



(12) 发明专利

(10) 授权公告号 CN 102549157 B

(45) 授权公告日 2015. 07. 01

(21) 申请号 201080044858. 8

(22) 申请日 2010. 10. 05

(30) 优先权数据

2009-231991 2009. 10. 05 JP

2009-232001 2009. 10. 05 JP

2009-232002 2009. 10. 05 JP

2009-232003 2009. 10. 05 JP

2009-232004 2009. 10. 05 JP

2009-232035 2009. 10. 05 JP

(85) PCT国际申请进入国家阶段日

2012. 04. 05

(86) PCT国际申请的申请数据

PCT/JP2010/067441 2010. 10. 05

(87) PCT国际申请的公布数据

W02011/043330 JA 2011. 04. 14

(73) 专利权人 花王株式会社

地址 日本东京都

(72) 发明人 田口浩之 吉田宽 布施千绘

有波忠雄

(74) 专利代理机构 北京尚诚知识产权代理有限公司 11322

代理人 龙淳

(51) Int. Cl.

C12N 15/09(2006. 01)

C12Q 1/02(2006. 01)

C12Q 1/68(2006. 01)

(56) 对比文件

WO 2008043644 A2, 2008. 04. 17,

Sabeti et al..Genome-wide detection and characterization of positive selection in human populations. 《Nature》.2007, 第 449 卷

Botchkarev et al..Edar Signaling in the Control of Hair Follicle Development. 《J Investig Dermatol Symp Proc》.2005, (第 10 期),

审查员 赵重甲

权利要求书6页 说明书71页

序列表138页 附图15页

(54) 发明名称

毛发形状易感性基因

(57) 摘要

本发明提供一种与毛发形状相关的基因多态性和毛发形状易感性基因,以及用于判定各个被检测者中对毛发形状的遗传易感性的方法。所述毛发形状易感性基因是与以下单倍域相重叠的基因,并且是包含该单倍域的碱基序列的一部分或者全部的基因,其中,所述单倍域是人 1 号染色体的 1q21. 3 区域 (D1S2696 ~ D1S2346) 上的、通过对等位基因频率在具有卷发特性的群体和具有非卷发特性的群体之间在统计学上显著不同的单核苷酸多态性 (SNP) 标记进行连锁不平衡分析而确定的、并且由序列号 1 ~ 5 中任意一个所表示的碱基序列构成的单倍域。并且,所述毛发形状易感性基因是包含所述单倍域的碱基序列的一部分或者全部的基因。

1. 一种非治疗目的的毛发形状的调节方法,其中,

包括控制以下基因的表达的工序,

所述基因是与以下单倍域相重叠的基因,并且是包含该单倍域的碱基序列的一部分或者全部的基因,

其中,所述单倍域是人 1 号染色体的 1q21.3 区域 D1S2696 ~ D1S2346 上的、通过对等位基因频率在具有卷发特性的群体和具有非卷发特性的群体之间在统计学上显著不同的单核苷酸多态性 (SNP) 标记进行连锁不平衡分析而确定的、并且由序列号 1 ~ 5 中任意一个所表示的碱基序列构成的单倍域,

所述基因选自 ANXA9、CRCT1 以及 IVL。

2. 如权利要求 1 所述的方法,其中,

所述毛发形状的调节为毛发的直发化。

3. 如权利要求 1 所述的方法,其中,

所述基因选自 ANXA9、CRCT1,并且所述毛发形状的调节为毛发的直发化。

4. 如权利要求 1 所述的方法,其中,

所述毛发形状的调节为毛发的卷发化。

5. 如权利要求 1 所述的方法,其中,

所述基因为 IVL,并且所述毛发形状的调节为毛发的卷发化。

6. 一种对毛发形状的遗传易感性的判定标记,其中,

所述对毛发形状的遗传易感性的判定标记是选自 (1)~(5) 组中的任意组,其中每组由组内所有 SNP 位点组成

(1) 序列号 1 所表示的碱基序列中的 rs3754211、其 SNP 位点对应于 A 或者 G, rs3754210、其 SNP 位点对应于 T 或者 G, rs16832604、其 SNP 位点对应于 G 或者 A, rs2305814、其 SNP 位点对应于 C 或者 T, rs7532008、其 SNP 位点对应于 C 或者 A, rs1673160、其 SNP 位点对应于 T 或者 A, rs771205、其 SNP 位点对应于 T 或者 C 所表示的碱基的组合;其分别位于序列号 1 所表示的碱基序列中的碱基号 1、2405、5874、7121、8494、18980、23252;

(2) 序列号 2 所表示的碱基序列中的 rs11581947、其 SNP 位点对应于 A 或者 G, rs6658925、其 SNP 位点对应于 G 或者 A, rs2105117、其 SNP 位点对应于 A 或者 G, rs1053590、其 SNP 位点对应于 C 或者 T, rs548252、其 SNP 位点对应于 T 或者 C, rs493133、其 SNP 位点对应于 C 或者 G, rs1970283、其 SNP 位点对应于 C 或者 G, rs1001834、其 SNP 位点对应于 A 或者 C, rs11205018、其 SNP 位点对应于 G 或者 T, rs545418、其 SNP 位点对应于 T 或者 C, rs12116609、其 SNP 位点对应于 T 或者 C, rs526099、其 SNP 位点对应于 C 或者 T, rs525960、其 SNP 位点对应于 A 或者 T, rs4845443、其 SNP 位点对应于 A 或者 G, rs569032、其 SNP 位点对应于 T 或者 C, rs528427、其 SNP 位点对应于 C 或者 G, rs478926、其 SNP 位点对应于 T 或者 G, rs1337338、其 SNP 位点对应于 A 或者 G, rs6587681、其 SNP 位点对应于 T 或者 C, rs1856120、其 SNP 位点对应于 A 或者 G, rs474086、其 SNP 位点对应于 T 或者 C, rs578382、其 SNP 位点对应于 A 或者 G, rs549044、其 SNP 位点对应于 T 或者 C, rs1123567、其 SNP 位点对应于 A 或者 G, rs1538083、其 SNP 位点对应于 G 或者 A, rs1538082、其 SNP 位点对应于 G 或者 A, rs7532535、其 SNP 位点对应于 A 或者 T, rs7518654、其 SNP 位点对应于 G 或者 C, rs533917、其 SNP 位点对应于 G 或者 A, rs564107、其 SNP 位点对应于 T 或者 C, rs7530609、

其 SNP 位点对应于 A 或者 C, rs4240885、其 SNP 位点对应于 C 或者 G, rs4240886、其 SNP 位点对应于 A 或者 T, rs4240887、其 SNP 位点对应于 G 或者 A, rs6687126、其 SNP 位点对应于 G 或者 T, rs6674451、其 SNP 位点对应于 T 或者 C, rs7550769、其 SNP 位点对应于 A 或者 G, rs7529157、其 SNP 位点对应于 A 或者 C, rs1988805、其 SNP 位点对应于 G 或者 T, rs7529441、其 SNP 位点对应于 T 或者 C 所表示的碱基的组合 ; 其分别位于序列号 2 所表示的碱基序列中的碱基号 2355、2569、3897、8196、9510、13643、15387、15708、16017、17106、17453、17579、17634、26924、28383、31275、31301、31653、31903、32209、33199、33822、34100、35791、36884、37072、37365、37613、38062、39063、46580、49618、50164、50278、50662、50822、50981、51133、51263、以及 51397 ;

(3) 序列号 3 所表示的碱基序列中的 rs11205072、其 SNP 位点对应于 G 或者 A, rs3753453、其 SNP 位点对应于 T 或者 C, rs3737859、其 SNP 位点对应于 T 或者 G, rs3904414、其 SNP 位点对应于 G 或者 A, rs12074783、其 SNP 位点对应于 T 或者 C, rs3908717、其 SNP 位点对应于 C 或者 G, rs3904415、其 SNP 位点对应于 C 或者 T, rs11205079、其 SNP 位点对应于 A 或者 T 所表示的碱基的组合 ; 其分别位于序列号 3 所表示的碱基序列中的碱基号 2509、5167、8449、17598、18481、20891、21734 以及 23382 ;

(4) 序列号 4 所表示的碱基序列中的 rs16834715、其 SNP 位点对应于 T 或者 C, rs12022319、其 SNP 位点对应于 C 或者 T, rs4845490、其 SNP 位点对应于 G 或者 A, rs4845491、其 SNP 位点对应于 C 或者 T, rs3737861、其 SNP 位点对应于 C 或者 A, rs16834728、其 SNP 位点对应于 C 或者 T, rs4845492、其 SNP 位点对应于 C 或者 G 所表示的碱基的组合 ; 其分别位于序列号 4 所表示的碱基序列中的碱基号 1、3308、4715、4985、6354、8553、8818 ; 以及

(5) 序列号 5 所表示的碱基序列中的 rs1854779、其 SNP 位点对应于 C 或者 T, rs16834751、其 SNP 位点对应于 C 或者 A, rs4523473、其 SNP 位点对应于 C 或者 T, rs11205131、其 SNP 位点对应于 A 或者 G, rs7528862、其 SNP 位点对应于 G 或者 A, rs7517189、其 SNP 位点对应于 G 或者 C, rs2229496、其 SNP 位点对应于 G 或者 A, rs913996、其 SNP 位点对应于 A 或者 G 所表示的碱基的组合 ; 其分别位于序列号 5 所表示的碱基序列中的碱基号 1、540、759、1007、1018、1075、1939、3440。

7. 一种对被检测者的毛发形状的遗传易感性的判定方法, 其中,

包括以下工序 (a) ~ (c) :

(a) 制备来自被检测者的基因组 DNA 的工序 ;

(b) 从该基因组 DNA 中, 检测出权利要求 6 所述的判定标记中的 (1) - (5) 组中任意一组的工序 ; 以及

(c) 在该检测出的 SNP 的等位基因频率在卷发者群体中比在非卷发者群体中在统计学上显著高的情况下, 判定为该被检测者具有卷发的遗传素质 ; 在该被检测出的 SNP 的等位基因频率在任意非卷发者群体中比在卷发者群体中在统计学上显著高的情况下, 判定为该被检测者不具有卷发的遗传素质的工序。

8. 一种对被检测者的毛发形状的遗传易感性的判定方法, 其中,

包括以下的工序 :

在来自被检测者的基因组 DNA 中的序列号 1 ~ 5 所表示的碱基序列中, 识别下述表中

所示的每个序列中碱基号对应的碱基所组成的组中该碱基是碱基 (i) 还是碱基 (ii), 在为碱基 (i) 的情况下, 判定为被检测者具有卷发的素质 ; 在为碱基 (ii) 的情况下, 判定为被检测者不具有卷发的素质的工序,

序列号	碱基号	碱基(i) (有素质)	碱基(ii) (无素质)
1	1	G	A
	2405	G	T
	5874	A	G
	7121	T	C
	8494	A	C
	18980	A	T
	23252	C	T
2	2355	G	A
	2569	A	G
	3897	G	A
	8196	T	C
	9510	C	T
	13643	G	C
	15387	G	G
	15708	C	A
	16017	T	G
	17106	C	T
	17453	C	T
	17579	T	C
	17634	T	A
	26924	G	A
	28383	C	T
	31275	G	C
	31301	G	T
	31653	G	A

	31903	C	T
	32209	G	A
	33199	C	T
	33822	G	A
	34100	C	T
	35791	G	A
	36884	A	G
	37072	A	G
	37365	T	A
	37613	C	G
	38062	A	G
	39063	C	T
	46580	C	A
	49618	G	C
	50164	T	A
	50278	A	G
	50662	T	G
	50822	C	T
	50981	G	A
	51133	C	A
	51263	T	G
	51397	C	T

序列号	碱基号	碱基(i) (有素质)	碱基(ii) (无素质)
3	2509	A	G
	5167	C	T
	8449	G	T
	17598	A	G
	18481	C	T
	20891	G	C
	21734	T	C
	23382	T	A
4	1	C	T
	3308	T	C
	4715	A	G
	4985	T	C
	6354	A	C
	8553	T	C
	8818	G	C
5	1	T	C
	540	A	C
	759	T	C
	1007	G	A
	1018	A	G
	1075	C	G
	1939	A	G
	3440	G	A

o

9. 一种对被检测者的毛发形状的遗传易感性的判定用试药, 其中,

所述判定用试药含有在严格的条件下与权利要求 6 所述的对毛发形状的遗传易感性的判定标记杂交的探针和 / 或引物。

10. 如权利要求 9 所述的试药, 其中,
所述探针和 / 或引物与所述标记中含有权利要求 6 中所述的 SNP 的区域杂交。
11. 一种对被检测者的毛发形状的遗传易感性的判定用试剂盒, 其中,
所述判定用试剂盒含有权利要求 9 或 10 所述的试药。
12. 一种使用毛发形状类型的标记的毛发形状类型的检测和 / 或判定方法, 其中,
所述毛发形状类型的标记由多核苷酸构成; 或者由多肽构成,
其中, 所述多核苷酸由序列号 44、序列号 50 或者序列号 52 所表示的碱基序列构成;
所述多肽由序列号 45、序列号 51 或者序列号 53 所表示的氨基酸序列构成。
13. 一种使用用于扩增权利要求 12 中所述的标记的引物的毛发形状类型的检测和 / 或判定方法, 其中,
所述引物扩增权利要求 12 中所述的多核苷酸,
其中, 所述多核苷酸由序列号 44、序列号 50 或者序列号 52 所示的碱基序列构成。
14. 一种使用用于检测权利要求 12 中所述的标记的探针的毛发形状类型的检测和 / 或判定方法, 其中,
所述探针能够检测权利要求 12 中所述的多核苷酸,
其中, 所述多核苷酸由序列号 44、序列号 50 或者序列号 52 所示的碱基序列构成。
15. 一种使用用于检测权利要求 12 中所述的标记的抗体的毛发形状类型的检测和 / 或判定方法, 其中,
所述抗体特异性识别由序列号 45、序列号 51 或者序列号 53 所示的氨基酸序列构成的多肽。
16. 一种非治疗目的且非外科手术的毛发形状类型的检测和 / 或判定方法, 其中,
包括以下工序 (a) ~ (c):
(a) 测定在来自被检测者的样品中权利要求 12 中所述的标记的表达量的工序;
(b) 将该 (a) 的测定结果与非卷发者的测定结果相比较的工序; 以及,
(c) 基于 (b) 的结果判断毛发形状类型的工序。
17. 如权利要求 16 所述的方法, 其中,
所述来自被检测者的样品是从被检测者采集的样品中制备的 RNA。
18. 如权利要求 16 所述的方法, 其中,
所述工序 (a) 是使从被检测者采集的样品和权利要求 15 中所述的抗体相接触, 测定与该抗体结合的该样品中的权利要求 12 中所述的标记的量的工序。
19. 如权利要求 16 ~ 18 中任一项所述的方法, 其中,
从被检测者采集的样品为来自上皮组织或者上皮细胞的样品。
20. 一种非治疗目的且非外科手术的毛发形状调节剂的评价或者选择方法, 其特征在于,
包括以下 (a) ~ (d) 的工序:
(a) 使被检测物质与权利要求 1 所述的基因能够表达的细胞相接触的工序;
(b) 测定被接触的细胞的该基因的表达量的工序;
(c) 将 (b) 中测定的表达量与不接触被检测物质的对照细胞的该基因的表达量进行比较的工序;

(d) 基于 (c) 的结果,选择使该基因的表达量减少或者增多的被检测物质作为毛发形状调节剂的工序。

21. 一种非治疗目的且非外科手术的毛发形状调节剂的评价或者选择方法,其特征在于,

包括以下的 (a) ~ (c) 的工序:

(a) 向表达权利要求 1 所述的基因的细胞中,导入该基因的表达控制区域和报告基因的融合基因,并在被检测物质的存在条件下以及不存在条件下培养该细胞的工序;

(b) 测定在被检测物质的存在条件下培养的细胞培养物中的报告基因表达产物的表达量,并将该量与在被检测物质的不存在条件下培养的细胞培养物中的报告基因表达产物的表达量进行比较的工序;

(c) 基于所述 (b) 的比较结果,选择使报告基因表达产物的表达量增减的被检测物质作为毛发形状调节剂的工序。

毛发形状易感性基因

技术领域

[0001] 本发明涉及与毛发形状相关的基因、以及对于毛发形状的遗传易感性的判定、毛发形状类型的检测和 / 或判定、以及用于对毛发形状的调节有效的成分的筛选的标记、以及该标记的用途。

背景技术

[0002] 根据毛发的卷曲程度,人的自然毛发形状一般可以分为直发、波浪发 (wavy hair)、卷发 (curly hair)、卷缩发 (kinky hair、coiled hair)。因为毛发形状和发型是作为人的身体特征而容易识别的特性之一,另外,也是决定人的第一印象的重要因素,所以成为不论男女老少在美容上都重要关心的事。在卷曲程度强的卷缩发和卷发的情况下,因为发型的自由度受到限制而出现不能形成期望的造型等烦恼。另一方面,即使在直发的情况下,又会出现没有发量,容易看出素颜等烦恼。

[0003] 作为改变毛发形状和发型的方法,广泛进行的有通过各种定型剂和吹风机 / 烫发器进行的理发、卷曲 / 拉直烫发处理等。但是虽然这些处理能够有效地改变毛发的形状,却对决定毛发形状的原因完全没有效果。因此,这些解决对策不是根本的对策,只是一时的方法,要维持毛发的形状以及发型的话需要频繁地重复这些处理,反而会增大对毛发的伤害,结果损坏美容上的价值。因此,期待开发出在毛发生长出来的时候就能改变毛发的形状,从根本上调节毛发形状的调节方法。

[0004] 探索决定毛发形状的原因,希望在开发从根本上调节毛发形状的调节方法上,认定作为其原因的基因可以提供有用的信息。对于有关毛发形状的因素和基因,在引起毛发形状变化的遗传病 (非专利文献 1 ~ 3) 和药剂产生的后天的卷发 (非专利文献 4),或者卷发模型动物 (非专利文献 5、6) 等中有报告。但是,其中展示的因素和基因只不过是表示对毛发形状产生影响的一个特殊例子,很难认为是决定人的自然毛发的形状的原因。

[0005] 但是,近些年来基因组分析技术也有飞跃的进步,疾病和基因之间的关联性也慢慢被解开。特别是,不仅仅是在由单基因发生变异或者异常所产生的所谓遗传病中,而且在以糖尿病、高血压等生活习惯病为首的频率高的常见疾病 (Common Disease) 这样以低的外显率 (Penetrance :在某种基因中具有变异的个体患有某疾病的比例) 为特征的多基因性疾病中,频繁进行使用了患病同胞对连锁分析等的非参数连锁分析方法的发病基因 (responsible gene) 的探索 (例如,参照非专利文献 7)。另外,世界上也在大力进行基于常见疾病的疾病关联基因的变异为频率高的基因多态性 (Common Variant),虽然在健康人中也存在,但是在患者身上保有率显著较高的假说 (Common Disease-Common Variant),通过使用了基因多态性 (例如, SNP (Single Nucleotide Polymorphism) :单核苷酸多态性) 的连锁不平衡分析来寻找发病基因 (例如,参照非专利文献 8)。

[0006] 进一步,最近根据国际人类基因单倍体型图计划 (The International HapMap Project) 的进展,整理了 4 个人群体中 100 万处以上的频率高的一般多态性 (SNP) 的数据库,不仅是对常见疾病,还对关系到取决于人种或者群体其表现型不同的一般特性、例如皮

肤颜色和头发颜色、眼睛颜色等的基因的研究也在进行中（例如，参照非专利文献 9、10）。

[0007] 认为关于人的自然毛发的形状也是同样，也是根据人种或者群体其表现型不同的一般的特性。概括来说，亚洲人种多为直发，非洲人种主要为卷缩发（或者卷发）。印欧人种则为这中间的波浪发 (wavy hair) 特性的比例较高。对于此遗传模式，Rostand, J 等首先观察，报告了相对于直发，卷发为常染色体（半）显性的特性（非专利文献 11）。另外，NCBI 的人的孟德尔式遗传数据库 (OMIM、<http://www.ncbi.nlm.nih.gov/omim/>) 中也有关于卷发特性的记载。但是，对于成为决定人的自然的头发形状的原因的基因，还没有彻底地进行系统的基因组分析研究，因此尚未发现。

[0008] 现有技术文献

[0009] 非专利文献

[0010] 非专利文献 1 :Norgett EE et al., Hum. Mol. Genet. 9(18), p. 2761-2766, 2000

[0011] 非专利文献 2 :Moller LB et al., Hum. Mutat. 26(2), p. 84-93, 2005

[0012] 非专利文献 3 :Kjaer KW et al., Am. J. Med. Genet. A. 127A(2), p. 152-157, 2004

[0013] 非专利文献 4 :Cullen SI et al., Arch. Dermatol. 125(2), p. 252-255, 1989

[0014] 非专利文献 5 :Du X et al. Genetics. 166(1), p. 331-340, 2004

[0015] 非专利文献 6 :Mann GB et al., Cell. 73(2), p. 249-61, 1993

[0016] 非专利文献 7 :Hanis CL et al., Nat. Genet. 13(2), p. 161-166, 1996

[0017] 非专利文献 8 :Altshuler D et al., Nat. Genet. 26(1), p. 76-80, 2000

[0018] 非专利文献 9 :Sulem P et al., Nat. Genet. 39(12), p. 1443-1452, 2007

[0019] 非专利文献 10 :Sabeti PC et al., Nature. 449(7164), p. 913-918, 2007

[0020] 非专利文献 11 :Rostand J et al., [An Atlas of Human Genetics], Hutchinson Scientific & Technical, London, pp. 26-29, 1964

发明内容

[0021] 本发明提供一种毛发形状易感性基因，所述毛发形状易感性基因是与以下单倍域相重叠的基因，并且是包含该单倍域的碱基序列的一部分或者全部的基因，其中，所述单倍域是人 1 号染色体的 1q21.3 区域 (D1S2696 ~ D1S2346) 上的、通过对等位基因频率在具有卷发特性的群体和具有非卷发特性的群体之间在统计学上显著不同的单核苷酸多态性 (SNP) 标记进行连锁不平衡分析而确定的、并且由序列号 1 ~ 5 中任意一个所表示的碱基序列构成的单倍域。

[0022] 另外，本发明还提供一种毛发形状判定标记，所述毛发形状判定标记是包含以下部分碱基序列的寡核苷酸或者多核苷酸或者它们的互补链，其中，所述部分碱基序列是上述单倍域的碱基序列中的部分碱基序列，并且，所述部分碱基序列由连续的碱基序列构成，所述连续的碱基序列含有一个以上的单核苷酸多态性，所述单核苷酸多态性包括等位基因频率在具有卷发特性的群体和具有非卷发特性的群体之间在统计学上显著不同的单核苷酸多态性 (SNP)、以及与该 SNP 连锁的 SNP。

[0023] 另外，本发明还提供一种对被检测者的毛发形状的遗传易感性的判定方法，其中，包括以下工序 (a) ~ (c)：

[0024] (a) 调制来自被检测者的基因组 DNA 的工序；

[0025] (b) 从该基因组 DNA 中,检测出存在于上述单倍域中,等位基因频率在具有卷发特性的群体和具有非卷发特性的群体之间在统计学上显著不同的单核苷酸多态性 (SNP) 以及和该 SNP 连锁的单核苷酸多态性 (SNP) 的工序;以及

[0026] (c) 在该检测出的 SNP 的等位基因频率在卷发者群体中比在非卷发者群体中在统计学上显著高的情况下,判定为该被检测者具有卷发的遗传素质,在该被检测出的 SNP 的等位基因频率在任意非卷发者群体中比在卷发者群体中在统计学上显著高的情况下,判定为该被检测者不具有卷发的遗传素质的工序。

[0027] 另外,本发明还提供一种对被检测者的毛发形状的遗传易感性的判定方法,其中,所述判定方法包含在来自被检测者的基因组 DNA 中的序列号 1~5 所表示的碱基序列中的下述表中所示的碱基号的碱基的任意一个以上中,识别该碱基是碱基 (i) 还是碱基 (ii),在为碱基 (i) 的情况下,判定为被检测者具有卷发的素质,在为碱基 (ii) 的情况下,判定为被检测者不具有卷发的素质的工序。

[0028] [表 1-1]

[0029]

序列号	碱基号	碱基(i) (有素质)	碱基(ii) (无素质)
1	1	G	A
	2405	G	T
	5874	A	G
	7121	T	C
	8494	A	C
	18980	A	T
	23252	C	T
2	2355	G	A
	2569	A	G
	3897	G	A
	8196	T	C
	9510	C	T
	13643	G	C
	15387	G	G
	15708	C	A
	16017	T	G
	17106	C	T
	17453	C	T
	17579	T	C
	17634	T	A
	26924	G	A
	28383	C	T
	31275	G	C
	31301	G	T
	31653	G	A
	31903	C	T
	32209	G	A
	33199	C	T
	33822	G	A
	34100	C	T
	35791	G	A
	36884	A	G
	37072	A	G
	37365	T	A
	37613	C	G
	38062	A	G
	39063	C	T
	46580	C	A
	49618	G	C
	50164	T	A
	50278	A	G
	50662	T	G
	50822	C	T
	50981	G	A
	51133	C	A
	51263	T	G
	51397	C	T

[0030] [表 1-2]

[0031]

序列号	碱基号	碱基(i) (有素质)	碱基(ii) (无素质)
3	2509	A	G
	5167	C	T
	8449	G	T
	17598	A	G
	18481	C	T
	20891	G	C
	21734	T	C
	23382	T	A
4	1	C	T
	3308	T	C
	4715	A	G
	4985	T	C
	6354	A	C
	8553	T	C
	8818	G	C
5	1	T	C
	540	A	C
	759	T	C
	1007	G	A
	1018	A	G
	1075	C	G
	1939	A	G
	3440	G	A

[0032] 另外,本发明还提供一种对被检测者的毛发形状的遗传易感性的判定用试药,其中,所述判定用试药含有在严格的条件下与本发明的毛发形状判定标记杂交的探针和 / 或引物。

[0033] 另外,本发明还提供一种对被检测者的毛发形状的遗传易感性的判定用试剂盒,其中,所述判定用试剂盒含有上述试药。

[0034] 另外,本发明还提供一种筛选毛发形状调节剂的方法,其中,包括以下工序 (a) 和 (b) :

[0035] (a) 用被检测物质向含有本发明的毛发形状易感性基因的细胞给药的工序 ;以及

[0036] (b) 从给药的被检测物质中,选择能够使存在于该毛发形状易感性基因上或者其附近、并且等位基因频率在具有卷发特性的群体和具有非卷发特性的群体之间在统计学上显著不同的单核苷酸多态性 (SNP) 以及和该 SNP 连锁的单核苷酸多态性 (SNP) 标记的碱基的多态性变换为其它多态性的物质作为毛发形状调节剂的工序。

[0037] 另外,本发明提供一种毛发形状类型的标记,其中,所述毛发形状类型的标记由多核苷酸或者其部分多核苷酸构成 ;或者由多肽或者其部分多肽构成,其中,所述多核苷酸由序列号 44、序列号 46、序列号 48、序列号 50 或者序列号 52 所表示的碱基序列、与这些碱基序列互补的碱基序列构成 ;所述多肽由序列号 45、序列号 47、序列号 49、序列号 51 或者序列号 53 所表示的氨基酸序列构成。

[0038] 另外,本发明提供一种用于扩增本发明的毛发形状类型的标记的引物,其中,所述引物由多核苷酸的部分多核苷酸构成,其中,所述多核苷酸由序列号 44、序列号 46、序列号 48、序列号 50 或者序列号 52 所示的碱基序列或者与其互补的碱基序列构成。

[0039] 另外,本发明提供一种用于检测本发明的毛发形状类型的标记的探针,其中,所述探针由多核苷酸或者其部分多核苷酸构成,其中,所述多核苷酸由序列号 44、序列号 46、序列号 48、序列号 50 或者序列号 52 所示的碱基序列或者与其互补的碱基序列构成。

[0040] 另外,本发明提供一种用于检测本发明的毛发形状类型的标记,抗体,其中,所述抗体特异性识别由序列号 45、序列号 47、序列号 49、序列号 51 或者序列号 53 所示的氨基酸序列构成的多肽或者其部分多肽。

[0041] 另外,本发明提供一种本发明毛发形状类型的检测和 / 或判定方法,其中,包括以下工序 (a) ~ (c) :

[0042] (a) 测定在来自被检测者的样品中本发明的毛发形状类型的标记的表达量的工序 ;

[0043] (b) 将该 (a) 的测定结果与非卷发者的测定结果相比较的工序 ;以及,

[0044] (c) 基于 (b) 的结果判断毛发形状类型的工序。

[0045] 另外,本发明还提供一种毛发形状调节剂的评价或者选择方法,其中,包括以下 (a) ~ (d) 的工序 :

[0046] (a) 使被检测物质与本发明的毛发形状易感性基因或者由该基因编码的蛋白质能够表达的细胞相接触的工序 ;

[0047] (b) 测定被接触的细胞的该基因或者该蛋白质的表达量的工序 ;

[0048] (c) 将 (b) 中测定的表达量与不接触被检测物质的对照细胞的该基因或者该蛋白质的表达量进行比较的工序 ;

[0049] (d) 基于 (c) 的结果,选择使该基因或者该蛋白质的表达量减少或者增多的被检测物质作为毛发形状调节剂的工序。

[0050] 另外,本发明还提供一种毛发形状调节剂的评价或者选择方法,其特征在于,包括以下的 (a) ~ (c) 的工序 :

[0051] (a) 向本发明的毛发形状类型的毛发形状易感性基因能够表达的细胞中,导入该毛发形状易感性基因的表达控制区域和报告基因的融合基因,并在被检测物质的存在条件下和不存在条件下培养该细胞的工序 ;

[0052] (b) 测定在被检测物质的存在条件下培养的细胞培养物中的报告基因表达产物的表达量,并将该量和在被检测物质的不存在条件下培养的细胞培养物中的报告基因表达产物的表达量进行比较的工序 ;

[0053] (c) 基于所述 (b) 的比较结果,选择使报告基因表达产物的表达量增加或者减少的被检测物质作为毛发形状调节剂的工序。

[0054] 另外,本发明提供一种毛发形状调节剂的评价或者选择方法,其特征在于,包括以下的 (a) ~ (c) 的工序 :

[0055] (a) 使被检测物质和含有本发明的毛发形状易感性基因编码的蛋白质的水溶液、细胞或者由该细胞调制的细胞组分接触的工序 ;

[0056] (b) 测定接触了该被检测物质的水溶液、细胞或者细胞组分中的该蛋白质的机能或者活性,并将该机能或者活性与不接触被检测物质的对照水溶液、对照细胞或者对照细胞组分中所述蛋白质的机能或者活性比较的工序 ;以及

[0057] (c) 基于所述 (b) 的比较结果,选择使该蛋白质的机能或者活性增加或者减少的

被检测物质作为毛发生态调节剂的工序。

[0058] 另外,本发明还提供一种毛发生态类型的调节方法,其特征在于,控制人的头发发根部的本发明的毛发生态易感性基因的表达。

[0059] 在一个实施方式中,本发明的毛发生态易感性基因选自 ANXA9、FAM63A、LCE5A、CRCT1、LCE2B、LCE2A、SMCP 以及 IVL。

[0060] 在本发明的毛发生态判定标记的一个实施方式中,所述 SNP 是选自以下的碱基中的 SNP:

[0061] (1) 序列号 1 所表示的碱基序列中,碱基号 1(dbSNP 数据库 ID:rs3754211、A 或者 G)、2405(rs3754210、T 或者 G)、5874(rs16832604、G 或者 A)、7121(rs2305814、C 或者 T)、8494(rs7532008、C 或者 A)、18980(rs1673160、T 或者 A)、以及 23252(rs771205、T 或者 C) 所表示的碱基;

[0062] (2) 序列号 2 所表示的碱基序列中,碱基号 2355(rs11581947、A 或者 G)、2569(rs6658925、G 或者 A)、3897(rs2105117、A 或者 G)、8196(rs1053590、C 或者 T)、9510(rs548252、T 或者 C)、13643(rs493133、C 或者 G)、15387(rs1970283、C 或者 G)、15708(rs1001834、A 或者 C)、16017(rs11205018、G 或者 T)、17106(rs545418、T 或者 C)、17453(rs12116609、T 或者 C)、17579(rs526099、C 或者 T)、17634(rs525960、A 或者 T)、26924(rs4845443、A 或者 G)、28383(rs569032、T 或者 C)、31275(rs528427、C 或者 G)、31301(rs478926、T 或者 G)、31653(rs1337338、A 或者 G)、31903(rs6587681、T 或者 C)、32209(rs1856120、A 或者 G)、33199(rs474086、T 或者 C)、33822(rs578382、A 或者 G)、34100(rs549044、T 或者 C)、35791(rs1123567、A 或者 G)、36884(rs1538083、G 或者 A)、37072(rs1538082、G 或者 A)、37365(rs7532535、A 或者 T)、37613(rs7518654、G 或者 C)、38062(rs533917、G 或者 A)、39063(rs564107、T 或者 C)、46580(rs7530609、A 或者 C)、49618(rs4240885、C 或者 G)、50164(rs4240886、A 或者 T)、50278(rs4240887、G 或者 A)、50662(rs6687126、G 或者 T)、50822(rs6674451、T 或者 C)、50981(rs7550769、A 或者 G)、51133(rs7529157、A 或者 C)、51263(rs1988805、G 或者 T)、以及 51397(rs7529441、T 或者 C) 所表示的碱基;

[0063] (3) 序列号 3 所表示的碱基序列中,碱基号 2509(rs11205072、G 或者 A)、5167(rs3753453、T 或者 C)、8449(rs3737859、T 或者 G)、17598(rs3904414、G 或者 A)、18481(rs12074783、T 或者 C)、20891(rs3908717、C 或者 G)、21734(rs3904415、C 或者 T) 以及 23382(rs11205079、A 或者 T) 所表示的碱基;

[0064] (4) 序列号 4 所表示的碱基序列中,碱基号 1(rs16834715、T 或者 C)、3308(rs12022319、C 或者 T)、4715(rs4845490、G 或者 A)、4985(rs4845491、C 或者 T)、6354(rs3737861、C 或者 A)、8553(rs16834728、C 或者 T)、以及 8818(rs4845492、C 或者 G) 所表示的碱基;以及

[0065] (5) 序列号 5 所表示的碱基序列中,碱基号 1(rs1854779、C 或者 T)、540(rs16834751、C 或者 A)、759(rs4523473、C 或者 T)、1007(rs11205131、A 或者 G)、1018(rs7528862、G 或者 A)、1075(rs7517189、G 或者 C)、1939(rs2229496、G 或者 A)、以及 3440(rs913996、A 或者 G) 所表示的碱基。

[0066] 在其它的实施方式中,上述毛发生态判定标记由 10 ~ 601 个碱基长度的连续碱基

序列构成。

[0067] 在对于本发明的被检测者的毛发形状的遗传易感性的判定用试药的一个实施方式中,上述探针和/或引物与上述(1)~(5)记载的碱基中含有 SNP 的区域杂交。

[0068] 在本发明的毛发形状类型的标记的一个实施方式中,部分多核苷酸为 15 个碱基以上的多核苷酸。

[0069] 在本发明的毛发形状类型的检测和/或判定方法的一个实施方式中,来自被检测者的样品是由从被检测者采集的活体样品中调制的 RNA 或者由该 RNA 转录的互补多核苷酸。

[0070] 在本发明的毛发形状类型的检测和/或判定方法的其它实施方式中,上述工序(a)是使从被检测者采集的活体样品和用于检测本发明的毛发形状类型的标记的抗体相接触,测定与该抗体结合的该活体样品中的本发明的毛发形状类型的标记的量的工序。

[0071] 在本发明的毛发形状类型的检测和/或判定方法的另外其它实施方式中,从被检测者采集的活体样品为来自上皮组织或者上皮细胞的活体样品。

附图说明

[0072] [图 1] 毛发形状的表现型的图。

[0073] [图 2] 1 号染色体中由患病同胞对连锁分析得到的微卫星标记和最大 LOD。

[0074] [图 3] 11 号染色体中由患病同胞对连锁分析得到的微卫星标记和最大 LOD。

[0075] [图 4] 1 号染色体中由患病同胞对连锁分析得到的微卫星标记和最大 LOD。

[0076] [图 5] 含有 SNP :rs2305814,从 SNP :rs3754211 到 SNP :rs771205 的由序列号 1 所表示的碱基序列所示的 23, 252bp 构成的单倍域的概念图。

[0077] [图 6] 含有 SNP :rs1053590,从 SNP :rs1885531 到 SNP :rs4845791 的由序列号 2 所表示的碱基序列所示的 56, 552bp 构成的单倍域的概念图。

[0078] [图 7] 含有 SNP :rs3753453、SNP :rs3737859、SNP :rs3904414、以及 SNP :rs3908717,从 SNP :rs2014369 到 SNP :rs11205079 的由序列号 3 所表示的碱基序列所示的 23, 382bp 构成的单倍域的概念图。

[0079] [图 8] 含有 SNP :rs3737861,从 SNP :rs16834715 到 SNP :rs4845492 的由序列号 4 所表示的碱基序列所示的 8, 818bp 构成的单倍域的概念图。

[0080] [图 9] 含有 SNP :rs4523473、SNP :rs2229496、以及 SNP :rs913996,从 SNP :rs1854779 到 SNP :rs913996 的由序列号 5 所表示的碱基序列所示的 3, 440bp 构成的单倍域的概念图。

[0081] [图 10-1] 表示卷发群和直发群的头发发根上的毛发形状易感性基因的表达量的图。A :ANXA9 基因、B :CRCT1 基因。

[0082] [图 10-2] 表示卷发群和直发群的头发发根上的毛发形状易感性基因的表达量的图。C :LCE2B 基因、D :LCE2A 基因。

[0083] [图 10-3] 表示卷发群和直发群的头发发根上的毛发形状易感性基因的表达量的图。E :IVL 基因。

[0084] [图 11] 表示各人种毛囊组织像的照片。箭头表示弯曲部位。

[0085] [图 12] 表示在人的毛囊器官培养体系中随着培养天数的经过的毛囊形态变化的

照片。

[0086] [图 13] 表示毛发形状易感性基因表达调节剂对于毛囊形态的效果的图。A:圆叶牵牛 (*Ipomoea purpurea*)、B:西番莲 (*Passiflora*)。

具体实施方式

[0087] 本发明涉及提供与卷发或直发这样的人的自然毛发形状相关的基因多态性以及毛发形状易感性基因,进一步提供以这些信息为基础,用于判定对各个被检测者的毛发形状的遗传易感性的方法。另外,本发明涉及提供用于简单实施该方法的有用的试药以及试药试剂盒。进一步,本发明涉及提供用于检测・判定卷发和直发这样的人的自然毛发形状类型的标记(多核苷酸、多肽),以及在使用该标记的毛发形状类型的检测和/或判定、和对毛发形状的调节有效的成分的评价・选择等中的该标记的用途。

[0088] 本发明者们以寻找作为决定人的自然毛发形状的原因的基因为目的,以日本人卷发家族、以及日本人卷发者群体和非卷发者群体为对象进行基因组分析,其结果在 1 号染色体 1q21.3 区域上认定与毛发形状相关的基因多态性、也就是毛发形状易感性 SNP 标记,并且认定毛发形状易感性基因。另外,本发明者们探索了毛发形状和毛发发根部处各种基因的基因表达之间的关系,结果发现在发根部非卷发者和卷发者之间上述毛发形状易感性基因的表达量显著不同。这些基因为毛发形状易感性基因,可以作为用于检测和/或判定毛发形状类型的标记。基于这些发现,本发明者们至此完成本发明。

[0089] 根据本发明,能够提供与卷发和直发的人的自然的毛发形状关联的毛发形状易感性基因以及毛发形状易感性 SNP 标记、以及利用了这些的毛发形状判定标记。通过对本发明的毛发形状易感性基因和 SNP 标记、以及毛发形状判定标记进行详细地分析,可以进行与毛发形状相关联的毛发形成的机理的研究、以及用于实现毛发形状的调节的确切方法的开发等的应用研究。

[0090] 另外,通过本发明的对被检测者的毛发形状的遗传易感性的判定方法,可以更加简便并且迅速地进行决定各个被检测者的毛发形状的主要原因的基因的探索,以及对各个被检测者的后天的毛发形状的变化易感性,也就是将来毛发形状变化的风险度的判定。进一步,基于此结果,可以提供对于各个人的确切的毛发形状的调节方法。另外,通过本发明的对被检测者的毛发形状的遗传易感性的判定用试药以及含有该试药的试剂盒,可以进一步简便并且迅速地实施上述判定方法。

[0091] 另外,根据本发明,可以不对毛发造成伤害而检测・判定卷发、卷缩发这样的毛发形状或者性状。另外,通过本发明的筛选对调节毛发形状有效的成分的方法而选择的物质,可以作为对毛发形状调节有效的毛发形状调节剂使用,另外,还可以用于含有该试剂的医药品、医药部外品、化妆品、或者健康食品等的制造中使用。另外,根据本发明,可以提供利用通过本发明所得到的毛发形状易感性 SNP 标记的毛发形状的调节方法。

[0092] 1. 本发明所使用的用语的定义

[0093] 本说明书中的碱基序列(核苷酸序列)、核酸等的通过略号的表示是根据 IUPAC-IUB 的规定(IUPAC-IUB communication on Biological Nomenclature, Eur. J. Biochem., 138, 9, 1984)、《用于制作包含碱基序列或者氨基酸序列的说明书等的指导原则》(日本特许厅编)、以及该技术领域中惯用记号而进行的表示。

[0094] 在本说明书中“DNA”不仅包含双链 DNA,还包含构成该双链 DNA 的有义链以及反义链这样的各单链 DNA。本说明书中“基因”,只要是没有特别提到,包含含有人类基因组 DNA 的双链 DNA、以及单链 DNA(有义链)、以及具有与该有义链互补的序列的单链 DNA(反义链)、以及这些 DNA 的片段中的任意种。另外,在本说明书中“基因”只要是没有特别提到,不区分表示调节区域、编码区域、外显子以及内含子。另外,在该“基因”或者“DNA”中,不仅包括以特定碱基序列表示的“基因”或者“DNA”,还以与其所编码的蛋白质在生物学机能上同等为限,包括编码其同源物(homologue)、衍生物以及变异体的“基因”或者“DNA”。

[0095] 另外,在本说明书中,“核苷酸”、“寡核苷酸”以及“多核苷酸”与核酸同义,包括 DNA、RNA 这两者。该 DNA 中含有 cDNA、基因组 DNA 以及合成 DNA 中的任意种。另外,该 RNA 中包含总 RNA(total RNA)、mRNA、rRNA、以及合成的 RNA 中的任意种。另外,“核苷酸”、“寡核苷酸”、以及“多核苷酸”可以为双链也可以为单链,在称作具有某种序列的“核苷酸”(或者“寡核苷酸”、“多核苷酸”)的情况下,只要不特别提到,也包含具有与其互补的序列的“核苷酸”(或者“寡核苷酸”、“多核苷酸”)的意思。进一步,在“核苷酸”(或者“寡核苷酸”、“多核苷酸”)为 RNA 的情况下,序列表所示的碱基记号“T”改称为“U”。

[0096] “由互补的碱基序列构成的多核苷酸”是指相对于由任意的碱基序列(有义链)构成的多核苷酸具有碱基互补关系的多核苷酸(互补链、反义链)。互补的碱基序列中,除了形成与作为对象的碱基序列完全互补的序列的情况之外,还包含与在严格条件下杂交所得到的碱基序列。在此,严格条件可以举通常“ $1\times\text{SSC}$ 、 0.1% SDS、 37°C ”左右的清洁条件,作为更加严格的杂交条件可以举“ $0.5\times\text{SSC}$ 、 0.1% SDS、 42°C ”左右,作为进一步严格的杂交条件可以举“ $0.1\times\text{SSC}$ 、 0.1% SDS、 65°C ”左右的条件。另外,本领域普通技术人员可以根据一般的教导(例如,Sambrook, J. & Russell, D., 2001, Molecular Cloning :a Laboratory Manual, 3rd edition. Cold Spring Harbor, NY :Cold Spring Harbor Laboratory)来决定严格的杂交条件。例如,作为在严格的条件下与作为对象的碱基序列杂交得到的碱基序列,可以举与作为对象的碱基序列具有 90% 以上、优选为 95% 以上的同一性的碱基序列。

[0097] “蛋白质”或者“多肽”中不仅包含以特定的碱基序列或者氨基酸序列表示的“蛋白质”或者“多肽”,还以与其生物学上机能同等为限,包含其片段、同源物、衍生物以及变异体。另外,上述变异体中包含天然存在的等位基因变异体、天然不存在的变异体以及具有通过人为的缺失、取代、添加以及插入而改变的氨基酸序列的变异体。另外,作为上述变异体,可以举与未变异的蛋白质或者多肽在氨基酸序列中具有 80% 以上、更加优选为 90% 以上、进一步优选为 95% 以上,更加进一步优选为 98% 以上的同一性的变异体。

[0098] 在本说明书中,氨基酸序列以及碱基序列的同一性是通过 Lipman-Pearson 法(Science, 227, 1435, 1985)计算的。具体来说,是使用遗传信息处理软件 Genetyx-Win(软件开发)的同源分析(Search homology)程序,并使参数 Unit size to compare(ktup) 为 2 进行分析而算出来的。

[0099] “抗体”中,包含多克隆抗体、单克隆抗体、嵌合抗体、单链抗体、以及 Fab 片段、Fab 表达库等生成的片段样的具有抗原结合性的上述抗体的一部分。

[0100] 在本说明书中,“基因多态性”是指在具有 2 个以上的遗传所决定的等位基因的情况下,指这些的等位基因。具体来说,在人的群体之中,以某个个体的基因组序列作为基准,在其它的一个或多个个体基因组中的特定部位上,存在一个或多个核苷酸的取代、缺失、插

入、转位、反位等变异的时候,此变异在统计学上确实不是在该 1 个或者多个个体内产生的突然变异,或者不是该个体内的特异变异,而是在遗传学上证明以 1% 以上的频率在群体内存在的情况,将该变异作为“基因多态性”。在本说明书中所用的“基因多态性”中,有例如,一个碱基被其它碱基取代的情况,也就是单核苷酸多态性 (SNP:Single Nucleotide Polymorphism)、1~数十个碱基缺失或者被插入的情况 (DIP)、2~数十个碱基作为 1 个单元的序列重复存在的部位中其重复次数不同的情况(将重复单元为 2~4 个碱基的情况称为微卫星多态性,将数个~数十个碱基的情况称为 VNTR(Variable Number of Tandem Repeat))等。

[0101] 在本说明书中“毛发形状”是指直发、波状发(波浪发(Wavy hair))、卷发、卷缩发(kinky hair、Coiled hair)等由于一根一根毛发的形状导致的头部的毛发整体的形状的倾向。

[0102] 在本说明书中“卷发”,如果没有特别的说明的话,是概括地指相比于直发的情况下直发以外的形状的用语。因此,在本说明书中,相比于“卷发”的情况,如果没有特别的说明,“直发”和“非卷发”是按相同意义来使用的。“卷发”、“非卷发”以及“直发”是相对的性质,可以按照后述的种种方法定义。“卷发特性”、“非卷发特性”以及“直发特性”是指表示“卷发”、“非卷发”以及“直发”的表现型。

[0103] 在本说明书中,“毛发形状易感性基因”是指成为决定作为多基因性的特性的毛发形状的原因的基因,“毛发形状易感性 SNP 标记”是指表示与该个体的毛发形状的特性具有关联性的 SNP 的部位碱基。

[0104] 在本说明书中,“对毛发形状的遗传易感性”、“毛发形状判定标记”、以及“毛发形状类型的标记”,分别是指与个体所保有的特定毛发形状相关的遗传素质、以及用于判定该素质的标记。

[0105] 在本说明书中,“患病同胞对连锁分析(Affected Sib-Pair Linkage Analysis)”是作为利用连锁来推定目的基因(疾病易感性基因等)的存在场所的方法之一,指不假定遗传方式(常染色体显性遗传、隐性遗传、性连锁基因等)和外显率的非参数连锁分析的代表分析方法。在患病同胞对连锁分析中,集合包括患病的(或者有特定的特性)的同胞(兄弟姐妹)的家族,以此家族的观察数据为基础进行似然(likelihood)计算,锁定和疾病(或者有特定的特性)连锁的标记基因位点区域。在一般的(没有患病,或者没有特定的特性)的同胞的群体中,如果对一个基因位点进行考虑,孩子得到双亲中的一方的 2 个等位基因内的 1 个(考虑双亲中的一方即使为同型接合体其各自的等位基因也不同)。因此,同胞在这种时候,会有与接受相同等位基因的情况不相同的接受等位基因的情况。因为孩子的 2 个等位基因来源为来自双亲中的各 1 个,因此如果考虑双亲的来源而认为同胞接受了哪个相同等位基因的话,有 0、1、2 的情况。将这 3 个情况分别称为 IBD(Identity By Descent)为 0、1、2。如果考虑多个同胞对的话,应该可以数出 $IBD = 0$ 的对、 $IBD = 1$ 的对、 $IBD = 2$ 的对的数,并且该数的比例根据概率法而成为一定的比例(1 : 2 : 1)。对此,集中患病(或者具有特定特性)的同胞对此群体进行同样的调查时,认为如果观察到的标记基因与疾病(或者特定特性)连锁的情况下会偏离此比例(1 : 2 : 1)($IBD = 2$ 的对的数量增加、 $IBD = 0$ 的对的数量减少)。同时,认为在与疾病(或者特定的特性)所关联的基因不连锁的标记基因中,和任意的同胞形成相同的分配(1 : 2 : 1)。在患病同胞对连锁分析中,利用此

假设,将在患病同胞对中等位基因的共有比例的偏离作为指标算出观察数据的似然。似然通过下式表示。

$$[0106] \quad L(Z) = \prod_{j=1}^N \sum_{i=0}^2 Z_i W_{ij}$$

[0107] 在此, W_{ij} 为第 j 号家族中患病同胞对为 $IBD = i$ 的概率。变数 $Z = (Z_0, Z_1, Z_2)$, 自由度为 2 ($Z_2 = 1 - Z_1 - Z_0$, 独立变数仅为 Z_0, Z_1 两个)。取得当标记基因与疾病 (或者特定的特性) 所关联的基因不连锁时 (也就是, $Z_0 = 0.25, Z_1 = 0.5, Z_2 = 0.25$) 的似然比, 通过似然最大化法 (最大似然法), 求得似然最大的 Z 。

[0108] 在本说明书中,“基因频率”是指对于一个基因位点,在群体中存在的全部基因数中,其等位基因所占的比例。

[0109] 在本说明书中,“单倍体型”是指在一个等位基因 (单倍体) 中存在的基因变异的组合。

[0110] 在本说明书中,“连锁不平衡分析”或者“单倍体型分析”是指分析在基因组区域中连锁不平衡的强度的程度。

[0111] 在本说明书中,“连锁不平衡”是指在群体中,发现多个基因位点的等位基因或者遗传标记 (多态性) 之间不是随机相关,也就是这些特定组合 (单倍体型) 的频率显著高的群体遗传学现象。这些一般是存在于同一染色体上的遗传连锁,但是有时候即使连锁也有不能发现连锁不平衡的情况,另外,例外的也会有在别的染色体上发现的情况。例如,在基因位点 X 上具有等位基因 a 和 b (它们以相同的频率存在),其附近的基因位点 Y 上有等位基因 c 和 d (它们也以相同的频率存在) 的情况下,期望作为各个基因多态性的组合的单倍体型 ac 在群体中以 0.25 的频率存在。在单倍体型 ac 大于此期望值的情况下,也就是在称作 ac 的特定基因型频繁出现的情况下,就称等位基因 ac 为连锁不平衡。连锁不平衡是等位基因的特定组合的自然选择或者从进化的角度看被导入群体的时期是最近的而产生的,是由于连锁的等位基因之间没有达到平衡而产生的。因此,如民族和人种等那样,在不同的群体中连锁不平衡的模式不一样,即使在某个群体中 ac 连锁不平衡的情况下,在其它群体中也有作为 ad 连锁不平衡的关系的情况。连锁不平衡的基因多态性检测,虽然该多态性本身不引起直接疾病,但却对检测对于疾病的易感性有效。例如,对于某基因位点 X 的等位基因 a 来说,虽然它不是作为疾病的原因的基因要素,但是由于和基因位点 Y 上的等位基因 c 连锁不平衡,因此可能显示疾病易感性。

[0112] 在本说明书中,“单倍域”是指作为分割基本没有发现历史重组的基因组的范围,规定在该范围内强的连锁不平衡存在的区域。本领域的技术人员可以根据连锁不平衡的强度来适当地确定单倍域,例如,可以根据 Gabriel 等的报告 (Gabriel, SB. et al., Science, 296 (5576), p2225-2229, 2002) 进行。在本说明书中,“强的连锁不平衡”是指在连锁不平衡分析中计算的连锁不平衡系数 D' 的 95% 置信区间的上限超过 0.98,其下限高于 0.7 的状态,“具有强的历史重组的证据”是指连锁不平衡系数 D' 的 95% 置信区间的上限不足 0.9 的状态。

[0113] 在本说明书中“次要等位基因”是指在一个基因的位点上存在 2 个等位基因的情况下,基因频率较低的等位基因 (allele)。

[0114] 在本说明书中,“基因频率”和“等位基因频率”用于同样意义,是指特定的等位基

因在任意的基因群体中所占的比例的用语。

[0115] 在本说明书中,“统计学上显著不同”是指在任意的统计学的方法中进行检测的情况下,风险率(p 值)小于 0.1%,优选为小于 0.07%,更加优选为小于 0.05%,更加进一步优选为小于 0.01%的状态。

[0116] 2. 毛发形状易感性基因以及毛发形状易感性 SNP 标记的认定

[0117] 寻找以及认定成为决定作为多因子性的一般特性的人的自然毛发形状的原因的基因(毛发形状易感性基因),可以通过使用性状作图(trait mapping)的方法进行遗传统计分析。即,通过患病同胞对连锁分析来认定卷发特性位点,以及根据以卷发特性位点为对象的病例・对对照例关联分析,可以有效地选择处于和毛发形状易感性基因连锁不平衡状态中的 SNP,并将含有此 SNP 的单倍域中存在的基因认定为毛发形状易感性基因。

[0118] 本发明的毛发形状易感性基因以及毛发形状易感性 SNP 标记的认定具体在后述的实施例中记载,但是可以通过实施具有下述的工序的认定方法来进行。

[0119] (i) 定义毛发形状,并收集卷发家族、以及具有卷发特质的人(病例)和具有直发特性的人(对照例)的工序、

[0120] (ii) 使用来自卷发家族的样品,实施以全基因组作为对象的患病同胞对连锁分析,认定卷发特性位点的工序、

[0121] (iii) 对在工序(ii)中确认的卷发特性位点,选定在该区域全体范围内均匀分布的多个 SNP 标记的工序、

[0122] (iv) 对在工序(iii)中选定的 SNP 标记,用来自病例和对照例的样品进行分型,根据统计学处理比较其结果,将认定有显著差的 SNP 标记认定为毛发形状易感性 SNP 标记的工序、以及

[0123] (v) 对毛发形状易感性 SNP 标记使用国际人类基因单倍体型图计划数据库(the International HapMap Project database)的 HapMap PHASE 数据,通过将作为在对象候选区域内确认连锁不平衡的区域,并且含有毛发形状易感性 SNP 标记的区域(单倍域)进行特定,从而认定毛发形状易感性基因的工序。

[0124] (vi) 对于在工序(v)中特定的单倍域中提取的单倍体型,使用国际人类基因单倍体型图计划数据库的 HapMap PHASE 数据,将与在工序(iv)中认定的毛发形状易感性 SNP 标记位点连锁的 SNP 位点进行特定,并将该 SNP 追加认定为新的毛发形状易感性 SNP 标记的工序。

[0125] 上述工序(i)为毛发形状(卷发或者直发)的定义和特性绘图的分析对象者的收集的工序。在特性绘图中,有必要在某种程度上定量地处理作为对象的特性,将对象者规定为卷发特性或者直发特性的毛发形状的定义,是在实施特性绘图中的重要工序。人的毛发形状多种多样,其计测方法、以及分类或者定义的方法也各种各样。例如,作为定义毛发形状的方法,可以举将卷发=1、直发=0的二值化方法,和用一些方法测定卷发的程度来数值化的方法,以及本领域的技术人员所公知的方法(例如,特开 2005-350801、特开 2008-268229、特许第 4159515 号等),但是不限于这些。作为定义方法的更加具体的例子,有将毛发形状按照整体的形状、毛发的卷曲程度(卷曲半径)、卷曲出现的频率、以及/或者与周围的毛发群的卷曲的协调性等特征分类为数个等级(例如,2~10 个等级,优选为 3~8 个等级,更加优选为 5~7 个等级),并将所述分类中卷缩发、以及卷发或者强的波浪

发等的卷曲半径有小的倾向的毛发形状定义为卷发特性,将波浪发、大略直发、或者微波浪发、直发等卷曲半径有大的倾向的毛发形状定义为直发特性的方法。

[0126] 上述工序(ii)为使用来自卷发家族的样品,以全基因组作为对象实施患病同胞对连锁分析的工序。作为实施患病同胞对连锁分析的卷发家族的构成成员,为在上述工序(i)中判定为卷发特性的同胞(兄弟姐妹对,2名)。更加优选为还加上同胞的双亲的家人4名(或者3名),进一步也可以加上其他的兄弟姐妹(不论毛发形状)和祖父母等。另外,为了实施患病同胞对连锁分析必需的卷发家族的数目,可以通过推测和/或观测卷发特性在母群体中的频率和构成其原因的基因频率(等位基因频率)、或者同胞相对风险率,并通过模拟来计算,但是一般的为50个家族到数百个家族。

[0127] 作为患病同胞对连锁分析使用的遗传标记,只要是基因多态性就没有特别的限定,但是优选使用在基因组中均匀存在的等位基因数较多的微卫星,扩增以及检测此微卫星的试剂盒(全基因组扫描、连锁分析试剂盒 Linkage Mapping Set)是从 Applied Biosystems, Inc. (ABI) 购买的。另外,在本发明中使用以平均 9.2cM 的间隔覆盖人染色体的 ABI PRISM Linkage Mapping Set-MD 10v2.5 (ABI 制造),以及以平均 5cM 的间隔覆盖人染色体的 ABI PRISM Linkage Mapping Set-MD 5v2.5 (ABI 制造)。

[0128] 另外,作为遗传标记的微卫星可以任意选定,也可以通过 The Mammalian Genotyping Service 的 Comprehensive human genetic maps (<http://research.marshfieldclinic.org/genetics/GeneticResearch/compMaps.asp>)、NCBI (<http://www.ncbi.nlm.nih.gov/>) 等检索。此时,优选基因组中以 0.1 ~ 数 cM 的间隔存在、且等位基因多的杂合度高的微卫星。进一步,对于认定连锁的染色体,可以追加微卫星标记,窄化连锁区域(详细作图)。另外,为了扩增和检测任意选定、追加的微卫星的 PCR 引物可以从 NCBI (<http://www.ncbi.nlm.nih.gov/>) 中检索碱基序列,基于此序列,例如可以通过市售的核苷酸合成机按常用方法制作。此时,为了能够迅速并且容易地检测扩增产物,该探针优选选用放射性物质、荧光物质、化学发光物质、或者酶等标识。

[0129] 在患病同胞对连锁分析中,将来自卷发家族的基因组 DNA 作为模板,使用全基因组扫描、连锁分析试剂盒 (ABI)、或者任意选定的微卫星标记的扩增引物进行 PCR,检测扩增产物(片段)。PCR 及其扩增产物的检测可以通过常用方法实施。此时,如果将各扩增引物用不同的荧光色素(例如,6-FAM(蓝色)、VIC(绿色)、或者 NED(黄色)等发出不同的荧光的色素中的任意种)进行标识,则即使是同一尺寸的扩增产物也可以通过分别辨别出各荧光颜色,从而能够迅速地检测出多个扩增引物。

[0130] 连锁的检测可以使用能够进行非参数分析的市售或者公开的遗传统计学软件(例如 Genehunter、Linkage package、Mapmaker/sibs 等)进行。

[0131] 认定为连锁的区域的判定是根据以下所示的 Lander 和 Kruglyak 的指导准则 (Nat. Genet., 11(3), 241-247, 1995), 基于得到假阳性的连锁的基准而进行的。虽然 Lander 和 Kruglyak 的指导准则(多因子疾病中在全基因组内的连锁分析)盛行,但是对于各个基因的连锁分析也要从此基因的功能判断此基因是否能成为其原因。但是因为在全基因组内的分析中不需要在此阶段考虑基因功能,能求得纯粹的数理遗传学上显著的判断标准(阈值)。在此,他们根据模拟,制定了下述表 2 所示的显著连锁基准。

[0132] [表 2]

[0133]

提示连锁 (Suggestive Linkage) (在全基因组内得到 1 个假阳性的连锁结果的基准)	$P < 7.4 \times 10^{-4}$ $LOD > 2.2$
显著连锁 (Significant Linkage) (在全基因组内得到 0.05 个假阳性的连锁结果的基准)	$P < 2.2 \times 10^{-5}$ $LOD > 3.6$
高度显著连锁 (High Significant Linkage) (在全基因组内得到 0.01 个假阳性的连锁结果的基准)	$P < 3.0 \times 10^{-7}$ $LOD > 5.4$

[0134] 通过本工序,可以筛选全部染色体,检测出能认定和卷发特性连锁的染色体上的区域。进一步,通过详细的作图,可以将染色体上的特定区域认定为卷发特性位点。这样认定的区域是强烈暗示毛发形状易感性基因存在的区域。

[0135] 上述工序 (iii) 是对在工序 (ii) 中确认的卷发特性位点,选定在该区域全体范围内均匀分布的多个 SNP 标记的工序。SNP 标记可以使用 dbSNP 数据库 (<http://www.ncbi.nlm.nih.gov/SNP/>)、JSNP 数据库 (http://snp.ims.u-tokyo.ac.jp/index_ja.html) 等有关 SNP 的各种数据库来选定。

[0136] 在 SNP 标记的选定中,选择对于认定毛发形状易感性基因有用的 SNP。具体来说,在日本人群体中,选择次要等位基因的基因频率为 10% 以上,更加优选为 15% 以上的 SNP。通过使用这样的基因频率的 SNP,可以选定置信度高的 SNP 标记。

[0137] 另外,在以基因频率作为指标选定 SNP 标记的情况下,有时候会出现 SNP 标记不均匀存在于特定的狭窄区域内的情况。这样的情况下,如果将选定的全部 SNP 标记用于认定毛发形状易感性基因的话,实验会变得复杂,另外,相互靠近的 SNP 也经常处于连锁不平衡状态而没有效率。因此,优选选定相互之间存在一定程度间隔的 SNP 标记来使用。这样,通过使存在一定间隔的标记均匀分布,可以进行包含全部对象候选区域的关联分析,使毛发形状易感性基因的认定容易进行。这样选定的相互邻接的 SNP 标记之间的距离优选为 5kb 以上,更加优选为 5kb ~ 10kb。如果此距离过长,可能出现和 SNP 标记之间的相互连锁不平衡的强度程度不能确认的区域。另外,此距离过短的话,确认相互具有强的连锁不平衡的 SNP 太多,因此没有效率。

[0138] 在对象候选区域的全部范围内彻底地选定 SNP 标记时,此 SNP 标记之间的距离,另外可以将标记在对象候选领域内的分布情况,即基因组的每单位距离上的标记的个数作为“标记密度”来表现。标记密度为每 10kb 基因组为 0.5SNP 以上,优选为 1SNP 以上,更加优选为 1SNP ~ 2SNP。如果标记密度过低的话,则标记之间的距离过长,有可能产生如前所述的 SNP 标记之间的连锁不平衡的强度程度不能确认的区域。另一方面,如果标记密度过高的话,则标记之间的距离过短,如前面所述,过密地选定标记而在认定毛发形状易感性基因的情况下实验量增多没有效率。

[0139] 上述工序 (iv) 是对在上述工序 (iii) 中选定的 SNP 标记进行病例・对照例关联分析的工序。病例・对照例关联分析是在病例 (患病者:具有卷发特性的人) 群体和对照例 (对照者:具有直发特性的人) 群体中比较某个遗传标记的等位基因频率,在两个群体之间能在等位基因频率中发现显著差的标记的方法。例如,使用从具有卷发特性的人 (病

例)以及具有直发特性的人(对照例)得到的样品进行分型,通过统计学处理比较其结果,认定发现显著差的 SNP 标记作为毛发形状易感性 SNP 标记。作为特性绘图必需的样品,只要是含有基因组 DNA 的都没有特别的限定,但是可以举例如末梢血等的血液、唾液、汗等体液、体细胞以及含有这些的组织或者器官等。实施用于病例・对照例关联分析所必需的病例以及对照例的数目虽然可以根据卷发特性在母群体中的频率和构成其原因的基因频率(等位基因频率)、基因型相对风险率等估计,但是一般的话为 50~数千人。另外,通过阶段性限定法,可以在限于样品大小和分型数等条件下得到较高的检测能力。另外,病例和对照例优选为由和特定毛发形状易感性基因的人种相同的人种构成,例如,在认定日本人的毛发形状易感性基因时,优选分析对象者为由日本人构成。

[0140] 作为 SNP 分型的方法,可以利用 PCR-SSCP、PCR-RLFP、PCR-SSO、PCR-ASP、定向测序(direct sequencing)、SNaPshot、dHPLC、Sniper 法、MALDI-TOF/MS 方法等本领域技术人员所公知的方法(例如,野岛博编,《基因组创药的最前线》(ゲノム創薬の最前線),p44-p54,羊土社,2001),采用利用例如 TaqMan SNP Genotyping Assays(注册商标)(ABI 制造)并且利用 TaqMan 系统的 SNP 分型法有效。

[0141] 关联分析,典型的可以通过在病例群体和对照例群体中比较各 SNP 标记的基因频率,针对频率之差在统计学上是不是显著而进行 χ^2 检测(东京大学教养学部统计学教室编,《统计学入门-基础统计学 I》进行分析,东京大学出版会,1991),也可以根据采用了各 SNP 标记的基因型频率、显性(或者隐性)模型的情况下的基因型频率、等位基因阳性率中的频率等进行分析。另外,除了 χ^2 检测之外,只要是能通过病例群体和对照例群体进行比较,也就是只要能检测多个群体所区分的特性、疾病等的表现特性和基因多态性的关联,也可以采用其它公知的统计学方法进行。

[0142] 另外,为了评价基因型的分型误差以及取样的妥当性而进行 Hardy-Weinberg 平衡检测。Hardy-Weinberg 平衡是在基因组统计学的领域内而公知的,存在如 SNP 等那样的 2 个等位基因(例如 C 和 T),它们在群体中各自的频率为 p 和 q 的时候,成为 $(p+q=1)$ 、C/C 同性的、C/T 异性的、T/T 同性的基因型频率分别为 p^2 、 $2pq$ 、 q^2 的 $(p^2+2pq+q^2=1)$ 。在进行关联分析的时候,希望在对照例群体中 Hardy-Weinberg 平衡成立,但是只要偏离 Hardy-Weinberg 平衡具有统计学上的显著差的个数在显著水平(典型的是 $p=0.01\sim0.05$)的预想范围内的话,可以评价选定的 SNP 标记为妥当。

[0143] 作为一个实施形态,针对由病例群体和对照群体得到的各个样品进行分型,用基因型、等位基因型、显性模型、隐性模型 4 个的方法通过 χ^2 检测来进行显著差检测。即,如果某个基因变异成为毛发形状变化的原因的话,可以预想在病例和对照例中此等位基因频率等中的差。在检测中,在比较少的对象者中实施该关联分析的情况下,或者在提高来自检测对象者的显著差的检测能力的情况下,可以宽松地设定显著水平,在对象者比较多数或者严格判断显著差的情况下,可以严格设定显著水平。通过检测可以将显示显著的基因频率差的 SNP 认定为毛发形状易感性 SNP 标记。

[0144] 接下来进行的工序(v)是对上述特定的毛发形状易感性 SNP 标记,使用国际人类基因单倍体型图计划数据库的 HapMap PHASE 数据,通过将作为在对象候选领域内确认连锁不平衡的区域,并且含有毛发形状易感性 SNP 标记的区域(单倍域)进行特定,从而认定毛发形状易感性基因的工序。

[0145] 单倍体型分析(连锁不平衡分析)为本领域工作人员公知的方法,可以通过以前进行的各种连锁不平衡分析(例如,镰鼓直之编, postgenome 时代的遗传统计学, p183-201, 羊土社, 2002)进行。单倍体型分析可以使用市售或者公开的各种遗传统计学软件(Haploview, Arlequin, SNP 疾病关联分析软件 SNPalyze(注册商标)(株式会社 DYNACOM 制造))等程序实施。更加具体来说,可以通过 EM Algorithm(Laird, N: "The EM Algorithm", Chap. 14, pp 509-520, Handbook of Statistics, Vol. 9, Computational Statistics, C. R. Rao(ed.)Elsevier Science Publishers B. V., 1993)进行连锁不平衡分析,算出连锁不平衡系数 D' (pair-wise LD coefficient) 进行分析。更加具体来说,在单倍体型分析中,分析上述特定的毛发形状易感性 SNP 标记和其它 SNP 标记之间是否存在连锁不平衡,认定连锁不平衡存在的区域为单倍域。连锁不平衡分析中所用的其它 SNP 标记可以从相对于上述毛发形状易感性 SNP 标记存在于基因组序列的上游以及下游的 SNP 中任意选择。例如,可以对上述毛发形状易感性 SNP 标记由近到远处存在的 SNP 依次进行连锁不平衡分析,或者,也可以对任意选择的远处的 SNP 实行分析以便将大致的单倍域区域特定之后,再对更近处的 SNP 将更加详细的单倍域区域进行特定。用于连锁不平衡分析中的其它 SNP 标记的个数,包括毛发形状易感性 SNP 标记在内为 4SNP 以上,优选为 20SNP 以上,更加优选为 32SNP 以上,对这些含有多个 SNP 标记的一系列 SNP 标记群进行分析。在此,对于 2 个 SNP,以第一 SNP 的各等位基因为 (A, a),第二 SNP 的各等位基因为 (B, b),4 个单倍体型 (AB, Ab, aB, ab) 的各频率为 P_{AB} 、 P_{Ab} 、 P_{aB} 、 P_{ab} 的话,连锁不平衡系数 D' 可以通过下式得到。另外,式中 $\text{Min}[(P_{AB}+P_{aB})(P_{ab}+P_{Ab}), (P_{AB}+P_{Ab})(P_{aB}+P_{ab})]$ 表示从 $(P_{AB}+P_{aB})(P_{ab}+P_{Ab})$ 和 $(P_{AB}+P_{Ab})(P_{aB}+P_{ab})$ 中取较小值的意思。

$$[0146] \quad D' = (P_{AB}P_{ab} - P_{Ab}P_{aB}) / \text{Min}[(P_{AB}+P_{aB})(P_{ab}+P_{Ab}), (P_{AB}+P_{Ab})(P_{aB}+P_{ab})]$$

[0147] 该 SNP 标记群的标记数会根据形成认定的和毛发形状易感性基因关联的单倍域(连锁不平衡区域)的区域大小而适当变更。另外,在预先能预想到区域的断开处的情况下,可以将此区域断开为 6SNP 左右的大小。进一步,也可以对开始在剪切毛发形状易感性 SNP 标记两端 5SNP 的共计 11SNP 进行连锁不平衡分析,也可以根据必要增加分析的标记数。

[0148] 通过进行连锁不平衡分析来特定对象候选区域中 SNP 连锁的区域(含有确认相互之间有强烈连锁不平衡的 SNP 标记群的单倍域)。例如,针对选定的 SNP 标记对全部的 2SNP 之间的组合算出连锁不平衡系数 D' ,选择表示 $D' > 0.9$ 的组合,其中,检测出含有更远的 SNP 所夹区域的一系列区域。接下来,算出该检测出的区域的外侧与该区域邻接的连续 3SNP 和该区域内的 SNP 之间的 D' 。在确认即使算出的任意组合中 D' 都为 0.9 以下的情况下,特定该区域作为“单倍域”。

[0149] 这样特定了单倍域的话,例如,对此区域,可以利用有关基因组的数据库等,特定在所观察的单倍域中存在的基因。另外,即使在不利用数据库的情况下,也可以根据常用方法决定在单倍域区域中存在的 SNP 标记附近的碱基序列,由此碱基序列来确定基因。

[0150] 上述工序(vi),是对工序(v)中特定的单倍域中提取的单倍体型,使用国际人类基因单倍体型图计划数据库的 HapMap PHASE 数据,特定在与工序(iv)中认定的毛发形状易感性 SNP 标记位点连锁的 SNP 位点,将该 SNP 追加认定为新的毛发形状易感性 SNP 标记的工序。

[0151] 在工序(v)中特定单倍域的同时,可以提取由用于单倍体型分析中的 SNP 标记群

的各个碱基构成的全部单倍体型,求出此单倍体型频率等。

[0152] 通过比较提取的单倍体型,也就是 SNP 标记群的各个碱基的组合,可以特定和工序 (iv) 中认定的毛发形状易感性 SNP 标记位点连锁的 SNP 位点,可以将这样认定的 SNP 位点作为新的毛发形状易感性 SNP 标记。

[0153] 通过上述工序 (i) ~ (vi),可以特定和卷发连锁所确认的染色体区域,接下来,可以从该染色体领域选定毛发形状易感性 SNP 标记。进一步,通过对选定的 SNP 标记的单倍体型分析,可以认定在该染色体区域中的与毛发形状关联的单倍域以及基因。之后,进一步,通过特定和毛发形状易感性 SNP 标记位点连锁的 SNP 位点,可以认定在该单倍域或者基因中存在的毛发形状易感性 SNP 标记。

