

[19] 中华人民共和国国家知识产权局

[51] Int. Cl<sup>7</sup>



# [12] 发明专利申请公开说明书

**C12Q 1/68**

G01N 33/68 C12N 15/85

C12N 15/63 C07K 14/56

C07K 16/00 C07K 16/24

A61K 38/21 A61P 9/00

A61P 31/12 A61P 35/00

A61P 37/00

[21] 申请号 02809055.1

[43] 公开日 2004 年 9 月 22 日

[11] 公开号 CN 1531600A

[22] 申请日 2002.3.29 [21] 申请号 02809055.1

[30] 优先权

[32] 2001. 3. 30 [33] FR [31] 01/04404

[86] 国际申请 PCT/EP2002/004082 2002. 3. 29

[87] 国际公布 WO2002/079249 英 2002. 10. 10

[85] 进入国家阶段日期 2003. 10. 28

[71] 申请人 吉恩奥迪塞公司

地址 法国库尔塔贝夫

[72] 发明人 J-L·伊斯卡利

[74] 专利代理机构 中国国际贸易促进委员会专利  
商标事务所

代理人 刘晓东

权利要求书 6 页 说明书 52 页 序列表 6 页  
附图 6 页

[54] 发明名称 IFN $\alpha$  -21 基因的新的多核苷酸和多肽

[57] 摘要

本发明涉及来源于 IFN $\alpha$  -21 基因核苷酸序列的包含新的 SNP 的新的多核苷酸、以及来源于天然野生型 IFN $\alpha$  -21 蛋白质的包含由至少一个本发明的 SNP 所致至少一个突变的新的多肽、及其治疗用途。

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

1. 分离的多核苷酸，其包含全部或部分的：
  - a) 核苷酸序列 SEQ ID NO. 1，只要所述核苷酸序列包含至少一个选自 c794g、c973a、g1011c、t1049a、t1155a、a1204g 和 t1265c 的 SNP；或
  - b) 与 a) 中核苷酸序列互补的核苷酸序列。
2. 权利要求 1 的分离的多核苷酸，其包含 SEQ ID NO. 1 的 670-1239 位核苷酸，只要该序列包含至少一个选自 c794g、c973a、g1011c、t1049a、t1155a 和 a1204g 的编码 SNP。
3. 权利要求 1 的分离的多核苷酸，其中所述的多核苷酸由至少 10 个核苷酸组成。
4. 分离的多核苷酸，其编码包含氨基酸序列 SEQ ID NO. 2 的全部或部分、并具有至少一个选自 A42G、Q102K、Q114H、V127D、C162stop 和 K179E 的编码 SNP 的多肽。
5. 分离的多核苷酸，其编码包含氨基酸序列 SEQ ID NO. 2 的全部或部分、并具有两个编码 SNP Q114H 和 V127D 的多肽。
6. 分离的多核苷酸，其编码包含氨基酸序列 SEQ ID NO. 2 的全部或部分、并具有编码 SNP K179E 的多肽。
7. 鉴定或扩增与核苷酸序列 SEQ ID NO. 1 具有 80-100% 同一性的多核苷酸的全部或部分的方法，其包含该多核苷酸在合适的杂交条件下与权利要求 1 的多核苷酸杂交。
8. 与核苷酸序列 SEQ ID NO. 1 具有 80-100% 同一性的多核苷酸的全部或部分的基因分型方法，其包含以下步骤：扩增受试者或受试者群体基因组 DNA 的目的区域，并确定位于核苷酸序列 SEQ ID NO. 1 中选自 794、973、1011、1049、1155、1204 和 1265 的至少一个位点的等位基因。
9. 权利要求 8 的方法，其中通过微型测序进行基因分型。
10. 包含权利要求 1 的多核苷酸的重组载体。

11. 包含权利要求 10 的重组载体的宿主细胞。
12. 分离多肽的方法，其包含在培养基中培养权利要求 11 的宿主细胞，并从培养基中分离所述多肽。
13. 由权利要求 1 的分离的多核苷酸编码的多肽。
14. 分离的多肽，其包含氨基酸序列 SEQ ID NO. 2 的全部或部分，并具有至少一个选自 A42G、Q102K、Q114H、V127D、C162stop 和 K179E 的编码 SNP。
15. 权利要求 13 的多肽，其包含氨基酸序列 SEQ ID NO. 2 的氨基酸 24-189、并具有至少一个选自 A42G、Q102K、Q114H、V127D、C162stop 和 K179E 的编码 SNP。
16. 权利要求 13 的多肽，其包含氨基酸序列 SEQ ID NO. 2 中 24-189 位氨基酸、并具有两个编码 SNP Q114H 和 V127D。
17. 权利要求 13 的多肽，其包含氨基酸序列 SEQ ID NO. 2 中 24-189 位氨基酸、并具有编码 SNP K179E。
18. 获得免疫特异性抗体的方法，其包含利用权利要求 13 的多肽免疫动物，并从该动物收集所述抗体。
19. 权利要求 18 的方法所产生的免疫特异性抗体。
20. 在一种或多种待测化合物中鉴定可活化或抑制分离多肽活性的试剂的方法，该分离的多肽包含氨基酸序列 SEQ ID NO. 2 的全部或部分、并具有至少一个选自 A42G、Q102K、Q114H、V127D、C162stop 和 K179E 的编码 SNP，所述方法包含：
  - a) 提供包含权利要求 10 的重组载体的宿主细胞；
  - b) 所述宿主细胞与待测化合物接触，
  - c) 确定对所述多肽活性的活化或抑制作用，从而鉴定所述的活化剂或抑制剂。
21. 在一种或多种待测化合物中鉴定其活性受分离多肽增强或抑制的试剂的方法，该分离的多肽包含氨基酸序列 SEQ ID NO. 2 的全部或部分、并具有至少一个选自 A42G、Q102K、Q114H、V127D、C162stop 和 K179E 的编码 SNP，所述方法包含：

- a) 提供包含权利要求 10 的重组载体的宿主细胞;
- b) 所述宿主细胞与待测化合物接触,
- c) 确定对该试剂活性的增强或抑制作用, 从而鉴定所述被增强或被抑制的试剂。

22. 分析受试者生物学特征的方法, 其包含进行至少一个下列步骤:

- a) 确定在受试者基因组中存在或缺少权利要求 1 的多核苷酸;
- b) 确定在受试者中权利要求 1 的多核苷酸的表达水平;
- c) 确定在受试者中存在或缺少权利要求 13 的多肽;
- d) 确定在受试者中权利要求 13 的多肽的浓度; 或
- e) 确定在受试者中权利要求 13 的多肽的功能性。

23. 包含一种或多种化合物的治疗剂, 所述化合物选自包含核苷酸序列 SEQ ID NO. 1 的全部或部分的分离的多核苷酸, 只要该核苷酸序列包含至少一个选自 c794g、c973a、g1011c、t1049a、t1155a、a1204g 和 t1265c 的 SNP, 或者与所述核苷酸序列互补的核苷酸序列; 包含所述多核苷酸的重组载体; 包含所述重组载体的宿主细胞; 包含氨基酸序列 SEQ ID NO. 2 的全部或部分、并具有至少一个选自 A42G、Q102K、Q114H、V127D、C162stop 和 K179E 的编码 SNP 的分离的多肽; 特异性针对所述多肽的抗体。

24. 用于预防或治疗个体疾病的方法, 所述疾病选自癌症和肿瘤、传染病、免疫相关疾病和/或自身免疫相关疾病、心血管疾病、代谢疾病、中枢神经系统疾病以及化学治疗相关疾病, 其包含给予个体治疗有效量的权利要求 23 的治疗剂, 外加药物可接受的赋形剂。

25. 权利要求 24 的方法, 其中所述癌症和肿瘤包含转移的肾癌、黑色素瘤、包含滤泡性淋巴瘤和皮肤 T 细胞淋巴瘤的淋巴瘤、包含毛样细胞白血病、慢性淋巴细胞性白血病和慢性骨髓性白血病的白血病、肝癌、颈癌、头癌和肾癌、多发性骨髓瘤、类癌瘤以及伴随免疫缺陷出现的肿瘤, 包含 AIDS 情况下的 Kaposi 肉瘤。

26. 权利要求 24 的方法, 其中所述代谢疾病包含非免疫相关性疾病, 例如肥胖症。

27. 权利要求 24 的方法，其中所述传染病包含病毒性传染，包括慢性乙型和丙型肝炎以及 HIV/AIDS，感染性肺炎以及性病，例如生殖器疣。

28. 权利要求 24 的方法，其中所述中枢神经系统疾病包含阿尔茨海默氏病、帕金森氏病、精神分裂症以及抑郁症。

29. 权利要求 24 的方法，其中所述免疫和自身免疫相关疾病包含组织或器官移植排斥、变态反应、哮喘、银屑病、类风湿性关节炎、多发性硬化、克罗恩氏病以及溃疡性结肠炎。

30. 预防或治疗个体疾病的方法，所述疾病选自创伤愈合、透析患者的贫血和/或骨质疏松症，包含给予该个体治疗有效量的权利要求 23 的治疗剂，外加药物可接受的赋形剂。

31. 增加或降低受试者中权利要求 13 的多肽活性的方法，其包含给予治疗有效量的一种或多种：包含核苷酸序列 SEQ ID NO. 1 的全部或部分的分离的多核苷酸，只要该核苷酸序列包含至少一个选自 c794g、c973a、g1011c、t1049a、t1155a、a1204g 和 t1265c 的 SNP，或与该核苷酸序列互补的核苷酸序列；包含该多核苷酸的重组载体；包含该重组载体的宿主细胞，其中该宿主细胞可得自待治疗的受试者；包含氨基酸序列 SEQ ID NO. 2 的全部或部分、并具有至少一个选自 A42G、Q102K、Q114H、V127D、C162stop 和 K179E 的编码 SNP 的分离的多肽；特异性针对该多肽的抗体；以及药物可接受的赋形剂。

32. 预防或治疗个体与其基因组中存在权利要求 1 的多核苷酸有关的病症或疾病的方法，其包含给予治疗有效量的一种或多种：包含核苷酸序列 SEQ ID NO. 1 的全部或部分、并具有至少一个选自 c794g、c973a、g1011c、t1049a、t1155a、a1204g 和 t1265c 的 SNP 的分离的多核苷酸，或与该核苷酸序列互补的核苷酸序列；包含一种该多核苷酸的重组载体；包含该重组载体的宿主细胞；包含氨基酸序列 SEQ ID NO. 2 的全部或部分、并具有至少一个选自 A42G、Q102K、Q114H、V127D、C162stop 和 K179E 的编码 SNP 的分离的多肽；特异性针对一种该多肽的抗体；以及药物可接受的赋形剂。

33. 确定至少一个选自 IFN $\alpha$ -21 基因中 c794g、c973a、g1011c、

t1049a、t1155a、a1204g 和 t1265c 的 SNP 与疾病或疾病抗性之间统计学相关性的方法，其包含：

- a) 对一组个体进行基因分型；
- b) 确定该疾病或疾病抗性在该组个体中的分布；
- c) 将基因型数据与该疾病或疾病抗性的分布作比较；以及
- d) 对上述比较进行统计学相关性分析。

34. 诊断或确定疾病或疾病抗性预后的方法，其包含检测 IFN $\alpha$ -21 基因中至少一个选自 c794g、c973a、g1011 c、t1049a、t1155a、a1204g 和 t1265c 的 SNP。

35. 在一种或多种待测化合物中鉴定与 Q114H/V127D 突变型 IFN $\alpha$ -21 基因产物具有基本类似的生物活性的化合物的方法，该方法包含以下步骤：

a) 确定所述化合物的生物活性，例如树突状细胞成熟、CD4<sup>+</sup>或 CD8<sup>+</sup>T 淋巴细胞释放细胞因子、单核细胞释放细胞因子、体外或体内抗病毒活性、对 Daudi Burkitt 细胞系的细胞抗增殖活性、对 TF-1 细胞系的细胞抗增殖活性；

b) 将由步骤 a) 确定的该化合物活性与 Q114H/V127D 突变型 IFN $\alpha$ -21 基因产物活性作比较。

c) 根据步骤 b) 进行的比较确定该化合物是否具有与 Q114H/V127D 突变型 IFN $\alpha$ -21 基因产物基本类似的活性。

36. 权利要求 35 的方法，其中所述待测化合物鉴定自合成肽组合文库、高通量筛选、或通过计算机辅助药物设计而设计，以具有与 SEQ ID NO. 2 的多肽、或包含氨基酸序列 SEQ ID NO. 2 的 24-189 位所含氨基酸的氨基酸序列相同的三维结构，只要该氨基酸序列包含 Q114H 和 V127D SNP。

37. 利用权利要求 35 的方法鉴定的化合物。

38. 从一种或多种待测化合物中鉴定与 K179E 突变型 IFN $\alpha$ -21 基因产物具有基本类似的生物活性的化合物的方法，该方法包含以下步骤：

a) 确定所述化合物的生物活性，例如树突状细胞成熟、CD4<sup>+</sup>或 CD8<sup>+</sup>T 淋巴细胞释放细胞因子、单核细胞释放细胞因子、体外或体内抗病毒活

性、对 Daudi Burkitt 细胞系的细胞抗增殖活性、对 TF-1 细胞系的细胞抗增殖活性;

b) 将由步骤 a) 确定的该化合物活性与 K179E 突变型 IFN $\alpha$ -21 基因产物活性作比较。

c) 根据步骤 b) 进行的比较, 确定该化合物是否具有与 K179E 突变型 IFN $\alpha$ -21 基因产物基本类似的活性。

39. 权利要求 38 的方法, 其中所述待测化合物鉴定自合成肽组合文库、高通量筛选、或通过计算机辅助药物设计而设计, 以具有与 SEQ ID NO. 2 的多肽、或包含氨基酸序列 SEQ ID NO. 2 的 24-189 位所含氨基酸的氨基酸序列相同的三维结构, 只要该氨基酸序列包含 K179E SNP。

40. 利用权利要求 38 的方法鉴定的化合物。

41. 用于预防或治疗个体疾病的方法, 所述疾病选自癌症和肿瘤、传染病、免疫相关疾病和/或自身免疫相关疾病、心血管疾病、代谢疾病、中枢神经系统疾病以及化学治疗相关疾病, 其包含给予该个体治疗有效量的权利要求 37 或 40 的试剂, 外加药物可接受的赋形剂。

42. 权利要求 41 的方法, 其中所述癌症和肿瘤包含转移的肾癌、黑色素瘤、包含滤泡性淋巴瘤和皮肤 T 细胞淋巴瘤的淋巴瘤、包含毛样细胞白血病、慢性淋巴细胞性白血病和慢性骨髓性白血病的白血病、肝癌、颈癌、头癌和肾癌、多发性骨髓瘤、类癌瘤以及伴随免疫缺陷出现的肿瘤, 包含 AIDS 情况下的 Kaposi 肉瘤。

43. 权利要求 41 的方法, 其中所述传染病包含病毒性传染, 包括慢性乙型和丙型肝炎以及 HIV/AIDS, 感染性肺炎以及性病, 例如生殖器疣。

44. 权利要求 41 的方法, 其中所述免疫和自身免疫相关疾病包含组织或器官移植排斥、变态反应、哮喘、银屑病、类风湿性关节炎、多发性硬化、克罗恩氏病以及溃疡性结肠炎。

45. 权利要求 41 的方法, 其中所述中枢神经系统疾病包含阿尔茨海默氏病、帕金森氏病、精神分裂症以及抑郁症。

46. 权利要求 41 的方法, 其中所述代谢疾病包含非免疫相关性疾病, 例如肥胖症。

## IFN $\alpha$ -21 基因的新的多核苷酸和多肽

### 相关申请

本发明要求对 2001 年 3 月 30 日提交的、名为《<Nouveaux polynucléotides comportant des polymorphismes de type SNP fonctionnels dans la séquence nucléique du gène IFNalpha21 ainsi que de nouveaux polypeptides codés par ces polynucléotides et leurs utilisations thérapeutiques.>>的法国专利申请 0104404 的优先权。

### 发明背景

#### 发明领域

本发明涉及来源于 IFN $\alpha$ -21 基因核苷酸序列的包含新的 SNP 的新的多核苷酸、以及来源于天然野生型 IFN $\alpha$ -21 蛋白质的包含该 SNP 所致突变的新的多肽、及其治疗用途。

#### 相关技术

干扰素  $\alpha$ 21 基因，以下简称 IFN $\alpha$ -21，如下列文献所述：

- Goeddel, D. V., Leung, D. W; "The structure of eight distinct cloned human leukocyte interferon cDNAs"; Nature 290 (5801), 20-26 (1981).

- Olopade OI., Bohlander SK.; "Mapping of the shortest region of overlap of deletions of the short arm of chromosome 9 associated with human neoplasia"; Genomics 14 (2), 437-443 (1992).

该基因的核苷酸序列可得自 GenBank 数据库的 HTG 部分，登录号为 AC009445。

IFN $\alpha$ -21 信使 RNA 序列如 NCBI 数据库登录号 NM\_002175 所述。

IFN $\alpha$ -21 基因的结构和功能同源性接近于人干扰素  $\alpha$  (IFN $\alpha$ ), 特别是 IFN $\alpha$ -2。

已知 IFN $\alpha$  具有抗细胞增殖作用, 并与抗病毒和抗寄生虫反应有关。

已知 IFN $\alpha$  还抑制其它几种细胞因子在造血干细胞水平的表达, 并抑制某些肿瘤的细胞增殖。

已知 IFN $\alpha$  还降低肾癌 EGF 受体的表达, 抑制某些线粒体基因的表达, 抑制成纤维细胞、单核细胞和 B 淋巴细胞的增殖, 特别是在体外, 并阻断 B 淋巴细胞的抗体合成。

已知 IFN $\alpha$  还诱导肿瘤细胞表面上的肿瘤特异性抗原的表达, 并通过作用于 ISRE (干扰素刺激反应元件) 的特定转录因子, 诱导位于 ISRE 型启动子区控制下的基因。

已知 IFN $\alpha$  与各种病症和/或人类疾病有关, 例如各种癌症, 如恶性肿瘤、黑色素瘤、淋巴瘤、白血病和肝癌、颈癌、头癌和肾癌, 心血管疾病, 与免疫系统无关的代谢疾病, 例如肥胖症, 传染病, 例如乙型肝炎、丙型肝炎和 AIDS, 肺炎, 溃疡性结肠炎, 中枢神经系统疾病, 例如阿尔茨海默氏病、精神分裂症和抑郁症, 组织或器官移植的排斥, 创伤愈合, 透析患者的贫血, 变态反应, 哮喘, 多发性硬化, 骨质疏松症, 银屑病, 类风湿性关节炎, 克罗恩氏病 (Crohn's disease), 自身免疫疾病和病症、胃肠疾病乃至与化学治疗有关的疾病。

IFN $\alpha$  特别用于治疗某些白血病、转移性肾癌以及伴随免疫缺陷出现的肿瘤, 例如爱滋病情况下的 Kaposi 肉瘤。IFN $\alpha$  还有效对抗其它类型肿瘤和某些病毒感染。FDA (食品和药品管理局) 确认 IFN $\alpha$  也可用于治疗生殖器疣或性病。

更具体而言, 利用原位杂交将 IFN $\alpha$ -21 定位于帕金森氏病或阿尔茨海默氏病患者的脑中。

与其它细胞相比, 小神经胶质 (microglial) 细胞大量表达 IFN $\alpha$ -21。

阿尔茨海默氏病患者的顶叶神经元显示存在 IFN $\alpha$ -21, 提示

IFN $\alpha$ -21 可能与其病理有关 (例如参见 Kawaguchi N, Yamada T, Yoshiyama Y. No To Shinkei. 1997 Jan.; 49(1): 69-73)。

然而, IFN $\alpha$  特别是 IFN $\alpha$ -21 用于药物组合物时具有很多副作用, 例如急性超敏反应(荨麻疹、支气管收缩、过敏性休克等)、心率不齐、低血压、癫痫发作、甲状腺功能问题、流感样综合症(发热、出汗、肌痛)等。

