



(12)发明专利

(10)授权公告号 CN 105358704 B

(45)授权公告日 2020.01.03

(21)申请号 201480025596.9

(22)申请日 2014.03.07

(65)同一申请的已公布的文献号
申请公布号 CN 105358704 A

(43)申请公布日 2016.02.24

(30)优先权数据
61/775,252 2013.03.08 US

(85)PCT国际申请进入国家阶段日
2015.11.05

(86)PCT国际申请的申请数据
PCT/US2014/022052 2014.03.07

(87)PCT国际申请的公布数据
W02014/138670 EN 2014.09.12

(73)专利权人 科罗拉多大学董事会
地址 美国科罗拉多州

(72)发明人 王晓京 张庆红 优素福·里法利

(74)专利代理机构 北京集佳知识产权代理有限公司 11227

代理人 彭鲲鹏 郑斌

(51)Int.Cl.

C12P 19/34(2006.01)

(56)对比文件

W0 2012040295 A2, 2012.03.29,
CN 101238215 A, 2008.08.06,
US 6251628 B1, 2001.06.26,

P. Andrew Chong等. An Expanded WW Domain Recognition Motif Revealed by the Interaction between Smad7 and the E3 Ubiquitin Ligase Smurf2.《The Journal of Biological Chemistry》.2006,第281卷(第25期),第17069-17075页.

Suntaek Hong等. Smad7 binds to the adaptors TAB2 and TAB3 to block recruitment of the kinase TAK1 to the adaptor TRAF2.《Nature Immunology》.2007,第8卷(第5期),第504-513页.

鄢飞等. 转化生长因子-β与软组织创伤愈合的研究进展.《中国民康医学》.2008,第20卷(第8期),第827-828页.

审查员 王雪

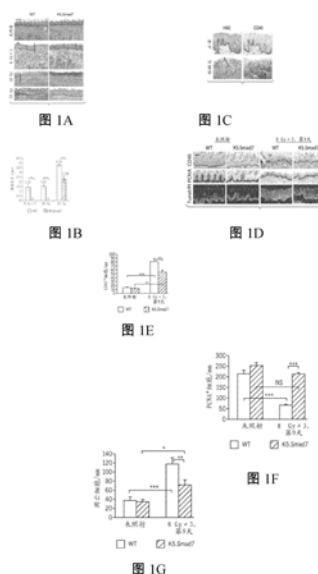
权利要求书1页 说明书79页
序列表52页 附图44页

(54)发明名称

PTD-SMAD7治疗剂

(57)摘要

本发明技术提供了用于治疗发炎和/或组织损伤病状的方法和组合物。特别地,描述了局部或全身递送至发炎和/或组织损伤部位的Smad7组合物的用途。其它特定实施方案涉及由辐射和/或化学疗法引起的副作用,包括(但不限于)口腔和胃粘膜的治疗或预防。还提供了编码Smad7融合蛋白质的密码子优化核酸。



1. 医药组合物在制造用于治疗创口或促进创口愈合之药物中的用途, 其中所述医药组合物包含一种蛋白质分子和一种或多种医药学上可接受的赋形剂, 所述蛋白质分子包含蛋白质转导域和由人类Smad7蛋白质的氨基酸2-258组成的人类Smad7蛋白质片段, 其中所述Smad7蛋白质片段保留功能性全长Smad7蛋白质的生物活性中的一者或多者, 并且其中所述蛋白质分子不包含Smad7蛋白质的C端。

2. 医药组合物在制造用于治疗创口或促进创口愈合之药物中的用途, 其中所述医药组合物包含一种蛋白质分子和一种或多种医药学上可接受的赋形剂, 所述蛋白质分子包含蛋白质转导域和由人类Smad7蛋白质的氨基酸259-426组成的人类Smad7蛋白质片段, 所述Smad7蛋白质片段保留功能性全长Smad7蛋白质的生物活性中的一者或多者, 并且其中所述蛋白质分子不包含Smad7蛋白质的MH1域。

3. 如权利要求2所述的用途, 其中所述蛋白质分子不包含Smad7蛋白质分子的N端。

4. 如权利要求1至3中任一项所述的用途, 其中所述蛋白质转导域是Tat。

5. 医药组合物在制造用于治疗创口或促进创口愈合之药物中的用途, 其中所述医药组合物包含一种核酸分子, 所述核酸分子包含人类Smad7cDNA核苷酸序列片段, 其中所述核苷酸序列编码权利要求1至3中任一项所述的蛋白质分子, 并且不编码全长Smad7蛋白质。

6. 如权利要求5所述的用途, 其中所述核苷酸序列进一步包含编码蛋白质转导域的核苷酸序列。

7. 如权利要求6所述的用途, 其中所述蛋白质转导域是Tat。

PTD-SMAD7治疗剂

[0001] 相关申请的交叉引用

[0002] 本申请要求2013年3月8日提交的美国临时专利申请USSN 61/775,252的优先权,这个专利申请以全文引用的方式并入本文中。

[0003] 政府利益的声明

[0004] 本发明是在政府支持下以国家卫生研究院 (National Institutes of Health) 授予的授权编号AR061796进行。政府对本发明具有一定权利。

[0005] 序列表

[0006] 本发明包括已经以ASCII格式电子提交的序列表,并且将其以整体通过参考并入本文。在2014年4月29日创建了所述ASCII拷贝,将其命名为089491-0303_SL.txt并且其大小为120,505字节。

背景技术

[0007] 口腔粘膜炎,一种重度口腔溃疡,是用于癌症的骨髓移植或颅面放射疗法的大剂量辐射的常见副作用。重度口腔粘膜炎可能需要饲管、重度疼痛的管理以及过早停止放射疗法。过度发炎和上皮消融是口腔粘膜炎的关键特征。

[0008] 帕利夫明 (Palifermin),一种KGF (人角质细胞生长因子) 重组蛋白质,经批准用于预防骨髓移植患者的口腔粘膜炎。在头颈癌患者中的两个帕利夫明临床试验显示,帕利夫明使重度口腔粘膜炎的发生率分别从67%和69%降至51%和54%。在临床试验或临床前研究中的其它口腔粘膜炎药物包括生长因子、辐射防护剂、消炎剂或免疫调节剂。

[0009] 帕利夫明和在上述类别中正在开发的药物的适度作用突出了对鉴别新颖疗法的生物标志物的需要。然而,来自口腔粘膜炎患者的常规诊断活检体或废弃组织的缺乏已经阻碍了这一努力。

[0010] 皮肤创口愈合进展经过三个重叠的阶段:发炎、组织形成以及组织重塑。这些是涉及表皮、白细胞、细胞外基质 (ECM) 以及皮肤成纤维细胞之间的相互作用的动态过程。响应于皮肤损伤,创口中的血凝块、浸润性发炎细胞以及其它细胞类型释放多种细胞因子和趋化因子。这些细胞因子起始成纤维细胞增殖和填充创口缺口并且引起创口闭合的ECM的合成。

[0011] 与此同时,在创口边缘处的角质细胞开始增殖并迁移以覆盖创口表面。在再上皮化的表皮下方,新基质,称为肉芽组织,开始填充创口间隙,这种基质含有临时的ECM、发炎细胞、成纤维细胞以及血管。一旦创口区域被肉芽组织填充并且由新再上皮化的表皮覆盖,创口闭合的过程即完成。稍后,创口经由组织重塑逐渐返回至正常强度和纹理。

[0012] 在已知影响创口愈合的许多分子当中,转化生长因子 β (TGF- β) 具有最广谱的作用,从而影响创口愈合的所有阶段中涉及的所有细胞类型 (Feng等, Annu Rev Cell Dev Biol 21:659-693, 2005)。TGF- β 的各种功能由包括Smad家族成员的许多信号传导分子介导。当配体结合至TGF- β I型和II型受体 (TGF β RI和TGF- β RII) 时, TGF- β RI使Smad2和Smad3磷酸化。磷酸化的Smad2和Smad3结合co-Smad,即Smad4,以形成异聚Smad复合体并且转位至核

中以调控TGF- β 靶基因的转录。

[0013] 已经报导了TGF- β 信号传导对创口愈合发挥正向与负向作用(Wang等, *J Invest Dermatol Symp Proc* 11:112-117, 2006)。举例来说, 部分消去TGF- β 信号传导的Smad3缺陷小鼠展现加速的创口愈合(Ashcroft等, *Nat Cell Biol* 1:260-266, 1999)。相比之下, 引入外源Smad3至创口部位以增强TGF- β 信号传导也加速兔皮肤溃疡模型中的创口愈合(Sumiyoshi等, *J Invest Dermatol* 123:229-236, 2004)。Smad4缺陷小鼠的皮肤创口具有发炎和血管生成的显著增加, 引起创口闭合的延迟并且形成过多疤痕(Owens等, *Am J Pathol* 176:122-133, 2010)。在角膜上皮和基质中TGF- β 信号传导的拮抗剂Smad7的瞬时腺病毒基因转移引起加速的角膜创口愈合伴有减轻的发炎(Saika等, *Am J Pathol* 166:1405-1418, 2005)。此外, Smad7基因转移至晶状体上皮和基质防止晶状体上皮细胞的损伤诱导的上皮-间质转变并且表明Smad7在预防囊纤维化中的潜在作用(Saika等, *Lab Invest* 84:1259-1270, 2004)。然而, Smad7经腺病毒载体递送至大鼠颈动脉中的球囊损伤处导致血管愈合下降(Mallawaarachchi等, *Arterioscler Thromb Vasc Biol* 25:1383-1387, 2005)。这些研究表明, TGF- β 信号传导组分(诸如Smad7)对创口愈合的影响是复杂的并且具有高度情形特异性。另外, Smad7的影响可能并不始终由其在TGF- β 信号传导中的作用来解释。举例来说, 还已经显示Smad7与Wnt/ β -连环蛋白(Han等, *Dev Cell Biol* 11:301-312, 2006)和TNF β /NF- κ B(Hong等, *Nat Immunol* 8:504-513, 2007)家族的组分相互作用。

发明内容

[0014] 本发明技术提供了一种核酸分子, 其包含密码子优化的人类Smad7cDNA核苷酸序列。在一些实施方案中, 密码子优化的人类Smad7核苷酸序列可以包括针对在细菌或酵母中的一者或多者中的表达进行优化的一个或多个精氨酸密码子, 包括针对在细菌或酵母中的一者或多者中的表达进行优化的一个或多个丝氨酸密码子, 和/或包括针对在细菌或酵母中的一者或多者中的表达进行优化的一个或多个组氨酸密码子。在一些实施方案中, 密码子优化的人类Smad7核苷酸序列可以包括针对在细菌或酵母中的一者或多者中的表达进行优化的28个丝氨酸密码子、6个组氨酸密码子以及9个精氨酸密码子。在一些实施方案中, 密码子优化的人类Smad7核苷酸序列可以选自由SEQ ID NO: 9、21、23、24、26、28、30、32-34、36、38、39、87、89、91、93、96、97、99和100组成的群组。在一些实施方案中, 密码子优化的人类Smad7核苷酸序列与人类Smad7cDNA可以具有约65%至75%的同源性, 可以包含编码N端片段SMAD7的核苷酸序列, 可以包含编码SMAD7的C端片段的核苷酸序列, 可以包含编码人类Smad7蛋白质的氨基酸2-258的核苷酸, 可以包含编码人类Smad7蛋白质的氨基酸259-426的核苷酸, 或可以包含编码人类Smad7蛋白质的氨基酸204-258的核苷酸。在一些实施方案中, 前述任一者可以进一步包含编码蛋白质转导域, 诸如Tat的核苷酸序列。在一些实施方案中, 前述任一者还可以进一步包含编码表位标签或纯化标签中的一者或多者, 诸如V5、谷胱甘肽-S-转移酶或6-组氨酸(SEQ ID NO: 40)的核苷酸序列。

[0015] 在一些实施方案中, 前述任一者可以经过分离和/或纯化。在一些实施方案中, 前述任一者还可以编码具有一种或多种选自由以下组成的群组的生物活性的多肽: 减少或消除Smad2的磷酸化、减少或消除NF κ B p50亚单元的核转位、增加细胞增殖、减少凋亡、减轻辐射诱导的DNA损伤、减轻发炎、减少血管生成、促进口腔粘膜炎的愈合、促进创口愈合以及治

疗自体免疫疾病。在一些实施方案中,提供了医药组合物,其包含上述核酸分子和一种或多种医药学上可接受的赋形剂。在一些实施方案中,提供了包含操作性地连接至启动子的上述核酸分子的表达载体,也提供了包含这类表达载体的宿主细胞,以及包含这类载体和宿主细胞与一种或多种医药学上可接受的赋形剂的医药组合物。

[0016] 在一个方面,提供了一种蛋白质分子,其包含在位置216处具有亮氨酸的人类Smad7蛋白质。在一些实施方案中,人类Smad7蛋白质可以在C端处被截短,或在N端处被截短。在一些实施方案中,截短的人类Smad7蛋白质可以包括约50%的全长Smad7序列,或可以包括约13%的全长Smad7序列。在一些实施方案中,人类Smad7蛋白质可以包含以下或由以下组成:人类Smad7蛋白质的氨基酸2-258、氨基酸204-258或氨基酸259-426。在一些实施方案中,蛋白质分子可以具有一种或多种选自以下组成的群组的生物活性:减少或消除Smad2的磷酸化、减少或消除NF κ B p50亚单元的核转位、增加细胞增殖、减少凋亡、减轻辐射诱导的DNA损伤、减轻发炎、减少血管生成、促进口腔粘膜炎的愈合、促进创口愈合以及治疗自体免疫疾病。在一些实施方案中,前述任一者可以进一步包含蛋白质转导域,诸如Tat。在一些实施方案中,前述任一者还可以进一步包含表位标签或纯化标签中的一者或多者,诸如V5、谷胱甘肽-S-转移酶或6-组氨酸(SEQ ID NO:40)。在一些实施方案中,提供了一种医药组合物,其包含前述任一者、蛋白质分子以及一种或多种医药学上可接受的赋形剂。

[0017] 在另一个方面,提供了一种治疗或预防受试者的发炎病状的方法,其包括向受试者提供治疗有效量的上文所描述的医药组合物。在一些实施方案中,发炎病状可以是慢性创口、皮肤发炎、牛皮癣或自体免疫疾病中的一者或多者。在一些实施方案中,组合物可以经由TGF- β 和NF- κ B信号传导的抑制而减轻发炎。

[0018] 在另一个方面,提供了一种预防或治疗受试者的疾病或病症的方法,其包括在受试者中增加细胞增殖或细胞迁移中的一者或多者或防止凋亡或DNA损伤中的一者或多者中的一项或多项,这种方法包括向受试者提供治疗有效量的如上文所描述的医药组合物,其中增加细胞增殖或细胞迁移中的一者或多者或防止凋亡或DNA损伤中的一者或多者中的一项或多项适用于预防或治疗疾病或病症。在一些实施方案中,疾病或病症可以包括慢性创口、急性创口或粘膜炎中的一者或多者。在一些实施方案中,慢性创口可以包括糖尿病性溃疡、压力性溃疡、静脉性溃疡或口腔溃疡中的一者或多者,急性创口可以包括创伤诱导的创口、手术创口或疤痕中的一者或多者,粘膜炎可以包括辐射诱导的粘膜炎或化学疗法诱导的粘膜炎中的一者或多者,并且粘膜炎可以包括口腔粘膜炎或消化道粘膜炎中的一者或多者。

[0019] 预期,本文所描述的任何方法或组合物可以针对本文所描述的任何其它方法或组合物来实施。

[0020] 词语“一个/种(a/an)”的使用当在权利要求书和/或说明书中结合术语“包含”使用时可以意指“一个/种”,但也与“一个/种或多个/种”、“至少一个/种”以及“一个/种或多于一个/种”的含义一致。词语“约”意指加上或减去规定数字的5%。

[0021] 本发明技术的其它目的、特征以及优点将从以下具体实施方式显而易见。然而,应了解,具体实施方式和特定实施例虽然指示本发明技术的特定实施方案,但仅作为说明而给出,这是因为在本发明技术的精神和范围内的各种变化和修改将从这个具体实施方式为本领域的技术人员显而易见。

附图说明

[0022] 以下图式构成本说明书的一部分并且被包括在内以进一步展示本发明技术的某些实施方案。实施方案可以通过参考这些图式中的一者或多者与本文所呈现的特定实施方案的具体实施方式组合来更好地理解。

[0023] 图1A-G提供了显示K5.Smad7小鼠对辐射诱导的口腔粘膜炎具抗性的数据的说明性实施方案。图1A提供了未照射和经过照射(在起始辐射之后第9天)的野生型(WT)和K5.Smad7舌头的H&E染色的说明性实施方案。来自WT小鼠的舌头的图像中的垂直线突出了溃疡边界并且图像中的虚线指示了上皮-基质边界(比例尺,50 μ m)。图1B提供了舌头溃疡的尺寸定量的图示(平均值 \pm s.e.m);在8Gy \times 3辐射中,对于WT小鼠n=8并且对于K5.Smad7小鼠n=7;在18-Gy辐射中,对于WT小鼠n=5并且对于K5.Smad7小鼠n=4;在22-Gy辐射中,对于WT和K5.Smad7小鼠每组n=5。图1C提供了使用H&E(左)和CD45染色(右)显现的舌头(上)和辐射诱导的舌头粘膜炎(下)的人腔室后部的说明性实施方案。实线指示了溃疡边界并且虚线指示了基底膜(比例尺,25 μ m)。图1D提供了在邻近于来自WT小鼠的溃疡的经过照射的切片中和来自K5.Smad7小鼠的损伤区域中的增殖细胞核抗原(PCNA)CD45的免疫染色和TUNEL分析法的说明性实施方案(PI,碘化丙锭)。虚线指示了基底膜(比例尺,25 μ m)。图1E-1G提供了图1D中染色定量的图示(每组n=3或4)。数据表示为平均值 \pm s.e.m(图1B)或平均值 \pm s.d(图1E-1G),并且双尾斯氏t检验(two-tail Student t-test)用于计算P值。*P<0.05,**P<0.01,***P<0.001,NS,通过双尾斯氏t检验确定无显著性。(图1A)、(图1C)以及(图1D)中的虚线突出了基底膜。比例尺:对于(图1A)和(图1C)中的所有画面为50 μ m,对于(图1D)中的所有画面为25 μ m。

[0024] 图2A-G提供了显示由Smad7削弱分子改变的数据的说明性实施方案。图2A提供了NF- κ B p50、TGF- β 1以及pSmad2的免疫染色的说明性实施方案。野生型(WT)的经过照射的舌头切片邻近于溃疡并且K5.Smad7的切片来自损伤区域。人类样品来自未照射的口腔粘膜和辐射诱导的粘膜炎。虚线描绘了上皮-基质边界。比例尺,对于所有画面为25 μ m。图2B提供了(图2A)中所示的NF- κ B p50和pSmad2的免疫染色的定量的图示。图2C提供了TGF- β 1的qRT-PCR的说明性实施方案(由角蛋白5标准化,对于第0天每组n=6,对于第7天和第9天n=4,并且对于第10天n=7)。图2D提供了人口腔角质细胞迁移的定量的图示(参看图8中的图像)。乱序,乱序siRNA;每组n=3。图2E提供了在Smad7敲落之后72小时,对于Smad7和对于Rac1的siSmad7-1和siSmad7-2的敲落效率的西方分析的说明性实施方案。M,分子标志物。图2F提供了总的和活化(GTP结合)的Rac1蛋白质的西方分析的说明性实施方案。M:分子标志物。图2G提供了Rac1敲落对Smad7介导的角质细胞迁移的影响的定量的图示(参看图9A中的敲落效率和图9D中的图像)。每组n=3。数据呈现为平均值 \pm s.d.并且双尾斯氏t检验用于计算(图2B-2D)和(图2G)的P值。*P<0.05,**P<0.01,***P<0.001.NS,无显著性。

[0025] 图3A-H提供了显示Smad7通过阻遏个别Smad和CtBP1结合至Rac1启动子的SBE来增加Rac1表达的数据的说明性实施方案。图3A提供了在野生型(WT)和Smad7转基因角质细胞中Rac1mRNA的定量的图示。每组n=4。图3B提供了在WT和Smad7角质细胞中GTP-Rac1和总Rac1的西方分析的说明性实施方案。通过用针对Smad7的抗体再探测微管蛋白西方墨点来测定WT和Smad7角质细胞中的Smad7蛋白质水平(参看图10A-B中的额外西方墨点和定量)。图3C提供了在敲落人角质细胞中的个别Smad2、Smad3或Smad4之后Rac1蛋白质水平的西方

分析的说明性实施方案(对于Smad敲落效率,参看图10C-10E)。图3D提供了对于在WT和Smad7转基因角质细胞中Smad-2、Smad-3、Smad-4以及Smad-7结合至Rac1启动子的-1.5Kb SBE位点的ChIP分析法的说明性实施方案。图3E提供了在小鼠角质细胞中Rac1荧光素酶报告基因分析法的定量的图示。乱序:乱序siRNA.n=6。图3F提供了在WT或Smad7转基因角质细胞中含有SBE或突变(mut) SBE的Rac1荧光素酶报告基因的活性定量的图示.n=6。图3G提供了在WT或K5.Smad7角质细胞中CtBP1结合至Rac1启动子的SBE-1.5Kb位点的ChIP分析法的图像的说明性实施方案。图3H提供了在WT和Smad7转基因角质细胞中图3G中所示的CtBP1结合至SBE的ChIP-qPCR定量的图示.n=4。数据呈现为平均值±s.d.并且双尾斯氏t检验用于计算图3A、3E、3F以及3H的P值。 $*P<0.05$, $**P<0.01$, $***P<0.001$ 。

[0026] 图4A-G提供了显示CtBP1相关的Rac1阻遏促成角质细胞迁移的抑制的数据的说明性实施方案。图4A提供了在敲落人口腔角质细胞中的CtBP1之后Rac1蛋白质的西方分析的说明性实施方案。图4B提供了含SBE的Rac1luc报告基因活性的定量的图示.n=6。图4C提供了CtBP1敲落对人口腔角质细胞迁移的影响的定量的图示。每组n=3。图4D提供了CtBP1的免疫染色的说明性实施方案。经过照射的切片邻近于溃疡(WT)或损伤区域(K5.Smad7)。虚线表示基底膜。比例尺,对于所有画面为50μm。图4E提供了在人类样本中的未照射的口腔粘膜和辐射诱导的口腔粘膜炎症中CtBP1的免疫染色的说明性实施方案。虚线表示基底膜。比例尺,对于两个画面为50μm。图4F提供了图4D-E中的CtBP1核阳性细胞的定量的图示。每组n=3或4。图4G提供了CtBP1的qRT-PCR的定量的图示(用角蛋白K5标准化)。对于第0天每组n=6,对于第7天和第9天n=4,并且对于第10天n=7。数据呈现为平均值±s.d.并且双尾斯氏t检验用于计算图4B、4C、4F以及4G的P值。 $*P<0.05$, $**P<0.01$, $***P<0.001$ 。

