

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
C12Q 1/68 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200680042670.3

[43] 公开日 2008 年 11 月 19 日

[11] 公开号 CN 101310026A

[22] 申请日 2006.11.9

[21] 申请号 200680042670.3

[30] 优先权

[32] 2005.11.15 [33] EP [31] 05110738.1

[86] 国际申请 PCT/EP2006/010726 2006.11.9

[87] 国际公布 WO2007/057119 英 2007.5.24

[85] 进入国家阶段日期 2008.5.15

[71] 申请人 弗·哈夫曼-拉罗切有限公司

地址 瑞士巴塞尔

[72] 发明人 R·G·克莱尔 G·迪沙托一源

L·埃塞奥克斯 D·拉加德

C-G·奥斯滕松

[74] 专利代理机构 北京市中咨律师事务所

代理人 黄革生 林柏楠

权利要求书 2 页 说明书 19 页 序列表 21 页
附图 6 页

[54] 发明名称

AZGP 基因单核苷酸多态性 (SNPs)

[57] 摘要

本发明提供了 AZGP1 基因中的单核苷酸多态性和单元型，其可以用于确定个体的肥胖症素质。

1. 分离的核酸，其包括 SEQ ID No. 2 或者含有位置 8047、8077-8083、8500、9556 或 12002 的 SEQ ID No. 2 的片段，但其在下述位置之一具有单多态性改变：

在位置 9556 的 zag15，其中在此位置的 T 被 C 替代，
在位置 8500 的 zag16，其中在此位置的 A 被 G 替代，
在位置 8047 的 zag17，其中在此位置的 A 被 G 替代，
在位置 8077-8083 的 zag_del，其中在这些位置的核酸发生缺失，
在位置 12002 的 zag35，其中在此位置的 T 被 C 替代。

2. 确定个体的肥胖症的素质的方法，包括下述步骤：

a) 从分离自患者的样品中分离核酸，和

b) 检测存在于一个或者多个如权利要求 1 列出的 Seq ID No. 2 中的多态性位点处的核苷酸，其中在根据权利要求 1 的多态性位点存在特定的核苷酸指示着患者患肥胖症的倾向。

3. 选自单元型 1 的分离的核酸分子，其中单元型 1-3 每个均包括 SEQ ID No. 2，但是在下述表中针对相应单元型给出的核苷酸存在于 Seq ID No. 2 的相应位置上：

位置	单元型 1	单元型 2	单元型 3
8047	A	G	G
8077-8083	del	wt	wt
8500	A	G	G
9556	T	C	T
12002	T	C	C

4. 用于对个体的 AZGP1 基因进行单元型分型的方法，包括下述步骤：

a) 从分离自个体的样品中分离核酸；

b) 确定在个体基因 X 的拷贝的位置 8047、8077-8083、8500、9556 和 12002 上存在的核苷酸，其中该位置编号通过与 SEQ ID No. 2 比较而确

定。

5. 权利要求 4 的方法，还包括步骤：

c) 通过将在所述位置上的核苷酸与权利要求 3 的表格中的相应于单元型给出的核苷酸比较而确定个体的特定单元型。

6. 权利要求 4 或 5 的方法，其中存在至少一个权利要求 3 的表格中列出的单元型指示着个体患肥胖症的倾向。

AZGP 基因单核苷酸多态性 (SNPs)

发明领域

本发明涉及与肥胖症相关的 *AZGP1* 基因中的 SNPs 和单元型, 以及涉及通过 *AZGP1* 基因中所述 SNPs 和/或单元型的存在和或不存在而确定个体的肥胖症素质的方法。

背景技术

诸如肥胖症的多因素质病由多于一种的基因中的突变引起, 并且环境因素在其中起到很大的作用。虽然在鉴定引起孟德尔式疾病的基因方面取得了引人注目的成功, 但是在发现涉及多因素质病的易感性基因方面至今仍存在困难。有证据表明人类遗传了在高脂肪膳食下增重的遗传素质。有利的是鉴定个体的肥胖症素质的标记。

AZGP1 是 Zn- α 2-糖蛋白, 其基因在肥胖症中被下调 (EP 1548445), 而在恶病质中被上调 (Russell 和 Tisdale, 2005, Brit. J. Cancer 92, 876-881; Russell 等人, 2004, Biochem. Biophys. Acta 1636, 59-68; Sanders 和 Tisdale, 2004, Cancer Lett. 212, 71-81; Bing 等人, 2004, Proc. Natl. Acad. Sci USA 101, 2500-2505)。

至今, 还没有有人将 *AZGP1* 单元型与肥胖症关联在一起。

发明内容

本发明要解决的问题是提供个体的肥胖症素质的标记。本发明通过鉴定 *AZGP1* 基因中与肥胖症相关的 SNPs 和单元型而解决该问题。使用从纤瘦的和肥胖的受试者获得的 DNA 样品进行 *AZGP1* 基因中的单元型和 SNPs 的鉴定。这些 SNPs 和单元型与肥胖症相关。如从文献已知的, 肥胖症与胰岛素抗性相关, 因此, 这些 SNPs 也可与胰岛素抗性关联。当采取

样品时，参与本研究的肥胖的受试者无糖尿病。*AZGP1* 基因的 DNA 片段通过 PCR 扩增并测序。测序之后，使用 Polyphred 软件（华盛顿大学）进行多态性分析。表 1 列出了在 *AZGP1* 中鉴定到的所有的标记。

表 2 显示了在 *AZGP1* DNA 样品中发现的所有的多态性位点的等位基因频率。对单元型频率的计算，仅仅使用具有高于 5% 的较小等位基因频率的 SNPs。更小频率的标记不可能被选择用于进一步的相关性研究，并且基本上对于常见单元型没有贡献。在表 2 中列出的 28 个标记中，粗体表示的 15 个标记进一步用于单元型分析。

表 3 提供了 *AZGP1* 的 15 个常见标记的单元型频率。如可以从表中看到的，一些标记对是完全冗余的（在不同的单元型中等位基因等同出现）：

- zag18 和 zag19
- zag17、zag16 和内含子序列缺失（zag del）。
- zag16、zag17 和内含子序列缺失（zag del）这一簇与 zag15 和 zag35 是近乎冗余的。通过察看表 6，从 zag17 至 zag35 仅仅有 3 个单元型存在：AdelATT、GwtGCC 和 GwtGTC。最后一个单元型仅仅存在于 H12 中，其是稀有单元型（ $f=0.005$ ）。

表 1. 在 DNA 样品中, AZGP1 中鉴定到的所有的标记的特征

标记	Pos.	等位基因	旁侧序列	在 AZGP1 中的定位	Seq ID No.
zag06	1090 ^{1,2}	G/A	AATAACAATACCTGCGGCTAGACTTTGGAGC	未知	3
zag05	1196 ^{1,2}	T/C	AACCAAAAGAGAGGGCTGGGCACAGTTGCTCA	未知	4
zag04	1216 ^{1,2}	T/A	ACAGTTGCTCACACTTGTAACCCAGCACTT	未知	5
zag03	1348 ^{1,2}	C/T	GCATGTGCCACCACGCGCAGCTAATTCCTGT	未知	6
zag07	2695 ^{1,2}	T/C	TAGGAACCATATGCCTGGAGCTGCTTCTGCT	内含子1	7
zag08	2760 ^{1,2}	T/G	CCTGCCTGACGCTGATGGAAAGAGAGAGCAG	内含子1	8
zag09	2762 ^{1,2}	A/G	TGCCTGACGCTGAGGAAAAGAGAGAGCACCC	内含子1	9
zag13	4528 ^{1,2}	G/A	TCAGCCTTCTGAGTCGCTGGGACTACAGGTG	内含子1	10
zag12	5013 ^{1,2}	T/C	ATTATGGAACATTATGGAAATGTCCCTCTC	内含子1	11
zag10	5369 ^{1,2}	G/T	TGCTTGGCTAATTTTGTGAATTCTTAGTAGA	内含子1	12
zag14	6561 ^{1,2}	C/T	GACCCTGAAAGACATCGTGGAGTATTACAAC	Ex02, 沉默	13
zag23	6730 ^{1,2}	A/G	AACACAGACATGTCCACATCCCACCCACCCC	内含子2	14
zag22	6894 ^{1,2}	C/T	GGAGGCTGATACCCCGTGAGAAGGCATCAG	内含子2	15
zag18	7202 ^{1,2}	C/A	GAAATTTGTGGAATCCACAGAGAAAAGCACC	内含子2	16
zag19	7219 ^{1,2}	G/A	CAGAGAAAAGCACCCGGCACACACCGTAGCC	内含子2	17
zag20	7454 ^{1,2}	T/C	CCAAGGCAGCCAACCTCAGGTCTGGTGAACCT	内含子2	18
zag21	7459 ^{1,2}	T/C	GCAGCCAACCTCAGGTCTGGTGAACCTGCTGG	内含子2	19
zag17	8047 ^{1,2}	A/G	TTGCACTACAGCCTGAGTGACAAGAGTGAAA	内含子2	20

zag del	8077- 8083 ²	AAAAAA C/.	TTGTCTAAAAACAAAAACAAAAACAAAAA	内含子2	21
zag16	8493 ¹ (8500 ²)	A/G	ATCAAACACCAGAAAAGTAGAAAGAAGTGA	内含子2	22
zag15	9549 ¹ (9556 ²)	T/C	GTAGTGGTGGGATTTTGCCATATCACCCCTGG	内含子2	23
zag24	10202 ¹ (10209 ²)	A/C	TGCTTCCTGCTCCCCAGTACTGAGCCCAGAA	内含子3	24
zag25	10439 ¹ (10446 ²)	G/A	CATCTCCAATTAACAGACAAGGAGCTTGAGG	内含子3	25
zag26	11020 ¹ (11027 ²)	G/T	GTCCACCTCAAGCCTGCAGTGTCACACTCTA	内含子3	26
zag35	11995 ¹ (12002 ²)	T/C	GGGAGAATATCTCTCTCAATATACAAGGGGT	未知	27
zag34	12385 ¹ (12392 ²)	G/T	TCCCAGTATCGCAGGGGGTGTGCACCCCCC	未知	28

¹ 标记在 EMBL 检索号 ac004977 的序列中的位置

² 标记在 Seq ID No. 2 中的位置

表 2 在 *AZGP1* 中发现的 SNPs 的等位基因频率

标记	区域	距离 (bp) ¹	All1	All2	N ²	f_All1	f_All2
zag06	5' 区	-	G	A	93	0.995	0.005
zag05	5' 区	110	T	C	92	0.549	0.451
zag04	5' 区	16	T	A	92	0.745	0.255
zag03	5' 区	132	T	C	92	0.022	0.978
zag07	内含子 1	1347	T	C	93	0.823	0.177
zag08	内含子 1	65	T	G	93	0.005	0.995
zag09	内含子 1	2	G	A	93	0.027	0.973
zag13	内含子 1	1766	G	A	93	0.995	0.005
zag12	内含子 1	485	T	C	91	0.973	0.027
zag10	内含子 1	355	T	G	91	0.709	0.291
zag14	外显子 2	1192	T	C	89	0.461	0.539
zag23	内含子 2	169	G	A	91	0.456	0.544
zag22	内含子 2	164	T	C	91	0.033	0.967
zag18	内含子 2	308	C	A	92	0.75	0.25
zag19	内含子 2	17	G	A	93	0.753	0.247
zag20	内含子 2	235	T	C	93	0.968	0.032
zag21	内含子 2	5	T	C	93	0.962	0.038
zag17	内含子 2	588	G	A	92	0.527	0.473
zag del	内含子 2	30	wt	del	91	0.527	0.473
zag16	内含子 2	423	G	A	90	0.533	0.467
zag15	内含子 2	1048	T	C	91	0.478	0.522
zag24	内含子 3	653	C	A	92	0.005	0.995
zag25	内含子 3	237	G	A	93	0.995	0.005
zag26	内含子 3	581	T	G	92	0.005	0.995
zag35	3' 区	975	T	C	93	0.478	0.522
zag34	3' 区	390	T	G	93	0.022	0.978

¹: 本 SNP 与前面一个 SNP 之间的距离

²: 具有测序数据的 DNA 样品的数目

Wt: 野生型序列

Del: 其中位置 8077-8083 发生缺失的序列

表 3 原始单元型频率表

标记	等位 基因	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16
zag05	T/C	T	C	C	T	T	C	T	C	T	T	T	C	T	T	T	T
zag04	T/A	T	T	T	A	A	T	T	A	A	T	A	T	A	A	A	A
zag07	T/C	T	T	C	T	T	C	T	T	T	C	T	T	T	T	C	C
zag10	T/G	G	T	T	T	T	T	T	T	T	G	T	T	T	T	T	T
zag14	T/C	C	T	T	C	C	T	C	C	C	C	C	T	T	T	C	C
zag23	G/A	A	G	G	A	A	G	A	A	A	A	G	G	A	G	A	A
zag18	C/A	C	C	C	A	A	C	A	A	C	C	C	C	A	A	A	C
zag19	G/A	G	G	G	A	A	G	A	A	G	G	G	G	A	A	G	G
zag17	G/A	A	G	G	A	G	A	A	A	G	A	G	G	A	A	A	A
zag_del	wt/del	del	wt	wt	del	wt	del	del	del	wt	del	wt	wt	del	del	del	del
zag16	G/A	A	G	G	A	G	A	A	A	G	A	G	G	A	A	A	A
zag15	T/C	T	C	C	T	C	T	T	T	C	T	C	T	T	T	T	T
zag35	T/C	T	C	C	T	C	T	T	T	C	T	C	C	T	T	T	T

单元型F¹ 0.285 0.274 0.161 0.161 0.07 0.006 0.006 0.006 0.006 0.005 0.005 0.005 0.005 0.003 0.003 0.003 0.003

F 是频率。对每个标记分别地进行 Hardy-Weinberg (H-W) 平衡检验。以 5% 的 I 型错误 (type I error)，未发现显著偏离 H-W 平衡。符合单元型频率评估条件 (Zhao 等人, 2003)。