此外, 用 IFN $\alpha$  治疗的患者可以产生中和该分子的抗体, 从而降低其效率。

本发明人发现 IFN $\alpha$ -21 基因的新的多肽和新的多核苷酸类似物, 其能够具有不同于天然野生型 IFN $\alpha$ -21 蛋白质的功能性。

这些新的多肽和多核苷酸可以特别用于治疗或预防前述病症或疾病, 并避免其所具有的全部或部分缺点。

#### 发明简述

本发明的首要目的在于新的多核苷酸, 其因包含一个或几个 SNP (单核苷酸多态性) 而不同于参考的野生型 IFN $\alpha$ -21 基因核苷酸序列。

参考的野生型人 IFN $\alpha$ -21 基因的核苷酸序列 SEQ ID NO. 1 由 2001 个核苷酸组成, 包含核苷酸 670 (起始密码子)- 1239 (终止密码子) 的 570 个核苷酸的编码序列。

申请人已经鉴定了参考的野生型 IFN $\alpha$ -21 基因核苷酸序列中的 8 个 SNP。这 8 个 SNP 如下: c794g、c973a、g1011c、t1049a、t1155a、a1204g、t1265c、t1277c。

应当理解, 在本发明意义上, 定位上文定义的 SNP 的相应编号是相对于核苷酸序列 SEQ ID NO. 1 的编号而言。

字母 a、t、c 和 g 分别对应于含氮碱基腺嘌呤、胸腺嘧啶、胞嘧啶和鸟嘌呤。

第一个字母对应于野生型核苷酸, 而最后一个字母对应于突变的核苷酸。

因此, 例如 SNP c794g 对应于参考的野生型 IFN $\alpha$ -21 基因核苷酸

序列 SEQ ID NO. 1 中的 794 位核苷酸胞嘧啶(c)突变为核苷酸鸟嘌呤(g)。

本申请人利用申请人于 2000 年 12 月 6 日提交的名为“Process for the determination of one or several functional polymorphism(s) in the nucleotide sequence of a preselected functional candidate gene and its applications”的专利申请 FR 00 22894(此处引用以供参考)中所述测定方法, 鉴定出了这些 SNP。

该专利申请中所述方法可以从随机个体群体中的至少一个个体中鉴定一个(或几个)已存在的 SNP。

在本发明范围内, 在随机选取的个体群体中的不同个体中分离例如包含编码序列的 IFN $\alpha$ -21 基因核苷酸序列片段。

在 DHPLC 分析(“变性高效液相层析”)之后, 对某些具有异质双链性质(即性质不同于参考的野生型 IFN $\alpha$ -21 基因序列)的样品进行这些片段的测序。

这样测序的片段然后与参考的野生型 IFN $\alpha$ -21 基因片段的核苷酸序列作比较, 并鉴定本发明的 SNP。

因此, 该 SNP 是天然的, 其中每一个均存在于世界群体的某些个体中。

参考的野生型 IFN $\alpha$ -21 基因编码对应于氨基酸序列 SEQ ID NO. 2 的 189 个氨基酸的未成熟蛋白质, 通过断裂包括前 23 个氨基酸的信号肽而转变为 166 个氨基酸的成熟蛋白质。

本发明的每个编码 SNP, 即: c794g、c973a、g1011c、t1049a、t1155a、a1204g, 引起 IFN $\alpha$ -21 基因核苷酸序列编码的蛋白质在氨基酸序列水平的改变。

这些氨基酸序列的改变如下:

SNP c794g 引起对应于氨基酸序列 SEQ ID NO. 2 的 IFN $\alpha$ -21 基因的未成熟蛋白质 42 位氨基酸丙氨酸(A)突变为甘氨酸(G), 其位于成熟蛋白质的 19 位。 在本发明的说明书中, 根据是否涉及成熟蛋白质抑或未成熟蛋白质, 而分别将该 SNP 编码的突变称为 A19G 或 A42G。

SNP c973a 引起对应于氨基酸序列 SEQ ID NO. 2 的 IFN $\alpha$ -21 基因的未成熟蛋白质 102 位氨基酸谷氨酰胺(Q)突变为赖氨酸(K), 其位于成熟蛋白质的 79 位。在本发明的说明书中, 根据是否涉及成熟蛋白质抑或未成熟蛋白质, 而分别将该 SNP 编码的突变称为 Q79K 或 Q102K。

SNP g1011c 引起对应于氨基酸序列 SEQ ID NO. 2 的 IFN $\alpha$ -21 基因的未成熟蛋白质 114 位氨基酸谷氨酰胺(Q)突变为组氨酸(H), 其位于成熟蛋白质的 91 位。在本发明的说明书中, 根据是否涉及成熟蛋白质抑或未成熟蛋白质, 而分别将该 SNP 编码的突变称为 Q91H 或 Q114H。

SNP t1049a 引起对应于氨基酸序列 SEQ ID NO. 2 的 IFN $\alpha$ -21 基因的未成熟蛋白质 127 位氨基酸缬氨酸(V)突变为天冬氨酸(D), 其位于成熟蛋白质的 104 位。在本发明的说明书中, 根据是否涉及成熟蛋白质抑或未成熟蛋白质, 而分别将该 SNP 编码的突变称为 V104D 或 V127D。

SNP t1155a 引起对应于氨基酸序列 SEQ ID NO. 2 的 IFN $\alpha$ -21 基因的未成熟蛋白质 162 位氨基酸半胱氨酸(C)突变为终止密码子(终止), 其位于成熟蛋白质的 139 位。在本发明的说明书中, 根据是否涉及成熟蛋白质抑或未成熟蛋白质, 而分别将该 SNP 编码的突变称为 C139stop 或 C162stop。

SNP a1204g 引起对应于氨基酸序列 SEQ ID NO. 2 的 IFN $\alpha$ -21 基因的未成熟蛋白质 179 位氨基酸赖氨酸(K)突变为谷氨酸(E), 其位于成熟蛋白质的 156 位。在本发明的说明书中, 根据是否涉及成熟蛋白质抑或未成熟蛋白质, 而分别将该 SNP 编码的突变称为 K156E 或 K179E。

SNP c794g、c973a、g1011c、t1049a、t1155a、a1204g 引起符合本发明的多肽较之野生型参考 IFN $\alpha$ -21 基因核苷酸序列编码的多肽, 其空间构象改变。

按照本领域技术人员熟知的方法, 通过计算机分子建模(modeling), 例如利用 de novo(例如 SEQFOLD/MSI)、同源性(例如

MODELER/MSI)、最小化势场(force field)(例如 DISCOVER、DELPHI/MSI)和/或分子动力学(例如 CFF/MSI)模建工具,可以观察到这些改变。

下文实验部分给出了上述模式的实例。

计算机分子模建表明,成熟突变蛋白质的 Q79K 突变包括因氢键破坏而使野生型 IFN $\alpha$ -21 蛋白质螺旋 C 的 N 端移位(displacement),如图 1A 和 1B 所示。

实际上,在 Q79K 突变型 IFN $\alpha$ -21 蛋白质中,野生型 IFN $\alpha$ -21 蛋白的 Q79 残基侧链氧原子、E83 残基的酸性基团以及螺旋 C 间的氢键消失。

因此, Q79K 突变蛋白质具有不同于天然野生型 IFN $\alpha$ -21 蛋白质的三维构象,其结构和功能显著变化。

计算机分子模建表明,成熟突变蛋白质的 Q91H 突变包括螺旋 C 在突变位置的移位,如图 2A 和 2B 所示。出现几个氢键和盐桥,特别位于 H91 和 D76 氨基酸的侧链之间,使螺旋更具刚性。

因此, Q91H 突变蛋白质具有不同于天然野生型 IFN $\alpha$ -21 蛋白质的三维构象,结构和功能显著变化。

计算机分子模建表明,成熟突变蛋白质的 V104D 突变包括螺旋 C 和 D 之间的环状结构在突变位置的改变,如图 3A 和 3B 所示。在突变结构中出现几个氢键(一方面为 Q102 和 G105,另一方面为 Q52、E107 和 T109),使螺旋 C 和 D 之间的环更具刚性。此外,还观察到螺旋 B N 端略微移位。

这些空间改变影响到与 IFN $\alpha$ -21 结合其受体有关的残基。

因此, V104D 突变蛋白质具有不同于天然野生型 IFN $\alpha$ -21 蛋白质的三维构象,结构和功能显著变化。

计算机分子模建表明,成熟突变蛋白质的 C139stop 突变引起蛋白质翻译过早停止,导致通常与野生型 IFN $\alpha$ -21 蛋白质螺旋 E 有关的多肽片段消失,如图 4 所示。

螺旋 E 对于 IFN $\alpha$ -21 与其受体结合必不可少。缺少螺旋 E 导致突

变蛋白质错误折叠，并引起蛋白质三维构象变化，其中蛋白质的疏水核心接触外部亲水性介质。因此，该突变蛋白质必须改变其三维构象，使其疏水核心为亲水性残基所覆盖，以免接触外部亲水性介质。

因此，C139stop 突变蛋白质具有不同于天然野生型 IFN $\alpha$ -21 蛋白质的三维构象，结构和功能显著变化。

计算机分子建模表明，成熟突变蛋白质的 K156E 突变包括螺旋 EC 端的伸展以及 C 末端环状的变化，如图 5A 和 5B 所示。

该突变增强了氢键网络，并在 E156 和 R161 残基之间形成盐桥。这些改变使 IFN $\alpha$ -21 蛋白质结构在此区域更为刚性。已知蛋白质该区域与其抗病毒活性有关。因此，可预测 K156E 突变型 IFN $\alpha$ -21 蛋白质的抗病毒活性被显著破坏，并预测 156 位谷氨酸导致成熟 IFN $\alpha$ -21 的结构和功能改变。

本发明的其它 SNP，即：t1265c、t1277c，并不包含 IFN $\alpha$ -21 基因核苷酸序列编码的蛋白质在氨基酸序列 SEQ ID NO. 2 水平的改变。SNP t1265c、t1277c 是非编码 SNP。

本发明的多核苷酸可以按此方式进行基因分型，以便确定这些多核苷酸在群体中的等位基因频率。下文实验部分给出了基因分型的实施例。

通过测定生物活性，可以同样进行本发明多肽的功能性测定。

在此方面，例如可测定本发明的多肽与天然野生型 IFN $\alpha$ -21 蛋白质相比，对 Daudi 细胞系的抗增殖作用 (Pielher 等人 J. Biol. Chem. ; 275 卷, 51 期, 40425-40433, 12 月 22, 2000; "New structural and Functional Aspects of the type I Interferon-Receptor interaction revealed by comprehensible mutational analysis of the binding interface")。

本发明目的还在于本发明的多核苷酸和多肽、以及从这些多核苷酸和多肽开始而获得和/或鉴定的治疗性分子的用途，特别是用于预防和/或治疗某些人类病症和/或疾病。

附图简述

图 1A 和 1B 表示包含 SNP c973a (Q79K) 的本发明编码蛋白质以及天然野生型 IFN $\alpha$ -21 蛋白质的模型。图 1B 表示图 1A 所示每种蛋白质下部模型的特写 (close up)。

在图 1A 和 1B 中, 黑色条带代表天然野生型 IFN $\alpha$ -21 蛋白质的结构, 而白色条带代表 Q79K 突变型 IFN $\alpha$ -21 蛋白质的结构。

图 2A 和 2B 表示包含 SNP g1011c (Q91H) 的本发明编码蛋白质以及天然野生型 IFN $\alpha$ -21 蛋白质的模型。图 2B 表示图 2A 所示每种蛋白质左部模型的特写。

在图 2A 和 2B 中, 黑色条带代表天然野生型 IFN $\alpha$ -21 蛋白质的结构, 而白色条带代表 Q91H 突变型 IFN $\alpha$ -21 蛋白质的结构。

图 3A 和 3B 表示包含 SNP t1049a (V104D) 的本发明编码蛋白质以及天然野生型 IFN $\alpha$ -21 蛋白质的模型。图 3B 表示图 3A 所示每种蛋白质上部模型的特写。

在图 3A 和 3B 中, 黑色条带代表天然野生型 IFN $\alpha$ -21 蛋白质的结构, 而白色条带代表 V104D 突变型 IFN $\alpha$ -21 蛋白质的结构。

图 4 表示包含 SNP t1155a (C139stop) 的本发明编码蛋白质(图 4B) 以及天然野生型 IFN $\alpha$ -21 蛋白质(图 4A) 的模型。图 4C 表示图 4A 和 4B 的两个蛋白质的叠加。在图 4 中, 黑色条带代表天然野生型 IFN $\alpha$ -21 的蛋白质的结构, 而白色条带代表 C139stop 突变型 IFN $\alpha$ -21 蛋白质的结构。

图 5A 和 5B 表示包含 SNP a1204g (K156E) 的本发明编码蛋白质以及天然野生型 IFN $\alpha$ -21 蛋白质的模型。

图 5B 表示图 5A 所示每种蛋白质上部模型的特写。在图 5 中, 黑色条带代表天然野生型 IFN $\alpha$ -21 蛋白质的结构, 而白色条带代表 K156E 突变型 IFN $\alpha$ -21 蛋白质的结构。

图 6 表示预先经 VSV 病毒感染以及 A42G、Q114H/V127D、或 K179E 突变型 IFN $\alpha$ -21 蛋白质处理的小鼠, 与利用野生型 IFN $\alpha$ -2 处理、或未经处理的小鼠相比的存活率。

在此图中, 横坐标对应存活时间(天), 纵坐标对应 VSV 感染小鼠

的相对存活率。黑色三角形、十字形、黑色菱形分别代表经 A42G 突变型 IFN $\alpha$ -21、Q114H/V127D 突变型 IFN $\alpha$ -21 以及 K179E 突变型 IFN $\alpha$ -21 处理的 VSV 感染小鼠的数据。黑色正方形代表经野生型 IFN $\alpha$ -2 处理的 VSV 感染小鼠的数据，空三角形代表未经处理的 VSV 感染小鼠的数据。

### 发明详述

#### 定义

“参考野生型基因的核苷酸序列”应理解为人 IFN $\alpha$ -21 基因的核苷酸序列 SEQ ID NO. 1。

该序列可得自 GenBank 登录号 AC009445，IFN $\alpha$ -21 信使 RNA 序列如 NCBI 数据库登录号 NM\_002175 所述。此外，人 IFN $\alpha$ -21 基因描述于 Goeddel, D. V., Leung, D. W “The structure of eight distinct cloned human leukocyte interferon cDNAs”; Nature 290 (5801), 20-26 (1981), 以及 Olopade OI., Bohlander SK. “Mapping of the shortest region of overlap of deletions of the short arm of chromosome 9 associated with human neoplasia”; Genomics 14 (2), 437-443 (1992)。

“天然野生型 IFN $\alpha$ -21 蛋白质”应理解为由参考的野生型 IFN $\alpha$ -21 基因的核苷酸序列编码的成熟蛋白质。天然野生型非成熟 IFN $\alpha$ -21 蛋白质对应于 SEQ ID NO. 2 所示肽序列。

“多核苷酸”应理解为多核糖核苷酸或多脱氧核糖核苷酸，可以是修饰或未修饰的 DNA 或 RNA。

术语多核苷酸包括，例如单链或双链 DNA、由一个或几个单链区以及一个或几个双链区的混合物组成的 DNA、单链或双链 RNA 以及由一个或几个单链区以及一个或几个双链区的混合物组成的 RNA。术语多核苷酸还可以包括包含一个或几个三链区的 RNA 和/或 DNA。多核苷酸同样可理解为因稳定性或其它原因包含一个或几个碱基修饰的而修饰骨架的 DNA 和 RNA。修饰碱基可理解为，例如诸如次黄苷等稀有碱

基。

“多肽”应理解为包含通过正常或修饰的肽键(例如, 诸如在同型空间配位肽的情况下)彼此互联的至少两个氨基酸的肽、寡肽、寡聚体或蛋白质。

多肽可以由除遗传密码定义的 20 个氨基酸以外的氨基酸组成。多肽同样可以由利用本领域技术人员熟知的天然过程(诸如翻译后成熟过程)或化学过程修饰的氨基酸组成。文献中充分详述了这些修饰。这些修饰可以出现在多肽的任何位置: 肽骨架、氨基酸链乃至羧基或氨基末端。

多肽可以是泛素化之后的分支状, 或者有或无分支的环状。这类修饰可以是本领域技术人员熟知的天然或合成的翻译后过程的结果。

例如, 多肽修饰可理解为包括乙酰化、酰基化、ADP-核糖基化、酰胺化、黄素的共价固定、血红素的共价固定、核苷酸或核苷酸衍生物的共价固定、脂质或脂质衍生物的共价固定、磷脂酰肌醇的共价固定、共价或非共价交联、环化、二硫键形成、脱甲基化、半胱氨酸形成、焦谷氨酸形成、甲酰化、 $\gamma$ -羧基化、糖基化、GPI 锚形成、羟基化、碘化、甲基化、十四酰化、氧化、蛋白水解过程、磷酸化、异戊二烯化、外消旋化、seneloylation、硫酸酯化(sulfatation)、氨基酸加成, 诸如精氨酸化或泛素化。文献中充分详述了这些修饰: PROTEINS-STRUCTURE AND MOLECULAR PROPERTIES, 第二版, T. E. Creighton, New York, 1993, POST-TRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, 1983, Seifter 等人 “Analysis for protein modifications and nonprotein cofactors”, Meth. Enzymol. (1990) 182: 626-646, 以及 Rattan 等人 “Protein Synthesis: Post-translational Modifications and Aging”, Ann NY Acad Sci (1992) 663: 48-62.

“分离的多核苷酸”或“分离的多肽”应理解为从人体中分离或利用技术方法产生的、分别如前文定义的多核苷酸或多肽。

“同一性”应理解为测定核苷酸或多肽序列的同一性。

同一性是本领域技术人员熟知的术语，在文献中有详述。参见 COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., Ed., Oxford University Press, New York, 1998; BIOCOMPUTING INFORMATICS AND GENOME PROJECT, Smith, D.W., Ed., Academic Press, New York, 1993; COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin, A.M. and Griffin H.G., Ed, Humana Press, New Jersey, 1994; 以及 SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heinje, G., Academic Press, 1987.

文献中同样详述了一般用于确定两个序列间同一性和相似性的方法。参见 GUIDE TO HUGE COMPUTER, Martin J. Bishop, Ed, Academic Press, San Diego, 1994, 以及 Carillo H. 和 Lipton D., Siam J Applied Math (1988) 48: 1073.

