗剂或抑制剂组合物或化合物。

较优选是精神分裂症的治疗方法,该方法包括给精神分裂症病人施加一有效治疗量的含有至少以下一种物质的组合物或化合物: D-氨基酸氧化酶拮抗剂或抑制剂、D-天冬氨酸氧化酶拮抗剂或抑制剂,或者 g34872 拮抗剂或抑制剂组合物或化合物。

较优选是双相性精神障碍的治疗方法,该方法包括给双相性精神障碍病人施加一有效治疗量的含有至少以下一种物质的组合物或化合物: D-氨基酸氧化酶拮抗剂或抑制剂、D-天冬氨酸氧化酶拮抗剂或抑制剂,或者 g34872 拮抗剂或抑制剂组合物或化合物。

应该指出的是,本发明优先排除对有需要的中枢神经系统障碍、精神病、精神分裂症或双相性精神障碍病人实施的治疗方法中所用的公知或固有起到抑制或

拮抗 D-氨基酸氧化酶、D-天冬氨酸氧化酶或 g34872 的本领域已知的组合物或化合物。

本发明一优选实施例涉及 D-氨基酸氧化酶活性的抑制方法。此外，本发明涉及通过抑制 D-氨基酸氧化酶活性来治疗精神分裂症的方法。较优选是通过用含有酮亚胺的组合物抑制 D-氨基酸氧化酶活性来治疗精神分裂症的方法。

本发明的另一优选实施例涉及 D-天冬氨酸氧化酶活性的抑制方法。此外，本发明涉及通过抑制 D-天冬氨酸氧化酶活性来治疗精神分裂症的方法。较优选是通过用含有酮亚胺的组合物来抑制 D-天冬氨酸氧化酶活性来治疗精神分裂症的方法。

本发明的另一优选实施例涉及抑制 D-氨基酸氧化酶与 g34872 之间相互作用的方法。

本发明的另一优选实施例涉及抑制 g34872 与 D-天冬氨酸氧化酶之间相互作用的方法。

本发明的另一优选实施例涉及 SEQ ID NO: 7、8、9、10 或 18 的 D-氨基酸氧化酶多肽的任何多肽片段，其拮抗所述 D-氨基酸氧化酶多肽与 SEQ ID NO: 14 的 g34872 多肽或其片段之间相互作用。较优选 D-氨基酸氧化酶多肽片段含有 SEQ ID NO: 7 的氨基酸 23-347。较优选 D-氨基酸氧化酶多肽片段含有 SEQ ID NO: 7 的氨基酸 227-347。较优选 D-氨基酸氧化酶多肽片段含有 SEQ ID NO: 7 的氨基酸 31-347。较优选 D-氨基酸氧化酶多肽片段含有 SEQ ID NO: 7 的氨基酸 51-347。较优选 D-氨基酸氧化酶多肽片段含有 SEQ ID NO: 7 的氨基酸 66-347。较优选 D-氨基酸氧化酶多肽片段含有 SEQ ID NO: 7 的氨基酸 101-347。较优选 D-氨基酸氧化酶多肽片段含有 SEQ ID NO: 7 的氨基酸 126-347。较优选 D-氨基酸氧化酶多肽片段含有 SEQ ID NO: 7 的氨基酸 146-347。较优选 D-氨基酸氧化酶多肽片段含有 SEQ ID NO: 7 的氨基酸 175-347。较优选 D-氨基酸氧化酶多肽片段含有 SEQ ID NO: 7 的氨基酸 180-347。较优选 D-氨基酸氧化酶多肽片段含有 SEQ ID NO: 7 的氨基酸 1-189。较优选 D-氨基酸氧化酶多肽片段含有 SEQ ID NO: 7 的氨基酸 1-205。较优选 D-氨基酸氧化酶多肽片段含有 SEQ ID NO: 7 的氨基酸 31-189。较优选 D-氨基酸氧化酶多肽片段含有 SEQ ID NO: 7 的氨基酸 31-205。较优选 D-氨基酸氧化酶多肽片段含有 SEQ ID NO: 7 的氨基酸 84-205。

本发明另一优选实施例涉及与 D-氨基酸氧化酶多肽或其片段结合的组合物。较优选与 D-氨基酸氧化酶多肽片段结合的组合物含有 SEQ ID NO: 7 的氨基酸 227-347。较优选与 D-氨基酸氧化酶多肽片段结合的组合物含有 SEQ ID NO: 7 的

氨基酸 227-347。较优选与 D-氨基酸氧化酶多肽片段结合的组合物含有 SEQ ID NO: 7 的氨基酸 31-347。较优选与 D-氨基酸氧化酶多肽片段结合的组合物含有 SEQ ID NO: 7 的氨基酸 51-347。较优选与 D-氨基酸氧化酶多肽片段结合的组合物含有 SEQ ID NO: 7 的氨基酸 66-347。较优选与 D-氨基酸氧化酶多肽片段结合的组合物含有 SEQ ID NO: 7 的氨基酸 101-347。较优选与 D-氨基酸氧化酶多肽片段结合的组合物含有 SEQ ID NO: 7 的氨基酸 126-347。较优选与 D-氨基酸氧化酶多肽片段结合的组合物含有 SEQ ID NO: 7 的氨基酸 146-347。较优选与 D-氨基酸氧化酶多肽片段结合的组合物含有 SEQ ID NO: 7 的氨基酸 175-347。较优选与 D-氨基酸氧化酶多肽片段结合的组合物含有 SEQ ID NO: 7 的氨基酸 180-347。较优选与 D-氨基酸氧化酶多肽片段结合的组合物含有 SEQ ID NO: 7 的氨基酸 1-189。较优选与 D-氨基酸氧化酶多肽片段结合的组合物含有 SEQ ID NO: 7 的氨基酸 1-205。较优选与 D-氨基酸氧化酶多肽片段结合的组合物含有 SEQ ID NO: 7 的氨基酸 31-189。较优选与 D-氨基酸氧化酶多肽片段结合的组合物含有 SEQ ID NO: 7 的氨基酸 31-205。较优选与 D-氨基酸氧化酶多肽片段结合的组合物含有 SEQ ID NO: 7 的氨基酸 84-205。

另一优选实施例涉及精神分裂症的治疗方法，所述方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与 D-氨基酸氧化酶多肽或其片段结合的组合物。较优选的精神分裂症的治疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 23-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的精神分裂症的治疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 227-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的精神分裂症的治疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 31-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的精神分裂症的治疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 51-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的精神分裂症的治疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 66-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的精神分裂症的治疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 101-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的精神分裂症的治疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 126-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的精神分裂症的治

疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 146-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的精神分裂症的治疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 175-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的精神分裂症的治疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 180-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的精神分裂症的治疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 1-189 的 D-氨基酸氧化酶多肽结合的组合物。较优选的精神分裂症的治疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 1-205 的 D-氨基酸氧化酶多肽结合的组合物。较优选的精神分裂症的治疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 31-189 的 D-氨基酸氧化酶多肽结合的组合物。较优选的精神分裂症的治疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 31-205 的 D-氨基酸氧化酶多肽结合的组合物。较优选的精神分裂症的治疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 84-205 的 D-氨基酸氧化酶多肽结合的组合物。

另一优选实施例涉及双极性精神障碍的治疗方法，所述方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与 D-氨基酸氧化酶多肽或其片段结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 23-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 227-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 31-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 51-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 66-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 101-347 的 D-氨

基酸氧化酶多肽结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 126-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 146-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 175-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 180-347 的 D-氨基酸氧化酶多肽结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 1-189 的 D-氨基酸氧化酶多肽结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 1-205 的 D-氨基酸氧化酶多肽结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 31-189 的 D-氨基酸氧化酶多肽结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 31-205 的 D-氨基酸氧化酶多肽结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 7 的氨基酸 84-205 的 D-氨基酸氧化酶多肽结合的组合物。

本发明另一优选实施例涉及与 SEQ ID NO: 14 的 g34872 多肽或其片段结合的组合物。较优选与 g34872 多肽结合的组合物含有 SEQ ID NO: 14 的氨基酸 65-153 或其片段。较优选的是与 SEQ ID NO: 16 的多肽或其片段结合的组合物。

另一优选实施例涉及精神分裂症的治疗方法，所述方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与 SEQ ID NO: 14 的 g34872 多肽或其片段结合的组合物。较优选的精神分裂症的治疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 14 的氨基酸 65-153 或其片段的 g34872 多肽结合的组合物。较优选的精神分裂症的治疗方法包括给精神分裂症病人施加一有效治疗量的组合物，其含有与 SEQ ID NO: 16 的多肽或其片段结合的组合物。

另一优选实施例涉及双极性精神障碍的治疗方法，所述方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与 SEQ ID NO: 14 的 g34872 多肽

或其片段结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与包含 SEQ ID NO: 14 的氨基酸 65-153 或其片段的 g34872 多肽结合的组合物。较优选的双极性精神障碍的治疗方法包括给双极性精神障碍病人施加一有效治疗量的组合物，其含有与 SEQ ID NO: 16 的多肽或其片段结合的组合物。

本发明另一优选实施例涉及 SEQ ID NO: 14 的 g34872 多肽的任何多肽片段，其拮抗所述 g34872 多肽或其片段与 D-氨基酸氧化酶多肽或其片段之间的相互作用。较优选的是通过 g34872 多肽拮抗 D-氨基酸氧化酶增加活性的任何 g34872 多肽片段。较优选的是含有 SEQ ID NO: 16 的氨基酸的 g34872 多肽片段。

本发明另一优选实施例涉及拮抗 SEQ ID NO: 14 的 g34872 多肽或其片段与 SEQ ID NO: 7-10 或 18 的 D-氨基酸氧化酶多肽或其片段之间的相互作用的组合物。

本发明另一优选实施例涉及拮抗 SEQ ID NO: 14 的 g34872 多肽或其片段与 SEQ ID NO: 21 或 22 的 D-天冬氨酸氧化酶多肽或其片段之间的相互作用的组合物。

本发明又一优选实施例涉及用 g34872 多肽或其片段增加 D-氨基酸氧化酶活性的方法。此外，本发明还涉及用 g34872 多肽或其片段增加 D-天冬氨酸氧化酶活性的方法。

本发明另一优选实施例涉及抑制 D-氨基酸氧化酶糖基化的方法。

本发明另一优选实施例涉及增强 D-氨基酸氧化酶多聚化的方法。

本发明另一优选实施例涉及抑制 D-氨基酸氧化酶翻译的方法。

本发明另一优选实施例涉及 D-氨基酸氧化酶变异型差别性鉴定的方法。

本发明一优选实施例涉及降低、抑制或拮抗 D-氨基酸氧化酶活性的组合物或化合物。较优选的组合物或化合物是 D-氨基酸氧化酶活性的竞争性抑制剂或拮抗剂。较优选的组合物或化合物是 D-氨基酸氧化酶活性的非竞争性抑制剂或拮抗剂。较优选的组合物或化合物是 D-氨基酸氧化酶活性的无竞争性抑制剂或拮抗剂。较优选的组合物或化合物是 D-氨基酸氧化酶活性的变构抑制剂或拮抗剂。较优选的组合物或化合物是 D-氨基酸氧化酶活性的可逆性抑制剂或拮抗剂。较优选的组合物或化合物是 D-氨基酸氧化酶活性的不可逆抑制剂或拮抗剂。

本发明一优选实施例涉及降低、抑制或拮抗 D-天冬氨酸氧化酶活性的组合物或化合物。较优选的组合物或化合物是 D-天冬氨酸氧化酶活性的竞争性抑制剂或拮抗剂。较优选的组合物或化合物是 D-天冬氨酸氧化酶活性的非竞争性抑制剂或拮抗剂。较优选的组合物或化合物是 D-天冬氨酸氧化酶活性的无竞争性抑制剂或拮抗剂。较优选的组合物或化合物是 D-天冬氨酸氧化酶活性的变构抑制剂或拮抗剂。

剂。较优选的组合物或化合物是 D-天冬氨酸氧化酶活性的可逆性抑制剂或拮抗剂。较优选的组合物或化合物是 D-天冬氨酸氧化酶活性的不可逆抑制剂或拮抗剂。较优选的组合物或化合物是能够降低、抑制或拮抗 D-氨基酸氧化酶和 D-天冬氨酸氧化酶活性。

较优选的是利用能降低、抑制或拮抗 D-氨基酸氧化酶活性的组合物或化合物来治疗中枢神经系统障碍的方法。较优选的是利用能降低、抑制或拮抗 D-天冬氨酸氧化酶活性的组合物或化合物来治疗中枢神经系统障碍的方法。较优选的是利用能降低、抑制或拮抗 D-氨基酸氧化酶或 D-天冬氨酸氧化酶活性的组合物或化合物来治疗中枢神经系统障碍的方法。较优选的是利用能降低、抑制或拮抗 D-氨基酸氧化酶或 D-天冬氨酸氧化酶活性的组合物或化合物来治疗中枢神经系统障碍的方法。较优选的是利用能降低、抑制或拮抗 D-氨基酸氧化酶活性的第一组合物或化合物与能降低、抑制或拮抗 D-天冬氨酸氧化酶活性的第二组合物或化合物结合来治疗中枢神经系统障碍的方法。较优选的是利用能降低、抑制或拮抗 D-氨基酸氧化酶活性的组合物或化合物与另一种组合物结合来治疗中枢神经系统障碍的方法。较优选的是利用能降低、抑制或拮抗 D-氨基酸氧化酶活性的组合物或化合物与另一种所述中枢神经系统障碍治疗所用的常规组合物结合来治疗中枢神经系统障碍的方法。较优选的是利用能降低、抑制或拮抗 D-氨基酸氧化酶活性的组合物或化合物与另一种治疗所述中枢神经系统障碍不相关的组合物结合来治疗中枢神经系统障碍的方法。较优选的是利用能降低、抑制或拮抗 D-天冬氨酸氧化酶活性的组合物或化合物与另一种组合物结合治疗中枢神经系统障碍的方法。较优选的是利用能降低、抑制或拮抗 D-天冬氨酸氧化酶活性的组合物或化合物与另一种治疗所述中枢神经系统障碍所用的常规组合物结合治疗中枢神经系统障碍的方法。较优选的是利用能降低、抑制或拮抗 D-天冬氨酸氧化酶活性的组合物或化合物与另一种治疗所述中枢神经系统障碍不相关的组合物结合治疗中枢神经系统障碍的方法。

本发明能降低、抑制或拮抗 D-氨基酸氧化酶或 D-天冬氨酸氧化酶活性的优选组合物或化合物选自但不限于下表：

- i. IRI, 2-氧代-3-戊炔酸盐;
- ii. CMI, 氨基胍(脒基胍; 氨甲基亚氨基胍; 匹马吉定(Pimagedine); GER 11; 胍基羧亚氨基酰胺)或盐酸盐(脒基胍盐酸盐), 碳酸氢盐, 硝酸盐, 硫酸盐(2: 1), 亚硫酸盐(1: 1), 及其半硫酸盐;
- iii. FI, 苯甲酸;
- iv. FI, 苯甲酸钠;

- v. FI, 2-氨基苯甲酸盐;
- vi. FI, 3-氨基苯甲酸盐;
- vii. FI, 4-氨基苯甲酸盐 (对氨基苯甲酸盐, PABA, 维生素 Bx, 维生素 H1);
- viii. CMI, 甲基乙二醛双(脒脞) (也称: 甲基 GAG; 米托脒脞; 1, 1'-((甲基联二亚甲基)双次氨基)二脒; 脒基羧亚氨基酰胺, 2, 2'-(1-甲基-1, 2-联二亚甲基)双-; 丙酮醛双(脒脞); Megag; 米托脒脞[INN-Spanish]; 脒, 1, 1'-((甲基联二亚甲基)双次氨基)二-; 1, 1'-((甲基联二亚甲基)双次氨基)二脒);
- ix. CMI, 甲基乙二醛双(脒基脞), 二盐酸盐;
- x. CMI, 苯基乙二醛双(脒基脞) (PhGBG);
- xi. CMI, 乙二醛双(脒基脞) (GBG; 脒, 1, 1'-((联二亚甲基双次氨基)二-(8CI); 脒基羧亚氨基酰胺, 2, 2'-(1, 2-联二亚甲基)二-(9CI));
- xii. CMI, 吡啶-丙酸 (IPA, 3-(3-吡啶基)丙酸);
- xiii. CMI, 3-吡啶-乙酸 (异植物生长素, IAA);
- xiv. CMI, 吡啶-3-乙酸钠盐;
- xv. CMI, 吡啶-3-丙酮;
- xvi. CMI, 吡啶-3-乙酰胺;
- xvii. CMI, 吡啶-3-乙酰基-L-天冬氨酸;
- xviii. CMI, 吡啶-3-乙酰基-L-丙氨酸;
- xix. CMI, 吡啶-3-乙酰基甘氨酸;
- xx. CMI, 吡啶-3-乙醛亚硫酸钠加成化合物;
- xxi. CMI, 吡啶-3-羧酸;
- xxii. CMI, 吡啶-3-丙酮酸 (3-(3-吡啶基)-2-氧代丙酸);
- xxiii. FI, 水杨酸 (2-羟基苯甲酸);
- xxiv. FI, 水杨酸钠盐;
- xxv. FI, 水杨酸钾盐;
- xxvi. IRI, 丹磺酰氯 (5-(二甲基氨基)萘-1-磺酰氯);
- xxvii. IRI, 丹磺酰氟 (5-(二甲基氨基)萘-1-磺酰氟);
- xxviii. CMI, 丹磺酰甘氨酸;
- xxix. CMI, 丙氨酸四唑;
- xxx. FI, 苯甲酸四唑;
- xxxi. CMI, 四唑;

- xxxii. CMI, 核黄素 5'-焦磷酸盐(RPP, 5-磷酸- α -D-核糖二磷酸盐, P-Rib-PP, P-RPP);
- xxxiii. IRI, DL-炔丙基甘氨酸 (DL-PG, 2-氨基-4-戊炔酸);
- xxxiv. IRI, L-C-炔丙基甘氨酸;
- xxxv. IRI, N-乙酰基-DL-炔丙基甘氨酸;
- xxxvi. FII, ($\pm$)-3-羟基丁酸钠;
- xxxvii. FI, 葫芦巴碱盐酸盐(1-甲基吡啶鎓-3-羧酸盐);
- xxxviii. FI, N-烟酸甲酯;
- xxxix. FI, 6-甲基烟酸甲酯;
- xl. FI, 2-甲基烟酸乙酯;
- xli. CMI, 曲酸(2-羧基甲基-5-羧基- γ -吡喃酮, 5-羧基-2-羧基甲基-4-吡喃酮);
- xlii. CMI, 曲酸衍生物, 例如: 6-(吡咯烷基甲基)-曲酸盐盐酸盐, 6-(吗啉代甲基)-曲酸, 6-(二乙基氨基甲基)-曲酸盐盐酸盐;
- xliii. IRI, O-(2, 4-二硝基苯基)羟胺;
- xliv. CMI, 2, 4-二硝基苯基甘氨酸;
- xlvi. IRI, 对硝基苯磺酸甲酯(4-硝基苯磺酸甲酯);
- xlvi. FIV, 氨基乙基半胱氨酸酮亚胺(AECK, 硫赖氨酸酮亚胺, 2H-1, 4-噻嗪-5, 6-二氢-3-羧酸, S-氨基乙基-L-半胱氨酸酮亚胺, 2H-1, 4-噻嗪-3-羧酸, 5, 6-二氢-);
- xlvi. FIV, 1, 4-噻嗪衍生物;
- xlvi. CMI, 4-苯基-1, 4-sulfonazan (四氢-4-苯基-4H-1, 4-噻嗪 1-氧化物, 4H-1, 4-噻嗪, 四氢-4-苯基-, 1-氧化物);
 - 1. CMI, 吩噻嗪(硫代二苯胺, 10H-吩噻嗪, AFI-Tiazin, Agrazine, Antiverm, 二苯并-1, 4-噻嗪);
 - li. CMI, 3, 4-二氢-2H-1, 4-噻嗪-3, 5-二羧酸(3, 4-Dhtca, CAS#86360-62-5);
 - lii. CMI, 硝呋莫司(Nifurtimox) (硝呋莫司[BAN: INN], 1-((5-硝基亚糠基)氨基)-2-甲基四氢-1, 4-噻嗪-4, 4-二氧化物, 3-甲基-4-(5'-硝基亚呋喃基-氨基)-四氢-4H-1, 4-噻嗪-1, 1-二氧化物, BAY 2502, 4-((5-硝基亚糠基)氨基)-3-甲基硫代码啉 1, 1-二氧化物, 等等);

- liii. FIV, 3-(1-吡咯烷基甲基)-4-(5, 6-二氯-1-二氢化茛羰基)-四氢-1, 4-噻嗪盐酸盐 (R 84760; R 84761; 硫代吗啉, 4-((5, 6-二氯-2, 3-二氢-1H-二氢化茛-1-基)羰基)-3-(1-吡咯烷基甲基)-, 单盐酸盐, (R-(R*, S*))-);
- liv. FIV, 酮亚胺还原形式;
- lv. CMI, 胱硫醚;
- lvi. FIII, 胱硫醚酮亚胺;
- lvii. FIV, 羊毛硫氨酸酮亚胺;
- lviii. FIV, 硫代吗啉-2-羧酸;
- lix. CMI, 硫代吗啉-2, 6-二羧酸;
- lx. FIV, TMDA (1, 4-硫代吗啉-3, 5-二羧酸);
- lxi. IRI, 1-氯-1-硝基乙烷;
- lxii. FI, 氨基苯甲酸盐;
- lxiii. FI, 2-氨基苯甲酸乙酯(氨基苯甲酸乙酯);
- lxiv. FI, 2-氨基苯甲酸甲酯(氨基苯甲酸甲酯);
- lxv. FI, 吡啶甲酸盐;
- lxvi. FI, 吡啶甲酸乙酯(2-(乙氧基羰基)吡啶, 2-吡啶羧酸乙酯,);
- lxvii. CMI, L-亮氨酸甲酯, 盐酸盐;
- lxviii. CMI, L-亮氨酸 ([(S)-(+)-亮氨酸]);
- lxix. IRI, 氟二硝基苯(1-氟-2, 4-二硝基苯, 2, 4-DNFB, 苯, 1-氟-2, 4-二硝基-, VAN, 等等);
- lxx. IRI, 二硝基氯苯(1-氯-2, 4-二硝基苯, 1, 3-二硝基-4-氯苯, 等等);
- lxxi. IRI, 1, 2-环己二酮;
- lxxii. IRI, 烯丙基甘氨酸 (D-烯丙基甘氨酸, 4-戊烯酸, 2-氨基-);
- lxxiii. CMI, 2-氨基-2, 4-戊二烯酸盐;
- lxxiv. CMI, 2-羧基-2, 4-戊二烯酸盐;
- lxxv. CMI, 2-氨基-4-酮-2-戊烯酸盐;
- lxxvi. FII, 2-羧基丁酸盐;
- lxxvii. FII, 2-羧基丁酸钠;
- lxxviii. IRI, N-氯-D-亮氨酸;
- lxxix. CMI, N-乙酰基-D-亮氨酸;
- lxxx. CMI, D-亮氨酸(D-2-氨基-4-甲基戊酸);

- lxxxix. IRI, D-炔丙基甘氨酸; 2-氨基-4-戊炔酸; D, L-炔丙基甘氨酸; L-2-氨基-4-戊炔酸;
- lxxxii. CMI, 孕酮(4-孕烯-3, 20-二酮);
- lxxxiii. CMI, FAD (黄素腺嘌呤二核苷酸, 1H-嘌呤-6-胺, 黄素二核苷酸, 腺苷 5'-(三氢焦磷酸盐), 带有核黄素的 5'-5'-酯, 等等);
- lxxxiv. CMI, 6-OH-FAD;
- lxxxv. IRI, 苯基乙二醛 (2, 2-二羧基苯乙酮);
- lxxxvi. IRI, 苯基乙二醛一水合物(2, 2-二羧基苯乙酮一水合物);
- lxxxvii. FIII, 环硫堇(全氢-1, 4-硫氮杂^草-3, 5-二羧酸, 1, 4-六氢硫氮杂^草-3, 5-二羧酸, 1, 4-硫氮杂^草-3, 5-二羧酸, 六氢-);
- lxxxviii. CMI, α - α' -亚氨基二丙酸(Alanopine; 2, 2'-亚氨基二丙酸基; L-丙氨酸, N-(1-羧乙基)-);
- lxxxix. CMI, 内消旋二氨基琥珀酸(3-氨基天冬氨酸; 二氨基琥珀酸; CAS RN: 921-52-8); 内消旋 2, 3-二氨基琥珀酸(CAS RN: 23220-52-2);
- xc. CMI, 氨基硫脲(氨基硫碳基酰肼);
- xc. CMI, 硫脲(硫代尿素; 硫脲);
- xcii. CMI, 甲基硫尿嘧啶(4(6)-甲基-2-硫尿嘧啶, 4-羧基-2-巯基 6-甲基嘧啶);
- xciii. CMI, 磺胺噻唑(N1-(2-噻唑基)磺胺, 4-氨基-N-2-噻唑基苯磺酰胺);
- xciv. CMI, 磺胺噻唑钠盐(4-氨基-N-2-噻唑基苯磺酰胺钠盐);
- xcv. CMI, 硫氰酸盐;
- xcvi. FI, 3-甲基苯甲基硫氰酸盐;
- xcvii. CMI, 甲巯咪唑(2-巯基-1-甲基咪唑, 1-甲基咪唑-2-硫醇);
- xcviii. FII, 二羧基羟酸;
- xcix. FII, 1, 3-丙酮二羧酸(3-氧代戊二酸);
- c. CMI, D-酒石酸([(2S, 3S)-(-)-酒石酸, 非天然酒石酸]);
- ci. CMI, L-酒石酸([(2R, 3R)-(+)-酒石酸, 天然酒石酸]);
- cii. CMI, DL-酒石酸;
- ciii. 酒石酸钾;
- civ. FII, D-苹果酸; [(R)-(+)-苹果酸, (R)-(+)-羟基琥珀酸];
- cv. FII, L-苹果酸; [(S)-(-)-苹果酸, (S)-(-)-羟基琥珀酸];
- cvi. FII, DL-苹果酸(DL-羟基琥珀酸);

- cvi. FII, α -酮酸, 其为氨基酸丙氨酸、亮氨酸、苯丙氨酸、苯甘氨酸, 酪氨酸、丝氨酸、天冬氨酸盐等及其盐和衍生物类似物;
- cvi. FII, 丙酮酸(2-氧代丙酸, α -酮丙酸);
- cix. FII, 丙酮酸钠;
- cx. FII, 丙酮酸甲酯(丙酮酸甲酯);
- cxi. FI, 苯基丙酮酸;
- cxii. FII, 苯基丙酮酸钙(丙酮酸钙);
- cxiii. FI, 苯基丙酮酸钠盐(苯基丙酮酸钠);
- cxiv. FII, 4-羟基苯基丙酮酸;
- cxv. FII, α -酮异戊酸钠(3-甲基-2-氧代丁酸钠盐, 3-甲基-2-氧代丁酸钠盐, α -酮异戊酸钠盐; 酮戊酸钠盐);
- cxvi. FI, 苯酰甲酸 (α -氧代苯乙酸, 苯基乙醛酸);
- cxvii. FII, 4-甲基硫代-2-氧代戊酸;
- cxviii. FII, 4-甲基-2-氧代戊酸(4-甲基-2-氧代戊酸; α -酮异己酸);
- cxix. FII, 4-甲基硫代-2-氧丁酸;
- cxx. FII, 2-氧丁酸(羟基丁酸盐; 2-羟基丁酸; α -羟基-正丁酸);
- cxxi. FII, DL- α -羟基丁酸钠盐(($\pm$)-2-羟基丁酸钠);
- cxxii. FII, 吲哚-3-丙酮酸(色氨酸的 α -酮类似物);
- cxxiii. 半胱胺与溴丙酮酸盐的反应产物;
- cxxiv. CMI, 半胱胺(2-氨基乙硫醇; 2-巯基乙胺);
- cxxv. CMI, 泛酰巯基乙胺;
- cxxvi. CMI, S-腺苷甲硫氨酸;
- cxxvii. IRI, 溴丙酮酸乙酯;
- cxxviii. IRI, 溴丙酮酸甲酯;
- cxxix. IRI, 溴丙酮酸盐; 以及
- cxix. CMI, 5-S-半胱氨酰基多巴胺,

其中 IRI 表示不可逆抑制剂组合物; CMI 表示竞争性抑制剂组合物但不包在以一般式 I-IV 表示的组合物中; FI 表示本文以一般式 I 表示的组合物; FII 表示本文以一般式 II 表示的组合物; FIII 表示本文以一般式 III 表示的组合物; 以及 FIV 表示本文以一般式 IV 表示的组合物。需要指出的是, 以一般式 I-IV 表示的组合物是 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的竞争性、非竞争性、无竞争性或变构抑制剂。

用于本发明方法能够在体内或体外降低、抑制或拮抗 D-氨基酸氧化酶催化活性的优选组合物选自以上列出的所有组合物“i”，直到并包括“cxxx”；较优选的组合物选自不可逆抑制剂组合物、一般式 I 表示的组合物、一般式 II 表示的组合物、一般式 III 表示的组合物和一般式 IV 表示的组合物；更优选的组合物选自一般式 I 表示的组合物、一般式 II 表示的组合物、一般式 III 表示的组合物和一般式 IV 表示的组合物；最优选的组合物选自一般式 I 表示的组合物和一般式 IV 表示的组合物。用于本发明方法能够在体内或体外降低、抑制或拮抗 D-氨基酸氧化酶或 D-天冬氨酸氧化酶催化活性的较优选组合物选自包括苯甲酸盐、氨基乙基半胱氨酸酮亚胺(AECK)及其衍生物的组。

在另一优选实施例中，用于本发明治疗中枢神经系统障碍的方法中优选组合物或化合物选自以上列出的所有组合物“i”，直到并包括“cxxx”；较优选的组合物选自不可逆抑制剂组合物、一般式 I 表示的组合物、一般式 II 表示的组合物、一般式 III 表示的组合物和一般式 IV 表示的组合物；更优选的组合物选自一般式 I 表示的组合物、一般式 II 表示的组合物、一般式 III 表示的组合物和一般式 IV 表示的组合物；最优选的组合物选自一般式 I 表示的组合物和一般式 IV 表示的组合物。用于本发明治疗中枢神经系统障碍的方法中较优选组合物选自包括苯甲酸盐、氨基乙基半胱氨酸酮亚胺(AECK)及其衍生物的组。

本发明能降低、抑制或拮抗 D-氨基酸氧化酶或 D-天冬氨酸氧化酶活性的高度优选化合物或组合物选自包括但不限于以下所列的：氨基乙基半胱氨酸-酮亚胺(AECK，硫赖氨酸酮亚胺，2H-1，4-噻嗪-5，6-二氢-3-羧酸，S-氨基乙基-L-半胱氨酸酮亚胺，2H-1，4-噻嗪-3-羧酸，5，6-二氢-)；氨基乙基半胱氨酸(硫赖氨酸)；半胱胺；泛酰巯基乙胺；胱硫醚和 S-腺苷甲硫氨酸。

本发明另一优选实施例涉及能降低、抑制或拮抗氨基酸或其衍生物氧化或降解的化合物或组合物。本发明的另一优选实施例涉及能降低、抑制或拮抗 L-氨基酸或其衍生物氧化或降解的化合物或组合物。本发明的另一优选实施例涉及能降低、抑制或拮抗 D-氨基酸或其衍生物氧化或降解的化合物或组合物。本发明的另一优选实施例涉及能降低、抑制或拮抗甘氨酸或其衍生物氧化或降解的化合物或组合物。本发明的又一优选实施例涉及能降低、抑制或拮抗至少一种 D-氨基酸氧化或降解的化合物或组合物，所述 D-氨基酸选自：D-Met、D-Pro、D-Phe、D-Tyr、D-Ile、D-Leu、D-Ala、D-Val、D-Ser、D-Arg、D-His、D-正亮氨酸、D-Trp、D-鸟氨酸、顺-4-羟基-D-脯氨酸、D-Thr、D-Trp-甲酯、N-乙酰基-D-Ala、D-Lys、D-Asp、D-Glu、D-Asn、D-Gln、D-Asp-二甲酯和 N-甲基-D-Asp。较优选的是能降低、抑制

或拮抗 D-丝氨酸氧化或降解的组合物。较优选的是能降低、抑制或拮抗 D-Ser、N-甲基-D-Asp、D-Asp 或 Gly 氧化或降解的组合物或化合物。本发明能降低、抑制或拮抗一种氨基酸或其衍生物氧化或降解的优选化合物或组合物选自包括但不限于以下所列的：氨基乙基半胱氨酸-酮亚胺(AECK, 硫赖氨酸酮亚胺, 2H-1, 4-噻嗪-5, 6-二氢-3-羧酸, S-氨基乙基-L-半胱氨酸酮亚胺, 2H-1, 4-噻嗪-3-羧酸, 5, 6-二氢-); 氨基乙基半胱氨酸(硫赖氨酸); 半胱胺; 泛酰巯基乙胺; 胱硫醚和 S-腺苷甲硫氨酸。本发明能降低、抑制或拮抗 D-Met、D-Pro、D-Phe、D-Tyr、D-Ile、D-Leu、D-Ala、D-Val、D-Ser、D-Arg、D-His、D-正亮氨酸、D-Trp、D-鸟氨酸、顺-4-羟基-D-脯氨酸、D-Thr、D-Trp-甲酯、N-乙酰基-D-Ala、D-Lys、D-Asp、D-Glu、D-Asn、D-Gln、D-Asp-二甲酯、N-甲基-D-Asp 或 Gly 氧化或降解的优选化合物或组合物选自包括但不限于以下所列的：氨基乙基半胱氨酸-酮亚胺(AECK, 硫赖氨酸酮亚胺, 2H-1, 4-噻嗪-5, 6-二氢-3-羧酸, S-氨基乙基-L-半胱氨酸酮亚胺, 2H-1, 4-噻嗪-3-羧酸, 5, 6-二氢-); 氨基乙基半胱氨酸(硫赖氨酸); 半胱胺; 泛酰巯基乙胺; 胱硫醚和 S-腺苷甲硫氨酸。本发明能降低、抑制或拮抗 D-丝氨酸氧化或降解的优选化合物或组合物选自包括但不限于以下所列的：氨基乙基半胱氨酸-酮亚胺(AECK, 硫赖氨酸酮亚胺, 2H-1, 4-噻嗪-5, 6-二氢-3-羧酸, S-氨基乙基-L-半胱氨酸酮亚胺, 2H-1, 4-噻嗪-3-羧酸, 5, 6-二氢-); 氨基乙基半胱氨酸(硫赖氨酸); 半胱胺; 泛酰巯基乙胺; 胱硫醚和 S-腺苷甲硫氨酸。

本发明另一优选实施例涉及能降低、抑制或拮抗还原的黄素腺嘌呤二核苷酸(Re-FAD)氧化的组合物。本发明另一优选实施例涉及能降低、抑制或拮抗氧化的黄素腺嘌呤二核苷酸(Ox-FAD)还原的组合物。本发明又一优选实施例涉及能降低、抑制或拮抗黄素激酶活性的组合物。本发明又一优选实施例涉及能降低、抑制或拮抗黄素腺嘌呤二核苷酸焦磷酸化酶活性的组合物。本发明又一优选实施例涉及能与还原的黄素腺嘌呤二核苷酸或氧化的黄素腺嘌呤二核苷酸结合或相互作用的组合物。本发明又一优选实施例涉及能与黄素激酶或黄素腺嘌呤二核苷酸焦磷酸化酶结合或相互作用的组合物。

本发明另一优选实施例涉及能增加、激动或促进胱硫醚 β -合成酶活性的组合物或化合物。能增加、促进或激动胱硫醚 β -合成酶活性的优选组合物包括 S-腺苷甲硫氨酸或高半胱氨酸。能增加、激动或促进胱硫醚 β -合成酶活性的另一优选组合物是维生素 B6 或其衍生物。

本发明另一优选实施例涉及一种能与 D-氨基酸氧化酶、D-天冬氨酸氧化酶、还原的黄素腺嘌呤二核苷酸、氧化的黄素腺嘌呤二核苷酸、黄素激酶、黄素腺嘌呤

二核苷酸焦磷酸化酶、胱硫醚 β -合成酶、L-氨基酸氧化酶、或者谷氨酰胺转氨酶结合或相互作用的组合物的筛选方法。本发明另一优选实施例涉及一种能降低、抑制或拮抗 D-氨基酸氧化酶、D-天冬氨酸氧化酶、黄素激酶、黄素腺嘌呤二核苷酸焦磷酸化酶、L-氨基酸氧化酶、或者谷氨酰胺转氨酶活性的组合物的筛选方法。本发明另一优选实施例涉及一种能激动、增加或促进胱硫醚 β -合成酶、L-氨基酸氧化酶或者谷氨酰胺转氨酶活性的组合物的筛选方法。

所以，本发明的一方面内容提供一种用于治疗疾病或用于提高氨基酸水平或减少氨基酸降解的候选分子的鉴定方法，所述方法包括：a)使 D-氨基酸氧化酶、D-天冬氨酸氧化酶、黄素激酶、黄素腺嘌呤二核苷酸焦磷酸化酶、L-氨基酸氧化酶、或者谷氨酰胺转氨酶多肽或其具有生物活性的片段与测试化合物接触；及 b)确定所述化合物是否与所述多肽选择性结合；其中所述化合物与所述多肽选择性结合确定表示所述化合物是治疗疾病或者提高氨基酸水平或减少氨基酸降解的候选分子。

本发明也提供一种用于治疗疾病或用于提高氨基酸水平或减少氨基酸降解的候选分子的鉴定方法，所述方法包括：a)使 D-氨基酸氧化酶、D-天冬氨酸氧化酶、黄素激酶、黄素腺嘌呤二核苷酸焦磷酸化酶、L-氨基酸氧化酶、或者谷氨酰胺转氨酶多肽或其具有生物活性的片段与测试化合物接触；及 b)确定所述化合物是否选择性抑制所述多肽的活性；其中所述化合物选择性抑制所述多肽的活性的确定表示所述化合物是治疗疾病或者提高氨基酸水平或减少氨基酸降解的候选分子。

本发明还提供一种用于治疗疾病或用于提高氨基酸水平或减少氨基酸降解的候选分子的鉴定方法，所述方法包括：a)提供含有 D-氨基酸氧化酶、D-天冬氨酸氧化酶、黄素激酶、黄素腺嘌呤二核苷酸焦磷酸化酶、L-氨基酸氧化酶、或者谷氨酰胺转氨酶多肽或其具有生物活性的片段的细胞；及 b)使所述细胞与测试化合物接触；及 c)确定所述化合物是否选择性抑制 D-氨基酸氧化酶、D-天冬氨酸氧化酶、黄素激酶、黄素腺嘌呤二核苷酸焦磷酸化酶、L-氨基酸氧化酶、或者谷氨酰胺转氨酶的活性；其中所述化合物选择性抑制所述多肽的活性的确定表示所述化合物是治疗疾病或者提高氨基酸水平或减少氨基酸降解的候选分子。

本发明还提供一种用于治疗疾病或用于提高氨基酸水平或减少氨基酸降解的候选分子的鉴定方法，所述方法包括：a)使胱硫醚 β -合成酶、L-氨基酸氧化酶或者谷氨酰胺转氨酶多肽或其具有生物活性的片段与测试化合物接触；及 b)确定所述化合物是否选择性增加所述多肽的活性；其中所述化合物选择性增加所述多肽的活性的确定表示所述化合物是治疗疾病或者提高氨基酸水平或减少氨基酸降解的

候选分子。

本发明另一实施例提供一种用于治疗疾病或用于提高氨基酸水平或减少氨基酸降解的候选分子的鉴定方法，所述方法包括：a) 提供含有胱硫醚 β -合成酶、L-氨基酸氧化酶或者谷氨酰胺转氨酶多肽或其具有生物活性的片段的细胞；b) 使所述细胞与测试化合物接触；及 c) 确定所述化合物是否选择性增加胱硫醚 β -合成酶、L-氨基酸氧化酶或者谷氨酰胺转氨酶的活性；其中所述化合物选择性增加所述多肽的活性的确定表示所述化合物是治疗疾病或者提高氨基酸水平或减少氨基酸降解的候选分子。

本发明另一优选实施例涉及一种拮抗、降低或抑制 D-氨基酸氧化酶体外活性的方法。较优选的是一种拮抗、降低或抑制 D-氨基酸氧化酶体内活性的方法。较优选的是一种拮抗、降低或抑制 D-氨基酸氧化酶体外或体内活性的方法，所述方法包括使 D-氨基酸氧化酶与能降低、抑制或拮抗 D-氨基酸氧化酶活性的组合物接触的步骤。待抑制的 D-氨基酸氧化酶优选活性是底物的氧化，该底物最好是 D-氨基酸，而该 D-氨基酸最好是 D-Ser、D-Asp、或 N-甲基-D-Asp。

本发明另一优选实施例涉及一种拮抗、降低或抑制 D-天冬氨酸氧化酶体外活性的方法。较优选的是一种拮抗、降低或抑制 D-天冬氨酸氧化酶体内活性的方法。较优选的是一种拮抗、降低或抑制 D-天冬氨酸氧化酶体外或体内活性的方法，所述方法包括使 D-天冬氨酸氧化酶与能降低、抑制或拮抗 D-氨基酸氧化酶活性的组合物接触的步骤。待抑制的 D-天冬氨酸氧化酶优选活性是底物的氧化，该底物最好是 D-氨基酸，而该 D-氨基酸最好是 D-Asp、D-Glu、D-Asn、D-Gln、D-Asp-二甲酯或 N-甲基-D-Asp。

本发明另一优选实施例涉及能提高至少一种 D-氨基酸在体外的水平的组合物。较优选的是能提高至少一种 D-氨基酸在体内的水平的组合物，优选在哺乳动物组织内，更优选在小鼠、大鼠、狗、牛、猪、猿、猴或人的组织内。较优选的是能提高至少一种 D-氨基酸在中枢神经系统最好在脑或脊髓的组织内的水平的组合物。较优选的是能提高至少一种 D-氨基酸在脑最好在海马、扁桃体、黑质、小脑、胼胝体、尾状核、大脑皮层、丘脑或垂体的组织内的水平的组合物。本发明的组合物能提高至少一种 D-氨基酸在其它优选组织的水平，所述的其它优选组织包括但不限于肾、肝脏、脂肪、肌肉和睾丸。

本发明一优选实施例涉及 SEQ ID NO: 15 的多肽或其片段在一种增加 D-氨基酸氧化酶活性的方法中的应用。较优选的是 SEQ ID NO: 15 的多肽或其片段在一种增加 D-天冬氨酸氧化酶活性的方法中的应用。较优选的是 SEQ ID NO: 15 的多肽

或其片段在一种降低丝氨酸消旋酶活性的方法中的应用。

本发明一优选实施例涉及 SEQ ID NO: 15 的多肽或其片段在一种提高化合物或组合物产量的方法中的应用, 所述化合物或组合物是以 D-氨基酸氧化酶为催化剂的反应产物。

本发明一优选实施例涉及一种与 g34872 多肽(SEQ ID NO: 15)或 g34872 多核苷酸(SEQ ID NO: 14)或其片段结合的组合物或化合物的筛选方法。较优选的是这样一种方法: 使 g34872 多肽或其片段与 D-氨基酸氧化酶接触, 从而使 D-氨基酸氧化酶活性增加到基本水平以上。较优选的是一种能降低、抑制、拮抗或阻断 D-氨基酸氧化酶与 g34872 相互作用的方法。较优选的是一种通过阻断 g34872 与 D-氨基酸氧化酶相互作用来治疗中枢神经系统障碍的方法。较优选的是一种用能降低、阻断、抑制或拮抗 g34872 与 D-氨基酸氧化酶相互作用的化合物或组合物来治疗中枢神经系统障碍的方法。

本发明优选的 D-氨基酸氧化酶多肽包括 SEQ ID NO: 7-10 和 19 的多肽及其片段以及编码它们的多核苷酸。本发明优选的 D-天冬氨酸氧化酶多肽包括 SEQ ID NO: 22 和 23 的多肽及其片段以及编码它们的多核苷酸。本发明优选的 D-氨基酸氧化酶多核苷酸包括 SEQ ID NO: 2-6 和 18 及其片段以及由它们编码的多肽。本发明优选的 D-天冬氨酸氧化酶多核苷酸包括 SEQ ID NO: 20 和 21 及其片段以及由它们编码的多肽。

SEQ ID NO: 1 所述以及以标记 24-1443-126 (SEQ ID NO: 24)、标记 24-1457-52 (SEQ ID NO: 26)和标记 24-1461-256 (SEQ ID NO: 29)的 47-mers 表示优选的 D-氨基酸氧化酶双等位标记。

本发明另一优选实施例涉及与 SEQ ID NO: 7 的多肽差异结合的组合物。本发明另一优选实施例涉及与 SEQ ID NO: 8 的多肽差异结合的组合物。本发明另一优选实施例涉及与 SEQ ID NO: 9 的多肽差异结合的组合物。本发明另一优选实施例涉及与 SEQ ID NO: 10 的多肽差异结合的组合物。较优选的是与 SEQ ID NO: 10 的多肽结合但不与 SEQ ID NO: 7、8 或 9 的多肽结合的组合物。较优选的是与 SEQ ID NO: 9 的多肽结合但不与 SEQ ID NO: 7、8 或 10 的多肽结合的组合物。较优选的是与 SEQ ID NO: 8 的多肽结合但不与 SEQ ID NO: 7、9 或 10 的多肽结合的组合物。较优选的是与 SEQ ID NO: 7 的多肽结合但不与 SEQ ID NO: 8、9 或 10 的多肽结合的组合物。较优选的是与 SEQ ID NO: 8、9 或 10 的多肽结合但不与 SEQ ID NO: 7 的多肽结合的组合物。

本发明另一优选实施例涉及与含有 SEQ ID NO: 7、8、9、10 或 15 的单体多

肽或其多肽片段差异结合的组合物。较优选的是与 SEQ ID NO: 7 的单体多肽或其片段结合但不与含有至少一种 SEQ ID NO: 7 的单体多肽或其片段的同或异多聚体形式结合的组合物。较优选的是与 SEQ ID NO: 8 的单体多肽或其片段结合但不与含有至少一种 SEQ ID NO: 8 的单体多肽或其片段的同或异多聚体形式结合的组合物。较优选的是与 SEQ ID NO: 9 的单体多肽或其片段结合但不与含有至少一种 SEQ ID NO: 9 的单体多肽或其片段的同或异多聚体形式结合的组合物。较优选的是与 SEQ ID NO: 10 的单体多肽或其片段结合但不与含有至少一种 SEQ ID NO: 10 的单体多肽或其片段的同或异多聚体形式结合的组合物。较优选的是与 SEQ ID NO: 15 的单体多肽或其片段结合但不与含有至少一种 SEQ ID NO: 15 的单体多肽或其片段的同或异多聚体形式结合的组合物。

本发明另一优选实施例涉及与含有 SEQ ID NO: 7、8、9、10 或 15 的至少一种多肽或其片段的多聚体多肽结合的组合物。较优选的是与含有 SEQ ID NO: 7 的至少一种单体多肽或其片段的同或异多聚体形式结合但不与 SEQ ID NO: 7 的单体多肽或其片段结合的组合物。较优选的是与含有 SEQ ID NO: 8 的至少一种单体多肽或其片段的同或异多聚体形式结合但不与 SEQ ID NO: 8 的单体多肽或其片段结合的组合物。较优选的是与含有 SEQ ID NO: 9 的至少一种单体多肽或其片段的同或异多聚体形式结合但不与 SEQ ID NO: 9 的单体多肽或其片段结合的组合物。较优选的是与含有 SEQ ID NO: 10 的至少一种单体多肽或其片段的同或异多聚体形式结合但不与 SEQ ID NO: 10 的单体多肽或其片段结合的组合物。较优选的是与含有 SEQ ID NO: 15 的至少一种单体多肽或其片段的同或异多聚体形式结合但不与 SEQ ID NO: 15 的单体多肽或其片段结合的组合物。

本发明另一实施例涉及与含有 SEQ ID NO: 2 的多核苷酸差异结合的组合物。本发明另一实施例涉及与含有 SEQ ID NO: 3 的多核苷酸差异结合的组合物。本发明另一实施例涉及与含有 SEQ ID NO: 4 的多核苷酸差异结合的组合物。本发明另一实施例涉及与含有 SEQ ID NO: 5 的多核苷酸差异结合的组合物。本发明另一实施例涉及与含有 SEQ ID NO: 6 的多核苷酸差异结合的组合物。较优选的是与 SEQ ID NO: 6 的多核苷酸结合但不与 SEQ ID NO: 2、3、4 或 5 的多核苷酸结合的组合物。较优选的是与 SEQ ID NO: 5 的多核苷酸结合但不与 SEQ ID NO: 2、3、4 或 6 的多核苷酸结合的组合物。较优选的是与 SEQ ID NO: 4 的多核苷酸结合但不与 SEQ ID NO: 2、3、5 或 6 的多核苷酸结合的组合物。较优选的是与 SEQ ID NO: 3 的多核苷酸结合但不与 SEQ ID NO: 2、4、5 或 6 的多核苷酸结合的组合物。较优选的是与 SEQ ID NO: 2 的多核苷酸结合但不与 SEQ ID NO: 3、4、5 或 6 的多核苷酸结合

的组合物。较优选的是与 SEQ ID NO: 3、4、5 或 6 的多核苷酸结合但不与 SEQ ID NO: 2 的多核苷酸结合的组合物。

本发明的另一优选实施例涉及含有 SEQ ID NO: 1 的多核苷酸的基因组序列。较优选的是使 SEQ ID NO: 1 的多核苷酸区基因分型的方法。

本发明一实施例涉及含有 SEQ ID NO: 1 的序列或其互补序列的纯化的或分离的核酸。较优选的是一种纯化的或分离的核酸，其含有至少 10 个 SEQ ID NO: 1 的序列或其互补序列的连续核苷酸。较优选的是一种核酸，其含有至少 15 个 SEQ ID NO: 1 的序列或其互补序列的连续核苷酸。

本发明一优选实施例涉及一种纯化的或分离的核酸，其含有至少 10 个 SEQ ID NO: 1 的序列或其互补序列或一个或多个外显子的连续核苷酸。较优选的是一种 SEQ ID NO: 1 的序列或其互补序列所示的纯化的或分离的核酸，其含有 SEQ ID NO: 1 的序列或其互补序列所示的核苷酸 40389 至 40670 中至少 10 个连续核苷酸。一种优选的 SEQ ID NO: 1 的序列或其互补序列所示的纯化的或分离的核酸含有 SEQ ID NO: 1 的序列或其互补序列所示的核苷酸 42666 至 42778 中至少 10 个连续核苷酸。另一种优选的 SEQ ID NO: 1 的序列或其互补序列所示的纯化的或分离的核酸含有 SEQ ID NO: 1 的序列或其互补序列所示的核苷酸 43416 至 43519 中至少 10 个连续核苷酸。另一种优选的 SEQ ID NO: 1 的序列或其互补序列所示的纯化的或分离的核酸含有 SEQ ID NO: 1 的序列或其互补序列所示的核苷酸 61159 至 61402 中至少 10 个连续核苷酸。另一种优选的 SEQ ID NO: 1 的序列或其互补序列所示的纯化的或分离的核酸含有 SEQ ID NO: 1 的序列或其互补序列所示的核苷酸 64050 至 64711 中至少 10 个连续核苷酸。另一种优选的 SEQ ID NO: 1 的序列或其互补序列所示的纯化的或分离的核酸含有 SEQ ID NO: 1 的序列或其互补序列所示的核苷酸 68126 至 68261 中至少 10 个连续核苷酸。另一种优选的 SEQ ID NO: 1 的序列或其互补序列所示的纯化的或分离的核酸含有 SEQ ID NO: 1 的序列或其互补序列所示的核苷酸 84906 至 85541 中至少 10 个连续核苷酸。

本发明另一优选实施例涉及一种含有 SEQ ID NO: 2 的序列或其互补序列的纯化的或分离的核酸。本发明又一优选实施例涉及一种含有 SEQ ID NO: 3 的序列或其互补序列的纯化的或分离的核酸。本发明另一优选实施例涉及一种含有 SEQ ID NO: 4 的序列或其互补序列的纯化的或分离的核酸。本发明另一优选实施例涉及一种含有 SEQ ID NO: 5 的序列或其互补序列的纯化的或分离的核酸。本发明另一优选实施例涉及一种含有 SEQ ID NO: 6 的序列或其互补序列的纯化的或分离的核酸。本发明另一优选实施例涉及一种含有 SEQ ID NO: 14 的序列或其互补序列的纯化的

或分离的核酸。本发明另一优选实施例涉及一种含有 SEQ ID NO: 16 的序列或其互补序列的纯化的或分离的核酸。本发明另一优选实施例涉及一种含有 SEQ ID NO: 18、20 或 21 的序列中任何一段序列或其互补序列的纯化的或分离的核酸。

本发明另一实施例涉及一种纯化的或分离的核酸，其含有 SEQ ID NO: 2-6 的序列中至少之一或其互补序列所示的至少 10 个连续核苷酸。较优选的是一种纯化的或分离的核酸，其含有 SEQ ID NO: 2-6 的序列中至少之一或其互补序列所示的至少 15 个连续核苷酸。

本发明另一实施例涉及一种纯化的或分离的核酸，其含有 SEQ ID NO: 14 的序列或其互补序列所示的至少 10 个连续核苷酸。较优选的是一种纯化的或分离的核酸，其含有 SEQ ID NO: 14 的序列或其互补序列所示的至少 15 个连续核苷酸。

本发明另一实施例涉及一种编码 SEQ ID NO: 7 的多肽的纯化的或分离的核酸。较优选的是一种编码 SEQ ID NO: 7 的多肽的至少 10 个连续氨基酸的纯化的或分离的核酸。较优选的是一种纯化的或分离的核酸，所述核酸编码 SEQ ID NO: 7 的多肽的至少 15 个连续氨基酸。

本发明另一实施例涉及一种编码 SEQ ID NO: 8 的多肽的纯化的或分离的核酸。较优选的是一种编码 SEQ ID NO: 8 的多肽的至少 10 个连续氨基酸的纯化的或分离的核酸。较优选的是一种纯化的或分离的核酸，所述核酸编码 SEQ ID NO: 8 的多肽的至少 15 个连续氨基酸。

本发明另一实施例涉及一种编码 SEQ ID NO: 9 的多肽的纯化的或分离的核酸。较优选的是一种编码 SEQ ID NO: 9 的多肽的至少 10 个连续氨基酸的纯化的或分离的核酸。较优选的是一种纯化的或分离的核酸，所述核酸编码 SEQ ID NO: 9 的多肽的至少 15 个连续氨基酸。

本发明另一实施例涉及一种编码 SEQ ID NO: 10 的多肽的纯化的或分离的核酸。较优选的是一种编码 SEQ ID NO: 10 的多肽的至少 10 个连续氨基酸的纯化的或分离的核酸。较优选的是一种纯化的或分离的核酸，所述核酸编码 SEQ ID NO: 10 的多肽的至少 15 个连续氨基酸。

本发明另一实施例涉及一种编码 SEQ ID NO: 15 的多肽的纯化的或分离的核酸。较优选的是一种编码 SEQ ID NO: 15 的多肽的至少 10 个连续氨基酸的纯化的或分离的核酸。较优选的是一种纯化的或分离的核酸，所述核酸编码 SEQ ID NO: 15 的多肽的至少 15 个连续氨基酸。

本发明另一实施例涉及一种编码 SEQ ID NO: 17 的多肽的纯化的或分离的核酸。较优选的是一种编码 SEQ ID NO: 17 的多肽的至少 10 个连续氨基酸的纯化的

或分离的核酸。较优选的是一种纯化的或分离的核酸，所述核酸编码 SEQ ID NO: 17 的多肽的至少 15 个连续氨基酸。

本发明另一实施例涉及一种编码 SEQ ID NO: 19 的多肽的纯化的或分离的核酸。较优选的是一种编码 SEQ ID NO: 19 的多肽的至少 10 个连续氨基酸的纯化的或分离的核酸。较优选的是一种纯化的或分离的核酸，所述核酸编码 SEQ ID NO: 19 的多肽的至少 15 个连续氨基酸。

本发明另一实施例涉及一种编码 SEQ ID NO: 22 的多肽的纯化的或分离的核酸。较优选的是一种编码 SEQ ID NO: 22 的多肽的至少 10 个连续氨基酸的纯化的或分离的核酸。较优选的是一种纯化的或分离的核酸，所述核酸编码 SEQ ID NO: 22 的多肽的至少 15 个连续氨基酸。

本发明另一实施例涉及一种编码 SEQ ID NO: 23 的多肽的纯化的或分离的核酸。较优选的是一种编码 SEQ ID NO: 23 的多肽的至少 10 个连续氨基酸的纯化的或分离的核酸。较优选的是一种纯化的或分离的核酸，所述核酸编码 SEQ ID NO: 23 的多肽的至少 15 个连续氨基酸。

本发明另一优选实施例涉及双等位标记。

附图说明

图 1 所示为表达酵母菌的重组 g34782 多肽和 D-氨基酸氧化酶多肽的活性。

图 2 所示为表达细菌的重组 g34782 多肽和 D-氨基酸氧化酶多肽的活性。

图 3 所示为用 D-丝氨酸作为底物通过 g34782 在体外活化纯化的 D-氨基酸氧化酶。

图 4 所示为 g34782 的剂量与 D-氨基酸氧化酶活性的依赖性关系。

图 5 所示为 g34782 多肽和 D-氨基酸氧化酶之间相互作用的动力学。

图 6 所示为法裔加拿大人精神分裂症病例与对照组之间的 D-氨基酸氧化酶双等位标记的关联性分析结果。

序列表所提供的序列的概述

SEQ ID NO: 1 D-氨基酸氧化酶的基因组序列；

SEQ ID NO: 2 D-氨基酸氧化酶的 cDNA；

SEQ ID NO: 3 具有外显子 U 2 3 4 5 6 7 8 9 10 11Long 的新 cDNA；

SEQ ID NO: 4 具有外显子 B C Ulong V 2 3 4 5 6 7 9 10 11Long 的新 cDNA；

SEQ ID NO: 5 具有外显子 U 2 4 5 6 7 8 9 10 11Long 的新 cDNA；

- SEQ ID NO: 6 具有外显子 B 2 3 7 8 9 10 11 的新 cDNA;
- SEQ ID NO: 7 来自 SEQ ID NO: 2 和 3 的 cDNA 的 D-氨基酸氧化酶多肽;
- SEQ ID NO: 8 来自 SEQ ID NO: 4 的 cDNA 的 D-氨基酸氧化酶多肽;
- SEQ ID NO: 9 来自 SEQ ID NO: 5 的 cDNA 的 D-氨基酸氧化酶多肽 5;
- SEQ ID NO: 10 来自 SEQ ID NO: 6 的 cDNA 的 D-氨基酸氧化酶多肽;
- SEQ ID NO: 11-12 含有 g34872 双等位标记 99/16105-152 和 99/5919-215 的多核苷酸;
- SEQ ID NO: 13 包括多态性的 g34872 多核苷酸;
- SEQ ID NO: 14 g34872 多肽, 其中位置 10 上的氨基酸是酪氨酸或丝氨酸, 位置 30 上的氨基酸是赖氨酸或精氨酸, 位置 50 上的氨基酸是谷氨酸或早熟终止子, 位置 60 上的氨基酸是精氨酸或甘氨酸, 以及位置 115 上的氨基酸是天冬氨酸或丙氨酸;
- SEQ ID NO: 15 编码双杂交实验所用的 SEQ ID NO: 16 的多肽的 g34872 多核苷酸;
- SEQ ID NO: 17 编码 SEQ ID NO: 18 的多肽的 D-氨基酸氧化酶多核苷酸;
- SEQ ID NO: 19 和 20 分别是编码 SEQ ID NO: 21 和 22 的多肽的 D-天冬氨酸氧化酶多核苷酸; 以及
- SEQ ID NO: 23-26 分别是含有 D-氨基酸氧化酶双等位标记 24-1443/126、24-1457/52、27-93/181 和 24-1461/256 的多核苷酸, 在位置 24 显示多态性碱基。

美国专利申请号 09/539, 333 与国际专利申请号 PCT/IB00/00435 都对 g34872 基因组序列和双等位标记进行了叙述, 以上两件申请均纳入其整体内容作为参考。

根据有关序列表的规定, 在本序列表中采用以下代号表示双等位标记在序列内的位置, 并确定多态性碱基含有的每个等位基因。序列中的代号“r”表示多态性碱基的其中一个等位基因是鸟嘌呤, 而另一个等位基因是腺嘌呤。序列中的代号“y”表示多态性碱基的其中一个等位基因是胸腺嘧啶, 而另一个等位基因是胞嘧啶。序列中的代号“m”表示多态性碱基的其中一个等位基因是腺嘌呤, 而另一个等位基因是胞嘧啶。序列中的代号“k”表示多态性碱基的其中一个等位基因是鸟嘌呤, 而另一个等位基因是胸腺嘧啶。序列中的代号“s”表示多态性碱基的其中一个等位基因是鸟嘌呤, 而另一个等位基因是胞嘧啶。序列中的代号“w”表示多态性碱基的其中一个等位基因是腺嘌呤, 而另一个等位基因是胸腺嘧啶。

具体实施方式

本发明涉及给容易患上中枢神经系统障碍的受试者提供这样一种精神障碍的预防方法以及给患有中枢神经系统障碍受试者提供治疗方法。具体地说,该方法包括给患者施加一定量的 D-氨基酸氧化酶和 D-天冬氨酸氧化酶拮抗剂化合物,一定程度上有效预防或改善中枢神经系统障碍的进展(即提供保护作用)、改善中枢神经系统障碍的症状,以及改善中枢神经系统障碍的复发。