[0027] 图5A-G提供了显示经口施加Tat-Smad7预防小鼠的辐射诱导的口腔粘膜炎症的数据的说明性实施方案。图5A提供了在起始8Gy×3辐射之后第9天口腔粘膜炎症溃疡尺寸的定量的图示。媒剂=盐水或50%甘油/PBS。图5B提供了在起始8Gy×3辐射的第9天的病理改变的说明性实施方案。媒剂=盐水或50%甘油/PBS。比例尺,对于H&E画面为50μm并且对于其余画面为25μm。虚线描绘了上皮-基质边界;实线突出了溃疡边界。图5C、5D、5E、5F以及5G提供了图5B中所示的免疫染色的定量的图示。每组n=3或4。数据呈现为平均值±s.em(图5A)或平均值±s.d.(图5C-5G)并且双尾斯氏t检验用于计算P值。 $*P<0.05$, $**P<0.01$, $***P<0.001$ 。NS,无显著性。

[0028] 图6A-G提供了显示对口腔粘膜炎症的Tat-Smad7处理的数据的说明性实施方案。图6A提供了在起始8Gy×3辐射之后第10天测量的溃疡尺寸的定量的图示。甘油=50%甘油/PBS。图6B提供了口腔粘膜的H&E染色的说明性实施方案。上画面:帕利夫明处理但未用Tat-Smad7处理的粘膜中的开放型溃疡。下画面:帕利夫明处理与Tat-Smad7处理的粘膜之间的上皮厚度的比较。虚线描绘了基底膜。垂直线突出了溃疡边界。比例尺,对于所有画面为50μm。图6C提供了在20Gy诱导的口腔粘膜炎症中在溃疡愈合之后Tat-Smad7处理的免疫染色的说明性实施方案。V5免疫染色显现了口腔上皮中的Tat-Smad7(切片远离损伤区)。K14免疫染色用作对比染色。虚线描绘了基底膜。比例尺,对于所有画面为25μm。图6D提供了在起始8Gy×3辐射之后第10天,Tat-Smad7处理的小鼠舌头的Rac1西方分析的说明性实施方案。图6E提供了在处理之后48小时,对Tat-Smad7处理的正常人口腔角质细胞的Rac1西方分析的说明性实施方案。图6F提供了Tat-Smad7处理对口腔人角质细胞迁移的影响的说明性实施方

案(NOK-SI,参看图13A中的图像)。每组 $n=4$ 。图6G提供了在有或没有Tat-Smad7处理的情况下NOK-SI角质细胞和SCC系(Ca127和MSK921)的存活曲线的定量的图示。对于每种辐射剂量每组 $n=4$ 。数据呈现为平均值 \pm s.e.m(图6A)或平均值 \pm s.d.(图6F、6G)并且双尾斯氏t检验用于计算P值。 $*P<0.05$, $**P<0.01$, $***P<0.001$ 。NS,无显著性。

[0029] 图7A-E提供了显示K5.Smad7口腔粘膜组织对辐射诱导的口腔粘膜炎症具抗性的数据的说明性实施方案。图7A提供了Smad7西方墨点的说明性实施方案:在未照射的野生型(WT)舌头中检测不到并且在辐射之后很少检测到。K5.Smad7舌头在辐射之前和之后具有可比的Smad7蛋白质水平。M:分子标志物。图7B提供了Smad7免疫染色的说明性实施方案。注意到一些经过照射的上皮细胞中的核是肥大的。虚线描绘了上皮-基质边界。图7C提供了在K5.Smad7小鼠中口腔粘膜炎症诱导的病态的发生率降低的定量的图示。费氏精确检验(Fisher's exact test)用于计算p值。 $**P=0.007$ 。图7D提供了显示与WT口腔粘膜炎症相比嗜中性粒细胞(Ly-6G)、巨噬细胞(BM8)以及活化T细胞(CD4)的浸润减少的K5.Smad7舌头的免疫染色的说明性实施方案。虚线描绘了上皮-基质边界。图7E提供了显示在辐射之前或之后在WT与K5.Smad7口腔粘膜之间的pSmad1/5/8-核阳性细胞(绿色)中无显著差异的免疫染色的说明性实施方案。角蛋白(K14)免疫染色(红色)突出了上皮区室。注意到经过照射的上皮细胞的核是肥大的。比例尺对于所有画面为 $50\mu\text{m}$ 。

[0030] 图8A-D提供了显示自发永生化的口腔上皮细胞(NOK-SI)的迁移通过敲落Smad7而延迟但通过敲落TGF- β 1而加速的数据的说明性实施方案。图8A和8B提供了细胞迁移的代表性图像的说明性实施方案。虚线对描绘了划痕创口。Smad7敲落的细胞迁移和效率的定量呈现于图2D和图2E(上述)中。乱序,乱序siRNA。图8C提供了来自3个独立实验的在TGF- β 1敲落之后细胞迁移的定量的图示。图8D提供了显示TGF- β 1敲落效率的qRT-PCR的图示。数据呈现为平均值 \pm s.d.并且双尾斯氏t检验用于计算P值。 $*P<0.05$, $**P<0.01$ 。NS,无显著性。

[0031] 图9A-D提供了显示敲落Rac1减少野生型(WT)和Smad7转基因角质细胞的增殖和迁移的数据的说明性实施方案。图9A提供了在Rac1siRNA(siRac1-1、siRac1-2)转染之后48小时对Rac1的西方墨点分析的说明性实施方案。对照,乱序siRNA。图9B提供了在有或没有Rac1敲落的情况下在BrdU并入分析法中WT和Smad7培养细胞中的BrdU标记细胞的百分比的图示。来自3个独立实验的数据呈现为平均值 \pm s.d。 $***P<0.001$ 。图9C提供了(图9B)中所呈现的BrdU阳性细胞的代表性免疫荧光的说明性实施方案。针对角蛋白14(K14,红色)的抗体用于对比染色。图9D提供了在Rac1敲落之后对Smad7转基因和WT角质细胞的体外细胞迁移分析法的说明性实施方案。虚线对描绘了划痕创口。细胞迁移的定量呈现于图2G中。

[0032] 图10A-F提供了显示Smad7通过阻遏Smad和CtBP1结合至Rac1启动子的SBE来增加Rac1表达的数据的说明性实施方案。图10A提供了在Smad7转基因角质细胞中对GTP-Rac1和总Rac1的西方墨点分析的说明性实施方案。额外样品示于图3B中。M,分子标志物。图10B提供了在图10A中和图3B中所示的WT和K5.Smad7角质细胞中的GTP-Rac1、总Rac1以及Smad7的定量的图示。每个墨点的WT角质细胞中的蛋白质水平标准化为“1”。数据呈现为平均值 \pm s.d.并且双尾斯氏t检验用于计算P值。 $**P<0.01$, $***P<0.001$ 。图10C和10D提供了在NOK-SI细胞中对Smad2、Smad3以及Smad4敲落的西方墨点分析的说明性实施方案。其对Rac1表达的影响示于图3C中。M,分子标志物。GAPDH,通过再探测相同墨点的内部蛋白质对照。图10F提供了显示CtBP1敲落促进NOK-SI细胞迁移的说明性实施方案。虚线对描绘了划痕创口。

CtBP1敲落的细胞迁移和效率的定量示于图4A和图4C中。

[0033] 图11A-G提供了显示Tat-Smad7和Tat-Cre蛋白质的纯化和表征的数据的说明性实施方案。图11A示出了Tat-Smad7蛋白质的示意图的说明性实施方案。图11A按出现的顺序分别公开了SEQ ID NO 49和101。图11B提供了纯化的Tat-Smad7蛋白质的西方墨点的说明性实施方案。图11C提供了在角质细胞中Tat-Smad7蛋白质转导的免疫染色的说明性实施方案。左侧和中间画面：使用V5抗体的Tat-Smad7染色(绿色)，用K14抗体进行对比染色(红色)。细胞显示在转导之后5分钟在核中以及在转导之后12小时在核与细胞质中的Tat-Smad7。右侧画面：Tat-Smad7消去Smad2磷酸化(pSmad2,绿色)。V5(红色)对比染色显现了Tat-Smad7转导的细胞。图11D提供了显示V5抗体染色检测在Tat-Smad7局部施加之后12小时颊粘膜中的Tat-Smad7转导的免疫染色的说明性实施方案。K14抗体用于对比染色。比例尺,对于两个画面为50 μ m。图11E提供了图11A中所示的具有相同Tat和V5标签的纯化的Tat-Cre蛋白质的西方墨点的说明性实施方案。图11F提供了显示Tat-Cre的活性的琼脂糖凝胶的说明性实施方案：Tat-Cre从7,650bp载体pLL3.7切出1,460bp侧接loxP(floxed)的片段。图11G提供了显示Tat-Smad7蛋白质预防性处理减轻20Gy辐射诱导的口腔溃疡的图示。数据表示为平均值 \pm s.e.m。双尾斯氏t检验用于计算P值。 $*P<0.05$, $***P<0.001$ 。

[0034] 图12A-I提供了显示Tat-Smad7处理对口腔粘膜炎的影响的数据的说明性实施方案。图12A提供了在Tat-Smad7(每天0.8 μ g,第6天至第9天)处理的口腔粘膜中溃疡尺寸减小的定量的图示。在第10天收集样品。每组n=8。图12B提供了来自图12A的样品的分子标志物的免疫染色的说明性实施方案。比例尺,对于顶部两个画面为50 μ m并且对于其它画面为25 μ m。碘化丙锭(PI)和K14用作对比染色。图12C-G提供了图12C中所示的免疫染色的定量的图示。使用3-4个样品。图12H提供了荧光素酶分析法的定量的图示。Tat-Smad7处理增加小鼠角质细胞中具有SBE而非突变SBE的Rac1启动子的活性。图12I提供了在Tat-Smad7处理的小鼠角质细胞中CtBP1结合至小鼠Rac1启动子的SBE的ChIP分析法的说明性实施方案。数据表示为平均值 \pm s.e.m(a)或平均值 \pm s.d(c-h)并且双尾斯氏t检验用于计算P值。 $*P<0.05$, $**P<0.01$, $***P<0.001$ 。NS,无显著性。

[0035] 图13A-H提供了显示Tat-Smad7处理对人角质细胞和肿瘤细胞系的迁移的影响的数据的说明性实施方案。图13A提供了显示Tat-Smad7加速NOK-SI细胞迁移的说明性实施方案。来自四个独立实验的定量示于图6F(上述)中。虚线对描绘了初始创口。图13B提供了在NOK-SI细胞中显示削弱的辐射诱导的pSmad2和NF- κ B p50核定位的Tat-Smad7处理的免疫染色的说明性实施方案。图13C提供了显示在Tat-Smad7处理之后2小时MSK921细胞的V5染色的说明性实施方案。K14染色用作对比染色。图13D提供了在Tat-Smad7处理之后60小时在MSK921中的Rac1西方分析的说明性实施方案。M,分子标志物。图13E提供了来自3个独立实验的MSK921细胞迁移的定量的图示。图13F提供了显示用Tat-Smad7和PBS处理的代表性MSK921细胞迁移分析法的说明性实施方案。实线对描绘了初始创口。虚线突出了迁移细胞的最前部。图13G提供了来自3个独立实验的Ca127细胞迁移的定量的图示。图13H提供了显示图13G的代表性图像的说明性实施方案。实线对描绘了初始创口。虚线突出了迁移细胞的最前部。数据表示为平均值 \pm s.d.并且双尾斯氏t检验用于计算P值。NS,无显著性。

[0036] 图14A-B示出了口腔粘膜炎的Smad7介导的保护和愈合的潜在机制的概括的说明性示意图。图14A示出了辐射如何活化NF- κ B、增加TGF- β 1和CtBP1的说明性示意图。NF- κ B和

TGF- β 1诱导发炎。TGF- β 1诱导凋亡、生长停滞并且活化Smad-2、Smad-3以及Smad-4,其将CtBP1募集至Rac1启动子以阻遏Rac1转录,导致再上皮化减弱。图14B示出了Smad7如何阻断NF- κ B和TGF- β 1诱导的发炎并且阻断TGF- β 1诱导的凋亡和生长停滞的说明性示意图。Smad7通过防止TGF- β 1介导的Smad活化(磷酸化)或与信号传导Smad/CtBP1转录阻遏复合体竞争结合至Rac1启动子来解除Rac1转录阻遏。由Smad7诱导的Rac1增加有助于在再上皮化期间角质细胞迁移。

[0037] 图15示出了与蛋白质伴侣、潜在靶点效应以及潜在生理效应相关的Smad 7域的说明性示意图。

[0038] 图16A-B是展示截短的Smad7蛋白质加速小鼠创口愈合模型中的创口愈合的能力的曲线图。图16A是显示C端截短(259-426aa)的Tat-C-Smad7相对于全长Tat-Smad7和对照(PBS)随着时间对平均创口愈合百分比的影响的曲线图。每组n=3。图16B是说明Tat-N-Smad7(1-258aa)相对于全长Tat-Smad7和对照(PBS)随着时间对平均创口愈合百分比的影响的曲线图。每组n=6。数据呈现为平均值 \pm s.d.并且双尾斯氏t检验用于计算P值。 $*p<0.05$,与对照(PBS)相比; $\#p<0.05$,与Tat-Smad7相比。

[0039] 图17A-C是展示Smad7加速受损创口愈合模型中的创口愈合的照片和曲线图。图17A是说明经过十三天的时段用PBS或Tat-Smad7处理的糖尿病(db/db)小鼠中的创口的大体外观的数码照片。图17B是显示Tat-Smad7相对于REGRANEX[®]和对照(PBS)随着时间对平均创口愈合百分比的影响的曲线图。每组n=6。数据呈现为平均值 \pm s.d.并且双尾斯氏t检验用于计算P值。 $*p<0.05$,与对照(PBS)相比。图17C是在开创口之后八天获取的创口样品的组织学比较。来自对照(PBS)db/db小鼠的图像中的垂直虚线(上画面)突出了创口边界。

具体实施方式

[0040] 如本文进一步描述,本公开提供了Smad7蛋白质及其生物活性片段和衍生物、编码这类蛋白质的核酸、包括这类核酸的载体,以及包涵载体、核酸和/或蛋白质的细胞,这些全部用于配制药剂并且用于治疗 and/或预防一种或多种疾病或病症。还提供了制造和筛选适用于治疗和/或预防一种或多种疾病或病症的Smad7蛋白质及其生物活性片段和衍生物的方法。还提供了使用一种或多种与暴露于Smad7相关的标志物来预测和/或评价对治疗的反应的方法。这类标志物可以包括(但不限于)对于细胞迁移的Rac1、对于发炎的NF- κ B,以及对于生长停滞和发炎的TGF- β 。

[0041] Smad7可治疗的疾病和病症可以包括有包括以下一者或多者的那些:减少的细胞增殖、减少的细胞迁移、增加的细胞死亡、过度发炎、和/或DNA损伤。Smad7可治疗的疾病和病症可以包括用具有以下一种或多种活性的Smad7蛋白质及其生物活性片段和衍生物治疗的那些:包括(但不限于)增加增殖、减少或抑制细胞死亡、减轻过度发炎、防止DNA损伤、和/或增加细胞迁移。这类疾病和/或病症可以包括(但不限于)急性(例如,经由手术、搏斗、创伤)和慢性创口(例如,溃疡,诸如糖尿病性、压力性、静脉性)、疤痕、纤维化和异常愈合、粘膜炎(例如,口腔和/或胃肠)、口炎、直肠炎、自体免疫疾病(例如,牛皮癣、关节炎),以及癌症。

[0042] 对于口腔粘膜炎的预防和治疗至关重要是克服因大量凋亡和减弱的角质细胞

增殖所致的上皮消融。当增加上皮细胞的强效生长抑制剂和凋亡诱导剂TGF- β 1时,Smad7的增殖和抗凋亡作用在口腔粘膜炎症中比在正常口腔粘膜中更加明显。

[0043] 尽管不希望受理论约束,但据信增加的Rac1活化在创口闭合中主要负责Smad7介导的角质细胞迁移。鉴于文献记载的TGF- β 信号传导经由Smad非依赖性机制在癌细胞中的Rho/Rac活化中的作用(Dernyck等,Nature 415:577-584,2003),这一发现是意想不到的。

[0044] 据信在口腔粘膜炎症期间,Smad依赖性Rac1阻遏克服了因增加的Smad信号传导(由增加的pSmad2证明)和Smad转录辅阻遏子CtBP1所致的Smad非依赖性Rac1活化(如果有的话)。当这种阻遏被Smad7消去时,其允许Rac1活化介导的角质细胞迁移。然而,在口腔癌细胞中,信号传导Smad损失或灭活,或其它机制独立地活化Rac1。因此,Rac1阻遏的Smad7介导的消去将不再发生。

[0045] 尽管Rac1活化也促成角质细胞增殖,但敲落Rac1仅部分削弱Smad7的增殖作用。因此,Rac1对增殖的贡献似乎是有限的,并且还需要阻断TGF- β 1诱导的生长停滞来克服辐射诱导的生长抑制作用。

[0046] 阻抑过度发炎创建了口腔粘膜炎症愈合的微环境。Smad7对TGF- β 与NF- κ B信号传导的拮抗作用使Smad7成为比仅靶向NF- κ B的其它药剂更有效的消炎分子。因为发炎细胞产生进一步活化TGF- β 和NF- κ B的细胞因子,所以在辐射之后在K5.Smad7或Tat-Smad7处理的口腔粘膜中所发现的减少的TGF- β 和NF- κ B信号传导反映了Smad7对这两个路径的直接拮抗作用和来自浸润性白细胞的发炎细胞因子减少的结果。然而,Smad7不会使NF- κ B或TGF- β 信号传导降低至低于其正常生理状态。NF- κ B或TGF- β 信号传导的这种不完全阻断可以有益于口腔粘膜炎症愈合,这是因为任一路径的完全损失可以诱导过度发炎。

[0047] 使用生长因子治疗癌症患者的口腔粘膜炎症的主要障碍是促进癌细胞生长的潜在风险。大多数人类口腔癌损失了肿瘤上皮细胞中的TGF- β 信号传导。因此,由Smad7达成的抗Smad相关的细胞增殖和迁移将在癌细胞中无效。在具有完整TGF- β 信号传导的肿瘤中,其它致癌路径的活化可以压制TGF- β 诱导的肿瘤抑制作用。这两种情形可以解释为何没有观测到Smad7增加具有突变或完整TGF- β 信号传导组分的口腔癌细胞的增殖和迁移。

[0048] 另外,TGF- β 信号传导在TGF- β 诱导的肿瘤抑制损失之后主要经由Smad非依赖性机制来促进肿瘤侵入。因此,在癌细胞中由Smad7阻断TGF- β 信号传导可以消去TGF- β 介导的肿瘤促进作用,其表现类似于当前用于晚期癌症的临床试验中的TGF- β 抑制剂。此外,Smad7的强效消炎作用可以降低肿瘤进展的风险。因此,长期施加Smad7也可以有助于癌症治疗。

[0049] 尚未观测到K5.Smad7小鼠中的自发肿瘤形成。因为Smad7不是分泌性蛋白质,所以在口腔粘膜炎症治疗中的局部和短期Smad7蛋白质递送应具有极少的全身作用。在口腔上皮不含癌细胞的骨髓移植患者中,Smad7局部施加可以适合于口腔粘膜炎症的预防与治疗。

[0050] 尽管不希望受任何理论约束,但Smad7介导的口腔粘膜炎症愈合似乎是靶向由一种或多种分子介导的多个发病过程的结果(参看例如图14A-B)。据信,这些分子(例如,TGF- β 、NF- κ B、CtBP1、Rac1)中的一者或多者也可以适用作患者的口腔粘膜炎症的预测和治疗反应标志物。

[0051] A. 核酸、载体以及宿主细胞

[0052] 在另一个实施方案中,本公开还提供了编码Smad7的基因。除了野生型SMAD7基因(SEQ ID NO:22,88)以及各种密码子优化的型式(SEQ ID NO:9、21、23、24、26、28、30、32-

34、36、38、39、87、89、91、93、96、97、99和100)之外,应该明确的是,本发明技术并不限于本文所公开的特定核酸。如下文所论述的“Smad7基因”可以含有多种不同碱基并且仍然产生在功能上与本文所公开的人类基因无法区分并且在一些情况下结构上与本文所公开的人类基因一致的对应多肽。

[0053] 1. 编码Smad7的核酸

[0054] 根据本发明技术的核酸可以代表表达具有一种或多种与Smad7相关的活性的多肽的Smad7的整个Smad7基因、截短部分和/或片段,所述活性诸如(但不限于)增加增殖、减少或抑制细胞死亡、减轻过度发炎、防止DNA损伤和/或增加细胞迁移,以及治疗或预防如本文进一步论述的这类治疗将有帮助的一种或多种疾病或病症。这类活性可以使用一种或多种分析法来评估,包括(但不限于)阻断Smad2的磷酸化和/或NF- κ B p50亚单元的核转位,增加细胞增殖,减少凋亡和/或辐射诱导的DNA损伤,减少发炎和/或血管生成,促进口腔粘膜炎、手术创口、糖尿病创口和/或与小鼠慢性发炎相关的创口的愈合的能力。核酸可以来源于基因组DNA,即,直接从特定生物体的基因组克隆。然而,在特定实施方案中,核酸将包含互补DNA(cDNA)。还提供了加上天然内含子或来源于另一种基因的内含子的cDNA;这类工程改造的分子有时被称作“微小基因”。在最低限度上,本发明技术的这些和其它核酸可以在例如凝胶电泳中用作分子量标准。

[0055] 术语“cDNA”意图指代使用信使RNA(mRNA)作为模板制备的DNA。与基因组DNA或从基因组、未加工或部分加工的RNA模板聚合的DNA相反,使用cDNA的优点在于cDNA主要含有对应蛋白质的编码序列。可能有时候完全或部分基因组序列是优选的,诸如在最佳表达需要非编码区的情况下或在反义策略中将要靶向诸如内含子的非编码区的情况下。

[0056] 如本申请中所用的术语“编码Smad7的核酸”可以指代已经从总细胞核酸中分离的核酸分子和/或可以指代编码Smad7多肽的cDNA。如本文所用的术语“从总细胞核酸中分离”意指如使用标准生物化学技术,诸如(但不限于)琼脂糖凝胶电泳所测定,核酸分子为约或至少约75%纯、80%纯、85%纯、90%纯、95%纯、96%纯、97%纯、98%纯、99%纯或100%纯的其它细胞核酸分子。如本文所用的术语“从总细胞蛋白质中分离”意指如使用标准生物化学技术,诸如(但不限于)西方墨点所测定,蛋白质分子为约或至少约75%纯、80%纯、85%纯、90%纯、95%纯、96%纯、97%纯、98%纯、99%纯或100%纯的其它细胞核酸分子。在某些实施方案中,本发明技术涉及基本上如SEQ ID NO:9、21、23、24、26、28、30、32-34、36、38、39、87、89、91、93、96、97、99和100中的任一者所陈述和/或包括SEQ ID NO:9、21、23、24、26、28、30、32-34、36、38、39、87、89、91、93、96、97、99和100中的任一者的核酸序列。

[0057] 分离的核酸分子可以使用重组DNA技术(例如,聚合酶链反应(PCR)扩增、克隆)或化学合成来产生。分离的核酸分子包括天然核酸分子和其同源物,包括(但不限于)天然等位基因变体和修饰的核酸分子,其中核苷酸已经按使得这类修饰提供所需作用(例如,在非人类表达系统中产生Smad7蛋白质)的方式插入、缺失、取代和/或倒转。