例如，与核苷酸序列 SEQ ID NO. 1 具有至少 95 % 同一性的多核苷酸是与该序列相比，每 100 个核苷酸中包含至多 5 个突变点的多核苷酸。

这些突变点可以是一个(或几个)核苷酸的一个(或几个)替换、添加和/或缺失。

同样，例如与氨基酸序列 SEQ ID NO. 2 具有至少 95 % 同一性的多肽是与该序列相比，每 100 个氨基酸中包含至多 5 个突变点的多肽。

这些突变点可以是一个(或几个)氨基酸的一个(或几个)替换、添加和/或缺失。

本发明的多核苷酸和多肽并不完全分别等同于核苷酸序列 SEQ ID NO. 1 或氨基酸序列 SEQ ID NO. 2, 应理解这些序列包含至少一个本发明的 SNP, 而被认为是这些序列的变体。

本发明的多核苷酸通常与包含至少一个本发明的 SNP 的核苷酸序列 SEQ ID NO. 1 具有相同或几乎相同的生物活性。

同样，本发明的多肽通常与包含至少一个本发明的编码 SNP 的氨基酸序列 SEQ ID NO. 2 具有相同或几乎相同的生物活性。

例如，通过定点诱变或直接合成，可以获得本发明的变体。

“SNP”可理解为核苷酸序列中碱基的任何天然变异。核苷酸序列的SNP可以是编码、沉默或非编码型。

编码 SNP 是包括在核苷酸序列的编码序列内、涉及该核苷酸序列所编码氨基酸序列中的氨基酸改变的一种多态性。在这种情况下，术语 SNP 经外延同样适用于氨基酸序列的突变。

沉默 SNP 是包括在核苷酸序列的编码序列内、不涉及该核苷酸序列所编码的氨基酸序列中的氨基酸改变的一种多态性。

非编码 SNP 是包括在核苷酸序列的非编码序列中的一种多态性。该多态性特别发现于内含子、拼接区、转录启动子或增强子位点序列中。

“功能性 SNP”可理解为如前文定义的、包括在核苷酸序列或氨基酸序列中的具有功能性的 SNP。

“功能性”可理解为多肽或多核苷酸的生物活性。

本发明的多肽或多核苷酸的功能性可以是保留、增加、降低或抑制野生型参考基因的核苷酸序列编码的多肽或后者核苷酸序列的生物活性。

本发明的多肽或多核苷酸的功能性同样可以是野生型参考基因的核苷酸序列编码的多肽或后者核苷酸序列的生物活性的本质改变。

生物活性可以特别涉及本发明的多肽与受体有或没有亲和力。

### 多核苷酸

本发明的首要目的在于分离的多核苷酸，其包含：

a) 与序列 SEQ ID NO. 1 或其编码序列(核苷酸 670-1239)具有至少 80 %同一性、优选至少 90 %同一性、更优选至少 95 %同一性以及更加优选至少 99 %同一性的核苷酸序列，应当理解该核苷酸序列包含至少一个以下编码 SNP c794g、c973a、g1011c、t1049a、t1155a、a1204g 或

b) 与 a) 中核苷酸序列互补的核苷酸序列。

应当理解，在本发明意义上，编号对应于该 SNP 在核苷酸序列 SEQ

ID NO. 1 中的定位。

本发明同样涉及分离的多核苷酸，其包含：

a) 核苷酸序列 SEQ ID NO. 1 或其编码序列，应当理解这些序列中每一个均包含至少一个以下编码 SNP：c794g、c973a、g1011c、t1049a、t1155a、a1204g，或

b) 与 a) 中核苷酸序列互补的核苷酸序列。

优选地，本发明的多核苷酸由序列 SEQ ID NO. 1 或其编码序列组成，应当理解这些序列中每一个均包含至少一个以下编码 SNP：c794g、c973a、g1011c、t1049a、t1155a、a1204g。

根据本发明，前面定义的多核苷酸包含选自 c794g、c973a、g1011c、t1049a、t1155a 和 a1204g 的单一编码 SNP。

如上定义的多核苷酸同样可以包括至少一个下列非编码 SNP：t1265c、t1277c。

本发明的目的同样在于分离的多核苷酸，其包含以下或由以下组成：

a) 核苷酸序列 SEQ ID NO. 1 或其编码序列，应当理解这些序列中每一个均包含至少一个以下非编码 SNP：t1265c、t1277c，或

b) 与 a) 中核苷酸序列互补的核苷酸序列。

本发明还涉及分离的多核苷酸，其由部分的：

a) 核苷酸序列 SEQ ID NO. 1 或其编码序列，应当理解这些序列中每一个均包含至少一个下列编码 SNP：c794g、c973a、g1011c、t1049a、t1155a、a1204g、t1265c、t1277c，或

b) 与 a) 中核苷酸序列互补的核苷酸序列组成。

该分离的多核苷酸由至少 10 个核苷酸组成。

优选地，如上定义的分开的多核苷酸由 10-40 个核苷酸组成。

本发明的目的还在于分离的多核苷酸，其编码的多肽包含：

a) 氨基酸序列 SEQ ID NO. 2，或

b) 包含氨基酸序列 SEQ ID NO. 2 中 24-189 位之间所含氨基酸的氨基酸序列，

应当理解 a) 和 b) 中的每个氨基酸序列均包含至少一个下列编码 SNP: A42G、Q102K、Q114H、V127D、C162stop、K179E。

应当理解, 在本发明意义上, 定位 A42G、Q102K、Q114H、V127D、C162stop、K179E SNP 的相应编号是相对于氨基酸序列 SEQ ID NO. 2 的编号而言。

根据本发明的优选目的, 前面定义的多肽包含单个如上定义的编码 SNP。

优选的, 本发明多核苷酸由 DNA 或 RNA 分子组成。

本发明多核苷酸可以利用标准 DNA 或 RNA 合成方法获得。

利用定点诱变从 IFN $\alpha$ -21 基因的核苷酸序列开始, 通过核苷酸序列 SEQ ID NO. 1 中每个 SNP 的突变核苷酸修饰野生型核苷酸, 同样可以获得本发明的多核苷酸。

例如, 利用定点诱变从 IFN $\alpha$ -21 基因的核苷酸序列开始, 将核苷酸序列 SEQ ID NO. 1 中 794 位核苷酸胞嘧啶(c)修饰为核苷酸鸟嘌呤(g), 可以获得包含 SNP c794g 的本发明多核苷酸。

可以这样应用的定点诱变方法为本领域技术人员熟知。特别提到的是文献 TA Kunkel, 1985 "Proc. Natl. Acad. Sci. USA" 82:488。

分离的多核苷酸同样可以包括, 例如编码前 (Pre-)、原 (Pro-) 或前原 (Pre-Pro-) 蛋白质氨基酸序列或诸如六组氨酸肽等标志氨基酸序列的核苷酸序列。

本发明的多核苷酸同样可以连接编码其它蛋白质或蛋白质片段的核苷酸序列, 以获得融合蛋白质或其它纯化产物。

本发明的多核苷酸同样可以包括诸如 5' 和/或 3' 非编码序列的核苷酸序列, 例如转录或非转录序列、翻译或非翻译序列、拼接信号序列、多腺苷酸化序列、核糖体结合序列乃至稳定 mRNA 的序列。

与核苷酸或多核苷酸序列互补的核苷酸序列可定义为可以在严格条件下与该核苷酸序列杂交的核苷酸序列。

"严格杂交条件"一般但不一定理解为可使具有至少 80%、优选大于或等于 90%、更优选大于或等于 95%、最优选大于或等于 97% 的

同一性的核苷酸序列杂交的化学条件。

可以按照本领域技术人员熟知的方法获得严格条件，例如将多核苷酸在包含 50 %甲酰胺、5xSSC ( 150 mM NaCl、15 mM 柠檬酸钠)、50 mM 磷酸钠 (pH = 7.6)、5x Denhardt 液、10 %硫酸葡聚糖和 20  $\mu$ g 变性鲑精 DNA 的溶液中，于 42°C 温育，然后于 0.1x SSC、65°C 洗涤滤膜。

在本发明范围内，如果该严格条件可使具有 100 %同一性的核苷酸序列杂交，则认为该核苷酸序列与如 a) 中所述核苷酸序列严格互补。

在本发明的涵义中可以理解，与一个核苷酸序列互补的核苷酸序列包含至少一个本发明的反义 SNP。

因此，举例来说，如果该核苷酸序列包含 SNP c794g，则其互补核苷酸序列在相当于 794 的位置包含胞嘧啶(c)核苷酸。

#### 包含 SNP 的多核苷酸的鉴定、杂交和/或扩增

本发明的目的还在于以下全部或部分的用途：

a) 与核苷酸序列 SEQ ID NO. 1 具有 80-100%同一性(优选至少 90 %同一性、更优选 95 %同一性、特别优选 100 %同一性)的多核苷酸，和/或

b) 包含至少一个 SNP 的本发明多核苷酸

以鉴定、杂交和/或扩增与核苷酸序列 SEQ ID NO. 1 或必要时其编码序列(核苷酸 670-1239)具有 80-100%同一性(优选至少 90 %同一性、更优选 95 %同一性、特别优选 100 %同一性)的多核苷酸的全部或部分，

应当理解，这些序列中每个均包含至少一个下列 SNP：c794g、c973a、g1011c、t1049a、t1155a、a1204g、t1265c、t1277c。

#### 基因分型并测定 SNP 频率

本发明的目的同样在于以下全部或部分的用途：

a) 与核苷酸序列 SEQ ID NO. 1 具有 80-100%同一性(优选至少 90 %同一性、更优选 95 %同一性、特别优选 100 %同一性)的多核苷酸，和/或

b) 包含至少一个 SNP 的本发明多核苷酸

用于与核苷酸序列 SEQ ID NO. 1 或必要时其编码序列(核苷酸 670-1239)具有 80-100%同一性(优选至少 90 %同一性、更优选 95 %同一性、特别优选 100 %同一性)的多核苷酸的全部或部分的基因分型, 应当理解, 这些序列中每个均包含至少一个下列 SNP: c794g、c973a、g1011c、t1049a、t1155a、a1204g、t1265c、t1277c。

根据本发明, 可以对个体或个体的群体进行基因分型。

在本发明的涵义中, 基因分型可定义为用于测定个体或个体群基因型的方法。基因型由存在于一个或多个特定座位的等位基因构成。

“个体群”可理解为以随机或非随机方式选择的一群个体。这些个体可以是人类、动物、微生物或植物。

个体群一般包含至少 10 个成员, 优选 100-300 个成员。

可以根据种族或表现型选择个体, 特别是受以下疾病和/或疾病影响的个体: 恶性肿瘤、黑色素瘤、淋巴瘤、白血病和肝癌、颈癌、头癌和肾癌、心血管疾病、诸如与免疫系统无关的代谢疾病, 例如肥胖症, 传染病, 特别是病毒性感染, 如乙型肝炎、丙型肝炎和 AIDS, 肺炎、溃疡性结肠炎、中枢神经系统疾病, 例如阿尔茨海默氏病、精神分裂症和抑郁症, 组织或器官移植排斥、创伤愈合、透析患者的贫血、变态反应、哮喘、多发性硬化、骨质疏松症、银屑病、类风湿性关节炎、克罗恩氏病、自身免疫疾病和异常、胃肠疾病乃至与化学治疗有关的疾病。

优选对个体群进行本发明功能性 SNP 的基因分型。

现有许多可以用来进行 SNP 基因分型的技术(特别参见 Kwok Pharmacogenomics, 2000, 卷 1, 95-100 页, “High-throughput genotyping assay approaches”)。这些技术基于以下四原理之一: 等位基因特异性寡核苷酸杂交、任选在脱氧核苷酸存在下通过双脱氧核苷酸的寡核苷酸延伸、等位基因特异性寡核苷酸的连接或等位基因特异性寡核苷酸的断裂。其中每个技术都可与检测系统相连, 例如直接或偏振荧光测定、或质谱分析。

利用热 ddNTP (由不同荧光团标记的两种不同的 ddNTP)和冷 ddNTP(两种不同的非标记的 ddNTP)的微型测序,结合偏振荧光扫描仪,可以特别进行基因分型。本领域技术人员熟知利用读取偏振荧光的微型测序方法(FP-TDI 技术或荧光偏振模板-直接染料-终止物参入)。

它可以对聚合酶链式反应(PCR)扩增每一个体 DNA 后获得的产物进行。选择包括含所研究 SNP 的多核苷酸基因区域的 PCR 产物。在 PCR 热循环仪的最后步骤之后,将平板置于偏振荧光扫描仪,利用荧光团特异性激发和发射滤器读取标记碱基。以图表报告标记碱基的强度值。

就本发明 SNP 的 PCR 扩增而言,本领域技术人员很容易根据本发明 SNP 的位置,分别选择有义和反义引物。

例如,用于 PCR 扩增引物的有义和反义核苷酸序列可以分别为:

SEQ ID NO. 3: 有义引物: GGTTC AAGGTTACCCATCTC

SEQ ID NO. 4: 反义引物: TTTGAAATGGCAGAAGTCAT

该核苷酸序列可以扩增具有核苷酸序列 SEQ ID NO. 1 中 620-1315 位核苷酸的 696 个核苷酸长度的片段。

然后对个体群中由包含 SNP 的基因编码的每个等位基因的频率(等位频率)进行统计分析,可确定其影响的重要性及其在不同亚群、特别是必要时在构成该个体群的不同种群中的分布。

分析基因型数据,估计在所研究群体中观察到的不同等位基因的分布频率。可借助于软件,例如 SAS-suite® (SAS) 或 SPLUS® (MathSoft),计算等位频率。利用软件 ARLEQUIN®和 SAS-suite®,可以比较本发明 SNP 在个体群中不同种群的等位分布。

#### 本发明 SNP 作为遗传标志

鉴于 SNP 改变的基因功能性序列(例如启动子、拼接位点、编码区)很可能与疾病易感性或抗性直接有关,所有 SNP (不论功能性与否)均可能提供用于鉴定与疾病状态有关的一个或几个基因的重要标志,从而可能与该疾病状态间接相关(参见 Cargill 等人(1999). Nature Genetics 22:231-238; Riley 等人(2000). Pharmacogenomics

1:39-47; Roberts L. (2000). Science 287: 1898-1899).

因此, 本发明还涉及包含在 IFN $\alpha$ -21 基因的多核苷酸的至少一个下列 SNP 的数据库: c794g、c973a、g1011c、t1049a、t1155a、a1204g、t1265c、t1277c。

应当理解, 所述 SNP 根据核苷酸序列 SEQ ID NO. 1 编号。

可分析该数据库, 确定以下两者之间的统计学相关性:

- (i) IFN $\alpha$ -21 基因多核苷酸中的至少一个下列 SNP: c794g、c973a、g1011c、t1049a、t1155a、a1204g、t1265c、t1277c, 以及
- (ii) 疾病或疾病抗性。

本发明还涉及 IFN $\alpha$ -21 基因多核苷酸中的至少一个下列 SNP 的用途: c794g、c973a、g1011c、t1049a、t1155a、a1204g、t1265c、t1277c, 用于开发疾病或疾病抗性的诊断/预后试剂盒。

如上定义的本发明 SNP 可直接或间接与疾病或疾病抗性有关。

优选地, 这些疾病可能是上文定义的疾病。

#### 表达载体和宿主细胞

本发明目的还在于包含至少一种本发明多核苷酸的重组载体。

可以使用多种表达系统, 包括而不限于染色体、附加体和衍生病毒。更特别而言, 所用重组载体可来自于细菌质粒、转座子、酵母附加体、插入元件、酵母染色体元件、诸如杆状病毒等病毒、诸如 SV40 等乳头状瘤病毒、痘苗病毒、腺病毒、福克斯痘病毒(fox pox viruses)、假狂犬病毒、逆转录病毒。

这些重组载体同样可以是粘粒或噬菌粒衍生物。可以利用本领域技术人员熟知的方法, 例如 MOLECULAR CLONING: A LABORATORY MANUAL, Sambrook 等人, 第四版, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2001) 中所述方法, 将核苷酸序列插入重组表达载体。

该重组载体可以包括控制多核苷酸表达调节的核苷酸序列、以及可表达和转录本发明多核苷酸、翻译本发明多肽的核苷酸序列, 根据所用宿主细胞选择这些序列。

因此，例如可以将合适的分泌信号引入重组载体，以便将本发明多核苷酸编码的多肽导向内质网腔、周质间隙、膜或细胞外环境。

本发明目的还在于包含本发明重组载体的宿主细胞。

可按照本领域技术人员公知的方法，例如 BASIC METHODS IN MOLECULAR BIOLOGY, Davis 等人，第二版，McGraw - Hill Professional Publishing, 1995 和 MOLECULAR CLONING : A LABORATORY MANUAL(同上)中所述方法，例如磷酸钙转染、DEAE 葡聚糖转染、转染、显微注射、阳离子脂质转染、电穿孔、转导或感染，将重组载体引入宿主细胞。

例如，宿主细胞可以是细菌细胞诸如链球菌、葡萄球菌、大肠杆菌(E. coli)或枯草杆菌(Bacillus subtilis)细胞，真菌细胞诸如酵母细胞和曲霉菌属、链霉菌属细胞，昆虫细胞诸如果蝇 S2 和 Spodoptera Sf9 细胞，动物细胞诸如 CHO、COS、HeLa、C127、BHK、HEK 293 细胞，以及所治疗受试者的人类细胞乃至植物细胞。

例如，可用来表达本发明多肽或作为药物组合物中活性产物的宿主细胞参见下文。

### 多肽

本发明目的还在于分离的多肽，其包含与以下序列具有至少 80 % 同一性、优选至少 90 % 同一性、更优选至少 95 % 同一性、更加优选至少 99 % 同一性的氨基酸序列：

a) 氨基酸序列 SEQ ID NO. 2, 或

b) 包含氨基酸序列 SEQ ID NO. 2 中 24-189 位之间所含氨基酸的氨基酸序列，

应当理解，a) 和 b) 中的每个氨基酸序列均包含至少一个下列编码 SNP: A42G、Q102K、Q114H、V127D、C162stop、K179E。

本发明的多肽同样可以包含：

a) 氨基酸序列 SEQ ID NO. 2, 或

b) 包含氨基酸序列 SEQ ID NO. 2 中 24-189 位之间所含氨基酸的氨基酸序列，

应当理解，a)和 b)中的每个氨基酸序列均包含至少一个下列编码 SNP: A42G、Q102K、Q114H、V127D、C162stop、K179E。

本发明的多肽可以更特别由以下序列组成：

- a) 氨基酸序列 SEQ ID NO. 2, 或
- b) 包含氨基酸序列 SEQ ID NO. 2 中 24-189 位之间所含氨基酸的氨基酸序列，

应当理解，a)和 b)中的每个氨基酸序列均包含至少一个下列编码 SNP: A42G、Q102K、Q114H、V127D、C162stop、K179E。

优选地，本发明的多肽包含选自 A42G、Q102K、Q114H、V127D、C162stop 和 K179E 的单个编码 SNP。

本发明的目的同样在于制备上述多肽的方法，其中在培养基中培养前面定义的宿主细胞，并从培养基中分离该多肽。

按照本领域技术人员熟知的方法，例如利用诸如盐、特别是硫酸铵、乙醇、丙酮或三氯乙酸等离液剂沉淀；酸提取；离子交换层析；磷酸纤维素层析；疏水作用层析；亲和层析；羟基磷灰石层析或排阻层析，可以从宿主细胞的培养基中纯化所述多肽。

“培养基”可理解为从中分离或纯化本发明多肽的培养基。该培养基可以由细胞外培养基和/或细胞裂解物组成。如果该多肽的构象在分离或纯化期间改变，同样可以利用本领域技术人员熟知的方法恢复该多肽的活性构象。

### 抗体

本发明的目的还在于获得免疫特异性抗体的方法。

“抗体”可理解为包括单克隆、多克隆、嵌合、单链、人源化抗体以及 Fab 片段，包括 Fab 或免疫球蛋白表达文库产物。

利用本发明的多肽免疫动物，可以获得免疫特异性抗体。

本发明还涉及前面定义的本发明多肽的免疫特异性抗体。

本发明的多肽、其片段之一、类似物、其变体之一或表达该多肽的细胞，也可以用来产生免疫特异性抗体。

术语“免疫特异性”是指该抗体对于本发明的多肽比对于现有技术

已知的其它多肽具有更好的亲合力。

按照本领域技术人员熟知的方法，将本发明的多肽、其片段之一、类似物或表位片段、或表达该多核苷酸的细胞给予哺乳动物、优选非人哺乳动物，可以获得免疫特异性抗体。

为制备单克隆抗体，可以从细胞系开始，使用抗体制备的典型方法，例如杂交瘤技术(Kohler 等人, Nature (1975) 256: 495-497)，三源杂交瘤(trioma)技术，人 B 细胞杂交瘤技术(Kozbor 等人, Immunology Today (1983) 4: 72) 以及 EBV 杂交瘤技术(Cole 等人, "The EBV-hybridoma technique and its application to human lung cancer," Monoclonal Antibodies and Cancer Therapy (27 卷, UCLA Symposia on Molecular and Cellular Biology, New Series) (R. A. Reisfeld 和 S. Sell 编著), 77-96 页, Alan R. Liss, Inc. N.Y., 1985, pp. 77-96)。