[0154] 作为通过上述工序而特定的和卷发的连锁被确认的染色体区域,可以举 1 号染色体和 11 号染色体,更加具体来说,可以举 1 号染色体的 1q21.3 区域(微卫星 D1S2696 ~ D1S2346 所夹的区域)(最大 LOD 分数 = 3.60)。将这些区域特定为卷发特性位点,强烈暗示此区域中存在毛发形状易感性基因。

[0155] 作为通过上述工序特定的单倍域,可以举在人 1 号染色体的基因组区域中,由序列号 1 所示的碱基序列所表示的 23,252bp 构成的区域、由序列号 2 所示的碱基序列所表示的 56,552bp 构成的区域、由序列号 3 所示的碱基序列所表示的 23,382bp 构成的区域、由序列号 4 所示的碱基序列所表示的 8,818bp 构成的区域、以及由序列号 5 所示的碱基序列所表示的 3,440bp 构成的区域。

[0156] 将与这些单倍域相重叠的基因,即含有该单倍域的碱基序列的一部分或者全部的基因,认定为毛发形状易感性基因。在此,“与单倍域相重叠的基因”是指具有和单倍域的一部分区域相同的碱基序列,或者具有和单倍域的整个区域的碱基序列相同的碱基序列的基因这两者。另外,将在这些单倍域中存在、并且等位基因频率在具有卷发特性的群体和具有非卷发特性的群体之间在统计学上显著不同的单核苷酸多态性 (SNP) 以及和该 SNP 连锁的 SNP,认定为毛发形状易感性 SNP 标记。

[0157] 作为与由序列号 1 所表示的碱基序列所表示的 23,252bp 构成的单倍域相重叠的基因,可以举人 1 号染色体上的 ANXA9 基因,以及 FAM63A 基因。ANXA9 基因为为在 Entrez Gene 数据库 (<http://www.ncbi.nlm.nih.gov/gene>) 中 GeneID :8416 表示的基因,如实施例 5 以及图 5 所示,该碱基序列的全部长度和上述单倍域重叠。另外,FAM63A 基因为为 Entrez Gene 数据库中 GeneID :55793 所示的基因,如实施例 5 以及图 5 所示,其碱基序列的一部分和上述单倍域重叠。

[0158] 作为序列号 1 所示的碱基序列中存在的毛发形状易感性 SNP 标记,可以举碱基号 1 (dbSNP 数据库 ID :rs3754211、A 或者 G)、2405 (rs3754210、T 或者 G)、5874 (rs16832604、G 或者 A)、7121 (rs2305814、C 或者 T)、8494 (rs7532008、C 或者 A)、18980 (rs1673160、T 或者 A)、以及 23252 (rs771205、T 或者 C) 所表示的碱基,优选碱基号 7121 (rs2305814、C 或者 T) 所表示的碱基。

[0159] 作为与由序列号 2 表示的碱基序列所表示的 56,552bp 构成的单倍域相重叠的基因,可以举人 1 号染色体上的 LCE5A 基因以及 CRCT1 基因。LCE5A 基因为为在 Entrez Gene 数据库中 GeneID :254910 所表示的基因,如实施例 5 以及图 6 所示,该碱基序列的全部长度和上述单倍域重叠。另外,CRCT1 基因为为在 Entrez Gene 数据库中 GeneID :54544 所表示的基

因,如实施例 5 以及图 6 所示,该碱基序列的全部长度和上述单倍域重叠。

[0160] 作为序列号 2 所示的碱基序列中存在的毛发形状易感性 SNP 标记,可以举碱基号 2355(rs11581947、A 或者 G)、2569(rs6658925、G 或者 A)、3897(rs2105117、A 或者 G)、8196(rs1053590、C 或者 T)、9510(rs548252、T 或者 C)、13643(rs493133、C 或者 G)、15387(rs1970283、C 或者 G)、15708(rs1001834、A 或者 C)、16017(rs11205018、G 或者 T)、17106(rs545418、T 或者 C)、17453(rs12116609、T 或者 C)、17579(rs526099、C 或者 T)、17634(rs525960、A 或者 T)、26924(rs4845443、A 或者 G)、28383(rs569032、T 或者 C)、31275(rs528427、C 或者 G)、31301(rs478926、T 或者 G)、31653(rs1337338、A 或者 G)、31903(rs6587681、T 或者 C)、32209(rs1856120、A 或者 G)、33199(rs474086、T 或者 C)、33822(rs578382、A 或者 G)、34100(rs549044、T 或者 C)、35791(rs1123567、A 或者 G)、36884(rs1538083、G 或者 A)、37072(rs1538082、G 或者 A)、37365(rs7532535、A 或者 T)、37613(rs7518654、G 或者 C)、38062(rs533917、G 或者 A)、39063(rs564107、T 或者 C)、46580(rs7530609、A 或者 C)、49618(rs4240885、C 或者 G)、50164(rs4240886、A 或者 T)、50278(rs4240887、G 或者 A)、50662(rs6687126、G 或者 T)、50822(rs6674451、T 或者 C)、50981(rs7550769、A 或者 G)、51133(rs7529157、A 或者 C)、51263(rs1988805、G 或者 T)、以及 51397(rs7529441、T 或者 C) 所表示的碱基,优选 8196(rs1053590、C 或者 T) 所表示的碱基。

[0161] 作为与由序列号 3 表示的碱基序列所表示的 23,382bp 构成的单倍域相重叠的基因,可以举人 1 号染色体上的 LCE2B 基因以及 LCE2A 基因。LCE2B 基因因为在 Entrez Gene 数据库中 GeneID:26239 所表示的基因,如实施例 5 以及图 7 所示,该碱基序列的全部长度和上述单倍域重叠。另外,LCE2A 基因因为在 Entrez Gene 数据库中 GeneID:353139 所表示的基因,如实施例 5 以及图 7 所示,该碱基序列的全部长度和上述单倍域重叠。

[0162] 作为序列号 3 所示的碱基序列中存在的毛发形状易感性 SNP 标记,可以举碱基号 2509(rs11205072、G 或者 A)、5167(rs3753453、T 或者 C)、8449(rs3737859、T 或者 G)、17598(rs3904414、G 或者 A)、18481(rs12074783、T 或者 C)、20891(rs3908717、C 或者 G)、21734(rs3904415、C 或者 T) 以及 23382(rs11205079、A 或者 T) 所表示的碱基,优选为碱基号 5167(rs3753453、T 或者 C)、8449(rs3737859、T 或者 G)、17598(rs3904414、G 或者 A)、以及 20891(rs3908717、C 或者 G) 所表示的碱基。

[0163] 作为与由序列号 4 表示的碱基序列所表示的 8,818bp 构成的单倍域相重叠的基因,可以举人 1 号染色体上的 SMCP 基因。SMCP 基因因为在 Entrez Gene 数据库中 GeneID:4184 所表示的基因,如实施例 5 以及图 8 所示,该碱基序列的一部分和上述单倍域重叠。

[0164] 作为序列号 4 所示的碱基序列中存在的毛发形状易感性 SNP 标记,可以举碱基号 1(rs16834715、T 或者 C)、3308(rs12022319、C 或者 T)、4715(rs4845490、G 或者 A)、4985(rs4845491、C 或者 T)、6354(rs3737861、C 或者 A)、8553(rs16834728、C 或者 T)、以及 8818(rs4845492、C 或者 G) 所表示的碱基,优选为 6354(rs3737861、C 或者 A) 所表示的碱基。

[0165] 作为与由序列号 5 表示的碱基序列所表示的 3,440bp 构成的单倍域相重叠的基因,可以举人 1 号染色体上的 IVL 基因。IVL 基因因为在 EntrezGene 数据库中 GeneID:3713 所表示的基因,如实施例 5 以及图 9 所示,该碱基序列的一部分和上述单倍域重叠。

[0166] 作为序列号 5 所示的碱基序列中存在的毛发形状易感性 SNP 标记,可以举碱基号 1(rs1854779、C 或者 T)、540(rs16834751、C 或者 A)、759(rs4523473、C 或者 T)、1007(rs11205131、A 或者 G)、1018(rs7528862、G 或者 A)、1075(rs7517189、G 或者 C)、1939(rs2229496、G 或者 A)、以及 3440(rs913996、A 或者 G) 所表示的碱基,优选为碱基号 759(rs4523473、C 或者 T)、1939(rs2229496、G 或者 A)、以及 3440(rs913996、A 或者 G) 所表示的碱基。

[0167] 3. 毛发形状判定标记

[0168] 本发明还提供一种毛发形状判定标记,其中,所述毛发形状判定标记是包含以下部分碱基序列的寡核苷酸或者多核苷酸或者它们的互补链,其中,所述部分碱基序列是在人 1 号染色体的 1q21.3 区域(D1S2696 ~ D1S2346)上,通过对等位基因频率在具有卷发特性的群体和具有非卷发特性的群体之间在统计学上显著不同的 SNP 标记进行连锁不平衡分析而确定的,并且是由序列号 1 ~ 5 中任意一个所表示的碱基序列所构成的单倍域的碱基序列的部分碱基序列,并且,所述部分碱基序列由连续的碱基序列构成,所述连续的碱基序列含有一个以上的单核苷酸多态性,所述单核苷酸多态性包括等位基因频率在具有卷发特性的群体和具有非卷发特性的群体之间在统计学上显著不同的单核苷酸多态性(SNP)、以及与该 SNP 连锁的 SNP。

[0169] 由这些碱基序列特定的寡核苷酸或者多核苷酸或者它们的互补链,存在于序列号为 1 ~ 5 所示的碱基序列所表示的单倍域中,含有一个以上等位基因频率在具有卷发特性的群体和具有非卷发特性的群体之间在统计学上显著不同的单核苷酸多态性(SNP)或者和该 SNP 连锁的 SNP 的毛发形状易感性 SNP 标记,通过检测这些寡核苷酸或者多核苷酸或者它们的互补链,可以检查和 / 或判定在被检测者中毛发形状的遗传素质(hereditary predisposition)。因此,可以将这些寡核苷酸或者多核苷酸或者它们的互补链规定为用于判定个体所具有的毛发形状的遗传素质的标记并进行使用。

[0170] 这些寡核苷酸或者多核苷酸或者它们的互补链的长度(碱基长度)只要是在人基因组上能够特异性识别的长度即可,在其限定上没有特别限定。通常为 10 个碱基长度以上并且 1000 个碱基以下,优选为 20 个碱基长度以上并且 500 个碱基长度以下,更加优选为 20 个碱基长度以上并且 100 个碱基长度以下。因此,根据必要,可以选择在包含例如上述序列号 1 ~ 5 所示的碱基序列所表示的单倍域中存在的毛发形状易感性 SNP 标记的 11 个碱基(优选为毛发形状易感性 SNP 标记的 5' 侧及 3' 侧各 5 个碱基)、21 个碱基(优选为毛发形状易感性 SNP 标记的 5' 侧及 3' 侧各 10 个碱基)、101 个碱基(优选为毛发形状易感性 SNP 标记的 5' 侧及 3' 侧各 50 个碱基)、或者 601 个碱基(优选为毛发形状易感性 SNP 标记的 5' 侧及 3' 侧各 300 个碱基)等。

[0171] 作为本发明的毛发形状判定标记中所应含有的本发明中利用的毛发形状易感性 SNP 标记的例子,可以举如下:

[0172] (1) 序列号 1 所表示的碱基序列中,碱基号 1(dbSNP 数据库 ID:rs3754211、A 或者 G)、2405(rs3754210、T 或者 G)、5874(rs16832604、G 或者 A)、7121(rs2305814、C 或者 T)、8494(rs7532008、C 或者 A)、18980(rs1673160、T 或者 A)、以及 23252(rs771205、T 或者 C) 所表示的碱基;

[0173] (2) 序列号 2 所表示的碱基序列中,碱基号 2355(rs11581947、A 或者 G)、

2569(rs6658925、G 或者 A)、3897(rs2105117、A 或者 G)、8196(rs1053590、C 或者 T)、9510(rs548252、T 或者 C)、13643(rs493133、C 或者 G)、15387(rs1970283、C 或者 G)、15708(rs1001834、A 或者 C)、16017(rs11205018、G 或者 T)、17106(rs545418、T 或者 C)、17453(rs12116609、T 或者 C)、17579(rs526099、C 或者 T)、17634(rs525960、A 或者 T)、26924(rs4845443、A 或者 G)、28383(rs569032、T 或者 C)、31275(rs528427、C 或者 G)、31301(rs478926、T 或者 G)、31653(rs1337338、A 或者 G)、31903(rs6587681、T 或者 C)、32209(rs1856120、A 或者 G)、33199(rs474086、T 或者 C)、33822(rs578382、A 或者 G)、34100(rs549044、T 或者 C)、35791(rs1123567、A 或者 G)、36884(rs1538083、G 或者 A)、37072(rs1538082、G 或者 A)、37365(rs7532535、A 或者 T)、37613(rs7518654、G 或者 C)、38062(rs533917、G 或者 A)、39063(rs564107、T 或者 C)、46580(rs7530609、A 或者 C)、49618(rs4240885、C 或者 G)、50164(rs4240886、A 或者 T)、50278(rs4240887、G 或者 A)、50662(rs6687126、G 或者 T)、50822(rs6674451、T 或者 C)、50981(rs7550769、A 或者 G)、51133(rs7529157、A 或者 C)、51263(rs1988805、G 或者 T)、以及 51397(rs7529441、T 或者 C) 所表示的碱基；

[0174] (3) 序列号 3 所表示的碱基序列中，碱基号 2509(rs11205072、G 或者 A)、5167(rs3753453、T 或者 C)、8449(rs3737859、T 或者 G)、17598(rs3904414、G 或者 A)、18481(rs12074783、T 或者 C)、20891(rs3908717、C 或者 G)、21734(rs3904415、C 或者 T) 以及 23382(rs11205079、A 或者 T) 所表示的碱基；

[0175] (4) 序列号 4 所表示的碱基序列中，碱基号 1(rs16834715、T 或者 C)、3308(rs12022319、C 或者 T)、4715(rs4845490、G 或者 A)、4985(rs4845491、C 或者 T)、6354(rs3737861、C 或者 A)、8553(rs16834728、C 或者 T)、以及 8818(rs4845492、C 或者 G) 所表示的碱基；以及

[0176] (5) 序列号 5 所表示的碱基序列中，碱基号 1(rs1854779、C 或者 T)、540(rs16834751、C 或者 A)、759(rs4523473、C 或者 T)、1007(rs11205131、A 或者 G)、1018(rs7528862、G 或者 A)、1075(rs7517189、G 或者 C)、1939(rs2229496、G 或者 A)、以及 3440(rs913996、A 或者 G) 所表示的碱基。

[0177] 上述碱基中，优选序列号 1 所表示的碱基序列中碱基号 7121(rs2305814、C 或者 T) 所表示的碱基、序列号 2 所表示的碱基序列中碱基号 8196(rs1053590、C 或者 T) 所表示的碱基、序列号 3 所表示的碱基序列中碱基号 5167(rs3753453、T 或者 C)、8449(rs3737859、T 或者 G)、17598(rs3904414、G 或者 A)、以及 20891(rs3908717、C 或者 G) 所表示的碱基、序列号 4 所表示的碱基序列中碱基号 6354(rs3737861、C 或者 A) 所表示的碱基，以及序列号 5 所表示的碱基序列中碱基号 759(rs4523473、C 或者 T)、1939(rs2229496、G 或者 A)、以及 3440(rs913996、A 或者 G) 所表示的碱基。

[0178] 毛发形状易感性 SNP 标记优选位于本发明的毛发形状判定标记的中央或者中央附近（例如，离中央 100 个碱基以内，优选为 50 个碱基以内，更加优选为 30 个碱基以内，进一步更加优选为 10 个碱基以内，再更加进一步优选为 5 个碱基以内），但是也没有必要必须如此。进一步，在本发明的毛发形状判定标记中含有 2 个以上的毛发形状易感性 SNP 标记的情况下，该毛发形状易感性 SNP 标记中可以全部都处于本发明的毛发形状易感性判定标记的中央或者中央附近，也可以该毛发形状易感性 SNP 标记中的 1 个处于中央或者中央附

近的位置并且其它处于任意位置,或者也可以是该毛发形状易感性 SNP 标记中全部都不处在中央或者中央附近。

[0179] 作为毛发形状易感性 SNP 标记处于中央的本发明的毛发形状判定标记的具体例子,例如在序列号 1 所表示的碱基序列中碱基号 7121 所表示的碱基中含有 SNP (dbSNP 数据库 ID :rs2305814、C 或者 T) 的情况下,可以举出由序列号 1 的碱基号 7116 到碱基号 7126 的碱基构成的 11 个碱基长度的多核苷酸、序列号 1 的碱基号 7111 到碱基号 7131 的碱基构成的 21 个碱基长度的多核苷酸、序列号 1 的碱基号 7021 到碱基号 7221 的碱基构成的 101 个碱基长度的多核苷酸、以及具有序列号 1 的碱基号 6821 到碱基号 7421 的碱基的 601 个碱基长度的多核苷酸。另外,还可以利用这些多核苷酸的互补链。同样,也可以决定含有其它 SNP 的标记的碱基序列。

[0180] 4. 对毛发形状的遗传易感性的判定方法

[0181] 本发明另外提供判定对被检测者的毛发形状的遗传易感性(遗传素质)的方法。本发明的对毛发形状的遗传易感性的判定方法为包括以下工序(a)、以及(b)的方法,并不限制于此限定:

[0182] (a) 调制来自被检测者的基因组 DNA 的工序;以及

[0183] (b) 从该基因组 DNA 中,检测出人 1 号染色体的 1q21.3 区域(D1S2696 ~ D1S2346)中的,通过对等位基因频率在具有卷发特性的群体和具有非卷发特性的群体之间在统计学上显著不同的单核苷酸多态性(SNP)标记的连锁不平衡分析而确定的,并且存在于序列号 1 ~ 5 的任意种所表示的碱基序列构成的单倍域中的,等位基因频率在具有卷发特性群体和具有非卷发特性群体之间在统计学上显著不同的单核苷酸多态性(SNP)以及和该 SNP 连锁的单核苷酸多态性(SNP)的工序。

[0184] 上述工序(a)(染色体 DNA 的提取)和工序(b)(SNP 的检测)可以使用公知的方法(例如,Birren Bruce et al.,GenomeAnalysis,Vol.4/A Laboratory Manual:Mapping Genomes,Cold Spring Harbor Laboratory, NY,1999)进行。

[0185] 在工序(a)中,来自被检测者的基因组 DNA 可以使用从被检测者及其临床标本等分离出来的所有细胞(包括培养细胞,但是不包括生殖细胞)、组织(包含培养组织)、器官、或者体液(例如,血液、唾液、淋巴液、气管粘膜、精液、汗、尿等)等作为材料。作为该材料,优选由末梢血分离得到的白血球或者单核细胞,更加优选白血球,这些材料可以根据在临床检查中通常所用的方法进行分离。

[0186] 例如,在将白血球作为材料的情况下,首先从被检测者身上分离出来的末梢血中按照常用方法分离白血球。接下来,向得到的白血球中加入蛋白酶 K 和十二烷基硫酸钠(SDS)来分解蛋白质,使之变性之后,进行苯酚/氯仿提取得到基因组 DNA(含有 RNA)。RNA 可以根据必要通过 RNase 除去。另外,基因组 DNA 的提取不限于上述方法,可以利用该技术领域中公知的方法(例如,Joseph Sambrook et al.,Molecular Cloning:A Laboratory Manual(3Vol. Set),Cold Spring Harbor Laboratory, NY,2001)、市售的 DNA 提取试剂盒等进行。根据进一步的需要,也可以分离含有人 1 号染色体的 1q21.3 区域、或者人 1 号染色体的基因组区域中的序列号 1 ~ 5 所表示的碱基序列所表示的单倍域的 DNA。该 DNA 的分离可以使用与 1q21.3 区域或者该单倍域杂交的引物,通过将基因组 DNA 作为模板的 PCR 等进行分离。

[0187] 在工序 (b) 中,从工序 (a) 中得到的基因组 DNA 中检测出等位基因频率在任意的卷发者群体中比在任意的非卷发者群体中高的 SNP 或者与该 SNP 连锁的 SNP,其中,所述 SNP 是在人 1 号染色体的 1q21.3 区域 (D1S2696 ~ D1S2346) 中,通过对等位基因频率在具有卷发特性的群体和具有非卷发特性的群体之间在统计学上显著不同的单核苷酸多态性 (SNP) 标记进行连锁不平衡分析而确定的,并且是存在于序列号 1 ~ 5 的任意种所表示的碱基序列构成的单倍域中的多态性。序列号 1 ~ 5 所表示的碱基序列为人 1 号染色体的基因组区域中序列号 1 所表示的 23, 252bp 构成的碱基序列,序列号 2 所表示的 56, 552bp 构成的碱基序列,序列号 3 所表示的 23, 382bp 构成的碱基序列,序列号 4 所表示的 8, 818bp 构成的碱基序列,以及序列号 5 所表示的 3, 440bp 构成的碱基序列。

[0188] 上述本发明的判定方法,进一步优选包括下述工序 (c) :

[0189] (c) 在该检测出的 SNP 的等位基因频率在卷发者群体中比在非卷发者群体中在统计学上显著高的情况下,判定为该被检测者具有卷发的遗传素质;在该检测出的 SNP 的等位基因频率在任意的非卷发者群体中比在卷发者群体中在统计学上显著高的情况下,判定为该被检测者不具有卷发的遗传素质的工序。

[0190] 作为上述工序 (c),例如可以举,在来自被检测者的基因组 DNA 中的序列号 1 ~ 5 所表示的碱基序列中的下述表中所示的碱基号的碱基中的任意一个以上中,识别该碱基为碱基 (i) 还是碱基 (ii),在为碱基 (i) 的情况下,判定为被检测者具有卷发的素质,在为碱基 (ii) 的情况下,判定为被检测者不具有卷发的素质的工序。

[0191] [表 3-1]

[0192]

序列号	碱基号	碱基(i) (有素质)	碱基(ii) (无素质)
1	1	G	A
	2405	G	T
	5874	A	G
	7121	T	C
	8494	A	C
	18980	A	T
	23252	C	T
2	2355	G	A
	2569	A	G
	3897	G	A
	8196	T	C
	9510	C	T
	13643	G	C
	15387	G	G
	15708	C	A
	16017	T	G
	17106	C	T
	17453	C	T
	17579	T	C
	17634	T	A
	26924	G	A
	28383	C	T
	31275	G	C
	31301	G	T
	31653	G	A
	31903	C	T
	32209	G	A
	33199	C	T
	33822	G	A
	34100	C	T
	35791	G	A
	36884	A	G
	37072	A	G
	37365	T	A
	37613	C	G
	38062	A	G
	39063	C	T
	46580	C	A
	49618	G	C
	50164	T	A
	50278	A	G
	50662	T	G
	50822	C	T
	50981	G	A
	51133	C	A
	51263	T	G
	51397	C	T

[0193] [表 3-2]

[0194]

序列号	碱基号	碱基(i) (有素质)	碱基(ii) (无素质)
3	2509	A	G
	5167	C	T
	8449	G	T
	17598	A	G
	18481	C	T
	20891	G	C
	21734	T	C
	23382	T	A
4	1	C	T
	3308	T	C
	4715	A	G
	4985	T	C
	6354	A	C
	8553	T	C
	8818	G	C
5	1	T	C
	540	A	C
	759	T	C
	1007	G	A
	1018	A	G
	1075	C	G
	1939	A	G
	3440	G	A

[0195] 更具体的说,本发明的对被检测者的毛发形状的遗传易感性的判定方法含有下述(1)～(70)中任意一个工序。

[0196] (1) 在序列号 1 所示的碱基序列中,识别碱基号 1 所表示的碱基为 A 还是为 G,在为 G 的情况下,判定为具有卷发的素质,另外,在为 A 的情况下,判定为不具有卷发的素质。

[0197] (2) 在序列号 1 所示的碱基序列中,识别碱基号 2405 所表示的碱基为 T 还是为 G,在为 G 的情况下,判定为具有卷发的素质,另外,在为 T 的情况下,判定为不具有卷发的素质。

[0198] (3) 在序列号 1 所示的碱基序列中,识别碱基号 5874 所表示的碱基为 G 还是为 A,在为 A 的情况下,判定为具有卷发的素质,另外,在为 G 的情况下,判定为不具有卷发的素质。

[0199] (4) 在序列号 1 所示的碱基序列中,识别碱基号 7121 所表示的碱基为 C 还是为 T,在为 T 的情况下,判定为具有卷发的素质,另外,在为 C 的情况下,判定为不具有卷发的素质。

[0200] (5) 在序列号 1 所示的碱基序列中,识别碱基号 8494 所表示的碱基为 C 还是为 A,在为 A 的情况下,判定为具有卷发的素质,另外,在为 C 的情况下,判定为不具有卷发的素质。

[0201] (6) 在序列号 1 所示的碱基序列中,识别碱基号 18980 所表示的碱基为 T 还是为 A,在为 A 的情况下,判定为具有卷发的素质,另外,在为 T 的情况下,判定为不具有卷发的素质。

[0202] (7) 在序列号 1 所示的碱基序列中,识别碱基号 23252 所表示的碱基为 T 还是为

C,在为C的情况下,判定为具有卷发的素质,另外,在为T的情况下,判定为不具有卷发的素质。

[0203] (8) 在序列号 2 所示的碱基序列中,识别碱基号 2355 所表示的碱基为 A 还是为 G,在为 G 的情况下,判定为具有卷发的素质,另外,在为 A 的情况下,判定为不具有卷发的素质。

[0204] (9) 在序列号 2 所示的碱基序列中,识别碱基号 2569 所表示的碱基为 G 还是为 A,在为 A 的情况下,判定为具有卷发的素质,另外,在为 G 的情况下,判定为不具有卷发的素质。

[0205] (10) 在序列号 2 所示的碱基序列中,识别碱基号 3897 所表示的碱基为 A 还是为 G,在为 G 的情况下,判定为具有卷发的素质,另外,在为 A 的情况下,判定为不具有卷发的素质。

[0206] (11) 在序列号 2 所示的碱基序列中,识别碱基号 8196 所表示的碱基为 C 还是为 T,在为 T 的情况下,判定为具有卷发的素质,另外,在为 C 的情况下,判定为不具有卷发的素质。

[0207] (12) 在序列号 2 所示的碱基序列中,识别碱基号 9510 所表示的碱基为 T 还是为 C,在为 C 的情况下,判定为具有卷发的素质,另外,在为 T 的情况下,判定为不具有卷发的素质。

[0208] (13) 在序列号 2 所示的碱基序列中,识别碱基号 13643 所表示的碱基为 C 还是为 G,在为 G 的情况下,判定为具有卷发的素质,另外,在为 C 的情况下,判定为不具有卷发的素质。

[0209] (14) 在序列号 2 所示的碱基序列中,识别碱基号 15387 所表示的碱基为 C 还是为 G,在为 G 的情况下,判定为具有卷发的素质,另外,在为 C 的情况下,判定为不具有卷发的素质。

[0210] (15) 在序列号 2 所示的碱基序列中,识别碱基号 15708 所表示的碱基为 A 还是为 C,在为 C 的情况下,判定为具有卷发的素质,另外,在为 A 的情况下,判定为不具有卷发的素质。

[0211] (16) 在序列号 2 所示的碱基序列中,识别碱基号 16017 所表示的碱基为 G 还是为 T,在为 T 的情况下,判定为具有卷发的素质,另外,在为 G 的情况下,判定为不具有卷发的素质。

[0212] (17) 在序列号 2 所示的碱基序列中,识别碱基号 17106 所表示的碱基为 T 还是为 C,在为 C 的情况下,判定为具有卷发的素质,另外,在为 T 的情况下,判定为不具有卷发的素质。

[0213] (18) 在序列号 2 所示的碱基序列中,识别碱基号 17453 所表示的碱基为 T 还是为 C,在为 C 的情况下,判定为具有卷发的素质,另外,在为 T 的情况下,判定为不具有卷发的素质。

[0214] (19) 在序列号 2 所示的碱基序列中,识别碱基号 17579 所表示的碱基为 C 还是为 T,在为 T 的情况下,判定为具有卷发的素质,另外,在为 C 的情况下,判定为不具有卷发的素质。

[0215] (20) 在序列号 2 所示的碱基序列中,识别碱基号 17634 所表示的碱基为 A 还是为

T,在为T的情况下,判定为具有卷发的素质,另外,在为A的情况下,判定为不具有卷发的素质。

[0216] (21) 在序列号 2 所示的碱基序列中,识别碱基号 26924 所表示的碱基为 A 还是为 G,在为G的情况下,判定为具有卷发的素质,另外,在为A的情况下,判定为不具有卷发的素质。

[0217] (22) 在序列号 2 所示的碱基序列中,识别碱基号 28383 所表示的碱基为 T 还是为 C,在为C的情况下,判定为具有卷发的素质,另外,在为T的情况下,判定为不具有卷发的素质。

[0218] (23) 在序列号 2 所示的碱基序列中,识别碱基号 31275 所表示的碱基为 C 还是为 G,在为G的情况下,判定为具有卷发的素质,另外,在为C的情况下,判定为不具有卷发的素质。

[0219] (24) 在序列号 2 所示的碱基序列中,识别碱基号 31301 所表示的碱基为 T 还是为 G,在为G的情况下,判定为具有卷发的素质,另外,在为T的情况下,判定为不具有卷发的素质。

[0220] (25) 在序列号 2 所示的碱基序列中,识别碱基号 31653 所表示的碱基为 A 还是为 G,在为G的情况下,判定为具有卷发的素质,另外,在为A的情况下,判定为不具有卷发的素质。

[0221] (26) 在序列号 2 所示的碱基序列中,识别碱基号 31903 所表示的碱基为 T 还是为 C,在为C的情况下,判定为具有卷发的素质,另外,在为T的情况下,判定为不具有卷发的素质。

[0222] (27) 在序列号 2 所示的碱基序列中,识别碱基号 32209 所表示的碱基为 A 还是为 G,在为G的情况下,判定为具有卷发的素质,另外,在为A的情况下,判定为不具有卷发的素质。

[0223] (28) 在序列号 2 所示的碱基序列中,识别碱基号 33199 所表示的碱基为 T 还是为 C,在为C的情况下,判定为具有卷发的素质,另外,在为T的情况下,判定为不具有卷发的素质。

[0224] (29) 在序列号 2 所示的碱基序列中,识别碱基号 33822 所表示的碱基为 A 还是为 G,在为G的情况下,判定为具有卷发的素质,另外,在为A的情况下,判定为不具有卷发的素质。

[0225] (30) 在序列号 2 所示的碱基序列中,识别碱基号 34100 所表示的碱基为 T 还是为 C,在为C的情况下,判定为具有卷发的素质,另外,在为T的情况下,判定为不具有卷发的素质。

[0226] (31) 在序列号 2 所示的碱基序列中,识别碱基号 35791 所表示的碱基为 A 还是为 G,在为G的情况下,判定为具有卷发的素质,另外,在为A的情况下,判定为不具有卷发的素质。

[0227] (32) 在序列号 2 所示的碱基序列中,识别碱基号 36884 所表示的碱基为 G 还是为 A,在为A的情况下,判定为具有卷发的素质,另外,在为G的情况下,判定为不具有卷发的素质。

[0228] (33) 在序列号 2 所示的碱基序列中,识别碱基号 37072 所表示的碱基为 G 还是为

A,在为A的情况下,判定为具有卷发的素质,另外,在为G的情况下,判定为不具有卷发的素质。

[0229] (34) 在序列号 2 所示的碱基序列中,识别碱基号 37365 所表示的碱基为 A 还是为 T,在为 T 的情况下,判定为具有卷发的素质,另外,在为 A 的情况下,判定为不具有卷发的素质。

[0230] (35) 在序列号 2 所示的碱基序列中,识别碱基号 37613 所表示的碱基为 G 还是为 C,在为 C 的情况下,判定为具有卷发的素质,另外,在为 G 的情况下,判定为不具有卷发的素质。

[0231] (36) 在序列号 2 所示的碱基序列中,识别碱基号 38062 所表示的碱基为 G 还是为 A,在为 A 的情况下,判定为具有卷发的素质,另外,在为 G 的情况下,判定为不具有卷发的素质。

[0232] (37) 在序列号 2 所示的碱基序列中,识别碱基号 39063 所表示的碱基为 T 还是为 C,在为 C 的情况下,判定为具有卷发的素质,另外,在为 T 的情况下,判定为不具有卷发的素质。

[0233] (38) 在序列号 2 所示的碱基序列中,识别碱基号 46580 所表示的碱基为 A 还是为 C,在为 C 的情况下,判定为具有卷发的素质,另外,在为 A 的情况下,判定为不具有卷发的素质。

[0234] (39) 在序列号 2 所示的碱基序列中,识别碱基号 49618 所表示的碱基为 C 还是为 G,在为 G 的情况下,判定为具有卷发的素质,另外,在为 C 的情况下,判定为不具有卷发的素质。

[0235] (40) 在序列号 2 所示的碱基序列中,识别碱基号 50164 所表示的碱基为 A 还是为 T,在为 T 的情况下,判定为具有卷发的素质,另外,在为 A 的情况下,判定为不具有卷发的素质。

[0236] (41) 在序列号 2 所示的碱基序列中,识别碱基号 50278 所表示的碱基为 G 还是为 A,在为 A 的情况下,判定为具有卷发的素质,另外,在为 G 的情况下,判定为不具有卷发的素质。

[0237] (42) 在序列号 2 所示的碱基序列中,识别碱基号 50662 所表示的碱基为 G 还是为 T,在为 T 的情况下,判定为具有卷发的素质,另外,在为 G 的情况下,判定为不具有卷发的素质。

[0238] (43) 在序列号 2 所示的碱基序列中,识别碱基号 50822 所表示的碱基为 T 还是为 C,在为 C 的情况下,判定为具有卷发的素质,另外,在为 T 的情况下,判定为不具有卷发的素质。

[0239] (44) 在序列号 2 所示的碱基序列中,识别碱基号 50981 所表示的碱基为 A 还是为 G,在为 G 的情况下,判定为具有卷发的素质,另外,在为 A 的情况下,判定为不具有卷发的素质。

[0240] (45) 在序列号 2 所示的碱基序列中,识别碱基号 51133 所表示的碱基为 A 还是为 C,在为 C 的情况下,判定为具有卷发的素质,另外,在为 A 的情况下,判定为不具有卷发的素质。

[0241] (46) 在序列号 2 所示的碱基序列中,识别碱基号 51263 所表示的碱基为 G 还是为

T,在为T的情况下,判定为具有卷发的素质,另外,在为G的情况下,判定为不具有卷发的素质。

[0242] (47) 在序列号 2 所示的碱基序列中,识别碱基号 51397 所表示的碱基为 T 还是为 C,在为C的情况下,判定为具有卷发的素质,另外,在为T的情况下,判定为不具有卷发的素质。

[0243] (48) 在序列号 3 所示的碱基序列中,识别碱基号 2509 所表示的碱基为 G 还是为 A,在为A的情况下,判定为具有卷发的素质,另外,在为G的情况下,判定为不具有卷发的素质。

[0244] (49) 在序列号 3 所示的碱基序列中,识别碱基号 5167 所表示的碱基为 T 还是为 C,在为C的情况下,判定为具有卷发的素质,另外,在为T的情况下,判定为不具有卷发的素质。

[0245] (50) 在序列号 3 所示的碱基序列中,识别碱基号 8449 所表示的碱基为 T 还是为 G,在为G的情况下,判定为具有卷发的素质,另外,在为T的情况下,判定为不具有卷发的素质。

[0246] (51) 在序列号 3 所示的碱基序列中,识别碱基号 17598 所表示的碱基为 G 还是为 A,在为A的情况下,判定为具有卷发的素质,另外,在为G的情况下,判定为不具有卷发的素质。

[0247] (52) 在序列号 3 所示的碱基序列中,识别碱基号 18481 所表示的碱基为 T 还是为 C,在为C的情况下,判定为具有卷发的素质,另外,在为T的情况下,判定为不具有卷发的素质。

[0248] (53) 在序列号 3 所示的碱基序列中,识别碱基号 20891 所表示的碱基为 C 还是为 G,在为G的情况下,判定为具有卷发的素质,另外,在为C的情况下,判定为不具有卷发的素质。

[0249] (54) 在序列号 3 所示的碱基序列中,识别碱基号 21734 所表示的碱基为 C 还是为 T,在为T的情况下,判定为具有卷发的素质,另外,在为C的情况下,判定为不具有卷发的素质。

[0250] (55) 在序列号 3 所示的碱基序列中,识别碱基号 23382 所表示的碱基为 A 还是为 T,在为T的情况下,判定为具有卷发的素质,另外,在为A的情况下,判定为不具有卷发的素质。

[0251] (56) 在序列号 4 所示的碱基序列中,识别碱基号 1 所表示的碱基为 T 还是为 C,在为C的情况下,判定为具有卷发的素质,另外,在为T的情况下,判定为不具有卷发的素质。

[0252] (57) 在序列号 4 所示的碱基序列中,识别碱基号 3308 所表示的碱基为 C 还是为 T,在为T的情况下,判定为具有卷发的素质,另外,在为C的情况下,判定为不具有卷发的素质。

[0253] (58) 在序列号 4 所示的碱基序列中,识别碱基号 4715 所表示的碱基为 G 还是为 A,在为A的情况下,判定为具有卷发的素质,另外,在为G的情况下,判定为不具有卷发的素质。

[0254] (59) 在序列号 4 所示的碱基序列中,识别碱基号 4985 所表示的碱基为 C 还是为 T,在为T的情况下,判定为具有卷发的素质,另外,在为C的情况下,判定为不具有卷发的素

质。

[0255] (60) 在序列号 4 所示的碱基序列中,识别碱基号 6354 所表示的碱基为 C 还是为 A,在为 A 的情况下,判定为具有卷发的素质,另外,在为 C 的情况下,判定为不具有卷发的素质。

[0256] (61) 在序列号 4 所示的碱基序列中,识别碱基号 8553 所表示的碱基为 C 还是为 T,在为 T 的情况下,判定为具有卷发的素质,另外,在为 C 的情况下,判定为不具有卷发的素质。

[0257] (62) 在序列号 4 所示的碱基序列中,识别碱基号 8818 所表示的碱基为 C 还是为 G,在为 G 的情况下,判定为具有卷发的素质,另外,在为 C 的情况下,判定为不具有卷发的素质。

[0258] (63) 在序列号 5 所示的碱基序列中,识别碱基号 1 所表示的碱基为 C 还是为 T,在为 T 的情况下,判定为具有卷发的素质,另外,在为 C 的情况下,判定为不具有卷发的素质。

[0259] (64) 在序列号 5 所示的碱基序列中,识别碱基号 540 所表示的碱基为 C 还是为 A,在为 A 的情况下,判定为具有卷发的素质,另外,在为 C 的情况下,判定为不具有卷发的素质。

[0260] (65) 在序列号 5 所示的碱基序列中,识别碱基号 759 所表示的碱基为 C 还是为 T,在为 T 的情况下,判定为具有卷发的素质,另外,在为 C 的情况下,判定为不具有卷发的素质。

[0261] (66) 在序列号 5 所示的碱基序列中,识别碱基号 1007 所表示的碱基为 A 还是为 G,在为 G 的情况下,判定为具有卷发的素质,另外,在为 A 的情况下,判定为不具有卷发的素质。

[0262] (67) 在序列号 5 所示的碱基序列中,识别碱基号 1018 所表示的碱基为 G 还是为 A,在为 A 的情况下,判定为具有卷发的素质,另外,在为 G 的情况下,判定为不具有卷发的素质。

[0263] (68) 在序列号 5 所示的碱基序列中,识别碱基号 1075 所表示的碱基为 G 还是为 C,在为 C 的情况下,判定为具有卷发的素质,另外,在为 G 的情况下,判定为不具有卷发的素质。

[0264] (69) 在序列号 5 所示的碱基序列中,识别碱基号 1939 所表示的碱基为 G 还是为 A,在为 A 的情况下,判定为具有卷发的素质,另外,在为 G 的情况下,判定为不具有卷发的素质。

[0265] (70) 在序列号 5 所示的碱基序列中,识别碱基号 3440 所表示的碱基为 A 还是为 G,在为 G 的情况下,判定为具有卷发的素质,另外,在为 A 的情况下,判定为不具有卷发的素质。

[0266] 另外,在本发明的判定对毛发形状的遗传易感性(遗传素质)的方法中检测出来的 SNP,可以为上述 SNP 中的任意 1 个,也可以是 2 个以上。优选为检测出 2 个以上的 SNP,这样关于作为多基因性的一般特性的毛发形状,可以明确被检测者的遗传素质的种类及其有无,可以以更高的精度检索构成确定被检测者的毛发形状的主要原因的基因。

[0267] 该 SNP 的检测,可以通过直接确定从含有基因组 DNA 的样品中分离出来的人 1 号染色体的 1q21.3 区域、或者人 1 号染色体的基因组区域中的序列号 1~5 所示的碱基序

列所表示的单倍域的碱基序列来进行。或者作为检测出多态性的方法,除了直接确定上述的该区域的基因序列的方法之外,还有在多态性序列为限制性内切酶识别部位的情况下,利用限制性内切酶切断方式的不同来确定基因型的方法(以下,称为RFLP);以使用多态性特异的探针杂交为基础的方法(例如,在芯片和载玻片、尼龙膜上贴上特定的探针,通过检测出对这些探针的杂交强度的差来确定多态性的种类,或者通过检测特异的探针的杂交效率,检测在扩增双链模板时聚合酶分解的探针的量来特定基因型的方法;或者通过某种双链特异的荧光色素发出的荧光随着温度变化而检测双链融解时候的温度差,这样来特定多态性的方法;在多态性部位特异的寡核苷酸探针的两端加上互补的序列,利用由于温度造成该探针在自身的分子内形成二级结构;或者在目标区域中杂交的差别来特定基因型的方法等)。或者进一步还有,由模板特异的引物通过聚合酶进行碱基延伸反应,特定此时在多态性部位所引入的碱基的方法(使用双脱氧核苷酸分别进行荧光标记,并检测出各荧光的方法;通过质谱分析检测引入的双脱氧核苷酸的方法),进一步,还有通过酶识别在模板特异的引物上所接的变异部位上互补的碱基对或者非互补的碱基对的有无的方法等。

[0268] 以下列出现有的公知的代表性的基因多态性的检测方法,但是本发明不限于这些。(a)RFLP(限制性片段长度多态性(restriction fragment length polymorphism))法、(b)PCR-SSCP法(单链构象多态性分析(single strand conformation polymorphism), Biotechniques,16, p296-297,1994, 以及 Biotechniques,21, p510-514,1996)、(c)ASO杂交法(Clin.Chim.Acta.,189, p153-157,1990)、(d)定向测序法(direct sequencing)(Biotechniques,11, p246-249,1991)、(e)ARMS法(Nuc.Acids Res.,19, p3561-3567, 1991、以及 Nuc.Acids Res.,20, p4831-4837,1992)、(f)变性梯度凝胶电泳(denatured gradient gel electrophoresis)(DGGE)法(Biotechniques,27, p1016-1018,1999)、(g)RNaseA切断法(DNA Cell Biol.,14, p87-94,1995)、(h)化学切断法(Biotechniques, 21, p216-218,1996)、(i)DOL法(Genome Res.,8, p549-556,1998)、(j)TaqMan-PCR法(Genet.Anal.,14, p143-149,1999、以及 J.Clin.Microbiol.,34, p2933-2936,1996)、(k)入侵(Invader)法(Science,5109, p778-783,1993、J.Bio.Chem.,30, p21387-21394, 1999、以及 Nat.Biotechnol.,17, p292-296,1999)、(l)MALDI-TOF/MS法(Genome Res., 7, p378-388,1997、以及 Eur.J.Clin.Chem.Clin.Biochem.,35, p545-548,1997)、(m)TDI法(Proc.Natl.Acad.Sci.USA,94, p10756-10761,1997)、(n)分子信标(molecular beacon)法(Nat.Biotechnol.,16, p 49-53,1998)、(o)动态等位基因特异性杂交(Dynamic Allele-Specific Hybridization)(DASH)法(Nat.Biotechnol.,17, p 87-88,1999)、(p)锁式探针法(Nat.Genet.,3, p225-232,1998)、(q)DNA芯片或者DNA微阵列(中村祐辅等,《SMP基因多态性的战略》(《SNP遗传子多型の戦略》),中山书店,p128-135,2000)、(r)ECA法(Anal.Chem.,72, p1334-1341,2000)。

[0269] 以上是代表性的基因多态性检测方法,但是对于本发明的毛发形状的遗传易感性(遗传素质)的判定方法,并不限于此,可以广泛使用其它公知或者将来开发的基因多态性检测方法。另外,在检测本发明的基因多态性的时候,可以单独使用这些基因多态性的检测方法,或者也可以组合2种以上的方法。下面作为代表的方法,针对在后述实施例中使用的TaqMan-PCR法以及入侵法,进行更详细地说明。

[0270] (1)TaqMan-PCR法

[0271] TaqMan-PCR法是利用了经过荧光标记的等位基因特异的寡核苷酸(TaqMan 探针)和Taq DNA聚合酶产生的PCR的方法。作为TaqMan 探针,使用含有在人1号染色体的基因组区域中,作为以序列号1~5所示的碱基序列表示的单倍域的部分碱基序列,并且含有上述任意的多态性部位的碱基的约15~约30个碱基连续的碱基序列的寡核苷酸(例如,在后述本发明的毛发形状判定用试药中含有的核酸探针)。该探针在其5'末端用FAM和VIC等荧光色素、3'末端用TAMRA等淬灭剂(quencher)(消光物质)分别标记,在这样的状态下,因为淬灭剂会吸收掉荧光能量所以不能检测出荧光。探针要针对双方的等位基因来调制,为了一次检测出来而优选用荧光波长互相不同的荧光色素(例如,一方的等位基因用FAM,另一方用VIC)标识。另外为了不引起TaqMan 探针导致的PCR延伸反应而将3'末端磷酸化。如果将为扩增含有和TaqMan 探针杂交的区域的基因组DNA的部分序列而设计的引物以及TaqDNA聚合酶同时进行PCR的话,TaqMan 探针和模板DNA杂交,同时引起PCR引物开始的延伸反应,进行延伸反应的话由TaqDNA聚合酶的5'核苷酸活性使杂交的TaqMan 探针被切断,荧光色素游离而不受淬灭剂的影响,从而检测出荧光。由于模板的扩增造成荧光强度呈指数规律地增大。例如,在序列号1所示的碱基序列中,在检测碱基号7121(rs2305814、C或者T)所示的碱基中的多态性时,将含有该碱基的等位基因特异的寡核苷酸(约15~约30个碱基长度;在5'末端C等位基因用FAM、T等位基因用VIC分别标识,3'末端则都用TAMRA标识)作为TaqMan 探针使用的情况下,被检测者的基因型如果为CC或者TT的话,则分别能识别出FAM或者VIC的强荧光强度,基本不能识别另一种荧光。另一方面,如果被检测者的基因型为CT的话,则能检测出FAM以及VIC两种荧光。

[0272] (2) 入侵法

[0273] 在入侵法中,与TaqMan-PCR法不一样,等位基因特异的寡核苷酸(等位基因探针)自身不被标识,在多态性部位的碱基的5'侧具有和模板DNA没有互补性的序列(折翼(flap)),3'侧具有和模板特异的互补序列。在入侵法中,进一步使用在模板的多态性部位的3'侧具有特异的互补序列的寡核苷酸(入侵探针,该探针的与作为5'末端的多态性部位相当的碱基为任意碱基);以及FRET(Fluorescence Resonance Energy Transfer)探针,其特征为在5'侧具有能得到发夹结构的序列,在形成发夹结构时从与5'末端的碱基成对的碱基开始连续到3'侧的序列是和等位基因探针的折翼互补的序列。FRET探针的5'末端上用了荧光标识(例如,FAM和VIC等),其附近结合有淬灭剂(例如,TAMRA等),在这样的状态(发夹结构)下不能检测出荧光。如果使作为模板的基因组DNA和等位基因探针以及入侵探针反应的话,三者互补结合的时候入侵探针的3'末端侵入多态性部位中。如果使用识别此多态性部位的结构酶(裂解酶(Cleavase))来切断等位基因探针的单链部分(即,多态性部位的碱基开始的5'侧的折翼部分)的话,折翼和FRET探针互补地结合,折翼的多态性部位侵入FRET探针的发夹结构中。通过裂解酶识别此结构并切断,FRET探针的末端标识的荧光色素游离不受淬灭剂的影响而能够检测出荧光。多态性部位的碱基和模板不匹配的等位基因探针不能通过裂解酶切断,但是没被切断的等位基因探针也能和FRET探针杂交,所以同样也能检测出荧光。但是,因为反应效率不同,所以在多态性部位的碱基所匹配的等位基因探针中,与不匹配的等位基因探针相比荧光强度显著较强。通常,在使其与3种探针以及裂解酶反应之前,优选将模板DNA使用能扩增含有等位基因探针及入侵探针所杂交的部分的区域的引物,通过PCR进行扩增。

[0274] 人的毛发形状能够通过烫发处理和定型剂处理、梳理等可以使之自在地变化,并且也可以由于年龄增加和代谢等变化而产生后天的变化。因此,人本来自然的头发形状仅仅从其表现型难以正确地判定或者分类。另外,考虑毛发形状为复杂的多基因性的一般特性,针对各个人,认为从上述所举的本发明的毛发形状易感性基因中,构成决定毛发形状的主要原因的基因分别不同。因此,通过检查和/或判定毛发形状的遗传素质,可以提供针对各个人的准确的毛发形状的调节方法。

[0275] 进一步,根据该方法,可以判定被检测者的对后天的毛发形状的变化易感性,即,毛发形状变化的风险度。毛发形状变化的风险度是将上述多态性作为基准(指标),不需要医生等具有专门知识的人的判断,可以机械进行。因此,本发明的方法也可以作为毛发形状变化的风险度的检测方法利用。

[0276] 通过本发明的对被检测者的毛发形状的遗传易感性(遗传素质)的判定方法,可以针对作为多基因性的一般特性的毛发形状,明确被检测者的遗传素质的种类和其有无,从本发明的毛发形状易感性基因中,可以探索作为决定该被检测者的毛发形状的主要原因的基因。进一步基于此结果,对于该被检测者,可以寻求用于调节毛发形状的准确对策。因此,本发明作为根本的用于毛发形状的调节的检查和/或判定方法非常有用。

[0277] 5. 对毛发形状的遗传易感性(遗传素质)的判定用试药以及含有该试药的试剂盒

[0278] 本发明另外还提供用于上述本发明的判定方法中使用的试药以及含有该试药的试剂盒。即,本发明的判定用试药以及含有该试药的试剂盒含有能检测出下面所述的 SNP 以及从与该 SNP 连锁的 SNP 中选择 1 个以上的 SNP 的核酸探针和/或引物,其中,所述 SNP 是在人 1 号染色体的 1q21.3 区域(D1S2696 ~ D1S2346)中,通过对等位基因在具有卷发特性的群体和具有非卷发特性的群体之间在统计学上显著不同的单核苷酸多态性(SNP)标记进行连锁不平衡分析而确定的,而且存在于由序列号 1 所表示的 23,252bp 的碱基序列、序列号 2 所表示的 56,552bp 的碱基序列、序列号 3 所表示的 23,382bp 的碱基序列、序列号 4 所表示的 8,818bp 的碱基序列、或者序列号 5 所表示的 3,440bp 的碱基序列构成的单倍域中的单核苷酸多态性(SNP),并且是一方的等位基因频率在任意的卷发者群体中比在任意的非卷发者群体中更高的 SNP。

[0279] 在一个实施方式中,在本发明的判定用试药和含有该试药的试剂盒中使用的核酸探针是与含有上述本发明的检查和/或判定方法中应该检测出的 SNP 部位的碱基的基因组 DNA 的区域特异杂交的核酸,例如是与本发明的毛发形状判定标记序列特异杂交的探针。该核酸探针只要对于应该杂交的目标部位是特异的并且能容易地检测出多态性,对其长度(和基因组 DNA 杂交的部分的碱基长度)没有特别的限定,例如为约 10 个碱基以上,优选为约 15 个碱基以上,更加优选为约 15 ~ 约 600 个碱基、更加进一步优选为约 15 ~ 约 200 个碱基,再进一步优选为约 15 ~ 约 50 个碱基。另外,“对于目标部位(序列)特异杂交”是指在通常的杂交条件下,优选在严格的杂交条件下(例如,Joseph Sambrook et al., Molecular Cloning: A Laboratory Manual (3Vol. Set), Cold Spring Harbor Laboratory, NY, 2001 中记载的条件),和其它 DNA 不明显地产生杂交。优选该核酸探针对于含有应该检测出多态性部位碱基的区域的碱基序列具有互补的碱基序列,但是只要可以进行相关的特异杂交,没有必要完全互补。

[0280] 该核酸探针也可以含有适合多态性检测的附加序列(和基因组 DNA 不互补的序

列)。例如,上述入侵法中使用的等位基因探针在多态性部位的碱基的 5' 末端具有被称为折翼的附加序列。另外,该探针也可以使用适当的标识剂,例如,放射性同位素(例如: ^{125}I 、 ^{131}I 、 ^3H 、 ^{14}C 等)、酶(例如: β -半乳糖苷酶、 β -葡萄糖苷酶、碱性磷酸酶、过氧化物酶、苹果酸脱氢酶等)、荧光物质(例如:荧光胺、异硫氰酸荧光素等)、发光物质(例如:发光氨、发光氨衍生物、荧光素、光泽精等)等标识。或者,也可以进一步在荧光物质(例如:FAM、VIC)的附近进一步结合吸收该荧光物质所发的荧光能量的淬灭剂(消光物质)。在相关的实施方式中,在检测反应的时候荧光物质和淬灭剂分离而能检测出荧光。

[0281] 该核酸探针可以固定于任意的固相上使用。因此,本发明的试药和含有该试药的试剂盒另外可以作为将上述探针固定于任意固相载体上而形成的固定化探针(例如固定有探针的基因芯片、cDNA 微阵列、寡 DNA 阵列、薄膜过滤器等)来提供。该固定化探针优选作为毛发形状易感性基因检测用的 DNA 芯片来提供。

[0282] 用于固定的固相载体,只要是能够固定核酸的都没有特别的限定,例如可以举玻璃板、尼龙薄膜、微珠、硅芯片、毛细管或者其它的载体等。核酸固定于固相载体上可以是预先合成的核酸载于固相上的方法,或者也可以是在固相上合成目标核酸的方法。固定方法,例如如果是 DNA 微阵列的话,可以利用市售的测位仪(Amersham 公司制造等)等,根据固定化探针的种类使用在该技术领域中公知的技术(例如,通过光刻蚀(photolithographic)技术(Affymetrix 公司)、喷墨技术(Rosetta Inpharmatics 公司)原位合成寡核苷酸等)。

[0283] 本发明的判定用试药以及含有该试药的试剂盒中使用的核酸引物,只要是能特异地杂交于含有上述本发明的检查和/或判定方法中应该检测出的 SNP 部位的碱基的基因组 DNA 的区域,能特异地扩增该核酸序列而设计的核酸引物即可。例如,该引物是特异地杂交于本发明的毛发形状判定标记的核酸序列上,并扩增该核酸序列的引物。在此,“对目标部位(序列)特异杂交”是指在通常的杂交条件之下,优选为严格的杂交条件之下(例如,Joseph Sambrook et al., Molecular Cloning: A Laboratory Manual (3Vol. Set), Cold Spring Harbor Laboratory, NY, 2001 中记载的条件),与其它 DNA 不会显著产生交叉杂交。

[0284] 作为该引物导致的核酸序列的扩增方法,只要是该领域中通常使用的方法都没有特别的限定。例如一般广泛使用 PCR 法,但是可以举 RCA(Rolling Circle Amplification; Proc. Natl. Acad. Sci, vol. 92, 4641-4645(1995))、ICAN(Isothermal and Chimeric primer-initiated Amplification of Nucleic acids)、LAMP(Loop-Mediated Isothermal Amplification of DNA; Bio Industry, 第 18 卷,第 2 期(2001))、NASBA(Nucleic acid Sequence-based Amplification method; Nature, 350, 91 ~ (1991))、TMA(Transcription mediated amplification method; J. Clin Microbiol. 第 31 卷, 3270 ~ (1993))等。扩增中所必需的该核酸引物的数量以及种类会依赖扩增方法而变化。例如,在使用 PCR 法的情况下,作为必要的引物可以举如下所述的一对核酸引物,其中,所述一对核酸引物是:含有作为人 1 号染色体的基因组区域中序列号 1 ~ 5 所示的碱基序列所表示的单倍域的部分碱基序列,与比应该检测出的多态性部位的碱基靠近 5' 侧的互补链序列的一部分特异地杂交的约 10 ~ 约 50 个碱基、优选为约 15 ~ 约 50 个碱基、更加优选为约 15 ~ 约 30 个碱基的碱基序列的核酸;和含有作为该部分碱基序列,并且与比该多态性部位的碱基靠近 3' 侧的互补链序列的一部分特异地杂交的约 10 ~ 约 50 个碱基、优选为约 15 ~ 约 50 个碱基、更加优

选为约 15 ~ 约 30 个碱基的碱基序列的核酸的组合, 并且由此扩增的核酸的片段长度为约 50 ~ 约 1000 个碱基、优选为约 50 ~ 约 500 个碱基、更加优选为约 50 ~ 约 200 个碱基。

[0285] 该引物也可以含有适合于多态性检测的附加序列(和基因组 DNA 不互补的序列), 例如也可以含有接头序列。另外, 该引物也可以使用适当的标识剂, 例如, 放射性同位素(例如, ^{125}I 、 ^{131}I 、 ^3H 、 ^{14}C 等)、酶(例如: β -半乳糖苷酶、 β -葡萄糖苷酶、碱性磷酸酶、过氧化物酶、苹果酸脱氢酶等)、荧光物质(例如: 荧光胺、异硫氰酸荧光素等)、发光物质(例如: 发光氨、发光氨衍生物、荧光素、光泽精等)等标识。

[0286] 优选本发明的判定用试药以及含有该试药的试剂盒中所用的核酸探针和/或引物含有本发明的毛发形状易感性 SNP 标记, 即以下所示的碱基:

[0287] (1) 序列号 1 所表示的碱基序列中, 碱基号 1(dbSNP 数据库 ID: rs3754211、A 或者 G)、2405(rs3754210、T 或者 G)、5874(rs16832604、G 或者 A)、7121(rs2305814、C 或者 T)、8494(rs7532008、C 或者 A)、18980(rs1673160、T 或者 A)、以及 23252(rs771205、T 或者 C) 所表示的碱基;

[0288] (2) 序列号 2 所表示的碱基序列中, 碱基号 2355(rs11581947、A 或者 G)、2569(rs6658925、G 或者 A)、3897(rs2105117、A 或者 G)、8196(rs1053590、C 或者 T)、9510(rs548252、T 或者 C)、13643(rs493133、C 或者 G)、15387(rs1970283、C 或者 G)、15708(rs1001834、A 或者 C)、16017(rs11205018、G 或者 T)、17106(rs545418、T 或者 C)、17453(rs12116609、T 或者 C)、17579(rs526099、C 或者 T)、17634(rs525960、A 或者 T)、26924(rs4845443、A 或者 G)、28383(rs569032、T 或者 C)、31275(rs528427、C 或者 G)、31301(rs478926、T 或者 G)、31653(rs1337338、A 或者 G)、31903(rs6587681、T 或者 C)、32209(rs1856120、A 或者 G)、33199(rs474086、T 或者 C)、33822(rs578382、A 或者 G)、34100(rs549044、T 或者 C)、35791(rs1123567、A 或者 G)、36884(rs1538083、G 或者 A)、37072(rs1538082、G 或者 A)、37365(rs7532535、A 或者 T)、37613(rs7518654、G 或者 C)、38062(rs533917、G 或者 A)、39063(rs564107、T 或者 C)、46580(rs7530609、A 或者 C)、49618(rs4240885、C 或者 G)、50164(rs4240886、A 或者 T)、50278(rs4240887、G 或者 A)、50662(rs6687126、G 或者 T)、50822(rs6674451、T 或者 C)、50981(rs7550769、A 或者 G)、51133(rs7529157、A 或者 C)、51263(rs1988805、G 或者 T)、以及 51397(rs7529441、T 或者 C) 所表示的碱基;

[0289] (3) 序列号 3 所表示的碱基序列中, 碱基号 2509(rs11205072、G 或者 A)、5167(rs3753453、T 或者 C)、8449(rs3737859、T 或者 G)、17598(rs3904414、G 或者 A)、18481(rs12074783、T 或者 C)、20891(rs3908717、C 或者 G)、21734(rs3904415、C 或者 T) 以及 23382(rs11205079、A 或者 T) 所表示的碱基;

[0290] (4) 序列号 4 所表示的碱基序列中, 碱基号 1(rs16834715、T 或者 C)、3308(rs12022319、C 或者 T)、4715(rs4845490、G 或者 A)、4985(rs4845491、C 或者 T)、6354(rs3737861、C 或者 A)、8553(rs16834728、C 或者 T)、以及 8818(rs4845492、C 或者 G) 所表示的碱基; 以及

[0291] (5) 序列号 5 所表示的碱基序列中, 碱基号 1(rs1854779、C 或者 T)、540(rs16834751、C 或者 A)、759(rs4523473、C 或者 T)、1007(rs11205131、A 或者 G)、1018(rs7528862、G 或者 A)、1075(rs7517189、G 或者 C)、1939(rs2229496、G 或者 A)、以及

3440(rs913996、A 或者 G) 所表示的碱基。

[0292] 更加优选,本发明的判定用试药以及含有该试药的试剂盒中所用的核酸探针和 / 或引物为序列号 1 所表示的碱基序列中,碱基号 7121(rs2305814、C 或者 T) 所表示的碱基 ;序列号 2 所表示的碱基序列中,碱基号 8196(rs1053590、C 或者 T) 所表示的碱基 ;序列号 3 所表示的碱基序列中,碱基号 5167(rs3753453、T 或者 C)、8449(rs3737859、T 或者 G)、17598(rs3904414、G 或者 A)、以及 20891(rs3908717、C 或者 G) 所表示的碱基 ;序列号 4 所表示的碱基序列中,碱基号 6354(rs3737861、C 或者 A) 所表示的碱基 ;以及序列号 5 所表示的碱基序列中,碱基号 759(rs4523473、C 或者 T)、1939(rs2229496、G 或者 A)、以及 3440(rs913996、A 或者 G) 所表示的碱基。

[0293] 作为具有上述的多态性部位的碱基的核酸探针,可以根据使用的多态性检测方法,对各多态性部位使用具有任意一种的等位基因的碱基的核酸,也可以使用具有分别对应于各等位基因的碱基的 2 种核酸。另外,对于在上述入侵法中使用的入侵探针,多态性部位的碱基(即,3' 末端的碱基)可以为任意的碱基。

[0294] 本发明的判定用试药以及含有该试药的试剂盒中所用的核酸探针和 / 或引物可以为 DNA 也可以为 RNA,另外,可以为单链也可以为双链。为双链的情况下,可以为双链 DNA、双链 RNA、DNA/RNA 杂交体中的任意种。上述核酸探针和 / 或引物,可以根据碱基序列的信息,例如利用市售的核苷酸合成机按照常用方法来制作。

[0295] 上述核酸探针和 / 或引物可以各自分别(或者可能的话以混合的状态)溶解于水或者适当的缓冲液(例如:TE 缓冲剂等)中至适当的浓度(例如:按照 2 ~ 20× 浓度,为 1 ~ 50 μM 等),在约 -20℃ 下保存。本发明的判定用试药以及含有该试药的试剂盒也可以根据使用的多态性检测方法,进一步含有该方法的实施中必要的其它成分,例如杂交反应用的缓冲剂、用于核酸扩增反应的酶、缓冲剂以及其它必要的试药、标识用试药、标识检测用试药、以及这些反应和步骤所需要的器具等来作为构成。例如,在该试药以及含有该试药的试剂盒为根据由 TaqMan-PCR 法的多态性检测用的情况下,该试药以及含有该试药的试剂盒进一步可以含有 10×PCR 反应缓冲液、10×MgCl₂水溶液、10×dNTPs 水溶液、Taq DNA 聚合酶(5U/μL)等。

[0296] 本发明的判定用试药和含有该试药的试剂盒可以在对毛发形状的遗传易感性(遗传素质)的检查和 / 或判定中使用。

[0297] 6. 毛发形状易感性基因或者其编码的蛋白质的使用

[0298] 按上述顺序认定的毛发形状易感性基因或者其表达产物,其表达和活性与毛发形状相关联而变化。因此,该毛发形状易感性基因及其表达产物可以作为用于检测和 / 或判定被检测者的毛发形状类型的毛发形状类型的标记使用。或者,通过测定和评价该毛发形状易感性基因或者其表达产物的表达量,可以进行人的毛发形状调节剂的评价或者选择。或者通过控制该毛发形状易感性基因或者其表达产物的表达量可以调节人的毛发形状。

[0299] 在本发明中,对于作为有必要检测和 / 或判定上述毛发形状类型,或者调节毛发形状的对象的人来说,虽然不限定特定的人种和群体,但是,优选为亚洲人种,更加优选日本人。

[0300] 作为用作毛发形状判定标记的毛发形状易感性基因及其表达产物来说,只要是与由序列号 1 ~ 5 中任意一个表示的碱基序列构成的单倍域相重叠的基因及其表达产物即

可,优选的可以举 ANXA9 基因、FAM63A 基因、LCE5A 基因、CRCT1 基因、LCE2B 基因、LCE2A 基因、SMCP 基因以及 IVL 基因、及其表达产物,其中更加优选 ANXA9 基因、LCE2B 基因、LCE2A 基因、IVL 基因以及 CRCT1 基因及其表达产物。

[0301] ANXA9 基因是由序列号 44 所示的多核苷酸构成的基因,其编码的 ANXA9 蛋白质由序列号 45 所示的氨基酸序列构成。报告有 ANXA9 基因作为钙依赖性磷脂结合蛋白质的一种 (Nguyen VT et al., J. Biol. Chem. 275(38), p29466-76, 2000)。该基因可以在 NCBI 的基因数据库中,以 GeneID :8416 查找。该基因可以通过公知的基因操作方法得到。ANXA9 蛋白质可以由序列号 44 所表示的多核苷酸构成的基因表达得到,另外,还可以根据序列号 45 中所示的氨基酸序列信息,通过通常的化学合成方法制造。

[0302] 如后述的实施例所示,分析了日本人的卷发者和非卷发者的毛发发根部中的基因表达,发现和非卷发群相比较,在卷发群中 ANXA9 基因的表达量显著降低,另外,通过将如圆叶牵牛那样具有直发作用的物质给药,从而卷发化被改善,ANXA9 基因的表达量也增加。

[0303] LCE2B 基因是序列号 46 所示的多核苷酸构成的基因,其编码的 LCE2B 蛋白质由序列号 47 所示的氨基酸序列构成。报告有 LCE2B 基因作为表皮细胞最后分化过程中表达的基因 (Jackson B et al., J. Invest. Dermatol. 124(5), p1062-70, 2005)。该基因可以在 NCBI 的基因数据库中,以 GeneID :26239 查找。该基因可以通过公知的基因操作方法得到。LCE2B 蛋白质可以由序列号 46 所表示的多核苷酸构成的基因表达得到,另外,还可以根据序列号 47 中所示的氨基酸序列信息,通过通常的化学合成方法制造。

[0304] 如后述的实施例所示,分析了日本人的卷发者和非卷发者的毛发发根部中的基因表达,发现和非卷发群相比较,在卷发群中 LCE2B 基因的表达量显著降低。

[0305] LCE2A 基因是序列号 48 所示的多核苷酸构成的基因,其编码的 LCE2A 蛋白质由序列号 49 所示的氨基酸序列构成。报告有 LCE2A 基因作为表皮细胞最后分化过程中表达的基因 (Jackson B et al., J. Invest. Dermatol. 124(5), p1062-70, 2005)。该基因可以在 NCBI 的基因数据库中,以 GeneID :353139 查找。该基因可以通过公知的基因操作方法得到。LCE2A 蛋白质可以由序列号 48 所表示的多核苷酸构成的基因表达得到,另外,还可以根据序列号 49 中所示的氨基酸序列信息,通过通常的化学合成方法制造。

[0306] 如后述的实施例所示,分析了日本人的卷发者和非卷发者的毛发发根部中的基因表达,发现和非卷发群相比较,在卷发群中 LCE2A 基因的表达量显著降低。

[0307] IVL 基因是序列号 50 所示的多核苷酸构成的基因,其编码的 IVL 蛋白质由序列号 51 所示的氨基酸序列构成。报告有 IVL 基因作为表皮细胞最后分化过程中表达的基因 (Eckert RL et al., J. Invest. Dermatol. 100(5), p613-7, 1993)。该基因可以在 NCBI 的基因数据库中,以 GeneID :3713 查找。该基因可以通过公知的基因操作方法得到。IVL 蛋白质可以由序列号 50 所表示的多核苷酸构成的基因表达得到,另外,还可以根据序列号 51 中所示的氨基酸序列信息,通过通常的化学合成方法制造。

[0308] 如后述的实施例所示,分析了日本人的卷发者和非卷发者的毛发发根部中的基因表达,发现和非卷发群相比较,在卷发群中 IVL 基因的表达量显著增加,另外,通过将如西番莲那样具有卷发作用的物质给药,从而促进卷发化,IVL 基因的表达量增加。

[0309] CRCT1 基因是序列号 52 所示的多核苷酸构成的基因,其编码的 CRCT1 蛋白质由序列号 53 所示的氨基酸序列构成。报告有 CRCT1 基因作为表皮细胞最后分化过程中表达的

基因 (Marenholz I et al., Genome Res. 11(3), p341-55, 2001)。该基因可以在 NCBI 的基因数据库中, 以 GeneID :54544 查找。该基因可以通过公知的基因操作方法得到。CRCT1 蛋白质可以由序列号 52 所表示的多核苷酸构成的基因表达得到, 另外, 还可以根据序列号 53 中所示的氨基酸序列信息, 通过通常的化学合成方法制造。

[0310] 如后述的实施例所示, 分析了日本人的卷发者和非卷发者的毛发发根部中的基因表达, 发现和非卷发群相比较, 在卷发群中 CRCT1 基因的表达量显著降低, 另外, 通过将如圆叶牵牛那样具有直发作用的物质给药, 从而卷发化被改善, CRCT1 基因的表达量增加。

[0311] (1) 用于检测和 / 或判定毛发形状类型的多核苷酸标记

[0312] 根据本发明, 得到用于检测和 / 判定毛发形状类型的标记 (毛发形状类型的标记) 为具有本发明的毛发形状易感性基因的碱基序列的多核苷酸或者其部分多核苷酸。例如, 作为本发明的毛发形状类型的标记, 可以举出由 ANXA9 基因、FAM63A 基因、LCE5A 基因、CRCT1 基因、LCE2B 基因、LCE2A 基因、SMCP 基因或者 IVL 基因的碱基序列构成的多核苷酸, 优选为由 ANXA9 基因、LCE2B 基因、LCE2A 基因、IVL 基因或者 CRCT1 基因的碱基序列构成的多核苷酸, 更加优选为由序列号 44、序列号 46、序列号 48、序列号 50 或者序列号 52 所示的碱基序列构成的多核苷酸; 和这些碱基序列互补的碱基序列构成的多核苷酸; 以及这些多核苷酸的部分多核苷酸。

[0313] 另外, 本发明的毛发形状类型的标记中, 可以含有相对于上述互补的碱基序列构成的多核苷酸或者其部分多核苷酸的碱基序列, 进一步具有互补关系的碱基序列构成的链。

[0314] 上述多核苷酸及其互补链可以以各自单链的形态作为本发明的标记使用, 也可以以双链的形态作为本发明的标记使用。

[0315] 作为部分多核苷酸来说, 在本发明的毛发形状易感性基因的碱基序列或者和其互补的碱基序列构成的多核苷酸中, 例如可以举连续的具有 15 个碱基以上的长度的部分多核苷酸。部分多核苷酸的长度可以根据其用途适当设定。

[0316] (2) 用于扩增毛发形状类型的标记的引物以及用于检测该标记的探针

[0317] 由本发明的毛发形状易感性基因的碱基序列或者与其互补的碱基序列构成的多核苷酸的部分多核苷酸能成为用于扩增所述毛发形状类型的标记的引物。优选引物扩增由序列号 44、序列号 46、序列号 48、序列号 50 或者序列号 52 所表示的碱基序列、或者与其互补的碱基序列所构成的多核苷酸, 或者这些多核苷酸的部分多核苷酸。

[0318] 另外, 由本发明的毛发形状易感性基因的碱基序列或者与其互补的碱基序列构成的多核苷酸、或者这些多核苷酸的部分多核苷酸能形成用于检测所述毛发形状类型的标记的探针。优选探针检测由序列号 44、序列号 46、序列号 48、序列号 50 或者序列号 52 所表示的碱基序列、或者与其互补的碱基序列所构成的多核苷酸, 或者这些多核苷酸的部分多核苷酸。

[0319] 即, 用于特异性识别并扩增由于 ANXA9 基因、LCE2B 基因、LCE2A 基因、IVL 基因或者 CRCT1 基因的表达所产生的 RNA 或者从其得到的多核苷酸的引物, 或者用于特异地检测出该 RNA 或者从其得到的多核苷酸的探针与此相符。

[0320] 具体来说, 可以在 Northern 杂交法 (Northern Blot)、RT-PCR 法、原位杂交法等特异地检测出特定基因的公知方法中, 根据常用方法用作引物或者探针。

[0321] 作为引物使用的情况下,其碱基长度通常为 15 ~ 100 个碱基、优选为 15 ~ 50 个碱基、更加优选为 15 ~ 35 个碱基。

[0322] 另外,作为检测探针使用的情况下,可以举通常为具有 15 个碱基以上,优选为 15 ~ 1000 个碱基、更加优选为 100 ~ 1000 个碱基的碱基长度的探针。