[0058] 术语“基本上如一个或多个核酸序列(例如,SEQ ID NO:9-11、21、23-41)所陈述”意指核酸序列实质上对应于一个或多个核酸序列(例如,SEQ ID NO:9、21、23、24、26、28、30、32-34、36、38、39、87、89、91、93、96、97、99和100)的至少一部分并且在一些情况下全部。在一些实施方案中,实质上对应于核酸序列的至少一部分的序列可以对应于本文所描述的序列中的一者或多者的约或至少约50个核酸、75个核酸、150个核酸、200个核酸、250个核

酸、300个核酸、350个核酸、400个核酸、450个核酸、500个核酸、550个核酸、600个核酸、650个核酸、700个核酸、750个核酸、800个核酸、900个核酸、1000个核酸、1100个核酸、1200个核酸或1250个核酸。在一些实施方案中,实质上对应于核酸序列的至少一部分的序列可以对应于约以下范围:本文所描述的序列中的一者或多者的约50-1250个核酸、75-1250个核酸、150-1250个核酸、200-1250个核酸、250-1250个核酸、300-1250个核酸、350-1250个核酸、400-1250个核酸、450-1250个核酸、500-1250个核酸、550-1250个核酸、600-1250个核酸、650-1250个核酸、700-1250个核酸、750-1250个核酸、800-1250个核酸、900-1250个核酸、1000-1250个核酸、1100-1250个核酸、1200-1250个核酸、至少约50-75个核酸、75-150个核酸、75-200个核酸、75-250个核酸、75-300个核酸、75-350个核酸、75-400个核酸、75-450个核酸、75-500个核酸、75-550个核酸、75-600个核酸、75-650个核酸、75-700个核酸、75-750个核酸、75-800个核酸、75-900个核酸、75-1000个核酸、75-1100个核酸、75-1200个核酸、或75-1250个核酸或1250个核酸。

[0059] 在一些实施方案中,实质上对应于核酸序列的至少一部分的序列包括与核酸序列的那个部分一致的序列。在一些实施方案中,实质上对应于核酸序列的至少一部分或核酸序列的全部的序列可以包括一个或多个功能等效密码子。术语“功能等效密码子”在本文中用于指代一个或多个编码相同氨基酸的密码子,诸如对于精氨酸或丝氨酸的六个密码子,并且在一些实施方案中指的是如以下页中所论述的编码生物等效氨基酸的密码子。术语“生物等效”氨基酸在本文中用于指代如下一个或多个氨基酸:当从存在于人类Smad7野生型蛋白质的氨基酸序列中的氨基酸变化时,并不改变本文所描述的Smad7的生物活性中的一者或多者(或在一些实施方案中,任何),诸如(但不限于)增加增殖、减少或抑制细胞死亡、减轻过度发炎、防止DNA损伤和/或增加细胞迁移,以及治疗或预防如本文进一步论述的这类治疗将有帮助的一种或多种疾病或病症。

[0060] 在允许遗传密码简并的一些实施方案中,具有约或至少约60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%和/或99%的与密码子优化核酸序列(例如,SEQ ID NO:9-11、21、23-41)中的任一者的核苷酸一致的核苷酸的序列可以被视为实质上对应的核酸序列。基本上与核酸序列(例如,SEQ ID NO:9-11、21、23-41)中的任一者所陈述的那些相同的序列也可以在功能上定义为能够在各种标准条件下与含有SEQ ID NO:9-11、21、23-41的补体的核酸区段杂交的序列。

[0061] 对于需要高选择性的应用,我们将通常希望采用相对高的严格度条件来形成杂交体。举例来说,相对低盐和/或高温条件,诸如在约50℃至约70℃的温度下由约0.02M至约0.10M NaCl所提供。这类高严格度条件容忍探针或引物与模板或目标链之间的极小(如果有的话)错配并且将特别适合于分离特定基因或检测特定mRNA转录物。一般应了解,可以通过添加越来越多量的甲酰胺而促使条件更加严格。

[0062] 对于某些应用,应了解,较低严格度条件是优选的。在这些条件下,杂交甚至可以在杂交链的序列并不完美互补,但在一个或多个位置处错配的情况下发生。可以通过增加盐浓度和/或降低温度而促使条件不太严格。举例来说,中等严格度条件可以在约37℃至约55℃的温度下由约0.1至0.25M NaCl提供,而低严格度条件可以在约20℃至约55℃范围内的温度下由约0.15M至约0.9M盐提供。取决于所需结果,可以容易地操纵杂交条件。

[0063] 在其它实施方案中,杂交可以在例如50mM Tris-HCl (pH 8.3)、75mM KCl、3mM

MgCl₂、1.0mM二硫苏糖醇的条件下,在介于约20℃至约37℃之间的温度下达成。所利用的其它杂交条件可以包括约10mM Tris-HCl (pH 8.3)、50mM KCl、1.5mM MgCl₂,在约40℃至约72℃范围内的温度下。

[0064] 为了确定两个氨基酸序列或两个核酸的同源性%,出于最佳比较的目的对序列进行比对(例如,将空位引入第一氨基酸或核酸序列的序列中用于与第二氨基酸或核酸序列进行最佳比对)。然后可以比较在对应的氨基酸位置或核苷酸位置处的氨基酸残基或核苷酸。当第一序列中的一个位置被与第二序列中的对应位置相同的氨基酸残基或核苷酸占据时,那么分子在那个位置处是一致的。两个序列之间的同源性%是由序列共用的相同位置的数目的函数(一致性%=相同位置数/位置总数(例如,重叠位置)×100)。在一些实施方案中,两个序列是相同的长度。

[0065] 为了确定两个序列之间的同源性%,可以使用Karlin和Altschul (1990) Proc.Natl.Acad.Sci.USA 87:2264-2268的算法,如在Karlin和Altschul (1993) Proc.Natl.Acad.Sci.USA 90:5873-5877中加以修改。将这样的算法并入Altschul等 (1990) J.Mol Biol.215:403-410的NBLAST和XBLAST程序中。使用NBLAST程序、得分=100、字长=12进行BLAST核苷酸搜索以获得与本文所描述或公开的核酸分子同源的核苷酸序列。使用XBLAST程序、得分=50、字长=3进行BLAST蛋白质搜索。为了获得空位比对用于比较的目的,可以利用如Altschul等 (1997) Nucleic Acids Res.25:3389-3402中所描述的空位BLAST。当利用BLAST和空位BLAST程序时,使用相应程序(例如,XBLAST和NBLAST)的默认参数。关于进一步的细节,参看国家生物技术信息中心(National Center for Biotechnology Information)的网站。适合用于本文所描述的方法中的蛋白质还包括如下蛋白质:与本文所描述的任何蛋白质的氨基酸序列相比,具有1个至15个之间的氨基酸变化,例如1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个或15个氨基酸取代、缺失或添加。在其它实施方案中,改变的氨基酸序列与本文所描述的任何蛋白质抑制剂的氨基酸序列至少75%一致,例如,77%、80%、82%、85%、88%、90%、92%、95%、97%、98%、99%或100%一致。这类序列变体蛋白质适合用于本文所描述的方法,只要改变的氨基酸序列保留足够的生物活性以在本文所描述的组合物和方法中发挥功能即可。在某些情况下,利用保守氨基酸取代。在氨基酸当中说明性的保守取代在以下群组中的每一者内:(1) 甘氨酸、丙氨酸、缬氨酸、亮氨酸以及异亮氨酸,(2) 苯丙氨酸、酪氨酸以及色氨酸,(3) 丝氨酸和苏氨酸,(4) 天冬氨酸和谷氨酸,(5) 谷氨酰胺和天冬酰胺,以及(6) 赖氨酸、精氨酸以及组氨酸。BLOSUM62表是来源于蛋白质序列区段的约2,000个局部多重比对的氨基酸取代矩阵,这些区段代表相关蛋白质的多于500个群组的高度保守区(Henikoff等 (1992),Proc.Natl Acad.Sci.USA,89:10915-10919)。BLOSUM62取代频率可以用于定义在一些实施方案中引入本文所描述或公开的氨基酸序列中的保守氨基酸取代。尽管有可能仅仅基于化学特性(如上文所论述)来设计氨基酸取代,但语言“保守氨基酸取代”优选地指代由大于-1的BLOSUM62值代表的取代。举例来说,如果氨基酸取代的特征为0、1、2或3的BLOSUM62值,那么这个取代是保守的。根据这个系统,优选的保守氨基酸取代的特征为至少1的BLOSUM62值(例如,1、2或3),而更优选的保守氨基酸取代的特征为至少2(例如,2或3)的BLOSUM62值。

[0066] 本发明技术的DNA区段包括如上文所描述的编码生物功能等效的Smad7蛋白质和

肽的那些。这类序列可以作为已知天然存在于核酸序列和由此编码的蛋白质内的密码子冗余和氨基酸功能等效的结果而产生。或者,功能等效蛋白质或肽可以经由应用重组DNA技术而创建,其中可以基于正在交换的氨基酸的特性的考虑因素对蛋白质结构中的变化进行工程改造。如别处所描述,由人所设计的变化可以经由应用定点诱变技术而引入或可以随机地引入并且稍后针对所需功能进行筛选。

[0067] 如下文更详细地描述,Smad7核酸序列已经针对在替代性宿主生物体(例如,非人类)中的表达进行优化。尽管如上文所描述,遗传密码是简并的,但如此频繁地,一个氨基酸可以由两个或更多个核苷酸密码子编码。因此,多个核酸序列可以编码一个氨基酸序列。尽管这创建了一致的蛋白质,但核酸本身是不同的并且可以具有不同的特性。如本文所描述,密码子使用的选择的一个方面可以是(但不限于)表达非原生细胞中的蛋白质(例如,细菌或酵母中的人类蛋白质)的能力,或在这类细胞中的表达水平。为了获得足够的蛋白质用于在体外分析法中、在动物模型中并且最终在临床发展中进行纯化、测试以及使用,需要在非人类系统中的有效蛋白质表达。

[0068] 已经鉴别了在由AGG(1.7%密码子利用率;9个残基)、AGA(2.8%密码子利用率;2个残基)、CGA(3.5%密码子利用率;4个残基)或CGG(5.4%密码子利用率;8个残基)中的一者或多者编码的人类Smad7蛋白质序列中的一系列的23个精氨酸氨基酸,并且已经确定,为了具有来自非人类来源(诸如(但不限于)细菌和/或酵母)的有效蛋白质表达,一个或多个并且潜在地所有精氨酸密码子应被修饰为CGT(20.6%密码子利用率)。因此,在一些实施方案中,Smad7密码子优化核酸序列包括至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少11个、至少12个、至少13个、至少14个、至少15个、至少16个、至少17个、至少18个、至少19个、至少20个、至少21个、至少22个、或23个已经变成CGT的精氨酸密码子。在一些实施方案中,Smad7密码子优化核酸序列在核酸序列位置7-9、43-45、169-171、403-405、490-492、526-528、526-528、823-825、1057-1059、16-18、136-138、199-201、598-600、31-33、112-114、316-318、772-774、940-942、973-975、1135-1137、1276-1278、637-639或814-816处包括一个或多个或所有变成CGT的精氨酸密码子。

[0069] 已经鉴别了在由TCC或TCG(9%)编码的人类Smad7蛋白质序列中的一系列的33个丝氨酸残基,并且已经确定,对于来自非人类来源(诸如(但不限于)细菌和/或酵母)的有效蛋白质表达和纯化可能有益的是,一个或多个并且潜在地所有丝氨酸密码子被修饰为AGC(15%密码子利用率)。因此,在一些实施方案中,Smad7密码子优化核酸序列包括至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少11个、至少12个、至少13个、至少14个、至少15个、至少16个、至少17个、至少18个、至少19个、至少20个、至少21个、至少22个、至少23个、至少24个、至少25个、至少26个、至少27个、至少28个、至少29个、至少30个、至少31个、至少32个或33个已经变成(AGC)的丝氨酸密码子。在一些实施方案中,Smad7密码子优化核酸序列在核酸序列位置19-21、46-48、133-135、292-294、349-351、451-453、454-456、460-462、511-513、514-516、544-546、595-597、616-618、634-636、691-693、694-696、739-741、745-747、775-777、847-849、907-909、919-921、943-945、1006-1008、1009-1101、1030-1032、1054-1056、1093-1095、1126-1128、1192-1194、1237-1239、1240-1242、1273-1275处包括一个或多个或所有丝氨酸密码子。在这些当中,可以改变23个密码子(19-21、292-294、349-351、451-453、454-456、460-462、511-513、514-

516-544-546、616-618、634-636、691-693、694-696、739-741、745-747、775-777、847-849、907-909、919-921、1009-1101、1030-1032、1054-1056、1093-1095), 而不引入潜在的替代性开放阅读框架。

[0070] 还已经鉴别了在由CAC (9.6%密码子使用率) 编码的人类Smad7蛋白质序列中的一系列的12个组氨酸残基, 并且已经确定, 对于来自非人类来源 (诸如 (但不限于) 细菌和/或酵母) 的有效蛋白质表达和纯化可能有益的是, 一个或多个并且潜在地所有丝氨酸密码子被修饰为CAT (任选地达到12.6%使用率)。因此, 在一些实施方案中, Smad7密码子优化核酸序列包括至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少11个、或12个已经变成 (CAT) 的组氨酸密码子。在一些实施方案中, Smad7密码子优化核酸序列在核酸序列位置142-144、214-216、217-219、220-222、226-228、289-291、589-591、778-780、1072-1074、1147-1149处包括一个或多个或所有丝氨酸密码子。在这些当中, 可以改变4个密码子 (核苷酸217-219、220-222、589-591、778-780), 而不引入潜在的替代性开放阅读框架。

[0071] 在一些实施方案中, 一个或多个密码子优化核酸可以包括至少一个和任何整数至22个修饰为CGT的其精氨酸密码子、至少一个和任何整数至28个修饰为AGC的其丝氨酸密码子 (任选地能够通过引入开放阅读框架而修饰)、或至少一个和任何整数至12个修饰为CAT的其组氨酸密码子 (任选地能够通过引入开放阅读框架而修饰) 中的一者或多者。在一些实施方案中, 一个或多个密码子优化核酸可以包括至少一个和任何整数至22个修饰为CGT的其精氨酸密码子、至少一个和任何整数至28个修饰为AGC的其丝氨酸密码子 (任选地能够通过引入开放阅读框架而修饰)、以及至少一个和任何整数至12个修饰为CAT的其组氨酸密码子 (任选地能够通过引入开放阅读框架而修饰)。在一些实施方案中, 一个或多个密码子优化核酸可以包括22个修饰为CGT的其精氨酸密码子、28个修饰为AGC的其丝氨酸密码子 (任选地能够通过引入开放阅读框架而修饰)、以及12个修饰为CAT的其组氨酸密码子 (任选地能够通过引入开放阅读框架而修饰)。在一些实施方案中, 一个或多个密码子优化核酸还可以在Met216 (ATG) 的密码子中具有核苷酸取代, 以形成Leu216 (CTG) 的密码子。

[0072] 在一些实施方案中, 一个或多个密码子优化核酸与人类Smad7野生型cDNA (SEQ ID NO: 12、22) 可以具有约65%至75%、约65%至68%、约68%至75%、或约68%至71%的同源性。在一些实施方案中, 一个或多个密码子优化核酸与人类Smad7野生型cDNA (SEQ ID NO: 22、88) 可以具有约65%、66%、67%、68%、69%、70%、71%、72%、73%、74%或75%的同源性。在一些实施方案中, 一个或多个密码子优化核酸还可以在Met216 (ATG) 的密码子中具有核苷酸取代, 以形成Leu216 (CTG) 的密码子。

[0073] 已经鉴别了具有被翻译机构 (诸如 (但不限于) 细菌或酵母) 认为是替代性开放阅读框架的潜力的甲硫氨酸密码子 (Met216; ATG)。尽管不打算受理论约束, 但据信第二潜在开放阅读框架的存在可以减少Smad7蛋白质的表达。在一些实施方案中, 在核苷酸位置 (646-648) 处修饰一个或多个Smad7核酸序列以编码Met216 (ATG) 被修饰为Leu216 (CTG) 的人类Smad7蛋白质。

[0074] 还已经发现, Smad7蛋白质的各种截短形式和片段保留全长人类Smad7的活性中的一者或多者, 诸如 (但不限于) 增加增殖、减少或抑制细胞死亡、减轻过度发炎、防止DNA损伤和/或增加细胞迁移, 以及治疗或预防如本文进一步论述的这类治疗将有帮助的一种或多

种疾病或病症。这类活性可以使用一种或多种分析法来评估,包括(但不限于)阻断Smad2的磷酸化和/或NF- κ B p50亚单元的核转位,增加细胞增殖,减少凋亡和/或辐射诱导的DNA损伤,减少发炎和/或血管生成,促进口腔粘膜炎、手术创口、糖尿病创口和/或与小鼠慢性发炎相关的创口的愈合的能力。

[0075] 此外,在一些实施方案中,Smad7蛋白质的各种截短形式和片段仅保留全长人类Smad7的活性中的一者或多者的子组。举例来说,Smad7的C端MH2域可以主要介导Smad7的消炎作用。具有这种消炎功能的Smad7肽可能是足够的并且任选地是一种改善,用于治疗慢性发炎相关的病状,尤其诸如(但不限于)口腔粘膜炎、口炎、关节炎以及牛皮癣。N端MH1域可以主要介导细胞迁移和/或阻断TGF- β 诱导的生长停滞和/或纤维化反应。具有这种细胞迁移和增殖功能的Smad7肽可能是足够的并且任选地是一种改善,用于增强与过度发炎不相关的愈合。可能得益于这种形式的治疗的创口类型尤其包括(但不限于)手术创口、纤维化疤痕以及糖尿病创口、缺陷性愈合和/或疤痕。

[0076] 在一些实施方案中,核酸分子(任选地,如上文和本文所描述的密码子优化核酸分子)编码Smad7蛋白质(任选地包括Leu216)的片段或截短形式。在一些实施方案中,Smad7蛋白质的这些片段和/或截短形式保留全长人类Smad7蛋白质的一种或多种或所有活性。在一些实施方案中,这类截短的核酸序列编码Smad7蛋白质的N端部分。在一些实施方案中,这类截短的核酸序列编码Smad7蛋白质的C端部分。在一些实施方案中,这类截短的核酸序列(核苷酸位置4-774)编码人类Smad7蛋白质的氨基酸2-258。在一些实施方案中,这类截短的核酸序列(核苷酸位置775-1278)编码人类Smad7蛋白质的氨基酸259-426。在一些实施方案中,核酸序列的这类片段(核苷酸位置610-774)编码人类Smad7蛋白质的氨基酸204-258。

[0077] 如本文中关于核酸分子所用的术语“截短”指的是含有编码对应蛋白质(具有或不具有裂解的前导序列)的天然N端的核苷酸序列,但缺乏一个或多个从分子的C端编码区开始的核苷酸的分子,或含有编码对应蛋白质(具有或不具有裂解的前导序列)的天然C端的核苷酸序列,但缺乏一个或多个从分子的N端编码区开始的核苷酸的分子。在一些实施方案中,特定地提供缺乏编码来自一端或另一端的至少约25个、至少约50个、至少约75个、至少约100个、至少约125个、至少约150个、至少约200个、至少约250个、至少约300个、或至少约350个、或至少约400个氨基酸的核苷酸的分子。类似地,术语“截短”也可以关于由截短的核酸分子编码的蛋白质分子而使用。在一些实施方案中,“截短”分子是生物活性的,其具有本文所描述的Smad7活性中的一者或多者(或编码具有这些活性的多肽)。

[0078] 如本文中关于核酸分子所用的术语“片段”指的是含有全长序列的相连残基,但缺乏全长序列的一些5'和/或3'序列的分子。在一些实施方案中,“片段”包括本文所描述的全长序列中的一者或多者的一部分。在一些实施方案中,“片段”不包括编码N端或C端的序列,而仅包括内部片段。在一些实施方案中,“片段”编码生物活性的多肽,这种多肽具有本文所描述的Smad7活性中的一者或多者。在一些实施方案中,核酸片段可以编码具有至少约25个、30个、35个、40个、45个、50个、55个、60个、65个、70个、75个、80个、85个、90个、95个、100个、150个、200个、250个、300个、350个、400个、450个、500个、550个、600个、650个、700个、750个、800个、850个、900个、950个、1000个、1050个、1100个、1150个氨基酸的蛋白质。类似地,“片段”也可以关于由Smad7核酸片段编码的蛋白质分子而使用。

[0079] 如本文所用的术语“N端部分”指的是含有蛋白质的N端,但缺乏内部残基的C端的

所有序列的对应蛋白质的片段。

[0080] 如本文所用的术语“C端部分”指的是含有蛋白质的C端,但缺乏内部残基的N端的所有序列的对应蛋白质的片段。

[0081] 尽管不打算受理论约束,但一般据信Smad7蛋白质活性是细胞的细胞质与核中的相互作用的结果。尤其出于那个原因,一般据信Smad7蛋白质不是治疗作用的候选物。然而,决定推行开发Smad7作为蛋白质治疗剂,并且修饰Smad7核酸序列以与Smad7核酸序列(例如,任选地本文所描述的编码Smad7蛋白质的任何核酸序列,包括人类野生型和密码子优化序列、全长与生物活性片段或截短部分)框内编码蛋白质转导域(PTD)。在一些实施方案中,PTD位于Smad7核酸序列的3'末端,并且在一些实施方案中,PTD位于Smad7核酸序列的5'末端。在一些实施方案中,存在连接PTD和Smad7核酸序列的编码1个、2个、3个、4个、5个或6个氨基酸的连接子序列。

[0082] 在一些实施方案中,PTD核酸序列是Tat核酸序列。编码GRKKRRQRRR (SEQ ID NO:2)的ggccgtaaaaaacgccgtcaacgccgccgt (SEQ ID NO:1)、编码YGRKKRRQRRR (SEQ ID NO:4)的tatggccgtaaaaaacgccgtcaacgccgccgt (SEQ ID NO:3)、或编码GRKKRRQ (SEQ ID NO:6)的ggccgtaaaaaacgccgtcaa (SEQ ID NO:5)。

[0083] 在一些实施方案中,核酸序列进一步包括编码表位标签或纯化标签中的一者或多者的核苷酸序列。在一些实施方案中,表位标签是V5。在一些实施方案中,纯化标签是谷胱甘肽-S-转移酶(GST)或6-组氨酸(H6) (SEQ ID NO:40)中的一者或多者。

[0084] 如本文中关于核酸分子所用的术语“表位标签”指的是编码由抗体或片段的可变区识别并结合的肽序列的核苷酸。在一些实施方案中,表位标签不是原生蛋白质的一部分。在一些实施方案中,表位标签是可去除的。在一些实施方案中,表位标签不是蛋白质的原生生物活性所固有的。表位标签的实例包括(但不限于)V5。

[0085] 如本文中关于核酸分子所用的术语“纯化标签”指的是编码有助于蛋白质的纯化,但一般并非蛋白质的生物活性所必需的肽序列的核苷酸。在一些实施方案中,可以在蛋白质纯化之后去除纯化标签。纯化标签的实例包括(但不限于)GST和H-6 (SEQ ID NO:40)。