同样可以使用单链抗体制备技术，例如美国专利 4,946,778 所述。诸如小鼠等转基因动物同样可以用来制备人源化抗体。

#### 与本发明多肽相互作用的试剂

本发明的目的同样在于鉴定活化或抑制本发明多肽的试剂的方法，其包含：

- a) 制备包含本发明的含有至少一个 SNP 的多核苷酸的重组载体，
- b) 制备包含 a) 中重组载体的宿主细胞，
- c) 将 b) 中宿主细胞接触待测试剂，以及
- d) 确定待测试剂产生的活化或抑制作用。

本发明的多肽还可以用于筛选与其相互作用的化合物的方法。

这些化合物可以是本发明多肽固有活性的活化剂(激动剂)或抑制剂(拮抗剂)。这些化合物同样可以是本发明多肽的配基或底物。参见 Coligan 等人, Current Protocols in Immunology 1 (2), Chapter 5 (1991)。

一般而言，为实施上述方法，最理想的是制备表达本发明多肽的合适的宿主细胞。上述细胞可以是，例如哺乳动物、酵母、昆虫(例

如果蝇)或细菌(例如大肠杆菌)的细胞。

然后将这些细胞或这些细胞的膜提取物置于待测化合物的存在下。

然后观察该待测化合物与本发明多肽的结合力以及对功能性反应的抑制或活化作用。

使用直接或间接标记的待测试剂, 可以进行上述方法的步骤 d)。还可以包括一个使用标记或非标记的试剂以及标记的竞争剂的竞争试验。

同样可以利用根据待检信号而适当选择的检测方法, 确定待检试剂是否对表达本发明多肽的细胞产生活化或抑制信号。

上述活化或抑制剂可以是多核苷酸、在某些情况下是寡核苷酸或多肽, 例如蛋白质或抗体。

本发明的目的还在于鉴定受本发明多肽活化或抑制的试剂的方法, 其包含:

- a) 制备包含本发明的含有至少一个 SNP 的多核苷酸的重组载体,
- b) 制备包含 a) 中重组载体的宿主细胞,
- c) 将 b) 中宿主细胞置于待测试剂的存在下, 以及
- d) 确定该多肽对待测试剂产生的活化或抑制作用。

受本发明多肽活化或抑制的试剂是在该多肽存在下, 分别对活化或抑制反应的试剂。受本发明多肽直接或间接活化或抑制的试剂可以由多肽例如膜或核受体、激酶并更优选酪氨酸激酶、转录因子或多核苷酸组成。

#### 疾病检测

本发明目的还在于分析受试者中本发明多核苷酸和/或本发明多肽的生物学特性的方法, 其包含以下至少一项:

- a) 确定受试者基因组中存在或缺少本发明的多核苷酸;
- b) 确定受试者中本发明的多核苷酸的表达水平;
- c) 确定在受试者中存在或缺少本发明的多肽;
- d) 确定在受试者中本发明多肽的浓度; 和/或

e) 确定受试者中本发明多肽的功能性。

可分析受试者或受试者样品中这些生物学特性。

这些生物学特性可进行遗传诊断和确定受试者是否患有与存在本发明多核苷酸和/或本发明多肽有关的疾病、不适或病症，或具有患有疾病、不适或病症的危险，或反之具有对疾病、不适或病症形成的部分抗性。

这些疾病可以是病症和/或人类疾病，例如癌症和肿瘤、传染病、性病、免疫相关疾病和/或自身免疫疾病和病症、心血管疾病、代谢疾病、中枢神经系统疾病、以及与化学治疗有关的疾病。

所述癌症和肿瘤包括，包含转移的肾癌、黑色素瘤的恶性肿瘤、包含滤泡性淋巴瘤和皮肤 T 细胞淋巴瘤的淋巴瘤、包含毛样细胞白血病、慢性淋巴细胞性白血病和慢性骨髓性白血病的白血病、肝癌、颈癌、头癌和肾癌、多发性骨髓瘤、类癌瘤以及包含 AIDS 情况下的 Kaposi 肉瘤在内的伴随免疫缺陷出现的肿瘤。

所述传染病包括病毒性传染，包含慢性乙型和丙型肝炎以及 HIV/AIDS，感染性肺炎以及性病，例如生殖器疣。

所述免疫和自身免疫相关疾病可能包括组织或器官移植排斥、变态反应、哮喘、银屑病、类风湿性关节炎、多发性硬化、克罗恩氏病以及溃疡性结肠炎。

所述代谢疾病可能包括上述非免疫相关性疾病，例如肥胖症。

所述中枢神经系统疾病可能包括阿尔茨海默氏病、帕金森氏病、精神分裂症和抑郁症。

所述疾病和病症还可能包括创伤愈合、透析患者的贫血以及骨质疏松症。

该方法还可遗传诊断受试者中与存在本发明 SNP 编码的突变等位基因有关的疾病或疾病抗性。

优选在步骤 a) 中检测包含前面定义的至少一个编码 SNP 的多核苷酸的存在与否。

可以从所研究受试者的生物样品(例如细胞、血液、尿、唾液)、

或其活检或尸检样品开始，检测该多核苷酸。例如，可使用基因组 DNA 直接或于 PCR 扩增后进行检测。同样可按类似方式使用 RNA 或 cDNA。

然后可将本发明多核苷酸的核苷酸序列与受试者基因组中检测到的核苷酸序列作比较。

可以通过测序、DNA 杂交方法、DNA 片段在有或无变性剂的电泳凝胶中的迁移率差异、或解链温度差异，进行核苷酸序列的比较。参见 Myers 等人，*Science* (1985) 230: 1242。利用核酸酶保护试验(例如 RNase 和 S1 核酸酶)或化学裂解剂，同样可以揭示核苷酸序列精确位点的上述结构改变。参见 Cotton 等人，*Proc. Nat. Acad. Sci. USA* (1985) 85: 4397-4401。同样可以利用包含本发明多核苷酸片段的寡核苷酸探针进行筛选。

许多为本领域技术人员所熟知的方法可用于确定本发明多核苷酸的表达，并鉴定这些多核苷酸的遗传变异性(参见 Chee 等人，*Science* (1996), Vol 274, pp 610-613)。

在步骤 b) 中，按照本领域技术人员熟知的方法，例如 PCR、RT-PCR、RNase 保护、Northern 印迹及其它杂交方法，定量该多核苷酸编码的 RNA(及编码多肽)的水平，可测定该多核苷酸的表达水平。

在步骤 c) 和 d) 中，利用公知的方法，例如放射免疫测定、竞争性结合试验、Western 印迹和 ELISA 试验，可以测定本发明多肽在受试者或来自受试者样品中的存在与否及其浓度。

步骤 d) 之后，测定的本发明多肽的浓度可与通常在受试者中发现的天然野生型蛋白质的浓度相比。

本领域技术人员借助于现有技术文献或诸如前述常规试验或测定，可以确定对以上所致疾病、不适或病症表现出敏感性或反之抗性的高低阈值。

在步骤 e) 中，利用本领域技术人员熟知的方法，例如上述体外试验，或使用表达本发明多肽的宿主细胞，可以测定本发明多肽的功能性。

#### 治疗性化合物及疾病治疗

本发明的目的还在于包含本发明多肽作为活性剂的治疗性化合物。

本发明还涉及本发明多肽的用途，用于制备意在预防或治疗各种人类病症和/或疾病的治疗性化合物。这些疾病可以是病症和/或人类疾病，例如癌症和肿瘤、传染病、性病、免疫相关疾病和/或自身免疫疾病和异常、心血管疾病、代谢疾病、中枢神经系统疾病、以及与化学治疗有关的病症。

所述癌症和肿瘤包括，包含转移的肾癌、黑色素瘤的恶性肿瘤、包含滤泡性淋巴瘤和皮肤 T 细胞淋巴瘤的淋巴瘤、包含毛样细胞白血病、慢性淋巴细胞性白血病和慢性骨髓性白血病的白血病、肝癌、颈癌、头癌和肾癌、多发性骨髓瘤、类癌瘤以及包含 AIDS 情况下的 Kaposi 肉瘤的伴随免疫缺陷出现的肿瘤。

所述传染病包括病毒性传染，包含慢性乙型和丙型肝炎以及 HIV/AIDS，感染性肺炎以及性病，例如生殖器疣。

所述免疫和自身免疫相关疾病可能包括组织或器官移植排斥、变态反应、哮喘、银屑病、类风湿性关节炎、多发性硬化、克罗恩氏病以及溃疡性结肠炎。

所述代谢疾病可能包括上述非免疫相关性疾病，例如肥胖症。

所述中枢神经系统疾病可能包括阿尔茨海默氏病、帕金森氏病、精神分裂症和抑郁症。

所述疾病和病症还可能包括创伤愈合、透析患者的贫血以及骨质疏松症。

优选地，本发明的多肽还可以用于制备意在预防或治疗各种人类病症和/或疾病的治疗性化合物，例如某些病毒性感染，如乙型和丙型肝炎，白血病例如毛样细胞白血病和慢性骨髓性白血病，多发性骨髓瘤、滤泡性淋巴瘤、类癌瘤、恶性黑色素瘤、转移性肾癌、阿尔茨海默氏病、帕金森氏病以及伴随免疫缺陷出现的肿瘤，例如在 AIDS 情况下的 Kaposi 肉瘤、以及生殖器疣或性病。

可以获得本发明多肽的某些化合物以及通过或由该多肽获得或鉴

定的化合物同样可以作为治疗性化合物，用于治疗人体。

正因为如此，本发明的目的还在于包含含有至少一个前面定义的编码 SNP 的本发明多核苷酸、前面定义的重组载体、前面定义的宿主细胞、和/或前面定义的抗体作为活性剂的药物。

本发明还涉及包含至少一个前面定义的编码 SNP 的本发明多核苷酸、前面定义的重组载体、前面定义的宿主细胞、和/或前面定义的抗体的用途，用于制备意在预防或治疗各种人类病症和/或疾病的药物。这些疾病可以是病症和/或人类疾病，例如癌症和肿瘤、传染病、性病、免疫相关疾病和/或自身免疫疾病和病症、心血管疾病、代谢疾病、中枢神经系统疾病、以及与化学治疗有关的疾病。

所述癌症和肿瘤包括，包含转移的肾癌、黑色素瘤的恶性肿瘤、包含滤泡性淋巴瘤和皮肤 T 细胞淋巴瘤的淋巴瘤、包含毛样细胞白血病、慢性淋巴细胞性白血病和慢性骨髓性白血病在内的白血病、肝癌、颈癌、头癌和肾癌、多发性骨髓瘤、类癌瘤以及包含 AIDS 情况下的 Kaposi 肉瘤的伴随免疫缺陷出现的肿瘤。

所述传染病包括病毒性传染，包含慢性乙型和丙型肝炎以及 HIV/AIDS，感染性肺炎以及性病，例如生殖器疣。

所述免疫和自身免疫相关疾病可能包括组织或器官移植排斥、变态反应、哮喘、银屑病、类风湿性关节炎、多发性硬化、克罗恩氏病以及溃疡性结肠炎。

所述代谢疾病可能包括上述非免疫相关性疾病，例如肥胖症。

所述中枢神经系统疾病可能包括阿尔茨海默氏病、帕金森氏病、精神分裂症和抑郁症。

所述疾病和病症还可能包括创伤愈合、透析患者的贫血以及骨质疏松症。

优选地，本发明涉及包含至少一个前面定义的 SNP 的本发明多核苷酸、前面定义的重组载体、前面定义的宿主细胞、和/或前面定义的抗体的用途，用于制备意在预防或治疗各种人类病症和/或疾病的药物，例如某些病毒性感染，如乙型和丙型慢性肝炎，白血病例如毛样

细胞白血病和慢性骨髓性白血病，多发性骨髓瘤、滤泡性淋巴瘤、类癌瘤、恶性黑色素瘤、转移性肾癌、阿尔茨海默氏病、帕金森氏病以及伴随免疫缺陷出现的肿瘤，例如在 AIDS 情况下的 Kaposi 肉瘤，以及生殖器疣或性病。

本发明用作活性剂的多肽和其它化合物的剂量取决于所选择的化合物、治疗适应症、给药方式、制剂性质、受试者性质及医生的判断。

用作活性剂时，一般按 1-100  $\mu\text{g}/\text{kg}$  受试者的剂量给予本发明的多肽。

本发明的目的还在于药物组合物，其包含至少一种上述化合物，例如本发明的多肽、包含至少一个前面定义的 SNP 的本发明多核苷酸、前面定义的重组载体、前面定义的宿主细胞、和/或前面定义的抗体作为活性剂，以及药物可接受的赋形剂。

该活性剂有利地以生理有效剂量存在于这些药物组合物中。

例如，这些药物组合物可以是固体或液体，并以目前用于人类医学的药物形式存在，例如简单或包衣片剂、软胶囊 (gelcap)、颗粒剂、焦糖剂、栓剂、优选可注射制剂和可注射用粉末。可以按照常用方法制备这些药物形式。

活性剂可以掺入通常用于药物组合物的赋形剂，例如滑石粉、阿拉伯胶、乳糖、淀粉、葡萄糖、甘油、乙醇、硬脂酸镁、可可脂、含水或无水载体、动物或植物来源的脂质、石蜡衍生物、乙二醇、各种湿润剂、分散剂或乳化剂、防腐剂。

本发明的活性剂可以单独使用或与其它化合物联用，例如治疗性化合物，诸如白细胞介素或干扰素等其它细胞因子。

根据给药方式，调整该药物组合物的不同剂型。

可以通过本领域技术人员已知的不同给药途径，给予该药物组合物。

本发明的目的同样在于诊断性组合物，其包含至少一种上述化合物（例如本发明的多肽、本发明多核苷酸的全部或部分、前面定义的重组载体、前面定义的宿主细胞、和/或前面定义的抗体）作为活性剂，

以及合适的药物可接受的赋形剂。

该诊断性组合物可包含，例如一般用于诊断性组合物的合适的赋形剂，诸如缓冲液和防腐剂。

本发明的目的同样在于下列的用途：

- a) 治疗有效量的本发明多肽，和/或
- b) 本发明的多核苷酸，和/或
- c) 前面定义的来自所需治疗受试者的宿主细胞，

以制备意在增加本发明多肽在受试者中表达或活性的治疗性化合物。

因此，为治疗需要增加本发明多肽表达或活性的受试者，可能有几种方法。

可给予受试者治疗有效量的本发明多肽，连同药物可接受的赋形剂。

还可通过给予受试者本发明的多核苷酸，增加本发明多肽的内源性生成。例如，可以将该多核苷酸插入逆转录病毒表达载体。利用包含编码本发明多肽的 RNA 的逆转录病毒质粒载体感染细胞，以使转导细胞产生包含目的基因的感染性病毒颗粒，可以从该细胞开始分离上述载体。参见 Human Molecular Genetics, Strachan 和 Read, 第 20 章, Gene Therapy and other Molecular Genetic-based Therapeutic Approaches, BIOS Scientific Publishers Ltd (1996)。

根据本发明，优选使用如上定义的包含至少一个编码 SNP 的多核苷酸。

同样可给予受试者属于自身的宿主细胞，预先取出这些宿主细胞，并如前所述进行修饰，使其表达本发明多肽。

本发明同样涉及：

- a) 治疗有效量的前面定义的免疫特异性抗体，和/或
- b) 可抑制本发明多核苷酸表达的多核苷酸的用途，

以便制备意在降低本发明多肽在受试者中表达或活性的治疗性化合物。

因此，可给予受试者治疗有效量的诸如前面定义的抑制剂和/或抗体，可能联用，连同药物可接受的赋形剂。

通过给予受试者可抑制本发明多核苷酸表达的本发明的互补多核苷酸，同样有可能降低本发明多肽的内源性生成。

优选使用如上定义的包含至少一个编码 SNP 的互补多核苷酸。

本发明还涉及 IFN $\alpha$ -21 蛋白质在制备药物中的用途，用于预防或治疗具有 IFN $\alpha$ -21 变体所致病症或疾病的患者，该变体与该患者基因组存在与核苷酸序列 SEQ ID NO. 1 具有至少 95%同一性(优选 97%同一性、更优选 99%同一性、特别优选 100%同一性)的核苷酸序列有关，只要该核苷酸序列包含下列 SNP 之一：c794g、c973a、g1011c、t1049a、t1155a、a1204g、t1265c、t1277c。

优选地，该药物用于预防或治疗选自下列之一的疾病：癌症和肿瘤、传染病、性病、免疫相关疾病和/或自身免疫疾病和病症、心血管疾病、代谢疾病、中枢神经系统疾病以及与化学治疗有关的疾病。

所述癌症和肿瘤包括，包含转移的肾癌、黑色素瘤的恶性肿瘤、包含滤泡性淋巴瘤和皮肤 T 细胞淋巴瘤的淋巴瘤、包含毛样细胞白血病、慢性淋巴细胞性白血病和慢性骨髓性白血病的白血病、肝癌、颈癌、头癌和肾癌、多发性骨髓瘤、类癌瘤以及包含 AIDS 情况下的 Kaposi 肉瘤的伴随免疫缺陷出现的肿瘤。

所述传染病包括病毒性传染，包含慢性乙型和丙型肝炎以及 HIV/AIDS，感染性肺炎以及性病，例如生殖器疣。

所述免疫和自身免疫相关疾病可能包括组织或器官移植排斥、变态反应、哮喘、银屑病、类风湿性关节炎、多发性硬化、克罗恩氏病以及溃疡性结肠炎。

所述代谢疾病可能包括上述非免疫相关性疾病，例如肥胖症。

所述中枢神经系统疾病可能包括阿尔茨海默氏病、帕金森氏病、精神分裂症和抑郁症。

所述疾病和病症还可能包括创伤愈合、透析患者的贫血以及骨质疏松症。

包含本发明 SNP 的 IFN $\alpha$ -21 多肽的类似 (mimetic) 化合物

本发明还涉及具有与下列多肽基本类似的生物活性的新化合物:

a) 氨基酸序列 SEQ ID NO. 2, 或

b) 包含氨基酸序列 SEQ ID NO. 2 中 24-189 位之间所含氨基酸的氨基酸序列;

只要 a) 和 b) 中所述氨基酸序列包含 SNP K179E。

例如, 如实验部分所述, 通过测定刺激树突状细胞成熟的能力、CD4<sup>+</sup>或 CD8<sup>+</sup> T 淋巴细胞释放细胞因子、单核细胞释放细胞因子、体外或体内抗病毒活性、对 Daudi Burkitt 细胞系的细胞抗增殖活性、对 TF-1 细胞系的细胞抗增殖活性, 可评价所述生物活性。

如实验部分所述, K179E 突变型 IFN $\alpha$ -21 在下列方面不同于野生型 IFN $\alpha$ -2:

\* K179E 突变型 IFN $\alpha$ -21 具有较强的刺激树突状细胞成熟的能力;

\* K179E 突变型 IFN $\alpha$ -21 具有较强的刺激 CD4<sup>+</sup>或 CD8<sup>+</sup> T 淋巴细胞释放 IFN- $\gamma$  的能力;

\* K179E 突变型 IFN $\alpha$ -21 在感染 VSV 的细胞培养物中具有低于野生型 IFN $\alpha$ -2 的抗病毒活性。

如实验部分所述, K179E 突变型 IFN $\alpha$ -21 对 Daudi Burkitt 细胞系具有略微低于天然野生型 IFN $\alpha$ -21 的细胞抗增殖活性。

同样如实验部分所述, K179E 突变型 IFN $\alpha$ -21 对 TF-1 细胞系具有类似于野生型 IFN $\alpha$ -2 的细胞抗增殖活性, 在 EMCV 小鼠模型中具有类似于野生型 IFN $\alpha$ -2 的抗病毒活性。