在白种人的其他人类基因中通常观察到 *AZGP1* 的此单元型特征: 少数常见单元型 (在此为 5 种)，以及一系列稀有单元型。

鉴定为与肥胖症相关的标记 (zag15、zag17 和 zag35) 的等位基因以杂合状态存在于 L3 和 L21 纤瘦的受试者中 (参见表 4)。这些等位基因在这两个具有最低 *AZGP1* 基因表达水平的受试者中的存在，提供了 *AZGP1* 在肥胖症状中具有重要性的一些证据。这个观察结果得到实施的基因组研究的支持，该研究清楚地表明当察看受试者的整个基因表达谱时，L3 和 L21 受试者接近于肥胖的受试者。

表 4: *AZGP1* 中鉴定的标记 (与肥胖症相关) 的特征和在纤瘦的和肥胖的患者中的 *AZGP1* mRNA 的表达水平 (参见 EP 1548445)

受试者	zag17	zag_del	zag16	zag15	zag35	<i>AZGP1</i> mRNA 表达水平
L3	AG	del/wt	AG	TC	TC	54
L7	GG	wt/wt	GG	CC	CC	257
L8	GG	wt/wt	GG	CC	CC	223
L10	GG	wt/wt	GG	CC	CC	115
L11	GG	wt/wt	GG	CC	CC	69
L17	GG	wt/wt	GG	CC	CC	98
L21	AG	del/wt	AG	TC	TC	41
O2	AG	del/wt	AG	TC	TC	18
O9	AG	del/wt	AG	TC	TC	10
O13	AG	del/wt	AG	TC	TC	8
O14	AA	del/del	AA	TT	TT	13
O15	AA	del/del	AA	TT	TT	12
O16	AG	del/wt	AG	TC	TC	9
O18	AA	del/del	AA	TT	TT	4
O19	GG	wt/wt	GG	CC	CC	14
20	GG	wt/wt	GG	CC	CC	6

表 5 中显示了从每个 Fisher 氏检验获得的 p 值。

表 5. 每个 SNP 和肥胖症状态之间的相关性结果

SNP	p 值 (显性编码)	p 值 (隐性编码)
zag04	1	0.476
zag05	0.361	0.361
zag07	0.198	0.214
zag10	0.08	0.476
zag14	0.361	0.361
zag15	0.311	0.03
zag17**	0.03	0.311
zag18	1	—*
zag19	1	1
zag23	0.361	0.361
zag35	0.311	0.03

*: 无信息编码, 因为所有的 21 个个体属于相同的类型。

**: zag16 和内含子序列缺失 (zag_del) 没有在表格中显示, 因为它们与 zag17 是完全冗余的 (参见表 3)。

三个标记具显著性: zag15、zag17 (其代表 zag16 和 zag_del) 和 zag35。

因此, 来自 *AZGP1* 的标记簇 zag15、zag16、zag17、zag_del 和 zag35 在来自 Oestensson 组的样品中与肥胖状态相关。由于这 5 个标记强烈的相关 (参见表 3), 这与它们提供相同强度的证据相符。

因此, 本发明提供了分离的核酸, 其除了在下文显示的位置之一上存在的单一多态性改变外, 包含 SEQ ID No. 2, 或者含有位置 8047、8077-8083、8500、9556 或 12002 的 SEQ ID No. 2 的片段:

在位置 9556 的 zag15, 其中在此位置的 T 被 C 置换,

在位置 8500 的 zag16, 其中在此位置的 A 被 G 置换,

在位置 8047 的 zag17, 其中在此位置的 A 被 G 置换,

在位置 8077-8083 的 zag_del, 其中在这些位置的核酸发生缺失,

在位置 12002 的 zag35, 其中在此位置的 T 被 C 置换。

这些多态性是确定个体肥胖症素质的方法的基础, 该方法包括下述步骤: a) 从分离自患者的样品中分离核酸, 和 b) 检测存在于如前文列出的 Seq ID No. 2 中的一个或者多个多态性位点上的核苷酸, 其中在前文列出

的多态性位点上存在的特定核苷酸指示着患者患肥胖症的倾向。

前文公开的多态性确定了在 *AZGP1* 基因中与肥胖症相关几种单元型。因此，本发明也提供了选自单元型 1 的分离的核酸分子，其中单元型 1-3 各自分别包含 SEQ ID No. 2，只是下表 6 中针对每个单元型所示的核苷酸存在于 Seq ID No. 2 的相应位置上：

表 6：目的标记的单元型

位置	单元型 1	单元型 2	单元型 3
8047	A	G	G
8077-8083	del	wt	wt
8500	A	G	G
9556	T	C	T
12002	T	C	C

如本文使用的，术语“del”涉及衍生自 Seq ID No. 1 的序列，其中 Seq ID No. 2 中的 8077 至 8083 的核酸从 Seq ID No. 1 中的相应位置上缺失。术语“wt”涉及衍生自 Seq ID No.2 的序列，其中存在从位置 8077 至 8083 的核酸。

而且，本发明还提供了用于对个体的 *AZGP1* 基因进行单元型分型的方法，包括下述步骤：a) 从分离自个体的样品中分离核酸；b) 确定存在于个体 *AZGP1* 基因拷贝的位置 8047、8077-8083、8500、9556 和 12002 上的核苷酸的存在，其中位置编号通过与 Seq ID No.2 比较而确定；c) 通过将存在于所述位置上的核苷酸与前文列出的在表 6 中的针对单元型提及的核苷酸比较而确定个体的特定单元型。优选地，存在至少一个表 6 中列出的单元型指示着个体有患肥胖症的倾向。

通过高密度寡核苷酸微阵列检测分离自 7 个纤瘦者和 9 个肥胖者的脂肪活组织检查样品中的 ~5000 个不同基因的表达水平。这组成它们的基因表达谱。然后在这些基因表达水平上进行对应的分析 (Benzecri JP. *L'analyse des données*. Dunod, Paris; 1979; Greenacre M. 对应分析的理论和应用 (Theory and application of Correspondence Analysis),

1984; Academic Press, London; Fellenberg K, Hauser N, Brors B, Neutzner A, Hoheisel JD,和 Vingron M. 应用于微阵列数据的对应分析 (Correspondence analysis applied to microarray data), *PNAS* 1998:10781-86)。如通过对应分析检测的, 图 2 中每个数据点代表一个受试者的整个基因表达谱在三维空间中的投射。受试者之间的距离反映了他们整个基因表达谱之间的距离。

除了 O16 患者, 所有肥胖的受试者都定位在 F3 轴的右侧, 而纤瘦的受试者定位于此相同轴的左侧 (除了四个纤瘦的受试者——L3、L11、L17 和 L21, 他们出现在肥胖的受试者中间)。

通过实施的统计学工作, 当比较肥胖的受试者与纤瘦的受试者时, 发现许多差异表达的基因。属于这些差异表达的基因之一的 *AZGP1* 基因, 相比较于纤瘦的受试者, 表现出在肥胖的受试者中的表达下调 (参见图 2, 倍数变化 (fold change) = -11.5, P 值 <5%)。具有最低的 *AZGP1* 基因表达水平的纤瘦的受试者 (L3 和 L21) 也是在图 2 中显示与肥胖的受试者接近的那些。同样的这两个纤瘦的受试者的临床参数显示他们躯干脂肪的百分比比展现出高水平的 *AZGP1* mRNA 的纤瘦的受试者的百分比高。相比较于在其他纤瘦的受试者中观察到的数值, L21 还具有非常低数值的能量消耗。

附图说明

图 1: 在用于发现 *AZGP1* 的 SNP 的基因组序列上作图定位的目的标记。下述序列来自 EMBL 检索号 ac004977。目的标记被加强显示 (统计学分析中描述为 SNPs 和缺失)。在这个序列中存在有 zag_del 缺失。

图 2: 在使用高密度寡核苷酸微阵列检测到的 7 个纤瘦的和 9 个肥胖的受试者的整个基因表达谱上进行的对应分析。每个数据点对应于一个受试者的整个基因表达谱。纤瘦的受试者用黑方格表示, 并且肥胖的受试者用灰方格表示。使用统计学软件包 X1Stat6.0(Addinsoft; New York, NY)进行此分析。

图 3: 使用高密度寡核苷酸微阵列检测的 *AZGP1* 表达谱(参见表 4 中的值)。

实施例:

实施例 1: DNA 样品

用于发现 SNP 的 DNA 样品来自两个不同的来源:

- 大部分样品直接作为 DNA 样品购自 Coriell Institute for Medical Research。
- 二十一份样品在 RCMG 从 Claes Oestenson 教授提供的全血中制备(参见 EP 1548445)。在采取样品的时候所有的患者都无糖尿病。使用基于硅胶的提取方法从 200 μ l 全血中提取 DNA (MagNA Pure LC DNA 分离试剂盒 I, Roche Molecular Biochemicals)。

实施例 2: 发现 SNP

在基因组序列(EMBL 检索号 ac004977, LocusLink 563)上对 *AZGP1* 的 mRNA 序列(NCBI 检索号 NM_001185)作图,以鉴定 *AZGP1* 的基因组组织方式(外显子和外显子/内含子边界)。设计引物以扩增 DNA 片段,这些片段将覆盖整个基因序列和额外的 *AZGP1* 起始密码子(ATG)上游 1.5 kb 和 *AZGP1* 终止密码子(TAG)下游 1 kb(表 7)。这些片段互相重叠。使用来自几个个体的 DNA 样品作为模板,通过 PCR 扩增片段。扩增条件如下,终体积为 20 μ l:

- 4 ng DNA
- 1x 缓冲液(详细内容参见表 8)
- dATP, dCTP, dGTP 和 dTTP 各 50 μ M
- 每种引物各 0.4 μ M
- 0.4 u 的聚合酶(详细内容参见表 8)

扩增反应在具有试样自动等分设备(Tecan biorobot)的 96 孔扩增板上准备。设立自动等分设备的操作参数以最小化交叉污染的可能性。热循环条件如下: 95 $^{\circ}$ C, 15 分钟用于 DNA 聚合酶激活,下述步骤的 35 个循环:

94℃变性 1 分钟，退火温度（表 8）下杂交 30 秒，72℃延伸 1 分钟，并且最终的延伸步骤为 72℃，5 分钟。扩增反应在 MJ Research PTC-200 DNA Engine 上运行。PCR 扩增之后，使用在 Tecan biorobot 上的 384 Cleanup Millipore 板纯化片段。根据厂商的说明书，使用 ABI Big 染料终止化学进行所有片段的双链 DNA 测序。用于测序的引物与用于片段扩增的引物相同。在 MJ Research PTC-200 DNA Engine 上实施测序反应并在 ABI 3730 测序仪上测序。测序后，使用 Polyphred 软件（得到华盛顿大学的授权许可）进行多态性分析。表 3 列出了在 *AZGP1* 中鉴定到的所有标记。这些标记在 *AZGP1* 基因组序列中的位置也在图 1 中被突出表示。

表 7 用于扩增和测序 *AZGP1* 的引物

引物名称	引物编号	开始的位置 1	结束的位置 ¹	序列
AZGP1-5Reg-F	3	1941	1961	GTCCAAAAACACACAAATGCC
AZGP1-5Reg-R	4	2441	2422	TTCCTCACCTCCTTCCAGTC
AZGP1-5Regc-F	5	694	713	TCCAACCAACAGCATGTAAG
AZGP1-5Regc-R	6	1539	1520	CCCTCCGAATACAAAGCAAC
AZGP1-ex01-F	7	2290	2309	AGAACCCTCCAAGCAGACAC
AZGP1-ex01-R	8	3107	3084	GGCACAGAATCAGATTAACATTCC
AZGP1-ex02-F	9	5815	5835	TTCTAACGCATGTCAGATTCC
AZGP1-ex02-R	10	6678	6658	CTATTTCATCCTGCTGATCC
AZGP1-ex03-F	11	9685	9704	TGAGACAAACCTGAAATGCC
AZGP1-ex03-R	12	10191	10173	AAGCAGTGAGTACCTTGCC
AZGP1-ex04-F	13	10614	10633	AAGAGCAAGCCAGTGTGAGC

AZGP1-ex04-R	14	11474	11453	AAATCCACCTCCTGTCTGTCCC
AZGP1-in03-F	15	10042	10060	AGCAGCCCAGATAACCAAG
AZGP1-in03-R	16	10832	10812	GCAATAAGTTGTGAATGCTCC
AZGP1-in02-F	17	8970	8989	GCTCACTACAACCTTCTGTCC
AZGP1-in02-R	18	9822	9801	GGCAACCCAAAAGAAATAAAGG
AZGP1-in02b-F	19	8245	8263	AGTTCAGGCAACACACCAG
AZGP1-in02b-R	20	9108	9089	GGCCAACATGGTAAGACCTC
AZGP1-in02c-F	21	7568	7587	GGCAAGAAAGAGATAGGCAG
AZGP1-in02c-R	22	8432	8412	CCACAACCTCTCAGAAATGGAC
AZGP1-in02d-F	23	6945	6964	AGCCACTCTCAAAGTCACTC
AZGP1-in02d-R	24	7798	7777	AGCCCTGCCTTCTATTATTTTC
AZGP1-in02e-F	25	6474	6493	ACAGGTGGAAGGAATGGAGG
AZGP1-in02e-R	26	7156	7137	TAGGTGATGGAGCTGCAAGG
AZGP1-in01-F	27	5094	5115	CTTACCCTGTGCTAATTCAGTC
AZGP1-in01-R	28	5967	5947	GTCCCTTTTGTCTTCTCATCCC
AZGP1-in01b-F	29	4767	4786	TACCCATTAACCACCCTCCC
AZGP1-in01b-R	30	5547	5527	GCTTGATGACAGAGTGAGAC
AZGP1-in01c-F	31	4087	4106	GGATTCTTGTTCTGTCACCC
AZGP1-in01c-R	32	4916	4895	CTTGCTCCTGAGTGTCTAAATG
AZGP1-in01d-F	33	3464	3483	GGATGAAGCCCACCACTATG
AZGP1-in01d-R	34	4272	4253	GGTCAAGAGGTCAAGACCAG

