

[19]中华人民共和国国家知识产权局

[51]Int. Cl⁷

[12] 发明专利申请公开说明书

[21] 申请号 98812216.2

A61K 38/17
A61K 31/00 A61K 31/47
A61K 31/495 A61K 31/55
A61K 31/70 C12N 5/00
//A61K48/00 C07K14/46
C12N15/11 C12N15/12
C07H21/04
[11]公开号 CN 1306436A

[43]公开日 2001 年 8 月 1 日

[22]申请日 1998.10.20 [21]申请号 98812216.2
[30]优先权
[32]1997.10.20US [33]US [31]08/955,552
[32]1998.9.11 [33]US [31]09/151,999
[86]国际申请 PCT/US98/22227 1998.10.20
[87]国际公布 WO99/20298 英 1999.4.29
[85]进入国家阶段日期 2000.6.15
[71]申请人 昂托基尼公司
地址 美国马萨诸塞州
[72]发明人 伊丽莎白·A·王

[74]专利代理机构 永新专利商标代理有限公司
代理人 林晓红

权利要求书 7 页 说明书 122 页 附图页数 3 页

[54]发明名称 HEDGEHOG 类多肽对上皮组织的调节及其相关配制品和用途

[57]摘要

本申请涉及通过将上皮细胞在体外或体内与 hedgehog 治疗剂或 ptc 治疗剂异位接触而调节上皮细胞的生长状态的方法,其中所述的 hedgehog 治疗剂或 ptc 治疗剂的用量是例如相对于不给予该治疗剂时足以改变上皮细胞的增殖速率(促进或抑制)的量。本发明的方法例如可用于调节上皮组织的生长状态,如诱导皮肤或其它皮组织的形成,或诱导毛发的生长。

ISSN 1008-4274

权 利 要 求 书

1、一种调节上皮细胞生长状态的方法，包括将上皮细胞与有效量的改变上皮细胞增殖速率的制剂异位接触，其中该制剂选自 hedgehog 治疗剂和 ptc 治疗剂。

2、一种调节上皮组织生长状态的方法，包括将组织与有效量的改变组织中的上皮细胞增殖速率的制剂异位接触，其中该制剂选自 hedgehog 治疗剂和 ptc 治疗剂。

3、一种诱导皮肤形成的方法，包括用有效量的诱导新皮肤组织形成的制剂处理皮肤，其中该制剂选自 hedgehog 治疗剂和 ptc 治疗剂。

4、一种诱导动物毛发生长的方法，包括用有效量的诱导毛发生长的制剂处理动物，其中该制剂选自 hedgehog 治疗剂和 ptc 治疗剂。

5、权利要求 1 或 2 的方法，其中该制剂增加上皮细胞的增殖速率。

6、权利要求 1 或 2 的方法，其中该制剂降低上皮细胞的增殖速率。

7、权利要求 1 或 2 的方法，其中上皮细胞是皮肤上皮细胞。

8、权利要求 7 的方法，其中上皮细胞是皮肤角质细胞。

9、权利要求 7 的方法，其中上皮细胞是粘膜上皮细胞。

10、权利要求 7 的方法，其中上皮细胞是上皮干细胞。

11、权利要求 7 的方法，其中上皮细胞是毛囊干细胞。

12、权利要求 1 的方法，其中该细胞处于培养中，该制剂以细胞培养添加剂的形式提供。

13、权利要求 1 的方法，其中该细胞在动物中被处理，该制剂以治疗组合物形式给予动物。

14、权利要求 1 的方法，其中上皮组织处于组织培养中，该制剂以组织培养添加剂形式提供。

15、权利要求 1 的方法，其中上皮组织在动物中被处理，该制剂以治疗组合物形式给予动物。

16、权利要求 3、4、13 或 15 的方法，其中该制剂是局部施用。

17、权利要求 1、2、3 或 4 的方法，其中该制剂是 hedgehog 治疗剂。

18、权利要求 17 的方法，其中 hedgehog 治疗剂是一种多肽，该多肽包括至少 hedgehog 蛋白的生物活性胞外部分的 *hedgehog* 多肽序列。

19、权利要求 18 的方法，其中该多肽包括至少 50 个 *hedgehog* 蛋白的 N-末端部分的氨基酸残基。

20、权利要求 18 的方法，其中该多肽包括至少 100 个 *hedgehog* 蛋白的胞外结构域的氨基酸。

21、权利要求 18 的方法，其中该多肽包括相应于 *hedgehog* 蛋白的胞外结构域的 19kd 片段的至少一部分 *hedgehog* 蛋白。

22、权利要求 18 的方法，其中 *hedgehog* 蛋白由脊椎动物的基因编码。

23、权利要求 18 的方法，其中该多肽包括 SEQ ID NO: 21 代表的 *hedgehog* 多肽序列。

24、权利要求 18 的方法，其中该多肽包括 SEQ ID NO: 22 代表的 *hedgehog* 多肽序列。

25、权利要求 18 的方法，其中 *hedgehog* 蛋白由人 *hedgehog* 基因编码。

26、权利要求 18 的方法，其中 *hedgehog* 多肽序列与选自 SEQ ID NO: 9、SEQ ID NO: 10、SEQ ID NO: 11、SEQ ID NO: 12、SEQ ID NO: 13、SEQ ID NO: 14、SEQ ID NO: 15 和 SEQ ID NO: 16 的 *hedgehog* 蛋白的氨基酸序列有至少 60% 相同性。

27、权利要求 26 的方法，其中 *hedgehog* 多肽序列与选自 SEQ ID NO: 9、SEQ ID NO: 10、SEQ ID NO: 11、SEQ ID NO: 12、SEQ ID NO: 13、SEQ ID NO: 14、SEQ ID NO: 15 和 SEQ ID NO: 16 的 *hedgehog* 蛋白的氨基酸序列有至少 75% 相同性。

28、权利要求 26 的方法，其中 *hedgehog* 多肽序列与选自 SEQ ID NO: 9、SEQ ID NO: 10、SEQ ID NO: 11、SEQ ID NO: 12、SEQ ID NO: 13、SEQ ID NO: 14、SEQ ID NO: 15 和 SEQ ID NO: 16 的 *hedgehog* 蛋白的氨基酸序列有至少 85% 相同性。

29、权利要求 26 的方法，其中 *hedgehog* 多肽序列由在严格条件下与选自 SEQ ID NO: 1、SEQ ID NO: 2、SEQ ID NO: 3、SEQ ID NO: 4、SEQ ID NO: 5、SEQ ID NO: 6、SEQ ID NO: 7 和 SEQ ID NO: 8 的序列杂交的核苷酸序列编码。

30、权利要求 23 的方法，其中 *hedgehog* 多肽序列是选自 SEQ ID NO: 9、SEQ ID NO: 10、SEQ ID NO: 11、SEQ ID NO: 12、SEQ ID NO: 13、SEQ ID NO: 14、SEQ ID NO: 15 和 SEQ ID NO: 16 的 *hedgehog* 蛋白的氨基酸序列。

31、权利要求 18 的方法，其中 *hedgehog* 多肽序列是 Sonic *hedgehog* 蛋白的氨基酸序列。

32、权利要求 18 的方法，其中 *hedgehog* 多肽序列是 Indian *hedgehog* 蛋白的氨基酸序列。

33、权利要求 18 的方法，其中 *hedgehog* 多肽序列是 Desert *hedgehog* 蛋白的氨基酸序列。

34、权利要求 18 的方法，其中 *hedgehog* 多肽序列包括约相应于 SEQ ID NO: 15 的 24—193 位残基的氨基酸序列。

35、权利要求 18 的方法，其中该多肽纯化至至少 80%（干重）。

36、权利要求 18 的方法，其中该多肽是重组产生的多肽。

- 37、权利要求 18 的方法，其中该多肽是化学合成的多肽。
- 38、权利要求 17 的方法，其中 hedgehog 治疗剂是 hedgehog 多肽序列的肽模拟物。
- 39、权利要求 1、2、3 或 4 的方法，其中该制剂是 ptc 治疗剂。
- 40、权利要求 39 的方法，其中 ptc 治疗剂是与 *patched* 蛋白结合并脱阻抑 *patched* 介导的对有丝分裂的抑制的一种小有机分子。
- 41、权利要求 39 的方法，其中 ptc 治疗剂与 *patched* 结合并模拟 hedgehog 介导的 *patched* 信号转导。
- 42、权利要求 41 的方法，其中 ptc 治疗剂是小有机分子。
- 43、权利要求 41 的方法，其中 ptc 治疗剂与 *patched* 结合导致 *patched* 和/或 *gli* 表达的正调节。
- 44、权利要求 39 的方法，其中 ptc 治疗剂是与上皮细胞相互作用以诱导 hedgehog 介导的 *patched* 信号转导的小有机分子。
- 45、权利要求 44 的方法，其中 ptc 治疗剂通过改变参与 *patched* 信号途径的一种胞内蛋白的定位、蛋白质-蛋白质结合和/或酶活性来诱导 hedgehog 介导的 *patched* 信号转导。
- 46、权利要求 39 的方法，其中 ptc 治疗剂改变 hedgehog 蛋白、*patched* 蛋白或参与 *patched* 的胞内信号转导途径的蛋白质的表达水平。
- 47、权利要求 46 的方法，其中 ptc 治疗剂是抑制一种蛋白质的表达的反义构建体，该蛋白参与 *patched* 的胞内信号转导途径并且其表达拮抗 hedgehog 介导的信号。
- 48、权利要求 47 的方法，其中该反义构建体是长度约 20—30 个核苷酸并且 GC 含量至少为 50% 的寡核苷酸。
- 49、权利要求 48 的方法，其中该反义寡核苷酸选自：
- 5'-GTCCTGGCGCCGCCGCCGCCGTCGCC;

5'-TTCCGGATGACCGGCCTTTCGCGGTGA; 和

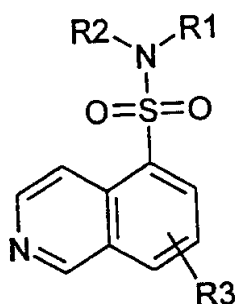
5'-GTGCACGGAAAGGTGCAGGCCACACT。

50、权利要求 44 的方法，其中 *ptc* 治疗剂是与 *patched* 结合并调节 *patched* 依赖性基因表达的小有机分子。

51、权利要求 44 的方法，其中 *ptc* 治疗剂是蛋白激酶 A (PKA) 的抑制剂。

52、权利要求 51 的方法，其中 PKA 抑制剂是 5-异喹啉磺酰胺。

53、权利要求 51 的方法，其中 PKA 抑制剂由下述通式代表：



其中，

R1 和 R2 各自独立地代表氢，并且若化合价和稳定性允许则代表低级烷基，低级链烯基，低级炔基，羰基（如羧基、酯、甲酸酯或酮），硫代羰基（如硫酯、硫代乙酸酯或硫代甲酸酯），氨基，酰基氨基，酰氨基，氰基，硝基，叠氮基，硫酸酯，磺酸酯，亚磺酰氨基， $-(CH_2)_m-R_8$ ， $-(CH_2)_m-OH$ ， $-(CH_2)_m-O$ -低级烷基， $-(CH_2)_m-O$ -低级链烯基， $-(CH_2)_n-O-(CH_2)_m-R_8$ ， $-(CH_2)_m-SH$ ， $-(CH_2)_m-S$ -低级烷基， $-(CH_2)_m-S$ -低级链烯基， $-(CH_2)_n-S-(CH_2)_m-R_8$ ，或者

R1 和 R2 一起与 N 形成一杂环（取代的或未取代的）；

R3 不存在或代表异喹啉环的一或多个取代，如低级烷基，低级链烯基，低级炔基，羰基（如羧基、酯、甲酸酯或酮），硫代羰基（如硫酯、硫代乙酸酯或硫代甲酸酯），氨基，酰基氨基，酰氨基，氰基，

硝基, 叠氮基, 硫酸酯, 磺酸酯, 亚磺酰氨基, $-(\text{CH}_2)_m\text{-R}_8$, $-(\text{CH}_2)_m\text{-OH}$, $-(\text{CH}_2)_m\text{-O-低级烷基}$, $-(\text{CH}_2)_m\text{-O-低级链烯基}$, $-(\text{CH}_2)_n\text{-O-}(\text{CH}_2)_m\text{-R}_8$, $-(\text{CH}_2)_m\text{-SH}$, $-(\text{CH}_2)_m\text{-S-低级烷基}$, $-(\text{CH}_2)_m\text{-S-低级链烯基}$, $-(\text{CH}_2)_n\text{-S-}(\text{CH}_2)_m\text{-R}_8$;

R_8 代表取代的或未取代的芳基, 芳烷基, 环烷基, 环烯基或杂环; 以及

n 和 m 各自独立为 0 或 1—6 的整数。

54、权利要求 51 的方法, 其中 PKA 抑制剂是环化 AMP 类似物。

55、权利要求 51 的方法, 其中 PKA 抑制剂选自 N-[2-((p-溴肉桂基)氨基)乙基]-5-异喹啉磺酰胺, 1-(5-异喹啉磺酰基)-2-甲基哌嗪, KT5720, 8-溴-cAMP, 二丁酰-cAMP 和 PKA 热稳定抑制剂同工型 α 。

56、一种促进皮肤上皮细胞增殖的方法, 包括将细胞与有效量的增加上皮细胞增殖速率的多肽异位接触, 该多肽包括含有至少 hedgehog 蛋白的 N-末端部分的生物活性部分的氨基酸序列。

57、权利要求 56 的方法, 该方法用作控制创伤愈合过程的治疗方法的一部分。

58、权利要求 56 的方法, 其中该疗法选自烧伤治疗、皮肤再生、皮肤移植、褥疮治疗、皮肤溃疡治疗、手术后疤痕减轻和溃疡性结肠炎治疗。

59、权利要求 56 的方法, 其中该方法用作秃头症治疗法的一部分。

60、一种抑制皮肤上皮细胞增殖的方法, 包括将细胞与有效量的降低上皮细胞增殖速率的制剂异位接触, 该制剂抑制 hedgehog 蛋白与 patched 结合。

61、权利要求 60 的方法, 其中上皮细胞是毛囊细胞。

62、权利要求 61 的方法，该方法抑制毛发生长。

63、一种多肽制剂，该多肽包括含有 hedgehog 蛋白的生物活性片段的 hedgehog 多肽序列，该多肽被配制用于局部施用至上皮组织。

64、权利要求 63 的制剂，其中该多肽被配制用于局部施用至皮肤。

65、权利要求 63 的制剂，其中该多肽包括至少 50 个 *hedgehog* 蛋白 N-末端部分的氨基酸残基。

66、权利要求 63 的制剂，其中该多肽包括至少 100 个 hedgehog 蛋白的胞外结构域的氨基酸。

67、权利要求 63 的制剂，其中该多肽包括相应于 hedgehog 蛋白的胞外结构域的 19kd 片段的至少一部分 hedgehog 蛋白。

68、权利要求 63 的制剂，其中 hedgehog 蛋白由脊椎动物的基因编码。

说明书

HEDGEHOG 类多肽对上皮组织的调节及其相关配制品和用途

发明背景

图式形成是胚胎细胞用以形成分化组织的有序空间排列的活性。高等生物体的物理复杂性在胚胎发生期间经细胞内在谱系和细胞外在信号的相互作用而产生。诱导性相互作用对于脊椎动物发育中从身体轮廓的最早建立到胚胎图式形成、器官系统的图式形成、组织分化过程中不同细胞类型的产生都是很重要的 (Davidson, E., (1990) 发育 108: 365—389; Gurdon, J.B., (1992) 细胞 68: 185—199; Jessell, T.M.等(1992) 细胞 68: 257—270)。发育细胞相互作用的作用是不同的。典型地, 通过诱导与应答细胞的未诱导和诱导状态均不同的细胞, 应答细胞从细胞分化的一个途径转向另一个途径 (诱导)。有时细胞诱导其相邻细胞象其一样分化 (同源诱导); 在另一些情况下, 细胞抑制其相邻细胞象其一样分化。早期发育中细胞相互作用可以是相继进行的, 从而两种细胞类型间的初始诱导导致多样性的渐进扩增。另外, 诱导性相互作用不仅在胚胎中发生, 而且在成熟细胞中也发生, 并且可以发挥作用建立和保持形态发生图式以及诱导分化 (J.G. Gurdon(1992) 细胞 68: 185—199)。

信号分子的 *Hedgehog* 家族成员在无脊椎动物和脊椎动物发育中介导许多重要的短程和长程图式形成过程。在苍蝇中, 一个单 *hedgehog* 基因调节节段和器官芽图式形成。相反, 在脊椎动物中, *hedgehog* 基因家族参与左—右不对称性、CNS 中的极性、体节和肢、器官发生、软骨发生以及精子发生的调控。

第一个 *hedgehog* 基因是通过在黑腹果蝇中的遗传筛选鉴定的 (Nusslein-Volhard, C.和 Wieschaus, E.(1980) 自然 287, 795—801),

这一筛选鉴定了若干个影响胚胎和幼虫发育的突变。在 1992 和 1993 年,报道了果蝇 *hedgehog(hh)* 基因的分子学性质 (C.F. Lee 等(1992) 细胞 71, 33—50), 从那时起, 从各种脊椎动物物种中分离了几个 *hedgehog* 同系物。尽管在果蝇和其它无脊椎动物中仅发现一个 *hedgehog* 基因, 但是在脊椎动物中存在多个 *Hedgehog* 基因。

各种 *Hedgehog* 蛋白由信号肽、高保守的 N—末端区和较分化的 C—末端结构域组成。除了在分泌途径的信号序列裂解 (Lee, J.J.等人 (1992) 细胞 71: 33—50; Tabata, T.等人(1992) 基因发育 2635—2645; Chang, D.E.等人(1994) 发育 120: 3339—3353) 之外, *Hedgehog* 前体蛋白发生内部自我蛋白酶解, 这种蛋白酶解取决于 C—末端部分的保守序列 (Lee 等人(1994) 科学 266: 1528—1537; Porter 等(1995) 自然 374: 363—366)。这一自我裂解导致产生一 19kD N—末端肽和一 26—28kD C—末端肽 (Lee 等人(1992), 文献同上; Tabata 等人(1992), 文献同上; Chang 等人(1994), 文献同上; Lee 等人(1994), 文献同上; Bumcrot, D.A.等人(1995) 分子细胞生物学 15: 2294—2303; Porter 等(1995), 文献同上; Ekker, S.C.等(1995) 当前生物学 5: 944—955; Lai, C.J.等(1995) 发育 121: 2349—2360)。N—末端肽与合成其的细胞表面紧密结合, 而 C—末端肽在体外和体内可自由扩散 (Lee 等人 (1994), 文献同上; Bumcrot 等人(1995), 文献同上; Mart', E.等(1995) 发育 121: 2537—2547; Roelink, H.等(1995) 细胞 81: 445—455)。令人感兴趣地, N—末端肽在细胞表面的保留依赖于自我裂解, 因为, 由在内部裂解的正常位置精确终止的一 RNA 编码的 HH 的一截短形式在体外 (Porter 等(1995), 文献同上) 和体内 (Porter, J.A.等(1996) 细胞 86, 21—34) 均是可扩散的。生物化学研究表明 HH 前体蛋白的自我蛋白酶解经一内部硫酯中间体, 其随后在一亲核取代中裂解。很可能亲核物质是一小亲核分子, 其与 N—肽的 C—末端共价结合

(Porter 等(1996), 文献同上), 使其粘连在细胞表面上。该生物学暗示是意义深远的。粘连的结果是在 Hedgehog 产生细胞的表面产生高局部浓度的 N-末端 Hedgehog 肽。这一 N-末端肽对于果蝇和脊椎动物中的短程和长程 Hedgehog 信号活动是必需的和充分的 (Porter 等(1995), 文献同上; Ekker 等(1995), 文献同上; Lai 等(1995), 文献同上; Roelink, H.等(1995) 细胞 81: 445—455; Porter 等(1996), 文献同上; Fietz, M.J.等(1995) 当前生物学 5: 643—651; Fan, C.-M.等(1995) 细胞 81: 457—465; Mart', E.等(1995) 自然 375: 322—325; Lopez-Martinez 等(1995) 当前生物学 5: 791—795; Ekker, S.C.等(1995) 发育 121: 2337—2347; Forbes, A.J.等(1996) 发育 122: 1125—1135)。

已暗示 HH 在果蝇发育中参与短程和长程图式形成的各个位点。在早期胚胎的节段极性的建立中, 其具有似乎被直接介导的短程效应, 而在器官芽的图式形成中, 其通过诱导二级信号而诱导长程效应。

在脊椎动物中, 在过去几年中已克隆了几个 *hedgehog* 基因 (见表 1), 在这些基因中, 对 *Shh* 进行了很多实验, 因为其在不同的组织中心中表达, 而这些组织中心是邻近组织图式形成的信号来源。最近的证据表明 *Shh* 参与这些相互作用。

hedgehog 蛋白与其相关受体之一 *patched* 的相互作用引发一级联, 参与下游效应器的活化和抑制, 在某些情况下, 其最终结果是在基因的转录和翻译中发生可检测的变化。Hedgehog 信号的转录靶是 *patched* 基因本身 (Hidalgo 和 Ingham, 1990 发育 110, 291—301; Marigo 等, 1996), 以及果蝇 *cubitus interruptus*(Ci)基因的脊椎动物同系物 *GLI* 基因 (Hui 等(1994) 发育生物学 162: 402—413)。已显示 *patched* 基因的表达在应答 *Shh* 的肢芽和神经板细胞中被诱导 (Marigo 等(1996) 发育 122: 1225—1233)。GLI 基因编码具有锌指

DNA 结合结构域的潜在的转录因子 (Orenic 等(1990) 基因和发育 4: 1053—1067; Kinzler 等(1990) 分子细胞生物学 10: 634—642)。已报道 *GLI* 基因的转录在肢芽中应答 *hedgehog* 呈正调节, 而 *GLI3* 基因的转录应答 *hedgehog* 诱导呈负调节 (Marigo 等(1996) 发育 122: 1225—1233)。另外, 已证实 *Ci* 的升高水平即使在缺乏 *hedgehog* 活性时也足以激活 *patched(ptc)* 和其它 *hedgehog* 靶基因。

发明概述

本申请的一个方面涉及通过将上皮细胞在体外或体内与 *hedgehog* 治疗剂或 *ptc* 治疗剂异位接触而调节上皮细胞的生长状态的方法, 其中所述的 *hedgehog* 治疗剂或 *ptc* 治疗剂的用量是例如相对于不给予该 *hedgehog* 治疗剂或 *ptc* 治疗剂时足以改变上皮细胞的增殖速率 (促进或抑制) 的量。本发明的方法例如可用于调节上皮组织的生长状态, 如诱导皮肤或其它皮组织的形成, 或诱导毛发的生长。

当本发明方法用 *hedgehog* 治疗剂进行时, 该 *hedgehog* 治疗剂优选是一种包括具有至少 *hedgehog* 蛋白的生物活性胞外部分的 *hedgehog* 部分的多肽, 例如, 该 *hedgehog* 部分包括 *hedgehog* 蛋白的 N-末端部分的至少 50, 100 或 150 个 (连续) 氨基酸残基。在一优选的实施方案中, *hedgehog* 部分包括至少相应于 *hedgehog* 蛋白胞外结构域的 19kd 片段的 *hedgehog* 蛋白的一部分。

在某些优选的实施方案中, 所述的 *hedgehog* 部分具有与 SEQ ID NO: 10—18 或 20 的任一个 *hedgehog* 蛋白至少 60%、75%、85% 或 95% 相同的氨基酸序列, 当然与这些序列相同的序列也适用于本发明方法。所述的 *hedgehog* 部分可由在严格条件下与 SEQ ID NO: 1—9 或 19 的任一个核酸序列杂交的核酸编码, 例如, 所述的 *hedgehog* 部分可由脊椎动物 *hedgehog* 基因、特别是人 *hedgehog* 基因编码。

在某些实施方案中，所述的 hedgehog 多肽用一或多个固醇部分例如胆固醇或其衍生物修饰。

在某些实施方案中，所述的 hedgehog 多肽用一或多个脂肪酸部分修饰，例如选自肉豆蔻酰基、棕榈酰基、硬脂酰基和花生四烯酰基的脂肪酸部分。

在某些实施方案中，所述的 hedgehog 多肽用一或多个芳烃修饰，例如苯、吡、菲、蒽、萘、茚或并四苯。

在某些实施方案中，所述的 hedgehog 多肽用 C7—C30 烷基或环烷基修饰一或多次。

在其它实施方案中，本发明方法可以通过给予基因活化构建体而进行，其中所述的基因活化构建体与患者的基因组 *hedgehog* 基因组合以提供与所述的 *hedgehog* 基因的编码序列可操纵连接的异源转录调控序列。

在其它的一些实施方案中，本发明方法可通过给予编码 *hedgehog* 多肽的基因治疗构建体而实施。例如，所述的基因治疗构建体可以提供在选自重组病毒颗粒、脂质体、多聚阳离子核酸结合剂的组合物中。

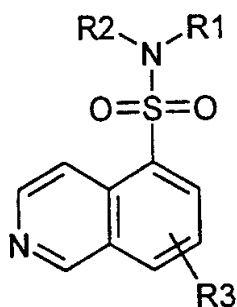
在其它的实施方案中，本发明的方法可以用 *ptc* 治疗剂进行。*ptc* 治疗剂的一个例子是与 *patched* 蛋白结合并使有丝分裂的 *patched* 介导的抑制脱阻抑的小有机分子，例如与 *patched* 结合并模拟 *hedgehog* 介导的 *patched* 信号转导的分子，与 *patched* 结合并调节 *patched* 依赖性基因表达的分子。例如，*ptc* 治疗剂与 *patched* 的结合可导致对 *patched* 和/或 *gli* 表达的正调节。

在更通用的意义上，*ptc* 治疗剂可以是与上皮细胞相互作用以例如通过改变参与 *patched* 信号途径的胞内蛋白的定位、蛋白-蛋白结合和/或酶活性而诱导 *hedgehog* 介导的 *patched* 信号转导的小有机分

子。例如，*ptc* 治疗剂可以改变 *hedgehog* 蛋白、*patched* 蛋白或参与 *patched* 胞内信号转导途径的蛋白的表达水平。

在某些实施方案中，*ptc* 治疗剂是一种反义构建体，其抑制参与 *patched* 信号转导途径的蛋白的表达，并且其表达拮抗 *hedgehog* 介导的信号。该反义构建体优选是长度为 20—30 个核苷酸的寡核苷酸并且 GC 含量至少为 50%。

在其它实施方案中，*ptc* 治疗剂是蛋白激酶 A (PKA) 的抑制剂，如 5-异喹啉磺酰胺。PKA 抑制剂可以是环化 AMP 类似物。PKA 抑制剂的例子包括 N-[2-((p-溴肉桂基)氨基)乙基]-5-异喹啉磺酰胺，1-(5-异喹啉磺酰基)-2-甲基哌嗪，KT5720，8-溴-



其中，

R1 和 R2 各自独立地代表氢，并且若化合价和稳定性允许则代表低级烷基，低级链烯基，低级炔基，羰基（如羧基、酯、甲酸酯或酮），硫代羰基（如硫酯、硫代乙酸酯或硫代甲酸酯），氨基，酰基氨基，酰氨基，氰基，硝基，叠氮基，硫酸酯，磺酸酯，亚磺酰氨基， $-(CH_2)_m-R_8$ ， $-(CH_2)_m-OH$ ， $-(CH_2)_m-O$ -低级烷基， $-(CH_2)_m-O$ -低级链烯基， $-(CH_2)_n-O-(CH_2)_m-R_8$ ， $-(CH_2)_m-SH$ ， $-(CH_2)_m-S$ -低级烷基， $-(CH_2)_m-S$ -低级链烯基， $-(CH_2)_n-S-(CH_2)_m-R_8$ ，或者

R1 和 R2 一起与 N 形成一杂环（取代的或未取代的）；

R3 不存在或代表异喹啉环的一或多个取代，如低级烷基，低级链烯基，低级炔基，羰基（如羧基、酯、甲酸酯或酮），硫代羰基（如

硫酯、硫代乙酸酯或硫代甲酸酯), 氨基, 酰基氨基, 酰氨基, 氰基, 硝基, 叠氮基, 硫酸酯, 磺酸酯, 亚磺酰氨基, $-(CH_2)_m-R_8$, $-(CH_2)_m-OH$, $-(CH_2)_m-O$ -低级烷基, $-(CH_2)_m-O$ -低级链烯基, $-(CH_2)_n-O-(CH_2)_m-R_8$, $-(CH_2)_m-SH$, $-(CH_2)_m-S$ -低级烷基, $-(CH_2)_m-S$ -低级链烯基, $-(CH_2)_n-S-(CH_2)_m-R_8$;

R_8 代表取代的或未取代的芳基, 芳烷基, 环烷基, 环烯基或杂环; 以及

n 和 m 各自独立为 0 或 1—6 的整数。

本发明方法可用于治疗例如上皮失调, 如控制创伤愈合过程。例如, 本发明方法可作为烧伤治疗、皮肤再生、皮肤移植、褥疮治疗、皮肤溃疡治疗、手术后减轻疤痕以及溃疡性结肠炎治疗等治疗的一部分。在控制毛发生长中, 本发明方法可用作秃头症治疗的一部分。

本发明的另一方面涉及配制用于上皮组织如皮肤局部应用的 *hedgehog* 或 *ptc* 治疗剂的制备。例如, 这种配制品可包括一种多肽, 该多肽具有包括 *hedgehog* 蛋白的生物活性片段的 *hedgehog* 多肽序列, 该多肽配制用于上皮组织局部应用。

附图的简要描述

图 1A, B 和 C 示出了在用各种 *hedgehog* 配制品处理的小鼠上对毛发生长的诱导。

发明的详细描述

正常皮肤表皮是一种复杂的上皮组织, 其含有处于增殖、分化和脱皮中的角质细胞, 并且其分层能使角质细胞中的形态学和功能变化以有序进程发生。因为角质细胞的增殖持续补偿从皮肤表面脱落的细胞的损失, 所以正常表皮保持一种动态稳态。在表皮内, 增殖发生在附着于下面的基膜的基底层角质细胞中, 并且当细胞迁移

通过上基层时，其发生终末分化，最终作为死的角质鳞从组织表面脱落。通过细胞动力学分析已鉴定了基层角质细胞的三个亚群：干细胞、转运扩增细胞和定型细胞。干细胞在成熟生命过程中保持高自我更新能力，并最终负责表皮的保持和修复。干细胞的子代细胞可以是干细胞本身或是已知为转运扩增细胞的细胞。转运扩增细胞分裂次数较少，但是产生不可逆退出细胞循环并进行终末分化的子代的可能性却很高。

I. 总述

本申请涉及这样一种发现，即 *hedgehog* 多肽的制剂可用于控制上皮组织的形成和/或保持。如所附实施例所述，*hedgehog* 蛋白质参与上皮干细胞的增殖并可提供调节干细胞分化成上皮组织的早期信号。一般地，本发明方法包括将上皮细胞与一定量的 *hedgehog* 治疗剂（见下述）接触，根据所述 *hedgehog* 治疗剂是有效的 *hedgehog* 兴奋剂还是有效的 *hedgehog* 拮抗剂，由所述细胞产生(i)诱导上皮组织形成或(ii)抑制上皮组织形成的非毒性应答。本发明方法可在分散在培养物中的上皮细胞或是完整组织或器官的一部分的上皮细胞上进行。另外，本发明方法可在培养物中提供的细胞（体外）或完整动物的细胞（体内）上进行。

在一个方面，本发明提供了用于控制上皮衍生组织的增殖的药物制剂和方法，其使用 *hedgehog* 多肽或其模拟物作为活性成分。本发明还涉及用本发明的药物制剂控制上皮衍生组织的增殖的方法。

本发明的 *hedgehog* 配制品可用作治疗上皮组织疾病的方案的一部分或者手术或美容修复上皮组织的方案的一部分，所述上皮组织例如皮肤和皮肤器官；角膜、晶状体和其它眼组织；粘膜和牙周上皮。本文公开的方法和组合物提供了治疗和预防各种损伤的上皮和粘膜组织的方法。例如，本发明方法可用于控制创伤愈合进程，例

如这在涉及上皮组织的外科手术如皮肤病学或牙周手术中是需要的。用 *hedgehog* 疗法作为候选治疗的外科修复的例子包括严重烧伤和皮肤再生，皮肤移植，褥疮，皮肤溃疡、皮裂、手术后减轻疤痕以及溃疡性结肠炎。

在本发明的另一方面，*hedgehog* 制剂可用于影响毛发的生长，例如用于治疗秃头症，其中毛发的生长被加强，或者例如可用于美容去毛（脱毛），其中毛发的生长被抑制。

在某些实施方案中，本发明的组合物可用于抑制而不是促进上皮衍生组织的生长。例如，本文所公开的一些组合物可用于治疗或防止各种增生或瘤形成症状。所述方法可用于治疗或预防例如银屑病、角化症、痤疮、粉刺损害、毛囊炎和假毛囊炎、角化棘皮瘤、老茧皮、Darier's 病、鱼鳞癣、扁平苔癣、传染性软疣、黑斑病、Fordyce 病、以及斑痕瘤或肥大疤痕。本发明的一些配制品还可用作自身免疫疾病的治疗方案的一部分以影响疾病增殖现象的愈合，例如用作治疗口腔溃疡、天疱疮如普通天疱疮、叶状天疱疮、增殖性天疱疮或红斑状天疱疮、表皮松解、狼疮损害或脱屑损害的一部分。

本发明的 *hedgehog* 治疗对患有这些症状的人和动物个体均有效。适用于本发明的动物个体延及家畜，包括宠物或商用家畜，例如狗、猫、牛、马、绵羊、猪和山羊。

本发明的另一方面提供了一种刺激组织培养物中的上皮组织的生长并调节其分化的方法。

不希望受任何特定理论的限制，*hedgehog* 蛋白对干细胞增殖的诱导至少部分由于这些蛋白质能拮抗（直接或间接）*patched* 介导的基因表达的调节以及由该蛋白介导的其他生理效应。据认为 *patched* 基因产物-一种细胞表面蛋白-发信号通过一途径，导致形态发生素的 Wnt 和 Dpp/BMP 家族成员的转录阻抑，形态发生素是赋予位置信息

的蛋白质。在脊椎动物的 CAN 和肢体图式形成的发育中, *hedgehog* 的导入消除 (脱阻抑) 由 *patched* 赋予的这种抑制, 使得特定基因程序的表达。

最近, 已报道人 *patched* 基因的突变导致遗传性皮肤癌并可能导致零星皮肤癌, *patched* 基因是在果蝇发育途径中鉴定的第一个基因, 例如, 参见 Hahn 等(1996) 细胞 86: 841-851; Johnson 等(1996) 科学 272: 1668-1671。痣样基细胞癌 (NBCC) 由人 *patched* 基因中的突变所致已得到证明, 这是 *patched* 在胚胎后发育中起作用的一个例子。这些发现使得现有技术中了解 *patched* 的一个活性是肿瘤抑制基因, 其可通过抑制来自 *hedgehog* 的增殖信号而起作用。我们下述观察揭示了 *hedgehog/patched* 途径在保持上皮细胞增殖和分化中的潜在新作用。因此, 本发明涉及能模拟 *hedgehog* 蛋白对 *patched* 信号传导的作用的其它制剂的用途, 例如, 所述其它制剂可从下述药物筛选分析中鉴别。

II. 定义

为方便起见, 说明书、实施例和权利要求书中采用的某些术语解释如下。

术语 “*hedgehog* 治疗剂” 指各种形式的 *hedgehog* 多肽以及肽模拟物, 其可通过在实施例中清楚示出的模拟或增强 (兴奋) 或抑制 (拮抗) 天然存在的 *hedgehog* 蛋白的作用而调节上皮细胞的增殖/分化状态。模拟或增强野生型 *hedgehog* 蛋白的活性的 *hedgehog* 治疗剂是 “*hedgehog* 兴奋剂”, 相反, 抑制野生型 *hedgehog* 蛋白的活性的 *hedgehog* 治疗剂是 “*hedgehog* 拮抗剂”。

特别地, 术语 “*hedgehog* 多肽” 包括 *hedgehog* 蛋白及其肽片段的制剂, 可以是兴奋剂和拮抗剂形式, 如下所述。

如本文所用术语 “*hedgehog* 蛋白的生物活性片段” 是指全长

hedgehog 多肽的一个片段，其中所述片段特异地刺激或拮抗由野生型 *hedgehog* 蛋白介导的诱导性事件。*hedgehog* 生物活性片段优选地是 *hedgehog* 蛋白的可溶性胞外部分，其中可溶性是参照生理学相容溶液而言。举例性的生物活性片段见 PCT 出版物 WO95/18856 和 WO96/17924 所述。

术语“*patched*”或“*ptc*”是指相关跨膜蛋白的一个家族，其已证实参与由细胞与 *hedgehog* 蛋白接触而诱导的信号转导中。例如，哺乳动物 *ptc* 家族包括 *ptc1* 和 *ptc2*。除了下述参考文献外，*ptc2* 的例子还参见 Takabatake 等（1997）FEBS 通信 410: 485 和 GenBank AB000847。除非另有说明，应理解针对 *ptc1*（或仅 *ptc*）所述的实施方案也指涉及其它 *ptc* 同系物如 *ptc2* 的等价实施方案。

术语“*ptc* 治疗剂”指这样的制剂，其(i)模拟 *hedgehog* 蛋白对 *patched* 信号转导的作用，例如，其拮抗 *patched* 的细胞周期抑制活性，或者(ii)活化或增强 *patched* 信号转导。在其它实施方案中，*ptc* 治疗剂可以是 *hedgehog* 拮抗剂。*ptc* 治疗剂可以是例如肽、核酸、碳水化合物、小有机分子或天然产物抽提物（或其部分）。

hedgehog 或 *ptc* 治疗剂的“增殖”形式是诱导上皮细胞特别是上皮干细胞的增殖的形式。相反，*hedgehog* 或 *ptc* 治疗剂的“抗增殖”形式是优选地以非毒性方式，例如通过促进或保持分化表型或以其它方式促进沉默而抑制上皮细胞的增殖的形式。

举例来说，而不希望受特定理论的限制，增殖性 *hedgehog* 多肽通常是脱阻抑 *patched* 介导的细胞循环抑制的蛋白形式，例如该多肽模拟天然存在的 *hedgehog* 蛋白对上皮细胞的作用。增殖性 *ptc* 治疗剂包括脱阻抑 *patched* 介导的细胞循环抑制的其它制剂，并可细胞外或细胞内起作用。