[0086] 2. 用于克隆、基因转移以及表达的载体

[0087] 在某些实施方案内,采用表达载体来表达Smad7多肽产物,然后可以纯化这种产物用于各种用途。在其它实施方案中,表达载体用于基因疗法中。表达需要在载体中提供适当信号,并且这些载体包括来自病毒与哺乳动物来源的各种调控元件(诸如增强子/启动子),其驱使所关注的基因在宿主细胞中表达。还定义了经过设计以优化宿主细胞中的信使RNA稳定性和可翻译性的元件。还提供了用于建立表达产物的永久的稳定细胞克隆体的许多显性药物选择标志物的使用条件,以及将药物选择标志物的表达与多肽的表达联系起来的元件。

[0088] 在本申请通篇中,术语“表达构建体”意图包括含有编码其中一部分或所有核酸编码序列能够被转录的基因产物的核酸的任何类型的基因构建体。可以将转录物翻译成蛋白质,但不必须如此。在某些实施方案中,表达包括基因转录与mRNA翻译成基因产物。在其它实施方案中,表达仅包括编码所关注的基因的核酸的转录。

[0089] 术语“载体”用于指代核酸序列可以插入其中以供引入可以复制这个核酸序列的细胞中的运载核酸分子。核酸序列可以是“外源的”,意指其对于引入载体的细胞是外来的,

或这个序列与细胞中的序列同源但处于宿主细胞核酸内通常未发现这个序列的位置处。载体包括质粒、粘粒、病毒(噬菌体、动物病毒以及植物病毒)以及人工染色体(例如,YAC)。本领域的技术人员将做充分准备以经由标准重组技术构建载体,这些技术描述于例如 Sambrook等,Molecular Cloning(Cold Spring Harbor Lab Press,1989)和Ausubel等,Current Protocols in Molecular Biology(Wiley,1994)中,两者都以引用的方式并入本文中。

[0090] 术语“表达载体”指的是含有编码能够被转录的基因产物的至少一部分的核酸序列的载体。在一些情况下,然后将RNA分子翻译成蛋白质、多肽或肽。在其它情况下,例如,在产生反义分子或核糖酶中,不翻译这些序列。表达载体可以含有多种“控制序列”,其指的是对于操作性地连接的编码序列在特定宿主生物体中的转录和可能翻译所必需的核酸序列,包括启动子和增强子。除了主导转录和翻译的控制序列之外,载体和表达载体可以含有发挥其它功能的核酸序列,诸如转录终止信号和多聚腺苷酸化位点。

[0091] 某些病毒载体有效地感染或进入细胞以整合至宿主细胞基因组中并且稳定表达病毒基因的能力已经促成许多不同病毒载体系统的开发和应用。Robbins等,Pharmacol.Ther.80:35-47(1998)。病毒系统当前用作离体和体内基因转移的载体。举例来说,当前正在评价腺病毒、单纯疱疹病毒、慢病毒、逆转录病毒以及腺相关病毒用于治疗诸如癌症、囊性纤维化、戈谢病(Gaucher disease)、肾病以及关节炎的疾病。Robbins等,Pharmacol.Ther.80:35-47(1998);Imai等,Nephrologie 19:379-402(1998);美国专利5,670,488。取决于特定的基因治疗应用,各种病毒载体呈现特定的优点和缺点。

[0092] 如本文所描述或如本领域的一般技术人员将知晓,据信用于本发明技术的适合用于细胞器、细胞、组织或生物体转化的核酸递送的非病毒方法包括可以将核酸(例如,DNA)引入细胞器、细胞、组织或生物体中的几乎任何方法。这类方法包括(但不限于)DNA的直接递送,诸如通过注射(美国专利5,994,624、5,981,274、5,945,100、5,780,448、5,736,524、5,702,932、5,656,610、5,589,466以及5,580,859,各自以引用的方式并入本文中),包括显微注射(Harland和Weintraub,1985;美国专利5,789,215,以引用的方式并入本文中);通过电穿孔(美国专利5,384,253,以引用的方式并入本文中);通过磷酸钙沉淀(Graham等,Virology 52:456-467(1973);Chen等,Mol.Cell Biol.7:2745-2752(1987);Rippe等,Mol.Cell Biol.10:689-695(1990));通过使用DEAE-葡聚糖、继之以聚乙二醇(Gopal,Mol.Cell Biol.5:1188-1190(1985));通过直接声波加载(Fechheimer等,PNAS 84:8463-8467(1987));通过脂质体介导的转染(Nicolau等,Biochim.Biophys.Acta 721:185-190(1982);Fraleigh等,PNAS 76:3348-3352(1979);Nicolau等,Methods Enzymol.149:157-176(1987);Wong等,Gene 10:87-94(1980);Kaneda等,J.Biol.Chem.264:12126-12129(1989);Kato等,J.Biol.Chem.266:3361-3364(1991));通过微弹轰击(PCT申请号W0 94/09699和95/06128;美国专利5,610,042、5,322,783、5,563,055、5,550,318、5,538,877以及5,538,880,并且各自以引用的方式并入本文中);通过与碳化硅纤维一起搅动(Kaeppler等,Plant Cell Rep.9:415-418(1990);美国专利5,302,523和5,464,765,各自以引用的方式并入本文中);或通过PEG介导的原生质体转化(Omirulלה等,Plant Mol.Biol.21:415-428(1993);美国专利4,684,611和4,952,500,各自以引用的方式并入本文中);通过干燥/抑制介导的DNA摄入(Potrykus等,Mol.Gen.Genet.199:169-177(1985))。通过应用诸如此类的

技术,可以稳定或瞬时转化细胞器、细胞、组织或生物体。

[0093] 3. 表达系统

[0094] 存在众多表达系统,其包含上文所论述的组合物的至少一部分或全部。可以采用基于原核生物和/或真核生物的系统用于本发明技术以产生核酸序列或其同源多肽、蛋白质以及肽。许多这类系统在商业上并且广泛可得。

[0095] 昆虫细胞/杆状病毒系统可以产生异源核酸区段的高水平的蛋白质表达,诸如美国专利5,871,986和4,879,236中所描述,两者都以引用的方式并入本文中,并且这种系统可以例如以名称**MaxBac[®] 2.0**从**Invitrogen[®]**并且以**BacPack[™]**杆状病毒表达系统从**Clontech[®]**获取。

[0096] 表达系统的其它实例包括**STRATAGENE[®]**的**COMPLETE CONTROL[™]**诱导性哺乳动物表达系统,这个系统涉及合成蜕皮激素诱导性受体或其pET表达系统,一种大肠杆菌表达系统。诱导性表达系统的另一个实例可获自**INVITROGEN[®]**,其携带T-REX[™](四环素调控表达)系统,一种使用全长CMV启动子的诱导性哺乳动物表达系统。**INVITROGEN[®]**还提供了酵母表达系统,称为甲醇毕赤酵母(*Pichia methanolica*)表达系统,其经过设计用于在甲醇营养型酵母甲醇毕赤酵母中以高水平产生重组蛋白质。本领域的技术人员将知晓如何表达载体,诸如表达构建体,以产生核酸序列或其同源多肽、蛋白质或肽。

[0097] 初级哺乳动物细胞培养物可以按各种方式制备。为了使细胞在体外并且与表达构建体接触时保持活的,必须要确保细胞维持与正确比率的氧气和二氧化碳以及营养素接触,但使其免受微生物污染。细胞培养技术有充分文献记载。

[0098] 前述内容的一个实施方案涉及使用基因转移以使细胞永生化用于产生蛋白质。如上文所描述可以将所关注的蛋白质的基因转移至适当宿主细胞中,继而在适当条件下培养细胞。可以按这种方式采用几乎任何多肽的基因。上文论述了重组表达载体和其中所包括的元件的产生。或者,将要产生的蛋白质可以是通常由所讨论的细胞合成的内源蛋白质。

[0099] 适用的哺乳动物宿主细胞系的实例是Vero和HeLa细胞以及中国仓鼠卵巢细胞系、W138、BHK、COS-7、293、HepG2、NIH3T3、RIN以及MDCK细胞。另外,可以选择宿主细胞株,其调节所插入的序列的表达,或以所需方式修饰并加工基因产物。蛋白质产物的这类修饰(例如,糖基化)和加工(例如,裂解)可能对于蛋白质的功能是重要的。不同宿主细胞具有蛋白质的翻译后加工和修饰的特征性和特定机制。可以选择适当细胞系或宿主系统以确保所表达的外来蛋白质的正确修饰和加工。

[0100] 可以使用许多选择系统,包括(但不限于)HSV胸苷激酶、次黄嘌呤-鸟嘌呤磷酸核糖转移酶以及腺嘌呤磷酸核糖转移酶基因,分别在tk细胞、hgp^{rt}细胞或ap^{rt}细胞中。而且,抗代谢物抗性可以用作对以下进行选择的基础:赋予抗性的dhfr;赋予对霉酚酸的抗性的gpt;赋予对氨基糖苷G418的抗性的neo;以及赋予对潮霉素的抗性的hygro。

[0101] 如本文所用的术语“细胞”、“细胞系”以及“细胞培养物”可以互换使用。所有这些术语还包括其子代,即任何和所有后代。应了解,所有子代由于有意或无意突变而可能不是相同的。在表达异源核酸序列的情形中,“宿主细胞”指的是原核或真核细胞,并且其包括能够复制载体和/或表达由载体编码的异源基因的任何可转化生物体。宿主细胞可以并且已

经用作载体的接受者。宿主细胞可以被“转染”或“转化”，其指的是将外源核酸转移或引入宿主细胞中的过程。转化的细胞包括初级主题细胞和其子代。

[0102] 宿主细胞可以来源于原核生物或真核生物(例如,细菌或酵母),这取决于所需结果是载体的复制还是载体编码的核酸序列的一部分或全部的表达。众多细胞系和培养物是可得的使用作宿主细胞,并且其可以经由美国菌种保藏中心(American Type Culture Collection, ATCC)获得,ATCC是充当活培养物和遗传材料的档案库的组织机构(atcc.org)。适当宿主可以由本领域的技术人员基于载体骨架和所需结果来确定。举例来说,可以将质粒或粘粒引入原核生物宿主细胞中用于许多载体的复制。用作载体复制和/或表达的宿主细胞的细菌细胞包括DH5 α 、JM109和KC8,以及许多可商购的细菌宿主,诸如SURE[®]感受态细胞和Solopack[™]黄金细胞(STRATAGENE[®], La Jolla)。或者,诸如大肠杆菌LE392的细菌细胞可以用作噬菌体病毒的宿主细胞。

[0103] 用于载体的复制和/或表达的真核宿主细胞的实例包括HeLa、NIH3T3、Jurkat、293、Cos、CHO、Saos以及PC12。来自各种细胞类型和生物体的许多宿主细胞是得得的并且将为本领域的技术人员所知。类似地,病毒载体可以与真核或原核宿主细胞联合使用,特别是允许用于载体的复制或表达的宿主细胞。

[0104] 一些载体可以采用允许其在原核与真核细胞中复制和/或表达的控制序列。本领域的技术人员将进一步了解孵育所有上述宿主细胞以维持其并且允许载体复制的条件。还了解和知晓将允许大规模产生载体以及产生由载体编码的核酸和其同源多肽、蛋白质或肽的技术和条件。

[0105] B. Smad7蛋白质和蛋白质片段

[0106] 母亲抗十五褶同源物7(mothers against decapentaplegic homolog 7, Smad7)先前被鉴别为经由若干机制的TGF- β 信号传导的拮抗剂,这些机制包括:(a) TGF- β 受体介导的信号传导Smad磷酸化和核转位的阻断;(b) 经由特定泛素-蛋白酶体路径而增加的TGF- β 受体和信号传导Smad降解;以及(c) 信号传导Smad对于其结合至Smad结合元件(SBE)的抑制。Smad7还拮抗其它信号传导路径,如NF- κ B路径。

[0107] Smad7蛋白质由上文所论述的SMAD7基因编码。如同许多其它TGF- β 家族成员一样,Smad7涉及于细胞信号传导中。其为TGF- β 1型受体拮抗剂。其阻断TGF- β 1和与受体相关的激活素(activin),从而阻断接近Smad2。其为抑制性Smad(I-SMAD)并且由SMURF2增强。Smad7还增强肌肉分化。

[0108] 术语“多肽”、“肽”以及“蛋白质”在本文中可互换使用以指代氨基酸残基的聚合物。这些术语适用于天然存在的氨基酸聚合物以及其中一个或多个氨基酸残基是非天然存在的氨基酸(例如,氨基酸类似物)的氨基酸聚合物。如本文所用的这些术语包涵任何长度的氨基酸链,包括全长蛋白质,其中氨基酸残基由共价肽键连接。

[0109] 在一个实施方案中,本发明技术涉及Smad7蛋白质组合物。除了整个Smad7分子之外,本发明技术还涉及多肽的截短部分和片段,其保留一种或多种与Smad7相关的活性,诸如(但不限于)增加增殖、减少或抑制细胞死亡、减轻过度发炎、防止DNA损伤和/或增加细胞迁移,以及治疗或预防如本文进一步论述的这类治疗将有帮助的一种或多种疾病或病症。这类活性可以使用一种或多种分析法来评估,包括(但不限于)阻断Smad2的磷酸化和/或NF- κ B p50亚单元的核转位,增加细胞增殖,减少凋亡和/或辐射诱导的DNA损伤,减少发炎

和/或血管生成,促进口腔粘膜炎、手术创口、糖尿病创口和/或与小鼠慢性发炎相关的创口的愈合的能力。

[0110] 蛋白质片段可以通过在编码区内翻译终止位点的遗传工程改造(下文所论述)来产生。或者,用蛋白水解酶(称为蛋白酶)处理Smad7分子可以产生多种N端、C端以及内部片段。这些片段可以根据已知方法来纯化,诸如沉淀(例如,硫酸铵)、HPLC、离子交换色谱法、亲和色谱法(包括免疫亲和色谱法)或各种尺寸分离(沉降、凝胶电泳、凝胶过滤)。

[0111] 如本文所用,在本发明的实施方案中提及分离的蛋白质或多肽包括全长蛋白质、融合蛋白质、嵌合蛋白质,或这类蛋白质的任何片段(截短形式、部分)或同源物。更具体地说,分离的蛋白质可以是已经从其天然环境中去除(即,已经经受人类操纵)的蛋白质(包括多肽或肽),并且可以包括(但不限于)纯化的蛋白质、部分纯化的蛋白质、重组产生的蛋白质、与脂质复合的蛋白质、可溶性蛋白质、合成产生的蛋白质,以及与其它蛋白质相关的分离蛋白质。这样的话,“分离”并不反映蛋白质已经纯化至哪种程度。优选地,重组产生分离的蛋白质。

[0112] 还提供了Smad7的变体-这些可以是取代、插入或缺失变体。缺失变体缺乏对于活性并非必需的原生蛋白质的一个或多个残基,包括上文和本文所描述的截短突变体。取代变体通常含有在蛋白质内的一个或多个位点处一个氨基酸与另一个交换,并且可以经过设计以调节多肽的一种或多种特性,诸如对蛋白水解性裂解和/或翻译和/或转录(蛋白质表达)的稳定性,而不损失其它功能或特性。这个种类的取代优选地是保守的,即,一个氨基酸被类似形状和电荷的氨基酸置换。保守取代在本领域中是众所周知的并且包括例如每个氨基酸可以被不同氨基酸更换或取代。在制造取代变体中,通常考虑亲水指数、亲水性、电荷以及大小。

[0113] 特定涵盖的Smad7的缺失变体包括截短和片段,例如,包括具有N端序列而没有C端序列、具有C端序列而没有N端序列、或具有内部序列而没有N端或C端序列的多肽分子。特定涵盖的Smad7多肽截短或片段包括(但不限于)包括对应于原生人类Smad7蛋白质序列的氨基酸残基2-258、259-426、204-258的分子。

[0114] 如本文中关于蛋白质序列所用的术语“截短”指的是含有对应蛋白质(具有或不具有裂解的前导序列)的天然N端,但缺乏一个或多个从分子的C端开始的氨基酸的分子,或含有对应蛋白质(具有或不具有裂解的前导序列)的天然C端,但缺乏一个或多个从分子的N端开始的氨基酸的分子。在一些实施方案中,特定地提供缺乏来自一端或另一端的至少约25个、至少约50个、至少约75个、至少约100个、至少约125个、至少约150个、至少约200个、至少约250个、至少约300个、或至少约350个、或至少约400个氨基酸的分子。在一些实施方案中,“截短”分子是生物活性的,其具有本文所描述的Smad7活性中的一者或多者。

[0115] 如本文中关于多肽序列所用的术语“片段”指的是含有全长序列的相连残基,但缺乏全长序列的一些N端和/或C端残基的分子。在一些实施方案中,“片段”包括本文所描述的全长序列中的一者或多者的一部分。在一些实施方案中,“片段”不包括编码N端或C端的序列,而仅包括内部片段。在一些实施方案中,“片段”编码生物活性的多肽,这种多肽具有本文所描述的Smad7活性中的一者或多者。在一些实施方案中,多肽片段编码具有至少约25个、30个、35个、40个、45个、50个、55个、60个、65个、70个、75个、80个、85个、90个、95个、100个、150个、200个、250个、300个、350个、400个、450个、500个、550个、600个、650个、700个、

750个、800个、850个、900个、950个、1000个、1050个、1100个、1150个氨基酸。

[0116] 一个特殊种类的变体是融合蛋白质。这种分子一般具有原生分子的全部或很大部分,其在N端或C端连接至第二多肽的全部或一部分。然而,在一些实施方案中,融合蛋白质可以包括在本公开通篇所描述的片段和/或截短(N端、C端) Smad7蛋白质中的任一者。举例来说,融合可以采用来自其它物种的前导序列以允许蛋白质在异源宿主中重组表达。另一种有用的融合包括添加任选的功能活性域,诸如(但不限于)抗体表位和/或纯化标签(例如,V5:GKPIPNPLLGLDST (SEQ ID NO:41);Flag:KYKDDDDK (SEQ ID NO:42);HA:YPYDVPDYA (SEQ ID NO:43))。另一种类型的融合包括附接可以充当活化或灭活配体的标靶的域,从而允许在递送至受试者后控制融合蛋白质的功能。这类域包括例如类固醇配体结合(例如,ER、PR、GR),其可以由小分子,例如4-羟基它莫西芬(4-hydroxyl tamoxifen)或RU486活化,这些小分子独特地能够活化那些类固醇配体结合域和/或在自然界中不存在并且将因此能够通过这些小分子的存在而完全控制Smad7功能。

[0117] 在本发明技术中有特定效用的融合蛋白质的另一种特定形式是包括蛋白质转导域(PTD),也称为细胞递送域或细胞转导域的融合体。这类域在本领域中已有描述并且一般特征为短的两亲或阳离子肽和肽衍生物,其常常含有多个赖氨酸和精氨酸残基(Fischer, Med.Res.Rev.27:755-795 (2007))。在一些实施方案中,PTD是来自HIV (GRKKRRQRRR (SEQ ID NO:2)、YGRKKRRQRRR (SEQ ID NO:4)或GRKKRRQ (SEQ ID NO:6))或替代地HSV VP16的TAT蛋白质的一个或多个变体。可以使用Tat的替代形式。在一些实施方案中,可以使用连接子以连接一个或多个PTD和SMad7。在一些实施方案中,PTD(任选地,Tat)框内融合或连接至本公开通篇所描述的Smad7全长、片段和/或截短(N端、C端)蛋白质中的任一者的N端和/或C端末端。本发明技术所提供的PTD的其它实例示于表1中。

表 1 - 蛋白质转导域

SEQ ID NO:		SEQ ID NO:	
GALFLGWLGAAGSTMGAKK KRKV	44	QAATATRGRSAASRPTEPR APARSASRPVRPVE	65
RQIKIWFQNRRMKWKK	45	MGLGLHLLVLAALQGAKS KRKV	66
RRMKWKK	46	AAVALLPAVLLALLAPAAAN YKKPKL	67
RRWRRWRRWRRWRR	47	MANLGYWLLALFVTMWTD VGLCKKRPKP	68
RGGRLSYSRRRFSTSTGR	48	LGTYTQDFNKFHTFPQTAIG VGAP	69
YGRKKRRQRRR	4	DPKGDPKGVTVTVTVTVTG KGDPXPD	70
RKKRRQRRR	49	PPPPPPPPPPPPPP	71
YARAAARQARA	50	VRLPPPVRLLPPPVRLLPPP	72
RRRRRRRR	51	PRPLPPRPG	73
KKKKKKKK	52	SVRRRPRPPYLPRRPPPPFFPP RLPPRIPP	74
GWTLSAGYLLGKINLKALA ALAKXIL	53	TRSSRAGLQFPVGRVHRLLR K	75
LLILLRRRIRKQANAHSK	54	GIGKFLHSAKKFGKAFVGEI MNS	76
SRRHHCRSKAKRSRHH	55	KWKLFKKIEKVGQNIRDGII KAGPAVAVVGQATQIAK	77
NRARRNRRRVR	56	ALWMTLLKKVLKAAAKAA LNAVLVGANA	78
RQLRIAGRRLRGRSR	57	GIGAVLKVLTTGLPALISWIK RKRQQ	79
KLIKGRTPIKFGK	58	INLKALAALAKKIL	80
RRIPNRRPRR	59	GFFALIPKIISSPLPKTLLSAV GSALGGSGGQE	81
KLALKLALKALKAALKLA	60	LAKWALKQGFAKLKS	82
KLAKLAKKLAKLAK	61	SMAQDIISTIGDLVKWIIQTV	83

表 1 - 蛋白质转导域

[0119]			NXFTKK	
	GALFLGFLGAAGSTNGAWSQ	62	LLGDFFRKSKEKIGKEFKRIV	84
	PKKKRKV		QRIKQRIKDFLANLVPRTES	
	KETWWETWWTEWSQPKKKR	63	PAWRKAFRWAWRMLKAA	85
	KV			
	LKKLLKKLLKKLLKKLLKKL	64	KLKLKLKLKLKLKLKLKL	86

[0120] 在特定实施方案中,本发明技术提供了其中一个或多个残基已经改变的Smad7的序列变体。举例来说,在一个实施方案中,在人类Smad7序列的位置216处发现的甲硫氨酸残基被修饰为亮氨酸残基(ATG至CTG)。

[0121] C. 治疗方法

[0122] Smad7可治疗的疾病和病症可以包括有包括以下一者或多者的那些:减少的细胞增殖、减少的细胞迁移、增加的细胞死亡、过度发炎、和/或DNA损伤。Smad7相关的疾病和病症可以包括用具有以下一种或多种活性的Smad7蛋白质及其生物活性片段和衍生物治疗有帮助的那些:包括(但不限于)增加增殖、减少或抑制细胞死亡、减轻过度发炎、防止DNA损伤、和/或增加细胞迁移。这类疾病和/或病症可以包括(但不限于)急性(例如,经由手术、搏斗、创伤)和慢性创口(例如,溃疡,诸如糖尿病性、压力性、静脉性)、疤痕、纤维化和异常愈合、粘膜炎(例如,口腔和/或胃肠)、口炎、直肠炎、自体免疫疾病(例如,牛皮癣、关节炎),以及癌症。

[0123] 在一些实施方案中,本文所描述的疾病和/或病症中的一者或多者可以通过向需要这类治疗的受试者提供治疗有效量的本公开所描述的Smad7蛋白质(例如,全长或生物活性截短(例如,N端或C端)或其片段)中的一者或多者来预防、治疗和/或改善。在一些实施方案中,一种或多种Smad7蛋白质是包括PTD域的融合蛋白质。在一些实施方案中,一种或多种Smad7蛋白质包括Leu216。在一些实施方案中,Smad7蛋白质构成包括一种或多种医药学上可接受的赋形剂的医药组合物的一部分。