[0323] 在此,“特异性识别”是指在例如 Northern 杂交法中,能够特异地检测出序列号 44、序列号 46、序列号 48、序列号 50 或者序列号 52 所示的碱基序列或者与其互补的碱基序列构成的多核苷酸,或者这些多核苷酸的部分多核苷酸,此外,在例如 RT-PCR 法中,如能使该多核苷酸特异生成等,能判断上述检测物或者生成物为序列号 44、序列号 46、序列号 48、序列号 50 或者序列号 52 所示的碱基序列或者与其互补的碱基序列所构成的多核苷酸,或者这些多核苷酸的部分多核苷酸。

[0324] 序列号 44、序列号 46、序列号 48、序列号 50 或者序列号 52 所表示的碱基序列或者与其互补的碱基序列构成的多核苷酸的部分多核苷酸可以基于上述序列号所表示的 ANXA9 基因、LCE2B 基因、LCE2A 基因、IVL 基因或者 CRCT1 基因的碱基序列,用例如 primer 3 或者 Vector NTI 的软件进行设计。能够将得到的引物或者探针的候选序列或者至少含有一部分该序列的序列设计为引物或者探针。

[0325] (3) 用于检测和 / 或判定毛发形状类型的多肽标记

[0326] 与上述列举的毛发形状易感性基因同样,这些基因的表达产物(毛发形状易感性基因编码的蛋白质、或者来自这些蛋白质的多肽、或者这些多肽的部分多肽)也能形成毛发形状类型的标记(多肽)。

[0327] 例如,作为上述表达产物,可以举出作为 ANXA9 基因、FAM63A 基因、LCE5A 基因、CRCT1 基因、LCE2B 基因、LCE2A 基因、SMCP 基因或者 IVL 基因编码的蛋白质的 ANXA 蛋白质 9、FAM63A 蛋白质、LCE5A 蛋白质、CRCT1 蛋白质、LCE2B 蛋白质、LCE2A 蛋白质、SMCP 蛋白质、以及 IVL 蛋白质(或者也称为 ANXA9、FAM63A、LCE5A、CRCT1、LCE2B、LCE2A、SMCP 或者 IVL),以及来自这些蛋白质的多肽,以及这些多肽的部分多肽,优选为 ANXA9、LCE2B、LCE2A、IVL 或者 CRCT1,以及来自这些蛋白质的多肽,以及这些多肽的部分多肽。

[0328] 更加优选上述表达产物是由序列号 44、序列号 46、序列号 48、序列号 50 或者序列号 52 所表示的碱基序列构成的多核苷酸编码的蛋白质,进一步优选为由序列号 45、序列号 47、序列号 49、序列号 51 或者序列号 53 所表示的氨基酸序列构成的蛋白质。

[0329] 另外,在上述表达产物中,还包含由序列号 45、序列号 47、序列号 49、序列号 51 或者序列号 53 所表示的氨基酸序列中 1 个或者多个氨基酸缺失、取代或者附加而得到的氨基酸序列构成的,而且和序列号 45、序列号 47、序列号 49、序列号 51 或者序列号 53 所表示的氨基酸构成的蛋白质具有同等的生物学机能的蛋白质和 / 或具有同等的免疫学活性的蛋白质(所谓的 ANXA9、LCE2B、LCE2A、IVL、或者 CRCT1 的相同物)。

[0330] 在此,作为具有同等的生物学机能的蛋白质,可以举和 ANXA9、LCE2B、LCE2A、IVL、或者 CRCT1 在生化学或者药理学的机能上相同的蛋白质。另外,作为具有同等的免疫学活性的蛋白质,可以举在适当的动物或者其细胞中诱发特定的免疫反应,并且具有和对 ANXA9、LCE2B、LCE2A、IVL、或者 CRCT1 的抗体特异结合的能力的蛋白质。

[0331] 另外,确定氨基酸残基的取代、插入或者缺失的指标,可以使用本领域技术人员公知的电脑程序,例如使用 DNA Star 软件来发现。例如,变异数典型的为总氨基酸的 10% 以

内,优选为总氨基酸的5%以内,更加优选为总氨基酸的1%以内。另外,从保持蛋白质的结构的观点出发,被取代的氨基酸优选为在氨基酸的极性、电荷、可溶性、疏水性、亲水性、两亲性等方面具有和取代前的氨基酸相似的性质的氨基酸。

[0332] 作为部分多肽,可以举由本发明的毛发形状易感性基因编码的氨基酸序列(例如,序列号45、序列号47、序列号49、序列号51或者序列号53中所示的氨基酸序列)中至少连续的5个氨基酸,优选为10~100个氨基酸构成的多肽,可以举和本发明的毛发形状易感性基因的表达产物(例如,ANXA9、LCE2B、LCE2A、IVL、或者CRCT1)具有同等的生物学机能和/或同等的免疫学活性的多肽。

[0333] 本发明的毛发形状易感性基因编码的多肽,可以通过基于该毛发形状易感性基因的碱基序列信息,通过DNA克隆、各质粒的构筑、到宿主的转染、转化体的培养以及从培养物中回收蛋白质的操作得到。这些操作可以按照公知的方法,例如,可以按照Molecular Cloning, T. Maniatis et al., CSH Laboratory(1983), DNA Cloning, DM. Glover, IRLPRESS(1985))等所记载的方法进行。

[0334] 具体来说,制作编码ANXA9、LCE2B、LCE2A、IVL、或者CRCT1的基因在所希望的宿主细胞中能表达的重组DNA(表达载体),将其导入到宿主细胞中进行转化,培养该转化体,从得到的培养物中,通过回收目标蛋白质而得到。

[0335] 另外,由本发明的毛发形状易感性基因编码的多肽,可以按照该毛发形状易感性基因编码的氨基酸序列,通过一般的化学合成法制造。

[0336] (4)特异性识别毛发形状类型的标记(多肽)的抗体

[0337] 特异性识别由本发明的毛发形状易感性基因编码的氨基酸序列构成的多肽或者其部分多肽的抗体,可以成为用于检测上述毛发形状类型的标记(多肽)的抗体。

[0338] 如后面所述,通过使用相关的抗体可以检测出被检测者的组织内的毛发形状类型的标记(多肽)(例如,ANXA9、LCE2B、LCE2A、IVL、CRCT1,或者来自这些蛋白质的多肽,或者这些多肽的部分多肽)的表达的有无、以及其表达的程度。具体来说,用活组织检查等采集被检测者的毛发发根部等的一部分,由此按照常用方法调制蛋白质,例如,在Western杂交(Western Blot)法、ELISA法等公知的检测方法中,通过按照常用方法使用本发明的抗体,可以检测出在该组织中存在的上述毛发形状类型的标记(多肽)。

[0339] 该毛发形状类型的检测用抗体可以是将毛发形状类型的标记(多肽)作为免疫抗原的多克隆抗体,也可以是单克隆抗体。

[0340] 这些抗体,可以按照公知的方法制造(Current protocols in Molecular Biology edit. Ausubel et al. (1987) Publish. John Wiley and Sons. Section 11.12-11.13)。具体来说,多克隆抗体可以按照常用方法使用在大肠杆菌等中表达并精制的本发明的毛发形状易感性基因编码的氨基酸序列构成的多肽(例如,ANXA9、LCE2B、LCE2A、IVL、或者CRCT1),或者按照常用方法合成该多肽的部分多肽,对家兔等非人的动物进行免疫,由该免疫动物的血清按照常用方法而得到。

[0341] 另一方面,单克隆抗体可以按照常用方法通过用在大肠杆菌等中表达并精制的上述多肽或者其部分多肽对小鼠等非人动物进行免疫,从将得到的脾脏细胞和骨髓细胞进行细胞融合而调制的杂交瘤细胞中得到(Current protocols in Molecular Biology edit. Ausubel et al. (1987) Publish. John Wiley and Sons. Section 11.4-11.11)。

[0342] 作为在此使用的部分多肽,为具有由本发明的毛发形状易感性基因编码的氨基酸序列构成的多肽(例如,ANXA9、LCE2B、LCE2A、IVL、或者 CRCT1)的部分氨基酸序列的寡肽,虽然不需要具有机能性的生物活性,但是优选具有与由上述氨基酸序列构成的蛋白质同等的免疫原特性。例如,可以举由上述氨基酸序列构成的蛋白质,优选具有和 ANXA9、LCE2B、LCE2A、IVL、或者 CRCT1 同样的免疫原特性而且由在上述氨基酸序列中至少连续的 8 个氨基酸、优选为 15 个氨基酸、更加优选为 20 个氨基酸构成的寡肽。

[0343] 对于相关部分多肽的抗体的制造,也可以通过根据宿主使用种种佐剂提高免疫学反应来进行。虽然没有限定,但是这样的佐剂中含有弗氏佐剂(Freund's adjuvant)、氢氧化铝这样的矿物胶;以及溶血卵磷脂、普卢兰尼克多元醇(Pluronic polyol)、聚阴离子(polyanion)、肽、油乳剂、钥孔血蓝蛋白(Keyhole limpet hemocyanin)、以及二硝基苯酚(dinitrophenol)这样的表面活性物质;BCG(Bacillus Calmette-Guérin、卡介苗)和棒状杆菌属(Corynebacterium)等人用佐剂。

[0344] (5) 毛发形状类型的检测和 / 或判定

[0345] 毛发形状类型的检测・判定是在进行活组织检查等中采集被检测者的毛发发根组织等的一部分,将其中含有的本发明的毛发形状类型的标记作为指标,检测和 / 或判定毛发形状类型。例如,在上述方法中,通过测定本发明的毛发形状易感性基因(ANXA9 基因、LCE2B 基因、LCE2A 基因、IVL 基因、或者 CRCT1 基因)、其互补链、或者其部分多核苷酸的表达水平(表达量)、或者由该基因得到的蛋白质(例如 ANXA9、LCE2B、LCE2A、IVL、或者 CRCT1)、其相同物、或者其部分多肽的表达量,从而检测和 / 或判定毛发形状类型。

[0346] 另外,本发明的检测・判定方法在例如卷发者中,在将改善该卷发的医药品和化妆品等进行给药的情况下,也可以用于判定该卷发改善的有无或者其程度。

[0347] 1) 活体样品

[0348] 在此使用的活体样品为含有被检测者的上皮组织或者上皮细胞、例如毛发发根部和皮肤等能够表达本发明的毛发形状易感性基因(例如,ANXA9 基因、LCE2B 基因、LCE2A 基因、IVL 基因、或者 CRCT1 基因)的细胞的组织,由这些组织调制的 RNA 或者由这些 RNA 进一步调制的多核苷酸或者由上述组织调制的蛋白质。这些 RNA、多核苷酸以及蛋白质可以是在例如在活组织检查等中采集被检测者的毛发发根部的一部分之后,按照常用方法调制。

[0349] 2) 标记的检测和 / 或测定

[0350] 标记的检测以及测定是根据作为测定对象使用的活体样品的种类,具体实施如下。

[0351] (i) 利用 RNA 作为测定对象的活体样品的情况

[0352] 利用 RNA 作为活体样品的情况下,通过检测、测定该 RNA 中的上述本发明的毛发形状类型的标记(多核苷酸),例如 ANXA9 基因、LCE2B 基因、LCE2A 基因、IVL 基因、CRCT1 基因、或者其部分多核苷酸等的表达水平来进行。

[0353] 在此,上述标记的表达量的测定,具体来说,可以使用用于扩增能形成前面所述的本发明标记的多核苷酸的引物或者用于检测该多核苷酸的探针,通过 Northern 杂交法、RT-PCR 法、DNA 芯片分析法、原位杂交分析法等公知的方法实施。

[0354] 在利用 Northern 杂交法的情况下,通过使用本发明的探针,可以检测、测定 RNA 中的上述标记(例如,ANXA9 基因、LCE2B 基因、LCE2A 基因、IVL 基因、CRCT1 基因、或者其部

分多核苷酸)的表达的有无和表达水平。

[0355] 具体来说,首先用放射性同位素(^{32}P 、 ^{33}P 等:RI)、荧光物质等标识探针DNA。接下来,将得到的标识疾病标记按照常用方法和转移到尼龙薄膜等上的来自被检测者的活体组织的RNA杂交。之后,可以例示如下方法,该方法是将形成的标识疾病标记(DNA)和RNA的双链用放射线检测器(BAS-1800II、富士胶卷公司制造)、荧光检测器等来检测、测定该标识疾病标记的标识物(RI、荧光物质等)发出的信号。

[0356] 另外,也可以采用如下方法,该方法是使用AlkPhos DirectTMLabelling and Detection System(Amersham Pharmacia Biotech公司制造),根据该规程来标识探针DNA,将其与来自被检测者的活体组织的RNA杂交之后,用Multibio Imager STORM860(Amersham Pharmacia Biotech公司制造)检测、测定由该探针DNA的标识物发出的信号。

[0357] 在利用RT-PCR法的情况下,使用本发明的引物,可以检测、测定RNA中的上述标记的表达的有无和表达水平。具体来说,首先从来自被检测者的活体组织的RNA中按照常用方法调制cDNA,以将其作为模板以扩增目标标记的区域的方式,将从本发明的标记多核苷酸调制的一对引物(上述cDNA(-链)上结合的正链、+链上结合的负链)与其杂交。之后,按照常用方法进行PCR法,检测得到的扩增双链DNA。

[0358] 在扩增的双链DNA的检测中,可以使用如下方法:检测通过使用预先用RI、荧光物质等标识的引物进行上述PCR而产生的标识双链DNA的方法;按照常用方法将产生的双链DNA转移到尼龙薄膜等上,将标识的疾病标记用作探针与其杂交来检测的方法。生成的标识双链DNA产物可以用Agilent 2100bioanalyzer(Yokogawa Analytical Systems Inc.制造)等测定。另外,在SYBR(注册商标)Green RT-PCR Reagents(Applied Biosystems公司制造)中按照该规程调制RT-PCR反应液,使其在ABI PRIME(注册商标)7700Sequence Detection System(Applied Biosystems公司制造)中反应,也能检测出该反应物。对于使用了所述的RT-PCR法的被检测者RNA中的本发明的毛发形状类型的标记(多核苷酸)的表达水平的检测、测定,在后面的实施例中详细叙述。

[0359] 在利用DNA芯片分析的情况下,准备附有本发明的DNA探针(单链或者双链)的DNA芯片,在其上杂交从来自被检测者的活体组织的RNA按照常用方法调制的cRNA,将形成的DNA和cRNA的双链与由本发明的标记多核苷酸调制的标识探针结合,从而能够检测、测定本发明的标记的表达的有无以及表达水平。

[0360] 另外,作为上述的DNA芯片,也可以使用能够检测、测定本发明的标记的表达水平的DNA芯片。作为该DNA芯片来说,可以举例如Affymetrix公司的GeneChip(注册商标)Human Genome U133plus 2。

[0361] (ii) 使用蛋白质作为测定对象的活体样品的情况

[0362] 在使用蛋白质作为测定对象的情况下,可以通过使本发明的抗体和活体样品相接触,检测 and 该抗体结合的活体样品中的本发明的毛发形状类型的标记(多肽),例如ANXA9、LCE2B、LCE2A、IVL、CRCT1、或者其部分多肽,测定其量(水平)来实施。

[0363] 在此,蛋白质的结合量的测定,可以通过使用Western杂交法等公知的方法进行。

[0364] Western杂交法可以通过在使用本发明的抗体作为第一抗体之后,使用与用 ^{125}I 等的放射性同位素、荧光物质、辣根过氧化物酶(horseradish peroxidase、HRP)等酶等标

识的第一抗体相结合的抗体作为第二抗体,标识其第一抗体,用放射线测定器、荧光检测器等测定这些标识物质发出的信号来实施。另外,也可以在使用本发明的抗体作为第一抗体之后,使用 ECL Plus Western Blotting Detection System(Amersham Pharmacia Biotech 公司制造),按照该规程检测,用 Multibio Imager STORM860(Amersham Pharmacia Biotech 公司制造)进行测定。

[0365] 3) 毛发形状类型的判定

[0366] 毛发形状类型的判定可以通过将按照上述测定的被检测者的活体样品中本发明的标记的水平(例如,ANXA9 基因、LCE2B 基因、LCE2A 基因、IVL 基因、CRCT1 基因的基因表达水平,或者 ANXA9、LCE2B、LCE2A、IVL、CRCT1 的量等),与非卷发者的该水平相比较,判定两者的不同来进行。

[0367] 被检测者的活体样品和非卷发者的活体样品之间的标记多核苷酸或者多肽的表达水平的比较,可以通过同时进行将被检测者的活体样品和非卷发者的活体样品作为对象的测定来实施。另外,即使不同时进行,也可以预先求得使用多个(至少 2 个,优选为 3 个以上,更加优选为 5 个以上)的非卷发者的组织在均匀的测定条件下测定得到的标记多核苷酸(ANXA9 基因、LCE2B 基因、LCE2A 基因、IVL 基因、CRCT1 基因、或者其部分多核苷酸等)的基因表达水平或者标记多肽(ANXA9、LCE2B、LCE2A、IVL、CRCT1 或者其部分多肽等)的表达水平的平均值或者统计的中间值,将此作为非卷发者的标记多核苷酸或者多肽的表达水平,用于被检测者中的测定值的比较。

[0368] 被检测者的毛发形状类型的判定可以将被检测者的组织中的标记多核苷酸(ANXA9 基因、LCE2B 基因、LCE2A 基因、IVL 基因、CRCT1 基因、或者其部分多核苷酸等)的基因表达水平或者标记多肽(ANXA9、LCE2B、LCE2A、IVL、CRCT1 或者其部分多肽等)的表达水平与非卷发者的这些比较的情况下的增加或者减少的程度(例如 2 倍以上,优选为 3 倍以上左右)作为指标进行。

[0369] 例如,如果被检测者的 ANXA9 基因或者 ANXA9 蛋白质的表达水平与非卷发者的这些水平相比少的话,则可以判断该被检测者为卷发者,或者怀疑将来可能会出现卷发。

[0370] 另外例如,如果被检测者的 LCE2B 基因或者 LCE2B 蛋白质的表达水平与非卷发者的这些水平相比少的话,则可以判断该被检测者为卷发者,或者怀疑将来可能会出现卷发。

[0371] 另外例如,如果被检测者的 LCE2A 基因或者 LCE2A 蛋白质的表达水平与非卷发者的这些水平相比少的话,则可以判断该被检测者为卷发者,或者怀疑将来可能会出现卷发。

[0372] 例如,如果被检测者的 IVL 基因或者 IVL 蛋白质的表达水平与非卷发者的这些水平相比多的话,则可以判断该被检测者为卷发者,或者怀疑将来可能会出现卷发。

[0373] 例如,如果被检测者的 CRCT1 基因或者 CRCT1 蛋白质的表达水平与非卷发者的这些水平相比少的话,则可以判断该被检测者为卷发者,或者怀疑将来可能会出现卷发。

[0374] 7. 毛发形状的调节方法

[0375] 通过变换位于本发明的毛发形状易感性 SNP 标记处的碱基可以调节个体的毛发形状。

[0376] 即,本发明另外还提供个体的毛发形状的调节方法。在一个实施方式中,该方法可以为以美容为目的的调节毛发形状的非治疗的方法,可以由美容师或理发师等进行。另外,在本说明书中,“非治疗的”是指不含医疗行为,即通过治疗对人体的处置行为的概念。

[0377] 该方法可以通过变换位于上述列举的本发明的毛发形状易感性 SNP 标记处的碱基来达成。作为具体的手法,只要是能够达成上述目的的方法都没有特别的限定,可以使用以前公知的方法以及将来开发的手法的任意方法,例如可以举例利用基因重组的方法等。

[0378] 或者,本发明的毛发形状的调节方法,可以通过控制有必要进行毛发形状调节的人的毛发发根部的本发明的毛发形状易感性基因的表达来进行。

[0379] 例如,对于困扰于卷发和卷缩发的人来说,通过诱导或者促进其表达对直发的表现型起作用的毛发形状易感性基因,例如 ANXA9 基因、LCE2B 基因、LCE2A 基因、CRCT1 基因的表达,可以抑制卷发和卷缩发。或者,通过阻碍其表达会对卷发或者卷缩发的表现型起作用的毛发形状易感性基因,例如 IVL 基因的表达,可以抑制卷发和卷缩发。另一方面对于希望头发波浪化的人来说,可以通过诱导或者促进其表达对卷发和卷缩发的表现型起作用的毛发易感性基因,例如 IVL 基因的表达,能够表达或者促进波浪化。或者通过阻碍其表达对直发的表现型起作用的毛发形状易感性基因,例如 ANXA9 基因、LCE2B 基因、LCE2A 基因、CRCT1 基因的表达,能够表达或者促进波浪化。

[0380] 例如,在抑制卷发和卷缩发的情况下,可以使人的毛发发根部的 ANXA9 基因、LCE2B 基因、LCE2A 基因、CRCT1 基因的表达水平达到非卷发者中该基因的 mRNA 表达水平以上,例如为约 3 ~ 10 倍以上。另一方面,在以促进波浪化为目的的情况下,可以使 ANXA9 基因、LCE2B 基因、LCE2A 基因、或者 CRCT1 基因的表达水平比非卷发者中该基因的 mRNA 表达水平低,例如为约 3 ~ 10 倍以下。

[0381] 另外例如,在抑制卷发和卷缩发的情况下,可以使人的毛发发根部的 IVL 基因的表达水平达到非卷发者中该基因的 mRNA 表达水平以下,例如为约 3 ~ 10 倍以下。另一方面,在以促进波浪化为目的的情况下,可以使 IVL 基因的表达水平比非卷发者中该基因的 mRNA 表达水平高,例如为约 3 ~ 10 倍以上。

[0382] 抑制、诱导或者促进人的毛发发根部上毛发形状易感性基因的表达,可以按照常用方法进行。例如,在基因抑制中,可以举出通过反义核苷酸的方法,例如阻碍从 mRNA 的翻译的方法等的手段,而在诱导或者促进中,则通过病毒载体等基因导入等来表达毛发形状易感性基因的手段等。另外,抑制毛发形状易感性基因编码的蛋白质的表达,基本上可以通过抑制该基因的表达的手段实现,另外,在诱导或者促进该蛋白质的表达中,除了使该基因高表达之外,可以举向皮内直接注入该蛋白质的人重组蛋白质的手段等。

[0383] 利用上述反义核苷酸的基因导入可以和通常基因治疗中使用的方法同样进行,例如,通过以反义寡核苷酸或者其化学修饰体直接对被检测者的体内给药来控制本发明的毛发形状易感性基因的表达的方法,或者通过向患者的目标细胞中导入反义 RNA 通过该细胞控制本发明的毛发形状易感性基因的表达的方法来实施。

[0384] 在此,“反义核苷酸”中包含对应于本发明的毛发形状易感性基因的至少 8 个碱基以上的部分的反义寡核苷酸、反义 RNA、反义 DNA 等。在其化学修饰体中,包含例如硫代磷酸酯、二硫代磷酸酯、烷基磷酸三酯、烷基磷酸酯、烷基亚磷酸酰胺等能提高向细胞内转移性和在细胞内的稳定性的衍生物(“Antisense RNA and DNA”, WILEY-LISS 期刊,1992 年, pp. 1-50, J. Med. Chem. 36, 1923-1937(1993))。

[0385] 反义核苷酸或者其化学修饰体在细胞内与有义链 mRNA 结合,可以控制毛发形状易感性基因的表达,即,毛发形状易感性基因编码的蛋白质的表达,并能控制相关的该蛋白

质的机能（活性）。

[0386] 在将反义寡核苷酸或者其化学修饰体直接对活体内给药的方法中，所用的反义寡核苷酸或者其化学修饰体优选具有 5～200 个碱基长度，更加优选为 8～25 个碱基长度，最优选为 12～25 个碱基的长度。在给药的时候，可以对反义寡核苷酸或者其化学修饰体使用通常使用的稳定剂、缓冲液、溶剂等进行制剂化。

[0387] 在将反义 RNA 导入被检测者的目标细胞中的方法中，所用的反义 RNA 优选为具有 100 个碱基以上，更加优选为具有 300 个碱基以上，更加优选为具有 500 个碱基以上的长度。另外，此方法包含向活体内的细胞中导入反义基因的体内（in vivo）法以及向一旦取出体外的细胞中导入反义基因，再将该细胞放回体内的体外（ex vivo）法（参考 NIKKEI SCIENCE, 1994 年 4 月号, 20-45 页、月刊药事, 36(1), 23-48(1994)、实验医学增刊, 12(15), 全页(1994) 等）。其中，优选体内法，体内法中有病毒导入法（使用重组病毒的方法）和非病毒导入法（参考前面所述各文献）。

[0388] 作为使用重组病毒的方法，例如，可以举向逆转录病毒（retrovirus）、腺病毒（adenovirus）、腺相关病毒（adeno-associated virus）、疱疹病毒（herpesvirus）、牛痘病毒（vaccinia virus）、脊髓灰质炎病毒（poliovirus）、辛德毕斯病毒（sindbis virus）等病毒基因组中重组 MLTK 基因的反义核苷酸而导入活体内的方法。其中，特别优选使用逆转录病毒、腺病毒、腺相关病毒等的方法。作为非病毒导入法，可以举脂质体（liposome）法、脂质体转染（lipofection）法等，特别优选使用脂质体法。作为其它的非病毒导入法，也可以举例如，微注射法、磷酸钙法、电穿孔法等。

[0389] 基因导入用制剂组合物是以上述反义核苷酸或者其化学修饰体，含有这些的重组病毒以及导入了这些病毒的感染细胞等为有效成分的组合物。

[0390] 将该组合物向被检测者的给药，可以以例如注射剂等的适当的给药方式，向静脉、动脉、皮下、肌肉内等给药，另外还可以从患者的皮肤直接给药、导入。在采用体内法的情况下，基因导入用组合物除了可以是含有毛发形状易感性基因的反义核苷酸的注射剂等的给药形态之外，还可以制成例如将含有毛发形状易感性基因的反义核苷酸的病毒载体包埋在脂质体或者膜融合脂质体中的形态（仙台病毒（Hemagglutinating virus of Japan, HVJ）-脂质体等）。这些脂质体制剂的形态中，含有悬浊液、冷冻剂、离心分离浓缩冷冻剂等。另外，基因导入用组合物也可以制成被导入了含有上述毛发形状易感性基因的反义核苷酸的载体的病毒所感染的细胞培养液的形态。这些各种形态的制剂中的有效成分的给药量可以根据作为治疗目的的疾病的程度、患者的年龄、体重等进行适当的调节。通常，在相对于毛发形状易感性基因的反义核苷酸的情况下，设定为被检测者成人每 1 人在数天～数月内一次给药约 0.0001～100mg，优选为约 0.001～10mg 的量即可。

[0391] 在含有反义核苷酸的逆转录病毒载体的情况下，作为逆转录病毒滴度，可以从 1 天患者体重每 1kg 约 1×10^3 pfu～ 1×10^{15} pfu 的量的范围内选择。在导入了反义核苷酸的细胞的情况下，可以以 1×10^4 细胞/body～ 1×10^{15} 细胞/body 左右给药。

[0392] 8. 毛发形状调节剂的评价或者选择方法

[0393] 另外，本发明提供用于评价或者选择毛发形状调节剂的方法（筛选方法）。

[0394] 该筛选方法可以按照例如下面所述的工序实施。

[0395] (a) 向含有本发明的毛发形状易感性基因的细胞将被检测物质给药的工序；以及

[0396] (b) 从被给药的被检测物质中,选择使在该毛发形状易感性基因之上或者其附近,例如存在于该基因所属的单倍域上的本发明的毛发形状易感性 SNP 标记的碱基的多态性变换为另一种多态性的物质作为毛发形状调节剂的工序。

[0397] 在上述工序(a)(被检测物质的给药工序)中使用的细胞,只要是能够导入人1号染色体的基因组区域中序列号1~5表示的碱基序列所表示的单倍域、或者至少是与该单倍域相重叠的基因,即本发明的毛发形状易感性基因,并且能够稳定地保持该基因的细胞即可,其来源(例如,原核细胞以及真核细胞的区别、昆虫细胞和动物细胞的区别等)没有特别的限制,另外,基因导入和细胞培养等可以任意使用在本领域中现有公知的方法来实施(例如,Joseph Sambrook et al., Molecular Cloning: A Laboratory Manual (3Vol. Set), Cold Spring Harbor Laboratory, NY, 2001、日本组织培养学会编,《组织培养的技术、第3版、基础编·应用编》,朝仓书店,1996等)。该细胞可以在用于评价或者选择对调节毛发形状有效的物质的方法(筛选方法)中,作为筛选工具有效利用。

[0398] 被给药的被检测物质没有特别的限制。例如,可以举天然化合物、有机化合物、无机化合物、蛋白质以及肽等单一化合物、以及化合物库、基因库的表达产物、细胞提取物、细胞培养上清液、发酵微生物产生物、海洋生物提取物、植物提取物等任意的化合物或者组合物。

[0399] 在上述的工序(b)(毛发形状调节剂的选择工序)中,检测出碱基的多态性的变换的有无以及变换后的碱基的种类。作为用于检测碱基多态性的变换的有无以及变换了的碱基的种类的方法,可以是直接测定碱基的种类的方法,另外也可以是能间接评价碱基变化的方法。例如,作为直接测定碱基的方法,可以举PCR-SSCP、PCR-RFLP、PCR-SSO、PCR-ASP、定向测序法、SNaPshot、dHPLC、Sniper法、MALDI-TOF/MS法等本领域的技术人员公知的方法,作为间接评价的方法,可以举通过测定作为该目的的碱基变换所产生/增加、或者消失/降低的机能、活性或者特定的mRNA量、蛋白质量的方法。

[0400] 用该方法选择的物质,可以作为对毛发形状的调节有效的毛发形状调节剂使用,另外,可以在用于制造含有该制剂的医药品、医药部外品、化妆品或者健康食品等的制造中使用。根据必要,通过将选择的物质进一步用于其它的药理试验和临床试验以及毒性试验,可以取得对人类更有效并安全的毛发形状调节剂。

[0401] 或者,上述筛选方法也可以在例如本发明的毛发形状易感性基因或者该基因编码的蛋白质能够表达的组织或细胞中,将该基因或者蛋白质的表达作为指标进行。

[0402] 具体可以通过下面的(a)~(d)的工序进行。

[0403] (a) 使被检测物质与本发明的毛发形状易感性基因或者该基因编码的蛋白质能够表达的组织或者细胞相接触的工序;

[0404] (b) 测定该组织或者细胞的该基因或者该蛋白质的表达量的工序;

[0405] (c) 将(b)中测定的表达量和不接触被检测物质的对照组织或者细胞的该基因或者该蛋白质的表达量进行比较的工序;

[0406] (d) 基于(c)的结果,选择使该基因或者该蛋白质的表达量减少或者增加的被检测物质作为毛发形状调节剂的工序。

[0407] 在此,作为被使用的本发明的毛发形状易感性基因或者该基因编码的蛋白质能够表达的组织或者细胞,只要是表达该基因或者该蛋白质的组织或者细胞,不问其种类,可以

是哺乳动物的组织或者细胞,例如可以举从人身上采集的皮肤组织、毛发发根部组织(毛囊组织)、表皮角化细胞、来自毛发发根部的细胞、以及上皮组织的细胞株等。该细胞中,也包含用本发明的毛发形状易感性基因(具有该基因的表达载体)转化而成的转化细胞。

[0408] 将该组织或者细胞与被检测物质接触,例如,可以在将被检测物质预先添加到培养液中达到所规定的浓度之后,将组织或者细胞载置于培养液中或者可以通过向载置有组织或者细胞的培养液中添加被检测物质到所规定的浓度来进行。

[0409] 培养液可以举例如 DMEM 培养基、MCDB 培养基、Willam's E 培养基、RPMI1640 培养基、DMEM/HamF12(1 : 1) 培养基、以及市售的各种上皮细胞用培养基等,也可以添加适当的琼脂和明胶。另外,也可以根据必要,添加抗生素、氨基酸、血清、生长因子、活体提取物等。

[0410] 组织培养可以通过例如向添加了培养液的 24 孔细胞培养板中加入采集的毛发发根部组织(毛囊组织),在 37℃ 的温度下,在含有 CO₂ 的空气气相条件下,通常培养 1 ~ 30 天,优选为 1 ~ 21 天时间来进行。

[0411] 另外,细胞培养可以通过例如向添加了培养液的 24 孔细胞培养板中加入细胞,在 37℃ 的温度下,在含有 CO₂ 的空气气相条件下,通常培养 1 ~ 7 天,优选为 1 ~ 3 天时间来进行。

[0412] 上述基因的表达的测定(定量)可以根据前面所述的毛发形状类型的标记的检测·测定((5)-2)-(i))中记载的方法进行。即,使用用于扩增能成为本发明的标记的多核苷酸的引物或用于检测该多核苷酸的探针,通过进行 Northern 杂交法、RT-PCR 法、DNA 芯片分析法、原位杂交分析法等公知的方法来实施。

[0413] 另外,上述蛋白质的表达的测定(定量)可以根据前面所述的毛发形状类型的标记的检测·测定((5)-2)-(ii))中记载的方法进行。即,使用了识别本发明的毛发形状类型的标记(多肽)的抗体的 Western 杂交法等公知的方法进行定量。

[0414] 2) 另外,本发明的毛发形状易感性基因的表达水平的测定可以在控制该基因的表达的基因区域(表达控制区域)上,使用导入了结合有例如萤光素酶基因等的报告基因的融合基因的细胞株,通过测定来自报告基因的蛋白质量或者其活性来实施。

[0415] 即,本发明中毛发形状调节剂的评价或者选择方法另外可以通过以下的(a) ~ (c) 的工序进行。

[0416] (a) 向本发明的毛发形状易感性基因能够表达的细胞中,导入该毛发形状易感性基因的表达控制区域和报告基因的融合基因,并在被检测物质的存在条件下和不存在条件下培养该细胞的工序;

[0417] (b) 测定在被检测物质的存在条件下培养的细胞培养物中的报告基因表达产物的表达量,并将该量与在被检测物质的不存在条件下培养的细胞培养物中的报告基因表达产物的表达量进行比较的工序;

[0418] (c) 基于上述(b)的比较结果,选择使该报告基因表达产物的表达量增减的被检测物质作为毛发形状调节剂的工序。

[0419] 作为上述报告基因,优选催化发光反应或者显色反应的酶的结构基因。具体来说,可以举上述萤光素酶基因、分泌型碱性磷酸酶基因、氯霉素乙酰基转移酶基因、β-葡萄糖苷酸酶基因、β-半乳糖苷酶基因、水母发光蛋白基因等。

[0420] 另外,作为毛发形状易感性基因的表达控制区域,可以使用例如该基因的转录开

始部位上游约 1kb ~ 约 10kb, 优选为约 2kb, 例如, 对于 ANXA9 基因、LCE2B 基因、LCE2A 基因、IVL 基因或者 CRCT1 基因, 分别可以举序列号 54 ~ 58 所示的碱基序列构成的区域。融合基因的制作和来自报告基因的活性测定可以按照公知的方法进行。

[0421] 所谓降低毛发形状易感性基因的表达量的物质, 是指抑制与构成该基因的多核苷酸有互补性的 mRNA 的表达或者促进其分解的物质, 所谓降低毛发形状易感性基因编码的蛋白质的表达量的物质, 是指抑制该毛发形状易感性基因或者其蛋白质的表达, 或者促进该基因或者蛋白质的分解, 作为结果减少其蛋白质的表达量的物质。

[0422] 所谓增加毛发形状易感性基因的表达量的物质, 是指促进与构成该基因的多核苷酸有互补性的 mRNA 的表达或者抑制其分解的物质, 所谓增加毛发形状易感性基因编码的蛋白质的表达量的物质, 是指促进该毛发形状易感性基因或者其蛋白质的表达, 或者抑制该基因或者蛋白质的分解, 作为结果增加其蛋白质的表达量的物质。

[0423] 增加毛发形状易感性基因或者该基因编码的蛋白质的表达量的物质成为卷发和卷缩发的降低剂或者促进剂。例如, 增加 ANXA9 基因、LCE2B 基因、LCE2A 基因或者 CRCT1 基因, 或者这些基因编码的蛋白质的表达量的物质, 可以作为卷发和卷缩发的降低剂或者改善剂, 另一方面, 降低这些的基因或者蛋白质的表达的物质, 可以作为卷发和卷缩发的促进剂或者波浪化促进剂。另外, 例如, 增加 IVL 基因或者由其编码的蛋白质的表达量的物质, 可以作为卷发和卷缩发的促进剂或者波浪化的促进剂, 另一方面, 降低该基因或者蛋白质的表达的物质, 可以作为卷发或者卷缩发的降低剂或者改善剂。通过以相关的毛发形状调节剂对人给药, 可以得到用于改善卷发或者卷缩发或者促进毛发的波浪化的医药品、化妆品等。

[0424] 3) 另外, 本发明的毛发形状调节剂的评价或者选择方法, 也可以将本发明的毛发形状易感性基因编码的蛋白质的机能 (活性) 作为指标进行。

[0425] 上述蛋白质的机能或者活性可以举例如, 乙酰胆碱受体活性 (Nguyen VT et al., J. Biol. Chem. 275 (38), p29466-76, 2000) 和磷脂酰丝氨酸结合能 (Goebeler V et al., FEBS Lett. 546 (2-3), p359-64, 2003), 该蛋白质的量和机能或者活性具有一定的相关关系。因此, 代替测定上述蛋白质的量, 也可以通过进行该蛋白质的机能或者活性的测定, 进行毛发形状调节剂的评价或者选择。

[0426] 具体来说, 通过下述的工序 (a)、(b) 以及 (c) 的工序进行。

[0427] (a) 使被检测物质和含有本发明的毛发形状易感性基因编码的蛋白质的水溶液、组织细胞或者由该组织细胞调制的细胞组分接触的工序;

[0428] (b) 测定接触了被检测物质的水溶液、组织细胞或者细胞组分中的该蛋白质的机能或者活性, 并将该机能或者活性与不接触被检测物质的对照水溶液、对照细胞或者对照细胞组分中的该蛋白质的机能或者活性相比较的工序;

[0429] (c) 基于上述 (b) 的比较结果, 选择使该蛋白质的机能或者活性增减的被检测物质的工序。

[0430] 作为含有毛发形状易感性基因编码的蛋白质的水溶液, 除了例如 ANXA9、LCE2B、LCE2A、IVL 或者 CRCT1 的水溶液之外, 可以举出例如含有这些蛋白质的组织细胞溶解液、核提取液或者细胞的培养上清液等。在此, 所使用的细胞可以举出表达本发明的毛发形状易感性基因 (例如, ANXA9 基因、LCE2B 基因、LCE2A 基因、IVL 基因或者 CRCT1 基因), 并且具

有由这些基因编码的蛋白质作为其表达产物的细胞。具体来说,可以使用哺乳动物的组织或细胞,例如,从人身上采集的皮肤组织、毛发发根部组织(毛囊组织)、表皮角化细胞、来自毛发发根部的细胞、以及上皮组织的细胞株等。该细胞中也包含用本发明的毛发形状易感性基因(或者具有其的表达载体)转化而成的转化细胞。作为该转化细胞中所用的宿主细胞,可以举例如 HeLa 细胞、COS 细胞、HEK293 细胞、MDCK 细胞、CHO 细胞、HL60 细胞等公知的细胞。另外,细胞组分是指来自上述细胞的各种组分,其中包含例如细胞膜组分、细胞质组分、细胞核组分等。

[0431] 本发明的毛发形状易感性基因编码的蛋白质的活性在例如测定乙酰胆碱受体活性和磷脂酰丝氨酸结合能的情况下,可以用结合分析法(binding assay)、免疫共沉淀法、pull-down 分析、Two-hybrid 法(Y2H)、荧光偏振法、时间分辨荧光共振能量转移(TR-FRET)法等公知的方法测定(例如,中内启光/编,免疫学的规程(免疫学的プロトコール),羊土社,2004 年、竹绳忠臣/编,解析蛋白质相互作用的最适方法(タンパク質相互作用を解明する最適メソッド),Biotechnology-Journal Vol. 5No. 6、羊土社、2005 年)。即,使用含有由毛发形状易感性基因编码的蛋白质的水溶液等将该蛋白质固化成薄膜或者板状,通过检测出用放射性同位素标识的乙酰胆碱或磷脂酰丝氨酸的结合量来测定。抑制(减少)该蛋白质的机能(活性)的物质称为降低乙酰胆碱受体活性或磷脂酰丝氨酸结合能的物质,亢进(增加)该蛋白质的机能(活性)的物质称为增加乙酰胆碱受体活性或磷脂酰丝氨酸结合能的物质。例如,亢进 ANXA9、LCE2B、LCE2A 或者 CRCT1 的机能(活性)的物质可以成为卷发和卷缩发的改善剂,抑制这些蛋白质的机能(活性)的物质能成为波浪化促进剂。另外,例如,亢进 IVL 的机能(活性)的物质,可以成为波浪化促进剂,抑制这些蛋白质的机能(活性)的物质,能成为卷发和卷缩发的改善剂。

[0432] 实施例

[0433] 以下举实施例来说明本发明。

[0434] 实施例 1 毛发形状的定义和卷发家族的收集

[0435] 在本实施例中,为了认定毛发形状易感性基因,以日本人群体为对象进行患病同胞对连锁分析,以及病例・对照例关联分析。

[0436] 一般,毛发形状根据人种而会有不同,亚洲人种相对直发较多,非洲人种卷缩发(或者卷发)是主流。印欧人种中为这中间的波状发(波浪发)的特性的比例较高。因为日本人群体为直发占优势的群体,所以将作为毛发形状具有卷发特性的人规定为患病者(病例),将具有直发特性的人规定为对照者(对照例)。在连锁分析等遗传分析中,有必要将作为对象的特性进行某种程度上的定量,例如,考虑卷发=1、直发=0 这样的二值化的方法,和将卷发的程度用某些方法测定而数值化的方法。但是,人的毛发形状多种多样,现在其测量和分类的方法也不十分确定。在此,首先进行了毛发形状的表现型的正确的分类。毛发形状根据毛发的总体特征和卷曲的程度(卷曲半径)规定。另外,不仅仅是一根毛发的卷曲特性,其和周围的毛发群的卷曲的同步性也是规定毛发形状的重要因素。在此,基于各种人种的毛发形状实际状态,将毛发形状的表现型分类如表 4 所示。本分类可以适用于以日本人群体为首的各种人种群体。另外,图 1 中显示毛发形状的表现型的图像。

[0437] [表 4]

[0438] 毛发形状的表现型的分类

[0439]

	特征	卷曲半径	毛发形状类型
类型 1	即使毛发长度改变也总体上呈现一个 C 卷曲, 或者只在发梢有一个卷曲的毛发	毛发总体上卷曲半径为 9.5cm 以上, 或者只在发梢有 3cm 以上的卷曲	直发
		毛发总体上卷曲半径小于 9.5cm, 或者只在发梢有小于 3cm 的卷曲	大略直发、或者微波浪发
类型 2	根据毛发的长度具有以其毛发特有的卷曲半径数次重复卷曲, 并且和周围的毛发卷曲周期同步的毛发	毛发总体上卷曲半径为 9.5cm 以上	大略直发、或者微波浪发
		毛发总体上卷曲半径为 3cm 以上且小于 9.5cm	波浪发
		毛发总体上卷曲半径小于 3cm	卷发、或者强波浪发
类型 3	具有毛发一根一根细密地重复的卷曲, 和周围的毛发卷曲周期不同步的毛发		卷缩发

[0440] 另一方面, 毛发形状的表现型是在群体中连续变化的量的特性, 由评价可知以何种程度判断为卷发特性或者直发特性。在本发明中, 基于毛发形状的实际状态, 将所述分类中的卷缩发、以及卷发或者强波浪发规定为卷发特性, 将波浪发、大略直发或者微波浪发、直发规定为直发 (非卷发) 特性。

[0441] 这样虽然可以对毛发形状的表现型进行正确的分类, 但是, 在收集遗传分析的对象的时候, 考虑应该解决以下问题。即, 在收集的时候毛发为明显的短发不能进行形状的评价的情况, 经过烫发和染发、各种定型剂等化学处理而改变原来的形状的情况。因此, 对于能成为遗传分析的对象者的所有候选者, 要求其提供在能判断其毛发形状的表现型的时候 (例如幼小时期) 拍摄的候选者本人的照片。即, 不是明显短发、并且毛发未经过化学处理的状态下的照片。同时要求所有的候选者提供多根头发。将提供的头发在化学处理的影响会消失的水浸条件下对毛发进行扭曲和拧曲、卷缩、卷曲特性等详细的形状评价。根据提供的候选者本人的照片进行的毛发形状评价、以及根据提供的毛发的形状评价、进而最后见面时进行的毛发形状调查来确定遗传分析的对象者。

[0442] 这样, 花了约 2 年时间, 从日本全国 3000 个以上的候选者中收集了 68 个家族 283 人的卷发家族。其具体为: 41 组的 2 人兄弟姐妹、22 组的 3 人兄弟姐妹、4 组的 4 人兄弟姐妹、1 组的 5 人兄弟姐妹, 最终的患病同胞对 (具有卷发特性的兄弟姐妹) 为 100 对。考虑遗传因素的强度、兄弟姐妹中的风险率, 能预测到以该同胞对数能充分地特定基因位点, 因此, 实施患病同胞对连锁分析。

[0443] 在从遗传分析的对象者身上收集检测体的时候, 事前得到了伦理委员会的承认, 知情同意实施担当者用书面对对象者说明了研究内容, 在得到了书面的同意的情况下收集了检测体。

[0444] 医生或者护士从遗传分析的对象者身上采集了约 20mL 的血液。从血液检测体中, 使用 PUREGENE Genomic DNA Purification Kit (Gentra Systems Inc. 制造) 根据指南提取了基因组 DNA。将基因组 DNA 溶解于 2mL 的 DNA 水合溶液 (DNA Hydration Solution)

中,测定浓度之后,在 4℃下保存。基因组 DNA 的平均收集量为 576.2g/20mL 血液。

[0445] 实施例 2 全基因组内的患病同胞对连锁分析

[0446] 在本实施例中,首先以日本人卷发家族作为对象,进行包含全基因组的患病同胞对连锁分析。如果简单叙述该方法的原理,是因为患病兄弟姐妹从父母身上继承了构成患病原因的等位基因,因此肯定共有该等位基因。另一方面,兄弟共有的等位基因的数目为 1(基于零假设(null hypothesis)的值)。通过探讨多个患病同胞对所共有的等位基因数,根据基于零假设的等位基因数,在能观察较多的等位基因共有的时候,认定为连锁。

[0447] 患病同胞对连锁分析使用 Applied Biosystems, Inc. (ABI) 制造的全基因组扫描、连锁分析试剂盒 (ABI PRISM Linkage Mapping Set-MD10v2.5) 进行。这是用于扩增作为平均分布于基因组中的多态性丰富的短重复序列的微卫星,并且通过共计 400 个的分型用荧光引物组以平均 9.2cM 的间隔来覆盖人染色体的全基因组扫描、连锁分析试剂盒。

[0448] 将在实施例 1 中调制的基因组 DNA 作为模板,使用 Linkage Mapping Set 进行 PCR (GeneAmp PCR System 9700G、ABI 制造),通过 ABI Prism 3100 Genetic Analyzer (ABI 制造) 进行扩增产物(片段)的检测。分型用荧光引物组为用 6-FAM(蓝色)、VIC(绿色)、NED(黄色)3 种荧光色素标识的引物,因此,即使是同一尺寸的片段也能分别判别 3 种颜色。这样,能迅速地处理大量的样品。

[0449] 片段的分型是通过 Genotyper 软件 v3.7 (ABI 制造) 和 GeneScan 软件 (ABI 制造) 进行的。

[0450] 连锁的检测使用作为非参数分析的 Genehunter v2.1_r5 软件 (Kruglyak, L. et al., Am. J. Hum. Genet., 58(6), 1347-1363, 1996) 进行。连锁被确认区域的判定是按照以下所示的 Lander 和 Kruglyak 的指导原则 (Nat. Genet., 11(3), 241-247, 1995), 以得到假阳性的连锁为基准。

[0451] Lander 和 Kruglyak 的指导原则,虽然现在盛行在多因素疾病中进行全基因组的连锁分析,但是对于各个基因的连锁分析来说,也加上由于该基因功能是否构成原因的判断。但是,因为在全基因组内的分析中不需要在此阶段考虑基因功能,因此,可以求得纯粹的数理遗传学上的显著的判断基准(阈值)。在此,他们根据模拟,设计了下述表 5 中所示的显著的连锁基准。

[0452] [表 5]

[0453]

提示连锁 (在全基因组内得到 1 个假阳性的连锁结果的基准)	$P < 7.4 \times 10^{-4}$ $LOD > 2.2$
显著连锁 (在全基因组内得到 0.05 个假阳性的连锁结果的基准)	$P < 2.2 \times 10^{-5}$ $LOD > 3.6$
高度显著连锁 (在全基因组内得到 0.01 个假阳性的连锁结果的基准)	$P < 3.0 \times 10^{-7}$ $LOD > 5.4$

[0454] 筛选全部染色体的结果识别出对 1 号染色体、以及 11 号染色体连锁。其结果分别

如图 2、图 3 所示。如图 2 所示,在 1 号染色体中,在 1q21 ~ 1q23.1 区域 (D1S498 附近) 中,最大 LOD 分数 = 3.49,1q32 ~ 1q41 区域 (D1S249 ~ D1S213) 中最大 LOD 分数 = 3.13。如图 3 所示,11 号染色体中,11q12 ~ 11q13.5 区域 (D11S905 ~ D11S937) 中,最大 LOD 分数 = 2.78。这样得到的值,满足上述 Lander 和 Kruglyak 的提示连锁。因此,可以将卷发特性位点特定于 1 号染色体、以及 11 号染色体上,强烈暗示此区域上存在毛发形状易感性基因。

[0455] 实施例 3 候选区域中的详细绘图

[0456] 接下来,针对实施例 2 中认定连锁的 1 号染色体,以窄化连锁区域为目的,进一步追加微卫星标记,进行患病同胞对连锁分析 (详细绘图)。

[0457] 构成详细绘图的标记的微卫星是使用 The Mammalian Genotyping Service 的完整人类基因组遗传图谱 (Comprehensive human genetic maps) (<http://research.marshfieldclinic.org/genetics/GeneticResearch/compMaps.asp>) 进行探索,选择基因组中以 1 ~ 2cM 的间隔存在且杂合度高的微卫星。另外,扩增微卫星的分型用引物是基于人类基因组数据库项目 (The Genome Database project) (GDB) (<http://www.gdb.org/>) 而设计的。另外,GDB 在结束运作之后,可以根据现在的 NCBI (<http://www.ncbi.nlm.nih.gov/>) 进行检索、设计。分型用荧光引物使用 ABI 制造的产品,一部分分型用引物使用全基因组扫描、连锁分析试剂盒 (ABI PRISM Linkage Mapping Set-MD 10v2.5,ABI 制造) 所包含的产品。作为详细绘图的标记的微卫星、以及分型用荧光引物如表 6-1 以及表 6-2 中所示 (参照序列号 6 ~ 43)。

[0458]

[表 6-1]作为详细绘图的标记的卫星和分型用荧光引物

ABI	微卫星	位置 (cM)	GenBank Accession	杂合度	扩增产物 (片段) 尺寸	标识	正向引物	反向引物
MD10	AFM249zg9	DIS252	Z17138	0.82	99-119	VIC 6-FAM	AGCACATAGCAGGCACTAGC (序列号 6)	CGATTGTGCCACTACACAGT (序列号 7)
	GATA12A07	DIS534	G07791	0.83	196-212		AAAAATGAGTCCAGTAGAAGCCT (序列号 8)	AGCCAGATTACATCCCAG (序列号 9)
	AFMa297xg9	DIS2696	Z52819	0.88	159-185			
	AFM336xb1	DIS498	Z24441	0.82	183-205			
IID5	AFM207yh6	DIS2346	Z51162	0.83	89-115	VIC	TATCTTGCCCTGCACC (序列号 10)	AAGTGGGTCTCCCCAG (序列号 11)
	AFMb009zb9	DIS2721	Z53073	0.74	233-247	VIC	TTGCTGGCCAGAGTCT (序列号 12)	AGGCATCACACCTGGCTAGT (序列号 13)
	AFMa127wh9	DIS506	Z24627	0.58	123-141	VIC	GGGCTATGGCTGGAA (序列号 14)	GGCTATGCTGGGGCAA (序列号 15)
	AFMa133ye5	DIS2635	Z52215	0.86	142-159	6-FAM		CATTGCTGATGCTGGAG (序列号 17)
MD10	AFMb334xb1	DIS2771	Z53685	0.72	243-259		TCAGTTCCATAGGCTGACG (序列号 16)	
MD10	AFM297wb9	DIS484	Z24182	0.64	136-142			
MD10	AFMa057ze5	DIS2878	Z51743	0.84	169-195			
	AFMb316zb9	DIS2762	Z53529	0.81	232-250	NED	CCTTAATTGGGTGTTGGT (序列号 18)	AAAAAICTGGAAGGCATAAA (序列号 19)
	AFM063xg9	DIS196	Z16503	0.73	267-279	6-FAM		TGGATAGCTTTCACCACCT (序列号 21)
	AFMb359xf5	DIS2799	Z53881	0.87	191-209		AGCAAGACCCCTGCTCAAAA (序列号 20)	
HD5	AFM248wg5	DIS452	Z23809	0.76	119-131			
MD10	AFM157xe7	DIS218	Z16701	0.83	266-286			
HD5	AFM123yc5	DIS460	Z23379	0.84	145-159	6-FAM	ACAAGGTGACCGGAAAAGACC (序列号 22)	AGCTCTGGCAAAGTTGAAGGA (序列号 23)
	AFMc025xb9	DIS2818	Z54047	0.70	258-268			TGGGCAAGGTAGAATATGTG (序列号 25)
	AFM348tg1	DIS2848	Z51502	0.82	105-123		ATCTGGGTTCACTATTAAACAGAGT (序列号 24)	
	AFM205xg1	DIS238	Z16920	0.86	272-302			
MD10	AFMa057vb5	DIS2877	Z51735	0.72	143-157	VIC		
HD5	AFM031xd12	DIS412	Z23298	0.71	129-147			
MD10	AFM165xe9	DIS413	Z23420	0.77	246-262			
	UT-492	DIS373	L16266	0.90	283-330		GGGTGACAGAGCAAGACTC (序列号 26)	CCCTGACCTCCCTTACAGA (序列号 27)
	AFM136xa7	DIS1723	Z51003	0.83	167-181	NED	AACTGTGTCCAGCAACT (序列号 28)	TATGTGCTGTGTGTGCAT (序列号 29)
	AFMa190xcd5	DIS2655	Z52412	0.90	224-260	VIC	AGGGTCCCAAAAGACCTTC (序列号 30)	ATGGCAGCACATCCTGCTTC (序列号 31)
	AFMa224xc1	DIS2668	Z52594	0.77	233-247	VIC	AATCACTTGAACCTGGGAG (序列号 32)	ACTGACTGGCTGTTCTGAG (序列号 33)

[0459]

[表 6-2]

ABI	微卫星	位置 (cM)	GenBank Accession	杂合度	扩增产物 (片段) 尺寸	标识	正向引物	反向引物
MD10 HD5	AFM234wf6	D1S249	Z17051	0.87	155-185	6-FAM	ACTGTATTTCGGAGTTGGA (序列号 34)	CCAAGAGTTTTCTTAGCAAAATCAC (序列号 35)
	AFMa290xd1	D1S2692	Z52805	0.87	276-316			
	AFMa082wf9	D1S2891	Z51920	0.75	211-273			
HD5	AFM224xc1	D1S245	Z17011	0.83	239-257	6-FAM	CTGAGCACAGCAGTGGTCTC (序列号 36)	AAGGCTTATCAAGAGCGCAGG (序列号 37)
	AFM108va3	D1S205	Z16585	0.80	94-112			
	AFM203/b6	D1S425	Z23538	0.81	92-108			
MD10	GATA87T04	D1S2141	G07856	0.82	236-263	6-FAM	AGACTTACAGCACTGGCTGC (序列号 38)	TGCTCCTAGGAAAGGAAACA (序列号 39)
	AFM297xc1	D1S2827	Z51306	0.78	142-152		GCTTCTGGCCTCTGTCA (序列号 40)	AAATTTGCCGTGTGTGTC (序列号 41)
	AFM184yf6	D1S227	Z16806	0.71	61-75			
HD5	AFMa052zd1	D1S2871	Z51685	0.84	215-241	NED		CGAGACATTTGCATCATCA (序列号 43)
	AFM147xf8	D1S213	Z16668	0.86	104-124		TGAAGTGTGCATTTCTNTACATCA (序列号 42)	

[0460] 针对 1 号染色体进行和实施例 2 同样的患病同胞对连锁分析（详细绘图），其结果如图 4 所示。如图 4 所示，1q21.3 区域（D1S2696 ~ D1S2346）中最大 LOD 分数 = 3.60，

1q32.1 ~ 1q32.2 区域 (D1S249 ~ D1S2891) 中最大 LOD 分数 = 2.14。考虑这样得到的值, 分别满足实施例 2 中所示的 Lander 和 Kruglyak 的显著连锁、提示连锁。可以窄化 1 号染色体的卷发特性位点, 强烈暗示此区域上存在毛发形状易感性基因。

[0461] 实施例 4 病例・对照例关联分析

[0462] 为了从上述实施例 3 中认定为强烈连锁的 1 号染色体 1q21.3 区域 (D1S2696 ~ D1S2346) 中认定毛发形状易感性基因, 针对在该区域中存在的单核苷酸多态性 (SNP) 标记通过病例・对照例关联分析进行等位基因频率的比较。

[0463] 病例 (患病者: 具有卷发特性的人) 以及对照例 (对照者: 具有直发特性的人) 有必要用与认定毛发形状易感性基因的人种相同的人种来构成, 因此, 本发明中以日本人非血缘关系的具有卷发特性的人和日本人非血缘关系的具有直发特性的人作为对象。基于实施例 1 中所示的基准同样收集对象者, 从 43 名的日本人非血缘关系的具有卷发特性的人、51 名日本人非血缘关系的具有直发特性的人身上分别得到基因组 DNA。

[0464] 进行分型的 SNP 参照 dbSNP 数据库 (<http://www.ncbi.nlm.nih.gov/SNP/>) 以及 JSNP 数据库 (http://snp.ims.u-tokyo.ac.jp/index_ja.html), 代表该分析区域内的一定区域, 并且选择在日本人专门小组中次要等位基因频率为 10% 以上的 SNP, 并从该分析区域内选择 32 个 SNP。

[0465] SNP 的分型中利用 TaqMan SNP 基因型分析 (TaqMan SNP Genotyping Assays) (ABI 制造、原名 Assays-on-Demand、Assays-by-Design), 通过 TaqMan PCR 法进行。另外, 使用 Applied Biosystems 7900HT 快速实时 PCR 系统 (Applied Biosystems 7900HT Fast Real time PCR System) (ABI 制造)、以及 Applied Biosystems 7500 实时 PCR 系统 (Applied Biosystems 7500 Real time PCR System) (ABI 制造) 的机器。该方法根据附加于各个机器上的指南进行。

[0466] 得到的分型数据按照病例、对照例分别总计, 以基因型、等位基因型、显性模型、隐性模型的 4 种方法用 χ^2 检测进行显著差检测。即, 如果某些基因变异成为毛发形状变化的原因的话, 可以预料到在病例和对照例中在其等位基因频率等上存在差异。另外, 在本实施例中, 为了在比较少数的对象者中实施该关联分析则将显著性水平定为 $p < 0.05$ 。进一步, 为了提高在一部分中的检测能力, 则缓和设定显著性水平 ($p < 0.07$)。

[0467] 其结果发现, 在以下 (a) ~ (i) 中所示的 9 个 SNP 中, 在病例和对照例之间具有统计学上的显著 ($p < 0.05$) 差。

[0468] (a) 对于 SNP: rs1053590 (序列号 2 表示的碱基序列中, 碱基号 8196 所表示的单核苷酸多态性) 来说, 相比于具有卷发特性的人, 具有直发特性的人平均具有 C 等位基因的比例显著高 (表 7-1)。

[0469] (b) 对于 SNP: rs3753453 (序列号 3 表示的碱基序列中, 碱基号 5167 所表示的单核苷酸多态性) 来说, 相比于具有直发特性的人, 具有卷发特性的人平均具有 C 等位基因的比例显著高, 即使根据基因型也能在具有直发特性的人和具有卷发特性的人之间观察到显著差 (表 7-2)。

[0470] (c) 对于 SNP: rs3737859 (序列号 3 表示的碱基序列中, 碱基号 8449 所表示的单核苷酸多态性) 来说, 相比于具有直发特性的人, 具有卷发特性的人平均具有 G 等位基因的比例显著高, 即使根据基因型也能在具有直发特性的人和具有卷发特性的人之间观察到显

著差（表 7-3）。

[0471] (d) 对于 SNP :rs3904414(序列号 3 表示的碱基序列中,碱基号 17598 所表示的单核苷酸多态性) 来说,相比于具有直发特性的人,具有卷发特性的人平均具有 A 等位基因的比例显著高,即使根据基因型也能在具有直发特性的人和具有卷发特性的人之间观察到显著差（表 7-4）。

[0472] (e) 对于 SNP :rs3908717(序列号 3 表示的碱基序列中,碱基号 20891 所表示的单核苷酸多态性) 来说,相比于具有直发特性的人,具有卷发特性的人平均具有 G 等位基因的比例显著高,即使根据基因型也能在具有直发特性的人和具有卷发特性的人之间观察到显著差（表 7-5）。

[0473] (f) 对于 SNP :rs3737861(序列号 4 表示的碱基序列中,碱基号 6354 所表示的单核苷酸多态性) 来说,相比于具有直发特性的人,具有卷发特性的人平均具有 A 等位基因的比例显著高,即使根据等位基因型也能在具有直发特性的人和具有卷发特性的人之间观察到显著差（表 7-6）。

[0474] (g) 对于 SNP :rs4523473(序列号 5 表示的碱基序列中,碱基号 759 所表示的单核苷酸多态性) 来说,相比于具有直发特性的人,具有卷发特性的人平均具有 T 等位基因的比例显著高,即使根据等位基因型也能在具有直发特性的人和具有卷发特性的人之间观察到显著差（表 7-7）。

[0475] (h) 对于 SNP :rs2229496(序列号 5 表示的碱基序列中,碱基号 1939 所表示的单核苷酸多态性) 来说,相比于具有直发特性的人,具有卷发特性的人平均具有 A 等位基因的比例显著高,即使根据等位基因型也能在具有直发特性的人和具有卷发特性的人之间观察到显著差（表 7-8）。

[0476] (i) 对于 SNP :rs913996(序列号 5 表示的碱基序列中,碱基号 3440 所表示的单核苷酸多态性) 来说,相比于具有直发特性的人,具有卷发特性的人平均具有 G 等位基因的比例显著高,即使根据等位基因型也能在具有直发特性的人和具有卷发特性的人之间观察到显著差（表 7-9）。

[0477] 另外,可知以下 (j) 所示的 1 的 SNP 中在病例和对照例之间也有差别。

[0478] (j) 对于 SNP :rs2305814(序列号 1 表示的碱基序列中,碱基号 7121 所表示的单核苷酸多态性) 来说,相比于具有直发特性的人,具有卷发特性的人平均具有 T 等位基因的比例高 ($p = 0.066$) (表 7-10)。对于这 10 个 SNP 任意一个都满足 Hardy-Weinberg 平衡状态。因此,判断这 10 个 SNP 为毛发形状易感性 SNP,认为与毛发形状相关联。

[0479] [表 7-1]SNP :rs1053590 中的关联分析

[0480]

SNP: rs1053590	等位基因型		基因型		
	C	T	CC	CT	TT
卷发特性	70.3%	29.7%	43.2%	54.1%	2.7%
直发特性 (对照)	80.4%	19.6%	64.7%	31.4%	3.9%
p 值 (χ^2 检测)	等位基因型		0.120		
	基因型		0.102		
	CCvsCT, TT		0.045		

[0481] [表 7-2] SNP :rs3753453 中的关联分析

[0482]

SNP: rs3753453	等位基因型		基因型		
	T	C	TT	TC	CC
卷发特性	66.3%	33.7%	46.5%	39.5%	14.0%
直发特性 (对照)	76.5%	23.5%	52.9%	47.1%	0.0%
p 值 (χ^2 检测)	等位基因型		0.122		
	基因型		0.022		
	TT, TCvs CC		0.006		

[0483] [表 7-3] SNP :rs3737859 中的关联分析

[0484]

SNP: rs3737859	等位基因型		基因型		
	G	T	GG	GT	TT
卷发特性	33.7%	66.3%	14.0%	39.5%	46.5%
直发特性 (对照)	23.5%	76.5%	0.0%	47.1%	52.9%
p 值 (χ^2 检测)	等位基因型		0.122		
	基因型		0.022		
	GGvsGT, TT		0.006		

[0485] [表 7-4] SNP :rs3904414 中的关联分析

[0486]

SNP: rs3904414	等位基因型		基因型		
	G	A	GG	GA	AA
卷发特性	67.1%	32.9%	46.3%	41.5%	12.2%
直发特性 (对照)	77.0%	23.0%	50.4%	46.0%	0.0%
p 值 (χ^2 检测)	等位基因型		0.136		
	基因型		0.039		
	GG, GAvsAA		0.011		

[0487] [表 7-5] SNP :rs3908717 中的关联分析

[0488]

SNP: rs3908717	等位基因型		基因型		
	C	G	CC	CG	GG
卷发特性	67.9%	32.1%	47.6%	40.5%	11.9%
直发特性 (对照)	76.5%	23.5%	52.9%	47.1%	0.0%
p 值 (χ^2 检测)	等位基因型		0.190		
	基因型		0.040		
	CC, CGvsGG		0.011		

[0489] [表 7-6] SNP :rs3737861 中的关联分析

[0490]

SNP: rs3737861	等位基因型		基因型		
	A	C	AA	AC	CC
卷发特性	48.8%	51.2%	31.0%	35.7%	33.3%
直发特性 (对照)	33.3%	66.7%	13.7%	39.2%	47.1%
p 值 (χ^2 检测)	等位基因型		0.032		
	基因型		0.116		
	AAvsAC, CC		0.044		

[0491] [表 7-7] SNP :rs4523473 中的关联分析

[0492]

SNP: rs4523473	等位基因型		基因型		
	T	C	TT	TC	CC
卷发特性	47.7%	52.3%	30.2%	34.9%	34.9%
直发特性 (对照)	32.4%	67.6%	11.8%	41.2%	47.1%
p 值 (χ^2 检测)	等位基因型		0.032		
	基因型		0.082		
	TTvsTC, CC		0.026		

[0493] [表 7-8] SNP :rs2229496 中的关联分析

[0494]

SNP: rs2229496	等位基因型		基因型		
	G	A	GG	GA	AA
卷发特性	53.6%	46.4%	35.7%	35.7%	26.8%
直发特性 (对照)	67.6%	32.4%	47.1%	41.2%	11.8%
p 值 (χ^2 检测)	等位基因型		0.050		
	基因型		0.120		
	GG, GAvsAA		0.041		

[0495] [表 7-9] SNP :rs913996 中的关联分析

[0496]

SNP: rs913996	等位基因型		基因型		
	G	A	GG	GA	AA
卷发特性	47.7%	52.3%	30.2%	34.9%	34.9%
直发特性 (对照)	32.4%	67.6%	11.8%	41.2%	47.1%
p 值 (χ^2 检测)	等位基因型		0.032		
	基因型		0.082		
	GGvsGA, AA		0.026		

[0497] [表 7-10] SNP :rs2305814 中的关联分析

[0498]

SNP: rs2305814	等位基因型		基因型		
	C	T	CC	CT	TT
卷发特性	55.4%	44.6%	35.1%	40.5%	24.3%
直发特性 (对照)	65.7%	34.3%	41.2%	49.0%	9.8%
p 值 (χ^2 检测)	等位基因型		0.167		
	基因型		0.184		
	CC, CTvsTT		0.066		

[0499] 实施例 5 单倍体型分析

[0500] 在实施例 4 中分析的结果,发现了 10 种毛发形状易感性 SNP,但是为了确认该 SNP 周边区域中存在的多态性、特别是没有进行分型的多态性也有相关,并认定毛发形状易感性基因而进行了单倍体型分析。

[0501] 分析使用 Haploview 4.1 软件 (Barrett, JC. et al., Bioinformatics, 21(2), 263-265, 2005),通过 EM 算法算出连锁不平衡系数 D' (pair-wise LD coefficient) 来分析。使用国际人类基因单倍体型图计划数据库的 HapMap PHASE 数据 (HapMap Data Rel 21/PhaseII Jul06, on NCBI Build 35 assembly, dbSNP b125),对该 SNP 和其周边区域中存在的 SNP 进行连锁不平衡分析。另外,分析专门小组为 JPT+CHB (日本,东京的日本人,以及中国北京的汉族中国人)。

[0502] 单倍域的推测方法使用置信区间 (Gabriel, SB. et al., Science, 296(5576), p2225-2229, 2002)。即,认为特定的单倍域是基本没有发现历史的重组的基因组范围,并且在其区域内存在强烈的连锁不平衡。通常在连锁不平衡系数 D' 的 95% 置信区间的上限小于 0.9 的情况下,判断此区域为有历史的重组证据的区域,另一方面,在 D' 的 95% 置信区间的上限超过 0.98,其下限高于 0.7 的情况下,判断此区域为存在强烈的连锁不平衡的区域。

[0503] 由此结果,发现含有以下所示的 10 个毛发形状易感性 SNP 的 (1) ~ (5) 的单倍域。

[0504] (1) 含有 SNP :rs2305814,且从 SNP :rs3754211 到 SNP :rs771205 的由序列号 1 表示的碱基序列所表示的 23, 252bp 构成的单倍域 (图 5)。此单倍域为含有 ANXA9 基因、以及 FAM63A 基因的区域。由此结果认定 ANXA9 基因、以及 FAM63A 基因为毛发形状易感性基因。

[0505] (2) 含有 SNP :rs1053590,且从 SNP :rs1885531 到 SNP :rs4845791 的由序列号 2 表示的碱基序列所表示的 56, 552bp 构成的单倍域 (图 6)。此单倍域为含有 LCE5A 基因、以及 CRCT1 基因的区域。由此结果认定 LCE5A 基因、以及 CRCT1 基因为毛发形状易感性基因。

[0506] (3) 含有 SNP :rs3753453、SNP :rs3737859、SNP :rs3904414,以及 SNP :rs3908717,且从 SNP :rs2014369 到 SNP :rs11205079 的由序列号 3 表示的碱基序列所表示的 23, 382bp 构成的单倍域 (图 7)。此单倍域为含有 LCE2B 基因、以及 LCE2A 基因的区域。由此结果认定 LCE2B 基因、以及 LCE2A 基因为毛发形状易感性基因。

[0507] (4) 含有 SNP :rs3737861,且从 SNP :rs16834715 到 SNP :rs4845492 的由序列号 4 表示的碱基序列所表示的 8, 818bp 构成的单倍域 (图 8)。此单倍域为含有 SMCP 基因的区域。由此结果认定 SMCP 基因为毛发形状易感性基因。

[0508] (5) 含有 SNP :rs4523473、SNP :rs2229496、以及 SNP :rs913996,且从 SNP :rs1854779 到 SNP :rs913996 的由序列号 5 表示的碱基序列所表示的 3, 440bp 构成的单倍域 (图 9)。此单倍域为含有 IVL 基因的区域。由此结果认定 IVL 基因为毛发形状易感性基因。

[0509] 实施例 6 毛发形状易感性 SNP 标记的认定

[0510] 在实施例 5 的单倍体型分析中发现单倍域的同时,同样使用 Haploview 4.1 软件 (Barrett, JC. et al., Bioinformatics, 21(2), 263-265, 2005),从各个单倍域中提取单倍体型。通过比较提取的单倍体型,即 SNP 标记群的各个碱基组合来认定和毛发形状易感性 SNP 标记位点连锁的 SNP 位点。这样认定的 SNP 位点可以作为新的毛发形状易感性 SNP 标记。

[0511] 由此结果,实施例 4 中所示的 (1) ~ (5) 的单倍域中,发现以下所示的各个新的毛发形状易感性 SNP 标记。

[0512] (1) 由序列号 1 表示的碱基序列所表示的 23, 252bp 构成的单倍域:此单倍域中存在着 9 种主要的单倍体型(表 8)。作为与毛发形状易感性 SNP 标记的 SNP :rs2305814 连锁的 SNP 位点,认定以下所示的 6 个新的毛发形状易感性 SNP 标记。

[0513] SNP :rs3754211(序列号 1 所表示的碱基序列中,碱基号 1 所表示的单核苷酸多态性)、SNP :rs3754210(同 2405 所示的单核苷酸多态性)、SNP :rs16832604(同 5874 所示的单核苷酸多态性)、SNP :rs7532008(同 8494 所示的单核苷酸多态性)、SNP :rs1673160(同 18980 所示的单核苷酸多态性)、以及 SNP :rs771205(同 23252 所示的单核苷酸多态性)。

[0514] [表 8]

SNP 标记	序列号 1 所表示的碱基序列的碱基序号	单倍体型									毛发形状易感性 SNP
		1	2	3	4	5	6	7	8	9	
rs3754211	1	A	G	G	A	G	A	A	A	A	○
rs3754210	2405	G	G	G	T	G	G	G	G	T	○
rs16832604	5874	A	A	A	G	A	A	A	A	G	○
rs2305814	7121	C	T	T	C	C	C	C	T	C	○(实施例 4)
rs7532008	8494	C	A	A	C	C	A	C	A	C	○
rs12033744	9072	G	A	G	G	G	G	G	G	G	
rs11204754	9398	C	C	T	C	C	T	C	T	C	
rs1722784	10013	A	A	G	A	A	G	A	G	A	
rs11587718	11422	C	T	T	C	C	T	C	T	C	
rs1673160	18980	T	A	A	T	T	A	A	A	T	○
rs771204	22306	A	A	A	G	A	A	A	A	G	
rs771205	23252	C	C	C	T	C	C	C	C	C	○