举例性的抗增殖性 *ptc* 治疗剂可增强 *patched* 介导的细胞循环抑

制。这种制剂可以是抑制例如 hedgehog 与 patched 结合的小分子，以及刺激和/或增强 patched 蛋白的信号转导途径的制剂。

术语“上皮”指机体内表面和外表面的细胞覆盖层（皮肤、粘膜和血浆），包括腺体和其衍生的其它结构，例如角膜、食管、表皮和毛囊上皮细胞。其它举例性的上皮组织包括：嗅觉上皮，其是鼻腔嗅觉区内层的假分层上皮并含有嗅觉受体；腺上皮，其指由分泌细胞组成的上皮；鳞状上皮，其指由平板状细胞组成的上皮。术语上皮还可指过渡上皮，其在经历由于收缩和膨胀所致的巨大机械变化的中空器官中呈特征性分布，例如代表分层的鳞状上皮和柱形上皮之间的过渡的组织。

术语“上皮化”指由覆盖裸露表面的上皮组织的生长所致的愈合。

术语“皮肤”指机体的外部保护层，由真皮和表皮组成，并理解为包括汗腺和皮脂腺以及毛囊结构。在本申请中，形容词“皮肤的”根据使用时的上下文，可以并且应该理解为通指皮肤的附属物。

术语“表皮”指皮肤的最外部的非血管层，其衍生自胚胎外胚层，厚度为 0.07-1.4mm。在手掌和脚底表面，其从外向内包括 5 层：由垂直排列的柱形细胞组成的基层；由具有短隆起或突起的扁平多面体细胞组成的棘细胞或突起层；由扁平颗粒细胞组成的颗粒层；由几层清晰透明的细胞组成的透明层，该细胞的细胞核不清楚或无细胞核；由扁平的角化的无核细胞组成的角状层。在一般的机体表面的表皮中，通常无透明层。

“真皮(corium)”或“真皮(dermis)”指表皮之下的皮肤层，其由致密的血管结缔组织层组成，并含有神经和感觉的终末器官。毛根和皮脂腺和汗腺是深埋在真皮中的表皮结构。

术语“指甲”指在手指或脚趾远端的背面上的角状皮板。

术语“表皮腺体”指与表皮相关的细胞的集合体，其特异分泌或排泄与其通常代谢需要无关的物质。例如“皮脂腺”是分泌油性物质和皮脂的真皮中的全泌腺体。术语“汗腺”指位于真皮或皮下组织中通过一导管在体表开放的分泌汗液的腺体。

术语“毛发”指线状结构，特别是由角蛋白组成的从真皮中的乳头凹陷中发育的特异表皮结构，其仅由哺乳动物并是该类动物的特征，另外，该术语还指这种毛发的集合体。“毛囊”指环绕毛发的表皮的一个个管状内陷，从中长出毛发；“毛囊上皮细胞”指环绕毛囊中的真皮乳头的上皮细胞，例如干细胞，外根鞘细胞，基质细胞和内根鞘细胞。这种细胞可以是正常的非恶性细胞或转化/永生细胞。

术语“鼻上皮组织”指鼻和嗅觉上皮。

“切割创伤”包括皮肤上皮层的撕裂伤、擦伤、切口、刺伤或破口并可延伸至真皮层，甚至延伸至皮下脂肪及以下。切割创伤可产生自手术程序或产生自皮肤的偶然刺破。

“烧伤”指由于热和/或化学试剂导致个体的大面积皮肤移位或丧失的情况。

“真皮溃疡”指组织的表面丧失导致的皮肤损害，通常伴有炎症。可以通过本发明方法治疗的真皮溃疡包括褥疮溃疡、糖尿病溃疡、静脉郁滞溃疡和动脉溃疡。褥疮创伤指长期施向皮肤区域的压力所致的慢性溃疡。这种创伤通常称为褥疮（bedsore 或 pressure sore）。静脉郁滞溃疡由缺陷静脉得到血液或其它流体的停滞导致。动脉溃疡指血液流动差的动脉周围区域中的坏死皮肤。

“牙组织”指口腔中的类似上皮组织的组织，例如牙龈组织。本发明的方法可用于治疗牙周病。

“内部上皮组织”指机体内侧的组织，其具有与皮肤的表皮层类似的特征。例子包括肠道内层。本发明方法可用于促进某些内伤

如手术创伤的愈合。

“眼组织的创伤”指严重的干眼综合征，角膜溃疡和擦伤及眼科手术创伤。

在本申请中，术语“增殖性皮肤失调”指以皮组织的不希望或异常增殖为标志的任何皮肤疾病/失调。典型地这些状态的特征在于表皮细胞增殖或不完全细胞分化，并包括，例如，X-连锁鱼鳞癣、银屑病、特应性皮炎、过敏性接触性皮炎、表皮溶解性角化过度症和脂溢性皮炎。例如，表皮发育不良是表皮发育不全的一种形式。另一个例子是“表皮溶解”，其指表皮的一种松解状态，自发或在外伤部位形成小泡或皱褶。

术语“癌”指由趋于渗透周围组织并产生转移的上皮细胞组成的恶性新的生长。癌的例子包括：“基细胞癌”，其是一种皮肤的上皮肿瘤，其尽管很少转移，但是具有局部侵袭和破坏的潜力；“鳞状细胞癌”，其指从鱼鳞状上皮产生的癌并具有立方形细胞；“癌肉瘤”，其包括由癌性和肉瘤性组织组成的恶性肿瘤；“腺囊性癌”，其是以由小上皮细胞巢或索分离或环绕的透明或粘质基质柱或条为标志的癌，其在乳腺和唾液腺以及呼吸道的粘液腺中发生；“表皮样癌”，其指趋于以与表皮相同方式分化的癌细胞，即它们易于形成棘细胞并发生角化；“鼻咽癌”，其指在鼻后空间的被覆上皮中产生的恶性肿瘤；“肾细胞癌”，其涉及由各种排列的管状细胞组成的肾实质的癌症。另一种癌性上皮生长是“乳头瘤”，其是从上皮衍生的良性肿瘤，乳头瘤病毒是其致病因子；以及“表皮样瘤”，其指在神经沟闭合时包含了外胚层元件而形成的脑肿瘤或脑膜肿瘤。

如本文所用，术语“银屑病”指改变皮肤的调控机制的过度增殖性皮肤失调。特别是，形成的损害涉及表皮增殖的初级和二级改变，皮肤的炎性应答以及调控分子如淋巴因子和炎性因子的表达。

银屑病皮肤的形态学特征是表皮细胞更新加快、表皮增厚、异常角质化、炎性细胞渗透进真皮层以及多形核白细胞渗透进表皮层导致基细胞循环增加。另外，存在角化过度细胞和角化不全细胞。

术语“角化病”指以表皮的角质层的增生为特征的增殖性皮肤病失调。举例性的角化病包括滤泡角化病（keratosis follicularis）、手掌脚底角化病、咽角化病、毛发角化病和光化性角化病。

如本文所用，“增殖”指进行有丝分裂的细胞。

如本文所用，“转化细胞”指自发转变成不受限制地生长状态的细胞，即细胞获得了在培养中无限次分裂而生长的能力。针对其丧失了生长控制，转化细胞可用如致瘤性、退行性和/或增生性来表征。

如本文所用，“无限增殖化细胞”指经化学和/或重组手段改变从而具有在培养中经无限次分裂而生长的能力的细胞。

用本发明方法治疗的“患者”或“个体”可以是人或非人动物。

术语“美容制剂”指配制用于局部施用的药物制剂形式。

针对本发明治疗方法，例如 hedgehog 治疗剂的“有效量”指制剂中例如 hedgehog 多肽的量，当作为所需的剂量方案的一部分施用时应能改变细胞的增殖速率和/或细胞的分化状态从而根据待治疗的失调的临床可接受的标准或美容目的产生一定量的上皮细胞增殖。

细胞的“生长状态”指细胞的增殖速率和细胞的分化状态。

“同源性”和“相同性”均指两个多肽序列间的序列相似性，相同性是更严格的比较。同源性和相同性可通过比较每个序列中的可用于对比的位置而确定。当比较序列中的位置由相同氨基酸残基占据时，则该两个多肽可以被称为在该位置相同；当等价位点由相同氨基酸（例如，相同）或相似氨基酸（例如，空间和/或电子性质相似）占据时，则该两个分子可被称为在该位置是同源的。序列间的同源性或相同性百分比是序列间共享的匹配或同源位置数的函

数。“不相关”或“非同源”序列与本发明的 AR 序列共享少于 40% 的相同性，优选少于 25% 相同性。

术语“相应于”，当指特定的多肽或核酸序列时，意味着感兴趣的序列与据称其与之相应的参照序列相同或同源。

术语“重组蛋白质”、“异源蛋白质”和“外源蛋白质”在说明书中互换使用，指由重组 DNA 技术产生的多肽，其中通常编码该多肽的 DNA 被插入合适的表达构建体中，该表达构建体被用于转化宿主细胞以产生异源蛋白质。即，多肽是从异源核酸表达的。

“嵌合蛋白”或“融合蛋白”是编码 *hedgehog* 多肽的第一个氨基酸序列与限定异于 *hh* 蛋白的任何结构域或基本上与之不同源的一个结构域的第二个氨基酸序列的融合体。嵌合蛋白可存在一外源结构域，该结构域在也表达第一个蛋白质的生物体中发现（但在不同的蛋白质中），或者其可以是由不同种生物体表达的蛋白质结构的“种间”、“基因间”等融合体。通常，融合蛋白可由通式 $(X)_n-(hh)_m-(Y)_n$ 表示，其中 *hh* 代表 *hedgehog* 代表的全部或部分，X 和 Y 各自独立地代表不是天然发现与 *hedgehog* 序列相邻的多肽链的氨基酸序列，m 是大于或等于 1 的整数，n 独立地代表 0 或大于或等于 1 的整数（n 和 m 优选地不大于 5 或 10）。

III. 方法和组合物的举例应用

本发明方法对于影响上皮组织的失调的治疗和预防以及美容应用具有广泛的实用性。概括地，本发明方法的特征在于包括给予动物有效改变所处理的上皮组织的增殖状态量的 *ptc* 或 *hedgehog* 治疗剂的步骤。给药方式和剂量方案将根据待处理的上皮组织而变化。例如，当所处理的组织是表皮组织如真皮或粘膜组织时，优选局部配制品。类似地，如下所详述的，特定 *ptc* 或 *hedgehog* 治疗剂的应用，例如兴奋剂或拮抗剂，将取决于所处理的组织的细胞的增殖是

所需的还是需防止的。

“促进创伤愈合”的方法由于治疗而导致比不治疗的类似创伤愈合更快的创伤愈合。“促进创伤愈合”也可指该方法导致特别是角质细胞的增殖和生长，或者创伤愈合伴有较小疤痕、较少创伤收缩、较少胶原沉积和较大表面积。在某些情况下，当与本发明的方法一起应用时，“促进创伤愈合”还可指创伤愈合的某些方法具有改进的成功率（例如，皮肤移植的愈合率）。

并发症是未完全愈合并仍开放的创伤的一经常的危险。尽管大多数创伤不经治疗会很快愈合，但是某些类型的创伤拒绝愈合。覆盖大面积的创伤也会长时间保持开放。在本发明的一个实施方案中，本发明方法可用于加速涉及上皮组织的如由于手术、烧伤、炎症或刺激产生的创伤的愈合。本发明的某些 hedgehog 和 ptc 治疗配方（例如，增殖形式）也可预防性施用，如以美容制剂形式施用以促进例如皮肤、毛发和/或指甲的组织再生过程。

尽管重建手术技术取得了显著的进展，但是疤痕仍然是愈合皮肤重新获得正常功能和表观的重要障碍。当病理性疤痕如手或面部的疤痕疙瘩或肥大疤痕导致功能障碍或生理变形时尤其如此。在最严重的情况下，这种疤痕可使人陷入心理忧伤和经济困境。创伤修复包括止血、炎症、增殖和重建阶段，增殖阶段涉及成纤维细胞、内皮细胞和上皮细胞的繁殖。通过使用本发明方法，创伤内和接近创伤的上皮细胞的增殖速率可被控制以加速创伤的闭合和/或使疤痕组织形成最小化。

全部和部分深度烧伤是经常覆盖大面积的创伤类型的例子并因此需要很长时间来愈合。结果是，经常发生危及生命的并发症如感染和丧失体液。另外，烧伤愈合通常是无序的，导致疤痕和破相。在某些情况下，由过量胶原沉积所致的创伤收缩导致创伤附近的肌

肉活动性降低。本发明的组合物和方法可被用于加速烧伤愈合速率并促进愈合过程，导致更希望的美容效果并降低创伤收缩，减少疤痕。

大面积严重烧伤经常通过从患者身体的未损坏部位进行皮肤自体移植而治疗。本发明方法也可以与皮肤移植一起使用以通过加速移植皮肤和接近移植物的患者皮肤的生长而改善移植的“愈合”率。

皮肤溃疡是适用于经本发明方法例如使溃疡愈合和/或防止溃疡成为慢性创伤而治疗的另一个创伤例子。例如，患有糖尿病的个体中 7 个中有 1 个会在其肢体发生皮肤溃疡，其很易于感染。患有感染性糖尿病性溃疡的个体经常需要住院、护理、昂贵的抗生素，并且在某些情况下需截肢。皮肤溃疡如由于静脉疾病（静脉郁滞性溃疡）、过度压迫（褥疮溃疡）和动脉溃疡所致皮肤溃疡也拒绝愈合。现有技术的治疗通常限制在保护创伤、避免感染以及在某些情况下经血管手术重建血流。根据本发明方法，皮肤受损区可通过包括 *hedgehog* 或 *ptc* 治疗剂的治疗方法，这种治疗剂促进创伤上皮化，例如加速皮肤溃疡的愈合速率。

本发明治疗法还可以作为治疗方案的一部分以有效治疗口腔溃疡和非口腔（paraoral）溃疡，例如由放疗和/或化疗导致的溃疡。这种溃疡通常在化疗或放疗后几天内发生，并且开始时通常是小的疼痛的无规则形状损害，通常被易破的灰白坏死膜覆盖并并炎症组织环绕。在许多情况下，缺少治疗会导致损害周围的组织呈炎症增殖。例如，溃疡周围的上皮通常具有增殖活性，导致表面上皮的完整性丧失。这些损害由于其大小和表皮完整性丧失使机体易于受潜在的继发感染。食物和水的常规摄入变成一非常疼痛的过程，并且如果溃疡增殖至消化道，通常发生痢疾，并伴有其所有并发症。根据本发明，包括施用 *hedgehog* 治疗剂的这种溃疡的治疗法可降低受损上

皮的异常增殖和分化，有助于降低后续炎症的严重性。

在另一个举例性实施方案中，本发明方法提供用于治疗或防止胃肠道疾病。简单地说，许多疾病与胃肠道上皮或绒毛的破坏有关，包括化疗和放疗诱导的肠炎（即肠道毒性）和粘膜炎，消化道溃疡病，胃肠炎和结肠炎，绒毛萎缩失调等。例如，在骨髓移植和癌症治疗中使用的化疗药剂和放疗损害造血组织和小肠中的快速增殖细胞，导致严重的和通常为剂量限制性的毒性。小肠粘膜屏障的损害导致严重并发症即出血和脓毒症。本发明方法可用于促进胃肠道上皮的增殖，从而增加放疗和化疗剂的耐受剂量。胃肠道疾病的有效治疗可以通过几种标准确定，包括肠炎分值，以及现有技术中已知的其它测试法。

本发明方法和组合物还可用于治疗皮肤病产生的创伤，如自身免疫疾病如银屑病产生的损害。特应性皮炎指与由过敏原如花粉、食物、头皮屑、昆虫毒液和植物毒素引起的免疫应答相关的过敏产生的皮肤损伤。

随着年龄的增长，表皮变薄，皮肤附件萎缩。毛发稀疏，皮脂分泌降低，结果易于干燥、皸裂和开裂。随着丧失弹性和胶原纤维，真皮变薄。另外，角质细胞增殖（其是皮肤厚度和皮肤增殖能力的指标）随年龄增加而降低。据信角质细胞增殖的增加能抵销皮肤老化即皱纹、厚度、弹性和修复。根据本发明，可治疗或美容应用 *hedgehog* 或 *ptc* 治疗剂的增殖形式以至少在一段时间抵销皮肤的老化。

本发明方法还可用于治疗眼组织创伤。通常，角膜组织的损伤，无论是疾病、手术或受伤，根据创伤的性质，其可影响上皮和/或内皮细胞。角膜上皮细胞是位于角膜外表面的非角质化上皮细胞并提供抗外界环境的保护屏障。角膜创伤愈合受到临床医生和研究人员

的关注。眼科医生经常面对角膜营养不良和疑难损伤，其导致持久复发性上皮糜烂，经常导致永久性内皮丧失。在这些情况下可应用本发明的 hedgehog 和/或其它 ptc 治疗剂的增殖形式以促进受损角膜组织的上皮化。

为进一步说明，与眼部的上皮细胞损害典型相关并且可用本发明方法有效治疗的具体疾病包括持续性角膜上皮缺陷，复发性糜烂，神经营养性角膜溃疡，干燥性角结膜炎，微生物性角膜溃疡，病毒性角膜溃疡等。典型地导致对上皮细胞层的损伤的手术程序包括在眼表面进行的激光程序，任何屈光矫正手术程序如放射状角膜切开术和散光角膜切开术，结膜瓣，结膜移植，表层角膜成形术（epikeratoplasty）和角膜刮除术。另外，浅层创伤如刮伤、表面糜烂、炎症等可导致上皮细胞的丧失。根据本发明，将角膜上皮与有效导致角膜上皮细胞增殖量的 ptc 或 hedgehog 治疗剂接触以合适地愈合创伤。

在其它实施方案中，hedgehog 或 ptc 治疗剂的抗增殖制剂可用于抑制晶状体上皮细胞增殖以防止白内障囊外摘除术后并发症。白内障是难治的眼疾，目前已进行了治疗白内障的各种研究。但是目前对白内障的治疗通过外科手术进行。白内障手术已应用了很长时间并且已检查了各种手术方法。囊外晶状体摘除术是去除白内障的一种方法。这一技术相对于囊内摘除术的主要医学优点是无晶状体性囊样黄斑水肿和视网膜脱离的发生率较低。囊外摘除术也是后房型人工晶状体移植所需的，该后房型人工晶状体目前被认为是大多数病例中所需的晶状体。

但是，白内障囊外摘除术的一个缺点是晶状体后囊混浊一通常称为继发性白内障一的发生率高，在手术后 3 年内发生率高达 50%。继发性白内障是由于囊外晶状体摘除后仍保留的赤道部和前囊晶状

体上皮细胞的增殖所致。这些细胞的增殖导致 Sommerling 环，其也与沉积和发生在后囊上的成纤维细胞一起导致后囊混浊，由此影响视力。防止继发性白内障是治疗所优选的。为防止白内障的二次形成，本发明方法提供了抑制保留的晶状体上皮细胞增殖的手段。例如，通过在晶状体摘除后向眼的前房注入含有抗增殖性 hedgehog 或 ptc 治疗剂制剂可诱导这些细胞保持静止。另外，可使该溶液等渗平衡以在注入眼的前房时提供最小的有效剂量，从而以一定的特异性抑制囊下上皮的生长。

本发明方法还可用于治疗以角膜上皮细胞增殖为标志的角膜病，例如治疗眼球上皮失调如眼球表面的上皮向下生长或鳞状细胞癌。

来自体内的组织和器官的保持也是非常希望的。在治疗人类疾病中已发展成熟组织置换疗法。例如，在 1996 年美国进行了超过 4 万例角膜移植。人表皮细胞可体外生长并用于植入烧伤部位和慢性皮肤溃疡和其它皮肤创伤。本发明方法可用于加速体外上皮组织的生长，以及加速培养的上皮组织向动物宿主的移植。

本发明方法可用于改善皮肤移植物的“愈合速率”。如果移植物的愈合速率很长，则表皮组织的移植物可起泡和剪切，降低自体移植物“愈合”，即粘附于创伤并形成下面为颗粒组织的基底膜的几率。通过诱导角质细胞的增殖本发明方法可增加愈合速率。增加愈合速率的方法包括将皮肤自体移植物与创伤愈合有效量的 hedgehog 或 ptc 治疗剂组合物接触，这种组合物如上所述在促进创伤愈合的方法和促进角质细胞的生长和增殖的方法中有描述。

皮肤等同物具有许多用途，不仅可作为人或动物皮肤的置换物用于皮肤移植，也可用作测试皮肤以确定药物物质和美容品在皮肤上的效果。药理学、化学和美容测试的一个主要困难是很难确定产

品在皮肤上的效果和安全性。本发明的皮肤等同物的一个优点在于它们可作为体外测试这些物质在测试皮肤上产生的效果的指示。

因此，在一个实施方案中，本发明方法可用作例如裸露区、肉芽创伤和烧伤的皮肤移植方案的一部分。增殖性 hedgehog 和/或 ptc 治疗剂的应用可促进移植技术如分裂厚度自体移植和表皮自体移植（培养的同源角质细胞）和表皮同种异体移植（培养的异源角质细胞）。在同种异体移植情况下，应用本发明方法以促进培养物中形成皮肤等同物有助于提供/保持这种移植物的稳定供应（例如，在组织库中），从而患者可在一个程序中用能产生永久愈合的材料覆盖。

为此，本发明还涉及复合活皮肤等同物，其包括培养的角质细胞的表皮层，该角质细胞已用 hedgehog 或其它 ptc 治疗剂处理而扩增。本发明方法可用作制备复合活皮肤等同物的方法的一部分。在一示意性实施方案中，这种方法包括获得一皮肤样品，酶处理该皮肤样品以将表皮与真皮分离，酶处理表皮以释放角质细胞，在 hedgehog 或 ptc 治疗剂的存在下培养表皮角质细胞直至铺满，同时，或单独地，酶处理真皮以释放成纤维细胞，培养该成纤维细胞直至亚铺满，用培养的成纤维细胞接种多孔交联胶原海绵膜，在其表面培养该接种的胶原海绵以使成纤维细胞生长贯通胶原海绵，然后用培养的角质细胞接种该胶原海绵，进一步在 hedgehog 或 ptc 治疗剂的存在下培养该复合皮肤等同物复合物以促进细胞的生长。

在其它实施方案中，可在培养物中分离含有上皮和间充质层的皮肤片并用补加增殖形式的 hedgehog 或 ptc 治疗剂的培养基扩增。

适于细胞培养技术的任何皮肤样品均可用于本发明。该皮肤样品可以是同源或异源的。

在本发明的另一方面，本发明方法可与各种希望控制牙周组织中或牙周组织周围的上皮细胞增殖的牙周程序结合应用。

在一个实施方案中，增殖形式的 hedgehog 和 ptc 治疗剂可用于促进天然牙和假牙周围的再上皮化，例如，促进牙龈组织的形成。

在另一实施方案中，抗增殖性 ptc 治疗剂可用于治疗牙周疾病。估计仅在美国就有超过 1 亿 2 千 5 百万成人患有各种形式的牙周疾病。牙周疾病的起因是由于位于牙龈与牙齿附着区域中的特殊细菌导致的炎症损害。通常最先发生的是下面的结缔组织的血管变化，结缔组织的炎症刺激牙沟的上皮层和上皮附件中的下列变化：在基底上皮层中的有丝分裂活性增加；增加产生脱落的角蛋白；牙齿表面相邻的细胞脱落易于加深缺口；牙沟底部的基层的上皮细胞和附件区的上皮细胞增殖进入结缔组织，且牙龈纤维开始发生断裂，其中结缔组织的溶解导致形成开放的损害。施用 hedgehog 制剂至牙周可用于一种上皮组织的增殖，因而进一步防止牙周溃坏的发展。

在另一方面，本发明方法可用于辅助控制定向组织再生，如当与可生物再吸收的材料一起使用时。例如，通过本方法可促进牙周移植物如假牙的掺入。牙齿的再附着涉及结缔组织纤维的形成和牙袋的再上皮化。本发明治疗方法可通过在创伤愈合的早期阶段控制基底上皮细胞的有丝分裂功能而用于加速组织再附着。

本发明的另一方面涉及用 hedgehog 治疗剂制剂控制毛发生长。毛发基本上由一种粗糙不溶的蛋白质角蛋白组成；其主要强度取决于其胱氨酸二硫键。单根毛发包括柱状毛干和毛根，并包含在毛囊中，毛囊是皮肤中的瓶状凹坑。毛囊底部含有指状突出物，称为乳头，乳头由长出毛发的结缔组织组成，通过乳头血管提供营养给细胞。毛干是延伸出皮肤表面的部分，而毛根则被描述为毛发的埋藏部分。毛根的基部延伸进位于乳头上的毛球中。产生毛发的细胞在毛囊的毛球中生长，随着细胞在毛囊中增殖，它们以纤维形式突出。毛发“生长”指通过分裂细胞所致的毛发纤维的形成和伸长。

如本领域所熟知的，通常的毛发循环分成 3 个阶段：生长期、退化期和末期。在活性期（生长期），真皮乳头的表皮干细胞快速分裂，子代细胞向上移动并分化形成毛发本身的同心层。过渡期即退化期的标志是毛囊中的干细胞停止有丝分裂。静止期又称为末期，在此期间毛发在头皮上保留几个星期，随后在其之下发育出现的新毛发将末期毛干从其毛囊中驱除出去。根据这一模式，很明显的是分化成毛发细胞的分裂干细胞群越大，毛发生长越多。因此，通过分别增强或抑制这些干细胞的增殖可进行增加或降低毛发生长的方法。

在一个实施方案中，本发明方法提供了改变毛发生长周期动力学以诱导毛囊细胞特别是毛囊的干细胞的增殖的手段。本发明的组合物和方法可用于例如通过促进毛囊干细胞的增殖而在温血动物如人中增加毛囊大小以及毛发生长速率。在一个实施方案中，所述方法包括向希望生长毛发的皮肤区域施用足以在动物中增加毛囊大小和/或毛发生长速率的量的 hedgehog 或 ptc 治疗剂。典型地，组合物以霜剂局部施用，并且每天施用直至观察到长出毛发，之后继续施用一段时间以足以保持所需量的毛发生长。这一方法可用于促进新发生长或刺激毛发生长的速率，例如在化疗后或用于治疗各种形式的秃头症，例如雄性模式秃头。例如，患有雄激素性秃头症的个体在生长期或退化期发生的几种生物化学细胞学和分子学失调中的一种可通过，例如在干细胞起作用或形成中，它们的迁移过程或在毛囊乳头和基质内产生角蛋白的有丝分裂期过程中，用本发明方法治疗而矫正或改善。

在其它实施方案中，相对于传统的剪、剃或拔去毛方法，可采用某些 hedgehog 和 ptc 治疗剂（例如抗增殖形式）作为降低人毛发生长的方式。例如，本发明方法可用于治疗以异常快速或浓密毛发生长的方式。

生长为特征的毛发病，例如多毛症。在一个举例性实施方案中，hedgehog 拮抗剂可用于处理多毛症，其是以异常多毛为标志的疾病。本发明方法还可提供延长脱毛时间的方法。

另外，由于 hedgehog 拮抗剂（或 ptc 兴奋剂）通常对于上皮细胞有细胞抑制性而非细胞毒性，因此这种制剂可用于保护毛囊细胞抗细胞毒性剂。细胞毒性剂需进入细胞周期的 S 一期才起作用，例如放射诱导的死亡。用本发明方法治疗可通过使毛囊细胞变为静止而提供保护，例如，通过抑制细胞不进入 S 期，从而防止毛囊细胞进行有丝分裂突变或程序性细胞死亡。例如，进行化疗或放疗的通常导致脱发的患者可使用 hedgehog 拮抗剂。通过在这类疗法中抑制细胞周期进程，本发明治疗法可保护毛囊细胞不死亡，否则会由于细胞死亡程序的激活而死亡。化疗或放疗结束后，可除去 hedgehog 或 ptc 疗法以解除对毛囊细胞增殖的抑制。

本发明方法还可用于治疗毛囊炎，如脱发性毛囊炎、网状萎缩性红斑性毛囊炎或疤痕疙瘩性毛囊炎。例如，在治疗假毛囊炎时可局部施用 hedgehog 治疗剂的一种美容制剂，假毛囊炎是一种通常在颈部的下颌下区发生并与剃须有关的慢性病，其特征性损害为红斑性丘疹和含有埋藏的毛发的脓疱。

在本发明的另一方面，本发明的 hedgehog 和 ptc 治疗剂的抗增殖形式可用于诱导上皮衍生组织的分化。这种形式的这些分子可提供用于治疗涉及上皮组织的增生和/或瘤形成的分化疗法的基础。例如，这种制剂可用于治疗皮肤细胞异常增殖或生长的皮肤病。

例如，本发明的药物制剂可用于治疗增生性表皮症状如角化病，以及用于治疗瘤形成表皮症状，如以高增殖速率为特征的症状，如各种皮肤癌，例如基层细胞癌或鳞状细胞癌。本发明方法还可用于治疗感染皮肤的自身免疫疾病，特别是涉及表皮病态增殖和/或角

质化的皮肤病，如由银屑病或特应性皮炎引起的皮肤病。

许多常见皮肤病如银屑病、鳞状细胞癌、角化棘皮瘤和光化性角化病的特征在于局部异常增殖和生长。例如，在银屑病中，该病的特征是皮肤上有鳞屑性红色凸出的斑，已知角质细胞的增殖比正常快许多，且分化不完全。

在一个实施方案中，本发明的制剂适于治疗与导致皮肤细胞的异常增殖的角化病连锁的皮肤病，所述角化病可有炎性组分或非炎性组分。举例来说，例如促进静止或分化的 ptc 兴奋剂的治疗制剂可用于治疗各种形式的银屑病，包括皮肤、粘膜或是指（趾）甲的银屑病。如上所述，银屑病的典型特征在于表皮角质细胞，其随“再生”途径显示显著的增殖活化和分化。用本发明的抗增殖性实施方案的治疗可用于逆转病理性表皮活化并可提供持续减轻该病的基础。

各种其它角化病损害也是用本发明抗增殖性制剂治疗的候选者。光化性角化病，例如，是在日晒和受辐射的皮肤上产生的浅层炎性恶化前肿瘤，该损害是红斑状至棕色，伴有可变的鳞屑。目前的疗法包括切除术和冷冻破坏法。但这些治疗是疼痛的，且经常产生美容上不能接受的疤痕。因此，角化病如光化性角化病的治疗可包括优选局部应用足以抑制损害的表皮/表皮样细胞的过度增殖的量的 ptc 兴奋剂组合物。

痤疮是另一种可用本发明方法的抗增殖性实施方案治疗的皮肤病。例如寻常痤疮是通常在十几岁少年和年轻人中发生的多因子疾病，其特征是在面部和上躯干上的炎性和非炎性损害表现。导致寻常痤疮的基本缺陷是活动过强的皮脂腺导管的角化过度，角化过度阻碍了皮肤和毛囊微生物的正常迁移，由此刺激痤疮丙酸杆菌和表皮葡萄球菌以及卵瓶形酵母释放脂酶。用 hedgehog 或 ptc 治疗剂的

抗增殖形式特别是局部制剂治疗可防止导管的暂时特征，例如导致损害形成的角化过度。本发明治疗方法还可进一步包括例如抗生素、类维生素 A 和抗雄激素。

本发明还提供了治疗各种皮炎的方法，皮炎是指不易划界的损害的一种描述性术语，这种损害是瘙痒的、红斑状、起鳞屑的、起疱的、有渗出的、开裂的或结痂的。这些损害的起因各种各样。最常见的皮炎类型是特应性、接触性和尿布皮炎。例如，皮脂溢性皮炎是在各种区域通常是头皮上的慢性通常瘙痒的皮炎，具有红斑，干、湿或滑腻的鳞屑以及黄色结痂块，过量干鳞屑表皮脱落的是停滞性皮炎，其是一种通常为慢性的经常是湿疹性皮炎。光化性皮炎是由于暴露于光化性辐射如日光、紫外光或 X 射线或 γ 射线导致的皮炎。根据本发明，本发明的 hedgehog 或 ptc 治疗剂制剂可用于治疗和/或防止由上皮细胞的不希望的增殖导致的皮炎的某些症状。各种皮炎的这种疗法还可包括局部和系统性应用皮质类固醇、止痒药和抗生素。

可用本发明方法治疗的疾病还包括非人类特异的疾病如兽疥癣。

IV. 举例性的 hedgehog 治疗剂化合物

本发明方法的 hedgehog 治疗剂组合物可通过各种技术产生，包括天然蛋白质的纯化，重组产生的蛋白质和合成化学。hedgehog 治疗剂的多肽形式优选地从脊椎动物 hedgehog 蛋白衍生，例如具有相应于来自脊椎动物的天然 hedgehog 蛋白质或其片段的序列。但是，应理解的是 hedgehog 多肽可相应于在任何后生动物中发生的 hedgehog 蛋白（或其片段）。

可衍生本发明的治疗剂的各种天然 hedgehog 蛋白的特征是一信号肽、一高度保守的 N-末端区以及一较不同的 C-末端结构域。除

了在分泌途径信号序列裂解之外 (Lee, J.J.等(1992)细胞 71: 33—50; Tabata, T.等(1992)基因发育 2635—2645; Chang, D.E.等(1994)发育 120: 3339—3353), *hedgehog* 前体蛋白天然发生基于 C—末端部分的保守序列的内部自我蛋白酶解 (Lee 等(1994)科学 266: 1528—1537; Porter 等(1995)自然 374: 363—366)。这一自我裂解导致产生一 19kD N-末端肽和一 26—28kD C-末端肽 (Lee 等(1992)文献同上; Tabata 等(1992)文献同上; Chang 等(1994)文献同上; Lee 等(1994)文献同上; Bumcrot, D.A.等(1995)分子细胞生物学 15: 2294—2303; Porter 等(1995)文献同上; Ekker, S.C.等(1995)当前生物学 5: 944—955; Lai, C.J.等(1995)发育 121: 2349—2360)。N-末端肽保持与合成其的细胞表面紧密相连, 而 C-末端肽可在体外和体内自由扩散 (Lee 等(1994)文献同上; Bumcrot, D.A.等(1995)文献同上; Mart', E.等(1995)发育 121: 2537—2547; Roelink, H.等(1995)细胞 81: 445—455)。N-末端肽在细胞表面的停留取决于自我裂解, 因为由在内部裂解的正常位置精确终止的 RNA 编码的一种截短形式的 *hedgehog* 在体外 (Porter 等(1995)文献同上) 和体内 (Porter, J.A.等(1996)细胞 86, 21—34) 是可扩散的。生物化学研究表明 *hedgehog* 前体蛋白的自我酶解经一内部硫酯中间体进行, 该中间体随后在一亲核取代中被裂解。提示亲核物质是一小的亲脂性分子, 更具体地是胆固醇, 其与 N-肽的 C-末端共价结合 (Porter 等(1996)文献同上), 使该肽与细胞表面粘连。

hedgehog 基因的脊椎动物家族包括至少 4 个成员, 例如单果蝇 *hedgehog* 基因 (SEQ ID NO: 19) 的平行进化同源基因。这些成员中的三个, 本文称为 Desert *hedgehog* (*Dhh*), Sonic *hedgehog* (*Shh*) 和 Indian *hedgehog* (*Ihh*), 明显存在于所有脊椎动物中, 包括鱼类、鸟类和哺乳类中。第四个成员, 本文称为 *tiggie-winkle hedgehog* (*Thh*), 看起来是特异于鱼类的。根据所附序列列表 (也见表 1), 鸡

的 *Shh* 多肽由 SEQ ID NO: 1 编码；小鼠的 *Dhh* 多肽由 SEQ ID NO: 2 编码；小鼠的 *Ihh* 多肽由 SEQ ID NO: 3 编码；小鼠的 *Shh* 多肽由 SEQ ID NO: 4 编码；斑马鱼的 *Shh* 多肽由 SEQ ID NO: 5 编码；人的 *Shh* 多肽由 SEQ ID NO: 6 编码；人的 *Ihh* 多肽由 SEQ ID NO: 7 编码；人的 *Dhh* 多肽由 SEQ ID NO: 8 编码；斑马鱼的 *Thh* 多肽由 SEQ ID NO: 9 编码。

表 1
序列表中 *hedgehog* 序列指南

	核苷酸	氨基酸
鸡 <i>Shh</i>	SEQ ID NO: 1	SEQ ID NO: 10
小鼠 <i>Dhh</i>	SEQ ID NO: 2	SEQ ID NO: 11
小鼠 <i>Ihh</i>	SEQ ID NO: 3	SEQ ID NO: 12
小鼠 <i>Shh</i>	SEQ ID NO: 4	SEQ ID NO: 13
斑马鱼 <i>Shh</i>	SEQ ID NO: 5	SEQ ID NO: 14
人 <i>Shh</i>	SEQ ID NO: 6	SEQ ID NO: 15
人 <i>Ihh</i>	SEQ ID NO: 7	SEQ ID NO: 16
人 <i>Dhh</i>	SEQ ID NO: 8	SEQ ID NO: 17
斑马鱼 <i>Thh</i>	SEQ ID NO: 9	SEQ ID NO: 18
果蝇 <i>HH</i>	SEQ ID NO: 19	SEQ ID NO: 20

除了各种 *hedgehog* 同系物之间有序列变异外，明显地 *hedgehog* 蛋白也以各种不同形式天然存在，包括蛋白原形式、全长成熟形式和其几种加工片段。蛋白原形式包括用于胞外结构域的定向分泌的 N-末端信号肽，而全长成熟形式缺少这一信号序列。

如上所述，在某些情况下发生成熟形式的进一步加工以产生该蛋白的生物学活性片段。例如，*sonic hedgehog* 进行进一步蛋白酶解

加工产生约 19kDa 和 27kDa 的两个肽，19kDa 片段相应于成熟蛋白的蛋白水解的 N-末端部分。

除了蛋白水解片段化之外，脊椎动物 *hedgehog* 蛋白也可被翻译后修饰，如经糖基化和/或添加亲脂性基元，如展伸（stents）、脂肪酸等，但是该蛋白的细菌产生（例如未修饰的）形式也保持某些天然蛋白的生物活性。本发明的 *hedgehog* 多肽的生物活性片段已产生并在例如 PCT 出版物 WO95/18856 和 WO96/17924 中有详细描述。