AZGP1-in01e-F	35	2837	2856	CCCAAATCCCACACTCAGAC
AZGP1-in01e-R	36	3652	3633	AGCTTGAAGGGATGGATACC
AZGP1-3REG-F	37	11199	11219	CACAATGGAAATGGCACTTAC
AZGP1-3REG-R	38	11990	11971	GATATTCTCCCCTCCCCAAC
AZGP1-3REGb-F	39	11780	11763	TGAACCCCTTTCCCTTG
AZGP1-3REGb-R	40	12500	12483	ATCTTCCTCTCCCCCTG

¹ 在 SEQ ID No. 1 中的位置

表 8 所有片段的扩增条件

片段	引物编号	退火 温度	缓冲液	聚合酶
5Reg	3/4	58	FastStart 缓冲液 (Roche)	AmpliTaqGold
5Regc	5/6	60	FastStart 缓冲液 (Roche)	AmpliTaqGold
ex01	7/8	62	FastStart 缓冲液 (Roche)	AmpliTaqGold
ex02	9/10	62	FastStart 缓冲液 (Roche)	AmpliTaqGold
ex03	11/12	58	FastStart 缓冲液 (Roche)	AmpliTaqGold
ex04	13/14	62	FastStart 缓冲液 (Roche)	AmpliTaqGold
in03	15/16	60	FastStart 缓冲液 (Roche)	AmpliTaqGold
in02	17/18	60	FastStart 缓冲液 (Roche)	AmpliTaqGold
in02b	19/20	62	Roche 缓冲液 ²	AmpliTaqGold
in02c	21/22	60	FastStart 缓冲液 (Roche)	AmpliTaqGold
in02d	23/24	60	FastStart 缓冲液 (Roche)	AmpliTaqGold
in02e	25/26	62	FastStart 缓冲液 (Roche)	AmpliTaqGold
in01	27/28	62	FastStart 缓冲液 (Roche)	AmpliTaqGold
in01b	29/30	62	FastStart 缓冲液 (Roche)	AmpliTaqGold

in01c	31/32	62	FastStart 缓冲液(Roche)	AmpliTaqGold
in01d	33/34	62	缓冲液 D (Invitrogen)³	Taq
in01e	35/36	60	FastStart 缓冲液(Roche)	AmpliTaqGold
3Reg	37/38	62	FastStart 缓冲液(Roche)	AmpliTaqGold
3Regb	39/40	58	FastStart 缓冲液(Roche)	AmpliTaqGold

¹FastStart 缓冲液 1x: 50 mM Tris-HCl, 10 mM KCl, 5 mM (NH₄)₂SO₄, 2 mM MgCl₂, pH 8.3, 25°C

²Roche 缓冲液 1x: 10 mM Tris-HCl, 50 mM KCl, 1.5 mM MgCl₂, pH 8.3, 25°C

³缓冲液 D 1x: 30 mM Tris-HCl, 7.5 mM (NH₄)₂SO₄, 3.5 mM MgCl₂, pH 8.5, 25°C.

实施例 3: 单元型频率的评估方法

如在 Genecounting 中实施的方式, 用 E-M 算法对单元型频率评估 (Zhao JH, Lissarrague S, Essioux L, Sham PC. GENECOUNTING: 具有缺失基因型的单元型分析 (GENECOUNTING: haplotype analysis with missing genotypes), Bioinformatics. 2002 Dec, 18 (12):1694-5)。该程序考虑具有未分型位点(untyped sites)的个体, 并且因此提供更准确的评估。

实施例 4: 在 10 个肥胖者/11 个纤瘦者中有关缺失发现的分析

自 Oestenson 教授的组 (EP 1548445) 的 10 个肥胖患者和 11 个纤瘦者样品中测序 *AZGP1* 基因组序列。表 3 中的所有常见 SNPs 都存在。在 11 个非冗余的常见 SNPs (zag05、zag04、zag07、zag10、zag14、zag23、zag18、zag19、zag17/zag16/zag del、zag15 和 zag35) 中进行肥胖状态与基因型之间的相关性分析。

比较于前面的分析, zag18 可以与 zag19 分开。因此其被处理为两个

非冗余的 SNPs。

使用两个编码方案：

- 显性编码 (dominant coding)，其中较稀少的等位基因的杂合子和纯合子被合并在一个类型中。
- 隐性编码 (recessive coding)，其中最多的常见等位基因的杂合子和纯合子被合并在一个类型中。

在每种编码方案下，每种基因型变量均是二元变量。对于产生的每个变量，实施确切 2x2 fisher 检验。显著性阈值取 0.05。

实施例 5：在纤瘦和肥胖的受试者中制作 *AZGP1* mRNA 谱

从来自 EP 1548445 中描述的组的二十一个受试者中获得皮下脂肪活组织检查样品。对于 5 个受试者 (L1、L5、L12、O2 和 O6)，未能使用对应的活组织检查样品实施微阵列分析。

使用 Affymetrix (Affymetrix GeneChip® Technology; Affymetrix, Inc.; Santa Clara, CA) 提供的高密度寡核苷酸微阵列基因技术，对剩余的十六个样品进行基因表达研究。

实施例 5.1：RNA 制备

根据厂商的说明，使用 TriZol 试剂 (Life Technologies) 和 Fast RNA green (BIO101) 试剂盒分离 500mg 皮下脂肪组织的总 RNA。使用 RNeasy 试剂盒 (Qiagen) 将总 RNA 从污染的 DNA 中纯化。

实施例 5.2：通过高密度寡核苷酸微阵列进行基因表达谱制作

使用 SuperScript Choice Gene Chip 试剂盒 (Life Technologies) 和来自 Gibco 的试剂进行 cDNA 第一链和第二链的合成。通过使用酚：氯仿：异戊醇的混合物 (v/v/v. 25/24/1, Life Technologies) 抽提而纯化含有加入的 T7 RNA 聚合酶结合位点的双链 cDNA。根据厂商的说明书，使用 Phase Lock Gel (Eppendorf) 分离有机相和水相，并且通过沉淀回收双链 cDNA，

然后重悬于水中。使用 T7 试剂盒 (Ambion) 和含有生物素的核糖核苷酸 (Enzo - LOXO GmbH), 通过体外转录 (IVT) 将双链 cDNA 转变为生物素标记的 cRNA。使用 RNeasy spin 柱 (Qiagen) 将 IVT 材料从未掺入的核糖核苷酸中纯化出来。提纯之后, 在 500 mM 乙酸钙, 150 mM 乙酸镁, pH 8.1 中, 95°C, 35 min, 将单链生物素标记的 cRNA 化学水解为较小的片段。通过在冰上冷却样品而终止反应。

每个样品与一个 U95-A Affymetrix GeneChip 微阵列杂交。每个微阵列含有代表~10,000 个基因的 12559 个探针集合。使用 GeneChip Fluidics Station (Affymetrix Inc.; Santa Clara, CA) 进行所有的洗涤、杂交、检测和信号扩增步骤。使用 GeneArray 扫描仪 (Affymetrix Inc.; Santa Clara, CA) 从杂交的 GeneArrays 收集荧光强度数据。使用 MAS 5.0 算法 (GCOS 1.0 软件的组成部分) 将含有荧光强度信息的原始文件转化为数据文件。仅仅有 45% 的作图于微阵列上的基因用于本分析, 因为其余基因被 MAS 5.0 算法称为不存在。使用 Roche Affymetrix 芯片实验分析 (RACE-A) 软件鉴定差异表达的基因。

实施例 6: zag14, zag15 和 zag16 的基因型分型

实施例 6.1: 组的描述

研究了 86 个葡萄糖耐量降低 (IGT) 的肥胖男性患者和 290 个葡萄糖耐量正常 (NGT) 的, 具有正常 BMI ($\text{BMI} < 25 \text{ Kg/m}^2$) 的男性对照受试者。所有的都是瑞典男性患者, 选择自斯德哥尔摩糖尿病预防计划 (Stockholm Diabetes Prevention Program)。IGT 肥胖受试者具有正常的出生体重, 正常的 BMI ($< 25 \text{ Kg/m}^2$) 并在口服葡萄糖耐量实验后 2 小时具有正常的血浆葡萄糖水平。如 Gu 等描述的进行血浆葡萄糖和血浆胰岛素的浓度和其他临床特征检测 (在瑞典白种人中脂连蛋白 (APM1) 基因的近端启动子区的单核苷酸多态性与 2 型糖尿病相关 (Single nucleotide polymorphisms in the proximal promoter region of the adiponectin (APM1) gene are associated with type 2 diabetes in Swedish

caucasians), *Diabetes* 53 Suppl 1: 31-5, 2004)。从所有的受试者获得知情同意, 并且本研究得到当地伦理委员会的认可。使用 Puregene DNA 纯化试剂盒 (Gentra) 从外周血提取基因组 DNA (Gu 等, 如上)。

实施例 6.2: PCR-动态等位基因特异杂交实验 (DASH) 设计和基因型分型

使用称为动态等位基因特异杂交 (DASH) 的高通量 SNP (单核苷酸多态性) 打分技术进行 SNP 基因型分型 (Howell 等人, 动态等位基因特异杂交: 一种新的对单核苷酸多态性打分的方法 (Dynamic allele-specific hybridisation: a new method for scoring single nucleotide polymorphisms), *Nat Biotech* 17: 87-88, 1999)。如在先文献描述, 进行 PCR-DASH 实验设计和 SNP 基因型分型 (Prince, 等人, 通过动态等位基因特异杂交 (DASH) 的可靠和准确的单核苷酸多态性基因型分型: 设计标准和实验的证实 (Robust and accurate single nucleotide polymorphism genotyping by dynamic allele-specific hybridization (DASH): design criteria and assay validation), *Genome Res* 11: 152-162, 2001)。

实施例 6.3: 统计学分析

统计学分析的目的是确认前面的结果: 对于 zag15 遗传多态性, 与纯合子 CC 患者比较, 纯合子 TT 患者和杂合子 CT 患者具有更高的成为 IGT 肥胖的风险。形成 2 乘 2 列联表。统计学检验假设是使用单侧备择假设 (unilateral alternatives hypotheses):

虚假设 (H_0): $p_1 = p_0$

备择假设 (H_1): $p_1 > p_0$

参数 p_1 和 p_0 分别是 IGT 肥胖患者和对照中在 zag15 携带至少一个 T 等位基因拷贝的患者的比例。用于比例比较的统计学检验是基于经反正弦转换后的比例的正态性 (normality of the arcsinus-transformed

proportions)。在虚假设下，该检验遵循正态分布 $N(0,1)$ 。也加入比例的确切检验 (Agresti, *Categorical data analysis*. New York: Wiley, pp. 59-66, 1990)。

以 5 % 的 I 型错误进行此检验。

计算了与 SNP zag15 位置的受试遗传特征相关的发展为葡萄糖耐量降低和肥胖症的比值比 (OR)。使用免费的统计软件 R 计算对应的 95 % 置信区间。

下表显示了 zag15 的每个基因型在两组患者之间的分布。

	肥胖的 IGT	正常的 NGT	总数
TT	28	67	95
CT	40	132	172
CC	18	87	105
总数	86	286	372

相比较于对照组中的 TT 和 CT 患者的比例是 0.7, 在肥胖的 IGT 组中 TT 患者和 CT 患者的比例是 0.79。携带至少一个 T 等位基因拷贝使成为 IGT 肥胖的比值增加 1.65 (95 % CI: [0.93; 2.94])。遗传模型和肥胖 IGT 状态之间无关的虚假设被拒绝，与之相对的是在 5 % 的水平下在肥胖的 IGT 群中较高的 TT/CT 患者比例 ($z=1.77$, $p=0.04$)。使用确切比例检验 (Agresti, 同上)，该结果是具有临界的显著性 ($p=0.055$)。

如实施例 1-5 中所描述的，利用根据这个来自相同人种和预防研究的扩展组，zag15 和肥胖症的葡萄糖耐量降低之间的遗传相关性被证实。考虑到多态性 zag15、zag16、zag_del、zag17 和 zag35 之间完全的遗传等价性，对于在这个簇 (即 zag16、zag17、zag_del 和 zag35) 中的所有多态性，此相关性都是正确的。

<110> 弗·哈夫曼-拉罗切有限公司 (F. Hoffmann-La Roche AG)

<120> AZGP 基因单核苷酸多态性 (SNPs)

<130> 23228

<160> 28

<170> PatentIn version 3.3

<210> 1

<211> 12550

<212> DNA

<213> 人 (Homo sapiens)

<220>

<221> 包含 AZGP1 的序列

<222> (1) .. (12550)

<223> EMBL 检索号 ac004977 的核苷酸 1 至 12550

<400> 1

```

gatcagataa ttgttttagtt atgccttcaa aaatgtctga gccaggaaca gtggataaat      60
gagcttgtga atcctcgaaa atttgctctt taagttttga aatatccaag gttaagttat      120
catcccaggc ttttaaagtgt cttgagacat tttcccagct atgttgatat ttattataag      180
cataaggcat tatgcaataa tcagaagtat tccaatcact ctgtaattgc atacggtggt      240
ccaaattcat atctcccagc cagattacac tttggcggag atcattaatt tgattagcta      300
atTTTTgatc aatttgagcc tgagaattcc aaagtctgga ggagtttttc tgccatgctt      360
caattcaaca tattgaacgg tttgaacaga attgtggatg gcaactccag cagttgctgc      420
tgttgcagta accgcaatta gtcctgcaat gatggcaata agagtaaaaa taaatctctt      480
tgttctttttg cggatacctt taagaacttc attgactatg tgaatagagg gggaagactc      540
ccatggacga tgtaaagaaa ctggtatcca taccctctcc ctagccctta ccaagagagt      600
acttgtagtg ggattaaaag tagcatcaat gcacgtgaac agcttacagt tatcacattc      660
tatagtttgt gtattgggaa tgataattat atttccaacc aacagcatgt aaggggggtt      720
gacatagctc ctgatgggta tcaccggttc agatatgagg gtgatgttga atgtgggtgt      780
tttgggtatta gtgtgaagga gttgataggt agtgttccat atccttattc ctgtcatagc      840
tgcagctaat tttcatagtt gaggatgttc tggggtaaca aagggatgaa tcatttttgg      900
tctaggagga gtaatgcctg catccatcca tttaaatagg taaggggaca cccattgtct      960
cagcctgtat gactgccatc catcctctac ataatctaac aaataattaa actctgagca     1020
tgaggtatTT ttgctggagc aatcttgcca ataatgcctt tttggagccc aatcaataac     1080
aatacctgcg gctagacttt ggagcacaac ggccttagga gcattacaat tattccataa     1140
aattgaagtc actataaaag gtccctttgt aggtttcttt aaccaaaga gaggctgggc     1200
acagttgctc acacttgtaa acccagcact ttgggagacc gaggcgggca gatcacgagg     1260