[0516] (2) 由序列号 2 表示的碱基序列所表示的 56, 552bp 构成的单倍域:此单倍域中存在着 10 种主要的单倍体型(表 9-1 ~ 表 9-3)。作为与毛发形状易感性 SNP 标记的 SNP :rs1053590 连锁的 SNP 位点,认定以下所示的 39 个新的毛发形状易感性 SNP 标记。

[0517] SNP :rs11581947(序列号 2 所表示的碱基序列中,碱基号 2355 所表示的单核苷酸多态性)、SNP :rs6658925(同 2569 所表示的单核苷酸多态性)、SNP :rs2105117(同 3897 所表示的单核苷酸多态性)、SNP :rs548252(同 9510 所表示的单核苷酸多态性)、SNP :rs493133(同 13643 所表示的单核苷酸多态性)、SNP :rs1970283(同 15387 所表示的单核苷酸多态性)、SNP :rs1001834(同 15708 所表示的单核苷酸多态性)、SNP :rs11205018(同 16017 所表示的单核苷酸多态性)、SNP :rs545418(同 17106 所表示的单核苷酸多态性)、SNP :rs12116609(同 17453 所表示的单核苷酸多态性)、SNP :rs526099(同 17579 所表示的单核苷酸多态性)、SNP :rs525960(同 17634 所表示的单核苷酸多态性)、SNP :rs4845443(同 26924 所表示的单核苷酸多态性)、SNP :rs569032(同 28383 所表示的单核苷酸多态性)、SNP :rs528427(同 31275 所表示的单核苷酸多态性)、SNP :rs478926(同 31301 所表示的单核苷酸多态性)、SNP :rs1337338(同 31653 所表示的单核苷酸多态性)、SNP :rs6587681(同 31903 所表示的单核苷酸多态性)、SNP :rs1856120(同 32209 所表示的单核苷酸多态性)、SNP :rs474086(同 33199 所表示的单核苷酸多态性)、SNP :rs578382(同 33822 所表示的单核苷酸多态性)、SNP :rs549044(同 34100 所表示的单核苷酸多态性)、SNP :rs1123567(同 35791 所表示的单核苷酸多态性)、SNP :rs1538083(同 36884 所表示的单核苷酸多态性)、SNP :rs1538082(同 37072 所表示的单核苷酸多态性)、

SNP :rs7532535 (同 37365 所表示的单核苷酸多态性)、SNP :rs7518654 (同 37613 所表示的单核苷酸多态性)、SNP :rs533917 (同 38062 所表示的单核苷酸多态性)、SNP :rs564107 (同 39063 所表示的单核苷酸多态性)、SNP :rs7530609 (同 46580 所表示的单核苷酸多态性)、SNP :rs4240885 (同 49618 所表示的单核苷酸多态性)、SNP :rs4240886 (同 50164 所表示的单核苷酸多态性)、SNP :rs4240887 (同 50278 所表示的单核苷酸多态性)、SNP :rs6687126 (同 50662 所表示的单核苷酸多态性)、SNP :rs6674451 (同 50822 所表示的单核苷酸多态性)、SNP :rs7550769 (同 50981 所表示的单核苷酸多态性)、SNP :rs7529157 (同 51133 所表示的单核苷酸多态性)、SNP :rs1988805 (同 51263 所表示的单核苷酸多态性)、SNP :rs7529441 (同 51397 所表示的单核苷酸多态性)。

[0518] [表 9-1]

[0519]

SNP标记	序列号2所表示的碱基序列的碱基序号	单倍体型										毛发形状易感性SNP
		1	2	3	4	5	6	7	8	9	10	
rs1885531	1	G	T	G	G	T	G	G	T	G	T	
rs1415432	422	A	G	A	A	G	A	G	A	G	G	
rs7521430	1464	G	G	G	T	G	G	T	G	G	G	
rs11581947	2355	A	G	A	G	G	A	G	G	A	G	○
rs6658925	2569	G	A	G	A	A	G	A	A	G	A	○
rs2105117	3897	A	G	A	G	G	A	G	G	A	G	○
rs2282298	4013											
rs4845441	5189	T	C	T	C	C	T	C	C	T	T	
rs3753450	5787	T	T	G	T	T	T	T	T	T	T	
rs2297760	6802	T	T	T	C	T	T	C	T	T	T	
rs16834168	7914	G	G	G	A	G	G	G	G	G	G	
rs1053590	8196	C	T	C	T	T	C	T	T	C	T	○(实施例4)
rs488393	9316	A	G	A	G	G	A	G	G	A	A	
rs548252	9510	T	C	T	C	C	T	C	C	T	C	○
rs4845783	12327	A	G	A	A	G	A	A	G	A	A	
rs10888491	12390	G	A	G	A	A	G	A	A	A	G	
rs499697	12922	A	G	A	A	G	A	A	G	A	G	
rs493133	13643	C	G	C	G	G	C	G	G	C	G	○
rs1970283	15387	C	G	C	G	G	C	G	G	C	G	○
rs1001834	15708	A	C	A	C	C	A	C	C	A	C	○
rs1538084	15780	C	A	C	C	A	C	C	C	C	A	
rs11205018	16017	G	T	G	T	T	G	T	T	G	T	○
rs545418	17106	T	C	T	C	C	T	C	C	T	C	○
rs12116609	17453	T	C	T	C	C	T	C	C	T	C	○
rs526099	17579	C	T	C	T	T	C	T	T	C	T	○
rs525960	17634	A	T	A	T	T	A	T	T	A	T	○
rs4845443	26924	A	G	A	G	G	A	G	G	A	G	○
rs569032	28383	T	C	T	C	C	T	C	C	T	C	○
rs16834184	29986	G	G	G	G	G	C	G	G	G	G	
rs6662637	30116	G	G	A	G	G	G	G	G	G	G	
rs528427	31275	C	G	C	G	G	C	G	G	C	G	○
rs478926	31301	T	G	T	G	G	T	G	G	T	G	○
rs1337338	31653	A	G	A	G	G	A	G	G	A	G	○
rs6587681	31903	T	C	T	C	C	T	C	C	T	C	○
rs1856120	32209	A	G	A	G	G	A	G	G	A	G	○
rs474086	33199	T	C	T	C	C	T	C	C	T	C	○
rs578382	33822	A	G	A	G	G	A	G	G	A	G	○
rs549044	34100	T	C	T	C	C	T	C	C	T	C	○
rs1123567	35791	A	G	A	G	G	A	G	G	A	G	○
rs1538083	36884	G	A	G	A	A	G	A	A	G	A	○
rs1538082	37072	G	A	G	A	A	G	A	A	G	A	○
rs7532535	37365	A	T	A	T	T	A	T	T	A	T	○
rs7518654	37613	G	C	G	C	C	G	C	C	G	C	○
rs533917	38062	G	A	G	A	A	G	A	A	G	A	○
rs564107	39063	T	C	T	C	C	T	C	C	T	C	○
rs7530609	46580	A	C	A	C	C	A	C	C	A	C	○
rs6587688	49390	C	T	C	C	T	C	C	T	C	T	

[0520] [表 9-2]

[0521]

SNP标记	序列号2所表示的碱基序列的碱基序号	单倍体型										毛发形状易感性 SNP
		1	2	3	4	5	6	7	8	9	10	
rs4240885	49618	C	G	C	G	G	C	G	G	C	G	○
rs4240886	50164	A	T	A	T	T	A	T	T	A	T	○
rs4240887	50278	G	A	G	A	A	G	A	A	G	A	○
rs6687126	50662	G	T	G	T	T	G	T	T	G	T	○
rs6674451	50822	T	C	T	C	C	T	C	C	T	C	○
rs7550769	50981	A	G	A	G	G	A	G	G	A	G	○
rs7529157	51133	A	C	A	C	C	A	C	C	A	C	○
rs1988805	51263	G	T	G	T	T	G	T	T	G	T	○
rs7529441	51397	T	C	T	C	C	T	C	C	T	C	○
rs4845790	56528	T	T	T	T	C	T	T	T	T	T	
rs4845791	56552	A	A	A	A	G	A	A	A	A	A	
rs1885531	1	G	T	G	G	T	G	G	T	G	T	
rs1415432	422	A	G	A	A	G	A	A	G	A	G	
rs7521430	1464	G	G	G	T	G	G	T	G	G	G	
rs11581947	2355	A	G	A	G	G	A	G	G	A	G	○
rs6658925	2569	G	A	G	A	A	G	A	A	G	A	○
rs2105117	3897	A	G	A	G	G	A	G	G	A	G	○
rs2282298	4013											
rs4845441	5189	T	C	T	C	C	T	C	C	T	T	
rs3753450	5787	T	T	G	T	T	T	T	T	T	T	
rs2297760	6802	T	T	T	C	T	T	C	T	T	T	
rs16834168	7914	G	G	G	A	G	G	G	G	G	G	
rs1053590	8196	C	T	C	T	T	C	T	T	C	T	○(实施例4)
rs488393	9316	A	G	A	G	G	A	G	G	A	A	
rs548252	9510	T	C	T	C	C	T	C	C	T	C	○
rs4845783	12327	A	G	A	A	G	A	A	G	A	A	
rs10888491	12390	G	A	G	A	A	G	A	A	A	G	
rs499697	12922	A	G	A	A	G	A	A	G	A	G	
rs493133	13643	C	G	C	G	G	C	G	G	C	G	○
rs1970283	15387	C	G	C	G	G	C	G	G	C	G	○
rs1001834	15708	A	C	A	C	C	A	C	C	A	C	○
rs1538084	15780	C	A	C	C	A	C	C	C	C	A	
rs11205018	16017	G	T	G	T	T	G	T	T	G	T	○
rs545418	17106	T	C	T	C	C	T	C	C	T	C	○
rs12116609	17453	T	C	T	C	C	T	C	C	T	C	○
rs526099	17579	C	T	C	T	T	C	T	T	C	T	○
rs525960	17634	A	T	A	T	T	A	T	T	A	T	○
rs4845443	26924	A	G	A	G	G	A	G	G	A	G	○
rs569032	28383	T	C	T	C	C	T	C	C	T	C	○
rs16834184	29986	G	G	G	G	G	C	G	G	G	G	
rs6662637	30116	G	G	A	G	G	G	G	G	G	G	
rs528427	31275	C	G	C	G	G	C	G	G	C	G	○
rs478926	31301	T	G	T	G	G	T	G	G	T	G	○
rs1337338	31653	A	G	A	G	G	A	G	G	A	G	○
rs6587681	31903	T	C	T	C	C	T	C	C	T	C	○
rs1856120	32209	A	G	A	G	G	A	G	G	A	G	○
rs474086	33199	T	C	T	C	C	T	C	C	T	C	○
rs578382	33822	A	G	A	G	G	A	G	G	A	G	○

[0522] [表 9-3]

[0523]

SNP标记	序列号2所表示的碱基序列的碱基序号	单倍体型										毛发形状易感性 SNP
		1	2	3	4	5	6	7	8	9	10	
rs549044	34100	T	C	T	C	C	T	C	C	T	C	○
rs1123567	35791	A	G	A	G	G	A	G	G	A	G	○
rs1538083	36884	G	A	G	A	A	G	A	A	G	A	○
rs1538082	37072	G	A	G	A	A	G	A	A	G	A	○
rs7532535	37365	A	T	A	T	T	A	T	T	A	T	○
rs7518654	37613	G	C	G	C	C	G	C	C	G	C	○
rs533917	38062	G	A	G	A	A	G	A	A	G	A	○
rs564107	39063	T	C	T	C	C	T	C	C	T	C	○
rs7530609	46580	A	C	A	C	C	A	C	C	A	C	○
rs6587688	49390	C	T	C	C	T	C	C	T	C	T	○
rs4240885	49618	C	G	C	G	G	C	G	G	C	G	○
rs4240886	50164	A	T	A	T	T	A	T	T	A	T	○
rs4240887	50278	G	A	G	A	A	G	A	A	G	A	○
rs6687126	50662	G	T	G	T	T	G	T	T	G	T	○
rs6674451	50822	T	C	T	C	C	T	C	C	T	C	○
rs7550769	50981	A	G	A	G	G	A	G	G	A	G	○
rs7529157	51133	A	C	A	C	C	A	C	C	A	C	○
rs1988805	51263	G	T	G	T	T	G	T	C	G	T	○
rs7529441	51397	T	C	T	C	C	T	C	C	T	C	○
rs4845790	56528	T	T	T	T	C	T	T	T	T	T	○
rs4845791	56552	A	A	A	A	G	A	A	A	A	A	○

[0524] (3) 由序列号 3 表示的碱基序列所表示的 23,382bp 构成的单倍域:此单倍域中存在着 9 种主要的单倍体型(表 10)。作为与毛发形状易感性 SNP 标记的 SNP:rs3753453、SNP:rs3737859、SNP:rs3904414、以及 SNP:rs3908717 连锁的 SNP 位点,认定以下所示的 4 个新的毛发形状易感性 SNP 标记。

[0525] SNP:rs11205072(序列号 3 所表示的碱基序列中,碱基号 2509 所表示的单核苷酸多态性)、SNP:rs12074783(同 18481 所表示的单核苷酸多态性)、SNP:rs3904415(同 21734 所表示的单核苷酸多态性)、SNP:rs11205079(同 23382 所表示的单核苷酸多态性)。

[0526] [表 10]

[0527]

SNP标记	序列号3所表示的碱基 序列的碱基序号	单倍体型									毛发形状易感性SNP
		1	2	3	4	5	6	7	8	9	
rs2014369	1	T	T	C	C	C	C	C	C	C	○
rs2014368	7	T	T	C	C	C	C	C	C	C	
rs12097462	1115	A	A	C	C	C	C	C	C	C	
rs11576287	1991	A	A	A	A	T	A	A	A	T	
rs10788844	2303	A	A	C	C	C	C	C	C	C	
rs1925663	2484	A	A	T	T	T	T	T	T	T	
rs11205072	2509	G	G	A	A	G	A	G	A	A	
rs11205073	3363	A	A	C	C	C	C	C	C	C	
rs12754873	4275	C	C	C	C	T	C	T	C	C	
rs3753453	5167	T	T	C	C	T	C	T	C	C	○(实施例4)
rs3753452	5560	C	T	C	C	C	C	C	C	C	
rs3737859	8449	T	T	G	G	T	G	T	G	G	○(实施例4)
rs10888508	9159	C	C	A	A	C	C	C	C	A	
rs1332509	9724	A	A	T	T	T	T	T	T	T	○(实施例4)
rs12040104	15127	C	C	C	A	C	C	C	A	C	
rs3904413	17476	T	T	C	C	C	C	C	C	C	
rs3904414	17598	G	G	A	A	G	A	G	A	A	
rs11205078	18437	G	G	A	A	A	A	A	A	A	○
rs12074783	18481	T	T	C	C	T	C	T	C	C	
rs4845317	19319	A	A	G	G	G	G	G	G	G	○(实施例4)
rs3908717	20891	C	C	G	G	C	G	C	G	G	
rs17661905	20983	T	T	T	T	C	T	C	T	T	○
rs16834343	21327	A	A	G	A	A	G	A	A	G	
rs3904415	21734	C	C	T	T	C	T	C	T	T	
rs11581506	22237	A	A	G	G	G	G	G	G	G	
rs11587581	23044	C	C	T	T	T	T	T	T	T	○
rs7543194	23198	T	T	C	C	C	C	C	C	C	
rs11205079	23382	A	A	T	T	A	T	A	T	T	○

[0528] (4) 由序列号 4 表示的碱基序列表示的 8,818bp 构成的单倍域:此单倍域中存在着 3 种主要的单倍体型(表 11)。作为与毛发形状易感性 SNP 标记的 SNP:rs3737861 连锁的 SNP 位点,认定以下所示的 6 个新的毛发形状易感性 SNP 标记。

[0529] SNP:rs16834715(序列号 4 所表示的碱基序列中,碱基号 1 所表示的单核苷酸多态性)、SNP:rs12022319(同 3308 所表示的单核苷酸多态性)、SNP:rs4845490(同 4715 所表示的单核苷酸多态性)、SNP:rs4845491(同 4895 所表示的单核苷酸多态性)、SNP:rs16834728(同 8553 所表示的单核苷酸多态性)、SNP:rs4845492(同 8818 所表示的单核苷酸多态性)。

[0530] [表 11]

SNP标记	序列号4所表示的碱基 序列的碱基序号	单倍体型			毛发形状易感性SNP
		1	2	3	
rs16834715	1	T	C	C	○
rs12022870	1816	A	G	G	
rs12048655	1967	G	T	T	
rs12022319	3308	C	T	C	○
rs2339383	4039	A	C	C	
rs4845490	4715	G	A	G	○
rs4845491	4985	C	T	C	○
rs12021601	5738	T	G	G	
rs3737862	6322	A	G	G	
rs3737861	6354	C	A	C	○(实施例4)
rs16834728	8553	C	T	C	○
rs16834734	8694	C	T	T	
rs4845492	8818	C	G	C	○

[0532] (5) 由序列号 5 表示的碱基序列表示的 3,440bp 构成的单倍域:此单倍域中存在着 3 种主要的单倍体型(表 12)。作为与毛发形状易感性 SNP 标记的 SNP:rs4523473、SNP:rs2229496 以及 SNP:rs913996 连锁的 SNP 位点,认定以下所示的 5 个新的毛发形状易感性 SNP 标记。

[0533] SNP:rs1854779(序列号 5 所表示的碱基序列中,碱基号 1 所表示的单核苷酸多态性)、SNP:rs16834751(同 540 所表示的单核苷酸多态性)、SNP:rs11205131(同 1007 所表示的单核苷酸多态性)、SNP:rs7528862(同 1018 所表示的单核苷酸多态性)、SNP:rs7517189(同 1075 所表示的单核苷酸多态性)。

[0534] [表 12]

SNP标记	序列号5所表示的碱基序列的碱基序号	单倍体型			毛发形状易感性SNP
		1	2	3	
rs1854779	1	C	T	C	○
rs16834751	540	C	A	C	○
rs4523473	759	C	T	C	○ (实施例4)
rs11205131	1007	A	G	G	○
rs7528862	1018	G	A	G	○
rs7517189	1075	G	C	G	○
rs2229496	1939	G	A	G	○ (实施例4)
rs913996	3440	A	G	A	○ (实施例4)

[0536] 实施例 7 卷发者、以及直发者的头发发根的基因表达分析

[0537] 根据实施例 1 的分类,收集 10 名卷发者和 10 名直发者,进行各被检测者的头发发根的毛发形状易感性基因的表达分析。另外,对于从被检测者身上收集的检测体,事前得到了伦理委员会的承认,在此之上,知情同意实施担当者通过书面对对象者说明了研究内容,并得到了书面的同意。

[0538] 从被检测者的整个头部均匀地拔下每个人约 60 根头发,仅仅将从发根部的形态判断为处于成长期的头发发根部,回收到装满冰冻的 PBS(Invitrogen 制造)的浅底

盘中。在实体显微镜下用镊子和注射针的刃从发根部尽量除去外毛根鞘和内毛根鞘,分离·调制仅仅发干的发根(发干角化区域)。将发干角化区域放入装有 0.5mL RNA 提取液 ISOGEN(NIPPON GENE CO., LTD 制造)的 1.5mL 容积的管中,用迷你无绳研磨机(Mini Cordless Grinder)和均质化用研棒充分破碎组织。加入 0.5mL 的 ISOGEN、200 μ L 氯仿,然后用漩涡式搅拌机(vortex mixer)充分搅拌之后,通过小型微量离心分离机离心分离(15000rpm、15 分钟),回收含有 RNA 的水相约 500 μ L。向回收的溶液中添加 50 μ L 的 3M 醋酸钠和 1 μ L 乙醇沉淀核酸载体(ethachinmate)(NIPPON GENECO., LTD 制造)来充分搅拌。进一步,加入 1mL 的异丙醇搅拌,通过小型微量离心分离机进行离心分离(15000rpm, 20 分钟),沉淀出总 RNA(total RNA)。舍去上清液之后,加入 75% 的乙醇,再次通过小型微量离心分离机进行离心分离(15000rpm, 10 分钟)。舍去上清液,使沉淀风干之后,溶解于 20 μ L 的无核酸酶水(Nuclease-free Water)(Invitrogen 制造)。使用其一部分,通过吸光度计(GeneQuant :Pharmacia 制造、或者 NonoDrop :Technologies 制造)或者 RiboGreen RNA 定量试剂盒(RiboGreen RNA Reagent and Kit)(Invitrogen 制造)测定 RNA 浓度。从得到的 1 μ g 总 RNA 中,使用 QuantiTect 逆转录试剂盒(QuantiTect Reverse Transcription Kit)(QIAGEN 制造),按照附带的规程来合成 cDNA,用于由 PCR 进行的基因表达量的定量。

[0539] 基因表达量的定量使用 Applied Biosystems 株式会社(ABI)制造的 TaqMan(注册商标)基因表达分析(TaqMan Gene Expression Assays)进行进行。根据附带的规程,混合合成的 cDNA、对于检测·定量化的基因特异的引物和探针组、以及实时 PCR(Real time PCR)试剂等(ABI 制造),用 Applied Biosystems 7500 实时 PCR 系统(Applied Biosystems 7500 Real time PCR System)(ABI 制造)来扩增检测·定量化的基因的片段。此时,使用来自自己知的标准发干角化区域样品的 cDNA 同样进行实时 PCR 来制作标准曲线进行基因表达量的标准化。另外,作为内部标准使用 GAPDH 基因的同时,还采用认定在样品的发干角化区域中均匀表达的 KRT31 基因、以及 KRT85 基因作为内部标准,进行检测·定量的基因表达量的标准化。

[0540] 作为用于检测·定量 ANXA9 基因的表达量的特异的引物和探针组,使用分析序号为 Hs00185977m1 的 TaqMan 基因表达分析(ABI 制造)。

[0541] 作为检测·定量 LCE2B 基因的表达量的特异的引物和探针组,使用分析序号为 Hs00863535g1 的 TaqMan 基因表达分析(ABI 制造)。

[0542] 作为检测·定量 LCE2A 基因的表达量的特异的引物和探针组,使用分析序号为 Hs00820278sH 的 TaqMan 基因表达分析(ABI 制造)。

[0543] 作为检测·定量 IVL 基因的表达量的特异的引物和探针试剂盒,使用分析序号为 Hs00846307s1 的 TaqMan 基因表达分析(ABI 制造)。

[0544] 作为检测·定量 CRCT1 基因的表达量的特异的引物和探针组,使用分析序号为 Hs00219416m1 的 TaqMan 基因表达分析(ABI 制造)。

[0545] 卷发群和直发群的头发发根上毛发形状易感性基因的表达量如图 10A ~ E 所示。由图 10 中所示的结果,发现和直发群相比较,在卷发群中的 ANXA9 基因、LCE2B 基因、LCE2A 以及 CRCT1 基因的表达量降低,而 IVL 基因的表达量增加。因此,判别 ANXA9 基因、LCE2B 基因、LCE2A 基因、IVL 基因、以及 CRCT1 基因是作为评价毛发形状的指标的毛发形状易感性基因,并且测定在发根部的这些基因的表达量是有意义的。

[0546] 实施例 8 调节毛发形状易感性基因的表达量的物质的筛选

[0547] 在筛选中使用正常人新生儿包皮表皮角化细胞(KK-4009, Kurabo 制造)。融解冷冻状态的正常人新生儿包皮表皮角化细胞之后, 将细胞以 2500 个细胞/cm²的密度播种于 75cm²烧瓶或者 25cm²烧瓶中, 在添加了补给物的人角化细胞培养用无血清培养基(Defined Keratinocyte-SFM、Invitrogen 制造), 在 37℃、CO₂浓度为 5%的条件下培养。在细胞达到分汇合的状态下的时刻进行传代, 以 2500 个细胞/cm²的细胞密度播种于 6 孔细胞培养板中。在细胞达到分汇合的状态的时候(第 0 天)将培养基和不含补给物的人角化细胞培养用无血清培养基交换, 将第 1 天的细胞作为筛选用细胞使用。

[0548] 向上述调制的筛选用细胞的培养基(不含补给物的人角化细胞培养用无血清培养基)中, 添加植物提取物至最终浓度达到 0.1%、以及 1%, 在 37℃、CO₂浓度为 5%的条件下培养细胞 24 小时。另外, 作为对照, 添加 50%的乙醇(对照例)同样至最终浓度为 0.1%、以及 1%来培养细胞。

[0549] 培养结束之后(第 2 天), 在吸取除去培养基的 PBS(Invitrogen 制造)中清洗 2 次细胞之后, 在每个孔中添加 1mL 的 ISOGEN(NIPPON GENE CO., LTD 制造)。在用移液管充分溶解混合细胞之后, 将溶液回收到 1.5mL 容量的管中。用和实施例 7 记载的方法同样的方法提取总 RNA, 得到通过 PCR 进行基因表达量的定量所用的 cDNA。对毛发形状易感性基因的表达量的定量也通过实施例 7 中记载的方法进行。

[0550] 作为调整基因的表达量的物质的判断标准, 例如和对照例比较, 基因的表达量高 10%, 优选高 30%, 更加优选高 50%以上的话为显著高, 可以选择该被检测物质作为毛发形状易感性基因的表达促进剂。另外, 例如与对照例相比, 基因的表达量低 10%, 优选低 30%, 更加优选低 50%以上的话为显著低, 可以选择该被检测物质作为毛发形状易感性基因的表达抑制剂。

[0551] 在上述筛选体系中评价约 700 种的植物提取物, 探索调节毛发形状易感性基因的表达量的物质。其结果, 分别发现表 13 所示的该基因的表达促进剂以及表达抑制剂。

[0552] [表 13]

[0553] 调节毛发形状易感性基因的表达量的物质

[0554]		植物提取物名	ANXA9 基因的表达量 (以对照例为 1)
	表达促进剂	鸡纳树 (树皮提取物)	6.04
		火炭母 (全草提取物)	4.82
		圆叶牵牛 (种子提取物)	2.03
	表达抑制剂	藁本 (根茎提取物)	0.38
		木蓝 (叶提取物)	0.22
		罗马甘菊 (花提取物)	0.18
		植物提取物名	IVL 基因表达量 (以对照例为 1)
	表达促进剂	黄樟 (树皮提取物)	6.00
		西番莲 (全草提取物)	3.69
		越橘 (全草提取物)	2.65
	表达抑制剂	穿心佩兰 (叶、花穗提取物)	0.42
		温州蜜柑 (果皮提取物)	0.29
		柚子 (果实提取物)	0.18
		植物提取物名	CRCT1 基因表达量 (以对照例为 1)
	表达促进剂	圆叶牵牛 (种子提取物)	7.72
		花水木 (树皮提取物)	3.34
		茵陈蒿 (花提取物)	2.13
	表达抑制剂	韭菜 (种子提取物)	0.38
		短毛堇菜 (全草提取物)	0.32
		欧白鲜 (根皮提取物)	0.20

[0555] 参考例毛发形状和毛囊的形态之间的关系

[0556] 一般毛发形状取决于人的种类的不同而有差别,亚洲人种中相对直发比较多,非洲人种中卷缩发 (或者卷发) 是主流。印欧人种中具有这两者之间的波状发 (波浪发) 的特性的比例比较高。作为与这样的毛发形状的不同相关联的特征,可以举毛发生根的毛囊的形态。即,毛囊的形态为弯曲的话,则毛发弯曲;毛囊的形态为直的话,则毛发为直发 (Thibaut, S. et al., Br. J. Dermatol., 152(4), p. 632-638, 2005)。

[0557] 为了更进一步详细调查毛发形状与毛囊的形态的关系,从各人种的人头皮组织来制作毛囊的组织标本,观察毛囊的形态。另外,对于从被检测者身上的检测体的收集,事前得到了伦理委员会的承认,在此之上,知情同意实施担当者通过书面对对象者说明了研究内容,并得到了书面的同意。将采集的毛囊包埋于冷冻组织切片制作用包埋剂 Tissue-Tek OCT Compound (Miles 制造) 冷冻,按照常规制作冷冻标本薄切片之后,施以 HE 染色,并用显微镜观察。

[0558] 图 11 中显示了各人种的毛囊组织像,由图 11 显示的结果可见直发的亚洲人的毛囊为直的,波浪发的高加索人的毛囊仅发根的最下部弯曲。进一步判别在卷发的美国黑人的情况下,毛囊组织总体为弯曲的。因此,可以确认毛发形状和毛囊的形态有密切的关系。

[0559] 实施例 9 由人毛囊器官培养的毛囊形态的评价

[0560] 作为评价毛发形状、以及毛囊形态的方法,针对人毛囊器官培养的评价方法进行探讨。拿到通过美容整形外科手术切除不要的 30 ~ 80 岁的男女的侧头部或者后头部的头皮组织,用于实验。另外,对于从被检测者身上的检测体的收集,事前得到了伦理委员会的

承认,在此之上,知情同意实施担当者通过书面对对象者说明了研究内容,并得到了书面的同意。

[0561] 将得到的人头皮组织,回收得到装满含有 1% 抗生素・抗真菌剂 (Invitrogen 制造) 的 Williams' E 培养基 (Sigma 制造) 的浅底盘中。在实体显微镜下用镊子和手术刀或者注射针的刃,在无菌条件下分离出每根毛囊。分离毛囊是在皮脂腺下部的位置处与表皮组织切断,并尽力除去毛囊下部所附着的多余的结合组织、脂肪细胞等。将调制的分离毛囊转移到添加了含有 400 μ L 的 10 μ g/mL 胰岛素 (Invitrogen 制造)、40ng/mL 氢化可的松 (Sigma 制造)、2mM L-谷氨酰胺 (Invitrogen 制造)、1% 抗生素・抗真菌剂 (Invitrogen 制造) 的 Williams' E 培养基 (Sigma 制造) 的 24 孔培养板中,每孔中一根,开始培养。培养为浮游培养,在 37℃、CO₂ 浓度为 5% 的条件下进行。以后,在每隔 2 ~ 3 天进行培养基交换的同时,拍摄毛囊的照片。

[0562] 随着培养天数的经过的毛囊的形态变化的照片如图 12 所示。毛囊中的发干在培养经过同时增殖伸长。另外,在培养经过同时,从培养开始 1 天之后 (1day) 原来为直的毛囊 (直发) 在培养天数经过的同时慢慢观察到了毛囊 (发干) 变弯曲的样子。

[0563] 为了将毛囊的 (发干) 的弯曲程度数值化,算出末端间距离率。末端间距离率是表示弯曲的程度的指标之一,通过以下的计算求得 (Hrady, D., Am. J. Phys. Anthropol., 39(1), p. 7-17, 1973)。

[0564] 对象物 (毛发、毛囊) 的末端间的直线长 / 沿对象物 (毛发、毛囊) 的轴的曲线长

[0565] 即,根据上述式子,末端间距离率表示 0 ~ 1 之间的值,如果直的话则接近 1,如果弯曲程度大的话则接近 0。

[0566] 使用图像分析软件 (Nexus NewQube Ver. 4. 23, IMAX System 制造) 分析图 11 所示的毛囊照片,求得毛囊 (发干) 的长度和末端间距离率 (表 14)。

[0567] 其结果,可以确认随着培养天数经过,毛囊 (发干) 在伸长的同时,也在慢慢变弯曲。因此,表示如果使用这个评价体系,可以探索毛发的卷发化剂、或者卷发改善剂 (直发化剂)。即向该人毛囊器官培养评价体系中添加被检测物质而培养,测定伸长到一定长度的毛囊 (发干) 的末端间距离率。如果与不添加被检测物质而培养的对照例相比较,在添加了被检测物质而培养的情况下末端间距离率变小的话,可以选择该被检测物质作为毛发的卷发化剂,如果与不添加被检测物质而培养的对照例相比较,在添加了被检测物质而培养的情况下末端间距离率变大的话,可以选择该被检测物质作为毛发的卷发改善剂 (直发化剂)。

[0568] [表 14]

[0569] 在毛囊中,随着培养天数的经过,毛囊 (发干) 的长度和末端距离率的变化

[0570]

培养天数 (day)	毛囊 (发干) 的长度 (mm)	末端间距离
1	3. 465	1. 005
3	4. 419	1. 002

6	5.732	0.997
8	6.748	0.988
10	7.571	0.973
12	8.131	0.958
14	8.758	0.901
16	9.433	0.825
18	9.720	0.818

[0571] 实施例 10 由毛发形状易感性基因的表达调节剂的人毛囊器官培养得到的评价

[0572] 以验证毛发形状易感性基因表达调节剂对毛囊形态产生的效果为目的,使用人毛囊器官培养评价体系进行评价。

[0573] 按照实施例 9 调制人毛囊。将分离毛囊以大小均匀的方式分为 2 组,每组 12 根,将一组用添加了表 13 中记载的作为 ANXA9 基因以及 CRCT1 基因的表达促进剂的圆叶牵牛提取物至最终浓度为 0.2% 的器官培养用培养基 (400 μ L) 进行浮游培养 15 天。另一组作为对照组用添加了 50% EtOH (添加到最终浓度为 0.83%) 的器官培养用培养基 (400 μ L) 进行浮游培养 15 天。以同样的顺序,制作表 13 中记载的作为 IVL 基因表达促进剂的西番莲提取物 (终浓度 0.2%) 添加组和对照组 (50% EtOH、最终浓度 0.83%) (各组 $n = 12$)。

[0574] 培养开始后,在每隔 2 ~ 3 天交换培养基的同时,拍摄毛囊的照片。从毛囊的照片画像上分别测定毛囊的伸长度和弯曲程度 (末端间距离率)。

[0575] 在毛囊 (发干) 的长度和培养开始的时候相比伸长达达到 1.5mm 以上的时候,测定毛囊 (发干) 的末端间距离率。其结果判定圆叶牵牛提取物添加组和添加 50% EtOH 的对照组相比,表示毛囊 (发干) 的弯曲程度的末端间距离率显著增加,并且,西番莲提取物添加组和添加 50% EtOH 的对照组相比,末端间距离率降低 (图 13)。由此,表示可以选择毛发形状易感性基因的表达调节剂作为毛发的卷发改善剂 (直发化剂)、或者卷发化剂。

[0001]

序列表

<110> 花王株式会社

<120> 毛发形状易感性基因

<130> KS1105

<150> JP2009-231991

<151> 2009-10-05

<150> JP2009-232001

<151> 2009-10-05

<150> JP2009-232002

<151> 2009-10-05

<150> JP2009-232003

<151> 2009-10-05

<150> JP2009-232004

<151> 2009-10-05

<150> JP2009-232035

<151> 2009-10-05

<160> 58

<170> PatentIn version 3.1

<210> 1

<211> 23252

<212> DNA

<213> 人类 (Homo sapiens)

<400> 1

gagggccagg gggaccccggtctgcggagct gcagtaccgt gtttaccag cctaactgcg 60

gtcaccagtc ctctgcatg tgtctgaggt cagcctcggg ggctcccctg agtaggtatc 120

tacctgaggc tcttcccctg atgctacctg cataccaatt ccaccagga acctggcgtc 180

ctcaggaggc tgcacttccc aagtgggctg agacagcccc ttaatatgtg gggtgggagc 240

[0002]

cagagacctc aaggggccac tagggggctc gagactctca gtggctgctg acgccctctt	300
gtggccattt tggggattgg ccagaaaagt ttcagaagcg agccaagatc ggaattctta	360
aggaggtaag taataacaac agtaacagtt accatataac aatacttate gttatatatt	420
tgataacatt gtgacatttc tacttaacag cgctgtgccc agaactcttc taagtgtttt	480
atatgcactg tagttattta attctcaca caacctataa cataaacact attacttgtc	540
tgttgatgat gaggaaactg agcacagaat agttgcccga ggtcacacag aagaccagaa	600
tcccgaactc gagaagtctg gctccagaag ctttgccttt agccaccacg ctgagactga	660
gttccatttt ggggttgtea ggtgaaatac aggacacact gttaaactat aatttcagat	720
aaacaaaaaa taataattac aggcgggtggc tcacgcctgt aatcccagca ctttgggagt	780
ccaaggcagg tggatcacga ggtcaggagt tcgagaccag cctggccaac atggtgaaac	840
cctatctcta ctaagaatac aaaaaattag ccaggcatga tggcaggcgc ctataatccc	900
agctactcgg gaggcggagg caggagaatc gcttgaaccc gggagggtga gcttgcagtg	960
agccaacata atgccactgc actccagcct gggcaacaga gagagactcc gcctcaaaaa	1020
aaaaaatgtt agtatattta atgaaatatt tgggtcatac ttgtaccaac aaatgatcc	1080
tcatttacgt gaaattctga ttaagctgca tgtctgtat atttatttat ttaatcta	1140
accctactt tacagggcac tttctctgc cacatgcaat tacacagatt actttttttt	1200
tttttttttt tttttgagac ggagtctcac tctgtcgecc aggetgtgga gtgcagtggc	1260
gccatctcag ctcaactgca cctctgcctc ccagggtcaa gtgattctcc tgcctcagcc	1320
tcccagtag ctgggattac aggcattgcgc cgccatgccc ggctaatttt ttgtattttt	1380
agtagagacg gggtttccact gtgtaagcea aggetgggtc cgatctcctg cctcatgate	1440
cacccgcctc ggccctcaca agtgctggga ttacaggcat gagccaccgt gcctggccac	1500
attaccttat ttatttattt atttatttat ttatttttat gaggttcaata gtttttgggg	1560

[0003]

acaaggtggt gtttggttac atggataagt tcttttagtgg acacattact tcattttaatc	1620
atcccaggaa acctgtcctc ggaggcccca gcatgtccca cagcctccgt ctctctcaat	1680
cctcccattt agagagctga aaggactagc gaagcttaca gaacataggt ggcagagctg	1740
gatatgaac cctcattctc cagttctgac acaggtgcat gttttccttt taacaggaga	1800
gaaaagtgag gagcagtgat taccgccaca gaaatacctt ttcattggcgt gtcctccggc	1860
atggtctact cagtgtggtg gagccactgc aggtccctcc tccagggatg taggaatgcc	1920
agtgaaggcc ttctgagtt gggagagagg cagaggagtg agtggaggca gggctgtagc	1980
ttctgtgccc cgcctctacc caggtctctc agggaggcag gaaagaggct gtgttttagac	2040
ctgagggagc cagctgtgag gctggagcag ttgctgcatg gcggggcggg ggctccacag	2100
ggctgttcac ctgctgctct gtgcagagac agcctcaagt ccagctgctg gggttgcatc	2160
acctgcagct aaaacagcca cagggccccag gctgccaaaa ggcaacagaa gcaatctaag	2220
aaaccaggca agggagcctc ccccatctcc ctgttgcttt ttgggagcat agagcttccc	2280
tctgaaggta acacccttg catttttctc tcttgactct gtcatcaac ttccactccc	2340
aaatacttgt gcccttcttt ctccacagat atacacaatc cctgggcaact ctccacctag	2400
gtcagctggc tctagtaagg cagcagtcaa gggagtttct gggtgggac cccctctggg	2460
gtgcaggctg ggggtgccctg ctggtatgtg cccaacagg tttttcccta tcagcettga	2520
aatcagcctg cagcctgtct gacctgcccc ccaggagct ggagccagag gcagagccag	2580
gtcagaaget ggggcagagc caggtcagaa gctggacttt gagactaggt ctctgtggtc	2640
ccacctggac ccacaggga ggcagtgcac aaaagcctcc tgtgtttgag gctgagccgc	2700
tgagaggctg agtggagtcc actcacatgg attgaggccc agttcctggg agaagagatg	2760
ctgggcagga aggtgtctgc atgtgggact ctgtacagcc cggctctctc ccacgtctgg	2820
gaggggccag agtcagacaa ctgctgggtt cgtccctaag agaggtcac tgactggctg	2880

[0004]

ttcagcctag gctgcacaca cccccacttt cctctaccag gccacaccgg aggcagtgt	2940
cacacaggca agctaccagg ccacaacaac gacacccacc tcacctctgg cacctctgag	3000
taagtactgg gccacaggcc catgccagg agggcttggt gattgcttgg gggagggaaa	3060
aatgagtttt ccttttccaa catgagagt cctcctttct gtccaggcat ccacgtactt	3120
gcaagaactc ttgtcacat cagctaagag attgcacctg ctgacctaga gattccggcc	3180
tgtgtctctg tgctgtgag caggtaggca gctgtacagg caggaattga gggttgcttg	3240
tcagctgggg ctgccgttg ggaagatgc tatttcgagt tgtcttccac tctgcgtggt	3300
taatctgcc acctcacctc agctccctg cactgccag gagacccag ccaggcccag	3360
gaaggagggc catagacctc agaaggagg catctaccgc ggaccaaggc ctgaggagct	3420
ttggtgtga ccctggcacc agggagggtt ggccctggt ccaggagcat tccgcatcaa	3480
tccccacttc ctaaaggagt ggggagaggt ggggaagggt ggggctcccc tcttacctca	3540
ccatttcagc tccattctg gtccaagcc taagcagggt gatctttttt tctgcagcaa	3600
ggaaaggaa acagtgactgg aaccagggt ccagggtctc gggggggtca taaggagtcc	3660
ctgaagggcc aggaattca gctctgtgg ctcctgttcc cttccccagg gcaaccagta	3720
gcaccatgtc tgtgactggc gggaagatgg caccgtccct caccaggag atctcagcc	3780
acctgggcct ggccagcaag gtaggggctg ttgggatttt agagatcact tttagtatct	3840
cagctcgggc tgtgagagga tgggaggcag gaaagtggag gacagccct gaggttatgc	3900
tgaggaggtg taggcagcct ggccctggcc cttgccaaact acctgtgct cccacagact	3960
gcagcgtggg ggacctggg caccctcagg accttcttga acttcagcgt ggacaaggat	4020
gcgcagaggc tactgagggc cattactggc caaggtgagc ccctttcccc cggcacttga	4080
gactgccttt tagagccaat ctgtagacca aggaggagcc aggaaggagg aactagtgt	4140
gttgagaaaa tgtttgtgga atgccttcat tcttgtatgt caggcgccac agtaggtgtg	4200

[0005]

tcaaacecat tctcacaaca cccctgtgag gcaaggetta ttatttccat tttccagatg	4260
aggaaactgg ggctcaggga gattaagagg cttgcccagg gtcactcagc tcatgacgga	4320
ggtggtattc tcatctaggt cagccgatcc ccccaagtcg tttttttttt taaccaccac	4380
accaggccac ctttgccctg gtctgggtca gatttgcatt ctgttttgac attttatttg	4440
caacttcttt tgagataaaa tggaggtggg ggtgatatag cccctgctcc ttctctccc	4500
agccctctcc tggagcgctt tccctctccc aaggagtagg aacatcccac tatgtaaaga	4560
ccctgctcac agcacagccc tcatctcggt tcaggcggtg accgcagtgc cattgtggac	4620
gtgtgacca accggagcag agagcaaagg cagctcatct cagaaaactt ccaggagcgc	4680
acccaacagg tgaggccatg ctgacctccc acagcagtgg actgggggtga gaaaccccct	4740
ggaggactcg agagacagca acagccgctg aggccagccc ctgcttctctg gctaaacagg	4800
gcctgagatg aaaaccagcc aaatcttccc aaacctcccc acacttctgg ggcccttccc	4860
ctcacccccg tccctctgct gtgaggacca tttattgtag gacctgatga agtctctaca	4920
ggcagcactt tccggcaacc tggagaggat tgtgatggct ctgctgcagc ccacagccca	4980
gtttgacgcc caggaattga ggacagctct gaaggtagca ggaggggaga cttgctgggg	5040
tgtctgggga agggagaagg ctgtcagcct tgcttttgca agacgagagt cccctctcg	5100
tcgtcccata ttgtttctta gaaggagaag cagccagcgt ctgccccctt gtctgccatc	5160
ctaatgacac ggccagtctg agagtagact cccaatttct ttaggcctc agattctget	5220
gtggacgtgg ccattgaaat tcttgccact cgaacccac cccagctgca ggagtgcctg	5280
gcagtctaca aacacagtaa gaatatagag ggaggggtcc cagatgctgg agaaggggct	5340
gtaggacagg ccactcctct cctgtactcc ccataggtga gctatctggc aacctggctg	5400
cctaagttct attcccttgg ggtctttcct cttgccccca ccagtgcctc agtctctcag	5460
gggagagacc cagatgtgat cattaaaaaa agagaagcag gccaggtacg gtgactcatg	5520

[0006]

cctgtcacc	caacactttg	ggaggtgag	gtgggcaggt	cacttgagcc	caggagtcca	5580
agaccagtca	caggcaataa	aggaaaacct	catctgtgct	aaaaattcaa	aaaaaccagc	5640
tggtcattggt	ggtgtgcgcc	tgtgggccca	gctactcagg	agtttgaggt	gggaggatca	5700
ccccagcttg	ggaagttgag	gctgcagtga	gctgagatcg	tgccactgcc	actgcacacc	5760
agcctgggtg	acaggagtga	gagcatgtct	ctctctctct	ctctctctct	cacacacaca	5820
cacacacaca	cacacacaca	aataaagaga	gagaagcaat	gattaacccat	cttaattctgg	5880
agctccaagg	cagattccaa	agtgacagtc	ccaaaggtae	tcattgtcacc	ttcccagccc	5940
cttgccatcc	ctgccatcac	accacattcc	aagctgaacc	cagccatgct	gcagcctgag	6000
gaaaagggtg	gaattctgcc	tctgagtcac	gggtgtcacc	tttcatgtcc	ctgatggctg	6060
gccttgcttg	gtacaagtg	tctcaccac	taatttgttg	ctttttccca	tectcatctt	6120
accacccctt	ccccccacc	ctacctcttt	ttcttgcttt	ccccaatttt	cctttgatgt	6180
ctaatcttgg	cctatgttct	tcttttgaga	caaggctctg	cctgttagcc	taggctgtgc	6240
agtgatgtca	tcattggtca	ctgcagcctt	gacctctctg	ggctcaatca	gtctctccac	6300
ctcagcctcc	taagtagctg	ggactacagg	tgtgccacca	ccatgccccag	ctaaattttg	6360
tattttttgt	agagatgggg	tttcgctgtg	ttgtccaggc	tggtctcgaa	cttctgggct	6420
caagccatct	cctgcctcag	cctctgaaag	tgtctgggatt	acaagcatga	gccaccgccc	6480
tggecggtgc	ctatgttctt	aatccctctc	ctaccttctt	tccttcttgc	atctgtttcc	6540
ctttctctct	tctttttctc	cctttctctg	tcttttctct	cctcatatcc	caactctcat	6600
atcatttttc	cttacctcca	atgagggtgc	caggctcctaa	ccatcttctg	ttataataacg	6660
ggtttcccta	gctcttcttc	actgcctggc	cccaactcta	ataatgggtg	cctctctgcc	6720
agcaggttgg	tgcttggtga	tgaagccact	gcacacctca	caacttagag	atgttttttag	6780
agtggccgtg	atggtacttg	tgctttacat	cctgggaatt	agcagccatc	aagctactcc	6840

[0007]

aagcaggggc tagcagggct ggcagggtat agcgggtcag gcatttggca ggtcctgggg	6900
atatctacac taggaaaact caggaggatt cagagagctc ctcacccac ccagatttc	6960
caggtggagg ctgtggatga catcacatct gagaccagtg gcattcttga ggacctgctg	7020
ttggccctgg ccaaggtgag gaggactgac ctgcaaggaa gggagtcacg ttcaccaggc	7080
aaggagacac gctggagcag ggagcattgc tgtcctgtaa tggacaagag agggctttag	7140
gggagaccag ccttcgcttg ttggtctgaa gataaaaagg tgccttcaa agagccctct	7200
aagaacagtt tctcctccta ggggggcccgt gacagctact ctggaatcat tgactataat	7260
ctggcagaac aagatgtcca ggtgagcagg gggtttagga gtgtgcacag ccgcatgca	7320
cataagatgc cttcctccac catatcttag agccggtagc ctacgtgcc atttttcccg	7380
caaaccctc cctgccttgg agaggagatt ctacgttatg tggtaggggc aggaggtgtc	7440
aggtagcgtc atagctgaag ccttttact catctagcct ctgactcctg agtcatcaa	7500
ctgacatatt ctcctctcat gacggtttgg gggcagaaag cctcctcaga ccttttgggg	7560
attatttaat cctcgggtac caggattttt ccaaatcagt gaatagagct gtgaggctgt	7620
aagggcagag ccctctctta gccctatttc tttcttaaaa aataaaaaac ataagataaa	7680
aaaacataaa taaaaacaac tagtaggcca ggcattggtg cttatacctg taatcccagc	7740
actttgggag gctgaaacaa gaggatcact tgagcccagg agtccaagac cagcctaggc	7800
aacatagtga gactctatct ccacaaaata tttttaaaaa ttagctgcac ctgtagttcc	7860
agctagctag gaggtgagg tgggaggatc gcttgagcct ggggaagtga ggcctaggta	7920
agccctggtc atgccactgc actccagcct gggtagacaga gagagacccc atctctctcg	7980
ctctctctct ctctctctct cacacacaca cacacacaca cagcacaca cacaaccta	8040
gtttacgggt taaaataaac tttctggctg gtcgtggtgg ctcacaccta taatcccaac	8100
aatttgggag gccgagacag gtggatcact tgaggtcagg agttcgaaac cagcctagcc	8160

[0008]

aacatggtga aaccccgctct ctactaaaaa tacaaaaatt agccaggcat gatggtgagt	8220
gcctgtaatc ccagctactc aggaggctga gcctggagaa tcacttgaac ctgggaggca	8280
gaggttgtgg tgagctggga tcacaccatt gcactccagc ctgggtgaca gaatgagagt	8340
ccttctcaaa aaaaaaaaaa aaaatccttt tccacctca atccatttc catcatgatg	8400
atgatagatt ctggaacgtc ataccatgg ctcctaggc cccaaccaat gatcctgatt	8460
gactcctccc tgactcattc ctccctecta ggcaactgcag cgggcagaag gacctagcag	8520
agaggaaaca tgggtcccag tcttcacca gcgaaatcct gaacacctca tccgaggtac	8580
acacaagcct tcttgteccc ctagcttget ctaatgatca gtttgggctg aggaagggtg	8640
ggagggecca tctttccag agataattaa tccccatcc atctttctaa ctgcctccta	8700
cacacacaag tgtttgatca gtaccagcgg agcaactggc aagagctgga ggaggtgtc	8760
cagaaccgtt tccatggaga tgctcaggtg gctctgctcg gcctaggtag gggcctgctc	8820
aggatttgtg aagtaagtct ctcttgggat gggagattgt ggtcctagtt gtgaacctcc	8880
atccttccat ctttgtttcc agcttcgggtg atcaagaaca caccgtgta ctttgctgac	8940
aaacttcac cagccctcca ggtgagaggg gcactcctt cctccccag aacagaaact	9000
ggggaggaga gaggaagtct cagcttgctg cttatgtaac catcctatat acaccgtcta	9060
aacctcaagc cgtctcctt cccagggtt acacaccagg gttacacag ggggtggagga	9120
atgagctgtg gtgagaacta acttagttgg ctcccaggg ccctgaaact aacattttgg	9180
ccttgtaat gtctaggtgg ccaaatagag atatgaatct gctctgtaag acagtgcctt	9240
gtgcaggctt agttgttgc agagtgatgg tgataaact agcaagaata atgacaagg	9300
ctggcatttt gcagggtctg attgtgtgcc acagcctatc ctgagcaggg cacacatatt	9360
tcaattaatc ctcaacaaca tatccatccg tatcttatag atgaaaaaac tgaagcccag	9420
agaagttaag taccttgctt gaggtcacgt agctagtaag ttatacagtc tccctacaca	9480

[0009]

acacttcacg atgtaagata ttgaccatcc ttgtcagcta tacgataagt gctcttagca	9540
cccaaggccc cctttctcca cgctggtcag attttgttcc acattateca tacttttttt	9600
tttttttttg agatggagtt tcactcttgt tgcccaggct ggaggctgga gtgcagtggc	9660
acaatctcgg ctcatggcaa cctccgcctc ctggattcaa gcgattttcc tgcctcagcc	9720
tcccaaatag ctgggattac aggcgcccgc caccacatta tcctaatttt ttgtattttt	9780
agcagatgcg gtttcaccat gttggtcagg ctggctcaca actcctgacc tcaggtcac	9840
aaactgtctc ggctcccgga agtgetagga ttacaggcgt gagccactgc gccagccca	9900
tattatecgt acttaaaactg ctgactgtgg cttgagtgt tccaccaata gaatecagtt	9960
attctgttga ttgaattagg tctcctatta gtagaactgt ccccaaaaat ctataaataa	10020
aaagttcagg aagggcgcag tggctcacgc ctataatccc agcactttgg gaggccaagg	10080
tggatggatc acttgaggtc aggagtcca gaccagcctg gccaacgtgg tgaaaccctg	10140
tctctactaa aaatacaaaa attagccagg catggtggca ggcacctgta attccagcta	10200
cttgggagtc tgaggcagga gaattgcttg aacctgggag gcggagggtg tagtgaacca	10260
agatcacgcc actgcactcc agcctgggtg acacagtcag actaaaaaag tagttatcat	10320
tttggcagat ggaatagctt tctgtagact ggttgagaag ccatcatgcc atgttgatga	10380
ttgccaaaa gttagtaaat ataagcaatt gttctgtcat ctagagattt ctatcttgac	10440
agtagattat tatacttgcc caataaagcc tgtatgatta taatgagtca aactaagat	10500
tcatttgaag agttgtactt atctgtccaa attcctgtca taaatacaaa ctateccagc	10560
agtagatgat tttgtagtgt atttgatta caagaccac actcttgagt gtagataaaa	10620
gatttatctt agaagtagta gtcaggggcc aagcatggtg gtcacacct gtaatcctag	10680
cactttggga agctaagggtg ggaggatcgc ttgaggtcag gaggccaaga ccaacttagg	10740
cgacatagca agaactcgtc tctacaaaaa taagaagaaa atttgccagg catggtggca	10800

[0010]

cctgcctcta gtcceaacta ctcaggagga tgtggtggga ggattgcttg agcccatgaa 10860
 ttagaagtta cagtgaactg tgattactcc actgcacttc agcctgggta acagagcaag 10920
 accctgtctc aaaaaaagaa aaaaagaagt ggtagtcaga aaataacata ggcccatttt 10980
 tctgaaaaaa aaaatttttt ttttgagata gagtctggct ctgttgccca ggctggagtg 11040
 cagtggcgtg atctcagctc actgcagcct tcacctcttg ggttcaagtg attctctggc 11100
 ctctgtgtct taagtagctg ggattacagg cgtagccac catatctggc taattctgaa 11160
 ttgtttctaa aaaggaaaaa ctcgggctgg gcgctgtggc gcatgcctgt aatcccagga 11220
 ctttggtagg ctgaggcggg tggattatct gaggccagga gttcaagacc agcctggcca 11280
 acatggtgaa accccatctc tactaaaaat acaaaaatta gccagggggg agtggtgagc 11340
 gcctgtaatc tcagctactt gggaggtga agcaggagaa tctcagcct gggaagcaga 11400
 ggttgccgtc agccgaggtg acgccactgc gctccagtct gggtgacaaa gtgagacctg 11460
 cctcaaaaaa aataaataaa caggccaggc gcggtggctc atgcctgtaa tctcagcact 11520
 ttgggaggcc aaggcgggcg gatcacgagg tcaggagatc gagaccatcc tggctaacaa 11580
 ggtaaaacc cgtctctact gaaaatacag aaattagcca ggcgtggtag cgggcgcctg 11640
 tagtcccagc tacgctggag gctaaggcag gagaatggct tgaaccagc aggcggagct 11700
 tgcagtgagc cgagatcgcg ccactgcact ccagcctggg caacagagt atgctccatc 11760
 tctaaaaaaa aaataataat aataaataaa taaacaaata taaaataaat aaataaggaa 11820
 aaaatcatga tacatctgta ttattctttt tattttttt tttttttatt tttttagatc 11880
 gagtctcact ctgtcgacca ggcctggagt cagtggcgcg atctcgctc attgcaacct 11940
 ccgcctcccg agttcaagcg attctctctc ctcagcctcc cgagtagctg ggattacagg 12000
 cgctcaccac catgcctggc taatttttgt gtttttagta gagacggagt ttcacctgt 12060
 tggtcaggct ggtccttaac tctgatctc gtgatccaca cgcctcagcc tcccagggtg 12120

[0011]

ctgggattac aggcctgagc caaggtgccc ggccttcttc ttcttctttt ttttctcttt 12180
 gagatggagt ctcactctgt cgctaggctg gagtgcagtg gcacgatctt ggctcactgc 12240
 aacctccacc tcctgggttc aagcaattct cctgcctcac cctcccaagt agctggaact 12300
 atagtgtccc gccaccacat cgggctaatt tttttgtttt tagtagaaac agggtttcac 12360
 cactgtggcc acgtgggtgg caaacTTTTT ttttttctt ttttaagatg gaggttgcct 12420
 ctgtcaccca ggctggagtg caatggcgcg atcttgctc actgcaacct ctgctcctg 12480
 ggttcagca attctcctgc ctccagctcc tgagtagctg agattacagg cgcgcgcac 12540
 cagcgcggc taattttgt atttttagta gagacagggt ttcacctgt tgaccaggt 12600
 ggtctcaaac tcctggcctc aggtgatcca ccagcctcag tcttccaaag tgctgggatt 12660
 acaggcgtga gccaccgtgt ccagcctgta ttattctttt atgagagata tgtagatata 12720
 tgaaaacata aattgtgaag aatatttatt cagaaaaaca accttcagaa aatcagcaat 12780
 ggctgattca ggacaaattt ccaccaatct gtagtaaaaa ttaatgttcc tgggtgaagat 12840
 cctagggccca tatgcaatgc caggtatacg attaagtatt ataattagc caggcatggt 12900
 ggctcatgcc tgtaatccca gcactttggg aggccaaggc aggtgaatca cccgaggtea 12960
 ggagtctgag accagactga ccaacatggt gaaacccgt ctctactaaa tatacaaat 13020
 tagctaggtg tgggtgtgtg cacctgtatt tccagctact tgggaggtg agataggaga 13080
 attgcttgaa cctgggagge agaggttaca gtgacctgag atcgaccac tgcacccag 13140
 cctgggtaac agagtgagac tccatctaaa caaacaacc atatatat atgctgaagc 13200
 gagcagatca cttgaggtca ggagtttgag accagcctgg ccaacatggt aaaacccac 13260
 ctgtactaaa aataaaaaaa ttagctgggc gtttggcaca tgctatagtc ccagctgctt 13320
 gggaggctga ggcaggagac ttgcttgaac ctgggaggcg gaggttcag tgagctgaat 13380
 tcacgccact gcattccact ctgggtgaca gagcaagaaa aaagaatact tgtacaaca 13440

[0012]

tgagagtatt cagttaatga gtgactcttg cccaattctg taactttaga gtaacaagcg	13500
ttggaagaag agaagaatgg atactgggga gcaattggca atctttatga caatacaaaa	13560
ctaatgatgc ataaaagtta aatgaagta gacattaggt ccactgacaa ttaggtaatg	13620
agtctaaatg aattggagca tctgtttta cttcattttc tttgttttg tcaaaggagt	13680
ttatgatcta ttgggatgcc ctaagtatgt gttgtaggca gttctgggaa gagtcttgtg	13740
agtaaagtga ctcaggaaac tctgaaaatg aggatattag gaattgatga caacactgg	13800
gcacagaaaa aaaaaaacac aggaatttta caattaaaag tttttattta aaaatctctc	13860
ataacaacac attacttagt ataatacaatt ataattcagg attttgatga cttgtacctg	13920
tgttatttct tttcttttct ttttttttga gacggagtct tgcctctgtcc caggctggag	13980
tgcagtgggtg tgatctcage ttactgcaac gtccgcctcc caggttcaag caattctcct	14040
acctcaccct cctcagtagc tgggattata ggcgcctacc accacacctg gctaattttt	14100
gtagttttaa tagagacggg ggtttcacca tgttgccag gctggctctg aactcctgac	14160
cataaatgat ctgcttgcc tggcctccca aaattctggg attacaggca tgagccacca	14220
cacctggcct gttatttctt tttcttttct ttcttttttt tttttttttt tgagacggag	14280
ttttgctctt cttgcccagg ttggagtga gtggcgcgat ctcagctcac tgcaacctcc	14340
gcctcctggg ttcaagcaat tctcctgcct cagccttcca agtagctggg attgcaggea	14400
cccaccacca tgcttggett atttttgtat ttttagtaga gacggggttt caccattttg	14460
gtcaggetgg tctegaacte ctacctaga tgatecaagc gctcctcct cccaaagtgc	14520
tgggattaca ggcatgagcc actgtgectg gectgttatt tctaaactga acctttcatt	14580
actttcttta gccaaaaaac aggcctgctg catgtacagg aggaagatag caatgtttca	14640
gcctaagggt ccaaccggag aaactaatat gcagagatgg ggctgtgaga tggccctatt	14700
gggttggaat cagaactaga agtcacacaa ccttgaacce tatcagagcc tcatcagctc	14760

[0013]

agtttaggga tagggcaggt atatectect tggatattca ctaacttgaa agttaaagc	14820
agttttectg gttttcatgc tggcagtcct tatatgcacc aaatcttttc tagccttcct	14880
gggtgtaccta tctttggagc tgctcaccac caaattctgg aggtacacag ctctaagcta	14940
atggtagggg caggaaagac atgtggaaag aaataaatac tcaagctgcc cccacagaga	15000
ctgatctttc ctgctgtgtc atcttccttg aggatggaca aacataactt ttgtcccgag	15060
ataatgtata atgtacactc ttateccttg actcccacct tgcctccct tectacaata	15120
accctccaag ccattcaacc cctactctca gtgagactct cactaactac ccccatctc	15180
tttttteccc tactaggaaa ctgagcccaa ttaccaagtc ctgattegca tccttatctc	15240
tcgatgtgag actgaccttc tgagtatcag agctgagttc aggaagaaat ttgggaagtc	15300
cctctactct tctctccagg tgaaacttgg ctacttctta gcctggagcc tcaggccttc	15360
actctcacc tccacctca ctccctgcac acagctgagc atattcttgc cccatagaaa	15420
accagtagt tagccaggca cagtggctca cgcctgtaat ccagcactt tgagaggctg	15480
aggtgggcgg atcacaggt caggagtcca agaccacct gaccaacatg gtgaaacccc	15540
gtctctacta aaaatacaaa aattagccag gcatgggtgc gtgcacctgt aatcccagct	15600
actcaggagg ctgaggcagg agagtcgctt gaaccagaa ggcagaggtt gcagtgagcc	15660
aagatcgtgc cactgcactc ccacctgggt gacagggcaa gactccatct taaaaagaa	15720
aaccagagag tctttgtta atgtagtcca ggactctgag ctcccgggag gaccttccc	15780
tcccagatga actgtgatgg accagcccaa aggaggggag agagcacttg ggccatagtg	15840
gtggtggatc ttttaaacac tgaattccct tgtctgcagg atgcagtgaaggaggattgc	15900
cagtcagecc tcttgccctt gtgcagggt gaagacatgt gagacttccc tgccccccc	15960
cacatgacat ccgaggatct gagatttccg tgtttggtg aacctgggag accagctggg	16020
cctccaagta ggataacccc tcactgagca ccacattctc tagcttcttg ttgaggctgg	16080

[0014]

aactgtttct ttaaaatccc ttaattttcc catctcaaaa ttatatctgt acctgggtca 16140
tccagctcct tcttgggtgt ggggaaatga gttttctttg atagtttctg cctcactcat 16200
ccctectgta ccctggccag aacatctcac tgatactega attcttttgg caaacttcgc 16260
tgttgtttgt gtccctgat tgaaggttgg gtggagcagg acatggaccg ggaagaggca 16320
ctggagtgg aagtgccttt gatgtgcact tggctattcc gcaggaatcc tgttttgcct 16380
tagtgetaca gtaatccaca cccaggtctc ccactccagg tctccctca cctcattct 16440
ccaggaaact tcaggcaaaa taattgagaa acaggcatta gaacaagtg agaaatagag 16500
ggagcaaagg accatctgta tgaactggga agcaagcgga gggccaagct accctccca 16560
gcatgggaat tcttgggtta gtggagagca caacctcaa attcatgtgt ccaagcagag 16620
atttgaaaac ccacctccca ggagagcact atttccata gaaaacaaaa acaaaaacag 16680
aaacaaaaac aaaaagtcag gctttggttc cccctgcagg ctatattaga aatgacaggt 16740
agctgggcgc agtggctcac gctgttaatt ccagcacttc gggaggccaa ggcaggcgga 16800
ttacttgagg tcaggaattc aagactagcc tgaccaaagt ggcgaaaccg tctctactaa 16860
aaatacaaaa attagctggg cgcagtggcg catgcctgta gtcccagcta ctcaggaggc 16920
tgaggtagga gaatcgttg aatccgggag gcggagggtg cagtgagctc agattgcacc 16980
actgcactcc agcctgggca atacagcgag gctccatctc aaaaaaaaaa aaaaaagaaa 17040
aaaagaaatg acaggttaacc accttctgc ctctgacceca cctcccgtt tttgcattct 17100
tcaattgggt caaatactta accctattgc ttcagggtta ggcaaacac agatctcagg 17160
aaagaaggtt ttttttttt ttagacagtt ttgctcttgt cgcccaggct ggagtgcaat 17220
ggtgcgatct cggtcactg caactttcac ctcccagggt caagtgatte tectgctca 17280
gcctcccgag tagctgggat tacaggcgcc caccatgacg ccagctaat ttcttgtatt 17340
ttagtagag acggggtttc accttttgg gcaggtggt ctcgaactcc taacctcagg 17400

[0015]

tgatccgccc accttggcct ctcaaattgt tgagattaca ggagtgagecc accgtgcccg	17460
gccagaagtt ttatttccaa accccaggaa ggcattacaa ataagagata gaaacccaaa	17520
ttaagctctg aaacaactgg agacaggcct gcctagggtga tcaggagcat ccaggcagca	17580
gggatgggaa gcagaagaga tgcattctgg ataggacact caccacagag cctcagtctg	17640
tacatactg tgactattca gggaccggga gttgagaacc agaaaccac caatcctagt	17700
gttgccctgg atgggaggca gagaaggcag cagcacgtga ggtcaaggac attaccaagt	17760
ctgaccttgg catttgttgc ctgctctcat ccccaacagt ccataaataa gttatccagc	17820
acatctcagg ggtggaggcg ggggagcaag ccaactagcc atagcctctg gaagaagggg	17880
caggccagcc tggcactggg gcagagctac agcagaatgc agtctgactc gtgcttcggc	17940
ctctgcccag gctccccg gcggcgctca gatgtggctc ctctccctg tateggattt	18000
aacaatttgt cactgaactt ccttggaaga gcacaagaag agattctaga agaaggggtg	18060
tcaggattgg gacagagagt agggagagga agacaggaag gggctccagg ggagagggac	18120
aggcagggga aggactctga gagaagaaac cttgaagaga ccagtgcagg aataacaggt	18180
gttctggacc cggcaaagca tgaggtggaa accgggagtg tgggcgccc ggcagcagcc	18240
cacctgcagt gacaggaccc gcgtccgcat ccgactggc tgcgtgcct gctgctgtt	18300
atactctct tgcagaagct gctgggccag ctccaagtgc gtaagccca gcgggcctgc	18360
tggtgttgc tgcagcagg acagagcaat caggtagtc tggggagaa aagagttgtg	18420
cagtgggtg aggccagag ccgtggctat tcaggcaagg ttccaaatc agttttccca	18480
catcctcaa attctctagc tctttctccc cacagagcag cacacacagc caaatgcatt	18540
atctgtgtg ccagagaata cacaccaagt gaggagacgg tctgaggtaa ccagagaga	18600
gaggcatggg ttcacatggc ctctagcagt taaaatgtat gcagtaaaaa gctctctgag	18660
gcacatgtgc tctctgact tttctaagt gcctacaca cctggctctac ctgcagctgc	18720

[0016]

gtttctgggg agccactccc accttctgct ccaggccctc tgcccaggga atgactcagg 18780

tgaagtcag agtcacaaaa gcagctgtct ccatccacat tgtgcaggct ctcccatacg 18840

acttgctcct cctgtagaaa gccctgggtca gtgaccagta ggtataagtg actctacaaa 18900

gaaaagaaca gagctcattg gggggacagt tgatacggag gcagttcact gccagaaatc 18960

tcggtatgag catgtggact ttttttttct tggagatgga gtctcgtctc gtcatccagg 19020

ctggagtgea gtggcacaat ctctgtctac cgcaagctct gcctcccagg ttcacaccat 19080

tctctgcct cagcctccca agtagctggg actacaggea cccgccacca tgcttggeta 19140

atgttttgta tttttttagt agagacgggg ttccaccgtg ttagccagga tggtctcgat 19200

ctcctgacct catgatctgc ccgcctcggc ctcccagggt gctgggatta caggcatgag 19260

ccaccgcacc cggaccctga gcatgcggac cttttacaaa tatgtgtggc cttgggggga 19320

agaaagggat gtctgagtgt atgtccagtt ggggtgtgagc acggtgggtca ttctggctaa 19380

gcgtggcatg gagtcgttgg ggctgagatg ccacataaga cctgtctccc cccatttccc 19440

attgttctct gggcttgctt gatggctctt aggataaatc ctttgagaac cagagacccc 19500

tccctctgct gctcaaatcc tagagagctt tggcagtttc taacttttga agaccagcat 19560

tgtttaaaaa ttccactgcc cctcctagg ggaggatcat actttcctgt cccatcgacg 19620

gcaggcttat cccttatggc ttgctctggc cagaagatgt gacagaaat gacatgtgtc 19680

atttccaaac agaagctcta agcatgagtg cacagtgecc ctctctctct ctgttcccac 19740

atatattcca gataaaggat cttttgtcag ctgggtccca gactgaagat ggtgtgggac 19800

agaaccgtag ccagtcacac gtggacacgc accaggagcg ggaaatgaac ctttgttgtg 19860

gtaaaccaca gggatttgag gggtcacttg ctaccacggc ttaatctagc tgatcccagc 19920

tggaccacag agaaaaaggc accaggaaat gacagcaccg cagcagccca ccccaaact 19980

cgcattacct tatgcttagt catggtgcta aagtgggtgt ttcggaaaaa gacgctaagt 20040

[0017]

tcaccctcct tagcagctgc tgcagctca cacagtcctg gtaggtcag ctgggccgcg	20100
gtggctctcca ggaactgctc tgcaatcagg cctgccagaa agggacgagt cgggggaaac	20160
ttggcttaaa ttcaaggtec acaacaggaa ggaccatcca gagagccct gtcaaaagcc	20220
cgggggggtc agtcccacct tctgacgtcc ctttcttgaa acctggctgt attcattcac	20280
ttaacaaata cttataatag agcagctacc ttgtgccagg cactggggat acctcaatga	20340
actaaacaga aattccccaa gctccttctt gtgaagctta tatctagtgg gatgtggtag	20400
gagatgacac ccaataaaaa ataagcctaa gtagctgtca ttccctccac atactatcca	20460
tgagaagggg aggaatgtca ccttctgtca cgaggttggg gtcactggag tgtttgcagg	20520
tgatgatcct ctccaccagc tggttgtaac tcagtttccc aactgcacgc acagcctcag	20580
gactctgctt gggtagtggg atgtgggata ccaaaaaaat tgggattttt ttttctttt	20640
taaggaagga aatgaacaaa gtcacaagg actacaaccc tgtgatgaat ggttcctaaa	20700
cacttttttt ttttggctc tcaaaaaaga aaaaaatttc aggaaccatt catcataggg	20760
ttgctatggt gccagactg gcctcgaact cctcccaagt agctgggaat acaggcatgt	20820
gccaccaaac ccatctgctt cctgaatttt tgatgtctaa agaactgacc tatttgcatg	20880
tagtttgctt gtatttgcac gttaccccca aaatagatta aaattcaagg aattctgtac	20940
tgggtccagt ctctaaacac tatttccctc ccaagccaat ctaacccta actcctaac	21000
ctctaccga caaaaatttg cctctctcat gtcagcttga gctggataag gtccagtg	21060
gccctccac cctcacctgt ggatcaacaa gccagccatg gtacagaggt atgcctagea	21120
ggtcaaagac actgcactcg ggtgtatact caaaatcaga gacgcctgtg aatgcacat	21180
tgacatccag acctgtggcc agtttaggea gcactgtcat tgcacatcc acattctggg	21240
ggtagaaaaa aaaatgatgg agattctage cttctctccc accactccac tcccactgc	21300
agaaattaa aattaacctt gtgctcttc ttccctgatt ctcttacaag ccaactcatt	21360