用于衍生 *hedgehog* 多肽的亲脂性基元范围很广。与 *hedgehog* 多肽相关时术语“亲脂性基团”指具有高碳氢含量从而赋予该基团对脂相的高亲和性的基团。例如，亲脂性基团可以是具有约 7-30 个碳的相对较长链烷基或环烷基（优选正烷基）基团。烷基基团可以以羟基或伯胺“尾”终止。进一步地，亲脂性分子包括天然存在的和合成的芳族和非芳族基元，如脂肪酸、甾醇、酯和醇、其它脂类分子、笼形结构如金刚烷和巴基明斯特富勒烯（buckminsterfullerenes）、以及芳烃如苯、萘、菲、蒽、蒘、蒹和并四苯。

在一个实施方案中，*hedgehog* 多肽被一或多个甾醇基元如胆固醇修饰，例如参见 PCT 出版物 WO96/17924。在某些实施方案中，如果 *hedgehog* 多肽相应于天然存在的 N-末端蛋白水解片段，胆固醇优选加到 C-末端甘氨酸上。

在另一实施方案中，*hedgehog* 多肽可被脂肪酸基元如肉豆蔻酰、棕榈酰、硬脂酰或花生四烯酰基元修饰，例如参见 Pepinsky 等（1998）生物化学杂志 273: 14037。

除了将胆固醇加到该蛋白胞外片段的 C-末端或将脂肪酸加到 N-末端观察的效果之外，经用亲脂性基元在蛋白的其它位点衍生该蛋白和/或用除胆固醇或脂肪酸之外的基元衍生该蛋白，*hedgehog* 基

因产物的至少某些生物学活性令人意想不到地得以增强。本发明的某些方面涉及在该蛋白的天然加工形式的 N-末端或 C-末端外的其它位点被修饰，和/或末端残基被其它亲脂性基元、即除了 C-末端用甾醇或 N-末端用脂肪酸修饰之外修饰的 hedgehog 多肽制剂的应用。

特别有用的亲脂性分子是脂环烃、饱和和不饱和脂肪酸和其它脂类和磷脂基元、蜡、胆固醇、类异戊二烯、萜和聚脂环烃，包括金刚烷和 buckminsterfullerenes、维生素、聚乙二醇或寡聚乙二醇、(C1-C18) 烷基磷酸二酯、-O-CH₂-CH(OH)-O-(C12-C18)-烷基，特别是与萜衍生物的共轭物。亲脂性基元可以是适用于本发明的亲脂性染料，包括但不限于，二苯基己三烯、尼罗红、N-苯基-1-萘胺、Prodan、Laurodan、芘、芘、若丹明、若丹明 B、四甲基若丹明、得克萨斯红、硫代若丹明 (sulforhodamine)、1,1'-双十二烷基-3,3',3'-四甲基靛羧花青过氯酸盐，十八烷基若丹明 B 和购自分子探针公司的 BODIPY 染料。

其它亲脂性基元的例子包括脂族羰基基团，包括 1-或 2-金刚烷基乙酰基，3-甲基金刚烷-1-基乙酰基，3-甲基-3-溴-1-金刚烷基乙酰基，1-萘烷基乙酰基，樟脑乙酰基，苡烷乙酰基，降金刚烷基乙酰基，降冰片烷乙酰基，二环[2.2.2]-辛-5-烯乙酰基，1-甲氧基二环[2.2.2]-辛-5-烯-2-羰基，顺-5-降冰片烯-内-2,3-二羰基，5-降冰片烯-2-基乙酰基，(1R)-(-)-myrtenane 乙酰基，2-降冰片烷乙酰基，抗-3-氧-三环[2.2.1.0<2,6>]-庚烷-7-羰基，癸酰基，十二酰基，十二碳烯酰基，十四碳二烯酰基，癸炔酰基或十二碳炔酰基。

hedgehog 多肽可与疏水性基元以各种方式连接，包括化学偶联手段或基因工程。

本领域技术人员已知有许多化学交联剂，本发明优选的交联剂是异双官能交联剂，其可用于以逐步方式连接 hedgehog 多肽和疏水性基元。异双官能交联剂提供了设计更特异的缀合蛋白质的偶联方法的能力，从而降低不希望的反源蛋白聚合物的发生。现有技术中已知各种异双官能交联剂，它们包括：琥珀酰亚胺基-4-(N-马来酰亚胺基甲基)环己烷-1-羧基酯 (SMCC)，m-马来酰亚胺基苯甲酰-N-羟基琥珀酰亚胺酯 (MBS)；N-琥珀酰亚胺基(4-碘乙酰基)氨基苯甲酸 (SIAB)，琥珀酰亚胺基-4-(p-马来酰亚胺基苯基)丁酸酯 (SMPB)，1-乙基-3-(3-二甲基氨基丙基)碳二亚胺盐酸盐 (EDC)；4-琥珀酰亚胺基氧基羰基-a-(2-吡啶基二硫代)-1-甲苯 (SMPT)，N-琥珀酰亚胺基-3-(2-吡啶基二硫代)丙酸酯 (SPDP)，琥珀酰亚胺基-6-[3-(2-吡啶基二硫代)丙酸酯]己酸酯 (LC-SPDP)。具有 N-羟基琥珀酰亚胺基元的交联剂可以 N-羟基磺化琥珀酰亚胺类似物形式获得，其通常具有较高的水溶性。另外，可以合成在连接的链内具有二硫键的交联剂替代烷基衍生物，从而降低体内连接物裂解的量。

除了异双官能交联剂，还存在各种其它交联剂，包括同双官能交联剂和光反应性交联剂。用于本发明的有用的同双官能交联剂的例子包括二琥珀酰亚胺基辛二酸酯 (DSS)，双马来酰亚胺基己烷 (BMH) 和二甲基庚二酰亚胺化物·2HCl (DMP)。用于本发明的有用的光反应性交联剂的例子包括双-[β -(4-叠氮水杨酰胺基)乙基]二硫化物 (BASED) 和 N-琥珀酰亚胺基-6(4'-叠氮-2'-硝基苯基氨基)己酸酯 (SANPAH)。蛋白质偶联技术的最近综述见 Means 等 (1990) 生物缀合物化学法 1: 2-12，引入本文作参考。

包括上述物质的一类特别有用的异双官能交联剂含有伯胺反应基团，N-羟基琥珀酰亚胺 (NHS)，或其水溶性类似物 N-羟基磺

化琥珀酰亚胺（磺化-NHS）。在碱性 pH 下的伯胺（赖氨酸 ϵ 基团）是非质子化的并经亲核攻击在 NHS 或磺化-NHS 酯上反应。这一反应导致形成酰胺键，并释放 NHS 或磺化-NHS 作为副产物。

用作异双官能交联剂一部分的另一反应基团是硫羟反应基团，常见的硫羟反应基团包括马来酰亚胺、卤素和吡啶二硫化物。在微酸至中性(pH6.5-7.5)条件下，马来酰亚胺在几分钟内特异地与游离巯基（半胱氨酸残基）反应。卤素（碘乙酰基官能团）在生理 pH 下与-SH 基团反应。这些反应基团均形成稳定的硫醚键。

异双官能交联剂的第三种组分是间隔臂或桥。桥是连接两个反应末端的结构。桥的最明显贡献是其对立体阻碍的作用。在某些情况下，较长的桥可以更容易地跨越连接两个复杂生物分子所需的距离，例如，SMPB 具有 14.5 埃的跨度。

用异双官能试剂制备蛋白质-蛋白质缀合物是涉及胺反应和巯基反应的两步方法。对于第一步胺反应，所选蛋白质应含有伯胺，这可以是赖氨酸 ϵ 胺或在大多数蛋白质 N-末端发现的 α 伯胺。该蛋白质不应含游离巯基基团。当待缀合的两个蛋白质含有游离巯基时，可修饰一个蛋白质以用例如 N-乙基马来酰亚胺阻断全部巯基（见 Partis 等(1983)蛋白质化学杂志 2:263, 引入本文作参考）。可用 Ellman's 试剂计算特定蛋白质中的巯基量（例如见 Ellman 等(1958)，生物化学生物物理档案 74: 443 和 Riddles 等(1979)分析生物化学 94: 75，引入本文作参考）。

反应缓冲液应不含外来胺和巯基，反应缓冲液的 pH 应为 7.0-7.5，这一 pH 范围防止了马来酰亚胺基团与胺反应，保存马来酰亚胺基团与巯基进行第二步反应。

含有 NHS 酯的交联剂具有有限的水溶性，在将交联剂导入反应混合物之前应将其溶解于少量有机溶剂（DMF 或 DMSO）中。交联

剂/溶剂形成乳液，使反应发生。

磺化-NHS 酯类似物水溶性较高，可以直接加入反应缓冲液中。应避免高离子强度的缓冲液，因为它们有“盐析”出磺化-NHS 酯的倾向。为避免由于水解所致的反应性的丧失，在溶解蛋白质溶液后立即将交联剂加入到反应混合物中。

在浓蛋白质溶液中反应可以更有效。反应混合物的 pH 越是碱性，反应速率越快。NHS 和磺化-NHS 酯的水解速率将随 pH 增加而提高。较高的温度将提高水解和酰化的反应速率。

反应完成后，第一个蛋白质被活化，具有巯基反应基元。经简单的凝胶过滤或透析可从反应混合物中分离活化的蛋白质。为进行第二步交联即巯基反应，选择用于与马来酰亚胺、活化卤素或吡啶基二硫化物反应的亲脂性基团必需含有游离的巯基。或者可修饰伯胺以添加一巯基。

在所有情况下，需将缓冲液脱气以防止巯基基团的氧化。可加入 EDTA 以螯合缓冲液中可能存在的任何氧化金属。缓冲液应不含任何含巯基化合物。马来酰亚胺在稍酸性至中性 pH 范围(6.5-7.5)特异与-SH 基团反应。中性 pH 对涉及卤素和吡啶基二硫化物的反应是足够的。在这些条件下，马来酰亚胺通常在几分钟内与-SH 基团反应，对于卤素和吡啶基二硫化物，需较长反应时间。

在胺反应步骤制备的第一个巯基反应蛋白质在合适的缓冲液条件下与含巯基的亲脂性基团混和。经凝胶过滤或透析等方法可从反应混合物中分离缀合物。

用于缀合的活化亲脂性基元的例子包括：N-(1-苵)马来酰亚胺；2,5-二甲氧基苵-4'-马来酰亚胺，曙红-5-马来酰亚胺；荧光素-5-马来酰亚胺；N-(4-(6-二甲基氨基-2-苯并呋喃基)苯基)马来酰亚胺；二苯酮-4-马来酰亚胺；4-二甲基氨基苯基氮杂苯

基-4'-马来酰亚胺 (DABMI), 四甲基若丹明-5-马来酰亚胺, 四甲基若丹明-6-马来酰亚胺, 若丹明红™ C2 马来酰亚胺, N-(5-氨基戊基) 马来酰亚胺, 三氟乙酸盐, N-(2-氨基乙基) 马来酰亚胺, 三氟乙酸盐, Oregon Green™ 488 马来酰亚胺, N-(2-((2-(((4-叠氮-2,3,5,6-四氟) 苯甲酰) 氨基) 乙基) 二硫代) 乙基) 马来酰亚胺 (TFPAM-SS1), 2-(1-(3-二甲基氨基丙基)-吡啶-3-基)-3-(吡啶-3-基) 马来酰亚胺 (双吡啶基马来酰亚胺; GF 109203X), BODIPY® FL N-(2-氨基乙基) 马来酰亚胺, N-(7-二甲基氨基-4-甲基香豆素-3-基) 马来酰亚胺 (DACM), Alexa™ 488 C5 马来酰亚胺, Alexa™ 594 C5 马来酰亚胺, 钠盐 N-(1-苌) 马来酰亚胺, 2,5-二甲氧基苌-4'-马来酰亚胺, 曙红-5-马来酰亚胺, 荧光素-5-马来酰亚胺, N-(4-(6-二甲基氨基-2-苯并呋喃基) 苯基) 马来酰亚胺, 二苯酮-4-马来酰亚胺; 4-二甲基氨基苯基氮杂苯基-4'-马来酰亚胺, 1-(2-马来酰亚胺基乙基)-4-(5-(4-甲氧基苯基) 噁唑-2-基) 甲磺酸吡啶鎓, 四甲基若丹明-5-马来酰亚胺, 四甲基若丹明-6-马来酰亚胺, 若丹明红™ C2 马来酰亚胺, N-(5-氨基戊基) 马来酰亚胺, N-(2-氨基乙基) 马来酰亚胺, N-(2-((2-(((4-叠氮-2,3,5,6-四氟) 苯甲酰) 氨基) 乙基) 二硫代) 乙基) 马来酰亚胺, 2-(1-(3-二甲基氨基丙基)-吡啶-3-基)-3-(吡啶-3-基) 马来酰亚胺, N-(7-二甲基氨基-4-甲基香豆素-3-基) 马来酰亚胺 (DACM), 11H-苯并[a]芴, 苯并[a]苌。

在一个实施方案中, hedgehog 多肽可用可购自分子探针 (Eugene, Oreg.) 的苌马来酰亚胺例如 N-(1-苌) 马来酰亚胺或 1-苌甲基碘乙酸酯 (PMIA 酯) 衍生化。

对于其中疏水性基元是多肽的实施方案, 本发明的修饰的

hedgehog 多肽可构建成融合蛋白，该融合蛋白含有 hedgehog 多肽和所述疏水性基元作为一个连续的多肽链。

在某些实施方案中，亲脂性基元是两亲性多肽，如爪蟾抗菌肽、杀菌肽、attacin、蜂毒肽、短杆菌肽 S、金黄色葡萄球菌的 α -毒素、丙甲甘肽或合成的两亲性多肽。来自病毒颗粒的融合包膜蛋白也是用于本发明的 hedgehog 蛋白的两亲性序列的方便来源。

另外，例如为了增强治疗或预防效果，或稳定性（例如来自体内的存放时间或抗体内蛋白降解的抗性），可使用诱变技术产生修饰的 *hh* 多肽。这种修饰的肽可通过例如氨基酸取代、缺失或添加二产生。修饰的 *hedgehog* 多肽也可包括相对于天然的 *hedgehog* 蛋白质具有改变的翻译后加工，例如改变的糖基化、胆固醇化、异戊二烯化等，的多肽。

在一个实施方案中，hedgehog 治疗剂是可被在严格条件下与 SEQ ID NO: 1-7 的一或多个序列所代表的 *hedgehog* 编码序列杂交的核苷酸序列编码的多肽。促进 DNA 杂交的合适的严格条件，例如 $6.0 \times$ 氯化钠/柠檬酸钠 (SSC) 在约 45°C ，之后在 50°C 于 $2.0 \times$ SSC 洗涤，是本领域技术人员公知的或可在分子生物学当前方案，John Wiley & Sons, N.Y.(1989), 6.3.1-6.3.6 中找到。例如，洗涤步骤的盐浓度可以选自低严格性的在 50°C 于约 $2.0 \times$ SSC 中洗涤至高严格性的在 50°C 于约 $0.2 \times$ SSC 中洗涤。另外，洗涤步骤的温度可从低严格条件的在室温即约 22°C 提高至高严格条件的在约 65°C 洗涤。

如在文献中所述的，例如来源于其它动物的其它 *hedgehog* 蛋白的基因可用本领域公知的技术从 mRNA 或基因组 DNA 样品中获得。例如，编码 *hedgehog* 蛋白的 cDNA 可通过从细胞、例如哺乳动物细胞、例如人细胞、包括胚胎细胞分离总 mRNA 而获得。随后可从总 mRNA 制备双链 cDNA，用各种已知技术插入到合适的质粒或噬菌

体载体中。编码 *hedgehog* 蛋白的基因也可用已知的聚合酶链反应技术克隆。

优选的核酸编码的 *hedgehog* 多肽包含与选自 SEQ ID NO: 8—14 的氨基酸序列具有至少 60% 同源性或相同性、更优选 70% 同源性或相同性、最优选 80% 同源性或相同性的氨基酸序列。编码与 SEQ ID NO: 8—14 之一所代表的氨基酸序列有至少约 90%、更优选至少约 95%、最优选至少约 98—99% 的同源性或相同性的多肽的核酸也属于本发明范围。

除了天然的 *hedgehog* 蛋白，本发明优选的 *hedgehog* 多肽与 SEQ ID NO: 8—14 任一所代表的氨基酸序列有至少 60%、更优选 70%、最优选 80% 的同源性或相同性。与选自 SEQ ID NO: 8—14 的序列有至少 90%、更优选至少 95%、最优选至少约 98—99% 的同源性或相同性的多肽也属于本发明范围。唯一的前提条件是 *hedgehog* 多肽能调节上皮细胞的生长。

术语“重组蛋白”是指本发明的多肽由重组 DNA 技术产生，其中通常编码 *hedgehog* 多肽的 DNA 插入一合适的表达载体中，该表达载体用于转化宿主细胞以产生异源蛋白质。另外，短语“衍生自”，当指重组 *hedgehog* 基因时，意味着在“重组蛋白”的含义里包括具有天然 *hedgehog* 蛋白的氨基酸序列或由该蛋白的天然形式的突变如取代和缺失（包括截短）产生的与其相似的氨基酸序列的蛋白质。

本发明方法可使用天然 *hedgehog* 多肽的变体形式，例如突变的变体进行。

如现有技术已知的，*hedgehog* 多肽可通过标准生物学技术或经化学合成产生。例如，用指导编码本发明多肽的核苷酸序列表达的核酸载体转染的宿主细胞可在合适的条件下培养以使肽表达。多肽 *hedgehog* 可被分泌或从含有重组 *hedgehog* 多肽的细胞和培养基混合

物中分离出。或者，可通过从重组 *hedgehog* 基因中除去信号肽序列而使该肽保留在细胞质中，收获细胞，裂解以及分离蛋白质。细胞培养物包括宿主细胞、培养基和其它副产物。用于细胞培养的合适的培养基是现有技术中已知的。重组 *hedgehog* 多肽可用本领域已知的纯化蛋白的技术如离子交换层析、凝胶过滤层析、超滤、电泳和用特异于该肽的抗体的免疫亲和纯化，从细胞培养物培养基、宿主细胞或两者中分离出来。在一优选的实施方案中，重组 *hedgehog* 多肽含有促进纯化的结构域的融合蛋白，如 *hedgehog*/GST 融合蛋白。宿主细胞可以是原核或真核细胞。

重组 *hedgehog* 基因可通过将编码 *hedgehog* 蛋白或其部分的核酸与适于在原核细胞、真核细胞或两种细胞中表达的载体连接二产生。用于产生重组形式的本发明 *hedgehog* 多肽的表达载体包括质粒和其它载体。例如，用于表达 *hedgehog* 多肽的合适的载体包括下列类型的质粒：pBR322 衍生的质粒、pEMBL 衍生的质粒、pEX 衍生的质粒、pBTac 衍生的质粒和 pUC 衍生的质粒，用于在原核细胞中表达，如在大肠杆菌中表达。

有各种载体可用于在酵母中表达重组蛋白，例如 YEP24、YIP5、YEP51、YEP52、pYES2 和 YRP17 是用于将基因构建体导入啤酒糖酵母的克隆和表达载体（见，例如，Broach 等(1983)，《基因表达的实验操作》，编辑 M.Inouye，学术出版社，83 页，引入本文作参考）。这些载体由于存在 pBR322 ori 而可在大肠杆菌中复制，由于存在酵母 2 μ 质粒的复制因子而可在啤酒糖酵母中复制。另外，可使用药物抗性标记如氨苄青霉素。在一示意性实施方案中，*hedgehog* 多肽用表达载体重组产生，该表达载体通过亚克隆 SEQ ID NO: 1-7 代表的一个 *hedgehog* 基因的编码序列而产生。

优选的哺乳动物表达载体即含有原核序列，以便载体在细菌中

繁殖，也含有一或多个在真核细胞中表达的真核转录单位。pcDNA1/amp、pcDNA1/neo、pRc/CMV、pSV2gpt、pSV2neo、pSV2-dhfr、pTk2、pRSVneo、pMSG、pSVT7、pko-neo 和 pHyg 衍生的载体是适于转染真核细胞的哺乳动物表达载体的例子。这些载体中的一些可用来自细菌质粒如 pBR322 的序列修饰以便在真核和原核细胞中复制和药物抗性筛选。或者，病毒衍生物如牛乳头瘤病毒（BPV-1）或 Epstein-Barr 病毒（pHEBo、pREP 衍生的和 p205）可用于蛋白质在真核细胞中的瞬时表达。制备质粒和转化宿主生物体的各种方法是本领域熟知的。对于其它合适的原核和真核细胞表达系统以及通用的重组程序，参见分子克隆实验手册，第 2 版，Sambrook, Fritsch 和 Maniatis 编辑（冷泉港实验室出版社，1989），第 16 和 17 章。

在某些情况下，希望的是用杆状病毒表达系统表达重组 *hedgehog* 多肽，这种杆状病毒表达系统的例子包括 pVL 衍生的载体（如 pVL1392, pVL1393 和 pVL941），pAcUW 衍生的载体（如 pAcUW1）和 pBlueBac 衍生的载体（如含有 β -gal 的 pBlueBac III）。

当希望仅表达 *hedgehog* 蛋白的一部分，如缺失 N-末端部分的形式即缺少信号肽的截短突变体时，可能需要在含有待表达的所需序列的寡核苷酸片段上加入一起始密码子（ATG）。本领域公知 N-末端位置的甲硫氨酸可用甲硫氨酸氨基肽酶（MAP）二酶切除去。MAP 已从大肠杆菌（Ben-Bassat 等(1987) 细菌学杂志 169: 751-757）和鼠伤寒沙门氏菌中克隆，其体外活性已在重组蛋白上证实（Miller 等(1987) PNAS 84: 2718-2722）。因此，如果需要除去 N-末端甲硫氨酸，在体内可通过在产生 MAP 的宿主（例如大肠杆菌或 CM89 或啤酒糖酵母）中表达 *hedgehog* 衍生的多肽，在体外可用纯化的 MAP 而实现（例如，Miller 等的程序，文献同上）。

或者，多肽的编码序列可作为包括编码不同多肽的核苷酸序列

的融合基因的一部分而掺入。公认的是融合蛋白也可促进蛋白质的表达，因而可用于表达本发明的 *hedgehog* 多肽。例如，*hedgehog* 多肽可作为谷胱甘肽-S-转移酶（GST-融合体）蛋白而产生。这种 GST-融合蛋白可使得 *hedgehog* 多肽易于纯化，例如通过使用谷胱甘肽衍生的基质纯化（例如，参见《分子生物学当前方案》，Ausubel 等编辑(N.Y.: John Wiley & Sons, 1991)）。在另一个实施方案中，编码纯化前导序列如聚（His）/肠激酶裂解位点序列的融合基因可用于替换 *hedgehog* 蛋白 N-末端天然存在的信号序列（例如，蛋白原形式的信号序列），以使得能用 Ni^{2+} 金属树脂经亲和层析纯化聚（His）-*hedgehog* 蛋白。随后可通过用肠激酶处理除去纯化前导序列（例如，参见 Hochuli 等(1987)层析杂志 411: 177; Janknecht 等, PNAS88: 8972）。

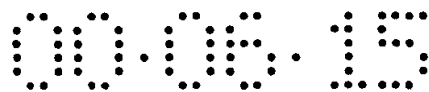
制备融合基因的技术是本领域技术人员公知的，简要地，根据常规技术，用钝端或交错切口末端进行连接，限制酶消化以提供合适的末端，如果需要补平粘末端，碱性磷酸酶处理以避免不希望的连接，以及酶连接来实现编码不同多肽序列的各 DNA 片段的连接。在另一个实施方案中，融合基因可经常规技术如自动化 DNA 合成仪合成。或者，可用锚定引物进行基因片段的 PCR 扩增，在两个连续基因片段之间产生互补突出端，随后可退火产生嵌合基因序列（例如，参见《分子生物学当前方案》，Ausubel 等编辑, John Wiley & Sons, 1992）。

通过与其它化学基元如糖基基团、胆固醇、类异戊二烯、脂类、磷酸根、乙酰基等形成共价或聚集缀合物，可化学修饰 *hedgehog* 多肽以产生 *hedgehog* 衍生物。*Hedgehog* 蛋白的共价衍生物可通过将化学基元与蛋白的氨基酸侧链上的官能团连接或将其连接在多肽的 N-末端或 C-末端而制备。

例如，可产生这样的 *hedgehog* 蛋白，其包括一不是与该蛋白天然相关的序列的基元，该基元结合胞外基质的组分并促进该类似物定位于细胞表面。例如，从纤连蛋白“III 型重复序列”衍生的序列如四肽序列 R-G-D-S (Pierschbacher 等(1984) 自然 309: 30-3; Kornblihtt 等(1985) EMBO 4:1755-9) 可加至 *hedgehog* 多肽以支持该嵌合分子通过结合 ECM 组分而附着于细胞 (Ruoslahti 等(1987) 科学 238: 491-497; Pierschbacher 等(1987) 生物化学杂志 262: 17294-8; Hynes (1987) 细胞 48: 549-54; Hynes (1992) 细胞 69: 11-25)。

在一优选的实施方案中，*hedgehog* 多肽是分离的形式，或者基本上不含其它细胞蛋白、特别是其它胞外蛋白或与 *hedgehog* 多肽天然结合的细胞表面结合蛋白，除非以与 *hedgehog* 多肽的融合蛋白形式提供。术语“基本上不含其它细胞或胞外蛋白”(本文也称为“污染蛋白”)或“基本上纯化的制备物”或“纯化的制备物”的定义是包括具有少于 20% (干重) 污染蛋白、优选具有少于 5% 污染蛋白的 *hedgehog* 多肽制备物。“纯化的”是指所指分子中基本上不存在其它生物学大分子如其它蛋白。本文所用术语“纯化的”优选指存在至少 80% (干重)、更优选 95-99% (重量)、最优选至少 99.8% (重量) 的同种类型生物学大分子 (但是水、缓冲液、其它小分子特别是分子量小于 5000 的分子可以存在)。本文所用术语“纯的”优选具有上述术语“纯化的”的相同数值限制。

如上针对重组多肽所述，分离的 *hedgehog* 多肽可包括 SEQ ID NO: 10-18 任一或 SEQ ID NO: 20 所代表的氨基酸序列的全部或一部分，或其同源序列。本发明的 *hedgehog* 蛋白的优选的片段相应于成熟蛋白的 N-末端和 C-末端蛋白水解片段。*hedgehog* 多肽的生物活性片段在 PCT 出版物 WO95/18856 和 WO96/17924 中有详述。



关于 *hedgehog* 多肽的生物活性片段，优选的 *hedgehog* 治疗剂包括 *hedgehog* 多肽的至少 50 个（连续）氨基酸残基，更优选至少 100 个（连续）氨基酸残基，更优选至少 150 个（连续）氨基酸残基。

可包括在 *hedgehog* 治疗剂中的另一种优选的 *hedgehog* 多肽是成熟蛋白的 N-末端片段，其分子量为约 19kDa。

优选的人 *hedgehog* 蛋白包括约相应于 SEQ ID NO: 15 的 24—197 位、SEQ ID NO: 16 的 28—202 位以及 SEQ ID NO: 17 的 23—198 位残基的 N-末端片段。“约相应于”是指感兴趣的序列与参照序列在长度上至多有 20 个氨基酸残基的不同，更优选在长度上至多 5、10 或 15 个氨基酸不同。

如上针对重组多肽所述，分离的 *hedgehog* 多肽可包括 SEQ ID NO: 8、SEQ ID NO: 9、SEQ ID NO: 10、SEQ ID NO: 11、SEQ ID NO: 12、SEQ ID NO: 13 或 SEQ ID NO: 14 所代表的氨基酸序列的全部或一部分，或其同源序列。本发明的 *hedgehog* 蛋白的优选的片段相应于成熟蛋白的 N-末端和 C-末端蛋白水解片段。*hedgehog* 多肽的生物活性片段在 PCT 出版物 WO95/18856 和 WO96/17924 中有详述。

其它优选的 *hedgehog* 多肽包括由式 A—B 代表的氨基酸序列，其中：(i) A 代表 SEQ ID NO: 21 的残基 1—168 指示的氨基酸序列的全部或一部分；B 代表 SEQ ID NO: 21 的残基 169—221 指示的氨基酸序列的至少一个氨基酸残基；(ii) A 代表 SEQ ID NO: 15 的残基 24—193 指示的氨基酸序列的全部或一部分；B 代表 SEQ ID NO: 15 的残基 194—250 指示的氨基酸序列的至少一个氨基酸残基；(iii) A 代表 SEQ ID NO: 13 的残基 25—193 指示的氨基酸序列的全部或一部分；B 代表 SEQ ID NO: 13 的残基 194—250 指示的氨基酸序列的至少一个氨基酸残基；(iv) A 代表 SEQ ID NO: 11 的残基 23—193

指示的氨基酸序列的全部或一部分；B 代表 SEQ ID NO: 11 的残基 194—250 指示的氨基酸序列的至少一个氨基酸残基；(v) A 代表 SEQ ID NO: 12 的残基 28—197 指示的氨基酸序列的全部或一部分；B 代表 SEQ ID NO: 12 的残基 198—250 指示的氨基酸序列的至少一个氨基酸残基；(vi) A 代表 SEQ ID NO: 16 的残基 29—197 指示的氨基酸序列的全部或一部分；B 代表 SEQ ID NO: 16 的残基 198—250 指示的氨基酸序列的至少一个氨基酸残基；或(vii) A 代表 SEQ ID NO: 17 的残基 23—193 指示的氨基酸序列的全部或一部分；B 代表 SEQ ID NO: 17 的残基 194—250 指示的氨基酸序列的至少一个氨基酸残基。在优选的实施方案中，A 和 B 一起代表所指序列的连续多肽序列，A 代表所指序列的至少 25、50、75、100 或 150 个（连续）氨基酸，B 代表序列表相应登录号所指氨基酸序列的至少 5、10 或 20 个（连续）氨基酸残基。A 和 B 一起优选代表相应于序列表登录号的连续序列。来自其它 *hedgehog* 的相似片段也包含在内，例如相应于上述数字表示的序列表登录号的优选片段的片段。在优选的实施方案中，*hedgehog* 多肽包括用胆固醇衍生的 C-末端甘氨酸（或其它合适的残基）。

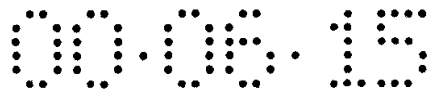
Hedgehog 蛋白的分离的肽部分可通过筛选从编码这种肽的相应核酸片段重组产生的肽而获得。另外，可用本领域已知的技术如常规的 Merrifield 固相 f-Moc 或 t-Boc 化学法化学合成片段。例如，本发明的 *hedgehog* 多肽可任意分成所需长度的不重叠的片段，或优选地分成所需长度的重叠的片段。可以（重组或经化学合成）产生片段并测试鉴别可作为野生型（例如“真的”）*hedgehog* 蛋白的兴奋剂或拮抗剂的肽片段。例如，Roman 等(1994)，欧洲生物化学杂志 222: 65—73 描述了用以鉴别结合结构域的应用来自较大蛋白的短的重叠的合成肽的竞争结合分析法。

本发明的重组 *hedgehog* 多肽还包括真的 *hedgehog* 蛋白的同系

物，如抗蛋白酶解的蛋白质，例如由于改变潜在的裂解序列或使与该蛋白相关的酶活性失活的突变所导致的蛋白质。本发明的 *hedgehog* 同系物也包括被翻译后修饰成不同于真的蛋白质的蛋白质。*Hedgehog* 蛋白的衍生物的例子包括缺少 N-糖基化位点的多肽（例如产生非糖基化蛋白），缺少胆固醇化位点的多肽，和/或缺少 N-末端和/或 C-末端序列的多肽。

本发明的 *hedgehog* 多肽结构的修饰的目的也可为了增强治疗或预防效果，或稳定性（例如来自体内的储存寿命和体内抗蛋白酶降解的抗性）。这种修饰的肽，当设计应用保持天然蛋白的至少一种活性时，被认为是本文详述的 *hedgehog* 多肽的功能等同物。这种修饰的肽例如可经氨基酸取代、缺失或添加而产生。

本领域公知的是，人们有理由预期可进行某些氨基酸的独立替换，例如用一个相关氨基酸替换另一个氨基酸残基（即同位和/或等电突变），而基本上不会影响所得分子的生物学活性。保守替换是发生在一族氨基酸中与其侧链有关的替换。遗传编码的氨基酸可分成四族：（1）酸性氨基酸＝天冬氨酸，谷氨酸；（2）碱性氨基酸＝赖氨酸，精氨酸，组氨酸；（3）非极性氨基酸＝丙氨酸，缬氨酸，亮氨酸，异亮氨酸，脯氨酸，苯丙氨酸，甲硫氨酸，色氨酸；（4）非带电荷极性氨基酸＝甘氨酸，天冬酰胺，谷氨酰胺，半胱氨酸，丝氨酸，苏氨酸，酪氨酸。苯丙氨酸，色氨酸和酪氨酸有时联合分类成芳香氨基酸。以相似的方式，氨基酸所有组成成分可分组为（1）酸性氨基酸＝天冬氨酸，谷氨酸；（2）碱性氨基酸＝赖氨酸，精氨酸，组氨酸；（3）脂族氨基酸＝甘氨酸，丙氨酸，缬氨酸，亮氨酸，异亮氨酸，丝氨酸，苏氨酸，丝氨酸和苏氨酸可任选地单独分为脂族羟基氨基酸组；（4）芳香氨基酸＝苯丙氨酸，酪氨酸，色氨酸；（5）酰胺氨基酸＝天冬酰胺，谷氨酰胺；（6）含硫氨基酸＝半胱氨酸，



甲硫氨酸。(参见, 例如, 《生物化学》, 第 2 版, L. Stryer 编辑, WH Freeman and Co.: 1981)。氨基酸序列的改变是否会导致有功能的 *hedgehog* 同系物 (例如, 有功能是指其模拟或拮抗野生型) 可以通过评价变体肽能否在细胞中以类似于野生型蛋白的方式产生应答或竞争性抑制这种应答而确定。发生一个以上替换的多肽可以相同方式容易地测定。

应特别指出的是本发明方法可用天然 *hedgehog* 蛋白的同系物进行。在一个实施方案中，本发明设计使用经组合诱变产生的 *hedgehog* 多肽。如本领域所知的，这种方法可方便地产生点突变体和截短突变体，并特别适于鉴别能结合 *hedgehog* 蛋白受体的潜在的变体序列（例如同系物）。筛选这种组合文库的目的是产生例如能作为兴奋剂或拮抗剂的新的 *hedgehog* 同系物。举例来说，可用本发明方法对 *hedgehog* 同系物工程化，以提供对关联受体如 *patched* 更有效的结合，并仍保持至少一部分 *hedgehog* 相关的活性。因此，可产生与该蛋白的天然形式相比具有增加的效力的组合衍生的同系物。类似地，可用组合途径产生作为拮抗剂的 *hedgehog* 同系物，即它们能够例如模拟与其它本文基质组分（如受体）的结合，但不诱导任何生物学应答，从而抑制真的 *hedgehog* 或 *hedgehog* 兴奋剂的作用。另外，用本发明方法对 *hedgehog* 某些结构域的操作可提供更适用于融合蛋白中的结构域，如掺入衍生自胞外基质和/或结合胞外基质组分的其它蛋白的一部分的结构域。

为进一步描述组合诱变的现有技术状态，应注意 Gallop 等 (1994)，医学化学杂志 37: 1233 的综述文章描述了 90 年代早期组合文库的现有技术状态。特别是 Gallop 等在第 1239 页指出“筛选类似物文库有助于确定活性序列的最小大小，以及有助于鉴别对结合和取代的不耐受关键的残基”。另外，Ladner 等的 PCT 出版物

WO90/02809, Goeddel 等的美国专利 5,223,408 和 Markland 等的 PCT 出版物 WO92/15679 描述了本领域技术人员可以用于产生 *hedgehog* 变体文库的具体技术, 该文库可被快速筛选以鉴别保持 *hedgehog* 多肽的特定活性的变体/片段。这些技术是对现有技术的举例并证实可产生相关变体/截短体的大文库并分析其以不需过度实验而分离特定变体。Gustin 等 (1993), 病毒学 193: 653, Bass 等 (1990), 蛋白质: 结构、功能和遗传 8: 309—314 也描述了可用于产生 *hedgehog* 多肽的诱变变体的现有技术的举例性技术。

实际上, 从组合诱变现有技术可以明显看出对 *hedgehog* 蛋白的大规模诱变, 而事先不指哪个残基是生物学功能关键残基, 将产生具有相等生物学活性的许多变体。实际上组合技术具有经高通量分析筛选数十亿个不同变体的能力, 从而不需事先知道关键残基。

举例而说, 优选地将一群 *hedgehog* 同系物或其它相关蛋白的氨基酸序列进行对比以促使最高同源性成为可能。这种变体群可包括, 例如, 来自一或多个物种的 *hedgehog* 同系物选择在对比序列的每个位置出现的氨基酸以产生一套简并的组合序列。在一优选的实施方案中, *hedgehog* 变体的多样化文库经核酸水平的组合诱变产生并由一多样化基因文库编码。例如, 合成寡核苷酸的混合物可与基因序列酶连接, 以便潜在的 *hedgehog* 序列的简并系列可表达为个体多肽, 或者, 表达为其中含有该系列 *hedgehog* 序列的一套较大的融合蛋白 (例如, 用于噬菌体展示)。

如 PCT 出版物 WO95/18856 所示, 为分析变体群的序列, 感兴趣的氨基酸序列可相对于序列同源性进行序列对比。特定变体的对比序列是否存在氨基酸与参照序列所选共有长度有关, 该参照序列可以是真序列或人工序列。

在一举例性实施方案中, 每个 *Shh* 克隆的外显子 1、2 和一部分

外显子 3 编码的序列（例如成熟蛋白的 N-末端约 221 个残基）的对比产生由下述通式代表的 *Shh* 多肽的简并系列：

C-G-P-G-R-G-X(1)-G-X(2)-R-R-H-P-K-K-L-T-P-L-A-Y-K-Q-F-I-P-N-V-A-E-K-T-L-G-A-S-G-R-Y-E-G-K-I-X(3)-R-N-S-E-R-F-K-E-L-T-P-N-Y-N-P-D-I-I-F-K-D-E-E-N-T-G-A-D-R-L-M-T-Q-R-C-K-D-K-L-N-X(4)-L-A-I-S-V-M-N-X(5)-W-P-G-V-X(6)-L-R-V-T-E-G-W-D-E-D-G-H-H-X(7)-E-E-S-L-H-Y-E-G-R-A-V-D-I-T-T-S-D-R-D-X(8)-S-K-Y-G-X(9)-L-X(10)-R-L-A-V-E-A-G-F-D-W-V-Y-Y-E-S-K-A-H-I-H-C-S-V-K-A-E-N-S-V-A-A-K-S-G-G-C-F-P-G-S-A-X(11)-V-X(12)-L-X(13)-X(14)-G-G-X(15)-K-X(16)-V-K-D-L-X(17)-P-G-D-X(18)-V-L-A-A-D-X(19)-X(20)-G-X(21)-L-X(22)-X(23)-S-D-F-X(24)-X(25)-F-X(26)-D-R (SEQ ID NO: 21)