```

tcaggagttc	aaaaccagct	tggccaacat	ggtgaaaccc	ctgcctcagc	ttcctgagtg	1320
ctgggattat	aggcatgtgc	caccacgcgc	agctaattct	tgtatttttt	cagtagagac	1380
gaggtttcat	catgttgccc	aggetggtct	cgagtacctg	acctcaacct	gaggtgatcc	1440
aaccacctca	gcctcccaaa	gtgctgggat	tacaagcatg	agtcaccgca	cccgggcccc	1500
agtcacttta	gaatagcatg	ttgctttgta	ttcggagggt	ctctctgcaa	atagcccatc	1560
aacactgagc	gtgcctggaa	agacctggtt	ttcaaataac	tggcttcgtc	tgtgtaaaaac	1620
gagtcttggt	gtatgcatta	aaaattatct	tggctgggcg	tggtaggtca	cgcctttaat	1680
cccagcattt	tgggaggctc	gttctgttgc	cagggaggct	gcaaaaggag	gatgctggcta	1740
tttctgcttc	catcatggaa	ggctgggatt	ccacaggcac	agaaaggaag	gagggagaga	1800
tgggaatgtg	actgtcctcc	agacacagcc	tctggcagga	tcggatgtgg	gagagtgagg	1860
gtcccccccc	agctgggggc	taccacaggc	catgtcttgg	acatgttgag	agtttttctg	1920
gaaggcaggg	atacagtgtg	gtccaaaaac	acacaaatgc	ccctactggc	ccaggggttg	1980
tcacaataga	ctggaagggt	gacacatccc	aggcgcttgc	cacccatcac	acgcacctcc	2040
taccacttgg	catccttcca	ccccaggcac	acacaaagcc	tcagtccaga	gatcaactct	2100
ggactcagct	ctgaatttgc	atatcctgtg	tgtagattca	ttcttcataa	cctctgcccc	2160
gcctagcttg	tgtatcattt	ttttttctct	attaggggag	gagcccgtcc	tggcactccc	2220
attggcctgt	agattcacct	cccctgggca	gggccccagg	accagagata	atatctgtgc	2280
ctcctgcccc	gaaccctcca	agcagacaca	atggtaagaa	tggtagcctgt	cctgctgtct	2340
ctgctgctgc	ttctgggtcc	tgctgtcccc	caggagaacc	aagatggtga	gtggggaaag	2400
caagggatgg	gtgctggaga	ggactggaag	gaggtgagga	acaggacatg	tggctgggag	2460
acaggctgga	tgcagctggg	ataccctggc	atacggcagg	aatgggtgcc	caaggctgtc	2520
aactccctca	gctcacacac	ttccaggagc	attcaggagg	cctctgcgct	ggccccgaaat	2580
aagaccttca	ggaatctgaa	tctaaaaccc	ctagtttaca	gtgaaaacaa	agactccaaa	2640
gaccaagcga	cctgcttggg	gtagacagtc	aggacggagt	aggaaccata	tgcctggagc	2700
tgcttctgct	cctgttcctt	ccctccttcc	gatggctggg	tacacctgcc	tgacgctgat	2760
gaaaagagag	agcagcccca	aggggaaagt	gggaaggcag	gttggctgga	gggatggtgc	2820
tagaaggaaa	cccgtgcccc	aatcccacac	tcagacacca	ctgcagtggg	tctggaaggc	2880
gagtggctgg	aagagaagag	agtgggagct	ccgggagatc	aagagtcact	cctaggataa	2940
gggaaggagg	ctgtttgttg	catgagaatg	tgcaggataa	agacatggaa	gcgaatggct	3000
tctcagttgt	gtgagtttaa	aattcatgac	atttacaat	tgtcagaaaa	ggtgttatat	3060
gtttgttata	taacaatcac	tttggaatgt	taatctgatt	ctgtgccaaa	atctgaatta	3120
ctcagggttc	tccagagaaa	cagaactaat	aggtggtaca	catatacata	tatatgtacg	3180
tacacataca	tacatacact	gtatacacat	ggatacacac	acacatagga	agagatttac	3240
atatatgtat	acaaaagaga	gagagagtag	agattttattt	taagaaattg	actcacacta	3300
ttgggaggag	taacaagtcc	taaatcttca	gagccggcca	gcaggctgga	gacccaggga	3360
agagttgatg	tcttagtctt	gattccaagg	gcagactgta	ggcagaattc	tttcctcttt	3420
aggggacatc	tgaggctttt	tctcttaagg	ccttcaactg	attggatgaa	gcccaccact	3480
atggagagta	atccacttta	ctcaaggctc	actgattttt	ttgtaaatta	aaaaaaaaac	3540

tgtgggtgca	tagtatgtgt	atatatttat	ggggtacatg	agaggttttg	attcaggcat	3600
gcaatgtgaa	ataatcacat	catcaaaaat	gaggtatcca	tcccttcaag	cttttatcgt	3660
ttgtgttaca	gacaatccaa	ttatactttt	ttgggtattt	tagtttttaa	aagtatttga	3720
ttatttattt	atttatttat	ttttgagaca	gagtctcact	ctgtcaccca	ggcaggagtg	3780
cagtggcatg	atctcggtc	actgcaacct	ccgcctccca	ggttcaagca	attttccctgc	3840
ctcagtctcc	tgagtagcta	ggactacagg	cacctgccac	cacacctggc	taattttttt	3900
gtattttttag	tagagacggg	tttcatcatg	ttggccaggc	tagtcttgat	atcctgacct	3960
cgtgatctgc	ccgccttgg	ctcccaaagt	gccgggatta	caggtgtcag	caactgcgcc	4020
tggcctctct	tttggttatt	taaaagtgtg	caattaaatt	atgattatta	ttattatttt	4080
tgagatggat	tcttgttctg	tcaccaggc	tggagtgcag	tggcgtgatc	ttggcttact	4140
gcaaacctcc	gcctgttggg	ttcaagcaat	tatcttgcc	cgggtgtaca	ctgccacaca	4200
cggctaactt	atgtattttt	aatagagata	gggtttcacc	atgttggcta	gactggctct	4260
gacctcttga	cctcaagtga	tccactcact	tcagcctccc	agagtgcctg	aattacaggc	4320
acgagccacc	acacctggcc	ccagttaaat	tattattgac	tatagtcacc	ctgttgtgct	4380
atcaaatagt	aggtcttatt	cattcttctt	tttttttttt	tttttgtgac	agagttgccc	4440
aggctggaat	gcagtgggtg	aatcttggct	cactgcaacc	tctgcctccc	gggcttaagc	4500
gattctcctg	cctcagcctt	ctgagtcgct	gggactacag	gtgtgtgcca	ccacgcccgg	4560
ctaatttatg	tatttttagt	agagatgggg	tttcaccatg	ttggccaggc	tggtttcgaa	4620
ctcctgacct	caagtgacct	acctgcctca	gcttcccaaa	gtgttggaat	tacaggcatg	4680
agccaccaca	cctggcccca	gttaaattat	tattcactgg	agtcactttg	ttgtgctatc	4740
aaatagtttt	ctaactattt	tttttgtacc	cattaaccac	cctcccaatt	ccccccaac	4800
cctgccacta	cccttcccag	cctttggtaa	ccatccttct	actctctatg	tccatgaatt	4860
caattgtagg	gtctactgat	ttaaaggcta	atcacattta	gacactcagg	agcaagaata	4920
attttagtaa	ttgaactagg	attctgccat	atgacctcca	acatcattag	cacctgtgta	4980
aattgtatca	taaaataatt	atggaactat	tatggaaatg	tccctctctc	ccagatccca	5040
ccttgtacca	aatgcaagg	tacaaccccg	ggaattctga	gtccatcct	agtcttacct	5100
tgtgctaatt	cagtctgggt	catttcttga	attttctgg	aaattctcct	ttctaccctt	5160
tctaactata	tgtatttgtc	aggttaagct	agaagtgtta	attttttttt	tttttgagat	5220
ggagccttgc	tttgtcacct	aggctgaagt	gcagtggcat	gatctcagct	cactgcaagc	5280
tccgcctccc	gggttcatgc	cattctcctg	cctcagcctc	ctgagtagct	gggactacag	5340
gcacccgcca	ccatgcttgg	ctaattttgt	gaattcttag	tagagacggg	gtttcaccat	5400
gttagccagg	atggtctcga	tctcctgacc	tcgtgatcca	ccgcctcgg	ccccctaaag	5460
tgctgggatt	acaggcgtga	gccactgagc	ccggacgaaa	tgtaattttg	ttttttttga	5520
gacggagtct	cactctgtca	tccaagctgg	agtgcagtgg	catgatcttg	gcttgttgca	5580
acctctgcct	ctctggttca	agtgattttc	ctgcctcagc	ctccagcatg	actgggatta	5640
caggcccgc	ccaccatgcc	cagctaattt	ttgtattttt	taatagagat	ggggtttcac	5700
catgttggcc	aggctgggtc	tcaactcctg	atctcaagta	atctgcctgc	cttggcctcc	5760
caaagtccctg	ggattacagg	catgagccac	ggagcccagc	ctagaaatgt	taatttctaa	5820

cgcatgtcag	attccatgca	cactgggcaa	ggttccattc	ctccatgggg	tgactcaggg	5880
atccaggcca	attgcatatt	gagactcttt	catattatcc	tgtggccttc	aaagtcgtca	5940
cctctagggg	tgagaaacaa	aagggacaag	ccagctggta	gggtcttgga	caagaagaaa	6000
gacatcactt	ctgctcacat	tctcttttga	caaaactcag	tcacatggtc	ccaatatatc	6060
ttcgaggtgg	ctgagtaatg	ttatcttcct	atgtgtcaag	cagaggaaat	aatgtagtga	6120
agacacagga	tgggtctctga	aatatcatct	caggcatgaa	agtagagcat	attcacttga	6180
gtgagcctcc	agtgggtgtga	agttgatggc	aggagaaaga	gctgggggaag	aaaaggccag	6240
tggcaggtct	ccccctctag	ccctatgcag	ccccacagtg	ggacccttgc	atggacctca	6300
accatcagaa	tcttttcttt	tgcaggtcgt	tactctctga	cctatatcta	cactgggctg	6360
tccaagcatg	ttgaagacgt	ccccgcgttt	caggcccttg	gctcactcaa	tgacctccag	6420
ttcttttagat	acaacagtaa	agacaggaag	tctcagccca	tgggactctg	gagacaggtg	6480
gaaggaatgg	aggattggaa	gcaggacagc	caacttcaga	aggccaggga	ggacatcttt	6540
atggagaccc	tgaagacat	cgtggagtat	tacaacgaca	gtaacgggtca	gtgaataaca	6600
gaccacaggg	gtggaaggtc	taaccaaga	ggcagccccc	ccagtgtgag	tggcaaggga	6660
tcagcaggat	ggaaatagtc	ccaatcccag	gggaagaaca	ggagacacag	cagaaacaca	6720
gacatgtcca	catcccaccc	acccacagc	acaggtgctc	cccgttccc	catcaattgc	6780
cccatectca	tcccaggcct	caggtcacac	aggaagtgat	ggcagagtca	cttccatatcc	6840
aggcacctat	gacctctcac	ctccacaccc	cacccatcgg	aggctgatac	ccccgtgaga	6900
aggcatcaga	ctcacccttg	tccagggagg	ttgcctggag	agtgagccac	tctcaaagtc	6960
actcagacct	gggctcacct	ggtggttctg	ccagtcctag	ctgttgacag	tgaaacgttc	7020
ccaaaatata	tggttgaaat	ctgcaaakat	tggagcactg	agacctacct	ccaaacaagt	7080
ctgtaatat	taactatgtc	tgttctatga	aggatgtcac	agtctgtcct	gatctccctt	7140
gcagctccat	cacctagcac	agggtacagc	caatattggc	tcaattgaaa	tttgtggaat	7200
ccacagagaa	aagcaccg	cacacaccgt	agcccatgct	gggggctcag	gaagtgtcgg	7260
attcaaaact	gtgggctggt	agagttcctt	ggagccctaa	agttcctcct	taccatacga	7320
tgcagaccca	ggaaggcca	cctgcgctat	ggtcagagga	gctgggtggca	gagcccgtgc	7380
agagatggtc	cctgtgcccc	cggcccagtg	ctctttctcc	taaaccacac	tgccagcccc	7440
aaggcagcca	acctcaggtc	tgggtgaactg	ctgggtgttaa	attatcatag	agtgggtgtc	7500
aaaagatggg	ctactaagta	caaaaatgcc	caagggtgcta	catgggatct	gaagattttc	7560
aaaaggaggc	aagaaagaga	taggcagatg	tttcaaggat	gtgggggtggg	ggaggtcctg	7620
gtaaggaaaa	tggcccaggc	tgtgtgtcag	caataggaga	ggagggggca	caggtgatca	7680
gaaaagacac	tgggggaagc	attgatggac	aggaatagaa	atggcaaagt	ggataattaa	7740
gaggaaggag	gatgaggaga	tgaacacagg	gtattagaaa	ataatagaag	gcagggcttg	7800
gtggctcact	cttgtaatcc	cagcactttg	ggaggctgag	gcaggcagat	cacctaaagg	7860
caggagtctg	agaccagccc	ggccaacatg	gtgaaacctt	gtctctacta	ataatacaaa	7920
aatagcctgg	catggtggca	cacgtctgtg	gtcccagcta	ctcaggaggc	tgaggcagga	7980
gaattgcttg	aaccagggag	gcagagggtta	cagtgaagcca	aaatcctacc	attgcactac	8040
agcctgagtg	acaagagtga	aacgttgtct	aaaaacaaaa	aacaaaaaaa	ggaaataata	8100