按照本发明的方法可治疗的中枢神经系统障碍包括早衰性痴呆(早期发作的阿耳茨海默氏病)、老年性痴呆(阿耳茨海默氏型痴呆)、帕金森氏症包括帕金森氏病、杭廷顿舞蹈病、迟发性运动障碍、运动过度、躁狂症、注意力缺陷活动过度障碍(ADHD)、注意力缺陷障碍(ADD)、焦虑症、诵读困难、精神病、精神分裂症、双相性精神障碍、重度抑郁症、躁狂症、轻度躁狂症、抑郁症、孤独症、滥用药物、过度攻击、抽搐症和图雷特氏综合征。本发明优选的中枢神经系统障碍包括精神分裂症和双相性精神障碍。精神分裂症和精神分裂症样精神障碍的较优选例子包括:精神分裂症(紧张型)、精神分裂症(无组织型)、精神分裂症(妄想型)、精神分裂症(混合型)、精神分裂症(残余型)、精神分裂症样精神障碍、短期反应性精神病、情感分裂性精神障碍、诱发性精神病、精神分裂型人格障碍、精神分裂样人格障碍、妄想型人格障碍和妄想(类偏执狂)性精神障碍。

以目前用于基因图谱的两种策略连锁分析和关联性研究,可进行鉴定与一具体性状有关的基因,例如,特定中枢神经系统障碍,象精神分裂症。连锁分析需要研究其成员有多人患病的家族,这种分析现在用于检测单个或寡基因固有性状。与之相反,关联性研究检查与阴性性状(T-)对照组相比无亲戚关系的性状(T+)的个体的标记等位基因频率,这种研究通常用于检测多基因遗传。

本申请揭示了位于与精神分裂症相关的 D-氨基酸氧化酶基因上的另外一些双等位标记。鉴定这些与精神分裂症相关的双等位标记可以进一步限定怀疑含有与发展精神分裂症易感因素有关的遗传决定子的染色体区,而且可导致鉴定本文所公开的与发展精神分裂症易感因素相关的新基因序列。此外,如前所述的 g34872 基因和现在所述的 D-氨基酸氧化酶基因上的双等位标记也可单独使用或联合使用以确定个体发展中枢神经系统障碍的风险。而且,如前所述的 g34872 基因和现在所述的 D-氨基酸氧化酶基因上的双等位标记都可单独使用或联合使用以确定个体通过本发明所述的治疗方法所获得的成效。另外,序列信息提供一种资源可以进一步鉴定那些区域内的新基因和标记。另外,含有精神分裂症相关基因的序列例如可用于分离推定基因族中的其它基因,从其它物种鉴定同系物,治疗疾病及用作本文所述

的诊断和筛选分析中的探针和引物。而且，所鉴定的多态性也可用来设计精神分裂症和双相性精神障碍遗传罹病倾向的可靠性检测分析。它们还可用来设计药物的筛选方案，准确和有效地评估新旧药物或治疗方案的治疗效果和潜在的副作用。

定义

术语“治疗”指的是改善、减轻症状，暂时或永久性消除症状的病因，或者预防或减慢指定障碍或疾病出现症状。

化合物的剂量是能防止所述疾病症状复发或者治疗病人所患疾病的某些症状的有效量。“有效量”、“有效治疗量”、“治疗量”或“有效剂量”指获得所希望的药理或治疗效果，从而有效预防或治疗该疾病的足够量。预防疾病在医学上表现为大大延迟疾病症状的发作时间。治疗疾病表现为减少与疾病相关的症状或者改善疾病症状的复发情况。本领域技术人员给予个体一定量的化合物并观察结果，可轻易确定本发明化合物的有效治疗量。另外，本领域技术人员也熟知鉴定个体患有中枢神经系统障碍的方法，并且能够轻易个体患有中枢神经系统障碍。

术语“拮抗剂”和“抑制剂”看成是同义的，本文可通用。本发明的“拮抗剂”化合物可与典型或非典型抗中枢神经系统障碍的药物如抗精神病药物一起给药。典型抗精神病药物包括：氟派啶醇(haloperidol)、氟奋乃静(fluphenazine)、奋乃静(perphenazine)、氯丙嗪、吗茛酮(molindone)、匹莫齐特(pimozide)、三氟拉嗪和硫利哒嗪、噻嗪化物、噻二唑、恶二唑等等。非典型抗精神病药物包括：氯扎平、利培酮、奥扎平、舍吲哚、M100907、齐拉西酮、思瑞康(seroquel)、佐替平(zotepine)、氨磺必利(amisulpride)、伊潘立酮(iloperidone)、苯乙肼等等。典型抗抑郁和抗焦虑症药物包括：杂环抗抑郁药(三环抗抑郁药、四环抗抑郁药等)、5-羟色胺再吸收抑制剂、混合型 5-羟色胺和去甲肾上腺素再吸收抑制剂、多巴胺再吸收抑制剂和单胺氧化酶抑制剂。拮抗剂也可用来治疗不能采用上述药物进行治疗的个体。本发明也提供一种精神分裂症、双相性精神障碍或其它中枢神经系统障碍的预防和治疗方案，该方法无需用其它抗精神病药物、抗抑郁药、抗焦虑症药或其它病人无反应的药物来治疗。抗精神病药物、抗抑郁药、抗焦虑症药或其它药物可以亚治疗剂量，即较单独用上述药进行治疗时通常所采用的剂量低的剂量给药。被认为是 D-氨基酸氧化酶或 D-天冬氨酸氧化酶抑制剂，或者固有地用作 D-氨基酸氧化酶或 D-天冬氨酸氧化酶抑制剂的治疗精神分裂症、双相性精神障碍、抑郁症及其它中枢神经系统障碍所用的药物明确排除于在 D-氨基酸氧化酶或 D-天冬氨酸氧化酶“拮抗剂”的定义之内，也明确排除于本发明的范围之内。此外，本文所公

开的任何一种分子、化合物或药物可明确排除于本发明的范围之内。

除非另有说明，“烷基”指含有 1 到 8 个碳原子的支链或无支链饱和烃链，例如甲基、乙基、丙基、异丙基、丁基、异丁基、叔丁基、正戊基、正己基等等。

“烷氧基”指-OR 基团，其中 R 如本文所定义。R 优选是含有 1 到 3 个碳原子的支链或无支链饱和烃链。

除非另有说明，“卤素”指氟、氯、溴或碘。

“苯基”包括以选自烷基、烷氧基、羟基、卤素和卤代烷基的取代基任意单取代或多取代的全部可能的异构体苯基。

优选的杂芳基环包括吡咯、呋喃、噻吩、吡啶、嘧啶、哒嗪、吡嗪、三唑、四唑、吡唑、咪唑、异噻唑、噻唑、异恶唑和恶唑。优选的“与苯基稠合的杂芳基”环包括吲哚、异吲哚、苯并呋喃、苯并噻吩、喹啉、异喹啉、喹喔啉、喹唑啉、苯并三唑、吡唑、苯并咪唑、苯并噻唑、苯并异恶唑和苯并恶唑。当用术语杂芳基环时假定包括“与苯基稠合的杂芳基”环在内。术语“饱和或部分不饱和杂环烷基环”指带有 5-7 个环原子并含有 1-3 个选自 N、O 或 S 的杂原子的饱和或部分不饱和(但不是芳香族或完全饱和)杂环。优选的饱和或部分不饱和杂环烷基环包括哌啶、哌嗪、吗啉、四氢吡喃、硫代吗啉或吡咯烷。

术语“药学上可接受的盐类”指该化合物的盐类具有所希望的药理活性并且其在生物上或者其他方面都不是不希望的。用无机酸可形成这些盐类，如乙酸盐、己二酸盐、藻酸盐、天冬氨酸盐、苯甲酸盐、苯磺酸盐、硫酸氢盐、丁酸盐、柠檬酸盐、樟脑酸盐、樟脑磺酸盐、环戊烷丙酸盐、二葡萄糖酸盐、十二烷基硫酸盐、乙烷磺酸盐、富马酸盐、葡萄糖庚酸盐、甘油磷酸盐、半硫酸盐、庚酸盐、己酸盐、盐酸盐、溴酸盐、碘酸盐、2-羟基乙烷磺酸盐、乳酸盐、马来酸盐、甲烷磺酸盐、2-萘磺酸盐、烟酸盐、草酸盐、硫氰酸盐、甲苯磺酸盐和十一烷酸盐。碱性盐包括铵盐、碱金属盐如钠盐和钾盐、碱土金属盐如钙盐和镁盐、带有有机碱的盐如二环己胺盐、N-甲基-D-葡萄糖胺、以及带有氨基酸的盐如精氨酸、赖氨酸等等。也可以用诸如短链卤代烷基如甲基、乙基、丙基和丁基氯化物、溴化物和碘化物，二烷基硫酸盐如二甲基、二乙基、二丁基和二戊基硫酸盐，长链卤化物如癸基、十二烷基、十四烷基和十八烷基氯化物、溴化物和碘化物，卤代芳烷基如苯甲基和苯乙基溴化物等试剂使含有碱性氮的基团季铵化。由此得到水溶性或油溶性或分散性产物。另外，以下叙述药学及药学上可接受的组合物。

本发明的化合物有非对称中心，因而可生成立体异构混合物或单个立体异构体。采用光学活性原料，在合成的某一适当阶段溶解外消旋或非外消旋中间混合物，

或者溶解以一般式(I)表示的化合物可得到单个立体异构体。需要明白的是,本发明的范围包括单个立体异构体以及立体异构混合物(外消旋和非外消旋)。本发明的化合物具有至少一个非对称中心,因而可生成立体异构混合物或单个 R-和 S-立体异构体。在合成的某一适当阶段溶解外消旋或非外消旋中间混合物可得到单个对映体。不用说,本发明的范围包括单个 R-和 S-立体异构体以及立体异构混合物。

“异构体”是具有同一个分子式的不同化合物。

“立体异构体”是原子在空间的排列方式不同的异构体。

“对映体”是一对相互间镜像不重叠的立体异构体。

“非对映异构体”是相互间非镜像的立体异构体。

“外消旋混合物”指含有相等部分的单个非对映异构体的混合物。“非外消旋混合物”是含有不相等部分的单个非对映异构体或立体异构体的混合物。

“取代的烷基”指羧基烷基如乙酰基、氨基烷基、二烷基氨基烷基、羟烷和巯基烷基、烷基甲硅烷基。

本发明涉及以一般式 I-VI 表示的化合物,包括但不限于本文所提供的具体实施例。此外,这些化合物中任何一种都可以制成药学上可接受的盐。

需要注意的是,有机化学领域技术人员可合成本文所述的本发明化合物。

本文所用的术语“精神状态”指病理学上的心理状态,这种状态是精神病或者可能与精神病特征相关。这些状态包括但不限于精神病的诊断和统计手册第四修订版(DSM-IV-R)(1994)所表述的精神障碍,包括精神分裂症和急性躁狂症。DSM-IV-R 由美国命名与统计学会专责小组制作,给出了诊断类别的清晰说明。本领域技术人员知道病理学上的心理状态有不同的命名、疾病分类及分类系统,而这些系统随医学的科学发展逐步形成。

术语“精神分裂症”包括或者可具体限于精神分裂症、精神分裂症样精神障碍、情感分裂性精神障碍、妄想精神障碍、短期精神障碍、由一般医学状况而引起的精神障碍、其它未注明或本文别处所述之精神障碍。这些疾病的症状大部分在 DSM-IV-R 有所定义。这本书涉及这些疾病的章节纳入其整体内容作为参考。

本文所用的术语“双相性精神障碍”指以 DSM-IV-R 及第三修订版(1994)中的类别 296. xx 所述的双相性精神障碍为特征的状态。为了更加清楚起见,申请人考虑了 DSM-IV-R 所述的双相性精神障碍 I 型和双相性精神障碍 II 型的治疗。该术语还包括循环性精神障碍。循环性精神障碍指交替出现抑郁症状和轻度躁狂症症状。本领域技术人员知道病理学上的心理状态有不同的命名、疾病分类及分类系统,而这些系统随医学的科学发展逐步形成。

本文所用与重度抑郁有关的术语“无反应性”指无合理的临床反应的病人(例如, 经过一个或多个常规抗抑郁药的临床疗程后, 病人的汉密尔顿抑郁量表(HAM-D)较基线评分减少 50%)。

“重度抑郁症发作”定义为至少两星期有抑郁情绪或者失去兴趣, 可能伴随其它抑郁症状。这些症状必须在至少连续两个星期内的几乎每天(即 14 天中至少有 10 天)白天大部分时间(即病人至少三分之二清醒时间)出现。“抑郁情绪”通常表现为病人感到悲伤、绝望、无助或无价值。旁人通过例如面部表情、姿势、声音及眼泪观察到病人看起来是悲伤的。儿童和青少年的情绪可以暴躁。“失去兴趣”通常表现为病人对业余爱好没有兴趣或者对以前认为很快乐的活动不再感到任何乐趣。

重度抑郁症发作可伴随其它抑郁症状, 包括不节食时体重急剧下降或体重增加(例如, 一个月内体重变化超过 5%), 或者食欲不振或大增; 失眠或嗜睡; 精神性运动激动或阻抑; 疲倦或精神不振; 无价值感或者过份或不适当自疚; 思考或集中力减退; 或难以作决定; 以及想死亡, 有或无特定计划的自杀念头, 或者试图自杀。

“躁狂症发作”定义为情绪异常及持续高涨、亢奋或易激动发怒的一段明显的时期。情绪异常的这段时期必须持续至少一个星期(或者如果需要接受住院治疗时间可短些)。这种情绪波动必须还要伴随以下列出的至少三种症状: 自我评价过高或变得夸大、对睡眠的需要减少、说话急速、思想活跃、注意力分散、增加投入有目标导向的活动或精神性运动激动、过度投入开心的活动, 但往往会有不愉快的结果。如果情绪容易动怒(而不是高涨或亢奋), 则必须要有以上至少四种症状。情绪波动的严重程度必须足以引起社交或工作能力显著下降或者需要住院治疗, 或者表现为有精神病特征。

“轻度躁狂症发作”比躁狂症发作的程度轻。轻度躁狂症发作的症状通常与躁狂症发作所限定的症状相同, 除了不存在错觉和幻觉外, 而且发作的严重程度不足以引起社交或工作能力显著下降或者个体不需要住院治疗。

本文所用的术语“孤独症”指以 DSM-IV-R 中的类别 296. xx, 包括 299. 00、299. 80 和 299. 10, 优选 299. 00 所述的状态。

术语“焦虑症”包括但不限于强迫症(obsessive-compulsive disorder)、精神活性物质焦虑症、创伤后压力症、广泛性焦虑症、其它未注明之焦虑症、以及器质性焦虑症(organic anxiety disorder)。

本文所用的术语“药物滥用”指不希望有的在身体和/或心理上对药物的依赖。

该术语指对可卡因、迷幻药、大麻、安他非明、致幻剂、苯环利定苯并二氧草类、酒精和烟碱等药物的依赖。

本文所用的术语“注意力缺陷活动过度障碍(ADHD)”指以持续出现疏忽大意、活动过度、冲动或其任意组合为特征的状态或障碍。

本文所用的术语“过度攻击”指以攻击过度以致影响个人的日常生活、人际关系，而且可能例如在企图进行暴力自杀的情况下威胁到个人安全为特征的情况。用本发明要求保护的方法可被治疗的过度攻击不依赖于精神状态，而且与采用的药物或其它物质没有直接关系。

痉挛是一种突然、快速复发、无规律、刻板的运动或发声，发作是不可抗拒的但可以抑制，持续时间不同。常见的简单运动痉挛包括眨眼、颈部抽搐、耸肩、扮鬼脸和咳嗽。常见的简单发声痉挛包括清喉咙、打呼噜、鼻吸气、喷鼻息和咆哮。常见的复杂运动痉挛包括面部表情、修饰行为、跳跃、触摸、跺脚及对一种事物微笑。常见的复杂发声痉挛包括重复上下文的字或词语、污言秽语(使用社会上不接受的词语，往往是很猥亵的)、重复语句(重复自己的发声或词语)，以及模仿语言(重复最后听到的声音、字或词语)。本文所用的术语“痉挛”指包括以一种或多种运动痉挛和一种或多种痉挛以及多种发声痉挛为特征的痉挛障碍。痉挛的例子包括短暂性痉挛障碍、图雷特氏障碍、慢性发声痉挛障碍、如 DSM-IV-R 所述其它未注明之痉挛障碍。

术语“包含”、“由...组成”、“基本上由...组成”有不同的含义，每个术语在本文可被另一个术语代替，改变本发明的范围。

术语“寡核苷酸”及“多核苷酸”在本文是互通的，包括 RNA、DNA 或 RNA/DNA 杂交序列，其具有一个以上单链或双螺旋型核苷酸。本文所用的术语“核苷酸”是形容词，描述含有任何长度单链或双螺旋型的 RNA、DNA 或 RNA/DNA 杂交序列的分子。术语“核苷酸”在本文也用作名词，指个体核苷酸或核苷酸簇，意味着分子、或大核酸分子中的个体单元，对于寡核苷酸或多核苷酸内的核酸，其包括嘌呤或嘧啶、核糖或脱氧核糖分子、以及磷酸盐基因、或磷酸二酯键。虽然术语“核苷酸”在本文也用于包括“修饰后的核苷酸”，其包括至少以下一种修饰：(a) 变换连接基，(b) 嘌呤的类似型，(c) 嘧啶的类似型，或 (d) 糖的类似型，例如，类似连接基、嘌呤、嘧啶及糖的例子可参看本文引用的参考文献 PCT 公开号 WO 95/04064。不过，本发明的多核苷酸优选包括大于 50% 常规脱氧核糖核苷酸，最好包括大于 90% 常规脱氧核糖核苷酸。本发明的多核苷酸序列可用任何一种已知的方法制备，包括合成、重组、体内产生、或其组合、以及用本技术领域已知的任何一种纯化方法。

本文的术语“纯化的”用于描述与其它化合物分离的本发明的多核苷酸或多核苷酸载体，包括但不限于：其它核酸、碳水化合物、脂质和蛋白质(例如合成多核苷酸所用的酶)、或者用于指共价密闭多核苷酸与线性多核苷酸分离的。当样本至少大约 50%，最好 60–75%具有单个多核苷酸序列和构象(线性对共价密闭)时，一种多核苷酸基本上是纯的。基本纯的多核苷酸通常占核酸样本的约 50%，最好为 60–90%(重量/重量)，更常见的是纯度约为 95%，最好是 99%以上。众所周知，多核苷酸纯度或均一性有许多方式表示，如先对样本进行琼脂糖或聚丙烯酰胺凝胶电泳，再观察凝胶染色后的单一条带。对于某些用途，用高效液相层析法或本技术领域公知的其它方法可获得高分辨率。当样本至少大约 50%，最好 60–75%具有单个多肽序列时，一种多肽基本上是纯的。基本纯的多肽通常占核酸样本的约 50%，最好为 60–90%(重量/重量)，更常见的是纯度约为 95%，最好是 99%以上。众所周知，多肽纯度或均一性有许多方式表示，如先对样本进行琼脂糖或聚丙烯酰胺凝胶电泳，再观察凝胶染色后的单一条带。对于某些用途，用高效液相层析法或本技术领域公知的其它方法可获得高分辨率。术语“纯化的”在本文用于描述与其它化合物分离的本发明的化学组合物。

术语“分离的”要求物质离开它原来的环境(例如，若物质是天然存在的，就离开天然环境)。例如，天然存在于有生命的动物的多核苷酸或多肽就不是分离的，但自天然系统中部分或全部共存物质分离出来的同一些多核苷酸或 DNA 或多肽就是分离的。这些多核苷酸可以是部分载体和/或这些多核苷酸或多肽可以是部分组合物，由于载体或组合物不属于天然环境的一部分，故这些多核苷酸仍然是分离的。

术语“引物”表示与靶核苷酸序列互补并用于与靶核苷酸序列杂交的特异性寡核苷酸序列。引物用作以 DNA 聚合酶、RNA 聚合酶或逆转录酶为催化剂的核苷酸聚合反应的引发点。

术语“探针”表示可用于鉴定样本所含的特异性多核苷酸序列的有限核酸节段(或者核苷酸类似物节段，例如本文所限定的多核苷酸)，所述核酸节段包括与待鉴定的特异性多核苷酸序列互补的核苷酸序列。

术语“性状”和“表型”在本文是互通的，指有机体在临床上任何看得见的、检测得到的或其它可测量的性质，如疾病症状或疾病易感性。本文所用的术语“性状”或“表型”通常指精神分裂症或双相性精神障碍的症状或易感性；或者指个体对作用于精神分裂症或双相性精神障碍的药剂有反应；或者指作用于精神分裂症或双相性精神障碍的药剂副作用的症状或易感性。

本文所用的术语“等位基因”指核苷酸序列的变异型。双等位基因多态性有

两种型式。通常将第一种鉴定的等位基因命名为原始型等位基因，而其它等位基因命名为变换型等位基因。二倍体有机体的等位基因型可以是纯合或杂合的。

本文所用的术语“杂合率”指个体在一以特定等位基因杂合的群体中的发生率。在双等位基因系统中，杂合率是平均来说等于 $2P_a(1-P_a)$ ，式中 P_a 是最小出现的等位基因的频率。为了用在遗传研究中，遗传标记应该有足够程度的杂合性，以便使随机选择一个人是杂合的概率比较高。

本文所用的术语“基因型”指个体或样本中所含的等位基因同一性。在本发明的上下文中，基因型最好用于描述个体或样本所含的双等位标记等位基因。术语使样本或个体的双等位标记“基因型分型”包括确定个体在双等位标记上携带的特异性等位基因或特异性核苷酸。

本文所用的术语“突变”指频率低于 1% 的不同基因组或个体之间的 DNA 序列差异。

本文所用的术语“单倍型”指个体或样本在单个染色体上所含有的等位基因组合。在本发明的上下文中，单倍型最好指在一给定个体中发现的并且与表型相关的双等位标记等位基因的组合。

本文所用的术语“多态性”指不同基因组或个体之间出现两种或多种变换基因组序列或等位基因。“多态的”指在一群序列中可找到两种或多种特异性基因组序列变异型的状况。“多态位点”是变异发生的基因座。多态性可包括一个或多个核苷酸的取代、缺失和插入。单核苷酸多态性是单一碱基对变化。单核苷酸多态性通常是在多态位点上用另一个核苷酸取代一个核苷酸。单核苷酸的缺失或单核苷酸的插入也会产生单核苷酸多态性。在本发明的上下文中，“单核苷酸多态性”最好指单核苷酸的取代。通常，不同基因组或不同个体之间的多态位点可以被两个不同核苷酸所置换。

术语“双等位基因多态性”和“双等位标记”在本文是互通的，指在一群体中具有两个频率相对较高的等位基因的多态性，优选指单核苷酸多态性。“双等位标记等位基因”指双等位标记位点所含有的核苷酸变异型。本发明双等位标记中最少出现的等位基因的频率被证实大于 1%，该频率宜大于 10%，较好的频率是至少 20% (即杂合率至少 0.32)，更好的频率是 30% (即杂合率至少 0.42)。最少出现的等位基因的频率为 30% 或以上的双等位标记称为“高质量双等位标记”。本发明所有的基因分型、单倍分型、关联性和相互作用的研究方法可任选地仅用高质量双等位标记进行。

本文以下列方式叙述多核苷酸中的核苷酸相对多核苷酸中心的位置。当多核

核苷酸有奇数核苷酸时，把与多核苷酸 3'和 5'端距离相等的核苷酸看成是多核苷酸的“中心”，紧邻在中心位置的核苷酸的任何一个多核苷酸，或者在中心位置的核苷酸本身认为“在中心的 1 个核苷酸以内”。在有奇数核苷酸的多核苷酸中，把在多核苷酸中心的 5 个核苷酸位置上的任何一个核苷酸认为在中心的 2 个核苷酸以内，如此类推。当多核苷酸有偶数核苷酸时，多核苷酸的中心是一个键而不是核苷酸。所以，把 2 个中心核苷酸中任何一个认为“在中心的 1 个核苷酸以内”，把在多核苷酸中心的 4 个核苷酸位置上的任何一个核苷酸认为“在中心的 2 个核苷酸以内”，如此类推。对于涉及一个或多个核苷酸的取代、插入或缺失，如果多态性的取代、插入或缺失的多核苷酸与多核苷酸的 3'端之间的距离与多态性的取代、插入或缺失的多核苷酸与多核苷酸的 5'端之间的距离之差等于零或 1 个核苷酸，则多态性、等位基因或双等位标记就在多核苷酸的“中心”上。如果这个差距等于零到 3 个核苷酸，就认为多态性“在中心的 1 个核苷酸以内”。如果这个差距等于零到 5 个核苷酸，就认为多态性“在中心的 2 个核苷酸以内”。如果这个差距等于零到 7 个核苷酸，就认为多态性“在中心的 3 个核苷酸以内”，如此类推。对于涉及一个或多个核苷酸的取代、插入或缺失，如果多态性的取代、插入或缺失的多核苷酸与多核苷酸的 3'端之间的距离与多态性的取代、插入或缺失的多核苷酸与多核苷酸的 5'端之间的距离之差等于零或 1 个核苷酸，则多态性、等位基因或双等位标记就在多核苷酸的“中心”上。如果这个差距等于零到 3 个核苷酸，就认为多态性“在中心的 1 个核苷酸以内”。如果这个差距等于零到 5 个核苷酸，就认为多态性“在中心的 2 个核苷酸以内”。如果这个差距等于零到 7 个核苷酸，就认为多态性“在中心的 3 个核苷酸以内”，如此类推。

本文所用的术语“上游”指从特定参考点面向多核苷酸 5'端的位置。

本文所用的术语“碱基配对”和“Watson & Crick 碱基配对”在本文是互通的，指由于序列同一性可通过氢键相互结合的核苷酸，例如在双螺旋 DNA 中可以发现胸腺嘌呤或尿嘧啶残基通过两个氢键与腺嘌呤残基连接，而胞嘧啶残基通过三个氢键与鸟嘌呤残基连接(参看 Stryer, L., Biochemistry, 第 4 版, 1995 年)。

本文所用的术语“互补的”或“其补体”指能够在整个互补区与另一个特异性多核苷酸形成 Watson & Crick 碱基配对的多核苷酸序列。该术语适用于多核苷酸对，它们仅基于其序列而非基于两个多核苷酸实际结合的任何一组特定条件。

术语“D-氨基酸氧化酶基因”，使用时包括编码本发明任何 D-氨基酸氧化酶蛋白质的基因组、mRNA 和 cDNA 序列，包括基因组 DNA 的非翻译调节区。

术语“g34872 基因”，使用时包括编码任何 g34872 蛋白质的基因组、mRNA 和

cDNA 序列, 包括基因组 DNA 的非翻译调节区。

术语“D-天冬氨酸氧化酶基因”, 使用时包括编码任何 D-天冬氨酸氧化酶蛋白质的基因组、mRNA 和 cDNA 序列, 包括基因组 DNA 的非翻译调节区。

本文所用的术语“13q31-q33 相关双等位标记”涉及一组位于人染色体 13q31-q33 区内的双等位标记。术语 13q31-q33 相关双等位标记包括美国专利申请号 09/539,333 与国际专利申请号 PCT/IB00/00435 中表 6b 所公开的全部双等位标记, 及与其连锁不平衡的任何双等位标记, 以及表 6c(同一个美国专利申请号 09/539,333 与国际专利申请号 PCT/IB00/00435)所公开的任何双等位标记, 及与其连锁不平衡的任何双等位标记, 两件申请资料纳入其整体内容作为参考。本发明优选的染色体 13q31-q33 相关双等位标记等位基因包括(同一个美国专利申请号 09/539,333 与国际专利申请号 PCT/IB00/00435)表 6b 各自公开的每一个等位基因或由所列等位基因全部可能的组合组成的组。

本文所用的术语“D 区相关双等位标记”涉及一组与染色体 13q31-q33 区的亚区(本文称为 D 区)连锁不平衡的双等位标记。术语 D 区相关双等位标记包括美国专利申请号 09/539,333 与国际专利申请号 PCT/IB00/00435 中表 6b 所公开的双等位标记 A1 到 A242、A249 到 A251、A257 到 A263、A269 到 A270、A278、A285 到 A299、A303 到 A307、A324、A330、A334 到 A335、A346 到 A357 和 A361 到 A489, 以及与其连锁不平衡的任何双等位标记, 两件申请公开的资料纳入其整体内容作为参考。

本文所用的术语“sbg1 相关双等位标记”涉及一组与 sbg1 基因或 sbg1 核苷酸序列连锁不平衡的双等位标记。术语 sbg1 相关双等位标记包括美国专利申请号 09/539,333 与国际专利申请号 PCT/IB00/00435 中表 6b 所公开的双等位标记 A85 到 A219, 以及与其连锁不平衡的任何双等位标记, 两件申请公开的资料纳入其整体内容作为参考。

本文所用的术语“g34665 相关双等位标记”涉及一组与 g34665 基因或 sbg1 核苷酸序列连锁不平衡的双等位标记。术语 g34665 相关双等位标记包括表 6b 所公开的双等位标记 A230 到 A236, 以及与其连锁不平衡的任何双等位标记。

术语“多肽”指氨基酸的聚合物, 与聚合物的长度无关; 所以, 肽、寡肽和蛋白质都包括在多肽的定义内。该术语也不指定或排除多肽的表达后修饰, 例如, 与糖基、乙酰基、磷酸基团、脂类基团等共价连接的多肽明确包括在术语多肽中。该定义还包括含有一个或多个氨基酸类似物(例如, 包括非天然存在的氨基酸、只

在不相关生物系统中天然存在的氨基酸、来自哺乳动物系统经过修饰的氨基酸等)的多肽、带有取代键以及本领域技术公知的其它修饰的多肽,无论是天然存在或非天然存在。

本文所用的术语“非人动物”指任何不是人的脊椎动物、鸟类,通常指哺乳动物,最好是灵长目动物、家畜,如猪、山羊、绵羊、驴和马、兔或者啮齿目动物,最好是大鼠或小鼠。本文所用的术语“动物”用于指任何一种脊椎动物,最好是哺乳动物。术语“动物”和“哺乳动物”明确包括人受试者,除非前面带有术语“非人动物”。

本文所用的术语“抗体”指包括至少一个结合结构域的多肽或多肽群,其中抗体结合结构域通过折迭抗体分子的可变区而形成,与内表面形状及电荷分布形成三维结合空间,该空间与抗原的抗原决定子的特征互补,可以与抗原进行免疫反应。抗体包括重组蛋白质,所述蛋白质包括结合结构域,以及片段,而片段包括 Fab、Fab'、F(ab)₂ 和 F(ab')₂ 片段。

本文所用的“抗原决定子”是决定抗原—抗体反应特异性的部分抗原分子,在本发明指 sbg1 多肽。“表位”指多肽的抗原决定子。一个表位在空间构象中可包括至少 3 个氨基酸,而一个表位就有唯一一个空间构象。通常,一个表位包括至少 6 个这样的氨基酸,较常见包括至少 8-10 个氨基酸。确定组成表位的氨基酸的方法包括 X 射线结晶学、二维核磁共振及表位作图法,例如 Geysen 等人(1984 年)、PCT 公开号 W0 84/03564 及 PCT 公开号 W0 84/03506 所述的肽扫描法。

在 2000 年 3 月 30 日提交、题为“与精神分裂症有关的基因、蛋白质及双等位标记”的共同待批美国专利申请号 09/539, 333 与共同待批国际专利申请号 PCT/IB00/00435 对“变异型和片段”、“核酸或多肽之间的同一性”、“严谨杂交条件”、“能够在重组的细胞宿主和转基因动物中使 sbg1 核酸序列的时空表达定向的 DNA 构造体”作了详细和完整的叙述,以上两件申请公开的资料均纳入其整体内容作为参考。

g34872 和 D-氨基酸氧化酶多核苷酸的基因组序列

本发明特别优选的 g34872 核酸包括分离的、纯化的或重组的多核苷酸,其包含,或者基本上由或由一段连续跨距组成,该跨距至少为 12、15、18、20、25、30、35、40、50、60、70、80、90、100、150、200、500 或 1000 个核苷酸或其补体,它们在美国专利申请号 09/539, 333 的 SEQ ID No: 1 的核苷酸位点 213818 到 243685 上(美国专利申请号 09/539, 333 与国际专利申请号 PCT/IB00/00435 公

开的资料纳入其整体内容作为参考)。

本发明的D-氨基酸氧化酶多核苷酸如本发明的SEQ ID No: 1所叙述。本发明特别优选的核酸包括分离的、纯化的或重组的多核苷酸,其包含,或者基本上由或由一段连续跨距组成,该跨距至少为12、15、18、20、25、30、35、40、50、60、70、80、90、100、150、200、500或1000个核苷酸,它们在SEQ ID No: 1在核苷酸位点6000到86600上。本发明核酸包括任何一种来源的D-氨基酸氧化酶核酸,包括灵长类动物、非人灵长类动物、哺乳动物和人D-氨基酸氧化酶核酸。

本发明的较优选核酸包括分离的、纯化的或重组的多核苷酸,其包含一段连续跨距,该跨距至少为12、15、18、20、25、30、35、40、50、60、70、80、90、100、150、200、500或1000个SEQ ID No: 1的核苷酸或其补体,其中所述连续跨距包含与D-氨基酸氧化酶相关的双等位标记。所述双等位标记任选地选自包含24-1443/126、24-1457/52或24-1461/256的组。所述双等位标记优选是24-1461/256。

需要注意的是,本章节所述的多核苷酸包括任何尺寸和序列的核酸片段。

所以,本发以分离的、纯化的或重组的多核苷酸进行实施例,所述多核苷酸含有一段核苷酸序列,其选自由D-氨基酸氧化酶基因(SEQ ID No: 1)的外显子或其互补序列组成的组。优选纯化的、分离的或重组的多核苷酸含有至少一个D-氨基酸氧化酶基因的外显子,或其互补序列或其片段或变异型。本发明还包括纯化的、分离的或重组的核酸,其含有至少2个D-氨基酸氧化酶基因的外显子的组合,所述外显子选自由外显子Z、A、B、C、Ulong、U、V、Z、2、3、4、5、6、7、8、9、10、11和11long组成的组,其中多核苷酸按照SEQ ID No: 1的同一对应次序排列在核酸内。

本发明特别优选的核酸包括分离的、纯化的或重组的多核苷酸,其包含一段连续跨距,该跨距至少为12、15、18、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100或200个SEQ ID No: 1的核苷酸或其补体。

本发明的另一个目的在于一种纯化的、分离的或重组的多核苷酸,在以上所述的严谨杂交条件下与SEQ ID No: 1的D-氨基酸氧化酶核苷酸序列、或其互补序列或其变异型杂交。

本发还包括纯化的、分离的或重组的多核苷酸,所述多核苷酸含有一核苷酸序列,其选自由D-氨基酸氧化酶基因(SEQ ID No: 1)的内含子或其互补序列组成的组。

在另一实施例中,本发明包括D-氨基酸氧化酶基因及D-氨基酸氧化酶基因组

序列，其由或基本上由 SEQ ID No: 1 的核苷酸位点的序列，其互补序列以及其片段和变异型组成，或者包含 SEQ ID No: 1 的核苷酸位点的序列，其互补序列以及其片段和变异型。

本发明还包括一种纯化的、分离的或重组的多核苷酸，其包含一段 D-氨基酸氧化酶核苷酸序列，所述序列与 SEQ ID No: 1 的核苷酸、或其互补序列或其片段具有至少 70、75、80、85、90 或 95% 同一性。

这些核酸及其片段和变异型可用作寡核苷酸引物或探针，以检测测试样本中所含的 g34782、D-氨基酸氧化酶或 D-天冬氨酸氧化酶核酸序列的拷贝，或用以扩增 g34782、D-氨基酸氧化酶或 D-天冬氨酸氧化酶核酸序列或相邻区内的靶核苷酸序列。

本发明的其它优选核酸包括分离的、纯化的或重组的 D-氨基酸氧化酶多核苷酸，其包含一段连续跨距，该跨距至少为 12、15、18、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100 或 200 个 SEQ ID No: 1 的核苷酸或其补体，其中所述的连续跨距含有至少一个双等位标记。所述的连续跨距任选地含有一个与 D-氨基酸氧化酶相关的双等位标记。需要注意的是，本章节所述的多核苷酸包括任何尺寸和序列的核酸片段。所述双等位标记可含有原始型或变换型等位基因。

本发明的其它核酸包括分离的、纯化的或重组的多核苷酸，其包含一段连续跨距，该跨距至少为 12、15、18、20、25、30、35、40、50、60、70、80、90、100、150、200 或 500 个核苷酸，达到这样的程度以致所述的跨距与 SEQ ID No: 1 的核苷酸位点范围一致，其中所述的连续跨距包含至少 1、2、3、5 或 10 个 SEQ ID No: 1 的下列核苷酸位点：215820 到 215941、216661 到 217009、230409 到 290721、231272 到 231411、234202 到 234321、240528 到 240567、240528 到 240827 和 240528 到 240996，或其补体，以及与所述的跨距有至少 70、75、80、85、90 或 95% 核苷酸同一性的多核苷酸和能够与所述的跨距杂交的多核苷酸。

本发明还包括一种纯化的、分离的或重组的多核苷酸，其编码具有 SEQ ID No: 7-10 任何之一的氨基酸序列的 D-氨基酸氧化酶蛋白质，或其多肽片段或变异型。

虽然本节题为“sbg1 基因组序列”，但要注意本节所述的多核苷酸也包括任何尺寸和序列的核酸片段，使基因组序列 sbg1 位于其中一侧或者两段或多段这样的基因组序列之间。

D-氨基酸氧化酶 cDNA 序列

业已表明, D-氨基酸氧化酶基因的表达可产生几种 mRNA。SEQ ID No: 2-6 所示为对应这些 mRNA 的几段 cDNA 序列。本发明包括一种纯化的、分离的或重组的核酸, 其包含一选自 SEQ ID No: 2-6 的核苷酸序列、其互补序列、其剪接变异型以及其等位基因变异型和片段。此外, 本发明优选的多核苷酸包括纯化的、分离的或重组的 D-氨基酸氧化酶 cDNA, 其由或基本上由一选自 SEQ ID No: 2-6 的核苷酸序列组成, 或者包含一选自 SEQ ID No: 2-6 的核苷酸序列。本发明特别优选的多核苷酸包括分离的、纯化的或重组的多核苷酸, 其包含一段连续跨距, 该跨距至少为 8、12、15、18、20、25、30、35、40、50、60、70、75、80、100、200 或 500 个核苷酸, 达到这样的程度以致所述的连续跨距与 SEQ ID No: 2-6 或其补体的长度一致。

需要注意的是, 本节所述的多核苷酸也包括任何尺寸和序列的核酸片段。

本发明还涉及纯化的或分离的核酸, 该核酸含有的多核苷酸与选自由 SEQ ID No: 2-6 组成的组的多核苷酸, 或其互补序列, 或其生物活性片段, 具有至少 70、80、85、90 或 95%核苷酸同一性, 有利的为 99%核苷酸同一性, 优选的为 99.5%核苷酸同一性, 最好的为 99.8%核苷酸同一性。

本发明的另一个目的涉及纯化的、分离的或重组的核酸, 该核酸含有一在本文所限定的严谨杂交条件下与选自由 SEQ ID No: 2-6 组成的组的多核苷酸、或其互补序列或其变异型或生物活性片段进行杂交的多核苷酸。SEQ ID No: 2-6 的 D-氨基酸氧化酶 cDNA 型式在序列列表中有进一步的叙述。

用来分离特定 D-氨基酸氧化酶 cDNAs 或进行基因分型的引物列在 SEQ ID No: 1 中。D-氨基酸氧化酶的双等位标记及其基因分型的引物列在 SEQ ID No: 1、24、26 和 29 中。g34872 多核苷酸列在 SEQ ID No: 14 和 16 中。g34872 双等位标记 99-16105-152 列在 SEQ ID No: 12, g34872 双等位标记 99-5919-215 列在 SEQ ID No: 13 中, 其中还叙述制备的引物。g34872 的 cDNA 列在 SEQ ID No: 14 中, 而用在双杂交实验中的多核苷酸列在 SEQ ID No: 16。

本发明还鉴定了取自各种组织的 cDNA 序列中的新外显子和变异, 它们列为 SEQ ID No: 1 的外显子 11long、Z、A、B、C 和 UL, 并列在 SEQ ID No: 2-6 的多核苷酸中。D-氨基酸氧化酶多肽的新型式列在 SEQ ID No: 8-10 中。

这些变异型表示 D-氨基酸氧化酶较罕见的新型式, 它们最好用来筛选中枢神经系统障碍的治疗方法所用的组合物。

需要注意的是, 本节所述的多核苷酸也包括任何尺寸和序列的核酸片段, 使 g34872、D-氨基酸氧化酶和 D-天冬氨酸氧化酶基因组序列位于其中一侧或者两段

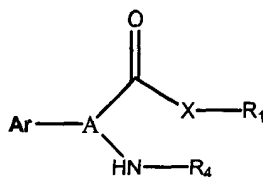
或多段这些基因组序列之间。

D-氨基酸氧化酶和 D-天冬氨酸氧化酶拮抗剂

本文所用的术语“拮抗剂”指抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶将 D-氨基酸底物转化为相应的 α -酮酸的酶反应。拮抗剂可分为竞争性、非竞争性、无竞争性、变构性或不可逆性 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的酶活性抑制剂。术语 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的“活性”或“酶活性”指上述的酶反应。拮抗剂可按 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的活性抑制程度进行分类。优选的拮抗剂使 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的活性下降至少 10%、20%、30%、40%、50%、60%、70%、80%、90%或 95%。抑制的效果也可以抑制常数或 K_i (M) 值表示。优选的拮抗剂的 K_i (M) 值少于 5×10^{-2} 、 1×10^{-2} 、 5×10^{-3} 、 1×10^{-3} 、 5×10^{-4} 、 1×10^{-4} 、 5×10^{-5} 、 1×10^{-6} 、 5×10^{-7} 、 1×10^{-7} 。需要注意的是， K_i (M) 值与抑制效果成反比，即 K_i (M) 值下降，抑制效果增大。所以，拮抗剂也可按它们对 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的特异性来分类。本发明包括抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的活性但不抑制其它人黄素蛋白(对羟基苯甲酸盐羟化酶、胆固醇氧化酶和葡萄糖氧化酶)或者其它人黄素蛋白的 K_i (M) 数值大于 1×10^{-2} 、 5×10^{-2} 的拮抗剂。由定义可知，通用术语“拮抗剂”和“抑制剂”可互通使用，表示以上所限定的抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶活性的任何一种组合物。另外，可按本说明书所述独立地列出拮抗剂或抑制剂的具体类别，例如竞争性抑制剂。

迄今为止，已研究了超过 200 种 D-氨基酸氧化酶和 D-天冬氨酸氧化酶抑制剂。D-氨基酸氧化酶和 D-天冬氨酸氧化酶拮抗剂可选自以上所提供的组合物，或者本领域已知的，或用本文所述或本领域已知的方法制备的其它拮抗剂。另外，D-氨基酸氧化酶和 D-天冬氨酸氧化酶拮抗剂可向市场上的供应商购买。没有任何限制作用的用于本发明的化合物的例子见表 I。D-氨基酸氧化酶和 D-天冬氨酸氧化酶拮抗剂还包括选自由包括以下组合物组的组合物家族：本文所提供的竞争性抑制剂组合物、不可逆抑制剂组合物、以一般式 I、一般式 II、一般式 III、一般式 IV、一般式 V 和一般式 VI 表示的组合物及其亚组。较优选的以一般式 I-VI 表示的典型组合物及其亚组包括但不限于下文详细说明。

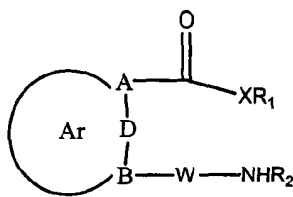
以一般式 I 表示的组合物或其药学上可接受的盐类以包括下列结构式的结构表示：



式中,

- a) A 是烷基, 例如甲基、乙基、丙基或丁基; 支链烷基, 例如异丁基、异丙基、异戊基或者环烷基, 例如环丙基、环戊基或环己基。这些基团本身可由 C_1 - C_6 烷基、卤素、羟基或氨基取代;
- b) X 是氧或氮;
- c) Ar 是芳香族单环、双环或三环稠合杂环, 其中环或者未取代或者 1 至 5 位上由氢、卤素、羟基、 $-\text{CN}$ 、 COR_2 、 $-\text{CONR}_2\text{R}_3$ 、 $-\text{S}(\text{O})_n\text{R}_2$ 、 $-\text{OP}(\text{OR}_2)\text{OR}_3$ 、 $-\text{PO}(\text{OR}_3)\text{R}_3$ 、 $-\text{OC}(\text{O})\text{NR}_2\text{R}_3$ 、 $-\text{COOR}_2$ 、 $-\text{CONR}_2\text{R}_3$ 、 $-\text{SO}_3\text{H}$ 、 $-\text{NR}_2\text{R}_3$ 、 $-\text{NR}_2\text{COR}_3$ 、 $-\text{NR}_3\text{COOR}_3$ 、 $-\text{SO}_2\text{NR}_2\text{R}_3$ 、 $-\text{N}(\text{R}_2)\text{SO}_2\text{R}_3$ 、 $-\text{NR}_2\text{CONR}_2\text{R}_2$ 、 $-\text{SO}_2\text{NHCOR}_2$ 、 $-\text{CONHSO}_2\text{R}_2$ 、 $-\text{SO}_2\text{NHCN}$ 、 $-\text{OR}_1$ 、 C_1 - C_6 直链或支链烷基或链烯基、或者由一个或多个卤素、羟基、氨基、羧基、羧酰胺、腈、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、磷酸盐、 Ar^1 、 N_3 或其组合取代的 C_1 - C_6 支链或直链烷基或链烯基, 其中杂环含有 1-6 个选自由氧、氮、硫及其组合组成的组的杂原子;
- d) R_4 是 H、烷基、 Ar^1 、O、取代的烷基;
- e) R^1 是 $(\text{C}_1 - \text{C}_6)$ 烷基、 Ar^1 、 $(\text{C}_1 - \text{C}_4)$ 烷氧基羰基甲基、取代的烷基;
- f) R_2 和 R_3 各自独立地是氢、 C_1 - C_6 直链或支链烷基或链烯基、或者由一个或多个卤素、羟基、氨基、羧基、羧酰胺、腈、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、磷酸盐、 Ar^1 、 N_3 取代的 C_1 - C_6 支链或直链烷基或链烯基; 以及
- g) Ar^1 是单环、双环或三环、碳环或杂环, 其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1 - C_6 直链或支链烷基或链烯基、 C_1 - C_4 烷氧基、 C_1 - C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代; 其中单个环的大小是 3-7 个原子; 而且杂环含有 1-6 个选自由氧、氮、硫及其组合组成的组的杂原子。

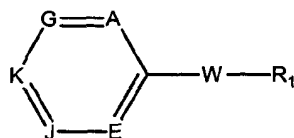
优选的以一般式 I 表示的组合物或其药学上可接受的盐类是一般式 Ia 的组合物或其药学上可接受的盐类, 其包括下列结构式:



式中,

- a) A 和 B 由碳或氮组成, D 可含有 0-2 个由碳或氮组成的原子;
- b) W 是 C_1-C_4 烷基, 如 $(CH_2)_n$ 、支链烷基;
- c) n 是 0-4。此外, 当 $n = 0$ 时假定 $-NHR_2$ 与 B 共价键合;
- d) X 是氧或氮;
- e) R_2 是氢、烷基、 Ar^1 、或氧取代的烷基;
- f) R^1 是 $(C_1 - C_6)$ 烷基、 Ar^1 、 $(C_1 - C_4)$ 烷氧羰基甲基或取代的烷基;
- g) Ar 是芳香族单环、双环或三环稠合杂环, 其中环或者未取代或者 1 至 6 位上由卤素、羟基、硝基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、苯氧基、苄氧基、氨基、 C_3-C_6 环烷基或其组合取代; 其中单个环的大小是 5-6 个原子; 而且杂环含有 1-6 个选自由氧、氮、硫及其组合组成的组的杂原子; 以及
- h) Ar^1 是单环、双环或三环、碳环或杂环, 其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代; 其中单个环的大小是 3-7 个原子; 而且杂环含有 1-6 个选自由氧、氮、硫及其组合组成的组的杂原子。

优选的以一般式 Ia 表示的组合物或其药学上可接受的盐类是一般式 Ib 的组合物或其药学上可接受的盐类, 其包括下列结构式:



式中,

- a) A、G、K、J、E 是六元碳环或杂环芳族环的原子, 其中杂环含有 1-6 个选自由碳、氮及其组合组成的组的杂原子;
- b) A、G、K、J、E 可各自独立地是未取代或由氢、卤素、羟基、 $-CN$ 、 COR_2 、 $-CONR_2R_3$ 、 $-S(O)_nR_2$ 、 $-OP(O)(OR_2)OR_3$ 、 $-PO(OR_3)R_3$ 、 $-OC(O)NR_2R_3$ 、 $-COOR_2$ 、 $-CONR_2R_3$ 、 $-SO_3H$ 、 $-NR_2R_3$ 、 $-NR_2COR_3$ 、 $-NR_3COOR_3$ 、 $-SO_2NR_2R_3$ 、 $-$

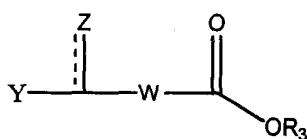
- $N(R_2)SO_2R_3$ 、 $--NR_2CONR_2R_2$ 、 $--SO_2NHCOR_2$ 、 $--CONHSO_2R_2$ 、 $--SO_2NHCN$ 、 $--OR_1$ 、 C_1-C_6 直链或支链烷基或链烯基取代，或者由一个或多个卤素、羟基、氨基、羧基、羧酰胺、腈、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、磷酸盐、 Ar^1 、 N_3 取代的 C_1-C_6 支链或直链烷基或链烯基；
- c) R_1 是 CN 、 COR_2 、 $--CONR_2R_3$ 、 $--S(O)_nR_2$ 、 $--OPO(OR_2)OR_3$ 、 $--PO(OR_3)R_3$ 、 $--OC(O)NR_2R_3$ 、 $--COOR_2$ 、 $--CONR_2R_3$ 、 $--SO_3H$ 、 $--NR_2R_3$ 、 $--NR_2COR_3$ 、 $--NR_3COOR_3$ 、 $--SO_2NR_2R_3$ 、 $--N(R_2)SO_2R_3$ 、 $--NR_2CONR_2R_2$ 、 $--SO_2NHCOR_2$ 、 $--CONHSO_2R_2$ 、 $--SO_2NHCN$ 、 SCN 、 $COCO_2H$ 、 C_1-C_6 直链或支链烷基或链烯基、或者由一个或多个卤素、羟基、氨基、羧基、羧酰胺、腈、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、磷酸盐、 Ar^1 、 N_3 取代的 C_1-C_6 支链或直链烷基或链烯基；
- d) W 是 N 、 $(CH_2)_x$ 、或 $-NCH_2$ ；
- e) $x=0-4$ ；
- f) $n=0-2$ ；
- g) R_2 和 R_3 各自独立地是氢、 C_1-C_6 直链或支链烷基或链烯基、或者由一个或多个卤素、羟基、氨基、羧基、羧酰胺、腈、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、磷酸盐、 Ar^1 、 N_3 取代的 C_1-C_6 支链或直链烷基或链烯基；以及
- h) Ar^1 是单环、双环或三环、碳环或杂环，其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代；其中单个环的大小是 5-6 个原子；而且杂环含有 1-6 个选自由氧、氮、硫及其组合组成的组的杂原子。

以一般式 I、Ia 和 Ib 表示的组合物或其药学上可接受的盐类的具体例子包括但不限于以下化合物：

- 苯甲酸；
- 2-氨基苯甲酸盐；
- 3-氨基苯甲酸盐；
- 4-氨基苯甲酸盐；
- 水杨酸；
- N-甲基烟酸盐；
- 6-甲基烟酸甲酯；

- h) 2-甲基烟酸乙酯;
- i) 氨基苯甲酸盐;
- j) 2-氨基苯甲酸乙酯;
- k) 2-氨基苯甲酸甲酯;
- l) 吡啶甲酸盐;
- m) 2-吡啶羧酸乙酯;
- n) 3-甲基苯甲基硫氰酸盐;
- o) 苯基丙酮酸;
- p) 苯基乙醛酸;
- q) 1-甲基吡啶鎓-3-羧酸盐;
- r) 贝氟沙酮(Befloxatone); (5R)-5-(甲氧基甲基)-3-[4-[(3R)-4,4,4-三氟-3-羟基丁氧基]苯基]-2-恶唑烷酮;
- s) 丁氨苯丙酮; 1-(3-氯苯基)-2-[(1,1-二甲基乙基)氨基]-1-丙酮;
- t) 可铁宁(Cotinine); 1-甲基-5-(3-吡啶基)-2-吡咯烷酮;
- u) 度洛西汀(Duloxetine); (γ S)-N-甲基- γ -(1-萘氧基)-2-噁吩丙胺;
- v) 芬戊二醇; 2-(4-氯苯基)-4-甲基-2,4-戊二醇;
- w) 氟伏沙明(Fluvoxamine); (E)-5-甲氧基-1-[4-(三氟甲基)苯基]-1-戊酮
O-(2-氨基乙基)肟;
- x) Iproclozide; 4-(氯苯氧基)乙酸 2-(1-甲基乙基)酰肼;
- y) 异丙烟肼; 4-吡啶羧酸 2-(1-甲基乙基)酰肼;
- z) 左法哌酯(levophacetoperane); α -苯基-2-哌啶甲醇乙酸盐;
- aa) 咯利普兰(Rolipram); 4-[3-(环戊氧基)-4-甲氧基苯基]-2-吡咯烷酮;
- bb) 反苯环丙胺(Tranlycypromine); (1R, 2S)-rel-2-苯基环丙胺; 及
- cc) 米那普仑(Milnacipran); (1R, 2S)-rel-2-(氨基甲基)-N,N-二乙基-1-苯基环丙烷羧酰胺。

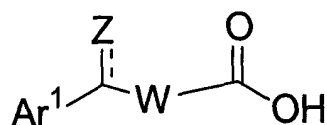
以一般式 II 表示的组合物或其药学上可接受的盐类以包括下列结构式的结构表示:



式中,

- a) $W = (CH_2)_n$;
- b) $n = 0-5$;
- c) Z 是氧或羟基;
- d) $Y = H, Ar^1, R_4(CH_2)_x, R_1S(CH_2)_x, R_1SO(CH_2)_x, R_1SO_2(CH_2)_x, R_1SO_3(CH_2)_x, HNR_1SO_2(CH_2)_x, R_1R_2N(CH_2)_x, R_1O(CH_2)_x, CF_3$ 或 OH ;
- e) $x = 0-6$;
- f) R_1, R_2 和 R_3 各自独立地是氢、 C_1-C_6 直链或支链烷基、或者由一个或多个卤素、羟基、氨基、羧基、羧酰胺、腈、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、磷酸盐、 Ar^1 取代的 C_1-C_6 支链或直链烷基;
- g) R_4 是卤素、 CN, N_3, C_1-C_6 直链或支链烷基、或者由一个或多个卤素、羟基、硝基、烷氧基、三氟甲基、磺酸盐、膦酸盐、磷酸盐、 $Ar^1, -COR_1, -COOR_1, -CONR_1R_2, CN, -NR_1, -NR_1R_2, -SR_1, -SO_2NHCN$ 或者 N_3 取代的 C_1-C_6 支链或直链烷基; 以及
- h) Ar^1 是单环、双环或三环、碳环或杂环, 其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代; 其中单个环的大小是 5-6 个原子; 而且杂环含有 1-6 个选自由氧、氮、硫及其组合组成的组的杂原子。

较优选的以一般式 II 表示的组合物或其药学上可接受的盐类是一般式 IIa 的组合物或其药学上可接受的盐类, 其包括下列结构式:



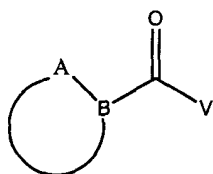
式中,

- a) Y 是 Ar^1 ;
- b) Z 是羰基或羟基;
- c) W 是 $(CH_2)_n$, 其中 ($n = 0, 1$ 或 2), $R_3 = H$; 以及
- d) Ar^1 是单环、双环或三环、碳环或杂环, 其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代; 其中单个环的大小是 5-6 个原子; 而且杂环含有 1-6 个选自由氧、氮、硫及其组合组成的组的杂原子。

以一般式 II 和 IIa 表示的组合物或其药学上可接受的盐类的具体例子包括但不限于以下化合物:

- a) 2-氧代-丙酸;
- b) 5-胍基-2-氧代-戊酸;
- c) 2-氧代-琥珀酰胺酸;
- d) 2-氧代-琥珀酸;
- e) 3-巯基-2-氧代-丙酸;
- f) 3-(1H-咪唑-4-基)-2-氧代-丙酸;
- g) 3-甲基-2-氧代-戊酸;
- h) 氧代-乙酸;
- i) 4-氨基甲酰-2-氧代丁酸;
- j) 2-氧代-戊二酸;
- k) 4-甲基-2-氧代-戊酸;
- l) 6-氨基-2-氧代-己酸;
- m) 4-甲基硫烷基-2-氧代-丁酸;
- n) 2-氧代-3-苯基丙酸;
- o) 3-羟基-2-氧代-丙酸;
- p) 3-羟基-2-氧代-丁酸;
- q) 3-(1H-吡啶-3-基)-2-氧代-丙酸;
- r) 3-(4-羟基-苯基)-2-氧代-丙酸;
- s) 3-甲基-2-氧代-丁酸;
- t) 2-羟基丁酸;
- u) 3-羟基丁酸; 以及
- v) 3-氧代戊二酸。

以一般式 III 表示的组合物或其药学上可接受的盐类以包括下列结构式的结构表示:

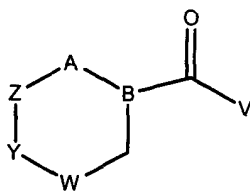


式中,

- a) A 和 B 一起形成一个 5-8 元饱和或部分饱和杂环，其含有至少一个处于化学稳定氧化状态的附加杂原子 O、S、SO、SO₂、NH、或 NR¹；
- b) V 是 O、OR₁、NR₂、NR₁R₂、CHR₁R₂、CH₂R₃、CHR₃R₄ 或 CH₂N₃；
- c) R₁ 和 R₂ 各自独立地是氢、C₁-C₆ 直链或支链烷基、或者由一个或多个卤素、羟基、氨基、羧基、羧酰胺、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、或 Ar¹ 取代的 C₁-C₆ 支链或直链烷基；
- d) R₃ 和 R₄ 或者是卤素、C₁-C₆ 直链或支链烷基、或者由一个或多个羟基、氨基、羧基、羧酰胺、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、Ar¹、-OC(O)R₁、-COOR₁、-CONR₁R₂、CN、NR₁、NR₁R₂、SR₁、SO₂NHCN 或 N₃ 取代的 C₁-C₆ 支链或直链烷基；以及
- e) Ar¹ 是单环、双环或三环、碳环或杂环，其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、C₁-C₆ 直链或支链烷基或链烯基、C₁-C₄ 烷氧基、C₁-C₄ 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代；其中单个环的大小是 5-6 个原子；而且杂环含有 1-6 个选自由氧、氮、硫及其组合组成的组的杂原子。

以一般式 III 表示的组合物的具体例子包括但不限于脘硫醚酮亚胺和环硫羧。

较优选的以一般式 III 表示的组合物或其药学上可接受的盐类是一般式 IV 的组合物或其药学上可接受的盐类，其以包括下列结构式的结构表示：



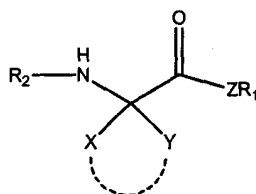
式中，

- a) W-Y-Z-A-B 包括六元饱和或部分饱和碳环或杂环，其中杂环含有选自由氧、氮、硫及其任意组合组成的组的杂原子；
- b) B 是 C、CH 或 N；
- c) A、W、Y、Z 各自独立地是 CH₂、CHR₃、CR₃R₄、O、S、SO、SO₂、NH、NR₁、NR₁R₂ 或 C=O；
- d) V 是 O、OR₁、NR₂、NR₁R₂、CHR₁R₂、CH₂R₃、CHR₃R₃ 或 CH₂N₃；
- e) R₁ 和 R₂ 各自独立地是氢、C₁-C₆ 直链或支链烷基、或者由一个或多个卤素、羟基、氨基、羧基、羧酰胺、腈、硝基、烷氧基、三氟甲基、硫、磺酸盐、膦酸盐、磷酸盐或 Ar¹ 取代的 C₁-C₆ 支链或直链烷基；

- f) R_3 和 R_4 各自独立地是卤素、 $-\text{OC}(\text{O})R_1$ 、 $-\text{COOR}_1$ 、 $-\text{CONR}_1R_2$ 、 CN 、 $-\text{NR}_1$ 、 $-\text{NR}_1R_2$ 、 $-\text{SR}_1$ 、 $-\text{SO}_2\text{NHCN}$ 、 N_3 、 $\text{C}_1\text{-C}_6$ 直链或支链烷基、或者由一个或多个卤素、羟基、硝基、烷氧基、三氟甲基、磺酸盐、膦酸盐、 Ar^1 、 $-\text{OC}(\text{O})R_1$ 、 $-\text{COOR}_1$ 、 $-\text{CONR}_1R_2$ 、 CN 、 $-\text{NR}_1$ 、 $-\text{NR}_1R_2$ 、 $-\text{SR}_1$ 、 $-\text{SO}_2\text{NHCN}$ 或 N_3 取代的 $\text{C}_1\text{-C}_6$ 支链或直链烷基；以及
- g) Ar^1 是单环、双环或三环、碳环或杂环，其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 $\text{C}_1\text{-C}_6$ 直链或支链烷基或链烯基、 $\text{C}_1\text{-C}_4$ 烷氧基、 $\text{C}_1\text{-C}_4$ 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代；其中单个环的大小是 5-6 个原子；而且杂环含有 1-6 个选自氧、氮、硫及其任意组合组成的组的杂原子。