其中每个简并位置“X”可以是在人、鼠、鸡或斑马鱼 *Shh* 克隆中该位置发生的氨基酸，或者，为扩大该文库，每个 X 也可选自是每个位置天然出现的氨基酸的保守取代的氨基酸。例如，Xaa(1)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、苯丙氨酸、酪氨酸或色氨酸；Xaa(2)代表精氨酸、组氨酸或赖氨酸；Xaa(3)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸或苏氨酸；Xaa(4)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸或苏氨酸；Xaa(5)代表赖氨酸、精氨酸、组氨酸、天冬酰胺或谷氨酰胺；Xaa(6)代表赖氨酸、精氨酸或组氨酸；Xaa(7)代表丝氨酸、苏氨酸、酪氨酸、色氨酸或苯丙氨酸；Xaa(8)代表赖氨酸、精氨酸或组氨酸；Xaa(9)代表甲硫氨酸、半胱氨酸、丝氨酸或苏氨酸；Xaa(10)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸或苏氨酸；Xaa(11)代表亮氨酸、缬氨酸、甲硫氨酸、苏氨酸或丝氨酸；Xaa(12)代表组氨酸、苯丙氨酸、酪氨酸、丝氨酸、苏氨酸、甲硫氨酸或半胱氨酸；Xaa(13)代表谷氨酰胺、天冬酰胺、谷氨酸或天冬氨酸；Xaa(14)代表组氨酸、苯丙氨酸、酪氨酸、苏氨酸、谷氨酰胺、天冬酰胺、谷氨酸或天冬氨酸。

酸；Xaa(15)代表谷氨酰胺、天冬酰胺、谷氨酸、天冬氨酸、苏氨酸、丝氨酸、甲硫氨酸或半胱氨酸；Xaa(16)代表丙氨酸、甘氨酸、半胱氨酸、亮氨酸、缬氨酸或甲硫氨酸；Xaa(17)代表精氨酸、赖氨酸、甲硫氨酸、异亮氨酸、天冬酰胺、天冬氨酸、谷氨酸、谷氨酰胺、丝氨酸、苏氨酸或半胱氨酸；Xaa(18)代表精氨酸、赖氨酸、甲硫氨酸或异亮氨酸；Xaa(19)代表丙氨酸、甘氨酸、半胱氨酸、天冬氨酸、谷氨酸、谷氨酰胺、天冬酰胺、丝氨酸、苏氨酸或甲硫氨酸；Xaa(20)代表丙氨酸、甘氨酸、半胱氨酸、天冬氨酸、天冬酰胺、谷氨酸或谷氨酰胺；Xaa(21)代表精氨酸、赖氨酸、甲硫氨酸、异亮氨酸、天冬酰胺、天冬氨酸、谷氨酸或谷氨酰胺；Xaa(22)代表亮氨酸、缬氨酸、甲硫氨酸或异亮氨酸；Xaa(23)代表苯丙氨酸、酪氨酸、苏氨酸、组氨酸或色氨酸；Xaa(24)代表异亮氨酸、缬氨酸、亮氨酸或甲硫氨酸；Xaa(25)代表甲硫氨酸、半胱氨酸、异亮氨酸、亮氨酸、缬氨酸、苏氨酸或丝氨酸；Xaa(26)代表亮氨酸、缬氨酸、甲硫氨酸、苏氨酸或丝氨酸。在一个更扩大的文库，每个X可以选自任何氨基酸。

以类似的方式，每个人、鼠、鸡或斑马鱼 *hedgehog* 克隆的序列对比可提供由下述通式代表的简并肽序列：

C-G-P-G-R-G-X(1)-X(2)-X(3)-R-R-X(4)-X(5)-X(6)-P-K-X(7)-L-X(8)-P-L-X(9)-Y-K-Q-F-X(10)-P-X(11)-X(12)-X(13)-E-X(14)-T-L-G-A-S-G-X(15)-X(16)-E-G-X(17)-X(18)-X(19)-R-X(20)-S-E-R-F-X(21)-X(22)-L-T-P-N-Y-N-P-D-I-I-F-K-D-E-E-N-X(23)-G-A-D-R-L-M-T-X(24)-R-C-K-X(25)-X(26)-X(27)-N-X(28)-L-A-I-S-V-M-N-X(29)-W-P-G-V-X(30)-L-R-V-T-E-G-X(31)-D-E-D-G-H-H-X(32)-X(33)-X(34)-S-L-H-Y-E-G-R-A-X(35)-D-I-T-T-S-D-R-D-X(36)-X(37)-K-Y-G-X(38)-L-X(39)-R-L-A-V-E-A-G-F-D-W-V-Y-Y-E-S-X(40)-X(41)-H-X(42)-H-X(43)-S-V-K-X(44)-X(45) (SEQ IDNo:22)

其中，如上所述，每个简并位置“X”可以是在一个野生型克隆

中相应位置发生的氨基酸，或者，可包括是保守取代的氨基酸残基，或者每个 X 可以是任何氨基酸残基。在一个举例性实施方案中，Xaa(1)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸或酪氨酸；Xaa(2)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸；Xaa(3)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、赖氨酸、组氨酸或精氨酸；Xaa(4)代表赖氨酸、精氨酸或组氨酸；Xaa(5)代表苯丙氨酸、色氨酸、酪氨酸或氨基酸间隔；Xaa(6)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸或氨基酸间隔；Xaa(7)代表天冬酰胺、谷氨酰胺、组氨酸、精氨酸或赖氨酸；Xaa(8)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸或苏氨酸；Xaa(9)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸或苏氨酸；Xaa(10)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸或苏氨酸；Xaa(11)代表丝氨酸、苏氨酸、谷氨酰胺或天冬酰胺；Xaa(12)代表甲硫氨酸、半胱氨酸、甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸或苏氨酸；Xaa(13)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸或脯氨酸；Xaa(14)代表精氨酸、组氨酸或赖氨酸；Xaa(15)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、精氨酸、组氨酸或赖氨酸；Xaa(16)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、苯丙氨酸或酪氨酸；Xaa(17)代表精氨酸、组氨酸或赖氨酸；Xaa(18)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸或苏氨酸；Xaa(19)代表苏氨酸或丝氨酸；Xaa(20)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、天冬酰胺或谷氨酰胺；Xaa(21)代表精氨酸、组氨酸或赖氨酸；Xaa(22)代表天冬氨酸或谷氨酸；Xaa(23)代表丝氨酸或苏氨酸；Xaa(24)代表谷氨酸、天冬氨酸、谷氨酰胺或天冬酰胺；Xaa(25)代表谷氨酸或天冬氨酸；Xaa(26)代表精氨酸、组氨酸或赖氨酸；Xaa(27)代表甘氨酸、丙氨酸、缬氨酸

酸、亮氨酸或异亮氨酸；Xaa(28) 代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸或苏氨酸；Xaa(29) 代表甲硫氨酸、半胱氨酸、谷氨酰胺、天冬酰胺、精氨酸、赖氨酸或组氨酸；Xaa(30)代表精氨酸、组氨酸或赖氨酸；Xaa(31)代表色氨酸、苯丙氨酸、酪氨酸、精氨酸、组氨酸或赖氨酸；Xaa(32) 代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸、苏氨酸、酪氨酸或苯丙氨酸；Xaa(33) 代表谷氨酰胺、天冬酰胺、天冬氨酸或谷氨酸；Xaa(34) 代表天冬氨酸或谷氨酸；Xaa(35)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸；Xaa(36)代表精氨酸、组氨酸或赖氨酸；Xaa(37) 代表谷氨酰胺、天冬酰胺、苏氨酸或丝氨酸；Xaa(38)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸、苏氨酸、甲硫氨酸或半胱氨酸；Xaa(39)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、苏氨酸或丝氨酸；Xaa(40)代表精氨酸、组氨酸或赖氨酸；Xaa(41)代表天冬酰胺、谷氨酰胺、甘氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸；Xaa(42)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸；Xaa(43)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸、苏氨酸或半胱氨酸；Xaa(44)代表甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、苏氨酸或丝氨酸；Xaa(45)代表天冬氨酸或谷氨酸。

有许多方式可用以从简并寡核苷酸序列产生潜在的 *hedgehog* 同系物的文库。简并基因序列的化学合成可在自动 DNA 合成仪上进行，然后将合成的基因连接进合适的表达载体中。简并系列基因的目的是提供在一个混合物中的编码所需的潜在 *hedgehog* 序列的所有序列。简并寡核苷酸的合成是本领域已知的(例如参见, Narang, SA(1983) *Tetrahedron* 39: 3; Itakura 等(1981) 重组 DNA, 第 3 届克里夫兰大分子研讨会论文汇编, 编辑 AG Walton, 阿姆斯特丹: Elsevier, 第 273—289 页; Itakura 等(1984)生物化学年度综述 53: 323; Itakura 等(1984)

科学 198: 1056; Ike 等(1983) 核酸研究 11: 477)。这种技术已在其它蛋白的定向进化中被采用 (例如参见 Scott 等(1990) 科学 249: 386—390; Roberts 等(1992) PNAS 89: 2429—2433; Devlin 等(1990) 科学 249: 404—406; Cwirla 等(1990) PNAS 87: 6378—6382; 以及美国专利 5,223,409, 5,198,346 和 5,096,815)。

本领域已知各种用于筛选由点突变制备的组合文库的基因产物以及筛选 cDNA 文库的具有某种性质的基因产物的技术。这种技术通常适于快速筛选由 *hedgehog* 同系物的组合诱变产生的基因文库。筛选大的基因文库的最常用技术典型地包括将基因文库克隆至可复制的表达载体上, 用得到的载体文库转化合适的细胞, 表达组合基因, 在所需活性的检测有利于编码所检测产物的基因的载体相对较易分离的条件下进行表达。下述每种举例性分析均适于高通量分析, 这是筛选由组合诱变技术产生的大量简并 *hedgehog* 序列所需的。

在一个实施方案中, 组合文库设计是能分泌的 (例如, 文库的多肽均包括信号序列但无跨膜或细胞质结构域), 并用于转染可与上皮干细胞共培养的真核细胞。由表达组合文库的细胞分泌的有功能的 *hedgehog* 蛋白将扩散进相邻的上皮细胞并诱导特定的生物学应答如增殖。增殖的检测模式为梯度函数, 并使得能分离 (通常在几次重复选择后) 产生是上皮细胞的增殖剂的 *hedgehog* 同系物的细胞。类似地, *hedgehog* 拮抗剂可通过产生有功能的拮抗剂的细胞保护相邻细胞 (例如抑制增殖) 不受加入到培养基中的野生型 *hedgehog* 的影响的能力来选择。

举例来说, 靶上皮细胞在 24 孔微滴板中培养, 用组合 *hedgehog* 基因文库转染其它真核细胞并在能插入到微滴板孔中的细胞培养插入物 (例如, Collaborative Biomedical Products, 货号#40446) 中培养。将该细胞培养插入物置于孔中, 从而由插入物中的细胞分泌的重组

hedgehog 同系物可以扩散通过插入物的孔状底部并接触微滴板孔中的靶细胞。在足以使有功能形式的 *hedgehog* 蛋白在靶细胞中产生可测的应答如增殖的时间后，除去插入物，测定变体 *hedgehog* 蛋白对靶细胞的作用。来自相应于活性为阳性的孔的插入物的细胞可以分成几份并在几个插入物上再培养，重复这一过程直至活性克隆被鉴别出来。

在另一种筛选分析中，候选的 *hedgehog* 基因产物展示在细胞或病毒颗粒表面，并且特定细胞或病毒颗粒经这一基因产物与 *hedgehog* 结合基元（如 *patched* 蛋白或其它 *hedgehog* 受体）相结合的能力在一“淘选分析”中被检测。这种淘选步骤可在从胚胎培养的细胞上进行。例如，基因文库可被克隆进编码细菌细胞表面膜蛋白的基因中，经淘选检测得到的融合蛋白（Ladner 等, WO88/06630; Fuchs 等(1991) 生物/技术 9: 1370—1371; Goward 等(1992) TIBS 18:136-140）。以类似的方式，结合 *hedgehog* 的荧光标记分子可用于鉴别潜在的有功能的 *hedgehog* 同系物。细胞可以在荧光显微镜下经肉眼检测并分离，或者，当细胞形态学允许时，用荧光活化的细胞分选仪分离。

在另一实施方案中，基因文库在病毒克隆上表达为融合蛋白。例如，在丝状噬菌体系统中，外来肽序列可在感染性噬菌体的表面上表达，从而赋予两种显著的优势。第一，由于这些噬菌体可以非常高的浓度加到亲和基质上，所以可一次筛选大量噬菌体。第二，由于每个感染性噬菌体在其表面展示组合基因产物，所以如果从亲和基质上回收的特定噬菌体量较低，该噬菌体可经另一轮感染二扩增。一组几乎相同的大肠杆菌丝状噬菌体 M13、fd 和 f1 最常用于噬菌体展示文库，因为噬菌体 gIII 或 gVIII 外壳蛋白可用于产生融合蛋白而不破坏病毒颗粒的最外层包装（Ladner 等, PCT 出版物 WO90/02909; Garrard 等, PCT 出版物 WO92/09690; Marks 等(1992) 生

物化学杂志 267: 16007—16010; Griffiths 等(1993) EMBO J 12: 725—734; Clackson 等(1991) 自然 352: 624—628; Barbas 等(1992) PNAS 89: 4457—4461)。

在一举例性实施方案中, 可容易地修饰重组噬菌体抗体系统 (RPAS, Pharmacia 货号 27—9400—01) 以用于表达和筛选 *hedgehog* 组合文库。例如, RPAS 试剂盒的 pCANTAB 5 噬菌粒含有编码噬菌体 gIII 外壳蛋白的基因。*Hedgehog* 组合基因文库可克隆进噬菌粒中邻近 gIII 信号序列, 从而其表达为 gIII 融合蛋白。连接后, 用噬菌粒转化感受态大肠杆菌 TG1 细胞, 转化的细胞随后用 M13KO7 辅助噬菌体感染以拯救噬菌粒及其候选的 *hedgehog* 基因插入子。得到的重组噬菌体含有编码特定的候选 *hedgehog* 的噬菌粒 DNA, 并展示一或多个拷贝的相应融合外壳蛋白。经淘选法选择并富集能结合 *hedgehog* 受体的噬菌体展示的候选 *hedgehog* 蛋白。例如, 将噬菌体文库加到表达 *patched* 蛋白的细胞中并洗涤以将未结合的噬菌体从细胞中除去。然后分离结合的噬菌体, 并且如果重组噬菌体表达至少一个拷贝的野生型 gIII 外壳蛋白, 则它们将保持其感染大肠杆菌的能力。因此, 继续进行再感染大肠杆菌及淘选将大大富集 *hedgehog* 同系物, 然后可筛选这些同系物进一步的生物学活性以区分兴奋剂和拮抗剂。

组合诱变具有产生非常大的突变蛋白文库的潜力, 例如 10^{26} 个分子级别的文库。这样大的组合文库即使用高通量筛选分析如噬菌体展示进行筛选在技术上也是富有挑战性的。为克服这一问题, 最近开发了一种新技术回归集团诱变 (recursive ensemble mutagenesis, REM), 其使得能在一随机文库中避免非常高比例的无功能蛋白质, 并增加有功能蛋白质的几率, 由此降低实现序列空间的有用取样所需的复杂性。REM 是一种算法, 当采用合适的选择或筛选方法时,

其增加文库中有功能的突变体的几率 (Arkin 和 Yourvan, 1992, PNAS USA 89:7811-7815; Yourvan 等, 1992, 从自然中解决的平行问题, 2, Maenner 和 Manderick 编辑, Elsevir 出版公司, 阿姆斯特丹, 401-410 页; Delgrave 等, 1993, 蛋白质工程 6(3):327-331)。

本发明还提供了降低 *hedgehog* 蛋白产生能破坏本发明的 *hedgehog* 多肽与 *hedgehog* 受体的结合的模拟物, 例如肽或非肽因子的方法。上述这种诱变技术也可用于定位参与蛋白质-蛋白质相互作用的 *hedgehog* 蛋白的决定簇, 这种相互作用例如参与本发明的 *hedgehog* 多肽与其它胞外基质组分的结合。举例来说, 可确定参与对 *hedgehog* 受体如 *patched* 的分子识别的本发明 *hedgehog* 多肽的关键残基并用于产生竞争性抑制真 *hedgehog* 蛋白与该基元的结合的 *hedgehog* 衍生的肽模拟物。通过采用, 例如, 扫描诱变以定位参与结合其它胞外蛋白的每种本发明 *hedgehog* 蛋白的氨基酸残基, 可产生模拟 *hedgehog* 蛋白的那些残基的肽模拟化合物以促进相互作用。这种模拟物随后可用于干扰 *hedgehog* 蛋白的正常功能。例如, 用苯并二氮杂 (例如, 参见 Freidinger 等, 见《肽: 化学和生物学》, G.R. Marshall 编辑, ESCOM 出版社: Leiden, 荷兰, 1988)、氮杂 (例如, 参见 Huffman 等, 见《肽: 化学和生物学》, G.R. Marshall 编辑, ESCOM 出版社: Leiden, 荷兰, 1988)、取代的 γ -内酰胺环 (Garvey 等, 见《肽: 化学和生物学》, G.R. Marshall 编辑, ESCOM 出版社: Leiden, 荷兰, 1988)、酮-亚甲基假肽 (Ewenson 等(1986) 医学化学杂志 29: 295; Ewenson 等, 见《肽: 结构和功能》(第 9 届美国肽研讨会论文汇编), Pierce 化学公司, Rockland, IL, 1985), β 转角二肽核心 (Nagai 等(1985) 四面体通信 26: 647; Sato 等(1986) J Chem Soc Perkin Trans 1:1231) 和 β -氨基醇 (Gordon 等(1985) 生物化学生物物理研究通讯 126: 419; Dann 等(1986) 生物化学生物物理研究通讯 134: 71), 可

以产生这种残基的不可水解的肽类似物。

重组产生形式的 *hedgehog* 蛋白可以用，例如，含有与至少一个转录调节序列可操纵相连的编码 *hedgehog* 多肽的核酸的表达载体来产生。可操纵相连是指核苷酸序列以能使该核苷酸序列表达的方式与一调节序列相连。调节序列是本领域已知的并选择用于指导 *hedgehog* 多肽的表达。因此，术语转录调节序列包括启动子、增强子和其它表达控制元件。这种调节序列在 Goeddel; 基因表达技术：酶学方法 185，学术出版社，圣地亚哥，CA（1990）中有描述。例如，任何表达控制序列，当与 DNA 序列可操纵相连时控制其表达的序列可用于这些载体中以表达编码 *hedgehog* 多肽的 DNA 序列。这种有用的表达控制序列包括，例如，病毒 LTR 如 Moloney 鼠白血病病毒的 LTR，SV40 的早期和晚期启动子，腺病毒或巨细胞病毒立即早期启动子，lac 系统，trp 系统，TAC 或 TRC 系统，其表达由 T7 RNA 聚合酶指导的 T7 启动子，噬菌体λ的主要操纵子和启动子区域，fd 外壳蛋白的控制区，3-磷酸甘油酯激酶或其它糖酵解酶的启动子，酸性磷酸酶的启动子，例如 Pho5，酵母α-交配因子的启动子，杆状病毒系统的多角体启动子以及已知控制原核细胞或真核细胞及其病毒的基因病毒的序列，及以上的各种组合。应理解的是表达载体的设计应基于这样的因素，如转化的宿主细胞的选择和/或需表达的蛋白质的类型。另外，也应考虑载体的拷贝数、控制该拷贝数的能力以及由载体编码的任何其它蛋白如抗生素标记的表达。

除了提供稳定的 *hedgehog* 多肽来源用于纯化外，本发明的基因构建体还可用作基因治疗方案的一部分以输送编码兴奋剂或拮抗剂形式的 *hedgehog* 多肽的核酸。因此，本发明的另一方面的特征是用在特定细胞类型中体内转染 *hedgehog* 多肽的表达载体以导致 *hedgehog* 多肽在上皮组织中的异位表达。

这种表达载体的配制品可以任何生物学有效的载体给予，例如能有效将重组基因体内输送至细胞的任何配制品或组合物。途径包括将 *hedgehog* 编码序列插入病毒载体如重组逆转录病毒、腺病毒、腺伴随病毒和单纯疱疹病毒-1，或重组细菌或真核质粒。病毒载体直接转染细胞；质粒 DNA 可借助于例如阴离子脂质体（脂转染）或衍化（例如缀合抗体），聚赖氨酸缀合物，短杆菌肽 S，人工病毒包膜或其它如此的胞内载体而输送，以及直接注射基因构建体或体内进行 CaPO_4 沉淀。应理解的是由于合适靶细胞的转导代表了基因治疗的关键的第一步，因而特定基因输送系统的选择依赖于这样的因素，如潜在的靶的表型和给药的途径，例如局部或系统给药。另外，应认识到提供用于 *hedgehog* 表达的体内转导的特定基因构建体也可用于细胞的体外转导，如用于下述的来自体内的组织培养系统。

将核酸体内导入细胞的优选的途径是利用含有编码所需的特定形式的 *hedgehog* 多肽的核酸如 cDNA 的病毒载体。用病毒载体感染细胞的优势是大部分靶细胞可接受核酸。另外，病毒载体内所编码的分子，例如由病毒载体内含的 cDNA 编码的分子可有效在已吸收病毒载体核酸的细胞中表达。

逆转录病毒载体和腺伴随病毒载体通常被认为是用于体内将外源基因转移至特别是人中应选择重组基因输送系统。这些载体能有效将基因输送至细胞中，并且转移的核酸能稳定地整合进宿主的染色体 DNA 中。应用逆转录病毒的一个主要前提条件是保证其应用的安全性，特别是应考虑野生型病毒在细胞群中传播的可能性。仅产生复制缺陷的逆转录病毒的特殊细胞系（称为“包装细胞”）的开发增加了逆转录病毒用于基因治疗的实用性，已很好地鉴定了能用于基因治疗目的的基因转移中的缺陷的逆转录病毒（综述见 Miller, A.D.(1990) 血液 76: 271）。因此，可以构建这样的重组逆转录病毒，

其中部分逆转录病毒编码序列 (*gag*, *pol*, *env*) 已被编码 *hedgehog* 多肽的核酸替换, 从而使逆转录病毒复制缺陷。然后将复制缺陷的逆转录病毒包装成病毒颗粒, 以用于通过标准技术借助辅助病毒感染靶细胞。产生重组逆转录病毒以及用这种病毒体外或体内感染细胞的方案可参见《分子生物学当前方案》, Ausubel, F.M.等(编辑), Greene 出版协会, (1989), 第 9.10-9.14 部分, 及其它标准实验室手册。合适的逆转录病毒的例子包括本领域公知的 pLJ, pZIP, pWE 和 pEM。用于制备亲嗜性和兼嗜性逆转录病毒系统的合适的包装病毒系的例子包括 Crip, Cre, 2 和 Am。逆转录病毒已用于将各种基因体内和/或体外导入许多不同的细胞类型, 包括上皮细胞 (例如参见 Eglitis 等(1985) 科学 230: 1395-1398; Danos 和 Mulligan(1988) 美国科学院院刊 85: 6460-6464; Wilson 等(1988) 美国科学院院刊 85: 3014-3018; Armentano 等(1990) 美国科学院院刊 87: 6141-6145; Huber 等(1991) 美国科学院院刊 88: 8039-8043; Ferry 等(1991) 美国科学院院刊 88: 8377-8381; Chowdhury 等(1991) 科学 254:1802-1805; van Beusechem 等(1992) 美国科学院院刊 89: 7640-7644; Kay 等(1992) 人类基因治疗 3:641-647; Dai 等(1992) 美国科学院院刊 89: 10892-10895; Hwu 等(1993) 免疫学杂志 150: 4104-4115; 美国专利 4,868,116; 美国专利 4,980,286; PCT 申请 WO89/07136; PCT 申请 WO89/02468; PCT 申请 WO89/05345 和 PCT 申请 WO92/07573)。

另外, 已表明可通过修饰病毒颗粒表面上的病毒包装蛋白来限制逆转录病毒的感染谱, 从而限制基于逆转录病毒的载体的感染谱 (例如参见 PCT 出版物 WO93/25234 和 WO94/06920)。例如, 修饰逆转录病毒载体的感染谱的策略包括: 将特异于细胞表面抗原的抗体与病毒 *env* 蛋白偶联 (Roux 等(1989) PNAS 86:9079-9083; Julan 等(1992) 普通病毒学杂志 73: 3251-3255; Goud 等(1983) 病毒学 163:

251—254); 或将细胞表面受体配体与病毒 *env* 蛋白偶联(Neda 等(1991) 生物化学杂志 266: 14143—14146)。偶联可以是与蛋白质或其它物质(例如乳糖以将 *env* 蛋白转化成脱唾液酸糖蛋白)的化学交联形式, 以及通过产生融合蛋白而偶联(例如单链抗体/*env* 融合蛋白)。这一技术尽管用于限制或指导对某些组织类型的感染, 但也可用于将亲嗜性载体转变成兼嗜性载体。

另外, 逆转录病毒基因输送的应用还可通过用控制逆转录病毒载体的 *hedgehog* 基因表达的组织特异性或细胞特异性转录调控序列来增强。

用于本发明方法的另一种病毒基因输送系统利用腺病毒衍生的载体。可操作腺病毒的基因组以使其编码并表达感兴趣的基因产物, 但使其在正常裂解性病毒生命周期中的复制能力失活。例如参见 Berkner 等(1988) 生物技术 6: 616; Rosenfeld 等(1991) 科学 252: 431—434; Rosenfeld 等(1992) 细胞 68: 143—155。衍生自腺病毒 Ad 株 5 型 dl324 或腺病毒其它株(例如 Ad2, Ad3, Ad7 等)的合适的腺病毒载体是本领域技术人员公知的。在某些情况下重组腺病毒是有利的, 因为它们可用于感染各种细胞类型, 包括上皮细胞(Rosenfeld 等(1992), 文献同上)。另外, 病毒颗粒相对较稳定并易于纯化和浓缩, 并且如上所述, 可被修饰以改变感染谱。另外, 导入的腺病毒 DNA(和其中所含的外源 DNA)不整合进宿主细胞的染色体, 而是保持附加体形式, 从而避免了导入的 DNA 整合进宿主基因组的情况(例如逆转录病毒 DNA)中由于插入诱变而可能发生的潜在问题。另外, 腺病毒基因组携带外源 DNA 的能力相对于其它基因输送载体也是较大的(高达 8kb)(Berkner 等, 文献同上; Haj-Ahmand and Graham(1986) 病毒学杂志 57: 267)。目前最常用的因而也是本发明优选的复制缺陷的腺病毒载体缺失了病毒 E1 和 E3 基因的全部或部

分但仍保持多达 80% 的腺病毒遗传物质(例如, 参见 Jones 等(1979) 细胞 16: 683; Berkner 等, 文献同上; Graham 等, 《分子生物学方法》, E.J. Murray 编辑(Humana, Clifton, NJ, 1991), 第 7 卷, 第 109—127 页)。插入的 *hedgehog* 基因的表达可在例如 E1A 启动子、主要晚期启动子 (MLP) 和相关前导序列、E3 启动子或外源添加的启动子序列的控制下。

除了如上所述的病毒转移方法, 也可采用非病毒方法以使 *hedgehog* 多肽在动物组织中表达。大多数非病毒基因转移方法依赖于哺乳动物细胞吸收及胞内转运大分子所用的正常机制。在优选的实施方案中, 本发明的非病毒基因输送系统依赖于胞吞途径已由靶细胞吸收 *hedgehog* 多肽基因。这一类型的基因输送系统的例子包括脂质体衍生的系统、聚赖氨酸缀合物以及人工病毒包膜。

在临床设置中, 用于治疗性 *hedgehog* 基因的基因输送系统可以通过各种方法导入患者, 每种方法均是本领域熟知的。例如, 基因输送系统的药物制剂可经例如静脉注射系统导入, 靶细胞中的蛋白质的特异性转导主要由基因输送载体提供的转染的特异性, 由于控制受体基因表达的转录调节序列导致的细胞型或组织型表达的特异性, 或其组合所产生。在其它实施方案中, 重组基因的最初输送更限制在导入相当局限的动物内。例如, 基因输送载体可经导管(见美国专利 5,328,470)或立体异构(stereotactic)注射(例如 Chen 等(1994) PNAS 91:3054-3057)而导入。*hedgehog* 表达构建体可在基因治疗构建体中经例如电穿孔被输送至皮肤细胞, 电穿孔所用技术例如如 Dev 等((1994) 癌症治疗综述 20: 105—115)所述。

基因治疗构建体的药物制剂可基本上由在可接受的稀释剂中的基因输送系统组成, 或者可包括一种缓释基质, 该基质中包埋基因输送载体。或者, 当可从重组细胞完整产生完整的基因输送系统,

例如逆转录病毒载体时，药物制剂可包括一或多种产生基因输送系统的细胞。

在另一个实施方案中，*hedgehog* 或 *ptc* 治疗剂可以是一种“基因活化”构建体，其通过与基因组 DNA 同源重组而改变内源基因的转录调节序列。例如，基因活化构建体可以将 *hedgehog* 基因的内源启动子用异源启动子替换，所述异源启动子例如是在不同于基因的正常表达模式的条件下能导致 *hedgehog* 基因的组成型表达或导致该基因的诱导型表达的启动子。*patched* 信号途径中的其它基因可以类似地靶击。已知基因活化构建体的各种不同形式。例如，参见 Transkaryotic 治疗公司的 PCT 出版物 WO93/09222, WO95/31560, WO96/29411, WO95/31560 和 WO94/12650。

在优选的实施方案中，用作基因活化构建体的核苷酸序列可以包括（1）来自内源性 *hedgehog* 基因的某些部分（外显子序列，内含子序列，启动子序列等）的指导重组的 DNA 和（2）异源转录调控序列，其在基因活化构建体的重组后与基因组 *hedgehog* 基因的编码序列可操纵相连。为用于产生 *hedgehog* 生产细胞的培养物，该构建体还可进一步包括报道基因以检测细胞中剔除构建体的存在。

基因活化构建体插入细胞，并与细胞的基因组 DNA 在如此的位置整合，以将异源调控序列可操纵地与天然 *hedgehog* 基因结合。这种插入经同源重组而发生，即与内源性 *hedgehog* 基因序列同源的活化构建体的重组区与基因组 DNA 杂交，并与基因组序列重组，从而该构建体掺入基因组 DNA 的相应位置。

术语“重组区”或“靶序列”是指基因活化构建体的区段（即一部分），其具有与基因组序列基本上相同或基本上互补的序列，例如，包括基因组基因的 5' 侧翼序列，并可促进基因组序列和靶转基

因构建体之间的同源重组。

如本文所用，术语“置换区”是指活化构建体的一部分，其在重组区和基因组序列之间的同源重组后整合进内源性染色体位置。

异源调控序列，例如，提供在置换区的异源调控序列，可包括一或多个不同元件，包括：启动子（如组成型或诱导型启动子）、增强子、阴性调控元件、基因座控制区、转录因子结合位点或其组合。可用于控制靶基因的体内表达的启动子/增强子包括但不限于巨细胞病毒（CMV）启动子/增强子（Karasuyama 等，1989，实验医学杂志，169: 13），人 β -肌动蛋白启动子（Gunning 等(1987) PNAS 84:4831-4835），存在于小鼠乳腺癌病毒长末端重复序列（MMTV LTR）中的糖皮质激素诱导型启动子（Klessig 等(1984) 分子细胞生物学 4: 1354—1362），Moloney 鼠白血病病毒的长末端重复序列（MuLV LTR）（Weiss 等(1985) RNA 肿瘤病毒，冷泉港实验室，冷泉港，纽约），SV40 早期或晚期区启动子（Bernoist 等(1981) 自然 290: 304—310；Templeton 等(1984) 分子细胞生物学 4: 817；Sprague 等(1983) 病毒学杂志 45: 773），Rous 肉瘤病毒（RSV）的 3' 长末端重复序列中所含的启动子（Yamamoto 等，1980，细胞 22: 787—797），单纯疱疹病毒（HSV）胸苷激酶启动子/增强子（Wagner 等(1981) PNAS 82:3567-71），以及单纯疱疹病毒 LAT 启动子（Wolfe 等(1992) 自然遗传学 1: 379—384）。

在一举例性实施方案中，部分人 Shh 基因的 5' 侧翼区用加入限制性位点的引物扩增，产生下述片段：

5' -gcgcgcttcgaaGCGAGGCAGCCAGCGAGGGAGAGAGCGAGCG
GGCGAGCCGGAGCGAGGAAatcgatgcgcgc(引物 1)

5' -gcgcgcagatctGGGAAAGCGCAAGAGAGAGCGCACACGCAC

ACACCCGCCGCGCGCACTCGggatccgcgcgc (引物 2)

如所示出的, 引物 1 包括人 Shh 基因的 5' 非编码区并且侧翼为 AsuII 和 ClaI 限制性位点。引物 2 包括位于引物 1 中存在的 5' 非编码区序列 3' 的 5' 非编码区的一部分。hedgehog 基因序列的侧翼为 XhoII 和 BamHI 限制性位点。用每种合适的酶切割纯化的扩增物。

载体 pCDNA1.1(Invitrogen)包括一 CMV 启动子, 用在 CMV 启动子序列 3' 裂解的 AsuII 切割质粒, 将引物 1 的 AsuII/ClaI 片段与 pcDNA 载体的 AsuII 裂解位点连接。ClaI/AsuII 连接破坏了在正确插入的引物 1 的 3' 末端的 AsuII 位点。

然后用 BamHI 切割载体, 并将引物 2 的 XhoII/BamHI 片段连接到该 BamHI 连接位点。如上所述, BamHI/XhoII 连接破坏了正确插入的引物 2 的 5' 末端的 BamHI 位点。

选择各个菌落, 用 AsuII 和 BamHI 切割, 并测定 AsuII/BamHI 片段的大小。其中引物 1 和引物 2 序列均正确插入的菌落进一步扩增, 并用 AsuII 和 BamHI 切割产生如下基因活化构建体:

```
cgaagcgagggcagccagcgaggggagagagcgagcgggcgagccggagcgaggaaATCGAAGGT
TCGAATCCTTCCCCACCACCATCACTTTCAAAGTCCGAAAGAATCTGCTCCCTGCTTGTGT
GTTGGAGGTCGCTGAGTAGTGCGCGAGTAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGA
CAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAG
ATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGT
TCATAGCCCATATATGGAGTTCGCGGTTACATAACTTACGGTAAATGGCCCCGCTGGCTGACC
GCCCCAACGACCCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGG
GACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCCCCACTTGGCAGTACATCA
AGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCA
TTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT
CGCTATTACCATGGTGATGCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTC
ACGGGGATTTCCAAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCA
ACGGGACTTTCCAAAATGTCGTAACAACCTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGT
```

ACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCT
TATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTTGGTACCGAGCTCGGATCgatct
gggaaagcgcaagagagagcgcacacgcacacacccgccgcgcgcactcgg

在这一构建体中，侧翼的引物 1 和引物 2 序列提供了重组区，其允许 CMV 启动子插入人 *Shh* 基因的编码序列之前。其它异源启动子（或其它转录调控序列）可通过类似方法插入基因组 *hedgehog* 基因中。

在其它实施方案中，置换区仅仅缺失了天然基因的阴性转录控制元件以例如活化表达或者除去（ablate）阳性控制元件，以例如抑制靶基因的表达。

V. 举例性的 *ptc* 治疗化合物

在另一实施方案中，本发明方法用 *ptc* 治疗组合物进行。这种组合物可以用例如与 *patched* 结合并改变其信号转导活性的化合物、改变参与 *patched* 信号途径的蛋白（例如胞内蛋白）的结合和/或酶学活性的化合物、以及改变 *hedgehog* 蛋白、*patched* 蛋白或参与 *patched* 的胞内信号转导途径的蛋白的表达水平的化合物来产生。

纯化的和重组 *hedgehog* 多肽的可得性促进了可用于筛选药物如小有机分子的分析系统的产生，该药物是 *hedgehog* 和/或 *patched* 蛋白的正常细胞功能、特别是它们在水平细胞增殖和/或分化的病理中的作用的兴奋剂或拮抗剂。在一个实施方案中，所述分析评价化合物调节 *hedgehog* 多肽和 *hedgehog* 受体如 *patched* 之间结合的能力。在其它实施方案中，所述分析仅仅记录测试化合物改变 *patched* 蛋白的信号转导活性的能力。以这种方式，可鉴别各种增殖性活性和抗增殖性活性的 *hedgehog* 和/或 *ptc* 治疗剂。有各种分析形式可提供给本领域技术人员，并且根据本发明的教导其是可由本领域技术人员理解的。

在测试化合物库和天然提取物的许多药物筛选程序中，希望的是高通量分析以便使给定时间内分析的化合物数达到最大。在无细胞系统中进行的分析，如可用纯化的或半纯化的蛋白质衍生的分析经常是“初级”筛选所优选的，因为它们可允许快速开发及容易检测由测试化合物介导的分子靶中的改变。另外，测试化合物的细胞毒性和/或生物利用性的效果可通常在体外系统中忽视，分析主要集中于药物对分子靶的作用效果，其可由与受体蛋白的结合亲和性的改变而证实。

因此，在一个举例性的筛选 *ptc* 治疗剂的分析中，感兴趣的化合物与包括 *hedgehog* 受体蛋白（例如表达 *patched* 受体的细胞）和 *hedgehog* 蛋白的混合物在通常该化合物能结合 *hedgehog* 蛋白的条件下接触，然后向混合物中加入含有测试化合物的组合物。受体/*hedgehog* 复合物的检测和定量提供了一种确定测试化合物抑制（或增强）受体蛋白和 *hedgehog* 多肽之间形成复合物的功效的手段。化合物的功效可通过从用各种浓度测试化合物获得的数据产生剂量应答曲线来评价。另外，还可进行对照分析以提供基线用于比较。在对照分析中，将分离的纯化的 *hedgehog* 多肽加入到受体蛋白，并在缺少测试化合物的条件下定量受体/*hedgehog* 复合物的形成。