gtagctgaca	tttactgagc	acttactttg	tgccaggccc	atctatgagc	atatataatg	8160
ctcagaatag	ccccctaaaa	cagtgtctct	ggcattgccca	tttcagaggt	gaggaaatag	8220
aggcacaggg	agttgagtgg	ctccagttca	ggcaacacac	caggtggggg	tggggggctg	8280
gggagagacc	tgggacgtga	gcccagacag	cttgagagct	ttcagagtct	atgccaacag	8340
caccaaccag	tgtctgggtaa	acacctgctt	ttatcatcag	aacaaagagg	ctgtgtcccc	8400
tgccttatga	ggtccatttc	tgagagttgt	ggctaattgg	caagaagggt	ggggcttttag	8460
agattttggga	taaagatatc	aaacaccaga	aaagtagaaa	gaagtgatca	gattaggggt	8520
acttaggtga	tgatatgaac	tcttcctaga	actgagagaa	aaagagagcc	ttcctttact	8580
catatgaaat	cacaaataat	ttctatccaa	tttggaaagta	cacttttggtg	tagttgtgac	8640
agcttcctca	ggactcagca	taaattcaaa	caaataattg	tccttagaag	agatgctata	8700
gaagagatag	aaatatattc	atattctgta	gctttttttt	ttttgagatg	gattttttgct	8760
cttgtcacc	aagctggagt	gcagtgatgc	aatctcagct	cactgcaaac	tttgccctct	8820
gggttcaagg	gattctcctg	cctcagcctc	ccgataactg	ggactacagg	ctacaggcat	8880
gtgtcactac	tcctgggttaa	tttttttttt	tttttttaag	actgagtctt	gctctgtctt	8940
tcaggctgat	gtacaatggc	tccatctcgg	ctcactacaa	cttctgtccc	ccaggttcaa	9000
gcgattctcc	tgcctcagcc	tcagttagtag	ctgggattac	aggcatgtgc	cagcacaccc	9060
agcaaatttt	tgtattttta	gtagagatga	ggccttacca	tgttggccag	gctggtctca	9120
aactcctgac	ctcaggtgat	cctttggcct	cagcctccct	aactgctggg	attacaggca	9180
tgagccactg	cgtccagcct	aattttatat	ttttggtaga	gatgggggtt	caccatattg	9240
gccaggctgg	tctcgaactc	atgacctaa	gtgatccatc	ctcctcagcc	tctcaaagtg	9300
ctgggattac	aagtgtgagc	cactgggcct	ggtgcttttt	tttttttttt	tttttttttt	9360
tttttttttt	gagataggg	ctcactctgt	caccagggt	gaaatgcagt	agtgtgattt	9420
tggctcattg	cagccttgac	ttcccaggct	gaagtgatcc	tcccacctca	gcctcctgag	9480
tagctggggc	tacaggcatg	caccaccatg	ctgcgcta	ttttatat	tttgtagtgg	9540
tgggattttg	ccatatcacc	ctggctggtc	tggaaacct	gggctcaagc	gatccactcg	9600
cttcagcttc	tcaaagtgtc	gggattacag	gcatgagcca	cagcgcccag	gctgtagctc	9660
tcttaaggag	gaacatatct	catctgagac	aaacctgaaa	tgccaaacca	aactgagtta	9720
gcccctctct	gtctgttgta	tatatggag	taataacct	tttgtcttga	taaagggatt	9780
gcatgcttga	attgcaaaaa	cctttatttc	ttttgggttg	cccaatgtgc	aagactaaga	9840
gttattttga	taaatttctc	accaggctga	ctgtctctct	gtgggggtcg	gggagttttc	9900
agggtctcac	gtattgcagg	gaagggttgg	ttgtgagatc	gagaataaca	gaagcagcgg	9960
agcattctgg	aaatattact	atgatggaaa	ggactacatt	gaattcaaca	aagaaatccc	10020
agcctgggtc	cccttcgacc	cagcagccca	gataaccaag	cagaagtggg	aggcagaacc	10080
agtctacgtg	cagcggggca	aggcttacct	ggaggaggag	tgcctgcga	ctctgcggaa	10140
atacctgaaa	tacagcaaaa	atatcctgga	ccggcaagg	actcactgct	tcctgctccc	10200
cagtactgag	cccagaataa	aagacgatct	caggctagga	gctcaggcaa	catcttagtc	10260
cggctctcatc	tgttcctgga	tgtccctcag	acccccagct	ttcatctttt	aggatttatt	10320
ccttcctctgg	gataatataa	tttgtgggtc	aaaaagaaca	tcatacaaat	ttcaggcaga	10380

atggggccagg	aaggccattc	tttcttgatg	agtgtcccca	aatcatctcc	aattaacaga	10440
caaggagctt	gaggttaggg	aggtgaggg	aacactgtct	gtaagaggca	gagctgggac	10500
tcaaattcca	gatttcagat	tccaaatccc	atcgtttttt	atctctacaa	tgatgcctcc	10560
catctgggtg	gtggagagaa	gggaggcgtg	taaaatgtca	gccccagaag	gacaagagca	10620
agccagtgtg	agcgggaattg	atggctgcaa	gctgagactt	ggattggaga	cgtagtgaga	10680
ctcaggattg	tgcaagtgtg	cagggaagtg	gttgctggat	agaggcatgg	gctgaaccaa	10740
gcagctggac	tgagactggg	ggacagaact	ccaaagccca	ctgagatgtg	ggaaaacatg	10800
gagaagcaca	cggagcattc	acaacttatt	gccgtcagag	tcaatacatg	ggtgaggtgg	10860
ggattgggca	agagggaaaag	cgtcagcctt	cctgatatt	ctggaaagtc	ttccggggct	10920
gggggtgggc	aggtacagag	cttcgagctc	tgctgatcgc	tgacatccag	gggtgggggt	10980
aggaagagac	ctgggccggg	agaagtccac	ctcaagcctg	cagtgtcaca	ctctatccct	11040
ccacagatcc	tccctctgtg	gtggtcacca	gccaccaggc	cccaggagaa	aagaagaaac	11100
tgaagtgcct	ggcctacgac	ttctaccag	ggaaaattga	tgtgcaactg	actcgggccg	11160
gcgaggtgca	ggagcctgag	ttacggggag	atgttcttca	caatggaaat	ggcacttacc	11220
agtcttgggt	ggtggtggca	gtgccccgc	aggacacagc	cccctactcc	tgccacgtgc	11280
agcacagcag	cctggccccag	cccctcgtgg	tgccctggga	ggccagctag	gaagcaaggg	11340
ttggaggcaa	tgtgggatct	cagacccagt	agctgccctt	cctgcctgat	gtgggagctg	11400
aaccacagaa	atcacagtca	atggatccac	aaggcctgag	gagcagtgtg	gggggacaga	11460
caggaggtgg	atgttgagac	cgaagactgg	gatgcctgtc	ttgagtagac	ttggacccaa	11520
aaaatcatct	caccttgagc	ccacccccac	cccattgtct	aatctgtaga	agctaataaa	11580
taatcatccc	tccttgccct	gcataacaga	gaatcctttt	tttaacggtg	atgcgctgta	11640
gaaatgtgac	tagattttct	cattggttct	gccctcaagc	actgaattca	tctgaaactc	11700
ttggtttccc	ctggaggcca	tggttcctgg	gcaccttgac	ctgggcaatc	ccaagtgtgg	11760
cctgaacccc	ctttcccttg	gggattgttc	aggtgtccct	agacgccttg	tggtattgta	11820
cctaataccc	atgaaggagg	aggatgat	tacttgccag	tgtacacccc	cctgtgat	11880
tgttcataat	gtccagagtg	aagaaagatg	atattactcc	caatatcaca	gaaggtgtac	11940
acccccctt	gatattgttc	ctaataccca	gttggggagg	ggagaatatc	tctctcaata	12000
tacaaggggt	gtttaaactc	tctgtgat	tgttccta	attcagggg	gacaaggatg	12060
atattaccca	aatattgcag	gggttgtaga	ccccccctt	gatattgttc	ctaatatcca	12120
ggggtggaga	ggatattact	cccaatattg	caggggtcta	catectcccc	ccgtgacatt	12180
gttcttaata	acaaaaaggt	gagaagctga	cattactccc	aataccacag	ggggtgtaca	12240
ccccctatga	gatattgttc	ttaatatcca	ggaggggaga	aatgatatt	actctcaata	12300
gcgcagggaa	tttcatccc	ccgtcgtaat	cttgttctta	atattcagga	agggagagga	12360
tgatacgact	ccagtatcg	cagggggtgt	gcaccccccc	gtgatttttt	tgctaataatc	12420
caggggtggga	gtggataata	cgcaggaagt	gtacaggtct	ctgtgatatt	tttcctaata	12480
tccagggggg	agaggaagat	attactttta	atagtgtacg	gggggtgtac	acccctctgt	12540
gatagtgttc						12550

<210> 2

<211> 12547

<212> DNA

<213> 人

<220>

<221> 包括 AZGP1 的序列

<222> (1)..(12547)

<223> 基于 ac004977, 没有 zag del

<400> 2

```

gatcagataa ttgttttagtt atgccttcaa aaatgtctga gccaggaaca gtggataaat      60
gagcttgtga atcctcgaaa atttgctctt taagttttga aatatccaag gttaagttat     120
catcccaggc ttttaaagtgt cttgagacat tttcccagct atgttgatat ttattataag     180
cataaggcat tatgcaataa tcagaagtat tccaatcact ctgtaattgc atacggtggt      240
ccaaattcat atctcccagc cagattacac tttggcggag atcattaatt tgattagcta     300
atTTTTgatc aatttgagcc tgagaattcc aaagtctgga ggagtttttc tgccatgctt      360
caattcaaca tattgaacgg tttgaacaga attgtggatg gcaactccag cagttgctgc     420
tgttgcagta accgcaatta gtcctgcaat gatggcaata agagtaaaaa taaatctctt     480
tgttcttttg cggatacctt taagaacttc attgactatg tgaatagagg gggaagactc     540
ccatggacga tgtaaagaaa ctggtatcca taccctctcc ctagccctta ccaagagagt     600
acttgtagtg ggattaaaag tagcatcaat gcacgtgaac agcttacagt tatcacattc     660
tatagtttgt gtattgggaa tgataattat atttccaacc aacagcatgt aaggggggtt      720
gacatagctc ctgatgggta tcacccttc agatatgagg gtgatgttga atgtgggtgt      780
tttggtatta gtgtgaagga gttgataggt agtgttccat atccttattc ctgtcatagc     840
tgcagctaat tttcatagtt gaggatgttc tggggtaaca aagggatgaa tcatttttgg     900
tctaggagga gtaatgcctg catccatcca tttaaataagg taaggggaca cccattgtct     960
cagcctgtat gactgccatc catcctctac ataatctaac aaataattaa actctgagca    1020
tgaggatatt ttgctggagc aatcttgcca ataatgccct tttggagccc aatcaataac    1080
aatacctgcg gctagacttt ggagcacaac ggccttagga gcattacaat tattccataa    1140
aattgaagtc actataaaag gtccctttgt aggtttcttt aacccaaaaga gaggctgggc    1200
acagttgctc acacttgtaa acccagcact ttgggagacc gaggcgggca gatcacgagg    1260
tcaggagttc aaaaccagct tggccaacat ggtgaaaccc ctgcctcagc ttcctgagtg    1320
ctgggattat aggcattgtc caccacgcgc agctaattct tgtatttttt cagtagagac    1380
gaggtttcat catgttgccc aggctggtct cgagtacctg acctcaacct gaggtgatcc    1440
aaccacctca gcctcccaa gtgctgggat tacaagcatg agtcaccgca cccgggcccc    1500
agtcacttta gaatagcatg ttgctttgta ttcggagggt ctctctgcaa atagcccatc    1560
aacactgagc gtgcctggaa agacctggtt ttcaaataac tggcttcgtc tgtgtaaaac    1620