[0018]

tctcccttgc ctcaaccctt ctcctatctta tataaatgaa cattgttttt ctttgagagt	21420
ctaagtctaa ttctttcctc caggctggag tgcagtgggtg tgatcctagc tcaactgcagc	21480
ctcaaaactct tggtttcaag tgaccctccc acctcagcct cctgagtagc tgggactaca	21540
ggcagcacc gccatacctg gctaattctt gaaaaaaaa atggtttttt ttagagacag	21600
gggtattatt atgttgccca ggcctggcctc gaactcctgg cctcaagtga tctctccgct	21660
ttggcatctc aaagtctgg gattacaggt gtgagcacc ggccaatgaa catattttct	21720
aggacaaaga taaactctcc tctgactcct gcagaatgaa atcatgggtc aaggaaacct	21780
ctgaagaagc aacaaaagct tatcttagta ttacaagatg tccccttaca aacctgcaaa	21840
cccctatgcc tgcagacctt ttgcctcacc tgctgaaaat taagctgaag tccctctgac	21900
ttctcctggg gcttgatgga caggagcag ttctctacaa gacagggecc cttatcagct	21960
tacccacca atcatccca aagctaagtt tctctctctc ctggagaatg cgggtaatgt	22020
cacagaacac agattttcag agctggaaaa tacttagagt ctaatccaga ctcttcattt	22080
ctcagaggag gcgacagcct cataaagggtg aacggccac ccagggtcag atacaaaagc	22140
taagggcaga attggtgtta caaccagga ctcccgacgc ctttgccttt cattctccag	22200
ccttgcagag aagccccata gaccttacc ctccctgtgt taccagaacc cccaaatgga	22260
gatecctgaa tagtctggat ttctggctga gtggatagct ctggagcccc ttgggtaatt	22320
cctcttcaac gccagtcac tcaccaagat gggccatgag ctcatccgat gtgatcaatt	22380
ccttctgcgg ggggagcttc acctggaggc agaaggggtg atcagtggtg tgccttctc	22440
caaagagcac tgaaggaaga acctatggaa aggcaagggg gagctccagt agtagggtgg	22500
ggtaggaaaa aaggcttgca actggcaca gggaatgaat gtgaaatctt gcaaggcatg	22560
tgcatttgct agggagtagg agatacagg gttaggacaga tcttctctt ttttttttt	22620
ctttgtttt aggaacatct tatctcattt gctcaagctg gagaaatagg gagaacaatc	22680

[0019]

tgaaaagttt tgactagttt gaggacatca ggatgatcaa ggaagtaagt cagggaagag	22740
tactccctga tatTTTTTgc acccctaact gcatgttctc ttacettcca ctgaagaaag	22800
aggatgttca tgatggcaag gagagggcaa gggccgttag tgctctgggt gatgatgggt	22860
gtctgttctc ctttccaagg gatccacttg acacagtaga aatctggctc aggctgtcgg	22920
gtcctggggg actgaggaag ctcttggggc atggagcatg cccttattgt ctctacttca	22980
ggcagtgtcc caagggttgg ccccggtgga gctgagctag cttcaggcag aggggactca	23040
aggttgtccc cacactggct aggcagcaaa gcttgggtctg ctggctcccg ttctctagcc	23100
tcccatcag catctcttgc atctgtgtcc tgagggtgct catctgggcc tgccagaacc	23160
tcatggtttt cagggatgac tgcttctgca gtcccgccct taccaggggc tggatctca	23220
ggctgatggt attccatggt caaaaggac tt	23252

<210> 2

<211> 56552

<212> DNA

<213> 人类

<400> 2

ttttgaattc catctctgct cccaactctc catgtgatcc agggaaagcc ttttccatc	60
tctaagctc aggttccitt tctgtcaaac aaggggcatg atctaaatca gcaattatca	120
atcccagtgg tgcagcagag ccacttgatg gacgtctgga gaatgtcaaa gcctagatct	180
caccgggggtt gagaaccacc ggctctgaat gatggccaag gtcttttcca caactattct	240
atcateccat gaaatgtcta gaatttgta atgaaaatct gggtttggat ttgtttgagg	300
cattcatctt gcttctgaa tagtaaattc aactgaacaa ggaaatgaaa actgagcacc	360
tgctgcctac aaggcactaa attagacccc ttagggattc ctgcactaaa gacettacca	420
tgtggtcaga taggtgggtc gtgtgcccac acaattataa ctaaaagggg atggacattg	480
ttccttttag aattcaccaa acccatgttt taaatgggta attcctttaa ttttactggg	540

[0020]

attccctttc cagaaattca ggcctgtgt tctaccttct tgcttatact tctccacaag	600
gagggcagag acagatccga cgccccaacc ctgtgggtctt tatctaggcc tgccctgagt	660
gctggaattt agaagtgaca cagcatcagt ttctatttta aacatttatt gaatgcttac	720
tatatgccgg gcacttttac aagagtttta ggtatgttaa ctcaatttaa tgccccatg	780
agataaacac tacttttttg accatttcac aggcgaggac agtgagggtg aaagtcata	840
tcacttggtc aaagtcacac cactggtaag tggggagttg ggattcaaac ccagagactg	900
gatttagaac tcacgtccc agccgggcac ggtggctcac acttgtaac ccagcacttt	960
gggaggecga ggtgggtgga tcacgaggtc aggagatcaa gaccatctg gctaacacag	1020
tgaacccccg tctctactaa aaatacaaaa agttagccgg gcgtgcgtgg tggcagggtc	1080
ctgtagtecc agctacttgg gaggtgagg cagaagaatg gcgtgtatcc aggtggcaga	1140
gcttgcaagt agcagagatc gatcgacca ctgcactcga gcctgggtga cagagcaaga	1200
ctcagtctca aaaaaaaaaa aaaagaactc acgtcccaa ccacttataa atagattctg	1260
tgcccttcat gaaatgttct tatggctacc aaaaataatg ctccaaaca acgaataaca	1320
tgggaaatgt tgatgttaaa atgttaagt aaaagttagt aaagtaaact catgtaagta	1380
ctatgagaaa gactatgtaa aatacaaaaca aaactcaa atacaaca acaacaaaac	1440
cactatatat agaaaaaaat gacgggaagg aaaatagcaa aatagtaacc atattttagg	1500
gatgggatta taggtgattt taaaattctt cctatttttt ggtgttttta aataatgagt	1560
atgctttctt tataatctca ttcctagtgc acaagcaaat actcttaata ttttgcgtga	1620
agaaaacatc gagtctaagt tctgatatca taaaattag ttttggtttc aaaattggca	1680
ggtcagtaat ttgaaattta gcatatattt tctcatggaa ataaggatag ccatagtgat	1740
tagttttttc ttttttaaac ttgcagttct aggaactgtg ctaagaactt catctgtatt	1800
tctttattca gtggattagg aatgttttca gcccccttt ccagtgatgg ttaagtaact	1860

[0021]

tgccagaggt tgcacaggtg gagagagcaa ggggtgggaac tgatgctgga tgtccctctc	1920
cagggtctgt gttataactg ggctgatget tctcactgca gccatttca aagagcatgg	1980
agctaaggta ctgccttttt ctcccagtct tattttctct cactgcagct ttcattgtga	2040
atatgccata ttttggccac cctttccagc taagtcaatt ttgataacta tcctggggac	2100
ccaactttcc ctgagaatat tgtttgctga gaaaatgttt tggatcccac catgggcttt	2160
taaaacctat atactcctcc tcttcacctc tccatccaaa gattaatctt ctctgaggt	2220
ccttcccagg ttggatgctc ttctcttatg aatgaccaga actgggtgect cacttctgga	2280
aacctccaga tgagtggaga gettttgcct caccctcagt atacacctgg tctttgctgc	2340
ctcacctcga gctggccaag ggcaccttct ttaacacgag taaataagcc attcctagac	2400
ttctgccctt ttaaaggcat tgtagtgct gaatcggtac tgtttgtgag cagttttgga	2460
gcatcaaatt caggcaagga gagaagccca accaccaca ctgtcagcat tagatctgcc	2520
taaaataaca tcactgagat atcaaacctt ccttgttctg gttatggcat catctagaaa	2580
tgaaaaaaaa atcatgttag attatttgc tgggcctaca attttatttt ggatctctat	2640
tattattcta aaaagggcc a ttaatttcc atcttgggaa tttctcctt cagttcctat	2700
accaaagaga aaacagaact tctgttcatt ccagaggtgt agagccgtat tgcccagct	2760
cagaaaatag aaggcatagc agagctgect gtgagataga taggtgceat gtgcaattca	2820
cagagctccc tcaccatcct agaggcctgc agttctcagg ggatcttctt gtgggttaat	2880
acttgagatg cttatgcagg gaggaactgg gtccctggga tgcgagggga gggagtcacc	2940
catgccttta tgaaagatgc ttctagcgtc ccgcacatgg tctgcacctc ctgatgetag	3000
gtcaggactg aattgtataa aaggcagaag cttctgacag caccctcagtc tacctgtctc	3060
ctgagtgate tgctgcagtg cctgaaccag gtagagtgt tctcaggacc aggatgaact	3120
cttggtgctg gtgttttggg cagaaagagc ccctgggtgg aggttgagge cattcttgaa	3180

[0022]

gaaagacaag gataaagaag gatttttgag gggaaggttc ttctggaagg ggagggtgga	3240
aggtaggtaa aaggaatgaa ggactgatgc ctttctgaag ttccaagtta ttcatttgaa	3300
aatatggagt tactaggtagg aaatcatttg tatcattctt gttactggtc ttgggagact	3360
gggtttttta aatctgggtc atcttaggcc aggagtagtg gctcatgcct gtaatcccag	3420
cactttggga ggctgaggtag gcgggtcaa gaggtcagga gttcgagact gaccaacatg	3480
ctgaaatccc gtctctacta aaaacacaaa aattagctgg gtgtggtagt gcacacctgt	3540
ctgaggcagg agaatcactt gaatccggga ggtggagggt gcagttagcc gagatagcac	3600
cactgcactc cagcctgagt gacagagtga gactccatcc caaataataa taataatcat	3660
aataataaat tctggtcac ttagagataa aatcttgtga aattatagtt caattttaag	3720
ggcaaggaag tgaccttggg cttgtattgt ctttcagctt tattcagctt ctacagagat	3780
gtcctgccag cagagccagc agcagtgcc gcctcctccc aaatgtaccc ctaaatgccc	3840
tccaagtgt actcctaagt gtccctccaa gtgtcccca aaatgcctc cccagtgttc	3900
agcccatgc ccacctccag tctcttctg ctgtgggtcc agctctgggg gctgctgcag	3960
ctctgagggt ggtggctgct gcctgagcca ccacaggccc cggcagtcct tccgacgccg	4020
acctcagagt tccagctgct gtggcagtag cagtggccag cagtctgggg gctccagctg	4080
ctgccacagc tctgggggct ctggctgctg ccacagctct ggaggtgct gctgacctgg	4140
gccatgagga gcacggagga gaaggactgg cagatcccag gtgtgaaga tgtgtgtcag	4200
cctgaggett cttttctctc atttccatg gaaggacttc ggaaatgcct taagtteccc	4260
tctttatcct gccatgttc actccattgt agggttgaag tctagcttgt gatattttct	4320
ggcctggett tccctctcag acatagctct ttggagaact aggtgttgta attcagttat	4380
gaagctatct tctctgtaac aataaagctt tttattcctg atatcactgc ctctgattt	4440
cattgtttgc ctctccact cctacctgc tgcataagc catcttccct tccctctct	4500

[0023]

taaatgcaag cgaacacttt acaaacttgt ttaggaagca cattttggca gtgaagctac	4560
ttaggcctct ggaaaacaat atttccttct atgcaactgga ttctggaact ttaagcttga	4620
aggaaagaat aatgtctgag accattacgc cagacacaag attgtatgag gagagccagg	4680
aatgggtttg aagagaagag gatggttgta ctcaaaaggt tcaggttctg agagggtttc	4740
cttcaagcat gggtcacagt ttgaaaaatt gcttgtcccc attcatgggg aataagattg	4800
cctgaagtgc tcctcctggg ctctaataaa tggacactct ttacttttct gattattata	4860
agctcaggtg tttccatcag ttggctagag ggagagctgg gtgggcagtg agtgagtttt	4920
ggaggcagat atgaaacgtc tgagcacatt ttccctgcat gctgaaatcc tggatttgaa	4980
tctgggtaaa gaactgagaa agacaccctt ctttgcaaag actgtctcta tgtggttggg	5040
acgtacgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgcagtgag aggtataaga	5100
gtacaggggt tttggtcctt gtgttgggtt ttccaagcaa catttcttta ggaatgggct	5160
ggttccttca ggagacactc ataatgtacc tagaaaagat aactagaaaa gttaattcca	5220
aatcagcctt taatataaaa gataaacaat tgatgcctcc ttgataaga gttaatagcg	5280
ataaataaag taagctggtt tatgagatca taggatgccc acagaatgag taggcacttg	5340
cccagattat ctccttatac tgattcctac caaacgttct tcattctttt tttttttttt	5400
taaaagaact ttgccattt tcccagagga gatcctgcct ctgtcttggt atttcattac	5460
ttccttctct gttctaatat tgactttact ttttggaatg tctcactat actaccaac	5520
aggtgctcta ggaactgga accgtttaat caggtcteta atccatgta aatgaaactg	5580
gtatcttttg tgacttatct ttccctgcct agtcccttga gccactceca gctcactcgc	5640
ccccaaacct gtctctcacc cttaattcct gaattcttag tcttggaact ctgttatagt	5700
tcacattcac agttgtgggc aatttgatta agaagcagtt ctgaagtga ctgccaaagt	5760
atcctccaat tcttaaaaat tagttattct gaaaacacta tctatcatga acactttgga	5820

[0024]

ccttttgaaa gaactttctt tctttcttct tctttttttt cttctttccc tccctccctt	5880
cttcttttcc cattttcttt tgcactgggc agtggcccta gggagaactt atcttcaaag	5940
gcagattggg ggttgggggt gggataagaa gcaggagtc cacaggctga ctagcctttt	6000
tgcaattggg ccactgtcct gcacacattt ttctctttgt cataccatgc tgtattatgg	6060
ggatatggca acctgactac attcacaact tttctggctt tatggttcct taactttgcc	6120
taggtttttc ttctgtctc cagtgcaga gagaaatcac ccctggggtt cccaaagggtg	6180
catecccgtc atcatcttca ctcaactgcat cttgaagcat atggatgaaa cccagcaggc	6240
ttatcgcata gcttgtgggc tgccttgggg agccaaggac caagagccag agagatgaga	6300
atcaagataa aaatagggtg actaccacct accaagtaga tgactacata ggaatggatg	6360
tgtgggcaag tggagccgtc tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg	6420
tgtattgaag acaagggcag gggttccaga cctccctgca tttagacccc atgccctctc	6480
ttagcataca cacagcctcc caggtagaaa cagtgccttg tttagaaagt aaacatgcc	6540
tagcaatcct ggcacctctg tgaagtgact tagatgccc catgcaccac ctacaatgct	6600
acctgagtga tgaatgtctc tgggatgtga gtgtcagctg agcacgtctg ttgggaaatg	6660
ggccccctcc aagagcctga acctgaatac ttcttataaa aaggttctct ggctggacgc	6720
tgtcaatcc actgcctagc aggtggccca ttccagttgg agaacgtagt gagtctttca	6780
gtggagccag ggtctggtaa gtctcctagc aggtctctgg ggagagggtg ggatgagttg	6840
agggcacagg ggaggagcaa acagaacaca ggcagcagat agatgtggaa ggggtgtccc	6900
gcaaggcacc tacgttctta gtgagccaac ctcttaaggt actaatgagg acataaaggc	6960
aacataggag gtggtttttg caagggtaca cagcaagtc gcagaagagc tgtgactaga	7020
accaggtct ttcagctccc tggatgggggt tatttatatt ggcacaactc tcctttccca	7080
catggttgtg gggcatgaga ggcgtaggca agaaagaatg atcaggacat gcttctggcc	7140

[0025]

tgtggattat ttagaccag caggtgggac cgcctgttg acttgggtgt tttctaactg	7200
atgggggtgt agtaactagg aagatacagt cgaaggtcgg taggaaagta aaatgaggga	7260
caagaaaaaa gccaaagtga gattttggac tggatcatac agataagttg ccgcagtgtg	7320
ctgtttatit tccactgttg agaacttgag cagtcttctt ctagtcagag aaaatggaaa	7380
tgcattgatt cattgcatgg attgtgtcct gctctgagca tgittaaact actggaaaca	7440
cctgcttga tttgtctctt tgcgtccgt ctttctgggt gcttttccct ccttgtecca	7500
ggccagctg acttgtacac taagccgaat ctcaaaagcc cagacagagt tccattgggtg	7560
gaaagaaaaa gccggtctgc ctttgaaagg ctgcctttct tccaggtttg tegtaggag	7620
ctcccgatg tctctcaac agagcgcgt ttccgcaaa ggttttcca aggggtcgtc	7680
ccagggeccc gtccgtgtc ccgcccggc gccaccccg gcgcccgt cctctctc	7740
ctgctgggc tccggcagg gctgctggc cgactcaggc tgctgggct ccagctccac	7800
cagttgctgc tgcttccaa ggagacgcg ccgacagcg agtagtggt gctgctgtg	7860
cggggcggc agccagaggt ccagcgtc caacaaccg agctcaggat gctgctccg	7920
ctgctgagag gccgcgaacc ccagcgtc cgctagagaa acccgcccag ccagagcgg	7980
gccgccccg ctgcggtcc cagcggggc tggcctcgg agtttgcgc gtaaagcgaa	8040
ttgcactttg atgttcagaa acccaattt ttctcagcca cgaataact cctgacccg	8100
atgtgatttt tctccccgg gattcgagag ccatgcgtg gacctggac cctactgtt	8160
acacgggctt gcacacagca ggtgtcagc aaatgtctat tgatttgatt gtcttttgaa	8220
gatgtcataa taaagctct acctctgag aacaccttt attgttctt cattcgttc	8280
tccaccagc tagaaaggct tgaaatgtt gtgaaaggc gaatactata attcatgct	8340
gaacaagaag gtggctttct gatagttaat tatgatagca gcgatcacat atcgactgtt	8400
tcgagtattt taaggatttt atgcgacaaa gtccttgaat attctcagga ctttatgtg	8460

[0026]

cacaatctga ttctcataat gacttgcgta ggagagtga attataacca ttttacagaa	8520
gagaaactga ggttaagtaa tttgttcaaa gtcaccacgc tggcgcatga taaagaagac	8580
aggacatctg attcaagggc ttcattgctc tgctgagttc tcacctgggc ttcceaagcc	8640
ttaccactga ggtgctcctc ccgaccacca ggtccaggca gctgtgtggt aaaagtttgt	8700
tttggaact ctttggggct ttggaatcaa gccagctgctc ttcttttttt ttttttttag	8760
ttttatgtct tattttattt tttatttttg gcatggaagt tcattgacct actgctttgc	8820
ttgctgggga ggaaattttt cctaagatga gccttccagg aaactgggag ctgcctctc	8880
caggctctcat ttgtgcagct gaagcaggac caacagccat gggtttctgg gagtctctgt	8940
ggtgcagcat gtgtgctcgg aaacatgggg atctcattct ttagagatgg tggtgtggc	9000
agtggaatg gatgggtcat cttgcatagg tcagaaacct atccctgtgg gcacagattt	9060
tatccactca tttcaggcac gaggaacag gagaacaata tcctctgcat ggtgtctgct	9120
gggttcagaa ggaagagata atcatggtaa taataatagt aacactagtt accatttatt	9180
tggcactata tttcaggcac aagtcaggc actttacctt tattatctca cttaatctc	9240
acagtaatct aaaaatatag acattcaatt cccactttt tatacgtgag cacatgggtt	9300
tagacagttc cagtggtttg gctgaaatca cagagctagt gaatgtccc ttgagtgagg	9360
gagatgagat tgaaactgga acctgtgtga ctcaaagtc atgcctcac actcaaagag	9420
tcattgcteta ctgtcttaga aatcaccacc acacaggga gagacactcg ggtctctgga	9480
cacacceagg ctctcagaaa attgcagctc ggtctcttg ctgtctctc cttggtgacc	9540
caatgagtgg ggagctctgt catggcatcc acccagcagc tggcggtga ggggtgtctc	9600
tcaaaggatga ggtcggaat gctctactga ttcacccagg atgtttgacc ttgcggccaa	9660
atgacccctt gccaggagg ctctggaggg cagtctgct ctgcctggga gtggccctc	9720
cttccatgct tgagattctg tgatttcag tgggaggtga tttgtggaag ctgtttacaa	9780

[0027]

acagatgata	tcacccgtgg	acctgcctcc	ccacactagt	aatgtgtgct	gaaacagcaa	9840
gccctgggac	ttggtgggca	gtgggtggga	aggaagaact	gcctgactca	gaacagacac	9900
actacgcagc	taatctgagg	gtgagggctc	aacctggagc	aggaagctga	ggcctgagct	9960
ctgggaaaga	gaggtggtga	ctcacggaga	gatattagag	tccaaatagg	agcacatgcc	10020
ctgtacacca	ggcatttggg	tctcatgtca	aatggctggc	cctgatggtg	gtggtacctc	10080
caggcccaca	tgctgctgaa	accagcagg	tcagcacaca	tggactttga	tgagagacat	10140
ttgagtattd	taagaatagt	ttcattttag	aaatctagca	agcgactttg	gaaggtctag	10200
agctctctgc	tcattgtcct	tcaggttttt	ctgatttcca	gcagtaggta	aagctgtatg	10260
tgtggcccag	gcactctctg	tgaggtgatg	gtggtggatg	ctgagcctgt	gggaaatttg	10320
gcaaagtga	gtctactca	gaggttgtea	cctggtacac	gtgacaggea	gaatcaggtt	10380
tcgtattatc	tctggaataa	atggacatta	ttgttaaata	tattttattc	atgatggctc	10440
caatttcagc	ttaaattata	ttattatact	ggtaataata	acaactgata	cttatgtagc	10500
atttagtggt	tatgagcggt	ttccacatat	attatctctt	tcataattct	gtaacataag	10560
tattaccatc	tctgtttaac	agatcaggga	gggactccta	cagtttaaac	aaccacctg	10620
aggttatcca	gctgatacgt	gtggtctctg	tctggaccag	gccttcta	gccttctaag	10680
aaggccaaca	gccacttctg	aatattcact	gaaataattg	aataatggca	cccagttgat	10740
tacttagaga	cagtgcctga	ttcactattg	catccctagt	atcatgccag	gtacatatta	10800
aaggtctata	aataaatgat	tggtttattg	tgcataatgt	gtaaggcatt	gtgctaaaga	10860
attccacaga	gatatacaaa	gcatttctgg	acctcaagga	gtttatagtc	tgattcagaa	10920
actatggtaa	ataaaaatac	aagatataaa	tgataactca	gatctgaaat	gcaaaaaatg	10980
tcacaagcag	aacatggcat	aggcaactgc	tctagggatt	caactgagtt	tagatgaggc	11040
attgtgatac	gactttgaat	gagggcaagt	tactgtggac	tggattattd	ataagccttc	11100

[0028]

atgggaggaa ggatctgggc tgggtgtgaa ggatggaaac acatcagacg ggcagagatt 11160
 cattcattta tgtatctgtt caactccatc agacactgtt tatggttctg ggacagaacc 11220
 agaaataaga ccaaggccct gctggcatgg acctcaattt ctggagtga gacagacaag 11280
 aaagtgagca aggaagtcag attatttcac accgtgaaaa gtgttgtgat gaaaacaaag 11340
 aaaataaagg tataaacatt aactagaggt gtatggggat tttagatgga gctaggccag 11400
 gcctctctga agacatgtta ttgagcagg agaaatagga gttatgacac agagcaggga 11460
 agcccatggt tacagggttc ctgcatccat ggcagcctag ggctgcaggt gtttgagggtg 11520
 tgttttcagt tatgttgect ccaccagaca gtttcaactac taggaaagta acttggaatt 11580
 ccagtttctt ggcagacaga ttgctagtgt tgattgatcg tgtccatga gggtggcatt 11640
 atttcattgg atcctcacia ctctcctaca agacaagttt tccccacatt ttgttaagaa 11700
 aactgaaaga tagagctgtt aggtagccat cctaggattt cccagcttgt aataggaaga 11760
 acaaggtttt gtactggaga cattttcctc tgaagtcaat gttctgaaca ttgtgtctaca 11820
 gtctgcattt ctgggaaaga ggggagtagt ctcaagcttg gactagggga ggagaaaggg 11880
 gtctcaaggg agtgggtgtg atattgaaag accctgagga cacatcttca gagaaggaga 11940
 taggaggcca gatcatggat ggitttgagt gccaaatcaa ggagtttcaa ctatctcata 12000
 tttgtggagt atggtggtag gaaaatggct ttgaaagaag tacattagag agataaaagt 12060
 tggtttatat aacatctaga gtactctctg acatgacct aagattaatt aggcagcaat 12120
 gattaagggt ctcagaaga ggaagttaat ggagagatta cctctgtgag cagggtgcaat 12180
 ggcccaggca tgaaatgaag atctgaacct ggcacatggc agtcattaca aaaaggaaat 12240
 gactgttgtg aaaagcattt gtccagaagg aagagggata gaatcaggtg atgtattggg 12300
 gttgagaaag aggaagcaac tgggtgtct ctgaggtttt gtgccttgtg cctgggaaat 12360
 gatgacagca ttgacaaaat agagaagtca ggagaaagaa caagtgtgag gaggagaggt 12420

[0029]

aagtttggtc caagatgtgt tgggtttgac gtgctggcag gacatgacca tgtgacagtg 12480

atgtggggct ggaacttcag actcagaatc cttgaacatt gacataaaca caattaagat 12540

aaatggtaat aagccttatg cagcttggag tcaaactctt gggaatgaat gcagtcctat 12600

ttctgcagtg aaggttatga gaaaatagac ttggtctgtc ttgttcattg cagcaacctc 12660

agtgcaaaga acagagacct aggcaagtaa aggcactcag tgaatatgtg ttgagtaagt 12720

caataaataa aacagcaaga cagagtgcc aagacttgag ttgagggaga aacgaagcag 12780

aaaggaggaa gttgaattcc gatcgaagg ggcaaagatg atgaagtcag agaaattact 12840

cagatatattg gectcaacaa ttcataataag gttacttta agtcttatgt tctaggaatc 12900

tgggcagaag ggaaccagca cgatgcaaga atgcagaagc taaagaatgt caaccagtca 12960

gggtcttct tcccactgcg cgatctatct ttggcatctt taaaccattg aacagttata 13020

gtaggaaatt ccatctgaag ggatagcagt ttcactctaa ctaggtagaa cgggttgaac 13080

aggaaaactg gcattcacat tgagtaaaaa aaaaaaatac ctctcttcca tggatgacac 13140

tgcagcaggt cagaatgatt tggccatgat aaatctgaaa catattatct tggctcaggt 13200

tgagtaatag tcaaccctag aaaagcaagc atctcttctt ctcttataaa cattgcatca 13260

gggcaagagt gaggggtatt aagccatctt atttaatgat atatctcatc ctttaaggatt 13320

tattagtcc attgccaata atgtagctat atcagctctg catgtttgga aacttaattt 13380

gcaattaagt tcaaaggcca tttttcatg gtatgaaata ggattatatg tacttgaaag 13440

tcagaaatta accatttccc actacceaaa ataattttct gttcattaat tactgcaatc 13500

ttccattcat aaacatttat tgtgtactta ctatatttta ggcactttgt gatacaaaca 13560

actcttctcc aatagtttag tgggaaagag tcaaatatca taatactgtg taatttgata 13620

aacgctttaa gtcagcaata cagagaggtg atggctcatt cggtgaggg actggcaaaa 13680

ggtattgttt gaagttagtc cagaatgatg tgtaggaatt gctctggctg agaacgtaga 13740

[0030]

aaagatattt aatttaaagt agacaacagc aagtgcaaag gttggaaagg tgatcatgta	13800
ttttatagaa agcaatattg cacggctgaa gatcatggaa gccagagtag gtgagattat	13860
atagctagtc ggggaataaa taaagcaggg tcttgtatga ctcttgaaaa ctcttggaac	13920
cttatcttgg aggcgacgtg gaccatggca tttttgtggc ttgataggtc tctctcttgt	13980
gtttttaatt atactgaagt taccttagct taaaaaatg agggcattag catcatttct	14040
agccaccatg ctacacctc ttcttttctt agtgtcttcc ccaaatttca aaaagtgtct	14100
ttcatagttt tatattttac ttgaatcttc cttaacctca ttgaataatg ctgtgtccac	14160
tagccattct gcttgatcat ttcateccag tcaccatgtc atctcttaag ttctgtgtct	14220
aacagtcttt ctttacttcc cagggtccca gccctctgga aagcattgcc cactcttggc	14280
cagctcctgc cctctattc tggacactgc actgtgttca gtgttttata attcactggg	14340
attctcccc gtctcttctc tgctagtatt tctgctagtt ccaattctaa atctgtgcat	14400
aataactaag gtttagccct tagctgtctg actttttctt tctacagaaa aatctcattc	14460
attcattcat ttactaatc ctacacctt tattgagaac ctaatatgtg ctaggtctt	14520
cattaggagt tcagatagaa atatgtatct gagatattgc tttatgtgtc tgtctctaaa	14580
attctatct ttaataatga tacttctga aaattaaggt taaaaacctc tgattctcat	14640
aaaatggttt tatttcaact attcagacgc atctcacatg aaacatcaact acctctcaaa	14700
ctagtgttgt tatatttttc tttagatttc tctactttag aggatgggga tactgtcacc	14760
aattacacta aaaactgcag tgaaatcttc gactcttccc tttcttttga ttttgcattc	14820
aatcatttgc caaagtgtat ttctttttct ttgtccctt tccccctata tctactacca	14880
tcattcttta atctctaaa agaaaatggg cacagaactt ggaagtaaag gatccaggca	14940
aggcatttgc catgtggtac cagacgaggc tcaagttgga aatcttaatc tatgcctaca	15000
tatggtggca acaagctatc tctgatgagg tcagaccaca tatgatggac agctgcttat	15060

[0031]

atttcccatg ctctggagca gtataacatg ttagagcact ttaactataa acatatactt 15120

cttgcccttc attttatfff tgtctaatat tgtagtttca tcttaacacc ctgaaactac 15180

ttctactatt ttatatgtc aatgcttggt tattttttatc cacattttca tattttttctc 15240

tgccccctta ccattgccta ctttagatta ttctctcctt atggattcaa ttttcttctt 15300

cctggagtaa atccttggtg gttcttttag ccactattag tagcaagatt aatcttcac 15360

tgaaaatacc ttaattgegc ccttacgctt aaataatatt ttagctagat atagaattcc 15420

aggttggtgca atattgtctc aaatcacttt gaagatatta ttccactatt ttctgacact 15480

aatggttgag aaaggtgcta tcagcctttt tttttctggg gaggtaatct gtttcatatt 15540

ccgatacttt ttacaatfff cccttcagat ttgtgtttgt gagaattcat gtctttcatt 15600

catttctaaa aatttagagc caatacctct tcaaatattg ctacttatte attatttctt 15660

atattccttc tggaaatctc attagatgca tttaggactt ttgcattctg taatatggaa 15720

atcactgcat tctaagtcac ttcaacatat ctctcagttc ataatttctc ttttcaacta 15780

ttttaaatat attattgttt gccacttaag tgtaatttta cctcaatate tatgtaacct 15840

tcctttcaca ttaataaaac ttctaagttt taattgggtc tatgacttcc aagaataaac 15900

attccaggat cccttcgagc taagcatcgc catatgactc gttctgggca ataggatgtc 15960

aatgggagta tcttcaagac acgttttcaa agataaaacc cacatccttt ttttctttc 16020

cttcttcatt ccttctctg cttagaatgt atactgatgg ctggaactgc gtattgagcc 16080

aaaataacaa gtgtcaccca ctaggagttg aaatagaaat gtagaagaat aatcagttcc 16140

agatggctat ggcgccctaa tgccaatgct gaacagtcaa gatctgggat ttttatgtga 16200

gaaacaaatt aattttatgt gagaaacttc taacttctta atctatcttt ttttggttt 16260

tattctcagg cacatctaat tttaacctat accatttatt tattgatttt caaatgacaa 16320

tgaatatact ttttaattct tatggttcta tttggttctt tttttataga gtcgatatat 16380

[0032]

tttcttatgt cttaaattct tatttgtgtt taatcatttt aaacatactt ttaaagattt 16440

aatTTTTtag agcagtttta ggtttacagc aaaattgaga ggaaggcatg gaaatctccc 16500

atacaccttc tgccccaaga tatgcataga ctcttgate accaatatcc cccaccagag 16560

tggtagattt gttacaattg ccgagcctac actgatactt cataatcacc caaatttaca 16620

atatttatct gataattctt ttatctgaaa tttcttgggtg gtctaattct actgcttata 16680

gtattcattg attctccctg aagatgaaat cacaattcaa aattagaaaa catacaagaa 16740

aacaagccac aagagtteat cttcagttag actttatgtg tgggcatcct aaggggatta 16800

tgttgaagac aaatccctct agattgattt tgaatttatt tctgccaggt atctcaggta 16860

tctgtcatct gcctggaacc aatttttatg taattatcta gtttaggtat tctcagatcc 16920

catgtgtagt gtcttttga aagagtggca aatatatatg aggcacagac ccttggttat 16980

acattttaag ggcttatttt ctccccctct tataaaccca agctaaggca gacaagcttc 17040

tttgtgtctt gtgggtgtgc ttctgtaate cacttttcc agtgaggatg tttcctttat 17100

tgtctccaat ctgtctctct tctcttcaag acctgagatt ccaaacctt agaaaataac 17160

attttctact ttcagacagc tgccacatca acctatgggc cttttatttt ggttttggtt 17220

ctcaattctc tattcaatta tggaacttgg gaattttttt aactttcatg caagtttaga 17280

tacacattaa aaatatattt attacattgt atctaacaca tgagatggta atagtataaa 17340

gattttaaag ttttctaata agcaatactg tcagaattgg tgtagttct tagaaagttc 17400

cagatgatca cagctgtgct aaggaaaacg aagggaaaac atttctgcta ttcttagggg 17460

caagtatgcc ctggagactg ctgaatcagt gatatttgat gtgagagaac cttaaaaaaa 17520

aggcaaaatc caggagcagg ttgttgatc tcagccatt gctaaatatg gcaaatttg 17580

tcatacaact gctggactct atgctccttg aggacagtgt catgtcccat acatcttact 17640

atttctctgt ctacttgga ctgcagtttg tatatggaag ttgtcaaca caactgatga 17700

[0033]

ttgattgttg acagtacaaa tctttgcttt gaactttgta ctcatcagaa aattacatca 17760

agggtgtgtg catttcagtg taaagtacca tattaactgt gcactccacc cagtgtagaa 17820

acatacaaaa aaatgccagg gctcattttt gacattatit ccatccata actaatatta 17880

tagtcaataa taaagtagag atgattgatg tattctgtct tagttcaaaa tagattctaa 17940

aaaacttttc agggtaaaca gtcgaaagtt cagcacatat ttataatttg agaaggttat 18000

tagagactat gtagtcagat cactcattaa atgcagaaaa atgtcttct gtaataacct 18060

gattagtgtc aaacacagct agatacagga tgttagctca ctgcctcaca taatcattta 18120

tctcattttg aaaaatttaa tcttttcaat aaatttattg caattctaga aaactttact 18180

taaaaaatct ttgtattagg tttttgattt taaaatatac ttttctctta ttattgaata 18240

attctagaat gttcagggca taacattttt atatattcag gaaagagact atgatgacct 18300

ccattataga ggaaggagca ctgccctgtt taaaccagg gcatataaaa ctgaattgta 18360

gagaaataca aatatacata aacaaaaaaa taaaacaact tataattcta ttgtctacag 18420

ataatcattt ttacacctaa ttgtgcattc ctccagatct ttatgcaaac tttttttct 18480

tttaacaaaa ttgtcttcag gttttgaacg ctattttttc atctgaactt gtcaccaat 18540

tatgtatcat agcatctttg catgacaata cacattcatt catagcaaac tttttatgcc 18600

ttaatcattt tctattttat ggatagtcca ctatttattt aagcaactcc ttattgttgg 18660

gcattgtctt gctccaatt ctttactata attaaatatg ctgtaatgca tttattcatt 18720

cagcaaatat ttgtcagga ccttaactct gctgcacca ttcagcatgg gaaacgaagt 18780

tctgtcttc agtgtttaact tctgaagggg gagatggaga ttgaaaataa acaataatc 18840

agagcacaaa tatcagataa tcacaaagag tgataattgc catgaaaaaa ataagatggg 18900

ttaatgagat agtgacaagg ctacgtctgt aatttaggag gggaaataac gttgtgtcac 18960

atatccatag gataaacctt agatgaacct gtagaattgg aattgctgca ttttctatta 19020

[0034]

agctatectc ccactgtggt ggagaaggcc tgcttcttgg agtgcttacc aacactgtct 19080
 tgcaccttgc tcccactctg ggccctagtc ctgccctctg caccacctca gcaccatttc 19140
 ttctctattt gctcataaaa aatattgtcc ctcacatgta taggacttta tactttacaa 19200
 acaagtttca catgctaact taatgcatta attattttgt ttctatatatt tattacgtaa 19260
 atgttcaaatt ctgcacaaaa agagagggtg ggataatgat cctcaattc ctgagcacc 19320
 agatgaaaaa gaaatcagca ttgttcccat attgtttcat ttgttctctg gcccctatt 19380
 ctttttttct ggactatttt aaagcaaagc ccacatataa cattttcccc caccactaa 19440
 tactggagta catgtctgta attgacatag acatttttct ttcttgcata tcttcaatgc 19500
 cattaccatt tcttttttgg ttgttaaac ttttaggttc aggggtacat gtgccggtt 19560
 gttatataag taaactcatg tcatgggggt ttgttgtgca gattatttt caccaggtg 19620
 ctaagcctag tacccaatag ttacttttct tgctccttct cttcctccca cctcggccc 19680
 tcaagtaggc cccaatgttg gttggtcct tctttgtgct catgagttct catggttag 19740
 ctcccactta taagtgagaa catgtggaat ttggtttct gttcctgcat tagtttctg 19800
 aggataatgg ccaccagctc catccatgtt ctgtaaaaga cataatctca ttcttttta 19860
 tggtgcata gtattccatg gtgtacatgt accacatttt cttatccag tctatcattg 19920
 atgggcattt aggttgatc catgtcttgg ctattgtgaa tagtctgca atgaatatgt 19980
 gtgcattgtt ctttatcata gaattacata tattctctg ggtacttacc cagtaatggc 20040
 aaatggcagt tccgatttta ggtctttgag taatcatcac attattttcc acaatggtg 20100
 aactaattta catgcccacc aagggtatat aagtattccc tttctctgc agcctacca 20160
 gcatctgtta tggttttgac tttttaatga tagecattct gtctgggtga agatggtatc 20220
 tcattgcggg tttgcatttc tccaatgac agtgatactg aactttctt catatgctt 20280
 ttggccacat gtatgtcttc tttgaaaag tgtctgttca tgcctttgc ccactttta 20340

[0035]

atgggggtgtt tgcttttttc ttgtaaattt atttaagttc cttatagatg ctggatttta 20400

gacctttgtc aggtggacag ttgcaaata ttctctccca ttctgtaggt tatctgttta 20460

ctctgttgat agtttctttt gctatgcaga agctcttttag tttaatcaga tcccatttat 20520

caatTTTTgc ttttgtgaa attgcttttg gcatcatgaa atctttgcct gtttctacgt 20580

cctgaatggg attgcctagg ttgtcttctg gggtttttat agtttagggg tttacattta 20640

aatctttaat ccactctgaa ttgagttttg tatatgggtg gaggaagggg tcccatttta 20700

atcttttgca tatgggtaac cgtttactga atagagagtc ctttcccat tgettgtttt 20760

tgtcagcttt gtcaaagatc agatgatcat aggtttgtgg ccttatttct tggtcccta 20820

ttctgtttca ttggtctatg tgtctgtttt tgtaccagta ctgtctgtt ttggttactg 20880

tagccctgta gtgtagttg aagtcaggta acatgatgcc tccagctttg ttctttttgc 20940

ttaggettgc tttgctatt tgggtccat atgaatttta aaatagttct ttctggttct 21000

gtgaaaaatg ccattggtag ttgatagga atagcattga atctgtaaat ttctttggga 21060

agtatggaca ttctaatac attgattctt tctatccatg agcatggaat ttttttttaa 21120

tttgtttgtg taacctctga ttctctgag aagtgttttg taactcttct thtagagctc 21180

ttcacctcc gtggttagct gtattcctag atattttact ctttgtgtgg tagttgtgaa 21240

tgggattgtg ttctgattt ggtctttggc ttgactgttg ttggtgtata gtagtgctag 21300

tgatttttgt acattgattt tgtatctga aactttgtg acattgttta tcagctgaag 21360

gagcttttga gccaaagacca tgggattttc cagatatgga atcacgtaat ctgtaaacag 21420

ggatagtttg acttctctc ttctatttt gatgccctt atttctttat ctccctgat 21480

tgtcttggcc aggacttcac atactgtgtt gaatagggat ggtgagagag ggcatccttg 21540

acttgtcca cttttaaggg gagtgttcc atcttttccc aattcagtat gatgttggt 21600

atgagtttgt catagatggg tcttgttatt ttgacatatg ttccttccat atctagtta 21660

[0036]

ttgagagttt ttaacatgaa ggaatgttgc accttatcaa aagcctttgc tacatttatt	21720
gagataatca tgtgtttttt gccttttagtt ctgtttatgt gatgaatcac atttattgtt	21780
ttgtgtatgt tgaaccaacc ttgcatccta gggatgaagc ctgcttgatc atggtggatt	21840
agctttttga tgtgctgctg gattcagttt gcaaatgttt tgttgatttt tgcacaaatt	21900
ttcatcaagg attttggcct gcagttttct tttttgttg tgtctctggc aggttttggc	21960
atcaggatga tgctagcctc atagaatgag ttgggaggaa tccctcctcc ttgaattttt	22020
gcgaatagtt ttagtaaaag tagcactagt tcttattttt acatctggta gaattcagct	22080
gtgaatctat ctggctctgg gcttttttga ttggcatgct atttgttact ggttcaattt	22140
ttgagctcaa tattggtctg cttagggcat caatttcttt cttcaattca gtcttgagag	22200
tgtgtgtgtc caggagtta cccatctcat ctaggcttcc tagttcatgt gcacagaggt	22260
gttcatagta gtttccagtg gttattttta tttctttggg gtcagtggta acatccctt	22320
tgtcatttct aattgtgttt acttggtctc ctctcttttc ttctttatta atctagttag	22380
tggcctatth atcttattaa ctttttcaaa aaaccacctc ctggatttgt tgatcttttg	22440
aatggctttt catgtcttga tttccttcag ttcagctctg attttggtaa tgtcttgtct	22500
tctgctagct ttgggttggg ttgctctgc ttctctaatt ctttcagttg tgacgttagg	22560
ttgttcattt gagatctttt taactttttg atgtgggcat ttagtgcctt taatttcctt	22620
gttaacactg ccttcattat gtccagggg ttctgggatg ttgtaacttt gttttcatta	22680
atttcaaaga tcttcttgat ttctgctta atttcattat ttactcaaaa gtcattcagg	22740
agcacgttgt ttaatttcca tgtaattgca tggttttgag agttttttta gtcttctatt	22800
tttaatgcgc tgtggcttga gagtgtgttt ggtataatth caattatatg atcagcttta	22860
gagtatgtac catgtggcaa tgaaaaaat gtatattctg ttgttttgtg gtggagagtt	22920
ctgtagaggt ctattatctc tatttgggtgc aattttgagt tcagatcctg aatatcttca	22980

[0037]

ttaatTTTTct tCctcggtta tccatctaataat aatgtcagtg gagtggttgaa gtctcccact 23040

attatTTTTgt gggagtcetaa gtctctttgt aggtctctaa gagcttgctt tatgaatctg 23100

ggtgctcctg tgttgggtgc atatatattt aggacagtta ggtcttcttg ttgaattgaa 23160

ccctttaccg ttatgcaatg cccttgtctt ttttcatctt tgttggttta cagtctcttt 23220

tgtctgaaat taggattgca acttctgctt gttatgattt ccatttgctt cgtaaatttc 23280

ctccatecct ttatTTtgag cctagggatg tcattacatg tgagacgagt ctctaaagac 23340

agcattccat tgagtcttgc ttttttattc agcttgcctt tctgtgcctt ttaagtggag 23400

catttagcaa atttatattc aaggttagta ttgatatgtg tagatttgat cctgtcattg 23460

tcttgttagt tggttattac gctggcttgt tgtgtggttg ctttatagtg ttactcatct 23520

gtgtatttaa atgtgcatta ccatttctaa ccgacctaac agtacttctt taatatcaca 23580

taacacatag ttcttgTTta catttccaca attatctcaa aaatgcagct tttgttgggg 23640

cgatcagacc caacaccagg tcgtgggggt gatgaagtc ccgaggagtc aaggaattag 23700

aaaaagacag tttgagaggg aaagtgggac cagggggcca ttgcaagtgt agaggctgcg 23760

aaggccctga gctctgggag ccatacaat ttattgggtg tcaaacaaaa aaacaggtgg 23820

tgaggatggg ggttgaaatg aaacagtgt tcaagtgaat aagaaacata tggctgcttg 23880

agataacggc agtgctagaa acaaggagcc agcaagtctg gcagacatgc tagccctgcc 23940

tcagcttctc tcccaacct cagctttact cccaacagct ttcaagttga tttctttgag 24000

ttaggatgca aacgaaatcc acacatttag ttggtttgct tcttaagttt cttctaattct 24060

gtagcaacc ctgtactctt tcacatcatt tacttttta ataaataggt tgtgcctcct 24120

gaaaaatttc tcacattttg agtttgctg attgcttctt tcttgtgtca cttgactttc 24180

ttgttatgtt tctatatccc ccagtatttc ttataaattg gcaattagat ttagaagctt 24240

aaactaaatt caagttcaat tttttttgtg ataagcataa atcggatgtg atgtgcactt 24300

[0038]

cttattatac catagaggag gcacatactg tctaggactt ctaagtttag tgatgttaag 24360

attcaactgc agatttaggt gtcagcacca tccatttate cattttaagg ttctccacaa 24420

tcctttcacc caatgatata agcacctctt ggttattggt tctagacca ctttttttc 24480

attatgggtt aataaacagt aactttctct cctctgtctt catttattaa ctggattttt 24540

taataaaaaa taactttect catcaattat ctggtaatcc aaattacagt tttataggaa 24600

atatagagta aatgcaaaac tctctctctt gccagtttcc agaataataa gttggtggcc 24660

caacaatttc caaagataac tctttttatt tttttaaaat agtatcatta tagattcatg 24720

ggtttctctt tttctttaat gcattctgat caattgcaat gattatttct ctcatttttt 24780

agacactggg agttctttaa aattgcctcc tgtgtttttt ttttttgac acaaccccca 24840

gagtctacca tagtttgctt gctttctggg acaataagat atcctgggct aattctgtat 24900

attacctgcc tcagatctag acccaggtat ttctccagtg gtttcttgcc tttttttttt 24960

attttttttt tttagtggta aatggtatgt agagacctta ctctagacac tatggagaca 25020

aaatgtgact gcgtttttat tgttgcgaga ttttttttgt gagcagttag aaaacacatt 25080

ttaaaagaat gaattaataa ccaattcaaa tttaacattt cagaataagc acttaacctc 25140

tttggttcta tacttgcaat tttttttctt atgctgaaaa ttttgtttcc tctttgtttt 25200

atcctacaat atggetatgt tttcttccaa gtgataatgc gaacattatt attatcagta 25260

agatcattca atgctattta cattttcttt gagaactcatt tgtccttgga atagaaatgt 25320

ggagtaaaaa tgtgctttag aattacttgc agtaattctt ttctgtatat agttatgcca 25380

ccaacctga tatgcatttg gacttatttt ggtttgtttt tactgtttgt agatttttat 25440

ttaataattt tgtttttaat tatgtatata acatttacet taatccaaag tcaaaactat 25500

caaacaagat gtattcaaca aagtctagct tacatctttg tcccccttct tcttcccttc 25560

tgttccataa gtaatcattt gtattagttt ttatgttttag cccacatttt ctttaaggca 25620

[0039]

ttatctctta actgtatfff ttaacactff tatgtaggca agagggatga tgcttaccta	25680
catcataaag gaagtagcac tgcctatfff aaacccccgc ccaggaatc caaatcctgt	25740
gtccttcct ctgtcccatg ctgccctcta cagacacctc ggggaatttc aaggagaggc	25800
agetgtcagt tcttcctgag gttagatgtc caaattaagc agactctttg tgcttcagt	25860
atactacctt ctactctaga attttgtaaa aatacgtatg cgaacctfff tggggctcca	25920
ggcagtgcct ggtgtccagg cgaaaatgtg tgaataagca agcactcgtg tgtgttgga	25980
aaacaaatta tccctaaatg ctgtcattff ctcttaacac cagctgcgat gatgtggaat	26040
tgaggtaaaa ttaaaacca tgaaatgttc tgcaagtta aaatgacct aggtttaaat	26100
atftftftft tctatctaat tcaatgtaat tttaaattgt cttgacaact aggcagaaaa	26160
tagaaactac aactttccta aagaaaattc ctggccgggt gcagtggctc acgcctgtaa	26220
tcccagcact ttgggaggct gaggcaggcg gatcacctga agtcaggagt ttgagaccag	26280
cctggccaac atggcgaaac cccgtctcta ctgaagtaca caaagtagcc ggctgtgttg	26340
ctggtgcctg taatcccgac tactcaggaa gttgaggcag gagaattgct tgaacccggg	26400
aggcggaggt tgcagtgagc cgagactggg ccaactgcact ccagcctggg cgacaagaga	26460
ctctgtctca aaaaaaaaaa aaaaaaaaaa aaaaaaaga aaaaaagaa agaaaattcc	26520
taactgggcg agtgcataag taagacagag ttaaaatcct acggtactgg gtattgttta	26580
ctcccaaagg gttttctcca agtcttctcc gctcccgctg ggcctctgc cggggcgga	26640
ggagcaaagg tggcgcggtt ctgtacacag acatgttcta gcgttctagc gteccagtc	26700
tcttcctct tegtctgct cctgaggca cctagaagca aaagaagggt ttgttgtgt	26760
tgttttgttt ttgtttgttt tgftttttga gaaagggcg cgaagcaggg tgcaatgga	26820
tccctaaatt caggctgtgg ctggctttcc cagcatccag tttctgctta gcaaagatcc	26880
aagagagctt ccccgtaact ttggatfff tgtctgagat aagggttcaa aagcagagca	26940

[0040]

ttttaatgtc atagatagtt cagatggata cattagaatt aagttatttg aaaaacaaac 27000

aaatgataca ggagctataa agaaattatt taggcagtta gtgagtgtaa gagagtcctc 27060

cgtaaggctt cccttttaac aaaaagcagc ctccaaataa tttcttttcc agcaaagaac 27120

agcttgtaaa actgagctgc agacctaaac aagcaagctg gaagcttgca tagttagatg 27180

ctgttagctg tgctgataga aaaaaggcta cctgggggcc aggcattgtc aacatggagg 27240

atccctcttc tcttttcttt gtcgccactt gggcagtaaa aaggcaggca acatggtccg 27300

gactggtaga gaccccatct gcataataaa atattagcgg gggatggcca gcttcttcac 27360

actctataca aacagcacac ctggctcgag caatctctca ggccgtatgt taatcagaca 27420

ctgcctcttc aagcttgctt ataaaaccct gtgcgtttca ccaggaaacc cgaagacca 27480

cttgggcacc cgtctctctc tgcaggagac agagctatc tcttttctct ttctttcacc 27540

tattaaacct ccactcttaa actcacttct tgtgtgtccg catcctcgat ttccttagca 27600

tgagacaatg aacctctggt atttacccta gactatgagg cctcttcaca aacaaacaaa 27660

cagaaaaacc aataacaaca gtaaaagtgc taaaatgcc gtaggaagtg gtgaaagcgg 27720

tgctgcaagt ggggtggggac cagaaacatc aactttgaga tgagggcagg cagggtagct 27780

ttaatttggt ccagtgagg caggccccct caggaggagc tgccaccggt ttggggagcc 27840

cacaggtcct cctcccagg ctggggcagg acccaggcag gttgcgcacc tgagcgacgg 27900

cccttgctct cttattttct gggttaagggg ctgccctttt ccgcacacac agtggtcagg 27960

agaggcagtc actgaattga ggtgaccagg gaagagcagg gccatccgtt tcttcggaaa 28020

aaaaaaatca gagtttttaa atcagagaca tgcccagggc tgaagcctga agctgtttct 28080

tcctctgtgc cagataacct caggatccat aattcttttg aatctcacc tgccgaatta 28140

ctttctttca aaaataaaat tattttctga taattacggt aatgtacatt ccctaaattt 28200

ttgaaaacca tggaacatg aaaaaggtgt gtgtgtacac acaatggaga cacactttac 28260

[0041]

attctgtaaa tgctgctttt tcacatcaca tgatgacagt ttgtccactt ttaaaagtgg	28320
ttgtaatgac tgaacaatat ttaccatt ttagggatgc acagttcttt agttaactat	28380
tcccttcgtg ttgaacatta aggcagttct acttcattgc tcctgtaaag aatgctcctt	28440
cacagaatac cctgggtctc agcttcaca ctggaggaga ggaagttgct gggaacatca	28500
gactgggaca ggtcacccat cctttgtttc ttccagtatt gtcctctgc aaaccttagg	28560
gccaccatca gcagcagcct ccaaaccagt attgtggggg aggggtgtca ctacataagg	28620
atatctggcc tatggggaaa gtggggactg taatccagct tgaatgttgc tggcctacct	28680
gtgtgtaagc ataagtctgt gtccattctt ctaattcact caaaggtgcc acatggatta	28740
gacgaacgag tggctcctgg agcaactggg gcctgagggg gtgcatacat agaaatggaa	28800
tactacgatg tcgttttgat tttttagttt tatttttagga atatcctggc ttagatttgg	28860
ctatgatttg tgctttactc tcatttcaac actttaccat cagtgggaaa tatcactgtt	28920
gctttatgtg caaagtgatg gttttaataa ggggtaaaga ggcatagtta gttgaccact	28980
gcactttcat ctgggacagc acctttctta cccaacagg atcatattga agctatgact	29040
tacatggttt gggatcttag ttgtatacat ttcaaaaagc gatgttaaac acattgagct	29100
aggetgaaaa aagtaattta gaaagaattt tacttcttgg ttcgaagtaa atcagtatag	29160
catggtggtt gtgttcatgg tctctgcac cagactgtg tatgacctg gcaagtgaag	29220
gcaacatgtg atgattatgt ttccataact gtaacatgaa agaactaagag tacctacccc	29280
aaagggttgt tgggaagatt aattgagaaa ctgtaggtaa gacacgtaga acaatgcctg	29340
gtcatcata agtgctaagt aaatgttaga ttttatgatt atatccattt ataaaaccag	29400
cagcatttct catttccact tttttcttct ttccagtctt cttacataag atactgttc	29460
atagaatccc agtctcattt ttctccctgc atccacaacc tcttcatgtt taactctccc	29520
taaggcactc cagatattta ctgaaaaaat tgagttaata gttattcctg ggttcttct	29580

[0042]

ttcccaccct ggctctctcag caacttttct tctagttttt ggattctttg cctccccatt 29640

ctctgtctta taggtgggga tcttgaata tctttcataa cctgatggta gcagttgcct 29700

tcaaccagag cctgggcaact acattacagg ggaaaatcag gcctcattgc tgtggctatg 29760

aactggtacc cctcctcttc ctgcccacac gtaatcgtec tacaggcccc ttcagagatg 29820

ttagaacttc tacttttctg gtctggagtt gagctctgtt tcttgtgcca ggcagtgtctg 29880

gctggagtag cagaggggag gagacactca cagagcagcc tgggctgtag caaaggcaga 29940

gagtaccacg tgggtctgat aaaacaaagg cttttatttt tcatagggtc agacacaggt 30000

gatgacaact ggctgtgcag atattgcaca cactggcaga taccagggtt gatatcctgt 30060

gggatcctgg agaatggaca ggaagagcca ctagggatac ctcagagctc tctgcggggg 30120

aacttctcag tgctgaccaa aggaaaggag atttgagtat gagagagcaa agcagaccag 30180

gatgtctctt ggactttctt tctctaaag ttgcttattt cagcataaag atccaggtea 30240

gtggcagcct gagccccac ctgtctgacc actgcccctg ccacagaagt tggagctgtg 30300

gcacttgcat tggcgggacc tgtggcacct gtggtggctc agggagcaca gacactttgt 30360

gtcccccaa agtgctctg tcaactactc tctcctctt gagatcaggt gtccacagtc 30420

ctcacacaca caggatttgt ctgtgatgtc ccctcaaaga ggggacacaa agtgttggaa 30480

aacacagtgc tattttcccc agttttgctg gtgatggagg ccaagagaat gtcagctggg 30540

cagaaagtca gtctaataa gctgtgtcat tgtgggctc tgggccttct aaagggtcca 30600

cctcagtcct tagtccatgg ctgtagatgg agacaccacc ttctgtgttt ctttateatt 30660

tggcttcagc cctcctccct tcctaattgc cagcaccage cctcctcact ccatgaaatt 30720

cctcctactt caagtccttc attcccatat tctttcatta ataaatttct attgagactt 30780

agcaatgttt cccatccaaa gagcaccaga aaagtaacta tagtagaatg actttatgtt 30840

tgttgacttg ttgcaatgat ggagatagga catgatatgg aagcattaag aggtgtcac 30900

[0043]

tgtaggattt	gggcttgtgt	gttaggtagt	ttgggggagt	gttgacggaa	ttaaagtttg	30960
ttctggattt	gatgctgtca	gaaagcaagg	acacttcaat	aagcaggtat	cttaattttt	31020
aactaggaag	ttggagcaat	gaggcaaggt	taaagctata	actagtaaaa	aggtagaagt	31080
cattcatatg	agctgagaca	agtgaatgtt	tggtcatttt	tgtggtttgg	gcaatgcttt	31140
tgettcatg	atgctcagag	tggtcttgtt	tttgtcttgc	tccctcgagg	tcacagagtt	31200
agcttatcta	aggtttagagt	tctgtaaaag	tatttatatt	caacaggaga	acatcagatc	31260
cttgctacta	gagcgaggcc	agcttgctgc	cagaaatctg	ggctactect	tcttttctca	31320
gcatcaatta	tgggccagge	atcatggtea	gtagtagtga	tagggtaggt	aagagaaaaat	31380
catggcatct	gtcttcaatg	agcttaaagt	ctagcagatc	taacaaacat	aaaagagtaa	31440
attgttaggg	ataatggacc	ccagaaggag	ttccatgtgc	tatgggcatg	tagagaaaag	31500
gtattgaagt	ctatactttg	ttgcaaactg	aatatttatt	tttcatatc	ttgcagttaa	31560
cataagcaag	ccttatcttc	cagacagggc	aaacatggca	gccattttcc	tccatggcag	31620
ccattttect	ccatagattg	caaccctgag	ccgttggcag	caggggagge	aagaggtagg	31680
agccagaatt	aaaaaggaaa	ataaatcaca	cattttgggc	aagtttcaaa	gaatcaaaat	31740
atcttttcca	atgactttca	gtacatcgtg	ttacctctgt	tttatgaaac	acagtggatg	31800
actattagge	ttttctccag	ccccattctt	gactacctgt	cttgaaagcc	tccccacaag	31860
acctggagag	ctcatthaate	tcctttctac	cttagtgatt	gcccttgtea	ctcactctcc	31920
tcctctgag	atcaggagtc	cacagtcctc	atatgcacag	gatttgteet	gtgatgacac	31980
atggtagcta	taagaccatg	atctggteac	ttcccttaca	gcacaagggc	acctccctga	32040
gagtcactag	tcccaggage	cccatcctga	ggagacttgg	cttctgttgg	tcccagata	32100
tcaaggaaat	acagagactg	ctctggctgg	atgaatccag	gaaaaatttc	atgacttagt	32160
cattccaatt	ggtactatag	gaacaggcac	tgcaggacct	caggtttage	tctttggagt	32220

[0044]

acatttttga cattccacat cagcaatcaa agatgagtaa acaaacagaa ttctttctccc	32280
tttgcttaca actcacagtt ccttctccat tctgttcttg ccatgactgt gacagtcaca	32340
ttccatggca agcagctgag tcattcttaa ggaaaatggg actcggtagc cagaaagtgc	32400
catatacctt caatgtgcct aagtttctta tatggattct aattttcccc tcaggttaag	32460
tagtattttg agctatttat acctctttag tctctgaga aagttctggt tggaccacgt	32520
gtgtaaaaga atgttggact cgaatttttt tcttggttg ggatgttggc tggagttgaa	32580
agataaaaat aatatgtgcc cttatcccc ttatactaaa gtaacctcag gataagctgc	32640
tgaaaatctc tgcaggactt gctaagaaaa aacataactc tatttctctg gtccagttat	32700
atatctggaa gtttgagttt tgataaaaat tttttaggta ggagagttat tgcctagaa	32760
gacagtacta tggtttcacc cttatcttgt ataggaaggg aactggttgc tctcttgtaa	32820
cagagatcaa tggaaagggg aggactcatg attcgaaatt gtagatgta gaggtccctt	32880
caaaaaagga agacaggatt ttcattccagc agcagaagtt ctgctgcagt agtagatttg	32940
attgtttaac ttctgagcct attgtttaat aaagaaaagg aattaaagag atagagcagg	33000
atgttgagac ataatagtag acaacatctg ggaaggaaga ccagcttgca cattcttctt	33060
agtgttgtgt ttagaacggt tttctgttga gcagtgagac ctcatcatca agaagctgtg	33120
gctgggtgtca atgaagggtg gaacttaagg agtatctgaa gggttaaatc tgactgtctc	33180
tcccatgtct gggagtagcg ctgtggggag aaccttgca gatgtccag aaggactcac	33240
aaaaacaccc tataagacag ccctctgca gacctgcca cacacagga gcaagcaacc	33300
aacagaggag gctagtggta cagtcacatg tgtaatgatt tcccctcccc tttccatct	33360
cccaatacaa tggaggagga agtgttatc aggaagaagg tggagagtg gtagtgaatg	33420
agacaaagag ccatacatgt tccctcccaa ctccagtctc ttgggagagc catgggcagt	33480
gtcactggg aagggtgaat ttgagtttga agatgcgtac cgaattaaat tcagctggtc	33540

[0045]

tggaactttaa tatccataaa attcactcca gaaatgtgga atttcctga gatattttca	33600
ggagagattc aacacagcat aggagaagat aatgatggga taagacaatg cttcttgttt	33660
ttatgagacc aagtttataa ggtttaatat actattttta aggttctaac aatgtcttta	33720
taacaaggga tcacttttgg gattctccac cttttctaac atcaatgaaa tcaataaatg	33780
ccaagattcc atgtcacaaa ggacaaaaaa aagcaggttc agggatgttg gttgagggt	33840
tcaaaatgac cacatgtggg taatgcaaag gcaataggaa gtagaaactg ggaagttctt	33900
gagggtggga gacatgttgg gatthttgggg gccagatagg aagtcttcta gggagaggaa	33960
tatcttaggt caagaggtgg aggagaggaa gttgggtgct ggatggtgag aatactgtgt	34020
gtactcacc tgtagtccat ggggagaaag aataaaggaa aacttttggg gagtcagtag	34080
aagaccctga tcaatctgac cettgccaaag ttctttttat tcaacatctg ccactctctt	34140
ccccaattct tggctacaat tcttttcttg taccatgttt tctcccaaca ttctgtctat	34200
cctgaatgta aggagaggag tgaatgtact gagacataaa agatctgtgc tagatcccag	34260
gcaggtgtat cacgttagtg actgatactt ggtactcaag aaataaactt tggctctgac	34320
agagttctcc tctccagggg gcatggaata ggtttattta ttgtggaatg gatagatgca	34380
gggtagttgg aagagaaaat aaatgggagc attgaaagaa cccaagaaca ttgacagaa	34440
tctggagtgg gcaccacaat attgtgttgg ctggcattac acctgggcaa tctgtctgag	34500
cagagtcttt ctggaaatta ggccaaagat ggggttatta ataccagcac aactgagtg	34560
ctggaatcca gatatttate aatgcaagtt tttatttaaa gattaggatt tatggtttat	34620
atgtgatttc caacttaaat ggtaaattct gtaagggaat cccacacct accatcttcc	34680
ctaatgttcc atggaacttc tccaccagg agatacagag gtgatggagg aggaaaagca	34740
tcacactgga gagaggttc tcttttgctt cactggatca tccagccctg ctctctgtg	34800
gttttcagag ccctgggtgc tgcattgagt acagcaccca gaactctgat gctgtgctta	34860

[0046]

gttggaggtg tgaagcttgt gatggctcag gcaggagcgt tctccagagc tgaggctatg 34920
 gcagcccca gagctggggc tccatgggaa gagactgcag tgggcactga agaccatgtt 34980
 tgggtgggtca ctttgcccc aaagtcattt gggagaggct tgaaactctt cctcgtttta 35040
 ctccaggac accttagggg aagtttagtg actctagaaa aagtaaaaag gggaataatg 35100
 gcatgtaatt gtatcagttt taaaaattgg ctattttctt atttacttca gagaatgtgg 35160
 ctttgaatcc acaacttagg gaaaattcct gaggagatga gagggggaaa tggaaagtta 35220
 acttgggtta aagggtcttc atccagttag gactgactat accttgagc tgttttctta 35280
 tttgttttg caggcaagaa atgagtgggt ggattgggag gggtaggagt gatgagtctt 35340
 gtttaggagt cagcttcttc cttctgttag cccaagccc tctgcatctc catggcaggt 35400
 gtcaggagtt atgaggagag cagggatcca agaaactcca atatccaaga tattcatata 35460
 tttctgctc aatctttctg taggtcttta gctctttatt tttcatttgt tatcatcatc 35520
 cgttatttaa acaatttata ttgagatgat ggtgatacga gaggcgggca gggaagtgtc 35580
 gggtagagaa gagggggtc cctggtgagg gctccacct caggcctatg cccatggacc 35640
 taagtgagga caggcactcc tggttttgca cccaaatgtt gcattttcta agccactct 35700
 gaccaccat gcccacatc ctgggcccac aaaatcttca agaccctagt gggcatggac 35760
 acaagtggct ggacgtggag aggagcagaa gacacactg actgatacca gcagacgcca 35820
 gcaggccatc aacagcagga cctcatggaa ttcagtgggg gcagttggag gccagctggc 35880
 cactgggtgg tccgaactca gtgaagcct atcttccac cccatccct tgetgcccc 35940
 cccatccatc tactgagag ctacttccac cactcaataa aaccttgcag tcattctcca 36000
 agcccatatg taatctgatt ttccagtac actagacct acagaatatg ggatacagaa 36060
 agctctctgt ccttgtgata agtcagaggg tctaactgag ctgattaaca caagctgcct 36120
 gcagacegca aagctgaaag agcacactgt aacacatgcc tactggggct ttgggagctg 36180

[0047]

taaacactca accctagaca ctgccatggg gtcacagccc aaaaatgctc ccaggacctg	36240
cccatttgca tgctccccct agcggttcga gcagcaggac atcaaagaag cgagtcacac	36300
ccctgttgca tgccctgcaa atgggataag ggaactcctc ccatttcaat ggcacactct	36360
gtgctgggca aggcatattc aagagtaaag atactatccc tacccttagg ctgatgaggt	36420
agtggggctg agagataagt aggaactaa gtctgagcca agtgagggtg ctgagataca	36480
cgtgtgtggg agtcctcagg cctgaggtea gagttgttgg aggttagtcc cttgttctc	36540
tgcagtgact tgtagtgaat gatgaataca gatggtgtca cctgctgtca cttgggatgg	36600
ttgacggggg gagtaagatg tggatcctgt gtccgtccaa ggacagtaaa ctcaggcatg	36660
gcaactaaga ggaacttgtg gacccaagaa gcatgaaggg ggtgatgagg aggaagagat	36720
aatggtgggg tgaggaagga tggggctgac agaaacattg ggacatctac aatcttgagg	36780
acaagatgat atgctcattc agagcatctg gtttggagac aggaacttcc tagtttaatc	36840
tttggccact tctcctaata gctgtgtcaa ggtagagatg cccaaggcag aaaaatctat	36900
atggagacaa gatgatccag ggatgtccca ggtaaaacca gagacagga taagaaatct	36960
aaaatggctg tgatggtaat gtgtcacaga ttaacataca ttaagacat tatgtcagga	37020
tcttagtgac tatagtattg attctaagag agccaggaaa gagagtgatt caaatggctg	37080
ggatggctaa gagagattag tccaacgatt aatggtaatt aatgggtgtg ttataatgag	37140
gttgatggga taaacgtgaa cacctgacat aaagtagact aagtttgcag agacagacct	37200
aaactctata ctattttcca tactatttat tttatatata tatacacaca catatatatg	37260
tatacacaca cacacacaca cacacacaca ttctatcctg aatcatgata tattaaaaat	37320
aagtaaaaat gaggccagga ttgggtggctc acatctgtga ttctgtgct ttgggaggct	37380
gaggtaggag gatcccttga gtctgggagt ttgaggccag cctgggcaac acaggaagac	37440
ctcatctcta ccaaaaaaaaa aggaaaaaag aaagtaagaa aaatatttaa acttcaata	37500

[0048]

tatgctcaaa	tcgcaagatt	ttgataaaat	attcttattt	attttggttt	aaaatggttg	37560
actagaagca	gctagtgtgc	actgctctta	tggagaggat	acaaagcagt	gactaaatac	37620
tagctcttca	agtgaattgt	caaagagatc	atattgggat	ttaccaagga	attgacagga	37680
cccacagaga	gcagagagga	gcaaagctgg	gcagccattc	atctgagact	ggcacagagc	37740
tgggggaagc	tcctaaagc	agagaaagag	tgagtgagag	tccccgggga	ctcacacttc	37800
tgccatgaac	ctttgcaata	ctaggcatag	gagagccctt	ctgactcett	gagcctccag	37860
accaacacag	agagcctgtg	cagaggcacc	agttaagtec	acttgagacc	ccataggcct	37920
tgtatcccta	acagccagtg	ccagctgctt	taggccagcc	aaaagggggc	ctgggctctt	37980
ttgcacaccc	ctgggataac	tgtgtgtgt	gcatggctga	ggagtggcca	ggattcagac	38040
tccatccctt	gctgtactt	aacagataag	gcctgcatgc	ctgggctccc	cgagcagtgg	38100
cctcaccccc	acctgaacag	tgcagttagt	catggtccta	tactcctctg	cagcaaaact	38160
cccagagtaa	cagacaaggc	ttggcacttt	tgcatacccc	caaccacaaa	atccagtggc	38220
agcaatggag	actggtggct	atgggggcac	atagcatgct	gaaaatgggt	ccccagtctc	38280
ttccagcttg	taaggtttct	gatgagtagt	ccactgttag	cctgatgggc	ttccctttgt	38340
acgtgatttg	acctttttct	ctaggagcct	ttaagtctga	ctatatgcct	gggtgatgtt	38400
catttgtata	tgcatacatt	ggtgggagag	aagtgtgagc	atgccatgtg	cccccatagc	38460
caccagtctc	cattgccacc	acggagaggc	cctgccctcc	ccagtgaaag	gcccacagca	38520
cagccacctg	ccttgcctga	acatttcagc	tacagcccag	agcccttcta	aaaactgaac	38580
cctcacaggc	ctataatact	tcctaagget	cttaccaccc	taagcattct	gtctgccact	38640
gcctaagagt	ttggcccatg	accagcccca	gcccttccca	tcacagccaa	cacctgaact	38700
gggctagcct	gaccttgtc	caaccctttt	aggactcata	cgcactgtcc	agggggccat	38760
gtaggaactg	gggaactccc	caccccatca	caactctgct	ggcacctgac	tacttgcttc	38820

[0049]

atggcctgaa gttgggctga cetaaccaac caacaccacc acaaccagta ccactcaca	38880
cagcccegggt gatgtagccc ctctacttat aagaagtggc tgtattgcct cattagagaa	38940
caggtgagca ataaagctat ctgttttagg ctgagtgaca acattatgct ctgaaaacac	39000
tcccatggag tatcacaaaa catgcatttc ccatggetgt caaccacatt gtggtcaaga	39060
gacagactac agtgtgettc tgaactagga gtcattgagtt ctgggtcaag ggtatgatag	39120
ggaaacagat caacttcctg cctacctaag atgaggaagc agtgcagett ccttgtctec	39180
ccacggagac ttcagtgtgc ttcaccagaa gttcctgtca gccacccttt tcagggett	39240
tgcctgtget tgccattgga gtattcatgg gcaagccagg gaetccagtt ctgcccagtg	39300
gtgtgcccct cccctgtgg agcaggaagc tcagggccct ggggtgetca ctgtccatcc	39360
tgtcacctga aataacaaag agcccttcac agtaaacaaa aattataccc acctgcttgt	39420
gtgcagctg gatcttacct gcaagcccta tctgtggcc ttaggttga actgcacagc	39480
ccaatataaa acctactaac ataagtgcac agagctataa aaacaaagcc aaaagacttt	39540
ttccaatata ctttccagtc ataccctaaa gggaaagaaa gaatatatag gaaaagagaa	39600
aaaaagtcta ttttcatgaa aataattaca ataattagaa gtgccagcct ttctagatga	39660
gaacaaactg gtataggatc tggcaccatg aaaaatctga atgttgacac accaccaaag	39720
tatcacacta gttctacagc aatggtcctt aaccaaatgg aaaaacagaa atgacagatt	39780
aagaattcaa aacatggatt tcaaggaagt tcaacaatat ctaagatgat gttgaaaaac	39840
aacacaaaga aaattctaaa gcaatgcagg aaacaaagga agagataaac atcttaaaaa	39900
taaatcaatt ggaggtaatg aaatggaaaa actcacttag gaatttcaaa ataaagtga	39960
aagctttatc aacagactca accaaacaga ataaataatt taatagcttg aagactggtc	40020
tttctttttt gtttttttaa ttatacttta agtttttagg tacatgtgca caacgtgcag	40080
gtttgtgaca tatatatata tacatatgcc atgttggtgt gctgcacca ttaactcgtc	40140