如上定义的本发明的新化合物可能具有基本上类似于 K179E 突变型 IFN $\alpha$ -21 的生物活性。

根据所研究的生物活性类型, 所述化合物可能还具有低于或高于 K179E 突变型 IFN $\alpha$ -21 的生物活性。

所述化合物可能是生物化学化合物, 例如多肽或肽, 或有机化学化合物, 例如合成肽类似物。

本发明还涉及包含 K179E SNP 的本发明多肽在鉴定如上定义化的

合物中的用途。

本发明还涉及鉴定本发明化合物的方法，其包含以下步骤：

a) 确定待测化合物的生物活性，例如树突状细胞成熟、CD4+或CD8+ T淋巴细胞释放细胞因子、单核细胞释放细胞因子、体外或体内抗病毒活性、对 Daudi Burkitt 细胞系的细胞抗增殖活性；

b) 比较：

i) 步骤 a)测定的待测化合物的活性，与

ii) 氨基酸序列 SEQ ID NO. 2 的多肽的活性，或包含氨基酸序列 SEQ ID NO. 2 中 24-189 位之间所含氨基酸的氨基酸序列的活性；只要该氨基酸序列包含 K179E SNP；以及

c) 根据步骤 b)中进行的比较，确定该待测化合物是否具有比氨基酸序列 SEQ ID NO. 2 的多肽、或与包含氨基酸序列 SEQ ID NO. 2 中 24-189 位所含氨基酸的氨基酸序列基本类似、或低、或高的活性；只要该氨基酸序列包含 K179E SNP。

优选该待测化合物可能预先鉴定自合成肽组合文库、高通量筛选、或由计算机辅助药物设计而设计，以具有与氨基酸序列 SEQ ID NO. 2 的多肽、或包含氨基酸序列 SEQ ID NO. 2 中 24-189 位所含氨基酸的氨基酸序列相同的三维结构；只要该氨基酸序列包含 K179E SNP。

本发明还涉及具有与下列多肽基本类似的生物活性的新化合物：

a) 氨基酸序列 SEQ ID NO. 2，或

b) 包含氨基酸序列 SEQ ID NO. 2 中 24-189 位之间所含氨基酸的氨基酸序列；

只要 a)和 b)中所述氨基酸序列包含 SNP Q114H 和 V127D。

例如，如实验部分所述，通过测定刺激树突状细胞成熟的能力、CD4+或 CD8+ T淋巴细胞释放细胞因子、单核细胞释放细胞因子、体外或体内抗病毒活性、对 Daudi Burkitt 细胞系的细胞抗增殖活性、对 TF-1 细胞系的细胞抗增殖活性，可评价所述生物活性。

如实验部分所述，包含 SNP Q114H 和 V127D 两者的 IFN $\alpha$ -21(所述双突变下文记作“Q114H/V127D”)在下列方面不同于野生型 IFN $\alpha$ -2:

\* Q114H/V127D 突变型 IFN $\alpha$ -21 具有较高的刺激 CD4<sup>+</sup>或 CD8<sup>+</sup> T 淋巴细胞释放 IFN- $\gamma$  的能力;

\* Q114H/V127D 突变型 IFN $\alpha$ -21 在感染 VSV 的细胞培养物中具有较低的抗病毒活性;

如实验部分所述, Q114H/V127D 突变型 IFN $\alpha$ -21 对 Daudi Burkitt 细胞系具有低于天然野生型 IFN $\alpha$ -21 的细胞抗增殖活性。

同样如实验部分所述, Q114H/V127D 突变型 IFN $\alpha$ -21 在 EMCV 小鼠模型中具有类似于野生型 IFN $\alpha$ -2 的抗病毒活性、类似于野生型 IFN $\alpha$ -2 的刺激单核细胞释放 IL-10、IL-12 和 TNF- $\alpha$  的能力、以及类似于野生型 IFN $\alpha$ -2 的对 TF-1 细胞系的细胞抗增殖活性。

如上定义的本发明新化合物可能具有基本类似于 Q114H/V127D 突变型 IFN $\alpha$ -21 的生物活性。

根据所研究的生物活性类型, 所述化合物可能还具有低于或高于 Q114H/V127D 突变型 IFN $\alpha$ -21 的生物活性。

所述化合物可能是生物化学化合物, 例如多肽或肽, 或有机化学化合物, 例如合成肽类似物。

本发明还涉及包含 Q114H/V127D SNP 的本发明多肽在鉴定如上定义的化合物中的用途。

本发明还涉及鉴定本发明化合物的方法, 其包含以下步骤:

a) 确定生物活性, 例如树突状细胞成熟、CD4<sup>+</sup>或 CD8<sup>+</sup> T 淋巴细胞释放细胞因子、单核细胞释放细胞因子、体外或体内抗病毒活性、对 Daudi Burkitt 细胞系的细胞抗增殖活性;

b) 比较:

i) 步骤 a)测定的待测化合物的活性, 与

ii) 氨基酸序列 SEQ ID NO. 2 的多肽的活性, 或包含氨基酸序列 SEQ ID NO. 2 中 24-189 位之间所含氨基酸的氨基酸序列的活性;

只要该氨基酸序列包含 Q114H/V127D SNP; 以及

c) 根据步骤 b)中进行的比较, 确定该待测化合物是否具有与氨基酸序列 SEQ ID NO. 2 的多肽、或包含氨基酸序列 SEQ ID NO. 2 中

24-189 位所含氨基酸的氨基酸序列基本类似、或更低、或更高的活性；只要该氨基酸序列包含 Q114H/V127D SNP。

优选该待测化合物可能预先鉴定自合成肽组合文库、高通量筛选、或由计算机辅助药物设计而设计，以具有与氨基酸序列 SEQ ID NO. 2 的多肽、或包含氨基酸序列 SEQ ID NO. 2 中 24-189 位所含氨基酸的氨基酸序列相同的三维结构；只要该氨基酸序列包含 Q114H/V127D SNP。

鉴定和设计化合物的方法为本领域技术人员所熟知。

例如，关于这些方法的文献可能有：

- Silverman R.B. (1992). "Organic Chemistry of Drug Design and Drug Action". Academic Press, 1st edition (January 15, 1992).

- Anderson S 和 Chiplin J. (2002). "Structural genomics; shaping the future of drug design" Drug Discov. Today. 7(2):105-107.

- Selick HE, Beresford AP, Tarbit MH. (2002). "The emerging importance of predictive ADME simulation in drug discovery". Drug Discov. Today. 7(2):109-116.

- Burbidge R, Trotter M, Buxton B, Holden S. (2001). "Drug design by machine learning: support vector machines for pharmaceutical data analysis". Comput. Chem. 26(1): 5-14.

- Kauvar L.M. (1996). "Peptide mimetic drugs: a comment on progress and prospects" 14(6): 709.

本发明的药物可用于制备意在预防或治疗选自下列疾病之一的药物：癌症和肿瘤、传染病、性病、免疫相关疾病和/或自身免疫疾病和病症、心血管疾病、代谢疾病、中枢神经系统疾病以及与化学治疗有关的疾病。

所述癌症和肿瘤包括，包含转移的肾癌、黑色素瘤的恶性肿瘤、包含滤泡性淋巴瘤和皮肤 T 细胞淋巴瘤的淋巴瘤、包含毛样细胞白血

病、慢性淋巴细胞性白血病和慢性骨髓性白血病的白血病、肝癌、颈癌、头癌和肾癌、多发性骨髓瘤、类癌瘤以及包含 AIDS 情况下的 Kaposi 肉瘤在内的伴随免疫缺陷出现的肿瘤。

所述传染病包括病毒性传染，包含慢性乙型和丙型肝炎以及 HIV/AIDS，感染性肺炎以及性病，例如生殖器疣。

所述免疫和自身免疫相关疾病可能包括组织或器官移植排斥、变态反应、哮喘、银屑病、类风湿性关节炎、多发性硬化、克罗恩氏病以及溃疡性结肠炎。

所述代谢疾病可能包括上述非免疫相关性疾病，例如肥胖症。

所述中枢神经系统疾病可能包括阿尔茨海默氏病、帕金森氏病、精神分裂症和抑郁症。

所述疾病和病症还可能包括创伤愈合、透析患者的贫血以及骨质疏松症。

本发明的化合物可能用于制备意在预防或治疗选自下列疾病的药物：某些病毒性感染，如乙型和丙型慢性肝炎，白血病例如毛样细胞白血病和慢性骨髓性白血病，多发性骨髓瘤、滤泡性淋巴瘤、类癌瘤、恶性黑色素瘤、转移性肾癌、阿尔茨海默氏病、帕金森氏病以及伴随免疫缺陷出现的肿瘤，例如在 AIDS 情况下的 Kaposi 肉瘤，以及生殖器疣或性病。

### 实验部分

实施例 1: 包含 SNP c794g、c973a、g1011c、t1049a、t1155a、a1204g 的核苷酸序列的多核苷酸编码的蛋白质以及野生型参考基因的核苷酸序列编码的蛋白质的模建。

第一步，从得自 PDB 数据库(编码 1ITF)的 IFN $\alpha$ -2 结构开始，利用软件 Modeler (MSI, San Diego, CA)，构建 IFN $\alpha$ -21 的三维结构。

然后修饰该成熟多肽的片段，以便复制突变 A19G、Q79K、Q91H、V104D、C139stop 和 K156E。

利用程序 AMBER 和 DISCOVER (MSI: Molecular Simulations Inc.)，对该突变片段进行 1000 次分子最小化步骤。

然后利用相同程序和相同势场，进行两次分子动态计算运行。

在所有情况下，于 300°K 计算 50,000 步，由 300 个平衡步终止。

模建结果由图 1、2、3、4 和 5 所示。

### 实施例 2: 个体群中 SNP c794g、c973a、g1011c、t1049a、t1155a、a1204g 的基因分型

SNP 的基因分型是根据微型测序原理，其中通过读取偏振荧光进行产物检测。该技术由荧光微型测序 (FP-TDI 技术或荧光偏振模板直接染料终止物掺入) 组成。

对群体中每一个体基因组 DNA 的 PCR 扩增产物进行微型测序。选择 PCR 产物，使其包括含待基因分型的 SNP 的基因区域。去除未使用的 PCR 引物以及未掺入的 dNTP，然后进行微型测序。

微型测序包含利用聚合酶和荧光标记的双脱氧核苷酸，延长位于 SNP 位点紧上游的寡核苷酸引物。通过读取偏振荧光，直接分析该延伸过程产生的产物。

所有这些步骤及读取，均在同一 PCR 平板上进行。

因此，基因分型需要 5 个步骤：

- 1) PCR 扩增
- 2) 利用酶促消化纯化 PCR 产物
- 3) 寡核苷酸引物延伸
- 4) 读取
- 5) 解读

每一 SNP c794g、c973a、g1011c、t1049a、t1155a、a1204g 均在相同条件下进行基因分型步骤 1 和 2。步骤 3、4 和 5 针对其中每一多态性是特异的。

1) 从来自不同种族来源的 268 位个体的基因组 DNA 开始，进行 IFN $\alpha$ -21 基因核苷酸序列的 PCR 扩增。

美国 Coriell Institute 提供了这些基因组 DNA。

该 268 位个体分布如下:

系统发生的群体	特定种族的群体	总计	%
非裔美国人	非裔美国人	50	100.0
	小计	50	18.7
美洲印第安人	南美安迪斯人	10	66.7
	西南美国印地安人	5	33.3
	小计	15	5.6
加勒比人	加勒比人	10	100.0
	小计	10	3.7
欧洲高加索人	北美高加索人	79	79.8
	伊比利亚人	10	10.1
	意大利人	10	10.1
	小计	99	36.9
墨西哥人	墨西哥人	10	100.0
	小计	10	3.7
东北亚人	中国人	10	50.0
	日本人	10	50.0
	小计	20	7.5
非欧洲高加索人	希腊人	8	21.6
	印度-巴基斯坦人	9	24.3
	中东人	20	54.1
	小计	37	13.8
东南亚人	太平洋岛民	7	41.2
	南亚人	10	58.8
	小计	17	6.3
南美人	南美人	10	100.0
	小计	10	3.7
总计		268	100

来自其中每一个体的基因组 DNA 构成一个样品。

所有 SNP 的 PCR 扩增均开始于下列引物:

SEQ ID NO. 5: 有义引物: GGTTCAAGGTTACCCATCTC

SEQ ID NO. 6: 反义引物: TTTGAAATGGCAGAAGTCAT

这些核苷酸序列可以扩增具有核苷酸序列 SEQ ID NO. 1 中 620-1315 位核苷酸的 696 个核苷酸长度的片段。

每个 SNP 均以 PCR 产物作为微型测序的模板。

每个样品 PCR 反应的总反应体积为 5  $\mu$ l。

该反应体积由下表所示试剂组成:

供应商	参考	反应物	初浓度	体积/管 ( $\mu\text{l}$ )	终浓度
Life Technology	连同 Taq 附送	缓冲液(X)	10	0.5	1
Life Technology	连同 Taq 附送	MgSO <sub>4</sub> (mM)	50	0.2	2
AP Biotech	27-2035-03	dNTPs (mM)	10	0.1	0.2
	函索	有义引物( $\mu\text{M}$ )	10	0.1	0.2
	函索	反义引物( $\mu\text{M}$ )	10	0.1	0.2
Life Technology	11304-029	Taq platinum	5U/ $\mu\text{l}$	0.02	0.1 U/反应
		H <sub>2</sub> O	Qsp 5 $\mu\text{l}$	1.98	
		DNA (样品)	2.5 ng/ $\mu\text{l}$	2	5 ng/反应
		总体积		5 $\mu\text{l}$	

这些试剂分布于 ABGene 提供的具有 384 孔的黑色 PCR 平板中 (ref :TF-0384-k)。将该平板密封、离心, 然后置于用于 384 孔平板的热循环仪(MJ Research 的 Tetrad), 进行下列温育: PCR 循环: 94°C 1 分钟, 继之以 3 个步骤的 36 个循环(94°C 15 秒、56°C 30 秒、68°C 1 分钟)。

2) 然后利用两种酶: 虾碱性磷酸酶(SAP)和核酸外切酶 I (Exo I) 纯化 PCR 扩增产物。其中第一种酶使 PCR 扩增期间未掺入的 dNTP 去磷酸化, 第二种酶去除单链 DNA 残基, 特别是 PCR 期间未使用的引物。

在 PCR 平板的每个孔中, 每种样品加入 5  $\mu\text{l}$  反应混合物, 进行消化。该反应混合物由以下试剂组成:

供应商	参考资料	反应物	初浓度	体积/管 ( $\mu\text{l}$ )	终浓度
AP Biotech	E70092X	SAP	1 U/ $\mu\text{l}$	0.5	0.5/反应
AP Biotech	070073Z	Exo I	10 U/ $\mu\text{l}$	0.1	1/反应
AP Biotech	连同 SAP 提供	缓冲液 SAP (X)	10	0.5	1
		H <sub>2</sub> O	Qsp 5 $\mu\text{l}$	3.9	
		PCR 产物		5 $\mu\text{l}$	
		总体积		10 $\mu\text{l}$	

加样后, 将平板密封、离心, 然后置于用于 384 孔平板的热循环仪 (MJ Research 的 Tetrad), 进行下列温育: SAP-EXO 消化: 37°C 45 分钟、80°C 15 分钟。

加入每个制备的样品 5  $\mu\text{l}$  反应混合物, 对消化的 PCR 产物进行延伸或微型测序步骤。

微型测序步骤 3)、读取步骤 4) 以及解读步骤 5) 针对每一 SNP c794g、c973a、g1011c、t1049a、t1155a、a1204g 是特异的。

下文描述了所有这些步骤, 详述了其中每个多态性所使用的特定条件。

### 3) 微型测序

根据对应于位于本发明 SNP 位点上游的核苷酸序列, 确定基因分型所需的两个微型测序引物的序列。包含该 SNP 的 PCR 产物是双链 DNA 产物, 因此可以对有义链或反义链进行基因分型。所选引物由 Life Technologies Inc 制备。

下表显示针对每一 SNP 已经测试的微型测序的引物序列以及用于基因分型的最佳条件:

SNP	待测引物	基因分型的最适条件
c794g	SEQ ID NO. 7: 正义 gagggccttgatactcctgg SEQ ID NO. 8: 反义 gagagattcttcccattgt	反义引物 + ddGTP-R110 + ddCTP-Tamra
e973a	SEQ ID NO. 9: 正义 actcatctgctacttgggaa SEQ ID NO. 10: 反义 aaatcttctaggaggctct	反义引物 + dGTP-R110 + ddTTP-Tamra
g1011c	SEQ ID NO. 11: 正义 ttctcactgaacttaacca SEQ ID NO. 12: 反义 gcttccaggctcattcagctg	反义引物 + ddGTP-R110 + ddCTP-Tamra
t1049a	SEQ ID NO. 13: 正义 agcctgcgtgatacaggagg SEQ ID NO. 14: 反义 ggggagctctctccacccca	反义引物 + ddATP-R110 + ddTTP-Tamra
t1155a	SEQ ID NO. 15: 正义 gagaagaaatacagccctg SEQ ID NO. 16: 反义 gctctgacaacctcccaggc	反义引物 + ddATP-R110 + ddTTP-Tamra
a1204g	SEQ ID NO. 17: 正义 tgagatccttctctttatca SEQ ID NO. 18: 反义 taatcttcttgaaaatt	正义引物 + ddGTP-R110 + ddATP-Tamra

首先利用 16 个样品验证 SNP 的微型测序，然后对由 268 位个体和 10 个对照组成的一组个体群进行基因分型。

然后如下表所示，进行延伸或微型测序步骤：

供应商	参考资料	反应物	初浓度	体积/管 ( $\mu$ l)	终浓度
自己制备		延伸缓冲液 <sup>1</sup> (X)	5	1	1
Life Technologies	函索	微型测序引物 ( $\mu$ M) A 或 B	10	0.5	1
AP Biotech	27-2051 (61, 71, 81)-01	ddNTPs <sup>2</sup> ( $\mu$ M) 2 个未 标记	各 2.5	0.25	各 0.125
NEN	Nel 472/5 和 Nel 492/5	ddNTPs <sup>2</sup> ( $\mu$ M) 2 个标 记 Tamra 和 R110	各 2.5	0.25	各 0.125
AP Biotech	E79000Z	热测序酶	3.2 U/ $\mu$ l	0.125	0.4 U/反 应
		H <sub>2</sub> O	Qsp 5 $\mu$ l	3.125	
		消化的 PCR 产物		10	
		总体积		15	

<sup>1</sup> 5X 延伸缓冲液由 250 mM Tris - HCl pH 9、250 mM KCl、25 mM NaCl、10 mM MgCl<sub>2</sub> 和 40 % 甘油组成。

<sup>2</sup> 对于 ddNTP，根据研究的多态性使用 4 种碱基的混合物。只有构成功能性 SNP 的 2 种目的碱基 (野生型核苷酸/突变的核苷酸) 标记

Tamra 或 R110。例如：

对于 SNP c973a, ddNTP 的混合物由以下组成：

- 2.5  $\mu\text{M}$  未标记 ddCTP,
- 2.5  $\mu\text{M}$  未标记 ddATP,
- 2.5  $\mu\text{M}$  ddTTP (1.875  $\mu\text{M}$  未标记 ddTTP 以及 0.625  $\mu\text{M}$  Tamra 标记的 ddTTP),
- 2.5  $\mu\text{M}$  ddGTP (1.875  $\mu\text{M}$  未标记 ddGTP 以及 0.625  $\mu\text{M}$  R110 标记的 ddGTP)。

加样后, 将平板密封、离心, 然后置于用于 384 孔平板的热循环仪 (MJ Research 的 Tetrad), 进行下列温育: 延伸循环: 93°C 1 分钟, 继之以 35 个由 2 个步骤组成的循环 (93°C 10 秒、55°C 30 秒)。

在热循环仪最后步骤之后, 将平板直接置于 LJL Biosystems Inc 的 Analyst® HT 型偏振荧光阅读仪。利用 Criterion Host® 软件, 通过两种方法读取平板。第一种方法可利用 Tamra 特异性的发射和激发滤波器 (激发 550-10nm、发射 580-10nm) 读取该荧光团标记的碱基, 第二种方法可利用 R110 特异性的发射和激发滤波器 (激发 490-10nm、发射 520-10nm) 读取该荧光团标记的碱基。两种情况均使用二色性的双反射镜 (dichroic double mirror) (R110/Tamra), 其它的阅读参数为:

Z-高度: 1.5 mm

衰减器: 关断

积分时间: 100,000 微妙

原始数据单位: 计数/秒

偏振转换: 按孔

平板稳定时间: 0 毫秒

PMT 设置: 精确读取(+), 灵敏度 2

动态起偏镜: 发射

静态起偏镜: S

从而获得包含 Tamra 滤波器和 R110 滤波器的 mP (毫偏振度) 计算值的文件结果。从平行面 (//) 和垂直面 ( $\perp$ ) 所得强度值开始, 按照下列公

式，计算这些 mP 值：

$$mP = 1000(// - g\perp) / (// + g\perp).$$

在该运算中，值 $\perp$ 利用因子 g 权重。这是必须由预实验确定的机器参数。

步骤 4) 和 5) 解读并确定基因型。

利用 Microsoft Inc. Excel 软件、和/或 LJL Biosystems Inc 开发的 Allele Caller® 软件，图解报告 mP 值。

横坐标表示 Tamra 标记碱基的 mP 值，纵座标表示 R110 标记碱基的 mP 值。强 mP 值表示掺入了该荧光团标记的碱基，反之，弱 mP 值表明该碱基未掺入。

可获得多达三个同源组的具有不同基因型的核苷酸序列。

一旦对不同组进行了鉴定，便可以使用 Allele Caller® 软件，以表格形式直接获取所确定的每一个体的基因型。

#### SNP c794g, c973a, g1011c, t1049a, t1155a, a1204g 的微型测序结果

完成基因分型步骤以后，针对此处所研究的 SNP，使用上述图形确定个体群中个体的基因型。

对于 SNP c794g，理论上所测试个体的基因型为纯合子 CC、或杂合子 CG、或纯合子 GG。实际上，如下所示，在个体群中未检测到纯合子基因型 GG。

对于 SNP c973a，理论上所测试个体的基因型为纯合子 CC、或杂合子 CA、或纯合子 AA。实际上，如下所示，在个体群中未检测到纯合子基因型 AA。

对于 SNP g1011c，理论上所测试个体的基因型为纯合子 GG、或杂合子 GC、或纯合子 CC。实际上，如下所示，在个体群中未检测到纯合子基因型 CC。

对于 SNP t1049a，理论上所测试个体的基因型为纯合子 TT、或杂合子 TA、或纯合子 AA。实际上，如下所示，在个体群中未检测到纯合子基因型 AA。

对于 SNP t1155a, 理论上所测试个体的基因型为纯合子 TT、或杂合子 TA、或纯合子 AA。实际上, 如下所示, 在个体群中未检测到纯合子基因型 AA。

对于 SNP a1204g, 理论上所测试个体的基因型为纯合子 AA、或杂合子 AG、或纯合子 GG。实际上, 如下所示, 在个体群中未检测到纯合子基因型 GG。

所测基因型在个体群中的分布以及计算所研究的 6 个 SNP 的不同等位频率的结果如下表所示:

系统发生的群体	总计	c794g (A42G)						总计
		f	(95% CI)	CC	%	CG	%	
非裔美国人	50	5,0	(0,7, 9,3)	45	90,0	5	10	50
美洲印第安人	15			15	100			15
加勒比人	10			10	100			10
欧洲高加索人	99			99	100			99
墨西哥人	10			10	100			10
非欧洲高加索人	37			37	100			37
东北亚人	20			20	100			20
南美人	10			10	100			10
东南亚人	17	2,9	(0, 8,6)	16	94,1	1	5,9	17
总计	268	1,1	(0,2, 2,0)	262	97,8	6	2,2	268

系统发生的群体	总计	c973a (Q102K)						总计
		f	(95% CI)	CC	%	CA	%	
非裔美国人	50			49	100			49
美洲印第安人	15			15	100			15
加勒比人	10			10	100			10
欧洲高加索人	99	1,5	(0, 3,2)	96	97,0	3	3,0	99
墨西哥人	10			9	100			9
非欧洲高加索人	37	1,4	(0, 4,0)	36	97,3	1	2,7	37
东北亚人	20			20	100			20
南美人	10			10	100			10
东南亚人	17			17	100			17
总计	268	0,8	(0, 1,5)	262	98,5	4	1,5	266

系统发生的群体	总计	g1011c (Q114H)						总计
		f	(95% CI)	GG	%	GC	%	
非裔美国人	50			49	100			49
美洲印第安人	15			15	100			15
加勒比人	10			10	100			10
欧洲高加索人	99	0,5	(0, 1,6)	92	98,9	1	1,1	93
墨西哥人	10			10	100			10
非欧洲高加索人	37			37	100			37
东北亚人	20			20	100			20
南美人	10			10	100			10
东南亚人	17			17	100			17
总计	268	0,2	(0, 0,6)	260	99,6	1	0,4	261

系统发生的群体		总计	t1049a (V127D)						总计
			f	(95% CI)	TT	%	TA	%	
非裔美国人	50	1,0	(0, 3.0)	48	98,0	1	2,0		49
美洲印第安人	15			14	100				14
加勒比人	10			10	100				10
欧洲高加索人	99			98	100				98
墨西哥人	10			10	100				10
非欧洲高加索人	37			37	100				37
东北亚人	20			20	100				20
南美人	10			10	100				10
东南亚人	17			17	100				17
总计	268	0,2	(0, 0.6)	264	99,6	1	0,4		265

系统发生的群体		总计	t1155a (C162STOP)						总计
			f	(95% CI)	TT	%	TA	%	
非裔美国人	50	1,0	(0, 3.0)	48	98,0	1	2,0		49
美洲印第安人	15			15	100				15
加勒比人	10			9	100				9
欧洲高加索人	99	3,6	(1.0, 6.2)	91	92,9	7	7,1		98
墨西哥人	10			10	100				10
非欧洲高加索人	37			37	100				37
东北亚人	20			20	100				20
南美人	10			10	100				10
东南亚人	17			17	100				17
总计	268	1,5	(0.5, 2.5)	257	97,0	8	3,0		264

系统发生的群体		总计	a1204g (K179E)						总计
			f	(95% CI)	AA	%	AG	%	
非裔美国人	50			50	100				50
美洲印第安人	15	3,6	(0, 10.4)	13	92,9	1	7,1		14
加勒比人	10			10	100				10
欧洲高加索人	99	1,0	(0, 2.4)	97	98,0	2	2,0		99
墨西哥人	10	10,0	(0, 23.1)	8	80,0	2	20,0		10
非欧洲高加索人	37	2,7	(0, 6.4)	35	94,6	2	5,4		37
东北亚人	20	17,5	(5.7, 29.3)	13	65,0	7	35,0		20
南美人	10	5,0	(0, 14.6)	9	90,0	1	10,0		10
东南亚人	17			17	100				17
总计	268	2,8	(1.4, 4.2)	252	94,4	15	5,6		267

在上表中

- N 代表个体数目。
- %代表特定亚群中个体的百分比。
- 等位频率代表特定亚群中突变等位基因的百分比。
- 95 % IC 代表最小和最大的 95 %置信度区间。

必须说明, 例如对于 SNP c973a, 反义读取的等位基因 g 对应于有义读取的等位基因 c, 并涉及未成熟 IFN $\alpha$ -21 蛋白质序列 102 位存在的谷氨酰胺(Q), 因而反义读取的等位基因 t 对应于有义读取的等位基因 a, 对应于相应蛋白质序列该位置的赖氨酸(K)。

通过按系统发生的群体以及按 SNP 考查这些结果, 可以看出:

- 对于 SNP c794g, 6 个 CG 杂合子个体来自非裔美洲人和东南亚人亚群。
- 对于 SNP c973a, 4 个 CA 杂合子个体来自欧洲人和非裔高加索人亚群。
- 对于 SNP g1011c, 唯一的 GC 杂合子个体来自欧裔高加索人亚群。
- 对于 SNP t1049a, 唯一的 TA 杂合子个体来自非裔美洲人亚群。
- 对于 SNP t1155a, 8 个 TA 杂合子个体来自非裔美洲人和欧裔高加索人亚群。
- 对于 SNP a1204g, 15 个 AG 杂合子个体来自美洲印第安人、欧洲人和非裔高加索人、墨西哥人、东北亚人和南美人亚群。

实施例 3. 在酵母中表达天然野生型 IFN $\alpha$ -21 以及 A42G、Q102K、Q114H/V127D、K179E 突变型 IFN $\alpha$ -21

a) 在真核生物表达载体 pPicZ $\alpha$ -topo 中克隆天然野生型 IFN $\alpha$ -21 和突变型 IFN $\alpha$ -21

利用来自相应 SNP 杂合子个体的基因组 DNA 为模板, PCR 扩增编码天然野生型 IFN $\alpha$ -21 成熟部分以及 A42G、Q102K、Q114H/V127D、K179E 突变型 IFN $\alpha$ -21 的核苷酸序列。

可以进行该扩增的 PCR 引物为:

SEQ ID NO. 19: 有义引物: TGTGATCTGCCTCAGACCCAC

SEQ ID NO. 20: 反义引物: TCATTCCTTCCTCCTTAATCTTTCTTG

将 PCR 产物插入真核生物表达载体 pPicZ $\alpha$ -TOPO, 位于可甲醇诱导的杂合启动子 AOX1 控制下 (TOPO<sup>TM</sup>-cloning; Invitrogen Corp.)。

该载体可使真核蛋白质在巴斯德毕赤氏酵母 (*Pichia pastoris*) 中异源表达。

在核对载体编码重组蛋白质区域的核苷酸序列之后, 利用 PmeI 限制酶将载体线性化, 并利用该重组表达载体转化巴斯德毕赤氏酵母菌株 (Invitrogen)。

#### b) 天然野生型 IFN $\alpha$ -21 和突变型 IFN $\alpha$ -21 蛋白质在巴斯德毕赤氏酵母中的异源表达及纯化

将包含编码天然野生型 IFN $\alpha$ -21 或编码 A42G、Q102K、Q114H/V127D 或 K179E 突变型 IFN $\alpha$ -21 的克隆的两个 50 ml BMGY 培养基 (2% 蛋白胨、1% 酵母抽提物、1.34% YNB、1% 甘油、100 mM 磷酸钾、0.4 mg/L 生物素 pH 6.0) 的饱和前培养物在 30°C、200 转/分 (rpm) 振荡培养 24-48 小时。

当培养物达到饱和细胞密度时 (相当于波长 600 nm 下所测光密度 12), 用来按 5 OD/mL 接种到 250 ml BMMY 培养基 (2% 蛋白胨、1% 酵母抽提物、1.34% YNB、0.5% 甲醇、100 mM 磷酸钾、0.4 mg/L 生物素 pH 6.0)。

然后利用终浓度 1% 的甲醇, 将培养瓶于 180 转/分振荡下 30°C 诱导蛋白质表达 24 小时。

由于在编码序列上游存在 " $\alpha$  因子" 信号肽序列, 酵母将蛋白质分泌于培养基中。 $\alpha$  因子在加工过程中自然断裂。

将悬浮液离心, 从所获上清开始, 利用 HPLC 纯化蛋白质。

在开始前步骤, 超滤 (Labscale, 截留 5000Da, Millipore) 后透析可使酵母上清在 50 mM Tris-Cl pH 9.0、25 mM NaCl 的缓冲液中浓缩 10 倍。

第一个层析步骤可利用于蓝色琼脂糖凝胶柱 (Amersham Pharmacia) 的亲和回收蛋白质。一方面通过 SDS PAGE 型电泳, 另一方面通过针对

IFN $\alpha$ -21 蛋白质的特异性抗体的免疫检测，证实在收集组分中存在该蛋白质。在此步骤，目的蛋白质的纯度高于 75%。

在第二个纯化步骤中，凝胶过滤可使对应于 IFN $\alpha$ -21 蛋白质的收集组分利用 50 mM Tris pH 9.0、25 mM NaCl 更换缓冲液。

最后的纯化步骤包括利用离子交换层析柱分离蛋白质。

将包含重组蛋白质的组分注入预先用 Tris 50 mM pH 9、NaCl 25 mM 缓冲液平衡的阴离子交换柱 (ResourceQ 6.0 mL, Pharmacia)。通过 0,025-1 M NaCl 梯度的 Tris (50 mM pH 9) 缓冲液的移动，进行蛋白质洗脱。

利用 SDS/PAGE 凝胶估计目的蛋白质的纯度，利用光密度法 (Quantity one, Biorad) 和 BCA 分析 (bicinchoninic 酸和硫酸铜, Sigma) 测定蛋白质浓度。

按照该方案获得纯化的天然野生型 IFN $\alpha$ -21 以及 A42G、Q102K、Q114H/V127D 或 K179E 突变型 IFN $\alpha$ -21 蛋白质，最后按比例放大，生产大量蛋白质，用于下述功能测试。

#### 实施例 4. 评价天然野生型 IFN $\alpha$ -21 以及 A42G、Q114H/V127D 或 K179E 突变型 IFN $\alpha$ -21 的免疫调节活性

IFN I 型 (IFN $\alpha$  和 IFN $\beta$ ) 能够调节免疫系统的某些功能。已经证明它们可增强树突状细胞 (DC) 成熟：增强 MHC I 类 (HLA-ABC) 和 II 类 (HLA-DR) 分子的表达，增强与 T 淋巴细胞共刺激有关的分子、CD80、CD86 和 CD83 分子的表达，增强 T 淋巴细胞的刺激功能。

a) A42G、Q114H/V127D 或 K179E 突变型 IFN $\alpha$ -21 对树突状细胞成熟的作用

首先研究 A42G、Q114H/V127D 或 K179E 突变型 IFN $\alpha$ -21 对树突状细胞成熟的免疫调节活性，并与选作商品化内含子 A (Intron A) 产物代表的野生型 IFN $\alpha$ -2 作比较。

为此，首先从利用 GM-CSF 和 IL-4 细胞因子培养的成人外周血单核细胞中生成树突状细胞。利用 CD 14+ 细胞纯化试剂盒纯化以后，将这些树突状细胞置于 100 ng/ml 野生型 IFN $\alpha$ -2 或 A42G、Q114H/V127D

或 K179E 突变型 IFN $\alpha$ -21 存在下, 通过 FACS 分析确定其表现型, 目的在于寻找 MHC I 类和 II 类分子以及 CD40、CD80、CD86、CD83 和 CD1a 标志的表达。这些树突状细胞的成熟状态也与未经 IFN $\alpha$  处理所获树突状细胞相比, 后者提供了未刺激树突状细胞的对照。

每个标志以及 3 次实验条件下测定的荧光强度平均值, 按任意单位表示, 如下表所示:

	HLA ABC	HLA DR	CD40	CD80	CD86	CD83	CD1a
无刺激	64	133	24	25	14	15	26
野生型 IFN $\alpha$ -2	87	281	331	76	45	15	155
A42G 突变型 IFN $\alpha$ -21	76	127	117	37	5	8	209
Q114H/V127D 突变型 IFN $\alpha$ -21	68	122	163	46	8	8	191
K179E 突变型 IFN $\alpha$ -21	181	322	491	87	44	14	127

该测试结果表明, K179E 突变型 IFN $\alpha$ -21 蛋白质具有较高的刺激树突状细胞成熟的能力, 该刺激效应高于野生型 IFN $\alpha$ -2。

b) Q114H/V127D 或 K179E 突变型 IFN $\alpha$ -21 对 T 淋巴细胞释放细胞因子的作用

将 T 淋巴细胞置于相应突变型 IFN $\alpha$ -21 蛋白质以及有或没有强抗原 (SEB) 的存在下, 以模拟抗侵袭免疫应答, 通过测定 T 淋巴细胞释放细胞因子, 同样研究 Q114H/V127D 或 K179E 突变型 IFN $\alpha$ -21 的免疫调节活性。同样在用作对照并选作内含子 A 商品化产物代表的野生型 IFN $\alpha$ -2 存在下进行该测试。

为此, 从健康供体分离外周血单核细胞 (PBMC), 在包含抗 CD3 和抗 CD28 抗体或 SEB 的合适培养基中刺激 16 小时。每个培养物中加入 4  $\mu$ g/mL 野生型 IFN $\alpha$ -2 或 Q114H/V127D、或 K179E 突变型 IFN $\alpha$ -21。刺激之后, T 淋巴细胞利用抗 CD3、抗 CD4 和抗 CD69 抗体或抗 CD3、抗 CD8 和抗 CD69 抗体细胞外标记, 利用针对 Th1 型细胞因子 (IFN- $\gamma$ ) 或 Th2 型细胞因子 (IL-10) 的特异性抗体进行细胞内标记。利用 FACScalibur 和 CellQuest 软件分析荧光细胞。

所得结果表明, 突变型 IFN $\alpha$ -21 蛋白质和野生型 IFN $\alpha$ -2 不刺激

IL-10 和 IFN- $\gamma$  的释放, 因而在缺少 SEB 时并不活化 T 淋巴细胞。相反, 突变型 IFN $\alpha$ -21 蛋白质和野生型 IFN $\alpha$ -2 刺激 SEB 活化的 T 淋巴细胞释放细胞因子 (IL-10 和 IFN- $\gamma$ ), 如下表所示。该表显示 T 淋巴细胞在 SEB 存在下释放细胞因子, 表示为 CD4+ CD69+ 细胞或 CD8+ CD69+ 细胞分别对于 CD4+ T 淋巴细胞和 CD8+ T 淋巴细胞的百分比, 以及 CD69+ 细胞占总细胞的百分比。

T 淋巴细胞		IFN $\gamma$	IL-10	CD69+ 细胞/总细胞
CD4+ CD69+	阴性对照	11.9	7.5	1.26
	野生型 IFN $\alpha$ -2	19.6	24.68	2.7
	Q114H/V127D IFN $\alpha$ -21	38.9	14.6	4.67
	K179E IFN $\alpha$ -21	29.5	15.1	3.84
CD8+ CD69+	阴性对照	8.73	0.65	4.69
	野生型 IFN $\alpha$ -2	16.37	4.26	10.02
	Q114H/V127D IFN $\alpha$ -21	32.24	4.91	14.98
	K179E IFN $\alpha$ -21	28.28	3.8	13.48

这些结果清楚表明, Q114H/V127D 突变型 IFN $\alpha$ -21 和 K179E 突变型 IFN $\alpha$ -21 强烈刺激经 SEB 抗原预先活化的 CD4+ 和 CD8+ T 淋巴细胞释放细胞因子 (IFN $\gamma$  和 IL-10)。在该测试中, T 淋巴细胞产生的干扰素  $\gamma$  在 Q114H/V127D 突变型 IFN $\alpha$ -21 或 K179E 突变型 IFN $\alpha$ -21 存在下高于野生型 IFN $\alpha$ -2 存在条件下。

c) Q114H/V127D 或 K179E 突变型 IFN $\alpha$ -21 对单核细胞释放细胞因子的作用

最后, 通过在有或无细菌性毒剂 (LPS) 存在下测定单核细胞释放的细胞因子, 研究 Q114H/V127D 或 K179E 突变型 IFN $\alpha$ -21 的免疫调节活性。同样在用作对照并选作内含子 A 商品化产物代表的野生型 IFN $\alpha$ -2 存在下进行该测试。

为此, 从健康供体分离人外周血单核细胞 (PBMC), 分析其表现型, 以确定 CD64+ CD4dim 细胞的相对量 (CD64 和 CD4dim 是血液单核细胞的标志)。过夜培养之后, 将这些 PBMC 在单独的培养基 (未刺激的细胞)

或 LPS 存在下(刺激的细胞)培养。在每个培养物中加入 4  $\mu\text{g}/\text{mL}$  野生型  $\text{IFN}\alpha\text{-2}$  或突变型  $\text{IFN}\alpha\text{-21}$ 。培养后,利用抗 CD64 和抗 CD4dim 细胞外标记细胞,并利用针对 Th1 型细胞因子( $\text{TNF-}\alpha$ )、IL-12 和 IL-10 的特异性抗体细胞内标记细胞。