在其它实施方案中，本发明的 *ptc* 治疗剂是破坏 *patched* 与 *smoothened* 的结合的治疗剂。

通过随后用例如培养的上皮细胞进行测试可以区分上皮细胞生长的兴奋剂和拮抗剂并可以评价化合物的功效。

在一举例性实施方案中，用作 *hedgehog* 受体的多肽可从 *patched* 蛋白产生。因此，一举例性的筛选分析包括完整的或合适部分的 *patched* 蛋白，其可从例如，人 *patched* 基因（GenBank U43148）或其它脊椎动物来源（鸡 *patched* 基因，GenBank 登录号 U40074，小

鼠 *patched* 基因, U46155), 以及果蝇 (GenBank 登录号 M28999) 或其它无脊椎动物来源获得。在筛选分析中, *patched* 蛋白可以完整蛋白形式提供 (优选地在细胞表面表达), 或者作为与 *hedgehog* 多肽结合的全长蛋白的片段提供, 例如, 作为一个或两个基本上为胞外的结构域提供 (例如, 相应于人 *patched* 蛋白的残基 Asn120-Ser438 和/或 Arg770-Trp1027, 其也是 *hedgehog* 依赖性信号转导的潜在拮抗剂)。例如, *patched* 蛋白可以可溶形式提供, 例如作为一个胞外结构域的制备物, 或经非结构接头共价连接的两个胞外结构域的制备物 (参见, 例如 Huston 等(1988)PNAS85: 4879; 美国专利 5, 091, 513)。在其它实施方案中, 该蛋白可以脂质体制剂的一部分提供或在细胞表面表达。*patched* 蛋白可衍生自例如在异源细胞中异位表达的重组基因。例如, 该蛋白可通过标准重组 DNA 技术在卵细胞、哺乳动物细胞 (例如, COS, CHO, 3T3 等) 或酵母细胞上表达。这些重组细胞可用于受体结合、信号转导或基因表达分析。Marigo 等 (1996) 发育 122: 1225-1233 描述了人 *hedgehog* 与在 *Xenopus laevis* 卵细胞中异位表达的鸡 *patched* 蛋白的结合分析。Marigo 等的分析系统可适于本发明的药物筛选分析。如该文献所述, *Shh* 以一种选择性、饱和的、剂量依赖性方式与 *patched* 蛋白结合, 由此证实 *patched* 是 *Shh* 的受体。

hedgehog 多肽和 *hedgehog* 受体间的复合物形成可通过各种技术检测, 例如, 可用例如可检测标记的蛋白质如放射标记的、荧光标记的或酶标记的 *hedgehog* 多肽经免疫分析或经色谱检测而定量对复合物形成的调控。

典型地, 对于无细胞分析, 希望的是固定化 *hedgehog* 受体或 *hedgehog* 多肽以利于从非复合形式的一种蛋白中分离受体/ *hedgehog* 复合物, 以及适应分析的自动化。在一个实施方案中, 可提供一融

合蛋白，其添加了一能使蛋白与基质结合的结构域。例如，可将谷胱甘肽-S-转移酶/受体（GST/受体）融合蛋白吸附到谷胱甘肽琼脂糖珠（Sigma 化学公司，St.Louis, MO）或谷胱甘肽衍生的微滴板上，然后将其与 *hedgehog* 多肽，例如 ^{35}S -标记的 *hedgehog* 多肽及测试化合物混和，并在利于复合物形成的条件下保温，例如在生理条件的盐和 pH 下保温，当然稍严格的条件可能是希望的。保温后，洗涤珠以除去任何未结合的 *hedgehog* 多肽，直接测定基质珠结合的放射标记（例如，置于闪烁剂中的珠），或在受体/ *hedgehog* 复合物解离后在上清液中测定。或者，可从珠中解离复合物，经 SDS-PAGE 凝胶分离，用标准电泳技术从凝胶中定量在珠组分中发现的 *hedgehog* 多肽水平。

还有其它的将蛋白质固定在基质上的技术可用于本发明分析。例如，*hedgehog* 受体蛋白的可溶部分用生物素和链亲和素的缀合而固定化。例如，用本领域熟知的技术（例如，生物素化试剂盒，Pierce 化学公司，Rockford, IL）可从生物素-NHS（N-羟基琥珀酰亚胺）制备生物素化受体分子，并固定在链亲和素包被的 96 孔平板（Pierce 化学公司）的孔中。或者，与 *hedgehog* 受体反应但不干扰 *hedgehog* 结合的抗体可衍生化至平板的孔中，并经抗体缀合在孔中捕获受体。如上所述，在平板的存在受体孔中保温 *hedgehog* 多肽的制备物和测试化合物，并可以定量孔中捕获的受体/ *hedgehog* 复合物的量。检测这种复合物的方法除了上述检测 GST-固定化复合物的方法外，还包括用与 *hedgehog* 多肽反应的抗体，或与受体蛋白反应并竞争结合 *hedgehog* 多肽的抗体免疫检测复合物，以及基于检测与 *hedgehog* 多肽相关的酶活性的酶联分析。在后者情况下，酶可以是化学缀合的或以与 *hedgehog* 多肽的融合蛋白形式提供。举例来说，*hedgehog* 多肽可与碱性磷酸酶化学交联或遗传融合，可用该酶的生色底物如对

硝基苯磷酸评价复合物中捕获的 *hedgehog* 多肽的量。类似地，可提供包含 *hedgehog* 多肽和谷胱甘肽-S-转移酶的融合蛋白，并用 1-氯-2,4-二硝基苯检测 GST 活性来定量复合物形成 (Habig 等(1974) 生物化学杂志 249: 7130)。

对于依赖免疫检测定量复合物中捕获的一种蛋白的方法，可使用抗该蛋白的抗体，如本文描述的抗 *hedgehog* 抗体。或者，复合物中待检测的蛋白质可以融合蛋白的形式进行“表位附加”，该融合蛋白除包括 *hedgehog* 多肽或 *hedgehog* 受体序列外，还包括可容易获得其抗体的第二种多肽（例如，商购抗体）。例如，用抗 GST 基元的抗体，上述 GST 融合蛋白还可用于定量结合。其它有用的附加表位包括 myc-表位（例如，参见 Ellison 等(1991) 生物化学杂志 266: 21150—21157），其包括来自 c-myc 的 10 残基序列，以及 pFLAG 系统 (International Biotechnologies, Inc.) 或 pEZZ-蛋白 A 系统 (Pharmacia, NJ)。

当希望的 *hedgehog* 受体部分（或其它结合 *hedgehog* 的分子）不能以可溶形式提供时，可使用脂质体小泡以提供可操作的和可分离的受体来源。例如，可将真的和重组形式的 *patched* 蛋白重组在人工脂质体小泡（例如磷脂酰胆碱脂质体）或细胞膜衍生的小泡中（例如，参见 Bear 等(1992) 细胞 68: 809—818; Newton 等(1983) 生物化学 22: 6110—6117; Reber 等(1987) 生物化学杂志 262: 11369—11374)。

除了上述无细胞分析外，现有技术提供的很易得到的 *hedgehog* 蛋白的来源也有利于产生基于细胞的分析以鉴别小分子兴奋剂/拮抗剂等。与上述筛选组合文库的基于细胞的分析类似，可将对 *hedgehog* 诱导敏感的细胞，例如 *patched* 表达细胞或对 *hedgehog* 诱导敏感的其它上皮衍生细胞与 *hedgehog* 蛋白和感兴趣的测试试剂接触，分析中的评分为在存在和不存在测试试剂时与细胞的简单结合至靶细胞

对 *hedgehog* 诱导性应答的调控。与无细胞分析一样，可鉴别产生 *hedgehog* 活性的统计学显著变化（抑制或增强）的试剂。

在其它实施方案中，基于细胞的分析评价破坏例如细胞表面膜或脂质体制备物中的 *patched* 和 *smoothed* 蛋白的结合的试剂。

除了鉴别天然表达 *patched* 蛋白的细胞外，已遗传工程化异位表达 *patched* 的细胞还可用于药物筛选分析。例如，表达低水平或不表达 *patched* 蛋白的细胞，例如 *Xenopus laevis* 卵细胞，COS 细胞或酵母细胞可用标准技术遗传修饰以异位表达 *patched* 蛋白（见 Marigo 等，文献同上）。

得到的重组细胞，例如表达有功能的 *patched* 受体的细胞可用于受体结合分析中以鉴别 *hedgehog* 结合的兴奋剂或拮抗剂。结合分析可用完整细胞进行。另外，本发明的重组细胞可以工程化以包括编码参与 *hedgehog* 依赖性信号途径的蛋白的其它异源基因。例如，*smoothed*、*costal-2* 和/或 *fused* 中的一或多种的基因产物可在反应物细胞中与 *patched* 共表达，从而分析对 *hedgehog* 信号转导级联的有功能重建敏感。

或者，可应用使用重组的 *patched* 蛋白的脂质体制备物。从来自真的和重组来源的去污剂提取物纯化的 *patched* 蛋白可重组在人工脂质体小泡（例如磷脂酰胆碱脂质体）或细胞膜衍生的小泡中（例如，参见 Bear 等(1992) 细胞 68: 809—818; Newton 等(1983)生物化学 22: 6110—6117; Reber 等(1987)生物化学杂志 262: 11369—11374)。得到的脂质体的薄层结构和大小可用电子显微镜鉴别。重组的膜中的 *patched* 蛋白的外部取向可通过例如免疫电子显微镜术证实。在候选试剂存在下含有 *patched* 的脂质体和没有该蛋白的脂质体的 *hedgehog* 蛋白结合活性可进行比较以鉴别 *hedgehog*—*patched* 相互作用的潜在调控物。

用于这些基于细胞的分析中的 *hedgehog* 蛋白可以纯化的来源(天然或重组的来源)提供,或以表达该蛋白并与靶细胞共培养的细胞/组织形式提供。如在全细胞分析中一样,当在分析中记录到 *hedgehog* 活性仅仅是简单的结合(而非诱导)时,可用任何上述技术标记蛋白,例如荧光标记,酶标记或放射标记,或用免疫分析检测。

除了结合分析,可使用功能分析鉴别 *hedgehog* 或 *patched* 活性的调控物,即兴奋剂或拮抗剂。通过检测在与测试试剂接触的表达式 *patched* 的细胞胞内信号的变化,如第二信使或基因表达的改变,可以鉴别 *patched* 信号的候选兴奋剂和拮抗剂。

有许多基因产物参与 *patched* 介导的信号转导,包括 *patched*, 转录因子 *cubitus interruptus*(ci), 丝氨酸/苏氨酸激酶 *fused*(fu)以及 *costal-2*, *smoothened* 和 *fused* 的抑制基因的基因产物。

Hedgehog 蛋白与 *patched* 的相互作用激活一级联,其参与下游效应器的活化和抑制,在某些情况下,其最终结果是,在基因的转录或翻译中导致一可检测的变化。*Patched* 信号转导的潜在转录靶是 *patched* 基因本身(Hidalgo and Ingham, 1990 发育 110: 291—301; Marigo 等, 1996)以及果蝇 *cubitus interruptus* 基因的脊椎动物同系物 *GLI* 基因(Hui 等(1994)发育生物学 162: 402—413)。*patched* 基因表达已证明在应答 *Shh* 的肢芽和神经板细胞中被诱导(Marigo 等(1996)PNAS, 出版中; Marigo 等(1996)发育 122: 1225—1233)。*GLI* 基因编码具有锌指 DNA 结合结构域的推测的转录因子(Orenic 等(1990)基因和发育 4: 1053—1067; Kinzler 等(1990)分子细胞生物学 10: 634—642)。已报道 *GLI* 基因的转录在肢芽中应答 *hedgehog* 而正调,而 *GLI3* 基因的转录应答 *hedgehog* 诱导而下调(Marigo 等(1996)发育 122: 1225—1233)。通过从这种靶基因选择转录调控序列,例如从 *patched* 或 *GLI* 基因选择转录调控序列,这些调控序列负责应答

patched 信号转导而正调或下调这些基因，并将这种启动子与报道基因可操纵相连，则可衍生一基于转录的分析，其对特异测试化合物修饰 *patched* 信号转导途径的能力敏感。因此，报道基因的表达提供了开发作为分化/静止的 *ptc* 诱导的兴奋剂或拮抗剂的化合物的有价值的筛选工具。

本发明的基于报道基因的分析测量上述级联的最末期，例如转录调控。因此，在实施该分析的一个实施方案时，报道基因构建体被插入反应细胞 (reagent cell) 以产生依赖于 *ptc* 信号转导的检测信号。为鉴别靶基因的转录调控序列中存在的潜在的应答 *ptc* 信号转导的调控元件，可用标准技术构建靶基因的基因组克隆的嵌套缺失。例如，参考《分子生物学当前方案》，Ausubel, FM 等编辑，Greene 出版协会，(1989)；美国专利 5,266,488；Sato 等(1995) 生物化学杂志 270: 10314—10322；Kube 等 (1995) 细胞因子 7: 1—7。来自基因的 5'侧翼区的嵌套系列 DNA 片段被置于报道基因如萤光素酶基因的上游，并分析其在表达 *patched* 的细胞中指导报道基因表达的能力。可以记录在 *hedgehog* 存在和不存在下用报道基因构建体瞬时转染的宿主细胞中报道基因的诱导表达，以确定应答 *patched* 依赖性信号转导的调控系列。

在实施该分析的一个实施方案时，报道基因构建体插入到反应细胞以产生依赖于用 *hedgehog* 蛋白诱导产生的第二信使的检测信号。典型地，报道基因构建体包括与应答 *hedgehog* 活性的一或多个转录调控元件可操纵相连的报道基因，报道基因的表达水平提供了 *hedgehog* 依赖的检测信号。从报道基因转录的量可用任何本领域技术人员已知的合适方法测量。例如，来自报道基因的 mRNA 的表达可用 RNase 保护或基于 RNA 的 PCR 检测，或者报道基因的蛋白产物可用特征性染色或固有活性鉴别。报道基因的表达量然后与在不

存在测试化合物（或 *hedgehog*）时相同细胞的表达量比较，或者其可与缺少靶受体蛋白的基本上相同的细胞中的转录量比较。转录量的任何统计学或其它方式显著的差异表示测试化合物以某种方式改变 *patched* 蛋白的信号转导，例如测试化合物是潜在的 *ptc* 治疗剂。

如下所进一步详述的，在优选实施方案中，报道基因的基因产物通过与该产物相关的固有活性检测。例如，报道基因可编码一基因产物，该产物通过酶活性给出基于颜色、荧光或发光的检测信号。在其它优选的实施方案中，报道基因或标记基因提供了选择性生长的优势，例如，报道基因可增强细胞存活性，降低细胞营养要求，和/或提供对药物的抗性。

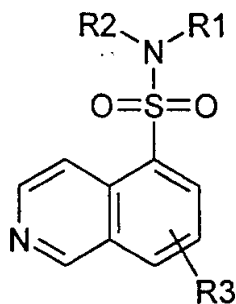
优选的报道基因是容易检测的报道基因。报道基因也可包括在构建体中，以与包括所需的转录调控系列或展示其它所需性质的基因形成融合基因的形式存在。报道基因的例子包括但不限于 CAT（氯霉素乙酰转移酶）（Alton and Vapnek(1979)，自然 282：864—869），萤光素酶和其它酶检测系统，如 β -半乳糖苷酶；萤火虫萤光素酶（deWet 等(1987) 分子细胞生物学 7：725—737）；细菌萤光素酶（Engbrecht and Silverman(1984)，PNAS1：4154—4158；Baldwin 等(1984)，生物化学 23：3663—3667）；碱性磷酸酶（Toh 等(1989)欧洲生物化学杂志 182：231—238，Hall 等(1983)分子应用遗传学杂志 2：101），人胎盘分泌的碱性磷酸酶（Cullen and Malim (1992)酶学方法 216：362—368）。

可包括在报道基因构建体中的转录控制元件包括但不限于启动子，增强子和阻遏蛋白和活化蛋白结合位点。合适的转录调控元件可衍生自在调控 *patched* 信号转导途径后其表达被诱导的基因的转录调控区。衍生转录控制元件的优选的基因的特征包括但不限于在静止细胞中表达很低或检测不到，在胞外模拟几分钟内快速诱导转录

水平，诱导是瞬时的且不依赖于新蛋白质的合成，随后转录的关闭需要新的蛋白质合成，从这些基因转录的 mRNA 具有短的半寿期。不需要所有这些性质均存在。

在其它实施方案中，第二信使的产生可直接在检测步骤测量，如定量胞内钙的移动，磷脂代谢或腺苷酸环化酶活性，例如，可测量磷脂水解的产物 IP_3 ，DAG 或 cAMP。例如，最近的研究表面蛋白激酶 A (PKA) 是 *hedgehog/patched* 信号途径的一个可能的组分 (Hammerschmidt 等(1996) 基因和发育 10: 647)。已表明高 PKA 活性在这些系统中拮抗 *hedgehog* 信号转导。尽管还不清楚 PKA 是直接作用于下游还是与 *hedgehog* 信号转导平行作用，但可能的是 *hedgehog* 信号转导通过抑制 PKA 活性而发生。因此，检测 PKA 活性提供了本分析的潜在的读出器。

在一个优选的实施方案中，*ptc* 治疗剂是 PKA 抑制剂。现有技术中已知各种 PKA 抑制剂，包括肽化合物和有机化合物。例如，*ptc* 治疗剂可以是 5-异喹啉磺酰胺，如由下述通式所代表的：



其中，

R1 和 R2 各自独立地代表氢，并且若化合价和稳定性允许则代表低级烷基，低级链烯基，低级炔基，羰基（如羧基、酯、甲酸酯或酮），硫代羰基（如硫酯、硫代乙酸酯或硫代甲酸酯），氨基，酰氨基，酰氨基，氰基，硝基，叠氮基，硫酸酯，磺酸酯，亚磺酰氨基， $-(CH_2)_m-R_8$ ， $-(CH_2)_m-OH$ ， $-(CH_2)_m-O$ -低级烷基， $-(CH_2)_m-O$ -低级

链烯基, $-(\text{CH}_2)_n\text{-O-(CH}_2)_m\text{-R}_8$, $-(\text{CH}_2)_m\text{-SH}$, $-(\text{CH}_2)_m\text{-S-低级烷基}$, $-(\text{CH}_2)_m\text{-S-低级链烯基}$, $-(\text{CH}_2)_n\text{-S-(CH}_2)_m\text{-R}_8$, 或者

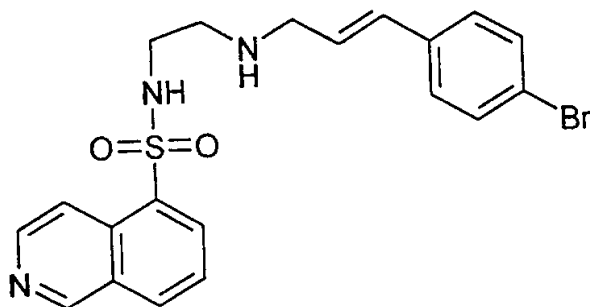
R1 和 R2 一起与 N 形成一杂环 (取代的或未取代的);

R3 不存在或代表异喹啉环的一或多个取代, 如低级烷基, 低级链烯基, 低级炔基, 羰基 (如羧基、酯、甲酸酯或酮), 硫代羰基 (如硫酯、硫代乙酸酯或硫代甲酸酯), 氨基, 酰基氨基, 酰氨基, 氰基, 硝基, 叠氮基, 硫酸酯, 磺酸酯, 亚磺酰氨基, $-(\text{CH}_2)_m\text{-R}_8$, $-(\text{CH}_2)_m\text{-OH}$, $-(\text{CH}_2)_m\text{-O-低级烷基}$, $-(\text{CH}_2)_m\text{-O-低级链烯基}$, $-(\text{CH}_2)_n\text{-O-(CH}_2)_m\text{-R}_8$, $-(\text{CH}_2)_m\text{-SH}$, $-(\text{CH}_2)_m\text{-S-低级烷基}$, $-(\text{CH}_2)_m\text{-S-低级链烯基}$, $-(\text{CH}_2)_n\text{-S-(CH}_2)_m\text{-R}_8$;

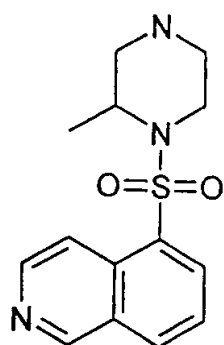
R_8 代表取代的或未取代的芳基, 芳烷基, 环烷基, 环烯基或杂环; 以及

n 和 m 各自独立为 0 或 1—6 的整数。

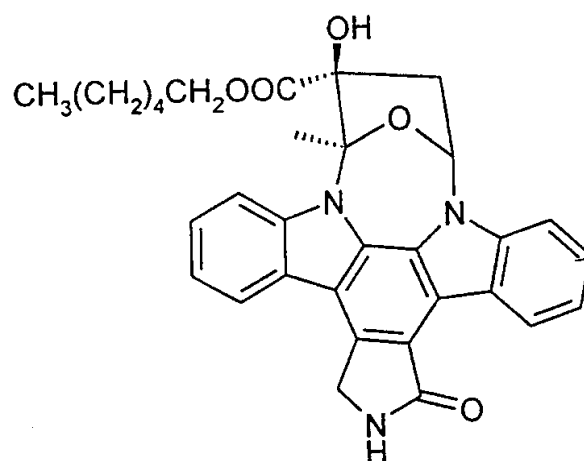
在一个优选的实施方案中, PKA 抑制剂是 N-[2-((p-溴肉桂基)氨基)乙基]-5-异喹啉磺酰胺 (H-89; Calbiochem, 货号 371963), 例如, 具有下式:



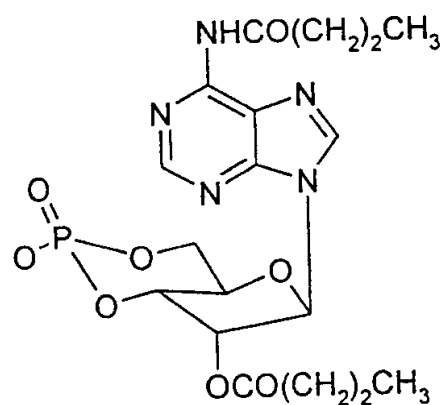
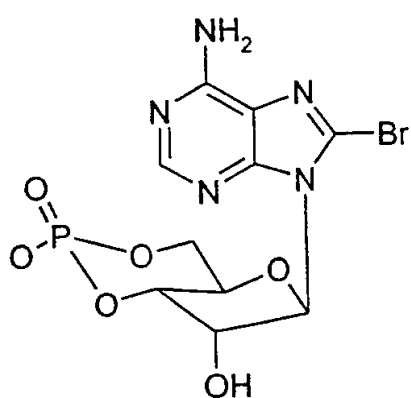
在另一个实施方案中, PKA 抑制剂是 1-(5-异喹啉磺酰基)-2-甲基哌嗪 (H-7; Calbiochem, 货号 371955), 例如, 具有下式:



在其它实施方案中，PKA 抑制剂是 KT5720 (Calbiochem, 货号 420315)，具有下述结构：



各种核苷类似物也可用作 PKA 抑制剂，例如本发明方法可用抑制 PKA 激酶活性的环化 AMP 类似物进行，例如，8-溴-cAMP 或二丁酰-cAMP



PKA 活性的肽抑制剂的例子包括 PKA 热稳定抑制剂（同工型 α ；例如参见 Calbiochem, 货号 539488; Wen 等(1995) 生物化学杂志 270: 2041)。

某些 *hedgehog* 受体可刺激磷脂酶活性。肌醇脂可用标准脂类提取技术提取并分析。所有三种肌醇脂 (IP_1 , IP_2 , IP_3) 的可溶性衍生物也可用放射标记技术或 HPLC 定量。

胞内钙的移动或来自细胞外的钙内向通量可以是对 *hedgehog* 刺激或缺乏的应答。反应细胞中的钙内向通量可用标准技术测量。合适的钙指示剂，荧光，生物发光，金属显色或钙离子敏感的微电极的选择取决于细胞类型以及所研究的事件的量级和时间常数 (Borle (1990) 环境健康预测 84: 45—56)。作为钙离子检测的举例方法，可用标准技术将细胞加载钙离子敏感的荧光染料 fura-2 或 indo-1，并用荧光剂测量钙离子的任何变化。

在该分析的某些实施方案中，可能希望的是筛选细胞磷酸化的变化。例如，编码丝氨酸/苏氨酸激酶的果蝇基因 *fused* (fu) 已被鉴别为 *hedgehog* 信号转导中潜在的下游靶 (Preat 等, 1990, 自然 347, 87—89; Therond 等, 1993, Mech. Dev. 44, 65-80)。化合物调控丝氨酸/苏氨酸激酶活化的能力可用抗磷酸化丝氨酸或苏氨酸残基的抗体经集落免疫印迹 (Lyons and Nelson(1984) 美国科学院院刊 81: 7426—7430) 筛选。进行这种分析的试剂是可商购的，例如测量丝氨酸和苏氨酸残基磷酸化增加的磷酸丝氨酸和磷酸苏氨酸特异性抗体可商购。

在另一个实施方案中，*ptc* 治疗剂是一种反义分子，其抑制参与 *patched* 介导的信号转导途径的蛋白质的表达。举例来说，通过抑制参与 *patched* 信号的蛋白质如 *fused*, *costal-2*, *smoothened* 和/或 Gli 基

因的表达, *patched* 信号途径抑制细胞增殖的能力可以被改变, 例如增强或阻遏。

如本文所用, “反义”治疗是指给予或原位产生在细胞条件下与编码 hedgehog 蛋白, *patched* 或参与 *patched* 介导的信号转导的蛋白的细胞 mRNA 和/或基因组 DNA 特异地杂交 (例如结合) 的寡核苷酸探针或其衍生物。杂交应例如通过抑制转录和/或翻译而抑制该蛋白的表达。结合应通过常规的碱基对互补, 或者, 例如在与 DNA 双螺旋结合的情况下, 通过在双螺旋的大沟中的特异相互作用进行。通常, “反义”治疗指现有技术中通常采用的技术范围, 并包括基于与寡核苷酸序列的特异结合的任何疗法。

本发明的反义构建体可以例如作为表达质粒而输送, 其在细胞中转录时, 产生与靶细胞 mRNA 的至少一独特部分互补的 RNA。或者, 反义构建体是一寡核苷酸探针, 其来自体内产生, 并且当导入细胞时, 通过与靶基因的 mRNA 和/或基因组序列的杂交导致表达的抑制。这种寡核苷酸探针优选地是修饰的寡核苷酸, 其对内源性核酸酶例如外切核酸酶和/或内切核酸酶有抗性, 并且因此是体内稳定的。用作反义寡核苷酸的核酸分子的例子是 DNA 的磷酰胺、硫代磷酸和甲基磷酸类似物 (也参见美国专利 5, 176, 996; 5, 264, 564 和 5, 256, 775)。另外, 构建用于反义治疗的寡聚物的通用方案已由例如 Van der Krol 等 (1988) 生物技术 6: 958—976; Stein 等 (1988) 癌症研究 48: 2659—2668 综述。

当构建反义寡核苷酸用于本发明方法时应考虑: (1) 寡聚物应具有 50% 或更高的 GC 含量; (2) 避免具有 3 个或更多个 G 的序列段; (3) 寡核苷酸应不超过 25—26 聚体。当测试反义寡核苷酸时, 应构建错配的对照。对照可通过反转相应的反义寡核苷酸的序列顺序而产生以保持相同的碱基比例。

在一举例性实施方案中，*ptc* 治疗剂可以是用于抑制 *patched* 表达的反义寡核苷酸，例如以模拟 *hedgehog* 对 *patched* 的抑制。举例性的反义构建体包括：

5'-GTCCTGGCGCCGCCGCCGCCGTCGCC

5'-TTCCGATGACCGGCCTTTCGCGGTGA

5'-GTGCACGGAAAGGTGCAGGCCACACT

VI. 举例性的 *hedgehog* 和 *ptc* 治疗剂的药物制剂

待配制的 *hedgehog* 和 *ptc* 治疗剂的来源将取决于试剂的特定形式。小有机分子和肽片段可以化学合成并以纯的形式提供以适合药用/美容用。天然提取物的产物可根据现有技术已知的技术提纯。例如，Cox 等的美国专利 5, 286, 654 描述了提纯天然形式的分泌蛋白的方法并可用于纯化 *hedgehog* 多肽。重组来源的 *hedgehog* 多肽也是可利用的。例如，从 PCT 出版物 WO95/18856 和 WO96/17924 特别地已知编码 *hedgehog* 多肽的基因。

治疗上皮组织的技术人员可以确定配制在药物或美容制剂中的 *hedgehog* 或 *ptc* 治疗剂的有效量。

用于本发明方法的 *hedgehog* 或 *ptc* 治疗剂配制品最优选地以合适的组合物形式施用。合适的组合物可以提及通常用来系统或局部施用药物的所有组合物。药物学可接受的载体应是基本上惰性的，从而不与活性成分作用。合适的惰性载体包括水、醇聚乙二醇、矿物油或石油凝胶，丙二醇等。

为制备本发明的药物组合物，将作为活性成分的有效量的特定 *hedgehog* 或 *ptc* 治疗剂与药物学可接受的载体混和，该载体根据给药所需的制剂形式可采用各种形式。希望的是这些药物组合物呈适合特别是口服、直肠、经皮或非肠道注射的给药方式的单一剂量形式。例如，在制备口服剂型的组合物时，可采用任何常用的药物介质，

如在口服液体制剂如悬液，糖浆，酏剂和溶液情况下为水，乙二醇，油，醇等；在粉剂，丸剂，胶囊剂和片剂情况下为固体载体如淀粉，蔗糖，高岭土，润滑剂，粘合剂，崩解剂等。由于其易于给药，片剂和胶囊剂为最有利的口服剂型，在此情况下显然采用固体药物载体。对于非肠道组合物，载体通常包括无菌水，无菌水至少占大部分，当然其它成分，例如有助于溶解性的成分也可包括。可制备可注射溶液，例如其中的载体包括盐水溶液，葡萄糖溶液或盐水和葡萄糖溶液的混合物。也可制备可注射悬液，其中可采用合适的液体载体，悬浮剂等。还包括固体形式的制剂，其在即将使用前可被转变成液体形式的制剂。在适于经皮施用的组合物中，载体任选地包括穿透促进剂和/或合适的润湿剂，其任选地与少量的任意性质的合适添加剂混和，该添加剂不会引起对皮肤的显著破坏作用。

除了制剂的直接局部施用外，它们可经其它方法局部施用，例如，包囊在温度和/或压力敏感基质中或包囊在可溶于体液的薄膜或固体载体中等以随后释放，优选地缓释活性成分。

用于局部施用的合适的组合物可包括所有通常用于局部施用治疗剂的组合物，例如，霜剂，啫哩（gellies），敷料，香波，酏剂，糊剂，软膏剂，油膏剂，粉剂，液体或半液体配制品等。所述组合物的施用可通过带有或不带有推进剂如氮气，二氧化碳，氟利昂的气雾剂如泵式喷雾器，滴剂，洗剂，或半固体如增稠的组合物，其可用拭子施用。在特定的组合物中，半固体组合物如油膏剂，霜剂，糊剂，啫哩，软膏剂等可方便地使用。

特别有利的是将本发明组合物配制成易于给药并且剂量单一的剂量单位形式。本文说明书和权利要求书中所用的剂量单位形式是指适于作为单元剂量的物理不连续的单位，每个单位含有经计算能产生所需治疗效果的预定量的活性成分及所需的药物载体。这种剂

量单位形式的例子是片剂（包括刻痕或包衣片剂），胶囊剂，丸剂，粉末包，糯米纸囊剂，注射溶液或悬液，一茶匙容量，一大汤匙容量等，及其分离的多份。

如上所述，本发明的药物制剂可用于被认为是美容用途的许多应用中。优选地低致敏性并 pH 控制的现有技术中已知的美容组合物是特别优选的，包括花露水，润肤膏，洗剂，乳液或乳状洗剂。该制剂除了 hedgehog 或 ptc 治疗剂外还含有这种制剂中通常采用的组分。这种组分的例子为油，脂肪，蜡，表面活性剂，湿润剂，增稠剂，抗氧化剂，粘度稳定剂，螯合剂，缓冲剂，防腐剂，香水，染料，低级烷醇等。如果需要，其它成分也可掺入组合物中，例如抗炎剂，抗细菌剂，抗真菌剂，消毒剂，维生素，遮光剂，抗生素或其它抗痤疮剂。

油的例子包括脂肪和如橄榄油和氢化油的油类；蜡如蜂蜡和羊毛脂；碳氢化合物如液体石蜡，地蜡，角鲨烷；脂肪酸如硬脂酸和油酸；醇如鲸蜡醇，十八烷醇，羊毛脂醇和十六烷醇；酯如异丙基肉豆蔻酸酯，异丙基棕榈酸酯和丁基硬脂酸酯。表面活性剂的例子包括阴离子表面活性剂如硬脂酸钠，十六烷基硫酸酯钠，聚氧乙烯十二烷基醚磷酸酯，N-酰基谷氨酸钠；阳离子表面活性剂如硬脂酰二甲基苄基氯化铵和硬脂酰三甲基氯化铵；两性表面活性剂如烷基氨基乙基甘氨酸盐酸溶液和卵磷脂；非离子表面活性剂如甘油单硬脂酸酯，山梨聚糖单硬脂酸酯，蔗糖脂肪酸酯，丙二醇单硬脂酸酯，聚氧乙烯油基醚，聚乙二醇单硬脂酸酯，聚氧乙烯山梨聚糖单棕榈酸酯，聚氧乙烯椰子油脂肪酸单乙醇酰胺，聚氧丙二醇（例如，商标为“Pluronic”的物质），聚氧乙烯海狸香油，聚氧乙烯羊毛脂。湿润剂的例子包括甘油，1，3-丁二醇和丙二醇；低级醇的例子包括乙醇和异丙醇；增稠剂的例子包括黄原胶，羟丙基纤维素，羟丙基

甲基纤维素，聚乙二醇和羧甲基纤维素钠；抗氧化剂的例子包括丁化羟基甲苯，丁基化羟基苯甲醚，没食子酸丙酯，柠檬酸和乙氧基喹；螯合剂的例子包括乙二胺四乙酸二钠和乙羟基二磷酸盐；缓冲剂的例子包括柠檬酸，柠檬酸钠，硼酸，硼沙和磷酸氢二钠；防腐剂的例子包括甲基对羟基苯甲酸酯，乙基对羟基苯甲酸酯，脱氢乙酸，水杨酸和苯甲酸。

为制备软膏剂，霜剂，花露水，乳液等，典型地将 0.01-10%、特别是 0.1-5%、更特别是 0.2-2.5% 活性成分例如 hedgehog 或 ptc 治疗剂掺入组合物中。在软膏剂或霜剂中，载体例如由 1—20%、特别是 5—15% 湿润剂，0.1-10%、特别是 0.5-5% 增稠剂和水组成；或者所述的载体可由 70-99%、特别是 20-95% 表面活性剂和 0—20%、特别是 2.5-15% 脂肪组织；或者由 80—99.9%、特别是 90—99% 增稠剂组成；或者由 5—15% 表面活性剂，1—25% 润湿剂，0—80% 油，非常少量 (<2%) 防腐剂，着色剂和/或香水，和水组成。在花露水中，载体例如由 2—10% 低级醇，0.1-10% 或特别是 0.5-1% 表面活性剂，1—20%，特别是 3—7% 润湿剂，0—5% 缓冲剂，水和非常少量 (<2%) 防腐剂，染料和/或香水组成。在乳液中，载体典型地由 10—50% 油，1—10% 表面活性剂，50—80% 水和 0—3% 防腐剂和/或香水组成。在上述制剂中，所有百分数均为重量百分比。

用于本发明方法中的特别的组合物是其中的 hedgehog 或 ptc 治疗剂配制在含有脂质体的组合物中的那些组合物。脂质体是由两亲性分子如极性脂形成的人工小泡，所述极性脂例如磷脂酰胆碱，乙醇胺和丝氨酸，鞘磷脂，心磷脂，缩醛磷脂，磷脂酸和脑苷脂。当将合适的两性性分子在水或水溶液中溶胀以形成通常有由经水性物质分开的许多双层组成的多层结构的液晶（通常成为粗脂质体）时，脂质体会形成。已知由包裹水性物质的单一双层组成的另一类型的

脂质体被称为单层小泡。如果在脂类的溶胀过程中水溶性物质包括在水相中，则它们被捕获在脂双层之间的水层中。

水溶性活性成分如各种盐形式的 hedgehog 多肽被包裹在分子层之间的水性空间内。Hedgehog 或 ptc 治疗剂的脂溶性活性成分如有机模拟物主要掺入脂质层中，但是极性头部基团从该层伸出进入水性空间。这些化合物的胶囊化可通过各种方法实现，最常用的方法涉及通过蒸发有机溶剂将磷脂薄膜铸型至烧瓶壁上，当将此薄膜在合适的水性介质中分散时，形成多层脂质体。经过合适的超声处理，粗脂质体形成较小的类似闭合的小泡。

水溶性活性成分通常通过用化合物的水溶液分散该铸型薄膜二掺入。然后经离心、层析、透析或其它现有技术已知的合适程序除去未胶囊化的化合物。脂溶性活性成分通常通过在铸型薄膜之前将其溶解于具有磷脂的有机溶剂中而掺入。如果脂相中的物质的溶解度或存在量没有超过可与脂类结合的量，上述方法制备的脂质体通常含有结合在脂双层中的大多数物质；不需分离脂质体和未胶囊化的物质。

制备脂质体配制形式的 hedgehog 和 ptc 治疗剂的特别方便的方法是 EP-A-253, 619 中描述的方法，该专利申请引入本文作参考。在该方法中，含有胶囊化的活性成分的单双层脂质体的制备方法为，将脂类成分溶解在有机介质中，将脂类成分的有机溶液在压力下注射进水性成分，同时用高速匀浆机或混和设备混和有机和水性成分，由此自发形成脂质体。

含有胶囊化的 hedgehog 或 ptc 治疗剂的单双层脂质体可直接采用，或者其可以与合适的药物学可接受载体一起用于局部给药。通过添加一或多种合适的增稠剂如黄原胶，羟丙基纤维素，羟丙基甲基纤维素及其混合物，可以增加脂质体的粘度。水性成分可仅由水