[0124] 在一些实施方案中,本文所描述的疾病和/或病症中的一者或多者可以通过向需要这类治疗的受试者提供治疗有效量的本公开所描述的编码一种或多种Smad7蛋白质(例如,全长或生物活性截短(例如,N端或C端)或其片段)的核酸分子中的一者或多者来预防、治疗和/或改善。在一些实施方案中,一种或多种核酸分子包括密码子优化核苷酸序列和/或编码Leu216的序列。在一些实施方案中,将一种或多种Smad7核酸分子在包括表达载体的构建体中提供给受试者。在一些实施方案中,Smad7核酸分子(任选地,表达载体的一部分)构成包括一种或多种医药学上可接受的赋形剂的医药组合物的一部分。

[0125] 如本文所用的术语“受试者”或“患者”指的是需要使用本文所描述的治疗中的一者或多者治疗和/或预防的人员或非人类动物。在一些实施方案中,非人类动物包括实验动物,诸如猴、小鼠、大鼠以及兔;家养宠物,诸如狗和猫;以及牲畜,诸如牛、马、猪、山羊以及绵羊。

[0126] 1. 慢性创口

[0127] 慢性创口是在一组有序的阶段中并且在可预测的时间量内以大多数创口愈合的

方式不会愈合的创口；在三个月内不会愈合的创口常常被视为慢性的。慢性创口似乎滞留在一个或多个创口愈合期中。举例来说，慢性创口常常在发炎阶段中保持太长时间。在急性创口中，诸如胶原蛋白的分子的产生与降解之间存在精确平衡；在慢性创口中，这种平衡失去并且降解起太大作用。

[0128] 如本文别处更详细地描述，已经显示PTD-Smad7增强小鼠皮肤模型和粘膜模型中的创口愈合。PTD-Smad7的施加经由局部途径有效，这是创口治疗所需要的。尽管不打算受理论约束，但据信PTD-Smad7可以用以经由多种途径治疗或改善慢性创口，尤其可以包括以下一者或多者：减轻发炎、增加细胞增殖（例如，角质细胞）、增加细胞迁移（例如，角质细胞）、或减少纤维化（例如，经由胶原蛋白的调节）。

[0129] 慢性创口可能永不愈合或可能需要数年来愈合。这些创口导致患者有严重的情绪和身体压力，以及对患者和整个卫生保健系统造成显著的财政负担。急性和慢性创口在以不同速率向愈合进展的一系列创口愈合类型的相对端。绝大多数慢性创口可以被分为三个类别：静脉性溃疡、糖尿病性以及压力性溃疡。少数不属于这些类别的创口可能是由于诸如辐射中毒或局部缺血的原因引起。

[0130] 静脉性和动脉性溃疡，通常在腿部出现的静脉性溃疡占慢性创口的约70%至90%并且主要影响老年人。其被认为是由于存在于静脉中以防止血液倒流的瓣膜的不当功能而引起的静脉高压。局部缺血因功能障碍所致，并且与再灌注损伤组合引起造成创口的组织损伤。

[0131] 糖尿病性溃疡，慢性创口的另一个主要原因，即糖尿病，正在日益流行。糖尿病患者因慢性溃疡所致的截肢风险比一般群体高15%。糖尿病导致神经病变，其抑制伤害感受和疼痛感知。因此，患者可能最初并未察觉腿和脚的小创口，并且可能因此无法预防感染或重复损伤。此外，糖尿病导致免疫损害和小血管损伤，从而阻止组织充分氧合，这可以造成慢性创口。压力也在糖尿病性溃疡的形成中起作用。

[0132] 压力性溃疡，慢性创口的另一个主要类型是压力性溃疡，其通常在患有抑制通常经受压力的身体部分（诸如脚后跟、肩胛骨以及骶骨）的运动的病状（诸如瘫痪）的人群中出现。压力性溃疡由局部缺血引起，局部缺血是在对组织的压力大于毛细血管中的压力时发生并且因此限制血液流入这个区域中。比皮肤需要更多氧气和营养素的肌肉组织显示出因长期压力所致的最坏影响。如同在其它慢性溃疡中一样，再灌注损伤损害组织。

[0133] 慢性创口可能仅影响表皮和真皮，或其可能影响组织一直至筋膜。其原先可能由造成急性创口的相同事物（诸如手术或意外创伤）形成，或其可能因全身感染、血管、免疫或神经功能不全或诸如瘤形成或代谢失调的并存病的结果而形成。尽管不打算受理论约束，但创口变成慢性的原因是身体处理损伤的能力被以下因素压垮：诸如重复创伤、持续压力、局部缺血、或疾病。导致慢性创口的一些主要因素包括（但不限于）局部缺血、再灌注损伤以及细菌定植。

[0134] 局部缺血，局部缺血是创口的形成和维持中的重要因素，尤其当其反复发生（如同往常一样）时或当与患者的高龄组合时。局部缺血导致组织变得发炎并且细胞释放吸引嗜中性粒细胞的因子，诸如白介素、趋化因子、白三烯以及补体因子。

[0135] 虽然其对抗病原体，但嗜中性粒细胞也释放破坏细胞的发炎细胞因子和酶。其重要功能之一在于产生反应性氧物质（ROS）来杀死细菌，为此其使用称为髓过氧化物酶的酶。

由嗜中性粒细胞和其它白细胞产生的酶和ROS破坏细胞并且通过破坏加速愈合的DNA、脂质、蛋白质、ECM以及细胞因子来阻止细胞增殖和创口闭合。嗜中性粒细胞在慢性创口中比其在急性创口中保持更长时间,并且促成慢性创口具有更高水平的发炎细胞因子和ROS这一事实。因为来自慢性创口的创口流体具有过量的蛋白酶和ROS,所以流体本身可以通过抑制细胞生长并且分解ECM中的生长因子和蛋白质来抑制愈合。

[0136] 细菌定植.由于创口环境中的更多氧气允许白血球产生ROS来杀死细菌,因此具有不足的组织氧合的患者,例如在手术期间经受体温过低的那些患者,处于更高的感染风险下。宿主对存在细菌的免疫反应延长发炎,延迟愈合,并且破坏组织。感染可以不仅导致慢性创口,而且导致坏疽、受感染的肢体损失以及患者死亡。

[0137] 如同局部缺血一样,细菌定植和感染通过使更多数目的嗜中性粒细胞进入创口部位而破坏组织。在具有慢性创口的患者中,对抗生素具抗性的细菌可以具有发育的时间。另外,携带耐药性细菌菌株,诸如耐甲氧西林金黄色葡萄球菌(methicillin-resistant *Staphylococcus aureus*, MRSA)的患者具有更多慢性创口。

[0138] 生长因子和蛋白水解酶.慢性创口在组成上也不同于急性创口,这在于其蛋白水解酶,诸如弹性蛋白酶和基质金属蛋白酶(MMP)的水平更高,而其生长因子,诸如血小板源性生长因子和角质细胞生长因子的浓度更低。

[0139] 由于生长因子(GF)在及时的创口愈合中是必要的,因此不足的GF水平可能是慢性创口形成中的重要因素。在慢性创口中,生长因子的形成和释放可能受阻,这些因子可能被隔绝并且不能执行其代谢作用,或由细胞或细菌蛋白酶过度降解。

[0140] 诸如糖尿病性和静脉性溃疡的慢性创口也由成纤维细胞无法产生足够的ECM蛋白质并且由角质细胞使创口上皮化而引起。慢性创口中的成纤维细胞基因表达与急性创口中不同。

[0141] 尽管所有创口需要一定水平的弹性蛋白酶和蛋白酶以供适当愈合,但过高的浓度有破坏性。创口区域中的白细胞释放弹性蛋白酶,其增加发炎,毁坏组织、蛋白多糖以及胶原蛋白,并且破坏生长因子、纤连蛋白以及抑制蛋白酶的因子。弹性蛋白酶的活性由人血清白蛋白增加,人血清白蛋白是慢性创口中所发现的最丰富的蛋白质。然而,具有不足白蛋白的慢性创口尤其不可能愈合,因此调控创口的那种蛋白质水平可能在将来证实有助于愈合慢性创口。

[0142] 由白细胞释放的过量基质金属蛋白酶也可能使创口变成慢性的。MMP分解ECM分子、生长因子以及蛋白酶抑制剂,并且因此增加降解同时减少构建,从而使产生与降解之间的精微折衷脱离平衡。

[0143] 口腔溃疡.口腔溃疡(也称为口腔溃疡或粘膜溃疡)是在口腔的粘膜上出现的溃疡。更简单地说,口腔溃疡是口中的疮或开放病灶。口腔溃疡是非常常见的,与许多疾病相关联并且通过许多不同机制而发生,但通常不存在严重的深层原因。口腔溃疡的两个最常见原因是局部创伤(例如,因饼馅上的尖锐边缘摩擦)和口疮性口炎(“口疮”),一种特征为出于大量未知的原因复发性形成口腔溃疡的病状。一些人考虑将嘴唇上或口腔周围的皮肤上的溃疡包括在一般术语口腔溃疡下(例如,由口唇疱疹,即冷疮引起的水疱破裂所留下的溃疡)。口腔溃疡常常引起疼痛和不适,并且可能在愈合发生的同时改变人们对食物的选择(例如,避免酸或辣的食物和饮料)。口腔溃疡可以单独出现或多个溃疡可以同时出现(一“批”溃疡)。一旦

形成,溃疡即可以通过发炎和/或继发性感染而维持。罕见地,持续多周没有愈合的口溃疡可能是口腔癌的征象。其它原因包括烧伤、化学损伤、或感染。

[0144] 粘膜溃疡是特定地在粘膜上出现的溃疡。溃疡是已经穿透上皮-结缔组织边界的组织缺损,其深层次的基底在粘膜下层中或甚至在肌肉或骨膜内。溃疡是比糜烂或脱皮更深的上皮裂口,并且涉及对上皮与固有层的损害。糜烂是上皮的表层裂口,而对深层的固有层有极少损害。粘膜糜烂是特定地在粘膜上出现的糜烂。只有表皮或粘膜的表层上皮细胞损失,并且病灶可以达到基底膜的深度。糜烂愈合而不形成疤痕。脱皮是有时用于描述比糜烂更深但比溃疡更浅的上皮裂口的术语。这种类型的病灶与钉突相切并且显示出由暴露的毛细血管回路引起的点状(小针头点)出血。

[0145] 2. 急性创口/创伤

[0146] 物理创伤是严重的并且改变身体的物理损伤,诸如肢体去除。钝力创伤是由来自或使用钝物体施加的冲击或其它力所致的物理创伤类型,而穿透性创伤是皮肤或组织被物体刺穿的物理创伤类型。创伤还可以被描述为无计划的,诸如意外,或在手术的情况下是计划好的。两者都可以表征为轻度至重度组织损伤、失血和/或休克,并且两者都可能导致后续感染,包括败血症。本发明技术提供了创伤的治疗,包括预治疗(在医疗程序的情况下)与在创伤性损伤已经发生之后的治疗。

[0147] 如本文别处更详细地描述(和上文简单地提及),已经显示PTD-Smad7增强小鼠皮肤模型和粘膜模型中的创口愈合。PTD-Smad7的施加经由局部途径有效,这是创口治疗所需要的。尽管不打算受理论约束,但据信PTD-Smad7可以用以经由多种途径治疗或改善创口,尤其可以包括以下一者或多者:减轻发炎、增加细胞增殖(例如,角质细胞)、增加细胞迁移(例如,角质细胞)、或减少纤维化(例如,经由胶原蛋白的调节)。如下文简单地描述,减轻的发炎可以显著促成加速的创口愈合,任选地经由减少的血管生成和胶原蛋白产生和/或减少的白细胞浸润,从而使得通常由白细胞释放的细胞因子和趋化因子减少,这些细胞因子和趋化因子是促血管生成和促纤维生成的。使用Smad7的暂时性治疗可以允许创口修复所需的早期血管生成和胶原蛋白产生,同时防止长期的血管生成和胶原蛋白产生。这些变化可以潜在地加速创口基质重塑并且防止因迁延性发炎或胶原蛋白过度产生所致的过多疤痕。对于手术程序(以及日常损伤),特别是在疤痕的潜在性成问题的情况下,使用Smad7的治疗可能是有益的。

[0148] 手术.手术对患者使用操作性的手动和仪器技术来研究和/或治疗诸如疾病或损伤的病理性病状,以帮助改善身体功能或外观,或有时是出于某种其它原因。如下文进一步定义,本发明技术可以处理由手术引起的创伤。

[0149] 一般来说,当程序涉及患者组织的切割或先前持续的创口的闭合时,这个程序被视为手术。不一定属于这一类目的其它程序,诸如血管成形术或内窥镜检查,如果涉及常用手术程序或设置,诸如使用无菌环境、麻醉、抗菌条件、典型手术仪器、以及缝合或缝折,那么这些程序可以被视为手术。所有形式的手术都被视为侵入性程序;所谓的非侵入性手术通常指的是不穿透正在处理的结构切除(例如,角膜的激光消融)或指的是放射性手术程序(例如,肿瘤照射)。手术可以持续数分钟至数小时。

[0150] 手术程序通常由紧急性、程序类型、所涉及的身体系统、侵入程度以及专用仪器来分类。完成选择性手术以矫正非危及生命的病状,并且在患者的要求下进行,受制于外科医

师和手术机构的可用性。紧急手术是必须快速完成以挽救生命、肢体或功能能力的手术。进行探查性手术以辅助或确认诊断。治疗性手术治疗先前诊断的病状。

[0151] 截肢涉及切掉身体部分,通常是肢体或指趾。再植涉及重接切断的身体部分。重建手术涉及身体的损伤、残缺或变形部分的重建。完成整容手术以改善其它方面正常的结构的外观。切除是从患者切去器官、组织或其它身体部分。移植手术是通过将来自不同的人(或动物)的另一个器官或身体部分插入患者中来置换器官或身体部分。从活的人或动物中移出器官或身体部分用于移植也是一种手术类型。

[0152] 当对一个器官系统或结构进行手术时,其可以由所涉及的器官、器官系统或组织来分类。实例包括心脏手术(对心脏进行)、胃肠手术(在消化道和其附属器官内进行)、以及矫形手术(对骨骼和/或肌肉进行)。

[0153] 最低侵入性手术涉及较小的外部切口以将微型化仪器插入体腔或结构内,如同在腹腔镜手术或血管成形术中一样。相比之下,开放手术程序需要大的切口以接近所关注的区域。激光手术涉及使用激光替代解剖刀或类似的手术仪器来切割组织。显微手术涉及使用手术显微镜以使外科医师观察到小的结构。机器人手术利用手术机器人,诸如Da Vinci或Zeus手术系统,以在外科医师的指导下控制仪器。

[0154] 3. 自体免疫/发炎疾病

[0155] 本发明技术涵盖多种自体免疫和/或发炎疾病病况的治疗,诸如脊椎关节病变、强直性脊椎炎、牛皮癣性关节炎、反应性关节炎、肠病性关节炎、溃疡性结肠炎、克罗恩氏病(Crohn's disease)、易激性肠病、发炎性肠病、类风湿性关节炎、青少年类风湿性关节炎、家族性地中海热、肌萎缩性侧索硬化、修格兰氏综合症(Sjogren's syndrome)、早期关节炎、病毒性关节炎、多发性硬化或牛皮癣。这些疾病的诊断和治疗在文献中有充分记载。

[0156] 一般来说,自体免疫疾病与身体对通常存在于身体中的物质和组织的过度活跃的免疫反应相关联,并且通常不是免疫反应的焦点。存在超过80种类型的自体免疫疾病,其中一些具有类似的症状,并且其可以由类似的深层原因引起。自体免疫疾病的典型征象是发炎,如本文所公开,其可经受Smad7(任选地,PTD-Smad7)组合物的治疗。

[0157] 4. 化学疗法、放射疗法以及细胞因子疗法的毒性

[0158] 各种形式的癌症疗法,包括化学疗法、辐射以及细胞因子,在癌症患者中与毒性相关联,有时是严重的。本发明技术寻求使用本发明技术的医药组合物降低这种毒性,从而减轻或缓和患者部分的不适,以及允许更高剂量的疗法。

[0159] 如本公开通篇详细地描述,已经发现PTD-Smad7用以愈合以及预防小鼠模型中的口腔粘膜炎。PTD-Smad7显示在直接比较中比帕利夫明更有效,帕利夫明是经批准用于预防口腔粘膜炎的现有药物。

[0160] 口腔癌,世界上第6位最常见癌症,是头颈癌的一种亚型,并且包括位于口腔中的任何癌组织生长。其可以作为起源于口腔组织中的任一者的原发性病灶,通过从远端起源部位转移或通过从邻近的解剖结构(诸如鼻腔)延伸而产生,或口腔癌可以起源于口中的任何组织,并且可以具有多变的组织学类型:畸胎瘤、来源于大唾液腺或小唾液腺的腺癌、来自扁桃体或其它淋巴组织的淋巴瘤、或来自口腔粘膜的产色素细胞的黑素瘤。存在几种类型的口腔癌,但约90%是起源于作为口和嘴唇内里的组织的鳞状细胞癌。口腔癌或口癌最常见地涉及舌头。其还可能在口底、脸颊内里、牙龈(齿龈)、嘴唇或上腭(口顶)上出现。大多

数口腔癌在显微镜下看起来非常相似并且被称为鳞状细胞癌。这些是恶性的并且倾向于快速扩散。

[0161] 超过80%的口腔癌患者用辐射疗法治疗并且至少75%的这些个体将发展出口腔粘膜炎症。口腔粘膜炎症是慢性口腔溃疡。这种疾病在所有癌症类型的辐射治疗患者中频繁出现,包括(但不限于)对器官移植患者进行辐射治疗(以消除移植物的排斥反应)的患者,和经历常规化学疗法的患者。重度口腔粘膜炎症是极度疼痛的并且削弱食物/液体摄入,因此常常是癌症疗法的最严重并发症。口腔粘膜炎症是确定对头颈区域的辐射和化学疗法可能的最大剂量的主要因素;其可以显著地使癌症治疗变复杂,延长住院,降低生活质量并且增加成本。

[0162] 当前,没有确立的疗法来有效地治疗重度口腔粘膜炎症。迄今,人角质细胞生长因子(KGF)的重组蛋白质帕利夫明(**Kepivance®**)是经FDA批准用于骨髓移植患者的重度口腔粘膜炎症的静脉内(i.v.)注射的唯一药物,并且其在癌症疗法中的用途仍有待确定。其还用于预防口腔粘膜炎症。因此,这种药物仅可用于4%的处于风险下的群体。其归因于静脉内施用途径还需要医疗服务提供者。其它潜在疗法包括局部冲洗剂,诸如粘性2%利多卡因(lidocaine)冲洗剂,或小苏打和盐水溶液,或鸡尾酒溶液,例如BAX(利多卡因、苯海拉明(diphenhydramine)、山梨糖醇以及**Mylanta®**)。其它研究性或粘膜保护性辅助疗法包括(但不限于) β 胡萝卜素、生育酚、激光照射、用硝酸银预防性冲刷口腔粘膜、米索前列醇(misoprostol)、亚叶酸、全身KGF、己酮可可碱(pentoxifylline)、别嘌呤醇漱口剂、全身硫酸铝、葡萄糖酸氯己定(chlorhexidine gluconate)以及冷冻疗法。

[0163] 化学疗法诱导和辐射诱导的消化道粘膜炎症是作为快速分化的肠上皮细胞的急性死亡的结果而产生的发炎病状。用于单独、以药物的组合形式或与辐射一起治疗实体肿瘤的大多数化学治疗药物将导致大量肠上皮细胞死亡。随之发生的粘膜炎症的临床表现包括消化道症状,诸如恶心和呕吐、严重腹泻、急性体重减轻和消瘦。这很快成为施用化学疗法用于许多癌症患者的限制因素之一。Tat-Smad7保护肠上皮细胞免受化学治疗剂、辐射或那些的组合损害的能力将显著减少癌症疗法的不当副作用,并且能够用现有工具以更侵袭的方式治疗疾病。

[0164] 骨髓衰竭综合症是当造血干细胞区室受损并且无法产生正常细胞类型时发展的一组病状。骨髓衰竭作为遗传性基因异常、暴露于有毒物质(诸如毒素、化学品或病毒)的结果而发生。尽管可以导致获得性骨髓衰竭发展的环境因素的性质和身份仍然并未完全了解,但少数因素已经与军事人员当中的获得性骨髓衰竭的发展联系起来,包括暴露于芥子气、电离辐射,以及传染剂,诸如内脏利什曼病或非洲锥虫病。用于管理骨髓衰竭综合症的最佳方法仍然是造血干细胞(HSC)的移植,除非足够数目的剩余驻留的骨髓HSC可以免受这些压力并且激励再植造血区室。如此处所描述的Smad 7的调节应能够有意保护展现与骨髓衰竭一致的临床征象的患者中的剩余驻留HSC。

[0165] 5. 癌症

[0166] TGF- β 和NF- κ B活化已知促进癌症侵入和转移。当前,TGF- β 抑制剂在临床试验中用于治疗转移癌并且NF- κ B抑制剂用于癌症预防。Smad7对阻断TGF- β 与NF- κ B信号传导所展示的作用呈现了其作为比仅抑制这两种路径之一的其它抑制剂甚至更强的抗癌/抗转移剂的

可能性。已经显示Smad7预防血管生成和纤维生成,并且可能因此特别适用于肿瘤需要发展供血和/或基质的情形。

[0167] 癌症可以选自由以下组成的群组:脑、肺、肝、脾、肾、淋巴结、小肠、胰脏、血细胞、结肠、胃、乳房、子宫内膜、前列腺、睾丸、子宫颈、子宫、卵巢、皮肤、头颈、食道、骨髓以及血液的癌症。癌症可以是转移性或原发性、复发性或多重耐药性的。在一些实施方案中,癌症是实体肿瘤(器官肿瘤)。实体肿瘤指的是在器官系统中生长并且可以在体内任何地方出现的细胞团。两种类型的实体肿瘤包括在器官内部或外部的上皮组织中出现的上皮肿瘤(癌瘤),和在结缔组织中出现的肉瘤(结缔组织肿瘤),这种结缔组织诸如(但不限于)肌肉、肌腱、脂肪、神经以及支持、围绕或连接体内的结构和器官的其它结缔组织。在一些实施方案中,癌症是血液、骨髓或淋巴结的液体肿瘤或癌症。这些肿瘤包括(但不限于)白血病、淋巴瘤以及骨髓瘤。

[0168] 6. 疤痕形成、纤维化以及异常愈合

[0169] 除了加速的再上皮化(例如,经由增加细胞增殖和/或增加细胞迁移)之外,Smad7对创口基质的作用尤其包括减少发炎、血管生成或胶原蛋白产生中的一者或多者。尽管不打算受理论约束,但这些作用可以经由NF- κ B信号传导的减少(由减少的p50证明)和阻断TGF- β 信号传导(由减少的pSmad2证明)来介导。因此,减轻的发炎可以显著促成加速的创口愈合,任选地经由减少的血管生成和胶原蛋白产生和/或减少的白细胞浸润,从而使得通常由白细胞释放的细胞因子和趋化因子减少,这些细胞因子和趋化因子是促血管生成和促纤维生成的。使用Smad7的暂时性治疗可以允许创口修复所需的早期血管生成和胶原蛋白产生,同时防止长期的血管生成和胶原蛋白产生。这些变化可以潜在地加速创口基质重塑并且防止因迁延性发炎或胶原蛋白过度产生所致的过多疤痕。