以一般式 IV 表示的组合物或其药学上可接受的盐类的具体例子包括但不限于：氨基乙基半胱氨酸酮亚胺 (2H-1, 4-噻嗪-5, 6-二氢-3-羧酸)、硫代吗啉-2-羧酸、羊毛硫氨酸酮亚胺和 1, 4-硫代吗啉-3, 5-二羧酸。

以一般式 V 表示的组合物或其药学上可接受的盐类以包括下列结构式的结构表示：

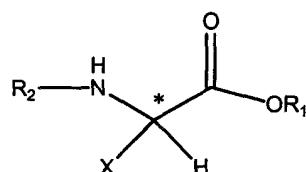


式中，

- Z 是 O 或 NH ；
- R^1 是 $(\text{C}_1\text{-C}_6)$ 烷基、 Ar^1 、或者 $(\text{C}_1\text{-C}_4)$ 烷氧基羰基甲基；
- X 、 Y 各自独立地是氢、 Ar^1 、 $(\text{C}_1\text{-C}_6)$ 烷基 (可由杂原子断开或取代，如氮、磷、氧、硫或硅，杂原子本身可由 $(\text{C}_1\text{-C}_3)$ 烷基取代一次或多次)、 $(\text{C}_2\text{-C}_6)$ 链烯基、 $(\text{C}_1\text{-C}_6)$ 卤代烷基、或卤素。当 X 和 Y 各自都是碳时，它们可共价结合形成一种饱和或部分不饱和独立地由碳、氮、氧和硫组成的 3-8 元碳环化合物，此外，环的原子本身可以是未取代或由卤素、羟基、羧基、硝基、三氟甲基、 $\text{C}_1\text{-C}_6$ 直链或支链烷基或链烯基、 $\text{C}_1\text{-C}_4$ 烷氧基、 $\text{C}_1\text{-C}_4$ 链烯氧基、苯氧基、苄氧基、氨基、取代的烷基、 Ar^1 或其组合取代；
- R_2 是氢、烷基、 Ar^1 或氧取代的烷基；以及
- Ar^1 是单环、双环或三环、碳环或杂环，其中环或者未取代或者 1 至 3 位上

由卤素、羟基、硝基、三氟甲基、 C_1 - C_6 直链或支链烷基或链烯基、 C_1 - C_4 烷氧基、 C_1 - C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代；其中单个环的大小是 3-7 个原子；而且杂环含有 1-6 个选自由氧、氮、硫及其任意组合组成的组的杂原子。

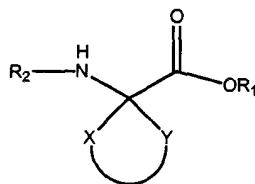
优选的以一般式 V 表示的组合物或其药学上可接受的盐类是以一般式 Va 表示的组合物或其药学上可接受的盐类，其包括下列结构式：



式中，

- a) * = 不对称中心，及
- b) $R^1 = (C_1 - C_6)$ 烷基、 Ar^1 、 $(C_1 - C_4)$ 烷氧羰基甲基及
- c) X 是氢、 $(C_1 - C_6)$ 烷基(可由杂原子断开或取代，如氮、磷、氧、硫或硅，杂原子本身可由 $(C_1 - C_3)$ 烷基取代一次或多次)、 $(C_2 - C_6)$ 链烯基、 $(C_1 - C_6)$ 卤代烷基、卤素或 Ar^1 ；
- d) R_2 是氢、烷基、 Ar^1 或氧取代的烷基；
- e) Ar^1 是单环、双环或三环、碳环或杂环，其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1 - C_6 直链或支链烷基或链烯基、 C_1 - C_4 烷氧基、 C_1 - C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代；其中单个环的大小是 3-7 个原子；而且杂环含有 1-6 个选自由氧、氮、硫及其任意组合组成的组的杂原子。

较优选的以一般式 V 表示的组合物或其药学上可接受的盐类包括以一般式 Vb 表示的组合物或其药学上可接受的盐类，其包括下列结构式：



式中，

- a) X 和 Y 各自是碳；
- b) X 和 Y 以 3-8 个碳原子的饱和或部分饱和环连接，这样一个环本身的 1 至 5

位上可由卤素、羟基、羧基、氨基、硝基、氰基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、或取代的烷基取代；

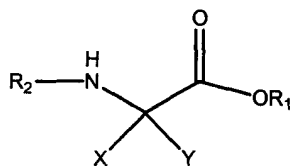
c) R^1 是 (C_1-C_6) 烷基、 Ar^1 或 (C_1-C_4) 烷氧羰基甲基；

d) R_2 是氢、烷基、 Ar^1 或氧取代的烷基；以及

e) Ar^1 是单环、双环或三环、碳环或杂环，其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代；其中单个环的大小是 3-7 个原子；而且杂环含有 1-6 个选自由氧、氮、硫及其任意组合组成的组的杂原子。

较优选的以一般式 Vb 表示的组合物包括连接 X 和 Y 的 3-6 元环。

较优选的以一般式 V 表示的组合物或其药学上可接受的盐类是以一般式 Vc 表示的组合物或其药学上可接受的盐类，其包括下列结构式：



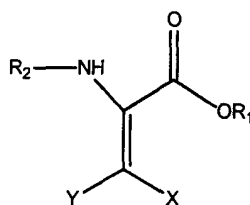
式中，

a) X、Y 各自独立地是氢、 Ar^1 、 (C_1-C_6) 烷基(可由杂原子断开或取代，如氮、磷、氧、硫或硅，杂原子本身可由 (C_1-C_3) 烷基取代一次或多次)、 (C_2-C_6) 链烯基、 (C_1-C_6) 卤代烷基、或卤素如萘基或苯基；

b) R_2 是氢、烷基、 Ar^1 或氧取代的烷基；以及

c) Ar^1 是单环、双环或三环、碳环或杂环，其中环或者未取代或者 1 至 3 位上由卤素、羟基、硝基、三氟甲基、 C_1-C_6 直链或支链烷基或链烯基、 C_1-C_4 烷氧基、 C_1-C_4 链烯氧基、苯氧基、苄氧基、氨基、或其组合取代；其中单个环的大小是 3-7 个原子；而且杂环含有 1-6 个选自由氧、氮、硫及其任意组合组成的组的杂原子。

较优选的以一般式 VI 表示的组合物或其药学上可接受的盐类以包括下列结构式的结构表示：



式中,

- a) R^1 是 (C_1-C_6) 烷基、 Ar^1 或 (C_1-C_4) 烷氧羰基甲基;
- b) R_2 是氢、烷基、 Ar^1 或氧取代的烷基;
- c) Y 是氢、 Ar^1 、 (C_1-C_6) 烷基(可由杂原子断开或取代, 如氮、磷、氧、硫或硅, 杂原子本身可由 (C_1-C_3) 烷基取代一次或多次)、 (C_2-C_6) 链烯基、 (C_1-C_6) 卤代烷基、或卤素; 以及
- d) X 是烷基或苯基。

在较优选的以一般式 VI 表示的组合物中, Y 是 Ar^1 。在另一较优选的以一般式 VI 表示的组合物中, Y 是苯基、萘基、3-甲酰吡啶、咪唑或吡唑。

本发明的其它化合物包括但不限于: (不可逆或自杀型抑制剂): 2-氧代-戊炔酸盐、丹酰氯、丹酰氟、炔丙基甘氨酸、N-乙酰基-炔丙基甘氨酸、O-(2, 4-二硝基苯基)羟胺、1-氯-2-硝基乙烷、1, 2-环己二酮、烯丙基甘氨酸、N-氯-D-亮氨酸、苯基乙二醛、溴丙酮酸乙酯、溴丙酮酸甲酯和溴丙酮酸盐。较优选的化合物包括但不限于: 甲基乙二醛双(脒基脒)、胍基羧亚氨基酰胺、丙酮醛双(脒基脒)、3-(3-吡啶基)丙酸、3-吡啶-乙酸、吡啶-3-丙酮、吡啶-3-乙酰胺、吡啶-3-乙酰基-L-天冬氨酸、吡啶-3-乙酰基-L-丙氨酸、吡啶-3-乙酰基甘氨酸、吡啶-3-羧酸、吡啶-3-丙酮酸、丹酰基甘氨酸、丙氨酸四唑、四唑、核黄素 5'-焦磷酸盐、5-羟基-2-羟甲基-4-吡喃酮、羟胺盐酸盐、四氢-4-苯基-4H-1, 4-噻嗪 1-氧化物、吩噻嗪、3, 4-二氢-2H-1, 4-噻嗪-3, 5-二羧酸、硝呋莫司(1-((5-硝基亚糠基)氨基)-2-甲基四氢-1, 4-噻嗪-4, 4-二氧化物)、2-氨基-2, 4-戊二烯酸盐、2-氨基-4-酮-2-戊烯酸盐、N-乙酰基-D-亮氨酸、D-Leu (D-2-氨基-4-甲基戊酸)、孕酮(4-孕烯-3, 20-二酮)、FAD(黄素腺嘌呤二核苷酸)、6-OH-FAD(黄素腺嘌呤二核苷酸)、N-(1-羧乙基)-L-丙氨酸、3-氨基天冬氨酸、硫代氨甲酰酰胍、5-S-半胱氨酰多巴胺、磷脂酰丝氨酸、4-羟基-2-巯基 6-甲基嘧啶、4-氨基-N-2-噻唑基苯磺酰胺、硫氰酸盐、2-巯基-1-甲基咪唑、酒石酸、2-氨基乙硫醇、S-腺苷甲硫氨酸和硫脲。

可用本领域已知的方法对本文或表 I 所列的没有限制作用的化合物作出变换或衍生, 生成下列一种或多种化合物: a) 前药; b) 具有高酶活性的化合物; c) 对 D-氨基酸氧化酶或 D-天冬氨酸氧化酶具有特异性的化合物; d) 毒性低的化合物; 或者 e) 没有不希望的副作用的化合物。测定 D-氨基酸氧化酶或 D-天冬氨酸氧化酶活性的方法已为本领域所公知, 可用本文所公开或本文引用的参考文献所公开的方法进行。以下所有引用作举例说明 D-氨基酸氧化酶或 D-天冬氨酸氧化酶拮抗剂的

参考文献均纳入其整体内容作为参考。

用作举例说明的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶拮抗剂

- a) 2-氧代-3-戊炔酸盐; (Biochemistry 1999 年 5 月 4 日; 38 (18): 5822-8)。
- b) 氨基胍; (J Neurochem 1998 年 3 月; 70(3): 1323-6)。
- c) 苯甲酸盐(苯甲酸)及其盐(例如, 苯甲酸钠); (Neurosci Lett 1996 年 11 月 8 日; 218(3): 145-8)。
- d) 邻、间和对氨基苯甲酸盐类(J. Biochem. (东京) 1976 年 11 月, 80(5): 1101-1108)。
- e) 甲基乙二醛双(脒基脲)(MGBG); 苯基乙二醛双(脒基脲)(PhGBG); 乙二醛双(脒基脲)(GBG); (Anticancer Drug Des 1996 年 10 月; 11(7): 493-508)。
- f) α -酮酸, 所述 α -酮酸是氨基酸丙氨酸、缬氨酸、亮氨酸、苯丙氨酸、苯甘氨酸, 酪氨酸、色氨酸、丝氨酸、天冬氨酸盐等的类似物(例如, 丙酮酸、 α -酮异戊酸、4-甲基硫代-2-氧代戊酸、4-甲基硫代-2-氧代丁酸、苯基丙酮酸、吡啶-3-丙酮酸、苯甲酰基甲酸、4-羟基苯基丙酮酸及其盐类和衍生物), 吡啶-丙酸、3-吡啶-乙酸、水杨酸及其盐类和衍生物 (Enzyme Microb Technol 1996 年 4 月; 18(5): 379-82)。
- g) 丹酰氯; (FEBS Lett 1995 年 4 月 25 日; 363(3): 307-10)。
- h) 丙氨酸四唑和苯甲酸四唑; (Res Commun Chem Pathol Pharmacol 1994 年 2 月; 83(2): 209-22)。
- i) 核黄素 5'-焦磷酸盐(RPP); (Anal Biochem 1992 年 5 月 1 日; 202(2): 348-55)。
- j) D-炔丙基甘氨酸(D-PG); (J Biochem (Tokyo) 1991 年 1 月; 109(1): 171-7)。
- k) D, L- β -羟基丁酸盐; (J Histochem Cytochem 1991 年 1 月; 39(1): 81-6)。
- l) 葫芦巴碱, 即 N-烟酸甲酯; (J Biochem (东京) 1990 年 5 月; 107(5): 726-31)。
- m) 曲酸及其盐类; (J Biol Chem 1989 年 2 月 15 日; 264(5): 2509-17)。
- n) O-(2, 4-二硝基苯基)羟胺; (Biochemistry 1987 年 3 月 24 日; 26(6): 1717-22)。
- o) 苯甲酸盐; (D'Silva, C., Willams, C.H., Jr., & Massey, V. (1986)

- Biochemistry 25, 5602-5608)。
- p) 对硝基苯磺酸甲酯; (J Biol Chem 1984 年 5 月 10 日; 259(9): 5585-90)。
- q) 氨基乙基半胱氨酸-酮亚胺(2H-1, 4-噻嗪-5, 6-二氢-3-羧酸); 1, 4-噻嗪衍生物, 酮亚胺还原形式(硫代吗啉-2-羧酸和硫代吗啉-2, 6-二羧酸); (Biochim Biophys Acta 1983 年 10 月 17 日; 748(1): 40-7)。
- r) 半胱氨酸和溴丙酮酸盐的反应产物; (J Appl Biochem 1983 年 8-10 月; 5(4-5): 320-9)。
- s) 1-氯-1-硝基乙烷; (J Biol Chem 1983 年 1 月 25 日; 258(2): 1136-41)。
- t) 苯甲酸盐; 氨基苯甲酸; 吡啶甲酸盐; L-亮氨酸; (J Biol Chem 1982 年 9 月 10 日; 257(17): 9958-62)。
- u) 氟二硝基苯; (Nishino, T., Massey, V., 和 Williams, C.H., Jr. (1980) J Biol Chem 255, 3610-3616)。
- v) 1, 2-环己烷二酮; (Eur J Biochem 1981 年 10 月; 119(3): 553-7)。
- w) 烯丙基甘氨酸; 2-氨基-2, 4-戊二烯酸盐; 2-羟基-2, 4-戊二烯酸盐; (Biochemistry 1978 年 12 月 26 日; 17(26): 5620-6)。
- x) 2-氨基-4-酮-2-戊烯酸盐; (Biochemistry 1978 年 12 月 26 日; 17(26): 5613-9)。
- y) D, L-2-羟基丁酸盐; (J Cell Biol 1978 年 4 月; 77(1): 59-71)。
- z) 对氨基苯甲酸盐; (J Biochem (Tokyo) 1976 年 11 月; 80(5): 1073-83)。
- aa) N-氯-D-亮氨酸; (J Biol Chem 1976 年 10 月 10 日; 251(19): 6150-3)。
- bb) D-炔丙基甘氨酸; (Biochemistry 1976 年 7 月 13 日; 15(14): 3070-6)。
- cc) D-2-氨基-4-戊炔酸 (D-炔丙基甘氨酸); (J Biochem (Tokyo) 1975 年 7 月; 78(1): 57-63)。
- dd) 孕酮(Biochim Biophys Acta 1978 年 1 月 12 日; 522(1): 43-8)。
- ee) 长链、中链和短链游离脂肪酸(Biochem Int 1990 年 12 月; 22(5): 837-42)。
- ff) 6-OH-FAD (Biochim Biophys Acta 1999 年 4 月 12 日; 1431(1): 212-22)。
- gg) 苯基乙二醛, L-酒石酸盐(Eur J Biochem 1992 年 4 月 1 日; 205(1): 127-32)。
- hh) 环硫基, TMDA, α - α' -亚氨基二丙酸(Physiol Chem Phys Med NMR 1986;

18(1): 71-4).

酸氧化酶的活性水平较标准低，显示候选化合物是 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的拮抗剂。D-氨基酸氧化酶或 D-天冬氨酸氧化酶的活性可在一种分离的或纯化的酶或在含有表达 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的细胞或组织的生物样本中测定。

此外，本领域技术人员能够鉴定抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶基因产物(mRNA 或多肽)的表达的化合物。使表达 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的细胞(例如，肝脏、肾或脑细胞)在有或没有测试化合物的情况下培育。通过测定有或没有测试化合物时 D-氨基酸氧化酶或 D-天冬氨酸氧化酶基因产物的表达水平或者在有或没有测试化合物时 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的活性水平，可鉴定抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶基因产物的表达的化合物。另外，可将含有可操作地连接于报导基因的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶调节序列的构造体(例如，萤光素、氯霉素乙酰转移酶、LacZ、绿萤光多肽、 β 牛乳糖等)引入宿主细胞，并检测测试化合物对报导基因表达的影响。适用于上述分析的细胞包括但不限于：与组织同来源的细胞或多肽已知得到表达的细胞系(例如，肾、肝脏和脑)。在 mRNA 水平上(例如采用 RNA 印迹、RT-PCR，优选对感兴趣的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶 mRNA 具有特异性的引物和探针进行定量 RT-PCR)或者在多肽水平上(通过测定 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的酶活性或在免疫分析如 ELISA 或放射性免疫分析、蛋白质印迹、免疫化学法中用多克隆或单克隆抗体)，可实现定量分析 D-氨基酸氧化酶或 D-天冬氨酸氧化酶多肽的表达。

另一方面，有一种分析基于细胞的分析，其中使表达 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分与测试化合物接触，确定测试化合物抑制、活化或增加 D-氨基酸氧化酶或 D-天冬氨酸氧化酶活性的能力。通过监测 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分的生物活性可以确定测试化合物抑制、活化或增加 D-氨基酸氧化酶或 D-天冬氨酸氧化酶活性的能力。最好监测氨基酸氧化。细胞可以源自例如哺乳动物、细菌或酵母菌细胞。例如，一些实施例的细胞可以是哺乳动物、细菌或酵母菌细胞。

采用本领域已知的组合库方法中众多方法的其中一种可得到本发明的测试化合物，这些方法包括：生物库、空间可寻址平行固相或液相文库、需要重叠合法(deconvolution)的合成化学库方法、“一珠粒一化合物(one-bead one-compound)”文库方法；以及应用亲和力层析选择的合成文库方法。生物库方法与肽库一起使用，而其它四种方法适用于肽、非肽低聚物或小分子化合物库(Lam, K. S.

(1997) *Anticancer Drug Des.* 12: 145, 它所公开的资料纳入其整体内容作为参考)。

在本领域可找到分子库的合成方法的例子, 例如, DeWitt 等人, (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90: 6909; Erb 等人, (1994) *Proc. Natl. Acad. Sci. USA* 91: 11422; Zuckermann 等人, (1994) *J. Med. Chem.* 37: 2678; Cho 等人, (1993) *Science* 261: 1303; Carrell 等人, (1994) *Angew. Chem. Int. Ed. Engl.* 33: 2059; Carrell 等人, (1994) *Angew. Chem. Int. Ed. Engl.* 33: 2061; 以及 Gallop 等人, (1994) *J. Med. Chem.* 37: 1233, 这些文献所公开的资料纳入其整体内容作为参考)。

溶液(例如, Houghten (1992) *Biotechniques* 13:412-421)、或珠粒(Lam (1991) *Nature* 354: 82-84)、芯片(Fodor (1993) *Nature* 364: 555-556)、细菌(Ladner U.S. Pat. No. 5, 223, 409)、孢子(Ladner 的美国专利号'409)、质粒(Cull 等人, (1992) *Proc Natl Acad Sci USA* 89: 1865-1869)或噬菌体(Scott 和 Smith (1990) *Science* 249: 386-390); (Devin (1990) *Science* 249: 404-406); (Cwirla 等人(1990) *Proc. Natl. Acad. Sci.* 87: 6378-6382); (Felici (1991) *J. Mol. Biol.* 222: 301-310); (上述提到的 Ladner)可含有化合物库, 这些文献所公开的资料纳入其整体内容作为参考。

例如, 将 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分与放射性同位素或酶标记偶联, 这样, 检测复合物中标记的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分可测定 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分与其同族靶分子的结合, 从而可确定测试化合物抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶活性的能力。D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子优选是一种与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质天然结合或相互作用的分子, 以致可实现 D-氨基酸氧化酶或 D-天冬氨酸氧化酶介导的功能。在一实施例中, D-氨基酸氧化酶靶分子是 g34872 多肽。例如, 化合物(如 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分)可用 ^{125}I 、 ^{35}S 、 ^{14}C 或 ^3H 直接或间接标记, 通过直接计算放射性放射或计算闪烁来检测放射性同位素。另一方面, 化合物也可用例如辣根过氧化酶、碱磷酸酶或萤光素作酶标记, 通过测定适当底物向产物的转化来检测酶标记。使标记的分子与其同族分子接触, 并测定复合物形成的程度。例如, 免疫沉淀复合物或进行凝胶电泳可测定复合物形成的程度。

无需标记任何一种相互作用的物质而确定化合物(例如 D-氨基酸氧化酶或 D-

天冬氨酸氧化酶蛋白质或其生物活性部分)与其同族靶分子相互作用的能力也在本发明的范围内。例如,可用微生理仪检测一种化合物与其同族靶分子的相互作用,而无需标记该化合物或靶分子(McConnell, H. M. 等人(1992) Science 257: 1906-1912, 其公开的资料纳入整体内容作为参考)。如细胞传感器等微生理仪是一种用光可寻址的电位器式传感器(LAPS)测量细胞酸化其环境速度的分析仪器。酸化速度的变化可作为化合物与受体之间相互作用的指标。

在一优选实施例中,该分析包括使表达或应答 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分的细胞与靶分子接触,形成分析混合物,使该分析混合物与测试化合物接触,确定测试化合物抑制或增加 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分的活性的能力,其中确定测试化合物抑制或增加 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分的活性的能力包括确定测试化合物抑制或增加表达细胞的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的生物活性的能力(例如,确定测试化合物抑制或增加转导、蛋白质与蛋白质相互作用、底物结合的能力)。

在另一实施例中,有一种分析基于细胞的分析,其包括使表达 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子(即与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶相互作用的分子)的细胞与测试化合物接触,确定测试化合物调节(例如分别刺激或抑制)D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子的活性的能力。例如,通过确定 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子结合或相互作用的能力可确定测试化合物调节 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子的活性的能力。

确定 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子结合或相互作用的能力可用上述测定直接结合的其中一种方法来实现。在一优选实施例中,测定靶分子活性可确定 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子结合或相互作用的能力。例如,使靶分子与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其片段接触,测定靶分子第二细胞信使(即细胞外 Ca^{2+} 、甘油二酯、IP3 等)的诱导,检测靶分子在一适当底物的催化/酶活性,检测报导基因(含有靶-反应调节元件,所述元件可操作地连接于编码可检测标记如萤光素的核酸)的诱导,或者检测靶-调节的细胞应答,例如信号转导或蛋白质与蛋白质相互作用,可以测定靶分子的活性。

在另一优选实施例中,本发明的一种分析是与细胞无关的分析,该分析使 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分与测试化合物接触,

确定该测试化合物与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分结合的能力。按照上述方法可直接或间接测定测试化合物与蛋白质的结合。在一优选实际例中, 该分析包括使 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分与已知的化合物接触, 该已知的化合物与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶(例如 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子)结合形成分析混合物, 使该分析混合物与测试化合物接触, 并确定该测试化合物与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质相互作用的能力, 其中确定测试化合物与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质相互作用的能力包括测定该测试化合物与已知的化合物相比, 与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶或其生物活性部分优先结合的能力。

在另一实施例中, 该分析是与细胞无关的分析, 其中使 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分与测试化合物接触, 确定该测试化合物调节(例如抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的活性或者激活 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的活性)该蛋白质或其生物活性片段的活性。例如, 通过上述测定直接结合的其中一种方法确定蛋白质与靶分子结合的能力可确定测试化合物调节蛋白质的活性的能力。例如, 这可用一种实时生物分子相互作用分析(BIA)等技术来实现(Sjolander, S. 和 Urbaniczky, C. (1991) *Anal. Chem.* 63: 2338-2345 以及 Szabo 等人, (1995) *Curr. Opin. Struct. Biol.* 5: 699-705, 这些文献所公开的资料纳入其整体内容作为参考)。本文所用的“BIA”是一种实时研究生物特异性相互作用的技术, 无需标记任何一种相互作用的物质(例如 BIAcore)。表面等离子共振仪光学现象的变化可作为生物分子之间实时作用的指标。

在一变换实施例中, 通过确定 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质进一步调节 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子下游效应子的活性的能力可确定测试化合物调节 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质的活性的能力。例如, 利用如前所述方法可测定效应子分子对合适的靶分子的活性或者可测定效应子与合适的靶分子的结合。

在另一实施例中, 该与细胞无关的分析包括使 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分与已知的化合物接触, 该已知的化合物与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶结合形成分析混合物, 使该分析混合物与测试化合物接触, 并确定该测试化合物与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质相互作用的能力, 其中确定该测试化合物与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质相互作用的能力包括确定 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质优先结合

或调节 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子的活性的能力。

本发明与细胞无关的分析可改为采用可溶性和/或膜结合形式的分离蛋白质(例如 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分或者与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子结合的分子)。在与细胞无关的分析中,若采用膜结合形式的分离蛋白质,就需要用增溶剂以使膜结合形式的分离蛋白质保持在溶液中。这些增溶剂的例子包括非离子去污剂,如正辛基葡糖苷、正十二烷基葡糖苷、正十二烷基麦芽糖苷、辛酰基-N-甲基葡糖酰胺、癸酰基-N-甲基葡糖酰胺、TritonTM X-100、TritonTM X-114、ThesitTM、异三癸基聚(亚乙基乙二醇醚)正 3-(3-胆酰氨基丙基)二甲基铵基-1-丙烷磺酸盐(CHAPS)、3-(3-胆酰氨基丙基)二甲基铵基-2-羟基-1-丙烷磺酸盐(CHAPS0)或者 N-十二烷基-N,N-二甲基-3-铵基-1-丙烷磺酸盐。

在一个以上本发明的上述分析方法实施例中,需要将 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或靶分子固定,以利于一个或多个蛋白质复合型式与非复合型式分离,以及适应分析的自动化操作。在任何适合容纳反应物的容器中,在候选化合物存在或不存在的条件下可完成测试化合物与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质的结合,或者 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质与靶分子的相互作用。这些容器的例子包括微滴定板、试管和微离心管。一实施例提供了一种融合蛋白,所述融合蛋白加入了一结构域,可以使一种或两种蛋白质与基质结合。例如,谷胱甘肽-S-转移酶/D-氨基酸氧化酶或 D-天冬氨酸氧化酶融合蛋白或谷胱甘肽-S-转移酶/靶分子融合蛋白可被吸入谷胱甘肽琼脂糖珠粒(Sigma Chemical, 密苏里州圣路易市),或者谷胱甘肽衍生化微滴定板,它们在有利复合物形成的条件(例如,盐和 pH 值的生理条件)下再与测试化合物或者没有被吸入的靶蛋白质或者 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质结合。培育后,洗涤珠粒或微滴定板,例如,如上所述,除去任何非结合组分、固定在珠粒的基质、直接或间接测定的复合物。此外,也可使复合物与基质解离,用标准技术测定 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的结合或活性水平。

将蛋白质固定在基质的其它技术也可用在本发明的筛选分析方法中。例如,用生物素和链霉生物素结合可将 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或者 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子固定。生物素化的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或靶分子可用本领域公知的技术(例如,生物素化试剂盒, Pierce Chemicals, Rockford, Ill.)由生物素-NHS(N-羟基-琥珀酰亚胺)制备,并固定在涂有链霉生物素的 96 孔板(Pierce Chemical)的孔中。另外,与 D-氨基酸

氧化酶或 D-天冬氨酸氧化酶蛋白质或靶分子反应但不会干扰 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质与其靶分子结合的抗体可衍生到孔板的孔中,而未结合的靶分子或 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质通过与抗体结合而被收集在孔中。这些复合物的检测方法除了以上所述关于 GST 固定化复合物的方法外,还包括复合物的免疫检测方法,其采用与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或靶分子反应的抗体,以及依靠检测与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或靶分子相关的酶活性的酶联分析法。

在本发明的另一方面,可用蛋白质作为双杂交或三杂交分析中的“诱饵蛋白质”(参见美国专利申请号 5,283,317; Zervos 等人, (1993) *Cell* 72: 223-232; Madura 等人, (1993) *J. Biol. Chem.* 268: 12046-12054; Bartel 等人, (1993) *Biotechniques* 14: 920-924; Iwabuchi 等人, (1993) *Oncogene* 8: 1693-1696; 以及 Brent 的 W094/10300, 这些文献所公开的资料纳入其整体内容作为参考),以便鉴定与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质结合或相互作用,和/或涉及 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质的活性的其它蛋白质。

双杂交系统基于大部分由分离的 DNA 结合和活化结构域组成的转录因子的模块性质。简单地说,这种分析利用两个不同的 DNA 构造体。在一个构造体中,编码 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其片段的基因与编码已知的转录因子(如 GAL-4)的 DNA 结合结构域的基因融合。在另一个构造体中,来自 DNA 序列库并编码未得到鉴定的蛋白质(“捕获物”或“样本”)的 DNA 序列与编码已知的转录因子的活化结构域的基因融合。如果“诱饵”和“捕获物”蛋白质能够在体内相互作用,形成依赖于 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的复合物,就使转录因子的 DNA 结合和活化结构域紧密接近。这种接近使得可操作地连接于应答转录因子的转录调节位点的报导基因(如 LacZ)转录。可检测到报导基因的表达,并且可以使含有功能性转录因子的细胞克隆分离,用来获得克隆基因,后者编码与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质相互作用的蛋白质。

本发明还涉及由上述筛选分析法所鉴定的新型试剂,以及用这些分析法生产所述试剂的方法。所以,本发明的一实施例包括由这样一种方法得到的化合物或试剂,该方法包括上述筛选分析(例如基于细胞的分析与细胞无关的分析)中任何一种方法的步骤。例如,本发明一实施例包括一种化合物或试剂,它是通过一种方法,即将表达 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子的细胞与测试化合物接触,并确定该测试化合物结合 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子或者调节 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子的活性的能力来而得到的。本发明的另

一个实施例包括一种化合物或试剂，它是通过另一种方法，即将表达 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子的细胞与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分接触，形成一种分析混合物，再将该分析混合物与测试化合物接触，并确定该测试化合物与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子相互作用或者调节 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子的活性的能力而得到的。本发明的另一个实施例包括一种化合物或试剂，它是通过另一种方法，即将 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分与测试化合物接触，并确定该测试化合物与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶分子结合或者抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分的活性的能力而得到的。本发明的又一个实施例包括一种化合物或试剂，它是通过另一种方法，即将 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质或其生物活性部分与已知的化合物接触，而所述已知的化合物与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质接触形成一种分析混合物，再将该分析混合物与测试化合物接触，并确定该测试化合物与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质相互作用或者调节 D-氨基酸氧化酶或 D-天冬氨酸氧化酶蛋白质的活性的能力而得到的。

采用体内筛选也可鉴定抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的活性或者抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶基因产物的表达的拮抗剂化合物。在这些分析中，给动物施加(如静脉、腹膜内、肌内、口服或其它途径)例如各种剂量水平的测试化合物。分别比较测试动物和表达 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的对照动物组织中 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的活性或基因产物的表达水平，以确定测试化合物对 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的活性或基因产物的表达的影响。合适的动物包括但不限于：啮齿动物(如小鼠和大鼠)和灵长类动物。人源化的非人动物如人源化的小鼠也可用作测试动物，这就是说，这些动物中内源性多肽被切除(剔除)并利用标准转基因方法加回同源性人多肽。这些动物只表达人型式的多肽。

体内分析还包括中枢神经系统障碍的动物模型。这些模型包括但不限于：如本文所述和本领域所公知的，大鼠模型中的条件性回避行为、沙鼠足踢踏模型、雪貂呕吐模型、分离诱导的发声模型、在 omnitech digiscan 动物活动监测器中小鼠和大鼠的行为活动评估、大鼠模型中安他非明刺激的运动的阻断、大鼠模型中声音惊吓的前冲动抑制、脱水吗啡诱导的攀爬行为模型的抑制、以及 DOI 诱导的头部颤搐和搔抓模型。

本发明的其它拮抗剂包括抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶基因产物

的表达的反义和三螺旋工具。在反义方法中,使与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶 mRNA 或基因组序列互补的核酸序列与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶 mRNA 或基因组 DNA 在细胞外杂交,从而阻断由 mRNA 编码的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶多肽的表达。用于 D-氨基酸氧化酶或 D-天冬氨酸氧化酶治疗的反义核酸分子可以是 DNA 或 RNA 序列。采用本发明的反义多核苷酸的优选方法是 Sczakiel 等人(1995)所述的步骤,它公开的资料纳入其整体内容作为参考。本发明其它优选的反义多核苷酸是或者与含有启动密码子 ATG 的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶 mRNA 序列或者与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶序列互补的序列。

反义多核苷酸最好含有一些序列,其与含剪接供体或受体位点的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶启动密码子(ATG)或基因组 DNA 互补。本发明的反义多核苷酸也最好具有被自切割核酶序列所置换的 3'聚腺苷酸化信号,以致在其 3'端无需聚腺苷酸化信号就可以产生 RNA 聚合酶 II 转录物,这些反义多核苷酸不能由胞核输出,如 Liu 等人(1994)所述,它公开的资料纳入其整体内容作为参考。D-氨基酸氧化酶或 D-天冬氨酸氧化酶的反义多核苷酸还可在核酶盒中含有一种组蛋白茎—环结构,以稳定切割转录物抗 3'-5'外切核酸酶降解,这样一种结构如 Eckner 等人(1991)所述,它公开的资料纳入其整体内容作为参考。

反义核酸的长度和解链温度要足以形成胞内双链体,其稳定性足以抑制双链体中 D-氨基酸氧化酶或 D-天冬氨酸氧化酶 mRNA 的表达。Green 等人(1986 年)以及 Izant 和 Weintraub(1984 年)揭示了设计适用于 D-氨基酸氧化酶或 D-天冬氨酸氧化酶治疗的反义核酸的策略,这些文献所公开的资料纳入其整体内容作为参考。

在一些策略中,反义分子可通过使 D-氨基酸氧化酶或 D-天冬氨酸氧化酶编码区的方向与启动子的方向相反来获得,以便转录在细胞中正常转录的反向链。另一种方法包括通过使含有反义序列的 DNA 与在适当表达载体的启动子可操作地连接,在体内转录 D-氨基酸氧化酶或 D-天冬氨酸氧化酶反义核酸。

另外,与在细胞中正常转录的链互补的寡核苷酸可在体外合成。所以,这些反义核酸与相应的 mRNA 互补,能够与 mRNA 杂交生成双链体。反义序列也可含有修饰的磷酸糖骨架以增加稳定性,并且使它们对 RNA 酶活性的灵敏度减少。适用于反义策略的修饰包括 2'-O-甲基 RNA 寡核苷酸和多肽—核酸(PNA)寡核苷酸。其它例子参见 Rossi 等人(1991 年)所述,它公开的资料纳入其整体内容作为参考。

也可采用与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶 cDNA 或基因组 DNA 互补的各种形式反义寡核苷酸。例如,可采用本文引用的国际申请 WO 94/23026 所述的稳

定和半稳定反义寡核苷酸。这些分子的 3'端或 3'和 5'端在互补碱基对之间的分子内氢键结合。这些分子能够更好地抵抗外切核酸酶攻击，并且具有较常规的反义寡核苷酸更高的稳定性。

在另一种使用反义技术抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶多肽的表达的方法中，采用了本文引用的国际申请 W0 96/31523 所述的共价交联反义寡核苷酸。这些双链或单链寡核苷酸分别含有一个或多个寡核苷酸间或寡核苷酸内的共价交联，其中这些交联由一条链的伯胺基团分别与另一条链或同一条链的羧基之间的酰胺键组成，伯胺基团在链核苷酸单糖环的位置 2'上直接取代，而羧基则分别由另一条链或同一条链的核苷酸或核苷酸类似物上取代的脂肪族分隔基团携带。

还可采用本文引用的国际申请 W0 92/18522 所公开的反义寡脱氧核苷酸和寡核苷酸。这些分子不易降解，含有至少一段与调控多肽结合的转录调控识别序列，作为它们的诱饵(decoy)十分有效。这些分子可含有“发夹”结构、“哑铃形”结构、“修饰的哑铃形”结构、“交联”诱饵结构及“环形”结构。

另外，还可采用本文引用的欧洲专利申请号 0 572 287 A2 所述的环形双链寡核苷酸。这些连接寡核苷酸“哑铃形”含有转录因子的结合位点，并通过遮蔽转录因子受到该因子的调控来抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的表达。

本文引用的国际申请 W0 92/19732 所公开的封闭型反义寡核苷酸也是一种变换。因为这些分子没有游离端，所以它们比常规寡核苷酸更能够抵抗外切核酸酶降解。这些寡核苷酸可以具有多种功能，与不与靶 mRNA 邻接的几个区相互作用。

应用体外表达分析可确定抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶表达所需的反义核酸的适当水平。用本领域已知的方法通过扩散、注射、感染或转染可将这些反义分子引入细胞中。例如，反义核酸可以裸露寡核苷酸、质粒中封装的寡核苷酸、用病毒多肽封装的寡核苷酸序列、或以可操作地连接于表达载体所含的启动子的寡核苷酸引入到身体中。表达载体可以是本领域已知的各种表达载体中任何一种，包括逆转录病毒或病毒性载体、能够进行染色体外复制的载体、或整合载体。这些载体可以是 DNA 或 RNA。

细胞样本可引入许多不同的浓度反义分子，优选为 $1 \times 10^{-10} \text{M}$ 与 $1 \times 10^{-4} \text{M}$ 之间。一旦确定了能足以调控 D-氨基酸氧化酶或 D-天冬氨酸氧化酶表达的最小浓度，就可以将最佳剂量转为适合体内服用的剂量。例如，培养基的抑制浓度为 $1 \times 10^{-7} \text{M}$ ，转化为剂量约等于 0.6 毫克/公斤体重。经过测试实验室动物的寡核苷酸毒性后，发现寡核苷酸水平接近 100 毫克/公斤体重或以上是可以的。此外，还要考虑除去脊椎动物的细胞，用反义寡核苷酸治疗，并重新引入到脊椎动物中。

在本发明的一优选实施例中, 首先鉴定由 D-氨基酸氧化酶或 D-天冬氨酸氧化酶编码的多肽, 或者测定酶活性, 以便可以用某些技术监测反义抑制对翻译的效果, 这些技术包括但不限于: 抗体介导的测试, 如放射性免疫分析、ELISA、功能性分析、放射性标记以及测定 D-氨基酸氧化酶或 D-天冬氨酸氧化酶活性的分析。

本发明所用抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶基因产物的表达的反义技术的一种变换包括使用核酶, 所述核酶通过其互补多核苷酸尾部与 D-氨基酸氧化酶或 D-天冬氨酸氧化酶靶序列结合, 而且通过水解其靶位点(即“锤头核酶”(hammerhead ribozymes))切割相应的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶 RNA。简单地说, 锤头核酶的简化循环包括(1)通过互补反义序列使序列特异性结合靶 D-氨基酸氧化酶或 D-天冬氨酸氧化酶 RNA; (2)使靶 D-氨基酸氧化酶或 D-天冬氨酸氧化酶链的可切割基序进行特异性位点水解; 以及(3)释放切割产物, 产生另一个催化循环。锤头核酶的构造和产生已为本领域所公知, Haseloff 和 Gerlach(Nature 20334: 585-591 (1988))作了较完整的叙述。事实上, 适宜使用长链反义多核苷酸(长至少 30 个碱基)或带有长反义臂的核酶。将这些反义核酶与亲脂基团通过共价键结合或用脂质体作为适当载体, 得到反义核酶的优选输送系统。按 Rossi 等人(1991)和 Sczakiel 等人(1995)所述的方法制备本发明的优选反义核酶, 具体制备步骤可参照本文引用的所述文章。

也可用 D-氨基酸氧化酶或 D-天冬氨酸氧化酶基因组 DNA 抑制基于胞内三螺旋形成的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶基因表达。用三螺旋寡核苷酸抑制基因组的转录。它们特别适用于研究细胞中 D-氨基酸氧化酶或 D-天冬氨酸氧化酶活性的变化。D-氨基酸氧化酶或 D-天冬氨酸氧化酶 cDNA 或基因组 DNA 或那些序列的片段可用来抑制患有与特定 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的表达相关的中枢神经系统障碍的个体的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶表达。类似地, 部分 D-氨基酸氧化酶或 D-天冬氨酸氧化酶基因组 DNA 可用来研究细胞内 D-氨基酸氧化酶或 D-天冬氨酸氧化酶转录的抑制效应。一般认为, 同源嘌呤序列在三螺旋策略中是最有用的。然而, 同源嘧啶序列也可用来抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶表达。这些同源嘧啶寡核苷酸与同源嘌呤的主沟(major groove)即同源嘧啶序列结合。

为了用三螺旋方法进行 D-氨基酸氧化酶或 D-天冬氨酸氧化酶治疗策略, 首先扫描 D-氨基酸氧化酶或 D-天冬氨酸氧化酶基因组 DNA 序列, 以鉴定一段从 10-mer 到 20-mer 的同源嘧啶或同源嘌呤延伸序列, 该序列可用在抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶表达并基于三螺旋的策略中。鉴定候选同源嘧啶或同源嘌呤延伸

序列后, 将含有候选序列的各种含量的寡核苷酸引入表达 D-氨基酸氧化酶或 D-天冬氨酸氧化酶的组织培养细胞中, 以评估它们抑制 D-氨基酸氧化酶或 D-天冬氨酸氧化酶表达的效率。按照上述方法监测经过处理的细胞中 D-氨基酸氧化酶或 D-天冬氨酸氧化酶酶活性的改变或 D-氨基酸氧化酶或 D-天冬氨酸氧化酶表达的减少。

如关于反义多核苷酸中所述, 利用所述的技术以基于体外结果计算出的剂量将能有效抑制组织培养细胞中 D-氨基酸氧化酶或 D-天冬氨酸氧化酶表达的寡核苷酸导入体内。

在一些实施例中, 用 α 端基异构体取代寡核苷酸单位的天然(β)端基异构体, 产生对核酶具有更大抗性的寡核苷酸。此外, 可使溴乙啶等嵌入剂附在 α 寡核苷酸 3'端上, 以稳定三螺旋体。要了解适合形成三螺旋的寡核苷酸的制备方法, 可参看本文的引用文献 Griffin 等人(Science 245: 967-71, 1989)。

药学和生理学上可接受的组合物及其给药

利用容易获得的原料, 以本领域技术人员公知的常规合成方法可制备本发明所用的化合物和组合物。另外, 本发明实验所用的化合物可购自市场供应商, 如 Sigma Chemical Company (密苏里州圣路易市)。

根据本文所述或本领域已知的 Nyberg 等人的方法(Psychopharmacology 119, 345-348, 1995), 对动物进行药理研究可确定 D-氨基酸氧化酶或 D-天冬氨酸氧化酶拮抗剂的相对活性、功效和特异性。这种测试能估计相对活性、功效, 并通过测量特异性可估计治疗指数。可采用的其它动物研究包括但不限于: 本文所述和本领域已知的涉及条件性回避、脱水吗啡诱导的攀爬、阻断 5-羟基-色氨酸诱导的头部颤搐和搔抓或其它动物模型的研究。虽然对人的临床研究可确定不同病人群体之间相差的新陈代谢, 但 Kerr 等人(Biochem. Pharmacol. 47, 1969-1979 (1994))) 和 Karam 等人(Drug Metab. Dispos. 24, 1081-1087, 1996)提供了一些更低廉和省时的替代方法。类似地, 按照 Leach 等人(Epilepsia 37, 1100-1106, 1996)的方法或按照 Kerr 等人(同上)以及 Turner 和 Renton(Can. J. Physiol. Pharmacol. 67, 582-586, 1989)的体外方法可在临床上评估药物与药物可能的相互作用。另外, 采用各种体外分析方法也可测试 D-氨基酸氧化酶或 D-天冬氨酸氧化酶拮抗剂的相对活性、功效和特异性。

有效剂量取决于各种因素而改变, 如病人的情况、疾病症状的严重程度、药物组合物的给药方式。对人患者而言, 典型化合物的有效剂量一般要求给每个病人施加的化合物至少约 1 毫克/24 小时, 通常至少约 10 毫克/24 小时, 往往至少约

25 毫克/24 小时。对人患者而言，典型化合物的有效剂量要求给每个病人施加的化合物一般不超过约 500 毫克/24 小时，通常不超过约 400 毫克/24 小时，往往不超过约 300 毫克/24 小时。另外，所施加的有效剂量使得化合物在病人血浆中的浓度在正常情况下不超过 500 纳克/毫升，往往不超过 100 纳克/毫升。

本发明的化合物和组合物每日可以约 0.1 到约 1,000 毫克的剂量水平施加给病人。对体重约 70 公斤的正常成年人而言，估计每日每公斤体重约 0.01 到约 100 毫克是足够的。不过，采用的具体剂量是可以改变的。例如，剂量可取决于多种因素，包括病人的要求、待治疗的病况严重性、所用化合物的药理活性。一个特定病人的最佳剂量的确定已为本领域技术人员所公知。有效成分可与载体材料结合配成单剂量形式，而有效成分的含量取决于待治疗的宿主和给药的具体方式。然而，需要明白的是，任何一个特定病人的具体剂量水平取决于各种因素，包括所用具体化合物的活性、年龄、体重、身体状况、性别、饮食、给药时间、给药途径、排泄速度、药物组合及待治疗的特定疾病的严重性。

在一些实施例中，本发明的实验采用各种组合的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶拮抗剂。所以，含有一种以上 D-氨基酸氧化酶拮抗剂的组合物可用在本发明的治疗方法中。此外，含有一种以上 D-天冬氨酸氧化酶拮抗剂的组合物也可用在本发明所公开的方法中。在另一实施例中，本文公开的治疗方法采用至少一种 D-氨基酸氧化酶拮抗剂和至少一种 D-天冬氨酸氧化酶拮抗剂的组合。

本发明方法所用的优选化合物能够通过病人的血脑屏障。故这些化合物能够进入病人的中枢神经系统。进行本发明所用的典型化合物的 $\log P$ 值一般大于 0，通常约大于 1，往往约大于 1.5。这些典型化合物的 $\log P$ 值一般约小于 4，通常约小于 3.5，往往约小于 3。 $\log P$ 值是衡量化合物通过扩散屏障如生物膜的能力。参见 Hansch 等人(J. Med. Chem., Vol. 11, p. 1, 1968)。此外，利用本领域已知的由旁路通过血脑屏障的组合物和方法(例如美国专利申请号 5,686,416; 5,994,392，纳入其整体内容作为参考)可使本发明的组合物由旁路通过血脑屏障，或者可直接注射入脑内。适当的注射部位包括大脑皮层、小脑、中脑、脑干、下丘脑、脊髓和心室组织、以及包括颈动脉体和肾上腺髓质在内的外周神经系统区域。组合物可以大丸剂或通过利用渗透泵等其它方法给药。

本发明的化合物可单独或作为组合物的一部分施加给病人，所述的组合物含有其它成分，例如本领域公知的赋形剂、稀释剂和载体。组合物可通过口服、直肠、肠胃外(静脉、肌肉内或皮下)、脑池内、阴道内、腹膜内、膀胱内、局部(粉末、药膏或液滴)，或者以口腔或鼻喷剂或吸入而给予人和动物。

适合肠胃外注射用的组合物可包括生理学上可接受的无菌或非水溶液、分散液、悬浮液、乳液、重构为无菌注射液或分散液的无菌粉末。合适的水溶性和非水溶性载体、稀释剂、溶剂或载色剂包括水、乙醇、多元醇(丙二醇、聚乙二醇、甘油等)及其合适的混合物,菜油(如橄榄油)和可注射的有机酯如油酸乙酯。例如,利用膜层如蛋黄素,保持分散液所需的颗粒大小以及利用表面活性剂都可保持适当的流质。

这些组合物还可含有辅助剂,如防腐剂、湿润剂、乳化剂和分散剂。通过各种抗菌剂和抗真菌剂,例如对羟苯甲酸酯、氯丁醇、苯酚、山梨酸等,可确保防止发生微生物作用。也希望组合物包括等渗剂,例如糖类、氯化钠等。采用延迟吸收的试剂,如单硬脂酸铝和明胶,可导致注射药物形式延长吸收。

口服的固体剂量形式包括胶囊、片剂、丸剂、粉末、和粒剂。这些固体剂量形式的活性化合物与至少一种常用惰性赋形剂(或载体),如柠檬酸钠或磷酸氢钙,或(a)填充剂或添加剂,如淀粉、乳糖、蔗糖、葡萄糖、甘露醇、硅酸;(b)结合剂,如羧甲基纤维素、藻酸盐、明胶、聚乙烯基吡咯烷酮、蔗糖和阿拉伯树胶;(c)保湿剂,如甘油;(d)崩解剂,如琼脂-琼脂、碳酸钙、马铃薯或木薯淀粉、褐藻酸、某些复合硅酸盐和碳酸钠;(e)阻溶剂,例如石蜡;(f)吸收加速剂,如季铵化合物;(g)湿润剂,如十六烷基醇和单硬脂酸甘油;(h)吸附剂,如高岭土和斑脱土;以及(i)润滑剂,如滑石、硬脂酸钙、硬脂酸镁、聚乙二醇固体、十二烷基硫酸钠、或其混合物一起掺合。胶囊、片剂和丸剂的剂量形式还可含有缓冲剂。

利用乳糖或奶糖以及高分子量聚乙二醇等赋形剂,也可用类似形式的固体组合物作为软填充和硬填充明胶胶囊的填充剂。以膜层和外壳,如肠溶包膜和本领域公知的其它膜层和外壳,可制备有壳剂量形式,如片剂、糖衣片、胶囊、丸剂和粒剂。它们可含有遮光剂,而且可以是这样一种组合物,即它们以延迟方式在肠道的某一部分释放出活性化合物。可采用的包埋组合物的例子是聚合物和蜡。如适用,活性化合物可与上述一种或多种赋形剂制成微被囊形式。

口服的液体剂量形式包括药学上可接受的乳液、溶液、悬浮液、糖剂和酏剂。除了活性化合物外,液体剂量形式可含有本领域常用的惰性稀释剂,如水或其它溶剂、增溶剂和乳化剂,例如乙醇、异丙醇、碳酸乙酯、乙酸乙酯、苯甲醇、苯甲基苯甲酸盐、丙二醇、1,3-丁二醇、二甲基甲酰胺、油类特别是棉花籽油、花生油、玉米胚芽油、橄榄树、调和油和芝麻油,甘油、四氢糠醇、聚乙二醇和山梨聚糖的脂肪酸酯或这些物质的混合物等等。除了这些惰性稀释剂外,组合物还可包括辅助剂,如湿润剂、乳化剂、悬浮剂、甜味剂、香味剂和芳香料。

悬浮液除了有活性化合物外，还可含有悬浮剂，例如乙氧基化异十八醇、聚氧乙烯山梨醇和山梨聚糖酯、微晶纤维素、偏氢氧化铝、斑脱土、琼脂-琼脂和西黄耆胶或这些物质的混合物等等。

直肠给药的组合物最好是栓剂，栓剂可通过将本发明的化合物与适当无刺激性赋形剂或载体如可可脂、聚乙二醇或栓剂蜡混合而制备，栓剂在常温下是固体，但在体温时是液体，因而会在直肠或阴道内融化并释放出活性组分。

本发明的化合物局部给药的剂量形式包括药膏、粉末、喷剂和吸入剂。在灭菌条件下将活性组分与生理学上可接受的载体和可能需要的任何一种防腐剂、缓冲液或推进剂混合。眼用的制剂、眼膏、粉末和眼药水也包括在本发明的范围之内。

另外，本发明的化合物可以与药学上可接受的溶剂如水、乙醇等配制成非溶剂化及溶剂化形式。一般认为，用于本发明用途的溶剂化形式相当于非溶剂化形式。

动物模型

大鼠模型中的条件性回避行为

条件性回避模型是预测抗精神病活性的标准行为测试。目前动物临床使用的抗精神病药物的一个重要药理性质是它们阻断条件性回避反应的能力。例如参见 Cook, L. 和 Davidson, A. B.: Behavioral pharmacology: Animal models involving aversive control of behavior; M. A. Lipton、A. Dimascio 和 K. Killam 编辑的 In Psychopharmacology, A Generation of Progress, 563-567, Raven 出版社, 纽约, 1978; Davidson, A. B. 和 Weidley, E. Differential Effects of Neuroleptic and other Psychotropic Agents on Acquisition of Avoidance in Rats, 18 Life Sci. 1279-1284 (1976), 这些文献公开的资料纳入其整体内容作为参考。抗精神病药物条件性回避测试的活性和效力与它们的临床疗效和效力密切相关。例如参见纳入其整体内容作为参考的文献 Creese, I., Burt, D. R. 和 Snyder, S. H.: Dopamine receptor binding predicts clinical and pharmacological properties of antischizophrenic drugs. Science (Washington D.C.) 192: 481-483, 1976。

在条件性回避测试中，动物学习对条件刺激的反应，以回避轻度电震。对条件刺激的反应称为回避反应，对电震的反应称为逃避反应；反应失败是当动物对条件性刺激或电震没有反应，这表示运动受损。在绝大多数场合下动物很快学会回避。抗精神病药物使回避反应的百分比下降，但由于动物确实发出逃避反应，故没有干扰动物作出反应的能力。将反应失败的百分比看成是衡量运动受损的一个指标。

要求大鼠按住实验箱的一反应杆，以回避或逃避足电震。每一个实验阶段由 50 次试验组成。每次试验实验箱都设有灯光和一个最长时间为 10 秒的调音器 (tone)。对音频有反应立即停止音频和灯光，结束试验。对单独的音频无反应，给予最长 10 秒的音频+足电震 (2.0 毫安)。对电震有反应立即终止电击、音频和灯光，结束试验。

要筛选药物，在实验开始前以适当的方式在适当的时间 30 分钟内，如腹膜内或皮下施加适当的剂量，如 3.0 毫克/公斤。待处理的一组可以只接受单剂量的 D-氨基酸氧化酶和 D-天冬氨酸氧化酶拮抗剂，或者可以在实验前每日处理 (例如 1 日 1 次、1 日 2 次或 1 日 3 次) 至少 1 日、3 日、1 星期、2 星期、1 个月或 2 个月。如果药物使回避反应的百分比下降到至少 50% 而反应失败又不会大于 50%，就认为该药物是有活性的。然后再测定活性药物的剂量—反应曲线。

沙鼠足踢踏 (Foot-tapping)

使雄性或雌性蒙古种沙鼠 (35-70 克) 吸入异氟烷/氧气混合物麻醉，露出颈静脉血管，以便静脉注射 5 毫升/公斤的测试或对照化合物或介质 (vehicle)。测试化合物还可经口服或皮下或腹膜内途径给予。待处理的一组可以只接受单剂量测试化合物，或者可以在分析前每日处理 (例如 1 日 1 次、1 日 2 次或 1 日 3 次) 至少 1 日、3 日、1 星期、2 星期、1 个月或 2 个月。然后在头皮中线切一个小口，露出头骨。垂直插入有套的 27 号针至前囟下深 4.5 毫米，将抗焦虑药 (如五肽胃泌素) 和/或对照试剂 (盐水、D-氨基酸氧化酶或 D-天冬氨酸氧化酶拮抗剂、D-Ser、D-Asp 等) 直接注入脑室 (例如脑室注射 5 μ l 中的 3 pmol，取决于测试用的物质)。缝合头皮切口，让动物在透明的有机玻璃箱 (25 厘米 x 20 厘米 x 20 厘米) 中从麻醉状态恢复过来。然后在大约 5 分钟内连续记录后足踢踏持续的时间及强度。此外，用类似的定量分析方法可研究测试化合物抑制由厌恶刺激，如足电震或独处，所引起的足踢踏的能力。本发明优选的拮抗剂能够抑制沙鼠的诱导性足踢踏。

雪貂呕吐模型

给单独关住的雄性雪貂喂食，口服测试化合物或对照化合物或介质。10 分钟后，给它们喂食约 100 克罐头猫食物。待处理的一组可以只接受单剂量测试化合物，或者可以在实验前每日处理 (例如 1 日 1 次、1 日 2 次或 1 日 3 次) 至少 1 日、3 日、1 星期、2 星期、1 个月或 2 个月。口服药剂后 60 分钟，以氟烷短暂麻醉，经插入的颈静脉导管静脉注射顺铂 (10 毫克/公斤)。移去导管，结扎颈静脉，缝合皮肤切

口。雪貂很快从麻醉状态恢复过来，10-20 分钟内就可以活动。注射顺铂后连续 4 小时观察动物的麻醉恢复情况。然后人工处死动物。由经过训练的观察人员记录给予顺铂后 4 小时期间出现的作呕和呕吐次数。

分离诱导的发声模型

整个研究期间以家庭为组别将雄性和雌性小天竺鼠与它们的母鼠及同窝出生的仔畜关在一起。当小天竺鼠 2 周大时断奶后开始实验。进行实验前筛选小天竺鼠，确保与母亲分隔后能重复发出强烈的声音反应。将小天竺鼠单独放在一个观察笼(55 厘米 x39 厘米 x19 厘米)中 15 分钟，观察笼置于实际上与家庭笼隔离的一间房内，记录此基线期间发声持续的时间。只选用发声超过 5 分钟的动物作药物攻击研究(所有小天竺鼠中约 50%未能达到这个标准)。待处理的一组可以只接受单剂量测试化合物，或者可以在实验前每日处理(例如 1 日 1 次、1 日 2 次或 1 日 3 次)至少 1 日、3 日、1 星期、2 星期、1 个月或 2 个月。如上所述，在测试日社交隔离 15 分钟前每只小天竺鼠接受口服剂量或者皮下或腹膜内注射测试化合物或介质，再立即送回家庭笼与母鼠和兄弟姐妹一起 30 到 60 分钟(或者长达口服剂量后 4 小时，取决于测试化合物的口服药物动力学)。每只动物在药物治疗日发声持续的时间以治疗前基线的百分比表示。每只受试动物每周再测试一次，最长达 6 周。有 6 至 8 只动物一次测试剂量接受一种测试化合物。本发明优选的拮抗剂能有效减弱下文所限定的小天竺鼠的分离诱导的发声。

在 omnitech digiscan 动物活动监测器中小鼠和大鼠的行为活动评估

本测试的目的在于评估用于抗精神样中枢神经系统效应和各种一般与中枢神经系统活动相关的其它行为效应的化合物。这种测试能够确定药物对啮齿动物活动能力许多方面的影响，包括水平活动(断开光束)、行进的总距离(以厘米计)、移动的次数、移动的时间(以秒计)、静止时间(以秒计)、垂直活动(断开光束)、垂直移动的次数、垂直时间(以秒计)、刻板性动作计数(Stereotypy counts)、刻板性动作发作的次数、刻板性动作的时间(以秒计)、边缘和中心时间(以秒计)、顺时针和逆时针方向转数以及在活动监测器各个角落所停留的时间(以秒计) 不过，一般来说，以行进的总距离(以厘米计)作为活动能力最准确的衡量指标来评估药物对行为的影响。。

这些研究用重 20 到 40 克的雄性 CD-1 白化病小鼠(Charles River 实验室)或重 150 到 300 克的雄性 Sprague-Dawley 大鼠(Harlan 实验室)。待处理的一组可以

在实验前只接受单剂量测试化合物,或者可以在实验前每日处理(例如1日1次、1日2次或1日3次)至少1日、3日、1星期、2星期、1个月或2个月。

Omnitech Digiscan 动物活动监测器由16英寸x16英寸x12英寸的有机玻璃室组成,内设两组16个红外线光束传感器,玻璃室底板的四侧每隔1英寸设1个传感器。另一组光束传感器直接置于较低光束传感器的上面,测量垂直活动。任何光束的截断应该会使位于监测器主框架中心的发光二极管指示灯发光。在实验开始前通常对24个监测器中每一个进行诊断性测试,测试时检查所有光束是否有任何截断。每个活动监测器可用适合在该有机玻璃室内的有机玻璃插板分成4个8英寸的四分部,其中2个四分部(前左和后右)可用来测试活动。一般来说,这种分割安排用于小鼠的活动研究(每个分割监测器有2只小鼠)与大鼠的研究(每个未分割监测器有1只大鼠)作对比。在长达999分钟的时间内可取得多达999个数据样本。通常,每持续10分钟收集6个小鼠数据样本(1小时的测试),或者每5分钟收集6个大鼠数据样本(30分钟的测试)。

一旦将动物放在活动室,就可以使这些活动室各自启动,开始收集数据。由于黑暗刺激得到的数据较为稳定,所以通常关掉高处的灯光监测活动水平。每次实验收集以下各种数据(附简单定义):