```

gagtcttggt	gtatgcatta	aaaattatct	tggctgggcg	tggctggctca	cgcctttaat	1680
cccagcattt	tgggaggctc	gttctgttgc	cagggaggct	gcaaaaggag	gatgcggcta	1740
tttctgcttc	catcatggaa	ggctgggatt	ccacaggcac	agaaaggaag	gagggagaga	1800
tgggaatgtg	actgtcctcc	agacacagcc	tctggcagga	tcggatgtgg	gagagtgagg	1860
gtcccccccc	agctggggtc	taccaggtc	catgtcttgg	acatgttgag	agtttttctg	1920
gaaggcaggg	atacagtgtg	gtccaaaaac	acacaaatgc	ccctactggc	ccaggggttg	1980
tcacaataga	ctggaagggt	gacacatccc	aggcgcttgc	cacccatcac	acgcacctcc	2040
taccactgg	catccttcca	ccccaggcac	acacaaagcc	tcagtccaga	gatcaactct	2100
ggactcagct	ctgaatttgc	atatcctgtg	tgtagattca	ttcttcataa	cctctgcccc	2160
gcctagcttg	tgtatcatth	ttttttctct	attaggggag	gagcccgtcc	tggcactccc	2220
attggcctgt	agattcacct	cccctgggca	gggccccagg	acccaggata	atatctgtgc	2280
ctcctgcccc	gaaccttcca	agcagacaca	atggtaagaa	tggctgctgt	cctgctgtct	2340
ctgctgctgc	ttctgggtcc	tgctgtcccc	caggagaacc	aagatgggtga	gtggggaaag	2400
caagggatgg	gtgctggaga	ggactggaag	gaggtgagga	acaggacatg	tggctgggag	2460
acaggctgga	tgcagctggg	ataccctggc	atacggcagg	aatgggtgcc	caaggctgtc	2520
aactccctca	gctcacacac	ttccaggagc	attcagggag	cctctgcgct	ggccccgaaat	2580
aagaccttca	ggaatctgaa	tctaaaaccc	ctagtttaca	gtgaaaacaa	agactccaaa	2640
gaccaagcga	cctgcttggg	gtagacagtc	aggacggagt	aggaaccata	tgcttgagac	2700
tgcttctgct	cctgttcctt	ccctccttcc	gatggctggg	tacacctgcc	tgacgctgat	2760
gaaaagagag	agcagcccca	aggggaaagt	gggaaggcag	gttggctgga	gggatggtgc	2820
tagaaggaaa	cccgtgcccc	aatccacac	tcagacacca	ctgcagtggg	tctggaaggc	2880
gagtggctgg	aagagaagag	agtgggagct	ccgggagatc	aagagtcact	cctaggataa	2940
gggaaggagg	ctgtttgtgg	catgagaatg	tgcaggataa	agacatggaa	gcgaatggct	3000
tctcagttgt	gtgagtttaa	aattcatgac	atttacaat	tgtcagaaaa	ggtgttatat	3060
gtttgtttata	taacaatcac	tttggaatgt	taatctgatt	ctgtgccaaa	atctgaatta	3120
ctcagggttc	tccagagaaa	cagaactaat	aggtggtaca	catatacata	tatatgtacg	3180
tacacataca	tacatacact	gtatacacat	ggatacacac	acacatagga	agagatttac	3240
atatatgtat	acaaaagaga	gagagagtag	agattttatt	taagaaattg	actcacacta	3300
ttgggaggag	taacaagtcc	taaatcttca	gagccggcca	gcaggctgga	gacccaggga	3360
agagttgatg	tcttagtctt	gattccaagg	gcagactgta	ggcagaattc	tttctctctt	3420
aggggacatc	tgaggctttt	tctcttaagg	ccttcaactg	attggatgaa	gcccaccact	3480
atggagagta	atccacttta	ctcaaggctc	actgattttt	ttgtaaatta	aaaaaaaaac	3540
tgtgggtgca	tagtatgtgt	atatatttat	ggggtacatg	agaggttttg	attcaggcat	3600
gcaatgtgaa	ataatcacat	catcaaaaat	gaggtatcca	tcccttcaag	cttttatcgt	3660
ttgtgtttaca	gacaatccaa	ttatactttt	ttggttatth	tagtttttaa	aagtatttga	3720
ttattttatt	attttatttat	ttttgagaca	gagtctcact	ctgtcaccca	ggcaggagtg	3780
cagtggcatg	atctcggctc	actgcaacct	ccgcctccca	ggttcaagca	attttcctgc	3840
ctcagtctcc	tgagtagcta	ggactacagg	cacctgccac	cacacctggc	taattttttt	3900

gtatTTTTtag	tagagacggg	tttcatcatg	ttggccaggc	tagtcttgat	atcctgacct	3960
cgtgatctgc	ccgccttgg	ctcccaaagt	gccgggatta	caggtgtcag	caactgcgcc	4020
tggcctctct	tttggttatt	taaaagtgt	caattaaatt	atgattatta	ttattatttt	4080
tgagatggat	tcttggtctg	tcacccaggc	tgagtgacag	tggcgtgatc	ttggcttact	4140
gcaaacctcc	gcctgttggg	ttcaagcaat	tatcttgcc	cgggtgtaca	ctgccacaca	4200
cggctaactt	atgtattttt	aatagagata	gggtttcacc	atgttggcta	gactggctct	4260
gacctcttga	cctcaagtga	tccactcact	tcagcctccc	agagtgtctg	aattacaggc	4320
acgagccacc	acacctggcc	ccagttaa	tattattgac	tatagtcacc	ctgttggtgt	4380
atcaaatagt	aggtcttatt	cattcttctt	tttttttttt	tttttgtgac	agagttgccc	4440
aggctggaat	gcagtgggtg	aatcttggct	cactgcaacc	tctgcctccc	gggcttaagc	4500
gattctcctg	cctcagcctt	ctgagtcgct	gggactacag	gtgtgtgcca	ccacgcccgg	4560
ctaatttatg	tatttttagt	agagatggg	tttcaccatg	ttggccaggc	tggtttcgaa	4620
ctcctgacct	caagtgaccc	acctgcctca	gcttcccaaa	gtgttggga	tacaggcatg	4680
agccaccaca	cctggcccca	gttaaattat	tattcactgg	agtcactttg	ttgtgctatc	4740
aaatagtttt	ctaactattt	tttttgtacc	cattaaccac	cctcccaatt	ttcccccaac	4800
cctgccacta	cccttcccag	cctttggtaa	ccatccttct	actctctatg	tccatgaatt	4860
caattgtagg	gtctactgat	ttaaaggcta	atcacattta	gacactcagg	agcaagaata	4920
atttttagtaa	ttgaactagg	attctgccat	atgacctcca	acatcattag	cacctgtgta	4980
aattgtatca	taaaataatt	atggaactat	tatggaaatg	tccctctctc	ccagatccca	5040
ccttgtacca	aatgcaagg	tacaaccccg	ggaattctga	gctccatcct	agtcttacc	5100
tgtgctaatt	cagtctgggt	catttcttga	attttctggt	aaattctcct	ttctaccctt	5160
tctaactata	tgtatttgtc	aggtaagct	agaagtgtta	attttttttt	tttttgagat	5220
ggagccttgc	tttgtcacct	aggctgaagt	gcagtggcat	gatctcagct	cactgcaagc	5280
tccgcctccc	gggttcatgc	cattctcctg	cctcagcctc	ctgagtagct	gggactacag	5340
gcacccgcca	ccatgcttgg	ctaattttgt	gaattcttag	tagagacggg	gtttcaccat	5400
gttagccagg	atggtctcga	tctcctgacc	tcgtgatcca	cccgctcgg	ccccctaaag	5460
tgtctgggatt	acaggcgtga	gccactgagc	ccggacgaaa	tgttaatttg	ttttttttga	5520
gacggagtct	cactctgtca	tccaagctgg	agtgcagtgg	catgatcttg	gcttggtgca	5580
acctctgcct	ctctggttca	agtgattttc	ctgcctcagc	ctccagcatg	actgggatta	5640
caggccccga	ccaccatgcc	cagctaattt	ttgtattttt	taatagagat	ggggtttcac	5700
catgttggcc	aggctggtct	tcaactcctg	atctcaagta	atctgcctgc	cttggcctcc	5760
caaagtccctg	ggattacagg	catgagccac	ggagcccagc	ctagaaatgt	taattttctaa	5820
cgcagtgcag	attccatgca	cactgggcaa	ggttccattc	ctccatgggg	tgactcaggg	5880
atccaggcca	attgcatatt	gagactcttt	catattatcc	tgtggccttc	aaagtcgtca	5940
cctctagggga	tgagaaacaa	aagggacaag	ccagctggta	gggtcttggga	caagaagaaa	6000
gacatcactt	ctgctcacat	tctcttttga	caaaactcag	tcacatgggtc	ccaatatatc	6060
ttcgaggtgg	ctgagtaatg	ttatcttctc	atgtgtcaag	cagaggaaat	aatgtagtga	6120
agacacagga	tggctctctga	aatatcatct	caggcatgaa	agtagagcat	attcacttga	6180

gtgagcctcc	agtgggtgtga	agttgatggc	aggagaaaga	gctggggaag	aaaaggccag	6240
tggcaggtct	cccctcctag	ccctatgcag	ccccacagt	ggacccttgc	atggacctca	6300
accatcagaa	tcttttcttt	tgcaggtcgt	tactctctga	cctatatcta	cactgggctg	6360
tccaagcatg	ttgaagacgt	ccccgcgttt	caggcccttg	gctcactcaa	tgacctccag	6420
ttcttttagat	acaacagtaa	agacaggaag	tctcagccca	tgggactctg	gagacaggtg	6480
gaaggaatgg	aggattggaa	gcaggacagc	caacttcaga	aggccaggga	ggacatcttt	6540
atggagaccc	tgaaagacat	cgtggagtat	tacaacgaca	gtaacgggtca	gtgaataaca	6600
gaccacaggg	gtggaagggtc	taacccaaga	ggcagccccc	ccagtgtgag	tggcaaggga	6660
tcagcaggat	ggaaatagtc	ccaatcccag	gggaagaaca	ggagacacag	cagaaacaca	6720
gacatgtcca	catcccaccc	accccacagc	acaggtgctc	cccgttccc	catcaattgc	6780
cccatactca	tcccaggcct	caggtcacac	aggaagtgat	ggcagagtca	cttccatatcc	6840
aggcacctat	gacctctcac	ctccacaccc	cacccatcgg	aggctgatac	ccccgtgaga	6900
aggcatcaga	ctcaccctcg	tccagggagg	ttgcctggag	agtgagccac	tctcaaagtc	6960
actcagacct	gggtcacct	ggtggttctg	ccagtcctag	ctgttgacag	tgaaacgttc	7020
ccaaaatata	tggttgaaat	ctgcaaacat	tggagcactg	agacctacct	ccaaacaagt	7080
ctgtaatat	taactatgtc	tgttctatga	aggatgtcac	agtctgtcct	gatctccctt	7140
gcagctccat	cacctagcac	agggtacagc	caatattggc	tcaattgaaa	tttgtggaat	7200
ccacagagaa	aagcaccctg	cacacaccgt	agcccatgct	gggggctcag	gaagtgtctg	7260
attcaaaact	gtgggctgtt	agagttcctt	ggagccctaa	agttcctcct	taccatacga	7320
tgcagaccca	ggaagggccca	cctgcgctat	ggtcagagga	gctgggtggca	gagccctgct	7380
agagatggtc	cctgtgcccc	cggcccagtg	ctctttctcc	taaaccacac	tgccagcccc	7440
aaggcagcca	acctcaggtc	tgggtgaactg	ctgggtgttaa	attatcatag	agtgggtgtc	7500
aaaagatggg	ctactaagta	caaaaatgcc	caaggtgcta	catgggatct	gaagattttc	7560
aaaaggaggc	aagaaagaga	taggcagatg	tttcaaggat	gtgggggtggg	ggagggtcttg	7620
gtaaggaaaa	tggcccaggc	tgtgtgtcag	caataggaga	ggagggggca	cagggtgatca	7680
gaaaagacac	tgggggaagc	attgatggac	aggaatagaa	atggcaaagt	ggataattaa	7740
gaggaaggag	gatgaggaga	tgaacacagg	gtattagaaa	ataatagaag	gcagggtctg	7800
gtggctcact	cttgtaatcc	cagcactttg	ggaggctgag	gcaggcagat	cacctaaagt	7860
caggagttcg	agaccagccc	ggccaacatg	gtgaaacctt	gtctctacta	ataatacaaa	7920
aatagcctgg	catgggtggca	cacgtctgtg	gtcccagcta	ctcaggaggc	tgaggcagga	7980
gaattgcttg	aaccagggag	gcagagggtta	cagttagcca	aaatcctacc	attgcactac	8040
agcctgagt	acaagagtga	aacgttgtct	aaaaacaaaa	aacaaaaaac	aaaaaaagga	8100
aataatagta	gctgacattt	actgagcact	tactttgtgc	caggcccatc	tatgagcata	8160
tataatgctc	agaatagccc	cctaaaacag	tgctcttggt	attgccattt	cagaggtgag	8220
gaaatagagg	cacagggagt	tgagtggctc	cagttcaggc	aacacaccag	gtgggggtgg	8280
ggggctgggg	agagacctgg	gacgtgagcc	cagacagctt	gagagctttc	agagtctatg	8340
ccaacagcac	caaccagtgc	tgggtaaaca	cctgctttta	tcatcagaac	aaagaggctg	8400
tgtcccctgc	cctatgaggt	ccatttctga	gagttgtggc	taatgggcaa	gaagggtggg	8460