[0050]

atgtaacatt gggatatatct ccgagtgcta tccctccccc gtccccccac cccacaacag 40200

gtccccgtgt gtgatgttcc ccttcctgtg tccatgtgtt ctcatgttgc aattcccacc 40260

tatgagtgag aacatgcggt gtttcgtttt ttgtccttgc aatagtttgc tgagaatgat 40320

ggtttccagc ttcatecatg tccctacaaa cgacatgaac tcatcatttt ttacggctgc 40380

gtagtattcc atgggtgtata tgtgccacat tttcttaate cagtctatca ttgttggaca 40440

tttggettgg ttccaagtct ttgctattgt gaatagtget gcaataaaca tacgtgtgeg 40500

tgtgtcttta tagcagcatg ctttaataaac ctttgggtat ataccagta atgggatggc 40560

tgggtcaaat ggtatttcta gttctagatc cctgaggaat cgcacactg acttccacaa 40620

tggttgaact agtttacagt cccaccaacg gtgtaaaagt ggtcctatth ctccacatcc 40680

tctccagcac ctgttgttgc ctgacttttt aatgattgcc attctaactg gtgtgagatg 40740

gtatctcatt gtggttttga tttgcatttc tctgatggcc agtgatgatg agcatttcct 40800

catgtgtttt ttagctgcat aaatgtcttc ttttgagaag tgtctgttca tatectttgc 40860

ccacttgttg atggggttgt ttgttttttt cttgtaaatt tatttgagct cattgtagat 40920

tctggatata agecctttgt cagatgagta gattggaaac attttctccc atcctatagg 40980

ttgcctgttc actctgatag tagtttcttt tgctgtgcag aagctcttaa gtttaattag 41040

atcccatttg tcaattttgg cttttgttgc cattgctttt ggtgttttag acatgaagac 41100

cttgcceatg cctatgtcct gaatggtaat gcctaggttt tcttctaggg tttttatggc 41160

tgtacateta acatttaaat ctttaateca tcttgaattg ctttttgtat aagggtgaag 41220

gaagggatcc agtttcagct ttctacatat ggctagccag ttttccagc accatttatt 41280

aaatagagaa tcccttccccc attgcttggt tttctcaggt ttgtcaaaga tcagatagtt 41340

gtagatatgc ggcattatth ctgagggtc tgttctgttc cattgatcta tatctctgtt 41400

ttggtaccag taccatgctg ttttggttac tgtagccttg tagtagagtt tgaagtcagt 41460

[0051]

tagcgtgatg cctccagctt tgttcttttg gcttaggatt gacttggcaa tgagggtctt 41520

tttttgggtc catatgaact ttaaagtagt ttttttccaa ttctgtgaag aaagtcattg 41580

gtagcttgat ggggatggca ttgaatgtat aaattacctt gggcagtatg gccattttca 41640

tgatattgat tcttcctgtc catgaatgtg gaatgttctt tcatttggtt gtatcctctt 41700

ttatttcatt gagcagtggg ttgtagtctt ccttgaagaa gtccttcaca tcccttggtg 41760

gttggaattcc taggtatttt attctctttg aagcaattgt gaatgggagt tcaactcatga 41820

tttggtcttg tgtttgtctg ttattggtgt ataagaatgc ttgtgatttt tgcacattga 41880

ttttgtatcc tgagaacttg ctgaagctgc ctatcagctt aaggagattt tgggctgaga 41940

cattgggggt ttctagatat acaatcatgt catctgtaaa cagggacaat ttgccttcct 42000

cttttcttaa ctgaataccc tttatttctt tctctgcct gattacctg gccagaactt 42060

ccaacactat gttgaatagg agtggtgaga gagggcatcc ctgtcttggt ccagttttca 42120

aagggaatgc ttccagtttt tgtccattca gtatgatatt ggctgtgggt ttgtcataaa 42180

taactcttat tattttgaga tacgtcccat caatacctaa tttattgaga gtttttatca 42240

tgaagggtg ttgaattttg tcaaaggcct tttctgcac tattgagata atcatatggt 42300

ttttgttgtt ggttctgttt atatgctgga ttacgtttat tgatttgtgt atgttgaacc 42360

agccttgcat ccagggatg aagcccactt gatcatggtg gataaacctt ttgacgtgt 42420

gttggaattcg gtttgccagt atcttactga ggatttttgc attgatgttc gtcagggata 42480

ttagtctaaa attctttttt tgtgtgtgtg tctctgcaa gctttggtat caggatgatg 42540

ctggcctcat aaaatgagtt agggaggatt cctctttttt ctattgattg gaatagtctc 42600

agaaggaatg gtaccagctc ctcttgtac ctctggtaga attcggtgt gaatccatct 42660

gtcctggac tttttgggt tggtaagcta ttaattttg cctcaatttc agagtctgtt 42720

attggtctat tcagagattc aacttcttcc tggtttagtc ttgggagggt gtatgtgttg 42780

[0052]

aggaatttat ccattttctt cagattttct agttttattg catagaggtg tttatagtat 42840

tctctgatgg tagtttgat tctgtggga tcggtggga tatccccctt atcatttttt 42900

attgtgtcta ttttattctt ctctcttttc ttctttatta ctcttgctag tggtttatca 42960

attttggtga tcttttcaaa aaaccagctc ctggattcat tgattttttg aagggttttt 43020

tgtgtctcta tttccttcag ttctgctctg atcttcgtta tttcttgect tctgctagct 43080

tttgaatgtg ttgctcttg ctctcttagt tcttttaatt gtgatgttag ggtgtcaatt 43140

ttagatcttt cctgctttct ctgtgggca tttagtcta taaatttccc tccacacact 43200

gttttgaatg tgcccagag attctggtat gttgtgtctt tgetctcatt ggtttcaaag 43260

aacatcttta tttctgcctt catttcgtta tgtaccaggt agtcattcat gagcagggtg 43320

ttcagtttcc atgtagttga gcggtttcca gtgagtttct taatcctgag ttctagtttg 43380

attgcactgt ggtctgagag acagttttatt ataatttctg ttcttttaca ttgctgggg 43440

agtgttttac ttccaacat gtgggtcaatt ttggaatagg tgtgggtgtga tgetgagaag 43500

aatatatatt ctgttgattt ggggtggaga gttctgtaga tgtctattag gtccacttgg 43560

tgcagagctg agttcaatc ctggatatcc ttgttaattt tgtctcgttg atctgtctaa 43620

tgttgacagt ggggtgttaa agtctccac tattattttg tgggagtcta tgtctcttg 43680

taggtctctg aggacttget ttatgaatct ggggtctcct gtattgggtg catatatatt 43740

taggatagtt agcttttctt gttgaattga tccctttacc attatataat ggccttcttt 43800

gtctcttttg ttctttgttg gtttaaagtc tgttttatca gagactagga ttgcaacccc 43860

tgcccttttt tgttttccat ttgcttgga gatcttctc tatccctctt ttttgagggt 43920

atgtgtgtct ctgcacgtga gatgggttct ctgaatacag cacactgatg gttcttgact 43980

ctttatecaa ttgcccagtc tgtgtctttt aattggagca tttagcccat atacatttaa 44040

gattaatatt gttatgtgtg aatttgatcc tgctgttatg atgttagctg gttattttgc 44100

[0053]

tcgttagttg atgcagtttc ttccatagcat cgatggcctt tacaatttgg catgtttttg 44160

cagtggctgg taccggttgt tcctttctat gtttagtgct tccttcagga gctcttttag 44220

ggcaggcctg ttgttgacaa aatctctcag catttgctcg tctgtaaagg attttatttc 44280

tccttcactt atgaagctta gtttggtgg atatgaaatt ctgggttgaa aattcttttc 44340

ttaagaatgt tgaatgttg tcttcagtct cttctggctt gcagagtttc tgctgagaga 44400

tcagctgcta gtctgatggg ctccctttg cgggtaaccc gacctttctc tctggctgea 44460

cttaacattt tttcgttcat ttcaactttg gtgaatctga caattatgtg tcttgagct 44520

getcttctca aggagtatct ttgtggggtt ctctgtattt cctgaatctg aatgtcagcc 44580

tgccttgcta gattggggaa gttctcctgg atgatacct gcagagtgtt ttccaacttg 44640

gttcattct cctgtcact ttcaggtaca ccaatcagat atagatttgg tcttttcaca 44700

tagtcccata tttcttggag gctttgttca tttcctttta ttcttttttc tctaaacttc 44760

tcttctcact tcatttcatt catttgatct tccatcactg atacccttcc ttccagtga 44820

tccaattggc tactgaggct tgtgcattca tcatgtagtt cttgtgceat ggttttcagc 44880

tccatcacat ccttttagga cttctgtgca ttggttattc tagttagcca ttcgtctaata 44940

ttttttcaag gtttttaact tctttgceat gggttcaaac ttcctccttt agcttggagt 45000

agtttgatca tctgaagcct tcttctctca gctcttcaaa gtcattctcc atccagcttt 45060

gtctgttgc tggtagaggag ctgcattcct ttggaggagg agaggtgctt tgatttttag 45120

atttttcagt tttctgctg ttttttcccc aactttgtgg tttatctac ctgtggtctt 45180

tgatgatggg gacgtacaga tggggttttg gtgtggatgt tctttctgtt tgttggtttt 45240

ccttctaaca gtcaggaccc tcagctgcag gtctcttgga gtttgccgga ggccaactcc 45300

agaccctgct tgccctgggtc tcagcagtgagg aggtgcaga acagtgagata ttggtgagca 45360

gcaaagtgtg ctgcctgacg gttcctctgg aagttttgtc tcagaggagt acctggccat 45420

[0054]

gtgaggtgtc agtctgcccc tacttgggggg tgcctcccag ttaggtact cgggggtcag 45480

ggacccactt gaggaggcag tctgtctgtt ctcagatctc aagctgcatg ctgggagaac 45540

cactactgtc ttccaagctg tcagacaggg acagttaagt ctgcagaggt ttctgtgcc 45600

ttttatttgg ctatgcctg cccccagagg tggagtctac agaggcaggt aggcctcctt 45660

gagctgtggt gggctccacc cagttcgagc ttcccagctg ctttgtttat ctattcaagc 45720

ctcaacaatg gcaggcgecc ctccccaagc ctggttctg ccttgcagtt tgatcttaga 45780

ctgctgtgct agcaatagtg aggtctctgt ggcgtaggac cttctgagcc atgcacggga 45840

tataatctcc tgggtgtgct tttgcttaga ccacagaaa agtgcagtat tagggtggga 45900

gtgacctgat ttccaggtg ccgtccatca cccctaggaa agggcattcc ccgaccctt 45960

gcacttcccg catgagggtg tgcctgccc tgcttcagct caggcttggt gcactgcacc 46020

cactatctg caccactgc ccaacaatcc ccagtgagat gcacccgata cctcagttgg 46080

aatgcagaa atcattcatc ttctgcattg ctcacactgg gagctgtaga ctggagctgt 46140

tcctgttcag ccactttggc tcaaaactaac ccactcagac tgtgttagtc tgaatgaaga 46200

atatgaagaa aaagaattta agaaattaac aaagtttttg agaaatatgg aattttgtaa 46260

agtgatcaaa cctatgaatt attggcattt ctgagagaga aggagaagga ctcaacaatt 46320

tggaaaatat atatgaggga ttaattcaag aaaattttct taatcttgct agagggatag 46380

acatccagat ataggaaaac cagagaacac ttgcagata ctgtacaaa tgaacatcac 46440

caaggcatat agccacagac tgtcaaaggt caatgctaaa gaaaaaaatc ttaaagccag 46500

ttagagaaaa aggtcaaacc atgtacaaag ggaagcccat caggctaaca gaggattact 46560

catgagaaac cttacaagcc agaagaaatt gggggcacat ttacagcatt cctaaaggaa 46620

agaaattcca tgcaagaatt ttatatccca tcgaactgaa cttcataagc aaaggagaaa 46680

cgaaaacttt tccaacaag caatcactaa tttgttacca ctagaccage cttaaaaaag 46740

[0055]

atccttaaga gatttataaa catagaaacg aaagaacaac atgtgcaacc acaaaatcac 46800

atgtaaatac atagtccaca gtccctatac agcaaccatg caatggaaac aacaaagcca 46860

tcagccatca acttcacaag aggattaaaa cttcacatat caatattaac cttgaatgta 46920

aatgatctaa tgccccacat aaatagcaca aagctgcaaa ttggaaaaaa aaaacaagac 46980

ccatttgtat actgtcttca agagactaat ttcatatgta atgacaccta tagtctcaaa 47040

gtaaatgggt tagagaaaga tctatcacgt aaactaaaaa caaaaaagaa cagggatata 47100

tatatatata tatatatata tatatatata tatatacgca tatcagataa aacagacttt 47160

aaatcaacaa gtaaagaagg gcaaagaagg gcactacata aactatagtt caacatgaag 47220

atttaactat cctagatata tatgcatcca acattggaac acttaaattt ataaaacaag 47280

tacttctaga cctattaaaa gacttagata gccatacaat aatagtgggg gacttcaata 47340

ccccactgac agcattagac agagcatcaa ggtagaaaac tagcaattct ggacttaaaa 47400

tcaacatttg accaattgga cccaatagac atctacagat cctccaccc atcaacaaca 47460

gaatacatct tcttatttgc gtgtagaaca tactctagga ttgaatacat gcttggccat 47520

aaagcaagtc tcaatatgtt caaaatcatt aaaattatgc caatcatact ttctgatcat 47580

agtggaataa aatggaaatc aatataaaaa gatttcttaa aaactacaac atgagttgga 47640

aattaaacaa cttgttccca aatgactttt ggggtgaacag taaaattaaa gcataagtea 47700

aacaattctt tgaaataaat gaaaacagac acacaacata ccaaaatatac tggcatgcag 47760

caaaagcagt ggtaagagga aagtttatag cactaaatgt ctacataaag aagttagaaa 47820

gatctcaaat taacaatcta acttcatccc taaaggaacc aggaagcagg aaaaactage 47880

cccaaagcta ggagaagaaa ataaatactt aaaatcagag caaaatcaaa tagaagagac 47940

ctaaaaattc atccaaagga tcaataaaac caaaagcttg ttgtttgaaa ggataaacia 48000

gatcaataga ctgctagcta gattagcaaa gaaagaaaag tacaattaca aatgacaaag 48060

[0056]

ctgacattat gacccagccc atagaaaaat aaaagatcct cagatactat tatgaccacc 48120
 tctatgcaca caagctagat aaactagaga aaatgaatac attcctagaa acacacaacc 48180
 ccctaagact gaatcaggaa caaaccaaaa ccctgaacag accaatattg agttccagaa 48240
 ttgaatcaat aacaaaaaaa cttaccaacc aaaatgagcc caggaccaga gggattcaca 48300
 gccaaattat actaaaggta cagagaagag ctggtacca tcccactgaa actattacaa 48360
 aaaaaacgag gatgaggggc tcctctcaac ctcattctac aaaccagta tcattttgat 48420
 accaaaatct ggaaaagaga caatgaaaag gaaaactatt gcaaatatcc ctgatgaaac 48480
 agatgcaaaa atccttaaca aacaaactga atgcagcagt agtcacaaaa gctaatgcac 48540
 catgatcaag tttgtaggct ttattcctgg aatgcagagt tggttcaaca aaaagcaaat 48600
 caataaatgt gattcaccag ctaaatggaa ttaaaaacaa aaaccatag ataattctca 48660
 tagatacaga aaaaactttt gatgtaatcc atcattgttt cataataaaa atcctcaaca 48720
 aactaggtat tcaaggaaca tacctcaaaa taatgagagc cgtctatgac aaaccacag 48780
 ccaacatcat actgaagggg caaaagctgg aaacatcccc ttaagaacag gaacaagaca 48840
 ttaattctca ctcttaccac acctattcaa catagtcta gaagtcctag tcagtgaat 48900
 cagecaagag aaagaatata aggcattcaa atagggaaaag aggatgccaa attctctctg 48960
 tttgctgaat gtaaaataaa agttgaaata atataataaa atcattattc aaaaaattat 49020
 taataggtat tttaaaaaca ctattattac attaaacatg tctgagtatc ctgatacagg 49080
 gaagtctgta ggtgtggcca tcttttttat ttgtcacat acattatagc caagtaggaa 49140
 acaaaagcac tttcaagaca cttgggtgaa tggggttgtt ccctgagaat gagactggga 49200
 aaggacgagg acttttcaac tttcaacaaa acatttcattg gttggaataa ccacctattc 49260
 cttgcctcac taaactctta cettgggtgt ggaagcaagc agatattaga tgtccttttt 49320
 gtatccacta agacaatgat atccaatgtc aaaagtttgc ccatgatggg tgccttcact 49380

[0057]

gggtttagct actctttacc agcatctgtc aagtgcccca gagtgggaat ccaattgaat 49440
 atagagecac ctttgtctaa atctgtctgc aagcaaggaa gccagaagat gcctccaatg 49500
 ggcccctgga acataagctc cctgagaggg gactttgtct tgtttcaaac tgatccaagt 49560
 acctagacca gtgccaaata agtatgtagt gaatgaataa atttttgaat gaatgaggct 49620
 tggectagaa gtcccctgat gaggactgaa aagtctgtct tcacacttac tgagacctca 49680
 cttgtaggaa tcagctteca gtggcgggtc accactgcca taacttctgg ctcccatgaa 49740
 ctgtactaat gtctcaacce aaactttctt agttgattct ggtgccacct ttctacatct 49800
 tggggtaatt cactgacagc aggagttctg gtggtacttt ggagcacaga tctgactcca 49860
 caggcacagc taggtctctg tgcccaggtc tagggatcct tttttttttt ttctatattt 49920
 cttaaaacat tttagaggaa ttcagggtac catggtggtc aggagcatgt ggtcagtgaa 49980
 actgtctggg tttgcacggc agcttctggg agacctggg caccattctt acattgattc 50040
 ccctctgtaa aatggggagaa ggctagtatc taactcaaag gactgcagtg aggattaagt 50100
 tatttactgt atgtcaagcc actggtacat tatgagtcct atataaatta tagatcctat 50160
 gaatacatat ccaatttgca actcctgaaa tgatctcatt tttcccctt tccctaagtt 50220
 cctctcagta tatgcccttg caaacacctc catttctctc tttgcacctg tcaccttaatt 50280
 agctttcaag gateccttcc taccccgta acaccagat ctccctata gaaataatat 50340
 gaatgttttc tctgcagcgg tgtgtccttc tgcaacctct ttccagtcct aacacataaa 50400
 ctctcaggg tctccctatg aatggagata agtgacattt ctcttcttc ttgtctgtg 50460
 aggcactcag cctgggagat aagaagggga agagcaatgg agcagctttg ggaggggcaa 50520
 aatagaaatg taggtggaga agagaagctt tatttttcac agagtgagac ataagtctg 50580
 ctaagcactg aacagaacat ggggtgtggag acaaccatga gggacacca gggagaccgc 50640
 tggagatcac tgacaacctt cttcaggatg aactcaggag ctgagacctt gggcaggagg 50700

[0058]

ttaggagtgg gtacatcaga ctgaggttaa tcctgggcac tttggtctct ggttccttta	50760
aaattgctca tctcaatgta ctgccaaggt cagcagtaat tctctagcaa tggctggagc	50820
ccctggcagg ctcaactaca gcagtagagg ctgtggtgcc tgtgccagta gaaaagatgg	50880
ggcctttgat aggtcaagca gcagtgcgcc actaaagtta aagtgacccc tgaactgggg	50940
ctgcagggaa aggctggagg tgggtatgag gcttagtact ggggtgtgca ctttgaaggc	51000
tttgggggtt atcactgetg ctggttttgc tggagggeat ctcagaagac atcaatagaa	51060
tcagatgtgc ttgtaatgca caaacctctg tcaatatctc ttctgtgatg ggtgtcatag	51120
accccatget ctctactga actaagaaat acacggaaga aatgaaaagt tacttggagt	51180
tagtttacct tggaaaaatg cgtttttctc tggctctcta ctagaaaagg atgctgagtc	51240
ttttctaact ggccccctca ggtatcttat ctccatgtaa tcacagagga gaataagcct	51300
gtttgggtaa accttccttg atgatgtctg gtgagatcag tgcattctgc agaacaatg	51360
taggaagaag atgagggtgt getgaaaacc cattaccttc acagtcacag gacacctgaa	51420
acatcctgag agggtgccca ttaacccac actcttctctg gtgcctccat ccagaccct	51480
ctgcttctct ctccctgat ggaagcctgg gcttgtcttc ctaggtcttc cagctcaggt	51540
cctgtctcca gctagacatc catctgattt catagaagcc agagaatgct gattttgaga	51600
agtcagtgga ctggggagcc caggacagga gctcctagca tgtgaggggc agcatgagtg	51660
ggctgctggg ggagttgaca gcaactaagta cccagccca gcttcaggac ttatccacct	51720
tccctggaac aggatggcat gtttgaccac aaagagcctg gtctcactt gtgcatcttc	51780
cacccaggcc tacatgtgtc cttaaaagag getgtgccct gtgctttgtg gtgaaaggta	51840
gagttcaata tgacatgtca cctgacctat agaaattaat tgtctaattg cacagacaaa	51900
caagtaaata gggagcctgc ttccatgaga aaattgctat aatgtaagtc actatgcaca	51960
aagtttgcag gtctttgtcc atatacttct ttctagcatg ttcccttagg aaaaccaaac	52020

[0059]

accagatgt ggtgttgggg tggtgacttg ctgttggggc ggtgactgag ctaaggacgc 52080

tggacattca gtatgggctg tgtgtaaaac agatggaagc ctaaagtcag gggaggtgag 52140

ggtccctgaa agagcccccg ggcaggagag gaggggagtc aaatgagtac ccgattgaag 52200

ggcaaatgta gagaaggaaa gtacaatagt tgcagtaata agagtcacaa ctgacgcatt 52260

ataaaaaagc atgaccttcc atcaaaatcc aaggaatgcg tgctctttat ttgaagacga 52320

tgaaaacact gagggaggag tcgttcatct gcttggactc ggtaaggaag ttgtgcaaat 52380

gtgcatgata atggatggcc ttgatgaaga tggctttgct gtacgtgttg ctgcttatga 52440

ggttgatggt agtgcagggt gcattctctt taatgaagct gacctgatgg tgcctatggt 52500

tgggtatgcc gaattgggtc aaactgaatg ctatttcctt gttaatttca ttttaatttt 52560

aaatgaattt atcattcaat taatctctaa tattaaaaga ccttttgcta gatatttttc 52620

aagataaatc acagtcaggc ctataaccaa actttagcca tattaaaatt ataattttca 52680

aagaaaatta aattatgttc attatcagtg ggtgtaattt ttctcttact ctgcctaaga 52740

gtttgccctt ttaagtacat agtgatgcct gtaggagcat aaatgcaccg ggctgtaagg 52800

aggacagtgg ctctgtgtgg atgggtagag ggtgggtgca gattttgata tcctcagcag 52860

ccttttcttg cctccagccc atggtacctg cctctgggtg aaaggtattt tcagaatcat 52920

ttccagaact actcctgagt ggcggttttt ctcatattgt ttagggtttc ttcagtagac 52980

cctggcctgc aaaaaggcac ttaagaaaaa tggaccaagt actttctcat ggtacctaga 53040

ttgtgaagct gaacaaagct ggttgccac ttttgettct taggtgtaga atctctagta 53100

tttgggatac aacaacagag gagtaagttc taaagtttca aaacacceca gaactgtagt 53160

gggggccctc cttctgtctc atggaatgca gcagatgtgt gcatatgggc ctctccacag 53220

ccccccccc gatgcctggc ttgggagget ccatacttgt cctggtgagg atgccacca 53280

aactttcttc cctccacga ggaactccca ctttttgccc cctctgtttg ttatcagagc 53340

[0060]

aggagctctg taggttcaag gctgagcagg tgtcctggga tgacagagge ttggctggag 53400

gaccactgcc ctgccactta cccagagcat gacatcagct ctctgagggt aatctctacc 53460

aggagtacca cccttagatc tctggctttc taaggttctg tgaggtcagg tgacaagaga 53520

acctggcctg gctggatata tctggataat ttcatgactt agtagtgcca atgggaatga 53580

tgggcacagg tggatatagtg gctcaggttt ggctctttgg agtgcatttt aatcaggtea 53640

tataaaatac acgtaaatga ctatctaaat aacaaaaaga tctttttcat tccctgcaga 53700

acacagcttc ctttcagttt aggaccttc catccccaaa ccattagcat gaccatggca 53760

agcagctgag acaatctggg ttctcaatgt tatccttggt tccctagtgg agacacaggt 53820

accacagaga gatgtgaagc cctccatgtg gtggttgga gtaaaagggg aatggaagag 53880

agccttacta atggaaaaga tccaggaatc ccccaactg tccaaatata tgttctttaa 53940

atcttctgta ggctaccat ctcttatttt tctttttgtc ctctctccc tcatcccttc 54000

tcttcttct tctgcccctc ctgcccctct ctttcccttc ctgcccctc cctctctatt 54060

cttccccat tcccttccc ttccctacc ttccctctct tctgtcatte tttccttctt 54120

tactttggaa tatgttgagc tgaacaactg taatctctat tactacgcc tctatatctt 54180

cttgggcaga tgctaagagt ttttttgtga ggaaagtcac aggactcttg aaattcctgg 54240

ggcccagcct tctgtaggc ctgtcacttt aattttttct ttcattcagt catcattta 54300

ttcaattatt caagtgagat gtgttggtct gatatggcac accaggcatt ttgctataca 54360

caggagagtt ataaatgaat aagatttggt ctctgttctt gagtagctca cagtgcagtg 54420

gggcagacag acaaataaca aaataagcat gateccatgt ggcaagtaca ggaatgcagg 54480

tgtgtggaat gtccttgggc tcaagttcag accctttggg ttcaaaacc cttttaactg 54540

gagtgactct gtgggggaaa atgtgtctag atgttgccac atgctatcat ttgaggatga 54600

gateactggc ctttgacaat agtatctgtg caagtagcag gagcagaggt ataaatcctt 54660

[0061]

gatctgtctg tgggcagcca gccaatgatg tgataccaag aggagctaag ggaacccagc 54720

tcattggaacc tcaccaggag gaagagataa tgctgttgct aaggagaaga gagctgaggg 54780

ggaccctgag ggatctaaga caaagtggag ggggcaggat aatgggcttg ttcagagcat 54840

tcatttttga ctcagacatt cctgagtttg aatcctgctt ccccaggtag cagctggtat 54900

caaggtggga agctataaag tgataagagg cacaatggac caaagatgtg aggggcaaaa 54960

tcagagaaaag agacaggaag tctatagagg ctgtgaggac atggtttaca aattaaagta 55020

catctgaggg ccattgtatc agtatcttag tgactgtggt atttactcta agagctccag 55080

agtagtttca gggagggagt gatccacett agttggggtg gctaaggga attagctgag 55140

ctgtgaataa tacttcatgg atgtagtgat gatgaggcca acagtgggga agcttgtgaa 55200

cactgaatgg gcagtgattg aacaggattt agcaaattgc aattacataa gctatgtagt 55260

attttttcca ttgcatttta tataacatta gagatgcatt gcctctgaaa caccaccttt 55320

aaaaatggga caaatctttt aaatcaaaaa cacgaattaa aatgatcaca aaatgttact 55380

gtcttttata aacgttgatg aaaattttgt gtctggttat cctgttacag ggaggcgtat 55440

aagtgatgaa atctccatta gatgatgaaa tagttaatat ttgtgggatg aaaagtgacc 55500

ctgtgagaag aggtttatc acaggcctgt ggagagaggg agggagatgg ggaaggcat 55560

gaggagagag gcctgcttgt gcaataactt tgtecccttg gtagtcatga tttatccttt 55620

cttctttctt ggctgttgte ttttgatgtg gaagctagaa aacacttgat gccctttcct 55680

tttgtcacca ccaaagccac tgtctcagag cctcaatatt ttgcctagge tgagtgtctt 55740

ctctgggttt attaaataaa caccagtgtc cactcacctc cagaggctga gaggccaaact 55800

gaggagaagg ccactgtgat tcaactctgc atgcatgcaa aatcggaac tgtctccttt 55860

ggaccagttt tgaatgtgag gtcccaaaag acagaaattt gtctctctga ttccttattg 55920

tgteccctagt gcctacaatc aagcacataa tcaatgctca ataaaaatgt gtcaagtga 55980

[0062]

taaatgggtc cagggtcag gacttggtat tgcccagatg cagtttggga ctttgttttt 56040
 ctctgctgct ggaaactccc ttattgggct ccatttttga gaggtccttt tatacttctt 56100
 tatagattcc caggatctca ttgatttgcc agcaatgttg actctgctcc attttctctc 56160
 cactcagatt cgateccaacc atageccagaa ccccgtcatt cttcagtgtg gtggaatgag 56220
 gtgtttctggg cctccaacct tgttctgcag cccctgacag gaggcaggt ctaattctcc 56280
 agttttgtcc agcctgtgtc atgctgcttc caggtctgca cagagatgct gggctctgcat 56340
 gtgggtatca ctcacattcc aagtgtggag ctttgcaggc tgtatcagcc atggcacact 56400
 cttccctggg ggcaggcctt tcacccctt cactgtccct gggtagagc tctctagtga 56460
 ttccaaccg aacttaaacc aatggcccaa agctgtgctt cctgcttcaa ttcaacaaa 56520
 tttctctctt tttttaaata cctattggtg cg 56552

<210> 3

<211> 23382

<212> DNA

<213> 人类

<400> 3

tcctgatecc aatagagcct cttacctgag ctttctaaa gtccaacaa tttctaaat 60
 ctcactggct gcttgatgt acagaaggga tagagaagta taaattttga tatctggaca 120
 tgggtgtggtt ctggacaaaa tttaaagatg taagttaaga atgaagtga gcacgcacat 180
 tgcgtacaca tccaatcaag tagttcttga tatagttcaa cctgctgagt aaagacagct 240
 ggattgctgg tttctggaaa tagcgacaca gactctctgg gtacacttca cttacctctt 300
 aatgcaacat tcaggaacgt cctgcccacg ttccacctt tatggtctct getgtatgca 360
 ctgaggcaaa tctgcctgga aggttgatta ggcaggggaa ttcactgcct gcagctgagg 420
 tctggagtgc tctggggggc tccctggaga caggcacaag acctggctta tggaagcctg 480

[0063]

gcggaggctc caggttactt ctgcctcctg tctccccacc tctaateccac ccatcaccca	540
gaagtgggcc aaggatcagg ggtctggatc atctctcaag ggcaggagga gggtagaggct	600
ctgagaagaa gtgtgctgtc caatcagaaa ccatacttgg tgttatggct gcctctcttt	660
caaaggtatc tagagtggga gggtagacaga tgattaaactc tgggttatct gaaatatgac	720
ccaagtgagg agagtcaaca gtaaccagaa cacattgtga aaaggaaaag gtggatagat	780
tgtcacctgg gccagaatcc ccagagagca ctttcacctc cacatcgaat gaaaaattct	840
acctggtcac cctcttgagt cccagaggg tcagcccagg aatagggtag agtagaaagg	900
agttttgggc caggatggat gcttgtgtgt gtggttatct caactcggac cctctgagat	960
ccacagcagt gcccagagac tgcgacgtg ctctctcaact gactctccag gctctgtcgc	1020
tgtctgtgca acagctctgc agccaaggtc tgcctgactgt gttttgcctg tgctaaatcc	1080
aggagcagct gctcccagag caggaccaac actaaccaca gtgctaagaa aaactccaga	1140
gcctgaggga ccagcatatt ggccggctct tggagggact ctactgttg ctggctcagt	1200
tggcaggata tcttaacatg agttcagctg gaagtgaatt cagctgaaca gaacacatgt	1260
gtaaactaga ataaggctac ttttcatttt ctaaacaatt ctcttctgc ttgtaaataa	1320
ttggcaaacc cactaatatt cagaaagaaa tccccatgt taccatttca tggttgctcc	1380
ttgaccttgc taggtcagat ttcaatccta gtccaccccc acgttggttg cattgctcca	1440
tcttcacacc taattcagaa ctatgtacag aaagagtaaa actatgtaat tctaaacttt	1500
tgtgctagtt ataagatatg cactcttata agtgtttcag gtgtaactca accaaaataa	1560
caattaattc acacaatatt gctgttagga aattacagct attttccaca ctcatagat	1620
gaaaagacag atgcacagac aatttcgggc acatttctaa ggetttaagg caagtaactt	1680
gttctcagaa tccaaatcta ggtgatttaa ctacattctg tggctcctaac aagttctctg	1740
gagttaaact gggaaggcat gcactattta ctgtgtgtga aattcaaaca gcttatggta	1800

[0064]

agttcttctc aacaagctag acttaaaaca gctattccct tcttttcttg cctgatatac	1860
taatgtttct gtgttagttc ccacactgca taaaatgagc ctctatttct catgcectct	1920
tcateccagt cctccctcc tttctgaaga atatagtcac agtgggcaga tattatattc	1980
ctttcatatc atgtagtata ccagagttg gccacatagt caacaatact tgctgagtgt	2040
gtataagagt gaattaataa actctgatta ttgaatttac taatgggaga cagtccaaga	2100
aaatcctcaa attgttttct ctttagagat gttcttttct ttaattcat ctgatcatat	2160
ccttctctca ttactcttg taagaactgc ttaatagtta cctaatttgc tcagcacaga	2220
attcttggaa aagttccttg ctccaggga tctaaggca tggatatgca gtgattgcaa	2280
atagaatatt taattaatgt gcactattca tctaactg acattcttct aattccaatg	2340
attagaaaga aatactggat ataaatggac ataggcagtg aggtgttgg tgtgtgtgtg	2400
tagtgatcat atgtaattaa cataatttct tccatgata gcctataaat agtgctgtct	2460
gggtcatgaa aaacgatggg ttcagtcttg gttttgatta agaagtaagt gcatagatat	2520
tgcatataag agaggtacac agacaatgtt ggcaggaat tggatatgga gagggattta	2580
gaaagcatag ttataaatta tctggaaaat tcatccacc aaacaaaata attgacaatc	2640
tggactagac agatggtctg acattgtgtg tgtgagtgtg tgtgtgtgtg tgtgtgtgtg	2700
tgtgtatggt agtgtatcta tgcgtgtgtg aaacttgttg tcagaaacct gttgacatta	2760
tactaaaata attgaaattt ttttggttaa attttctgaa ctttatccc agctgtgtga	2820
acttgggeta ttgttttagc tetaggagcc ttgttttct cattagtaca gtgggaatat	2880
taataccacc tacatcacat gattagtatg aggattaaat gagataatgt gtataaaaca	2940
cttaaataat acctagaaaa aaatagatgg tcagtaaaca tcagttatta tattacttcc	3000
tgttattttt atcattgtac acacatgtat gtatgcatgt atacacacac acacacacac	3060
acacacacac acacacacac ggaggcattt gacccaaaag cacaactcca ctcaacctat	3120

[0065]

acctcaggga aagatgggga agtaggtaga cctgaattgg ctagggtttt gcagagaaac	3180
agaagagaga gataaattta ttttaagaaa ttggctcatg agatgactgt gtgggctggc	3240
aagtccaaaa tcttggtctg aaattcagat aacagttgat gttgcagcat tgagtctaaa	3300
ttctgcaacg cagtagcttg gaaacaggca gagtttctac attgcagtct tgaggataat	3360
ttattctttg ggaaacctca gtgtccatta ttaaagactg cagctgattg gatgagctc	3420
atgttaggga agataatctg cttactcagt ctactgatta aatacttaca gagaacatg	3480
tggattgcta tttgatcaaa caactgggta acacagcata gcctgggcaa aattaaccat	3540
cacagagact gagaggatac atggaggtea ctggagatat gaccaagag acagaagagg	3600
gtcaacagtg acgaggacac actaagaaca tgccatgatg gttttatfff cacttaggat	3660
aaaggcaaat ttctggacag acctttggag aatattggga aaacagaagc ttctacctgg	3720
taaccccatg gtaagcccaa caaagctgtg atgttaggag gactggagaa agggagtgca	3780
gcatgcggtt gaggagccc ttggccttcg cactttggtt tcacagaagt gtagtcagta	3840
cagcacggac aaggctgctg accactgcac agtagctgga gctctgggga ctggtgattg	3900
ctgcagcttc ttctatgca ggggtcacac aagaagattc ggaggtgaac atgggctgtg	3960
cactggagtg gcactgtgga gggccctttg aaggcattgg cattgctggt ggttttttgg	4020
aaaggtatct gagtttcaga atgaattgaa taataagtca gataaattag agccaggacg	4080
ggcataatff tgccctttta aacctattct tgggtgtgetg gcagacagtt tataaaatac	4140
atttgtatcg aattagtacc tatgaagaaa cgtataaccc caagttaaaa gtgttagaac	4200
tgacatgctt aaaaacagaa ttaatgctaa atcaatgtca attctgtctc tcatttcatt	4260
tcatatctac gtaccactt cctacctaca gtatttcatg agctttaaga ggctatcaat	4320
tgtaggga acccttattt catgcaccac aaggaaagaa aaaaatgctg caaattcaac	4380
aattacgtgc caccaatcat atggtatata ctgatttaag ggctgttaac aagtattaga	4440

[0066]

agactgtatc atggaattga tgaattaaga taagtataaa accgatccac tttaggtcat	4500
ataaatatgt atgtaaatac tagacgggcc aattcttgct ttgccagaa caccctctg	4560
ctgtgctcct ccctatcact atgtctagca ctagacactg ctagacagta ttgacaaaa	4620
agatcaaaag gcctacaaag atatggtaca ggaaaatctt ggattggtca atctattaaa	4680
taaagggaga caaataaaaa gacatggtaa ggctttttct tgtactgtag ataaaataaa	4740
gttcccatat ttggtaggga gtatttttct ggaatatatg gattttaagt tcccaggag	4800
acaaatcaag gagctacttg gagagaccag cctgcaggag agtgacttgc cagatgacat	4860
gaacttcaga tgaaatgttt gaagtcaccc gcagggcatt cttctcatga cacctacatg	4920
ctcgtccatg tggctcctaa atgatgcact gatcctgcc aactcatgt ggcaattaag	4980
tatgtgggtg gtcctgagg gagcatgggc aatgtgaccc actgggtctc ctggctcaga	5040
gagatcacac atgatgatgt gaagactgtg gagacagaag caagaatatc tatagtcagg	5100
tgtttcttca ctgacttggt aaccaattcc caccattttt tgagatataa agggaaaaag	5160
tgggcatgtt aaagagctgt ttatctccca gttctggttt cccagagggt cggttcccat	5220
tagtccttcc attgcttatt caaccactca atacagatgc tcatgtatgt gcatgttcat	5280
ttcacctgca cttttatatg aatacagcat acatgcacac agtgcataca tgcacacagt	5340
gcacacatgc acacacacac acaggettgc cteccctatt tctcactta caacattgta	5400
acaaatatct ttgtcaaaga gtatgaggaa gctgagggtg atatcaccat gaagacagt	5460
gaatgaacta acatctttta aacacattgg gaaatagtga ctcaggaggt tcaggcctca	5520
aacatcatct tctttggtat ccagtttgct tgcagtcate ttcaaaaagg ccactaaatt	5580
ctacagtctt tgaggtcatg ggtacttat tgttcacctc tatatcccta tgtttggaca	5640
tacatgctcc aagaatgttg taggctttca aagaatgcct ttggaattca ctgctctaat	5700
cacactgtgg acttagttta tcttgagggt aaaagctctc tgagaacagg actcgggtct	5760

[0067]

gaaaatagga ctgtgtaaat aggtgaatag aatccctcag gtagaagcag acaatgggaa	5820
actatgcttt ctacctaacc accctccacc accaagtaaa cattaggcaa tgtcaaggca	5880
cggctctgttt ctgttttaaat cagaggtgtg aataactatc tttttcaatt ctaatttcaa	5940
catcttcatt gagccataat tcacttatta tacagctcat ttatttaaag tgttaattca	6000
gtgggtttca atatatcat agagttggac aatcaccacc aaattttaga acattttcat	6060
tactccacaa agaaaaccta tatccatttg ttgtaatttt ctaatttctc ccagcctget	6120
aaagtacctg gaagccactg atatttctc agaggatgtg cacaaaagta aaaggcgata	6180
cagggaacag aatcttgtec aggetctggt tctcagatg ccetaagtea gctcttcagg	6240
cgtccagcct gagagagget gtaactggct aaagcctcat tccctttcta gttctacaga	6300
tttgctact ctacacattt catataaata ggatcataag aaatgcagtc ttttgtggtt	6360
ggctttcttc atttagcgta aggttttcaa agttcatcct ccatatatct gtttaaata	6420
agtgaataaa tgaatgggta actctgtcta gtgcaacct gaatcagaac tctgagtaag	6480
tagaccctt tctgaccctc tacttcatga taggaaagga attccctctc cacagaagcc	6540
acaataaatg atgtctactt cctaaggag ttttcaaata aactgtcttg taagggttct	6600
atatggacaa tgaatgccac ataagacctg ggtaagtatc caatacatgg cattagtttt	6660
tcttttaaac acatatttat taagggtcta tcaagcacat agtgcacac taagctcttt	6720
cttacttctc atttaatat aaaaacagct cagtgaagata agtatactg gtttataacg	6780
gccattacat atgagaattg atactcattg tgagtatcat agtgacactc tagcaattgt	6840
tgataattgt tgtagcaatt tccagagatc actcagttat catatgcaag tttctgatt	6900
ccagtagcac attcaatata attcctccga ctaagggcaa tgttttgatt ttagagggaa	6960
ttgtttattc tcaatcatat tctttgctct agaccacaga atcaactcag atttaatcca	7020
ctgaatctgc tggccatggt gagttggagg aaaagcagaa acacagagge caggaggaag	7080

[0068]

ggctgagagc ctaggcaagg gagggatgcc tagaaaaagc caaggctgta gaagccccc	7140
gggatggctg gatccacagc aggacatgga agagaaactg ctcagggtgg agggagtcat	7200
aatggaacat ttccacaagg tcagggaag ctgtggctca gaactgatgt ttcctcagag	7260
gacacgaata caaaagatgc agggaacagg atcttgtcca ggcatgggtc ccttgggtaa	7320
cctaagttaa ctcattaggg ttcccacctg gtggaggtc tgactggcta aagcttctct	7380
agaccgcagc ccagagagg agtcctggaa gctcctgagg cctctttgga ttgcaggagc	7440
taggaggcag gagaatgtca ggaatgacc agcacgtgcc tggttgtgca aaagccacct	7500
cctcctcttt ccagagccaa gcctgggccc tataaaagac acatccagct tcagcacctc	7560
atctgctctg actccccagg gacgtgtctg tgctcctgcg tgtgaccagg gtgagtggca	7620
acctgggata ccagaggggt atgagcaagg cagagggatg gggaggatgg aggtaggctg	7680
gagaaagagt tgtgtgcttg tttctgcca tactgaggag gaatgggact cagtctgcct	7740
cgaatctgtt gaggetgtga ggccctaccg aggtcgtata ggtggggaaa gctaactgtg	7800
cacatctttt ccaattctgt gttcagtgc ttcatgttgg taacttcaa actgctgtgt	7860
tgggcatatt tacacatgg aaactggcaa atacttcaag ccaggtctt ttatctggag	7920
gtttcattgc aaaatctttt tccagaggac ctctgggtgt gtctgcaac acgcagctgg	7980
tccttagggc ttgtattaag ggaatcgta ctcatggag tgtgaatact ctattcaggc	8040
ttgtgaggat agagtaatct cagcttcccc tgaaaggaag aatacgaagc tcattcggtt	8100
ggttttttaa tgtatgtgt tttatttaga ggaaagatc aagtgtttc tttcaggta	8160
ctgatgtgac tgtatgatct aaatatcatt cttcttctac acatgcattg attttagaag	8220
aggtataata tgatccttgg ttgaaatat ttaaagagtt tcattattat ctttcaggtt	8280
gactaaactc tgccaggatg tcttgccagc aaaaccagca gcagtgccag cccctccca	8340
agtgctctcc caagtgtacc ccaaatgtc cacctaagtg tccccctaaa tgcctgcccc	8400

[0069]

agtgeccagc tccatgttcc cctgcagtct cttcttgcgtg tggteccatc tctgggggct	8460
gctgtgggtcc cagctctggg ggctgctgca actctggggc tgggtggctgc tgcctgagcc	8520
accacaggcc ccgtctcttc caccggcgcc ggcaccagag ccccgactgc tgtgagagtg	8580
aaccttctgg gggtcttggc tgcctgccaca gctctggggg ctgctgctga cctgggctaa	8640
gaagaactct ttggacagaa tgtttaagaa cctcctacag cctgatgctt aacctttcc	8700
atttcctctc attccattca tgggtggaca ggcaccacaa agactcatgg ggcttccttg	8760
ggagaacttt gcacttgatg gagcaacctca attgcaggtt ttgttttctt cctttacctc	8820
atgttataat aaagctctga tttctgactc acaaattgtc ttggctgttc tcttctcttc	8880
tgagatttca aacgtcgtcc aaaggtcagg gcctcagaga aattctcttc atgggtcaag	8940
gacaggacct gagtaggtga tacatgaaaa gctgctccaa ggaactgatt cctcaaatg	9000
cagggtggat ggagctcgaa gagggtcacc tctggctatg gcagctttgc ctttctctcc	9060
actaagecag acacatgttc taattttctg tgtattccag acattgggaa ggttgggaag	9120
ctctgtagac ttgttcttct ctagggagaa gaacacatct ctatgaagca aaacttcct	9180
gctgttagaa tcaatagtcg gataaaatgc ttgaaatgca tgttgagacc cttecttctg	9240
caaagtacag cccacagtca ctgaactcct tctcaggag ggtgagtctg gttaagatca	9300
gcccattgac ttgctgcctc agacacctag ctteagttcc tcaaagtgtt agatctgagt	9360
cccaggatgc tgtcctgcca cttggcctat gtgttgggca ggtgagtgat gagtgatga	9420
gatattttatt agaaaaaaca tggaggtecc actttgaaaa tacaacgaat ttatgttaca	9480
cgacactgta gggaagatta gacgtttcct ccttcaatac taaccttctt tgtccagcca	9540
ccctcagatg tgagggggga gaagtgccag gatggttggg ggatctttag tattaatggg	9600
agaaggggag agaagcaggg aaacagcagg ggaaataagg actagtgtt cttaggttca	9660
ggctgacagg tgtcaacatg aagataatgg attcagcgaa gagagaatct gacagagatg	9720

[0070]

gagactgtta ttaagatggt cagctgcctc ctcatccttg cagctgctaa gaaaaacacc	9780
tagctcttta accagctctga gattctgtac tctctgtgtt ttaaattagc ctttgcactg	9840
agtcatttta tttgatgtaa tgctctttat tatttttgcc agtctcattt tattttattt	9900
tttttactta agttctggga tacatgagct gaatgtgcag gtttgttaca taggtatata	9960
tgtgccatgg tagtttgeta tgcctattaa cccgtcatct aagttttaag ctccacatgc	10020
attacgtatt tgcctaatag ctctccctcc tctcacctcc accccaata tgecccgatg	10080
tgtgatgttc cctgccctgt gtctatgtct tctcattget cagctcccac ttattagtga	10140
gaacatgegg tgtttggttt tetgttcttg ttttagtttg ctgaagatga tggtttccag	10200
cttcatecat gtgcctgtaa aggacctaaa cttattcttt ttatggctgc atagtattcc	10260
atgggtgtata tgtgccacat tttctttatc cactctatca ttgatgggca tttgggttgg	10320
ttccaagtct gctattgtaa atagttgtgc tataaacata agtatgcatg tctttatagt	10380
agcataattt ataatccttt gggatatatac ccagaaatgg gattgctggg tcaaatagta	10440
ttcttggttc tatatccttg aggaatcacc aactgtctc ccacatgggt gatctaattt	10500
acagtcccac caacagtgtg aaagcattcc tatttctccg catcctcttg agcatctgtt	10560
gtttccagac tttttaatga tcgccattct aactgggttg agatgggtatc tcatttttgt	10620
tttgatttgc acatctctaa tgaccagtga ttatgagctt ttttcatat gtttgttggc	10680
cgcataaatg tcttcttttg agaagtgtct gttcagatcc tttgcccact ttttgatggg	10740
gttgtttgtt tttctcttga aaatttggtt aagttccttg tagattcttg atattagacc	10800
tttgtcagat ggatagattg caaaaatttt ctcccactct gtaggttgcc tgttgactct	10860
gatgatattg tcttttgaaa agattaacaa aatagataga ccactageca gactaataaa	10920
gaataaatga gagaagaatc aaatagacac aataaaaaat gataaagcat atatcaccac	10980
tgataacaca gaaatacaaa caaaccacca tcaaagaata atataaacac ctctatgcaa	11040

[0071]

ataaactaga aaatcttgaa gaaatggata aatttctgga cacagaaacc ctcccaacac 11100

taaactagga agaaattgaa tccttgaata gaccaataac aagttctgaa attgaggcag 11160

taattaatag cctaccaagc aaaaaaaac ccaggaata gagagattca cagtcaatt 11220

ttaccagagg tacaaaggagg agctgatacc attccttctg aaactattcc aaacaataga 11280

aaaagaggga ctctcccta actcatttaa tgaggccagc atcatcctga ttttaaaatc 11340

tggcagagac acaacaaaaa aagaaaattt caggccaata tcaatgatga acatcaatgc 11400

aaaaatectc aataaaatac tggcaactg attccagcag cacatcaaaa agcttatcca 11460

ccacattcaa gtcggettca tccttgggat gcaaggetgg ttcaacatac acaaatcaat 11520

aaacgtaatc catcacataa acagaactaa tgacaaaaac cacatgatta tctcgatgga 11580

tgcagaaaag gcctttgata aaattcaaca ctgcttcacg ctaaaaactc tcaataaact 11640

aggtattgat ggaacatata tcaaaataat aagagccatt tatgacaatc ctgtagccaa 11700

tatcatactg aatgggcaaa agctggaagc attccctttg aaaaccagca caagacaaag 11760

atgccttctc tcaccactcc tattcaacat agtatcagaa gttctggcca gagcaatcag 11820

gcaagagaag gaagtaaagg ttattcaaat aggaagagag gacatcaaat cgtatctctt 11880

tgcagatgac atgattgtat attcagaaaa cccatcgtc tcagccaaa aactccttaa 11940

gctgataagc aattcagca aagtctccgg atacaaaatc aatgtgcaa aattacaagc 12000

attcctctac accaataata gacaagcaga cagccaaatt atgagtaaac tccattcac 12060

aattgctaca aagaaaataa aatacctagg aatataactt acaaggagg tgaaggaatt 12120

cttcaaggag aactacaaac cactgctcaa ggaaataaga gagggcaca acaaatggaa 12180

aaacattcca tgctcatgga taggaagaat caatatcttg aaaatggcca tactgccaa 12240

agtaatttat agattcaagc tatcccatc aagctaccac tgactttctt cacagaacta 12300

caaaaaacta ctttaaattt catatggaat caaaaaaga gccatatag ccaagacaat 12360

[0072]

cctaagcaaa aagaacaaag ctggaagcat caagctacct gacttcaaac tgtactacag 12420
 agctacagta accaaaatag catggtactg gtacccaaac agatatatag accaatggaa 12480
 caaacagag accccagaaa taacatcaca cgtctacaac catctgatct tcaacaaacc 12540
 tgaccaaaaa agcaatgggg aaaggattcc ctatttaata aatggtgctg ggaaaactgg 12600
 ctaatcatat gcagaaaaca gaaactcgac cccttcctta caccttatat aaaaattaac 12660
 tcaaatggg tgtaaagac ttaaatgtaa aacctgaaag tataaaaacc ctagaagaaa 12720
 acctaggcaa tacaattcag gacataggca tgggcaaaga cttcatgact aaaacagcaa 12780
 aagcaattgc aacaaaagcc aaagttgaca aatgggatct aattaaacta aagagctcag 12840
 ccagtctcat tattaactga aatttcaagt ccctgaacag tagtaatttt gttcttctaa 12900
 aaacacgtac attggccatt tctacaaatt tcataacatg atttcaaatt ttcttttcaa 12960
 ttttcaaggg ctcttgatat atcacagata ttaaaactga tccgccttat taaaatattt 13020
 aatgcaggga cccacttgag gaggcagtct gtgggtccct gacccccgag tagcctaact 13080
 gggaggcacc cccagtagg ggcatactga cgcctcccat ggtcgggtac tcctctgaga 13140
 caaaacttcc agaggaacga tcaggcagca acatttctg ctcaccaata tccgtgttc 13200
 tgcagcctct gctgctgata cccaggcaaa cagagtctgg agtggacctc cagcaaactc 13260
 caacagacct gcagctgagg gtcttgactg ttagatggaa aactaacaaa cagaaaggac 13320
 atccaaacca aaaccccatc tgtacgtcac catcatcaaa gaccaaaggt agataaaacc 13380
 acaaagatgg ggaaaaaaca gcagaaaaac tggagattct aaaaatcaga gtgcctctcc 13440
 tcctccaaag gaacgcagct cctctctagc aacagaacaa agctggacag agaatgactt 13500
 tgaggagtgg agagaagaag gcttcagatg atcaaactgc tctgcgctaa aggaggaagt 13560
 ttgaacccat ggcaaagaag ttaaaaacct tgaaaaaaa ttagacaaat ggctaactag 13620
 aataaccaat gcagagaagt ccttaaagga cctgatggag ctgaaaacca tggcatgaga 13680

[0073]

actatgtgac gaatgcacaa gcctcagtag ccaattggat caactggaag aaagggtatc 13740

agtgatggaa gatcaaatga atgaaatgaa gtgagaagag aagtttagag aaaaaagaat 13800

aaaaggaaat gaacaaatcc tccaagaaat atgggactat gtgaaaagac caagcctaag 13860

tctgattggt gtacctgaaa gtgatgggga gaatggaacc aagttggaac acactctgca 13920

ggatatcatc caggagaact tcccgaatct actaaggcag gccaacattc aaattcagga 13980

aatacagaaa acgtcacgaa gatactctc gagaagagca gctccaagac acataattgt 14040

cagattcacc aaagttgaaa tgaacgaaaa aatgttaagt gcagccagag agaaaggtcg 14100

ggttaccac aaaggaagc ccatcaaact aacagctgat ctcttgccag aaactctaca 14160

agccagaaga gactgaagac caatattcaa cgttcttaa gaaaagaatt ttcaaccag 14220

aatttcatat ccagccaaac taaccatcat aagtgaagga gaaataaaat cctttacaca 14280

caaacaaatg ctgagagatt ttgtcaccac gaggcctgcc ctaaaagagc tcctgaagga 14340

agcactaaac atggaaagga acaaccggta tcagccactg caaaaacatg ccaaattgta 14400

aagaccatcg aggctaggaa gaaactgcat caactaacga gaaaataac cagctaacag 14460

cataatgata ggatcaaatt cacacataac aatattaacc ttcaatgtaa atgggctaaa 14520

tgtccaatt aaaaggcaca gactggcaaa ttggataaag agtcaagacc catcagtgtg 14580

ctgtattcag gaaaccatc tcacgtgcag agatacatat aggetcaaaa taaaaggata 14640

gaggaagatc tactaaggaa atggaaaaca aaaaaaggca ggggttgcaa tcctagtctc 14700

ggataaaaca gactttaaac caacaaagag acaagaagg ccattatata atggtaaagg 14760

gatcaattca acaagaagag ctaactatcc caaatatata tgcacgcaat acaggagcac 14820

ccagattcat aaagcaagtc cttagtgcac tacaaagaga cttagactcc cataaaataa 14880

taatgggaga ctttaacacc cactgtcaa cattagacag atcaacgaga cagaaagtta 14940

acaatgatat ccaggaattt aactcagctc tgcaccaagt ggacctaata gacatctaca 15000

[0074]

gaactttcca ccccaaatca acagaatata cattcttctc agcaccacac cacacttatt 15060
 ctaaaattga ccacatagtt ggaagtaaag cactcctcag caaatgtaaa agaataataa 15120
 ctataacaaa ctgtctctca gaccacagtg caatcaaact agaactcagg attaagaaac 15180
 acactcaaaa ccactcaact acatggaaac tgaaaaacct gctcctgact actgggtaca 15240
 tatgaaatga aggcagaaat aaagatgttc ttgaaaccaa cgagaccaca acatactaga 15300
 atctctggga cacattcaaa gcagtgtgta gagggaaatt tatagcacta aatgcecaaa 15360
 agagaaagca ggaaagacct aaaattgaca ccctaacatc acaattaaaa gaactagaga 15420
 agcaagagca aacacattca aaagctagca gaagacaaga aataactaac atcagagcag 15480
 aactgaagga aatagagaca caagaaaccc ttcaaaaaat caatgaatcc aggagctggc 15540
 tttttgaaaa gatccacaaa attgatagac cactagcaag actaataaag aagaaaagag 15600
 agaagaatca aacagacaca ataaaagatg ataaagggga tatcatcacc aatcccacag 15660
 caatacaaac taccatcaga gaatactata aacacctcta cgcaaataaa ctagaaaatc 15720
 tagaagaaat ggataaatc ctggacacat acacctccc aagactaac caggaagaag 15780
 ttgaatctct gaatagacca ataacaggct ctgaaattga ggcaataatt aatagcttac 15840
 caacccaaaa aagtccagga ccagatggat tcacagccga attctaccag aggtacaagg 15900
 aggagctggc accattcett ctgaaactat tccaatcaat agaaaaagag agaatectec 15960
 ctaactcatt ttatgaggec agcatcatcc tgataccaaa gtctggcaga gatacaacaa 16020
 aaaaagagaa ttttagacca ataccctga tgaacatga tgcaaaaatc ctcaataaaa 16080
 tactggcaaa ccaaaccag cagcacatca aaaagcttat ccaccatgat caagtgggct 16140
 tcatccctgg gatgcaagge tggttcaaca tacacaaatc aataaacata attcagcata 16200
 taaacagaac caacaacaaa aaccatatga ttatctcaat agattcagaa aaggcctttg 16260
 acaaaattca acaacacttc atgctaaaaa ctctcaataa attaggtatc gatgggacat 16320

[0075]

atctcaaaat aatctaagag ctatctatga caaaccacaca gccaatataa tacctaattgg	16380
acaaaaactg gaagcattcc ctttgaaaac tggcacaaga cagggatgcc ctctctcaac	16440
actcctattc aacatagtgt tggaagttct agccagggca atcaggcagg agaaggaaat	16500
aaagggcatt cagttaggaa aagaggaagt caaattgtcc ctgtttgcag ttgacatgat	16560
tgtatatcta gaaaacccca tcattctcagc ccaaaatctc cttaaagtga taggcaagtt	16620
cagcaaaagtc tcaagataca aaatcaatgt gcaaaaatca caagcattct tatacaccaa	16680
taacagacag ccaaatcatg agtgaactcc cattcacaat tgcttcaaag gaataaaata	16740
cctaggaatc caacttaca gggatgtgaa ggacctcttc aaggagaact acaaatcaact	16800
gtcaacgaa ataaaagagg atacaaacaa atggaagaac attccatgtt catgggtagg	16860
aagaatcaat attgtgaaaa tggccatact gcccaaggta atttatagat tcaatgetat	16920
ccccatcaag ctaccattgc ctttcttcac agaattggga aaaactactt taaagttcat	16980
atagaaccaa aaaagagccc gcattgccaa gacaatccta agccaaaaga acaaagctgg	17040
aggcatcaca ctatctgact tcaaactcta ctacaaggct acagtaacca aaacagcatg	17100
gtactggtac caaaacagag atatagacca atggaacaga acacagccct cagaaataat	17160
gtgcatatc tacaaccatc tgaccacgca tgcttagata ttcttgatat tctttctec	17220
tcttttatct ctttgaaaa tcatgggaga ttccaggata cccaccagt ttcttctget	17280
caaaagttat aaatggettc agcaggaatc tagccactgg gcaatagttt cagggacca	17340
gttcaactatg gagtcaactc ttgttacgaa acttggttct cacctcaact ccaggtctgc	17400
tctatgagct ggtggettag gactgggatt gtctgttttc agagggaagg agaggttgaa	17460
atgagctaag agatgtgttt tacaacaaga aactataaaa taacctactg cttgcagaag	17520
ctgaaggaga gggatgagtg acattggtaa ggtcccatgc ctggtgaata agagaggaaa	17580
taccacaaag agagagcgtg ttttgagat caggaggttt tgtgatagag ttagtttgag	17640

[0076]

gtcacagcaa aaatatttgc ctcttagaag acatttggca ggatcaaact gcaggtcagt 17700

gactttctag gcaaagcaga ggtggcttaa aggaagtatt ttaaataact cccaggtaat 17760

tcttctaata acatttaatt tattcattca tccatctaata caatagtcta tgggtcaactg 17820

ctagactccc caggaaaaga aacaatgggt aagacctgat tctgagctca tgctctgggtg 17880

acttaataatc cgtggacata gagccaagat tgtctagaat tcttactgggt gtcaaaggct 17940

gtaagcctga tctctaacag tgactgacca tattttatca taaaaaatat ggaaggacat 18000

tttgtctgct gagactgget tcctgtgaac gaggcacata tcagctttgg agtcttctct 18060

ccaacccttg gccacgetgc tgtgcttgct cctggacttg caagcacacc ttcattgccc 18120

tgctgcccac acacacatac acacacattc ataaaggcac agcccatecc attctcttcc 18180

ttcatgggta ctatcatctt tgtaaggag aatgaggaac ctagcatgac attccctcag 18240

gcattatcag gtgatccagt gtcccacaa tctcattgga aaatagactc atttggtatg 18300

cacataaatt tcacagaggt ctccagaata tccactgaat tgtgaattcc tagaggtcac 18360

aaaacacaca cagtgtttgc ctctataaac ctaagtgttt ggatagaaaa tccactgaata 18420

cctaggagct gacaaagcag tatcttcttg attactgct atgtttggat atcaacattt 18480

ttcactcttg acctgaacac tctcaagggt aggggaactac atctgagact agaaaagaaa 18540

actattggag atggggcacc tcttagtatg acagaaatgt aaagaccata tactcctacc 18600

ttctagacca aggttatatg aaagaaggga ctatcttttt caaaagctgt attctaactt 18660

atgccataca gtaggttttc agtaaatgtt ggatggatgg atgaaataaa ttgattaatg 18720

gatgaatgta tctaatacaa aattgtatca gattccttac tctcctgtca ttcactaaga 18780

agattaaata aatcatttga tcccatcac cttcttgaca agaaataatt agacttctca 18840

aaatcctgaa tgtacactgt acattcctaa gagagccttc gctagcctac ctatattttt 18900

aaataaaaatt ttcaaagaca gggactgaat attgtccatt tttctgtctc acacactgcc 18960

[0077]

ctcgttcttt taaaaaatga cattgtttta gtgtcttctt tgtgccttat catgtgcaaa 19020

agcctcactt gatggaacca acaattcaga gaggaagggtg ctaataggcc tgttttacag 19080

ttgaggattt agattgtcta tatatagtct ttgaggtcac tgggctcata aacacagagc 19140

tcttctgatt ccaattcaaa aaactattgg cttattttaa aaactgcctt gttacacagt 19200

agacatttta atttctaagt gaagaagttt agtcacaatg atcatcacca ccaaattctga 19260

ggaaaccatc ttgattcact tattaatcct cctagcaggg ttctgcaggc cacacagcaa 19320

catggggtea ggaatgaggg ctgagcagaa ggggaggagt gtatcctagg aaggaaatgg 19380

tctctgactc aatgccacaa agaaacctgc ctagctgtgg aagcagggtt tggaanaacta 19440

tacttgggga tgggaagaag ctggaatagc atacttccac aggaaagcta tggcaaggat 19500

ctctggtgtt ctcaagaagc actctttgag cctttgtcca ggcacaagct ccttaagaat 19560

cccctggcag ctccacgggg accagcctgg ggggggctgt gactggttaa agtccttta 19620

gaccccttc ccagaggag gtcttgggag ctcccagtgc ttcttggat ttagggaca 19680

aggaggcagg ggaatgtgag gaatgacca gcacgtgcct ggttgtgcaa gagccacttc 19740

ccgtcttcc cagagccaag cctgagccta taaaagacat attcagctcc agcacctcat 19800

ctgctctgac ttcccagga cgtgtctgtg ctctgtgtg tgaccagggt gattggcaac 19860

ctgggagcca gagggcacag gggaatggga agattaaggg aggttcaga gatttatttc 19920

tttgattctg tgctattata tgcagaaaag agatacagtc tgacttgagt ttctggaaat 19980

cttgccttgg gctgtgagaa atgtgtatct gacagtttca tcaagggtt attcttatat 20040

ttcaaaaatg gatacttgtg tgggtatgaa tactttatc aagctcatga agattagaaa 20100

atactgcta tcttctaaaa gaataagaag cttcttatta tatctgattt tcaagtatat 20160

attttcattc agcaggattt gttttcttct acttgttgat atatatgtgt gtatgtgtat 20220

atatatacac aactgagta ttattcacca tcaccatag atttagtttt aaaagagata 20280

[0078]

aaattgcecat ccttggtttt tcacagctaa agggtttgac tatttcgaat ttcaggttga 20340
 aaaagtegeca ctgagatgtc ctgccagcaa aaccagcage agtgccagcc cctcccaag 20400
 tgccccccaa aatgcccacc caagtgtcct ccaaagtgcc gacctcagtg cccagcccca 20460
 tgcccacctc cagtctcttc ctgctgttgt cccagctctg ggggctgtg cggtccagc 20520
 tctgggggct gctgcagctc tgggggtggc ggctgtgcc tgagccacca caggccccgt 20580
 ctcttcacc ggcaccggca ccagagcccc gattgttg agtgtgaacc ttctgggggc 20640
 tctggctgtc gccacagctc tggggactgc tgctgaccag acctegaaca tcacagagca 20700
 acccttatgg agaaacttgc aaccaggacc tgteccagag tgatgcttct cctgcccctt 20760
 ttctccttt ccttgggctg acacacctg tgagggttt tgtctgttg catggccaa 20820
 gagcccatcc tggatcctga tcttacctc ccactttacc tcatacaaca ataaagctct 20880
 ttgcctctt cgtgaagtac cctggccatt gtcaattctt ctattctttt caaatccctt 20940
 aaagtccagt gcccagccg tacactttta aggatgcttt ggtaggagtg gagtgtgaga 21000
 aagaggtatg ctggaagcag tctccagagg gtcaatcctg aaagagcagc ctcacagtag 21060
 ggcttctcaa tgggactgcc tgtgtgcact tggagactgt gagatgtgga ccctctcagc 21120
 ccaggcaggg gcctgaaggg actgacagcc ccagggtga gaaccgctgg ccatgtgtg 21180
 ctgtctcagt ttaactcacc atgtgaatg taacttattt tcattttctc ttcataccat 21240
 gatgacaaag ctagaaaaat ctttggtact ccaagcaaaa aggtatattt aacagaattc 21300
 ttttcactg atttaaaatg cttggtatga gaatcagaac ccttaggggc aagtttaaat 21360
 gacaactcat tgccaactaa ctactctcca gggaggaaga gtctgtccag ttcttgettt 21420
 ctttcatgtt ccagctctc cacgaaaca ctgttaaata aatacgtaag cagtgtgaag 21480
 ctgaaggaac acctcagtat gggcctgctc agttcctgtg actcaatcca tggctaagt 21540
 aagcagagta tggaacaggg gtccgggcca tgteccactt ttattgctac gcctgctca 21600

[0079]

ctgcaaagga tcaggcctgg aggggtgtgaa ggagaattat gagaaatacc actctcaaac 21660

aatatatatg aaacaaaaca taaacaaaaa taaacaaaaa agaaaaacca ctcttatgta 21720

gctgtaaaat aagctgcatg tagcccagcc tgggtctagc cctccttagg ttcaggctctc 21780

ctcagctcag cactcatatt aaaagaatth taggacggag aagcagcttg tctggatgac 21840

tggggctctg tgccttgggg aggaagtgtg aagaagagtg agggcaccac taggggaatt 21900

aaaaatagtt agtctatcag cagactgtga ggtgcttgcg gtgagatgat gagctcagtt 21960

tagagcaagg ccctgctgaa aatgggctaa taaagcatcc tgtcccttct tctgcatgat 22020

acggatgggt accaattaac atctccactt gtgtttcttc actctagtec ctccctatca 22080

tgcttttaggc ttctttcttt gctccttttt gtaattattg caactcctat aggtaattat 22140

atgtcattgt tagtttaatt tgaaatctca tagtactaat tctgagtcta tgttcatata 22200

tgtattgccc acacttttct ttttttgttt ttctgcaaac tacatattca tatactctga 22260

cttatttttc actatactcc ttgttttttt ttttttcttc tggctctatct taagtatttt 22320

ccctctgcat tagccagggt tctccaaaga aacagaacca ataggagata tcggtctcga 22380

cagaaatata gatatagtta tagatatagg tagagataca aatttatgac aagaaattgg 22440

aaatggattt atgattatga aggctgagaa atctcaagat ctgtagttgg aaaactggaa 22500

acctagaaga gctgatagtt ctcatctgag tctgaaggca tggaaaccag gagagctgat 22560

ggtggaaatt gcagtctgag cacatgagaa gatcaatate ccagcacaga gccgtcaggc 22620

agaagaaatt tctcttact cagccctatt gttctattct ggtcatcaat tgattggatg 22680

aagcccactc atataggga gagtaatcta ctttaatcgg tctacaaatt caaatgttaa 22740

tttcatccaa aaacatctct acagaaacag cccgagtaat gtttgaccaa atgtctgggc 22800

accttatggc ccagtcaagt ctccataaaa ttageccagca tacttctccc ctctctcctt 22860

atattttaat gagccatatt ggcttatagt aacattgttt gctataagac attttaaatt 22920

[0080]

ttctttttatc ttgaaagttt tagactttctc ggtaaagaga ttttttttca tgtttcttaa	22980
ttatacatgt agtcttcaag actttctgct aaggttttat ttgcttgctg ttttacttcc	23040
taacttggtta gtatattcat ttttattgtg agatacaacc tatatacaaa agaaaaaata	23100
aaatatattt atgtacttta aaaagaaaca caaataatac caaccacagt gaggataaga	23160
aagagaagac tgctagaact ccagagggtgc cccatgttgc ccaatcaca gactcttctc	23220
cacccctgcc aaagtaagca tgtcacaatt ttttaatttctc attgtagttt taccatcgat	23280
gtatgcttct cagttgagtt tttaagcttt ataaacaagg agtcacattt ttatgtaatt	23340
ttttctctt tgttcaatgt ccagttgaga aggtaaacac ta	23382
<210> 4	
<211> 8818	
<212> DNA	
<213> 人类	
<400> 4	
cttagggaca gcaatcagag caggcaagga gagcagccag gaggcttgcc tgaaagcaag	60
cacctttgtt ttctcttaga ttattcagtg acagttctgt aaaacagcaa agtttgtag	120
tcaggtcagt cttcttaca tattgagaat gttgcaattg cccataatga acacaaaggt	180
gcctcagcag agacacaatc tctgatttta aatccttacc ctgaagagac agcaccggg	240
aaaacaagtg ttctcaggt ctaaggccca tgggaaggga ccgcacaga tgtgggaagc	300
aaggtagcac ttctggaagc attggagatg aatcatattc acttcagtgg gcaaaagcca	360
agcagatgca gggacaacca gaccataaat ggaaggaggg tctgagtctg acaaagagac	420
aacgtgcacc cttttgtaga agttcactaa aacctgacca agagaaagct caacttcaga	480
gcctttctgc ttttattgcc agagatagct ctaactttgc cagtgtacc cttgctaatt	540
cagttcctag aactctgtgg tggagaagtt gtgcagaggg ttgtcattct atttctggc	600
ttaggtggta gtcacacaga tgttttactt tgtttaaatg cattggctga aatgcattgg	660

[0081]

gctgtattct cgatgcatgt atgttacagt tcaacaaaac cacattaata taatacacac	720
aatgaataa atataaaaat atccaggaca attcttaaaa ataaaattca tgtgtaggag	780
ctagttegcac tagatgtata acatattata aagctattat agttagacag tttgtactgg	840
accatgaata tacagacaga tcaatagagc atgataagtt cagagctggg aaaaaatata	900
cagaagaagc tagtacatga taagcacacc ctcaaatctg tgaaaaaaga tcagttaata	960
aatgatgttg ctacaactgg gaagccatta agaaaataaa aactctcatc ttatactagg	1020
atgagttatt taagggttaa aacattaataa gttaactctt aagaaaatct gggggtattc	1080
ttctacactt tcaaaatggg gatcactttt ataattataa tacaacaaac aagaactaga	1140
aaataaaaaa attttaaaac gtgtttacat aattatcatt ttctatatgg ctgaaatata	1200
acaaagtcaa aggtcaaaaa acacagagga agaaatttag ttctcatag cacaaaaata	1260
acttaatttc tccaatatta tataagtgtc acaaatcaat aataaagtca tcaagaatac	1320
aattaagtat taggtaagtt atgtcatcag gcagttccca gaaaatgaat gttgtggttc	1380
actaaatgta ataaaatata gttgagctta gttactaaga gcaactcaaa actactttga	1440
gateccattt ttaacctacc agatgggcc aagaacaatta gtttaaaaat acatcatttt	1500
gatgagtttg tggggaaaaa aatgcagtct tatacattgt agggagtga ataatcagt	1560
gaaactcata tggaagaaac tggaataact gatcaaatca caaaagtatt tagcctctga	1620
gtcagcaatt gcacttaaat aaatatattc tatctcctga ggaatgatgt ataaataatg	1680
tgttcaactgc agcattgttt gcaagagcca aaattgggaa ctgctatata tacatagggt	1740
attatccaga ctgggaaaca aagagagaaa gtcttaatgc actgaagtgg aatgctctcc	1800
aacagatgct gttgtgaaaa aagcaatgag caaaatgatg tggagttact cttatttcat	1860
taaaagaaga aaaaatatac atgtttgtgt ggggtttttt gtgcataaaa taacctgaa	1920
aggatatttc agaaattgat ctctgtgaac tgaagtaatt ggagtttcaa gttggggatt	1980