变量1: 水平活动—水平传感器出现光束断开的总次数。

变量2: 行进的总距离(以厘米计)—由于考虑到任何一次对角线运动,所以它是一个更为准确的移动性活动指标。

变量3: 移动的次数—间隔至少1秒内不连续移动的次数。

变量4: 移动的时间(以秒计)—移动的时间段

变量5: 静止时间(以秒计)—取样时间与移动所用的时间之差。

变量6: 垂直活动—动物直立时在垂直传感器出现光束断开的总次数。

变量7: 垂直移动的次数—动物每次直立并断开垂直传感器的次数(间隔至少1秒)。

变量8、9、10和11: 在角落处所停留的时间(左和右前、左和右后)—动物在很接近笼两面邻接墙所停留的时间。

变量12: 垂直时间(以秒计)—直立时断开垂直光束所用的时间。

变量13: 刻板性动作计数—重复(刻板性动作)断开同一束光束(或同一组光束)期间出现光束断开的次数。

变量14: 刻板性动作的次数—监测器在间隔至少1秒内观察刻板性动作行为的次数。

变量 15: 刻板性动作时间(以秒计)—显示刻板性动作行为的总时间段。

变量 16: 顺时针方向转数—动物绕至少 2 英寸直径旋转的次数(不计算较急转的旋转运动)

变量 17: 逆时针方向转数—动物绕至少 2 英寸直径旋转的次数

变量 18: 边缘时间(以秒计)—动物在很接近有机玻璃笼墙壁处所停留的时间。

变量 19: 中心时间(以秒计)—动物远离有机玻璃笼墙壁处所停留的时间。

数据可以实际计数、时间(以秒计)、行进的厘米、或者活动能力抑制百分比相对同时测试的介质处理的对照动物来表示。通过 t 检验或方差分析和纽曼-库尔变域测试法(Newman-Keul's multiple-range test)确定活性相对对照动物的(即行进的厘米)差异显著。活性水平的刺激以负数表示。用落在剂量—效应曲线直线部分的至少三个数据点以回归分析法估计预期可能使活性水平下降 50%(ED. sub. 50)的剂量以及在该数值左右的 95%置信区间。

大鼠模型中安他非明刺激的运动的阻断

阻断安他非明刺激的运动步骤是上述在 omnitech digiscan 活动监测器中活动能力方案的一种变化。阻断安他非明刺激的运动步骤使用中枢神经系统刺激剂 d-安他非明来评估多巴胺能试剂的抗精神病活性。

这些研究采用雄性 Sprague-Dawley 大鼠(Harlan 实验室)。待处理的一组可以在实验前只接受单剂量测试化合物,或者可以在实验前每日处理(例如 1 日 1 次、1 日 2 次或 1 日 3 次)至少 1 日、3 日、1 星期、2 星期、1 个月或 2 个月。对于腹膜内注射研究,在给药前 20 分钟给予安他非明,然后进行 30 分钟活动能力测试。对于口服研究,在测试前 30 分钟给药,而在测试前 15 分钟给予安他非明,这样可以有充足的时间口服吸收。在 16 英寸 x16 英寸的敞开式小室内测量活动能力(每 30 分钟测试所行进的厘米)。安他非明产生的运动一般比盐水对照组高 2 到 3 倍。药物的影响以安他非明刺激的运动的反向百分比来表示。用 t 检验确定与安他非明刺激的运动相对用安他非明处理的对照动物的差异显著。以回归分析法估计使安他非明刺激的运动下降 50%(ED. sub. 50)的剂量及 95%置信区间。

大鼠模型中声音惊吓的前脉冲抑制方案

声音惊吓的前脉冲抑制是感觉运动选通(sensorimotor gating)的一种形式,当一个弱刺激先于惊吓刺激就会发生声音惊吓的前脉冲抑制,导致惊吓反应幅度减少。精神分裂症患者显示声音惊吓的前脉冲抑制比对照受试者低,与感觉运动选通

的损失一致。所以利用这种现象的动物模型在已知的和潜在的抗精神病药物研究中十分有用。例如，用诸如脱水吗啡等直接多巴胺激动剂，或者间接多巴胺激动剂安他非明都可以阻断大鼠的前脉冲抑制，这种效应可用多巴胺拮抗剂如氟哌啶醇拮抗。

来自 Harlan 实验室的雄性 Sprague-Dawley 大鼠(180-280 克)以 5 只为一组关在一个笼中，使它们处于一个 12 小时的光/暗循环，能够自由进食和饮水。待处理的一组可以在实验前只接受单剂量测试化合物，或者可以在实验前每日处理(例如 1 日 1 次、1 日 2 次或 1 日 3 次)至少 1 日、3 日、1 星期、2 星期、1 个月或 2 个月。

本研究使用的惊吓室(SR-LAB, San Diego Instruments)由有机玻璃圆柱筒组成，它放在有通风设备、声音衰减的密封环境内一个有机玻璃框架上。由设在大鼠上面的扬声器发出声刺激。有机玻璃框架下面装上一个压电装置，检测和转换开始声刺激后 100 毫秒期间圆柱筒内发生的运动。由微电脑和界面组合(San Diego Instruments)记录 100 毫秒记录窗口(100x1 毫秒读数)期间的平均反应。每个惊吓室之间要相互校准，确保扬声器的性能水平在较大范围的分贝水平(67 至 125 dB)内一致。用分贝计(例如 Radio Shack)评估声水平。调节每个稳定计(它装有压电装置)，使恒定振动校准器的反应灵敏度相等。

用测试化合物处理的动物可以在实验前只接受单剂量测试化合物，或者可以在实验前每日处理(例如 1 日 1 次、1 日 2 次或 1 日 3 次)至少 1 日、3 日、1 星期、2 星期、1 个月或 2 个月。每只动物实验前用盐水、测试化合物或对照化合物(例如脱水吗啡、氟哌啶醇、氯扎平等)预处理。

每个测试阶段开始是 5 分钟 70 dB 白色噪音的测试适应期。测试阶段总共持续 30 分钟；连续进行数个测试，使每一处理组有足够数量的大鼠。第一个和最后一个试验是只发出 120 dB 脉冲的试验，相隔 7 到 23 秒，这期间大鼠很快就适应发出的噪音。前脉冲抑制的计算不包括这些数据。中间试验由 120 dB 脉冲的单一试验和以下 5 种试验组成，其中 5 种之中的每一种试验以伪随机次序：(1)没有刺激，(2)120 dB 惊吓前 72 dB 预脉冲 100 毫秒，(3)120 dB 惊吓前 74 dB 预脉冲 100 毫秒，(4)120 dB 惊吓前 78 dB 预脉冲 100 毫秒，以及(5)120 dB 惊吓前 86 dB 预脉冲 100 毫秒。预脉冲(2、4、8 和 16 dB，背景噪音为 70 dB)持续 20 毫秒，而惊吓刺激持续 40 毫秒。当预脉冲与 120 dB 脉冲配对时，人耳检测不到与只有 120 dB 脉冲时有明显的声差异。声音惊吓的前脉冲抑制的反射以当 2 至 16 dB(相对背景)预脉冲先于惊吓刺激时所产生的 120 dB 惊吓幅度的抑制百分比来表示。

脱水吗啡诱导的攀爬行为的抑制

在动物药理研究中，通过脱水吗啡诱导的攀爬行为可测试测试化合物的抗精神病活性(P. Protais 等人, "Psychopharmacology", 50, 1-6, 1976)。研究使用重 22-24 克的雄性瑞士小鼠。用测试化合物处理过的动物可以在实验前只接受单剂量测试化合物，或者可以在实验前每日处理(例如 1 日 1 次、1 日 2 次或 1 日 3 次)至少 1 日、3 日、1 星期、2 星期、1 个月或 2 个月。在时间 0 点给动物口服测试药物或 25%琼脂。60 分钟后，皮下注射 1 毫克/公斤剂量的脱水吗啡，再过 70 分钟，评估动物的行为。每隔 10 分钟进行 2 次附加评估。评估时将动物放在垂直小箱(11x7.5x4.5 厘米)底上。除了一个侧表面(宽 7.5 厘米)是 3 毫米金属丝网外，箱的其它墙壁由半透明的甲基丙烯酸酯制成。按照下列标准在 2 分钟内给动物的位置计分：0=4 只爪在地上；1=3 只爪在地上；2=2 只爪在地上；3=1 只爪在地上；以及 4=四只爪握住金属丝网。如果一只动物在 2 分钟观察期内保持几个位置，就记录每个位置所保持的秒数。最后，计算平均分数。在这些实验条件下，计算有效剂量 50%(ED. sub. 50)的数值。

DOI 诱导的头部颤搐和搔抓的抑制

也可以通过 1-(2, 5-二甲氧基-4-碘苯基)-2-氨基丙烷(DOI)诱导的头部颤搐和搔抓的抑制试验来测试测试化合物的抗精神病活性(M. Oka 等人, "J. Pharm. Exp. Ther.", 264(1), 158-165, 1993)。研究使用重 22-26 克的雄性 N. M. R. I. 小鼠。动物称重后，实验前 2 小时将它们单独放在透明的笼中。用测试化合物处理过的动物可以在实验前只接受单剂量测试化合物，或者可以在实验前每日处理(例如 1 日 1 次、1 日 2 次或 1 日 3 次)至少 1 日、3 日、1 星期、2 星期、1 个月或 2 个月。在时间 0 点给动物口服测试化合物。60 分钟时，腹膜内注射 3 毫克/公斤剂量溶于盐水的 DOI。评估动物头部颤搐和搔抓的次数以及试图逃跑和不逃跑的次数。计算上述实验条件下所获得的有效剂量 50%(ED. sub. 50)的数值。

人的临床试验

人的临床试验证实，本发明的 D-氨基酸氧化酶和 D-天冬氨酸氧化酶拮抗剂对治疗或减缓精神分裂症、双相性精神障碍或其它中枢神经系统障碍具有活性。例如，可将一项研究设计为双盲、平行、安慰剂为对照的多中心试验。受试者随机分为 4 组，安慰剂和 3 组测试化合物，3 组测试化合物 1 日 3 次给药，剂量逐渐增加，例如 25、50 和 75 毫克。给药的方式按本文所述或本领域技术人员常用的方法，如与

食物一起口服。在 4 次访问中观察受试者，得到基线测量值。其它访问，如第 5-33 次访问，作为研究的治疗阶段。

在访问期间，观察受试者精神病行为或双相性行为的迹象，如兴奋、情绪波动、震颤、精神错乱、不去社交以及精神力集中的能力。以受试者的数目和百分比对治疗组进行比较，而受试者在双盲部分研究(访问 5 到 33 次)中曾经有过症状，而严重性则差过在基线访问(访问 1 到 4 次)期间。

遗传诊断学方法中的 D-氨基酸氧化酶、D-天冬氨酸氧化酶及其双等位标记

本发明的 D-氨基酸氧化酶和 D-天冬氨酸氧化酶基因组和 cDNA 序列以及双等位标记也可用来开发一些诊断试验，所述试验能够鉴定表达由特异性基因型导致的可检测性状的个体，或者其基因型使他们以后有可能发展一种检测性状的个体。采用本发明诊断学方法来分析的性状可用来诊断任何一种可检测的性状，包括精神分裂症或双相性精神障碍易感性、可检测性状病发的年龄、对精神分裂症或双相性精神障碍的治疗有益的反应或与所述治疗相关的副作用。这样一种诊断可用于精神分裂症或双相性精神障碍的监测、预后和/或预防性或治愈性治疗。

本发明的诊断技术可采用各种方法判断测试受试者是否有与可能逐渐发展一种可检测性状相关的基因型，或者个体是否由于一种特定的突变而带有一种可检测的性状，所述方法包括能够分析个体染色体的单倍分型的方法，如家族研究、单个精子 DNA 分析或体细胞杂交。

诊断技术涉及人 D-氨基酸氧化酶或 D-天冬氨酸氧化酶基因内，优选为 D-氨基酸氧化酶或 D-天冬氨酸氧化酶外显子或编码序列内，所含的特异性等位基因的检测。更具体地说，本发明涉及一种核酸的检测，所述核酸含有至少一段 SEQ ID No: 1 至 6 的核苷酸序列或其片段或其包括多态性碱基的互补序列。

这些方法包括从个体采集核酸样本，确定该核酸样本是否含有至少一个等位基因或至少一个双等位标记单倍型，这表示有发展性状的风险或者表示该个体表达由于具有特定的人 D-氨基酸氧化酶或 D-天冬氨酸氧化酶相关多态性或突变(引起性状的等位基因)而产生的性状。

在这些诊断方法中，核酸样本最好取自个体，而该样本用本领域已知的或本文引用文献 PCT/IB00/00435 的实施例所述的方法进行基因分型。诊断方法也可基于单个双等位标记或一组双等位标记。

在这些方法中，核酸样本均取自测试受试者，确定一个或以上本发明的双等位标记的双等位标记模式。

在一实施例中，对核酸样本进行 PCR 扩增，以便扩增那些与可检测的表型相关的多态性已被识别的区域。对扩增产物作序列分析，以确定该个体是否具有一种或多种与可检测的表型相关的人 D-氨基酸氧化酶或 D-天冬氨酸氧化酶相关多态性。另外，使核酸样本进行微序列分析反应，以确定个体是否具有一种或多种与可检测的表型相关的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶相关多态性，所述的表型是由于 D-氨基酸氧化酶或 D-天冬氨酸氧化酶基因组序列的突变或多态性而产生的。在另一实施例中，使核酸样本与一个或多个等位基因特异性寡核苷酸探针接触，所述探针和一个或多个与可检测的表型相关的人染色体 D-氨基酸氧化酶或 D-天冬氨酸氧化酶相关等位基因特异性杂交。在另一实施例中，使核酸样本与第二个寡核苷酸接触，所述寡核苷酸与等位基因特异性寡核苷酸一起用在扩增反应时能够产生扩增产物。扩增反应有扩增产物表示该个体具有一个或多个与可检测的表型相关的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶相关等位基因。在一优选实施例中，可检测的性状是精神分裂症或双相性精神障碍。诊断试剂盒包括本发明的任何一种多核苷酸。

这些诊断方法非常有用，这是因为它们在某些情况中可用于启动预防性治疗，或者可使携带显著单倍型的个体预知征兆，如不严重的症状等。

能够分析或预测药物反应或药物副作用的诊断方法可用来确定个体是否应该用特定药物进行治疗。例如，如果诊断显示个体可能会对用特定药物进行的治疗呈阳性反应，就可以给个体施加该药物。相反，如果诊断显示个体可能会对用特定药物进行的治疗呈阴性反应，就要换另一种疗法。阴性反应定义为缺乏有效应答或存在有毒副作用。

本发明的标记的另一种应用是临床药物试验。用上述方法可鉴定一个或多个标记，所述标记显示对精神分裂症起作用的药剂或对精神分裂症起作用的药剂副作用的反应。然后，在这样一种药剂的临床试验中筛选潜在的参与物，以鉴定那些对药物的反应最有可能呈阳性的个体，排除那些可能有副作用的个体。以这种方式测量对药物反应呈阳性的个体中药物治疗的效果，由于研究包括不可能呈阳性反应的个体，所以测量的结果是不会下降的，也没有不希望看到的安全问题的风险。

在防治疾病中的 D-氨基酸氧化酶或 D-天冬氨酸氧化酶

由于大部分人有自杀倾向，所以对精神分裂症、双相性精神障碍及其它精神病易感性的检测变得非常重要。故本发明涉及一种精神分裂症或双相性精神障碍或相关精神障碍的治疗方法，所述方法包括以下步骤：

• 选择个体,该个体的 DNA 含有与精神分裂症或双相性精神障碍相关的一个 D-氨基酸氧化酶或 D-天冬氨酸氧化酶相关双等位标记或一组 D-氨基酸氧化酶或 D-天冬氨酸氧化酶相关标记等位基因;

• 跟踪所述个体表现出来与精神分裂症或双相性精神障碍相关的症状; 以及

• 在疾病的适当阶段给所述个体施加对精神分裂症或双相性精神障碍或其症状起作用的治疗。

本发明的另一个实施例包括一种精神分裂症或双相性精神障碍的治疗方法,所述方法包括以下步骤:

• 选择个体,该个体的 DNA 含有与精神分裂症或双相性精神障碍相关的一个 D-氨基酸氧化酶或 D-天冬氨酸氧化酶相关双等位标记或一组 D-氨基酸氧化酶或 D-天冬氨酸氧化酶相关标记等位基因;

• 给所述个体施加精神分裂症或双相性精神障碍的预防性治疗。

在又一个实施例中,本发明涉及一种精神分裂症或双相性精神障碍的治疗方法,所述方法包括以下步骤:

• 选择个体,该个体的 DNA 含有与精神分裂症或双相性精神障碍相关的一个 D-氨基酸氧化酶或 D-天冬氨酸氧化酶相关双等位标记或一组 D-氨基酸氧化酶或 D-天冬氨酸氧化酶相关标记等位基因;

• 给所述个体施加精神分裂症或双相性精神障碍的预防性治疗;

• 跟踪所述个体的精神分裂症或双相性精神障碍症状的表现和发展; 以及可选择地

• 在疾病的适当阶段给所述个体施加对精神分裂症或双相性精神障碍或其症状起作用的治疗。

要用来确定患病个体的疗程,本发明还涉及一种精神分裂症或双相性精神障碍的治疗方法,所述方法包括以下步骤:

• 选择有精神分裂症或双相性精神障碍的个体,该个体的 DNA 含有与精神分裂症或双相性精神障碍或其症状的严重性相关的一个 D-氨基酸氧化酶或 D-天冬氨酸氧化酶相关双等位标记或一组 D-氨基酸氧化酶或 D-天冬氨酸氧化酶相关标记等位基因; 以及

• 给所述个体施加对精神分裂症或双相性精神障碍或其症状起作用的治疗。

本发明还涉及一群所选的个体的精神分裂症或双相性精神障碍的治疗方法,所述方法包括以下步骤:

• 选择有精神分裂症或双相性精神障碍的个体,该个体的 DNA 含有一个 D-氨基

酸氧化酶或D-天冬氨酸氧化酶相关双等位标记或一组D-氨基酸氧化酶或D-天冬氨酸氧化酶相关标记等位基因,所述双等位标记与用对精神分裂症或双相性精神障碍或其症状起作用的有效治疗量的药物进行的治疗呈阳性反应相关;

- 和/或该个体的DNA不包含一个D-氨基酸氧化酶或D-天冬氨酸氧化酶相关双等位标记或一组D-氨基酸氧化酶或D-天冬氨酸氧化酶相关标记等位基因,所述双等位标记与用所述药物进行的治疗呈阴性反应相关;

- 每隔一段适当时间给所选择的个体施加有效量的所述药物。

在本发明的上下文中,对药物呈“阳性反应”定义为包括疾病相关症状有所减缓。在本发明的上下文中,对药物呈“阴性反应”定义为包括对药物缺乏阳性反应,导致症状没有减缓或在施加药物后出现副作用。

本发明也涉及一种受试者是否可能对药物治疗呈阳性反应的判断方法。该方法包括鉴定对所述药物呈阳性反应的第一组个体以及鉴定对所述药物呈阴性反应的第二组个体。在对所述药物呈阳性反应的第一组个体中鉴定一个或多个双等位标记,或者在对所述药物呈阴性反应的第二组个体中鉴定一个或多个双等位标记。这些双等位标记可用本文所述的技术进行鉴定。

然后,从待测受试者采集DNA样本。对该DNA样本进行分析,确定它是否含有一个或多个与对药物治疗呈阳性反应相关的双等位标记等位基因,和/或含有一个或多个与对药物治疗呈阴性反应相关的双等位标记等位基因。

在一些实施例中,如果DNA样本含有一个或多个与对药物治疗呈阳性反应相关的双等位标记等位基因和/或如果DNA样本缺乏一个或多个与对药物治疗呈阴性反应相关的双等位标记等位基因,就可以在临床试验中给受试者施加该药物。在优选实施例中,该药物是对精神分裂症或双相性精神障碍起作用的一种药物。

利用本发明的方法,在可能对药物呈阳性反应的一群个体中评估药物的功效。

本发明的另一方面内容是一种药物的使用方法,该方法包括从受试者采集DNA样本,判断DNA样本是否含有一个或多个与对药物呈阳性反应相关的双等位标记等位基因和/或判断DNA样本是否含有一个或多个与对药物呈阴性反应相关的双等位标记等位基因,如果DNA样本含有一个或多个与对药物呈阳性反应相关的双等位标记等位基因和/或如果DNA样本缺乏一个或多个与对药物呈阴性反应相关的双等位标记等位基因,就给受试者施加该药物。

本发明也涉及一种药物最好是对精神分裂症或双相性精神障碍或其症状起作用的药物的临床试验方法。该方法包括以下步骤:

- 给一组异源个体施加一种药物,最好是一种容易对精神分裂症或双相性精神

障碍或其症状起作用的药物；

- 鉴定对所述药物呈阳性反应的第一组个体以及鉴定对所述药物呈阴性反应的第二组个体；
- 鉴定对所述药物呈阳性反应的所述第一组个体的双等位标记；
- 选择一些个体，其 DNA 含有与对所述药物呈阳性反应相关的双等位标记；以及
- 给所述个体施加所述药物。

精神分裂症或双相性精神障碍的任何一种预防、诊断和治疗方法包括药物的使用方法、药物的临床试验方法、判断受试者是否可能对药物治疗呈阳性反应的方法。

一般认为这样一些方法对提高某些药物给予的效益/风险比值有显著效果，而这些药物的给予可能会引起不希望看到的副作用和/或即使正常给予对部分病人也不起作用。

一旦个体被诊断为有精神分裂症或双相性精神障碍，就进行选择测试以确定该个体的 DNA 是否含有一个双等位标记或一组双等位标记等位基因，而所述双等位标记与对治疗呈阳性反应或包括副作用或无反应在内的对治疗呈阴性反应相关。

通过以上所述的检测方法选择用本发明方法进行治疗的病人。选择的病人最好是那些 DNA 不含有与对治疗呈阴性反应相关的一个双等位标记或一组双等位标记等位基因。弄清个体对特定药物无反应或有副作用的基因倾向，医生就可以给予适当药物直接治疗精神分裂症或双相性精神障碍或其症状。

一旦确定了病人的基因倾向，医生就可以为病人选择不会出现或者只在边上出现的阴性反应特别是副作用的适当治疗。

本发明的双等位标记已被证实与精神分裂症或双相性精神障碍相关。不过，本发明的相关精神障碍，特别是相关的中枢神经系统障碍的预防、诊断、管理和治疗方法还包括本文所述任何一种利用本发明的双等位标记的预防、诊断、预后和治疗方法。

实施例

用于细菌和酵母菌中蛋白质表达的质粒的构建

用 pET11a 载体(Stratagen)进行没有标志(tag)的重组蛋白质的表达。以对应 ORF 极限的引物利用 PCR 技术(TaqPlusPrecision System, Stratagen)得到具有适当位点(5'是 NdeI, 3'是 HindIII)的编码序列。

使所产生的 PCR 产物(Qiaquick PCR, Qiagen)纯化, 用 NdeI 和 HindIII 消化, 凝胶纯化(Microspin, PolyLabo), 并接合到以相同的酶开始的载体中。将这些构造体转染到 DH10B 细菌宿主(Gibco BRL), 提取质粒 DNAs, 并作序列分析以选择合适的编码序列。

以 pESC-LEU 穿梭载体(Stratagen)构建用于酵母菌中人 D-氨基酸氧化酶和 g34872 表达的质粒。

没有标志的重组 g34872 的表达和纯化

然后, 使质粒转染到 BL21(DE3) CodonPlus RIL 细菌宿主(Stratagen), 让细菌在 0.8 升 LB 培养基中生长, 直至 A600 吸光度达到 0.7。加入 1 mM 异丙基- β -D-吡喃葡萄糖苷诱导融合蛋白的表达, 再培养多 3 小时。制备细菌沉淀, 立即急冻(-80°C), 再在 30°C 水浴解冻; 加入 2mM AEBSF。将细菌细胞悬浮于 25 毫升 BugBuster 提取剂(Novagen)中, 补充蛋白酶抑制剂混合物(SetIII, Calbiochem)和 10 mM EGTA。悬浮液室温培养 30 分钟, 再加入苯甲酸酶(Novagen), 并连续培养 15 分钟。溶胞产物在 4°C 以 10,000 $\times$ g 离心 30 分钟。通过盐沉淀使细菌蛋白质与上清液分离。将对应 35-55%硫酸铵饱和溶液的蛋白质沉淀溶解在 2 毫升 50 mM pH8 TrisHCl 缓冲液/含有 10 mM DTT 的 50 mM 氯化钠中, 离心澄清溶液, 并装在 Ultragel AcA44 (Pharmacia)层析柱(1.6x 65 厘米)中, 用 20 mM pH8 TrisCl 缓冲液/50 mM 氯化钠缓冲液平衡。以电泳分析洗脱的蛋白质, 合并含有 MN2R 蛋白质的馏分, 超滤浓缩(截留分子量为 10K, Biomax-15, Sigma)。再将蛋白质装在 DEAE-Macroprep(Bio-Rad)层析柱(1x2 厘米)中, 用 20 mM pH8 TrisCl 平衡, 用线性盐梯度洗脱(0 到 1M 氯化钠, 20 柱体积)。合并含有 MN2R 蛋白质的馏分, 超滤浓缩(截留分子量为 10K), 并装在 Superdex 75 (Pharmacia)层析柱(1x27 厘米)中, 用 20 mM pH8 TrisCl 缓冲液平衡。合并单一个主峰的馏分, 浓缩至 5 毫克/毫升, 4°C 储存。电泳纯化的同源蛋白质的产率通常是每升细菌培养基 5 毫克。根据制造商(NuPage by Novagen)的建议, 使蛋白质在 10% NuPage 订制凝胶中进行变性电泳, 使用 MES/SDS 电泳凝胶缓冲液。分子量标记 See-Blue 来自 Invitrogen。用考马斯亮蓝 G 胶体溶液(Sigma)染色后观察蛋白质。

猪肾的天然 D-氨基酸氧化酶的纯化

从 Sigma 购买猪肾 D-氨基酸氧化酶的粗制剂。蛋白质溶于 50 mM pH8 TrisCl(10 毫升配 1 克); 离心澄清溶液并装在 Sephadex G-50 培养层析柱(2.6 x 40 厘米)中,

用 10 mM pH8 TrisCl/100 mM 氯化钠平衡。再将脱盐后的蛋白质超滤(截留分子量为 30K, Biomax-15, Sigma)浓缩 3 倍, 用 10 mM pH8 TrisCl/100 mM 氯化钠/1 mM DTT/10 mM ATP 平衡, 并装在不含 ATP 的同一种缓冲液中的 DEAE-琼脂糖层析柱(1.6x7 厘米)中。所述层析柱先以 2 柱体积 10 mM pH8 TrisCl/100 mM 氯化钠洗涤, 再用 1 柱体积 10 mM pH8 TrisCl/125 mM 氯化钠洗涤, 然后用 10 mM pH8 TrisCl/150 mM 氯化钠缓冲液洗脱。分析馏分的 D-氨基酸氧化酶的酶活性, 合并馏分, 超滤浓缩。接着, 将蛋白质装在 Ultragel AcA44 层析柱(1.6 x 65 厘米), 用 10 mM pH8 TrisCl/100 mM 氯化钠/1 mM DTT 平衡, 并用同一种缓冲液洗脱。超滤浓缩含有电泳纯化的 D-氨基酸氧化酶的馏分, 4°C 储存。

重组的人 D-氨基酸氧化酶的表达和纯化

使质粒转染到 BL21(DE3) CodonPlus RIL 细菌宿主(Stratagen), 让细菌在 3 升 LB 培养基中生长, 直至 A600 吸光度达到 0.7。加入 1 mM 异丙基- β -D-吡喃葡萄糖苷诱导融合蛋白的表达, 再培养多 5 小时。用 BugBuster 提取剂(Novagen)在 2mM AEBSF 存在下提取细菌沉淀, 通常使用苯甲酸酶。溶胞产物用 50 mM TrisCl 调至 pH8, 并在 4°C 以 10,000 $\times$ g 离心 20 分钟。用硫酸铵(30 到 50%饱和溶液)使蛋白质从提取液沉淀下来, 在 4°C 以 10,000 $\times$ g 离心 60 分钟收集沉淀物, 将沉淀物溶于 50 mM pH8 TrisHCl(10 毫升)中; 离心澄清溶液, 并装在 Sephadex G-75 层析柱(2.6x 40 厘米)中, 用 10 mM pH8 TrisCl/100 mM 氯化钠平衡。接下来的纯化步骤与上述关于猪肾 D-氨基酸氧化酶的纯化步骤几乎一样。唯一差别在于蛋白质从 DEAE 树脂洗脱出来: 它是用 10 mM pH8 TrisCl/300 mM 氯化钠缓冲液进行的。电泳纯化的同源 D-氨基酸氧化酶蛋白质的产率通常是每升细菌培养基 0.7 毫克。

酵母菌酿酒酵母中 g34872 和人 D-氨基酸氧化酶蛋白质的表达和提取

使酵母菌酿酒酵母 YPH499 和 FY1679-18B 菌株在 YPD 丰富培养基中生长。用标准乙酸锂方法使质粒转染在酵母菌细胞中; 在 YNG 合成培养基中选择转染株, 30°C 以 2%蜜三糖为碳源在 1 升不含亮氨酸的合成培养基中生长, 直至培养基密度高达 1 u A600/毫升。在 20°C 离心收集细胞, 用 YNGal(含 2% D-半乳糖)置换培养基, 并连续培养 20 小时。沉淀细胞, 用冰水洗涤, 重悬于 20 mM pH8 TrisCl 缓冲液/2 mM AEBSF 中, 用玻璃珠(Sigma)每分钟搅动 8 次, 提取蛋白质。溶胞产物在 4°C 以 10,000 $\times$ g 离心 30 分钟, 收集上清液(S1), 储存于 4°C。沉淀物重悬于 20 mM pH8 TrisCl 缓冲液/2 毫克/毫升皂苷/0.3%肌氨酰胺中, 每分钟搅动 3 次。离心澄

清沉淀提取液(S2)，于-20℃立即急冻。用Bradford试剂(Bio-Rad)检测蛋白质浓度，用兔抗 g34872-his6 血清(稀释浓度为 1/5000)通过蛋白质印迹步骤以及以 D-丝氨酸为底物通过检测 D-氨基酸氧化酶活性确认表达。

D-氨基酸氧化酶活性的检测

分析混合物通常由 200 mM D-丝氨酸(Aldrich)、0.1 mM FAD、75 mM pH8 磷酸钠缓冲液、1U/毫升 HR-过氧化酶(Sigma)组成。混合物使用之前才充入空气饱和。向混合物加入邻苯二茴香胺(Sigma)。在常规分析中 25 微升分析混合物加入 5 微升酶(D-氨基酸氧化酶和混合物)，加入 50 微升 20%硫酸终止培养。以 540 纳米处的过氧化酶氧化的邻苯二茴香胺的吸收率来观察活性。测量吸收率之前以 14000 rpm 的转速离心含高浓度蛋白质的反应物 15 分钟。

进行对照实验以确定 g34872 蛋白质不会影响过氧化酶的酶活性。对过氧化酶作分析，条件与对 D-氨基酸氧化酶的分析完全一样；用过氧化氢为底物，确认 g34872 对 HRP 没有影响。

实施例 1

用在 pESC-Leu 表达载体中所构建的质粒转染酵母菌细胞。一个质粒可以表达人 D-氨基酸氧化酶，第二个质粒 C 末端标记 g34872，第三个质粒是没有插入的载体(对照)。这些细胞诱导表达后(在培养基有 2%半乳糖)再培养 24 小时，然后按下述方法混合和组合提取液：不同体积的 D-氨基酸氧化酶提取液或者与 g34872-cHis6 或者载体提取液混合。相同体积的 D-氨基酸氧化酶提取液也可与 BSA 混合(外部对照)。预培养 30 分钟后，用该混合提取液测量 D-氨基酸氧化酶活性。所有酵母菌提取液和 BSA 溶液的总蛋白质浓度相同。在 37℃ 用 D-丝氨酸测定 D-氨基酸氧化酶活性。图 1 显示 g34872 活化 DAO。

实施例 2

在含有表达 g34872 的大肠杆菌提取液和 BSA 溶液中加入纯化的重组人 D-氨基酸氧化酶。D-氨基酸氧化酶的浓度是 0.5 和 0.3 毫克/毫升。细胞提取液和 BSA 的总蛋白质浓度是 12 毫克/毫升。预培养 30 分钟后，在 37℃ 用该混合提取液测量 D-氨基酸氧化酶活性，底物为 D-丝氨酸(图 2)。

实施例 3

在 g34872 蛋白质存在下体外活化纯化的 D-氨基酸氧化酶:活化效果取决于 g34872 浓度

将纯化的 D-氨基酸氧化酶和 g34872 混合,在活性分析之前在环境温度下培养 50 分钟。所有混合物的总蛋白质浓度相等。D-氨基酸氧化酶的底物采用 D-丝氨酸,反应的 pH 值为 8.0。所用的蛋白质:纯化的猪 D-氨基酸氧化酶在混合物的浓度始终保持在 50 纳克/微升,纯化的重组 g34872 在混合物的浓度是 0 到 450 纳克/微升,牛血清白蛋白(BSA)在混合物的浓度是 0 到 450 纳克/微升。可以认为, g34872 蛋白质的浓度范围在生理学上对应蛋白域腔的高尔基蛋白(lumenal Golgi proteins)的相关数据。见图 3。

实施例 4

D-氨基酸氧化酶体外活化所需的 g34872 浓度极限的估计

将猪肾 D-氨基酸氧化酶与 g34872 在 PBS 中混合,并在 20°C 培养 50 分钟。所有混合物的 D-氨基酸氧化酶浓度是 50 纳克/微升。用 200 mM D-丝氨酸在 20°C 测定 D-氨基酸氧化酶的酶活性, pH 等于 8.0。将猪肾 D-氨基酸氧化酶与 g34872 在 PBS 中混合,并在 20°C 培养 1 小时。所有混合物的 D-氨基酸氧化酶浓度是 50 纳克/微升。用 200 mM D-丝氨酸在 20°C 测定 D-氨基酸氧化酶的酶活性, pH 等于 8.0。见图 4。

实施例 5

D-氨基酸氧化酶在 g34872 蛋白质存在下的动力学: g34872 是 D-氨基酸氧化酶的变构活化剂

将猪肾 D-氨基酸氧化酶与 g34872 在 PBS 中混合,并在 20°C 培养 30 分钟。在蛋白质混合物中 D-氨基酸氧化酶浓度是 200 纳克/微升, g34872 浓度是 2 微克/微升。对照混合物(不含 g34872)由 200 纳克/微升 D-氨基酸氧化酶和 2 微克/微升 BSA 组成。用 D-丝氨酸在 20°C 测定 D-氨基酸氧化酶的酶活性,所用的底物浓度是 0 到 100 mM,混合物的其它圣爱尔莫光(corposant)和 pH 值为标准值。

观察到 g34872 和 D-氨基酸氧化酶混合物的 V_{max} 对应 $K_m = 4$ mM,观察到 D-氨基酸氧化酶和 BSA 混合物的 V_{max} 对应 $K_m = 4$ mM。这个结果显示, D-氨基酸氧化酶对其底物(D-丝氨酸)的亲和力没有改变,并且提示 g34872 与 D-氨基酸氧化酶相互作用的位点不是酶的活化位点。见图 5。

实施例 6

本发明的双等位标记

在 D-氨基酸氧化酶基因 (SEQ ID NO: 1) 内发现了有效多态性 (一般群体出现的频率大于 5%)。这些多态性也称为双等位标记, 以 SEQ ID NO: 23-26 表示, 分别编号为 24/1443-126、24/1457-52、27/93-181 和 24/1461-256, 其中多态性碱基位于位置 24。SEQ ID NO: 1 所示为检测本发明各个 D-氨基酸氧化酶双等位标记并含有扩增子和微量测序引物的多核苷酸。如图 6 所示, 已经确定标记 27-93/181 (SEQ ID NO: 25) 和 24-1461/256 (SEQ ID NO: 26) 分别与精神分裂症 $p=0.0066$ 和 0.0111 密切相关。本发明的标记还可用来判断个体是否有精神分裂症的风险, 如图 6 所示, 以及是否有其它相关的中枢神经系统障碍的风险, 优选是抑郁症和双相性精神障碍。

实施例 7

本发明化合物或组合物的制备

本发明的 D-氨基酸氧化酶和 D-天冬氨酸氧化酶拮抗剂组合物和化合物可以通过本领域技术人员公知的各种方法来制备。通用方案包括但不限于以下所述的那些方案。

以一般式 I、Ia、Ib 表示的化合物的制备

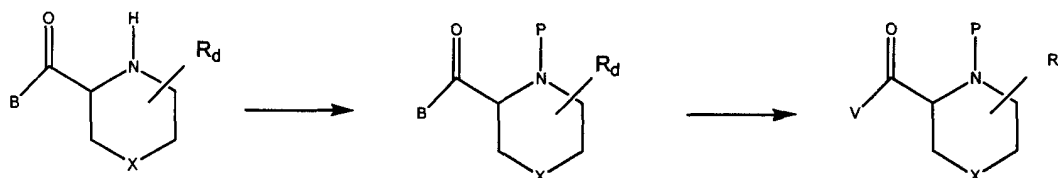
以一般式 I、Ia、Ib 表示的大量化合物可在市场上购买, 或用本领域技术人员所知的常规方法由市售化合物合成。具体地说, 利用亲电子取代反应, 如 Friedel Crafts 烷基化反应、芳基化反应或在浓硝酸中的硝基化反应, 可将取代基引入芳族环中, 如苯基、萘基或者取代的萘基或苯基。修饰芳族环的其它常用方法是将芳族基团转化为有机金属盐类, 如格林亚试剂 (Grignard reagents), 或者通过重氮芳基化合物引入取代基。这些操作方法和其它有关转化方法的例子在标准教科书中已有讨论, 如 March, Advanced Organic Chemistry (Wiley), Carey 和 Sundberg, Advanced Organic Chemistry (Vol 2.) 以及 Keeting, Heterocyclic Chemistry (全部 17 卷)。

以一般式 II 表示的化合物的制备

以一般式 II 表示的化合物可在市场上购买, 或由本领域技术人员用已知的合成技术轻易地合成。

在位置 Z 上被取代以一般式 IV 表示的化合物的制备

为了控制 R_1 , 本领域技术人员要明白可以选择在制备杂环之前、之后或同时来制备 R_1 。对于式中 A 是氮的化合物, 制备这些化合物的优选方法是:

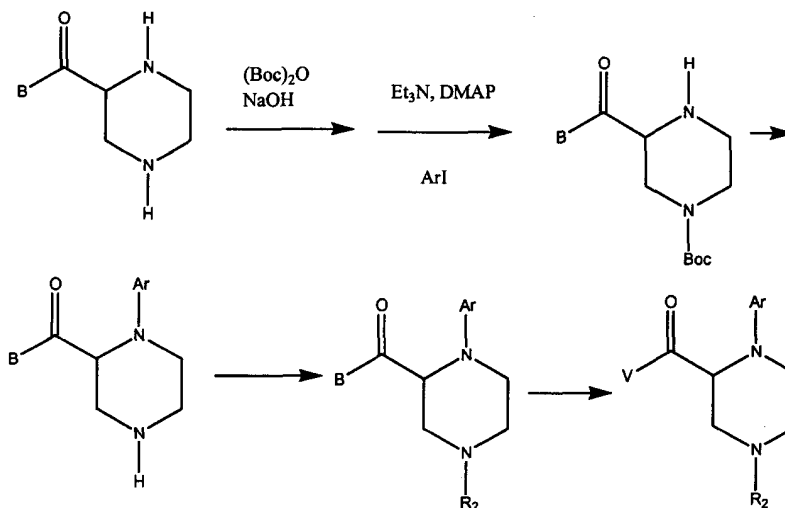


方案 4a

式中,

R_d 是衍生基团或者可被控制或取代, 这些化合物是已知的或可用已知方法制备。(P) 是保护基, 例如芳基, 而 (B) 是合适的阻断基团。为了清楚起见, 未示出一般式 IV 的位置 (Y) 上的基团。

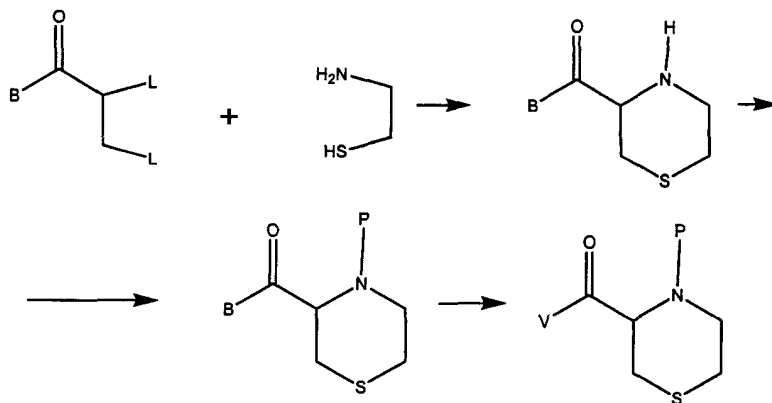
要制备和加工杂环, 本领域技术人员要明白可以选择在制备杂环之前、之后或同时来制备 R_1 。为了清楚起见, 未示出位置 Z 和 Y 上的取代基。对于式中 X 是氮的化合物, 下面示出 R_2 的一种优选控制方法。在以下方案中, L 是任何一个可接受的离去基团, B 是如上所述的阻断基团。Boc 是一个得到本领域公认的优选阻断基团 (blocking group)。本领域技术人员知道有机化学领域的技术人员能够选择阻断基团。



方案 4b

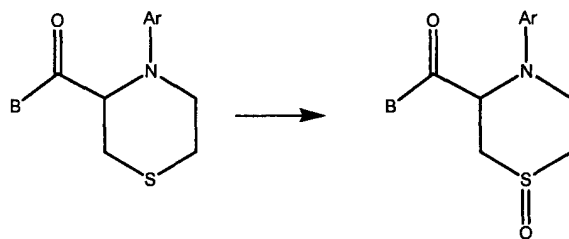
对于杂环含有硫的化合物, 下面示出这种杂环形成的优选方法。要制备和加工杂环, 本领域技术人员要明白可以选择在制备杂环之前、之后或同时来制备 R_1 。

为了清楚起见, 未示出位置 Z 和 Y 上的基团。



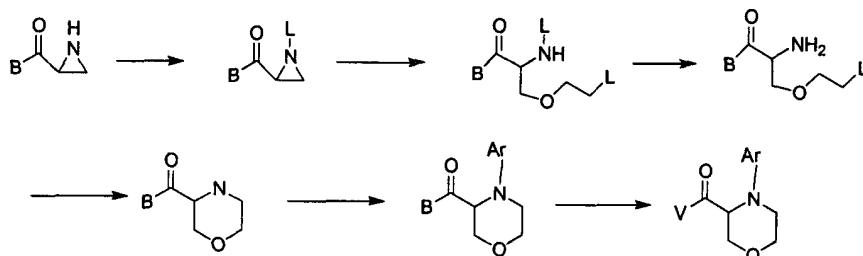
方案 4c

当 X 是硫时, 形成杂环后还可进一步对杂环进行加工。例如, 用已知的方法氧化环的硫原子, 可以得到相应的亚砷和砷, 如下图所示。



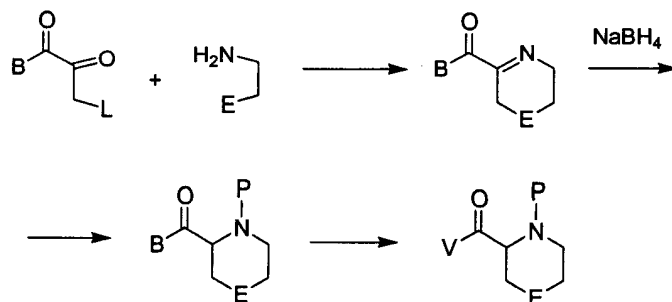
方案 4d

对于杂环含有氧的化合物, 下面示出这种杂环形成的优选方法。使具有双功能的部分, 例如卤代羟基物质与下面的氮丙啶反应。卤素部分作为离去基团, 用于环的闭合反应。杂环形成后, 按上述方法进行本发明的加工。



方案 4e

另一个可接受的制备本发明杂环, 其中 E 是硫、氮或氧, 的策略包括以下方案。这是一条优选的路径, 由此也可制备式中 A 是氮且 A-B 是不饱和的化合物。

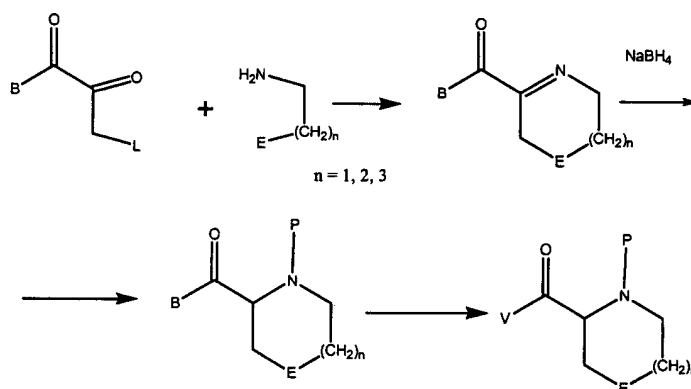


方案 4f

以一般式 IV 和一般式 III 表示的其它优选化合物的制备

当然，本领域技术人员知道，对于全部所述的基团方案 I 可适用于位置(Y)上的取代基。当 Z 是酮缩醇或硫代酮缩醇时，本发明的化合物可由环上带有羰基的化合物来制备。这些化合物由已知的方法制备，其中很多化合物是已知的或可在市场上购买。所以，本领域技术人员知道可以将羟基、氨基、亚氨基、烷氧基或其它基团转化为羰基化合物。

本领域技术人员也知道以一般式 IV 表示的化合物的上述合成路径可适用于以一般式 III 表示的化合物，其中环的大小是 7 和 8 个原子。符号 B、L、P 和 V 如上所定义。以下例子只起举例说明作用，但不具有任何限制性。

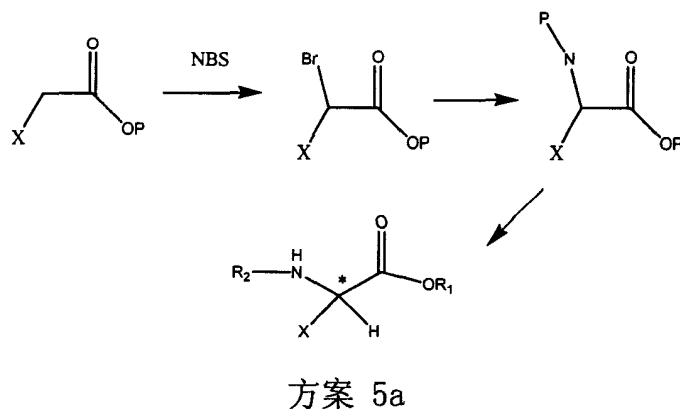


方案 4g

以一般式 Va 表示的化合物的合成

以一般式 Va 表示的化合物可用各种方法合成。可用于不同 α 氨基酸的公知最佳途径是 Strecker 合成路径。该方法用氨水和氢氰酸处理适当的乙醛，以便形成 α 氨基腈，然后再使 α 氨基腈进行水解反应。

以一般式 Va 表示的化合物的另一种可接受的合成策略是以下方案：

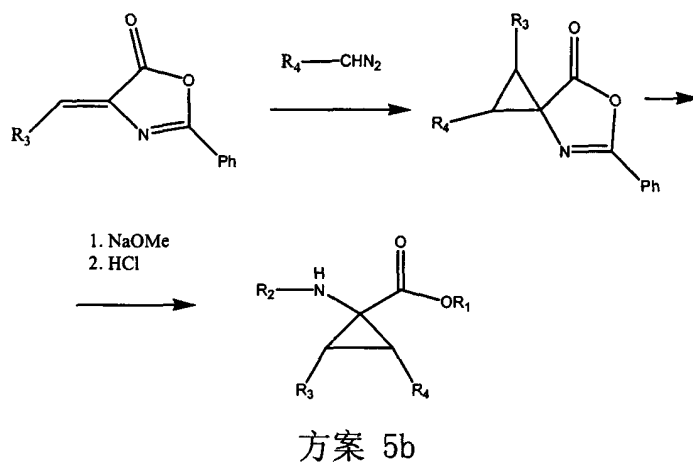


式中,

P 是保护基团, 如叔丁基, 可以与 R_1 相同。X 是如上所述的基团。受保护的化合物用卤化剂如 PBr_3 、NBS 或 CBr_4 进行溴化处理, 然后用氨或保护胺衍生物如苯邻二甲酰亚胺钾置换卤素。本领域技术人员可轻易引入 R_1 和 R_2 。

以一般式 Vb 表示其中 X 和 Y 含有环丙烷环的化合物的合成

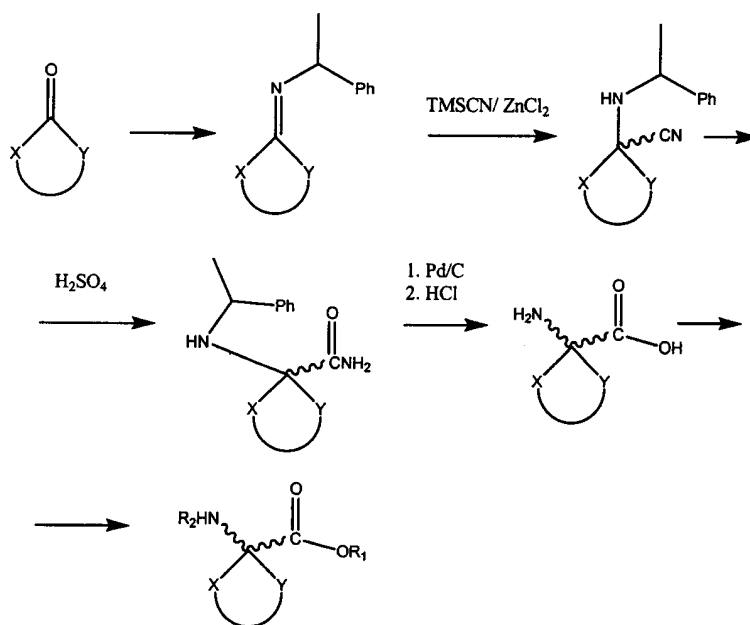
在构建 α 氨基酸的各种途径中, 普遍采用以 α, β -脱氢氨基酸衍生物来 1, 3-双极性环加成重氮烷烃。所以, 以下方案证实脱氢氨基酸的 R_3 取代基, 优选是烷基或 Ar^1 , 可以作为亚氨基酯得到保护, 其中 Ar^1 如上所定义。本领域技术人员知道这些化合物可以与重氮取代的化合物, 优选是烷基或 Ar^1 , 起反应生成希望得到的保护环丙烷衍生物。这些化合物与碱性醇类溶液如甲醇钠反应, 再进行酸性水解, 可生成相应的 R_3 、 R_4 取代的环丙烷氨基酸。利用已知的方还可以轻易得到 R_1 和 R_2 的其它衍生物。



以一般式 Vb 表示其中 X 和 Y 含有 5-8 元环的化合物的合成

可将含有优选 5、6、7、8 个原子的取代碳环或杂原子转化为与以一般式 Vb

表示的化合物一致的氨基酸衍生物。有一些已公认的途径, 其中一条途径是将含有化合物的环酮转化为相应的氨基酸衍生物。这些环酮化合物有很多已在文献中有所叙述, 本领域技术人员也可轻易合成。原料化合物可以是得到保护的或无保护的。向原料酮的亚胺衍生物加入三甲基硅氰化物可生成氰加成产物。水解和还原裂解保护胺产生氨基酸。利用已知的方还可以轻易得到 R_1 和 R_2 的其它衍生物。

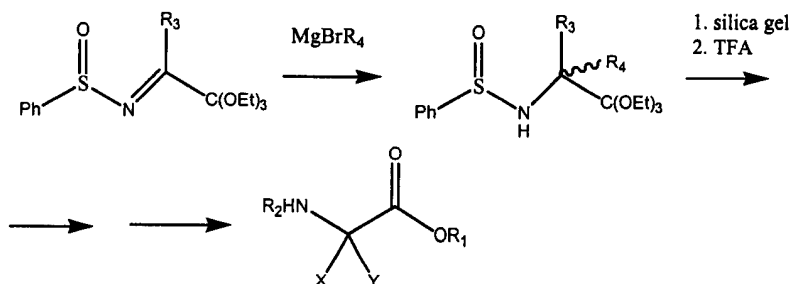


方案 5c

以一般式 Vc 表示的化合物的合成

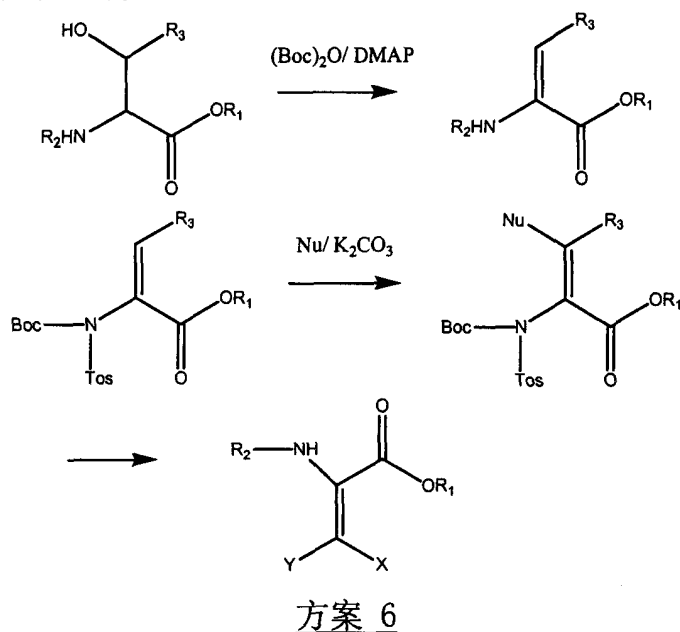
以一般式 Vc 表示的化合物可由 R_3 取代的化合物的亚磺酰亚胺衍生物合成, 其中 R_3 优选是烷基或芳基。有多种路径制备取代的亚磺酰亚胺的, 本领域技术人员可轻易合成这些化合物。

加入以有机金属试剂如烷基溴化镁形式存在的 R_4 , 再以三氟乙酸处理, 生成相应的二取代氨基酸, 它们用已知的方法可在 R_1 和 R_2 上作进一步衍生。



方案 5d

以一般式 VI 表示的化合物的合成



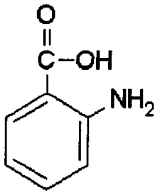
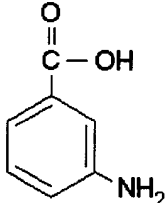
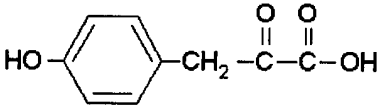
由取代的氨基醇可合成单取代或双取代的脱氢氨基衍生物。本领域技术人员采用与前述步骤相似的方法可轻易合成这些氨基醇。以 $(\text{Boc})_2\text{O}/\text{DMAP}$ 使单取代的氨基醇脱氢生成脱氢氨基衍生物。在碱存在下加入亲核试剂(Nu)生成双取代的脱氢氨基衍生物。

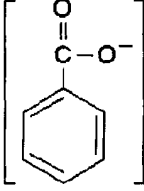
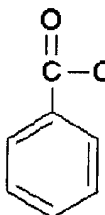
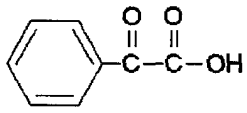
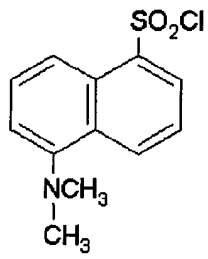
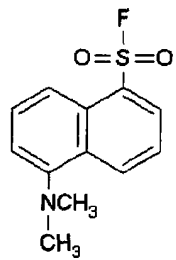
可以改变这些步骤以提高所希望的产物的产率。本领域技术人员也知道合理选择反应物、溶剂和温度是合成获得成功的一个重要组成部分。当确定最佳条件等是常规条件时,要明白的是,参照以上方案的指导可以类似的方式制备各种化合物。

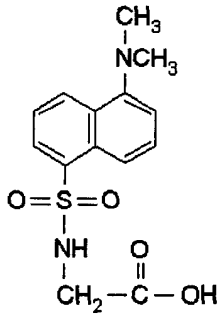
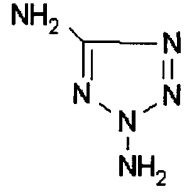
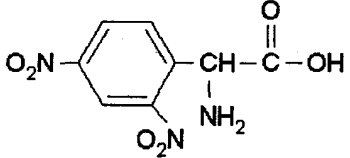
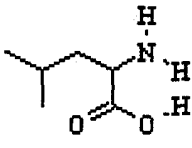
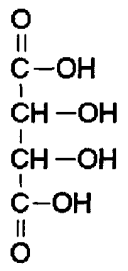
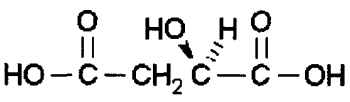
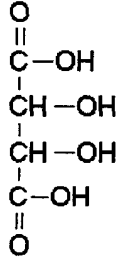
业已公认,有机化学领域的技术人员无需进一步指导就可轻易进行有机化合物的标准操作,也就是说,技术人员清楚知道这些操作步骤的实施。这些操作包括但不限于:羰基化合物还原为它们相应的醇、羟基等的氧化、酰化反应、芳族取代、亲电子和亲核反应、醚化反应、酯化反应和皂化反应等等。这些操作方法的例子在标准教科书中已有讨论,如 March, Advanced Organic Chemistry(Wiley), Carey 和 Sundberg, Advanced Organic Chemistry(Vol 2.) 以及 Keeting, Heterocyclic Chemistry (全部 17 卷)。

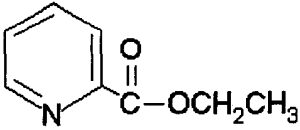
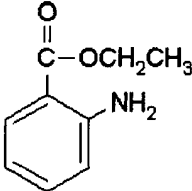
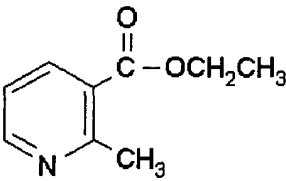
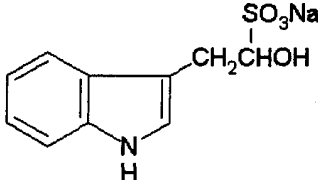
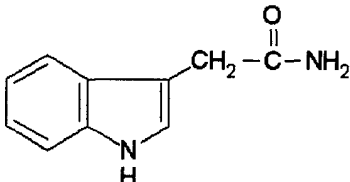
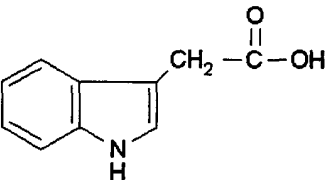
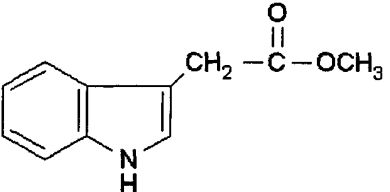
本领域技术人员容易地知道,某些反应最好在分子的其它官能团被遮蔽或被保护的时候进行,以避免任何不希望看到的副反应和/或提高反应的产率。这些反应可在文献中找到,而且已为本领域技术人员所知。这些操作方法的许多例子可以在 T. Greene 的 Protecting Groups in Organic Synthesis 中找到。当然,最好阻断用作原料的氨基酸与侧链反应,以防止不希望的副反应。

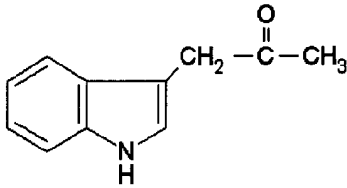
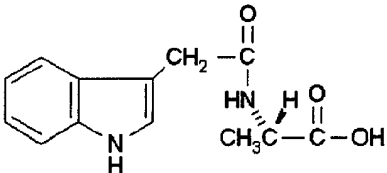
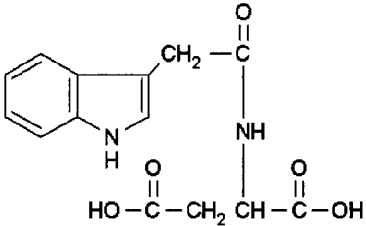
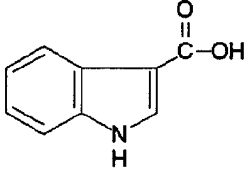
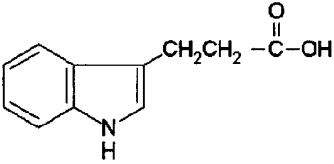
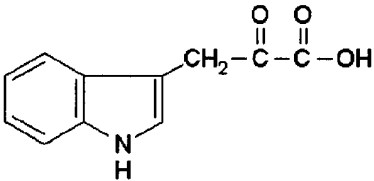
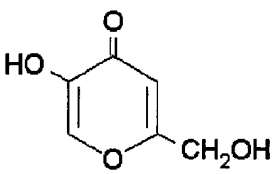
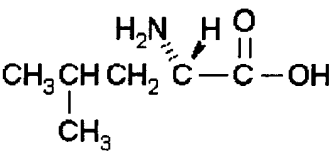
表 I