[0170] 7. 口炎

[0171] 口炎是口中任何结构的粘膜内里的发炎,这些结构可以涉及脸颊、齿龈、舌头、嘴唇、咽喉以及口顶或口底。发炎可以由口中本身的状况,诸如口腔卫生不良、膳食蛋白质缺乏、不良安装的假牙引起,或由热的食物或饮料所致的口烧伤、有毒植物引起,或由影响整个身体的状况,诸如药物、过敏反应、辐射疗法或感染引起。重度缺铁性贫血可以导致口炎。铁是用于细胞复制和修复的转录元件的上调所必需的。铁缺乏可以引起这些元件的基因下调,从而导致上皮细胞的无效修复和再生,尤其在口和嘴唇中。这种病状也在维生素B₂(核黄素)、B₃(烟酸)、B₆(吡哆醇)、B₉(叶酸)或B₁₂(钴胺素)缺乏的人群中流行。当口炎还涉及牙龈(齿龈)的发炎时,其被称为龈口炎。口炎还可以见于核黄素缺乏症(核黄素缺乏)或嗜中性白血球减少症中。

[0172] 在唇角处的刺激和开裂被称为口角炎或唇角炎。在儿童中,口角炎是反复舔嘴唇的常见原因,并且在成人中,其可能是深层缺铁性贫血或维生素B缺乏(例如,B₂-核黄素、B₉-叶酸盐或B₁₂-钴胺素)的征象,继而可能证明不良饮食或营养不良(例如,乳糜泻)。而且,唇角炎可以由患者的颌部在静态时因缺牙或牙磨损所致的“过度闭合”而引起,这使得颌部比存在完全/未受影响的牙列时更紧密地靠在一起。这引起口角周围的皮肤皱襞,其因唾液而保持潮湿,继而倾向于感染;主要受白色念珠菌(*Candida albicans*)或类似物种感染。治疗通常涉及施用局部制霉菌素或类似的抗真菌剂。另一种治疗可以用牙科治疗(例如,假牙或咬合调整)矫正颌位关系。

[0173] 游走性口炎是口腔粘膜中的广泛区域受薄的白边所包围的环形萎缩性红色病灶影响的病状。这是地图样舌病状的相对不常见的形式,与游走性口炎相反,其仅局限于舌粘膜的脊部和侧部方位。

[0174] 8. 直肠炎

[0175] 直肠炎是直肠,即通向肛门的大肠下端的内里的发炎。直肠炎,即直肠内里-称为直肠粘膜的发炎,是不舒服的并且有时有疼痛。这种病状可能导致从直肠出血或粘膜排出,以及其它症状。直肠炎的一些原因包括(但不限于):性传播疾病(STD),诸如可以在肛交期间传播的那些(例如,淋病、衣原体、梅毒以及疱疹);因例如食源性细菌(例如,沙门氏菌(*Salmonella*)和志贺氏菌(*Shigella*))所致的非STD感染;因例如肛交或将物体或物质插入直肠中(例如,来自灌肠剂的化学品)所致的肛门直肠创伤;溃疡性结肠炎和克罗恩氏病或其它发炎性肠病,可能在结肠和直肠的内衬中引起溃疡(例如,疮);辐射疗法,特别是骨盆区域(例如,直肠、卵巢或前列腺癌)的辐射疗法,其可能导致直肠出血;抗生素,其导致共生菌流失,从而允许有害细菌(例如,艰难梭状芽胞杆菌(*Clostridium difficile*))引起疾病。

[0176] 9. 配方和施用途径

[0177] 在涵盖临床应用的情况下,将有必要制备医药组合物-蛋白质、表达载体、病毒储备液、蛋白质以及药物-呈适于预期应用的形式。一般来说,这将要求制备基本上不含热原质以及可能对人类或动物有害的其它杂质的组合物。

[0178] 在用于动物模型中之前广泛地纯化PTD-Smad7(和截短变体)。使用甘油与PBS的混合物制备PTD-Smad7(和截短型式)用于局部和透粘膜施加。

[0179] 我们将一般需要采用适当的盐和缓冲液以促使递送载体稳定并且允许被靶细胞摄入。当将重组细胞引入患者中时,也将采用缓冲液。本发明技术的水性组合物包含有效量的至细胞的载体,其溶解或分散于医药学上可接受的载剂或水性介质中。这类组合物也被称作接种物。短语“医药学上或药理学上可接受”指的是当施用于动物或人类时不会产生不利、过敏或其它不良反应的分子实体和组合物。如本文所用的“医药学上可接受的载剂”包括任何和所有溶剂、分散介质、包衣、抗菌剂和抗真菌剂、等张剂和吸收延迟剂,等等。这类介质和试剂用于医药活性物质的用途在本领域中是众所周知的。除非任何常规介质或试剂与本发明技术的载体或细胞不相容,否则涵盖其在治疗性组合物中的用途。还可以将补充的活性成分并入组合物中。

[0180] 本发明技术的活性组合物可以包括典型医药制剂。根据本发明技术的这些组合物的施用将经由任何常用途径,只要靶组织经由那个途径可用即可。这类施用途径可以包括经口、肠道外(包括静脉内、肌肉内、皮下、皮内、关节内、滑膜内、鞘内、动脉内、心内、皮下、眶内、囊内、脊柱内、胸骨内以及透皮)、经鼻、经颊、经尿道、经直肠、经阴道、经粘膜、经皮、或局部(包括经皮、经颊以及舌下)。或者,施用可以通过原位、皮内、皮下、肌肉内、腹膜内或静脉内注射。这类组合物将通常以上文所描述的医药学上可接受的组合物施用。特别关注的是直接瘤内施用、肿瘤灌注,或对肿瘤的局部或区域性施用,例如,在局部或区域性血管结构或淋巴系统中,或在切除的肿瘤床中。施用还可以经由鼻用喷雾、手术植入物、内用手术涂剂(internal surgical paint)、输注泵,或经由导管、支架、气囊或其它递送装置。最适用和/或有益的施用模式可以有变化,这尤其取决于接受者的状况和所治疗的病症。

[0181] 呈游离碱或药理学上可接受的盐形式的活性化合物的溶液可以在与表面活性剂，诸如羟丙基纤维素适当混合的水中制备。分散体也可以在甘油、液体聚乙二醇以及其混合物中和油中制备。在储存和使用的一般条件下，这些制剂含有防腐剂以防止微生物生长。

[0182] 适合于可注射使用的医药形式包括无菌水性溶液或分散体以及用于临时制备无菌可注射溶液或分散体的无菌粉末。在所有情况下，这种形式必须是无菌的并且必须是达到易于注射的程度的流体。其必须在制造和储存条件下稳定并且必须针对诸如细菌和真菌的微生物的污染作用进行防腐。载剂可以是溶剂或分散介质，其含有例如水、乙醇、多元醇（例如，甘油、丙二醇以及液体聚乙二醇，等等）、其适合的混合物、以及植物油。适当流动性可以例如通过使用诸如卵磷脂的包衣，在分散体的情况下通过维持所需的粒度，以及通过使用表面活性剂来维持。防止微生物的作用可以由各种抗细菌、抗真菌剂来达成，例如，对羟基苯甲酸酯、氯丁醇、苯酚、山梨酸、硫柳汞（thimerosal），等等。在许多情况下，将优选地包括等张剂，例如，糖或氯化钠。可注射组合物的延长吸收可以通过在组合物中使用延迟吸收剂来达成，例如，单硬脂酸铝和明胶。

[0183] 无菌可注射溶液是通过将活性化合物以所需量与上文所列举的各种其它成分一起并入适当溶剂中来制备，需要时继之以过滤灭菌。一般来说，分散体是通过将各种已灭菌的活性成分并入含有基础分散介质和来自上文所列举的那些的所需其它成分的无菌媒剂中来制备。在用于制备无菌可注射溶液的无菌粉末的情况下，优选的制备方法是真空干燥和冷冻干燥技术，这些技术得到来自其先前无菌过滤溶液的活性成分加上任何额外的所需成分的粉末。

[0184] 如本文所用的“医药学上可接受的载剂”包括任何和所有溶剂、分散介质、包衣、抗细菌剂和抗真菌剂、等张剂和吸收延迟剂，等等。这类介质和试剂用于医药活性物质的用途在本领域中是众所周知的。除非任何常规介质或试剂与活性成分不相容，否则涵盖其在治疗性组合物中的用途。还可以将补充的活性成分并入组合物中。

[0185] 本发明技术的组合物可以按中性或盐形式配制。医药学上可接受的盐包括酸加成盐（与蛋白质的游离氨基形成）并且与以下酸形成：无机酸，诸如盐酸或磷酸，或有机酸，诸如乙酸、草酸、酒石酸、扁桃酸，等等。与游离羧基形成的盐也可以来源于无机碱，诸如氢氧化钠、氢氧化钾、氢氧化铵、氢氧化钙或氢氧化铁，以及有机碱，诸如异丙胺、三甲胺、组氨酸、普鲁卡因（procaine），等等。

[0186] 配方易于以多种剂型施用。取决于所治疗的受试者的病状，剂量将必然发生一些变化。负责施用的人员将在任何情况下决定用于个别受试者的适当剂量。此外，对于人类施用，制剂应符合FDA生物制剂标准办公室（FDA Office of Biologics standards）所要求的无菌性、致热原性、一般安全性以及纯度标准。

[0187] 对于经口施用，可以将本发明技术的多肽与赋形剂合并并且以不可摄取的漱口剂和洁齿剂的形式使用。预期本领域的技术人员已知的几乎任何丸剂或胶囊形式，包括例如包覆包衣和延时、缓慢释放等等，可以用于本发明技术。漱口剂可以通过将活性成分以所需量并入适当溶剂，诸如硼酸钠溶液（多贝耳氏溶液（Dobell's Solution））中来制备。或者，可以将活性成分并入含有硼酸钠、甘油以及碳酸氢钾的抗菌洗涤剂中。还可以使活性成分分散于洁齿剂中，包括：凝胶、糊剂、乳膏、粉末以及浆液。可以将活性成分以治疗有效量添加至糊状洁齿剂中，这种洁齿剂可以包括水、粘合剂、研磨剂、调味剂、发泡剂以及保湿剂。

[0188] 适合于口服剂量的医药组合物可以采用各种形式,诸如片剂、胶囊、囊片以及糯米纸(包括速溶或泡腾),各自含有预定量的活性剂。组合物还可以呈粉末或颗粒、于水性或非水性液体中的溶液或悬浮液、以及液体乳液(水包油和油包水)的形式。活性剂还可以作为大丸剂、药糖剂或糊剂递送。一般应了解,上述剂型的制备方法在本领域中一般是已知的,并且任何这类方法将适合于制备用于组合物递送的相应剂型。

[0189] 在一个实施方案中,活性剂化合物可以与医药学上可接受的媒剂,诸如惰性稀释剂或可食用载剂组合经口施用。可以将口服组合物封入硬壳或软壳明胶胶囊中,可以压成片剂,或可以与患者膳食的食物直接合并。组合物和制剂的百分比可以有变化;然而,这类治疗上有用的组合物中物质的量优选地将获得有效剂量水平。

[0190] 含有活性剂化合物的硬胶囊可以使用生理学上可降解的组合物,诸如明胶来制得。这类硬胶囊包含化合物,并且可以进一步包含额外成分,包括例如惰性固体稀释剂,诸如碳酸钙、磷酸钙或高岭土。含有化合物的软明胶胶囊可以使用生理学上可降解的组合物,诸如明胶来制得。这类软胶囊包含化合物,这种化合物可以与水或油介质,诸如花生油、液体石蜡或橄榄油混合。

[0191] 舌下片剂经过设计以极快速地溶解。这类组合物的实例包括酒石酸麦角胺、二硝酸异山梨醇酯以及盐酸异丙肾上腺素。除药物之外,这些片剂的组合物含有各种可溶性赋形剂,诸如乳糖、粉末状蔗糖、葡萄糖以及甘露糖醇。本发明技术的固体剂型可以任选地被包覆包衣,并且适合的包覆材料的实例包括(但不限于)纤维素聚合物(诸如醋酸邻苯二甲酸纤维素、羟丙基纤维素、羟丙基甲基纤维素、邻苯二甲酸羟丙基甲基纤维素以及醋酸琥珀酸羟丙基甲基纤维素)、聚醋酸乙烯邻苯二甲酸酯、丙烯酸聚合物和共聚物、以及甲基丙烯酸树脂(诸如以商品名EUDRAGIT[®]可商购的那些)、玉米蛋白、虫胶以及多糖。

[0192] 医药制剂的粉末状和颗粒状组合物可以使用已知方法制备。这类组合物可以直接施用于患者或在进一步制备剂型中使用,以便形成片剂,填充胶囊,或通过向其中添加水性或油性媒剂来制备水性或油性悬浮液或溶液。这些组合物中的每一者可以进一步包含一种或多种添加剂,诸如分散剂或湿润剂、悬浮剂以及防腐剂。额外的赋形剂(例如,填充剂、甜味剂、调味剂或着色剂)也可以被包括于这些组合物中。

[0193] 适合于经口施用的医药组合物的液体组合物可以按液体形式或预期在使用前用水或另一种适合的媒剂复原的干燥产品的形式制备、包装和出售。

[0194] 含有本文所描述的一种或多种活性剂化合物的片剂可以通过容易为本领域的技术人员所知的任何标准工艺来制造,诸如通过压缩或模制,任选地与一种或多种佐剂或附属成分一起。片剂可以任选地被包覆包衣或刻痕并且可以经过配制以提供活性剂的缓慢释放或控制释放。

[0195] 可以配制固体剂型以提供活性剂的延迟释放,诸如通过施加包衣。延迟释放包衣在本领域中是已知的,并且含有其的剂型可以通过任何已知的适合方法来制备。这类方法一般包括在制备固体剂型(例如,片剂或囊片)之后,施加延迟释放包衣组合物。施加可以通过诸如无气喷雾、流化床包衣、使用包衣锅等方法。用作延迟释放包衣的材料在性质上可以是聚合的,诸如纤维素材料(例如,丁酸邻苯二甲酸纤维素、邻苯二甲酸羟丙基甲基纤维素以及羧甲基乙基纤维素),以及丙烯酸、甲基丙烯酸以及其酯的聚合物和共聚物。

[0196] 根据本发明技术的固体剂型也可以是持续释放(即,经过一段长期的时间释放活

性剂),并且可能也是或可能不是延迟释放。持续释放组合物在本领域中是已知的并且一般通过使药物分散于逐渐可降解或可水解的材料的基质内来制备,这种材料诸如不溶性塑料、亲水性聚合物或脂肪族化合物。或者,固体剂型可以用这样的材料包覆。

[0197] 用于肠道外施用的组合物包括水性和非水性无菌注射溶液,其可以进一步含有额外的试剂,诸如抗氧化剂、缓冲剂、抑菌剂以及溶质,这些试剂促使组合物与预期接受者的血液等张。组合物可以包括水性和非水性无菌悬浮液,其含有悬浮剂和增稠剂。用于肠道外施用的这类组合物可以存在于单位剂量或多剂量容器中,诸如密封安瓿和小瓶,并且可以在冷冻干燥(冻干)条件下储存,这仅需要在即将使用前添加无菌液体载剂,例如水(注射用)。临时注射溶液和悬浮液可以从先前所描述的种类的无菌粉末、颗粒以及片剂制备。

[0198] 用于经直肠递送的组合物包括直肠栓剂、乳膏、软膏以及液体。栓剂可以作为活性剂与本领域中一般已知的载剂,诸如聚乙二醇组合呈现。可以设计这类剂型以快速地或经过一段长期的时间崩解,并且完全崩解的时间可以在诸如约10分钟的短时间至诸如约6小时的一段长期时间的范围内。

[0199] 局部组合物可以呈在本领域中适合并且易知用于递送活性剂至体表的任何形式,包括经皮、经颊以及舌下。局部组合物的典型实例包括软膏、乳膏、凝胶、糊剂以及溶液。用于口中施用的组合物包括糖锭。

[0200] 根据这些实施方案,经口(局部、经粘膜和/或经皮)递送材料还可以包括乳膏、油膏、软膏、贴片、脂质体、纳米粒子、微粒、定时释放配方以及本领域中已知的其它材料以供递送至受试者的口腔、粘膜和/或皮肤用于治疗 and/或预防本文所公开的病状。某些实施方案涉及使用生物可降解的经口(局部、经粘膜和/或经皮)贴片递送系统或胶凝材料。这些组合物可以是液体配方或用这些组合物的配方处理的医药学上可接受的递送系统,并且还可以包括活化剂/诱导剂。

[0201] 用于本发明技术的方法中的组合物还可以透皮施用,其中将活性剂并入层状结构(一般被称作“贴片”)中,这种层状结构适于保持与接受者的表皮紧密接触一段长期的时间。通常,这类贴片可用作单层“含药物胶粘剂”贴片或多层贴片,其中活性剂含于与胶粘剂层隔开的层中。两种类型的贴片一般还含有背层和在附着至接受者的皮肤之前加以去除的衬里。透皮药物递送贴片还可以由在背层下面的储集层组成,这个储集层由半透膜和胶粘剂层与接受者的皮肤隔开。透皮药物递送可以经由被动扩散、电转运或离子导入而发生。

[0202] 在某些实施方案中,本文所涵盖的贴片可以是缓慢溶解或延时释放贴片。根据这些实施方案,缓慢溶解贴片可以是海藻酸盐贴片。在某些实例中,贴片可以含有可检测的指示剂染料或试剂,诸如荧光剂。在其它实施方案中,标签(例如,可检测标签,诸如生物素或加荧光标签的试剂)可以与治疗分子相关联以检测递送至受试者之后的分子。在某些实施方案中,取决于如由保健专家评估的受试者的需求,可以每天三次、每天两次、每天一次、每隔一天、每周等将本文所涵盖的一个或多个经口递送贴片或其它治疗施用于受试者。本文所涵盖的贴片可以是经口生物可降解的贴片或可能或可能不降解的外用贴片。视需要而定,本文所涵盖的贴片的尺寸可以为1mm、2mm、3mm、4mm至5mm或更大。另外,本文涵盖了用于例如罹患牛皮癣的受试者的皮肤贴片。在治疗牛皮癣和慢性创口中,可以使用诸如甘油、羧甲基纤维素的媒剂局部递送Smad7。还可以使用透皮系统(例如,从3M可商购)以供递送。还可以使用皮下注射至病灶中(在生理盐水或PBS中)。

[0203] 在一些实施方案中,组合物可以包括短期、快速起效、快速失效、控制释放、持续释放、延迟释放以及脉冲释放组合物,条件是这些组合物达成如本文所描述的化合物的施用。参看Remington's Pharmaceutical Sciences (第18版;Mack Publishing Company,Eaton,Pennsylvania,1990),以全文引用的方式并入本文中。

[0204] 在某些实施方案中,本文所公开的化合物和组合物可以经由医疗装置递送。这类递送一般可以经由任何可插入或可植入的医疗装置,包括(但不限于)支架、导管、气囊导管、分流器或线圈。在一个实施方案中,本发明技术提供了医疗装置,诸如支架,其表面包覆有如本文所描述的化合物或组合物。这项技术的医疗装置可以用于例如治疗、预防或以其它方式影响疾病或病状(诸如本文所公开的那些)的进程的任何应用中。

[0205] 预期本领域中已知的任何分子生物学、细胞生物学或生物化学技术可以用于产生和/或测试本文所提供的治疗。另外,预期蛋白质化学技术用于评估治疗在本文所公开的模型系统(例如,小鼠模型系统)中的效用。

[0206] 10. 组合疗法

[0207] 在许多医学领域中常见的是用多种治疗模态,通常被称为“组合疗法”治疗疾病。本文所描述的许多疾病(例如,发炎疾病和癌症)无一例外。在一些实施方案中,为了使用本发明技术的方法和组合物治疗发炎病症,我们将使靶细胞、器官或受试者与Smad7蛋白质、表达构建体或活化剂和至少一种其它疗法接触。这些疗法将以有效达成一个或多个疾病参数的减小的组合量提供。这个过程可以涉及使细胞/受试者同时与两种药剂/疗法接触,例如使用包括两种药剂的单一组合物或药理配方,或通过使细胞/受试者同时与两种不同的组合物或配方接触,其中一种组合物包括Smad7药剂并且另一种包括其它药剂。

[0208] 或者,Smad7药剂可以在其它治疗之前或之后,间隔时间在数分钟至数周的范围内。我们一般将确保在每次递送的时间之间不会中断一段显著的时间,以使得疗法将仍然能够对细胞/受试者发挥有利组合的作用。在这类情况下,预期我们将使细胞与两种模态在相隔约12-24小时内、相隔约6-12小时内或以仅约12小时的延迟时间接触。在一些情况下,可能需要显著延长治疗的时段;然而,在相应的施用之间推移几天(2、3、4、5、6或7)至几周(1、2、3、4、5、6、7或8)。

[0209] 还设想到将需要Smad7药剂或其它疗法的多于一次施用。可以采用各种组合,其中Smad7药剂为“A”,并且其它疗法为“B”,如以下所例示:

A/B/A B/A/B B/B/A A/A/B B/A/A A/B/B B/B/B/A B/B/A/B

[0210] A/A/B/B A/B/A/B A/B/B/A B/B/A/A B/A/B/A B/A/A/B B/B/B/A

A/A/A/B B/A/A/A A/B/A/A A/A/B/A A/B/B/B B/A/B/B B/B/A/B

[0211] 提供了其它组合。适合用于针对发炎病症的组合疗法中的其它药剂包括类固醇、糖皮质激素、非类固醇消炎药(NSAID;包括COX-1和COX-2抑制剂)、阿斯匹林(aspirin)、布洛芬(ibuprofen)以及萘普生(naproxen)。镇痛剂通常与消炎药相关,但其没有消炎作用。一个实例是扑热息痛(paracetamol),在美国被称为醋氨酚并且以商标名Tylenol出售。与通过抑制COX酶而减轻疼痛和发炎的NSAID相反,最近已经显示扑热息痛阻断内源性大麻素(endocannabinoid)的再吸收,这仅减轻疼痛,从而可能解释其对发炎具有最小作用的原因。用于组合使用的特定药剂是抗TGF- β 抗体。

[0212] 技术人员受Remington's Pharmaceutical Sciences,第15版,第33章,特别是第624-652页,1990的指导。取决于所治疗的受试者的病状,剂量将必然发生一些变化。负责施用的人员将在任何情况下决定用于个别受试者的适当剂量。此外,对于人类施用,制剂应符合FDA生物制剂标准办公室所要求的无菌性、致热原性、一般安全性以及纯度标准。

[0213] 还应指出,前述疗法中的任一者可以证实本身适用于治疗发炎。

[0214] 如上文所论述,本发明技术特别关于治疗由某些抗癌疗法所致的DNA损伤和/或发炎,以及用于治疗癌症。因此,特别地,本发明技术可以作为与癌症疗法的组合来应用。这个过程可以涉及使细胞、器官或患者同时与药剂/疗法接触,包括通过使细胞、器官或患者与包括两种药剂的单一组合物或药理配方接触,或同时与两种不同的组合物或配方接触,其中一种组合物包括Smad7药剂并且另一种包括其它药剂。或者,类似于上文所陈述的计划,组合物可以在不同时间递送,包括一种或两种药剂的重复剂量。