gcttttagaga	tttgggataa	agatatcaaa	caccagaaaa	gtagaaagaa	gtgatcagat	8520
taggggttact	taggtgatga	tatgaactct	tcctagaact	gagagaaaaa	gagagccttc	8580
ctttactcat	atgaaatcac	aaataatttc	tatccaatth	ggaagtacac	tttgggtgtag	8640
ttgtgacagc	ttcctcagga	ctcagcataa	attcaaacia	ataattgtcc	ttagaagaga	8700
tgctatagaa	gagatagaaa	tatattcata	ttctgtagct	tttttttttt	tgagatggat	8760
ttttgctctt	gtcaccacaag	ctggagtgc	gtgatgcaat	ctcagctcac	tgcaaacttt	8820
gcctcctggg	ttcaagggat	tctcctgcct	cagcctcccg	ataactggga	ctacaggcta	8880
caggcatgtg	tcactactcc	tggttaattt	tttttttttt	ttttaagact	gagtcttgct	8940
ctgtctttca	ggctgatgta	caatggctcc	atctcggctc	actacaactt	ctgtccccc	9000
ggttcaagcg	attctcctgc	ctcagcctca	tgagtagctg	ggattacagg	catgtgccag	9060
cacaccacagc	aaatttttgt	atttttagta	gagatgaggt	cttaccatgt	tggtccaggct	9120
ggtctcaaac	tcctgacctc	aggtgatcct	ttggcctcag	cctccctaac	tgctgggatt	9180
acaggcatga	gccactgcgt	ccagccta	tttatatttt	tggtagagat	gggttttcac	9240
catattggcc	aggctgggtct	cgaactcatg	acctaagggtg	atccatcctc	ctcagcctct	9300
caaagtgtctg	ggattacaag	tgtgagccac	tggtcctggg	gctttttttt	tttttttttt	9360
tttttttttt	tttttttgag	ataggggtctc	actctgtcac	ccaggctgaa	atgcagtagt	9420
gtgattttgg	ctcattgcag	ccttgacttc	ccaggctgaa	gtgatcctcc	cacctcagcc	9480
tcctgagtag	ctggggctac	aggcatgcac	caccatgctg	cgctaatttt	tatatatttt	9540
gtagtgggtg	gattttgcca	tatcaccctg	gctgggtctg	aaccctggg	ctcaagcgat	9600
ccactcgctt	cagcttctca	aagtgtctgg	attacaggca	tgagccacag	cgcccaggct	9660
gtagctctct	taaggaggaa	catatctcat	ctgagacaaa	cctgaaatgc	caaaccacac	9720
tgagttagcc	cctctctgtc	tgttgtatat	attggagtaa	taacctattt	gtcttgataa	9780
agggattgca	tgcttgaatt	gcaaaaacct	ttatttcttt	tggtttgccc	aatgtgcaag	9840
actaagagtt	attttgataa	atttctcacc	aggctgactg	tctctctgtg	gggtcggggg	9900
agttttcagg	gtctcacgta	ttgcaggga	ggtttggttg	tgagatcgag	aataacagaa	9960
gcagcggagc	attctggaaa	tattactatg	atggaaagga	ctacattgaa	ttcaacaaag	10020
aaatcccagc	ctgggtcccc	ttcgaccag	cagcccagat	aaccaagcag	aagtgggagg	10080
cagaaccagt	ctacgtgcag	cggtccaagg	cttacctgga	ggaggagtgc	cctgcgactc	10140
tgcggaata	cctgaaatac	agcaaaaata	tcctggaccg	gcaaggtagt	cactgcttcc	10200
tgctccccag	tactgagccc	agaataaaa	acgatctcag	gctaggagct	caggcaacat	10260
cttagtccgg	tctcatctgt	tcctggatgt	ccctcagacc	cccagctttc	atcttttagg	10320
atttattcct	tcctggggat	aatataattt	gtgggtccaa	aagaacatca	tcaaaatttc	10380
aggcagaatg	ggccaggga	gccattcttt	cttgatgagt	gtccccaat	catctccaat	10440
taacagacaa	ggagcttgag	gttagggagg	tgagggtaac	actgtctgta	agaggcagag	10500
ctgggactca	aattccagat	ttcagattcc	aaatcccata	gtttttttat	tctacaatga	10560
tgctcccat	ctgggtggtg	gagagaagg	aggcgtgtaa	aatgtcagcc	ccagaaggac	10620
aagagcaagc	cagtgtgagc	ggaattgatg	gctgcaagct	gagacttgga	ttggagacgt	10680
agtgtgactc	aggattgtgc	agtgtgtcag	ggaagtgggt	gctggataga	ggcatgggct	10740

```

gaaccaagca gctggactga gactggggga cagaactcca aagcccactg agatgtggga 10800
aaacatggag aagcacacgg agcattcaca acttattgcc gtcagagtca atacatgggt 10860
gaggtgggga ttgggcaaga gggaaagcgt cagccttccc tgatattctg gaaagtctcc 10920
cggggctggg ggtgggcagg tacagagctt cgagctctgc tgatcgctga catccagggg 10980
tgggggtagg aagagacctg ggccgggaga agtccacctc aagcctgcag tgtcacactc 11040
tatccctcca cagatcctcc ctctgtggtg gtcaccagcc accaggcccc aggagaaaag 11100
aagaaactga agtgacctggc ctacgacttc taccagggga aaattgatgt gcactggact 11160
cgggcccggcg aggtgcagga gcctgagtta cggggagatg ttcttcacaa tggaaatggc 11220
acttaccagt cctgggtggt ggtggcagtg ccccgagcagg acacagcccc ctactcctgc 11280
cacgtgcagc acagcagcct ggcccagccc ctctgtggtg cctgggagggc cagctaggaa 11340
gcaagggttg gaggcaatgt gggatctcag acccagtagc tgcccttcct gcctgatgtg 11400
ggagctgaac cacagaaatc acagtcaatg gatccacaag gcctgaggag cagtgtgggg 11460
ggacagacag gaggtggatt tggagaccga agactgggat gcctgtcttg agtagacttg 11520
gacccaaaaa atcatctcac cttgagccca cccccacccc attgtctaata ctgtagaagc 11580
taataaataa tcatccctcc ttgcctagca taacagagaa tccttttttt aacgggtgatg 11640
cgctgtagaa atgtgactag attttctcat tggttctgcc ctcaagcact gaattcatct 11700
gaaactcttg gtttcccctg gaggccatgg ttcttgggca ccttgacctg ggcaatccca 11760
agtgtggcct gaacccccct tcccttgggg attgttcagg tgcccttaga cgccttgtgg 11820
tattgtacct aatacccatg aaggagagg atgatattac ttgccagtgt acacccccct 11880
gtgatattgt tcataatgtc cagagtgaag aaagatgata ttactcccaa tatcacagaa 11940
ggtgtacacc ccccttgat attgttccta ataccagtt ggggagggga gaatatctct 12000
ctcaatatat aaggggtgtt taaactctct gtgatattgt tcctaataatt cagggggggac 12060
aaggatgata ttacccaaat attgcagggg ttgtacaccc ccccttgat attgttccta 12120
atatccaggg gtggagagga tattactccc aatattgcag ggggtctacat cctccccccg 12180
tgacattggt cttaataacc aaaagggtgag aagctgacat tactcccaat accacagggg 12240
gtgtacaccc cctatgagat attgttctta atatccagga ggggagaaaa tgatattact 12300
ctcaatagcg cagggaattt acatcccccg tcgtaatctt gttcttaata ttcaggaagg 12360
gagaggatga tacgactccc agtatcgag ggggtgtgca ccccccggtg atttttttgc 12420
taatatccag ggtgggagtg gataatacgc aggaagtgtg caggtctctg tgatattttt 12480
cctaatatcc aggggggaga ggaagatatt acttttaata gtgtacgggg ggtgtacacc 12540
cctctgt
12547

```

<210> 3

<211> 31

<212> DNA

<213> 人

<220>

<221> zag06
<222> (16) .. (16)
<223> 在 Seq ID No. 1 或 2 的位置 1090 的 G/A 多态性

<400> 3
aataacaata cctgcgggcta gactttggag c 31

<210> 4
<211> 31
<212> DNA
<213> 人

<220>
<221> zag05
<222> (16) .. (16)
<223> 在 Seq ID No. 1 或 2 的位置 1196 的 T/C 多态性

<400> 4
aaccaaaaga gaggctgggc acagttgctc a 31

<210> 5
<211> 31
<212> DNA
<213> 人

<220>
<221> zag04
<222> (16) .. (16)
<223> 在 Seq ID No. 1 或 2 的位置 1216 的 T/A 多态性

<400> 5
acagttgctc acacttgtaa acccagcact t 31

<210> 6
<211> 31
<212> DNA
<213> 人

<220>	
<221>	zag03
<222>	(16) .. (16)
<223>	在 Seq ID No. 1 或 2 的位置 1348 的 C/T 多态性
<400>	6
	gcatgtgccca ccacgcgcag ctaattcttg t 31
<210>	7
<211>	31
<212>	DNA
<213>	人
<220>	
<221>	zag07
<222>	(16) .. (16)
<223>	在 Seq ID No. 1 或 2 的位置 2695 的 T/C 多态性
<400>	7
	taggaaccat atgcctggag ctgcttctgc t 31
<210>	8
<211>	31
<212>	DNA
<213>	人
<220>	
<221>	zag08
<222>	(16) .. (16)
<223>	在 Seq ID No. 1 或 2 的位置 2760 的 T/G 多态性
<400>	8
	cctgcctgac gctgatggaa agagagagca g 31
<210>	9
<211>	31
<212>	DNA
<213>	人

<220>	
<221>	zag09
<222>	(16)..(16)
<223>	在 Seq ID No. 1 或 2 的位置 2762 的 A/G 多态性
<400>	9
tgcttgacgc	tgaggaaaag agagagcacc c 31
<210>	10
<211>	31
<212>	DNA
<213>	人
<220>	
<221>	zag13
<222>	(16)..(16)
<223>	在 Seq ID No. 1 或 2 的位置 4528 的 G/A 多态性
<400>	10
tcagccttct	gagtcgctgg gactacaggt g 31
<210>	11
<211>	31
<212>	DNA
<213>	人
<220>	
<221>	zag12
<222>	(16)..(16)
<223>	在 Seq ID No. 1 或 2 的位置 5013 的 T/C 多态性
<400>	11
attatggaac	tattatggaa atgtccctct c 31
<210>	12
<211>	31
<212>	DNA

<213>	人	
<220>		
<221>	zag10	
<222>	(16) .. (16)	
<223>	在 Seq ID No. 1 或 2 的位置 5369 的 G/T 多态性	
<400>	12	
	tgcttggcta attttgtgaa ttcttagtag a	31
<210>	13	
<211>	31	
<212>	DNA	
<213>	人	
<220>		
<221>	zag14	
<222>	(16) .. (16)	
<223>	在 Seq ID No. 1 或 2 的位置 6561 的 C/T 多态性	
<400>	13	
	gaccctgaaa gacatcgtgg agtattacaa c	31
<210>	14	
<211>	31	
<212>	DNA	
<213>	人	
<220>		
<221>	zag23	
<222>	(16) .. (16)	
<223>	在 Seq ID No. 1 或 2 的位置 6730 的 A/G 多态性	
<400>	14	
	aacacagaca tgtccacatc ccacccaccc c	31
<210>	15	
<211>	31	

<212> DNA

<213> 人

<220>

<221> zag22

<222> (16) .. (16)

<223> 在 Seq ID No. 1 或 2 的位置 6894 的 C/T 多态性

<400> 15

ggaggctgat acccccgtga gaaggcatca g

31

<210> 16

<211> 31

<212> DNA

<213> 人

<220>

<221> zag18

<222> (16) .. (16)

<223> 在 Seq ID No. 1 或 2 的位置 7202 的 C/A 多态性

<400> 16

gaaatttgtg gaatccacag agaaaagcac c

31

<210> 17

<211> 31

<212> DNA

<213> 人

<220>

<221> zag19

<222> (16) .. (16)

<223> 在 Seq ID No. 1 或 2 的位置 7219 的 G/A 多态性

<400> 17

cagagaaaag caccggtcac acaccgtagc c

31

<210> 18

<211> 31
<212> DNA
<213> 人

<220>
<221> zag20
<222> (16)..(16)
<223> 在 Seq ID No. 1 和 2 的位置 7219 的 G/A 多态性

<400> 18
ccaaggcagc caacctcagg tctggtgaac t 31

<210> 19
<211> 31
<212> DNA
<213> 人

<220>
<221> zag21
<222> (16)..(16)
<223> 在 Seq ID No. 1 或 2 的位置 7459 的 T/C 多态性

<400> 19
gcagccaacc tcaggtctgg tgaactgctg g 31

<210> 20
<211> 31
<212> DNA
<213> 人

<220>
<221> zag17
<222> (16)..(16)
<223> 在 Seq ID No. 1 或 2 的位置 8047 的 A/G 多态性

<400> 20
ttgcactaca gcctgagtga caagagtgaa a 31

<210>	21	
<211>	31	
<212>	DNA	
<213>	人	
<220>		
<221>	zag_del	
<222>	(13)..(19)	
<223>	在 Seq ID No. 2 的位置 8077 至 8083 的 AAAAAAC/. 多态性	
<400>	21	
	ttgtctaaaa acaaaaaaca aaaaacaaaa a	31
<210>	22	
<211>	30	
<212>	DNA	
<213>	人	
<220>		
<221>	zag16	
<222>	(16)..(16)	
<223>	在 Seq ID No. 1 的位置 8493 和在 Seq ID No. 2 的位置 8500 的 A/G 多态性	
<400>	22	
	atcaaacacc agaaaagtag aaagaagtga	30
<210>	23	
<211>	31	
<212>	DNA	
<213>	人	
<220>		
<221>	zag15	
<222>	(16)..(16)	
<223>	在 Seq ID No. 1 的位置 9549 和在 Seq ID No. 2 的位置 9556 的 T/C 多态性	
<400>	23	
	gtagtggtgg gattttgccca tatcacctg g	31
<210>	24	

<211> 31
<212> DNA
<213> 人

<220>
<221> zag24
<222> (16) .. (16)
<223> 在 Seq ID No. 1 的位置 10202 和在 Seq ID No. 2 的位置 10209 的 A/C 多态性

<400> 24
tgcttcctgc tccccagtac tgagcccaga a 31

<210> 25
<211> 31
<212> DNA
<213> 人

<220>
<221> zag25
<222> (16) .. (16)
<223> 在 Seq ID No. 1 的位置 10439 和在 Seq ID No. 2 的位置 10446 的 G/A 多态性

<400> 25
catctccaat taacagacaa ggagcttgag g 31

<210> 26
<211> 31
<212> DNA
<213> 人

<220>
<221> zag26
<222> (16) .. (16)
<223>在 Seq ID No. 1 的位置 11020 和在 Seq ID No. 2 的位置 11027 的 G/T 多态性

<400> 26
gtccacctca agcctgcagt gtcacactct a 31

<210> 27
<211> 31
<212> DNA
<213> 人

<220>
<221> zag35
<222> (16)..(16)
<223> 在 Seq ID No. 1 的位置 11995 或在 Seq ID No. 2 的位置 12002 的 T/C 多态性

<400> 27
gggagaatat ctctctcaat atacaagggg t 31

<210> 28
<211> 31
<212> DNA
<213> 人

<220>
<221> zag34
<222> (16)..(16)
<223> 在 Seq ID No. 1 的位置 12385 和在 Seq ID No. 2 的位置 12392 的 G/T 的多态性