名称	结构式
2-氨基苯甲酸盐	
2-羟基丁酸盐	$\text{CH}_3 \text{CH}_2 \underset{\text{OH}}{\text{CH}} - \overset{\text{O}}{\underset{\text{ }}{\text{C}}} - \text{ONa}$
2-羟基丁酸盐	$\text{CH}_3 \text{CH}_2 \underset{\text{OH}}{\text{CH}} - \overset{\text{O}}{\underset{\text{ }}{\text{C}}} - \text{ONa}$
3-氨基苯甲酸盐	
3-羟基丁酸盐	$\text{CH}_3 \underset{\text{OH}}{\text{CH}} \text{CH}_2 - \overset{\text{O}}{\underset{\text{ }}{\text{C}}} - \text{ONa}$
4-羟基苯基丙酮酸盐	
丙酮二羧酸盐	$\text{HO} - \overset{\text{O}}{\underset{\text{ }}{\text{C}}} - \text{CH}_2 - \overset{\text{O}}{\underset{\text{ }}{\text{C}}} - \text{CH}_2 - \overset{\text{O}}{\underset{\text{ }}{\text{C}}} - \text{OH}$
氨基胍二羧酸盐	$\text{H}_2\text{NNH} - \overset{\text{NH}}{\underset{\text{ }}{\text{C}}} - \text{NH}_2 \quad \cdot \text{H}_2\text{CO}_3$
氨基胍盐酸盐	$\text{H}_2\text{NNH} - \overset{\text{NH}}{\underset{\text{ }}{\text{C}}} - \text{NH}_2 \quad \cdot \text{HCl}$
氨基胍半硫酸盐	$\text{H}_2\text{NNH} - \overset{\text{NH}}{\underset{\text{ }}{\text{C}}} - \text{NH}_2 \quad \cdot \frac{1}{2}\text{H}_2\text{SO}_4$
氨基胍硝酸盐	$\text{H}_2\text{NNH} - \overset{\text{NH}}{\underset{\text{ }}{\text{C}}} - \text{NH}_2 \quad \cdot \text{HNO}_3$

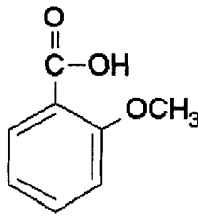
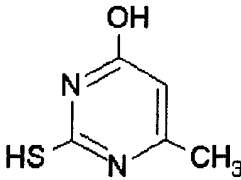
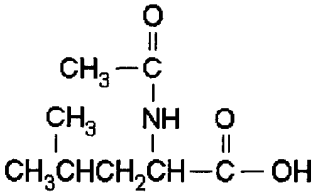
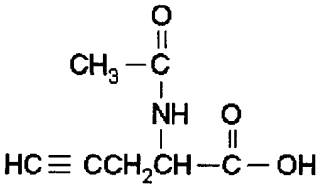
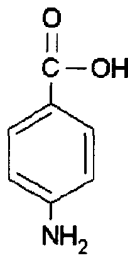
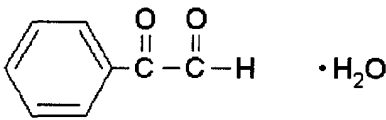
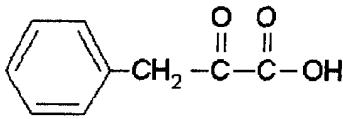
氨基胍硫酸盐	$\text{NH}_2\text{NH}-\overset{\text{NH}}{\underset{\text{ }}{\text{C}}}-\text{NH}_2 \quad \cdot \text{H}_2\text{SO}_4$ $\cdot \text{H}_2\text{O}$
苯甲酸铜盐	 $\left[\text{C}_6\text{H}_5\text{COO}^- \right]_2 \quad \text{Cu}^{2+}$ $\cdot 2\text{H}_2\text{O}$
苯甲酸	
苯甲酰甲酸	
半胱胺	$\text{H}_2\text{NCH}_2\text{CH}_2\text{SH}$
丹酰基氯化物	
丹酰基氟化物	

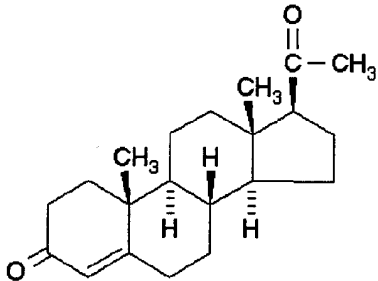
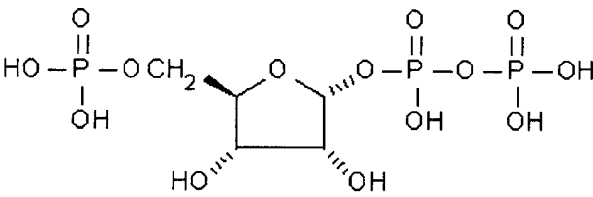
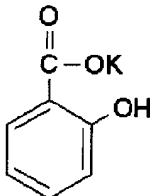
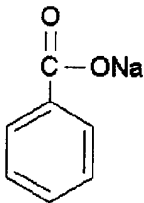
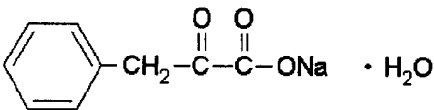
丹酰基氯化物甘氨酸	
二胺四唑	
二硝基苯基甘氨酸	
D-亮氨酸	
DL-酒石酸盐	
D-苹果酸	
D-酒石酸盐	

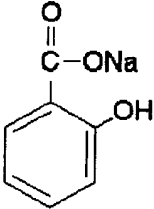
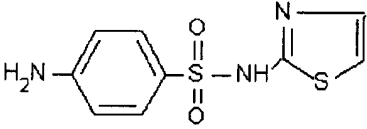
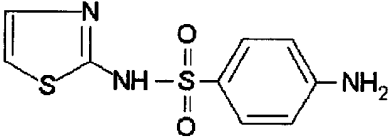
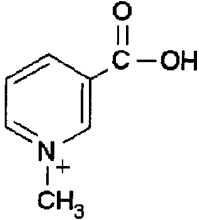
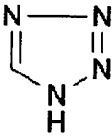
2-吡啶甲酸乙酯	
氨基苯甲酸乙酯	
溴化丙酮酸乙酯	$\text{BrCH}_2-\overset{\text{O}}{\parallel}{\text{C}}-\overset{\text{O}}{\parallel}{\text{C}}-\text{OCH}_2\text{CH}_3$
甲基烟酸乙酯	
羟胺盐酸盐	$\text{H}_2\text{NOH} \cdot \text{HCl}$
吲哚-3-乙醛亚硫酸氢钠	
吲哚-3-乙酰胺	
吲哚-3-乙酸盐	
吲哚-3-乙酸甲酯	

吲哚-3-丙酮	
吲哚-3-乙酰基丙氨酸	
吲哚-3-乙酰基 ASP	
吲哚-3-羧酸盐	
吲哚-3-丙酸盐	
吲哚-3-丙酮酸盐	
曲酸	
L-亮氨酸	

L-亮氨酸甲酯盐酸盐	$\begin{array}{c} \text{H}_2\text{N} \quad \text{H} \quad \text{O} \\ \quad \quad \\ \text{CH}_3\text{CHCH}_2\text{C}-\text{C}-\text{OCH}_3 \quad \cdot \text{HCl} \\ \\ \text{CH}_3 \end{array}$
L-苹果酸	$\begin{array}{c} \text{O} \quad \text{HO} \quad \text{H} \quad \text{O} \\ \quad \quad \quad \\ \text{HO}-\text{C}-\text{CH}_2-\text{C}-\text{C}-\text{OH} \end{array}$
苯甲酸镁	$\left[\begin{array}{c} \text{O} \\ \\ \text{C}-\text{O}^- \\ \\ \text{C}_6\text{H}_5 \end{array} \right]_2 \text{Mg}^{2+}$
甲硫咪唑	$\begin{array}{c} \text{N} \\ // \\ \text{C} \\ \quad \\ \text{N} \quad \text{SH} \\ \\ \text{CH}_3 \end{array}$
4-硝基苯磺酸甲酯	$\begin{array}{c} \text{O}_2\text{N}-\text{C}_6\text{H}_4-\text{S}(=\text{O})_2-\text{OCH}_3 \end{array}$
氨基苯甲酸甲酯	$\begin{array}{c} \text{O} \\ \\ \text{C}-\text{OCH}_3 \\ \\ \text{C}_6\text{H}_4-\text{NH}_2 \end{array}$
苯甲基硫氰酸甲酯	$\begin{array}{c} \text{CH}_2\text{SCN} \\ \\ \text{C}_6\text{H}_4-\text{CH}_3 \end{array}$
甲基烟酸甲酯	$\begin{array}{c} \text{O} \\ \\ \text{C}-\text{OCH}_3 \\ \\ \text{C}_5\text{H}_4\text{N}-\text{CH}_3 \end{array}$
丙酮酸甲酯	$\begin{array}{c} \text{O} \quad \text{O} \\ \quad \\ \text{CH}_3-\text{C}-\text{C}-\text{OCH}_3 \end{array}$

水杨酸甲酯	
甲基尿嘧啶	
MGAG 二盐酸盐	$\text{H}_2\text{N}-\overset{\text{NH}}{\underset{\text{ }}{\text{C}}}-\text{NH}-\text{N}=\overset{\text{CH}_3}{\underset{ }{\text{C}}}-\text{CH}=\text{N}-\text{NH}-\overset{\text{NH}}{\underset{\text{ }}{\text{C}}}-\text{NH}_2 \quad \cdot 2\text{HCl}$ $\cdot x\text{H}_2\text{O}$
N-乙酰基-D-亮氨酸	
N-乙酰基-D-L-炔丙基甘氨酸	
硝基乙烷	$\text{CH}_3\text{CH}_2\text{NO}_2$
对氨基苯甲酸盐	
苯基乙二醛一水合物	
苯基丙酮酸盐	

酒石酸钾	$\begin{array}{ccccccc} & \text{O} & \text{OH} & \text{OH} & \text{O} & & \\ & & & & & & \\ \text{HO} & -\text{C} & -\text{CH} & -\text{CH} & -\text{C} & -\text{OK} & \end{array}$
孕酮	
炔丙基甘氨酸	$\begin{array}{ccccccc} & & & \text{NH}_2 & \text{O} & & \\ & & & & & & \\ \text{HC} \equiv & \text{C} & \text{CH}_2 & -\text{CH} & -\text{C} & -\text{OH} & \end{array}$
丙酮酸盐	$\begin{array}{ccccc} & \text{O} & & \text{O} & \\ & & & & \\ \text{CH}_3 & -\text{C} & - & \text{C} & -\text{OH} \end{array}$
RPP	
水杨酸盐	
苯甲酸钠	
α-酮异戊酸钠	$\begin{array}{ccccccc} & \text{CH}_3 & \text{O} & \text{O} & & & \\ & & & & & & \\ \text{CH}_3 & -\text{CH} & -\text{C} & -\text{C} & -\text{ONa} & & \end{array}$
苯基丙酮酸钠	

丙酮酸钠	$\text{CH}_3-\overset{\text{O}}{\parallel}\text{C}-\overset{\text{O}}{\parallel}\text{C}-\text{ONa}$
水杨酸钠	
磺胺噻唑钠	 (Na)
磺胺噻唑	
氨基硫脲	$\text{H}_2\text{NNH}-\overset{\text{S}}{\parallel}\text{C}-\text{NH}_2$
硫脲	$\text{H}_2\text{N}-\overset{\text{S}}{\parallel}\text{C}-\text{NH}_2$
葫芦巴碱	 Cl⁻
未取代的四唑	

<110> 根瑟特公司 (GENSET)

<120> 用D-氨基酸氧化酶和D-天冬氨酸氧化酶拮抗剂治疗中枢神经系统障碍

<130> 101.W01

<150> 60/261,883

<151> 2001-01-16

<150> 60/305,445

<151> 2001-07-13

<150> 60/

<151> 2001-10-22

<150> 60/333,881

<151> 2001-11-19

<160> 26

<170> Patent.pm

<210> 1

<211> 86592

<212> DNA

<213> 智人 (Homo sapiens)

<220>

<221> misc_feature

<222> 38388..40388

<223> 5' 调控区

<220>

<221> exon

<222> 40389..40670

<223> 外显子 Z

<220>

<221> exon

<222> 42666..42778

<223> 外显子 A

<220>

<221> exon

<222> 43416..43519

<223> 外显子 B

<220>

<221> exon

<222> 61159..61402

<223> 外显子 C

<220>
<221> exon
<222> 64050..64711
<223> 外显子 Ulong

<220>
<221> exon
<222> 64578..64711
<223> 外显子 U

<220>
<221> exon
<222> 68126..68261
<223> 外显子 V

<220>
<221> exon
<222> 69488..69690
<223> 外显子 2

<220>
<221> exon
<222> 71942..72056
<223> 外显子 3

<220>
<221> exon
<222> 73962..74038
<223> 外显子 4

<220>
<221> exon
<222> 74701..74766
<223> 外显子 5

<220>
<221> exon
<222> 77478..77532
<223> 外显子 6

<220>
<221> exon
<222> 78762..78866
<223> 外显子 7

<220>
<221> exon
<222> 81507..81589
<223> 外显子 8

<220>
<221> exon
<222> 83181..83298
<223> 外显子 9

<220>
<221> exon
<222> 83879..83977
<223> 外显子 10

<220>
<221> exon
<222> 84906..85423
<223> 外显子 11

<220>
<221> exon
<222> 84906..85541
<223> 外显子11长序列

<220>
<221> misc_feature
<222> 85542..86592
<223> 3' 调控区

<220>
<221> allele
<222> 6199
<223> 24-1443-126:多态性碱基C或T

<220>
<221> allele
<222> 44814
<223> 24-1457-52:多态性碱基A或C

<220>
<221> allele
<222> 72600
<223> 27-93-181:多态性碱基A或G

<220>
<221> allele
<222> 84046
<223> 24-1461-256:多态性碱基A或G

<220>
<221> primer_bind
<222> 6074..6094
<223> 24-1443.pu

<220>
<221> primer_bind
<222> 6559..6579
<223> 24-1443.rp 补体

<220>
<221> primer_bind
<222> 44763..44782
<223> 24-1457.pu

<220>
<221> primer_bind
<222> 45257..45277
<223> 24-1457.rp 补体

<220>
<221> primer_bind
<222> 72359..72378
<223> 27-93.rp

<220>
<221> primer_bind
<222> 72763..72780
<223> 27-93.pu 补体

<220>
<221> primer_bind
<222> 83791..83811
<223> 24-1461.pu

<220>
<221> primer_bind
<222> 84307..84324
<223> 24-1461.rp 补体

<220>
<221> primer_bind
<222> 6180..6198
<223> 24-1443-126.mis

<220>
<221> primer_bind
<222> 6200..6218
<223> 24-1443-126.mis补体

<220>
<221> primer_bind
<222> 44795..44813
<223> 24-1457-52.mis

<220>
<221> primer_bind
<222> 44815..44833
<223> 24-1457-52.mis补体

<220>
<221> primer_bind
<222> 72581..72599
<223> 27-93-181.mis

<220>
<221> primer_bind
<222> 72601..72619
<223> 27-93-181.mis补体

<220>
 <221> primer_bind
 <222> 84027..84045
 <223> 24-1461-256.mis

 <220>
 <221> primer_bind
 <222> 84047..84065
 <223> 24-1461-256.mis补体

<220>
 <221> misc_binding
 <222> 6187..6211
 <223> 24-1443-126. 探针

<220>
 <221> misc_binding
 <222> 44802..44826
 <223> 24-1457-52. 探针

<220>
 <221> misc_binding
 <222> 72588..72612
 <223> 27-93-181. 探针

<220>
 <221> misc_binding
 <222> 84034..84058
 <223> 24-1461-256. 探针

<400> 1
 attattggaa caggccacac ttgcgaggga agtccctgcc tcagaaagat tcagaaaagc 60
 tagacagtca ctggaagaac aattacaacc gcaagacggt caaacactaa acaccgctat 120
 gcctcagaac cgtacagata atggccaaat agatggggct ctgggcattt ctgagagcac 180
 ctgcctgggt gcaccccatc ctaatggacc atgccctcca gtctccaagt ggctcttcag 240
 agctcacatc cgaacacctc ctatgttaca ggttcttcta gccccaggtt cccaaccacc 300
 ccaaggccac agaggccagc cccaactcca tcttctacat gtgtcacagg aaactttctc 360
 atagtctat ttattatgta ctgcgggggt gggggccatg tcataaaaaga aatgtcctcc 420
 cttttttatt catctccttc taacaagcat caaagtctca gtcgctagca tgtgacttac 480
 agaagctctc atgggaacaa gacaagacca tactgttacc gtgacactca cggcctccct 540
 gactggtttc tgctgttgat tctgectcaa atgctcctca aatgcacctt gctgctccgc 600
 ctccacccta gagctgcctt gactgcccac ttgcccgta agagtcggct taggcttcac 660
 tcctgccaga aaggctcctg cagggtgctc caacagtcac cccctcctgt ggtctcacia 720
 aacccagaca cctctcggtc actctctccc tcctatctgg ttgtgactgt cttccatgct 780
 cacttagaag ctctctgagg ccaagaactg tgtgtactgt tgcttctttg tttacctggg 840
 cctagcccat tgccctcaca acaggagaat gcaaataaat catatgctta atgaatgagt 900
 cgatgaatga atgatgaata aagggaatct aatctagttt taacaaatcc aggttttgca 960
 atgatctcac aggcattcat ttatcttggt atgtcagggg agtgactcca cctcatttc 1020
 acacgcattc tgggggtcaat gctctaactt acttgccctc cagttagtgg gaaattacaa 1080
 gctacacttc aagcctctga ctaggacctg ccatgaagta cttgggaatc agtggagtat 1140
 cactgtgggg tgagggtgtc gaggcgaggc ccaccaatct ccatacttct ccccgggccc 1200
 ctctgcctga gaggtctctc ctgcttccct tggcagactc tggtttgccc ttctgggttc 1260
 ggcgttggtg tcacctcctt caggaagcat ttctggctaa ggtgcccac tctatagcag 1320
 ctggtgtaaa acctctctaa gcaaacagca taactttctg tcctctcaat tgactctgag 1380

ttctgagagc	acagcctgga	gctggcacgg	tgcctggcac	agagagctga	aatggcacac	1440
cctagtgttc	ccagtggctc	gactccccag	gtctccatc	aggacgcagc	cctctcccac	1500
ctctgatgga	tatgggacca	tggaatgctt	tgccagcag	caactcttgc	ctccctcaca	1560
gaagggaaca	cctagcccat	cagactcacc	tttcttact	ggaaaagtcc	actcccagca	1620
agatattctc	ctcgggtgtc	tggcgcccgc	tgctgtacac	caccaccatg	taccggaccc	1680
ggtccgcccc	ggcgctctcc	aggcgccactg	cctggaacag	ggcagacatg	ctctcactaa	1740
cctgcctttg	gaggtgggtc	ctccctccca	tctccaatgc	aagatcaaca	ctttcagtgt	1800
tctacctttc	cctctgggag	ttaaaaatga	agagaaaatt	cttggctggg	catgggtggt	1860
caggcctgta	atcccagcac	tttgggaggc	caaggtaggc	agatcacttg	aggtcaggag	1920
ttcgagacca	gcctggccaa	catggcaaaa	ccccatctct	tactaaaaat	acaaaaatta	1980
gctgggcatg	gtggcgctgc	cctgtaatcc	cagctactca	ggggactgag	gcacgagaat	2040
ctcttgaaac	cgggaagcgg	aggttgagct	gagctgagat	catgccacca	catgccagcc	2100
agagcgacag	agtgagattc	tgctcacaac	aacaacaaca	caacaaaaca	caaagcggaa	2160
gttcttgaca	gcaggaacca	ggcctcggtt	ctctctgtag	caccagggac	gccgcctggc	2220
tcagaggaat	caccccaaat	gcaagaaatc	agtgaacaca	tgaatccaa	agaaagtctg	2280
tatttagctt	atttaactgc	cgtggagacc	tgtttcatcc	ctctctccgc	ccctctgggg	2340
aactgaggag	tcaacctggc	tttggcttta	gtgcacaatt	tgagaatttg	ttgtaaccta	2400
aaagcttttc	cccttatcat	tcacgaatgg	ttccccacca	ggtttcacaa	ttaaaaatta	2460
aaacttgctg	gctgggcacg	gtggccttaca	cctataactc	caggactttg	ggaggcagag	2520
gcaggagaat	catttgaggc	caggagtcca	agaccagcct	gggcaaaata	gcaagacccc	2580
atatccacaa	aattttttta	aaaataaggc	agggtggtac	acacttgtag	tccagctat	2640
ccaggaggct	gaggtgggaa	gattgcttga	acccaggagt	ttgaggctgt	agtgagctaa	2700
gatcatgccca	ctgcactcca	gcctgggcaa	cagagcaaga	ccctcatctc	acaaaaatta	2760
aaaaaaaaatt	ttttaacttg	acattctcac	tgcttcttac	cagcttgatt	ctgtcttcgc	2820
aacgcagaag	gttgatcatc	acctgaagat	gttgaggcag	atcacctgtt	ggaccaataa	2880
agaaagcttt	aaaaggtctc	ttacctactc	tctaggaaaa	aaaacctctg	aaaggctgac	2940
tttgagggtc	tggaaaaaga	ttgagaagtt	aaaatttgtc	tacctacacc	acaggagaat	3000
caccacaaaa	acttcaagtc	tgaatttctc	ttacaccact	ctgaatactg	tgcgacgtgg	3060
atgggtgaca	tggagcttac	tgctcatgtg	ttaaaagttg	ctcttatttc	ctgaaataca	3120
tacagtatag	gtttccaaat	acaaaatgtg	aaaaatacag	gcaagcctag	agaaaaatgt	3180
tatttcatc	aagccaatgt	tactcggcag	gttgggggtc	ctagaaacga	cagctgtggc	3240
tggaaagtaag	gcatttgcta	agagttaatc	attagagaaa	aaggacagag	catcacgttt	3300
ctcttcaaaa	caacttcttc	ttctatacag	agtctcgcac	tgtcacccag	gctggagtgc	3360
agtggtacga	tctcagctca	ctgcaacctc	cgcctcctgg	attcaacaga	ttcttctgct	3420
tcagcctcct	aagtagctgg	gattacaggt	gcccataccc	agaccggtct	aatttttgta	3480
tttatagtag	agatggagtt	tcaccatgtt	ggccaggctg	gtctcgaact	cctgacctca	3540
agtgatctgc	acgcctcagc	ctcccaaagt	gctggggatta	caggcataag	ctaccccacc	3600
caggccccac	ttcaaaactc	tgcatitttc	actggaggca	gacattattt	ccataaccgg	3660
gggggcgggg	ggaatgtttt	aagtgactct	acagatagca	gctgtatgct	ggttgcccag	3720
agaaataatt	tgaatagaaa	ccaatctgtc	attttctctt	ttcttgctaa	aaattatgta	3780
ctcttttttc	ttcactatgt	aaaacaggca	gtaaccaggg	acggcttctg	aacttctctg	3840
agctgcccc	gggttcagga	ggtgttcctg	gagtgcagtg	aggaaagtct	cttactggcc	3900
atgagctctg	cgcgaagcag	agaccctgtc	agaagaagcg	cacactttca	cggaggggaa	3960
agttgtaagg	gaggtgcata	attagtaagt	agcaggtgtg	actccaaggt	tgcttttttt	4020
ctctagctta	cacatttttc	ttatatctg	caaggatttc	tttctgaaga	aagggtcatc	4080
tgtagagatg	ctaataatcag	cctgggtgtg	tggtcacac	ctgtaatccc	agtgccttgg	4140
gaggccgagg	caggagactc	acttgaggcc	aggcattcaa	gaccagctcg	ggcaacatgg	4200
caagacccca	tctctacaga	aaagtaaaaa	attagctggg	ctttctggtt	cacatctgta	4260
gtcccagcta	cttggggaggc	caaggcagga	ggatcgctgg	agcccaggag	tttgagatca	4320
ccctgggcaa	cacgataaga	ccctgtctct	acaaaggaaa	aaaaattact	ctatacatca	4380
caattacaac	cccaaaaagga	tcaataatgc	ttacacactc	aaatgctcca	aaaggaaaca	4440
ttgtgtttgt	tccttttgca	aaagcatctt	tttcatttta	agggagaagg	acagatgatg	4500
tccaaattgc	acttctgtc	tcagagagga	attgggtcat	tagaaaattg	tgcctctagc	4560
caggagggta	gatctcatgt	taagcgttct	ttctttttct	tttttttcca	atagagacag	4620
ggttttgcc	tggtggccag	gctggtctcg	aactcctgga	ctcaaatgat	cctctcacct	4680

```

cagcctccca aagtgtctggg attataggca tgagccacca aacccaggcc aattaagcat 4740
tctttccaca ataagtaaaa tttaaaaaag aaaagaacca tgccctcttt atctgtcttc 4800
tccagttata caattccaca gtgtataaca cctgtgttg accctgtctt ctatgatgag 4860
cgatttggag ataagggttc acattaaaga aagccataga cctccccagc cccttctctc 4920
acccgtcatg tcaccaatgc aacacaacga caacgaccat gagctgtgtc ttcacctgcc 4980
tgggcccctc caccatctac ccgagtcaca gaactgcatt ggggaaagca aaaacaaacc 5040
cctgtctgat aaatgcctaa atgaaaggga catttccac acagataaac ttctttcagt 5100
gggattgttt gctgagatat ggaactgtcg acagacagaa atccaaacc cagtctgaca 5160
tccacacaca aaaaaatcag agaataaag ccctagaaag ggtctcaa at tgactggact 5220
ggctgaaaca aactgaacta cttttccaag gacagaatta accctcaatt gtactcagct 5280
ctgcacagtg gttactgggg ggccctctgt acattcagga gacttgatgg taattctagg 5340
gaaaaaaagg aactaacgta agtctagtct gcgtctgtcc caaggtcatt tacagaccaa 5400
ctgtggacag ctggcgggcc cttgtccttc cgacctcctc gtccactcca gacctcagg 5460
cacaagagtc agccagctgg tggcttgcat cctacccttc tagtctttgg attagaggaa 5520
ggaggtatct gacacttagt gagcagagct tgagcatitt ctttgtcata tgtgttacia 5580
ttaaaacatg aacaacagct acatttctaa gagggcagaa taattagcaa attcaagaac 5640
gaagaatctg gctgagtagt gcacctcaaa cctataatcc caatgctttg agaggctgag 5700
gtggggagat ggcttgaggc caggagtggg agaccagcct gagcaacata gtgagtgaga 5760
cctcatctac acacacacac acacacacac acacaactag ctgggtgtgg tgacacgcct 5820
ctatggctcc agctactcag gaagctaagg ctggacaate acttgagccc aggaggtcga 5880
ggctgcagtg agctatgac aggccactgc acgccagcct gggcaagaga atgagatccg 5940
tctctaaaaa aacttttcat ataattaaaa aaaaaaaag aatgaagggt ctgtttatag 6000
ctgtattgta ctagaagtca tcgtaataac aatgatagtt acccatatat atatacagca 6060
cctactacag gtaggtatgt tacacgcata actctaaatt tccatattgt ctgaggcacc 6120
agtatttgat gccattgta aagactagga aactgaggct tagaagtcca cctgttacgg 6180
cttagtaagt tggagaacya ggatcagaag acaggctctc ctggcttcaa aacaaatact 6240
atttccacaa accacactgc ctcctgttac aggacagtta ttttctttgc ttaaacaga 6300
cctaaatatt atcaacatca gtatgtgaaa atactgactg agccttgggtg tttgtataa 6360
attgcatggt gtagaattct aacctgagca ctcagatcta aaatgaagct gaatgacttg 6420
aggttaaaca aacaaaatgt tcacaagaaa actggccaca atagctggtt ggtttcacct 6480
gctgtgttgc tgaagggtaa aggccttctc agtcacaga cattcaatta tgcactgcct 6540
ctccaagaaa tgccttgaga tgtgtccac ctacgacaaa gatccactta catgcaagca 6600
cttttctctc tttctttctt tttgagatag ggtccttttc tttgtcacc caggctggag 6660
tgcagtggcg caatcgtggc tcaactgagca acacagttag caacatagtg agacctcatt 6720
tacatacaca cccacaaaaa actagctggc tgtgtgaca catcagcctc gacctcctgg 6780
gctcaagcaa tcctccccc accctgctgg gattacagcg atgcgccacc 6840
acgcccagct aatttttgta tttttgtag agatagggtt tcaactgtgtt gttcgggtta 6900
gtctggaact cctgggttca agcgagatct gccaccttg gcctcccaaa tcctggaatt 6960
acaggcaaga gccaccgtgc ctggccataa gtgtgttttg ttgttattgt ttttaagaaa 7020
cagagtctct cttgttcacc caggctggag tgcagtggcg tgatcctagc ttgctgcagc 7080
ctcaaactcc tgggctcaag cgatcctccc aactcagcct cccaaagcac tgggattaca 7140
ggtgtgagat accatgcagg gccacgcaag catttcttga attctcttt ctaactgcct 7200
tcagctctga gtcaagtctc ctaagaaaac cagtcttact acttagtagg cacttcttat 7260
ttaaactcag tttgatctc accctattac ttctgtctac ttcttaaaaa caaactatta 7320
cagaatcaag acttctact acagtgtcta tctcagagtt ggagccaaag gcccttcaag 7380
aaattctcca aatgagtgtt ttccaatgc ttggagaaat ccatcccaag attaggtata 7440
cagcactcca gatggttatt ttcaagtga cgacatctgg ctataattca ttttggtgca 7500
tttgttaaaa agtcaggctg taacttacag cctgcaatta actgataaac tacagagagg 7560
aaatctttgc atcccagcag gatgtgtctg acctactcc tgacgcagac agacatgaca 7620
taaaagggtg gaaaatgtgc gtgtgtctgt caagagagag catctgagcc tctgcctgca 7680
ctggtcactg caaacctgcg tccactatgt ctaaggcctt caaactcagc aacatcacca 7740
acaatggaag tttctctgc tgtccagaaa agaagctcca atgtaagagt atcaacttag 7800
agccctcacc tgcagtcttg tgggggtgct gaagactccg ctggccttga gggctgcttc 7860
cctgtttaa gaagagggt gcgcctttca ccatgaaaaa gctctcactt aagctgggaa 7920
ggataagacc agagcacagt tagaccggaa ttcagacagg aaaatggaca aagaattact 7980

```

```

gcaggggaaa aagcttttagc gtggacaaat ggcatgtaaa atgcaaatag gatgaaactg 8040
cttttataat aattccacgt agtacttttc tcaaaccttg cttttgctaa aagcttgctg 8100
ctggagaatt ttcgtgacaa aataatgctt ctgtgacaac acccaaaagt ctacataggc 8160
tctccagggc ccccttctgc agaatactgg acagggatct cactgtcata taacattttc 8220
ttctttcttt ttttttgag acgcagtttc actctgtcat ccaggctgga gtacagtggg 8280
gtgatctcaa ctactgcaa cctctgcctc ctgggttcaa gcgattctca tgtctcagcc 8340
tccccagaag ctgagattac aggcattgtc caccatggcc agctaattat tgtattatta 8400
gtagagacat ggttttacca tgttgccag gctggctca aactcctgac ctccagcaat 8460
ctgcctgctt aggcctcctg gagtgtctg attacaggcg tgaccacgcc cagccataac 8520
attttctaag aaaagagAAC aactccctga ttaggagagg gcagtctact ttgtgaattc 8580
tcatgctctt gctgttgatc tctgcttcta actctctggc ttttaacaac tccattgttt 8640
cttggtgact tcccttgatg gaatacaagg atgaattac actttcacta gttgtttgca 8700
ttttaagaaa agtggggagg gccgggtgg ctcaagcctg taatcccagc attttgggag 8760
gccaaggcag tggatcactt gaggtcagga gttcgagacc agtctggcca acatggtgaa 8820
acctgtctc tactaaaaat gcaaaaatta gccaggcgtg gtggcacatg cctgtaatcc 8880
cagctactca gaaggctgag gcacaagaat cgcttacttg agccccaggc acggaggttg 8940
gagtgagcca agatgcacc acgcaccact gcactccagc ctggcgaca gagcaagact 9000
gtgtctcaaa aaaaaaaaaa gaaagaaaga aagaagaaa aaagtgggtg gatactgact 9060
tgtgatttaa cttagtcaag gtgtcctgt ccactattct tgaggaaaac ctcaagtgg 9120
cccaatgaat ttctcagcag aatgaatctt tggcctttgt tattttagct agcaataaca 9180
tttataacta cctataactt taaaaattac aattaaaaaa tgtttatttg ggaggctggg 9240
gtggaaggat catttgagcc caggagttcg agaccagcct gggcaacggt gtgagacccc 9300
gtcgtacatc aaaagttttt atttttaatt tcactttcat gacttggtca tcaagtctgg 9360
cttttgcaaa aattaagaca taagaaaaga atgcttcagc tatgaattac tatcaattgt 9420
tcaaaaatac catcaactct caaaattatg cataaaatac accaaaatta ttaacaacgg 9480
ctttgcggga ggtgggggag gaggaggaat agattatctt ctagtatttt ccaaatgttc 9540
tatattaaac atatattaac ctttaaaaca tctacttttg ttgtattctc aaaataatat 9600
aaaacactac tatataattt aaaaagaaca ttctaactct aataatttca taaaaggagg 9660
tcacagtcca aattgtaggc aactataaaa atttcgctct tgaacaacca atgaacatat 9720
acatgatttg aaggaaaaat ccctaagaaa aagcagctct ctaattaaag agaacctga 9780
aattaagtaa atcaattcct gacagaaaga cgaagatgtt ttctgtaata caagaaagca 9840
agatcacctt tccccagac atctaattgt agtagttaa cgttcgaatt ctggaataaa 9900
aaactcagca aagctaaag tatgactctg ggtgccaaaga aaatgccaca ggaactagca 9960
tttccaatca gcagctcctg agatcaggaa gactgttatg ttctatgata taaagtccac 10020
aataaaatct gttagttttt ctggttaaat gctcatgcta aaaatagtga ctgctcaaat 10080
attaagtaag aagacttagt tttgcctct tgttcagtc tctgaattcc aggcaattgg 10140
ttttcgatat cttgtgacac caatacttga catctaacag cattttgtcc actactgcag 10200
atgcactgcc gagtcactct tccaccctc tcacaggcat atatttgtgc tgcaaggttc 10260
aagtgttgag gagctcagga ttataaataa cgaaagaaac gagaagcagc ctttctttgc 10320
tgtctcacc tcactcatag gaagtaaaaa gctcttttag atccatctgg ccgatctcat 10380
ttcacaggct gcagaatcac ctaacccttt ccacctgcaa agcttgctac tctctccttc 10440
cttagaatct cacagctgag tatgttttca gaactgttct tagacacaga tcatttacta 10500
tttattctca tcaaatctg aaacagctat gcgagaggtt ccaaactcat gaaacctaaa 10560
acaaccatca gttcatcgaa gcagctggga aaatcttttc gagacaacat caactgcttt 10620
tgttcatgag attaaaaaaa aaaaattcat actgaccaga aaccaagca cgctggaaac 10680
agccaacat taacgatgac ctttgccttg gaaacctga gcaaaaattc cccttggttt 10740
cccttatatt tcctttggaa aaaaaagga acaatgcaac agactaggct ggtttcactc 10800
tgtgatcact tacaaggcca gctgttcctc ctccatgttc ctactgat aagaatcagg 10860
gactcctgct ctacgcatga agtcaggatg gcattgatgg gggccctgga acactctgcc 10920
tctgttcccc cagacaatc aagtaacagg cattactgt aaaaagcaag actggaagct 10980
gcaggggaag ccaagtagca gcgcattatc ccgaagctgt gagatcaccc tgcgtcctgc 11040
aataacagtc aggagataca gccagaggaa accgcagac atgactctcc ggggtggggg 11100
tggggtggga ggccgcagag catggtcagt cacaggattt atgaaaacaa gatgcagaaa 11160
gtctctgtga cccggtctcc tggcttctct tctgagctca ctctgggcc agagcctcat 11220
gcgcccctg cgtggctgac ctgaatactg tatctgacga ctgcagcttc tgatgccag 11280

```

```

aggcacaggc tcccattca tcagaccctc aaagtgtccc actggggaag tccatgaaga 11340
aatccacatt ggtgatggca cgtcacttt accaggtgtc tggggccagg aagcccaaac 11400
ccacaagcca tccatcccag ccacccagaa gtcactcttc tcacaaaaga tctgagtgtc 11460
ctaaaaggag tgactaaagt tacaaaaggt cagacgcaga cagacaaaac ggaaatgtct 11520
tctccaccg ctgtaagaaa aatcttgatg agggataaaa aaaaaaaag ccgtgccct 11580
ctctaccgc caactggaat gttttatct ccaccacaca gatctgttct cggacactga 11640
ttactgcat tcgggaagct tcataagatt aaagtttctc caaagcattg aagacagaca 11700
aaaaacctca atcaatgtc ctcaaaaac ccaggcccc caaatataa acagccagt 11760
tcattccaga accaagccat ggcaggaac cagtaatcag ggtggtcata cgtactaatt 11820
tgagctggaa acctctggac agcagaagca gtgggtggc tgaaggaaga tgcagaagtc 11880
ggtaaaataa aagaggttcg tggctgcagt gtcacatct ctaacgctcc ctacaactgc 11940
cctccgagct ctggccatct gctccctatg gagatcagga aaagccagga ggctgccgag 12000
tgctccacg agggctgggg agccaactcc tctcagagt cctaccgaa aagcaaatgg 12060
ctcttgtaga actctgtct tctctgata ttttggtga aaaaggccct tgtccagca 12120
catctgatg aaagagggcc attcagaaa acagctgagg ttctctaat cactgcactc 12180
ctacgggctt ttctgtaggc cggagaaca agcaccgggg tgtgattcg acattgtgag 12240
ggcaacaac tgccccaag gaaccaacc caagcaaca gaccccttc cgattcaat 12300
caacattctg aaggatgact cttctttca aatcagcatc cattaccca acggtgacgg 12360
tgacgtgggc agctgccga gttagtatt ctgcgtactc aaagcacggt tacatcctga 12420
aaattcttca gtcattgtaa cagctatctg aggggacacg ccaggtagag gggaccacat 12480
gcacacctat gaggagctct gggatacgca cgggtcccaa ggcaggtcag gctgcaagg 12540
tcctaaaggt tggaggtgtg atcccaaac ctccaggcac aagccagcca agagctgtgt 12600
ttttagctgt tcttcagtg agagaaataa gttcaggatg tgaataacca tgacgcagga 12660
gagaatggaa taagtacct aagaaagggg ctcggtagg gttacaaga gggaggaggg 12720
agcatttaac tgggtacttc tggacaatt cctgaaggaa gcagcactga gtaggggctt 12780
ctctccctc ggctcacaag tgaccaagcg atctcccta cggattaagt gaaacacaca 12840
ttaccatgat tctggtttt cagggtagga aacccagct tgcctaggag cacatatctc 12900
tacaagatgg ggctggactc acatctatct gccccagcc cacctgctta accctgtta 12960
agcagctgtt ctactcatcc agaataaaaa tcagagccat tatgtcggg tcacatccgc 13020
tcattgctgc ccaggtgctt aatggcaaa ccactaaggc actgagaagt cagaatgtgg 13080
atcacatctt ccgtccttct tcccagtgtg tgaatgcac atgctggtga aagagagaga 13140
aggaaccatt caagcaaca gaactccagg aagacgagac tgtgccgggg ttcttccatc 13200
tgcccaagta gaaatcagaa gggcagggga cccacagct tctctacc accactgccg 13260
tcattggtgg gggacaggac acatcctttg gcccttctgc actgcataga ggctaaggag 13320
ttctgtaaac cacacagcca cgctgacca gaagtcgctt tcaagtaag ttctcatca 13380
acaggactat tatttactga gcatctccca tgtggccaag gctgtaggag gtacttagct 13440
acgccacgtc attgaactct ggcagtctg cagggtagg tattttctcc atatgacaaa 13500
cgaaggaaag ccgtcaacaa ttccaaaata gaatcaccag ggatagcatg gacaacgcc 13560
atggtgactg ccgcgcttta aggtttaaga aaagtaaaaa ctgggggtga tgactcattc 13620
ctgtaatccc agcactttgg gaggctgaga tgggtggatc atttgaggtc aggagttcga 13680
gactagcatg gtcaacatgg caaaacctg tcttactaa aaatacaaaa attagccagg 13740
tgttggtgtg catgcctgta atcccagcta ctcaggaggc tgaggcagag aatcacttga 13800
accagaggag cggaggttgc agtgagcaa gatcgacca ctgtactcca gcttggcgga 13860
caaagtgaga cactgtcttg ggcaggggcg gtggggaaca aaagtaaaaa caatggtctg 13920
ggaattcata ttctgggtt ccaatttaca ttctaccata tatactctga ttaaccccta 13980
gaattaaccc ctagaattcc ttacagggtt ctgttcattc atccaaacag gcaaacttt 14040
gcagagcatg gagcacagg taagccaagc cagcccaagc tctgataagg gcaaagacag 14100
ccatcctctt taaggaatgg gtatatgtgc tgggtatctg ggtgtctgcc ctgctgcata 14160
gaaacagcat ttcttgaaga acaaaaatag taggtataga aacatcacag tatggaatat 14220
ccaacacccc ctgaattcca actctggtca tacattgaaa caacctatca aactcctaaa 14280
acacattcat gccagggtcc agcctcagca gagtctaatt cggaaggctc gtgatgagtc 14340
ctgggcatct acttttttaa aaagtccag ggagctgggc atggtagctc atgcctgtaa 14400
tcacagcact ttgggaggtc aaagtgggag aatcagttga cccctggagt tcaagattaa 14460
cctgggcaac gtaacaagat cccatctcta caaaaaata aaaataaaat tagctaggct 14520
tgggtgtgtg tgctgtagt ccagctgct caggaggtg aggtgggaga atcacttgag 14580

```

```

cctgggtgagg tcaaggctac agtgagctgt gaccacacca ctgcactcca acctgggaga 14640
cagatcttgt ctccagaag ttccagggg tgcttctgat gcacagccaa gttttaaaaa 14700
cctcagaatc aaataacatc atggccaggc atggtggctc acgcctgtaa tcctggcact 14760
ttgggaggcc aaggtgggtg gatcacttga ggtcaggagt tcaagaccag cctggaaaac 14820
atggtgaaac ccagctctta ctaaaaatac aaaaattagc tgagcgtggt gacgcacact 14880
tgtagtccca gcttcttggg aagctgaggc acgagaatca cttgtaccct ggaggtcgag 14940
gctgcactga gtggagattg tgatcctgga gtccccactg cactccagcc tggatgagag 15000
tgagactgtc tcaaaaacaa acacacaaac aaacaacatc agaagacaca gagaaaacag 15060
tcttctccat gggtctcata aagatacctc tcacataggt acacgtcgat gtttctgct 15120
ggtaaaaggt aacaccaaca aaaaggcatg gtgctctcag aaggtgggtg atgtgattag 15180
gtgcaataaa gggaggtcat gctaggttca aaaacaaaat aatactctct ttggaagcag 15240
taaaacagat gctagtcttc tactacacac ttccagagac ctgaatgttc ttctggccct 15300
ctaagggaga cgctgcatac tgacaatacg aaatgatgac agtgaaagca aaaacagatc 15360
agacctgtgc tgttgaaac agacatgggg tctcgctatg ttcccaggc tggctctcaa 15420
ctcctgagct caagcgattg ttccgccttg gcctcccaaa gtgctgggtg gacagctgtg 15480
agccaccgag accaacctca gatcagacct ttgacaaact ctgctgtgga caaagcattc 15540
tgggtaatgt caactcatct gatcttcaca aaaccgtgtg gaagaccaga caggcattat 15600
tacactaatt tatgcctaag gaaacaggga gttaaatagt acaaatttag gatttctgat 15660
gctgtatctc gaaaaaaaag tagagaatat gagcctgaag aagaggccct gtaaagggtc 15720
ccagattgat gggacaggct gagacaaacg gaatcacttt tccctggata gaactaacc 15780
tcaatggtac ccactctgc atggtgatta ctgaggggac tgtcaattgt ccagcgaact 15840
tgatggtaat tctaggagaa aaaggaaact atgtaatgct gtcagcatag aaagatgggt 15900
gccaacgagc attccaaaaa ggaggtcttg ttaattcggg ttcatcaac aagtatttgc 15960
tgagtgtcta ttgtgtccgg tcagtgtctaa ggctgagaa tttagaagtg aaacagaccc 16020
ggtttccacc catgccacag accactccac acctggtctg gagtgcact ggagggccag 16080
gcaggcacag gacagtaact tcgatataag gcagcaagtt ccacggtgga aggaggtgga 16140
aggtgcagat gcacgtacac acagggggtc agggaggcct cctggaaga aatgaagcct 16200
gcgagggcct gaaggatcag taaacagaga ggcataaggg gcaggagagt aagatgatta 16260
tgctacatgt acctatttgt gaaccagga ggatttggcc tctgtcataa aaggccccc 16320
tgtgggttca taaacctcaa ttacaaaatt gtgctttata tctcagttcc ttataagttt 16380
ggttagcgta aattggtttc ttagaacttg atcatccctg agtgaactca caaattcaag 16440
tttcagaatg tgcaaaccta agaaacaac ctcattgctt gggttgagac atcgactgt 16500
caacatcaca aattctcagc acctgaatgc ctggtatact atcaacatat attgttttaa 16560
atatgtaaat aatagctttc tagttataga gagtttgtcc ctacattttt ccacttaatt 16620
tttacaatcc cattcccttg atgaacaac ccagcctgg gcaacatggc aaaatcctgc 16680
ctctacaaaa aatacaaaaa ttagctgggt gtggtggcgt gcacctgag tcccgggtat 16740
ttgggaggct gaagtgggag aatcacttga gcccgggagg cagaggttgc agtgagccaa 16800
gattgtaccg ctgcactcca gcctgggaga cagagggaga ctctgtccca cccacccccg 16860
cctccaaaaa gaaaagaaaa gaaagaaaag aaaatgaaac caccaagact gggagaagat 16920
aaatgacttg tctgtgttca tctggctaat aagaggtaga atggggctga aaaagttcgg 16980
tgctcttctt gaagaatcca taggtcagaa agcagcacca tctgacctgc agcaatagca 17040
gcaacgtgga aagctaatac actgacctca aaaccactct cagtggagct ctggatggat 17100
tcagaacccc aggcctagca aagtgaagt gataaagatg taaaggagat cgaaaattca 17160
ccatttggag agagattagc taaagactgc aggtcggatg gaaaattctt tccatggttc 17220
tcccacaggt tcttccctca tttggaactc gtgtttaaaa gtcacaaaga cctgagttg 17280
ggccaaggct tcgttcttct tcaactgtggg ccttgcaagt caacatggca gggcctcgtt 17340
ccaaatgtca ctcttcagag cctaagaaaa caagtaactt tagggacaca cctgtcaacc 17400
ggagctccca aattgtaccc ccctaaacac ataatgctga gcatagaaaa attccagctc 17460
tgcagagcgt tatacttagg gaaaggggtc acagacaagg aatgctggca gggtcatta 17520
caaatactct tgctgtgga acatgtattg tttggctaga aggcgtagc ttctctcaga 17580
gagaaggaat gtccaaaagt atttcagaca gtaagagaca ttctctgagc cagctacaca 17640
gtctctcttc aaaccaacgg gtagcggcaa gcagctgaac tgaccagcga gctcgcaaaa 17700
gcaagctttt ttttttttt ctcctaaat aagacagcaa gtgatgtgtc ttggcttgggt 17760
ttagcaaat ttaagatagt tcctgatga cccaagagc cctcagccc catggaagct 17820
ggagctaatg catcttctc caagcatcat ctgctctacc aggatctaag ccccttcacg 17880

```

```

agggcagaag gtataaaggc tgcactgtgc gggaaatgct atggcagcaa agacagccaa 17940
acacgccaga aataacaggc acatgaagga aatgtttctg agacagctca aaaattccga 18000
gaagagatta tcccactgt cccaggttct cagccctgtc tatggtatgc agccccatac 18060
cacagtcatt tgcaccgag tcctaacttt gtcagaggcc cctcctttca ggtctctcag 18120
gcaccaccca gttctggccc tcctcacccc cgtgagccag gcgacatcca agcagcccca 18180
cgggtcaccc ggctctgtgc tgcattctct gaatgtccct gaaggccagg gctgttgtat 18240
tctctaccca ctctctgtct agtatgggag ccaactggcca gatgtgttga ctgaacactt 18300
aagatgcagt aagtgtgacc aagaaactgg gtttgtcatt ttatttcatt ttagttaatt 18360
taaatttaag tttaattagc tacatgaggc tatcagctgt ggtataggac agcagagctc 18420
cggaagcttt tggcctgggt agaagaatca ggacaagcgc ctccctggcc tctcgcccac 18480
tctgcacagc cgctaacct tgctctcatg acattcttcc ccagccccag aacttttagc 18540
catgtgacat catctattga ttagagtcca aacttcttgt gctaactctc tatgggttgc 18600
cacaattagc cattgtatgt cgtaaacctt aatttcattc atctgctatg tctgacctt 18660
aggggcttag aatatagtta gaaaacagta ttcagaata aaaaaccatt cttgtattac 18720
ctctcgact attccccctg ttctccatgc ttcgacttct ctgttctatc cccagctata 18780
gcactgtccc cataaagcct actgtgggtc tcggctcatg gtgtccttcc tcccatctgc 18840
ctcccgacgt catgcctgtc ttccagtgtt tatgccttct ccaggaagct ttctcttggt 18900
gcccctcgtg tgaagtatag ctccctccct tcaacatctt ctgacacctc ctcttactgt 18960
gcaggtgagg acaactgaggc ttaagggtta agtcacttgc tcaaggtcac atacagtcgg 19020
tctctgtat ccgcggttcc ctcatcagt gattcaacca actacagaca gaaaatacag 19080
tatttgagg atgctgaact ctttgaatta gtgggttctg cgggtgctta agcatccatg 19140
gattttgta tctcggcaa aggcgggggt cctgaaacca atcccttgg atactgaagg 19200
aaagaccacc cttagtata ggaacctagg aaccacagtt ccctcatctc caaatcgtgt 19260
tccctgaccc acttatttac taactagtgg tgaagccatc ttctgcccag tatattttaa 19320
cttcacaatg ggatgtgagg gccaggatgc acatgctttt taaatctccc tctgtgcttg 19380
acatacagta gattgaaagt aagtgtgtgt agatacactg gccaaactgt gctcttctct 19440
gaagtcaagta ttccaggagt aactcacctt ggtcatctct gtgccctggg cacactgggc 19500
actcccaca cacagggtta acctggcaaa taagactcac agcatcatgc cacgtgcgag 19560
ttaagccac ctggagggtc ggtcagggtt tctgacaac tgagtgttc aaataacaca 19620
acagcagcta agttccccac atcaccttga gtgtctggag agctaggcct atgacttctc 19680
tgtctcagga tccctctcag tcccagaaa acagtggaca tcaataaatg taacaccaat 19740
aacatcttcc ttgagcgcta tgctaagcac atcagggtatg ttaactcatt tattccccag 19800
tgtccatctc tcagtgtttt atacatacgg gaactgaggc tcaattagcc gagcgtggtg 19860
tcgtgtcct gtaatcccag ctattgggag gcacaagaat cactcgaacc caggagatgg 19920
aggttgacgt gagccgagat tgtgccacag cactgcaaca gagtaagact ccgtctttaa 19980
aaaacaaaaa aacagaaaac aaaacaaaca aacactgagg ctccaggagg ttaagtcacc 20040
tgcccaagtt catgagacca aggagccggg aagcaggaag gggaaggcag gagtgtact 20100
ctgaaacctc tgctcttagg cactggcttt cagctgaact gatacctctg gaaaacagtc 20160
tcaaaaaagt ccacttctcc tcccaacaat tcagacctaa aaaccatttg gcggggaagg 20220
gcagggaag cttctgagtt ggggaggggg tgtgggatcc caagctgagg tgtctgttgg 20280
caagcagggt gcaagggca tctgtgcagg gaggggctg caaggagagc agagactgct 20340
cacaggcaag gaatgaaata ttaacatta atgttaatat taatatttat aattaatata 20400
tttatgatat atagcatata tacatattat attaatatt tataactata ttaatatata 20460
taattataac tatattaata taatcaatta taactatatt aatataatcg attataacta 20520
tattaatata atcgattata actatattaa tataatcgat tataactata ttaatatata 20580
cgattatatt aatattaata taatcgatta tattaatatt aatataatcg attatattaa 20640
tattaatata atcaatatta acaaatatat actatataat ataaataata cctaagttaa 20700
tataatatgc ataattgtta tattatttaa tatttcaggg acaatgggag tcatgaatat 20760
ggagagacaa aactagaatg aaccccaagg tgcgcatca gaattgaagg taccagtctg 20820
aactcatagt ttcaacctt ttgaaataaa tatagatgca cgtgtgtgtg tatgcacgta 20880
catacaaatg ttccctaatt ctgccattg agaggcctgt ggttagcaac accccaacag 20940
caataagcag acctagcttg gctcctaaat ttcattttcc actaaaagga accagagccc 21000
cttggataaa ggactgattc cacagggtgg tagggagcat ctgttgccag aaagcaagaa 21060
agcacttaaa gaatgatgtg gacatgtcaa agggacacag aagccagcct ggatgagatc 21120
ccactggccc taactgtcca caaggacaat ttgagcaagg atgtcaacaa ttaagagca 21180

```

gattataaac cactgaataa aacagaaaaa tacaaagaat tgaaacggac attgatggca 21240
 gacagatat taacataatt ttaaagtatc tctccaagga atgcttctga atgatgaagg 21300
 ggaaaaaat aactgtacag tgaaaaagcc tggtaaaacc caccitagtgc accaaagtga 21360
 atgtcaccat agtgggacaa aaggaaatca agtgccacct tatgggattc aacgaggacg 21420
 cagcatccct tgggtgatgt tccagccaaa tacacgtgcc cgggtggaac acacaagaac 21480
 atcagacaca ctcacactga gggacactct gcaaaactgac agtactgggc acaaacatgt 21540
 ccaggctatg gtcgaccgca gtggctcatg cctgtaatcc cagcattttg ggaggctgag 21600
 gtgggcggat cacttgaggt caggggttcg agaccagcct ggccaacatg gcaacaccct 21660
 attctctact aaaaatacaa aaattagccg agcgtgtgg agcatgccc taatcccagc 21720
 tacttggggc gctaaggcac aagaatcgct tgaacccggg aggtggagggt tgcagcgagc 21780
 tgagatatca ccgctgcact ccagcttggg cgacagagtg agtttccaac tcaaaaaata 21840
 aaaaaataaa ataaaaatcca ggcacacaaga gtcaaagaaa gactgaggaa ggttccagac 21900
 tgcaggagag ccaagagaca ggataactag atgcaatggg cagtcctgaa ttggatcttt 21960
 tgttatgaag gacaacgctg ggacatatgg tgactcttga atggggtagg aggactagac 22020
 ggtgggaatg catcagatgc agtgtcccgc gtggatggct gtgttgcggt tctgtgggag 22080
 aatgcccttg tctgtattcc aagggtaatg gagtagcagg ttgacaaatt actttcaaat 22140
 ggttcaaaaa agaaagtctt ttactctga cttgcaattc ttaigtgaagc tggaaattat 22200
 ctcaaaatta acgagaattt ttatcgacg tagtatttta catatttatg gaaaacatgt 22260
 aagtatttgt tacatgcata aactgtgtaa tgaccaagtc agagtatctg gggtatccat 22320
 gaccttgagt attaatcatt tgtatgtgtt gggagcatta caagttttcg agttaccaat 22380
 tttttttttt ttcctttgag acaggtcttt actctgtcgc ccaggctgga gtgcagtggg 22440
 acgaccacgg ctcacgcagc acagcctcca cctcccagc tcaagcgatc cttccacctc 22500
 aaccacccaa gtagctggga ctacaggtgt gtgctgccac cccagctaa ttttttaatt 22560
 tttttgtaga gacagggtct cactatgctg ccagggtcgg tactgaactc ctaggctcaa 22620
 gagatcctcc caccctcggtc tcccaaagtg ctgggatcat aggcagcagc caccataccc 22680
 agccaaattt tttaaagtta ttttttaaat ctccacttaa ttcgattttg gtaaaacacg 22740
 acctgtaatt tttctttatc ggtaggtaat aaaagcttca gatgatttta ctgatcactg 22800
 gtatgggcat atttcatgac ttgccccttt catctcttgc atagttttac cctcaccaag 22860
 caagaccttc cctgcctcag cactgtttgc cctcttcgtg tttccagaa cagaagtggc 22920
 cctgtttcgt gccagagca gaagagaacg atgaagagct ctgctctccc aggtcttctc 22980
 ggtctgtgtg tgtccaggtt ttgagggcct ctcacataca cggctctgga ccacgtaaga 23040
 tctaatttta gcattttct gctcggagac cacaatgttt ggaacagcag gggtgacct 23100
 gccctgacag gcctcctatt gtgaagggca cgcgaagcca ggataccgca gccctgcagg 23160
 atgtgactca gcatacctgtc tcagtctgtg ggccggccagc agctctggca ccaagtgtg 23220
 ctgctgacct caccctctaa gaccacaaat acccagggta attggtggga taggcatgca 23280
 gcatcagctc tccctgttaa gacaacttgc ttgtccatcc attatgtctg gcttccttgt 23340
 gaacaccaca ggatctctatc aggaagaggt cttccgagga actgatctgc tggatttttc 23400
 aggacaccaa gaatcaagag attggtcttg tttctctctt tgctttgact accaggaaac 23460
 tcaaagtcag atctgtggcc aaattctggt aaccatacca atgctatgtc atgtattaca 23520
 tgtacaaacc tcccctttac ttcattttat tttctctgc tttcttctg tcccgatatt 23580
 ctactaatg ttacattcta ttgttctcta tgaatgttgt aagggttttc aaatcctttt 23640
 tggagggcac tactgtagat acaacacaca ttacccttga gggataagga ctctttttga 23700
 ctccacacag aatccctggc atttggcaaa gaaccatata ttaggcacta aatacacatg 23760
 ggctgaatgg aaaaagccaa tagctaagta aaaaccacct ccattaccat attgtttcac 23820
 aagaggttct tttcccttcc atctcatgag gtggggcctg gtcaggagtc cccagggcct 23880
 gggaattagg ttccttaggg agccttcttg ctgtaggggc agccaacagg tcagtggcct 23940
 tgactccaga cctaaagagc cactcctaga ctcccagctg caacagacac agcgtggcac 24000
 ggggtggcct ggccactggg gaagtgacaa gtgatttcca gatgtgcag ccagcctggc 24060
 tctttccaga ccacactgaa ggccccttcc tgtgggaatt ctgatggggc ccagatttgg 24120
 ggaacacagc ctcgaggact cttggcaagt gctgtgccag cctggaccag gaatgacttc 24180
 tgtgggcaca gggagagacc aggcatttcc taacacagga ccttgaacag ctttctctga 24240
 aacaaagtct ttctaaaaat agcttcaaaa gtaaccattc aagaaaagaa agaaaaaaa 24300
 aactgtaaaa gttaaaggc tcaagaatga tatttcccag ataaaagcct ggcacagggt 24360
 tcagagggaac ttgcaggaaa acaggtcaag gctgggtttt tctcttagg tgtcacttgg 24420
 tiaacattgg tctttggagg ggaacaagt cggcaggaag ggctggcact gaaaatgatg 24480

```

gccactgggt ataggccagg gccagacact gtacacagaa caagactctc tggaggcctc 24540
aggaggcccc tgaggaggagg aaggcagggt gtgggcccag ggtcagacat gcaagtgagc 24600
taagtggcaa ggccgatgcc ccattccagaa gccccgctct gaccacacgc aggcctctccc 24660
ggcatgtcct catttatgcg gcagtccttt gtatctcact gcaattctgc cccacactg 24720
caggctggcc agcgtggctt cctcataagc acatcacctt gcatcccgac actgactaca 24780
cccacaaagc aggagccccc gcaccctcca gcccattcgc tcagtctgct ttgaaaatgg 24840
ctcctctcgg gggctgggcg cagtggctca tgcctgtaat cccagcactt tgggaggccg 24900
agggtggttg attgcttggt gtcaggaatt caagaccagc ctggccaaca tgggtgaaac 24960
ccatctctac taaaaatata aaaaattcgc caggcatggt ggtacaagcc tatagtcta 25020
gttacccaag aggcctgagg aggagaatca cttgaacca ggaggcagag gtgacagta 25080
ggcgagatcg tggcactgca ctctagcctg ggtgacggag caagactgtc tcaaaaaaaa 25140
gaaaaaaaaa aagaaaatgg ctctcttgga ttttgattaa tcctattttg attaatcctg 25200
gtttctcatt ttcagccttc cttgaagcag catgacccat ctggatgtcc tcctcatctc 25260
aggaattttc taataagctg tctaaatcca gagatccgac cacagaacaa tgaatgcaa 25320
agatgagttc taaagatgcg agtactttct ttctaaacgg acgctgcttt gtgtatggct 25380
ctgctcctgg gggcagacgc ggcaggctaa gccctgcgga ggaggagggt agtcccagca 25440
gagggtcact tctctcagc agcccgctg gttttctcca ctgcagggtc agaccatagc 25500
cctgacccag ctgaccccc ataagcgcat gacctgtctc tcaccgtggg aataaaactc 25560
gtgatagtca gttacaaata cacagcaaat gatgagcagc acaatataaa cacagatcta 25620
gattggtggg tctgaggact cattcttaaa ttggaggcc atcacctaatt ctgtctttt 25680
cactttacat agcaggagac agggaccagc agaagtgaag aggcgttgcc ttaggttgca 25740
cagcagatga cgcctctcaa gatggacct aggttgtctg actccgtctc acagctttgc 25800
cccatttatc atgaagatga acgctggtaa cactgctacc tacgagctga gcttgacgcg 25860
acattcctgg tgtgtacatg catgcgtgca cgctcacgca atgtgctaag tgcacaggaa 25920
ggagaccaga gcctgaggc gttcttttga agtctaagta ctgggtgttc gaaagttaa 25980
tgaaacctac tagactctga gcaaaattcg ttttacgtta acctaatga aaagttaa 26040
taagtctga cagaattaac tcttcacgct tctgtcctca ttgtcccca ttctagaatg 26100
agttttctaa ttaaaaaaaa tatatagggc cgggtgcagt ggctcacgcc tgtaatcca 26160
gcactttggg aggccgaggc gagtggatca cctgaggta ggagttcgag accaacctgg 26220
ccaacatggt gaaaccccg ctctactaaa aatacgaaaa attagctagg ggtggtggcg 26280
catgctgta atcccagcta ctccggaggc tgaggcagga gaatcactgg aacctggag 26340
gcagaggttg cagtgagcca agatcgtgcc actgcactcc agcctggtga cagagcaagt 26400
actcccatcc cccccacaa aaaaaagta tataatgtgt tgtgtgtata tatatatata 26460
tagctaggca cagtggctca tgcctggaat cccagcactt tgggaggccg atgtgggcag 26520
atcacttgag tccaggagtt caagatcagc ctgggcaaca cagtgagacc ctgtctctac 26580
caaaaatata aggtggtgtg cactgtggt cccagctact tgggaggctg aggtgggagg 26640
accaattgag ccaggagggt cggggctgca gtgagctgta atcatgccac tgtactccag 26700
tctgggcaac agagcaagac tctgtctcaa aaagaagaaa agagagagag agggaaaaaa 26760
aattgaagc aaattctgat tttcaaatca aacgttcaa caaactgcag aaataaaacc 26820
cgagttaaac caaaaggaac agccaaacag cacaatgacc ccaatgttta aatatgcccc 26880
aatgtttaaa agtgggagtc aatgggaggc cactacctac aaggccacag gggttagggc 26940
aggactcagg tccctgaatc acagcagcct gcattcaaac cctggctcag gcctcccacc 27000
agcctcgtgg aactggtttc ctaaaatgag gagagtcctt actttgcagg cttgtgacaa 27060
caagatgaca gcaagtgcaa aagttccaag cccagagcct gcagcctgca gaagctggcc 27120
tcattaccac ccggatgttc tccgggctgc agcacatgaa ggggatacgt gacaatccct 27180
gttttaagta cagctcaggg agttgacggg acctgcccc gcacatagtg atgccgctaa 27240
tggtcacca ggaagaatgg actgcaaaag ctggttcttc tgataaactc cattctgtct 27300
ccagtggtgg gttctgatgc atagggagga ggaagagaca gtgcttggt tttggggtga 27360
agagcacagg ttttgagtc aatgagacat ggagtatgag ggtctcagct ctaccgttta 27420
ctactaaata aaaacaggcc actgacctct ctggggttta gtcttctcct ccagggaatg 27480
ggaattcaaa tgtccttaca gggttttcac aaagattaac tgaataaatg cacacaaggc 27540
aatcacagag tggagtatgg gtgctccctt ttctctctc catccctgct ttattttttc 27600
gcctgggcac ttaccaacac acgattatg cgcttggtta tttattttac tgtctgtct 27660
cctcaacaga atgtcagctt ccagagcagg aatttttatt ttgtttgttg ctatattccc 27720
agccccataa acaggccttg gcacacagta ggagctcaaa aaatatttgt tgaatgaata 27780