[0082]

tcactgtgta tcctttgtaa cttctgattt taggcccag tgaatgtgtg ataatgatca	2040
ctattaataa tattaaagga gatgggcat gctccgtggt gcatgcctat aatcccagca	2100
gtttgggaag ctgaggcagg aggatctctt gagctcagaa gttcaagacc agcctgggca	2160
atgtggcaaa atcccatctc tacaaaagat acaaaaatta gccaggtgta gtggcacgca	2220
cctgtagtcc cagctacttg ggagactgag gtggaaggat cacttgagcc caggaggtca	2280
aggetgcagt gaggcaagat cacaccactg cactacaacc tgggcaacag ggtgagaccc	2340
tgtctcaaaa aaaaaaaaaa ggacaaaaaa aattgcatta aaggagacct gcaggaagca	2400
gagcctcaaa tgctggaacc caggaagaca gcaccataag gaetgctcc gctccgttac	2460
tggatgtaat cacatcatta cattagcaaa ttctgcaaa ctggcctcaa cttttctcac	2520
catcaattaa catagcccat gcagcaaaga aagatggtaa tagtcttccc cacttttgat	2580
aacagaggca gaaaaacata aggaggtagg agtgtggctg ccgcgggaca cattctgcta	2640
gtttggatga tgcatgaaa acagccttaa aaaatgagtt cccaatttg gacaagcaag	2700
ttgggagtta ttggcagaaa aagtttgatt taagaaaaaa agaccaacaa tcaccattca	2760
tgcatttaac taatatTTTT gagtatttac tatgtattag gaactattct gactgagagc	2820
ctgagatata tcagtaaaca acatttgcct ttttggcaat gacattctaa tgagagaaga	2880
caatacataa aactataata aataaaaatt agctcatttt agtgetatga gatactggtc	2940
tttatattat acttatatta tattgtatgt gaaaagtact accaaaggaa aacaaaatag	3000
gactgtgtaa aggtgactgg agtgagggca agtgtggaag attgtcattt taattaggaa	3060
aatcagtgtg tacctcaccg agaaattgat atttaagcaa agacttgaaa tacaacaaa	3120
aggcccttag tggcatgctt gtgcattcgt atgcatgtac acatgtgcac acacctggcc	3180
tgctgcagga acagcaagga ggtgagtgtg gctgcagcag agcaagtgc aggaagaaca	3240
gtaggagatg aggtccaagg tgtgtgtgga gttgggggtc aataatgcag ggccttgtgg	3300

[0083]

gctattgtga ggacttcaat ttaccacagca aggcattggaa agccacttca ggtttttgag	3360
caggagaatt acatgatatg acttatgttt taaagtccac cccaactgct ctgctgaaaa	3420
cagacaatag aggaccaaga atagagcaga gataccagtt aaaagctgat tagaaaagtc	3480
agcaagaaat catggtgtct cagagtaggt tggccacagt tacagatgga ttcattttga	3540
agttaacaga atttactgat attaactgag tgtgagttat acaagacaga agccaaggag	3600
gactcaaaag tctttagctt gagcaaatgg aagaatagag ttagcatctt ctgggctgag	3660
atgtctaggg gtagaggatg ttttcaggaa aatatcaata gctcagttgt gagcacatta	3720
aatatgagat gtctattaga tatccatgtg agaatactaa ataaataagc agttgaatat	3780
gaaagtctga agttccggag agattttctca gctgttatta acaaagtaga caccatcatc	3840
aatgacatat agatggcctt catggtcagg agactgaatg ttctaacca gaaaatgagt	3900
attgatggag aagaaaagag aaccaaggac tccagcagga agtagccagg cagaagaaca	3960
ggaacacaaa gagttccaag ttaatgtact ggctaaatta ttcttccatg tattttatta	4020
ctaaatgtgg atttagtacc taggatgttc catacattgt gttggaaata gaatttaa	4080
ggtgagcaga aagtatatct atcctcacct tctagatagc cttcattagc ttaatttta	4140
caaaactaata tgcattcatgg ttaaaaattc aaatagtttg tgtgtatttg attcttctag	4200
gaaaatattt atgtatacta acatatatgt ttatcttatt ttgtatatat ctccgtatat	4260
acattctttt tataagaaaa ggaatatact aacatatgtt ttactggatt gctttctttg	4320
attacatagt tgggagaatt tttcatagea tcatctacaa ctggactctg tttttattag	4380
ctatatacta tctattaca tgattgtaaa atgaaatctc attaacagaa tatgaaaata	4440
cctgtttctc aacatatcca ttaacaacta atgtgatcaa aatttcatat ttgttgccag	4500
actgtcatag aaaagccagt gtctcattga tcttctgggt tgcattcctt tagttatgaa	4560
taagcatttc ttcataagatt tgttgcccat ttccacctct ctctctcact ctacctctct	4620

[0084]

ctccctctcc gtactcttcc tagaaagaac ttaattctta tagattaagg ctttggctcc	4680
aactacaatg ttagagactt tgggagatgt gaataagagt agacatttta ctgagaataa	4740
attttagaga taactataag catgaagacc tgaagaagac aagtaaaaga gaaatacttg	4800
aataggctat tatagaatga aaagcacaat tcacatactt cccaagagtg ctacccttg	4860
gccgaggtgg ttctgtcctg gacttcagac tttaaagagt cttctcttg ctctctcca	4920
gccatgaccc acccccatag gatgattcaa taggtccagg gtacaggecc agagctgctg	4980
tctctctcac tccagctatc caggacceca gaaatctctg tcaactggcc ctgagatcac	5040
ctttatgtgt ctatctctt gaagactcaa acageccaata ctgtgctcac ttctaggtct	5100
caatgaccaa gagaggggca gtttgagaga tccatgtatg aaaagaacat ggacggtatg	5160
gactaggata tttaactca tgcacgccc aggtccctcg tgacacaaga tgaagccagg	5220
gagtaggaag agaaccagg gtgaaccaca ggggtcagaa gcctctgttt gtgctcttgc	5280
ctggtgtcct acaaagtgtt atggcaggca tgcttccag aagaacagaa ttctctagtc	5340
actgtcctca aagaggccat gctacaggaa gaaaaggaaa gagtaaacad aatgaaatgc	5400
ttctctctta tcattatctt agtcaattca ctattcaatt gattcagttt cagaatcttt	5460
tatgtgccaa ttgtcagcc agacactggg aatatcaaag tgaataaaat agggctcctg	5520
cccctaaggc atcccagttt ctcccctgtt taaaaccttt cagtgggtcc cattgtctc	5580
aggataaagt ccaaactctt ttcttatgg cataaaggtc ctaccaaca cagccttacc	5640
tacctaggta ggtgaatctc ccccttctcc ttactcccca ctctgctcca gccacaactga	5700
gtttttccac caaatagcct atactttcac atttctggat ctgaaaatgt cctccagac	5760
acaccctat gtccaacttc ctatttccaa ttccaagtc tcagcttaaa tactactccc	5820
tccaggaagc atatctgacc cccaaacctg acccctgtgc ttactctacc atcattatta	5880
tcacatggta ttgttgcta ttgacttgc tccatattga attagatctt tgagaatggg	5940

[0085]

aactacatcc cattgagcta cacttcactt tttaggacag agcctggtag ataggagaca	6000
atcaatacag atttgtgttt taggttattg catgcttatg ggtagctag ggagctttct	6060
ttttatcttt caggaaagaa accaaagaga cagactctcc agcaggcagg gggcgccagt	6120
gagttagcta ggcatagagt cttgtggaga gaattatggt gtcatagagg cttagaatct	6180
tggcattagc agggaaggga acccatctgg ttcagaaggc tttggcttct gatagtcatg	6240
gactcactag gctgctgagg aagatcaata atacctactg gaatcagtca tgagaagtea	6300
agcatgaaa ttgtgaattg tgtgtgtggc cagaccaggt aacacagaga aacattgtct	6360
tttgggtcct ctctccatgg ctttttcatg tgttcattag ttcattacgt atttacetac	6420
tgagacccat gcaatgccat tcagtgtgtt ggatgagagg attccaggat tcacctatc	6480
tttcaggaaa aagctcctgc ttctatttaa gaggaagaca tgggagccac tgggggtggg	6540
gatagtgaga atgcctgtgt ttatggaact aacagaccac aaaagtgaga caattgggaa	6600
aataataacc actataaaaa tgtaatggga agtacagggg atccaagaaa gcctttcagg	6660
gagatggagg ttttgggtta aggaagtttt cctcgaggag tgatgtgtaa gcagagactt	6720
caggtttgaa cgatattcag tgagcatgtg gggacagaga cccaggaagg gtggtttaag	6780
ctgggaaaga gcacattcaa agacctggaa agaggtgaga aaattcagca ggtgtctgct	6840
ctgagccagt ctggtgcctg agggaacagt gtttggacat gaggetgggc ctgtgggcag	6900
agggagaagg tgccttgcca gagecctcct ggtggtttgg gctctgtatc aatcctctct	6960
gaatagagca ataagageta ggaagatggc cctggccttg caccacaagt gtttgacatg	7020
attaggcaga tttcctgagg caggetgtgc acagcatggt tgggaaggaa gggaaatgga	7080
gettgagag aatatttatt tatttgtttg tttcttttta tttatttatt gagatggagt	7140
ctcactctgt caccaggtt ggagtgcagt ggtgaggtct cggtcactg taacttctac	7200
ctcccggatt caagtgattc tectgcctca gcttcccaag tagctggtat tacaggtgca	7260

[0086]

cgccaccaca cccagctaat ttttgtat	7320
caggttggtc tcaaactcct gacctcaagt gattcacccg tcccagcctc ccaaagcgct	7380
gggacaacag gtgtgagcca ttgtgcccg ccaaggagag aatcttaaga agaaactctt	7440
gcaggaggca gaggggtcca ggaggcaggt ttgttggaca ggagccaggg ttctgaaagt	7500
ggaggatata gatactattc accaggcaga aggaagacag tggagacagc actggctcag	7560
aggggcagga cctgcatcca tgggggtgag gttgcagaag gacttgatgt tgcctccct	7620
atgtacagt caacaagact ggtttccaac ctgctgtcct tctctgatat tgcacaacct	7680
ccacattcag gacctattcc caggatctca ctatggggga catgcttga gtggggacat	7740
tggtgggact ggtgaggcaa gcccctggac ctgggctcca ggtgcctctg atgcctctgc	7800
acctctgac actggaatga gaattgcttc tgtagaaaca gtcagctctg gacttgtaaa	7860
gatggacact aacaaactgg acttccaagt ggtgatgaga ctctgcactt tgggtcaattc	7920
agggcactga ccttcagagt agagaaaggt aaaccagaga caaaagacag aaaggtagag	7980
acagaatgaa gaccagagaa tgatagagag gcagagagga aaaaaaata gcaaaaggca	8040
aaaataaaca cagagaggaa caacagaaaa taaaaacaga aggggctcac tctcctctcg	8100
tccctctcaa aacctagtct taggtctgac ctctaaggga gcagattcaa acaacaagga	8160
gatacggaca gatactgggc tccctgcata ccccgacaca cccaaattca ccactgtaca	8220
gtgcagggg cacccaagtg tgtgcttggt tgtgcgtgtg gacacacaca cacacttct	8280
gtcaaacagc ttgaaacct gcaaagggtg tagagaacat ggcctgtac tctctcact	8340
cttccctatt ttatactgg tctgtctcgg tcttatttcc atttttggtg tgcctctcca	8400
ttgtcacct ctctccttcc tgttctctct cgggcaagtc aaggcctgc agggcctgta	8460
tctgccacct cagcctcttt gcagcttctc ctgacctaga aggagcaata ggggtgacttg	8520
tccccacccc catcctttag gagtctttgc tctagtgtct tctctttaa tggaagcagg	8580

[0087]

ggccttttcc tgaaagatag gatgaatcca ggaataactg gcagagatag agtcatgact	8640
gaattaaagc caacatccta acccaagtca ctctccccctt tttagggtggg aactgccctt	8700
gaccagttc aggetgcagc catatagcca gggacttagg atcatgaaag gtttatacca	8760
ggtggcaggg ccttaagaga taactcttcc ctcccaatct tctctaccc taatgctg	8818
<210> 5	
<211> 3440	
<212> DNA	
<213> Homo sapeins	
<400> 5	
tggcctaata atgagaacaa actcattttg aaagtggaaa aattgagatt cagagcagaa	60
gtttgactaa ggtcacaaaa cagtaggatg cctcactcag ctccctgtgc ctaggtcaga	120
aaagcatcac aggaatagtt gagctaccag aatcctctgg ccaggcagga gctgtgtgtc	180
cctgggaaat ggggccctaa agggtttget gcttaagatg cctgtggtga gtcaggaagg	240
ggttagagga agttgaccaa ctagagtggg gaaacctgtc catcaccttc aacctggagg	300
gaggccagge tgcagaatga tataaagagt gccctgactc ctgetcagct cagcactcca	360
ccaaagcctc tgcctcagcc ttactgtgag tctggtaagt gtcggatggg agaaccaggg	420
ttgggactcg ggacctcaa cagcatacga tgtggtgggg gtgggcagcc tgggtggggg	480
tgggcattac tctggggctg gattcagctg gactttcatt ctagggggac tcgagtcaga	540
gtactgagag aaaagtgcct tggcacagaa gtgcagaaca gagagtaatc atcctatgtc	600
ccatcttttc ttgtgaccat atttttggat ttgtgtgtga gagagaatta tggaaggag	660
gaggggaata gcattcaact tctttcctaa acctcttggg tttgacaga ccatcatttt	720
gccttcttta tggagggaga ggttcagga agagcttcta ccttttggct atgctgcaca	780
gagggatggc agaatgggga aacctttcta tttggagaaa cctaggcaga gctgggacag	840

[0088]

gaaaactcaa cttagaagta taagacttgg aagaacaacc tccaactctc agcaaccttc	900
cagctcccgc agccccaccc cagacacaag gactgcagct aaacctcaga aggtcaggag	960
agaaagcagc cctggggttg aataggccaa cctgctggct ttacaggggg gaaaacaaa	1020
tcccaggaga ctaagtgaca tgcccagaaa cacacagcat tccaatggga gattcaggcc	1080
tagagcatgt cctgtggctc cagtctggag gtcacaccat gacctcttag gtcctctctg	1140
gcacggccta ttggttttct aggacttggg gttctccaag agacatttca ttcctaagg	1200
cettactcct cactgtgaca taateccaga acgcctctct gctccttggg cagtgaagcg	1260
atgagggtgg acacaaggac tagacaagag cagacagtga gctggcacct gacccaccct	1320
tgcagaacag ccctgcagac agatctcctt gttggctctc acctgggaac aaggagctc	1380
ctaggaggac ctttctctgc ccctccacat ttccaccctt ctctctctgc tgcttttggg	1440
aaatgatagt ccagagggtg tagaacagta ccctgcccaa gggaagaggg gatgctaaaa	1500
aaccagatac ttctgcagat tcccaagggt tcctctatct cctttgcctt cagcctgtgc	1560
atcagacctc ttctgtcttt cagggtgaca gtagcttcta agatgtccca gcaacacaca	1620
ctgccagtga ccctctcccc tgccctcagt caggagctcc tcaagactgt tcctctcca	1680
gtcaataccc atcaggagca aatgaaacag ccaactccac tgccctcccc atgccagaag	1740
gtgectgtcg agctcccagt ggagggtccca tcaaagcaag aggaaaagca catgactget	1800
gtaaagggac tgctgagca agaattgtgag caacagcaga aggagccaca ggagcaggag	1860
ctgcagcaac agcaactggga acagcatgag gaatatcaga aagcagaaaa ccagagcag	1920
cagcttaagc aggagaaaac acaaagggat cagcagctaa acaaacagct ggaagaagag	1980
aagaagctct tagaccagca actggatcaa gagctagtca agagagatga gcaactggga	2040
atgaagaaag agcaactgtt ggagctccca gagcagcagg aggggcacct gaagcaccta	2100
gagcagcagg agggacagct gaagcacccg gagcagcagg aggggcagct ggagctccca	2160

[0089]

gagcagcagg aggggcagct ggagctccca gagcagcagg aggggcagct ggagctccca	2220
gagcagcagg aggggcagct ggagctccca gagcagcagg aggggcagct ggagctccca	2280
gagcagcagg aggggcagct ggagctccca cagcagcagg aggggcagct ggagctctct	2340
gagcagcagg aggggcagct ggagctctct gagcagcagg agggacagct gaagcacctg	2400
gagcaccagg aggggcagct ggaggtccca gaggagcaga tggggcagct gaagtacctg	2460
gaacagcagg aggggcagct gaagcacctg gatcagcagg agaagcagcc agagctccca	2520
gagcagcaga tggggcagct gaagcacctg gagcagcagg aggggcagcc taagcatctg	2580
gagcagcagg aggggcagct ggagcagctg gaggagcagg aggggcagct gaagcacctg	2640
gagcagcagg aggggcagct ggagcacctg gagcaccagg aagggcagct ggggctccca	2700
gagcagcagg tgctgcagct gaagcagcta gagaagcagc aggggcagcc aaagcacctg	2760
gaggaggagg aggggcagct gaagcacctg gtgcagcagg aggggcagct gaagcatctg	2820
gtgcagcagg aggggcagct ggagcagcag gagaggcagg tggagcacct ggagcagcag	2880
gtggggcagc tgaagcacct agaggagcag gagggacaac tgaagcatct ggagcagcag	2940
caggggcagt tggaggtccc agagcagcag gtggggcagc caaagaacct ggagcaggag	3000
gagaagcaac tggagctccc agagcagcaa gagggccagg tgaagcacct ggagaagcag	3060
gaggcacagc tggagctccc agagcagcag gtaggacagc caaagcacct ggaacagcag	3120
gaaaagcacc tagagcacc agagcagcag gacggacaac taaaacatct ggagcagcag	3180
gaggggcagc tgaaggacct ggagcagcag aaggggcagc tggagcagcc tgtgtttgcc	3240
ccagctccag gccaggtcca agacattcaa ccagccctgc ccacaaaggg agaagtattg	3300
cttcctgtag agcaccagca gcagaagcag gaggtgcagt ggccacccaa acataaataa	3360
ccacccgcag tgtccagagg cctcagatc gtctcataca agggaagaga gagccactgg	3420
ctccacttat ttcgggtccg	3440

[0090]

<210>	6	
<211>	20	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	
<400>	6	
	agcacatagc aggcaactagc	20
<210>	7	
<211>	20	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	
<400>	7	
	cgattgtgcc actacacagt	20
<210>	8	
<211>	23	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	
<400>	8	
	aaaaatgagt ccagtagaag cct	23
<210>	9	
<211>	19	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	

[0091]

<400> 9
agccagattt acatcccag 19

<210> 10
<211> 16
<212> DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸

<400> 10
tatcttgccc tgcacc 16

<210> 11
<211> 16
<212> DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸

<400> 11
aagtgggtct ccccag 16

<210> 12
<211> 17
<212> DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸

<400> 12
ttgctcggcc agagtct 17

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

[0092]

<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸

<400> 13

acgcatacaca cctggctagt

20

<210> 14

<211> 16

<212> DNA

<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸

<400> 14

gggcctatgg ctggaa

16

<210> 15

<211> 16

<212> DNA

<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸

<400> 15

ggctatgctg gggcaa

16

<210> 16

<211> 19

<212> DNA

<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸

<400> 16

tcagttccat aggctgacg

19

[0093]

<210>	17	
<211>	18	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	
<400>	17	
	cattgctgat gctggagg	18
<210>	18	
<211>	19	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	
<400>	18	
	ccttaattgt ggtgttggt	19
<210>	19	
<211>	20	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	
<400>	19	
	aaaaatctgg aaggcataaa	20
<210>	20	
<211>	20	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	

[0094]

<400> 20	
agcaagaccc tgtctcaaaa	20
<210> 21	
<211> 19	
<212> DNA	
<213> 人工序列	
<220>	
<223> 作为 PCR 引物的寡核苷酸	
<400> 21	
tggatagctt tccaccact	19
<210> 22	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<223> 作为 PCR 引物的寡核苷酸	
<400> 22	
acaaggtgac cggaaagacc	20
<210> 23	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<223> 作为 PCR 引物的寡核苷酸	
<400> 23	
agctctggca agttgaagga	20
<210> 24	
<211> 25	
<212> DNA	

[0095]

<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸

<400> 24

atctgggttc actattaaac agagt

25

<210> 25

<211> 20

<212> DNA

<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸

<400> 25

tgggcaaggt agaatatgtg

20

<210> 26

<211> 19

<212> DNA

<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸

<400> 26

gggtgacaga gcaagactc

19

<210> 27

<211> 19

<212> DNA

<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸

<400> 27

ccctgacctc ccttacaga

19

[0096]

<210>	28	
<211>	20	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	
<400>	28	
	aactgtgtcc agcagcaact	20
<210>	29	
<211>	20	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	
<400>	29	
	tatgtgectg ttgtgtgcat	20
<210>	30	
<211>	20	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	
<400>	30	
	agggtcecca aagagccttc	20
<210>	31	
<211>	20	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	

[0097]

<400> 31	
atgcagcac atcctgcttc	20
<210> 32	
<211> 19	
<212> DNA	
<213> 人工序列	
<220>	
<223> 作为 PCR 引物的寡核苷酸	
<400> 32	
aatcacttga acctgggag	19
<210> 33	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<223> 作为 PCR 引物的寡核苷酸	
<400> 33	
actgactggc tgtttctgag	20
<210> 34	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<223> 作为 PCR 引物的寡核苷酸	
<400> 34	
actgcttatt cggagtggga	20
<210> 35	
<211> 24	
<212> DNA	

[0098]

<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸

<400> 35

ccaagagttt tcttagcaaa tcac

24

<210> 36

<211> 20

<212> DNA

<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸

<400> 36

ctgagcacag cagtggctctc

20

<210> 37

<211> 20

<212> DNA

<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸

<400> 37

aaggettatac aagagcgagg

20

<210> 38

<211> 20

<212> DNA

<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸

<400> 38

agacttacag cactggctgc

20

[0099]

<210>	39	
<211>	20	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	
<400>	39	
	tgctcctagg aaaggaaaca	20
<210>	40	
<211>	17	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	
<400>	40	
	gcttctggcc tctgtca	17
<210>	41	
<211>	18	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	
<400>	41	
	aattttgcgt gtgtgtgc	18
<210>	42	
<211>	23	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	

[0100]

<220>		
<221>	无法归类的特征 (misc_feature)	
<222>	(16)..(16)	
<223>	n	
<400>	42	
	tgaagtgtgc attctntaca tca	23
<210>	43	
<211>	19	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	作为 PCR 引物的寡核苷酸	
<400>	43	
	cgagacattt gcatcatca	19
<210>	44	
<211>	1762	
<212>	DNA	
<213>	人类	
<220>		
<221>	CDS	
<222>	(416)..(1453)	
<223>		
<400>	44	
	cgctgagagg ctgagtggag ttcactcaca tggattgagg cccagttcct gggagaagag	60
	atgctgggca ggaaggtgtc tgcattgtggg actctgtaca gcccggtcct ctcccacgtc	120
	tgggaggggc cagagtcaga caactgctgg gttegtccct aagagaggtc atctgactgg	180
	ctgttcagcc taggctgcac acacccccac ttctctctac caggccacac cggaggcagt	240
	gtcacacag gcaagctacc aggccacaac aacgacaccc acctcacctc tggcacctct	300

[0101]

gagcatccac gtacttgcaa gaactcttgc tcacatcage taagagattg cacctgctga	360
cctagagatt ceggcctgtg ctcctgtgct gctgagcagg gcaaccagta gcacc atg	418
Met	
1	
tct gtg act ggc ggg aag atg gca ccg tcc ctc acc cag gag atc etc	466
Ser Val Thr Gly Gly Lys Met Ala Pro Ser Leu Thr Gln Glu Ile Leu	
5 10 15	
agc cac ctg ggc ctg gcc agc aag act gca gcg tgg ggg acc ctg ggc	514
Ser His Leu Gly Leu Ala Ser Lys Thr Ala Ala Trp Gly Thr Leu Gly	
20 25 30	
acc etc agg acc ttc ttg aac ttc agc gtg gac aag gat gcg cag agg	562
Thr Leu Arg Thr Phe Leu Asn Phe Ser Val Asp Lys Asp Ala Gln Arg	
35 40 45	
cta ctg agg gcc att act ggc caa ggc gtg gac cgc agt gcc att gtg	610
Leu Leu Arg Ala Ile Thr Gly Gln Gly Val Asp Arg Ser Ala Ile Val	
50 55 60 65	
gac gtg ctg acc aac cgg agc aga gag caa agg cag etc atc tca cga	658
Asp Val Leu Thr Asn Arg Ser Arg Glu Gln Arg Gln Leu Ile Ser Arg	
70 75 80	
aac ttc cag gag cgc acc caa cag gac ctg atg aag tct cta cag gca	706
Asn Phe Gln Glu Arg Thr Gln Gln Asp Leu Met Lys Ser Leu Gln Ala	
85 90 95	
gca ctt tcc ggc aac ctg gag agg att gtg atg get ctg ctg cag ccc	754
Ala Leu Ser Gly Asn Leu Glu Arg Ile Val Met Ala Leu Leu Gln Pro	
100 105 110	
aca gcc cag ttt gac gcc cag gaa ttg agg aca get ctg aag gcc tca	802
Thr Ala Gln Phe Asp Ala Gln Glu Leu Arg Thr Ala Leu Lys Ala Ser	
115 120 125	
gat tct get gtg gac gtg gcc att gaa att ctt gcc act cga acc cca	850
Asp Ser Ala Val Asp Val Ala Ile Glu Ile Leu Ala Thr Arg Thr Pro	
130 135 140 145	
ccc cag ctg cag gag tgc ctg gca gtc tac aaa cac aat ttc cag gtg	898
Pro Gln Leu Gln Glu Cys Leu Ala Val Tyr Lys His Asn Phe Gln Val	

[0102]

150	155	160	
gag gct gtg gat gac atc aca tct gag acc agt ggc atc ttg cag gac			946
Glu Ala Val Asp Asp Ile Thr Ser Glu Thr Ser Gly Ile Leu Gln Asp			
165	170	175	
ctg ctg ttg gcc ctg gcc aag ggg ggc cgt gac agc tac tct gga atc			994
Leu Leu Leu Ala Leu Ala Lys Gly Gly Arg Asp Ser Tyr Ser Gly Ile			
180	185	190	
att gac tat aat ctg gca gaa caa gat gtc cag gca ctg cag cgg gca			1042
Ile Asp Tyr Asn Leu Ala Glu Gln Asp Val Gln Ala Leu Gln Arg Ala			
195	200	205	
gaa gga cct agc aga gag gaa aca tgg gtc cca gtc ttc acc cag cga			1090
Glu Gly Pro Ser Arg Glu Glu Thr Trp Val Pro Val Phe Thr Gln Arg			
210	215	220	225
aat cct gaa cac ctc atc cga gtg ttt gat cag tac cag cgg agc act			1138
Asn Pro Glu His Leu Ile Arg Val Phe Asp Gln Tyr Gln Arg Ser Thr			
230	235	240	
ggg caa gag ctg gag gag gct gtc cag aac cgt ttc cat gga gat gct			1186
Gly Gln Glu Leu Glu Glu Ala Val Gln Asn Arg Phe His Gly Asp Ala			
245	250	255	
cag gtg gct ctg ctc ggc cta gct tcg gtg atc aag aac aca ccg ctg			1234
Gln Val Ala Leu Leu Gly Leu Ala Ser Val Ile Lys Asn Thr Pro Leu			
260	265	270	
tac ttt gct gac aaa ctt cat caa gcc ctc cag gaa act gag ccc aat			1282
Tyr Phe Ala Asp Lys Leu His Gln Ala Leu Gln Glu Thr Glu Pro Asn			
275	280	285	
tac caa gtc ctg att cgc atc ctt atc tct cga tgt gag act gac ctt			1330
Tyr Gln Val Leu Ile Arg Ile Leu Ile Ser Arg Cys Glu Thr Asp Leu			
290	295	300	305
ctg agt atc aga gct gag ttc agg aag aaa ttt ggg aag tcc ctc tac			1378
Leu Ser Ile Arg Ala Glu Phe Arg Lys Lys Phe Gly Lys Ser Leu Tyr			
310	315	320	
tct tct ctc cag gat gca gtg aaa ggg gat tgc cag tca gcc ctc ctg			1426
Ser Ser Leu Gln Asp Ala Val Lys Gly Asp Cys Gln Ser Ala Leu Leu			

[0103]

325	330	335	
gcc ttg tgc agg gct gaa gac atg tga gacttcctg cccacccca			1473
Ala Leu Cys Arg Ala Glu Asp Met			
340	345		
catgacatcc gaggatctga gatttcctg tttggtgaa cctgggagac cagctgggcc			1533
tccaagtagg ataaccctc actgagcacc acattctcta gcttcttgtt gaggttgaa			1593
ctgtttcttt aaaatccctt aattttccca tctcaaaatt atatctgtac ctgggtcatc			1653
cagctccttc ttgggtgtgg ggaaatgagt tttctttgat agtttctgcc tctctcatcc			1713
ctctgtacc ctggccagaa catctcactg atactcgaat tcttttggc			1762
<210>	45		
<211>	345		
<212>	PRT		
<213>	人类		
<400>	45		
Met Ser Val Thr Gly Gly Lys Met Ala Pro Ser Leu Thr Gln Glu Ile			
1	5	10	15
Leu Ser His Leu Gly Leu Ala Ser Lys Thr Ala Ala Trp Gly Thr Leu			
20	25	30	
Gly Thr Leu Arg Thr Phe Leu Asn Phe Ser Val Asp Lys Asp Ala Gln			
35	40	45	
Arg Leu Leu Arg Ala Ile Thr Gly Gln Gly Val Asp Arg Ser Ala Ile			
50	55	60	
Val Asp Val Leu Thr Asn Arg Ser Arg Glu Gln Arg Gln Leu Ile Ser			
65	70	75	80

[0104]

Arg Asn Phe Gln Glu Arg Thr Gln Gln Asp Leu Met Lys Ser Leu Gln
85 90 95

Ala Ala Leu Ser Gly Asn Leu Glu Arg Ile Val Met Ala Leu Leu Gln
100 105 110

Pro Thr Ala Gln Phe Asp Ala Gln Glu Leu Arg Thr Ala Leu Lys Ala
115 120 125

Ser Asp Ser Ala Val Asp Val Ala Ile Glu Ile Leu Ala Thr Arg Thr
130 135 140

Pro Pro Gln Leu Gln Glu Cys Leu Ala Val Tyr Lys His Asn Phe Gln
145 150 155 160

Val Glu Ala Val Asp Asp Ile Thr Ser Glu Thr Ser Gly Ile Leu Gln
165 170 175

Asp Leu Leu Leu Ala Leu Ala Lys Gly Gly Arg Asp Ser Tyr Ser Gly
180 185 190

Ile Ile Asp Tyr Asn Leu Ala Glu Gln Asp Val Gln Ala Leu Gln Arg
195 200 205

Ala Glu Gly Pro Ser Arg Glu Glu Thr Trp Val Pro Val Phe Thr Gln
210 215 220

Arg Asn Pro Glu His Leu Ile Arg Val Phe Asp Gln Tyr Gln Arg Ser
225 230 235 240

Thr Gly Gln Glu Leu Glu Glu Ala Val Gln Asn Arg Phe His Gly Asp
245 250 255

[0105]

Ala Gln Val Ala Leu Leu Gly Leu Ala Ser Val Ile Lys Asn Thr Pro
260 265 270

Leu Tyr Phe Ala Asp Lys Leu His Gln Ala Leu Gln Glu Thr Glu Pro
275 280 285

Asn Tyr Gln Val Leu Ile Arg Ile Leu Ile Ser Arg Cys Glu Thr Asp
290 295 300

Leu Leu Ser Ile Arg Ala Glu Phe Arg Lys Lys Phe Gly Lys Ser Leu
305 310 315 320

Tyr Ser Ser Leu Gln Asp Ala Val Lys Gly Asp Cys Gln Ser Ala Leu
325 330 335

Leu Ala Leu Cys Arg Ala Glu Asp Met
340 345

<210> 46

<211> 607

<212> DNA

<213> 人类

<220>

<221> CDS

<222> (55)..(387)

<223>

<400> 46

cagggacgtg tctgtgetcc tgcgtgtgac cagggttgac taaactctgc cagg atg 57
Met
1

tct tgc cag caa aac cag cag cag tgc cag ccc cct ccc aag tgt cct 105
Ser Cys Gln Gln Asn Gln Gln Gln Cys Gln Pro Pro Pro Lys Cys Pro
5 10 15

ccc aag tgt acc cca aaa tgt cca cct aag tgt ccc cct aaa tgc ctg 153

[0106]

Pro Lys Cys Thr Pro Lys Cys Pro Pro Lys Cys Pro Pro Lys Cys Leu	
20 25 30	
ccc cag tgc cca gct cca tgt tcc cct gca gtc tct tct tgc tgt ggt	201
Pro Gln Cys Pro Ala Pro Cys Ser Pro Ala Val Ser Ser Cys Cys Gly	
35 40 45	
ccc atc tct ggg ggc tgc tgt ggt ccc agc tct ggg ggc tgc tgc aac	249
Pro Ile Ser Gly Gly Cys Cys Gly Pro Ser Ser Gly Gly Cys Cys Asn	
50 55 60 65	
tct ggg gct ggt ggc tgc tgc ctg agc cac cac agg ccc cgt ctc ttc	297
Ser Gly Ala Gly Gly Cys Cys Leu Ser His His Arg Pro Arg Leu Phe	
70 75 80	
cac cgg cgc cgg cac cag agc ccc gac tgc tgt gag agt gaa cct tct	345
His Arg Arg Arg His Gln Ser Pro Asp Cys Cys Glu Ser Glu Pro Ser	
85 90 95	
ggg ggc tct ggc tgc tgc cac agc tct ggg ggc tgc tgc tga	387
Gly Gly Ser Gly Cys Cys His Ser Ser Gly Gly Cys Cys	
100 105 110	
cctgggctaa gaagaactct ttggacagaa tgtttaagaa cctcctacag cctgatgctt	447
aaccctttcc atttctctc attccattca tgggtggaca gegaccacaa agactcatgg	507
ggcttccctg ggagaacttt gcaattgatg gagcacctca attgcaggtt ttgttttct	567
cctttacctc atgttataat aaagctctga tttctgactc	607

<210> 47

<211> 110

<212> PRT

<213> 人类

<400> 47

Met Ser Cys Gln Gln Asn Gln Gln Gln Cys Gln Pro Pro Pro Lys Cys
1 5 10 15

Pro Pro Lys Cys Thr Pro Lys Cys Pro Pro Lys Cys Pro Pro Lys Cys

[0107]

20	25	30
Leu Pro Gln Cys Pro Ala Pro Cys Ser Pro Ala Val Ser Ser Cys Cys		
35	40	45
Gly Pro Ile Ser Gly Gly Cys Cys Gly Pro Ser Ser Gly Gly Cys Cys		
50	55	60
Asn Ser Gly Ala Gly Gly Cys Cys Leu Ser His His Arg Pro Arg Leu		
65	70	75 80
Phe His Arg Arg Arg His Gln Ser Pro Asp Cys Cys Glu Ser Glu Pro		
85	90	95
Ser Gly Gly Ser Gly Cys Cys His Ser Ser Gly Gly Cys Cys		
100	105	110
<210> 48		
<211> 592		
<212> DNA		
<213> 人类		
<220>		
<221> CDS		
<222> (52).. (372)		
<223>		
<400> 48		
ggacgtgtct gtgtcctgt gtgtgaccag ggttgaaaaa gtgcactga g atg tcc		57
	Met Ser	
	1	
tgc cag caa aac cag cag cag tgc cag ccc cct ccc aag tgc ccc cca		105
Cys Gln Gln Asn Gln Gln Gln Cys Gln Pro Pro Pro Lys Cys Pro Pro		
5	10	15
aaa tgc cca ccc aag tgt cct cca aag tgc cga cct cag tgc cca gcc		153
Lys Cys Pro Pro Lys Cys Pro Pro Lys Cys Arg Pro Gln Cys Pro Ala		
20	25	30

[0108]

cca tgc cca cct cca gtc tct tcc tgc tgt ggt ccc agc tct ggg ggc	201
Pro Cys Pro Pro Pro Val Ser Ser Cys Cys Gly Pro Ser Ser Gly Gly	
35 40 45 50	
tgc tgc ggc tcc agc tct ggg ggc tgc tgc agc tct ggg ggt ggc ggc	249
Cys Cys Gly Ser Ser Ser Gly Gly Cys Cys Ser Ser Gly Gly Gly Gly	
55 60 65	
tgc tgc ctg agc cac cac agg ccc cgt ctc ttc cac cgg cac cgg cac	297
Cys Cys Leu Ser His His Arg Pro Arg Leu Phe His Arg His Arg His	
70 75 80	
cag agc ccc gat tgt tgt gag tgt gaa cct tct ggg ggc tct ggc tgc	345
Gln Ser Pro Asp Cys Cys Glu Cys Glu Pro Ser Gly Gly Ser Gly Cys	
85 90 95	
tgc cac agc tct ggg gac tgc tgc tga ccagacctg aacatcacag	392
Cys His Ser Ser Gly Asp Cys Cys	
100 105	
agcaaccctt atggagaaac ttgcaaccag gacctgtccc agagtgatgc ttctcctgcc	452
cctttttctc ctttccttgg gctgacacac cttgtgaggt gttttgtctg ttgtcatggc	512
ccaagagccc atcctggatc ctgatcttac ctcccactt tacctcatac aacaataaag	572
ctcttttgcc tcttcgtgaa	592

<210> 49

<211> 106

<212> PRT

<213> 人类

<400> 49

Met Ser Cys Gln Gln Asn Gln Gln Gln Cys Gln Pro Pro Pro Lys Cys
 1 5 10 15

Pro Pro Lys Cys Pro Pro Lys Cys Pro Pro Lys Cys Arg Pro Gln Cys
 20 25 30

[0109]

Pro Ala Pro Cys Pro Pro Pro Val Ser Ser Cys Cys Gly Pro Ser Ser
35 40 45

Gly Gly Cys Cys Gly Ser Ser Ser Gly Gly Cys Cys Ser Ser Gly Gly
50 55 60

Gly Gly Cys Cys Leu Ser His His Arg Pro Arg Leu Phe His Arg His
65 70 75 80

Arg His Gln Ser Pro Asp Cys Cys Glu Cys Glu Pro Ser Gly Gly Ser
85 90 95

Gly Cys Cys His Ser Ser Gly Asp Cys Cys
100 105

<210> 50

<211> 2134

<212> DNA

<213> 人类

<220>

<221> CDS

<222> (47)..(1804)

<223>

<400> 50

ctctgcctca gccttactgt gagtctggtt gacagtagct tctaag atg tcc cag 55
Met Ser Gln
1

caa cac aca ctg cca gtg acc ctc tcc cct gcc ctc agt cag gag ctc 103
Gln His Thr Leu Pro Val Thr Leu Ser Pro Ala Leu Ser Gln Glu Leu
5 10 15

ctc aag act gtt cct cct cca gtc aat acc cat cag gag caa atg aaa 151
Leu Lys Thr Val Pro Pro Pro Val Asn Thr His Gln Glu Gln Met Lys
20 25 30 35

[0110]

cag cca act cca ctg cct ccc cca tgc cag aag gtg cct gtc gag ctc Gln Pro Thr Pro Leu Pro Pro Pro Cys Gln Lys Val Pro Val Glu Leu 40 45 50	199
cca gtg gag gtc cca tca aag caa gag gaa aag cac atg act gct gta Pro Val Glu Val Pro Ser Lys Gln Glu Glu Lys His Met Thr Ala Val 55 60 65	247
aag gga ctg cct gag caa gaa tgt gag caa cag cag aag gag cca cag Lys Gly Leu Pro Glu Gln Glu Cys Glu Gln Gln Gln Lys Glu Pro Gln 70 75 80	295
gag cag gag ctg cag caa cag cac tgg gaa cag cat gag gaa tat cag Glu Gln Glu Leu Gln Gln Gln His Trp Glu Gln His Glu Glu Tyr Gln 85 90 95	343
aaa gca gaa aac cca gag cag cag ctt aag cag gag aaa aca caa agg Lys Ala Glu Asn Pro Glu Gln Gln Leu Lys Gln Glu Lys Thr Gln Arg 100 105 110 115	391
gat cag cag cta aac aaa cag ctg gaa gaa gag aag aag ctc tta gac Asp Gln Gln Leu Asn Lys Gln Leu Glu Glu Glu Lys Lys Leu Leu Asp 120 125 130	439
cag caa ctg gat caa gag cta gtc aag aga gat gag caa ctg gga atg Gln Gln Leu Asp Gln Glu Leu Val Lys Arg Asp Glu Gln Leu Gly Met 135 140 145	487
aag aaa gag caa ctg ttg gag ctc cca gag cag cag gag ggg cac ctg Lys Lys Glu Gln Leu Leu Glu Leu Pro Glu Gln Gln Glu Gly His Leu 150 155 160	535
aag cac cta gag cag cag gag gga cag ctg aag cac ccg gag cag cag Lys His Leu Glu Gln Gln Glu Gly Gln Leu Lys His Pro Glu Gln Gln 165 170 175	583
gag ggg cag ctg gag ctc cca gag cag cag gag ggg cag ctg gag ctc Glu Gly Gln Leu Glu Leu Pro Glu Gln Gln Glu Gly Gln Leu Glu Leu 180 185 190 195	631
cca gag cag cag gag ggg cag ctg gag ctc cca gag cag cag gag ggg Pro Glu Gln Gln Glu Gly Gln Leu Glu Leu Pro Glu Gln Gln Glu Gly 200 205 210	679

[0111]

cag ctg gag ctc cca gag cag cag gag ggg cag ctg gag ctc cca gag Gln Leu Glu Leu Pro Glu Gln Gln Glu Gly Gln Leu Glu Leu Pro Glu 215 220 225	727
cag cag gag ggg cag ctg gag ctc cca cag cag cag gag ggg cag ctg Gln Gln Glu Gly Gln Leu Glu Leu Pro Gln Gln Gln Glu Gly Gln Leu 230 235 240	775
gag ctc tct gag cag cag gag ggg cag ctg gag ctc tct gag cag cag Glu Leu Ser Glu Gln Gln Glu Gly Gln Leu Glu Leu Ser Glu Gln Gln 245 250 255	823
gag gga cag ctg aag cac ctg gag cac cag gag ggg cag ctg gag gtc Glu Gly Gln Leu Lys His Leu Glu His Gln Glu Gly Gln Leu Glu Val 260 265 270 275	871
cca gag gag cag atg ggg cag ctg aag tac ctg gaa cag cag gag ggg Pro Glu Glu Gln Met Gly Gln Leu Lys Tyr Leu Glu Gln Gln Glu Gly 280 285 290	919
cag ctg aag cac ctg gat cag cag gag aag cag cca gag ctc cca gag Gln Leu Lys His Leu Asp Gln Gln Glu Lys Gln Pro Glu Leu Pro Glu 295 300 305	967
cag cag atg ggg cag ctg aag cac ctg gag cag cag gag ggg cag cct Gln Gln Met Gly Gln Leu Lys His Leu Glu Gln Gln Glu Gly Gln Pro 310 315 320	1015
aag cat ctg gag cag cag gag ggg caa ctg gag cag ctg gag gag cag Lys His Leu Glu Gln Gln Glu Gly Gln Leu Glu Gln Leu Glu Glu Gln 325 330 335	1063
gag ggg cag ctg aag cac ctg gag cag cag gag ggg cag ctg gag cac Glu Gly Gln Leu Lys His Leu Glu Gln Gln Glu Gly Gln Leu Glu His 340 345 350 355	1111
ctg gag cac cag gaa ggg cag ctg ggg ctc cca gag cag cag gtg ctg Leu Glu His Gln Glu Gly Gln Leu Gly Leu Pro Glu Gln Gln Val Leu 360 365 370	1159
cag ctg aag cag cta gag aag cag cag ggg cag cca aag cac ctg gag Gln Leu Lys Gln Leu Glu Lys Gln Gln Gly Gln Pro Lys His Leu Glu 375 380 385	1207

[0112]

gag gag gag ggg cag ctg aag cac ctg gtg cag cag gag ggg cag ctg Glu Glu Glu Gly Gln Leu Lys His Leu Val Gln Gln Glu Gly Gln Leu 390 395 400	1255
aag cat ctg gtg cag cag gag ggg cag ctg gag cag cag gag agg cag Lys His Leu Val Gln Gln Glu Gly Gln Leu Glu Gln Gln Glu Arg Gln 405 410 415	1303
gtg gag cac ctg gag cag cag gtg ggg cag ctg aag cac cta gag gag Val Glu His Leu Glu Gln Gln Val Gly Gln Leu Lys His Leu Glu Glu 420 425 430 435	1351
cag gag gga caa ctg aag cat ctg gag cag cag cag ggg cag ttg gag Gln Glu Gly Gln Leu Lys His Leu Glu Gln Gln Gln Gly Gln Leu Glu 440 445 450	1399
gtc cca gag cag cag gtg ggg cag cca aag aac ctg gag cag gag gag Val Pro Glu Gln Gln Val Gly Gln Pro Lys Asn Leu Glu Gln Glu Glu 455 460 465	1447
aag caa ctg gag ctc cca gag cag caa gag ggc cag gtg aag cac ctg Lys Gln Leu Glu Leu Pro Glu Gln Gln Glu Gly Gln Val Lys His Leu 470 475 480	1495
gag aag cag gag gca cag ctg gag ctc cca gag cag cag gta gga cag Glu Lys Gln Glu Ala Gln Leu Glu Leu Pro Glu Gln Gln Val Gly Gln 485 490 495	1543
cca aag cac ctg gaa cag cag gaa aag cac cta gag cac cca gag cag Pro Lys His Leu Glu Gln Gln Glu Lys His Leu Glu His Pro Glu Gln 500 505 510 515	1591
cag gac gga caa cta aaa cat ctg gag cag cag gag ggg cag ctg aag Gln Asp Gly Gln Leu Lys His Leu Glu Gln Gln Glu Gly Gln Leu Lys 520 525 530	1639
gac ctg gag cag cag aag ggg cag ctg gag cag cct gtg ttt gcc cca Asp Leu Glu Gln Gln Lys Gly Gln Leu Glu Gln Pro Val Phe Ala Pro 535 540 545	1687
gct cca ggc cag gtc caa gac att caa cca gcc ctg ccc aca aag gga Ala Pro Gly Gln Val Gln Asp Ile Gln Pro Ala Leu Pro Thr Lys Gly 550 555 560	1735

[0113]

gaa gta ttg ctt cct gta gag cac cag cag cag aag cag gag gtg cag 1783
 Glu Val Leu Leu Pro Val Glu His Gln Gln Gln Lys Gln Glu Val Gln
 565 570 575

tgg cca ccc aaa cat aaa taa ccacccgcag tgccagagg cctcagatc 1834
 Trp Pro Pro Lys His Lys
 580 585

gtctcataca agggaagaga gagccactgg ctccacttat ttgggtccg ctaggtggcc 1891

cgtctcatct gtgaacttga ctctgtccct ctacatgtct ctttaatggg gtgagggtgg 1954

gggagagagg gaattattgt ccagtgccaa cccaatgac cccaatcca acctcaggtg 2014

agcagagcct ctacttgagg gactattgtt actataggaa tccttacttc cccagtattg 2074

aagctgaatc agtgagtggt tacaatgata cataataaat cctggaagtc ttgggaccc 2134

<210> 51

<211> 585

<212> PRT

<213> 人类

<400> 51

Met Ser Gln Gln His Thr Leu Pro Val Thr Leu Ser Pro Ala Leu Ser
 1 5 10 15

Gln Glu Leu Leu Lys Thr Val Pro Pro Pro Val Asn Thr His Gln Glu
 20 25 30

Gln Met Lys Gln Pro Thr Pro Leu Pro Pro Pro Cys Gln Lys Val Pro
 35 40 45

Val Glu Leu Pro Val Glu Val Pro Ser Lys Gln Glu Glu Lys His Met
 50 55 60

Thr Ala Val Lys Gly Leu Pro Glu Gln Glu Cys Glu Gln Gln Gln Lys
 65 70 75 80

[0114]

Glu Pro Gln Glu Gln Glu Leu Gln Gln Gln His Trp Glu Gln His Glu
85 90 95

Glu Tyr Gln Lys Ala Glu Asn Pro Glu Gln Gln Leu Lys Gln Glu Lys
100 105 110

Thr Gln Arg Asp Gln Gln Leu Asn Lys Gln Leu Glu Glu Glu Lys Lys
115 120 125

Leu Leu Asp Gln Gln Leu Asp Gln Glu Leu Val Lys Arg Asp Glu Gln
130 135 140

Leu Gly Met Lys Lys Glu Gln Leu Leu Glu Leu Pro Glu Gln Gln Glu
145 150 155 160

Gly His Leu Lys His Leu Glu Gln Gln Glu Gly Gln Leu Lys His Pro
165 170 175

Glu Gln Gln Glu Gly Gln Leu Glu Leu Pro Glu Gln Gln Glu Gly Gln
180 185 190

Leu Glu Leu Pro Glu Gln Gln Glu Gly Gln Leu Glu Leu Pro Glu Gln
195 200 205

Gln Glu Gly Gln Leu Glu Leu Pro Glu Gln Gln Glu Gly Gln Leu Glu
210 215 220

Leu Pro Glu Gln Gln Glu Gly Gln Leu Glu Leu Pro Gln Gln Gln Glu
225 230 235 240

Gly Gln Leu Glu Leu Ser Glu Gln Gln Glu Gly Gln Leu Glu Leu Ser
245 250 255

[0115]

Glu Gln Gln Glu Gly Gln Leu Lys His Leu Glu His Gln Glu Gly Gln
260 265 270

Leu Glu Val Pro Glu Glu Gln Met Gly Gln Leu Lys Tyr Leu Glu Gln
275 280 285

Gln Glu Gly Gln Leu Lys His Leu Asp Gln Gln Glu Lys Gln Pro Glu
290 295 300

Leu Pro Glu Gln Gln Met Gly Gln Leu Lys His Leu Glu Gln Gln Glu
305 310 315 320

Gly Gln Pro Lys His Leu Glu Gln Gln Glu Gly Gln Leu Glu Gln Leu
325 330 335

Glu Glu Gln Glu Gly Gln Leu Lys His Leu Glu Gln Gln Glu Gly Gln
340 345 350

Leu Glu His Leu Glu His Gln Glu Gly Gln Leu Gly Leu Pro Glu Gln
355 360 365

Gln Val Leu Gln Leu Lys Gln Leu Glu Lys Gln Gln Gly Gln Pro Lys
370 375 380

His Leu Glu Glu Glu Glu Gly Gln Leu Lys His Leu Val Gln Gln Glu
385 390 395 400

Gly Gln Leu Lys His Leu Val Gln Gln Glu Gly Gln Leu Glu Gln Gln
405 410 415

Glu Arg Gln Val Glu His Leu Glu Gln Gln Val Gly Gln Leu Lys His
420 425 430

[0116]

Leu Glu Glu Gln Glu Gly Gln Leu Lys His Leu Glu Gln Gln Gln Gly
435 440 445

Gln Leu Glu Val Pro Glu Gln Gln Val Gly Gln Pro Lys Asn Leu Glu
450 455 460

Gln Glu Glu Lys Gln Leu Glu Leu Pro Glu Gln Gln Glu Gly Gln Val
465 470 475 480

Lys His Leu Glu Lys Gln Glu Ala Gln Leu Glu Leu Pro Glu Gln Gln
485 490 495

Val Gly Gln Pro Lys His Leu Glu Gln Gln Glu Lys His Leu Glu His
500 505 510

Pro Glu Gln Gln Asp Gly Gln Leu Lys His Leu Glu Gln Gln Glu Gly
515 520 525

Gln Leu Lys Asp Leu Glu Gln Gln Lys Gly Gln Leu Glu Gln Pro Val
530 535 540

Phe Ala Pro Ala Pro Gly Gln Val Gln Asp Ile Gln Pro Ala Leu Pro
545 550 555 560

Thr Lys Gly Glu Val Leu Leu Pro Val Glu His Gln Gln Gln Lys Gln
565 570 575

Glu Val Gln Trp Pro Pro Lys His Lys
580 585

<210> 52

<211> 694

[0117]

<212>	DNA	
<213>	人类	
<220>		
<221>	CDS	
<222>	(74).. (373)	
<223>		
<400>	52	
gccattcca gttggagaac gtagtgagtc ttccagtga gccaggtct ggtttgtcgt	60	
gaggagctcc gcg atg tcc tct caa cag agc gcc gtt tcc gcc aaa ggc	109	
Met Ser Ser Gln Gln Ser Ala Val Ser Ala Lys Gly		
1 5 10		
ttt tcc aag ggg tcg tcc cag ggc ccc gct ccg tgt ccc gcc ccg gcg	157	
Phe Ser Lys Gly Ser Ser Gln Gly Pro Ala Pro Cys Pro Ala Pro Ala		
15 20 25		
ccc acc ccg gcg ccc gcc tcc tcc tcc tcc tgc tgc ggc tcc ggc agg	205	
Pro Thr Pro Ala Pro Ala Ser Ser Ser Ser Cys Cys Gly Ser Gly Arg		
30 35 40		
ggc tgc tgc ggc gac tca ggc tgc tgc ggc tcc agc tcc acc agt tgc	253	
Gly Cys Cys Gly Asp Ser Gly Cys Cys Gly Ser Ser Ser Thr Ser Cys		
45 50 55 60		
tgc tgc ttc cca agg aga cgc cgc cga cag cgg agt agt ggt tgc tgc	301	
Cys Cys Phe Pro Arg Arg Arg Arg Arg Gln Arg Ser Ser Gly Cys Cys		
65 70 75		
tgc tgc ggg ggc ggc agc cag agg tcc cag cgc tcc aac aac cgg agc	349	
Cys Cys Gly Gly Gly Ser Gln Arg Ser Gln Arg Ser Asn Asn Arg Ser		
80 85 90		
tca gga tgc tgc tcc ggc tgc tga gaggcccgca acccccagcg ctgcgctaga	403	
Ser Gly Cys Cys Ser Gly Cys		
95		
gaaaccgccc cagcccagag cgggcccgccc ccgctgcggc tcccacgcgg ggctgggcct	463	
cggagtttgc cccgtaaagc gaattgcact ttgatgttca gaaaccact ttgttctcag	523	
ccacgcaaaa ctccctgacc ccgatgtgat ttttctcccc ggggattcga gagccatgcg	583	

[0118]

tgggacactg gaccctactg tctacacggg cttgcacaca gcaggtgctc agcaaattgc 643

tattgatttg attgtctttt gaagatgtca taataaagct tctacctct g 694

<210> 53

<211> 99

<212> PRT

<213> 人类

<400> 53

Met Ser Ser Gln Gln Ser Ala Val Ser Ala Lys Gly Phe Ser Lys Gly
1 5 10 15

Ser Ser Gln Gly Pro Ala Pro Cys Pro Ala Pro Ala Pro Thr Pro Ala
20 25 30

Pro Ala Ser Ser Ser Ser Cys Cys Gly Ser Gly Arg Gly Cys Cys Gly
35 40 45

Asp Ser Gly Cys Cys Gly Ser Ser Ser Thr Ser Cys Cys Cys Phe Pro
50 55 60

Arg Arg Arg Arg Arg Gln Arg Ser Ser Gly Cys Cys Cys Cys Gly Gly
65 70 75 80

Gly Ser Gln Arg Ser Gln Arg Ser Asn Asn Arg Ser Ser Gly Cys Cys
85 90 95

Ser Gly Cys

<210> 54

<211> 5000

<212> DNA

<213> 人类

[0119]

<400> 54	
tcttccttct ttttagttctg aatccctttc ttgtggtttt cctaagtgcc aggettttct	60
gtagtccta ccagactttt caaatgcttc ctcaggagaa ctgaagactg gttactttt	120
ttttcttttt tttaagatag cagtggggag gggaagtctt ggggcagaag ttggataacc	180
agggcctgag aaataaagat aagagggtat attcctcccc caggagacaa aagagaaggt	240
ggatggagaa ggggaaactg ggcctcagcc agcaccggga tgatgcgacc agctttgtca	300
gcacgcttgg ggagcataca caccgcttgc cttggctgga cgggagactg tgatcaccac	360
ccttccaacc catccccaa cagcagget ggcatggcca cgcacacagt gcccgaagt	420
ctgcctgcct tgcagtgact cacttctctt gtatgggctg ctgaactgtc ccaaggagag	480
gacaggccat ctccaggcct tcagcccca gtttagatat ctgggcaggc tttgtggatg	540
ggcaagcata caggagccc aaatgtggaa ggtaggtttt ttcaggtcag gtggaactgt	600
taacactgaa gcctatttgg gtaagaagca ggcctagct agtacctaag ggagattcag	660
gggccaggag tatcggggag gaaacatttc ctaacagaga ggtaaaatag ccaagaggag	720
tttctatgta ttctgaaatc aggcaatgga aagtatttgc agaaaagcaa ttagggacc	780
aggcgtggtg gcttatgcct gtaatccaa cactttggga ggccgaggtg ggtggatcac	840
ctgcggtcag gagtcaaga ccagcctgcc caacatggtg aaacctcatc tctactaaa	900
ataaaaagt tagctaggca tgggtgcaca tgcctgtagt tccagctact ttggaggctg	960
aggcaggaga attgcttgaa cctaggaggg tgaggttgca gtgagccgag atcatgccac	1020
tgcctccag cctggcaaca gagtgagact ccatctcaaa aataaataaa aggetgggtg	1080
cagtgattca tgcctataat tccagcactt tgggaggccg aggtgggcgg ataacctgag	1140
gtcaagagtt tgggaccagc ctggccaaca tggtgaaacc ctgtctctac taaaaataca	1200
aaaattagct ggggtgtggtg gtgggccctg taatccagc tactcgggag gctgaggaag	1260

[0120]

gagaatcact tgaacctggg aggcagaggt tacagtgage tgagatcaca ccactgcact	1320
ccagcctggg tgacagagta agactcggtc tcaataaata aataaataaa agcaatgttg	1380
gacttggagg ctgaagaggt ggacatgcat gaacagtgga ggcgggtgtat gtggctgcct	1440
ttcccctgta ttccactgag agggagccca aatctgcctc gcctgcattg agggatgaag	1500
aaaacactga gaggcagctg agaaacatgg atgctagtgc caactccatg gcctcgggaa	1560
gtcactcaga ctctttagt cagtaatgaa aaggtagtag attagaggct gggcgcagtg	1620
gtcatgcct gtaatccca cactttggga ggccgaggtg ggcggatcac aaggtcagga	1680
gttcaagacc agcctggcca acatggtgaa acactgtctc tactaaaaat acaaaaatta	1740
gttgggcatg gtggtgcatg cctgtaatcc cagctactca ggaggttgag gcaggagaat	1800
tgcttgaacc agggagtcgg aggttgcaat gagecgagat cgcgccattg cactctaacc	1860
tgggcaacag agtgagactc tgtctcaaaa aaaagaaaag aaaaggtggt agattagatc	1920
ctctctaaga tttctccaa ctctgcaatt ctaaaatgta caacatgggg aaggaaggtg	1980
atcagtacaa ctgtaacaa ctgcaatccg gccctgaagt tgaacaaaac aaaacctagg	2040
gagaggaatg agttttagc aggtttcaat gcagegattg tgaacgcac acacacacac	2100
tcctcacagg caaggacttc gttctgcca tgtactgag actctgttg acttctccaa	2160
attcagagtt cagcctggca tgccggccag gagecccttg gagagattcg gcgtgaggaa	2220
acaggaggcc tcacgttcc ttcggttctc cctaaatctg ctgggcaggt agccagaaga	2280
aactctgaag ggtacagaaa gccgagggcc agggggaccc cggtgcgga gctgcagtac	2340
cgtgtttacc cagcctaact gcggtcacc agtctcctgc atgtgtctga ggtcagctc	2400
gggggctccc ctgagtaggt atctacctga ggctcttccc ctgatgctac ctgcatacca	2460
attccacca ggaacctggc gtccctcagga ggctgcactt cccaagtggg ctgagacagc	2520
cccttaatat gtggggctgg agccagagac ctcaaggggc cactaggggg ctcgagactc	2580

[0121]

tcagtggctg ctgacgccct cttgtggcca ttttggggat tggccagaaa agtttcagaa	2640
gcgagccaag atcggaattc ttaaggaggt aagtaataac aacagtaaca gttaccatat	2700
aacaatactt atcgttatat atttgataac attgtgacat ttctacttaa cagcgctgtg	2760
cccagaactc ttctaagtgc tttatatgca ctgtagttat ttaattctca caacaaccta	2820
taacataaac actattactt gtctgttgat gatgaggaaa ctgagcacag aatagttgcc	2880
caaggteaca cagaagacca gaatcccgac tctgagaagt ctggetccag aagctttget	2940
tttagccacc acgctgagac tgagttccat tttggggttg tcaggtgaaa tacaggacac	3000
actgttaaac tataatttca gataaacaaa aaataataat tacaggcggt ggetcacgcc	3060
tgtaatecca gcactttggg agtccaaggc aggtggatca cgaggtcagg agttcgagac	3120
cagcctggcc aacatggtga aaccctatct ctactaagaa tacaaaaaat tagccaggca	3180
tgatggcagg cgcctataat ccagctact cgggaggcgg aggcaggaga atcgcttgaa	3240
cccgggaggt ggagcttgca gtgagccaac ataatgccac tgcactccag cctgggcaac	3300
agagagagac tccgcctcaa aaaaaaaaaat gttagtatat ttaatgaaat atttgggtca	3360
tacttgtacc aacaaatgat tectcattta cgtgaaattc tgattaagct gcatgtcctg	3420
tatattttatt tatttaatct aataccccta ctttacaggg cacttttctc tgccacatgc	3480
aattacacag attacttttt tttttttttt ttttttttga gacggagtct cactctgtcg	3540
cccaggetgt ggagtgcagt ggcgccatct cagctcactg caacctctgc cteccaggtt	3600
caagtgattc tectgcctca gcctcccgag tagctgggat tacaggcatg cgcgcccatg	3660
cccggctaatt tttttgtatt tttagtagag acgggggttc actgtgtaag ccaaggetgg	3720
tctcgatctc ctgcctcatg atccaccgc ctcggcctca caaagtctg ggattacagg	3780
catgagccac cgtgcctggc cacattacct tatttttta tttattttatt tattttttt	3840
tatgagttca atagtttttg gggacaaggt ggtgtttggt tacatggata agttctttag	3900

[0122]

tggacacatt acttcattta atcatcccag gaaacctgtc ctcgagggcc ccagcatgtc	3960
ccacagcctc cgtctctctc aatcctccca tttagagagc tgaaaggact agcgaagctt	4020
acagaacata ggtggcagag ctggatatgt aaccctcatt ctccagttct gacacaggtg	4080
catgttttcc ttttaacagg agagaaaagt gaggagcagt gattaccgcc acagaaatat	4140
cctttcatgg cgtgtcctcc ggcatggtct actcagttgt gtggagccac tgcaggtccc	4200
tcctccaggg atgtaggaat gccagtgaag gccttctga gttgggagag aggcagagga	4260
gtgagtggag gcagggtgt agcttctgtg ccccgccct acccagctc tccaggagg	4320
caggaaagag gctgtgttta gacctgaggg agccagctgt gaggtggag cagttgtgc	4380
atggcggggc gggggctcca cagggtgtt cacctgtgc tctgtcaga gacagctca	4440
agtccagctg ctggggttgc atcacctgea gctaaaacag ccacagggcc caggctgcca	4500
aaaggcaaca gaagcaatct aagaaaccag gcaagggage ctccccatc tccctgttgc	4560
cttttgggag catagagctt cctctgaag gtaacacct ttgcattttt ctctcttgac	4620
tctgttcac aacttccact cccaaatact tgtgccctc cttctccaca gatatacaca	4680
atccctgggc actctccacc taggtcagct ggctctagta aggcagcagt caaggagtt	4740
tctgggctgg gacccctct ggggtgcagg ctgggggtgc ctgctggtat gtgcccac	4800
aggtttttcc ctatcagct tgaaatcagc ctgcagctg tctgacctgc ccccaggga	4860
gtggagcca gaggcagagc caggtcagaa gctggggcag agccaggtca gaagctggac	4920
tttgagacta ggtctctgtg gtcccacctg gaccacagg gaaggcagtg cataaaagcc	4980
tcctgtgttt gaggtgagc	5000

<210> 55

<211> 5000

<212> DNA

<213> 人类

[0123]

<400> 55	
tttagaaagc atagttataa attatctgga aaattcatcc caccaaaca aataattgac	60
aatctggact agacagatgg tctgacattg tgtgtgtgag tgtgtgtgtg tgtgtgtgtg	120
tgtgtgtgta tggtagtgta tctatgctgt gtgtaaactt gttgtcagaa acctgttgac	180
attatactaa aataattgaa atttttttgg ttaaattttc tgaactttat tcccagctgt	240
gtgaacttgg gctatttggt tagctctagg agcctttgtt ttctcattag tacagtggga	300
atattaatac cacctacatc acatgattag tatgaggatt aaatgagata atgtgtataa	360
aacacttaaa taatacctag aaaaaaatag atggtcagta aacatcagtt attatattac	420
tttctgttat ttttatcatt gtacacacat gtatgtatgc atgtatacac acacacacac	480
acacacacac acacacacac acacggaggc atttgacceca aaagcacaac tccactcaac	540
ctatacctca gggaaagatg gggaagtagg tagacctgaa ttggctaggg ttttcagag	600
aaacagaaga gagagataaa tttattttaa gaaattggct catgagatga ctgtgtgggc	660
tggcaagtcc aaaatcttgg ctggaaattc agataacagt tgatgttgca gcattgagtc	720
taaattctgc aacgcagtag tctggaaaca ggcagagttt ctacattgca gtcttgagga	780
taatttatte tttgggaaac ctcagtgtcc attattaaag actgcagctg attggatgag	840
gctcatgtta gggaagataa tctgcttact cagtctactg attaaatact tacagagaaa	900
catgtggatt gctatttgat caaacaactg ggtaacacag catagcctgg gcaaaattaa	960
ccatcacaga gactgagagg atacatggag gtcactggag atatgacceca agagacagaa	1020
gagggtaac agtgacgagg acacactaag aacatgccat gatggtttta ttttactta	1080
ggataaaggc aaatttctgg acagaccttt ggagaatatt gggaaaacag aagcttctac	1140
ctggtaacce catggtaagc ccaacaaagc tgtgatgtta ggaggactgg agaaaggag	1200
tgcagcatgc ggctgaggag gcccttggcc ttcgcacttt ggtttcacag aagtgtagtc	1260
agtacagcac ggacaaggct gctgaccact gcacagtagc tggagctctg gggactggtg	1320

[0124]

attgctgcag cttcttccta tgcaggggtc acacaagaag attcggaggt gaacatgggc	1380
tgtgcactgg agtggcactg tggagggccc tttgaaggca ttggcattgc tgggtggttt	1440
ttggaaaggt atctgagttt cagaatgaat tgaataataa gtcagataaa ttagagccag	1500
gacgggcata attttgcct tttaaaccta ttcttgggtg gctggcagac agtttataaa	1560
atacatttgt atcgaattag tacctatgaa gaaacgtata accccaagtt aaaagtgtta	1620
gaactgacat gcttaaaaac agaattaatg ctaaataaat gtcaattctg tctctcattt	1680
catttcatat ctacgtacce acttctacc tacagtattt catgagcttt aagaggtat	1740
caattgtaag ggacaccctt atttcatgca ccacaaggaa agaaaaaat gctgcaaatt	1800
caacaattac gtgccaccaa tcatatggta tatactgatt taagggtgt taacaagtat	1860
tagaagactg tatcatggaa ttgatgaatt aagataagta taaaaccgat ccactttagg	1920
tcatataaat atgtatgtaa atactagacg ggccaattct tgctttgccc agaacacccc	1980
tctgctgtgc tctccctat cactatgtct agcactagac actgctagac agtattgaca	2040
aaaaagatca aaaggcctac aaagatatgg tacaggaaaa ttttggattg gtcaatctat	2100
taaataaagg gagacaaata aaaagacatg gtaaggcttt ttcttgtagt gtagataaaa	2160
taaagttccc acatttggtg gggagtattt ttctggaata tatggagttt aagttcccag	2220
ggagacaaat caaggagcta cttggagaga ccagcctgca ggagagtgc ttgccagatg	2280
acatgaactt cagatgaaat gtttgaagtc accgcaggg cattcttctc atgacaccta	2340
catgctegtc catgtggtec taaaatgatg cactgacct gccaaacactc atgtggcaat	2400
taagtatgtg ggtggtcctt gagggagcat gggcaatgtg acccactggg tctctggct	2460
cagagagatc acacatgatg atgtgaagac tgtggagaca gaagcaagaa tatctatagt	2520
caggtgtttc ttcactgact tgtgaaccaa ttcccaccat tttttgagat ataaagggaa	2580
aaagtgggca tgttaaagag ctgtttatct ccagttctg gtttccaga gggtcggttc	2640

[0125]

ccattagtcc ttccattgct tattcaacca ctcaatacag atgctcatgt atgtgcatgt	2700
tcatttcacc tgcactttta tatgaatata gcatacatgc acacagtga tacatgcaca	2760
cagtgcacac atgcacacac acacacagge ttgcctcccc tatttctca cttacaacat	2820
tgtacaaaat atctttgtca aagagtatga ggaagctgag ggtgatatca ccatgaagac	2880
agtggaatga actaacatct tttaaacaca ttgggaaata gtgactcagg aggttcagge	2940
ctcaaacatc atcttctttg gtatccagtt tgtctgcagt catcttcaaa aaggccacta	3000
aattctacag tctttgaggt catgggctac ttattgttca cctctatata cctatgtttg	3060
gacatacatg ctccaagaat gttgtagget ttcaaagaat gccttttgaa ttcactgtc	3120
taatcacact gtggacttag tttatccttg aggtaaaage tctctgagaa caggactcgg	3180
gtctgaaaat aggactgtgt aataggtga atagaatccc tcaggtagaa gcagacaatg	3240
ggaaactatg ctttctacct aaccaccctc caccaccaag taaacattag gcaatgtcaa	3300
ggcacggctc gtttctgttt taatcagagg tgtgaataac tatcttttct aattctaatt	3360
tcaacatctt cattgagcca taattcactt attatacagc tcatttatct aaagtgttaa	3420
ttcagtgggt ttcaatatat tcatagagtt ggacaatcac caccaaattt tagaacattt	3480
tcattactcc acaaagaaaa cctatatcca ttgtttgtaa ttttctaatt tctcccagcc	3540
tgctaaagta cctggaagcc actgatattt cctcagagga tgtgcacaaa agtaaaagge	3600
gatacaggga acagaatctt gtccaggtc tggttcctca gatgccctaa gtcagctctt	3660
caggcgtcca gcctgagaga ggetgtaact ggctaaagcc tcattccctt tctagttcta	3720
cagatttgcc tactctacac atttcatata aataggatca taagaaatgc agtcttttgt	3780
ggttggettc ttccatttag cgtaaggttt tcaaagttca tcttccatat atctgtttaa	3840
aataagtga taaatgaatg ggtaactctg tctagtcaa cctgaatca gaactctgag	3900
taagtagacc cctttctgac cctctacttc atgataggaa aggaattccc tctccacaga	3960

[0126]

agccacaata aatgatgict acttccctaag ggagttttca aataaaactgt cttgtaaggg	4020
ttctatatgg acaatgaatg ccacataaga cctgggtaag tatccaatac atggcattag	4080
ttttttcttt aaacacatat ttattaaggg ttatatcaagc acatagtgtc acactaagct	4140
ctttcttact tctcatTTaa tattaaaaac agctcagtga gataagtata cctggtttat	4200
aacggccatt acatatgaga attgatactc attgtgagta tcatagtgc actctagcaa	4260
ttgttgataa ttgttgtagc aatttccaga gatcactcag ttatcatatg caagttttct	4320
gattccagta gcacattcaa tacaattcct ccgactaagg gcaatgtttt gatttttagag	4380
ggaattgttt atttcaatc atattctttg ctctagacca cagaatcaac tcagatttaa	4440
tccactgaat ctgctggcca tggtagttg gaggaagaagc agaaacacag aggccaggag	4500
gaagggtga gagcctaggc aaggaggga tgcctagaaa aagccaaggc tgtagaagcc	4560
ccctgggatg gctggatcca cagcaggaca tggaagagaa actgctcagg gtggaggag	4620
tcataatgga acatttccac aaggtcaggg aaagctgtgg ctcagaactg atgtttctc	4680
agaggacacg aatacaaaag atgcaggga caggatcttg tccaggcatg ggtcccttg	4740
gtaacctaa g ttaactcatt agggttccca cctggtggag gctctgactg gctaaagctt	4800
ctctagaccg cagccccaga gaggagtctt ggaagctctt gaggcctctt tggattgcag	4860
ggactaggag gcaggagaat gtcaggaatg acccagcacg tgcttggttg tgcaaaagcc	4920
acctctctct ctttcagag ccaagcctgg gccctataaa agacacatcc agcttcagca	4980
cctcatctgc tctgactccc	5000