组成，或者其可含有电解质，缓冲系统和其它成分如防腐剂。可采用的合适的电解质包括金属盐如碱金属和碱土金属盐。优选的金属盐是氯化钙，氯化钠和氯化钾。电解质的浓度范围为 0—260mM，优选 5mM—160mM。水性成分置于合适的容器中，该容器可适于经在注射有机成分过程中产生大湍流而实现均质化。两种成分的均质化可在该容器内完成，或者，水性成分和有机成分可分别注射进置于容器外的混和设备中。在后一情况下，在混和设备中形成脂质体，然后被转移至另一容器中收集。

有机成分由合适的非毒性药物学可接受的溶剂以及溶于该溶剂中的合适的磷脂组成，所述溶剂例如为乙醇，甘油，丙二醇和聚乙二醇。可采用的合适的磷脂例如包括卵磷脂，磷脂酰胆碱，磷脂酰丝氨酸，磷脂酰乙醇胺，磷脂酰肌醇，溶血磷脂酰胆碱和磷脂酰甘油。还可采用其它亲脂性添加剂以选择性修饰脂质体的性质。这种其它添加剂的例子包括硬脂酰胺，磷脂酸，生育酚，胆固醇和羊毛脂提取物。

另外，可防止磷脂氧化的其它成分也可加入到有机成分中，这种其它成分的例子包括生育酚，丁基化羟基苯甲醚，丁化羟基甲苯，棕榈酸抗坏血酸酯和油酸抗坏血酸酯。也可加入防腐剂如苯甲酸，甲基对羟基苯甲酸酯和丙基对羟基苯甲酸酯。

除了上述组合物外，还可利用含有合适量的 hedgehog 或 ptc 治疗剂的包裹物，例如膏药，绷带，敷料，纱布垫等。在某些情况下可利用已用含有治疗剂配制品的局部用配制品浸渍的膏药，绷带，敷料，纱布垫等。

举例

本发明已一般性地描述，通过参照下述实施例其可更容易理解，实施例仅是举例说明本发明的某些方面和实施例，而非限制本发明。

Hedgehog 蛋白的纯化

在杆状病毒/昆虫细胞系统 (Roelink 等(1995) 细胞 81: 445—455) 中表达人 sonic hedgehog 蛋白 (残基 24—197)。将条件培养基加至用 50mM 磷酸钾, 0.5mM DTT, pH7.0 的 Fast Flo SP 琼脂糖上。用这一缓冲液洗柱, 然后用 10M NaCl 梯度洗脱。分析组分在间充质干细胞 (C3H10T1/2 细胞, 例如参见 Wang 等(1993) 生长因子 9: 57—71) 上对碱性磷酸酶活性的诱导, 然后基于这一活性合并组分并在 SDS 凝胶上纯化。在 Amicon 超滤装置 (PM10 膜) 上浓缩合并的物质并对 10mM Tris, pH7.4, 0.5mM DTT 透析。用 γ -球蛋白作为标准经 Bradford 方法估计蛋白质的量。

胶原海绵的制备

用 MilliQ 水仔细洗涤胶原海绵以除去来自厂商的任何表面活性剂和添加剂。然后用 70% 乙醇洗海绵, 真空干燥。

移植物的制备

将蛋白质加入到 1.0-1.5mm $\times$ 8-10mm 的胶原海绵片 (重量 1.5-3.0mg) 中。在某些情况下在将 hedgehog 蛋白加入胶原海绵之前加入硫酸锌至终浓度为 0.2mM。然后冻干该重组的海绵。

移植

将海绵植入 Sprague Dawley 大鼠 (4—8 周龄) 的胸部皮下或兔 (11—14 周龄) 的股骨肌肉中。饲养动物 2—5 周, 之后取出移植物。然后将移植物在 4% 福尔马林或 4% 低聚甲醛中固定然后包埋在 JB-4 树脂中。用甲苯胺蓝染色 (Wang 等(1988) PNAS85: 9484—9488)。

新毛囊、皮脂腺和其它皮结构的诱导经其明显的形态学鉴别。我们将移植物仔细地植入皮下或肌肉内并将这些移植物小心地取出排除了现存皮结构的污染的可能性。另外, 在较短 (2 周) 期间在移植物中可见不太成熟的毛囊。

从在 3 周时从兔肌肉中取出的肌内移植物获得活检切片并用苏木精和曙红染色。类似检测了用甲苯胺蓝染色的 3 周兔肌肉肌内移植物的切片。某些样品的切片揭示了组织形态学，表明存在在肌内组织形成的毛囊和毛发类结构。

Shh 对毛发的诱导

作为上述实验的继续，将载有 hedgehog 的胶原海绵植入小鼠剃毛的皮下。如图 1A—C 所示，hedgehog 制剂可在移植物上诱导毛发生长。另外，具有 Von Willebrand 因子胶原结合位点的在 C 末端修饰的 Ihh 蛋白在毛发生长中有活性，表明移植的蛋白的局部诱导活性。

所有上述参考文献和出版物均引入本文作参考。

等价物

本领域技术人员不经创造性劳动能理解或能确定本文所述的特异多肽、核酸、方法、分析和试剂的各种等价物。这种等价物被认为属于本发明范围。

序列表

(2) SEQ ID NO:1 的信息:

(i) 序列特征

(A) 长度: 1277碱基对

(B) 类型: 核酸

(C) 链性: 双链

(D) 拓扑学: 线性

(ii) 分子类型: cDNA

(ix) 特征:

(A) 名称/关键词: CDS

(B) 位置: 1..1275

(xi) 序列描述: SEQ ID NO:1:

ATG GTC GAA ATG CTG CTG TTG ACA AGA ATT CTC TTG GTG GGC TTC ATC	48
Met Val Glu Met Leu Leu Leu Thr Arg Ile Leu Leu Val Gly Phe Ile	
1 5 10 15	
TGC GCT CTT TTA GTC TCC TCT GGG CTG ACT TGT GGA CCA GGC AGG GGC	96
Cys Ala Leu Leu Val Ser Ser Gly Leu Thr Cys Gly Pro Gly Arg Gly	
20 25 30	
ATT GGA AAA AGG AGG CAC CCC AAA AAG CTG ACC CCG TTA GCC TAT AAG	144
Ile Gly Lys Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys	
35 40 45	
CAG TTT ATT CCC AAT GTG GCA GAG AAG ACC CTA GGG GCC AGT GGA AGA	192
Gln Phe Ile Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg	
50 55 60	
TAT GAA GGG AAG ATC ACA AGA AAC TCC GAG AGA TTT AAA GAA CTA ACC	240
Tyr Glu Gly Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr	
65 70 75 80	
CCA AAT TAC AAC CCT GAC ATT ATT TTT AAG GAT GAA GAG AAC ACG GGA	288
Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly	
85 90 95	
GCT GAC AGA CTG ATG ACT CAG CGC TGC AAG GAC AAG CTG AAT GCC CTG	336
Ala Asp Arg Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu	
100 105 110	
GCG ATC TCG GTG ATG AAC CAG TGG CCC GGG GTG AAG CTG CGG GTG ACC	384
Ala Ile Ser Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg Val Thr	
115 120 125	
GAG GGC TGG GAC GAG GAT GGC CAT CAC TCC GAG GAA TCG CTG CAC TAC	432
Glu Gly Trp Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr	
130 135 140	
GAG GGT CGC GCC GTG GAC ATC ACC ACG TCG GAT CGG GAC CGC AGC AAG	480
Glu Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Ser Lys	
145 150 155 160	
TAC GGA ATG CTG GCC CGC CTC GCC GTC GAG GCC GGC TTC GAC TGG GTC	528

00.06.15

Tyr	Gly	Met	Leu	Ala	Arg	Leu	Ala	Val	Glu	Ala	Gly	Phe	Asp	Trp	Val		
				165					170					175			
TAC	TAC	GAG	TCC	AAG	GCG	CAC	ATC	CAC	TGC	TCC	GTC	AAA	GCA	GAA	AAC	576	
Tyr	Tyr	Glu	Ser	Lys	Ala	His	Ile	His	Cys	Ser	Val	Lys	Ala	Glu	Asn		
			180					185					190				
TCA	GTG	GCA	GCG	AAA	TCA	GGA	GGC	TGC	TTC	CCT	GGC	TCA	GCC	ACA	GTG	624	
Ser	Val	Ala	Ala	Lys	Ser	Gly	Gly	Cys	Phe	Pro	Gly	Ser	Ala	Thr	Val		
			195				200					205					
CAC	CTG	GAG	CAT	GGA	GGC	ACC	AAG	CTG	GTG	AAG	GAC	CTG	AGC	CCT	GGG	672	
His	Leu	Glu	His	Gly	Gly	Thr	Lys	Leu	Val	Lys	Asp	Leu	Ser	Pro	Gly		
	210					215					220						
GAC	CGC	GTG	CTG	GCT	GCT	GAC	GCG	GAC	GGC	CGG	CTG	CTC	TAC	AGT	GAC	720	
Asp	Arg	Val	Leu	Ala	Ala	Asp	Ala	Asp	Gly	Arg	Leu	Leu	Tyr	Ser	Asp		
					230				235						240		
TTC	CTC	ACC	TTC	CTC	GAC	CGG	ATG	GAC	AGC	TCC	CGA	AAG	CTC	TTC	TAC	768	
Phe	Leu	Thr	Phe	Leu	Asp	Arg	Met	Asp	Ser	Ser	Arg	Lys	Leu	Phe	Tyr		
				245				250						255			
GTC	ATC	GAG	ACG	CGG	CAG	CCC	CGG	GCC	CGG	CTG	CTA	CTG	ACG	GCG	GCC	816	
Val	Ile	Glu	Thr	Arg	Gln	Pro	Arg	Ala	Arg	Leu	Leu	Leu	Thr	Ala	Ala		
				260				265					270				
CAC	CTG	CTC	TTT	GTG	GCC	CCC	CAG	CAC	AAC	CAG	TCG	GAG	GCC	ACA	GGG	864	
His	Leu	Leu	Phe	Val	Ala	Pro	Gln	His	Asn	Gln	Ser	Glu	Ala	Thr	Gly		
			275				280					285					
TCC	ACC	AGT	GGC	CAG	GCG	CTC	TTC	GCC	AGC	AAC	GTG	AAG	CCT	GGC	CAA	912	
Ser	Thr	Ser	Gly	Gln	Ala	Leu	Phe	Ala	Ser	Asn	Val	Lys	Pro	Gly	Gln		
			290			295					300						
CGT	GTC	TAT	GTG	CTG	GGC	GAG	GGC	GGG	CAG	CAG	CTG	CTG	CCG	GCG	TCT	960	
Arg	Val	Tyr	Val	Leu	Gly	Glu	Gly	Gly	Gln	Gln	Leu	Leu	Pro	Ala	Ser		
					310				315						320		
GTC	CAC	AGC	GTC	TCA	TTG	CGG	GAG	GAG	GCG	TCC	GGA	GCC	TAC	GCC	CCA	1008	
Val	His	Ser	Val	Ser	Leu	Arg	Glu	Glu	Ala	Ser	Gly	Ala	Tyr	Ala	Pro		
				325					330					335			
CTC	ACC	GCC	CAG	GGC	ACC	ATC	CTC	ATC	AAC	CGG	GTG	TTG	GCC	TCC	TGC	1056	
Leu	Thr	Ala	Gln	Gly	Thr	Ile	Leu	Ile	Asn	Arg	Val	Leu	Ala	Ser	Cys		
				340				345					350				
TAC	GCC	GTC	ATC	GAG	GAG	CAC	AGT	TGG	GCC	CAT	TGG	GCC	TTC	GCA	CCA	1104	
Tyr	Ala	Val	Ile	Glu	Glu	His	Ser	Trp	Ala	His	Trp	Ala	Phe	Ala	Pro		
				355			360					365					
TTC	CGC	TTG	GCT	CAG	GGG	CTG	CTG	GCC	GCC	CTC	TGC	CCA	GAT	GGG	GCC	1152	
Phe	Arg	Leu	Ala	Gln	Gly	Leu	Leu	Ala	Ala	Leu	Cys	Pro	Asp	Gly	Ala		
						375					380						
ATC	CCT	ACT	GCC	GCC	ACC	ACC	ACC	ACT	GGC	ATC	CAT	TGG	TAC	TCA	CGG	1200	
Ile	Pro	Thr	Ala	Ala	Thr	Thr	Thr	Thr	Gly	Ile	His	Trp	Tyr	Ser	Arg		
						390				395					400		
CTC	CTC	TAC	CGC	ATC	GGC	AGC	TGG	GTG	CTG	GAT	GGT	GAC	GCG	CTG	CAT	1248	
Leu	Leu	Tyr	Arg	Ile	Gly	Ser	Trp	Val	Leu	Asp	Gly	Asp	Ala	Leu	His		
				405					410					415			

CCG CTG GGC ATG GTG GCA CCG GCC AGC TG
 Pro Leu Gly Met Val Ala Pro Ala Ser
 420 425

1277

(2) SEQ ID NO:2 的信息:

(i) 序列特征:

- (A) 长度: 1190碱基对
- (B) 类型: 核酸
- (C) 链性: 双链
- (D) 拓扑学: 线性

(ii) 分子类型: cDNA

(ix) 特征:

- (A) 名称/关键词: CDS
- (B) 位置: 1..1191

(xi) 序列描述: SEQ ID NO:2:

ATG GCT CTG CCG GCC AGT CTG TTG CCC CTG TGC TGC TTG GCA CTC TTG	48
Met Ala Leu Pro Ala Ser Leu Leu Pro Leu Cys Cys Leu Ala Leu Leu	
1 5 10 15	
GCA CTA TCT GCC CAG AGC TGC GGG CCG GGC CGA GGA CCG GTT GGC CGG	96
Ala Leu Ser Ala Gln Ser Cys Gly Pro Gly Arg Gly Pro Val Gly Arg	
20 25 30	
CGG CGT TAT GTG CGC AAG CAA CTT GTG CCT CTG CTA TAC AAG CAG TTT	144
Arg Arg Tyr Val Arg Lys Gln Leu Val Pro Leu Leu Tyr Lys Gln Phe	
35 40 45	
GTG CCC AGT ATG CCC GAG CGG ACC CTG GGC GCG AGT GGG CCA GCG GAG	192
Val Pro Ser Met Pro Glu Arg Thr Leu Gly Ala Ser Gly Pro Ala Glu	
50 55 60	
GGG AGG GTA ACA AGG GGG TCG GAG CGC TTC CGG GAC CTC GTA CCC AAC	240
Gly Arg Val Thr Arg Gly Ser Glu Arg Phe Arg Asp Leu Val Pro Asn	
65 70 75 80	
TAC AAC CCC GAC ATA ATC TTC AAG GAT GAG GAG AAC AGC GGC GCA GAC	288
Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Ser Gly Ala Asp	
85 90 95	
CGC CTG ATG ACA GAG CGT TGC AAA GAG CGG GTG AAC GCT CTA GCC ATC	336
Arg Leu Met Thr Glu Arg Cys Lys Glu Arg Val Asn Ala Leu Ala Ile	
100 105 110	
GCG GTG ATG AAC ATG TGG CCC GGA GTA CGC CTA CGT GTG ACT GAA GGC	384
Ala Val Met Asn Met Trp Pro Gly Val Arg Leu Arg Val Thr Glu Gly	
115 120 125	
TGG GAC GAG GAC GGC CAC CAC GCA CAG GAT TCA CTC CAC TAC GAA GGC	432
Trp Asp Glu Asp Gly His His Ala Gln Asp Ser Leu His Tyr Glu Gly	
130 135 140	
CGT GCC TTG GAC ATC ACC ACG TCT GAC CGT GAC CGT AAT AAG TAT GGT	480
Arg Ala Leu Asp Ile Thr Thr Ser Asp Arg Asp Arg Asn Lys Tyr Gly	

00:06:15

145					150					155					160	
TTG	TTG	GCG	CGC	CTA	GCT	GTG	GAA	GCC	GGA	TTC	GAC	TGG	GTC	TAC	TAC	528
Leu	Leu	Ala	Arg	Leu	Ala	Val	Glu	Ala	Gly	Phe	Asp	Trp	Val	Tyr	Tyr	
				165					170					175		
GAG	TCC	CGC	AAC	CAC	ATC	CAC	GTA	TCG	GTC	AAA	GCT	GAT	AAC	TCA	CTG	576
Glu	Ser	Arg	Asn	His	Ile	His	Val	Ser	Val	Lys	Ala	Asp	Asn	Ser	Leu	
			180					185					190			
GCG	GTC	CGA	GCC	GGA	GGC	TGC	TTT	CCG	GGA	AAT	GCC	ACG	GTG	CGC	TTG	624
Ala	Val	Arg	Ala	Gly	Gly	Cys	Phe	Pro	Gly	Asn	Ala	Thr	Val	Arg	Leu	
		195					200					205				
CGG	AGC	GGC	GAA	CGG	AAG	GGG	CTG	AGG	GAA	CTA	CAT	CGT	GGT	GAC	TGG	672
Arg	Ser	Gly	Glu	Arg	Lys	Gly	Leu	Arg	Glu	Leu	His	Arg	Gly	Asp	Trp	
	210					215					220					
GTA	CTG	GCC	GCT	GAT	GCA	GCG	GGC	CGA	GTG	GTA	CCC	ACG	CCA	GTG	CTG	720
Val	Leu	Ala	Ala	Asp	Ala	Ala	Gly	Arg	Val	Val	Pro	Thr	Pro	Val	Leu	
225					230				235						240	
CTC	TTC	CTG	GAC	CGG	GAT	CTG	CAG	CGC	CGC	GCC	TCG	TTC	GTG	GCT	GTG	768
Leu	Phe	Leu	Asp	Arg	Asp	Leu	Gln	Arg	Arg	Ala	Ser	Phe	Val	Ala	Val	
				245					250					255		
GAG	ACC	GAG	CGG	CCT	CCG	CGC	AAA	CTG	TTG	CTC	ACA	CCC	TGG	CAT	CTG	816
Glu	Thr	Glu	Arg	Pro	Pro	Arg	Lys	Leu	Leu	Leu	Thr	Pro	Trp	His	Leu	
			260					265					270			
GTG	TTC	GCT	GCT	CGC	GGG	CCA	GCG	CCT	GCT	CCA	GGT	GAC	TTT	GCA	CCG	864
Val	Phe	Ala	Ala	Arg	Gly	Pro	Ala	Pro	Ala	Pro	Gly	Asp	Phe	Ala	Pro	
		275				280						285				
GTG	TTC	GCG	CGC	CGC	TTA	CGT	GCT	GGC	GAC	TCG	GTG	CTG	GCT	CCC	GGC	912
Val	Phe	Ala	Arg	Arg	Leu	Arg	Ala	Gly	Asp	Ser	Val	Leu	Ala	Pro	Gly	
	290					295					300					
GGG	GAC	GCG	CTC	CAG	CCG	GCG	CGC	GTA	GCC	CGC	GTG	GCG	CGC	GAG	GAA	960
Gly	Asp	Ala	Leu	Gln	Pro	Ala	Arg	Val	Ala	Arg	Val	Ala	Arg	Glu	Glu	
305					310				315						320	
GCC	GTG	GGC	GTG	TTC	GCA	CCG	CTC	ACT	GCG	CAC	GGG	ACG	CTG	CTG	GTC	1008
Ala	Val	Gly	Val	Phe	Ala	Pro	Leu	Thr	Ala	His	Gly	Thr	Leu	Leu	Val	
				325					330					335		
AAC	GAC	GTC	CTC	GCC	TCC	TGC	TAC	GCG	GTT	CTA	GAG	AGT	CAC	CAG	TGG	1056
Asn	Asp	Val	Leu	Ala	Ser	Cys	Tyr	Ala	Val	Leu	Glu	Ser	His	Gln	Trp	
			340					345					350			
GCC	CAC	CGC	GCC	TTC	GCC	CCT	TTG	CGG	CTG	CTG	CAC	GCG	CTC	GGG	GCT	1104
Ala	His	Arg	Ala	Phe	Ala	Pro	Leu	Arg	Leu	Leu	His	Ala	Leu	Gly	Ala	
		355				360						365				
CTG	CTC	CCT	GGG	GGT	GCA	GTC	CAG	CCG	ACT	GGC	ATG	CAT	TGG	TAC	TCT	1152
Leu	Leu	Pro	Gly	Gly	Ala	Val	Gln	Pro	Thr	Gly	Met	His	Trp	Tyr	Ser	
	370					375					380					
CGC	CTC	CTT	TAC	CGC	TTG	GCC	GAG	GAG	TTA	ATG	GGC	TG				1190
Arg	Leu	Leu	Tyr	Arg	Leu	Ala	Glu	Glu	Leu	Met	Gly					
385					390				395							

(2) SEQ ID NO:3 的信息:

(i) 序列特征:

- (A) 长度: 1281碱基对
 (B) 类型: 核酸
 (C) 链性: 双链
 (D) 拓扑学: 线性

(ii) 分子类型: cDNA

(ix) 特征:

- (A) 名称/关键词: CDS
 (B) 位置: 1..1233

(xi) 序列描述: SEQ ID NO:3:

ATG TCT CCC GCC TGG CTC CGG CCC CGA CTG CGG TTC TGT CTG TTC CTG	48
Met Ser Pro Ala Trp Leu Arg Pro Arg Leu Arg Phe Cys Leu Phe Leu	
1 5 10 15	
CTG CTG CTG CTT CTG GTG CCG GCG GCG CGG GGC TGC GGG CCG GGC CGG	96
Leu Leu Leu Leu Leu Val Pro Ala Ala Arg Gly Cys Gly Pro Gly Arg	
20 25 30	
GTG GTG GGC AGC CGC CGG AGG CCG CCT CGC AAG CTC GTG CCT CTT GCC	144
Val Val Gly Ser Arg Arg Arg Pro Pro Arg Lys Leu Val Pro Leu Ala	
35 40 45	
TAC AAG CAG TTC AGC CCC AAC GTG CCG GAG AAG ACC CTG GGC GCC AGC	192
Tyr Lys Gln Phe Ser Pro Asn Val Pro Glu Lys Thr Leu Gly Ala Ser	
50 55 60	
GGG CGC TAC GAA GGC AAG ATC GCG CGC AGC TCT GAG CGC TTC AAA GAG	240
Gly Arg Tyr Glu Gly Lys Ile Ala Arg Ser Ser Glu Arg Phe Lys Glu	
65 70 75 80	
CTC ACC CCC AAC TAC AAT CCC GAC ATC ATC TTC AAG GAC GAG GAG AAC	288
Leu Thr Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn	
85 90 95	
ACG GGT GCC GAC CGC CTC ATG ACC CAG CGC TGC AAG GAC CGT CTG AAC	336
Thr Gly Ala Asp Arg Leu Met Thr Gln Arg Cys Lys Asp Arg Leu Asn	
100 105 110	
TCA CTG GCC ATC TCT GTC ATG AAC CAG TGG CCT GGT GTG AAA CTG CGG	384
Ser Leu Ala Ile Ser Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg	
115 120 125	
GTG ACC GAA GGC CGG GAT GAA GAT GGC CAT CAC TCA GAG GAG TCT TTA	432
Val Thr Glu Gly Arg Asp Glu Asp Gly His His Ser Glu Glu Ser Leu	
130 135 140	
CAC TAT GAG GGC CGC GCG GTG GAT ATC ACC ACC TCA GAC CGT GAC CGA	480
His Tyr Glu Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg	
145 150 155 160	
AAT AAG TAT GGA CTG CTG GCG CGC TTA GCA GTG GAG GCC GGC TTC GAC	528
Asn Lys Tyr Gly Leu Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp	
165 170 175	
TGG GTG TAT TAC GAG TCC AAG GCC CAC GTG CAT TGC TCT GTC AAG TCT	576

00.05.15

Trp	Val	Tyr	Tyr	Glu	Ser	Lys	Ala	His	Val	His	Cys	Ser	Val	Lys	Ser		
			180					185					190				
GAG	CAT	TCG	GCC	GCT	GCC	AAG	ACA	GGT	GGC	TGC	TTT	CCT	GCC	GGA	GCC	624	
Glu	His	Ser	Ala	Ala	Ala	Lys	Thr	Gly	Gly	Cys	Phe	Pro	Ala	Gly	Ala		
		195					200					205					
CAG	GTG	CGC	CTA	GAG	AAC	GGG	GAG	CGT	GTG	GCC	CTG	TCA	GCT	GTA	AAG	672	
Gln	Val	Arg	Leu	Glu	Asn	Gly	Glu	Arg	Val	Ala	Leu	Ser	Ala	Val	Lys		
	210					215					220						
CCA	GGA	GAC	CGG	GTG	CTG	GCC	ATG	GGG	GAG	GAT	GGG	ACC	CCC	ACC	TTC	720	
Pro	Gly	Asp	Arg	Val	Leu	Ala	Met	Gly	Glu	Asp	Gly	Thr	Pro	Thr	Phe		
225					230					235					240		
AGT	GAT	GTG	CTT	ATT	TTC	CTG	GAC	CGC	GAG	CCA	AAC	CGG	CTG	AGA	GCT	768	
Ser	Asp	Val	Leu	Ile	Phe	Leu	Asp	Arg	Glu	Pro	Asn	Arg	Leu	Arg	Ala		
				245					250					255			
TTC	CAG	GTC	ATC	GAG	ACT	CAG	GAT	CCT	CCG	CGT	CGG	CTG	GCG	CTC	ACG	816	
Phe	Gln	Val	Ile	Glu	Thr	Gln	Asp	Pro	Pro	Arg	Arg	Leu	Ala	Leu	Thr		
			260					265					270				
CCT	GCC	CAC	CTG	CTC	TTC	ATT	GCG	GAC	AAT	CAT	ACA	GAA	CCA	GCA	GCC	864	
Pro	Ala	His	Leu	Leu	Phe	Ile	Ala	Asp	Asn	His	Thr	Glu	Pro	Ala	Ala		
		275					280					285					
CAC	TTC	CGG	GCC	ACA	TTT	GCC	AGC	CAT	GTG	CAA	CCA	GGC	CAA	TAT	GTG	912	
His	Phe	Arg	Ala	Thr	Phe	Ala	Ser	His	Val	Gln	Pro	Gly	Gln	Tyr	Val		
	290					295					300						
CTG	GTA	TCA	GGG	GTA	CCA	GGC	CTC	CAG	CCT	GCT	CGG	GTG	GCA	GCT	GTC	960	
Leu	Val	Ser	Gly	Val	Pro	Gly	Leu	Gln	Pro	Ala	Arg	Val	Ala	Ala	Val		
305					310					315					320		
TCC	ACC	CAC	GTG	GCC	CTT	GGG	TCC	TAT	GCT	CCT	CTC	ACA	AGG	CAT	GGG	1008	
Ser	Thr	His	Val	Ala	Leu	Gly	Ser	Tyr	Ala	Pro	Leu	Thr	Arg	His	Gly		
				325					330					335			
ACA	CTT	GTG	GTG	GAG	GAT	GTG	GTG	GCC	TCC	TGC	TTT	GCA	GCT	GTG	GCT	1056	
Thr	Leu	Val	Val	Glu	Asp	Val	Val	Ala	Ser	Cys	Phe	Ala	Ala	Val	Ala		
			340					345					350				
GAC	CAC	CAT	CTG	GCT	CAG	TTG	GCC	TTC	TGG	CCC	CTG	CGA	CTG	TTT	CCC	1104	
Asp	His	His	Leu	Ala	Gln	Leu	Ala	Phe	Trp	Pro	Leu	Arg	Leu	Phe	Pro		
		355					360					365					
AGT	TTG	GCA	TGG	GGC	AGC	TGG	ACC	CCA	AGT	GAG	GGT	GTT	CAC	TCC	TAC	1152	
Ser	Leu	Ala	Trp	Gly	Ser	Trp	Thr	Pro	Ser	Glu	Gly	Val	His	Ser	Tyr		
	370					375					380						
CCT	CAG	ATG	CTC	TAC	CGC	CTG	GGG	CGT	CTC	TTG	CTA	GAA	GAG	AGC	ACC	1200	
Pro	Gln	Met	Leu	Tyr	Arg	Leu	Gly	Arg	Leu	Leu	Leu	Glu	Glu	Ser	Thr		
385					390					395					400		
TTC	CAT	CCA	CTG	GGC	ATG	TCT	GGG	GCA	GGA	AGC	TGAAGGGACT	CTAACCACTG	1253				
Phe	His	Pro	Leu	Gly	Met	Ser	Gly	Ala	Gly	Ser							
				405					410								
CCCTCCTGGA	ACTGCTGTGC	GTGGATCC														1281	

(2) SEQ ID NO:4 的信息:

(i) 序列特征:

- (A) 长度: 1313碱基对
- (B) 类型: 核酸
- (C) 链性: 双链
- (D) 拓扑学: 线性

(ii) 分子类型: cDNA

(ix) 特征:

- (A) 名称/关键词: CDS
- (B) 位置: 1..1314

(xi) 序列描述: SEQ ID NO:4:

ATG CTG CTG CTG CTG GCC AGA TGT TTT CTG GTG ATC CTT GCT TCC TCG	48
Met Leu Leu Leu Leu Ala Arg Cys Phe Leu Val Ile Leu Ala Ser Ser	
1 5 10 15	
CTG CTG GTG TGC CCC GGG CTG GCC TGT GGG CCC GGC AGG GGG TTT GGA	96
Leu Leu Val Cys Pro Gly Leu Ala Cys Gly Pro Gly Arg Gly Phe Gly	
20 25 30	
AAG AGG CGG CAC CCC AAA AAG CTG ACC CCT TTA GCC TAC AAG CAG TTT	144
Lys Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe	
35 40 45	
ATT CCC AAC GTA GCC GAG AAG ACC CTA GGG GCC AGC GGC AGA TAT GAA	192
Ile Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu	
50 55 60	
GGG AAG ATC ACA AGA AAC TCC GAA CGA TTT AAG GAA CTC ACC CCC AAT	240
Gly Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn	
65 70 75 80	
TAC AAC CCC GAC ATC ATA TTT AAG GAT GAG GAA AAC ACG GGA GCA GAC	288
Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp	
85 90 95	
CGG CTG ATC ACT CAG AGG TGC AAA GAC AAG TTA AAT GCC TTG GCC ATC	336
Arg Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu Ala Ile	
100 105 110	
TCT GTG ATG AAC CAG TGG CCT GGA GTG AGG CTG CGA GTG ACC GAG GGC	384
Ser Val Met Asn Gln Trp Pro Gly Val Arg Leu Arg Val Thr Glu Gly	
115 120 125	
TGG GAT GAG GAC GGC CAT CAT TCA GAG GAG TCT CTA CAC TAT GAG GGT	432
Trp Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr Glu Gly	
130 135 140	
CGA GCA GTG GAC ATC ACC ACG TCC GAC CGG GAC CGC AGC AAG TAC GGC	480
Arg Ala Val Asp Ile Thr Ser Asp Arg Asp Arg Ser Lys Tyr Gly	
145 150 155 160	
ATG CTG GCT CGC CTG GCT GTG GAA GCA GGT TTC GAC TGG GTC TAC TAT	528
Met Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr	
165 170 175	

00.05.15

GAA TCC AAA GCT CAC ATC CAC TGT TCT GTG AAA GCA GAG AAC TCC GTG Glu Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn Ser Val 180 185 190	576
GCG GCC AAA TCC GGC GGC TGT TTC CCG GGA TCC GCC ACC GTG CAC CTG Ala Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr Val His Leu 195 200 205	624
GAG CAG GGC GGC ACC AAG CTG GTG AAG GAC TTA CGT CCC GGA GAC CGC Glu Gln Gly Gly Thr Lys Leu Val Lys Asp Leu Arg Pro Gly Asp Arg 210 215 220	672
GTG CTG GCG GCT GAC GAC CAG GGC CGG CTG CTG TAC AGC GAC TTC CTC Val Leu Ala Ala Asp Asp Gln Gly Arg Leu Leu Tyr Ser Asp Phe Leu 225 230 235 240	720
ACC TTC CTG GAC CGC GAC GAA GGC GCC AAG AAG GTC TTC TAC GTG ATC Thr Phe Leu Asp Arg Asp Glu Gly Ala Lys Lys Val Phe Tyr Val Ile 245 250 255	768
GAG ACG CTG GAG CCG CGC GAG CGC CTG CTG CTC ACC GCC GCG CAC CTG Glu Thr Leu Glu Pro Arg Glu Arg Leu Leu Leu Thr Ala Ala His Leu 260 265 270	816
CTC TTC GTG GCG CCG CAC AAC GAC TCG GGG CCC ACG CCC GGG CCA AGC Leu Phe Val Ala Pro His Asn Asp Ser Gly Pro Thr Pro Gly Pro Ser 275 280 285	864
GCG CTC TTT GCC AGC CGC GTG CGC CCC GGG CAG CGC GTG TAC GTG GTG Ala Leu Phe Ala Ser Arg Val Arg Pro Gly Gln Arg Val Tyr Val Val 290 295 300	912
GCT GAA CGC GGC GGG GAC CGC CGG CTG CTG CCC GCC GCG GTG CAC AGC Ala Glu Arg Gly Gly Asp Arg Arg Leu Leu Pro Ala Ala Val His Ser 305 310 315 320	960
GTG ACG CTG CGA GAG GAG GAG GCG GGC GCG TAC GCG CCG CTC ACG GCG Val Thr Leu Arg Glu Glu Glu Ala Gly Ala Tyr Ala Pro Leu Thr Ala 325 330 335	1008
CAC GGC ACC ATT CTC ATC AAC CGG GTG CTC GCC TCG TGC TAC GCT GTC His Gly Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys Tyr Ala Val 340 345 350	1056
ATC GAG GAG CAC AGC TGG GCA CAC CGG GCC TTC GCG CCT TTC CGC CTG Ile Glu Glu His Ser Trp Ala His Arg Ala Phe Ala Pro Phe Arg Leu 355 360 365	1104
GCG CAC GCG CTG CTG GCC GCG CTG GCA CCC GCC CGC ACG GAC GGC GGG Ala His Ala Leu Leu Ala Ala Leu Ala Pro Ala Arg Thr Asp Gly Gly 370 375 380	1152
GGC GGG GGC AGC ATC CCT GCA GCG CAA TCT GCA ACG GAA GCG AGG GGC Gly Gly Gly Ser Ile Pro Ala Ala Gln Ser Ala Thr Glu Ala Arg Gly 385 390 395 400	1200
GCG GAG CCG ACT GCG GGC ATC CAC TGG TAC TCG CAG CTG CTC TAC CAC Ala Glu Pro Thr Ala Gly Ile His Trp Tyr Ser Gln Leu Leu Tyr His 405 410 415	1248
ATT GGC ACC TGG CTG TTG GAC AGC GAG ACC ATG CAT CCC TTG GGA ATG Ile Gly Thr Trp Leu Leu Asp Ser Glu Thr Met His Pro Leu Gly Met	1296

420

425

430

GCG GTC AAG TCC AGC TG
Ala Val Lys Ser Ser
435

1313

(2) SEQ ID NO:5 的信息:

(i) 序列特征:

- (A) 长度: 1256碱基对
- (B) 类型: 核酸
- (C) 链性: 双链
- (D) 拓扑学: 线性

(ii) 分子类型: cDNA

(ix) 特征:

- (A) 名称/关键词: CDS
- (B) 位置: 1..1257

(xi) 序列描述: SEQ ID NO:5:

ATG CGG CTT TTG ACG AGA GTG CTG CTG GTG TCT CTT CTC ACT CTG TCC	48
Met Arg Leu Leu Thr Arg Val Leu Leu Val Ser Leu Leu Thr Leu Ser	
1 5 10 15	
TTG GTG GTG TCC GGA CTG GCC TGC GGT CCT GGC AGA GGC TAC GGC AGA	96
Leu Val Val Ser Gly Leu Ala Cys Gly Pro Gly Arg Gly Tyr Gly Arg	
20 25 30	
AGA AGA CAT CCG AAG AAG CTG ACA CCT CTC GCC TAC AAG CAG TTC ATA	144
Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe Ile	
35 40 45	
CCT AAT GTC GCG GAG AAG ACC TTA GGG GCC AGC GGC AGA TAC GAG GGC	192
Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu Gly	
50 55 60	
AAG ATA ACG CGC AAT TCG GAG AGA TTT AAA GAA CTT ACT CCA AAT TAC	240
Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn Tyr	
65 70 75 80	
AAT CCC GAC ATT ATC TTT AAG GAT GAG GAG AAC ACG GGA GCG GAC AGG	288
Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp Arg	
85 90 95	
CTC ATG ACA CAG AGA TGC AAA GAC AAG CTG AAC TCG CTG GCC ATC TCT	336
Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ser Leu Ala Ile Ser	
100 105 110	
GTA ATG AAC CAC TGG CCA GGG GTT AAG CTG CGT GTG ACA GAG GGC TGG	384
Val Met Asn His Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp	
115 120 125	
GAT GAG GAC GGT CAC CAT TTT GAA GAA TCA CTC CAC TAC GAG GGA AGA	432
Asp Glu Asp Gly His His Phe Glu Glu Ser Leu His Tyr Glu Gly Arg	
130 135 140	
GCT GTT GAT ATT ACC ACC TCT GAC CGA GAC AAG AGC AAA TAC GGG ACA	480