<400> 28
tcccagtatc gcaggggggtg tgcaccccc c 31

```

7951  gtcccagcta ctcaggaggc tgaggcagga gaattgcttg aaccaggag
| 8001  gcagaggtta cagtgaacca aaatcctacc attgcactac agcctgagtg
| 8051  acaagagtga aacgttgtct aaaaacaaaa aadaaaaaaa ggaaataata
8101  gtagctgaca ttactgagc acttactttg tgccaggccc atctatgagc
8151  atatataatg ctcagaatag cccctaaaa cagtgtcttt ggcattgcc
8201  tttcagaggt gaggaaatag aggcacaggg agttgagtgg ctccagttca
8251  ggcaacacac caggtggggg tggggggctg gggagagacc tgggacgtga
8301  gccagacag cttgagagct ttcagagtct atgccaacag caccaaccag
8351  tgctgggtaa acacctgctt ttatcatcag aacaaagagg ctgtgtcccc
8401  tgccctatga ggtccatttc tgagagttgt ggctaattgg caagaagggt
| 8451  ggggctttag agatttgga taaagatatc aaacaccaga aagtagaaa
8501  gaagtgatca gattagggtt acttaggtga tgatatgaac tcttcctaga
8551  actgagagaa aaagagagcc ttcctttact catatgaaat cacaataat
8601  ttctatccaa tttggaagta cactttggtg tagttgtgac agcttcctca
8651  ggactcagca taaattcaaa caaataattg tccttagaag agatgctata
8701  gaagagatag aaatatattc atattctgta gctttttttt ttttgagatg
8751  gatttttgct cttgtcacc aagctggagt gcagtgatgc aatctcagct
8801  cactgcaaac tttgcctcct gggttcaagg gattctcctg cctcagcctc
8851  ccgataactg ggactacagg ctacaggcat gtgtcactac tcttggttaa
8901  tttttttttt tttttttaag actgagtctt gctctgtctt tcaggctgat
8951  gtacaatggc tccatctcgg ctactacaa cttctgtccc ccaggttcaa
9001  gcgattctcc tgctcagcc tcatgagtag ctgggattac aggcatgtgc
9051  cagcacaccc agcaaatttt tgtattttta gtagagatga ggtcttacca
9101  tgttggccag gctggtctca aactcctgac ctcaggtgat cctttggcct
9151  cagcctccct aactgctggg attacaggca tgagccactg cgtccagcct
9201  aattttatat ttttggtaga gatgggggtt caccatattg gccaggctgg
9251  tctcgaactc atgacctaa gtgatccatc ctctcagcc tctcaaagtg

```

图 1. 目的标记在用于发现 AZGP1 中的
SNP 的基因组序列上的作图

```

9301 ctgggattac aagtgtgagc cactgggcct ggtgcttttt tttttttttt
9351 tttttttttt tttttttttt gagatagggc ctactctgt caccagget
9401 gaaatgcagt agtgtgattt tggctcattg cagccttgac ttccagget
9451 gaagtgatcc tcccacctca gcctcctgag tagctggggc tacaggcatg
9501 caccaccatg ctgcgctaatt ttttatattt tttgtagtgg tgggattttg
9551 ccatatcacc ctggctgggc tgggaaccct gggctcaagc gatccactcg
9601 cttcagcttc tcaaagtgcct gggattacag gcatgagcca cagcgcccag
9651 gctgtagctc tcttaaggag gaacatatct catctgagac aaacctgaaa
9701 tgccaaacca aactgagtta gccctctctt gtctgttgta tatattggag
9751 taataacctt tttgtcttga taaagggtt gcatgcttga attgcaaaaa
9801 cctttatttc ttttgggttg cccaatgtgc aagactaaga gttattttga
9851 taaattttct accaggctga ctgtctctct gtggggctcg gggagttttc
9901 agggctctac gtattgcagg gaaggtttgg ttgtgagatc gagaataaca
9951 gaagcagcgg agcattctgg aatatattact atgatggaaa ggactacatt
10001 gaattcaaca aagaaatccc agcctgggtc cccttcgacc cagcagccca
10051 gataaccaag cagaagtggg aggcagaacc agtctacgtg cagcgggcca
10101 aggccttacct ggaggaggag tgccctgcga ctctgcgaa atacctgaaa
10151 tacagcaaaa atatcctgga cgggcaaggt actcactgct tctgtctccc
10201 cagtactgag ccagaataa aagacgatct caggctagga gctcaggcaa
10251 catcttagtc cggctctatc tgttcctgga tgtccctcag acccccagct
10301 ttcattctttt aggatttatt ccttccctgg gataatataa tttgtggtcc
10351 aaaaagaaca tcatcaaaat ttcaggcaga atgggccagg aaggccattc
10401 tttcttgatg agtgtcccca aatcatctcc aattaacaga caaggagctt
10451 gaggttaggg aggtgagggg aacactgtct gtaagaggca gagctgggac
10501 tcaaattcca gatttcagat tccaaatccc atcgtttttt atctctacaa
10551 tgatgcctcc catctgggtg gtggagagaa gggaggcgtg taaaatgtca
10601 gccccagaag gacaagagca agccagtgtg agcggaattg atggctgcaa
10651 gctgagactt ggattggaga cgtagtgaga ctcaggattg tgcagtgtg
10701 caggggaagtg gttgctggat agaggcatgg gctgaaccaa gcagctggac

```

图 1 续

```

10751  tgagactggg ggacagaact ccaaagocca ctgagatgtg ggaaaacatg
10801  gagaagcaca cggagcattc acaacttatt gccgtcagag tcaatacatg
10851  ggtgaggtgg ggattgggca agagggaag cgtcagcctt ccctgatatt
10901  ctggaaagtc tcccggggct gggggtgggc aggtacagag cttcgagctc
10951  tgctgatcgc tgacatccag ggggtggggg aggaagagac ctgggccggg
11001  agaagtcac ctcaagcctg cagtgtcaca ctctatccct ccacagatcc
11051  tccctctgtg gtggtcacca gccaccaggc cccaggagaa aagaagaaac
11101  tgaagtgcct ggcctacgac ttctaccag ggaaaattga tgtgactgg
11151  actcgggccg gcgaggtgca ggagcctgag ttacggggag atgttcttca
11201  caatggaaat ggcacttacc agtctgggt ggtggtggca gtgccccgc
11251  aggacacagc cccctactcc tgccacgtgc agcacagcag cctggcccag
11301  cccctcgtgg tgccctggga ggccagctag gaagcaaggg ttggaggcaa
11351  tgtgggatct cagaccaggt agctgccctt cctgcctgat gtgggagctg
11401  aaccacagaa atcacagtca atggatccac aaggcctgag gagcagtgtg
11451  gggggacaga caggaggtgg atttgagac cgaagactgg gatgcctgtc
11501  ttgagtagac ttggaccaa aaaatcatct caccttgagc ccacccccac
11551  cccattgtct aatctgtaga agctaataaa taatcatccc tccttgccca
11601  gcataacaga gaatcctttt ttaacggtg atgcgctgta gaaatgtgac
11651  tagattttct cattggttct gccctcaagc actgaattca tctgaaactc
11701  ttggtttccc ctggaggcca tggttcctgg gcaccttgac ctgggcaatc
11751  ccaagtgtgg cctgaacccc cttcccttg gggattgttc aggtgtccct
11801  agacgccttg tggattgta cctaataccc atgaaggag aggatgatat
11851  tacttgccag tgtacacccc cctgtgatat tgttcataat gtccagagtg
11901  aagaaagatg atattactcc caatatcaca gaaggtgtac accccccctt
11951  gatattgttc ctaataccca gttggggagg ggagaatata tcttcaata
12001  tacaaggggt gtttaaactc tctgtgatat tgttcctaata attcaggggg
12051  gacaaggatg atattaccca aatattgcag gggttgtaca ccccccttt
12101  gatattgttc ctaatatcca ggggtggaga ggatattact cccaatattg
12151  caggggtcta catcctcccc ccgtgacatt gttcttaata accaaaaggt

```

图 1 续

```
12201  gagaagctga cattactccc aataccacag ggggtgtaca cccctatga
12251  gatattgttc ttaatatcca ggaggggaga aaatgatatt actctcaata
12301  gcgcaggga tttacatccc ccgtcgtaat cttgttctta atattcagga
12351  agggagagga tgatacgact ccagtatcg cagggggtgt gcaccccccc
12401  gtgatttttt tgctaataatc cagggtgga gtggataata cgcaggaagt
12451  gtacaggtct ctgtgatatt tttcctaata tccagggggg agaggaagat
12501  attactttta atagtgtacg gggggtgtac acccctctgt gatagtgttc
```

图 1 续

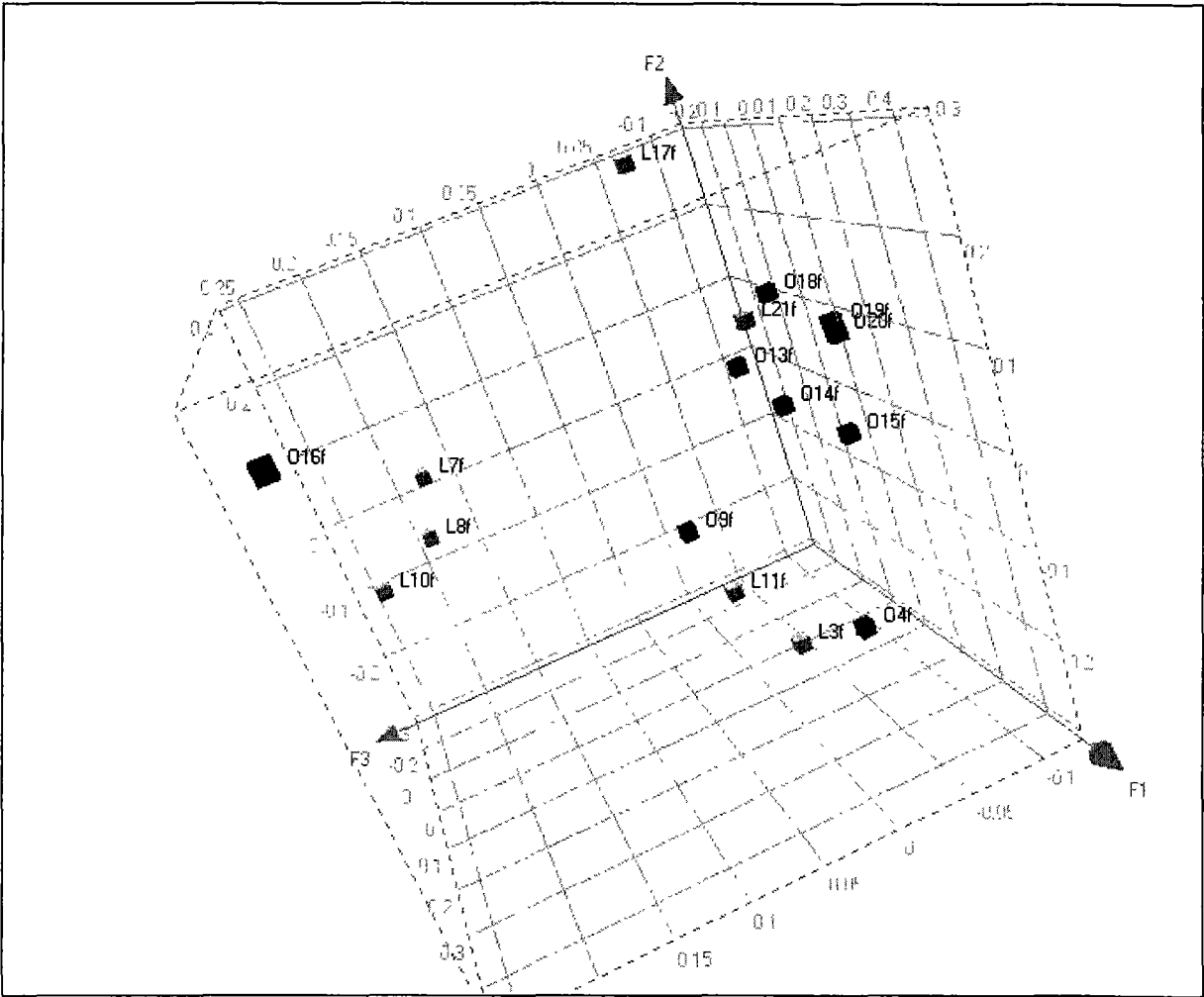


图 2 对应分析

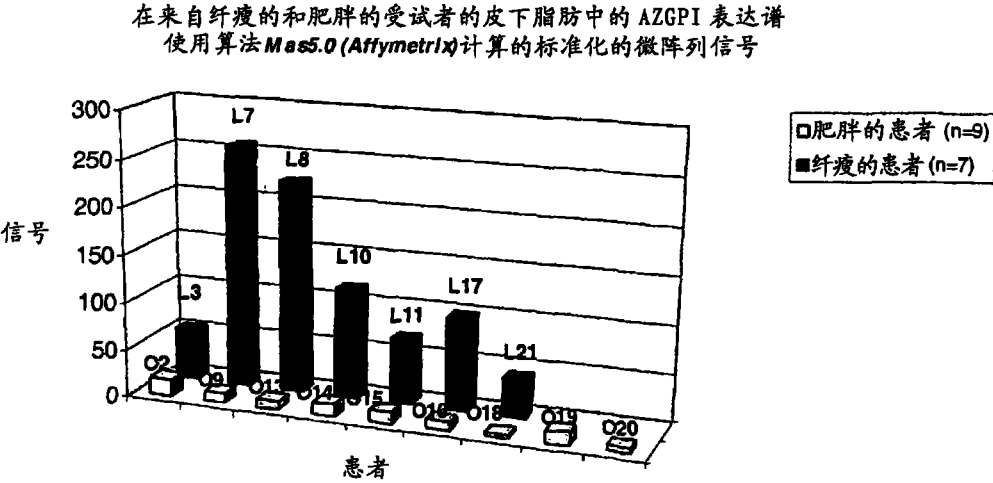


图 3