- <210> 56
 <211> 5000
 <212> DNA
 <213> 人类

 <400> 56

[0127]

caccagatt cataaagcaa gtccttagtg acctacaaag agacttagac tcccataaaa	60
taataatggg agactttaac accccactgt caacattaga cagatcaacg agacagaaag	120
ttaacaatga tatccaggaa tttaaactcag ctctgcacca agtggaccta atagacatct	180
acagaacttt ccaccccaaa tcaacagaat atacattctt ctacgacca caccacactt	240
attctaaaat tgaccacata gttggaagta aagcactcct cagcaaagtgt aaaagaatat	300
aaactataac aaactgtctc tcagaccaca gtgcaatcaa actagaactc aggattaaga	360
aacacactca aaaccactca actacatgga aactgaaaaa cctgctcctg actactgggt	420
acatatgaaa tgaaggcaga aataaagatg ttcttgaaac caacgagacc acaacatact	480
agaatctctg ggacacattc aaagcagtgt gtagaggga atttatagca ctaaattgcc	540
aaaagagaaa gcaggaaaga cctaaaattg acaccctaac atcacaatta aaagaactag	600
agaagcaaga gcaaacacat tcaaaagcta gcagaagaca agaaataact aacatcagag	660
cagaactgaa ggaaatagag acacaagaaa cccttcaaaa aatcaatgaa tccaggagct	720
ggttttttga aaagatccac aaaattgata gaccactagc aagactaata aagaagaaaa	780
gagagaagaa tcaaacagac acaataaaag atgataaagg ggatatcatc accaatccca	840
cagcaatata aactaccatc agagaatact ataaacacct ctacgcaaata aaactagaaa	900
atctagaaga aatggataaa ttcttgaca catacacct cccaagacta aaccaggaag	960
aagttgaatc tctgaataga ccaataacag gctctgaaat tgaggcaata attaataget	1020
taccaacca aaaaagtcca ggaccagatg gattcacagc cgaattctac cagaggtaca	1080
aggaggagct ggtaccattc cttctgaaac tattccaatc aatagaaaaa gagagaatcc	1140
tccctaactc attttatgag gccagcatca tcttgatacc aaagtctggc agagatacaa	1200
caaaaaaaga gaattttaga ccaatacccc tgatgaacat cgatgcaaaa atcctcaata	1260
aaatactggc aaaccaaac cagcagcaca tcaaaaagct tatccaccat gatcaagtgg	1320

[0128]

gcttcacccc tgggatgcaa ggctgggttca acatacacia atcaataaac ataattcagc	1380
atataaacag aaccaacaac aaaaaccata tgattatctc aatagattca gaaaaggcct	1440
ttgacaaaat tcaacaacac ttcattgctaa aaactctcaa taaattaggt atcgatggga	1500
catatctcaa aataatctaa gagctatcta tgacaaaccc acagccaata taatacctaa	1560
tggacaaaaa ctggaagcat tccctttgaa aactggcaca agacagggat gccctctctc	1620
aacactecta ttcaacatag tgttggaagt tctagccagg gcaatcaggc aggagaagga	1680
aataaagggc attcagttag gaaaagagga agtcaaattg tccctgtttg cagttgacat	1740
gattgtatat ctagaaaacc ccatcatctc agcecaaaat ctcttaage tgataggcaa	1800
gttcagcaaa gtctcaagat acaaaatcaa tgtgcaaaaa tcacaagcat tcttatacac	1860
caataacaga cagccaaatc atgagtgaac tccattcac aattgcttca aaggaataaa	1920
atacctagga atccaactta caagggatgt gaaggacctc ttcaaggaga actacaaatc	1980
actgctcaac gaaataaaag aggatacaaa caaatggaag aacattccat gttcatgggt	2040
aggaagaatc aatattgtga aaatggccat actgccaag gtaatttata gattcaatgc	2100
tatcccacac aagctacatc tgcctttctt cacagaattg ggaaaaacta ctttaaagtt	2160
catatagaac caaaaaagag cccgcattgc caagacaatc ctaagccaaa agaacaagc	2220
tggaggcatc acactatctg acttcaaact ctactacaag gctacagtaa ccaaacagc	2280
atggtactgg taccaaaaca gagatataga ccaatggaac agaacacagc cctcagaaat	2340
aatgctgcat atctacaacc atctgacceca gcattgcctag atattcttga tattcttttc	2400
tcctctttta tctcttttga aaatcatggg agattccagg ataccccacc agtttcttct	2460
gtcaciaaag tataaatggc ttcagcagga atctagccac tgggcaatag tttcagggac	2520
ccagttcact atggagtcaa ctcttggttac gaaacttggc tctcacctca cctccaggtc	2580
tgtcttatga gctggtggct taggactggg attgtctgtt ttcagaggga aggagaggtt	2640

[0129]

gaaatgagct aagagatgtg ttttacaaca agaaactata aaataaccta ctgcttgcag	2700
aagctgaagg agagggatga gtgacattgg taaggtecca tgcctggtga ataagagagg	2760
aaataccaca aagagagagc gtgttttgga gatcaggagg tttgtgata gagttagttt	2820
gaggtcacag caaaaatatt tgctcttag aagacatttg gcaggatcaa actgcaggtc	2880
agtgactttc taggcaaagc agaggtggct taaaggaagt attttaaata actcccaggt	2940
aattcttcta ataacattta atttattcat tcatccatct aatcaatagt ctatggtcaa	3000
ctgetagact ccccaggaaa agaaacaatg ggtaagacct gattctgagc tcatgctctg	3060
gtgacttaat atccgtggac atagagccaa gattgtctag aattcttact ggtgtcaaag	3120
gtgtgaagcc tgatctctaa cagtgactga ccatatttta tcataaaaaa tatggaagga	3180
cattttgtct gctgagactg gcttcctgtg aacgaggcac atatcagctt tggagtcttc	3240
tctccaaccc ttggccacgc tgctgtgctt gtccctggac ttgcaagcac accttcatgc	3300
ccatgctgcc cacacacaca tacacacaca ttcataaagg cacagcccat cccattctct	3360
tccttcatgg ttactatcat ctttgtcaag gagaatgagg aacctagcat gacattccct	3420
caggcattat caggtgatcc agtgtccaa caatctcatt ggaaaataga ctcatttggc	3480
atgcacataa atttcacaga ggtcttcaga atatccactg aattgtgaat tcctagaggt	3540
cacaaaacac acacagtgtt tgctctata aacctaatg tttggataga aaatcactga	3600
atacctagga gctgacaaag cagtatcttc ttgattaact gctatgtttg gatatcaaca	3660
tttttactc ttgacctgaa cactctcaag ggtagggaac tacatctgag actagaaaag	3720
aaaactattg gagatggggc atctcttagt atgacagaaa tgtaaagacc atatactcct	3780
accttctaga ccaagggtat atgaaagaag ggactatctt tttcaaaagc tgtattctaa	3840
cttatgccat acagtaggtt ttcagtaaag gttggatgga tggatgaaat aaattgatta	3900
atggatgaat gtatctaate caaaattgta tcagattcct tactctcctg tcattcacta	3960

[0130]

agaagattaa ataaatcatt tgattcccat caccctcttg acaagaaata attagacttc	4020
tcaaaatcct gaatgtacac tgiacattcc taagagagcc ttcgctagcc tacctatatt	4080
tttaaataaa attttcaaag acagggactg aatattgtcc atttttctgt ctcacacact	4140
gccctcgttc ttttaaaaaa tgacattggt taagtgtctt ccttgtgcct tatcatgtgc	4200
aaaagcctca cttgatggaa ccaacaattc agagaggaag gtgctaatac gcctgtttta	4260
cagttgagga tttagattgt ctatatatag tctttgaggt cactgggctc ataaacacag	4320
agctctcttg attccaattc aaaaaactat tggtctatct taaaaactgc cttgttacac	4380
agtagacatt ttaatttcta agtgaagaag tttagtcaca atgatcatca ccaccaaatc	4440
tgaggaaacc atcttgattc acttattaat cctcctagca gggttctgca ggccacacag	4500
caacatgggg tcaggaatga gggtctagca gaaggggagg agtgtatcct aggaaggaaa	4560
tggtctctga ctcaatgcca caaagaaacc tgcctagctg tggaagcagg gcttgaaaaa	4620
ctatacttgg ggatgggaag aagctggaat agcatacttc cacaggaaag ctatggcaag	4680
gatctctggt gttctcagaa gtcactcttt gagectttgt ccaggcaca gctccttaag	4740
aatccccctg cagctccacg gggaccagcc tggggggggc tgtgactggt taaagctcct	4800
ttagaccccc tcccagagg gaggtcttgg gagctcccag tgcctccttg gattgtaggg	4860
acaaggaggc aggggaatgt gaggaatgac ccagcacgtg cctggttgtg caagagccac	4920
tcccgcctct tcccagagcc aagcctgagc ctataaaaga catattcagc tccagcacct	4980
catctgcctc gaattcccca	5000

<210> 57

<211> 5000

<212> DNA

<213> 人类

<400> 57

atggactgtg tgactgtggg gcttcatgtt tttcatgggc aaaatgagtt agttaaagct	60
---	----

[0131]

gatgaacttt aaagaactta caccttcaaa actatcttta ggtaggtaac acactgtaat	120
ccatcagcaa aactttgcaa tggccccag cgataaaata gccaaacttc tgactacaac	180
ctatgtgcta agcacatata ttccttcacc aaattctcac agaattctaa ggaattgtta	240
gtatcccat tttacagaca aggaaactga aattcataga ggtgaaatgt cttgtccaga	300
taaaaagacg tgcaggtgca attcaaacc aggctgtct gacagcaaag ccagtgccca	360
ccattgaaca gtaacattct ttcagattgg attaggtga aataaactta aacaccagac	420
ccagccagga agtggagctc cctattccca ggaccaaaga gggtcatecc aagggcattt	480
tggetgaatg agagccagag aaaactgaca ctccctccag aaaaccaga actggccatg	540
ttgaggtaag tgactgactc acagagctag aatgtgtctt tgaagtcacc tagtggaacc	600
gtcacagcc atcagatgct caagcaccta gcaaagcagc ctgatgggga ggagagcccc	660
tgactggagg tcaggacaca tgggttccag cccagctcta tcacagctc gctgtaggct	720
cttgggcaac tcatgttctc tccagacttc aattctgtc agaagacaga tgagccccaa	780
tggtgagaag taggcatagg tgtgttaaga agcacctgca ttctaagcta cttgaaggca	840
gggatttgtt gttttctatt cataacccta tctccagtgc cttgcacata gtagaggctc	900
agtaagtgtt ggttgaatga atacatcaat tatattaaca actttacctt gtgacctggt	960
attaggcaaa ggcattactt ctttttttgt tgaggtaaca ttcataaac ataaagttac	1020
cattttaaag tgaacagttc agtgggtgtt agtacattca tatattatgc aaccaccaac	1080
tetatgtcgt tetataacat tttcatcctc cccattagea gtacatecc attctctcct	1140
cctctcactt cctggcaacc atcagtttgc tttctgttct tatgaattta catattgtga	1200
atttaagtta ttattattag ttattattag atagataaat agagaagctc ctatgcacca	1260
ggcgcagtgt aaggcatttc ccatgcattc tctgaggact cctccatctc agacattctg	1320
ctgctggaca gagaatcaac tgtacaaaat gtgatttccc caacaatgag cccaaagggg	1380

[0132]

agggaaatttc ctttagttgg agaaggggga actgagtgtc tctttcctct actactggca	1440
gcctttatgc ctctgtcctt cttgttgcca ccaagagcag agctgtggct gcccttcac	1500
agctctcaga tccaagaaca aaggccagaa ctgggtactg agatccagtg atccaacaag	1560
cccaggagag aaagctataa gaagatgtaa gagaggaaag caggtcagga cagacagatc	1620
caccagaag cccaaaagg atgagtcagc cctcttgcc ctttcttggt ctcactccc	1680
ttccttgaat agcagacggg cacaaaaggt gagagaagag ccaaggtcag gtccaggaag	1740
caatggaaaa agatgaggag gattcaaagt ttgtgggaa ggtcgggcac cgtggctcac	1800
gcctgtaatc tcagcacttt gggaggctga ggccggcaga ccatgaggtt aggatttcga	1860
aaccagcctg gccaacatag tgaaacccca tcactactaa aaatacaaaa attatccagg	1920
catgttggca catgcctgta gtcccagcta ctccaggagc tgaggcagga gaatcgttg	1980
aattggggag gcagaggttg cagtgagcca agattgcacc actgcactca agcctgagag	2040
acagagtgga ctctgtctca aaaaaaaaa aaaaaaaga agtgtgtggg gagacaacta	2100
gtgcaggttg gcaggttgag gactgagagc cgctgatgta ggagacagtg ccctggggag	2160
aggaggaggc ttgtgagagg gagaggagc tagcagctc agagcatggg gcagtgccat	2220
accaggaagc agatgaggca tggtatttca agatgagaat gcatgatttt ctggcccaag	2280
tcaccttcca ggaccatctg ctatgacccc tctggcaatc tgtaagagc tgtctcagc	2340
atccctcccc acaggacatt gaatcaccat aaccaatttc ttctgcccc taccactaag	2400
caacagctca aataggttgc ctcactcagg agagagacaa actagccagc ttctgcctcc	2460
aaatttgaaa tagccagtgc atccttaggc tggggcagag ttggggccaa agcttctcca	2520
tgtgtcatgg gatatgagct catccttatt atgttgggtg ggggttgga agttaccag	2580
actgtcatg tggacctgga gcttatgagg tcattcacat aggcagtga agaacctctc	2640
ccatatacgt gaatgcctgt ctcccaaatg gggcaacctg tgggcagaat aagggacttc	2700

[0133]

tcagccctag aatgttgagg ttcccccaac ccctcccttg catacacaca cacacaaaca	2760
ctccctcagc tgtatccact gccctctttc ccacacccta gctttgcca gcagtcaaag	2820
gctcacacat accatcttct ccttaaggct cttattatgc cgtgagtcag agggcgggag	2880
gcagatctgg cagatactga gcccctgcta acccataaga ccggtgtgac ttccttgatc	2940
tgagtctgct gcccagact gactgtcacg ggctgggaag aggcagattc ccccagatg	3000
aagtcagcag cagagcaca gggcatcagc gccaaagtaa ggatgcttga ttagttcttc	3060
agggcagagt gggctgtgct tcctctgccc cagaaaatgg cacagtcctt gttctatggg	3120
aaaaagaatg tgaggtcctt ggggtgggctc agggacacaga gaggtcatga ggaggggata	3180
gcactgcaga aaccaagggt gccttgtagg tcttcctct gtcttttttag gcatgatcca	3240
ggaacatgac aaaattagtg ctttaaatag atttacttgg ggctaagaga aatgtgcctg	3300
tcaggaaaac tatggggaat caggacactt ctcaaaatta gcccactga gtattgtctt	3360
tataattcct tctttttgga ttagattgta aaaaagagag tgtaaatgaa tgatgtccat	3420
ataataagtt attagccaac cattaaaaag aaagggaaga aataaatcag tttggttttt	3480
acacacacat acagacacac acatataaac attgatcaac actgaaatgt ttaatagtca	3540
ttattttcgg gtcgtaaaat tcactgttct tcaatgaata cttgtagagc acatattata	3600
tgcagtagtt ttgatagggt ctaggggtat agtggaacac ataccaggta tacgtgctc	3660
ttagcttatt ttccagtggg aaagatagac aataagcaag tgaacaaatg caaataaatt	3720
actctagatt gttataagtg aaattaagta ccaatcctt agatatggtc cacagagaag	3780
gatctctgac agacccaac attgacactg aagctgaaag gcataaaaga accagagacc	3840
tggggagggg ccggtgggca gaaggagagc aggtgccaag ccccaggtg gagagctctg	3900
ggctcatctc aggaaccgaa ggccctcagt gaggttaagaa tatacctctc agggagagat	3960
tgacatgaat tggggcccca gaagaaggca gaagccaggt accagggtc ttttaacca	4020

[0134]

cggcagtgag ttggaatggt atttcaagtg tgctgggtgca ctgttggcac gggggagaga	4080
tggtgctcaaa tccccactct gaaagatttc ttaagctatt tctagagtat gatttacaac	4140
aggaaatgga tgatttgatt ctgatcttta taccttcatt catttaaaaa agtacttaag	4200
aaagtagttt gggttgatcat tataaaaagc aatacttatt tttatattgt gtagattcaa	4260
tcttgtttcc ttgcctagag tgggacctgc ttggagttc ttatgagcat ggcattcctg	4320
agaacttctc taactgcagc ctgaggcata gaggtgggc agcaagtggc agcagcagag	4380
gactcctaga agccttctac ttgactctac ttggcctaaa gtcaaacctc ctccaccaa	4440
gacagagttt atttccacat aggatggagt taaaaaatat attctgagag aggaagggt	4500
tggtggccaa gagaacaccc cagaaatacc acccttcat gggaagtgc tctatcttca	4560
aacatataac ccagcctgga catccccgaa agacacataa ctttccattt catgccttg	4620
aaagtgaatc ttttggccta ataatgagaa caaactcatt ttgaaagtgg aaaaattgag	4680
attcagagca gaagtttgac taaggtcaca aaacagtagg atgcctcact cagctccctg	4740
tgcctaggtc agaaaagcat cacaggaata gttgagctac cagaatctc tggccaggca	4800
ggagctgtgt gtccctggga aatggggccc taaagggtt gctgcttaag atgcctgttg	4860
tgagtcagga aggggttaga ggaagttgac caactagagt ggtgaaacct gtccatcacc	4920
ttcaacctgg agggaggcca ggctgcagaa tgatataaag agtgcctga ctcctgtca	4980
gctcagcact ccaccaaagc	5000

<210> 58

<211> 5000

<212> DNA

<213> 人类

<400> 58

ttttcttttt taaacttgca gttctaggaa ctgtgctaag aacttcattt gtattttttt 60

[0135]

attcagtgga ttaggaatgt tttcagcccc cttttccagt gatggttaag taacttgcca	120
gaggttgcac aggtggagag agcaagggtg ggaactgatg ctggatgtcc ctctccaggg	180
tctgtgttat aactgggctg atgctttctca ctgcagccca tttcaaagag catggagcta	240
aggtaactgcc tttttctccc agtcttattt tctctcactg cagctttcat tgtgaatatg	300
ccatattttg gccacccttt ccagctaagt caattttgat aactatcctg gggacceaac	360
tttccctcag aatattgttt gctgagaaaa tgttttggat cccaccatgg gcttttaaaa	420
cctatatact cctcctcttc acctctccat ccaaagatta atcttctect gaggtccttc	480
ccaggttgga tgctcttctc ttatgaatga ccagaaactgg tgcctcactt ctggaaacct	540
ccagatgagt ggagagcttt tgctccaccc tcagtataca cctgggtcttt gctgcctcac	600
ctcgagctgg ccaagggcac cttctttaac acgagtaaat aagccattcc tagacttctg	660
cccttttaaa ggcatgttta gtgctgaatc ggtactgttt gtgagcagtt ttggagcacc	720
aaattcaggc aaggagagaa geccaaccac ccacactgtc agcattagat ctgcctaaaa	780
taacatcact gagatatcaa accttctctg ttctgggttat ggcatcatct agaaatgaaa	840
aaaaaatcat gttagattat ttgcttgggc ctacaatttt attttggatc tctattatta	900
ttctaaaaag ggccatttaa ttccatctt gggaattttc tccttcagtt cctataccaa	960
agagaaaaca gaacttctgt tcattccaga ggtgtagagc cgtattggcc cagctcagaa	1020
aatagaaggc atagcagagc tgcctgtgag atagataggt gccatgtgca attcacagag	1080
ctccctcacc atctagagg cctgcagttc tcaggggac tcctctgtggg ttaatacttg	1140
agatgcttat gcagggagga actgggtccc tgggatgcga ggggaggag tcacctatgc	1200
ctttatgaaa gatgcttcta gcgtcccga catggtctgc acctcctgat gctaggtcag	1260
gactgaattg tataaaagge agaagcttct gacagcacct cagtctacct gtctcctgag	1320
tgatctgctg cagtgcctga accaggtaga gtgcttctca ggaccaggat gaactcttgg	1380

[0136]

tgctggtggtt ttgggcagaa agagccctg ggtggagggtt gaggccattc ttgaagaaag	1440
acaaggataa agaaggattt tggaggggaa ggttctctctg gaaggggagg gtggaaggta	1500
ggtaaaagga atgaaggact gatgccttct tgaagttcca agttattcat ttgaaaatat	1560
ggagttacta ggtggaaatc atttgtatca ttcttggtac tggctctggg agactggggt	1620
tttaaaatct ggttcatctt aggccaggag tagtggtcca tgctgtaat ccagcactt	1680
tgggaggctg aggtggggcg gtcaagaggt caggagtctg agactgacca acatgctgaa	1740
atcccgctct tactaaaaac acaaaaatta gctgggtgtg gtggtgcaca cctgtctgag	1800
gcaggagaat cacttgaatc cgggagggtg aggttgcagt gagccgagat agcaccactg	1860
cactccagcc tgagtgcag agtgagactc catcccaaat aataataata atcataataa	1920
taaattctgg tcattctaga gataaaatct tgtgaaatta tagttcaatt ttaagggcaa	1980
ggaagtgacc ttggccttgt attgtcttct agctttatct agcttctaca gagatgtcct	2040
gccagcagag ccagcagcag tgccagcctc ctcccaaatg taccctaaa tgcctccca	2100
agtggtactcc taagtgtcct cccaagtgtc ccccaaatg ccctccccag tgttcagccc	2160
catgccacc tccagtctct tctgtctgtg gttccagctc tgggggctgc tgcagctctg	2220
agggtggtgg ctgtgcctg agccaccaca ggccccgcca gtccctccga cgccgacctc	2280
agagttccag ctgtgtggc agtggcagt gccagcagtc tgggggctcc agctgtgcc	2340
acagctctgg gggtctggc tgtgccaca gctctggagg ctgtgtctga cctgggcat	2400
gaggagcacg gaggagaagg actggcagat ccaggtgtct gaagatgtgt gtcagctga	2460
ggcttctttt ctctcatctt ccatggaagg acttcggaaa tgccttaagt tcccctttt	2520
atcctgcccc tgttcactcc attgtagggt tgaagtctag cttgtgatat tttctggcct	2580
ggctttccct ctgagacata gctctttgga gaactagggt ttgtaattca gttatgaagc	2640
tattttctct gtaacaataa agctttttat tctgatata actgcctctt gatttcattg	2700

[0137]

tttgectctc ccactcccta cctgctgcat caagccatct tcccttccct tctcttaaat	2760
gcaagcgaac actttacaaa cttgttttagg aagcacattt tggcagtga gctacttagg	2820
cctctggaac acaatatttc cttctatgca ctggattctg gaactttaag cttgaaggaa	2880
agaataatgt ctgagaccat tacgccagac acaagattgt atgaggagag ccaggaatgg	2940
gtttgaagag aagaggatgg ttgtactcaa aaggttcagg ttctgagagg gtttccitca	3000
agcatgggtc acagtttgaa aaattgcttg tccccattca tggggaataa gattgcctga	3060
agtgcctatc ctgggcteta ataaatggac actctttact tttctgatta ttataagctc	3120
agggttttcc atcagttggt tagagggaga gctgggtggg cagtgagtga gttttggagg	3180
cagatatgaa acgtctgagc acattttccc tgcattctga aatcctggat ttgaatctgg	3240
gtaaagaact gagaaagaca cccttctttg caaagactgt ctctatgtgg ttgggacgta	3300
cgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgca gtgagaggta taagagtaca	3360
ggggttttgg tccctgtgtt gggttttcca agcaacattt ctttaggaat gggctggttc	3420
cttcaggaga cactcataat gtacctagaa aagataacta gaaaagttaa ttccaaatca	3480
gcctttaata taaaagataa acaattgatg cctcctttga taagagttaa tagcgataaa	3540
taaagtaage tggtttatga gatcatagga tgcccacaga atgagtaggc acttgcccag	3600
attatctcct tatactgatt cctaccaaac gttcttcatt cttttttttt ttttttaaaa	3660
gaactttgcc cattttccca gaggagatcc tgcctctgtc ttgttatttc attacttctc	3720
tctctgttct aatattgact ttactttttg gaatgtctca cctatactca ccaacaggtg	3780
ctctaggaac tggcaacctg ttaatcaggc tctaaatcca tgttaaatga aactggatc	3840
ttttgtgact tatctttcct cgcctagtc cttgagccac tcccagctca ctgcceccca	3900
aacctgtctc tcactcttaa ttccctgaatt cttagttctg gacttctgtt atagttcaca	3960
ttcacagttg tgggcaattt gattaagaag cagttctgaa gttgactgcc aaagtatcct	4020

[0138]

ccaattctta aaaattagtt attctgaaaa cactatctat catgaacact ttggaccttt	4080
ggaaagaact ttctttcttt ctttctcttt tttttcttct ttcctccct cccttcttcc	4140
tttcccatth tcttttgca c tgggcagtgg ccctagggag aacttatctt caaaggcaga	4200
ttggtggttg ggggtgggat aagaagcagg agtcccacag gctgactagc ctttttgcaa	4260
ttgggccact gtccctgcaca catthttctc ttgtcatac catgctgtat tatggggata	4320
tggcaacctg actacattca caactthtct ggtcttatgg ttccttaact ttgcctaggt	4380
ttttcttctt gtctccagtg acagagagaa atcacccctg gggttcccaa aggtgcatcc	4440
ccgtcatcat cttcactcac tgcatttga agcatatgga tgaaaccag caggcttacc	4500
gcatagett tggctgccc tggggagcca aggaccaaga gccagagaga tgagaatcaa	4560
gataaaaata ggtgaactac cacctacca gtagatgact acataggaat ggatgtgtgg	4620
gcaagtggag ccgtctgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtat	4680
tgaagacaag ggcaggggtt ccagacctcc ctgcatttga gcccattgcc ctctcttagc	4740
atacacacag cctcccaggt acaaacagt cctgttttg aaagtaaaca tgccatagca	4800
atcctggcac ctctgtgaag tgacttagat gccacatgc accacctaca atgetacctg	4860
agtgatgaat gtctctggga tgtgagtgt agctgagcac gtctgttggg aaatgggccc	4920
ctcccaagag cctgaacctg aatacttctt ataaaaagg tctctggctg gacgtgctc	4980
aatccactgc ctagcaggtg	5000

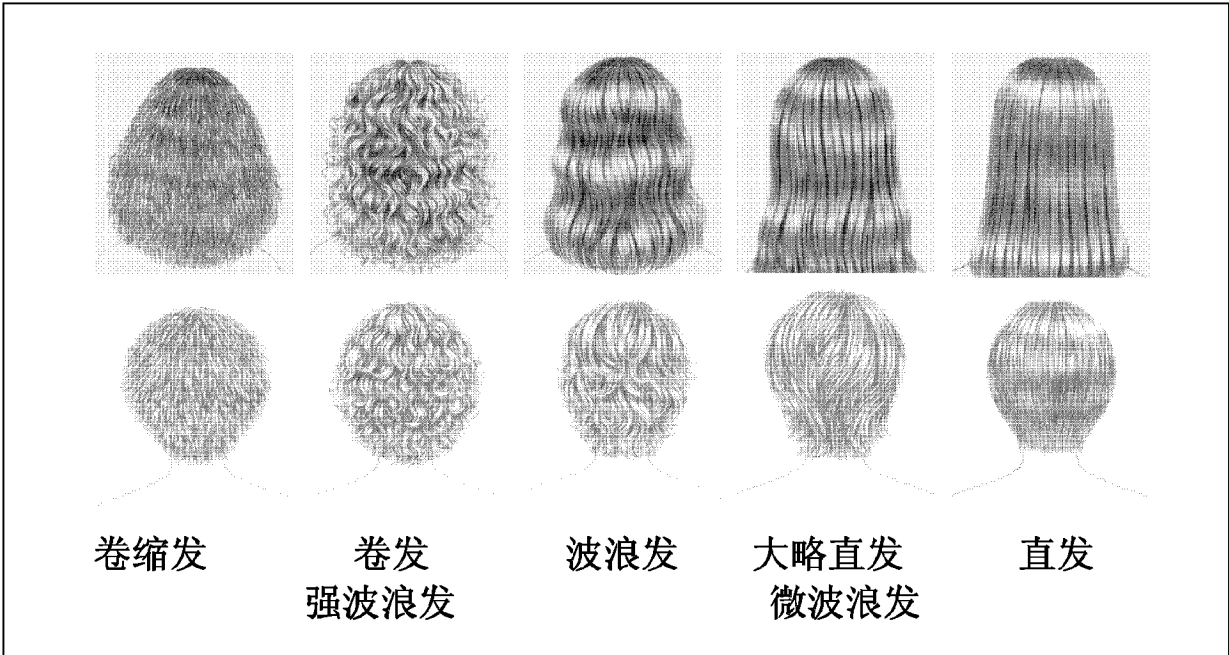


图 1

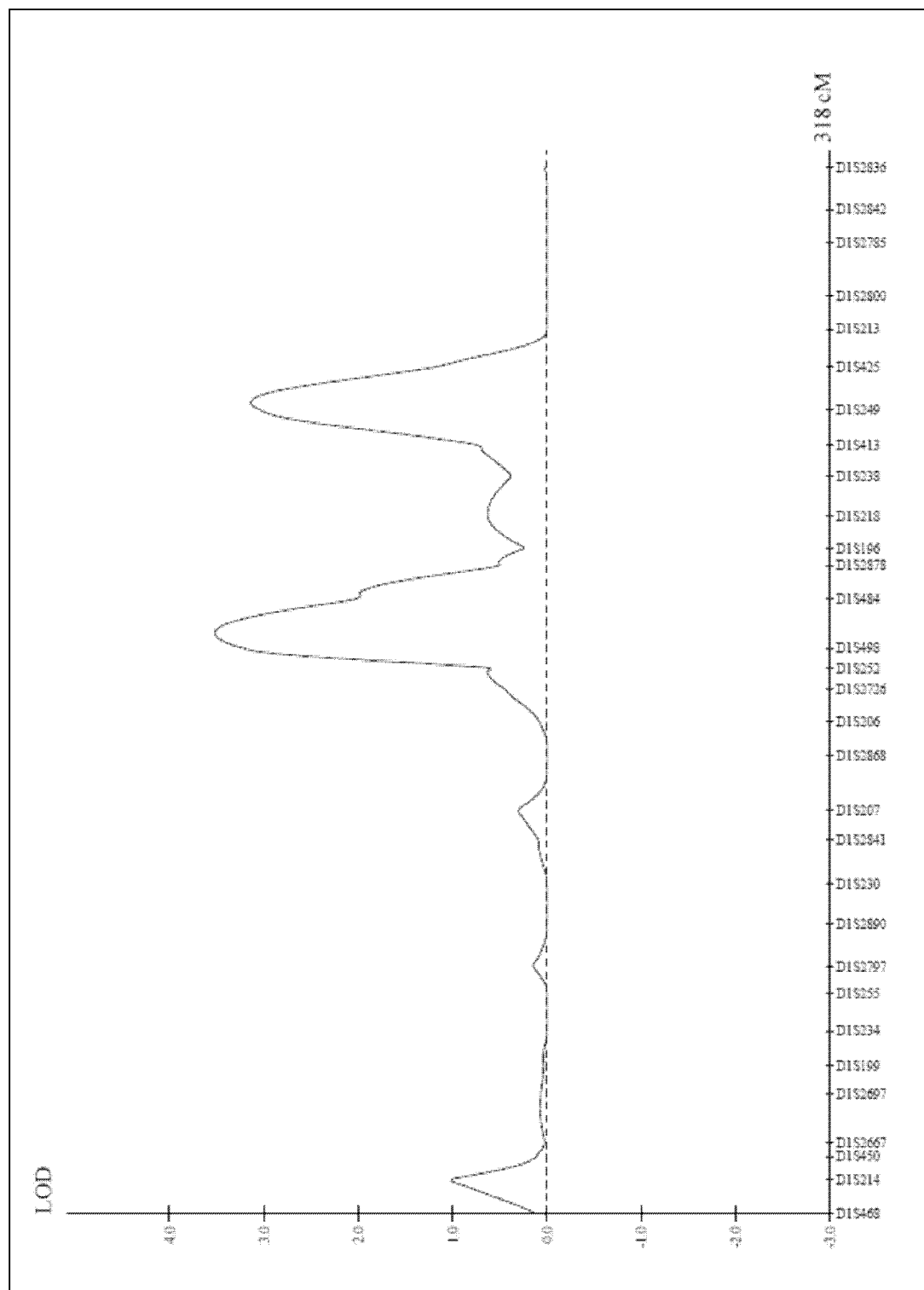


图 2

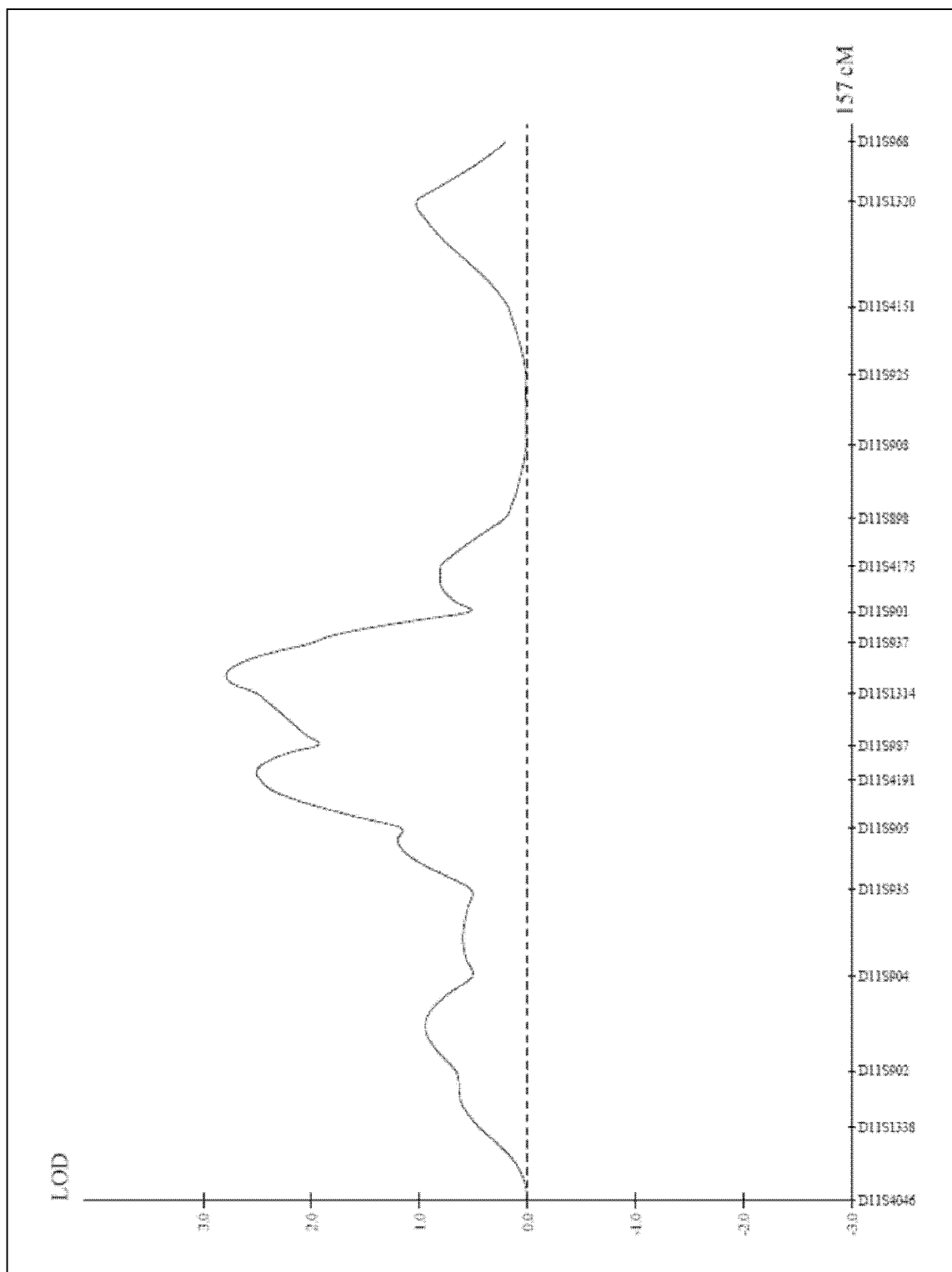


图 3

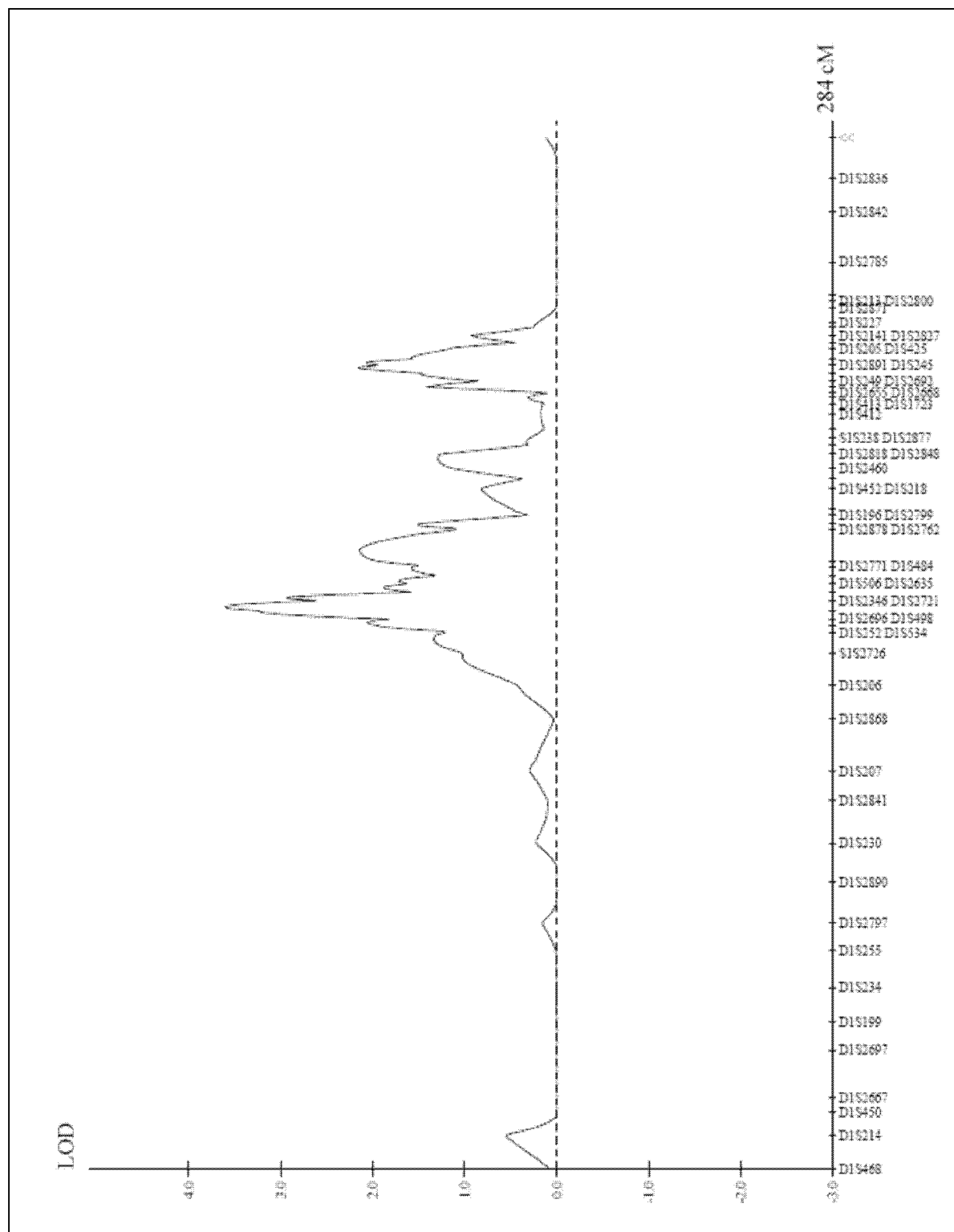


图 4

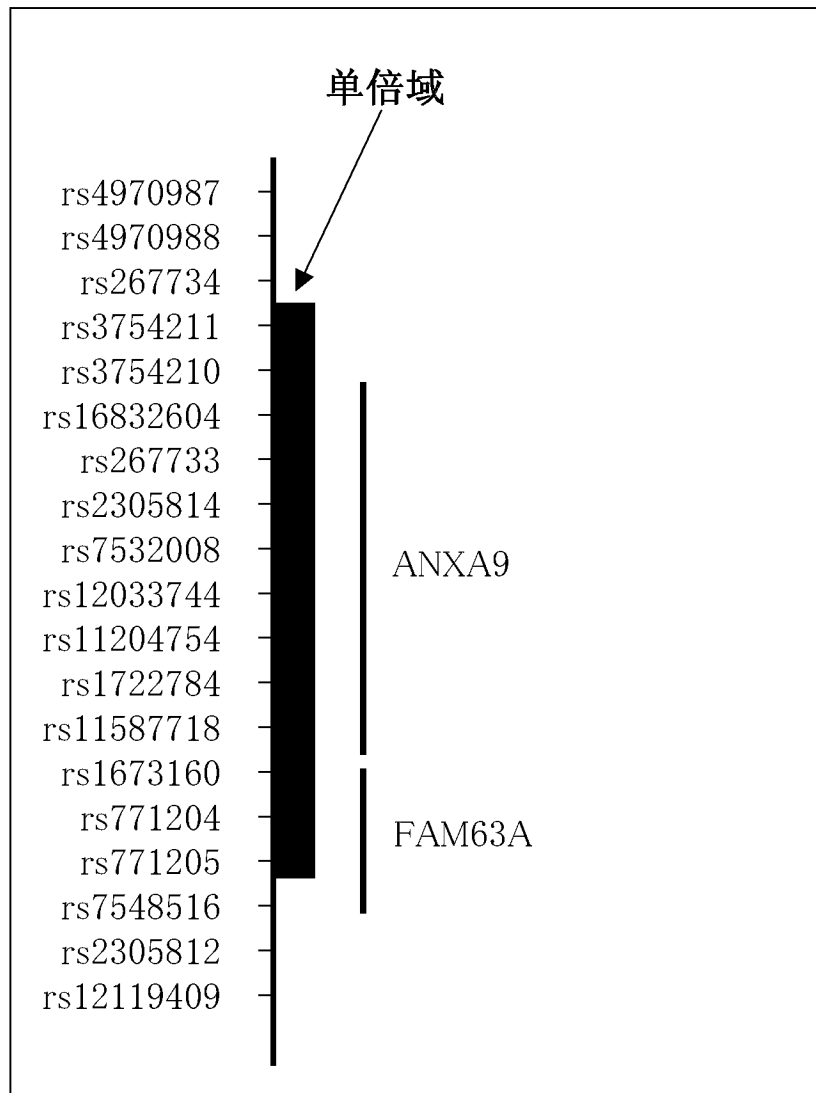


图 5

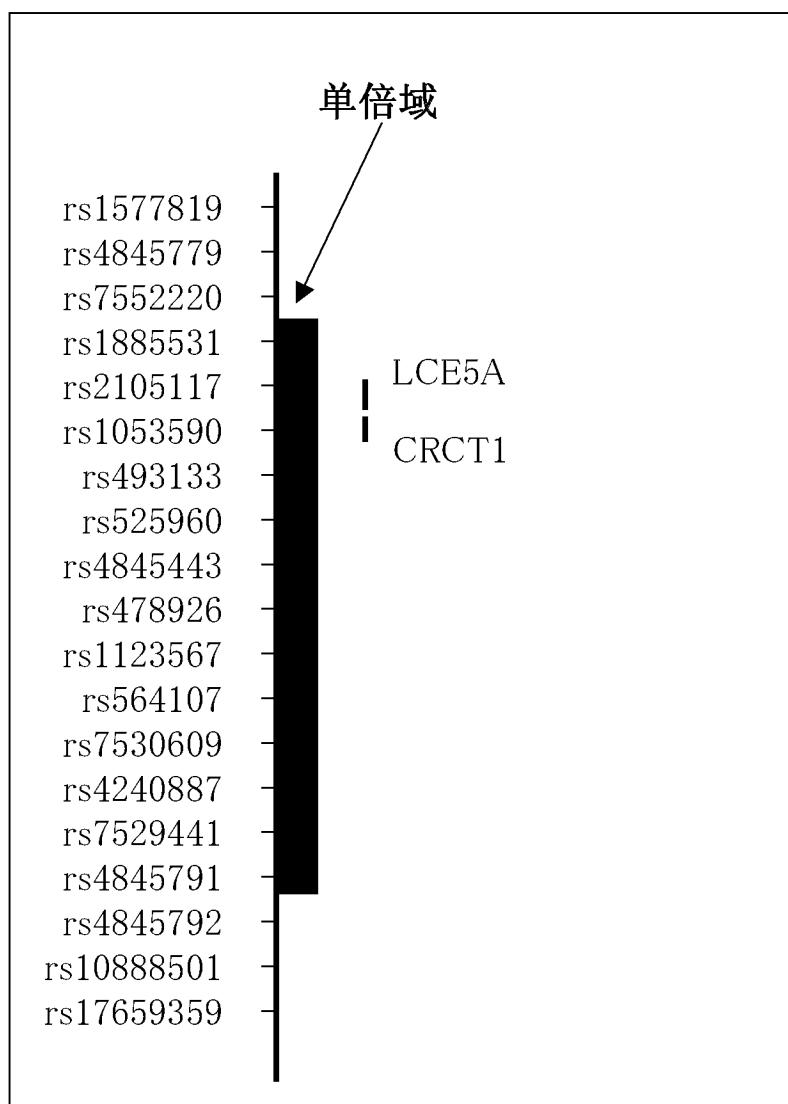


图 6

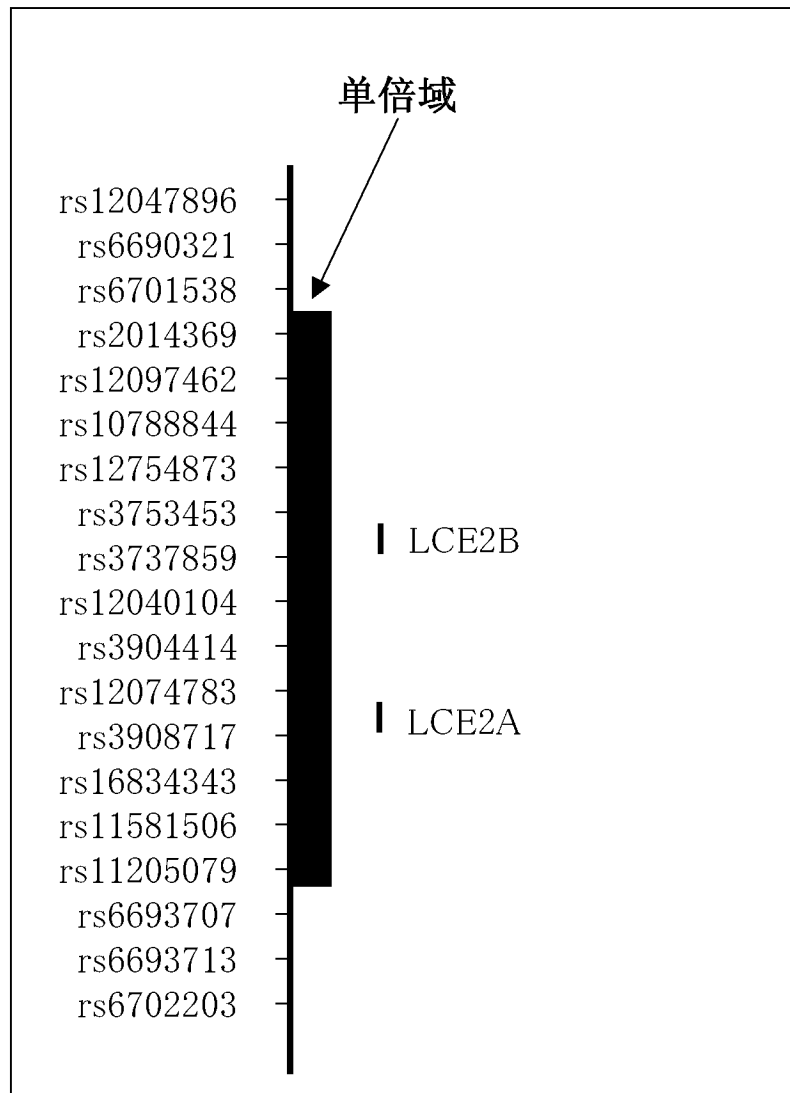


图 7

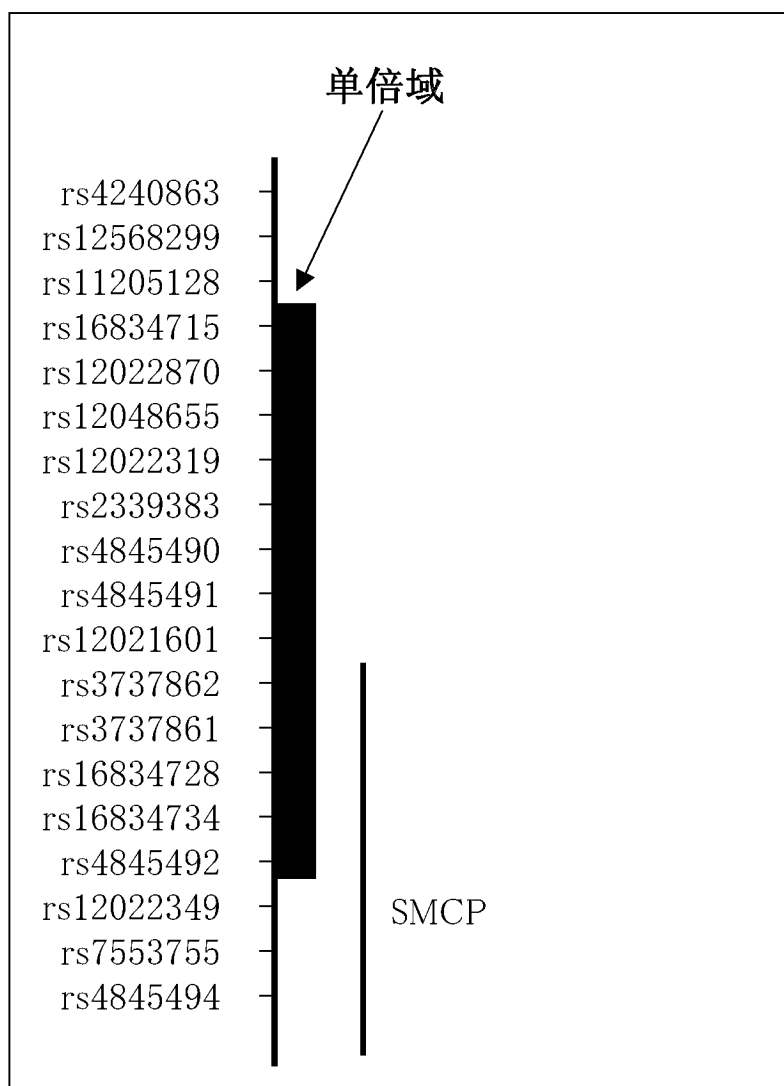


图 8

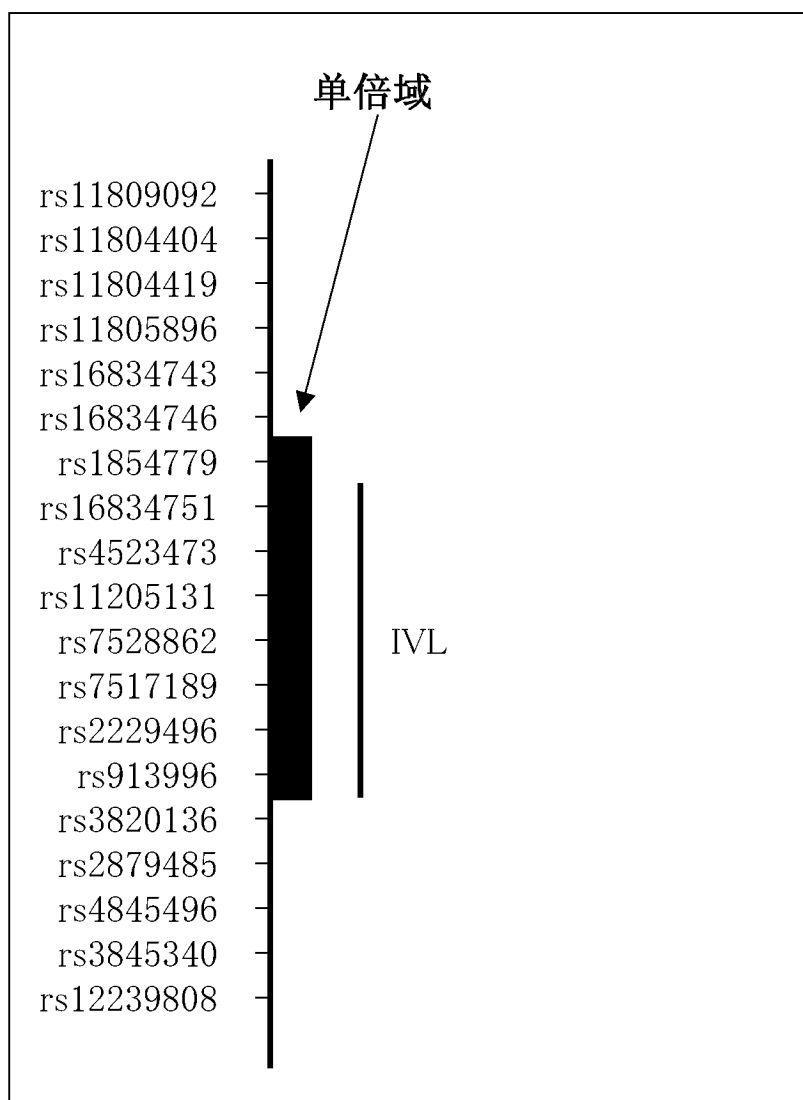


图 9

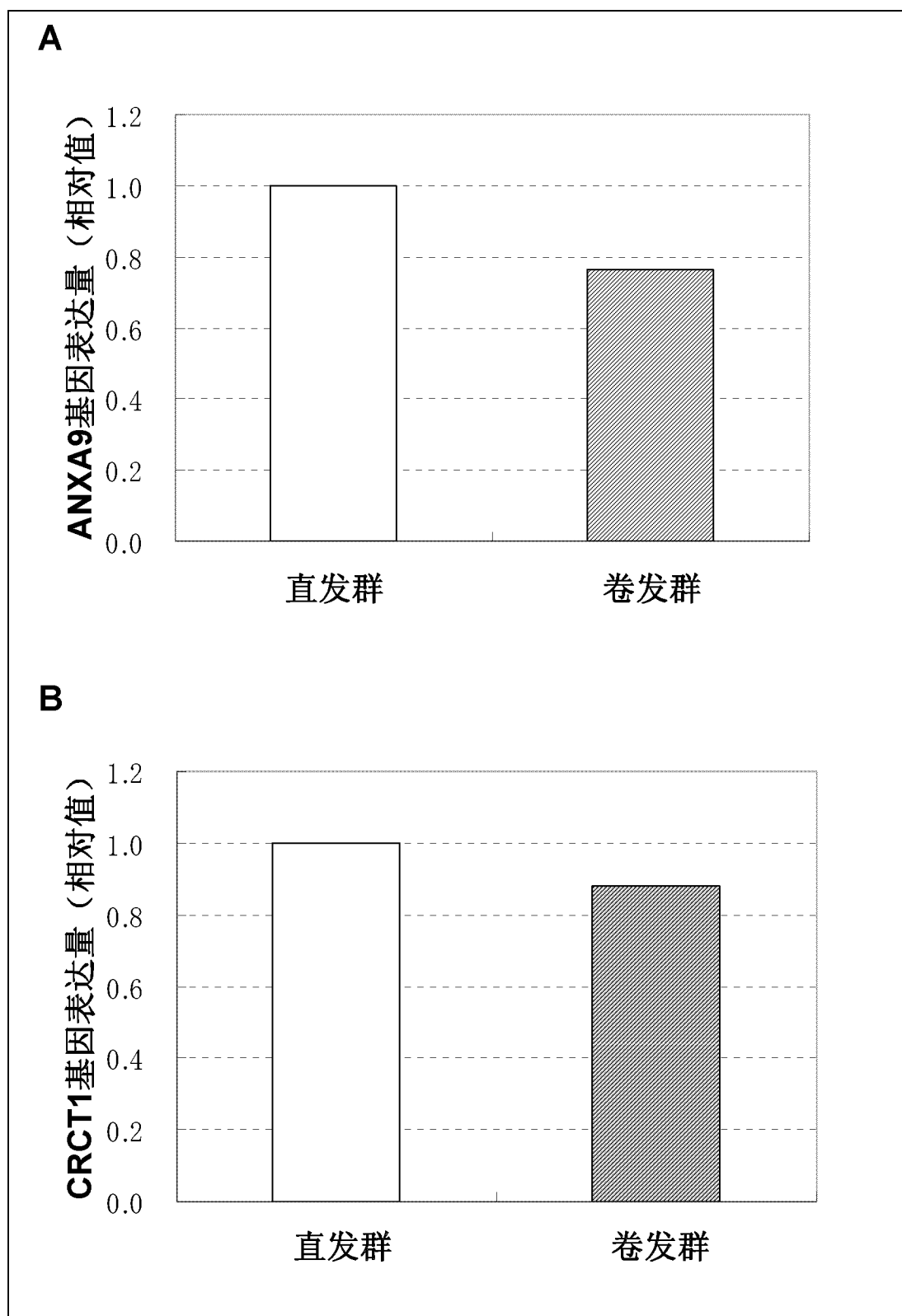


图 10-1

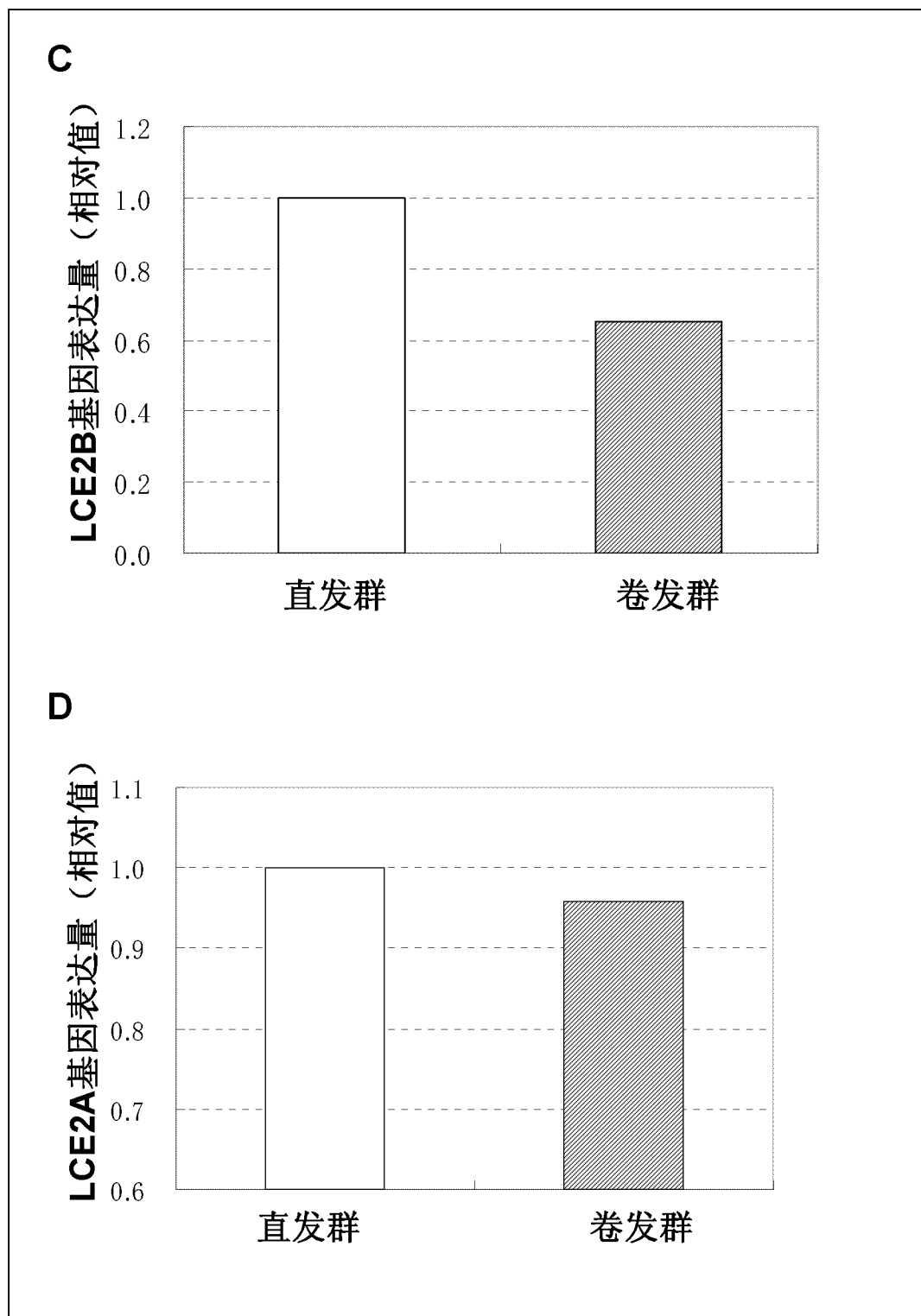


图 10-2

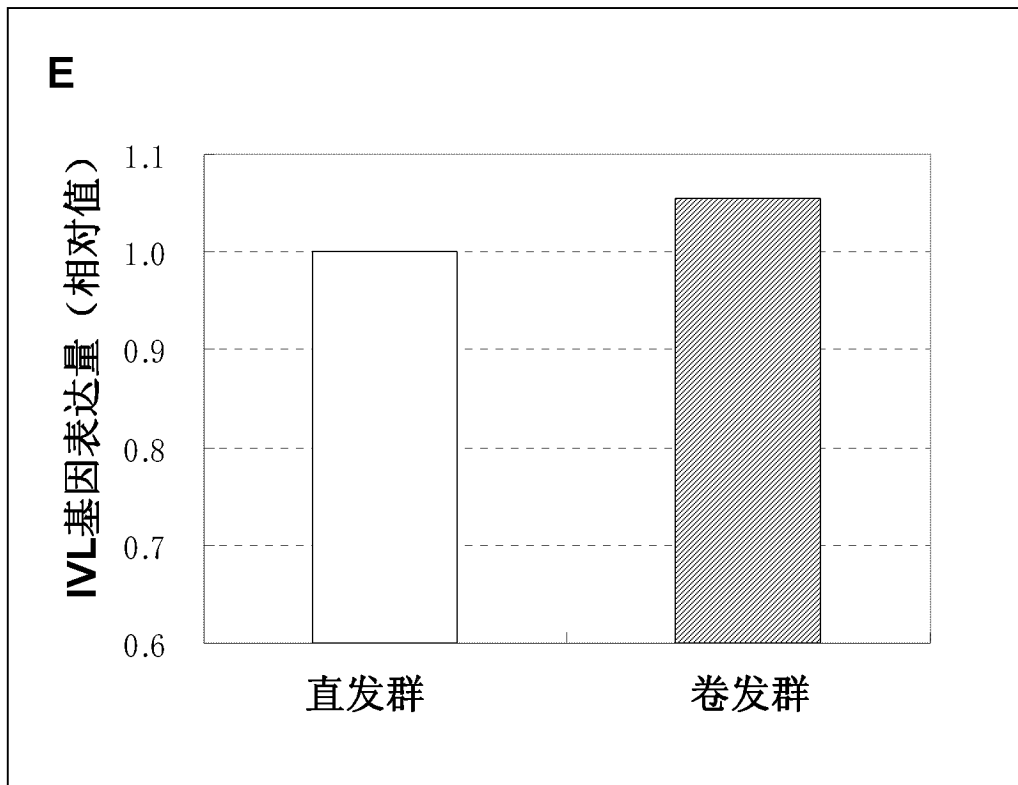


图 10-3

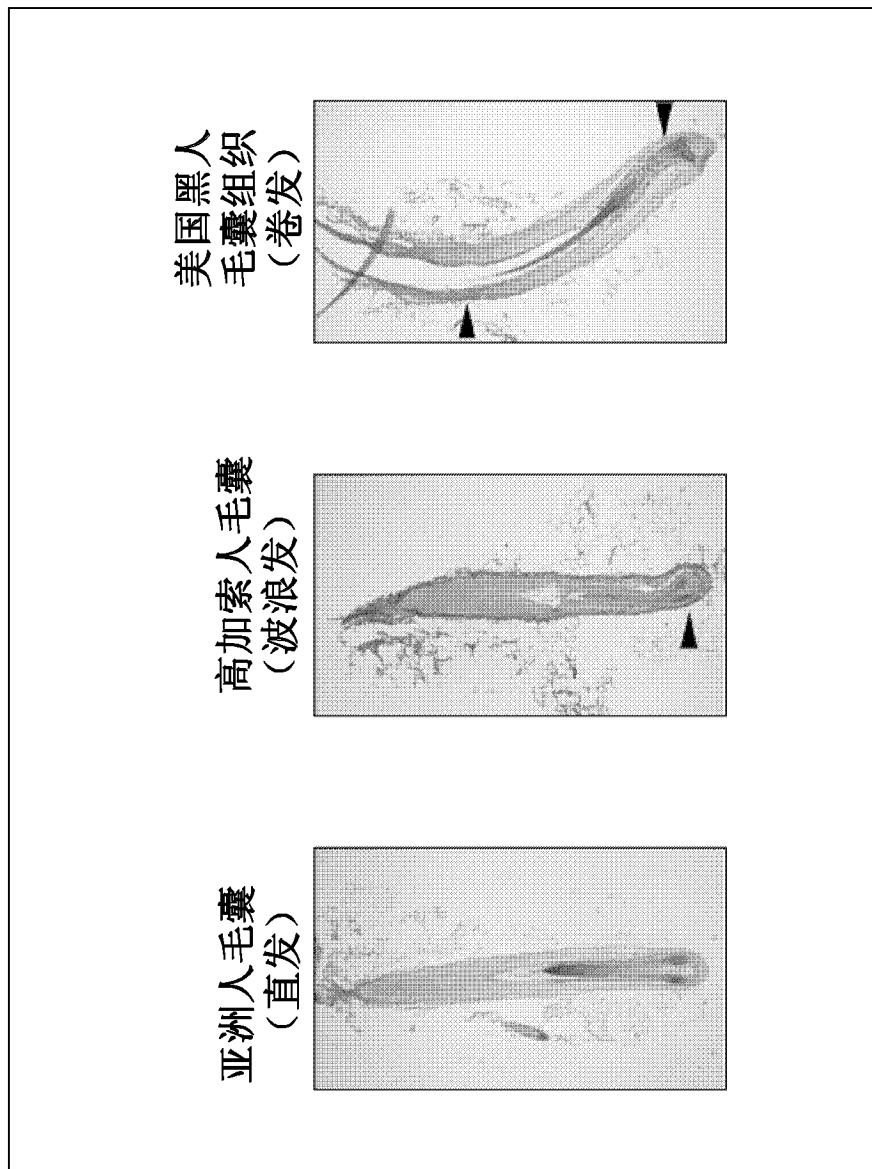


图 11

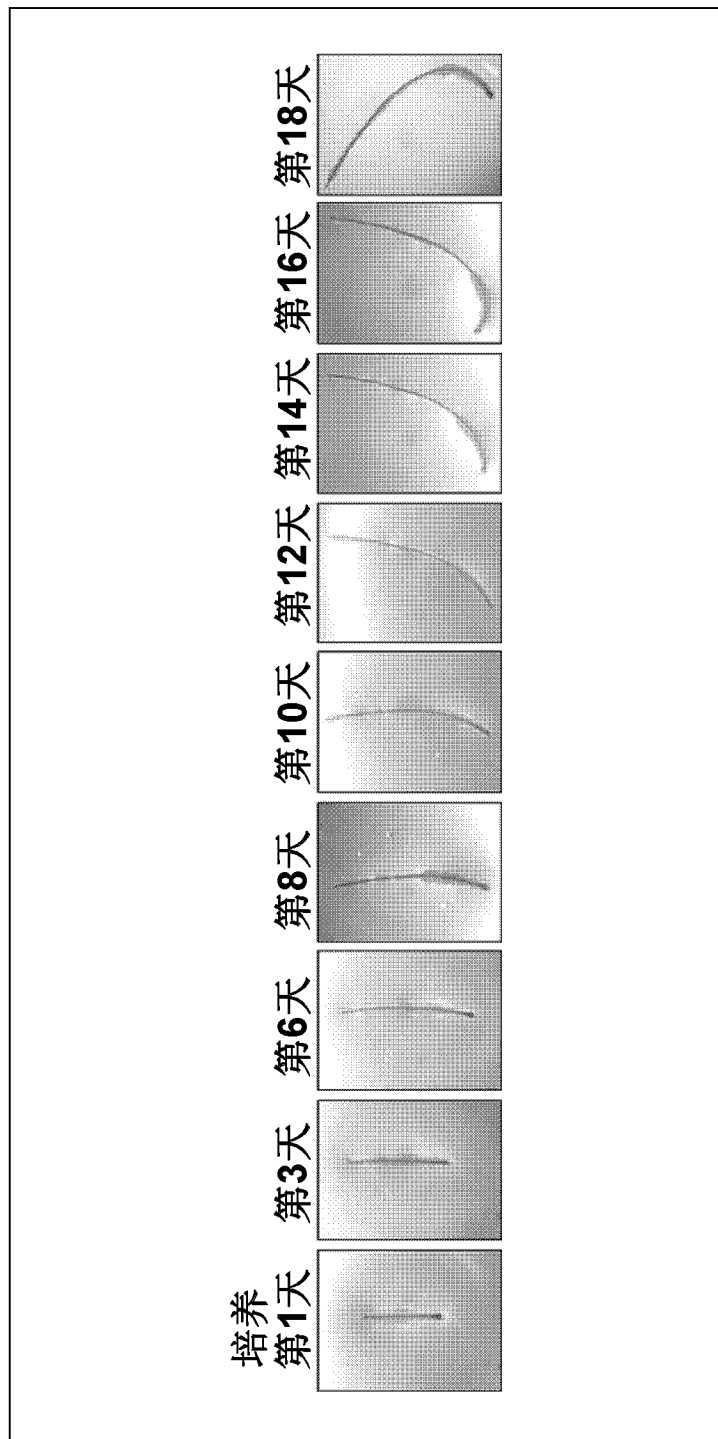


图 12

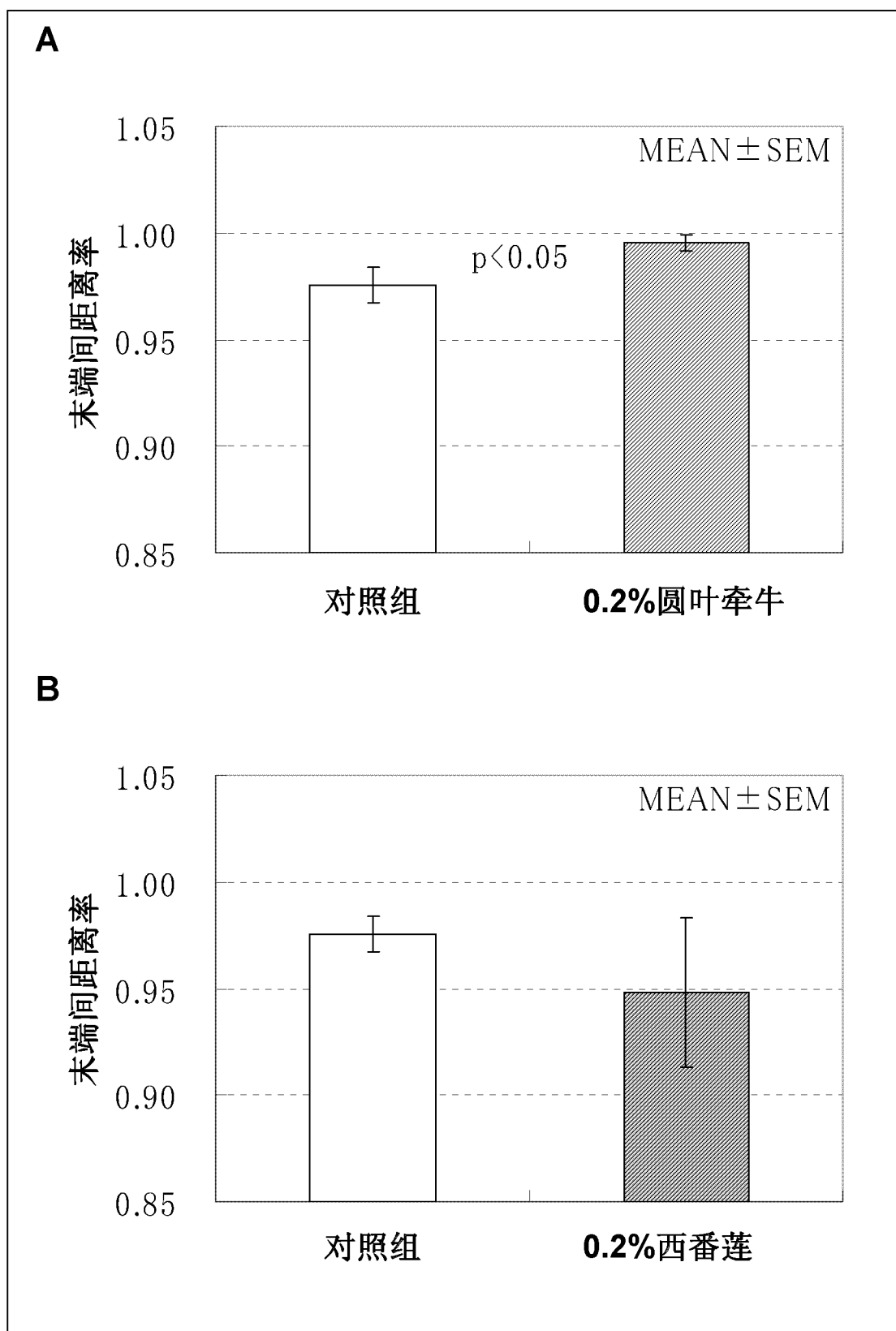


图 13