00.05.15

Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Lys Ser Lys Tyr Gly Thr	
145 150 155 160	
CTG TCT CGC CTA GCT GTG GAG GCT GGA TTT GAC TGG GTC TAT TAC GAG	528
Leu Ser Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu	
165 170 175	
TCC AAA GCC CAC ATT CAT TGC TCT GTC AAA GCA GAA AAT TCG GTT GCT	576
Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn Ser Val Ala	
180 185 190	
GCG AAA TCT GGG GGC TGT TTC CCA GGT TCG GCT CTG GTC TCG CTC CAG	624
Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Leu Val Ser Leu Gln	
195 200 205	
GAC GGA GGA CAG AAG GCC GTG AAG GAC CTG AAC CCC GGA GAC AAG GTG	672
Asp Gly Gly Gln Lys Ala Val Lys Asp Leu Asn Pro Gly Asp Lys Val	
210 215 220	
CTG GCG GCA GAC AGC GCG GGA AAC CTG GTG TTC AGC GAC TTC ATC ATG	720
Leu Ala Ala Asp Ser Ala Gly Asn Leu Val Phe Ser Asp Phe Ile Met	
225 230 235 240	
TTC ACA GAC CGA GAC TCC ACG ACG CGA CGT GTG TTT TAC GTC ATA GAA	768
Phe Thr Asp Arg Asp Ser Thr Thr Arg Arg Val Phe Tyr Val Ile Glu	
245 250 255	
ACG CAA GAA CCC GTT GAA AAG ATC ACC CTC ACC GCC GCT CAC CTC CTT	816
Thr Gln Glu Pro Val Glu Lys Ile Thr Leu Thr Ala Ala His Leu Leu	
260 265 270	
TTT GTC CTC GAC AAC TCA ACG GAA GAT CTC CAC ACC ATG ACC GCC GCG	864
Phe Val Leu Asp Asn Ser Thr Glu Asp Leu His Thr Met Thr Ala Ala	
275 280 285	
TAT GCC AGC AGT GTC AGA GCC GGA CAA AAG GTG ATG GTT GTT GAT GAT	912
Tyr Ala Ser Ser Val Arg Ala Gly Gln Lys Val Met Val Val Asp Asp	
290 295 300	
AGC GGT CAG CTT AAA TCT GTC ATC GTG CAG CGG ATA TAC ACG GAG GAG	960
Ser Gly Gln Leu Lys Ser Val Ile Val Gln Arg Ile Tyr Thr Glu Glu	
305 310 315 320	
CAG CGG GGC TCG TTC GCA CCA GTG ACT GCA CAT GGG ACC ATT GTG GTC	1008
Gln Arg Gly Ser Phe Ala Pro Val Thr Ala His Gly Thr Ile Val Val	
325 330 335	
GAC AGA ATA CTG GCG TCC TGT TAC GCC GTA ATA GAG GAC CAG GGG CTT	1056
Asp Arg Ile Leu Ala Ser Cys Tyr Ala Val Ile Glu Asp Gln Gly Leu	
340 345 350	
GCG CAT TTG GCC TTC GCG CCC GCC AGG CTC TAT TAT TAC GTG TCA TCA	1104
Ala His Leu Ala Phe Ala Pro Ala Arg Leu Tyr Tyr Tyr Val Ser Ser	
355 360 365	
TTC CTG TCC CCC AAA ACT CCA GCA GTC GGT CCA ATG CGA CTT TAC AAC	1152
Phe Leu Ser Pro Lys Thr Pro Ala Val Gly Pro Met Arg Leu Tyr Asn	
370 375 380	
AGG AGG GGG TCC ACT GGT ACT CCA GGC TCC TGT CAT CAA ATG GGA ACG	1200
Arg Arg Gly Ser Thr Gly Thr Pro Gly Ser Cys His Gln Met Gly Thr	
385 390 395 400	

00:05:15

TGG CTT TTG GAC AGC AAC ATG CTT CAT CCT TTG GGG ATG TCA GTA AAC	1248
Trp Leu Leu Asp Ser Asn Met Leu His Pro Leu Gly Met Ser Val Asn	
405 410 415	
 TCA AGC TG	 1256
Ser Ser	

(2) SEQ ID NO:6 的信息:

(i) 序列特征:

- (A) 长度: 1425碱基对
- (B) 类型: 核酸
- (C) 链性: 单链
- (D) 拓扑学: 线性

(ii) 分子类型: cDNA

(ix) 特征:

- (A) 名称/关键词: CDS
- (B) 位置: 1..1425

(xi) 序列描述: SEQ ID NO:6:

ATG CTG CTG CTG GCG AGA TGT CTG CTG CTA GTC CTC GTC TCC TCG CTG	48
Met Leu Leu Leu Ala Arg Cys Leu Leu Leu Val Leu Val Ser Ser Leu	
1 5 10 15	
CTG GTA TGC TCG GGA CTG GCG TGC GGA CCG GGC AGG GGG TTC GGG AAG	96
Leu Val Cys Ser Gly Leu Ala Cys Gly Pro Gly Arg Gly Phe Gly Lys	
20 25 30	
AGG AGG CAC CCC AAA AAG CTG ACC CCT TTA GCC TAC AAG CAG TTT ATC	144
Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe Ile	
35 40 45	
CCC AAT GTG GCC GAG AAG ACC CTA GGC GCC AGC GGA AGG TAT GAA GGG	192
Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu Gly	
50 55 60	
AAG ATC TCC AGA AAC TCC GAG CGA TTT AAG GAA CTC ACC CCC AAT TAC	240
Lys Ile Ser Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn Tyr	
65 70 75 80	
AAC CCC GAC ATC ATA TTT AAG GAT GAA GAA AAC ACC GGA GCG GAC AGG	288
Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp Arg	
85 90 95	
CTG ATG ACT CAG AGG TGT AAG GAC AAG TTG AAC GCT TTG GCC ATC TCG	336
Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu Ala Ile Ser	
100 105 110	
GTG ATG AAC CAG TGG CCA GGA GTG AAA CTG CGG GTG ACC GAG GGC TGG	384
Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp	
115 120 125	
GAC GAA GAT GGC CAC CAC TCA GAG GAG TCT CTG CAC TAC GAG GGC CGC	432
Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr Glu Gly Arg	
130 135 140	

00-05-15

GCA GTG GAC ATC ACC ACG TCT GAC CGC GAC CGC AGC AAG TAC GGC ATG Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Ser Lys Tyr Gly Met 145 150 155 160	480
CTG GCC CGC CTG GCG GTG GAG GCC GGC TTC GAC TGG GTG TAC TAC GAG Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu 165 170 175	528
TCC AAG GCA CAT ATC CAC TGC TCG GTG AAA GCA GAG AAC TCG GTG GCG Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn Ser Val Ala 180 185 190	576
GCC AAA TCG GGA GGC TGC TTC CCG GGC TCG GCC ACG GTG CAC CTG GAG Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr Val His Leu Glu 195 200 205	624
CAG GGC GGC ACC AAG CTG GTG AAG GAC CTG AGC CCC GGG GAC CGC GTG Gln Gly Gly Thr Lys Leu Val Lys Asp Leu Ser Pro Gly Asp Arg Val 210 215 220	672
CTG GCG GCG GAC GAC CAG GGC CGG CTG CTC TAC AGC GAC TTC CTC ACT Leu Ala Ala Asp Asp Gln Gly Arg Leu Leu Tyr Ser Asp Phe Leu Thr 225 230 235 240	720
TTC CTG GAC CGC GAC GAC GGC GCC AAG AAG GTC TTC TAC GTG ATC GAG Phe Leu Asp Arg Asp Asp Gly Ala Lys Lys Val Phe Tyr Val Ile Glu 245 250 255	768
ACG CGG GAG CCG CGC GAG CGC CTG CTG CTC ACC GCC GCG CAC CTG CTC Thr Arg Glu Pro Arg Glu Arg Leu Leu Leu Thr Ala Ala His Leu Leu 260 265 270	816
TTT GTG GCG CCG CAC AAC GAC TCG GCC ACC GGG GAG CCC GAG GCG TCC Phe Val Ala Pro His Asn Asp Ser Ala Thr Gly Glu Pro Glu Ala Ser 275 280 285	864
TCG GGC TCG GGG CCG CCT TCC GGG GGC GCA CTG GGG CCT CGG GCG CTG Ser Gly Ser Gly Pro Pro Ser Gly Gly Ala Leu Gly Pro Arg Ala Leu 290 295 300	912
TTC GCC AGC CGC GTG CGC CCG GGC CAG CGC GTG TAC GTG GTG GCC GAG Phe Ala Ser Arg Val Arg Pro Gly Gln Arg Val Tyr Val Val Ala Glu 305 310 315 320	960
CGT GAC GGG GAC CGC CGG CTC CTG CCC GCC GCT GTG CAC AGC GTG ACC Arg Asp Gly Asp Arg Arg Leu Leu Pro Ala Ala Val His Ser Val Thr 325 330 335	1008
CTA AGC GAG GAG GCC GCG GGC GCC TAC GCG CCG CTC ACG GCC CAG GGC Leu Ser Glu Glu Ala Ala Gly Ala Tyr Ala Pro Leu Thr Ala Gln Gly 340 345 350	1056
ACC ATT CTC ATC AAC CGG GTG CTG GCC TCG TGC TAC GCG GTC ATC GAG Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys Tyr Ala Val Ile Glu 355 360 365	1104
GAG CAC AGC TGG GCG CAC CGG GCC TTC GCG CCC TTC CGC CTG GCG CAC Glu His Ser Trp Ala His Arg Ala Phe Ala Pro Phe Arg Leu Ala His 370 375 380	1152
GCG CTC CTG GCT GCA CTG GCG CCC GCG CGC ACG GAC CGC GGC GGG GAC Ala Leu Leu Ala Ala Leu Ala Pro Ala Arg Thr Asp Arg Gly Gly Asp 1200	1200

385	390	395	400	
AGC GGC GGC GGG GAC CGC GGG GGC GGC GGC GGC AGA GTA GCC CTA ACC				1248
Ser Gly Gly Gly Asp Arg Gly Gly Gly Gly Gly Gly Arg Val Ala Leu Thr	405	410	415	
GCT CCA GGT GCT GCC GAC GCT CCG GGT GCG GGG GCC ACC GCG GGC ATC				1296
Ala Pro Gly Ala Ala Asp Ala Pro Gly Ala Gly Ala Thr Ala Gly Ile	420	425	430	
CAC TGG TAC TCG CAG CTG CTC TAC CAA ATA GGC ACC TGG CTC CTG GAC				1344
His Trp Tyr Ser Gln Leu Leu Tyr Gln Ile Gly Thr Trp Leu Leu Asp	435	440	445	
AGC GAG GCC CTG CAC CCG CTG GGC ATG GCG GTC AAG TCC AGC NNN AGC				1392
Ser Glu Ala Leu His Pro Leu Gly Met Ala Val Lys Ser Ser Xaa Ser	450	455	460	
CGG GGG GCC GGG GGA GGG GCG CGG GAG GGG GCC				1425
Arg Gly Ala Gly Gly Gly Ala Arg Glu Gly Ala	465	470	475	

(2) SEQ ID NO:7 的信息:

(i) 序列特征:

- (A) 长度: 1622碱基对
- (B) 类型: 核酸
- (C) 链性: 双链
- (D) 拓扑学: 线性

(ii) 分子类型: cDNA

(ix) 特征:

- (A) 名称/关键词: CDS
- (B) 位置: 51..1283

(xi) 序列描述: SEQ ID NO:7:

CATCAGCCCA CCAGGAGACC TCGCCCGCCG CTCCCCGGG CTCCCCGGCC ATG TCT	56
Met Ser	
1	
CCC GCC CGG CTC CGG CCC CGA CTG CAC TTC TGC CTG GTC CTG TTG CTG	104
Pro Ala Arg Leu Arg Pro Arg Leu His Phe Cys Leu Val Leu Leu	
5 10 15	
CTG CTG GTG GTG CCC GCG GCA TGG GGC TGC GGG CCG GGT CGG GTG GTG	152
Leu Leu Val Val Pro Ala Ala Trp Gly Cys Gly Pro Gly Arg Val Val	
20 25 30	
GGC AGC CGC CGG CGA CCG CCA CGC AAA CTC GTG CCG CTC GCC TAC AAG	200
Gly Ser Arg Arg Arg Pro Pro Arg Lys Leu Val Pro Leu Ala Tyr Lys	
35 40 45 50	
CAG TTC AGC CCC AAT GTG CCC GAG AAG ACC CTG GGC GCC AGC GGA CGC	248
Gln Phe Ser Pro Asn Val Pro Glu Lys Thr Leu Gly Ala Ser Gly Arg	
55 60 65	
TAT GAA GGC AAG ATC GCT CGC AGC TCC GAG CGC TTC AAG GAG CTC ACC	296

00.06.15

Tyr	Glu	Gly	Lys	Ile	Ala	Arg	Ser	Ser	Glu	Arg	Phe	Lys	Glu	Leu	Thr		
			70					75					80				
CCC	AAT	TAC	AAT	CCA	GAC	ATC	ATC	TTC	AAG	GAC	GAG	GAG	AAC	ACA	GGC	344	
Pro	Asn	Tyr	Asn	Pro	Asp	Ile	Ile	Phe	Lys	Asp	Glu	Glu	Asn	Thr	Gly		
		85					90					95					
GCC	GAC	CGC	CTC	ATG	ACC	CAG	CGC	TGC	AAG	GAC	CGC	CTG	AAC	TCG	CTG	392	
Ala	Asp	Arg	Leu	Met	Thr	Gln	Arg	Cys	Lys	Asp	Arg	Leu	Asn	Ser	Leu		
	100					105					110						
GCT	ATC	TCG	GTG	ATG	AAC	CAG	TGG	CCC	GGT	GTG	AAG	CTG	CGG	GTG	ACC	440	
Ala	Ile	Ser	Val	Met	Asn	Gln	Trp	Pro	Gly	Val	Lys	Leu	Arg	Val	Thr		
	115				120					125					130		
GAG	GGC	TGG	GAC	GAG	GAC	GGC	CAC	CAC	TCA	GAG	GAG	TCC	CTG	CAT	TAT	488	
Glu	Gly	Trp	Asp	Glu	Asp	Gly	His	His	Ser	Glu	Glu	Ser	Leu	His	Tyr		
				135					140					145			
GAG	GGC	CGC	GCG	GTG	GAC	ATC	ACC	ACA	TCA	GAC	CGC	GAC	CGC	AAT	AAG	536	
Glu	Gly	Arg	Ala	Val	Asp	Ile	Thr	Thr	Ser	Asp	Arg	Asp	Arg	Asn	Lys		
			150					155					160				
TAT	GGA	CTG	CTG	GCG	CGC	TTG	GCA	GTG	GAG	GCC	GGC	TTT	GAC	TGG	GTG	584	
Tyr	Gly	Leu	Leu	Ala	Arg	Leu	Ala	Val	Glu	Ala	Gly	Phe	Asp	Trp	Val		
		165					170					175					
TAT	TAC	GAG	TCA	AAG	GCC	CAC	GTG	CAT	TGC	TCC	GTC	AAG	TCC	GAG	CAC	632	
Tyr	Tyr	Glu	Ser	Lys	Ala	His	Val	His	Cys	Ser	Val	Lys	Ser	Glu	His		
		180				185					190						
TCG	GCC	GCA	GCC	AAG	ACG	GGC	GGC	TGC	TTC	CCT	GCC	GGA	GCC	CAG	GTA	680	
Ser	Ala	Ala	Ala	Lys	Thr	Gly	Gly	Cys	Phe	Pro	Ala	Gly	Ala	Gln	Val		
	195				200					205				210			
CGC	CTG	GAG	AGT	GGG	GCG	CGT	GTG	GCC	TTG	TCA	GCC	GTG	AGG	CCG	GGA	728	
Arg	Leu	Glu	Ser	Gly	Ala	Arg	Val	Ala	Leu	Ser	Ala	Val	Arg	Pro	Gly		
				215					220					225			
GAC	CGT	GTG	CTG	GCC	ATG	GGG	GAG	GAT	GGG	AGC	CCC	ACC	TTC	AGC	GAT	776	
Asp	Arg	Val	Leu	Ala	Met	Gly	Glu	Asp	Gly	Ser	Pro	Thr	Phe	Ser	Asp		
			230					235					240				
GTG	CTC	ATT	TTC	CTG	GAC	CGC	GAG	CCC	CAC	AGG	CTG	AGA	GCC	TTC	CAG	824	
Val	Leu	Ile	Phe	Leu	Asp	Arg	Glu	Pro	His	Arg	Leu	Arg	Ala	Phe	Gln		
		245					250					255					
GTC	ATC	GAG	ACT	CAG	GAC	CCC	CCA	CGC	CGC	CTG	GCA	CTC	ACA	CCC	GCT	872	
Val	Ile	Glu	Thr	Gln	Asp	Pro	Pro	Arg	Arg	Leu	Ala	Leu	Thr	Pro	Ala		
		260				265					270						
CAC	CTG	CTC	TTT	ACG	GCT	GAC	AAT	CAC	ACG	GAG	CCG	GCA	GCC	CGC	TTC	920	
His	Leu	Leu	Phe	Thr	Ala	Asp	Asn	His	Thr	Glu	Pro	Ala	Ala	Arg	Phe		
				280						285				290			
CGG	GCC	ACA	TTT	GCC	AGC	CAC	GTG	CAG	CCT	GGC	CAG	TAC	GTG	CTG	GTG	968	
Arg	Ala	Thr	Phe	Ala	Ser	His	Val	Gln	Pro	Gly	Gln	Tyr	Val	Leu	Val		
				295					300					305			
GCT	GGG	GTG	CCA	GGC	CTG	CAG	CCT	GCC	CGC	GTG	GCA	GCT	GTC	TCT	ACA	1016	
Ala	Gly	Val	Pro	Gly	Leu	Gln	Pro	Ala	Arg	Val	Ala	Ala	Val	Ser	Thr		
			310					315					320				

CAC GTG GCC CTC GGG GCC TAC GCC CCG CTC ACA AAG CAT GGG ACA CTG His Val Ala Leu Gly Ala Tyr Ala Pro Leu Thr Lys His Gly Thr Leu 325 330 335	1064
GTG GTG GAG GAT GTG GTG GCA TCC TGC TTC GCG GCC GTG GCT GAC CAC Val Val Glu Asp Val Val Ala Ser Cys Phe Ala Ala Val Ala Asp His 340 345 350	1112
CAC CTG GCT CAG TTG GCC TTC TGG CCC CTG AGA CTC TTT CAC AGC TTG His Leu Ala Gln Leu Ala Phe Trp Pro Leu Arg Leu Phe His Ser Leu 355 360 365 370	1160
GCA TGG GGC AGC TGG ACC CCG GGG GAG GGT GTG CAT TGG TAC CCC CAG Ala Trp Gly Ser Trp Thr Pro Gly Glu Gly Val His Trp Tyr Pro Gln 375 380 385	1208
CTG CTC TAC CGC CTG GGG CGT CTC CTG CTA GAA GAG GGC AGC TTC CAC Leu Leu Tyr Arg Leu Gly Arg Leu Leu Leu Glu Glu Gly Ser Phe His 390 395 400	1256
CCA CTG GGC ATG TCC GGG GCA GGG AGC TGAAAGGACT CCACCGCTGC Pro Leu Gly Met Ser Gly Ala Gly Ser 405 410	1303
CCTCCTGGAA CTGCTGTACT GGGTCCAGAA GCCTCTCAGC CAGGAGGGAG CTGGCCCTGG	1363
AAGGGACCTG AGCTGGGGGA CACTGGCTCC TGCCATCTCC TCTGCCATGA AGATACACCA	1423
TTGAGACTTG ACTGGGCAAC ACCAGCGTCC CCCACCCGCG TCGTGGTGTA GTCATAGAGC	1483
TGCAAGCTGA GCTGGCGAGG GGATGGTTGT TGACCCCTCT CTCCTAGAGA CCTTGAGGCT	1543
GGCACGGCGA CTCCCAACTC AGCCTGCTCT CACTACGAGT TTTCATACTC TGCCTCCCCC	1603
ATTGGGAGGG CCCATTCCC	1622

(2) SEQ ID NO:8 的信息:

(i) 序列特征:

- (A) 长度: 1191碱基对
- (B) 类型: 核酸
- (C) 链性: 双链
- (D) 拓扑学: 线性

(ii) 分子类型: cDNA

(ix) 特征:

- (A) 名称/关键词: CDS
- (B) 位置: 1..1191

(xi) 序列描述: SEQ ID NO:8:

ATG GCT CTC CTG ACC AAT CTA CTG CCC TTG TGC TGC TTG GCA CTT CTG Met Ala Leu Leu Thr Asn Leu Leu Pro Leu Cys Cys Leu Ala Leu Leu 1 5 10 15	48
GCG CTG CCA GCC CAG AGC TGC GGG CCG GGC CGG GGG CCG GTT GGC CGG Ala Leu Pro Ala Gln Ser Cys Gly Pro Gly Arg Gly Pro Val Gly Arg 20 25 30	96

00.06.15

CGC CGC TAT GCG CGC AAG CAG CTC GTG CCG CTA CTC TAC AAG CAA TTT	144
Arg Arg Tyr Ala Arg Lys Gln Leu Val Pro Leu Leu Tyr Lys Gln Phe	
35 40 45	
GTG CCC GGC GTG CCA GAG CGG ACC CTG GGC GCC AGT GGG CCA GCG GAG	192
Val Pro Gly Val Pro Glu Arg Thr Leu Gly Ala Ser Gly Pro Ala Glu	
50 55 60	
GGG AGG GTG GCA AGG GGC TCC GAG CGC TTC CGG GAC CTC GTG CCC AAC	240
Gly Arg Val Ala Arg Gly Ser Glu Arg Phe Arg Asp Leu Val Pro Asn	
65 70 75 80	
TAC AAC CCC GAC ATC ATC TTC AAG GAT GAG GAG AAC AGT GGA GCC GAC	288
Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Ser Gly Ala Asp	
85 90 95	
CGC CTG ATG ACC GAG CGT TGC AAG GAG AGG GTG AAC GCT TTG GCC ATT	336
Arg Leu Met Thr Glu Arg Cys Lys Glu Arg Val Asn Ala Leu Ala Ile	
100 105 110	
GCC GTG ATG AAC ATG TGG CCC GGA GTG CGC CTA CGA GTG ACT GAG GGC	384
Ala Val Met Asn Met Trp Pro Gly Val Arg Leu Arg Val Thr Glu Gly	
115 120 125	
TGG GAC GAG GAC GGC CAC CAC GCT CAG GAT TCA CTC CAC TAC GAA GGC	432
Trp Asp Glu Asp Gly His His Ala Gln Asp Ser Leu His Tyr Glu Gly	
130 135 140	
CGT GCT TTG GAC ATC ACT ACG TCT GAC CGC GAC CGC AAC AAG TAT GGG	480
Arg Ala Leu Asp Ile Thr Thr Ser Asp Arg Asp Arg Asn Lys Tyr Gly	
145 150 155 160	
TTG CTG GCG CGC CTC GCA GTG GAA GCC GGC TTC GAC TGG GTC TAC TAC	528
Leu Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr	
165 170 175	
GAG TCC CGC AAC CAC GTC CAC GTG TCG GTC AAA GCT GAT AAC TCA CTG	576
Glu Ser Arg Asn His Val His Val Ser Val Lys Ala Asp Asn Ser Leu	
180 185 190	
GCG GTC CGG GCG GGC GGC TGC TTT CCG GGA AAT GCA ACT GTG CGC CTG	624
Ala Val Arg Ala Gly Gly Cys Phe Pro Gly Asn Ala Thr Val Arg Leu	
195 200 205	
TGG AGC GGC GAG CGG AAA GGG CTG CGG GAA CTG CAC CGC GGA GAC TGG	672
Trp Ser Gly Glu Arg Lys Gly Leu Arg Glu Leu His Arg Gly Asp Trp	
210 215 220	
GTT TTG GCG GCC GAT GCG TCA GGC CGG GTG GTG CCC ACG CCG GTG CTG	720
Val Leu Ala Ala Asp Ala Ser Gly Arg Val Val Pro Thr Pro Val Leu	
225 230 235 240	
CTC TTC CTG GAC CGG GAC TTG CAG CGC CGG GCT TCA TTT GTG GCT GTG	768
Leu Phe Leu Asp Arg Asp Leu Gln Arg Arg Ala Ser Phe Val Ala Val	
245 250 255	
GAG ACC GAG TGG CCT CCA CGC AAA CTG TTG CTC ACG CCC TGG CAC CTG	816
Glu Thr Glu Trp Pro Pro Arg Lys Leu Leu Leu Thr Pro Trp His Leu	
260 265 270	
GTG TTT GCC GCT CGA GGG CCG GCG CCC GCG CCA GGC GAC TTT GCA CCG	864
Val Phe Ala Ala Arg Gly Pro Ala Pro Ala Pro Gly Asp Phe Ala Pro	

275	280	285	
GTG TTC GCG CGC CGG CTA CGC GCT GGG GAC TCG GTG CTG GCG CCC GGC			912
Val Phe Ala Arg Arg Leu Arg Ala Gly Asp Ser Val Leu Ala Pro Gly			
290	295	300	
GGG GAT GCG CTT CGG CCA GCG CGC GTG GCC CGT GTG GCG CGG GAG GAA			960
Gly Asp Ala Leu Arg Pro Ala Arg Val Ala Arg Val Ala Arg Glu Glu			
305	310	315	320
GCC GTG GGC GTG TTC GCG CCG CTC ACC GCG CAC GGG ACG CTG CTG GTG			1008
Ala Val Gly Val Phe Ala Pro Leu Thr Ala His Gly Thr Leu Leu Val			
325	330	335	
AAC GAT GTC CTG GCC TCT TGC TAC GCG GTT CTG GAG AGT CAC CAG TGG			1056
Asn Asp Val Leu Ala Ser Cys Tyr Ala Val Leu Glu Ser His Gln Trp			
340	345	350	
GCG CAC CGC GCT TTT GCC CCC TTG AGA CTG CTG CAC GCG CTA GGG GCG			1104
Ala His Arg Ala Phe Ala Pro Leu Arg Leu Leu His Ala Leu Gly Ala			
355	360	365	
CTG CTC CCC GGC GGG GCC GTC CAG CCG ACT GGC ATG CAT TGG TAC TCT			1152
Leu Leu Pro Gly Gly Ala Val Gln Pro Thr Gly Met His Trp Tyr Ser			
370	375	380	
CGG CTC CTC TAC CGC TTA GCG GAG GAG CTA CTG GGC TG			1191
Arg Leu Leu Tyr Arg Leu Ala Glu Glu Leu Leu Gly			
385	390	395	

(2) SEQ ID NO:9 的信息:

(i) 序列特征:

- (A) 长度: 1251碱基对
- (B) 类型: 核酸
- (C) 链性: 双链
- (D) 拓扑学: 线性

(ii) 分子类型: cDNA

(ix) 特征:

- (A) 名称/关键词: CDS
- (B) 位置: 1..1248

(xi) 序列描述: SEQ ID NO:9 :

ATG GAC GTA AGG CTG CAT CTG AAG CAA TTT GCT TTA CTG TGT TTT ATC	48
Met Asp Val Arg Leu His Leu Lys Gln Phe Ala Leu Leu Cys Phe Ile	
1 5 10 15	
AGC TTG CTT CTG ACG CCT TGT GGA TTA GCC TGT GGT CCT GGT AGA GGT	96
Ser Leu Leu Leu Thr Pro Cys Gly Leu Ala Cys Gly Pro Gly Arg Gly	
20 25 30	
TAT GGA AAA CGA AGA CAC CCA AAG AAA TTA ACC CCG TTG GCT TAC AAG	144
Tyr Gly Lys Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys	
35 40 45	

00.05.15

CAA TTC ATC CCC AAC GTT GCT GAG AAA ACG CTT GGA GCC AGC GGC AAA Gln Phe Ile Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Lys 50 55 60	192
TAC GAA GGC AAA ATC ACA AGG AAT TCA GAG AGA TTT AAA GAG CTG ATT Tyr Glu Gly Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Ile 65 70 75 80	240
CCG AAT TAT AAT CCC GAT ATC ATC TTT AAG GAC GAG GAA AAC ACA AAC Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Asn 85 90 95	288
GCT GAC AGG CTG ATG ACC AAG CGC TGT AAG GAC AAG TTA AAT TCG TTG Ala Asp Arg Leu Met Thr Lys Arg Cys Lys Asp Lys Leu Asn Ser Leu 100 105 110	336
GCC ATA TCC GTC ATG AAC CAC TGG CCC GGC GTG AAA CTG CGC GTC ACT Ala Ile Ser Val Met Asn His Trp Pro Gly Val Lys Leu Arg Val Thr 115 120 125	384
GAA GGC TGG GAT GAG GAT GGT CAC CAT TTA GAA GAA TCT TTG CAC TAT Glu Gly Trp Asp Glu Asp Gly His His Leu Glu Glu Ser Leu His Tyr 130 135 140	432
GAG GGA CGG GCA GTG GAC ATC ACT ACC TCA GAC AGG GAT AAA AGC AAG Glu Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Lys Ser Lys 145 150 155 160	480
TAT GGG ATG CTA TCC AGG CTT GCA GTG GAG GCA GGA TTC GAC TGG GTC Tyr Gly Met Leu Ser Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val 165 170 175	528
TAT TAT GAA TCT AAA GCC CAC ATA CAC TGC TCT GTC AAA GCA GAA AAT Tyr Tyr Glu Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn 180 185 190	576
TCA GTG GCT GCT AAA TCA GGA GGA TGT TTT CCT GGG TCT GGG ACG GTG Ser Val Ala Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Gly Thr Val 195 200 205	624
ACA CTT GGT GAT GGG ACG AGG AAA CCC ATC AAA GAT CTT AAA GTG GGC Thr Leu Gly Asp Gly Thr Arg Lys Pro Ile Lys Asp Leu Lys Val Gly 210 215 220	672
GAC CGG GTT TTG GCT GCA GAC GAG AAG GGA AAT GTC TTA ATA AGC GAC Asp Arg Val Leu Ala Ala Asp Glu Lys Gly Asn Val Leu Ile Ser Asp 225 230 235 240	720
TTT ATT ATG TTT ATA GAC CAC GAT CCG ACA ACG AGA AGG CAA TTC ATC Phe Ile Met Phe Ile Asp His Asp Pro Thr Thr Arg Arg Gln Phe Ile 245 250 255	768
GTC ATC GAG ACG TCA GAA CCT TTC ACC AAG CTC ACC CTC ACT GCC GCG Val Ile Glu Thr Ser Glu Pro Phe Thr Lys Leu Thr Leu Thr Ala Ala 260 265 270	816
CAC CTA GTT TTC GTT GGA AAC TCT TCA GCA GCT TCG GGT ATA ACA GCA His Leu Val Phe Val Gly Asn Ser Ser Ala Ala Ser Gly Ile Thr Ala 275 280 285	864
ACA TTT GCC AGC AAC GTG AAG CCT GGA GAT ACA GTT TTA GTG TGG GAA Thr Phe Val Ala Ser Asn Val Lys Pro Gly Asp Thr Val Leu Val Trp Glu 290 295 300	912

GAC ACA TGC GAG AGC CTC AAG AGC GTT ACA GTG AAA AGG ATT TAC ACT Asp Thr Cys Glu Ser Leu Lys Ser Val Thr Val Lys Arg Ile Tyr Thr 305 310 315 320	960
GAG GAG CAC GAG GGC TCT TTT GCG CCA GTC ACC GCG CAC GGA ACC ATA Glu Glu His Glu Gly Ser Phe Ala Pro Val Thr Ala His Gly Thr Ile 325 330 335	1008
ATA GTG GAT CAG GTG TTG GCA TCG TGC TAC GCG GTC ATT GAG AAC CAC Ile Val Asp Gln Val Leu Ala Ser Cys Tyr Ala Val Ile Glu Asn His 340 345 350	1056
AAA TGG GCA CAT TGG GCT TTT GCG CCG GTC AGG TTG TGT CAC AAG CTG Lys Trp Ala His Trp Ala Phe Ala Pro Val Arg Leu Cys His Lys Leu 355 360 365	1104
ATG ACG TGG CTT TTT CCG GCT CGT GAA TCA AAC GTC AAT TTT CAG GAG Met Thr Trp Leu Phe Pro Ala Arg Glu Ser Asn Val Asn Phe Gln Glu 370 375 380	1152
GAT GGT ATC CAC TGG TAC TCA AAT ATG CTG TTT CAC ATC GGC TCT TGG Asp Gly Ile His Trp Tyr Ser Asn Met Leu Phe His Ile Gly Ser Trp 385 390 395 400	1200
CTG CTG GAC AGA GAC TCT TTC CAT CCA CTC GGG ATT TTA CAC TTA AGT Leu Leu Asp Arg Asp Ser Phe His Pro Leu Gly Ile Leu His Leu Ser 405 410 415	1248
TGA	1251

(2) SEQ ID NO:10 的信息:

(i) 序列特征:

(A) 长度: 425氨基酸

(B) 类型: 氨基酸

(D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO:10:

Met Val Glu Met Leu Leu Leu Thr Arg Ile Leu Leu Val Gly Phe Ile 1 5 10 15
Cys Ala Leu Leu Val Ser Ser Gly Leu Thr Cys Gly Pro Gly Arg Gly 20 25 30
Ile Gly Lys Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys 35 40 45
Gln Phe Ile Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg 50 55 60
Tyr Glu Gly Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr 65 70 75 80
Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly 85 90 95
Ala Asp Arg Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu

A 10x10 grid of dots forming the numbers 00015. The first three digits are '000' and the last two are '15'. The dots are arranged in a way that the numbers are easily recognizable.

100					105					110					
Ala	Ile	Ser	Val	Met	Asn	Gln	Trp	Pro	Gly	Val	Lys	Leu	Arg	Val	Thr
		115					120					125			
Glu	Gly	Trp	Asp	Glu	Asp	Gly	His	His	Ser	Glu	Glu	Ser	Leu	His	Tyr
	130					135					140				
Glu	Gly	Arg	Ala	Val	Asp	Ile	Thr	Thr	Ser	Asp	Arg	Asp	Arg	Ser	Lys
145					150					155					160
Tyr	Gly	Met	Leu	Ala	Arg	Leu	Ala	Val	Glu	Ala	Gly	Phe	Asp	Trp	Val
				165					170					175	
Tyr	Tyr	Glu	Ser	Lys	Ala	His	Ile	His	Cys	Ser	Val	Lys	Ala	Glu	Asn
			180					185					190		
Ser	Val	Ala	Ala	Lys	Ser	Gly	Gly	Cys	Phe	Pro	Gly	Ser	Ala	Thr	Val
		195					200						205	-	
His	Leu	Glu	His	Gly	Gly	Thr	Lys	Leu	Val	Lys	Asp	Leu	Ser	Pro	Gly
	210					215					220				
Asp	Arg	Val	Leu	Ala	Ala	Asp	Ala	Asp	Gly	Arg	Leu	Leu	Tyr	Ser	Asp
225					230					235					240
Phe	Leu	Thr	Phe	Leu	Asp	Arg	Met	Asp	Ser	Ser	Arg	Lys	Leu	Phe	Tyr
				245					250					255	
Val	Ile	Glu	Thr	Arg	Gln	Pro	Arg	Ala	Arg	Leu	Leu	Leu	Thr	Ala	Ala
			260					265					270		
His	Leu	Leu	Phe	Val	Ala	Pro	Gln	His	Asn	Gln	Ser	Glu	Ala	Thr	Gly
		275					280					285			
Ser	Thr	Ser	Gly	Gln	Ala	Leu	Phe	Ala	Ser	Asn	Val	Lys	Pro	Gly	Gln
	290					295					300				
Arg	Val	Tyr	Val	Leu	Gly	Glu	Gly	Gly	Gln	Gln	Leu	Leu	Pro	Ala	Ser
305					310					315					320
Val	His	Ser	Val	Ser	Leu	Arg	Glu	Glu	Ala	Ser	Gly	Ala	Tyr	Ala	Pro
				325					330					335	
Leu	Thr	Ala	Gln	Gly	Thr	Ile	Leu	Ile	Asn	Arg	Val	Leu	Ala	Ser	Cys
			340					345					350		
Tyr	Ala	Val	Ile	Glu	Glu	His	Ser	Trp	Ala	His	Trp	Ala	Phe	Ala	Pro
		355					360					365			
Phe	Arg	Leu	Ala	Gln	Gly	Leu	Leu	Ala	Ala	Leu	Cys	Pro	Asp	Gly	Ala
	370					375					380				
Ile	Pro	Thr	Ala	Ala	Thr	Thr	Thr	Thr	Gly	Ile	His	Trp	Tyr	Ser	Arg
385					390					395					400
Leu	Leu	Tyr	Arg	Ile	Gly	Ser	Trp	Val	Leu	Asp	Gly	Asp	Ala	Leu	His
				405					410					415	
Pro	Leu	Gly	Met	Val	Ala	Pro	Ala	Ser							
			420					425							

00:06:15

(2) SEQ ID NO:11 的信息:

(i) 序列特征:

(A) 长度: 396氨基酸

(B) 类型: 氨基酸

(D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO:11:

```

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

```

00:06:15

```

                260                265                270
Val Phe Ala Ala Arg Gly Pro Ala Pro Ala Pro Gly Asp Phe Ala Pro
   275                                280                                285

Val Phe Ala Arg Arg Leu Arg Ala Gly Asp Ser Val Leu Ala Pro Gly
   290                                295                                300

Gly Asp Ala Leu Gln Pro Ala Arg Val Ala Arg Val Ala Arg Glu Glu
   305                                310                                315                                320

Ala Val Gly Val Phe Ala Pro Leu Thr Ala His Gly Thr Leu Leu Val
   325                                330                                335

Asn Asp Val Leu Ala Ser Cys Tyr Ala Val Leu Glu Ser His Gln Trp
   340                                345                                350

Ala His Arg Ala Phe Ala Pro Leu Arg Leu Leu His Ala Leu Gly Ala
   355                                360                                365

Leu Leu Pro Gly Gly Ala Val Gln Pro Thr Gly Met His Trp Tyr Ser
   370                                375                                380

Arg Leu Leu Tyr Arg Leu Ala Glu Glu Leu Met Gly
   385                                390                                395

```

(2) SEQ ID NO:12 的信息:

(i) 序列特征:

- (A) 长度: 411氨基酸
- (B) 类型: 氨基酸
- (D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO:12:

```

Met Ser Pro Ala Trp Leu Arg Pro Arg Leu Arg Phe Cys Leu Phe Leu
  1                    5                    10                    15

Leu Leu Leu Leu Leu Val Pro Ala Ala Arg Gly Cys Gly Pro Gly Arg
   20                    25                    30

Val Val Gly Ser Arg Arg Arg Pro Arg Lys Leu Val Pro Leu Ala
   35                    40                    45

Tyr Lys Gln Phe Ser Pro Asn Val Pro Glu Lys Thr Leu Gly Ala Ser
   50                    55                    60

Gly Arg Tyr Glu Gly Lys Ile Ala Arg Ser Ser Glu Arg Phe Lys Glu
   65                    70                    75                    80

Leu Thr Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn
   85                    90                    95

Thr Gly Ala Asp Arg Leu Met Thr Gln Arg Cys Lys Asp Arg Leu Asn
  100                    105                    110

Ser Leu Ala Ile Ser Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg
  115                    120                    125