```

gctcacaagc agacagatga ggacagaggg gtcttgagac tgatctaaca gcaccgatat 27840
 tactaaactg caacggaggc aacgggtgga agaatttctc tgcctttgt ttcctgaaag 27900
 tccaagacca cttttagttg ctcaacagga aacaatactc aacttacaag acctctaggg 27960
 cctatccagg gcaaaactggg cactgtgagg caggaggtca ggccagccctg tccctaggg 28020
 ggctcacggt ctagtgggca ggccagctt ctcatatgt gctcagaggg gcccctgtct 28080
 tggtttaata ctctgttggt gccatcttga aattcttaat aatgtttgtt gttgttggtt 28140
 gttgttggt ttgagacaga gtctcactct gtcccccagg gtggaatgca gtggtgtgat 28200
 ctgagctcac tgtaacctcc acctcccggg ttccagtgtat tctcctgcct cagcctccca 28260
 agtagctggg attacaggca cgcgccacca tactcggcta atttttgtat ttttagcaga 28320
 gacagggttt caccatgttg ccaggatgg tctcaaacct ctgacctcaa gtgatccgcc 28380
 cgctcggcc tccc aaagt ctaggattac aggcgtgagc cactgagccc agcctcttaa 28440
 taatgttttt aaaaggggct ctcccatgtt catittgcac tgggcttcac aaattacgca 28500
 gccagtctg cattacagga aatatttctg tacctaagta catatactac aaagcaagta 28560
 ccaaacacca aggaacact aaggagagaa aaacgcctgt gagaagaaaa aggaagacac 28620
 gaatcattcc caacagaagc tgttaccatg aaggaagtac gggcaggggc attgttgaa 28680
 tgtctactat gggagaagg gttcgcata tgagcacatt taattctgac aaccacccta 28740
 caagctgtgt actatactgg ccatttgaaa ctaaggcctg cccgagatca tataatagcc 28800
 taggagtgca caaaggacag acacaggagc caaacccatg cccatccctc cctaagtcca 28860
 aaatcataga aaaaaaaaaa taagaatcaa catgggcgtt tatitttaag gccagcatgt 28920
 tcaaggtggg ggcaaatcca agagacacta agcctcagag catgaacaag catgtgggtg 28980
 ctgagtgagg gggaccagtg tttaccaggg tgatgtcaga ctctgcaagg ctgctcccc 29040
 gtgtttctgg tctcttccca tgagcaccag gcaccctta ccatcccaa actaggcaca 29100
 tctgtaacgc tgaatggaag cctacttgtt tacatgtgtt ctatgttaga ctgggggcat 29160
 ccctagaaca cacacagatt gactgggtgg cagaattctg ctaggtgcat gcaccgagt 29220
 gagcctttct ctttgaatgt gggagggacc agagaacaag atgggagagc tgttccctta 29280
 attaggtgt gctgcacatt aaaggcgtg agacagtcatt tccagtata acaatctgtc 29340
 ataagaccct acagaagcag actctcctgt tggccttgaa gaagcaagca ccacgaattc 29400
 tccacagctg caagaaaatg aattcaggcc aggcgtgggt gctcacgcct gtaatcccag 29460
 cactttggga ggcagaggcg ggtggatcac ctgaagtcag gagtttgaga ccagcctgac 29520
 caatacgggt aaaccccatc tctactaaaa atacaaaaat ttgcagggtta cacctacagt 29580
 cccagctact cgggaggtg agacaggaca aaaatttgca gggtagacct acagtcccag 29640
 ctactcggga ggcagagaca ggacaaaaat ttgcagggtta cacctacagt cccagctact 29700
 cgggaggtg agacaggaga attacttgaa cccaggaggc agaggctgca gtgagccgag 29760
 atcgaccac tgactccag cctgggcaac agagcagaaa aaaaaaaaaa aaagtaaaaa 29820
 aaaaaagaaa atgaattcag ccaagaacca cgtgagctta aaaggagacc ctggggttca 29880
 caaagacct cagccccgc cagcaagcct tgtgagttcc cgaacagaga acccagctat 29940
 accgtgtcca gattcctgac ccatggaagc tgtgagataa taaacatggc ctgggtgcgg 30000
 tggctggcca ggcatgatgg ctcatgcctg taatcccagc actttgggag gccaaaggcgg 30060
 gcagatcacc tgagttcagg tgcctgagac caacctgccc gacatgatga aacctgtct 30120
 ctactaaaaa taaaaaatg gccaggtgtg gtgttatgag cctgtaatcc cagatacttg 30180
 ggaggtgtag gcaggagaat cgcttgaaac cgggagcgcg aggttgacag gagctgagat 30240
 tgcgccattg cactctagcc tgagcaacgt gagcgaaact ccatttcaaa tttaaacaaa 30300
 ataaacatat attgtttaac tgttaagtta gtggttaact gtcatgcagc aggcattgac 30360
 tgatacagta acctatgcac acatccatct ccagtacgga cacagaactt ggatgcacgg 30420
 ggtgcatgac acctcttggc aggaactaac tggacagaca agcaacaaag acaataaagc 30480
 ccaggctaag atggactgcc aagggcaggg aggaacccca gagtgtggac aggtgcaaaa 30540
 gtagggtgt tcaatgaaga ggggaagcat ggtctgcagg gcaatgacat gccaaacccc 30600
 atccactctg aactgttagg ggagggggtg aaggcaaac cacacttcaa aaggctgtag 30660
 ggagaatggg gtccctgggg gacttccaag tggagaccaa aagggaagg gagtgcggag 30720
 agaaaggcag aggagttagg gatttcacag tttaccactg aaaccaaata aaacagaaga 30780
 gacaaaatcc tgcagctcgc tctggcccaa acctttgcta gggcaggcaa tcacaaatga 30840
 gcaaattata ataattctaa tgaccacgtt cccgcaattg acttggaat gctggattaa 30900
 aaaaaaaaaa ctctactcct gatccacacc ctggggacaa tattatctcc ccagtgtcct 30960
 acctagccca caactactta tgtctcatgc cagactgagc cagctcccg gatggcaagg 31020
 gagccaggag ctgctgccag cagggccatc tctcaccaca tcccacagt ctgaacggca 31080

```

cagcttccaa agagggacta cgagcggcca gcagcagcct gcacatgcag aaggcagggg 31140
agagcgaggg aaatggatct atactgctct gctgcaatca tctgcatgct ggggtgtgaga 31200
tgatcagttc ttgagacact tcccagaagg ccttcagaaa tactgtctga gttacaacac 31260
tgcttctctc aagtctgtat tcttatttgc atcttatagg aatgtagccg ggtaaaggag 31320
gaaggctgct tcaagtcaaa gggcatccat ggtggcgccc ctctcaggcc tggaccagc 31380
acctgcagga gtcggccctt ttaattctcc tctgccgtga actaactctg cacatcagca 31440
atactttgtg aagaccgagc acagcaacca agcccactgt ggacctgatt ctaggcaagg 31500
aacttttttt ttttttgaga cagggtcttg ctgtcaccga ggatggagtg cagcggcaca 31560
atctcagctc actgcagtct cgacctcttg ggttcaagtg attctcctac ctcaacctac 31620
ttgagtagct gggactacag gcgtgcgcca ccatgccag ctaatttttc tatttttttt 31680
gcggagatgg ggtcttgcca tgttgtgtag gctgtctca aactacgggg ctcaaagcaa 31740
tccactcacc ttgcttccc aaagtattgg gattacaggc gtgagccact ggggctggcc 31800
tagacaagga acatcacaca actactccac aacctgaaa ggtcacagtg tcatccctgt 31860
tttataggtg gaacaattga gaccacaga gctgtaagaa cacacaagg taaaggaact 31920
cagccaccag cacaggagcc agatgccaaa cttaggtctg cctgactcag aacccactg 31980
ccttccctct acacctggct gttctccta catgtctgga atttactgca gggtaaagg 32040
ttcatccatt taaactgttc acttttatca acttacttat ttgagacag agtctcgctc 32100
tgttgccag gctggagtgc agtgacgcaa actcggtcca ctgcaacctc ccgggttcaa 32160
acaattctcc tgcctcagcc tcccagtag ctgggattac aggagcgac caccaggccg 32220
gctgattttt gtatttttag tagaaagggg gtttcacat gttggccagg ctggtctcga 32280
actccggagc tccagtgatc cgtcccgctt tggccttcca aagtgtctgg attagagatg 32340
tgagccaccg tcccagcca tttaaactgt ttaaatgcta cacaaaggca gagaaatgag 32400
gccgtcacta agggatttga gacagttag ggatacaaca agggcacaca gacctgcatt 32460
gtaaggcggg tgtggcact gtcacagata gggatgccag gggctccctg ctttctctga 32520
agagagggaa atcacaaata tctggggcag gcgcactttt agctggctat gaggactaca 32580
gccaggtgaa aaggaactgg cctagggaa cgtgtgtacg ggggagcagg gtagtctcc 32640
aatggactgg aaaaggcaca tgcgagaggg gaggttgaa aggccaccaa cgccggtgac 32700
gctggagctc agaaaagact ccgaggacca gaaggaagaa gcataaagg accaggggt 32760
gatatgccaa ggtagagagg atgggtctga ggtgttcctc tgtgacggga cagaagagat 32820
gtgaggacct gaagaggcgc caccagggaa cttgagaaaa agaaggcagg tggctcggaa 32880
gacccgatcc atgtgacct gcaatttatt ggatatggat gaatcagagc tgactttttc 32940
gatgacccaa aagatgaaac tattaattaa ggctagacgg aggggagaaa agagagagaa 33000
atacctgct acctccagtt tttctccta cagcacctcc tagagatgag gtgatggcag 33060
ctcaccttca ggatccattg gaaacaaaga gaaatctctc cctaattctc ctgaccccaa 33120
acagaaagca accaactatt ccataatttt cttctctagg agagattcag agagaagagg 33180
cttgaaaatg caaattaaca cacgtagtg aatgagttc agtcaaagct taaaccagga 33240
cacaactttt ctgtgtagc gaggaaggct agggaggcagg atggtgtcct gtgcctaaaa 33300
aggtgagtg gacgtcaagg tggctgaaga ggggtttaaa agggcaccta gggggacaag 33360
cccagagccc agcatccac cctaaatgag aacacagtc tccaactcca gccagggtc 33420
tcctgttgt cacaatgggt ttggggacct gctgacagt gcacggaagg actctcgggt 33480
gtggtcagaa tgaccaaca tcccaggag cacctgccac cagttggcat gagtcccttg 33540
tgctggccct ggcgctgctg ctaatccacc ccagtggact taggcatgct cctcacctg 33600
tatgtccaag acagctaatt caacagtact actaacctgg tcccagaaa ggcggcagag 33660
taggatcaac ttggcattag aggtctcgct tcaatacag gatttcccaa ttccaatctt 33720
gggcctcagc caccaaccgg gaaaaccccc ctccaagggc tgttttgaga atatggaaag 33780
ctgccaagc tttatttccc atccctttgt aaggtccct gcagctccca actagaagag 33840
aaagggacct ttatttagca agaaaagggc caggtgcagt gactgtcct acgctgtaa 33900
tcccagcact ttgggaggct gaggtgggag gatcacttaa gcccagagt ttgagaccag 33960
cctcaacaac acagtgtgat ctcatctctt caaacacat ttaaaaaaat tagccaggtg 34020
tggtggcgca cgctgtagt cccagctact tgggaggctg aggtgggaag atagcttgcc 34080
cccaggagtt caaggctgca gtgagctatg atcacaccat tgcactccgg cctggacaac 34140
aaagaccctg tctctaaaaa ataaaaataa aaagtittaa ttttaaaaaa ggttaataac 34200
taatgagaaa ggtccacata caaattttca tgtcactgct atttataagt tcttattaac 34260
tgaaaacact atgctataag ctattaaagg taactaaaaa ataaaaatag ctctttgccc 34320
caaagacagt ctaagtgaac gagctcagta ctactcatc aacaacagcc ctgagagaga 34380

```

```

caagagatgg aagagattcc aaccccaaga aagggtctgga ggggagccag gggaggaggc 34440
atgggggagg ggacctccaa gagggagcta ggggagggca gatggggagg agagcccaga 34500
aggagagctgg gggaggggaa catgggggag gggagcatgg gggaggggag ctccaagggg 34560
gagcatgggg gaggggagat ggggagggga gctctaggag ggagctgggg gaggggagat 34620
ggggagggga gctccaggaa ggagctgggg gaggggagca ttggggaggg gagatggggg 34680
aggggagctc caagagttag ctgggggagg ggagatgggg aggggagctc caggaggggag 34740
ctgggggagg ggagatgggg aggggagctc caggagggag ctgggggaaa agggcttggg 34800
gagggttagct caatgacggg acgtggggat ggggagctaa ggaggagatc tgggagaggg 34860
gagcttgggg aggggagatg ggggagggga gatgggggag gaaagctggg gaaggggatc 34920
tgggagaggg aagcttgggg agaggatcta gggaggggag ctgggggttg ggagatgggg 34980
aggagaaatg ggagaggagc ttggggaggg ggatctaggg aggaatctg gggcaaggga 35040
gctgagggag gggaaattgg ggaggggatt tagggagggg agctggggga aggggagctc 35100
agagagggga cttcagggag ggagatggg agaggggatt ggggagggat ggttagggat 35160
gggatctagg gaggggattt gggggaggga agcttgggga ggggatctgg gggaggggag 35220
tggggaagag agatggggag tcgggggagg ggaacctgga gggagggatc tggggaaggg 35280
gattttgggg aggagaacag gtggagagag gagctggtgg ggagggcagt tggggcagg 35340
catctggggg agatttgggg ggaagggaag ctgggcgccc acaggagccg ctgtgaggtg 35400
ggcaagcccc tcttctcagt cctcctcagc agtcagtctc cagacttcca ctccaccct 35460
cctgcttcc acccagacag tctgactgc aactcgccc atgactgcc ccattgggaa 35520
tccagctgct tctagcttgg gaacctgac gtgggccttg acctgaccaa tcaaaaacc 35580
cagggtgatg gagcaaatgt gtcctgtatc ttgagcataa cattaaaagt gaggaccag 35640
cagaagtcce ccagcgagga ccagaaata aggaatctct ttgattctg caggctagt 35700
tttccctacc cacataatct ttagaaatca tgtgtgccg aataaaagt agtatttccc 35760
ctcccttacc tcaagcacac agaacatcg gaaaaagct gagcatatt ctaccagtc 35820
tgcatatgag ttgaccaga acacctgct gtcgtaatg aatggttag cccaatttct 35880
gaacacatat ttcttttcc aattaattt cttcccctc atgagataa acagactatt 35940
tttttttaa gaacaatatt cctgaaaatt tattacttt ttttaaaact atgaggtcag 36000
agttaaagac tggtctcttg gtatgaagga atacatgata ttaatataac aaagggtga 36060
atcttccata aatcaacaaa acacccaaac aaaggcagaa ctaattttt ggcaagaaa 36120
aaacaaaaat gtttttggg tccattagt aatacatcag ctgaggactg ccattcttga 36180
atcttttaa tgagcagagc taaagattc tcataagcac aattaaagca ccctgaattg 36240
atacctttag ggggttagt atctgttca aatcagcaa gtgcttacg caaaaggaac 36300
accttaccaa aagcaagatg aaaaagttag ggcagagtgt catgattatt acttttttt 36360
ttaagcagaa gaatagtctg caaaaaata cataaaatg ctcaagttag gccgggcaca 36420
gtagctcatg actgtaatcc cagtacttca ggaggccaag caggagatg tcttgaggcc 36480
aggagtccaa gaccagcctg ggcaacacag caagacctg tctctattag aaaataataa 36540
gttaccaaa aatgctcaaa atggtaatgt aagggtgta gaataatggc aattattttc 36600
cttcttttcc aagcgttata taattattt aaagtggcta gacatatgta tggattttaa 36660
agcacttcag ttttatgtgt tttaggata atttctaaag cactaaaaaa ttggcatatt 36720
ctttttttt tttttttaag acggagtgtt gctctgtcgc cagctggag tgcagtggcg 36780
caatcttggc tcaacttgc tccgggttc aagcgattcc cccgctcag cccccaagt 36840
agctgggact acaagcacac actaccacgc ccggttaatt tttctgttt ttttagtaga 36900
gacagggttt caccatgttg gccaggatgg tctcgatctc ctgacctgt gatccaccg 36960
cctcagcctc ccaaagtgt ggaattacag gcgtgaccca ccgacccag cccaaaattg 37020
gcatattctt ttgaacgtt ttcccttgg gagaggaaca agagcattcc ttacctgctt 37080
gggagaaaga cttaggaaca agaattgaaa gtctgttac ctgaggttta attttcgatc 37140
ttcttcgtc ccagcttcca actacagaga aagaaagaga atatcacacc acaggcacca 37200
ctgtcaacac gcctcgggag gcagtctcac atttctacc ccgtactgg aaaaagataa 37260
agatatccag gaaacctagc tacttctaaa cagccgtgcc ctttctcac caatcccgg 37320
ctgtcccttg gagtcatttc cgtgggggaa ttttcaggtt tccaaatgt gaccacatt 37380
cctgccgag tccagggat ggagtctgt tagctcaaca tttctatct ggtgtgtta 37440
cccagcacgg tcttttagc ctacgacctc aactttccga ggtgttctg gacattatc 37500
tgtttttct ttttaaggg aggggtcat gtttaaagag aatccactc ctccgagag 37560
ccaggcaata acagctgagt gatgaacacc attttcaaaa aaccaacca ggcaagactt 37620
gcacagtgga aggtggccag gaatcagggc gtctgttgtt ggtcttgaa agctcttgat 37680

```

```

ggttctcgaa aagacttaaa catttgatac gaaacatcct aggctatcgg tttatttata 37740
taaatgcaag aaagagatat ttaatatattt ctgaaatcta aaaggccacg agtttgggct 37800
ccagaagtac ctatgactta tttttatttt ttttctttca gagagcaaac tgaaaataag 37860
aaggaaacac atacacaccc cccaacaac tccgcaccgc tgggacttgg catgtttttt 37920
atgttgacac gaggcgcccc ttgaatggga aagagaaacc tggaaagctg tgatggctgg 37980
gagagatgca gggctgatcg aggacagaaa tgaggcagga gccaaggcg aaggaaaaag 38040
ggtccagaga taatgtaggg aggggcctgg gcagcaagg acacccacca ggaggtggca 38100
acttaacca agaagtagta caccagcccc gcgcagtggc tcacacctgg aatcccagca 38160
ctgcaggagg ccgaggtggg cggatcacct gaggtcagga gttcagagc agcctagcca 38220
acaaggtgac acgctgtctc tactaaaaat acaaaaatta gccaggcacg gtgacatgta 38280
cctgtaatcc cagctacctg ggaggctgag gcaggagaat cacttgaacc caggaggcgg 38340
agggtgcagt gagccgagat tgcaccactg cactccagcc tggtgaaaga gcaatacttc 38400
gtttcaaaaa aaaaaaaaaa gcgtacacca gagggcctgg gagtccctac atcatattaa 38460
gatgaagtac atacaagatt ctgcagaggc acctacccca cactgaagga gaagtggaaa 38520
gagcagggaa ggcttcaatg accacaaaac aagtcacaag agcaacaat tgacaagag 38580
tatgttgggg tctaaccag tggttctcaa tgcggagcaa gttctctcaa caggagacat 38640
tggaagatgt ctggaggcat tttctggag gtcactactg gcacctagt ggtagaggac 38700
aggatccac aacacacagg atgttcccc cacaagagag aatgttctgg tcccaagtat 38760
caacagtggg ttgagaaact ctggtccaat ccaaaaaagt gcctggagat ctgcccata 38820
aaattagta ctttgaatgt ttctcagaaa taacaatgtt atgatccatt cctgaaaatt 38880
attgatttat ctattcttgt gctctgctg ttcacacaaa ggaactgaga taagattcac 38940
agggaaatga cataagtac ataataaag actgggaaaa aaagaaaatt aaaagagaaa 39000
agagagcctt aggactgagg gctgtgatcc cccgtttccc acgccggcag caggcctggc 39060
tgtgtcagga aagcactgcc ctaagtgtct gactcattat gaagtgtcaa ttgaaagat 39120
gatgacgtga cttgggggta cggacttcac aatcatttaa ctctcggtca ctctctgagg 39180
ttctcagatg aaaggccatc tcaggtcagt tattccaggg aaactacatc tgccaaggaa 39240
cacatgaag aggtaatcca gtccttttag atgagccagg gccacacac aggaagcaac 39300
tcaagcgagg gcggaccagg gcagaaccgg cctggcctag gtctcctgac cccatacaca 39360
cttctgtct ccatccccacc ttgcttctca cctcaacaca tctgaacgag ggccttgct 39420
tcgggaaaca tcccagcgca ttcaaagcca agcaatgaat gctgcagctt tgctatgatc 39480
aaataaaagc tgcttgatgt ttactttatg tttatcaggg tcattggcac ttggtaaaaa 39540
taatgcttta tataataatt gaaaaatgtat ctagtagaca caacacaaat gtccaacaaa 39600
atgtgttaca tccatacaat ggagtattat acagccatga aaaggaatga agtactgcca 39660
catgtacaaa tatgatgaa ctttgaaaac gtgatgtga gtcaaagaag ccagacacaa 39720
aaggccacac aggtgtgat tccatttata taaaatgtcc agaataagca aatccataga 39780
tacagaaagt agattagtgg ttgcctaagg ttaggagaga tggggcaggg ggaggctgca 39840
ggtgagggt aacgggtaca gggttcctat ttggggagg atgaaaatgt tctggaatta 39900
gatggtggtg gttgcacaat ctcatgtata tactaaaaac cactgaattg tacacttaa 39960
aatggcaaat ttatggtatg taactaaaaa ataataagac cttaaaatgc gtaagacaag 40020
aacagattag gttgcaagta actctaggaa catgggggtt tgaatcagaa atctgggctg 40080
acaggttcag cctgaagcca acctctcccc ttacctcaca taaacttttg tgatgagaac 40140
ctgagattaa gtgcataaat tgccagacag cagtaccgg aacagaaaac agcccctcct 40200
cgctgtggga aggaaggcg ctctgcaac ctaacttctc agtagcaggc tattgatcgc 40260
cagtgttctt ttgctcttaa tcagcgtgta gagggggatt actagaacct tctgtgtata 40320
gataactcat gaatggctc tcctctccaa ggagggggct gtgaagggtc aacttcccag 40380
ccactctgaa aatgtccctg ccaatcccag caaaacaagg ctgaagaact accctaccag 40440
gagacagggc tgtcaagcca aatgcaaaca ttattctctt gtctctctca gacacacaaa 40500
ctccccctg gttatcagtc aacttcccc actccctccc acaaagaaag gggctgaaga 40560
gcccagatgc tggctgcgga acttctggg cctgggaccg caggccgct cctccagtct 40620
tctctaaca cagctaagg tctgcaggcg gacactcagc cttgttatag gtaagagttt 40680
agaccagagg ccttgacggg ttcttcaaga gatgtgggc aagattgcgc gaccagaggg 40740
tcacccctgc agctacagag ggctgacctg ctgagaggcc caaggcccca gcctaggaca 40800
agccaggcca accctgcagg ctaagagggc aacagtgcgc tcaatcaacc ccagaggaaa 40860
aagtggccag gcaaacggac ctgggccaca cacagaccga caaaacgcg cacagtgcc 40920
ggacacgcaa cccaggaatg cacctatgca atcaccaga atgggtcaca gccacacaga 40980

```

aagatagatg cacataaaca cacaggcctg agtgatgtta cagaaaggaa aagccagact 41040
aaggctgcac gcacagacgt gaaacacagc cacacagagc ccacagcacg ctcggtcacc 41100
gtcacacagt gacacgggca cgcctacaga cagaactcca gaggcggcag gcggggaac 41160
aatctcacac gttttaggg gcactcccag atgcctgtct cacgctggca cagtccccgg 41220
tacggcaggt cagcaacagt cacatctcac atcgacagc caggcataca ggcaaagggc 41280
ctagaactac cccggccaca ggtctcagaa ccagcggctc acgcagtcac ccaatcaagg 41340
gtcccagttg cacatccagt cacccttga ccctggtcac actgcagagt cactcacaaa 41400
tgggagtcac gacagacgca cagtccctcc cagacagagg tcaacccaag atgggggtca 41460
cacctgaaat cacagtcccc acacaatcac gaggtcacat ctgcacacac agtctctgca 41520
cagtcaccct taggggtcac aacgcacaca gtctccgct aacaggggtc accccaagat 41580
gggggtacac cctgatgtg ggtcacagc cacacacagt atctgcaga cactcccaga 41640
ggggtcgcac tgcatacaca gtccctgcaa agtcgcccc cgatagggtt cacaccgcac 41700
acaaagtccc cgcacagctc cccaagacag ggtcacatcg cacacagtcc tcgcacggtc 41760
accccggtcc ggctgcccgg ctctgttctt acggcggggc cccgaggagc ccgcgcagcc 41820
gccccctgc cccgcacgc cgccccagc tccggcggcc tcggcgggc gtccggcggc 41880
ccaggccggg cgcggcgagc cgggggtca cctcgtgtt gctggccgag gaggaggcgg 41940
cgctgggctg gggcgagcgc tgcagggtca ccagggccat ggctgcggcg cgggtgcgag 42000
gcgccacaga cgtctcgagc tagagccgcc accgccaccg ccgccgggc cggggccggg 42060
gcctcctgga gccgcgcgc ggcggccggg ccgagccggg ccggggccgc cctccccct 42120
cgcgctgcc accgccccg cccccagctc ccgctccc cgcggcgcg cgccagcctc 42180
agtgcgcgga gtggcgggg aagcgggcag ggcgggacga ggaggcgcgc gtgcgcgggg 42240
gccttgaggg ctgcccaggg cctcggtgg tcgatcacgt cctcgcgcg ccgacacac 42300
gcgccccgc ccgcgcgcc cgctatcagg cctgggactc gggggcgcgc gcgccccc 42360
gagcccgta cccccagggg ccctccccg tgcctgcct ggggaaactg agggccggcg 42420
accgtgcaga caggactgta cagcgaccag gaaataaag acgtcctggg gccggcgcg 42480
gtggtcacg cctgtaatcc cagcactttg ggaggctgag gcgggcggat tacgaggtca 42540
agagatcgag accatcctgg ccaacatggt gaaaccccg ctctactaaa aagacaaaaa 42600
ttagctgggc gcagtgtgc gcgcctgtg tcccagctac tcgggaggct gaggcaagag 42660
aatcgcttga atctgggagg cggagggttc aatgagctga gatcgcgcca ctgactcca 42720
gcctgggoga cagagcgaga ctgcgttca aaaaaacaa aaacaaaaa caaaaacagt 42780
aagcaaaata gattcgctg attttgcaga ggttaataca gttattagc acgttttta 42840
aaaagtattt tgctaattt ttcaatgaa ttcttcttg gtgttctgaa acccagccaa 42900
ctccttgag gtcagggaag gcttcccaga agagctttat tctgaggctt gggcttgagc 42960
ataagcagga ttaacagtg aaagaacaga gagacagctc tccaagcagg ggggatcagc 43020
gtgccctgaa gcaggaagaa gtttgtcaac cggaggccag cactcaggga agggaagagg 43080
ggaggaatgg ctggagtct catcctctct ggaaagatcg ctccggctgc tgcgtggatg 43140
agggaaccag gggcagagg ctgagggaga ccaggagga ggctgctgct gttgtcccg 43200
ggagagtgta ccagttatgg gtagggagag gggaacatgg aataagatac caagaaggca 43260
attctggctt gacttagtag taggaaactt ttcttttagc caaaatctca tctcccggct 43320
cccaccccca acctctgcat gttgcacaag cactcgcaaa cgcagtggtc ccagcctgcc 43380
ccgcagctta gcaaatgtt ctactgccc aacaggaaac ccacgcagcc tctggattc 43440
ttccccgtcc ctcctctgt cctgggctg tgacctctc catgttattc acagggtctc 43500
agcacgattc atctcaaagg tgattctagt ggggggcaact gtagcttcta cggagcgttt 43560
ctaagagggg atttgggga atgtttgtg ttgtcttct gatggagggg gagagctcct 43620
ggcatttaga gtgcaagagc cttggatgct aaatgtctc caatgcactg gacagtctcc 43680
ccaacaagaa ttgtccatt cccacaaaat gtttcttgg tgaanaacc atttatagta 43740
atttgaagcc agaacctaac tccatttcat gcatcaacac tagtcttctc tcttctctc 43800
cttcttctc tcttctctc ctgcttctc tcttctctc ctctcttct ctcactttt 43860
ttctgaacaa ggtctcact cccgtcacc aggtgaagt gcaatgtcac aatcatagct 43920
cactgcagcc tccatctccc aggtcacaat catctctct cttagctct ctgagtacaa 43980
cgggtacaca ccaccacacc cagctcctt aaaaaaaag tttaactatg ttgccaggc 44040
aatcctctg cttccgctt ccaaagtgt gggattacag acagaagcca ccatggctag 44100
cctggtattt ttactgaat ttccagaaag gtgactatgt tgaaaccctg tctctcctaa 44160
aaatacaaaa aattagccag gcattggtgc gggcacctat aatctcagct actcaggagg 44220
ctgaggcagg agaactcatt gaacccggga ggcagagggt gcagcaatct gagatcgtgc 44280

cactgcactc cagcctgtgt gacacagcaa gacagagaga aagagagaag ggaaggagg 44340
 ggagggagg ggagaggagg ggagaggagg ggagaggagg ggaggggaga ggagggagg 44400
 ggagaggagg ggagggagg ggaggggaga ggaggggaga ggagggagg ggagaggagg 44460
 ggaggggaga ggagggagg ggagggagg ggacgggaga ggagggagg ggagggaag 44520
 gaaggaaaa tacactttgt ttgtcttgag agttttgtca agagtgttc atccatcctt 44580
 agggaaaagg aggtaatgga tggcaacgcc tctgctaata ttagagcacc ccacacaagg 44640
 tgcccacaac tgtagctgca ctctaggtag acagacagtc ataggtactt aaatgtcaaa 44700
 tataagggaa aattgtggac aaaattcagt tgagtagaga atattttatt tctcaaatcc 44760
 aagcacattg attattggca ggcccatgct tctgagatgc ccctgtgtcc tctmagggag 44820
 tagtggctga gcatttccac attgtaatgc atgtgtttc attatgattt atttttctt 44880
 tatgtctctc ttacattagt ttcaaattt gagagtttga gaatccccg gagaaaatac 44940
 agattgctag accccacctc ccagagtttc gaattcaca gggttctgt agggtggaa 45000
 aatttgacg tctaacaat tcacaggcaa tgcctgatgct tctgtctggg gacgacagtc 45060
 tgagaactac tgcctataca aatgcaatgg cctcttcacc aagaaattcc tacctagatc 45120
 tgatcctggt acccgtccgt ggccccaat cctaattccc ctgctctggc cggcctgctt 45180
 ttccactcac cccaactttt ttggaggcag tctccacccc ttctcacttc ctcttagagc 45240
 tgagagccct tttcttccc acaactaact ctgtctagaa atcacctcca aaaagctttc 45300
 cctgccctt aagcagtgtc atttccagga tctctagacc ctcacctac ccttaaacac 45360
 acagcaagtg tcagtctgcc ttatcataat ggggtccatct ctctgtctt tcccattacc 45420
 gttaggccag gaacggctcc taagaaaagc ctcaggaatc aggttgggac cagcgtgagg 45480
 gtgcaaaatg taagagggtg ccccaaaaa ctcaatgatt aagataaata gtattttaat 45540
 gcaatatttt agaaaatcaa aattaatgcc aaatccatga tgaataaaat atttttaaaa 45600
 tttgttttt ttttttttt ttaattgaga cagagtcttg ctctgttgc caggctggag 45660
 tgcagtgtgg cacaatctct gcctcttggg ttcaagcagt tctctgcct cagcctccc 45720
 agtagctggg attacagacc cccaccacca tgaccggcta attttgtat ttttagtaga 45780
 gatggggttt caccatgttg gccaggctgg tctcaaatc ctgaaatcag tgatctgcct 45840
 gcctcggcct cccaaaatgc tgggattaca ggtgtgagcc actgcacctg gtcaaaatat 45900
 ttcaaaaaat ttttaagag ccaaggcttc attctgtcac ccaggactgg gtgtagtgtt 45960
 gcaatcctag ctacttcag ccttgaactc tgggtcgaag ccattcctct gcctctgcct 46020
 ccggagtacc tgagactaca ggtgtacacc accacgcctg gctgacttta tttttgccag 46080
 aaactgggtg ttgctatgtt gccaggctg gtttcaaact cctggaggca ctcaatcccc 46140
 cgaccttggc ctcccaaagc tttgggatta ccggcatgag ccaccacacc tggccaaagt 46200
 atcaaatttt taagtaaaat tggcatcagt attgtgtcac tgattcttcc acttacttca 46260
 gacttcagt tagctcagca aagcactttt attgactctg tctttatttg attcttttac 46320
 aactttggcc attctaaagc cttttgtgaa aatggcctgt ggttcagctg ggcattgttg 46380
 cgtgcacctg taatccagc tactcgggag gctgtggcag gagaatgcc tgaaccagg 46440
 aggtggaggc tgcagtgggc tgagatctgt ccacttttga cactctgtct caaaaaaaa 46500
 aaaaaaaa aaaaaggag cctgtcggct tgactccagt agcctctgat ggggtggagt 46560
 ggacaagggg aagtgaagc tccaggcct cagtcagggc aggtcccaag aagccctgag 46620
 catggaggag ggaacaatc cagtagaggc agctctgaag ttttctcca tgcattagag 46680
 ccctttccaa tcagtatcat gatttttcat catataatag tttatttaat catctttgac 46740
 ctctctcttg tagtccagc tcaactttgt aactaataaa aaacagttag ttattgagct 46800
 atttgcctc tgctaaggca caatgcaaag tgccttgtga gtgtgtggg gacatgattt 46860
 attaacatgt gactgtcccc ccacttatac tccaagatca cctcctccag gaagccttcc 46920
 ttgccccgtg gctgggttag gcaccccttc tctgtgtcc tacagccct gtgcattagt 46980
 gacaatggca ttgtgatct gccttaggcc cattctggg ttgggacact ttaggtacat 47040
 tcattcttgt caccctgtga ttctatttc atgggtgagg aaattgatgc acagagtgtt 47100
 taaggcactg gcccaagtt atgtaactaa ggagtgtga acctggttca cccatgtttt 47160
 tctgctttag aactcaggca aagacaggtt cttccaggac agcctcagaa agtgttggtg 47220
 caaattaggt tgggtcaaaa gtaattgcgg tttttgtcat tttttttt ttaattgtg 47280
 caaaagtaat tgcggttttg tcattaatga ccaactatta taagtaatag tcccttttt 47340
 ttttttttt gagatggaat cttgctctgt tgcccaggct ggagtgcagt ggcttgatct 47400
 tggctctctg caaactccgc ttcttgggtt caagtgttc tcctgcctca gcctcccaag 47460
 tagctgggat tacagtgcc caccacctg cccagctaat tttgtattt ttagtagaaa 47520
 cggggtttca ccatgttggc caggctggc ccgaactcct gacctcaagt gatccacca 47580

```

cctcggcctc cccaaagtgc tgggattaca ggtgtgagcc actgcacctg gccagtagtt 47640
tgccctgttaa agcaaataac ttgtaatttc tccttaatta ttcatccaa aatgatattc 47700
agaggtaata aagctctgat aggctgaata atggcctgca aagatgtcca tattccaaat 47760
ccctagaatc cctgcctatg ttaccttgca tgctaagagg gttttacaga tgtgattaaa 47820
ctcaggatgt ttagatgggg aaattttcct ggaggaggcc caagaggtec taatgtaatc 47880
acaagggtcc ttataagagg gaggtgagaa ggtcagagtc agtagtaaga gatgtgacaa 47940
cggaactgag ggattagagt gaaggaagag gccacaatcc aaggaatgca ggcagttgct 48000
aaaagtggaa aaacaccaa aaatgaattc tcctttcaga gcctccagaa agaattggagc 48060
cctgctgata tctttttctt ttcttttttg agttagggtc ttgctcacag agctgtcacc 48120
caggctggag tgcagtgga tcatcatagc tcacagcagc ctgcacctcc agggctcaag 48180
ggattctccc acctcagcct cctgagtagc tgcgactaca gacacacacc actatgcccg 48240
gttgactttt ttttaattatt attatacttt aagttctggg gtacatgtgc agaattgtga 48300
ggcttggttac ataggtatag acgtgccatg gtggtttgct gcacctatca accgctcatc 48360
tgcattagat atttctccta atgttatccc tcccctggcc caccaccccc tgactggccc 48420
cggtgtgtga tgttccccat gcccggttga tttttaaggg tttgtttgt ttgtttgtt 48480
ttttagagac gaggtgtcga gctgggtgca gtggctcatg cctgtaattc cagcactttg 48540
ggaggtaagg cgggcagatt gcttcagccc aggagttcaa gaccagcctg ggcaacatgg 48600
cgaaacaaaa aaatgcaaaa aattaactgg gcatggtggc acatgcctga ggcctgaggtg 48660
ggagtatcgt ctgagcctgg gagatcaagg ctgcagttag ccatgatcat gccactgtgc 48720
tccagcctgg ttgatggggt gagacctgt gtctaaaaaa taaaagaaat gaaggtcttg 48780
ctgtgtttcc taggctgttc ttgaactcct aggctcaagc aatcctcctg cctcagccac 48840
cccagttgct tggattacag gcacaagcca ccatgtccaa tcctggcaac gtcttgattt 48900
tagacttctg atctctacaa ttgcaagaga ataaatttat gtgttttaa gccacgaaat 48960
ctctgggaat ttgttacagc agccatacga aatgaatata aaactcaacc tccatttggg 49020
ctttaaaaaa catatcata taatgccatt acccagtata ttccaggtgc ttccaagcg 49080
ttgtgtcatt ttctcattca ctcaactcat ccaataaact atgtttgttg ctctcctggg 49140
cactagtcta ggaatctggg ttccatcagt gaacaaaatg gaatcactgc ccttgaagag 49200
cattcaatca agtgggaaat atagtaaaaa tataatatata tgcaaatatg tttaaatca 49260
tatgtggtaa atatattgca tttaaatgaa ttaataggcc gggcacgggtg gctcatgcct 49320
gtaateccag cactttggga ggccgagggc agtggatcac ttgagggcag gagttcgaga 49380
ccagcctggc caacatggcg aaaccccgtc tctactaaaa gtacaaaaat tagccagttg 49440
tggtggtggg tgccgtgtaat cccaggtact cgggagggtg aggcacaaaa atcgcttgaa 49500
ctgagggggt gcggagggtg cagtgagccg agatcatgcc actgcactcc agcctgggtg 49560
acagagttag actgtctcaa aataataata ataataatta attaaatgaa ttaatatgg 49620
taagggtcct tagaacaaga taggcactga tatgtgtcaa ataaatgaaa tatgatgtcc 49680
aatcatgaaa aagcttggga gaaaaacaaa gcaggctaag ggcagagtaa tggaggaggc 49740
cacttagaca aatggctcagg gaagcttctg ggtgaggtga tatttgagca gaggaatcac 49800
catgacagca ccaccaggga ggtgtagaaa ccctgggata tccttggttc attcaactg 49860
gcctccccac taaggaaact tgaggtaact tttctgagac ccattttctt tctgtctgtg 49920
tcaccaggc tggagcgcag tggcgcgata tcggctcact gcaacctcct cccccaggc 49980
tcaagtatc ctcccacctc agcctcctga gtagctagga ttacaggtgt gtgccacat 50040
accagctaa tttttgtatt tttagtagag tcggcggttc accatgtttg ccaggccagg 50100
ctgccacctt ggcttcttac agtctggga ttacaggtgt gacacctcag acccagccga 50160
gaccactgt ctttctctgt aaaattgata tgaagtgtat agtgcctggc cgggcatagt 50220
ggctcacgcc tgtaatccca gcactttggg aggccaaagt gggcagataa cctgaggtca 50280
ggagtccaag accagcctgt ccaagacggt gaaacctgt ctctactgaa aatacaaaaa 50340
ttagccaggt gtggtgtgtg gtgcctataa tctcagctac tcaggaggct gaggcaggag 50400
aatcgcttga acccaggaag cagaggttac agtgagtcga ggtcccgcca ctactcca 50460
gcctggacaa caaagcaaga ctccatctca aaaaaaaaaa aaaaaagaaa agaaaagaaa 50520
agaaagtggg agtgcagacc tcagagcttg gttgtgtcaa ttgaacagca tactatgcag 50580
gaaaggcaga gcgtgctgtc ctatttacta atagtacctt aggtattggg ttgaattgtg 50640
tccccacaaa attcactagt cctgtgaat gggaccttat ttggaaatga ggtctttgca 50700
gctgatcaag ttaagatgag gtcattaggg cggggcccta ttcgcatatg actgtgtccg 50760
tatgaaaagg gggaaatttg ctgggcggcg tagctcatgc ctataatccc agcacttttg 50820
gagaccaagg cgggtggatc acctgaggtc aggagttcga gaccagcctg acaaacatgg 50880

```

```

agaaacccctg tctctattac aaatacaaaa ttagccaggc gtggtggtgc atgtctgtaa 50940
tcccagctac ttcggaggct gaggcaggag aatcacttga acccgggggg tggagggtgc 51000
agtgaactga gattgcgcca ttgactcca gcctgggcaa caagagcga actgcatctc 51060
aaaaaataaa caaacaaaca aataaataaa taaataataa aaggggaaat ttggaccag 51120
agccaagggg aaaatgcttc ctgaagggtg tagttgtgct gccacaagcc aaagagcacc 51180
cgagatggtc agcaaaccac cagagctagg agtgagaagt gaggagcaga tttgcgtggc 51240
cttctgaaga aaccagcaac tcgatttcag agttccaggc tccagaactg agagagtaaa 51300
tgccgtgtgt ttaagcctcc cagtttgtgg cactttgtta cagcagccac aggaaaggaa 51360
cgcatctaac atgatcattt catcagctgc agaaaatgag gctcagagca aggctagggt 51420
ttgaaccag gccaaactaga cccagacca catacatggg ttgttggcct ttctagcctg 51480
ggagggtgaca gtttggatgt tccactattt gcagggaacg gtgttcagag gactcaaagc 51540
ttctgccacc tgggccaggg tgtccagtgt tagatatgga agtcaggtat ctggggcttc 51600
tgacaaagca tctctctggg tgggtcaata ggcaccagca gccaggcagt tgagggatct 51660
ctgcccctgt cgggagttgg ctgaagcctc ccttcctcta cccatccctt cattaacct 51720
gcttgccaga tcaaggtgtt ccttgccctc tgccagggggt gtattcacct gaatttcct 51780
tttatttact atgacacaa gcaacacact gacccttgct gggcctcaga atctcaactc 51840
ttggtgggggt gcggtggccc atgctggtaa tcccagcact gggaggccaa ggtgggtgaa 51900
tcacttgagg ccacgagtta gagaccagcc tggccaacat ggcaaaaacc tgtcttact 51960
aaaaatacaa aaattagcca ggcatggtgg catgcacctg tagtccagc tactcaggag 52020
gctgatgcac aagaatcact tgaatccagg agatggagggt tgcaatgagc caagatcaca 52080
ccactgcact ccagcctggg tgacggagtg agactctgtc tcaaaaaaca aacaaaagaa 52140
tctcaactct taatatggaa tgattaatgt cctataagaa agcattggtt agatcaattg 52200
ggtatatgca tgtgatatcc ctctagata gctctcagt gtcctcatct cctgatattc 52260
atgcctgtg gagtctcttc acacaataca tagaactgac ctgtgtaacc actaggatat 52320
cacagatact acagcatgtg gcttctgagg ctagggttca taaaagacac tgaagctgct 52380
gtcttctgt ctcttgatt tctcattctg ggggaatcca gctaccatgt catggggaca 52440
tttaggacat ttaagcagcc aaagagagag ggccgcatgg caagaaactg aggtctcctg 52500
ccaatgacca gcactaacct attgtcatgt gaatgcacca ccttgaaaat ggatcctcca 52560
gccccagtc ggcctttatt tattatttta cttattaatt gagaccgggt ctcatctgt 52620
ctcccaggct ggagtacagt ggcaccatct tggctcgctg taacctctgc ctctgggtt 52680
caagcgattc tcattcttca gcctcccag tagctgggat tacaggcgtg cgctaccatg 52740
cccagctagt tttttgtat ttttagtaga gacagggtt cgccatgttg ccagcctgg 52800
tctcaaacct ctggcctcca gtgatctgcc tatctcgac ccccaaagt ctgagattac 52860
aggcaagagc cattgtgcca ggcccccat tcaagcctc agatgagatc acagccatgg 52920
ccaacatctg gagtgcaacc tcatgagaca ctctgagcca gagtgccca gctgagctgc 52980
ttccagattc caggcccaaa gaaaatgtat gagataataa atgtttattg ttttaagctg 53040
ctaaatttta aggttaactt ttatgcagca atagataact tttatatgct gccataaaaa 53100
tattataaaa ccatgcacta gtacagaaag atttttataa aatattaagt ggaagaaaag 53160
aaaagcaggc caccaaacag cgtaggacag tagacccat ttttgaaaga aaaatgtgaa 53220
gagttaaaaa actctacca aaggggaaaa aaaagaggc atcaatggag agatggagaa 53280
gctttgtttt tggatgggaa gactcagtat ggtagagtta acaacctctc gaaattaaat 53340
taaatggaaa tgctattgaa atcccaactt gattctttt agtgtaggga tctttacaac 53400
agataactgg acaagctaac atttattggg tatatatgtg cgttgcatca cgtgacagtc 53460
actgtttcat ctttaattcca ccataggaga aagtcctctt ttatttaatt tttctgagag 53520
taaagtactg ctattacctg ttcctctcc cattttactt aggaggtttc aagaggggac 53580
ttgtctgaga tcttggaac cgtggagggt agatgacatc aagcatgttt gatatttaca 53640
tgtgtgccct tgggcctcct gccacatggc ctccccactg tgccctggtt tccctaagta 53700
ccagcccaag gacacatgga taggaaaggt ggagctgggg caccagccca gtctgcctga 53760
ctccagagtc cctggtctta atcactaac caccacagaa aagtaaccgt gggagaagag 53820
acctgcaaac taggaaaaag aagattaaag ggaaggaatc tgttctgcta gatattaaaa 53880
catatgacaa agctgtagaa attaaaacag aatggggccg ggtgcagcag ctatgcctg 53940
taatcccagc actttgggag gccaaagtga gtggatcacc tgaggtcagg agttctagat 54000
cagtgtgacc aatatggtaa aaccctgtct ctactaaaag tataaaaatt agctgggcat 54060
agtgggtgac acctgtagtc ccagctactc tgcaggctga gccaggagaa ttacttgaac 54120
ctgggaggca gaggttcag tgagccaagc tcacactact acactccagc ctgggctaca 54180

```

```

gagcgagact ccagttcaaa aaaaaaaaaa aagaaagaaa aaaaaagaag aagaaaaaaa 54240
aaaaccgggc gcagtggtgc atgcctgtaa tcccagcact ttgggaggcc gagtggtgtg 54300
gatcacctga ggtcaggaat tcaagaccag cctggccaac atggtgaaac cctgtctcta 54360
ctaaaaacac aaaatcagca ggggtgtgtg ctgcatgcgt ataatcccag ctacttggga 54420
gactgaggca agagaatccc ttgaacctgg aaggcagagg ttgcagtga ccaagactgc 54480
gccactgcac tccagcctgg gcaacaagag caaaattcca tctcaaaac aaaacaaaac 54540
aaaaacaaaa acaaaaacaa acatagaatg ggtgctggcc tagaaagcca caaacagatg 54600
aatggagcac aacattaagt ccaggaataa actcaaacac acaggaaggt ctggtgaacg 54660
ataataggag gtaggtatct gtgaagcgtg gagtaagagt gatttactca accaacagtt 54720
ctgctgccac tggctagcaa gtaaatgcgg atccctaccc cactgctagt ctccaaaatt 54780
aattccaaat gagtcagtta aatgtttaa aaaccttcaa aagttgccag gcatggtggc 54840
tcacgcctgt aatcccagca ctttgggagg ccgaggcggg tggatcacct gaggtcagga 54900
gttctagatc agtgtgacca atatagcaaa accccacctc tactaaaaac aaaaaatta 54960
gctgggcatg gtcgaggcgc cctatcgtcc cagctactca ggaggctgag ccaggagatt 55020
tactagaacc caggaggcag aggttgacgt gggccaagat cacaccaccc aactccagc 55080
ctgggcaaca gactgagact cgttctcaga aaaaaaaaaa aaaaaaacct taaaagtta 55140
tagaaagtct gtgtgagtaa tttttaatat gaagcaaggg agagagatga agcagggttt 55200
ctaaacatga ctatagccat ataaaagtat gttataaagc tggcgctggt ggctcacgcc 55260
tgtaatecca gcactttggg aggttgaggc ggggtgatca cctgaggtea ggagtttgag 55320
accagcctga ccaacatgga gaaacccctg ctctactaaa aatacaaaaa ctagctgggt 55380
atggtggcgc atgtctgtaa tcccagctcc tcaggaggct gaggcaggag aattgcttga 55440
agtcgggagg tggaggttgc agtgagccga gatcgacca ttgactcca gcctgggcaa 55500
caagagcgaa actccgactc aaaaaaaaaa aatgttataa aaccacacac cactataaag 55560
aaatgataat gcaaaaaatc taaaggacga aaaaaaaag gaaatagatt taactacaaa 55620
aaagtttttg ttttgccttt gtttttttag agttagagtc ttgttctttt tcccaggctg 55680
gtacaatcat agtcactgc cacctgaac tcttgggctc aagcaatcct cctgcctctg 55740
aaactgcgtt tgcaaaaatt ataactgaga aaacgatgac agtgaagag atctgacct 55800
actgactcca tcttgcctct aacctccaag ctgtccgtgt tcattcctgg gtgtaggcca 55860
aactaacttt gggaggaatt tagtttatag tttaactttg tcaaagttta actaagatgt 55920
taatagccca ttttccaaaa caaacccctt tccgtcctgg ggactagact gcctttgcag 55980
gactaacaaa ttattatagc taccagatta gaaattatgg ttaggagtc atgcagctga 56040
agcctacaag attctgaatc tccaaaattg ctctggaga taacatcacc attgtaaaac 56100
ctaagatcag tgcttgacat attttgaga cctcgcactc gatggatcag ctggcactac 56160
ccaatggat aaacaggctc atctgatctg tggccccac ccagaactg acccagcata 56220
agaggacgcg tcaactcctc ataactttgt ctccaacctg aacaatcaac actcccctac 56280
tttctgacct cctaccacc aaattaccct taaaaacct agccaggcgc agtggctcat 56340
gcctgtaatc ccagcacttt gggaggctaa ggcaggcggg tcacctgagg tcagggttcg 56400
agaccaacca tggccaacat agtgaacccc catctctact aaaaatacaa aattagccag 56460
gtgtggtagt gtgcgcctgt aatcccagct actcaggagg ctgaggcagg agaatcgct 56520
gaaccgggga cacagagggt gcagtgaacc aagatcactc cactgcactc cagcctgtat 56580
gacaagagca aaactcggtc tcaaaaacaa aacaaaacaa aaaaccaca gaaaaaaacc 56640
ctgaacctat atctaaact ctttactat tgcaattccc ctgacttgat acattggctc 56700
tgtctaggca ggggcaagg ataaccatt gggcagttgc acctcagcct cctgagtagc 56760
tgggattaca aatgcaagcc acagctaaaa aaattttcaa acctttgtag gacagacaaa 56820
ttgggaaaaa catttgcaac aaagggatga tacacataca tataaagagt tctttcaagg 56880
ctgggcacag tggctcacgc ctgtaatccc agcactttgg gaggccgagg caggcagatc 56940
acgagggtcag gatttcaaga ccagcctggc caatatggtg aaaccccatc tgtactaaaa 57000
atacaaaaat tagccgggtg tggtagcatg cgcctgtaat ccagttact caggaggctg 57060
aggcaggaga attgctttaa cccgggaagc agagattgca gtgagccgag atcgaccac 57120
tgactccag cctgggtgac agagtgaagc tccatctcaa aaaaaaaaaa aaaaaaaag 57180
agttatttca aattaataag aaaaaatacc aacacaattc aatagaaaaa tggggaaaaa 57240
agaataggca ctttacaag aaataaatc aaagcccagt ggacatgaaa ctttgccatc 57300
tccttagcag ggcttgaggt aagatgggga gaaggaagga tgcaaaattt aaggaggctg 57360
tcactctcag ggcatgtaa gtacaaagt ggcatatgag ggtaagtgc tccttaaatg 57420
tgtaaatgac tagagccctg ctggatggca gtctggcaaa atggatcaat attttaaatg 57480