利用 FACScalibur 和 CellQuest 软件分析荧光细胞。

所得结果表明,在缺少 LPS 时,突变型  $\text{IFN}\alpha\text{-21}$  蛋白质和野生型  $\text{IFN}\alpha\text{-2}$  不刺激细胞因子(IL-10、IL-12 和  $\text{TNF-}\alpha$ )的释放。相反,存在 LPS 时, Q114H/V127D 和 K179E 突变型  $\text{IFN}\alpha\text{-21}$  蛋白质以及野生型  $\text{IFN}\alpha\text{-2}$  刺激单核细胞释放细胞因子(IL-10、IL-12 和  $\text{TNF-}\alpha$ ),如下表所示。该表显示单核细胞在存在 LPS 时释放的细胞因子,表示为 CD64+ CD4dim 细胞的百分比以及 CD4dim CD64+细胞占总细胞的百分比。

	IL-10	IL-12	$\text{TNF-}\alpha$	CD4dim CD64+ 细胞 /总细胞
无刺激	16.21	8.52	13.88	3.1
野生型 $\text{IFN}\alpha\text{-2}$	49.34	34.48	50.87	2.71
Q114H/V127D $\text{IFN}\alpha\text{-21}$	50.63	31.81	56.5	2.31
K179E $\text{IFN}\alpha\text{-21}$	60.14	36.42	60.16	4.43

#### 实施例 5. 评价 A42G、Q102K、Q114H/V127D 和 K179E 突变型 $\text{IFN}\alpha\text{-21}$ 的体外抗增殖活性

##### a) 对 Daudi Burkitt 细胞系人淋巴瘤细胞的增殖

对 A42G、Q102K、Q114H/V127D 和 K179E 突变型  $\text{IFN}\alpha\text{-21}$  蛋白质以及野生型  $\text{IFN}\alpha\text{-21}$  蛋白质进行这些测试。将预先在 RPMI 1640 培养基中(补充有 10%胎牛血清和 2 mM L-谷氨酰胺)培养的细胞(人 Daudi Burkitt 淋巴瘤细胞系,下文称为“Daudi 细胞”)按细胞密度  $4 \cdot 10^4$  细胞/孔接种于 96 孔板。

将每孔中的 Daudi 细胞与浓度递增的突变型或野生型  $\text{IFN}\alpha\text{-21}$  蛋白质接触。每个待鉴定的  $\text{IFN}\alpha\text{-21}$  按终浓度为 0.003  $\mu\text{M}$  - 600 nM 进行测试。

然后将 Daudi 细胞在 37°C、5% CO<sub>2</sub> 下培养 66 小时，其后在培养物中加入 Uptiblu 试剂 (Uptima)。再温育 4 小时以后，测定 590nm 处发射的荧光(激发 560nm)，定量细胞增殖速率。

突变型或野生型 IFN $\alpha$ -21 的抗增殖活性是根据 IC<sub>50</sub> 的测定，其对应于抑制 50% 细胞生长的 IFN $\alpha$ -21 浓度。

每个实验条件进行至少 3 次实验，一式三份，可以测定每个 IFN $\alpha$ -21 的平均 IC<sub>50</sub> 值。可以比较对应于突变型与野生型蛋白质 IC<sub>50</sub> 值的比值。结果汇总于下表(标准差如括号内所示)：

	野生型 IFN $\alpha$ -21	A42G IFN $\alpha$ -21	Q102K IFN $\alpha$ -21	Q114H/V127D IFN $\alpha$ -21	K179E IFN $\alpha$ -21
IC <sub>50</sub> (pM)	1.02	2.55	1.05	13.92	3.73
比值 野生型/突变型	-	2.15 (0.78)	1.30 (0.24)	13.10 (3.06)	3.72(0.83)

该测试表明，A42G、Q114H/V127D 和 K179E 突变型 IFN $\alpha$ -21 蛋白质对 Daudi 细胞的细胞抗增殖活性低于野生型 IFN $\alpha$ -21。特别而言，Q114H/V127D IFN $\alpha$ -21 与野生型 IFN $\alpha$ -21 相比，对 Daudi 细胞的细胞抗增殖活性低约 10-16 倍。

#### b) 对 TF-1 红白血病细胞系

还评价了 A42G、Q114H/V127D 和 K179E 突变型 IFN $\alpha$ -21 对红白血病细胞系的作用。同样在用作对照并选作内含子 A 商品化产物代表的野生型 IFN $\alpha$ -2 存在下进行该测试。

为此，将 TF-1 细胞与浓度递增的突变型 IFN $\alpha$ -21 或野生型 IFN $\alpha$ -2(0.001 - 1000 ng/ml)接触，测定细胞增殖。

结果可以表示为 IC<sub>30</sub>，即对应于抑制 30% 细胞增殖的 IFN $\alpha$  浓度，汇总于下表：

	野生型 IFN $\alpha$ -2	A42G IFN $\alpha$ -21	Q114H/V127D IFN $\alpha$ -21	K179E IFN $\alpha$ -21
IC <sub>30</sub> (ng/mL)	0.66	2.33	1.97	1.62

这些数据表明，这 3 个突变型 IFN $\alpha$ -21 对 TF-1 细胞具有弱的抗增

殖作用，该作用类似于野生型 IFN $\alpha$ -2，提示 A42G、Q114H/V127D 和 K179E 突变型 IFN $\alpha$ -21 的血液学毒性并不高于野生型 IFN $\alpha$ -2。

### 实施例 6. 评价 A42G、Q114H/V127D 和 K179E 突变型 IFN $\alpha$ -21 的抗病毒活性

IFN 在抗病毒防护中具有重要作用。IFN 的抗病毒活性部分应归于 IFN 诱导的酶系统，例如：

- 2' 5' 寡腺苷酸合成酶，催化腺苷寡聚物合成。这些寡聚物可活化 RNase L，一种一旦活化便可破坏病毒 RNA 的内切核糖核酸酶。

- 抑制病毒转录本合成和/或成熟的 Mx 蛋白质(GTP 酶)。该活性主要作用于流感病毒。

- 由双链 RNA 活化的 PKR 蛋白质(或 p68 激酶)。活化的 PKR 可抑制蛋白质合成。

也可通过其它机制诱导 IFN 的抗病毒活性，例如对于逆转录病毒而言，抑制病毒颗粒进入细胞、复制、结合、颗粒排出以及病毒颗粒的感染力。

最后，IFN 通过调节免疫系统的某些功能、特别是促进细胞介导的应答(包括增加 MHC I 类和 II 类分子、增加 IL-12 和 IFN- $\gamma$  生成、增加 CTL 活性等)，而发挥间接的抗病毒活性，

利用细胞培养物体外以及小鼠模型体内评价了 A42G、Q114H/V127D 和 K179E 突变型 IFN $\alpha$ -21 的抗病毒活性。两个测试均利用用作对照并选作内含子 A 商品化产物代表的野生型 IFN $\alpha$ -2 平行进行。

#### a) 细胞培养物体外抗病毒活性

该测定利用水泡性口膜炎病毒(VSV)可以评价 A42G、Q114H/V127D 和 K179E 突变型 IFN $\alpha$ -21 以及野生型 IFN $\alpha$ -2 在细胞培养物中的抗病毒活性。

为此，在浓度递减的突变型 IFN $\alpha$ -21 和野生型 IFN $\alpha$ -2 存在下培养 WISH 人上皮细胞 24 小时。然后利用水泡性口膜炎病毒(VSV)感染细胞 24-48 小时，测定细胞溶解作用。

通过比较 IC50 值，即对应于抑制 50% VSV 诱导的细胞溶解作用的

IFN 浓度，确定所测各种 IFN $\alpha$  的抗病毒作用。

同样实验进行两次，测定的 IC<sub>50</sub> 平均值如下表所示：

	野生型 IFN $\alpha$ -2	A42G IFN $\alpha$ -21	Q114H/V127D IFN $\alpha$ -21	K179E IFN $\alpha$ -21
IC <sub>50</sub> (ng/mL)	4	14	22	25

因此，在 VSV 感染的细胞培养物中，A42G、Q114H/V127D 和 K179E 突变型 IFN $\alpha$ -21 具有低于野生型 IFN $\alpha$ -2 的抗病毒活性。

#### b) 小鼠模型体内抗病毒活性

在 EMCV (脑心肌炎病毒) 小鼠模型中进行体内测试。

人 IFN 在小鼠中显示剂量依赖性抗病毒活性，一般比相同数量的小鼠 IFN 低 100 - 1,000 倍 (Meister 等人 (1986). J. Gen. Virol. 67, 1633-1644)。

利用脑心肌炎病毒 (EMCV) 腹膜内注射小鼠，引起以有关中枢神经系统和脑炎为特征的快速进行性致命疾病 (Finter NB (1973). Front Biol. 2: 295-360)。已经证明，小鼠和人干扰素- $\alpha$  均有效保护小鼠抵抗致命的 EMCV 感染 (Tovey 和 Maury (1999). J. IFN Cytokine Res. 19: 145-155)。

利用 100 x LD<sub>50</sub> EMCV 腹膜内 (ip) 感染 20 只 6 周龄的 Swiss 小鼠组，然后 1 小时后利用 2  $\mu$ g A42G、Q114H/V127D、K179E 突变型 IFN $\alpha$ -21 或野生型 IFN $\alpha$ -2 制剂处理，以后每日一次，进行 3 天。对照组动物只利用赋形剂处理。每日跟踪动物存活率，进行 21 天。

结果如图 6 所示，表明 A42G、Q114H/V127D 或 K179E 突变型 IFN $\alpha$ -21 处理小鼠的相对存活率相比未处理小鼠高很多，但是仍类似于野生型 IFN $\alpha$ -2 处理小鼠所观察的结果。

所有这些结果表明，A42G、Q114H/V127D 和 K179E 突变型 IFN $\alpha$ -21 具有独特的生物学特性。

<110> Escary, Jean-Louis  
 <120> IFN $\alpha$ -21 基因的新多核苷酸和多肽  
 <130> 021349/0012  
 <150> FR 0104404  
 <151> 2001-03-30  
 <160> 20  
 <170> PatentIn version 3.1  
 <210> 1  
 <211> 2001  
 <212> DNA  
 <213> 人类 (Homo sapiens)  
 <400> 1  
 attttgtcat ataaagcaaa attcaaagct tcatatatat actatgagaa aaattttaaa 60  
 aaattattga ttcatatatt tagcagtttt gaatgattaa ctatgtaatt atattcatat 120  
 tattaatgtg tatttatata gattttttatt ttgcatatgt aatttcataa aacaaaattt 180  
 acatgaacaa attacattaa aagttattcc acaaatatac ttatcaaatt aagttaaatg 240  
 tcaatagctt ttaaacttag attttagttt aacatttctg tcattcttta ctttgaataa 300  
 aaagagcaaa ctttatagtt tttatctgtg aagtagagat atacatatta tacatacata 360  
 gataagccaa atctgtgtta ctaaaatttc atgaagattt caattagaaa aaaataccat 420  
 aaaatgtttt gagtgcaggg gaaaaatagg caatgatgaa aaaaaatgaa aaacatctgt 480  
 aaacacatgt agagagtgca taaagaaagc aaaaacagag atagaaagta aaactagggc 540  
 atttagaaaa tggaaattag tatgttccact atttaagacc taogcacaga gcaaagtctt 600  
 cagaaaacct agaggccaag gttcaagggt acccatotca agtagcctag caatattggc 660  
 aacatcccaa tggccctgtc cttttcttta ctgatggcog tgctgggtgct cagctacaaa 720  
 tccatctgtt ctctgggtg tgatctgcct cagacccaca gcctgggtaa taggagggcc 780  
 ttgatactcc tggcacaaat ggaagaatc tctcctttct cctgcctgaa ggacagacat 840  
 gactttggat tccccagga ggagtttgat ggcaaccagt tccagaaggc tcaagccatc 900  
 tctgtcctcc atgagatgat ccagcagacc ttcaatctct tcagcacaaa ggactcatct 960

```

gctacttggg aacagagcct cctagaaaaa ttttccactg aacttaacca gcagctgaat 1020
gacctggaag cctgcgtgat acaggagggtt ggggtggaag agactcccct gatgaatgtg 1080
gactccatcc tggctgtgaa gaaatacttc caaagaatca ctctttatct gacagagaag 1140
aaatacagcc cttgtgcctg ggagggttgc agagcagaaa tcatgagatc cttctcttta 1200
tcaaaaattt ttcaagaaag attaaggagg aaggaatgaa acctgtttca acatggaaat 1260
gatctgtatt gactaataca ccagtcocaca cttctatgac ttctgccatt tcaagactc 1320
atctctccta taaccaccgc atgagttgaa tcaaaaattt cagatctttt caggagtgtg 1380
aggaaacatc atgtttacct gtgcaggcac tagtccttta cagatgacca tgctgataga 1440
tctaattatc tatctattga aatatttatt tatttattag atttaaatta tttttgtcca 1500
tgtaatatta tgtgtacttt tacatttgtt tatatcaaaa tatgtgattt atatttagtc 1560
aatatattat ttcttttaac taaattttac tattaaaaact tcttatatta tttgtttatt 1620
ctttaataaa gaaataccaa gcccaaatgt gcaatctcat taaagaatgg atgggtacaat 1680
tcatttacc c atcatcattg tatccaaatt gtaagtaaaa attgactttc tctaagcgag 1740
gttttatatt gcccttagga tatccagggtg aacataacaa ataccgtttt cgctttcttg 1800
tatcttttat ttttgtaagg aaaataataa ctatactttc taatacctgt tacattaaat 1860
gctatagtga gaagaaataa aaacaaatga aattcagtaa aactgaagca aggcatatca 1920
aaatTTTTTT taaaaaagta gtagatatcc tctatagcag acaagtagac atctaagtgc 1980
aagtgtccat tggtaacctg a 2001