```

00:06:15

```

Val Thr Glu Gly Arg Asp Glu Asp Gly His His Ser Glu Glu Ser Leu
 130                               135                       140

His Tyr Glu Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg
 145                               150                       155                       160

Asn Lys Tyr Gly Leu Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp
                               165                       170                       175

Trp Val Tyr Tyr Glu Ser Lys Ala His Val His Cys Ser Val Lys Ser
                               180                       185                       190

Glu His Ser Ala Ala Ala Lys Thr Gly Gly Cys Phe Pro Ala Gly Ala
 195                               200                       205

Gln Val Arg Leu Glu Asn Gly Glu Arg Val Ala Leu Ser Ala Val Lys
 210                               215                       220

Pro Gly Asp Arg Val Leu Ala Met Gly Glu Asp Gly Thr Pro Thr Phe
 225                               230                       235                       240

Ser Asp Val Leu Ile Phe Leu Asp Arg Glu Pro Asn Arg Leu Arg Ala
                               245                       250                       255

Phe Gln Val Ile Glu Thr Gln Asp Pro Pro Arg Arg Leu Ala Leu Thr
                               260                       265                       270

Pro Ala His Leu Leu Phe Ile Ala Asp Asn His Thr Glu Pro Ala Ala
 275                               280                       285

His Phe Arg Ala Thr Phe Ala Ser His Val Gln Pro Gly Gln Tyr Val
 290                               295                       300

Leu Val Ser Gly Val Pro Gly Leu Gln Pro Ala Arg Val Ala Ala Val
 305                               310                       315                       320

Ser Thr His Val Ala Leu Gly Ser Tyr Ala Pro Leu Thr Arg His Gly
                               325                       330                       335

Thr Leu Val Val Glu Asp Val Val Ala Ser Cys Phe Ala Ala Val Ala
 340                               345                       350

Asp His His Leu Ala Gln Leu Ala Phe Trp Pro Leu Arg Leu Phe Pro
 355                               360                       365

Ser Leu Ala Trp Gly Ser Trp Thr Pro Ser Glu Gly Val His Ser Tyr
 370                               375                       380

Pro Gln Met Leu Tyr Arg Leu Gly Arg Leu Leu Leu Glu Glu Ser Thr
 385                               390                       395                       400

Phe His Pro Leu Gly Met Ser Gly Ala Gly Ser
                               405                       410

```

(2) SEQ ID NO:13 的信息:

(i) 序列特征:

- (A) 长度: 437氨基酸
- (B) 类型: 氨基酸
- (D) 拓扑学: 线性

00-06-15

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO:13 :

```

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

```

00.06.15

```

Ala Glu Arg Gly Gly Asp Arg Arg Leu Leu Pro Ala Ala Val His Ser
305                      310                      315                      320

Val Thr Leu Arg Glu Glu Glu Ala Gly Ala Tyr Ala Pro Leu Thr Ala
                      325                      330                      335

His Gly Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys Tyr Ala Val
                      340                      345                      350

Ile Glu Glu His Ser Trp Ala His Arg-Ala Phe Ala Pro Phe Arg Leu
                      355                      360                      365

Ala His Ala Leu Leu Ala Ala Leu Ala Pro Ala Arg Thr Asp Gly Gly
                      370                      375                      380

Gly Gly Gly Ser Ile Pro Ala Ala Gln Ser Ala Thr Glu Ala Arg Gly
385                      390                      395                      400

Ala Glu Pro Thr Ala Gly Ile His Trp Tyr Ser Gln Leu Leu Tyr His
                      405                      410                      415

Ile Gly Thr Trp Leu Leu Asp Ser Glu Thr Met His Pro Leu Gly Met
                      420                      425                      430

Ala Val Lys Ser Ser
                      435

```

(2) SEQ ID NO:14 的信息:

(i) 序列特征:

- (A) 长度: 418氨基酸
- (B) 类型: 氨基酸
- (D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO:14:

```

Met Arg Leu Leu Thr Arg Val Leu Leu Val Ser Leu Leu Thr Leu Ser
 1                      5                      10                      15

Leu Val Val Ser Gly Leu Ala Cys Gly Pro Gly Arg Gly Tyr Gly Arg
                20                      25                      30

Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe Ile
 35                      40                      45

Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu Gly
 50                      55                      60

Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn Tyr
 65                      70                      75                      80

Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp Arg
                85                      90                      95

Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ser Leu Ala Ile Ser
 100                      105                      110

Val Met Asn His Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp

```

00.05.15

115					120					125					
Asp	Glu	Asp	Gly	His	His	Phe	Glu	Glu	Ser	Leu	His	Tyr	Glu	Gly	Arg
130					135						140				
Ala	Val	Asp	Ile	Thr	Thr	Ser	Asp	Arg	Asp	Lys	Ser	Lys	Tyr	Gly	Thr
145					150					155					160
Leu	Ser	Arg	Leu	Ala	Val	Glu	Ala	Gly	Phe	Asp	Trp	Val	Tyr	Tyr	Glu
				165					-170					175	
Ser	Lys	Ala	His	Ile	His	Cys	Ser	Val	Lys	Ala	Glu	Asn	Ser	Val	Ala
			180					185					190		
Ala	Lys	Ser	Gly	Gly	Cys	Phe	Pro	Gly	Ser	Ala	Leu	Val	Ser	Leu	Gln
			195				200					205			
Asp	Gly	Gly	Gln	Lys	Ala	Val	Lys	Asp	Leu	Asn	Pro	Gly	Asp	Lys	Val
	210					215					220				
Leu	Ala	Ala	Asp	Ser	Ala	Gly	Asn	Leu	Val	Phe	Ser	Asp	Phe	Ile	Met
225					230					235					240
Phe	Thr	Asp	Arg	Asp	Ser	Thr	Thr	Arg	Arg	Val	Phe	Tyr	Val	Ile	Glu
				245					250					255	
Thr	Gln	Glu	Pro	Val	Glu	Lys	Ile	Thr	Leu	Thr	Ala	Ala	His	Leu	Leu
			260					265					270		
Phe	Val	Leu	Asp	Asn	Ser	Thr	Glu	Asp	Leu	His	Thr	Met	Thr	Ala	Ala
		275					280					285			
Tyr	Ala	Ser	Ser	Val	Arg	Ala	Gly	Gln	Lys	Val	Met	Val	Val	Asp	Asp
	290					295					300				
Ser	Gly	Gln	Leu	Lys	Ser	Val	Ile	Val	Gln	Arg	Ile	Tyr	Thr	Glu	Glu
305					310					315					320
Gln	Arg	Gly	Ser	Phe	Ala	Pro	Val	Thr	Ala	His	Gly	Thr	Ile	Val	Val
				325					330					335	
Asp	Arg	Ile	Leu	Ala	Ser	Cys	Tyr	Ala	Val	Ile	Glu	Asp	Gln	Gly	Leu
			340					345					350		
Ala	His	Leu	Ala	Phe	Ala	Pro	Ala	Arg	Leu	Tyr	Tyr	Tyr	Val	Ser	Ser
		355					360					365			
Phe	Leu	Ser	Pro	Lys	Thr	Pro	Ala	Val	Gly	Pro	Met	Arg	Leu	Tyr	Asn
	370					375					380				
Arg	Arg	Gly	Ser	Thr	Gly	Thr	Pro	Gly	Ser	Cys	His	Gln	Met	Gly	Thr
385					390					395					400
Trp	Leu	Leu	Asp	Ser	Asn	Met	Leu	His	Pro	Leu	Gly	Met	Ser	Val	Asn
				405					410					415	
Ser	Ser														

(2) SEQ ID NO:15 的信息:

(i) 序列特征:

(A) 长度: 475氨基酸

(B) 类型: 氨基酸

(D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ- ID NO:15 :

```

Met Leu Leu Leu Ala Arg Cys Leu Leu Leu Val Leu Val Ser Ser Leu
 1           5           10           15

Leu Val Cys Ser Gly Leu Ala Cys Gly Pro Gly Arg Gly Phe Gly Lys
 20           25           30

Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe Ile
 35           40           45

Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu Gly
 50           55           60

Lys Ile Ser Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn Tyr
 65           70           75           80

Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp Arg
 85           90           95

Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu Ala Ile Ser
100           105           110

Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp
115           120           125

Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr Glu Gly Arg
130           135           140

Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Ser Lys Tyr Gly Met
145           150           155           160

Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu
165           170           175

Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn Ser Val Ala
180           185           190

Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr Val His Leu Glu
195           200           205

Gln Gly Gly Thr Lys Leu Val Lys Asp Leu Ser Pro Gly Asp Arg Val
210           215           220

Leu Ala Ala Asp Asp Gln Gly Arg Leu Leu Tyr Ser Asp Phe Leu Thr
225           230           235           240

Phe Leu Asp Arg Asp Asp Gly Ala Lys Lys Val Phe Tyr Val Ile Glu
245           250           255

Thr Arg Glu Pro Arg Glu Arg Leu Leu Leu Thr Ala Ala His Leu Leu
260           265           270

```

Phe Val Ala Pro His Asn Asp Ser Ala Thr Gly Glu Pro Glu Ala Ser
 275 280 285
 Ser Gly Ser Gly Pro Pro Ser Gly Gly Ala Leu Gly Pro Arg Ala Leu
 290 295 300
 Phe Ala Ser Arg Val Arg Pro Gly Gln Arg Val Tyr Val Val Ala Glu
 305 310 315 320
 Arg Asp Gly Asp Arg Arg Leu Leu Pro Ala Ala Val His Ser Val Thr
 325 330 335
 Leu Ser Glu Glu Ala Ala Gly Ala Tyr Ala Pro Leu Thr Ala Gln Gly
 340 345 350
 Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys Tyr Ala Val Ile Glu
 355 360 365
 Glu His Ser Trp Ala His Arg Ala Phe Ala Pro Phe Arg Leu Ala His
 370 375 380
 Ala Leu Leu Ala Ala Leu Ala Pro Ala Arg Thr Asp Arg Gly Gly Asp
 385 390 395 400
 Ser Gly Gly Gly Asp Arg Gly Gly Gly Gly Gly Arg Val Ala Leu Thr
 405 410 415
 Ala Pro Gly Ala Ala Asp Ala Pro Gly Ala Gly Ala Thr Ala Gly Ile
 420 425 430
 His Trp Tyr Ser Gln Leu Leu Tyr Gln Ile Gly Thr Trp Leu Leu Asp
 435 440 445
 Ser Glu Ala Leu His Pro Leu Gly Met Ala Val Lys Ser Ser Xaa Ser
 450 455 460
 Arg Gly Ala Gly Gly Gly Ala Arg Glu Gly Ala
 465 470 475

(2) SEQ ID NO:16 的信息:

(i) 序列特征:

- (A) 长度: 411 氨基酸
- (B) 类型: 氨基酸
- (D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO:16:

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

00.06.15

Gly	Arg	Tyr	Glu	Gly	Lys	Ile	Ala	Arg	Ser	Ser	Glu	Arg	Phe	Lys	Glu	65	70	75	80
Leu	Thr	Pro	Asn	Tyr	Asn	Pro	Asp	Ile	Ile	Phe	Lys	Asp	Glu	Glu	Asn	85	90	95	
Thr	Gly	Ala	Asp	Arg	Leu	Met	Thr	Gln	Arg	Cys	Lys	Asp	Arg	Leu	Asn	100	105	110	
Ser	Leu	Ala	Ile	Ser	Val	Met	Asn	Gln-Trp	Pro	Gly	Val	Lys	Leu	Arg		115	120	125	
Val	Thr	Glu	Gly	Trp	Asp	Glu	Asp	Gly	His	His	Ser	Glu	Glu	Ser	Leu	130	135	140	
His	Tyr	Glu	Gly	Arg	Ala	Val	Asp	Ile	Thr	Thr	Ser	Asp	Arg	Asp	Arg	145	150	155	160
Asn	Lys	Tyr	Gly	Leu	Leu	Ala	Arg	Leu	Ala	Val	Glu	Ala	Gly	Phe	Asp	165	170	175	
Trp	Val	Tyr	Tyr	Glu	Ser	Lys	Ala	His	Val	His	Cys	Ser	Val	Lys	Ser	180	185	190	
Glu	His	Ser	Ala	Ala	Ala	Lys	Thr	Gly	Gly	Cys	Phe	Pro	Ala	Gly	Ala	195	200	205	
Gln	Val	Arg	Leu	Glu	Ser	Gly	Ala	Arg	Val	Ala	Leu	Ser	Ala	Val	Arg	210	215	220	
Pro	Gly	Asp	Arg	Val	Leu	Ala	Met	Gly	Glu	Asp	Gly	Ser	Pro	Thr	Phe	225	230	235	240
Ser	Asp	Val	Leu	Ile	Phe	Leu	Asp	Arg	Glu	Pro	His	Arg	Leu	Arg	Ala	245	250	255	
Phe	Gln	Val	Ile	Glu	Thr	Gln	Asp	Pro	Pro	Arg	Arg	Leu	Ala	Leu	Thr	260	265	270	
Pro	Ala	His	Leu	Leu	Phe	Thr	Ala	Asp	Asn	His	Thr	Glu	Pro	Ala	Ala	275	280	285	
Arg	Phe	Arg	Ala	Thr	Phe	Ala	Ser	His	Val	Gln	Pro	Gly	Gln	Tyr	Val	290	295	300	
Leu	Val	Ala	Gly	Val	Pro	Gly	Leu	Gln	Pro	Ala	Arg	Val	Ala	Ala	Val	305	310	315	320
Ser	Thr	His	Val	Ala	Leu	Gly	Ala	Tyr	Ala	Pro	Leu	Thr	Lys	His	Gly	325	330	335	
Thr	Leu	Val	Val	Glu	Asp	Val	Val	Ala	Ser	Cys	Phe	Ala	Ala	Val	Ala	340	345	350	
Asp	His	His	Leu	Ala	Gln	Leu	Ala	Phe	Trp	Pro	Leu	Arg	Leu	Phe	His	355	360	365	
Ser	Leu	Ala	Trp	Gly	Ser	Trp	Thr	Pro	Gly	Glu	Gly	Val	His	Trp	Tyr	370	375	380	
Pro	Gln	Leu	Leu	Tyr	Arg	Leu	Gly	Arg	Leu	Leu	Leu	Glu	Glu	Gly	Ser	385	390	395	400

00.06.15

Phe His Pro Leu Gly Met Ser Gly Ala Gly Ser
405 410

(2) SEQ ID NO:17 的信息:

(i) 序列特征:

(A) 长度: 396氨基酸

(B) 类型: 氨基酸

(D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO:17 :

```

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

```

00.05.15

```

Glu Thr Glu Trp Pro Pro Arg Lys Leu Leu Leu Thr Pro Trp His Leu
      260                      265                      270
Val Phe Ala Ala Arg Gly Pro Ala Pro Ala Pro Gly Asp Phe Ala Pro
      275                      280                      285
Val Phe Ala Arg Arg Leu Arg Ala Gly Asp Ser Val Leu Ala Pro Gly
      290                      295                      300
Gly Asp Ala Leu Arg Pro Ala Arg Val-Ala Arg Val Ala Arg Glu Glu
      305                      310                      315                      320
Ala Val Gly Val Phe Ala Pro Leu Thr Ala His Gly Thr Leu Leu Val
      325                      330                      335
Asn Asp Val Leu Ala Ser Cys Tyr Ala Val Leu Glu Ser His Gln Trp
      340                      345                      350
Ala His Arg Ala Phe Ala Pro Leu Arg Leu Leu His Ala Leu Gly Ala
      355                      360                      365
Leu Leu Pro Gly Gly Ala Val Gln Pro Thr Gly Met His Trp Tyr Ser
      370                      375                      380
Arg Leu Leu Tyr Arg Leu Ala Glu Glu Leu Leu Gly
      385                      390                      395

```

(2) SEQ ID NO:18 的信息:

(i) 序列特征:

(A) 长度: 416氨基酸

(B) 类型: 氨基酸

(D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO:18:

```

Met Asp Val Arg Leu His Leu Lys Gln Phe Ala Leu Leu Cys Phe Ile
  1          5          10          15
Ser Leu Leu Leu Thr Pro Cys Gly Leu Ala Cys Gly Pro Gly Arg Gly
  20          25          30
Tyr Gly Lys Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys
  35          40          45
Gln Phe Ile Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Lys
  50          55          60
Tyr Glu Gly Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Ile
  65          70          75          80
Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Asn
      85          90          95
Ala Asp Arg Leu Met Thr Lys Arg Cys Lys Asp Lys Leu Asn Ser Leu
      100          105          110
Ala Ile Ser Val Met Asn His Trp Pro Gly Val Lys Leu Arg Val Thr

```

00.05.15

115	120	125
Glu Gly Trp Asp Glu Asp Gly His His Leu Glu Glu Ser Leu His Tyr		
130	135	140
Glu Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Lys Ser Lys		
145	150	155
Tyr Gly Met Leu Ser Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val		
165	170	175
Tyr Tyr Glu Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn		
180	185	190
Ser Val Ala Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Gly Thr Val		
195	200	205
Thr Leu Gly Asp Gly Thr Arg Lys Pro Ile Lys Asp Leu Lys Val Gly		
210	215	220
Asp Arg Val Leu Ala Ala Asp Glu Lys Gly Asn Val Leu Ile Ser Asp		
225	230	235
Phe Ile Met Phe Ile Asp His Asp Pro Thr Thr Arg Arg Gln Phe Ile		
245	250	255
Val Ile Glu Thr Ser Glu Pro Phe Thr Lys Leu Thr Leu Thr Ala Ala		
260	265	270
His Leu Val Phe Val Gly Asn Ser Ser Ala Ala Ser Gly Ile Thr Ala		
275	280	285
Thr Phe Ala Ser Asn Val Lys Pro Gly Asp Thr Val Leu Val Trp Glu		
290	295	300
Asp Thr Cys Glu Ser Leu Lys Ser Val Thr Val Lys Arg Ile Tyr Thr		
305	310	315
Glu Glu His Glu Gly Ser Phe Ala Pro Val Thr Ala His Gly Thr Ile		
325	330	335
Ile Val Asp Gln Val Leu Ala Ser Cys Tyr Ala Val Ile Glu Asn His		
340	345	350
Lys Trp Ala His Trp Ala Phe Ala Pro Val Arg Leu Cys His Lys Leu		
355	360	365
Met Thr Trp Leu Phe Pro Ala Arg Glu Ser Asn Val Asn Phe Gln Glu		
370	375	380
Asp Gly Ile His Trp Tyr Ser Asn Met Leu Phe His Ile Gly Ser Trp		
385	390	395
Leu Leu Asp Arg Asp Ser Phe His Pro Leu Gly Ile Leu His Leu Ser		
405	410	415

(2) SEQ ID NO:19 的信息:

(i) 序列特征:

- (A) 长度: 1416碱基对
- (B) 类型: 核酸
- (C) 链性: 双链

(D) 拓扑学: 线性

(ii) 分子类型: cDNA

(ix) 特征:

(A) 名称/关键词: CDS

(B) 位置: 1..1413

(xi) 序列描述: SEQ ID NO:19:

ATG GAT AAC CAC AGC TCA GTG CCT TGG GCC AGT GCC GCC AGT GTC ACC	48
Met Asp Asn His Ser Ser Val Pro Trp Ala Ser Ala Ala Ser Val Thr	
1 5 10 15	
TGT CTC TCC CTG GGA TGC CAA ATG CCA CAG TTC CAG TTC CAG TTC CAG	96
Cys Leu Ser Leu Gly Cys Gln Met Pro Gln Phe Gln Phe Gln Phe Gln	
20 25 30	
CTC CAA ATC CGC AGC GAG CTC CAT CTC CGC AAG CCC GCA AGA AGA ACG	144
Leu Gln Ile Arg Ser Glu Leu His Leu Arg Lys Pro Ala Arg Arg Thr	
35 40 45	
CAA ACG ATG CGC CAC ATT GCG CAT ACG CAG CGT TGC CTC AGC AGG CTG	192
Gln Thr Met Arg His Ile Ala His Thr Gln Arg Cys Leu Ser Arg Leu	
50 55 60	
ACC TCT CTG GTG GCC CTG CTG CTG ATC GTC TTG CCG ATG GTC TTT AGC	240
Thr Ser Leu Val Ala Leu Leu Leu Ile Val Leu Pro Met Val Phe Ser	
65 70 75 80	
CCG GCT CAC AGC TGC GGT CCT GGC CGA GGA TTG GGT CGT CAT AGG GCG	288
Pro Ala His Ser Cys Gly Pro Gly Arg Gly Leu Gly Arg His Arg Ala	
85 90 95	
CGC AAC CTG TAT CCG CTG GTC CTC AAG CAG ACA ATT CCC AAT CTA TCC	336
Arg Asn Leu Tyr Pro Leu Val Leu Lys Gln Thr Ile Pro Asn Leu Ser	
100 105 110	
GAG TAC ACG AAC AGC GCC TCC GGA CCT CTG GAG GGT GTG ATC CGT CGG	384
Glu Tyr Thr Asn Ser Ala Ser Gly Pro Leu Glu Gly Val Ile Arg Arg	
115 120 125	
GAT TCG CCC AAA TTC AAG GAC CTC GTG CCC AAC TAC AAC AGG GAC ATC	432
Asp Ser Pro Lys Phe Lys Asp Leu Val Pro Asn Tyr Asn Arg Asp Ile	
130 135 140	
CTT TTC CGT GAC GAG GAA GGC ACC GGA GCG GAT GGC TTG ATG AGC AAG	480
Leu Phe Arg Asp Glu Glu Gly Thr Gly Ala Asp Gly Leu Met Ser Lys	
145 150 155 160	
CGC TGC AAG GAG AAG CTA AAC GTG CTG GCC TAC TCG GTG ATG AAC GAA	528
Arg Cys Lys Glu Lys Leu Asn Val Leu Ala Tyr Ser Val Met Asn Glu	
165 170 175	
TGG CCC GGC ATC CGG CTG CTG GTC ACC GAG AGC TGG GAC GAG GAC TAC	576
Trp Pro Gly Ile Arg Leu Leu Val Thr Glu Ser Trp Asp Glu Asp Tyr	
180 185 190	
CAT CAC GGC CAG GAG TCG CTC CAC TAC GAG GGC CGA GCG GTG ACC ATT	624
His His Gly Gln Glu Ser Leu His Tyr Glu Gly Arg Ala Val Thr Ile	

00.06.15

195	200	205	
GCC ACC TCC GAT CGC GAC CAG TCC AAA TAC GGC ATG CTC GCT CGC CTG Ala Thr Ser Asp Arg Asp Gln Ser Lys Tyr Gly Met Leu Ala Arg Leu 210 215 220			672
GCC GTC GAG GCT GGA TTC GAT TGG GTC TCC TAC GTC AGC AGG CGC CAC Ala Val Glu Ala Gly Phe Asp Trp Val Ser Tyr Val Ser Arg Arg His 225 230 235 240			720
ATC TAC TGC TCC GTC AAG TCA GAT TCG TCG ATC AGT TCC CAC GTG CAC Ile Tyr Cys Ser Val Lys Ser Asp Ser Ser Ile Ser Ser His Val His 245 250 255			768
GGC TGC TTC ACG CCG GAG AGC ACA GCG CTG CTG GAG AGT GGA GTC CGG Gly Cys Phe Thr Pro Glu Ser Thr Ala Leu Leu Glu Ser Gly Val Arg 260 265 270			816
AAG CCG CTC GGC GAG CTC TCT ATC GGA GAT CGT GTT TTG AGC ATG ACC Lys Pro Leu Gly Glu Leu Ser Ile Gly Asp Arg Val Leu Ser Met Thr 275 280 285			864
GCC AAC GGA CAG GCC GTC TAC AGC GAA GTG ATC CTC TTC ATG GAC CGC Ala Asn Gly Gln Ala Val Tyr Ser Glu Val Ile Leu Phe Met Asp Arg 290 295 300			912
AAC CTC GAG CAG ATG CAA AAC TTT GTG CAG CTG CAC ACG GAC GGT GGA Asn Leu Glu Gln Met Gln Asn Phe Val Gln Leu His Thr Asp Gly Gly 305 310 315 320			960
GCA GTG CTC ACG GTG ACG CCG GCT CAC CTG GTT AGC GTT TGG CAG CCG Ala Val Leu Thr Val Thr Pro Ala His Leu Val Ser Val Trp Gln Pro 325 330 335			1008
GAG AGC CAG AAG CTC ACG TTT GTG TTT GCG CAT CGC ATC GAG GAG AAG Glu Ser Gln Lys Leu Thr Phe Val Phe Ala His Arg Ile Glu Glu Lys 340 345 350			1056
AAC CAG GTG CTC GTA CCG GAT GTG GAG ACG GGC GAG CTG AGG CCC CAG Asn Gln Val Leu Val Arg Asp Val Glu Thr Gly Glu Leu Arg Pro Gln 355 360 365			1104
CGA GTG GTC AAG TTG GGC AGT GTG CGC AGT AAG GGC GTG GTC GCG CCG Arg Val Val Lys Leu Gly Ser Val Arg Ser Lys Gly Val Val Ala Pro 370 375 380			1152
CTG ACC CGC GAG GGC ACC ATT GTG GTC AAC TCG GTG GCC GCC AGT TGC Leu Thr Arg Glu Gly Thr Ile Val Val Asn Ser Val Ala Ala Ser Cys 385 390 395 400			1200
TAT GCG GTG ATC AAC AGT CAG TCG CTG GCC CAC TGG GGA CTG GCT CCC Tyr Ala Val Ile Asn Ser Gln Ser Leu Ala His Trp Gly Leu Ala Pro 405 410 415			1248
ATG CGC CTG CTG TCC ACG CTG GAG GCG TGG CTG CCC GCC AAG GAG CAG Met Arg Leu Leu Ser Thr Leu Glu Ala Trp Leu Pro Ala Lys Glu Gln 420 425 430			1296
TTG CAC AGT TCG CCG AAG GTG GTG AGC TCG GCG CAG CAG CAG AAT GGC Leu His Ser Ser Pro Lys Val Val Ser Ser Ala Gln Gln Gln Asn Gly 435 440 445			1344

ATC CAT TGG TAT GCC AAT GCG CTC TAC AAG GTC AAG GAC TAC GTG CTG 1392
 Ile His Trp Tyr Ala Asn Ala Leu Tyr Lys Val Lys Asp Tyr Val Leu
 450 455 460

CCG CAG AGC TGG CGC CAC GAT TGA 1416
 Pro Gln Ser Trp Arg His Asp
 465 470

(2) SEQ ID NO:20 的信息:

(i) 序列特征:

- (A) 长度: 471氨基酸
- (B) 类型: 氨基酸
- (D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO:20:

Met Asp Asn His Ser Ser Val Pro Trp Ala Ser Ala Ala Ser Val Thr
 1 5 10 15

Cys Leu Ser Leu Gly Cys Gln Met Pro Gln Phe Gln Phe Gln Phe Gln
 20 25 30

Leu Gln Ile Arg Ser Glu Leu His Leu Arg Lys Pro Ala Arg Arg Thr
 35 40 45

Gln Thr Met Arg His Ile Ala His Thr Gln Arg Cys Leu Ser Arg Leu
 50 55 60

Thr Ser Leu Val Ala Leu Leu Ile Val Leu Pro Met Val Phe Ser
 65 70 75 80

Pro Ala His Ser Cys Gly Pro Gly Arg Gly Leu Gly Arg His Arg Ala
 85 90 95

Arg Asn Leu Tyr Pro Leu Val Leu Lys Gln Thr Ile Pro Asn Leu Ser
 100 105 110

Glu Tyr Thr Asn Ser Ala Ser Gly Pro Leu Glu Gly Val Ile Arg Arg
 115 120 125

Asp Ser Pro Lys Phe Lys Asp Leu Val Pro Asn Tyr Asn Arg Asp Ile
 130 135 140

Leu Phe Arg Asp Glu Glu Gly Thr Gly Ala Asp Gly Leu Met Ser Lys
 145 150 155 160

Arg Cys Lys Glu Lys Leu Asn Val Leu Ala Tyr Ser Val Met Asn Glu
 165 170 175

Trp Pro Gly Ile Arg Leu Leu Val Thr Glu Ser Trp Asp Glu Asp Tyr
 180 185 190

His His Gly Gln Glu Ser Leu His Tyr Glu Gly Arg Ala Val Thr Ile
 195 200 205

Ala Thr Ser Asp Arg Asp Gln Ser Lys Tyr Gly Met Leu Ala Arg Leu
 210 215 220

Ala Val Glu Ala Gly Phe Asp Trp Val Ser Tyr Val Ser Arg Arg His
 225 230 235 240
 Ile Tyr Cys Ser Val Lys Ser Asp Ser Ser Ile Ser Ser His Val His
 245 250 255
 Gly Cys Phe Thr Pro Glu Ser Thr Ala Leu Leu Glu Ser Gly Val Arg
 260 265 270
 Lys Pro Leu Gly Glu Leu Ser Ile Gly Asp Arg Val Leu Ser Met Thr
 275 280 285
 Ala Asn Gly Gln Ala Val Tyr Ser Glu Val Ile Leu Phe Met Asp Arg
 290 295 300
 Asn Leu Glu Gln Met Gln Asn Phe Val Gln Leu His Thr Asp Gly Gly
 305 310 315 320
 Ala Val Leu Thr Val Thr Pro Ala His Leu Val Ser Val Trp Gln Pro
 325 330 335
 Glu Ser Gln Lys Leu Thr Phe Val Phe Ala His Arg Ile Glu Glu Lys
 340 345 350
 Asn Gln Val Leu Val Arg Asp Val Glu Thr Gly Glu Leu Arg Pro Gln
 355 360 365
 Arg Val Val Lys Leu Gly Ser Val Arg Ser Lys Gly Val Val Ala Pro
 370 375 380
 Leu Thr Arg Glu Gly Thr Ile Val Val Asn Ser Val Ala Ala Ser Cys
 385 390 395 400
 Tyr Ala Val Ile Asn Ser Gln Ser Leu Ala His Trp Gly Leu Ala Pro
 405 410 415
 Met Arg Leu Leu Ser Thr Leu Glu Ala Trp Leu Pro Ala Lys Glu Gln
 420 425 430
 Leu His Ser Ser Pro Lys Val Val Ser Ser Ala Gln Gln Gln Asn Gly
 435 440 445
 Ile His Trp Tyr Ala Asn Ala Leu Tyr Lys Val Lys Asp Tyr Val Leu
 450 455 460
 Pro Gln Ser Trp Arg His Asp
 465 470

(2) SEQ ID NO:21 的信息:

(i) 序列特征:

- (A) 长度: 221氨基酸
- (B) 类型: 氨基酸
- (D) 拓扑学: 线性

(ii) 分子类型: 肽

(v) 片段类型: 内部

(xi) 序列描述: SEQ ID NO:21:

00.05.15

```

Cys Gly Pro Gly Arg Gly Xaa Gly Xaa Arg Arg His Pro Lys Lys Leu
1          5          10          15
Thr Pro Leu Ala Tyr Lys Gln Phe Ile Pro Asn Val Ala Glu Lys Thr
20          25          30
Leu Gly Ala Ser Gly Arg Tyr Glu Gly Lys Ile Xaa Arg Asn Ser Glu
35          40          45
Arg Phe Lys Glu Leu Thr Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys
50          55          60
Asp Glu Glu Asn Thr Gly Ala Asp Arg Leu Met Thr Gln Arg Cys Lys
65          70          75          80
Asp Lys Leu Asn Xaa Leu Ala Ile Ser Val Met Asn Xaa Trp Pro Gly
85          90          95
Val Xaa Leu Arg Val Thr Glu Gly Trp Asp Glu Asp Gly His His Xaa
100         105         110
Glu Glu Ser Leu His Tyr Glu Gly Arg Ala Val Asp Ile Thr Thr Ser
115         120         125
Asp Arg Asp Xaa Ser Lys Tyr Gly Xaa Leu Xaa Arg Leu Ala Val Glu
130         135         140
Ala Gly Phe Asp Trp Val Tyr Tyr Glu Ser Lys Ala His Ile His Cys
145         150         155         160
Ser Val Lys Ala Glu Asn Ser Val Ala Ala Lys Ser Gly Gly Cys Phe
165         170         175
Pro Gly Ser Ala Xaa Val Xaa Leu Xaa Xaa Gly Gly Xaa Lys Xaa Val
180         185         190
Lys Asp Leu Xaa Pro Gly Asp Xaa Val Leu Ala Ala Asp Xaa Xaa Gly
195         200         205

Xaa Leu Xaa Xaa Ser Asp Phe Xaa Xaa Phe Xaa Asp Arg
210         215         220

```

(2) SEQ ID NO:22 的信息:

(i) 序列特征:

- (A) 长度: 167氨基酸
- (B) 类型: 氨基酸
- (D) 拓扑学: 线性

(ii) 分子类型: 肽

(v) 片段类型: 内部

(xi) 序列描述: SEQ ID NO:22:

```

Cys Gly Pro Gly Arg Gly Xaa Xaa Xaa Arg Arg Xaa Xaa Xaa Pro Lys
1          5          10          15

```

00.05.15

Xaa	Leu	Xaa	Pro	Leu	Xaa	Tyr	Lys	Gln	Phe	Xaa	Pro	Xaa	Xaa	Xaa	Glu	20	25	30
Xaa	Thr	Leu	Gly	Ala	Ser	Gly	Xaa	Xaa	Glu	Gly	Xaa	Xaa	Xaa	Arg	Xaa	35	40	45
Ser	Glu	Arg	Phe	Xaa	Xaa	Leu	Thr	Pro	Asn	Tyr	Asn	Pro	Asp	Ile	Ile	50	55	60
Phe	Lys	Asp	Glu	Glu	Asn	Xaa	Gly	Ala	Asp	Arg	Leu	Met	Thr	Xaa	Arg	65	70	75
Cys	Lys	Xaa	Xaa	Xaa	Asn	Xaa	Leu	Ala	Ile	Ser	Val	Met	Asn	Xaa	Trp	85	90	95
Pro	Gly	Val	Xaa	Leu	Arg	Val	Thr	Glu	Gly	Xaa	Asp	Glu	Asp	Gly	His	100	105	110
His	Xaa	Xaa	Xaa	Ser	Leu	His	Tyr	Glu	Gly	Arg	Ala	Xaa	Asp	Ile	Thr	115	120	125
Thr	Ser	Asp	Arg	Asp	Xaa	Xaa	Lys	Tyr	Gly	Xaa	Leu	Xaa	Arg	Leu	Ala	130	135	140
Val	Glu	Ala	Gly	Phe	Asp	Trp	Val	Tyr	Tyr	Glu	Ser	Xaa	Xaa	His	Xaa	145	150	155
His	Xaa	Ser	Val	Lys	Xaa	Xaa										165		

说明书附图

棕榈酰化Shh蛋白诱导毛发生长

12 天

对照

0.5 ug

1.0 ug

2.0 ug

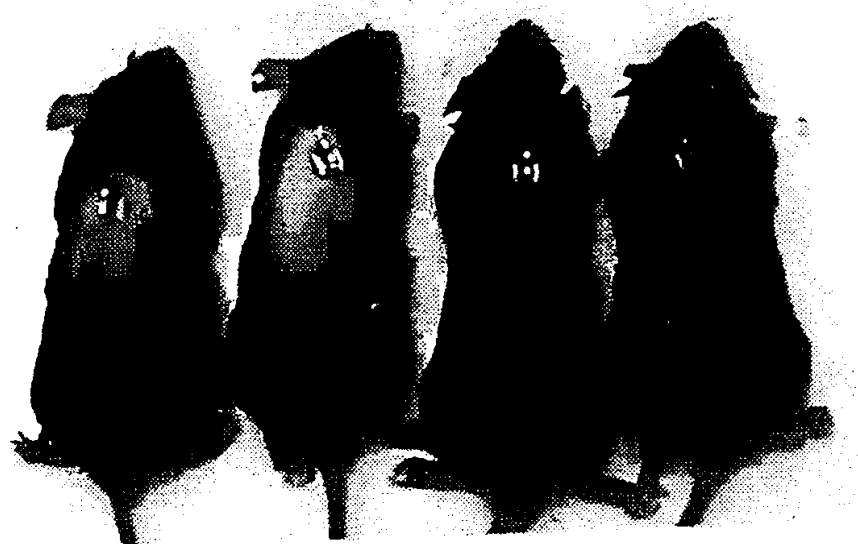


图1A

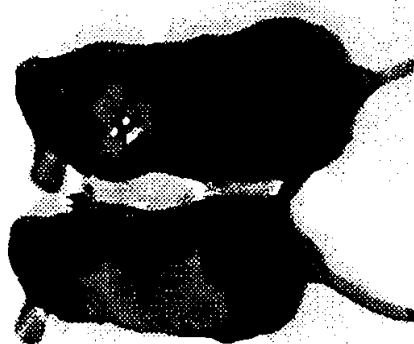
Shh诱导小鼠的毛发生长

13 天

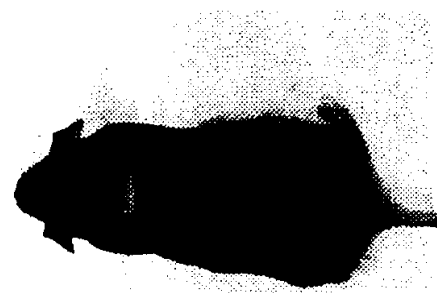
对照

BMP 10 ug

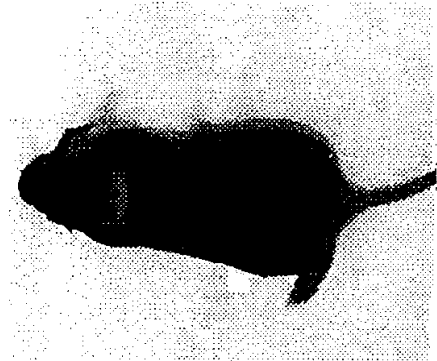
Bac-Shh 50 ug



Pal-Shh 0.5 ug



Pal-Shh 1.0 ug



Pal-Shh 2.0 ug

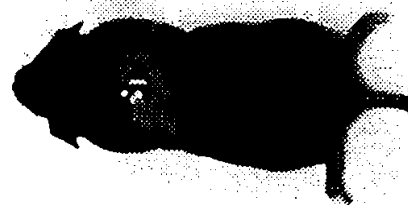


图1B

棕榈酰化Shh蛋白诱导毛发生长

12 天

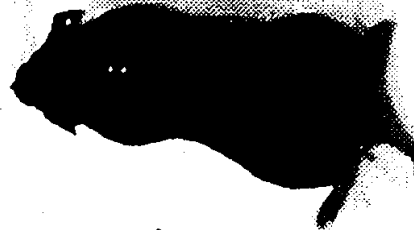
对照



0.5 ug



1.0 ug



2.0 ug

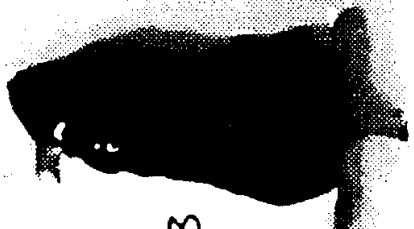


图1C