```

tacaaacctt ggcacaatga ttccattttt aggaactgac cttatggaaa cgatcaggca 57540
 agtgtgccaa gaaacacatc taggatgttt ttaatgtcga caaattagaa atgacaggta 57600
 aattcaaccc tacggactga ctttaaaaat tgttacatct ggctgggcat ggtggctcac 57660
 gcctgtaatc ccagcatttt gggagaccaa catgggagga tcgcttgagc ccaggagttc 57720
 aagaccagtc tgggcaacat agggagaccc cgtcgtaca aaaaaaaaaa aaagtaaaaa 57780
 ttagccaggt tgggtgtgca tgcctgtagt tctaactact caggaggctg aggaggagg 57840
 atcacttgag ccctagaggt caagactaca gtgaactgtg attgcgccac tgtactccag 57900
 cctgggcaat agagtggagg cctgtctcaa aaaagaaaaa aaaatgttac atccaggtac 57960
 attggcatcc tgtgtaaaaa ggatgccatc ctgtagtccc agctgcttgg gaggtgagg 58020
 caggagaatc gtttgagccc aggaattcga ggcttcagt agctatgttc acaccactgc 58080
 acttcagcct aggcaacaga gcaagacttt gtcaataaat taaaagaaaa ataaaaagta 58140
 cgtcactgtt ctatagtgtt cgtggaaaga ggctcagggt atgcatattt ggtacaagt 58200
 aacagaggaa ccaacatatg tcacatgata ctatttttgc caccctgccg tgtttatgtt 58260
 cacatttggg aatatgttgc caaagtaatg gtccctattt cctgggtgtg ggattaattg 58320
 caggggatcc ttactttctt ctttatgcct gctgcataaa tacttgaaat ccttaataac 58380
 tgcttaatac ttgaaaaagt gattaaagct aattttgtct gagaaagaga gtgggagtta 58440
 acctgttatt ctgtaacttc ctggccccac cagggttgac tcctgcagag cattctccag 58500
 gtaaatgttt ttgcccctgg ctgactgtat ttcagaacta ccaggaggtc gttttgttta 58560
 tcaaccaccc agtgggggtc aaaagacct taacttctac aattccagcc aaataaacag 58620
 aagtgtcttt cgaagtctta gggcctccca ttactaggat cagtgtgttt aggacttcag 58680
 ggtagtggaa agggccttgg tcccacagag ctgtctcagg gcacttaaat ttccctaagt 58740
 gtaaaatgga cagcttcaac cgtatcagt tttctcacct ttctcttttc tttcttttg 58800
 agacagggtc ttgtctgttt acccaggctg gagtgtcaat gcaagatctc agctcagtgc 58860
 cgctcacc acccaggcta atcaatcctt ctacctcagc ctcccagta actgggacta 58920
 caggcctgtg ccaccatgct tggctaattt tttgtagaga tggggtttca tcatgttgcc 58980
 caggctggta tcaaacgctt gggctcaaga gatcctcctg cccagcctc tcaaagtgt 59040
 gcgattacag gcgtgagcca ctgtgcctgg ctttttctta aactcactct cttttttat 59100
 aaagataaaa ttcttacacc ctctctagt ggtaccttct tccttattcc aatagccgag 59160
 aagatactgt ggaactttac tttctgtaga ttatatcacg aaaacaatag ttgtccccc 59220
 agctcatttt ccaaaattaa ataataatc taagtatgct tgtttgtaca cagtacagga 59280
 ctttctgaag ccacaggcca cctccagtc tggtcactga tgcctggggt ctttctctgg 59340
 ctctcaatta aaagctatag ttagtgact gagtacccca gtctggggac acaacctggg 59400
 tgagggtgc caggtaaaat acagggcgtt ctggggaggg tggcccacgc ctgtaatcct 59460
 agcactttgg gagggcaaaag tgggaggatc aggagttaa gaccagcctg gccaacatgg 59520
 caaacaatgt ctatactaaa aataaaaaaa ttagcctggt gcagtggcac atgcctataa 59580
 tggcagctac ttgggaggct gaggcacaag aatcacttga accaggaggg cggagtttgc 59640
 agtgagccaa gaccacgcca ctgactcca gcctgggcaa cagagcgaga ccctatctca 59700
 aaaaaaaaaa aaatatagat acacacacac acacacacac acacacacac acacacacac 59760
 acacacacat atggtgtcct ggaatctatt tcctagatct ggcaacccta acctagtcca 59820
 catttgggct tctgttcca ggcagtgtga ctataagcac agtctgtctt tccttttttc 59880
 tttgtctcac cctctttctt cttctttctt tcttccctcc ttgectgcct gctttctctt 59940
 tctttcattt ttctctctc cttctctccc ctccactccc tctctcttcc tctctttatt 60000
 cttacttcc tctctcttt tctctctct tttcttccct aattgtgtca agtgcataca 60060
 tcttaatttt aaatatgcag cttgatgaat tttacatat gcataaactc ctgcaaccac 60120
 taccagatt aaggagcacg tttccagcat cccaggaaat tttctcatgc ctcttgctgg 60180
 tcagtatctc cccagagggt aacctctt ctcacagcct gttattgtca attaatgtt 60240
 tatgttcttg aatttcataa aagtggaaat atgcaatatg agctcttaag tgtctggctg 60300
 cttcttctta acctaatgac tgagattcat tcaggttgct atataataca gtattttccc 60360
 ttttcattgc tgtataatat tccatttgtt gaatttttt ttggaggggg gagttttgtt 60420
 tcttgaatac accacaattt gtttatccat cctctgtctc atagatatat ggttgtttcc 60480
 agtttggggt gtaaatcaaa aataaaatcc taagggtcca ctaaatgaac acccttcttg 60540
 gcaaagggaa cccagaaaaa actttaaaaa ctttgtttcc agccatgatg agacaggagg 60600
 tcaggcacac cacattacac tcccttctt ccttttggg tttagataca agaaaagatc 60660
 agcatcaatg ctaaaataga gggtgagta tggtagtca cactgtaat cccagtcctc 60720
 tgggagactg aggaaggcag atcacttgag gccagaagtt cgagaccagc ctgggcaaca 60780

```

tggtgaaact ctgtctctac aaaataaaat aaaataaaat aaaataatta gccaggcacg 60840
gtggtgcgtg tctgtgtgtc ccagctactg gggaggctga ggtgagagga tcgcttgagc 60900
ccaggaaagca gaggtgcag tgagtcatga tctttccact gcactccagc atgggtaata 60960
gagtgagact ctgtctcaaa aaaaaaaaaa aaagagagag agattataag actgacagaa 61020
cagacttttt gtggcaataa gataccaaat tataaacaca gcctaaggcc atgtcaggca 61080
agggttaagt caggtgcccc tactcttaag gaataaacta tgttctaatt atgttacaag 61140
atttttcttt ttctctagca gcgaaacaag cactggcctc agaagaagca atattaaaac 61200
agttacaact catctagcac acagacaccc aactgacacc ctgttcctcc agtcataaca 61260
acaactacag ctttgattga acaagagact gagtttgta actttctcct aataaaaaga 61320
tcctgaaagc tggactgctt ctgggtgggt tacgaaaccg caacctcatg tgctgcatt 61380
tcctgaaagc acattttgat gtgtaggttc taattgtaat acattgattg attgattgat 61440
caattgattg attgagatag ggtcttactc tgttgcccag gctggagtgc agtggcacga 61500
tcacaactca ctgcaacctc tgctctctgg gctcaagcaa tctctccacc tcagcctccc 61560
aagtagctgg gactacaggt gcacgcaact gcgcccggct actttttgta tttttttag 61620
agacagggtt ttcgcatagt tgcccagct ggtctcaaac tctgggctc aagcgatcca 61680
cccaccttgg actccaaaag tgctagtatt ataggcatga gccaccatgg ctggcctaata 61740
tgtaatacat ttaaatgtta agtctccacc ccaaagtga catgggttgt atgttacatg 61800
cacatttgtt catacacatg tgttggggcc acctcataa atattcatag ctctctctgt 61860
aacctgctgg atatatcatt cagccaaccc cttcagcaca aagctcctaa cccaacccct 61920
cctccttcaa agtgcccgtc tctgttcttg gtaggaggca tacttcccag gccatggact 61980
ggtcaccttg tgggtcataa cccctataa gaaataagat ttcttctcct ctctgaattt 62040
acacatttgt gatttttttt ttttttttt ttagttaaca ggggctatga acattcttac 62100
agaagccttt tgattgatgt gtgttttcat ttatcttggg tatatatata ggcgtgggca 62160
tgatagatat taggatagcc atctttaact tcagtggatg ctggagcaag tttctgaatt 62220
tcaactctga agtggggatg ataataacag cacctgcctt acagggtctg ttcgagattc 62280
aaagagaaaa tctgggtaag gcagggtgcg gtggctcacg cctataatcc caccactttg 62340
ggaggccaag gtgggcagat cacctgaggt caggagtcca agaccagcct ggccaacatg 62400
gtgaaacctt gtcctacta aaaatagaaa aacaatgagc caggtgaggt ggtatgtgcc 62460
tgtaaaccca gccactcggg agtctgaggc aggagaattg ctgaaatctg ggaggcagat 62520
gttgacagta gttgagatgg caccactgca ctccagcctg ggcgacagag tgagactctg 62580
tctcagaaaa aaataaaaaa gaaaaaaaga aaatccaggt atttagaatt ggtacaccgc 62640
aatttacaaa acgtaaatta ttgctgtgat ggcagtgggg agcatgaaga tattggacta 62700
acttttatga atgttcaagt gctcccatga tgaattaaac acacagggaa ctttataagg 62760
gccatatgtt atataagtga tacatgacta ttgtattaaa attcaaaacta gttagatata 62820
aagtataaag tgggtttcac cctatccatt ttttattatt gaagaaaaaa aaatatgtca 62880
tagcgtgggt gcttatgcct gtaatcccaa cctttggga ggtcgggggt ggatgattgc 62940
ttgaggccag gagtgtgaga ccagcttggg caaaatagca agaccctgtc tttaaaaaa 63000
gtaagtaatt tggctgggtg ttatggcatg catctgtagt cctggctagg ctgaagcaga 63060
aggattgctt gagcgagga gttcaaggcg cactgcact ctggcctggg tgacagagtg 63120
agatcctctc tctctctctc tctctttttt tttttttttt ttttgtttt tgagactggg 63180
tctcactctg tcaccagagc tagagtgcag tggcttgatc ttggttact gcaagctccg 63240
cctcccagtt caagtgattt tcttgcctca gcctcccag tagctgagat tacggacatg 63300
tgccaccacg gccggttaatt tttgtattt ttagtagaga tagggtttca ccaacatgtt 63360
agccaggctg gtcctaaacg cctaacctca agtgatccat ccacctcggc ctcccaaagt 63420
gctgcgatta caggcaagag cactgcgcc tggcctgacc ctgtctgtta tcttttcttt 63480
ttcttttttt ttgttttctt ttttttttt agacagagta tcgctctgta gccaggctg 63540
gagtgtgcag tgggtccatc ttggtcact gctacctcca cccaccaggt tcaagcaatt 63600
ctcctgcctc agcctcctgt gtagccagga ttacaggcac accccaccac tctggctga 63660
ttttttgtaa ttttagtaga gacggggttt cgccatgttg gccaggctgg tctcgaactc 63720
ctgacctcag gtgatccacc caccatggcc tccc aaagtg tcagaattac aggtgtgagg 63780
cactgtgccc agccgacctt cttttaaaaa aggaaaaaat actatgcagt gagtattttg 63840
catgcatttt cttatttcat ctctgtcttt ttatttgatg atactaaagg cagggtttag 63900
aggctggatt gctaaagctg acccaaagaa tgctccctc agggctgggt ggtccctctc 63960
tctcaggcct cagtcttccc atctgtacag tgaggctcct gcagatctct gggctctaaa 64020
aatcacagct ccatgtttat ccttggcaga ggaaggcct ggagtctctc tgcttgcgtc 64080

```

```

tctgggatac gggagcaaag agccacgcat cctcatggcc cacacaggcg tcacctccag 64140
tctctccttg gcctcatctc ccagcgctcc tggaatggca tcgggctggc ccaggggagcc 64200
cctgtcctgt gcctctcctt tccccacagg ggctgccagg ctgaccaccc ccaccgcagg 64260
ccaggccctac agtgccccat ggaacgtcct gaccctcccc cagggtggca gcaggaagaa 64320
ggaagaaagg g gatcctctc cagctggcca gagagacaga cttctttgtg ctcatcaacc 64380
ctccaagaat gcctgccctc cctccttccc ccaaggcctg tcacagggg cttgagatca 64440
gccagaaaag tcaggcaact tttaggggac tgggagcgag gtctcccgcc cgggcctggg 64500
tccagtctct gtgggcagtg cagtgcggag cccccccct caagccgtgc cctgtccata 64560
gtccagact ttgacctgc actccagtc gggtggcg acagagggtt ggaacaaga 64620
cgctccagaa tcaggagctt cccctcagga aatagcatcc tgttcccc cactgcagtt 64680
gtctggcttc tccagcagtt tggtaacttc ggtgagtgcc agatgcacct ttgagctggg 64740
gacaggggtt gggagagggg agaggcaaag gatttcatgt cctcccaatg tcaaagacag 64800
ggctcaacat tacagcctaa agcaggtgac aggaaggag agatccagcc tctcaacat 64860
ccagcagaga gacctaggt aagtgtttt tccctccca agcctcagtt tcttcacctg 64920
gaacatggg atcataact cctcttaca gcgtgagtct gagtgttaaa agaggtggg 64980
catgtaaagt gcttagagca gatctaggca catagcaagt actcaaatgg tagttattat 65040
tatttttggg gggggagttg taggctggt tctcaactt ttatagcttc tgttccattt 65100
caaggataaa ctctgcaaat aacttcatga gaagtagccg tgtggtgcaa ccaggggagaa 65160
ctaattatgt tcattcaaat gcctcatctc tggttactg atttttttt ttaaaaagaa 65220
gtctttcata ttctttgcta tgggcacata gcaatcaaag gcatcagctg tctcagattg 65280
ccttctaggg gacaaggag gtcttaggca gataaatgca agactgaaag acaagcagaa 65340
agcatcaagt ggcaactgca tgcaactgc ctaaatattt ttttgagca gtgcagaaag 65400
cgccgataga actgggtcta ggtccgaatg ctgtcccata ctgactgcgt aaccttgggt 65460
gggtgacttc tctccctaa acctcagtc cagcctccag aatgaggcg gtaaccttc 65520
ctacttccta gagcagttga gaggattgag aggattatgt cggtactgca tctacaggtg 65580
tctggcaagt ggcagagacc aaaatacatt ggttcccttc ctgtccaca cttacacaga 65640
cattctaate acacacacac acacacacac acacacacac acacacacaa atataataat 65700
cccagctgtt tgcattctct gggatacata ctccaagctt gctgggttga agtaatgatg 65760
taaacagag gagaacggca acactaataa aaacatcagc aacaacacga aaatgtccaa 65820
ccgaataact gagctgggtg cgtttaagtc caaaagctca ttacctacac gcatgaatga 65880
ttttacctaa ggctggatct gccacatctg acaatctgtc tctggcttgt catgaggacc 65940
tcatgcattt attttgtatt ttaaaacaca cacacacaca cacacacaca cacacacacg 66000
ttgtataat cagtgtcaac ttgactcat atcttgaatt tttttaaaa aagataattg 66060
acttaggact cacactttt tctttttaa ttttttttt ttttttttt ttgacagagt 66120
ttactcttg tcacctgggc tggagtcaa tggcatgatt tctgcccact gcaatctcca 66180
cctcccaggt tcaagggtt ctctgcctc ggctcccg gtagctgaa tttaggcgt 66240
gcaccacat gccaaagtaa ttttttcta tttttttaga gacaggttt caccatattg 66300
gccaggtggt tcttgaactc ccgacctcaa gtgatctgcc agcctcgacc tccaaagtg 66360
ctggaattaa agacgtgagc cactgtgccc ggctttttg attttccatt ctattctac 66420
caacactcta aaaattccta caggcatttt attttattt attttatctt attttatatt 66480
atattttatg tttgaaatgc aggactctga agcttcagct gttcctattt accggcttga 66540
ttctcagatt tttcaaacca tgtgatttac tggcaagcat ggcatttaag cacctaggct 66600
tatgagtcag gctggcctgg gctctgcctc tcaccacctg ggtgtccagg agctgatatt 66660
ccagttagga gacaataagg caaggagctt gtgcagctct cataaaagt tatagatgag 66720
gtcgggcatg gtggctcacg cctgtaatcc tagcactttg agagtctgag gccagcaaat 66780
cacctgaggt cagaagtttg agaccagcct ggccaacatg gtgaaacct gtctctacta 66840
aaaatacaaa aattagccag gcatgttggg gcatgcctgt aatcccagct actcaggagg 66900
ctgaggcagg agaatacct gaacccggga ggcagagtct gcagttagcc aagattgtac 66960
cattgcactc cagcctgggc gacaagagtg aaaattcctg ctgaaaaaa taaagtttat 67020
agatgaggaa actgaggttc gattaggatt aaccaactca tctgtgttg cctgggactc 67080
tgatgcactg acttttagtc tgaaagtctg catcctggga ggacctcag cctgggcaa 67140
gtcggggagg ttgtcacc tcactcagtc aagttgagca acttcccag ggttacatgg 67200
ctggtgtgtg ccaagttag gctcgaacc tgggtctgtc tgactctcag cctgggcat 67260
actgtctctt agattcttca tggagaatta ggaaaaaac agaaagccct ttattctct 67320
gccttctcat tgttaacata taaaatggg caagcggcg ggtgcagtgg cacacacctg 67380

```

```

aaagcccagc gctttgggag gctgaggggg gaggattgct tgagcctagg aattggagggt 67440
ggcagtgagc tatgattgtg cactgcact ccagcctggg tgacagagtg agacctgtc 67500
tcttaaaaaa aagaaaaaga gtggtcagct ctccgaaat tatgcagaca gtcaaaaagc 67560
ccagagaggg gaattaactt agccaaggtc gcacagcaag gcagaagtga agccaggctt 67620
gactctgcct ttctcttctc ctctttttt tttgaggca gaatttcgct ctgttgccta 67680
gactggaatg cagtgggtgc aactcgactc gctgcaacct ctgctgcca ggttcaagcg 67740
attctctgc ctcagcctcc cgagtagctg ggattacagg cgctgccac cgcgcctggc 67800
taatttttgt agttttcagt agagatgggg ttccaccatc ttggccagac tggctttgaa 67860
gtcctgacct cgtgatccac ccgcctcggc ctcccagggt attgggatta caggcgtaag 67920
ccactgcagc tggctctccc tctctcttt tgttctgca atgtctttgt tctatgtgat 67980
ttttcaaaat gctaggagac aggaaggagg ctgctgtgtg ttgaggccct actctgtgcc 68040
aggcgtggta ccaagaactt ttgctaaact tcttatttaa tccttaaaat gaccctgtga 68100
gattgggatt aacctgttt tgcatgaa gagcttgtgt ctccagaggc aaagtatggg 68160
ggaagaggga agagagaaga ccaagggtcc ctgagagggg ctgtcaccta agccccagta 68220
tccaagctcg ggctcgaagc tggaaggaga attgcctaga ggaacgatac ctttctgtt 68280
gttggttcta tctcaactt ggcttctgaa accccaacag agtccagttc ttgtgggctg 68340
gagccgttt cctctcttta taaaactagg ccatattaag aatgtccgc tgtccagggc 68400
cacaggcccg agttgccagg agctgaggtc tgcgggagga gatttgtgag tgaagaggag 68460
ggaagattga atttggctct tctgggcaca aataattctc ttgttctgcc tcagcaggag 68520
cctgcagaat atttcctgc tgtcgggct taagtagctt caaggtaaa agctggtagg 68580
ccttctaacc ttctcagggc ccaatcagcc ctgtgcccc aggcaagggtg agttctgtgc 68640
tggaagacca agttctgagg ccagacactg cgtctgtcat gtcatagct gcatttgcta 68700
gctgccagcc tggcacatgg taggtgtgca ttaagcgtgt gttgagttta ctcaaattga 68760
aattaagtca cagctgtacc atttaactgg ctgtgtgact tcaggtaagt cacatcacct 68820
ctctgaacca cagtttctc ctctgtaaga cgggactgat aacagcagcc cctacctcat 68880
gagagtgttg ggagacttgg atgaatggat gcttgtgaag cacttagtgc cggggccagc 68940
tggtcacag taggtgtcc acaaatgtca gtatattact tcttttgcat caggcagctt 69000
gttaaatttg ttacgttttg catcttgttc aatttcccat ccateccctc aagcataggt 69060
tattagaggt tgaagcatct tgccaaaagt taacggccag taggtggcag agctgagtc 69120
tgaagccaga gcccatcgca ctaaccaccg gcctaccag cctacagttg gtcgtgccct 69180
ctgctgggtc ttttctattc ccagcccaga aactgggtgt ctggggacgc tcccagaga 69240
aagttgcatc attcaccagc cgtgtgactg tggccaagtc tcggtcactt ctccatacct 69300
cagtgatttc atttgcaaaa cgggaacaat gatattcct cctcctaggg gtcctcgga 69360
aggtcaata taaaagggtc ttgttggtgt ctggcacct ctaagcctc agtggatggt 69420
ggcaatggcg ctaaggatga tggagatgat ggtgatgatg gtgtgcctca acccttctt 69480
cccacaggct gctcaatgc gtgtgggtgt gattggagca ggagtcacg ggctgtccac 69540
cgccctctgc atccatgagc gctaccactc agtctgcag ccactggaca taaaggtcta 69600
cgcggaccgc ttcacccac tcaccaccac cgactggct gccggcctct ggcagcccta 69660
cctttctgac ccaacaacc cacaggaggc gtgagtagg gtcacatagg gtagcctggg 69720
gtgcccagtg acctaatgt gcagaggag tcagggttcc catcaccaag agcaagcccc 69780
ttgtggaagc tactgatcta gcataaaata aagaaaatgc caggcgtggt ggttcacgcc 69840
tttaatccta gcactttggg aggtcgaggt gggaggatca cttgaggcca ggagtccag 69900
atcagcctgg gcaacgtggt gaaaccccat ctctaccaa aatacaaaaa attagccggg 69960
catggtggcg cacacctgta atcccagcta ctccggaggc tgaggcagga aaaccattt 70020
agcctaggag gtgaagggtg cagttagctg agattccgc actgcactc tgacagagt 70080
agactctgtt tcaaaaagaa aaaaataaag aaaagattca taaatattaa gcccttgct 70140
ctgtgccaga tactaggagg ctttgtctcg tcttccctaa actgggtgcc tgtcaatacc 70200
acatgatttg tgaatctgga aaacttctc tgttttaatt tatacattt tatttattt 70260
ttgagattgt gtttactct tgtgccccag actggagtgc aatggcgtga tcctggctca 70320
ctgcactctc tgctccag gttcaagcga ttctctgcc tcagcttccc aagtaactgg 70380
gattacaggc atctgccacc acgcctggct aatttttga ttttagtag agatggagt 70440
tcatgttggc cagactggtc tcgaactct gacctcaagt gatctgcca ccttgacct 70500
ccaaagtgtc gggattacag gcattgagcca tcattgctt ccaaattta tctttttaa 70560
tagagatagg gtctactat gttccccggg ctggtcttga actcctgggc tcaagtgate 70620
tgcctctctt ggcctcccaa aagtgtgga attacaggca tgagccatca tgccttgcca 70680

```

```

aattttatct ttttaaatag agatagggtc tcaactatgtt gcccgggctg gtcttgaact 70740
cctgggctca agtgaatctgc cctccttggc ctcccaaagt gctgggatta cagggtgtgag 70800
cccttgacac cagctgaatc tagaaaactt ctaagtgggt gaacatctaa gtgggtggat 70860
ggatgcacag atttatcaaa taaattgcaa aggtcattat ggtagtttag aaactgccag 70920
atggttcagc aaatggaaca cccaatgaat agcagctcaa acagattaaa aaaaaatttt 70980
taagaggcat cctgtcacc caggctgaagt gcagtacat gatcatagct cattgcagcc 71040
ttgacctcct gggtcaagt gatcctccca cctcagcctc ccaagtagcg aggacacaca 71100
tgcattgctat catgcctgga taattttctt tattttttgt agagccaggg tcttctatg 71160
ttaccagggc ttgtctcaaa ctctgacct caagtgacct tcctgcctca gcctcctgaa 71220
gagctgggat tataggcatg agccactgca ccagccaga attttaattc acacagctgt 71280
aaaaaaccaa tgattctgat cagtgggcag tgattcgggg ccagcgctcc ttccatctaa 71340
tggtctctgcc gtttttccac gtgcttttga ggtcacctca atgtcacctc tcacatgggc 71400
tggtgactgc agaaggatca tgcaggacca cacgtgcaga gtcttttagag tcccttggcc 71460
agaacaagt cacacagtca catttagctg cgagagggtc taggaaatgt aggtgagctg 71520
tgtgcccagg gggaggagga aaagttgtag gagcggaag tcgatgtctg ccaccagtgt 71580
ttacaaggag ggggtccttg agccagactg aacagtgtgg ctcataatcc ccaaagccag 71640
gtcaaggact tcaactgaac tcatcagcca tgtaatccca tgctggaggt gcactccata 71700
tggttatgat ggggcatcct tcattccctc tcttctttat tctattaatg gggaaatatt 71760
ggaaaattta ggaggagaa gacccaaggc atttggggag ttgcaggagt gaacgtggtg 71820
gatttctggg ttttgacac accccaagct cctgatcatg ccacagcccc atgccagctg 71880
acctaagggt ttttcccag ctccaggcat tgggtgatcg aactcttcat gacccttcca 71940
gggactggag ccaacagacc ttgactatc tcctgagcca tgtccattct cccaacgctg 72000
aaaacctggg cctgttccca atctcgggt acaacctctt ccatgaagcc attccggtgg 72060
gtgaacagtt cttgacctg agggatgagc acccagggtc ggggtagtga ggggtgggtg 72120
agcagagcct taatcacaga tgaggcgagg gtgctttgag tctcgtaggc aacagactcc 72180
tggtttcaaa acaggttttg tttaaattct atttttgctt tgaaaattat ttttgtttta 72240
cattttgctg ttaaatggc agaggacaaa gaatcttctg atgccaggg gaaactagcc 72300
tttgattagc atggctaaaa tacaacatg ttctgcagt acgggcactt ggtgctgaag 72360
ccaaaagggt tcaagtggc ctgaaggctc caaggtttt tatcaagaag gaataaaata 72420
ctcatcaag caaaaactgc caaagcattt attatgtgcc aggtccagtc ctaactaatt 72480
tacagctagc gactaattta attctcttta taaccgggag gtaagggtg tccttatcct 72540
cacttaataa atgagaaaac ggaggtctca agaatggag taacttgccc aaggccacar 72600
agctcgccag tggcagagct gggatttgaa ccagggccat ctgtgactcc atggtgtcca 72660
gtgtgctaac aggaacagca cagccctggg acggtttgct caggctcctt ggagagggtg 72720
gtctggcgct gtgcccagac ccccggtgca gctctcaagg ttcattcaac ctttggcact 72780
gtgctaaggg ctttatccac attatctgtt acccttcatg ggaccaagag tatttttttt 72840
tttttgagac aggtctcac tgtattgcc aggtggagt gcattggcat gatctcggct 72900
cactgcaacc tctgcttctt gggttcaagc cattctcctg tctcagcctc ctgagtaatt 72960
gggattacag gtgcgacta ccacgcctgg ctaattttt ttttttaag agatggggtt 73020
tcaactatgac gcccggtgct gtctcgaact cctgacctca agtgaatctg ctgccttggc 73080
ctcccaaagt gctgggatta caggcgtgaa ccaactgcacc cggccaagag tgattattaa 73140
ctccatgata cagacaagga aactgagtct cagagaattc aagtagcaag tgatagggt 73200
ggggtctctg acactatgct ctgctgtctg acactatgct ctgttgcttt ctctcatccc 73260
cggggactct cactgtttct gctttctctc ccctatttct gacttttccc ctataactca 73320
ccctcggtct tactcttacc cttaccataa atagggggtta agaacatgaa ctctggaact 73380
aagctgtatg ggttaaaatc tcaacaccac catttattag ctgtgtaatc ttagacaagt 73440
tatttaactt ttctaagcct caattggtcc atctgtaaac tggggaaaga atagcatcca 73500
cccaaatggc ttcttgtga gattaaatgg accagtataa gaaaatgctt ggaacagtgc 73560
cttatatgca cttagcatta cataagtctc tgtcattatc attttttttt ttttttgaga 73620
tgaagtctcg ctctgtggcc caggctggag tgcagcgga caatttcggc ttactgcaac 73680
ctccagctcc cagattcaag caattctcct gcctcaatct cctgagtatc taggattaca 73740
ggcatgcacc accatattct gctaattttt gtattattat ttagtataaa cagggtttca 73800
ccatgttggc cagactggtc tggagctcct gacctcaggt gatccacca tctcagcttc 73860
ccaaagtgtc gggattacag gtgtgagcca cctcgcctgg ccattatca ttattattga 73920
cttccatccc acccagtgcc cccttgtcc ttctcttcca ggaccttcc tggagggaca 73980

```

```

cagttctggg atttcggaag ctgaccccca gagagctgga tatgttccca gattacgggt 74040
gagtttattg tcacaggcaa aggggactgg ggccctgacga gtttagcagac ctgtccagaa 74100
ggcagcagag ggtagaggca ccagatttec tgcctaccc aggccttggg accctgggtct 74160
cctgggtcctt ggtccagctc ctccagagag gctaccact caaacctggc cttgggctgg 74220
gaaggtaggg ggtatgaaat cacagatctc aagcccagaa gctccatc accatattgt 74280
ttttagatg aagatactga gtttcagaga ggctaagtga cttcctaagg tcacacagcc 74340
aagtggccaa actgggattc caaccagtct gtatgacccc acaccctcc tttctttct 74400
ctacagcctg atgcctctct ggtcttctcc tcaccccacc ccacaccaca cctgaatccc 74460
ctctacgaat gcacattcaa tctccacttg cattttccaa tgtcagatat ggctttttct 74520
gatagaaaaa ttttccttgc attgagctca aaaccacgtc ccccttgaa cttcacgtag 74580
tggctcctgc actacccttt ggcccacag aacaacattg cttccacctc catttcacag 74640
ccttcaaata tagctctgat ttttaccctt atttccacct ttgtcttact gtgactctag 74700
ctatggcttg ttccacacaa gcctaattct ggagggaag aactatctac agtggctgac 74760
tgaaagtgat gattttaagc ttcactttga gggaggtacc tcccagagac caagttgtag 74820
tggaagatgg ttcgtgggct tccctcagca tggactaacc cccaggtttg aagaataccc 74880
ttaggcctgg tgtgggagct atccttggtc ctgatcaccg ctgggcacag aggcaatgga 74940
tcctgagcct agctgagcat cagaaccacc tgggcagctg ttacacatg atgtccatta 75000
acaacctctt tcaaatccct aatgtttgtg taatagtttt agatgtattc ttttaaggtt 75060
tccagataca ttacatcat ctgcaaaaca taagctgcct tttattttta tctccctctc 75120
tctttttttt tcttctctaa ctgctttggc caatacctct agaacaatgt tactcataca 75180
gatgatagtg gatgtctttg acttgttctc catgttaata ggaatcttgc agtgttctaa 75240
attagcaaac actcacttga gagatacatt ggtatttata ttacataca ttcatattaa 75300
gggagattcc ataccttttt tgtgtgtgtg agatggagtc tcgctctgtc acccaggctg 75360
gagtgcagtg gtgcgatctt ggctcactgc aagctctgcc tcctgggttc atgtcattct 75420
cctgcctcag cctcctgagt agctgcgact acagggtcct gccaccacca cacctggcta 75480
atTTTTtGta tttttagtag agatgggttt tcaccatgtt agccaggatg gtctcaatct 75540
cctgacctcg taatctgccc gcctcggcct cccaaaatgc tgggattaca ggtgtcagcc 75600
accacgcccc gcctgattcc atacattttt tatatattac tgttttttaa agatttttag 75660
gccaggcatg gtggttcata cctgtaatcc tagcactttg agaggccgag gtgggcagat 75720
cacttgagcc caggagtcca agaccagcct gggcaacatg gcaaaacct gtctctacag 75780
aaaaattcaa aaatcagcca ggtataatgg tgcattgctg tagtcacagc tacttaggag 75840
gctgaggtgg gaggatggct ttatcccggg aaggagaggc tgcagtgagc tgtgatcatg 75900
ccactgcact ccagcctggg gggacagggc gagaccctgt ctcaaaaaa aaaaaaaga 75960
ttaaaaaaat atggaatata tagtggcttt tatcagatga cctcagaaga ttttttttaa 76020
atgtagattt taggaccccc ctctatacct gctgaattag aacttctggg ataaggttca 76080
taaatttgct ttttttcatt tttttgagac aaaatcttac ttgtcacc aggctggagt 76140
gggatgtagt ggtatgaaca caactcacag cagcctcaac ttctgggct caaatgatcc 76200
tcccacctca gcctccaaag tagctgggac cacatgcatg tgccacaatg cctatctaat 76260
ttttaaatat tttttagag atagggtctc actatgttgc ccaggctggt ctcaaaccct 76320
tgggctcaag caatcttctc gcctcagcct cccaaagtgc tgggattaca ggcgtgagca 76380
aacaggccta gcaaaaattt gcattttaag aagcttctg gcgattctaa ttatcagcca 76440
tgtttgggaa tcattgtact aagacatggc tatttctctc aacctgggga cacatgacct 76500
ttgtccagtc tttccagga aaaacatgcc ctcaagatgt ttttctatct tgaggaaatg 76560
atggaaatga gatagttcca aggttatgct tcaccttctt ttggcttat ttcctgttct 76620
ttggatgttt ctagtgtatt tcttcttctc ttctttttt ttttttttt ttgagacag 76680
agtcttgctc tgtcaccag gctggagtgc agtggcgcaa tctcggtca ctgcaagctc 76740
cgctccttgg gttcatgcct ttctcctgcc tcagcctccc gagtagcttg gaatacaggc 76800
gcctgccacc acgtctggct aattttttt ttgtattgtt agtagagac gggtttcacc 76860
gtgttagcca ggatggtctc aatctcctga cctcgtgatc gcccacctc ggctcccaa 76920
agtactggga ttacaggcgt gagccaccgc gcctcgtgt ttctagtgtt ttctaatcg 76980
tgatagatgt ttttctatg ggatgtttaa aaggagggtg gatgtcctca gccacctcc 77040
ctcctcatgc cggcttctg acaaggggga atttggcact ggtacaactc tccccttctc 77100
tactctgaat ctcatgctt ttgctgttac aaagcaatgt ggtggtcata ggaagtgtg 77160
ggggctaaga ggctgggtt tgagtccaa ctccatcatt gactcactct atggccttca 77220
gcaaggccct tccccactc catctgccc acaaggggct tggaccatct ctggtttctc 77280

```

```

aaaggagatt ttgtggacca ccagtcacgt aggtgctcat gagctgattt gatgacacag 77340
ccatcttctc aagcagcacc ctgtgcaact aacgtccgca gaaggttgtt tggggaaagg 77400
tccctgtgcc acccttcttg gtgggatggg ggcagatagc tgaatactgg gctttttgat 77460
gtgtttgatc atcccagggt aactgagagg ggagtgaagt tcttccagcg gaaagtggag 77520
tcttttgagg aggtgagttg cagggtgatg gcggtggatg gggcagggaa gaagtaggga 77580
ggcctctgct tcttctgctt gagtcggggg ctcccttctc aggcctcctag ggtccccaca 77640
ggcctgcctc agcacccttg cccagaagc actcaggtat tctgaaggga ggaagtctct 77700
gccttcatgt tggtagtggg aacaaaggaa cactgggatc atggtggcca ttaggagctg 77760
atttatatct gagactcaat gagttttggg tctagagagc tggccgcatt ttctcagtgt 77820
cagctgcact ccaaggctag aacttgggtg ctccctagcc ctaccgacat ctgtgttggg 77880
ctttctgcaa agtccaggcc ctcagctgac tcacctctaa agaagcacca ccaccaataa 77940
taatgacagg aaaagccacc atctccagc accagcaaaa agagctttac tgtatggctt 78000
cattcaatcc cagcatctaa aacctgctt ggcacaagga aggcgctccg tacatgtagc 78060
tactagtgtt atgtcatgaa gactaacctg ctctggtcag gccctgatgg acaccgaaga 78120
tacatggctg acccaatgca gtccctattc tcagtcattc actcaggaac aatagtagcg 78180
tcttgcaatg tgtgtgtccc ttaacttact cgtggtgaga gtcactgggg ctgggttggg 78240
gagcttaggg gctcacgatg cgtgcttgag atgagatcat ctcatctgta gacagagctg 78300
gggttccaac gtgtcttctg caaatgtctt ggcagagtag aaggcaagag aataaagtta 78360
aaaggagtca gaaggagaaa gagaactctc tctgttctt ttctgacttc ttttgggagg 78420
ttccaggaag atttccccca tccaaagaac tgttttacia ccacttttat attcagagtt 78480
gtgcaggagc ctcataacag cctatgaaca gccatgggca gcctcatitt acaggggcag 78540
ctgagaatta aggaggtaac cagacatttt caaggtcaca cgtcagataa atggcagtat 78600
gaaaatttga agccaggccc ctctgattcc tcattgagac ctctccccac tgttcatcag 78660
ggagtagaca gattgagggg agaagaaagg ggaagagaga acaggggata ccagggtctt 78720
ccccaccttt catccccac tacctgttg gtgtctacca ggtggcaaga gaaggcgag 78780
acgtgattgt caactgcact ggggtatggg ctggggcgct acaacgagac cccctgctgc 78840
agccaggccg ggggcagatc atgaagggtg gtgtgagggg gagacccta ctttttgtta 78900
ataggaagat cattctgcat gcttatttca tccctcaaga tcatggacaa atcaggaaca 78960
tctgttagag gaaccccccg gactgcaggg aattgacatg taaaaaaac aaacctgtcc 79020
cacccecatg gctctcttcc aggatcttct ctgtatcgtg aagcatgcat gtatgcgctt 79080
gtacctatgt gggagcagca tatgcttcta ttgcaataaa aatagcaaac attagagtgt 79140
ttaccaagcg cgagatacag tccaaagcac tttattgtgt ttattattat tattaattat 79200
taattgtgtt attattatta tcattgttat tattattttt gagacagggt atcaactccat 79260
tgcccagggt agagtgcagt atcttgatca tggctcactg tagccttgac ctcccaggct 79320
cccaccttag cctactgagt agctgagact acaagcgcat gccaccacca tgcctagcca 79380
atttttttat tttttgtaga gaaaggattt caccatattg ctgaggttgg tctcaaaact 79440
ctgggtctaa gtgatcccc cacttggcc tgtcaaagtg ctgggattac aggcgtgagc 79500
caccagcttc agctattgt gtttaattat ttagtgtgg ccacagccct tcgagctggg 79560
tactaccata tcttattgt catcttacag atgaagaaat tgaggcacag aggagttaag 79620
taacaggcac aagttcacac ggtagtacgc agtgcaattg ggattggaat ccaggcaacc 79680
tggcttaaga gcctgtgcgt gcaagcattg ttccatgcct cctcttctg tgtgtgtgca 79740
tatgagggtg tgtgtgtgtg catgtatgtg tgtgtgtgta tgaagggtg tgtgtgcata 79800
tgtgtgtgtg catgtgtgtg ggtgtgtgtg catgtgtgtg ggtgtgtgtg catgtgtgtg 79860
agggtgtgtg catatgtgat ggtgtgtgca catatgtgat ggtgtgtgtg catgtgtgtg 79920
agggtgtgtg catatgtgtg atgtgtgtg tgcacgtatg tgggggtgat tgtgcatgta 79980
tgtgagggtg tatgtgcata tgtgtgatgg tgtgcgtgca tgcacacat gtgagggtgta 80040
tgtgtgtgtg catgtgtgtg aaggtgtgtg cgcattgtgag agtgtatgta cgtgtgtgat 80100
ggtgtgtgtg tgtgagggtg tgtatgcatt tgtgtgaggg tatgtgtgtg catgtgtgtg 80160
agggtgtgtg catgtatgta aggtgtgtgt tacatgcatt tgtgtgaggg gtatatgtgt 80220
ggatgcattg gagggtgtgt gtgtgcatgt gtgtgaggg gtgtgtctga ggggtgtgtgt 80280
gtgcatgtgc acctgtgagt gttcataggt gtgcaggtgt gtgtgcttct gtgtgtaggg 80340
gtgcgtgtgt gtgttcttaa tgtgggtgta tgggtgtaac aaccaaatga gtgactgaag 80400
cataagtctc aaatcatcga ggtttatgga gccagcttga gggcgacccc agggaaaacg 80460
cgagtcacag atgcacctgt gactcctttt tccaaagagg ttctcaggag atttagtctt 80520
tatacatttt ctttaaaaaa aaaaaagtga gagaagggtg tagcagcgag agaattgattg 80580

```

```

catacttgtg aaacttttagt tagtcccag taaatctaca ttttacataa gatgaaggtt 80640
tgggccaggc gtggtgactc acacctgtaa tcccagcact ttgggaggct gaggcaggtg 80700
gatcacgagg tcaggagttc gagaccagcc tggccaacgt ggtgaacccc catctctact 80760
aaaaatacaa aaaattagct ggggtgtggtg gcgggtgcct gtaatcccag ctactcggga 80820
ggctgacgca ggagaatcgc ttgaacccgg gaggcagagg ctgcagttag ccgagactac 80880
accactgcac tccagcctgg caacagagcg aggcctgtctc aaaaaaaaaa aaaaaaaaaa 80940
aaattgaagg tttgaaggaa aaaggaatgg aggaagtctt gtatctggga agataagctt 81000
gtcattgatg ttatcagtgt ggagctctgtt gaaagggctg gtttctgtt accccttagg 81060
gaagaaagcc taactttggt caggctcattg agggagggga tataatgaga cgtgtcggac 81120
ctcccttccc cccgcagctg tgaactcagc tccaaggttt ctctggggct cctggggcca 81180
agagggggctc tgttcagctg gttggggact tagaatTTTA tttttatttc tcatgtgtat 81240
gcattttacat ggtgttactg gtgcttttct tcggacatgt gggtagaggag aaacaatgct 81300
tcaggggagca ggggtggctg ccaattaggg cagctcttcc tgcaagaggc aagcagtcag 81360
gtgcagactt gggccatagt gtcagttagg gtcttataag gaatcagcct ggccactctt 81420
gtcaggacat ctggccacag aggggagcaa gggcagccac attgactcac ctccgctgat 81480
gagactttcc tgccctgaat caacaggtgg acgccccctg gatgaagcac ttcattctca 81540
cccatgaccc agagagaggc atctacaatt ccccgtagat catcccaggg taaaattgga 81600
ctgttctcgg gcagaagagt ggtccccctc atgccctctt catgacctg ctgctcccc 81660
caagctcctt actccctgca gttgttcctt ttcaatgttt ttatgtactt agctattttt 81720
tattattatt ttttgagaca gagtttactt cttattgccc aggcctggagt gtaatgggtg 81780
gatcttggct cactgcaacc tctgcctccc aggttcaagc aattatcctg cctcagcctc 81840
ccaagtagct gagattacag gtgcccacca ccacatccag ctaatttttt gtatttttag 81900
tagagacagg gtttcacat gttggccagg ctggtcttga actcctgact cagggtgatcc 81960
acctaccctt gcctcccaaa gtgctgggat tacaggcgtg agccaccgtg cctggccccc 82020
tttcaatggt tttagttagt ttgagctact gaatccctgg gaaggcagac tcagcctcga 82080
ctgagggtcta ccgtgaacat tcttttggat gacaatagtg gtgatgctgg agacaaggc 82140
agtggatgta atgtggtgac actaaaagtg gtatgtaggt ggctcacgcc tgtaatccca 82200
gcacttttgc aggccaatgt gggaggattt cttgagccca ggagttcaag accagcttgg 82260
gcaacatggc aagacccctg ctctacaaaa atacaaaaat tagccgggcg tgatggtgta 82320
tgccatggtt cccagctatt cgagaggctg agatgggagg attgcttgaa cctgggagggt 82380
tgaggatgca gtgagccatg ttcacaccac tgtactccag ctggggccac agagcgagac 82440
cccatctcaa aaaaaaaaaa aagtgggtgtg aatggcaata atgggagtg gaatgggaat 82500
ggtgatggg gctgatggtg atgataatgt taacgggtgga gatgacaatg tcaactgaaac 82560
cagtgggtgt gttcatggga tgacaatatt gttgatagcg gaatgggtgt attagggata 82620
atatgtgatt ttgggggaag acagcgttca tgggggtggt gattagcgta agagtgttag 82680
agtgggtgat ttaattggagg tggctctggtg ctgatgagga gatcaatgtt gatgaagggtg 82740
tgattgggag tggggatggt agctgggtgt gatggaaatg acactatcaa tgatgttaat 82800
actgtagcag agctgacagt ctcaaaggca atgttaataa catggttgca ccaaccatgt 82860
tatctcaatg gcgatgttac tgggtctgtg gagatgacaa tatcaatggc aatgttagtg 82920
gtggtggtga aatgatgaat gcagtgtgtg gtgatgacct attaatgata gtagcaaaga 82980
caatgttgtt gatggagatg acaacattga tggaaagtgt gatggaagag ttcgttgttg 83040
gtgttgatgg tgatgacagt ggcaatttag gtatgtatgg tgggtggtgtt agcagagggtg 83100
acaaggttga tggtaatgac ctttattcat ctgagagcct tcattttcct tcatccttga 83160
ccctctcat ttttatctag gacccagaca gttactcttg gaggcactct ccagtgggga 83220
aactggagtg aactaaacaa tatccaggac cacaacacca tttgggaagg ctgctgcaga 83280
ctggagccca cactgaaggt aaggtaggga ggagtagcag tgccttaaac caaggtcgtg 83340
ggagcttgggt aatgaggaca cttcaggacg ggaagatgcc accgctggga taactgggca 83400
aattaattcc agcaagggat gtggaacata acagaatttg ataatgtaca gggaagttct 83460
tgctatgggc taatgaatcc tgtctggcca tggctgagag cccttggttt tcacatttgt 83520
ctgcgagtga tgatgacagt agttaggtgt atgaggatga gttgttactg atggtgagga 83580
aaatgctgag aatggtaata gtgatgtgta taagggtgtg acagttgtta aaattatggt 83640
ggtggctgat ggtgagggtg gtggttgatg atggaattgg tggaaagggt gaagcagtaa 83700
tggtaatgat gttggttagt gataaagatg gtgttggtgg tagtggtgat tgataaagat 83760
gactgtgatt atattagtgg tgggtgtgat gagattctaa aagctaactc cctactacct 83820
aaaaatggca gcaggaaaaa aaaatccaga aatgagtgat cagcactttt ctttccagaa 83880

```

```

tgcaagaatt attggtgaac gaactggcctt ccggccagta cgccccaga ttcggctaga 83940
aagagaacag cttcgcactg gaccttcaaa cacagaggta tgctcccatg gcaaggaaag 84000
taatgccctc ttccactcct cagatggctc tggcattttc agggarcagt catgtctgat 84060
ctcaagtccc acacaggctc catagcaggc agggggcagt gtggctaata tcccctcctc 84120
tataaatggg gaaactgagg ctcaatgatg gttaaggacc tgctcaaggt tacatagagg 84180
ggcagtgggt atgttaatgg aggtgggtgt gatgagatca atgttgataa tgggtgtgact 84240
gggagtgggg atggtagctg gtgctgatgg aaatgacact atcaagtatg ttagtaccac 84300
agcagagggt acgatctcaa aggcagtgtt aacatggctg cactaactgt ctcatggca 84360
atattaatcg tgtggcagag atgacagtat caatggcagt gttaatgatg gtggtgaaat 84420
gggtgaatggg gttggttttc taaagtctgt ggtcaaataa caggaaaatg tgtacttact 84480
ggatgtgtac ttcgtgtcag acacagcagc aagtccatta catgaatgac cttattaaat 84540
ctctcttggg gctctttggg atagggacag ttctccctat gcttcggatg aggaaactgg 84600
gggtgaattaa gaggtgaagt cacttgccca agtcagacca ctggtggaag gcagggctgg 84660
gatgtgattt gaatttgact ccaaggctat ttccagatat ccattttgtg gctgccccat 84720
catctcttgc aactgttcca ggggggtcccc accattccac cccggtgcca agagaagctc 84780
agggtggcatc tggttttggc caggactctt cggggaggctc ctgagtcttc cagggcagaa 84840
gagcttcac tattctttcc actgtccctc toggacctgg ccacctctc tcttgctct 84900
cctaggtcat ccacaactat ggccatggag gctacgggct caccatccac tggggatgtg 84960
ccctggaggc agccaagctc tttgggagaa tcctggaaga aaagaaattg tccagaatgc 85020
caccatccca cctctgaaga ctccagtac tgctgcctcc cccacaaga actccttct 85080
cccctcagcc aatgaatcaa tgtgtcctt cataagccat tgcttctccc tcaactcttt 85140
cctcaaagaa gcatgaggtg agagaaagcc acaaagtcag tgcttgaga agggttcagc 85200
ccaacatggg gccctctca tcaactgaaat ccctctacct tctctgggtc tggcattata 85260
aagaacagct gaggtgtca ttccatgagt ctccagaaga aaggacagct cagaaaatca 85320
aagaggccaa ctgcccagag ccacagaaaa tggaggataa ttgaggctaa gtaacctgat 85380
tacaagttgt actaacatat taaaggttct gaaaagtcct gcagcaaaga caactatctg 85440
atgttgttta acccagtgct tgctaaacct atctggctat ggaactctt tgcccagagc 85500
acccatgaat gccatgacac aaatctgaga aaatgctgga acagattttg ttgtatctgt 85560
tgtgtttgtt gtaggaggtt atacatacaa ctgggggtgtg gagggggcag agaggtgagg 85620
cactgaacta gtaacacatg gtgtttgttc cacatctaga attccaaatg gcatcagcta 85680
ttcaccgagt ggcccatga gccaccagta acctttgagg aggggccact ggagggatca 85740
tcccacaagg aaccccttca tagagaactg ttttagtcca ttttctgtt cttataacag 85800
aatatctgaa actggagatt ttttttttt ttttttgaga caggatctca ctctgtcacc 85860
caggctgggtg tgcagtggca tgattttggc tcaactgaac ctccgcctcc caggctcaaa 85920
tgatcctccc tctcagcca cccgagtagc tgggactaca ggcgcttgct accatgcca 85980
gctaattttg tgtgtgtgtg tgtgtgtgtg tgtgtttgt agagagtgtt ttgtagagac 86040
tgggtttggc catgttgtcc aggttggtg tgaactcctg ggatcaagt atcctcctgc 86100
ctcagcctcc aaagtgttg gattataggc ataagccacc acgcctggcg gaaactgtgg 86160
aattaataga gaaaaggaat ttatttatta ccgttataga gtctgagaag tccaagggtg 86220
aggggccaca tctggtgaga gccttctctc tggtgtgtgc agagggtggg actctctgca 86280
gagtcaccag gaggttagg gcatcacgtg gtgagggggc tgattgtgct aatgtgctag 86340
ctcagctctg tcccttgtct tagaaagcca ccattttcct tcccagatg acccattaat 86400
ccattaacct aataacccat taattgataa atggattaat ccatttatga gagcagcgt 86460
cttaggatcc aatcacctct taaaggccc acctctccag accaccacta aggtgttgga 86520
ctaaggacta agtctcaacg tgagttttgg cagggacgtt taagcaatag caagaactaa 86580
actaccaag ca 86592

```

<210> 2

<211> 1573

<212> DNA

<213> 智人 (Homo sapiens)

<220>

<221> 5' UTR

<222> 1..143

```

<220>
<221> CDS
<222> 144..1187

<220>
<221> 3' UTR
<222> 1188..1573

<220>
<221> polyA_signal
<222> 1549..1554

<400> 2
tgcactccag tccgggctgg cggacagagg gctggaaaca agacgtccca gaatcaggag 60
cttccccctca ggaaatagca tcctgtgtcc ccgcactgca gttgtctggt ctctccagca 120
gttttggtact tccggctgct gca atg cgt gtg gtg gtg att gga gca gga gtc 173
Met Arg Val Val Val Ile Gly Ala Gly Val
1 5 10
atc ggg ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc 221
Ile Gly Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val
15 20 25
ctg cag cca ctg gac ata aag gtc tac gcg gac cgc ttc acc cca ctc 269
Leu Gln Pro Leu Asp Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu
30 35 40
acc acc acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac 317
Thr Thr Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp
45 50 55
ccc aac aac cca cag gag gcg gac tgg agc caa cag acc ttt gac tat 365
Pro Asn Asn Pro Gln Glu Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr
60 65 70
ctc ctg agc cat gtc cat tct ccc aac gct gaa aac ctg ggc ctg ttc 413
Leu Leu Ser His Val His Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe
75 80 85 90
cta atc tcg ggc tac aac ctc ttc cat gaa gcc att ccg gac cct tcc 461
Leu Ile Ser Gly Tyr Asn Leu Phe His Glu Ala Ile Pro Asp Pro Ser
95 100 105
tgg aag gac aca gtt ctg gga ttt cgg aag ctg acc ccc aga gag ctg 509
Trp Lys Asp Thr Val Leu Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu
110 115 120
gat atg ttc cca gat tac ggc tat ggc tgg ttc cac aca agc cta att 557
Asp Met Phe Pro Asp Tyr Gly Tyr Gly Trp Phe His Thr Ser Leu Ile
125 130 135
ctg gag gga aag aac tat cta cag tgg ctg act gaa agg tta act gag 605
Leu Glu Gly Lys Asn Tyr Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu
140 145 150
agg gga gtg aag ttc ttc cag cgg aaa gtg gag tct ttt gag gag gtg 653
Arg Gly Val Lys Phe Phe Gln Arg Lys Val Glu Ser Phe Glu Glu Val
155 160 165 170
gca aga gaa ggc gca gac gtg att gtc aac tgc act ggg gta tgg gct 701
Ala Arg Glu Gly Ala Asp Val Ile Val Asn Cys Thr Gly Val Trp Ala
175 180 185
ggg gcg cta caa cga gac ccc ctg ctg cag cca ggc cgg ggg cag atc 749
Gly Ala Leu Gln Arg Asp Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile
190 195 200

```

```

atg aag gtg gac gcc cct tgg atg aag cac ttc att ctc acc cat gac      797
Met Lys Val Asp Ala Pro Trp Met Lys His Phe Ile Leu Thr His Asp
    205                210                215
cca gag aga ggc atc tac aat tcc ccg tac atc atc cca ggg acc cag      845
Pro Glu Arg Gly Ile Tyr Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln
    220                225                230
aca gtt act ctt gga ggc atc ttc cag ttg gga aac tgg agt gaa cta      893
Thr Val Thr Leu Gly Gly Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu
    235                240                245                250
aac aat atc cag gac cac aac acc att tgg gaa ggc tgc tgc aga ctg      941
Asn Asn Ile Gln Asp His Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu
    255                260                265
gag ccc aca ctg aag aat gca aga att att ggt gaa cga act ggc ttc      989
Glu Pro Thr Leu Lys Asn Ala Arg Ile Ile Gly Glu Arg Thr Gly Phe
    270                275                280
cgg cca gta cgc ccc cag att cgg cta gaa aga gaa cag ctt cgc act     1037
Arg Pro Val Arg Pro Gln Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr
    285                290                295
gga cct tca aac aca gag gtc atc cac aac tat ggc cat gga ggc tac     1085
Gly Pro Ser Asn Thr Glu Val Ile His Asn Tyr Gly His Gly Gly Tyr
    300                305                310
ggg ctc acc atc cac tgg gga tgt gcc ctg gag gca gcc aag ctc ttt     1133
Gly Leu Thr Ile His Trp Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe
    315                320                325                330
ggg aga atc ctg gaa gaa aag aaa ttg tcc aga atg cca cca tcc cac     1181
Gly Arg Ile Leu Glu Glu Lys Lys Leu Ser Arg Met Pro Pro Ser His
    335                340                345
ctc tga agactccagt gactgctgcc tccccccaca agaactccct tctccctca     1237
Leu *
gccaatgaat caatgtgctc cttcataagc cattgcttct ccctcacttc tttcctcaaa 1297
gaagcatgag gtgagagaaa gccacaaagt cagtgcctgg agaagggttc agcccaacat 1357
ggggccctc tcactactga aatccctcta cttctcttgg gtctggcatt ataaagaaca 1417
gctgaggctg tcattccatg agtcttcaga agaaaggaca gctcagaaaa tcaaagaggc 1477
caactgcccc gagccacaga aaatggagga taattgaggc taagtaacct gattacaagt 1537
tgtactaaca tattaaaggt tctgaaaagt cctgca 1573

```

<210> 3

<211> 1691

<212> DNA

<213> 智人 (Homo sapiens)

<220>

<221> 5' UTR

<222> 1..143

<220>

<221> CDS

<222> 144..1187

<220>

<221> 3' UTR

<222> 1188..1691

<400> 3

```

tgcactccag tccgggctgg cggacagagg gctggaacaa agacgtcca gaatcaggag      60

```

```

cttccccctca ggaaatagca tccgtgtgcc ccgcactgca gttgtctggt ctctccagca 120
gttttggtact tccggctgct gca atg cgt gtg gtg gtg att gga gca gga gtc 173
      Met Arg Val Val Val Ile Gly Ala Gly Val
      1          5          10
atc ggg ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc 221
Ile Gly Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val
      15          20          25
ctg cag cca ctg gac ata aag gtc tac gcg gac cgc ttc acc cca ctc 269
Leu Gln Pro Leu Asp Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu
      30          35          40
acc acc acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac 317
Thr Thr Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp
      45          50          55
ccc aac aac cca cag gag gcg gac tgg agc caa cag acc ttt gac tat 365
Pro Asn Asn Pro Gln Glu Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr
      60          65          70
ctc ctg agc cat gtc cat tct ccc aac gct gaa aac ctg ggc ctg ttc 413
Leu Leu Ser His Val His Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe
      75          80          85          90
cta atc tcg ggc tac aac ctc ttc cat gaa gcc att ccg gac cct tcc 461
Leu Ile Ser Gly Tyr Asn Leu Phe His Glu Ala Ile Pro Asp Pro Ser
      95          100          105
tgg aag gac aca gtt ctg gga ttt cgg aag ctg acc ccc aga gag ctg 509
Trp Lys Asp Thr Val Leu Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu
      110          115          120
gat atg ttc cca gat tac ggc tat ggc tgg ttc cac aca agc cta att 557
Asp Met Phe Pro Asp Tyr Gly Tyr Gly Trp Phe His Thr Ser Leu Ile
      125          130          135
ctg gag gga aag aac tat cta cag tgg ctg act gaa agg tta act gag 605
Leu Glu Gly Lys Asn Tyr Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu
      140          145          150
agg gga gtg aag ttc ttc cag cgg aaa gtg gag tct ttt gag gag gtg 653
Arg Gly Val Lys Phe Phe Gln Arg Lys Val Glu Ser Phe Glu Glu Val
      155          160          165          170
gca aga gaa ggc gca gac gtg att gtc aac tgc act ggg gta tgg gct 701
Ala Arg Glu Gly Ala Asp Val Ile Val Asn Cys Thr Gly Val Trp Ala
      175          180          185
ggg gcg cta caa cga gac ccc ctg ctg cag cca ggc cgg ggg cag atc 749
Gly Ala Leu Gln Arg Asp Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile
      190          195          200
atg aag gtg gac gcc cct tgg atg aag cac ttc att ctc acc cat gac 797
Met Lys Val Asp Ala Pro Trp Met Lys His Phe Ile Leu Thr His Asp
      205          210          215
cca gag aga ggc atc tac aat tcc ccg tac atc atc cca ggg acc cag 845
Pro Glu Arg Gly Ile Tyr Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln
      220          225          230
aca gtt act ctt gga ggc atc ttc cag ttg gga aac tgg agt gaa cta 893
Thr Val Thr Leu Gly Gly Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu
      235          240          245          250
aac aat atc cag gac cac aac acc att tgg gaa ggc tgc tgc aga ctg 941
Asn Asn Ile Gln Asp His Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu
      255          260          265
gag ccc aca ctg aag aat gca aga att att ggt gaa cga act ggc ttc 989
Glu Pro Thr Leu Lys Asn Ala Arg Ile Ile Gly Glu Arg Thr Gly Phe
      270          275          280

```

```

cgg cca gta cgc ccc cag att cgg cta gaa aga gaa cag ctt cgc act 1037
Arg Pro Val Arg Pro Gln Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr
      285              290              295
gga cct tca aac aca gag gtc atc cac aac tat ggc cat gga ggc tac 1085
Gly Pro Ser Asn Thr Glu Val Ile His Asn Tyr Gly His Gly Gly Tyr
      300              305              310
ggg ctc acc atc cac tgg gga tgt gcc ctg gag gca gcc aag ctc ttt 1133
Gly Leu Thr Ile His Trp Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe
      315              320              325              330
ggg aga atc ctg gaa gaa aag aaa ttg tcc aga atg cca cca tcc cac 1181
Gly Arg Ile Leu Glu Glu Lys Lys Leu Ser Arg Met Pro Pro Ser His
      335              340              345
ctc tga agactccagt gactgctgcc tccccccaca agaactccct tctccctca 1237
Leu *
gccaatgaat caatgtgctc cttcataagc cattgcttct cctcacttc tttcctcaaa 1297
gaagcatgag gtgagagaaa gccacaaagt cagtgcctgg agaagggttc agcccaacat 1357
ggggccctc tcactactga aatccctcta ctttctctgg gtctggcatt ataaagaaca 1417
gctgaggctg tcattccatg agtcttcaga agaaaggaca gctcagaaaa tcaaagaggc 1477
caactgccca gagccacaga aaatggagga taattgaggc taagtaacct gattacaagt 1537
tgtactaaca tattaaaggt tctgaaaagt cctgcagcaa agacaactat ctgatgttgt 1597
ttaacccagt gcttgctaaa cctatctggc tatggaactc tttgcccag agcacccatg 1657
aatgccatga cacaaatctg agaaaatgct ggaa 1691

<210> 4
<211> 2620
<212> DNA
<213> 智人 (Homo sapiens)

<220>
<221> 5' UTR
<222> 1..1155

<220>
<221> CDS
<222> 1156..1818

<220>
<221> 3' UTR
<222> 1819..2620

<400> 4
gaaacccacg cagcctcctg gattcttccc cgtccctccc tctgtcctgg ggctgtgacc 60
tcttccatgt tattcacagg gtctcagcac gattcatctc aaagcagcga aacaagcact 120
ggcctcagaa gaagcaatat taaaacagtt acaactcatc tagcgcacag acaccaact 180
gacaccctgt tctcctcagtc ataacaacaa ctacagcttt gattgaacaa gagactgagt 240
ttggtaactt tctcctaata aaaagatcac tgactatgga ctgcttctgg tggggttacg 300
aaaccgcaac ctcatgtgcc tgcatttcct gaaaagacat tttgatgtag gaagggcctg 360
gagtcctgct gcttcgtct ctgggatacg ggagcaaaga gccacgcac ctcatggccc 420
acacaggcgt cacctccagt ctctccttgg cctcatctcc ccagcgtcct ggaatggcat 480
cgggctggcc cagggagccc ctgtcctgtg cctctccttt cccctcaggg gctgccaggc 540
tgaccacccc caccgcaggc caggcctaca gtgccccatg gaacgtcctg accctcccc 600
agggtggcag caggaagaag gaagaaaggg gatectctcc agctggccag agagacagac 660
cttcttgtgc tcactcaacc tccaagaatg cctgccctcc ctcttcccc caaggcctgt 720
ccacaggggc ttgagatcag ccagaaaagt caggcaactt ttcagggact gggagcgagg 780
tctcccgccc gggcctgggt ccagtctctg tgggcagtgc agtgccgagc cccacccctc 840

```

aagccgtgcc ctgtccatag ctccagactt tgacctgca ctccagtcgg ggctggcgga	900
cagagggctg gaaacaagac gctccagaat caggagcttc cctcaggaa atagcatcct	960
gtgtccccgc actgcagttg tctggctctt ccagcagttt ggtacttccg atgaagagct	1020
tgtgtctcca gagggaaagt atgggggaag aggggaagaga gaagaccaag ggtccctgag	1080
aggggctgtc ccctaagccc cagtatccaa gctcgggctc gaagctggaa ggagaattgc	1140
ctagaggctg ctgca atg cgt gtg gtg gtg att gga gca gga gtc atc ggg	1191
Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly	
1 5 10	
ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc ctg cag	1239
Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln	
15 20 25	
cca ctg gac ata aag gtc tac gcg gac cgc ttc acc cca ctc acc acc	1287
Pro Leu Asp Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr	
30 35 40	
acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac ccc aac	1335
Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn	
45 50 55 60	
aac cca cag gag gcg gac tgg agc caa cag acc ttt gac tat ctc ctg	1383
Asn Pro Gln Glu Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu	
65 70 75	
agc cat gtc cat tct ccc aac gct gaa aac ctg ggc ctg ttc cta atc	1431
Ser His Val His Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile	
80 85 90	
tcg ggc tac aac ctc ttc cat gaa gcc att ccg gac cct tcc tgg aag	1479
Ser Gly Tyr Asn Leu Phe His Glu Ala Ile Pro Asp Pro Ser Trp Lys	
95 100 105	
gac aca gtt ctg gga ttt cgg aag ctg acc ccc aga gag ctg gat atg	1527
Asp Thr Val Leu Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu Asp Met	
110 115 120	
ttc cca gat tac ggc tat ggc tgg ttc cac aca agc cta att ctg gag	1575
Phe Pro Asp Tyr Gly Tyr Gly Trp Phe His Thr Ser Leu Ile Leu Glu	
125 130 135 140	
gga aag aac tat cta cag tgg ctg act gaa agg tta act gag agg gga	1623
Gly Lys Asn Tyr Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu Arg Gly	
145 150 155	
gtg aag ttc ttc cag cgg aaa gtg gag tct ttt gag gag gtg gca aga	1671
Val Lys Phe Phe Gln Arg Lys Val Glu Ser Phe Glu Glu Val Ala Arg	
160 165 170	
gaa ggc gca gac gtg att gtc aac tgc act ggg gta tgg gct ggg gcg	1719
Glu Gly Ala Asp Val Ile Val Asn Cys Thr Gly Val Trp Ala Gly Ala	
175 180 185	
cta caa cga gac ccc ctg ctg cag cca ggc cgg ggg cag atc atg aag	1767
Leu Gln Arg Asp Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile Met Lys	
190 195 200	
gac cca gac agt tac tct tgg agg cat ctt cca gtt ggg aaa ctg gag	1815
Asp Pro Asp Ser Tyr Ser Trp Arg His Leu Pro Val Gly Lys Leu Glu	
205 210 215 220	
tga actaaacaat atccaggacc acaacacccat ttgggaaggc tgctgcagac	1868
*	
tgagagccac actgaagaat gcaagaatta ttggtgaacg aactggcttc cgccagctac	1928
gccccagat tcggctagaa agagaacagc ttgcactgg accttcaaac acagaggtca	1988
tccacaacta tggccatgga ggctacgggc tcaccatcca ctggggatgt gccttggagg	2048
cagccaagct ctttgggaga atcctggaag aaaagaaatt gtccagaatg ccaccatccc	2108
acctctgaag actccagtga ctgtgcctc cccccacaag aactcccttc tccctcagc	2168
caatgaatca atgtgtcct tcataagcca ttgtttctcc ctcacttctt tctcaaaga	2228

```

agcatgaggt gagagaaagc cacaaagtca gtgcctggag aagggttcag cccaacatgg 2288
ggcccctctc atcactgaaa tccctctacc ttctctgggt ctggcattat aaagaacagc 2348
tgaggctgtc attccatgag tcttcagaag aaaggacagc tcagaaaatc aaagaggcca 2408
actgcccaga gccacagaaa atggaggata attgaggcta agtaacctga ttacaagttg 2468
tactaacata ttaaaggttc tgaaaagtc tgcagcaaag acaactatct gatgttggtt 2528
aaccagtgct ttgctaaacc tatctggcta tggaactctt ttgccagag caccatgaa 2588
tgccatgaca caaatctgag aaaatgctgg aa 2620

<210> 5
<211> 1576
<212> DNA
<213> 智人 (Homo sapiens)

<220>
<221> 5' UTR
<222> 1..143

<220>
<221> CDS
<222> 144..380

<220>
<221> 3' UTR
<222> 381..1576

<400> 5
tgactccag tccgggctgg cggacagagg gctggaaaca agacgtcca gaatcaggag 60
cttcccctca ggaatatgca tctgtgtcc cgcactgca gtgtctgtg ctctccagca 120
gtttggtact tccggtgct gca atg cgt gtg gtg gtg att gga gca gga gtc 173
Met Arg Val Val Val Ile Gly Ala Gly Val
1 5 10
atc ggg ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc 221
Ile Gly Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val
15 20 25
ctg cag cca ctg gac ata aag gtc tac gcg gac cgc ttc acc cca ctc 269
Leu Gln Pro Leu Asp Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu
30 35 40
acc acc acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac 317
Thr Thr Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp
45 50 55
ccc aac aac cca cag gag gcg acc ctt cct gga agg aca cag ttc tgg 365
Pro Asn Asn Pro Gln Glu Ala Thr Leu Pro Gly Arg Thr Gln Phe Trp
60 65 70
gat ttc gga agc tga cccacagaga gctggatatg ttccagatt acggctatgg 420
Asp Phe Gly Ser *
75
ctggttccac acaagcctaa ttctggaggg aaagaactat ctacagtggc tgactgaaag 480
gttaactgag aggggagtga agttcttcca gcggaaagtg gactcttttg aggaggtggc 540
aagagaaggc gcagacgtga ttgtcaactg cactggggta tgggctgggg cgctacaacg 600
agaccccctg ctgcagccag gccgggggca gatcatgaag gtggacgccc cttggatgaa 660
gcatttcatt ctacccatg acccagagag aggcattctac aattccccgt acatcatccc 720
agggaaccag acagttactc ttggaggcat ctccagttg ggaactgga gtgaactaaa 780
caataccag gaccacaaca ccatttggga aggtctgtgc agactggagc ccacactgaa 840
gaatgcaaga attattgtg aacgaactgg ctccggcca gtacgcccc agattcggct 900
agaaagagaa cagcttcgca ctggacctc aaacacagag gtcattccaca actatggcca 960