```

<210> 2

<211> 189

<212> PRT

<213> 人类

<400> 2

Met Ala Leu Ser Phe Ser Leu Leu Met Ala Val Leu Val Leu Ser Tyr  
1 5 10 15

Lys Ser Ile Cys Ser Leu Gly Cys Asp Leu Pro Gln Thr His Ser Leu  
20 25 30

Gly Asn Arg Arg Ala Leu Ile Leu Leu Ala Gln Met Gly Arg Ile Ser  
35 40 45

Pro Phe Ser Cys Leu Lys Asp Arg His Asp Phe Gly Phe Pro Gln Glu  
50 55 60

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

<210> 3  
 <211> 20  
 <212> DNA  
 <213> 人类

<400> 3  
 ggttcaaggt tacccatctc 20

<210> 4  
 <211> 20  
 <212> DNA  
 <213> 人类

<400> 4  
 ggttcaaggt tacccatctc 20

<210> 5  
 <211> 20  
 <212> DNA  
 <213> 人类

<400> 5  
 ggttcaaggt tacccatctc 20

<210> 6  
 <211> 20

---

<212> DNA  
<213> 人类

<400> 6  
tttgaaatgg cagaagtcac 20

<210> 7  
<211> 20  
<212> DNA  
<213> 人类

<400> 7  
gagggccttg atactcctgg 20

<210> 8  
<211> 20  
<212> DNA  
<213> 人类

<400> 8  
gagagattct tcccatattgt 20

<210> 9  
<211> 20  
<212> DNA  
<213> 人类

<400> 9  
actcatctgc tacttgggaa 20

<210> 10  
<211> 20  
<212> DNA  
<213> 人类

<400> 10  
aaatttttct aggaggctct 20

<210> 11  
<211> 20  
<212> DNA  
<213> 人类

<400> 11  
ttttccactg aacttaacca 20

<210> 12  
<211> 20  
<212> DNA

---

<213> 人类

<400> 12  
gcttccaggt cattcagctg 20

<210> 13  
<211> 20  
<212> DNA  
<213> 人类

<400> 13  
agcctgcgtg atacaggagg 20

<210> 14  
<211> 20  
<212> DNA  
<213> 人类

<400> 14  
ggggagtctc ttccacccca 20

<210> 15  
<211> 20  
<212> DNA  
<213> 人类

<400> 15  
gagaagaaat acagcccttg 20

<210> 16  
<211> 20  
<212> DNA  
<213> 人类

<400> 16  
gctctgacaa cctcccaggc 20

<210> 17  
<211> 20  
<212> DNA  
<213> 人类

<400> 17  
tgagatcctt ctctttatca 20

<210> 18  
<211> 20  
<212> DNA

<213> 人类

<400> 18

taatccttct tgaaaaattt

20

<210> 19

<211> 21

<212> DNA

<213> 人类

<400> 19

tgtgatctgc ctcagaccca c

21

<210> 20

<211> 27

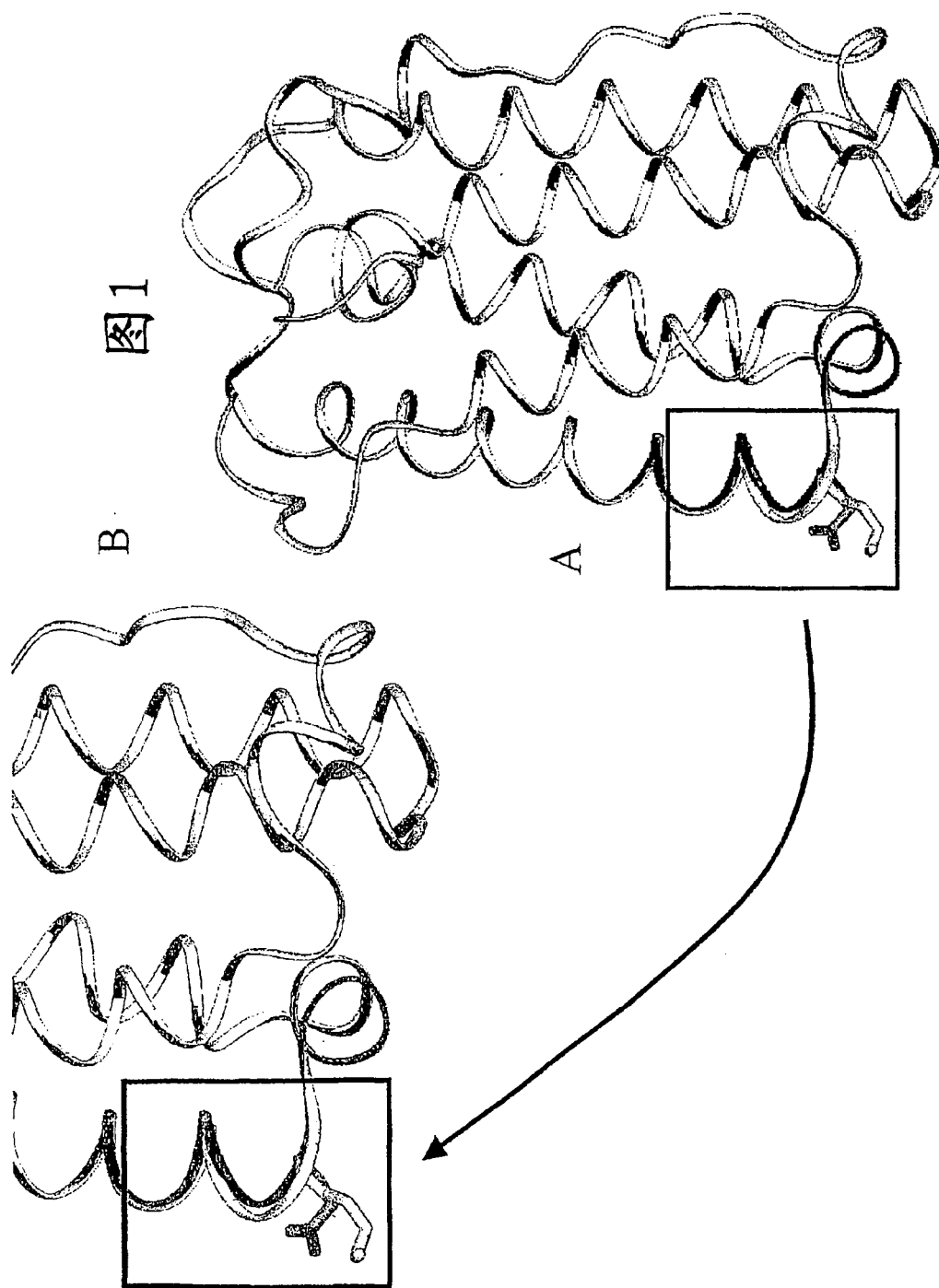
<212> DNA

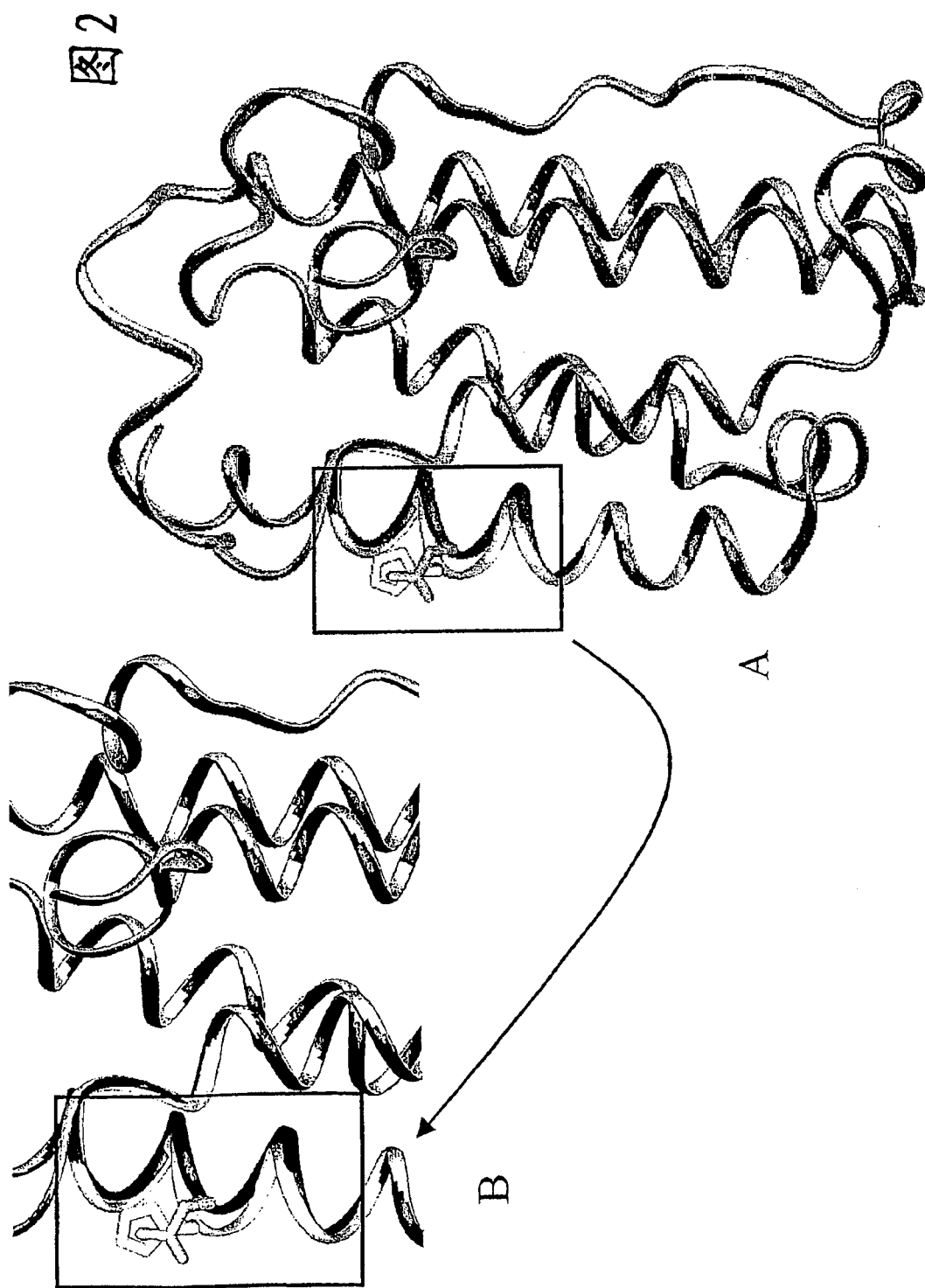
<213> 人类

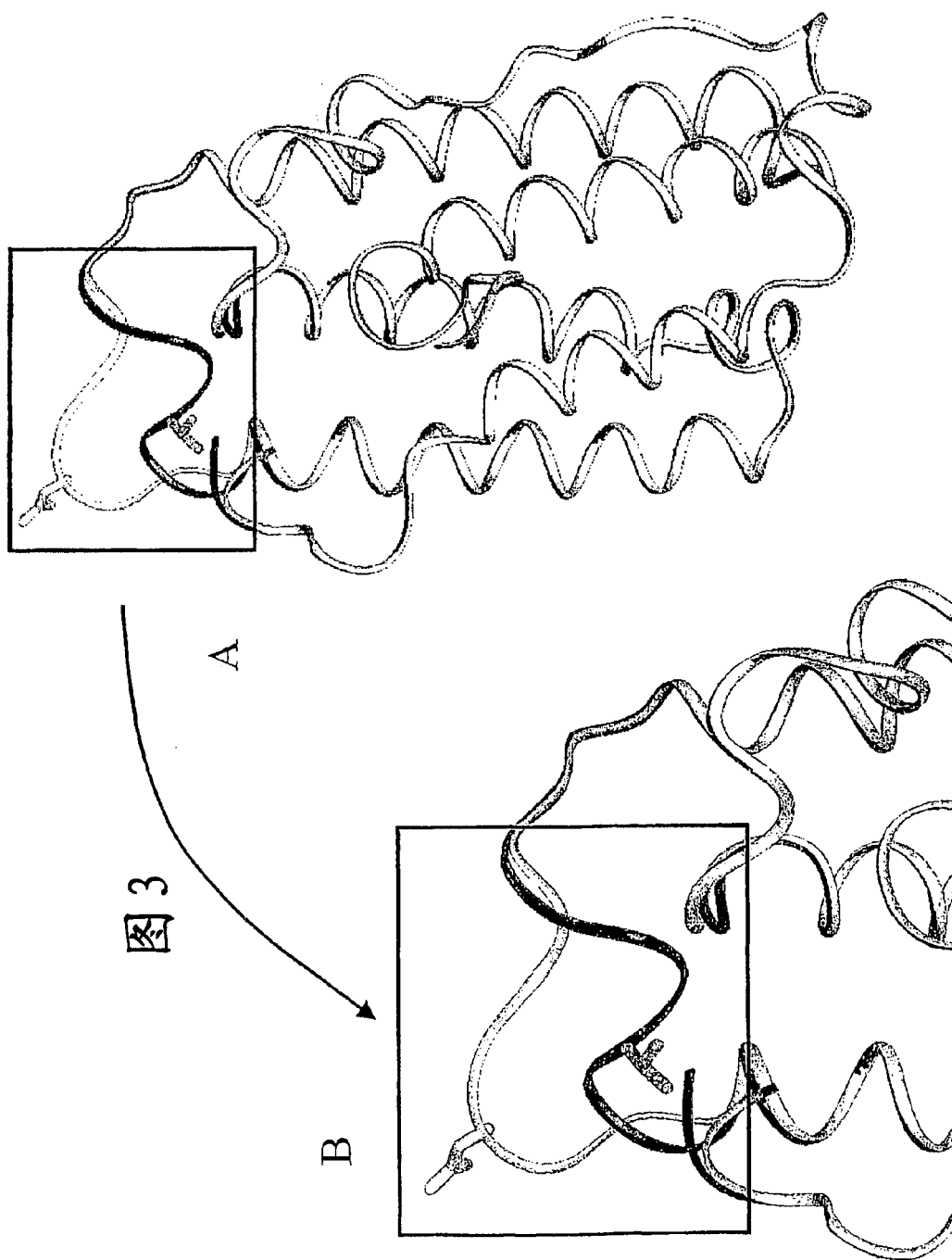
<400> 20

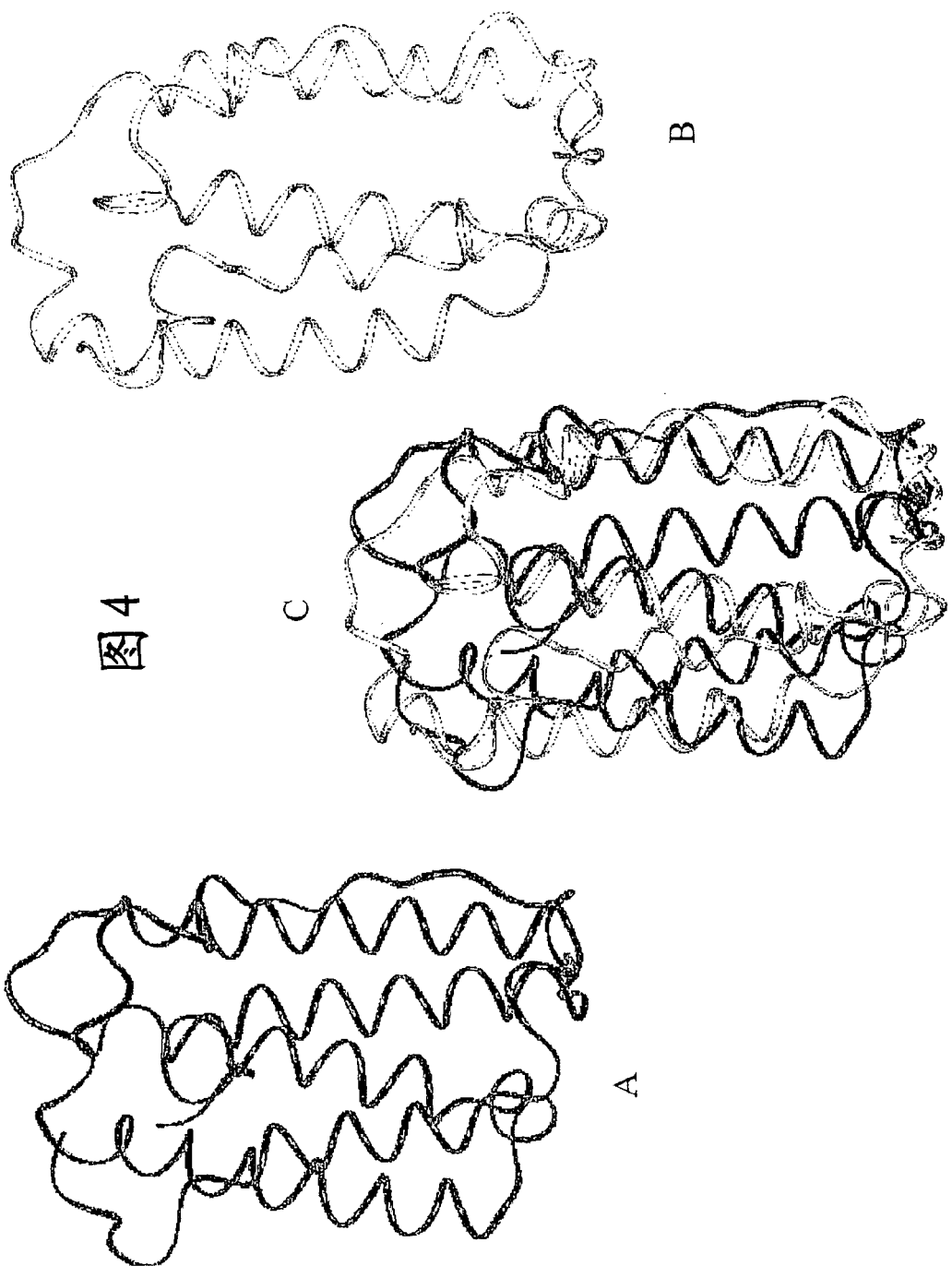
tcattccttc ctccttaatc tttcttg

27









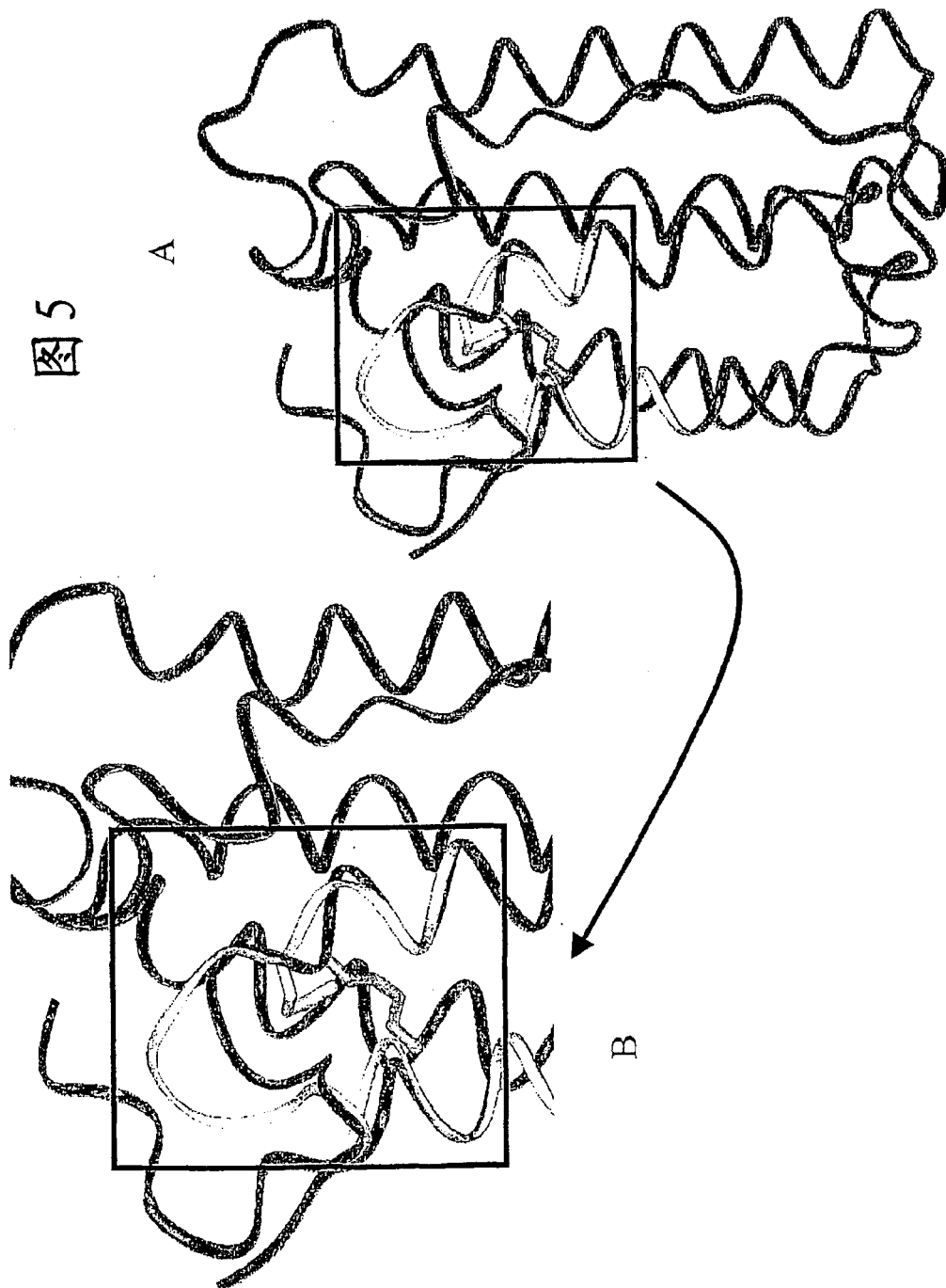
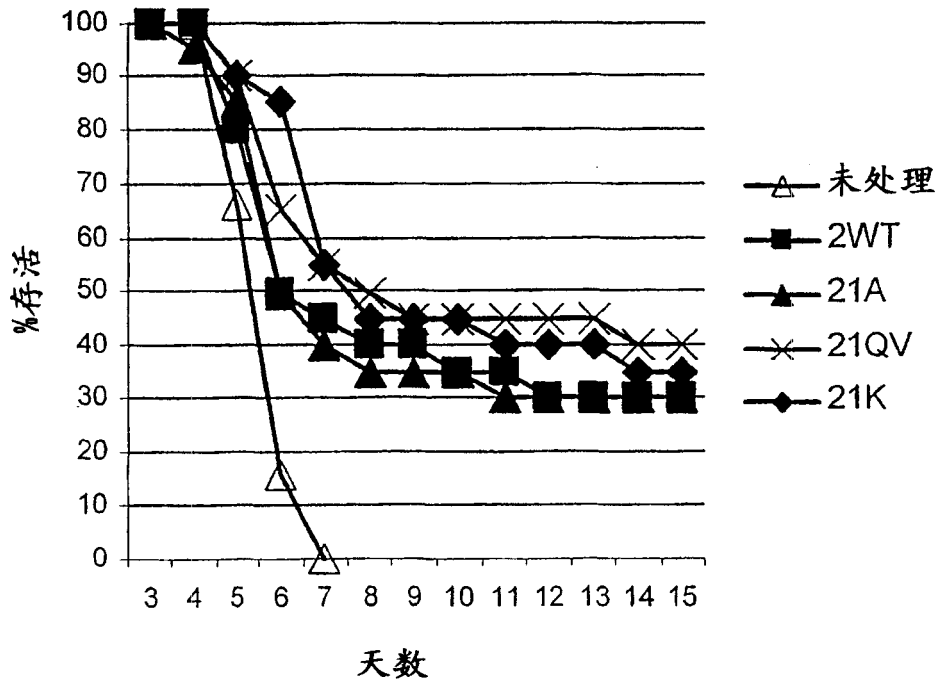


图6



专利名称(译)	IFN $\alpha$ - 21基因的新的多核苷酸和多肽		
公开(公告)号	<a href="#">CN1531600A</a>	公开(公告)日	2004-09-22
申请号	CN02809055.1	申请日	2002-03-29
发明人	J-L·伊斯卡利		
IPC分类号	G01N33/50 A61K31/7088 A61K38/21 A61K39/00 A61K39/395 A61K48/00 A61P1/04 A61P3/00 A61P3/04 A61P7/06 A61P9/00 A61P11/06 A61P17/02 A61P17/06 A61P17/12 A61P19/00 A61P19/02 A61P19/10 A61P25/00 A61P25/16 A61P25/24 A61P25/28 A61P31/12 A61P31/14 A61P31/18 A61P35/00 A61P35/02 A61P37/00 A61P37/02 A61P37/06 A61P37/08 C07K14/56 C07K16/24 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12N15/10 C12N15/21 C12N15/85 C12P21/02 C12P21/08 C12Q1/02 C12Q1/68 C12Q1/6886 G01N33/15 G01N33/53 G01N33/566 G01N33/68 C12N15/63 C07K16/00		
CPC分类号	C12N2830/15 C12N15/1034 A61K2039/53 C12Q1/6886 C07K14/56 C12N2830/002 A61K39/00 C12N15/85 A61P1/04 A61P3/00 A61P3/04 A61P7/06 A61P9/00 A61P11/06 A61P17/02 A61P17/06 A61P17/12 A61P19/00 A61P19/02 A61P19/10 A61P25/00 A61P25/16 A61P25/24 A61P25/28 A61P31/00 A61P31/12 A61P31/14 A61P31/18 A61P35/00 A61P35/02 A61P37/00 A61P37/02 A61P37/06 A61P37/08 C12Q2600/136 C12Q2600/156		
代理人(译)	刘晓东		
优先权	2001004404 2001-03-30 FR		
外部链接	<a href="#">Espacenet</a> <a href="#">SIPO</a>		

摘要(译)

本发明涉及来源于IFN $\alpha$  - 21基因核苷酸序列的包含新的SNP的新的多核苷酸、以及来源于天然野生型IFN $\alpha$  - 21蛋白质的包含由至少一个本发明的SNP所致至少一个突变的新的多肽、及其治疗用途。