[0215] 适合用于组合疗法中的药剂或因子包括当施加至细胞时诱导DNA损伤的任何化合物或治疗方法。这类药剂和因子包括诱导DNA损伤的辐射和波,诸如照射、微波、电子发射,等等。多种化合物,也描述为“化学治疗剂”或“遗传毒性剂”,意图用于本文所公开的组合治疗方法中。在根据本发明技术治疗癌症中,我们将使肿瘤细胞与除表达构建体之外的药剂接触。这可以通过照射局部化肿瘤部位来达成;或者,可以通过向受试者施用治疗有效量的医药组合物而使肿瘤细胞与药剂接触。

[0216] 提供各种种类的化学治疗剂用于与本发明技术的肽组合使用。举例来说,选择性雌激素受体拮抗剂(“SERM”),诸如它莫西芬、4-羟基它莫西芬(Afimoxifene)、氟维司群(Falsodex)、雷洛昔芬(Raloxifene)、巴多昔芬(Bazedoxifene)、氯米芬(Clomifene)、芙婷宝(Femarelle)、拉索昔芬(Lasofoxifene)、奥美昔芬(Ormeloxifene)以及托瑞米芬(Toremifene)。

[0217] 预期有用的化学治疗剂包括例如喜树碱、放线菌素D、丝裂霉素C。本发明技术还包涵了使用一种或多种DNA损伤剂(无论是基于辐射的还是实际化合物)的组合,诸如使用X射线与顺铂(cisplatin)或使用顺铂与依托泊苷(etoposide)。如上文所描述,可以制备药剂并且通过将其与MUC1肽组合而以组合的治疗性组合物或试剂盒使用。

[0218] 热休克蛋白90是在许多真核细胞中发现的调控蛋白质。已经显示HSP90抑制剂适用于治疗癌症。这类抑制剂包括格尔德霉素、17-(烯丙基氨基)-17-去甲氧基格尔德霉素、PU-H71以及利福布丁(Rifabutin)。

[0219] 还设想直接交联DNA或形成加合物的药剂。可以使用诸如顺铂的药剂,和其它DNA烷化剂。顺铂已经广泛用于治疗癌症,其在临床应用中所用的有效剂量为20mg/m²持续5天,每三周一剂,共计三个疗程。顺铂经口不吸收并且因此必须经由静脉内、皮下、瘤内或腹膜内注射来递送。

[0220] 损害DNA的药剂还包括干扰DNA复制、有丝分裂以及染色体分离的化合物。这类化学治疗性化合物包括阿霉素(也称为多柔比星(doxorubicin))、依托泊苷、维拉帕米(verapamil)、鬼臼毒素(podophyllotoxin),等等。在用于治疗赘生物的临床环境中广泛使用的这些化合物经由静脉内大丸剂注射,以25-75mg/m²(21天时间间隔,对于多柔比星)至35-50mg/m²(对于依托泊苷)的范围内的剂量静脉内施用,或经口加倍静脉内剂量。还涵盖微管抑制剂,诸如紫杉烷(taxane)。这些分子是由紫杉属植物产生的双萜,并且包括太平洋

紫杉醇 (paclitaxel) 和多烯紫杉醇 (docetaxel)。

[0221] 表皮生长因子受体抑制剂, 诸如Iressa, 哺乳动物雷帕霉素靶蛋白mTOR, 也称为FK506-结合蛋白12-雷帕霉素相关蛋白1 (FRAP1), 是调控细胞生长、细胞增殖、细胞运动、细胞存活、蛋白质合成、以及转录的丝氨酸/苏氨酸蛋白质激酶。因此提供雷帕霉素和其类似物(“类雷帕霉素(rapalog)”)与根据本发明技术的癌症疗法组合使用。

[0222] 另一种可能的与本文所要求的肽的组合疗法是TNF- α (肿瘤坏死因子- α), 其为全身发炎中所涉及的细胞因子和刺激急性期反应的一组细胞因子的成员。TNF的主要作用在于调控免疫细胞。TNF还能够诱导凋亡性细胞死亡, 诱导发炎, 以及抑制肿瘤发生和病毒复制。

[0223] 破坏核酸前体和亚单元的合成和保真度的药剂也导致DNA损伤。这样的话, 已经开发了许多核酸前体。特别适用的是已经经历广泛测试并且容易可用的药剂。这样的话, 诸如5-氟尿嘧啶(5-FU)的药剂优先被增生组织使用, 使得这种药剂特别适用于靶向增生细胞。尽管毒性相当大, 但5-FU适用于大范围的载剂中, 包括局部, 然而, 通常使用剂量在3至15毫克/公斤/天的范围内的静脉内施用。

[0224] 引起DNA损伤并且已经广泛使用的其它因子包括通常被称为 γ 射线、x射线的那些和/或放射性同位素至肿瘤细胞的定向递送。还涵盖其它形式的DNA损伤因子, 诸如微波和UV照射。最有可能的是所有这些因子对DNA、DNA的前体、DNA的复制和修复、以及染色体的组装和维持实现广泛范围的损伤。x射线的剂量范围在对于长期时间段(3至4周)的50至200伦琴的日剂量至2000至6000伦琴的单剂量的范围内。放射性同位素的剂量范围广泛变化, 并且取决于同位素的半衰期、所发射的辐射的强度和类型、以及由增生细胞的摄入。

[0225] 技术人员受Remington's Pharmaceutical Sciences, 第15版, 第33章, 特别是第624-652页的指导。取决于所治疗的受试者的病状, 剂量将必然发生一些变化。负责施用的人员将在任何情况下决定用于个别受试者的适当剂量。此外, 对于人类施用, 制剂应符合FDA生物制剂标准办公室所要求的无菌性、致热原性、一般安全性以及纯度标准。

[0226] 除了将Smad7疗法与化学疗法和放射疗法组合之外, 还预期与免疫疗法、激素疗法、毒素疗法以及手术组合。特别地, 我们可以采用靶向疗法, 诸如Avastin®、Erbix®、Gleevec®、Herceptin®以及Rituxan®。

[0227] 在其它实施方案中, 为了评估在口腔粘膜炎症的情形中Smad7的作用和机制, 开发了“基因开关”转基因小鼠模型以允许控制特定地在口腔上皮中的Smad7转基因表达的水平 and 持续时间。根据这些实施方案, 这些模型可以用于测试其它基因或下游分子对口腔上皮和口腔粘膜的影响。因此, 这些模型可以用于(但不限于)口腔创口愈合生物学的进一步分析以及测试对口腔创口愈合的治疗方法。在这些研究中鉴别的分子Smad7靶标可以向罹患口腔粘膜炎症的受试者提供额外的治疗靶标。本文所开发的模型和资源可以提供用于分析研究的独特工具以鉴别与Smad7过度表达和控制相关的生物标志物和治疗靶标, 例如, 由Smad7开启或结合的下游分子可以被鉴别为例如额外的治疗靶标, 以治疗口腔粘膜炎症、牛皮癣以及因TGF- β 活性和NF- κ B活性加重的其它病状。

[0228] D. 试剂盒

[0229] 在某些实施方案中, 本文所提供的试剂盒可以包括上文所论述的用于治疗患有本文所提供的病状的受试者的组合物, 这种病状诸如(但不限于)口腔粘膜炎症、牛皮癣或创口

愈合。试剂盒可以包括一个或多个含有本发明技术的治疗性Smad7组合物的容器。这些试剂盒中的任一者一般将包括至少一个小瓶、试管、烧瓶、罐、注射器或其它容器,组合物可以优选地和/或适当地等分至其中。本文中的试剂盒还可以包括用于评估促成本文所提供的病状的生物标靶的试剂盒。

[0230] E. 预测或评价反应的方法

[0231] 还提供了通过评估一种或多种与暴露于Smad7相关的标志物的表达水平来预测和/或评价对使用Smad7治疗的反应的方法。这类标志物可以包括(但不限于)对于细胞迁移的Rac1、对于发炎的NF- κ B,以及对于生长停滞和发炎的TGF- β 。如实施例中所论述,在本领域中提供和/或已知一种或多种与Smad7活性相关的标志物的水平的检测方法和/或变化。在一些实施方案中,可以评估受试者中Smad7标志物中的一者或多者的表达水平,并且基于所检测的水平,可以决定用Smad7治疗(或决定继续或中断治疗),或采用替代性治疗。

[0232] 如本文所用的术语“检测”指的是以一定可重复和可控的水平测量标志物的存在或不存在的能力。通常,在背景值上进行检测,这些背景值可以包括测试系统所固有的噪声(或检测极限)。这样的话,通常存在与分析法相关的检测“下限”,并且为了检测,变化可能需要高于例如某个截止水平。这类极限的确定在本领域中是众所周知的。

[0233] 在一些实施方案中,与对照相比较进行检测,其可以包括(但不限于)与来自缺乏受试者(或所测试的受试者的特定组织)中所存在的疾病或病症的正常受试者和/或可比的正常组织(在相同或不同受试者中)的数据相比较。在一些实施方案中,可以在患者中以多种时间间隔(和/或位置处)检测的水平之间进行比较。在一些实施方案中,检测需要与背景或对照水平相比是统计上显著的;评估显著性的能力在本领域中是众所周知的,并且在实施例中例示。

[0234] 如本文所用的术语“水平的变化”指的是从对照或背景水平和/或先前所检测的水平的可检测变化。在一些实施方案中,变化是与另一种水平相比的增加,并且在一些实施方案中,变化是与另一种水平相比的减小。在一些实施方案中,可检测的变化(增加或减小)是统计上显著的。在一些实施方案中,这类变化可以定量地评估为至少约5%、10%、25%、50%、100%、200%、500%或更大变化,和/或约5-10%、10-25%、10-50%、25-50%、50-75%、50-100%、100-150%、100-200%、200-300%、300-500%或500-1000%的变化。

[0235] F. 筛选额外的生物活性片段的方法

[0236] 在另一个方面,涵盖了筛选Smad7的额外的生物活性片段(包括(但不限于)截短)的方法。在一些实施方案中,生物活性可以使用本文所描述的方法之一来评估,包括下文实施例5和8中所描述的那些。可以评估的一些生物活性包括(但不限于)增加细胞增殖、减少或抑制细胞死亡、减轻过度发炎、防止DNA损伤和/或增加细胞迁移,以及治疗或预防如本文进一步论述的这类治疗将有帮助的一种或多种疾病或病症的动物模型。这类活性可以使用一种或多种分析法来评估,包括(但不限于)阻断Smad2的磷酸化和/或NF- κ B p50亚单元的核转位,增加细胞增殖,减少凋亡和/或辐射诱导的DNA损伤,减少发炎和/或血管生成,促进口腔粘膜炎、手术创口、糖尿病创口和/或与小鼠和其它实验室模型的慢性发炎相关的创口的愈合的能力。一些特定实例包括(但不限于)免疫荧光(IF)、免疫组织化学(IHC)以及用于凋亡的TUNEL分析法。

[0237] 在一些实施方案中,生物活性片段是经过选择以包括一种或多种或所有本文所描

述的活性的那些。在一些实施方案中,生物活性片段经过选择以包括仅仅或主要1种、仅仅或主要2种、仅仅或主要3种、仅仅或主要4种、或仅仅或主要5种本文所描述的活性。在一些实施方案中,生物活性片段经过选择以排除仅仅或主要1种、仅仅或主要2种、仅仅或主要3种、仅仅或主要4种、或仅仅或主要5种本文所描述的活性。在一些实施方案中,生物活性片段经过选择以包括或排除本文所描述的活性的特定子组。举例来说,增加的增殖和迁移可能足以治疗糖尿病性创口,而在慢性发炎创口中需要消炎。减少的凋亡和DNA损伤活性对于治疗口腔粘膜炎是需要的,但对于治疗手术创口不是需要的。

[0238] 如本文所用的术语“主要包括”指的是如下片段:尽管可以保留一定水平的其它生物活性,但那种活性与全长片段相比降低,而被视为“主要”的活性保持在如全长原生蛋白质中所观测到的大约相同或增加的水平下。类似地,如本文所用的术语“主要排除”指的是如下片段:尽管可以保留一定水平的特定生物活性,但那种特定活性的水平与全长片段相比降低(任选地,显著地和/或统计上显著地降低),而一种或多种其它生物活性保持在如全长原生蛋白质中所观测到的大约相同或增加的水平下。

[0239] 在涉及生物活性片段的选择的一些实施方案中,方法包括评估一种或多种生物活性的表达水平的变化,包括在所选片段中一种或多种活性的增加和减小,如关于全长蛋白质中所观测到的活性的变化来评估。在一些实施方案中,选择一种或多种生物活性以保持与全长片段中所观测到的相同,而其它活性可以增加或减小或甚至消除(例如,这类片段将缺乏所论述的一种或多种活性)。在一些实施方案中,变化是与另一种水平相比的增加,并且在一些实施方案中,变化是与另一种水平相比的减小。在一些实施方案中,可检测的变化(增加或减小)是统计上显著的。在一些实施方案中,这类变化可以定量地评估为至少约5%、10%、25%、50%、100%、200%、500%或更大变化,和/或约5-10%、10-25%、10-50%、25-50%、50-75%、50-100%、100-150%、100-200%、200-300%、300-500%或500-1000%的变化。在一些实施方案中,仍然可以观测到“保持相同”的活性相比全长蛋白质的活性有一些变化,但这类变化可能限于例如约1%、2%、5%、10%或20%的变化或更小。

[0240] 在一个非限制性实例中,所关注的片段可以包括主要介导Smad7的消炎作用的那些。具有这种消炎功能的Smad7肽可能是足够的并且任选地是一种改善,用于治疗慢性发炎相关的病状,尤其诸如(但不限于)口腔粘膜炎、口炎以及牛皮癣。在另一个非限制性实例中,所关注的片段可以包括主要介导细胞迁移和/或阻断TGF- β 诱导的生长停滞和/或纤维化反应的那些。具有这种细胞迁移和增殖功能的Smad7肽可能是足够的并且任选地是一种改善,用于增强与过度发炎不相关的愈合。可能得益于这种形式的治疗的创口类型尤其包括(但不限于)手术创口、纤维化疤痕以及糖尿病创口、缺陷性愈合和/或疤痕。

[0241] G. 产生Smad7蛋白质的方法

[0242] 在另一个方面,涵盖了产生Smad7蛋白质,包括本文所描述的Smad7变体、片段、截短、融合蛋白质(例如,PTD-Smad7)中的任一者的方法。本发明者已经发现了在足以用于研究、开发或商业化的水平和纯度下产生Smad7蛋白质的方法,其包括核酸密码子优化。因此,明确涵盖了产生Smad7的方法,其包括使用本文所描述的密码子优化的Smad7核酸分子中的一者或多者(例如,在实施例)。

[0243] 实施例

[0244] 包括以下实施例以说明各种实施方案。本领域的技术人员应了解,以下实施例中

所公开的技术代表了发现在实施所要求的方法、组合物以及设备中发挥良好功能的技术。然而,本领域的技术人员鉴于本公开应了解,在不脱离本发明技术的精神和范围的情况下可以在所公开的特定实施方案中作出许多改变并且仍然获得相同或相似的结果。

[0245] 实施例1:K5.Smad7小鼠对口腔粘膜炎症具抗性

[0246] 如先前所描述(Han等,Dev.Cell,11:301-312,2006)产生在角质细胞中表达人类Smad7蛋白质的转基因小鼠模型(K5.Smad7)。确认口腔上皮中的转基因表达(图7A-B)。将小鼠育种至C57BL/6背景中,并且8-10周大的雄性和雌性转基因小鼠以及野生型同窝仔畜用于研究中。这些小鼠显示出切除的皮肤创口(Han等,Am.J.Pathol.,179:1768-1779,2011)和辐射诱导的口腔粘膜炎症的愈合改善。

[0247] 使K5.Smad7小鼠和野生型同窝仔畜暴露于头颅辐射以确定诱导小鼠的口腔粘膜炎症所需的生物等效剂量(BED)。确定 $8\text{Gy} \times 3$ ($\text{BED}=43.2$),一种在临床上与低分次放射疗法有关的方案,是诱导口腔粘膜炎症所需的最小剂量(图1A-B)。为了评价Smad7作用的效力,还测试了头颅辐射的单剂量并且发现口腔粘膜炎症的严重程度与介于 18Gy ($\text{BED}=50.4$)与 22Gy ($\text{BED}=70.4$)之间的BED值相关联(图1A-B、图7C)。截止起始辐射之后第9天,野生型小鼠发展出口腔溃疡(图1A-B)。

[0248] 在照射之前的K5.Smad7口腔粘膜具有类似于野生型小鼠的形态,但展现出对辐射诱导的口腔粘膜炎症的抗性(图1A-B)。组织学分析揭示,野生型小鼠发展出口腔粘膜炎症(图1A),其类似于人类的口腔粘膜炎症(图1C)。中国昆明医科大学第一附属医院提供了去识别的存档的人类组织石蜡切片并且批准人类受试者免除这项研究。口腔粘膜炎症病灶是来自邻近于已经经历放射疗法的复发性口腔癌的舌头、颊或口咽粘膜。未照射的口腔粘膜切片是来自手术移出的睡眠呼吸暂停口腔组织和邻近于囊肿(粘液囊肿)的舌头活检体。

[0249] K5.Smad7口腔上皮通常显示出辐射剂量依赖性损伤,即,在 $8\text{Gy} \times 3$ 辐射之后上皮变薄并且舌乳头变平,并且在 18Gy 和 22Gy 辐射之后上皮细胞更加损伤(发育不良或肥大)(图1A)。与人类口腔粘膜炎症病灶中增加的白细胞浸润一致(图1C),野生型小鼠中的病灶具有众多浸润性白细胞(图1D-E),其由嗜中性粒细胞、巨噬细胞以及淋巴细胞(图7D)组成;所有细胞在K5.Smad7口腔粘膜中都实质上减少(图1D-E和7D)。

[0250] 因为难以在急性期捕捉人类口腔粘膜炎症病态,所以利用小鼠模型来评估当正形成溃疡时的增殖和凋亡。类似于先前的报导,增殖细胞在经过照射的野生型口腔上皮中是稀少的,但在经过照射的K5.Smad7口腔上皮中观察到更多(图1D和1F)。相反地,凋亡细胞在经过照射的K5.Smad7口腔粘膜中与野生型小鼠相比显著减少(图1D和1G)。

[0251] 正如预期,具有核NF- κ B p50亚单元的细胞在口腔粘膜炎症中与未照射的野生型口腔粘膜相比显著增加(图2A-B)。令人感兴趣地,在内脏器官中是免疫抑制剂,但在口腔粘膜中促发炎症的TGF- β 1,连同其活化的信号传导介体磷酸化(p) Smad2一起,在口腔粘膜炎症中与野生型小鼠中的未照射的口腔粘膜相比也增加(图2A-B)。在人类口腔粘膜炎症病灶中也检测到类似变化(图2A-B)。

[0252] 经过照射的K5.Smad7口腔上皮显著减少对于核NF- κ B p50和pSmad2呈阳性的细胞,即使其仍然具有丰富的TGF- β 1蛋白质(图2A-B)。经过照射的野生型口腔粘膜中的TGF- β 1 mRNA在第9天和第10天显著增加(图2C)。K5.Smad7粘膜中的TGF- β 1 mRNA水平在较早的时间点类似于野生型粘膜,但截止第10天返回至正常(图2C)。尽管不希望受任何理论约束,但这

些数据表明TGF- β 1转录不受Smad7抑制,而其在K5.Smad7粘膜中的更快速下降可能是加速愈合的结果。

[0253] 活化的BMP信号传导的标志物磷酸化Smad1/5/8在辐射之前和之后不受Smad7影响(图7E)。这个结果与Smad7优先抑制TGF- β 信号传导的能力一致。

[0254] 实施例2: Rac1促成Smad7介导的角质细胞迁移

[0255] 为了判定Smad7是否促成人口腔角质细胞的愈合,敲落自发永生化的人口腔角质细胞(NOK-SI)中的Smad7。Smad7敲落减弱在开创口之后的角质细胞迁移(图2D和图8A)。相反地,敲落TGF- β 1加速角质细胞迁移(图8B-8D),这与不存在TGF- β 1或Smad3的小鼠中所观察到的加速创口愈合一致。

[0256] 为了研究与Smad7介导的角质细胞迁移相关的分子机制,检查了对于口腔创口愈合必不可少的蛋白质Rac1。在Smad7敲落之后Rac1减少(图2E)。预期在口腔粘膜炎症中TGF- β 1过度表达将经由Smad非依赖性机制活化Rac1。然而,尽管在照射之后总Rac1蛋白质增加2倍,但在野生型舌头中活化的Rac1蛋白质并不明显改变(图2F)。

[0257] 在K5.Smad7口腔粘膜中,与野生型口腔粘膜相比总Rac1与活化的Rac1分别显著增加4倍和8倍(图2F)。为了确定Smad7诱导的Rac1活化的功能意义,在分离自野生型和Smad7转基因新生皮肤的初级角质细胞中敲落Rac1,并且进行细胞增殖和迁移的分析法。Rac1敲落显示野生型和Smad7角质细胞的增殖适度减少(图9A-9C),但Smad7诱导的迁移几乎完全消去(图2G和图9D),这表明增加的Rac1促成Smad7介导的细胞迁移。

[0258] 观测到在Smad7转基因角质细胞中增加的Rac1 mRNA水平与总的和活性Rac1蛋白质水平相关联(图3A-B和图10A-B),这表明Smad7角质细胞中增加的Rac1活化至少部分是增加的Rac1转录物的结果。此外,在敲落NOK-SI细胞中的个别Smad(图10C-10E)之后Rac1蛋白质增加约3倍(图3C)。这些数据表明正常Smad信号传导阻遏Rac1转录。

[0259] 在小鼠Rac1启动子中的两个假定Smad结合元件(SBE)(编码序列上游-2.1Kb和-1.5Kb)当中,这两个结合元件在人类Rac1启动子的类似区域中,染色质免疫沉淀(ChIP)鉴别出Smad-2、Smad-3、Smad-4以及Smad-7结合至-1.5Kb位点(图3D),但不结合至野生型角质细胞中的-2.1Kb位点;Smad-2、Smad-3以及Smad-4的结合在Smad7转基因角质细胞中显著降低(图3D)。

[0260] 使用含SBE的Rac1-Luc构建体的荧光素酶报告基因分析法显示,野生型角质细胞中Smad7的敲落显著降低荧光素酶活性(图3E)。相反地,Smad7转基因细胞与野生型细胞相比具有增加的荧光素酶活性,并且使SBE突变削弱了这种增加(图3F)。因此,Smad7结合至SBE似乎有必要驱除信号传导Smad以消去Rac1阻遏。

[0261] 在已知的Smad转录辅阻遏子当中,发现CtBP1结合至野生型角质细胞中的Rac1启动子SBE-1.5Kb位点(图3G),并且Smad7转基因表达显著减少CtBP1结合至SBE(图3G-H)。当在NOK-SI细胞中敲落CtBP1时,与用乱序siRNA转染的角质细胞相比Rac1蛋白质和Rac1-Luc活性增加(图4A-B),这表明CtBP1结合至SBE-1.5Kb阻遏Rac1表达。此外,敲落NOK-SI细胞中的CtBP1增加其迁移(图4C和图10F)。