```

```

tggaggctac gggctcacca tccactgggg atgtgccctg gaggcagcca agctctttgg 1020
gagaatcctg gaagaaaaga aattgtccag aatgccacca tcccacctct gaagactcca 1080
gtgactgctg cctcccccca caagaactcc cttctccct cagccaatga atcaatgtgc 1140
tccttcataa gccattgctt ctccctcact tctttcctca aagaagcatg aggtgagaga 1200
aagccacaaa gtcagtgcct ggagaagggt tcagcccaac atggggcccc tctcatcact 1260
gaaatccctc taccttctct gggtctggca ttataaagaa cagctgaggc tgtcattcca 1320
tgagtcttca gaagaaagga cagctcagaa aatcaaagag gccaactgcc cagagccaca 1380
gaaaatggag gataattgag gctaagtaac ctgattacaa gttgtactaa catattaaag 1440
gttctgaaaa gtcctgcagc aaagacaact atctgatgtt gtttaacca gtgcttgcta 1500
aacctatctg gctatggaac tcttttgccc agagcaccca tgaatgcat gacacaaatc 1560
tgagaaaatg ctggaa 1576

```

<210> 6
 <211> 1345
 <212> DNA
 <213> 智人 (Homo sapiens)

<220>
 <221> 5' UTR
 <222> 1..113

<220>
 <221> CDS
 <222> 114..959

<220>
 <221> 3' UTR
 <222> 960..1345

<220>
 <221> polyA_signal
 <222> 1321..1326

```

<400> 6
gaaacccacg cagcctcctg gattcttccc cgtccctccc tctgtcctgg ggctgtgacc 60
tcttccatgt tattcacagg gtctcagcac gattcatctc aaaggctgct gca atg 116
                                     Met
                                     1
cgt gtg gtg gtg att gga gca gga gtc atc ggg ctg tcc acc gcc ctc 164
Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala Leu
      5              10              15
tgc atc cat gag cgc tac cac tca gtc ctg cag cca ctg gac ata aag 212
Cys Ile His Glu Arg Tyr His Ser Val Leu Gln Pro Leu Asp Ile Lys
      20              25              30
gtc tac gcg gac cgc ttc acc cca ctc acc acc acc gac gtg gct gcc 260
Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr Thr Asp Val Ala Ala
      35              40              45
ggc ctc tgg cag ccc tac ctt tct gac ccc aac aac cca cag gag gcg 308
Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn Asn Pro Gln Glu Ala
      50              55              60              65
gac tgg agc caa cag acc ttt gac tat ctc ctg agc cat gtc cat tct 356
Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu Ser His Val His Ser
      70              75              80
ccc aac gct gaa aac ctg ggc ctg ttc cta atc tcg ggc tac aac ctc 404
Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile Ser Gly Tyr Asn Leu

```

```

      85              90              95
ttc cat gaa gcc att ccg gtg gca aga gaa ggc gca gac gtg att gtc 452
Phe His Glu Ala Ile Pro Val Ala Arg Glu Gly Ala Asp Val Ile Val
      100              105              110
aac tgc act ggg gta tgg gct ggg gcg cta caa cga gac ccc ctg ctg 500
Asn Cys Thr Gly Val Trp Ala Gly Ala Leu Gln Arg Asp Pro Leu Leu
      115              120              125
cag cca ggc cgg ggg cag atc atg aag gtg gac gcc cct tgg atg aag 548
Gln Pro Gly Arg Gly Gln Ile Met Lys Val Asp Ala Pro Trp Met Lys
      130              135              140              145
cac ttc att ctc acc cat gac cca gag aga ggc atc tac aat tcc ccg 596
His Phe Ile Leu Thr His Asp Pro Glu Arg Gly Ile Tyr Asn Ser Pro
      150              155              160
tac atc atc cca ggg acc cag aca gtt act ctt gga ggc atc ttc cag 644
Tyr Ile Ile Pro Gly Thr Gln Thr Val Thr Leu Gly Gly Ile Phe Gln
      165              170              175
ttg gga aac tgg agt gaa cta aac aat atc cag gac cac aac acc att 692
Leu Gly Asn Trp Ser Glu Leu Asn Asn Ile Gln Asp His Asn Thr Ile
      180              185              190
tgg gaa ggc tgc tgc aga ctg gag ccc aca ctg aag aat gca aga att 740
Trp Glu Gly Cys Cys Arg Leu Glu Pro Thr Leu Lys Asn Ala Arg Ile
      195              200              205
att ggt gaa cga act ggc ttc cgg cca gta cgc ccc cag att cgg cta 788
Ile Gly Glu Arg Thr Gly Phe Arg Pro Val Arg Pro Gln Ile Arg Leu
      210              215              220              225
gaa aga gaa cag ctt cgc act gga cct tca aac aca gag gtc atc cac 836
Glu Arg Glu Gln Leu Arg Thr Gly Pro Ser Asn Thr Glu Val Ile His
      230              235              240
aac tat ggc cat gga ggc tac ggg ctc acc atc cac tgg gga tgt gcc 884
Asn Tyr Gly His Gly Gly Tyr Gly Leu Thr Ile His Trp Gly Cys Ala
      245              250              255
ctg gag gca gcc aag ctc ttt ggg aga atc ctg gaa gaa aag aaa ttg 932
Leu Glu Ala Ala Lys Leu Phe Gly Arg Ile Leu Glu Glu Lys Lys Leu
      260              265              270
tcc aga atg cca cca tcc cac ctc tga agactccagt gactgctgcc 979
Ser Arg Met Pro Pro Ser His Leu *
      275              280
tccccccaca agaactccct tctcccctca gccaatgaat caatgtgctc cttcataagc 1039
cattgcttct ccttcacttc ttctctcaaa gaagcatgag gtgagagaaa gccacaaaagt 1099
cagtgccctgg agaagggttc agcccaacat ggggcccttc tcatcactga aatccctcta 1159
ccttctctgg gtctggcatt ataaagaaca gctgaggctg tcattccatg agtcttcaga 1219
agaaaggaca gctcagaaaa tcaaagaggc caactgccca gagccacaga aaatggagga 1279
taattgaggc taagtaacct gattacaagt tgtactaaca tattaaaggt tctgaaaagt 1339
cctgca 1345

```

<210> 7

<211> 347

<212> PRT

<213> 智人 (Homo sapiens)

<400> 7

```

Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala
1              5              10              15
Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln Pro Leu Asp Ile
      20              25              30

```

```

Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr Thr Asp Val Ala
  35          40          45
Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn Asn Pro Gln Glu
  50          55          60
Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu Ser His Val His
  65          70          75          80
Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile Ser Gly Tyr Asn
      85          90          95
Leu Phe His Glu Ala Ile Pro Asp Pro Ser Trp Lys Asp Thr Val Leu
      100          105          110
Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu Asp Met Phe Pro Asp Tyr
      115          120          125
Gly Tyr Gly Trp Phe His Thr Ser Leu Ile Leu Glu Gly Lys Asn Tyr
      130          135          140
Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu Arg Gly Val Lys Phe Phe
      145          150          155          160
Gln Arg Lys Val Glu Ser Phe Glu Glu Val Ala Arg Glu Gly Ala Asp
      165          170          175
Val Ile Val Asn Cys Thr Gly Val Trp Ala Gly Ala Leu Gln Arg Asp
      180          185          190
Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile Met Lys Val Asp Ala Pro
      195          200          205
Trp Met Lys His Phe Ile Leu Thr His Asp Pro Glu Arg Gly Ile Tyr
      210          215          220
Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln Thr Val Thr Leu Gly Gly
      225          230          235          240
Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu Asn Asn Ile Gln Asp His
      245          250          255
Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu Glu Pro Thr Leu Lys Asn
      260          265          270
Ala Arg Ile Ile Gly Glu Arg Thr Gly Phe Arg Pro Val Arg Pro Gln
      275          280          285
Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr Gly Pro Ser Asn Thr Glu
      290          295          300
Val Ile His Asn Tyr Gly His Gly Gly Tyr Gly Leu Thr Ile His Trp
      305          310          315          320
Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe Gly Arg Ile Leu Glu Glu
      325          330          335
Lys Lys Leu Ser Arg Met Pro Pro Ser His Leu
      340          345

```

<210> 8

<211> 220

<212> PRT

<213> 智人 (Homo sapiens)

<400> 8

```

Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala
  1          5          10          15
Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln Pro Leu Asp Ile
      20          25          30
Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr Thr Asp Val Ala
      35          40          45
Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn Asn Pro Gln Glu
      50          55          60

```

Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu Ser His Val His
 65 70 75 80
 Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile Ser Gly Tyr Asn
 85 90 95
 Leu Phe His Glu Ala Ile Pro Asp Pro Ser Trp Lys Asp Thr Val Leu
 100 105 110
 Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu Asp Met Phe Pro Asp Tyr
 115 120 125
 Gly Tyr Gly Trp Phe His Thr Ser Leu Ile Leu Glu Gly Lys Asn Tyr
 130 135 140
 Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu Arg Gly Val Lys Phe Phe
 145 150 155 160
 Gln Arg Lys Val Glu Ser Phe Glu Glu Val Ala Arg Glu Gly Ala Asp
 165 170 175
 Val Ile Val Asn Cys Thr Gly Val Trp Ala Gly Ala Leu Gln Arg Asp
 180 185 190
 Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile Met Lys Asp Pro Asp Ser
 195 200 205
 Tyr Ser Trp Arg His Leu Pro Val Gly Lys Leu Glu
 210 215 220

<210> 9
 <211> 78
 <212> PRT
 <213> 智人 (Homo sapiens)

<400> 9
 Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala
 1 5 10 15
 Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln Pro Leu Asp Ile
 20 25 30
 Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr Thr Asp Val Ala
 35 40 45
 Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn Asn Pro Gln Glu
 50 55 60
 Ala Thr Leu Pro Gly Arg Thr Gln Phe Trp Asp Phe Gly Ser
 65 70 75

<210> 10
 <211> 281
 <212> PRT
 <213> 智人 (Homo sapiens)

<400> 10
 Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala
 1 5 10 15
 Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln Pro Leu Asp Ile
 20 25 30
 Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr Thr Asp Val Ala
 35 40 45
 Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn Asn Pro Gln Glu
 50 55 60
 Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu Ser His Val His
 65 70 75 80
 Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile Ser Gly Tyr Asn

```

      85      90      95
Leu Phe His Glu Ala Ile Pro Val Ala Arg Glu Gly Ala Asp Val Ile
      100      105      110
Val Asn Cys Thr Gly Val Trp Ala Gly Ala Leu Gln Arg Asp Pro Leu
      115      120      125
Leu Gln Pro Gly Arg Gly Gln Ile Met Lys Val Asp Ala Pro Trp Met
      130      135      140
Lys His Phe Ile Leu Thr His Asp Pro Glu Arg Gly Ile Tyr Asn Ser
      145      150      155      160
Pro Tyr Ile Ile Pro Gly Thr Gln Thr Val Thr Leu Gly Gly Ile Phe
      165      170      175
Gln Leu Gly Asn Trp Ser Glu Leu Asn Asn Ile Gln Asp His Asn Thr
      180      185      190
Ile Trp Glu Gly Cys Cys Arg Leu Glu Pro Thr Leu Lys Asn Ala Arg
      195      200      205
Ile Ile Gly Glu Arg Thr Gly Phe Arg Pro Val Arg Pro Gln Ile Arg
      210      215      220
Leu Glu Arg Glu Gln Leu Arg Thr Gly Pro Ser Asn Thr Glu Val Ile
      225      230      235      240
His Asn Tyr Gly His Gly Gly Tyr Gly Leu Thr Ile His Trp Gly Cys
      245      250      255
Ala Leu Glu Ala Ala Lys Leu Phe Gly Arg Ile Leu Glu Glu Lys Lys
      260      265      270
Leu Ser Arg Met Pro Pro Ser His Leu
      275      280

```

<210> 11
 <211> 456
 <212> DNA
 <213> 智人 (Homo sapiens)

<220>
 <221> allele
 <222> 152
 <223> 99-16105-152:多态性碱基A或G

<220>
 <221> misc_binding
 <222> 140..164
 <223> 99-16105-152. 探针

<220>
 <221> primer_bind
 <222> 133..151
 <223> 99-16105-152.mis

<220>
 <221> primer_bind
 <222> 153..171
 <223> 99-16105-152.mis补体

<220>
 <221> primer_bind
 <222> 1..19
 <223> 99-16105.pu

<220>
 <221> primer_bind
 <222> 437..456
 <223> 99-16105.rp补体

<400> 11
 cgctttgttg tattctttgt tatttatcca ttttgccaaa ttatctgcaa gtagaaatat 60
 cgaaataaga agctcttttag caatttactt tggatattgg ttttcttttg aaggacagtt 120
 attaaaatag ctgttaggat tactcatttt crtittttctt ctttttaaat ataaagcaat 180
 gtcatactt ttttcctgt attatatttc tcctcaataa ttgatatgct acattaaagg 240
 aacacaaaat ggtcttaatt atgcaataat gatcaaggca aagagtgttt cctgggaact 300
 aatggttgcc tgagaggagg tgatggcttg aggtccagct gggtattaag ccgcaggaaa 360
 tgctgcaggc caagatttgt attatttctc tgagatgaaa atgaacccea aaaaaggcaa 420
 aatgggtttt tctccactaa tgggtaaaat gaactc 456

<210> 12
 <211> 463
 <212> DNA
 <213> 智人 (Homo sapiens)

<220>
 <221> allele
 <222> 215
 <223> 99-5919-215:多态性碱基A或G

<220>
 <221> misc_binding
 <222> 203..227
 <223> 99-5919-215.探针

<220>
 <221> primer_bind
 <222> 196..214
 <223> 99-5919-215.mis

<220>
 <221> primer_bind
 <222> 216..234
 <223> 99-5919-215.mis补体

<220>
 <221> primer_bind
 <222> 11..29
 <223> 99-5919.pu

<220>
 <221> primer_bind
 <222> 445..465
 <223> 99-5919.rp补体

<400> 12
 ttctcttga ctacagcaat gcagatttca attctgccaat tgaattccca gacatattcg 60
 tcatcccat tttcatcccc caccacctg ccattttctt cgtgttaact tgttttctg 120
 actcacagaa atcacctttt cctgtatata tttttaggat gtcagacttt atttcaatga 180

```

tttctcctag ttgcccccca aaattgtatt ctacrgtgtg attttaaagc tgaattttca 240
agatgatatt tcatatctat attttcacaa gcttttcttc tatgaatgtt attgtcagct 300
gtcagggtgt gagatggtac ttgatactac attctttcca agctgttgcc tgaatcggtt 360
taagacaaag tcattactag gctgtaaact gttgctctgc aaaattgagc agcacgtatt 420
taaccactca tacttcttag ctctccaaca ctttgagtcr ata 463

```

<210> 13
 <211> 742
 <212> DNA
 <213> 智人 (Homo sapiens)

<220>
 <221> 5' UTR
 <222> 1..46

<220>
 <221> CDS
 <222> 47..508

<220>
 <221> 3' UTR
 <222> 509..742

<220>
 <221> polyA_signal
 <222> 718..723

<220>
 <221> allele
 <222> 21
 <223> 8-135-112:多态性碱基C或T

<220>
 <221> allele
 <222> 75
 <223> 8-135-166:多态性碱基A或C

<220>
 <221> allele
 <222> 135
 <223> 99-16038-118:多态性碱基A或G

<220>
 <221> allele
 <222> 194
 <223> 8-137-152:多态性碱基G或T

<220>
 <221> allele
 <222> 224
 <223> 8-137-182:多态性碱基A或G

<220>
 <221> allele
 <222> 390

<223> 8-130-220:多态性碱基A或C

<220>

<221> allele

<222> 406

<223> 8-130-236:多态性碱基A或G

<220>

<221> allele

<222> 578

<223> 8-131-199:多态性碱基A或C

<220>

<221> allele

<222> 641

<223> 8-132-97:多态性碱基C或T

<220>

<221> allele

<222> 708

<223> 8-132-164:多态性碱基C或T

<220>

<221> allele

<222> 723

<223> 8-132-179:多态性碱基A或T

<400> 13

```

tcattctctgc ttcacaatgc ygatgattta gctgggagga cccaaa atg ctg gaa      55
                                     Met Leu Glu
                                     1
aag ctg atg ggt gct gat tmt ctc cag ctt ttc aga tcc aga tat aca      103
Lys Leu Met Gly Ala Asp Xaa Leu Gln Leu Phe Arg Ser Arg Tyr Thr
   5              10              15
ttg ggt aaa atc tac ttc ata ggt ttt caa arg agc att ctt ctg agc      151
Leu Gly Lys Ile Tyr Phe Ile Gly Phe Gln Xaa Ser Ile Leu Leu Ser
  20              25              30              35
aaa tct gaa aac tct cta aac tct att gca aag gag aca gaa kaa gga      199
Lys Ser Glu Asn Ser Leu Asn Ser Ile Ala Lys Glu Thr Glu Xaa Gly
              40              45              50
aga gag acg gta aca agg aaa gaa rga tgg aag aga agg cat gag gac      247
Arg Glu Thr Val Thr Arg Lys Glu Xaa Trp Lys Arg Arg His Glu Asp
              55              60              65
ggc tat ttg gaa atg gca cag agg cat tta cag aga tca tta tgt cct      295
Gly Tyr Leu Glu Met Ala Gln Arg His Leu Gln Arg Ser Leu Cys Pro
              70              75              80
tgg gtc tct tac ctt cct cag ccc tat gca gag ctt gaa gaa gta agc      343
Trp Val Ser Tyr Leu Pro Gln Pro Tyr Ala Glu Leu Glu Glu Val Ser
              85              90              95
agc cat gtt gga aaa gtc ttc atg gca aga aac tat gag ttc ctt gmc      391
Ser His Val Gly Lys Val Phe Met Ala Arg Asn Tyr Glu Phe Leu Xaa
  100              105              110              115
tat gag gcc tct aar gac cgc agg cag cct cta gaa cga atg tgg acc      439
Tyr Glu Ala Ser Lys Asp Arg Arg Gln Pro Leu Glu Arg Met Trp Thr
              120              125              130

```

tgc aac tac aac cag caa aaa gac cag tca tgc aac cac aag gaa ata 487
 Cys Asn Tyr Asn Gln Gln Lys Asp Gln Ser Cys Asn His Lys Glu Ile
 135 140 145
 act tct acc aaa gct gaa tga gtttgaagc agattcttcc cagccaatcc 538
 Thr Ser Thr Lys Ala Glu *
 150
 ttctgatgac aatgtagtct ggccaacatc ttacttggam tctgacggac tctgtgtctg 598
 ggaccacagct gataacacgt ggtgatggga ttgtatttgc aaytctctgg tcagtaagtg 658
 ataaaatgcc atttctatgc acccacctgg cctgtgtgac tgggagaaty tctcttttta 718
 ttaawtgtgc ttcaagtttt aaca 742

 <210> 14
 <211> 153
 <212> PRT
 <213> 智人 (Homo sapiens)

 <400> 14
 Met Leu Glu Lys Leu Met Gly Ala Asp Xaa Leu Gln Leu Phe Arg Ser
 1 5 10 15
 Arg Tyr Thr Leu Gly Lys Ile Tyr Phe Ile Gly Phe Gln Xaa Ser Ile
 20 25 30
 Leu Leu Ser Lys Ser Glu Asn Ser Leu Asn Ser Ile Ala Lys Glu Thr
 35 40 45
 Glu Xaa Gly Arg Glu Thr Val Thr Arg Lys Glu Xaa Trp Lys Arg Arg
 50 55 60
 His Glu Asp Gly Tyr Leu Glu Met Ala Gln Arg His Leu Gln Arg Ser
 65 70 75 80
 Leu Cys Pro Trp Val Ser Tyr Leu Pro Gln Pro Tyr Ala Glu Leu Glu
 85 90 95
 Glu Val Ser Ser His Val Gly Lys Val Phe Met Ala Arg Asn Tyr Glu
 100 105 110
 Phe Leu Xaa Tyr Glu Ala Ser Lys Asp Arg Arg Gln Pro Leu Glu Arg
 115 120 125
 Met Trp Thr Cys Asn Tyr Asn Gln Gln Lys Asp Gln Ser Cys Asn His
 130 135 140
 Lys Glu Ile Thr Ser Thr Lys Ala Glu
 145 150

 <210> 15
 <211> 476
 <212> DNA
 <213> 智人 (Homo sapiens)

 <400> 15
 cat gag gac ggc tat ttg gaa atg gca cag agg cat tta cag aga tca 48
 His Glu Asp Gly Tyr Leu Glu Met Ala Gln Arg His Leu Gln Arg Ser
 1 5 10 15
 tta tgt cct tgg gtc tct tac ctt cct cag ccc tat gca gag ctt gaa 96
 Leu Cys Pro Trp Val Ser Tyr Leu Pro Gln Pro Tyr Ala Glu Leu Glu
 20 25 30
 gaa gta agc agc cat gtt gga aaa gtc ttc atg gca aga aac tat gag 144
 Glu Val Ser Ser His Val Gly Lys Val Phe Met Ala Arg Asn Tyr Glu
 35 40 45
 ttc ctt gcc tat gag gcc tct aag gac cgc agg cag cct cta gaa cga 192
 Phe Leu Ala Tyr Glu Ala Ser Lys Asp Arg Arg Gln Pro Leu Glu Arg

```

      50              55              60
atg tgg acc tgc aac tac aac cag caa aaa gac cag tca tgc aac cac    240
Met Trp Thr Cys Asn Tyr Asn Gln Gln Lys Asp Gln Ser Cys Asn His
65              70              75              80
aag gaa ata act tct acc aaa gct gaa tga gtttggaagc agattcttcc    290
Lys Glu Ile Thr Ser Thr Lys Ala Glu *
      85              90
cagccaatcc ttctgatgac aatgtagtct ggccaacatc ttacttgac tctgacggac    350
tctgtgtctg ggaccagct gataacacgt ggtgatggga ttgtatttgc aactctctgg    410
tcagtaagtg ataaaatgcc atttctatgc acccacctgg cctgtgtgac tgggagaatc    470
tctctt                                         476

<210> 16
<211> 89
<212> PRT
<213> 智人 (Homo sapiens)

<400> 16
His Glu Asp Gly Tyr Leu Glu Met Ala Gln Arg His Leu Gln Arg Ser
1              5              10              15
Leu Cys Pro Trp Val Ser Tyr Leu Pro Gln Pro Tyr Ala Glu Leu Glu
      20              25              30
Glu Val Ser Ser His Val Gly Lys Val Phe Met Ala Arg Asn Tyr Glu
      35              40              45
Phe Leu Ala Tyr Glu Ala Ser Lys Asp Arg Arg Gln Pro Leu Glu Arg
      50              55              60
Met Trp Thr Cys Asn Tyr Asn Gln Gln Lys Asp Gln Ser Cys Asn His
65              70              75              80
Lys Glu Ile Thr Ser Thr Lys Ala Glu
      85

<210> 17
<211> 1633
<212> DNA
<213> 智人 (Homo sapiens)

<400> 17
ttgggggtcca ttgcaacccg aggcgagact agagttccca agcgagaagg gaagaggcag    60
tgggtgcacg tggaaggcgg acagagggct ggaacaaga cgctccagaa tcaggagctt    120
ccccctagga aatagcatcc tgtgtcccg cactgcagtt gtctggtctc tccagcagtt    180
tggtacttcc ggctgctgca atg cgt gtg gtg gtg att gga gca gga gtc atc    233
      Met Arg Val Val Val Ile Gly Ala Gly Val Ile
              1              5              10
ggg ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc ctg    281
Gly Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val Leu
              15              20              25
cag cca ctg cac ata aag gtc tac gcg gac cgc ttc acc cca ctc acc    329
Gln Pro Leu His Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr
              30              35              40
acc acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac ccc    377
Thr Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro
      45              50              55
aac aac cca cag gag gcg gac tgg agc caa cag acc ttt gac tat ctc    425
Asn Asn Pro Gln Glu Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu
60              65              70              75

```

ctg agc cat gtc cat tct ccc aac gct gaa aac ctg ggc ctg ttc cta	473
Leu Ser His Val His Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu	
80 85 90	
atc tcg ggc tac aac ctc ttc cat gaa gcc att ccg gac cct tcc tgg	521
Ile Ser Gly Tyr Asn Leu Phe His Glu Ala Ile Pro Asp Pro Ser Trp	
95 100 105	
aag gac aca gtt ctg gga ttt cgg aag ctg acc ccc aga gag ctg gat	569
Lys Asp Thr Val Leu Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu Asp	
110 115 120	
atg ttc cca gat tac ggc tat ggc tgg ttc cac aca agc cta att ctg	617
Met Phe Pro Asp Tyr Gly Tyr Gly Trp Phe His Thr Ser Leu Ile Leu	
125 130 135	
gag gga aag aac tat cta cag tgg ctg act gaa agg tta act gag agg	665
Glu Gly Lys Asn Tyr Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu Arg	
140 145 150 155	
gga gtg aag ttc ttc cag cgg aaa gtg gag tct ttt gag gag gtg gca	713
Gly Val Lys Phe Phe Gln Arg Lys Val Glu Ser Phe Glu Glu Val Ala	
160 165 170	
aga gaa ggc gca gac gtg att gtc aac tgc act ggg gta tgg gct ggg	761
Arg Glu Gly Ala Asp Val Ile Val Asn Cys Thr Gly Val Trp Ala Gly	
175 180 185	
gcg cta caa cga gac ccc ctg ctg cag cca ggc cgg ggg cag atc atg	809
Ala Leu Gln Arg Asp Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile Met	
190 195 200	
aag gtg gac gcc cct tgg atg aag cac ttc att ctc acc cat gac cca	857
Lys Val Asp Ala Pro Trp Met Lys His Phe Ile Leu Thr His Asp Pro	
205 210 215	
gag aga ggc atc tac aat tcc ccg tac atc atc cca ggg acc cag aca	905
Glu Arg Gly Ile Tyr Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln Thr	
220 225 230 235	
gtt act ctt gga ggc atc ttc cag ttg gga aac tgg agt gaa cta aac	953
Val Thr Leu Gly Gly Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu Asn	
240 245 250	
aat atc cag gac cac aac acc att tgg gaa ggc tgc tgc aga ctg gag	1001
Asn Ile Gln Asp His Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu Glu	
255 260 265	
ccc aca ctg aag aat gca aga att att ggt gaa gca act ggc ttc cgg	1049
Pro Thr Leu Lys Asn Ala Arg Ile Ile Gly Glu Ala Thr Gly Phe Arg	
270 275 280	
cca gta cgc ccc cag att cgg cta gaa aga gaa cag ctt cgc act gga	1097
Pro Val Arg Pro Gln Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr Gly	
285 290 295	
cct tca aac aca gag gtc atc cac aac tat ggc cat gga ggc tac ggg	1145
Pro Ser Asn Thr Glu Val Ile His Asn Tyr Gly His Gly Gly Tyr Gly	
300 305 310 315	
ctc acc atc cac tgg gga tgt gcc ctg gag gca gcc aag ctc ttt ggg	1193
Leu Thr Ile His Trp Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe Gly	
320 325 330	
aga atc ctg gaa gaa aag aaa ttg tcc aga atg cca cca tcc cac ctc	1241
Arg Ile Leu Glu Glu Lys Lys Leu Ser Arg Met Pro Pro Ser His Leu	
335 340 345	
tga agactccagt gactgtgcc tccccccaca agaactccct tctccccca	1294
*	
gccaatgaat caatgtgtc cttcataagc cattgtctct ccctcacttc tttctcaaa	1354
gaagcatgag gtgagagaaa gccacraagt cagtgcctgg agaagggttc agcccaacat	1414

```

ggggccctc tcactactga aatccctcta cttctctgg gtctggcatt ataaagaaca 1474
gctgaggctg tcattccatg agtcttcaga agaaaggaca gtcagaaag tcaaagaggc 1534
caactgccca gagccacaga aaatggagga taattgaggc taagtaacct gattacaagt 1594
tgtactaaca tattaaaggt tctgaaaagt cctgcaaaa 1633

```

<210> 18

<211> 347

<212> PRT

<213> 智人 (Homo sapiens)

<400> 18

```

Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala
1      5      10     15
Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln Pro Leu Asp Ile
20     25     30
Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr Thr Asp Val Ala
35     40     45
Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn Asn Pro Gln Glu
50     55     60
Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu Ser His Val His
65     70     75     80
Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile Ser Gly Tyr Asn
85     90     95
Leu Phe His Glu Ala Ile Pro Asp Pro Ser Trp Lys Asp Thr Val Leu
100    105    110
Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu Asp Met Phe Pro Asp Tyr
115    120    125
Gly Tyr Gly Trp Phe His Thr Ser Leu Ile Leu Glu Gly Lys Asn Tyr
130    135    140
Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu Arg Gly Val Lys Phe Phe
145    150    155    160
Gln Arg Lys Val Glu Ser Phe Glu Glu Val Ala Arg Glu Gly Ala Asp
165    170    175
Val Ile Val Asn Cys Thr Gly Val Trp Ala Gly Ala Leu Gln Arg Asp
180    185    190
Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile Met Lys Val Asp Ala Pro
195    200    205
Trp Met Lys His Phe Ile Leu Thr His Asp Pro Glu Arg Gly Ile Tyr
210    215    220
Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln Thr Val Thr Leu Gly Gly
225    230    235    240
Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu Asn Asn Ile Gln Asp His
245    250    255
Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu Glu Pro Thr Leu Lys Asn
260    265    270
Ala Arg Ile Ile Gly Glu Ala Thr Gly Phe Arg Pro Val Arg Pro Gln
275    280    285
Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr Gly Pro Ser Asn Thr Glu
290    295    300
Val Ile His Asn Tyr Gly His Gly Gly Tyr Gly Leu Thr Ile His Trp
305    310    315    320
Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe Gly Arg Ile Leu Glu Glu
325    330    335
Lys Lys Leu Ser Arg Met Pro Pro Ser His Leu
340    345

```

<210> 19
 <211> 1200
 <212> DNA
 <213> 智人 (Homo sapiens)

<400> 19

atg gac aca gca cgg att gca gtt gtc ggg gca ggt gtg gtg ggg ctc	48
Met Asp Thr Ala Arg Ile Ala Val Val Gly Ala Gly Val Val Gly Leu	
1 5 10 15	
tcc acg gct gtg tgc atc tcc aaa ctg gtg ccc cga tgc tcc gtt acc	96
Ser Thr Ala Val Cys Ile Ser Lys Leu Val Pro Arg Cys Ser Val Thr	
20 25 30	
atc att tca gac aag ttt act cca gat acc acc agt gat gtg gca gcc	144
Ile Ile Ser Asp Lys Phe Thr Pro Asp Thr Thr Ser Asp Val Ala Ala	
35 40 45	
gga atg ctt att cct cac act tat cca gat aca ccc att cac acg cag	192
Gly Met Leu Ile Pro His Thr Tyr Pro Asp Thr Pro Ile His Thr Gln	
50 55 60	
aag cag tgg ttc aga gaa acc ttt aat cac ctc ttt gca att gcc aat	240
Lys Gln Trp Phe Arg Glu Thr Phe Asn His Leu Phe Ala Ile Ala Asn	
65 70 75 80	
tct gca gaa gct gga gat gct ggt gtt cat ttg gta tca ggt tgg cag	288
Ser Ala Glu Ala Gly Asp Ala Gly Val His Leu Val Ser Gly Trp Gln	
85 90 95	
ata ttt cag agc act ccg act gaa gaa gtg cca ttc tgg gct gac gtg	336
Ile Phe Gln Ser Thr Pro Thr Glu Glu Val Pro Phe Trp Ala Asp Val	
100 105 110	
gtt ctg gga ttt cga aag atg act gag gct gag ctg aag aaa ttc ccc	384
Val Leu Gly Phe Arg Lys Met Thr Glu Ala Glu Leu Lys Lys Phe Pro	
115 120 125	
cag tat gtg ttt ggt cag gct ttt aca acc ctg aaa tgt gaa tgc cct	432
Gln Tyr Val Phe Gly Gln Ala Phe Thr Thr Leu Lys Cys Glu Cys Pro	
130 135 140	
gcc tac ctc ccg tgg ttg gag aaa agg ata aag gga agt gga ggc tgg	480
Ala Tyr Leu Pro Trp Leu Glu Lys Arg Ile Lys Gly Ser Gly Gly Trp	
145 150 155 160	
aca ctc act ccg cga ata gaa gac ctg tgg gaa ctt cat ccg tcc ttt	528
Thr Leu Thr Arg Arg Ile Glu Asp Leu Trp Glu Leu His Pro Ser Phe	
165 170 175	
gac atc gtg gtc aac tgt tca ggc ctt gga agc aga cag ctt gca gga	576
Asp Ile Val Val Asn Cys Ser Gly Leu Gly Ser Arg Gln Leu Ala Gly	
180 185 190	
gac tca aag att ttc cct gta agg ggc caa gtc ctc caa gtt cag gct	624
Asp Ser Lys Ile Phe Pro Val Arg Gly Gln Val Leu Gln Val Gln Ala	
195 200 205	
ccc tgg gtg gag cat ttt atc cga gat ggc agt ggg ctg aca tat att	672
Pro Trp Val Glu His Phe Ile Arg Asp Gly Ser Gly Leu Thr Tyr Ile	
210 215 220	
tat cct ggt aca tcc cat gta acc cta ggt gga act agg caa aaa ggg	720
Tyr Pro Gly Thr Ser His Val Thr Leu Gly Gly Thr Arg Gln Lys Gly	
225 230 235 240	
gac tgg aat ctg tcc ccg gat gca gaa aat agc aga gag att ctt tcc	768
Asp Trp Asn Leu Ser Pro Asp Ala Glu Asn Ser Arg Glu Ile Leu Ser	
245 250 255	

cga tgc tgt gct ctg gag ccc tcc ctc cac gga gcc tgc aac atc agg	816
Arg Cys Cys Ala Leu Glu Pro Ser Leu His Gly Ala Cys Asn Ile Arg	
260 265 270	
gag aag gtg ggc ttg agg ccc tac agg cca ggc gtg cga ctg cag aca	864
Glu Lys Val Gly Leu Arg Pro Tyr Arg Pro Gly Val Arg Leu Gln Thr	
275 280 285	
gag ctc ctt gcg cga gat gga cag agg ctg cct gta gtc cac cac tat	912
Glu Leu Leu Ala Arg Asp Gly Gln Arg Leu Pro Val Val His His Tyr	
290 295 300	
ggc cat ggg agt ggg ggc atc tca gtg cac tgg ggc act gct ctg gag	960
Gly His Gly Ser Gly Gly Ile Ser Val His Trp Gly Thr Ala Leu Glu	
305 310 315 320	
gcc gcc agg ctg gtg agc gag tgt gtc cat gcc ctc agg acc ccc att	1008
Ala Ala Arg Leu Val Ser Glu Cys Val His Ala Leu Arg Thr Pro Ile	
325 330 335	
ccc aag tca aac ctg tag atgacataaa atgacagcaa agagactgag	1056
Pro Lys Ser Asn Leu *	
340	
agactgttga tcaaagcaca gaacagggttc aaataacttt tccactgcat gaaagtttaa	1116
ttagacatttt cttgtttttc aacattagaa gtggtgtaac atgtaagctg agcacggtag	1176
catgcctata gtcccagcta ctg	1200
<210> 20	
<211> 2056	
<212> DNA	
<213> 智人 (Homo sapiens)	
<220>	
<221> misc_feature	
<222> 1025	
<223> n=a, g, c或t	
<400> 20	
atg gag aca gca cgg att gca gtt gtc ggg gca ggt gtg gtg ggg ctc	48
Met Asp Thr Ala Arg Ile Ala Val Val Gly Ala Gly Val Val Gly Leu	
1 5 10 15	
tcc acg gct gtg tgc atc tcc aaa ctg gtg ccc cga tgc tcc gtt acc	96
Ser Thr Ala Val Cys Ile Ser Lys Leu Val Pro Arg Cys Ser Val Thr	
20 25 30	
atc att tca gac aag ttt act cca gat acc acc agt gat gtg gca gcc	144
Ile Ile Ser Asp Lys Phe Thr Pro Asp Thr Thr Ser Asp Val Ala Ala	
35 40 45	
gga atg ctt att cct cac act tat cca gat aca ccc att cac acg cag	192
Gly Met Leu Ile Pro His Thr Tyr Pro Asp Thr Pro Ile His Thr Gln	
50 55 60	
aag cag tgg ttc aga gaa acc ttt aat cac ctc ttt gca att gcc aat	240
Lys Gln Trp Phe Arg Glu Thr Phe Asn His Leu Phe Ala Ile Ala Asn	
65 70 75 80	
tct gca gaa gct gga gat gct ggt gtt cat ttg gta tca ggg ata aag	288
Ser Ala Glu Ala Gly Asp Ala Gly Val His Leu Val Ser Gly Ile Lys	
85 90 95	
gga agt gga ggc tgg aca ctc act cgg cga ata gaa gac ctg tgg gaa	336
Gly Ser Gly Gly Trp Thr Leu Thr Arg Arg Ile Glu Asp Leu Trp Glu	
100 105 110	
ctt cat ccg tcc ttt gac atc gtg gtc aac tgt tca ggc ctt gga agc	384

```

Leu His Pro Ser Phe Asp Ile Val Val Asn Cys Ser Gly Leu Gly Ser
115 120 125
aga cag ctt gca gga gac tca aag att ttc cct gta agg ggc caa gtc 432
Arg Gln Leu Ala Gly Asp Ser Lys Ile Phe Pro Val Arg Gly Gln Val
130 135 140
ctc caa gtt cag gct ccc tgg gtg gag cat ttt atc cga gat ggc agt 480
Leu Gln Val Gln Ala Pro Trp Val Glu His Phe Ile Arg Asp Gly Ser
145 150 155 160
ggg ctg aca tat att tat cct ggt aca tcc cat gta acc cta ggt gga 528
Gly Leu Thr Tyr Ile Tyr Pro Gly Thr Ser His Val Thr Leu Gly Gly
165 170 175
act agg caa aaa ggg gac tgg aat ctg tcc ccg gat gca gaa aat agc 576
Thr Arg Gln Lys Gly Asp Trp Asn Leu Ser Pro Asp Ala Glu Asn Ser
180 185 190
aga gag att ctt tcc cga tgc tgt gct ctg gag ccc tcc ctc cac gga 624
Arg Glu Ile Leu Ser Arg Cys Cys Ala Leu Glu Pro Ser Leu His Gly
195 200 205
gcc tgc aac atc agg gag aag gtg ggc ttg agg ccc tac agg cca ggc 672
Ala Cys Asn Ile Arg Glu Lys Val Gly Leu Arg Pro Tyr Arg Pro Gly
210 215 220
gtg cga ctg cag aca gag ctc ctt gcg cga gat gga cag agg ctg cct 720
Val Arg Leu Gln Thr Glu Leu Leu Ala Arg Asp Gly Gln Arg Leu Pro
225 230 235 240
gta gtc cac cac tat ggc cat ggg agt ggg ggc atc tca gtg cac tgg 768
Val Val His His Tyr Gly His Gly Ser Gly Gly Ile Ser Val His Trp
245 250 255
ggc act gct ctg gag gcc gcc agg ctg gtg agc gag tgt gtc cat gcc 816
Gly Thr Ala Leu Glu Ala Ala Arg Leu Val Ser Glu Cys Val His Ala
260 265 270
ctc agg acc ccc att ccc aag tca aac ctg tag atgacataaa atgacagcaa 869
Leu Arg Thr Pro Ile Pro Lys Ser Asn Leu *
275 280
agagactgag agactgttga tcaaagcaca gaacaggttc aaataacttt tccactgcat 929
gaaagttaa ttagacattt cttgttttc aacattagaa gtggtgtaac atgtaagctg 989
agcacggtag catgcctata gtcccagcta cttgatggac acagcacgga 1041
ttgcagttgt cggggcaggt gtggtggggc tctccacggc tgtgtgcatc tccaaactgg 1101
tgccccgatg ctccgttacc atcatttcag acaagtttac tccagatacc accagtgatg 1161
tggcagccgg aatgcttatt cctcacactt atccagatac acccattcac acgcagaagc 1221
agtgtttcag agaaaccttt aatcacctct ttgcaattgc caattctgca gaagctggag 1281
atgctggtgt tcatttggtg tcagggataa agggaagtgg aggctggaca ctactcggc 1341
gaatagaaga cctgtgggaa cttcatccgt ctttgacat cgttgtcaac tgttcaggcc 1401
ttggaagcag acagcttgca ggagactcaa agattttccc tgtaaggggc caagtcctcc 1461
aagttcagcg tccctgggtg gagcatttta tccgagatgg cagtgggctg acatatattt 1521
atcctgtgtac atcccatgta accctaggtg gaactaggca aaaaggggac tggaatctgt 1581
ccccggatgc agaaaatagc agagagattc tttcccgatg ctgtgctctg gagccctccc 1641
tccacggagc ctgcaacatc agggagaagg tgggcttgag gccctacagg ccaggcgtgc 1701
gactcgagac agagctcctt gcgcgagatg gacagaggct gcctgtagtc caccactatg 1761
gccatgggag tgggggcacg tcagtgcact ggggcactgc tctggaggcc gccaggctgg 1821
tgagcgatg tgccatgcc ctcaggaccc ccattcccaa gtcaaacctg tagatgacat 1881
aaaatgacag caaagagact gagagactgt tgatcaaagc acagaacagg ttcaaataac 1941
ttttccactg catgaaagt taattagaca tttctttgtt ttcaacatta gaagtgggtg 2001
aacatgtaag ctgagcacgg tagcatgcct atagtccag ctacttg 2048

```

<210> 21

<211> 341

<212> PRT

<213> 智人 (Homo sapiens)

<400> 21

```

Met Asp Thr Ala Arg Ile Ala Val Val Gly Ala Gly Val Val Gly Leu
1      5      10      15
Ser Thr Ala Val Cys Ile Ser Lys Leu Val Pro Arg Cys Ser Val Thr
      20      25      30
Ile Ile Ser Asp Lys Phe Thr Pro Asp Thr Thr Ser Asp Val Ala Ala
      35      40      45
Gly Met Leu Ile Pro His Thr Tyr Pro Asp Thr Pro Ile His Thr Gln
      50      55      60
Lys Gln Trp Phe Arg Glu Thr Phe Asn His Leu Phe Ala Ile Ala Asn
65      70      75      80
Ser Ala Glu Ala Gly Asp Ala Gly Val His Leu Val Ser Gly Trp Gln
      85      90      95
Ile Phe Gln Ser Thr Pro Thr Glu Glu Val Pro Phe Trp Ala Asp Val
      100     105     110
Val Leu Gly Phe Arg Lys Met Thr Glu Ala Glu Leu Lys Lys Phe Pro
      115     120     125
Gln Tyr Val Phe Gly Gln Ala Phe Thr Thr Leu Lys Cys Glu Cys Pro
      130     135     140
Ala Tyr Leu Pro Trp Leu Glu Lys Arg Ile Lys Gly Ser Gly Gly Trp
145     150     155     160
Thr Leu Thr Arg Arg Ile Glu Asp Leu Trp Glu Leu His Pro Ser Phe
      165     170     175
Asp Ile Val Val Asn Cys Ser Gly Leu Gly Ser Arg Gln Leu Ala Gly
      180     185     190
Asp Ser Lys Ile Phe Pro Val Arg Gly Gln Val Leu Gln Val Gln Ala
      195     200     205
Pro Trp Val Glu His Phe Ile Arg Asp Gly Ser Gly Leu Thr Tyr Ile
      210     215     220
Tyr Pro Gly Thr Ser His Val Thr Leu Gly Gly Thr Arg Gln Lys Gly
225     230     235     240
Asp Trp Asn Leu Ser Pro Asp Ala Glu Asn Ser Arg Glu Ile Leu Ser
      245     250     255
Arg Cys Cys Ala Leu Glu Pro Ser Leu His Gly Ala Cys Asn Ile Arg
      260     265     270
Glu Lys Val Gly Leu Arg Pro Tyr Arg Pro Gly Val Arg Leu Gln Thr
      275     280     285
Glu Leu Leu Ala Arg Asp Gly Gln Arg Leu Pro Val Val His His Tyr
      290     295     300
Gly His Gly Ser Gly Gly Ile Ser Val His Trp Gly Thr Ala Leu Glu
305     310     315     320
Ala Ala Arg Leu Val Ser Glu Cys Val His Ala Leu Arg Thr Pro Ile
      325     330     335
Pro Lys Ser Asn Leu
      340

```

<210> 22

<211> 282

<212> PRT

<213> 智人 (Homo sapiens)

<400> 22

```

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

```

<210> 23

<211> 47

<212> DNA

<213> 人工序列

<220>

<223> 寡核苷酸 24-1443-126

<400> 23

tacggcttag taagttggag aacyaggatc agaagacagg tctgcct

47

<220>

<221> allele

<222> 24

<223> 多态性碱基C或T

<210> 24

<211> 47

<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸 24-1457-52

<220>
<221> allele
<222> 24
<223> 多态性碱基C或A

<400> 24
tctgagatgc ccctgtgtcc tctmagggag tagtggctga gcatttc 47

<210> 25
<211> 47
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸 27-93-181

<220>
<221> allele
<222> 24
<223> 多态性碱基C或T

<400> 25
cccagctctg ccaactggcga gctytgtggc cttgggcaag ttactcc 47

<210> 26
<211> 47
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸 24-1461-256

<220>
<221> allele
<222> 24
<223> 多态性碱基A或G

<400> 26
gatggctctg gcattttcag ggarcagtca tgtctgatct caagttc 47

重组的人D-氨基酸氧化酶
在酵母菌酿酒酵母混合提取液中的活性

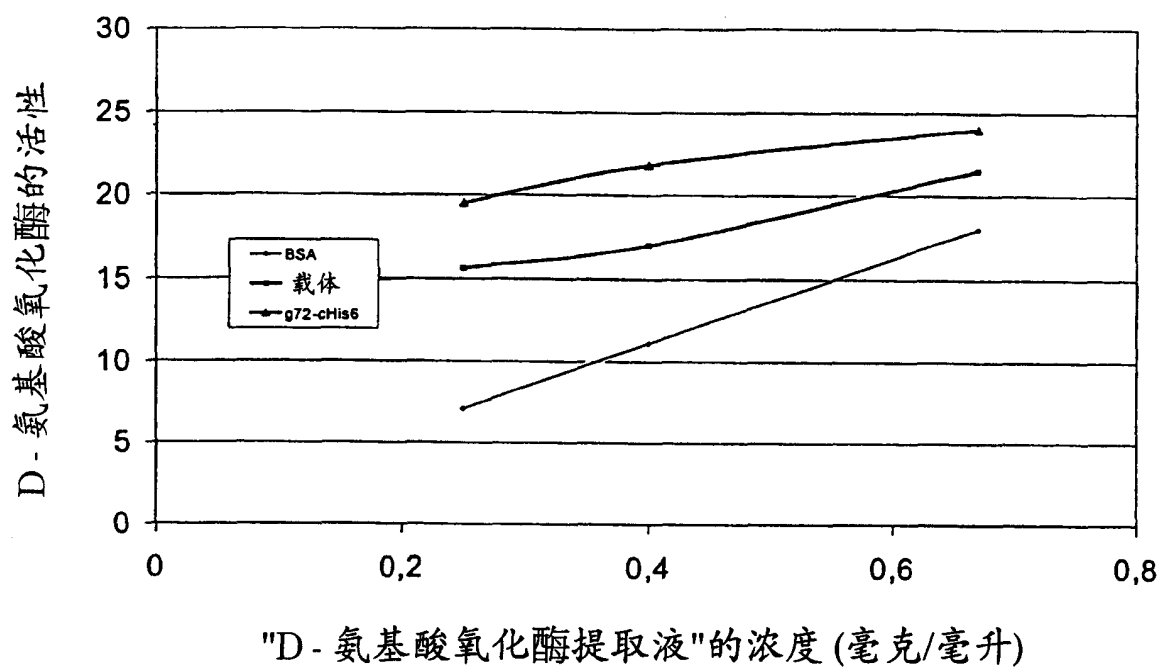


图 1

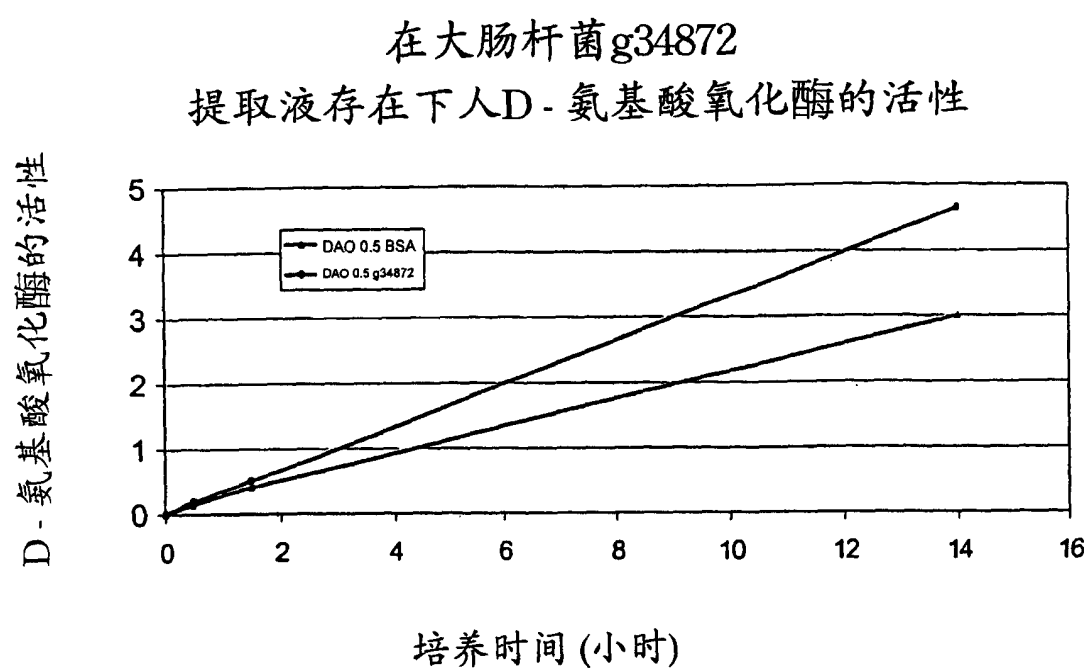


图 2

不同浓度的g34872下D-氨基酸氧化酶的活性

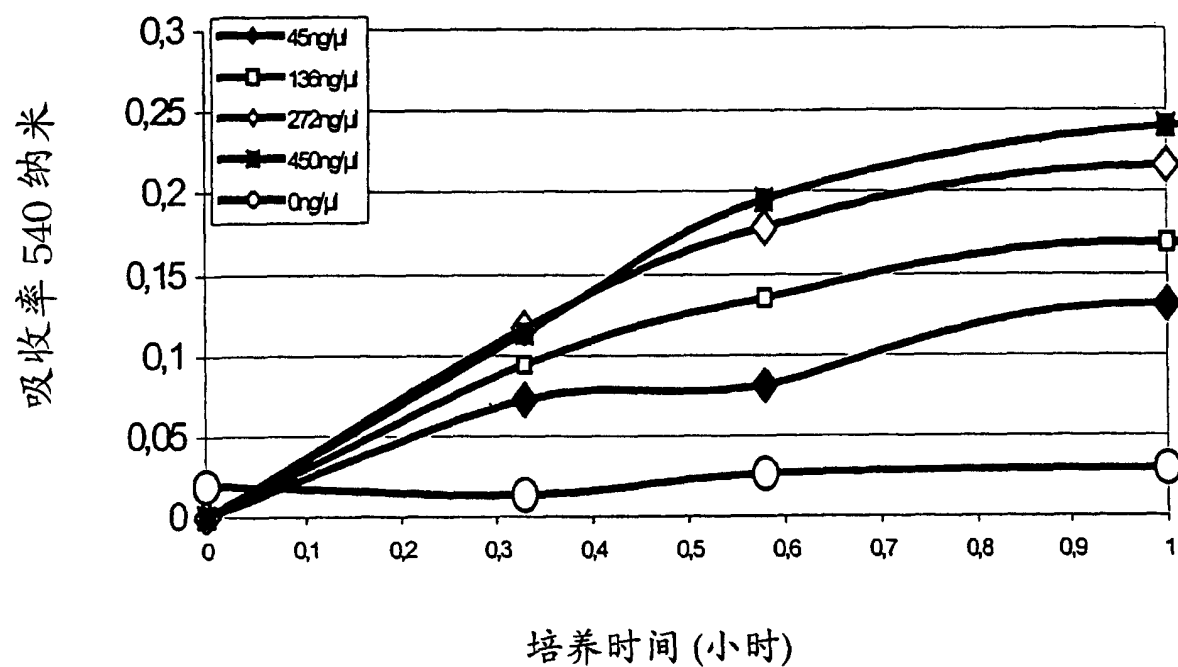


图 3

不同浓度的g34872下D-氨基酸氧化酶的活性

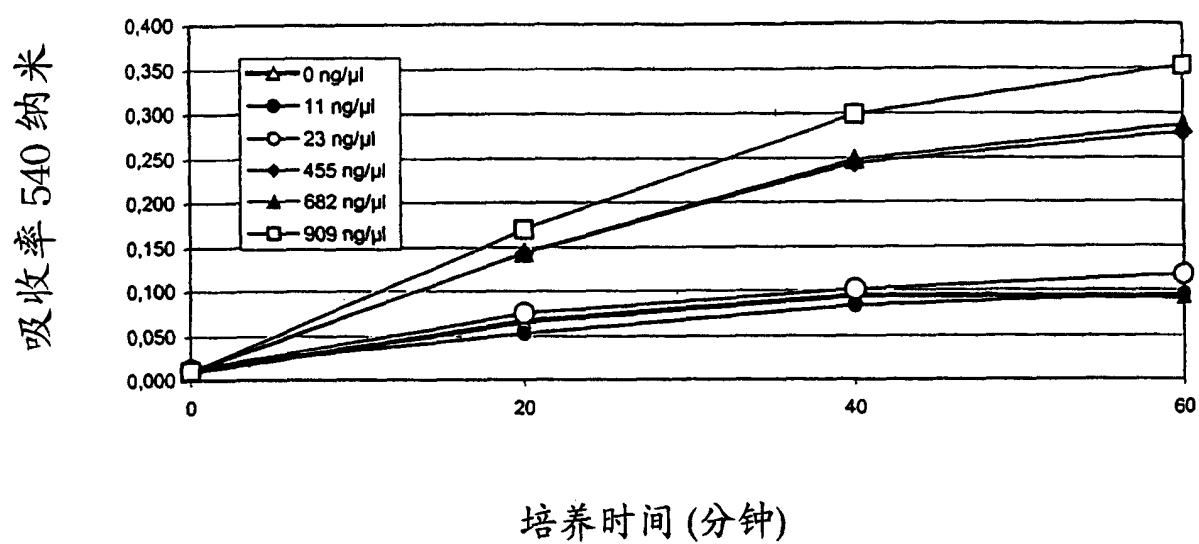


图 4

D-氨基酸氧化酶动力学

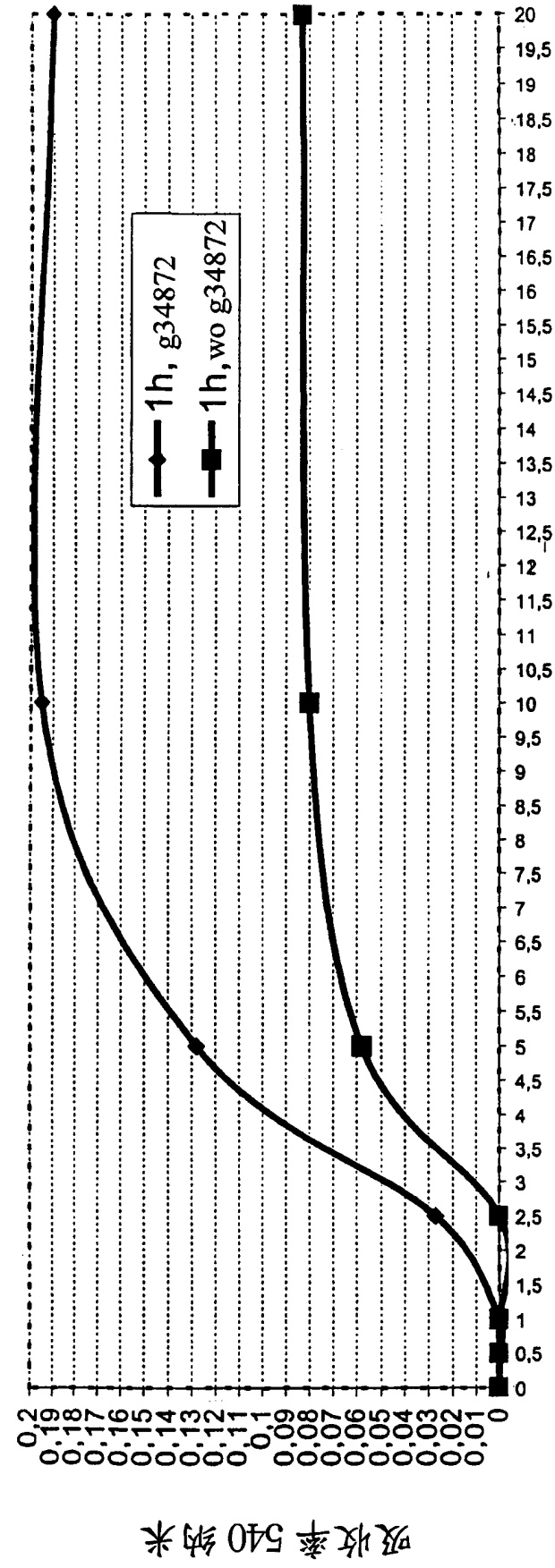


图 5 D-丝氨酸 (mM)

D-氨基酸氧化酶-PJ 27 Algene 样本
的关联性分析结果 (213个病例, 241个对照)

标记	位置	位点	等位基因试验				基因型试验	
			所选的 等位基因	等位基因 频率差值	卡方值	P 值	卡方值	P 值
24- 1443/1 26	基因 5'	132905	T	0.004	0.0171	8.96E-01	0.0206	9.90E-01
24- 1457/5 2	内含子 B	94290	C	0.029	0.6981	4.03E-01	1.0478	5.92E-01
27- 93/181	内含子 3	65365	T	0.079	7.3783	6.60E-03	8.9084	1.16E-02
24- 1461/2 56	内含子 10	55058	G	0.077	6.4495	1.11E-02	6.5728	3.74E-02
(*)：精确检验								

图 6