[0262] 在检查辐射诱导的口腔粘膜炎症中的CtBP1蛋白质时,发现CtBP1在未照射的小鼠和人类口腔粘膜中很少检测到(图4D-4F);然而,CtBP1阳性细胞在野生型和K5.Smad7小鼠的经过照射的口腔粘膜中以及人类口腔粘膜炎症中显著增加(图4D-4F)。另外,经过照射的野生

型口腔粘膜中的CtBP1mRNA在第9天和第10天显著增加(图4G)。K5.Smad7粘膜中的CtBP1mRNA水平在较早的时间点类似于野生型粘膜,但截止第10天下降至正常(图4G)。这些结果指示,Smad7并不减少CtBP1mRNA,而是通过将Smad/CtBP1复合体从SBE结合位点排除来抑制CtBP1结合至Rac1启动子;此外,K5.Smad7粘膜中更快速的CtBP1减少充当愈合的标志。

[0263] 实施例3:Tat-Smad7缓和辐射诱导的口腔粘膜炎

[0264] Smad7转基因阻断口腔粘膜炎的多个病理过程的能力促使我们探索局部化Smad7递送是否可以用于预防和治疗口腔粘膜炎。因为Smad7是核蛋白,所以局部Smad7递送需要在唾液洗去蛋白质之前允许Smad7快速进入细胞中。因此,产生了具有N端Tat标签的重组人类Smad7,其允许蛋白质快速渗透细胞膜并且进入核中。将V5表位添加至Tat-Smad7蛋白质的C端末端以追踪Tat-Smad7细胞穿透(图11A-11D)。

[0265] 使用其阻断Smad2磷酸化的能力,测试Tat-Smad7生物活性(图11C)。产生了具有与对照相同的标签的Tat-Cre重组蛋白质(图11E-F),并且将其克隆至含有编码C端6XHis的序列(SEQ ID NO:40)的pET101-Topo蛋白质表达载体(Invitrogen)中。将Tat-Cre转化至BL-21StarTM大肠杆菌(Invitrogen)中以产生Tat-Cre蛋白质并且用Ni-NTA柱纯化。

[0266] 使用SDS-PAGE电泳来验证两种蛋白质的纯度和大小。为了评价体外Tat-Smad7蛋白质的转导和活性,将Tat-Smad7添加至初级小鼠角质细胞中。将载片固定于冷甲醇中5分钟并且针对V5和pSmad2进行染色。通过消化来自7,650bp载体pLL3.7(Addgene)的1,460bp侧接loxP的片段来验证Tat-Cre活性。对于体内处理,使用30 μ L 50%甘油/PBS作为媒介对照和Tat-Cre作为相关的蛋白质对照。将Tat-Smad7或Tat-Cre(在30 μ L 50%甘油/PBS中,剂量和方案在每个图中指定)局部施加至小鼠口腔并且限制小鼠经口进食1小时。

[0267] 对于口腔粘膜炎预防,在辐射之前24小时开始至起始辐射之后第8天,将Tat-Smad7与Tat-Cre(在50%甘油/PBS中)每天局部施加至8-10周大的C3H雌性(Jackson Laboratory)或C57BL/6小鼠的口腔。第9天检查经过处理的组织。收集小鼠舌头,固定于10%福尔马林中,包埋于石蜡中,并且切成5 μ m切片。分析组织学变化并且使用H&E染色的载片测量溃疡。另一组使用临床方案接受帕利夫明处理,即在照射之前每天6.25mg kg⁻¹(腹膜内)持续3天,以及在辐射的最后一次剂量之后24小时每天持续3天。

[0268] Tat-Cre显示与媒介对照相比没有作用(图5A-B)。Tat-Smad7处理显示对溃疡形成的预防作用类似于帕利夫明(图5A)。Tat-Smad7的剂量依赖性作用当用于给予诱导比分次辐射更大的口腔溃疡的20Gy(BED=60)单剂量的辐射的动物时更加明显(图11G)。在显微镜下,帕利夫明与Tat-Smad7处理的口腔粘膜在大多数情况下防止开放型溃疡形成(图5B)。帕利夫明处理的粘膜相比Tat-Smad7处理的粘膜展现出更多的角质细胞生长减退,以及更加受损的角质细胞(凝结的核或木炭状核、膨胀的单核或多核细胞以及在角质层中破碎的核片段)(图5B)。免疫染色揭示,帕利夫明比Tat-Smad7更显著地增加增殖。Tat-Smad7减少凋亡、白细胞浸润、核pSmad2以及NF- κ B p50,但帕利夫明并不如此(图5B-5G)。

[0269] 为了测试Tat-Smad7是否可以用于治疗现有的口腔粘膜炎,使小鼠暴露于分次(8Gy \times 3)头颅辐射,并且在起始辐射(当粘膜损伤明显可见时)之后第6天直至第9天,每天施加Tat-Smad7(局部)或帕利夫明(6.25mg kg⁻¹,腹膜内)。在第10天检查经过处理的组织。尽管在比当前方案更早的时间点开始辐射后施用帕利夫明减轻小鼠的口腔粘膜炎,但使用当前方案施用帕利夫明并不加速溃疡闭合(图6A),与其对整个口腔粘膜的高增殖作用无关

(图6B)。这并不令人惊讶,因为帕利夫明经批准用于预防而非治疗口腔粘膜炎。

[0270] Tat-Smad7处理的口腔粘膜炎减小在分次与单剂量辐射之后的溃疡尺寸和病理改变(图6A-B和图12A-12G)。远离溃疡,Tat-Smad7处理的口腔粘膜相比帕利夫明处理的口腔粘膜展现出更少的增生和更加分化的上皮(图6B)。在比分次辐射引起更缓慢愈合的20Gy单剂量辐射下,Tat-Smad7对创口闭合之后的恢复的作用更加明显。当媒介处理的溃疡刚刚再上皮化时,Tat-Smad7处理的粘膜已经几乎恢复至正常形态(图6C)。

[0271] 与K5.Smad7小鼠中的观测结果一致,Tat-Smad7增加Rac1启动子活性并且减少CtBP1结合至小鼠Rac1启动子的SBE(图12G和12I),并且增加小鼠口腔粘膜炎和人口腔角质细胞中的Rac1蛋白质(图6D-E)。

[0272] 在创口划痕之后Tat-Smad7处理的人口腔角质细胞已经加速创口闭合(图6F和图13A)。此外,经过照射的人口腔角质细胞增加核pSmad2和NF- κ B p50,其通过Tat-Smad7处理而削弱(图13B)。相比之下,尽管Tat-Smad7有效地穿透口腔癌细胞(图13C),但其并不进一步升高在癌细胞中已经很丰富的Rac1蛋白质水平(图13D)。这个结果可以说明癌细胞比正常角质细胞更快迁移(图6F和图13A、13E-13H),以及Tat-Smad7对两种口腔癌细胞系的迁移的作用缺乏:不含TGF- β 信号传导组分的遗传损失的MSK921;和具有突变的Smad4的Cal27(图13E-13H)。

[0273] 集落分析法显示,人口腔角质细胞的存活在有或没有辐射的情况下通过Tat-Smad7处理而略微增加(图6G)。与在照射之后存活减少在癌细胞中比在正常细胞中更加突出的观念一致,SCC细胞显示在辐射之后细胞存活的实质性减少。用Tat-Smad7处理并不影响在有或没有辐射的情况下SCC细胞的存活(图6G)。

[0274] 实施例4:设计细胞穿透Smad7蛋白质

[0275] 假设为了有效作为治疗剂,SMAD7需要能够有效地穿透细胞。为了达成这一目的,修饰Smad7序列以包括蛋白质转导域。

[0276] 选择来自HIV的Tat序列以与作为蛋白质转导域的Smad7一起测试。与Smad7和Smad7片段一起用于融合蛋白质中的Tat的核苷酸和蛋白质序列来源于Cardarelli等, Traffic Apr 9(4):528-39(2008)。Tat核苷酸和氨基酸序列提供如下:

[0277] ggccgtaaaaaacgccgtcaacgccgccgt (SEQ ID NO:1)

[0278] G R K K R R Q R R R (SEQ ID NO:2)

[0279] 制备融合蛋白质,其具有如下所示在Smad7的5'或3'末端直接框内连接至人类Smad7互补DNA(cDNA)的Tat:

[0280] 5'Tat:Ggccgtaaaaaacgccgtcaacgccgccgt (SEQ ID NO:7)-Smad7

[0281] 3'Tat:Smad7-Ggccgtaaaaaacgccgtcaacgccgccgt (SEQ ID NO:8)

[0282] 5'Tat-Smad7构建体包括3'V5标签序列,并且将其克隆至pGEX-6p-1蛋白质表达载体(New England Biolabs)中以制造GST-Tat-Smad7融合蛋白质。将Tat-Smad7基因转化至BL-21Star大肠杆菌(Invitrogen)中以产生Tat-Smad7蛋白质。通过谷胱甘肽柱纯化和洗脱,使用由谷胱甘肽S转移酶(GST)融合体(Precision酶,GE Life Sciences)引起的酶促裂解来纯化蛋白质。

[0283] 在创建PTD-Smad7融合蛋白质的时候,在3'末端包括V5标签以通过使用V5抗体进行免疫染色来监测Tat-Smad7穿透至细胞中。适当时,可以删去这个表位标签以供临床上使

用(例如,通过在不存在V5标签的情况下再克隆序列)。

[0284] 还创建了具有6-组氨酸(6-H)标签(SEQ ID NO:40)用于蛋白质纯化的PTD-Smad7融合蛋白质(Tat-Smad7-V5-6H)(如SEQ ID NO:40公开的“6H”),并且如下所示。Tat-Smad7-V5-6H(如SEQ ID NO:40公开的“6H”)具有以下核苷酸序列:1-53包括pET-TOP0的5'序列;54-1365包括Tat-Smad7;1366-1497包括含有V5表位和6xHis标签(SEQ ID NO:40)的3'pET-TOP0(V5包括1393-1434,His标签包括1444-1461,并且终止包括1462-1464)。

[0285] 克隆至pET101/D-Topo载体的针对蛋白质产生进行密码子优化的Tat-人类Smad7如下所示:

ttccctctagaaataattttgtttaactttaagaaggaattcaggagcccttcacatg

M

cgtaaaaaacgccgtcaacgccgccgtggtttccgtacgaaacgctcggccctggtccgt

R K K R R Q R R R G F R T K R S A L V R

cgcctgtggcgctcccgtgctccgggtggtgaagatgaagaagaaggtgctggcgggcggt

[0286]

R L W R S R A P G G E D E E E G A G G G

ggcggtggcggtgaactgcgtggcgagggtgcaaccgatagtcgtgcacacggtgcaggc

G G G G E L R G E G A T D S R A H G A G

ggtggcggtccgggtcgtgctggttgctgtctgggtaaagctgtgcgcggcgcgaaaggt

G G G P G R A G C C L G K A V R G A K G

catcaccatccgcacccgccggcagcaggtgcaggtgcagctggcggtgcggaagccgat

H H H P H P P A A G A G A A G G A E A D

[0287]

ctgaaagccctgaccatagtgctcctgaaaaactgaaagaacgtcagctggagctgctg
L K A L T H S V L K K L K E R Q L E L L
ctgcaagcagtagaatcccgtggcggtacccgtacggcttgctgctgctgccgggtcgt
L Q A V E S R G G T R T A C L L L P G R
ctggattgccgtctgggtccgggtgcaccggctggtgcgcagccggcacaaccgccgagc
L D C R L G P G A P A G A Q P A Q P P S
tcttacagcctgccgctgctgctgtgtaaagtgttcggtggccggacctgcgccacagt
S Y S L P L L L C K V F R W P D L R H S
tccgaagttaaaccgctgtgctgttgcgagagctatggcaaaattaaccgggaactggtt
S E V K R L C C C E S Y G K I N P E L V
tgttgcaatccgcaccatctgtctcgtctgtgtgaactggagagcccgccgccgctat
C C N P H H L S R L C E L E S P P P P Y
tctcgttaccgatggatttcctgaaaccgactgcagattgcccggacgcagtcctcgtca
S R Y P M D F L K P T A D C P D A V P S
tcggctgagaccggcgccaccaactatctggcaccggggcggtctgagtgattcccagctg
S A E T G G T N Y L A P G G L S D S Q L
ctgctggaaccggggcgaccgttcacattggtgtgtggtgcctattgggaagagaaaacg
L L E P G D R S H W C V V A Y W E E K T
cgtgtcggctgcctgtactgcgtacaggaaccgtcgtcggatatctttatgacctgccg
R V G R L Y C V Q E P S L D I F Y D L P
cagggcaatggtttctgtctgggccaactgaactcagataataaatcgcagctggtgcaa
Q G N G F C L G Q L N S D N K S Q L V Q
aaagttcgctcaaaaattggctgcggtatccagctgacctggaagttgacggtgtctgg
K V R S K I G C G I Q L T R E V D G V W
gtatataaccgcagctcttaccgatttttatcaaaagtgccaccctggataatccggac
V Y N R S S Y P I F I K S A T L D N P D
tcccgtacgtcgtcgtccacaaagtatttccgggcttctcaatcaaagcgttcgattac
S R T L L V H K V F P G F S I K A F D Y

gagaaagcctactcgctgcagcgcccgaacgacatgaattcatgcagcaaccgtggacg

E K A Y S L Q R P N D H E F M Q Q P W T

ggttttactgtgcagatctctttcgtaaaggctgggggtcaatgctacacccgtcagttt

G F T V Q I S F V K G W G Q C Y T R Q F

atctcgctcgtgccgtgctggctggaagtgattttcaatagccgcaagggcgagctcaat

[0288]

I S S C P C W L E V I F N S R K G E L N

tcgaagcttgaaggaagcctatccctaaccctctcctcggtctcgattctacgcgtacc

S K L E G K P I P N P L L G L D S T R T

ggatcatcatcaccatcaccattgagtttgatccggctgctaacaaagcccgaaagga (SEQ ID NO: 9)

G H H H H H H - (SEQ ID NO: 10)

[0289] 下面提供了Tat-Smad7-v5与Smad7的蛋白质序列的比较。Tat-Smad7中Smad7的第一氨基酸不是M(不同于Smad7),因为Tat-Smad7经过设计以与Tat和/或GST框内形成GST融合蛋白质。然后,在纯化之后使Tat-Smad7从GST融合蛋白质裂解。大写的核苷酸鉴别V5标签。加下划线的斜体指示来自任选的pET101-Topo骨架载体的氨基酸。

[0290] 以下呈现了Tat-Smad7-v5与Smad7的比较:

Tat-Smad7-V5 1 gsgrrkrrrrrrgfrtkrsalvrrlwrsrapggedeegagggggggelr

人类 Smad7 1 -----mfrtkrsalvrrlwrsrapggedeegagggggggelr

Tat-Smad7-V5 51 gegatdsrahgaggggppragccclgkavrgakghhhphppaagagaagga

人类 Smad7 39 gegatdsrahgaggggppragccclgkavrgakghhhphppaagagaagga

Tat-Smad7-V5 101 eadlkalthsvlkkklkerqllellqavesrggtrtaclllpgrldcrlgp

人类 Smad7 89 eadlkalthsvlkkklkerqllellqavesrggtrtaclllpgrldcrlgp

[0291]

Tat-Smad7-V5 151 gapagaqpaqppssyslplllckvfrwpdlrshssevkrccccesygkinp

人类 Smad7 139 gapagaqpaqppssyslplllckvfrwpdlrshssevkrccccesygkinp

Tat-Smad7-V5 201 elvccnphhlsrlcelesppppysrypmdflkptadcpdavpssaetggt

人类 Smad7 189 elvccnphhlsrlcelesppppysrypmdflkptadcpdavpssaetggt

Tat-Smad7-V5 251 nylapgglsdsqlllepgdrshwcvvayweektrvgrylcvcqepsldify

人类 Smad7 239 nylapgglsdsqlllepgdrshwcvvayweektrvgrylcvcqepsldify

Tat-Smad7-V5	301	dlpqngfclgqlnsdnksqlvqkvrskigcqiqltrevdgvwvynrssi
人类 Smad7	289	dlpqngfclgqlnsdnksqlvqkvrskigcqiqltrevdgvwvynrssi
Tat-Smad7-V5	351	pifiksatldnpdsrtllvhkvfpgfsikafdyekayslqrpndhefmqq
人类 Smad7	339	pifiksatldnpdsrtllvhkvfpgfsikafdyekayslqrpndhefmqq
[0292] Tat-Smad7-V5	401	pwtgftvqisfvkgwgqcytrqfisscpwlevifnsr <u>kgelnskle</u> GKP
人类 Smad7	389	pwtgftvqisfvkgwgqcytrqfisscpwlevifnsr-----
Tat-Smad7-V5	451	IPNPLLGLDST (SEQ ID NO: 11)
人类 Smad7	427	----- (SEQ ID NO: 12)

[0293] 实施例5:用于PTD-Smad7蛋白质活性的额外分析法

[0294] 免疫荧光 (IF)、免疫组织化学 (IHC) 以及用于凋亡的TUNEL分析法.如先前所描述 (Han,G.,Li,F.,Ten Dijke,P.以及Wang,X.J.小鼠表皮中的暂时性smad7转基因诱导加速皮肤创口愈合 (Temporal smad7transgene induction in mouse epidermis accelerates skin wound healing).Am J Pathol 179,1768-1779 (2011)) 进行IF和IHC.所使用的初级抗体是针对K14的天竺鼠抗体 (1:400,Fitzgerald,20R-CP200)、针对CD4的大鼠抗体 (1:20,BD Bioscience,550278)、Ly-6G (1:20,BD Bioscience,550291)、BM8 (针对F4/80的抗体,1:20,Invitrogen,MF48000)、针对BrdU的FITC标记抗体 (BD Bioscience,347583)、用于小鼠样品的针对CD45的大鼠抗体 (1:50,BD Bioscience,550539)、用于人类样品的针对CD45的小鼠抗体 (1:50,Abcam,Ab781)、针对TGF- β 1的鸡抗体 (1:50,R&D,AF-101-NA)、针对CtBP1的兔抗体 (1:100,Millipore,07-306)、针对NF- κ B p50的兔抗体 (1:200,Santa Cruz Biotechnology,SC-7178)、针对PCNA的兔抗体 (1:200,Santa Cruz Biotechnology,SC-7907)、针对pSmad2的兔抗体 (1:100,Cell Signaling Technology,3101),以及针对V5的小鼠抗体 (1:500,Invitrogen,460705)。对于IF,针对不同种类IgG的二级抗体是Alexa Fluor[®] 594 (红色) 或488 (绿色) 接合的 (所有都是1:200,Invitrogen) o对于IHC,使用针对不同种类IgG的二级生物素化抗体 (1:300,Vector Labs) 并且使用Vectastain ABC试剂盒 (Vector Labs) 显色。在福尔马林固定的组织切片上使用末端脱氧核苷酸转移酶尿苷缺口末端标记 (TUNEL,G3250) 试剂盒 (Promega) 以检测凋亡细胞。通过在安乐死之前1小时腹膜内注射0.125mg g⁻¹BrdU在体内进行BrdU标记。将PCNA或BrdU定量为每毫米包括所有上皮细胞的上皮长度的细胞,TUNEL或CD45阳性细胞定量为每毫米包括肌肉层上面的所有上皮层和基质的上皮长度的细胞,核pSmad2或NF- κ B p50阳性细胞定量为阳性细胞/现有的总剩余上皮细胞 (即,排除由照射诱导的脱落的上皮细胞) 的数目。载片的连续视场用于使用MetaMorph软件对BrdU标记的细胞进行计数。

[0295] 细胞培养.如先前所描述 (Han,G.,Li,F.,Ten Dijke,P.以及Wang,X.J.小鼠表皮中的暂时性smad7转基因诱导加速皮肤创口愈合 (Temporal smad7transgene induction in mouse epidermis accelerates skin wound healing).Am J Pathol 179,1768-1779 (2011)) 从新生小鼠皮肤制备Smad7转基因和野生型初级角质细胞,并且在PCT培养基 (CELLnTEC) 中培养。在限定角质细胞培养基中培养和维持来源于健康志愿者的牙龈组织的

自发永生化的正常口腔角质细胞(NOK-SI)(Castilho,R.M.等Rac1是在皮肤和口腔粘膜创口愈合期间的上皮干细胞功能所需的而不是小鼠中的组织内稳态所需的(Rac 1 is required for epithelial stem cell function during dermal and oral mucosal wound healing but not for tissue homeostasis in mice).PloS one 5,e10503 (2010))。在补充有10%胎牛血清(GIBCO®;Invitrogen)的达尔伯克改良伊格尔培养基(Dulbecco Modified Eagle Medium)中培养口腔癌细胞Cal27(ATCC)和MSK921(D.Raben实验室,由科罗拉多大学癌症中心组织培养核心(University of Colorado Cancer Center Tissue Culture Core)采集指纹)。为了评估Tat-Smad7在经过照射的细胞中的作用,在腔室载片(BD Bioscience,354108)中培养上述人类细胞系,用3Gy照射,并且在照射之后立即将Tat-Smad7($1\mu\text{g mL}^{-1}$)添加至培养基中。在Tat-Smad7处理之后4小时将细胞固定于100%冷甲醇中用于pSmad2、NF- κ B p50以及V5的免疫染色。

[0296] 用siRNA转染.当经过培养的角质细胞达到70%汇合时,使用LIPOFECTAMINE® 2000(Invitrogen)转染100nM靶siRNA或乱序siRNA(Dharmacon)。在转染之后48-72小时收集细胞并且经受西方分析以测定敲落效率。对于迁移分析法,当将细胞涂板时转染siRNA。在这项研究中所包括的靶siRNA是:小鼠siRac1-1(Invitrogen,MSS237708)和siRac1-2(IDT,MMC.RNAI.N009007.12.3);人类siSmad2(Dharmacon,L-003561-00-0005)、siSmad3(Invitrogen,HSS106252)以及siSmad4(Invitrogen,HSS118066);人类siCtBP1-1和siCtBP1-2;人类siSmad7-1和siSmad7-2;人类TGF- β 1(Dharmacon,J-012562-08-0005);小鼠siSmad7。

[0297] 体外角质细胞增殖分析法.通过在野生型和Smad7转基因角质细胞中的BrdU并入来测定体外角质细胞增殖。用Rac1siRNA转染处于70%汇合的细胞,并且在24小时后更换成常规培养基。原位细胞增殖试剂盒(Roche Applied Science)用于进行体外BrdU标记和检测,并且MetaMorph软件用于对BrdU标记的细胞进行计数。

[0298] 体外细胞迁移分析法.当细胞达到100%汇合时,用 $10\mu\text{g mL}^{-1}$ 的丝裂霉素C(Sigma)处理细胞2小时以抑制细胞增殖,并且用Fisherbrand吸管尖端引入划痕创口。每天为细胞迁移拍照。当在siRNA转染24至36小时之后细胞达到汇合时进行迁移分析法,并且使用Image-J软件以被迁移细胞占据的创口区域记录细胞迁移。对于Tat-Smad7处理,在创口划痕之后使细胞暴露于培养基中的 $1\mu\text{g mL}^{-1}$ Tat-Smad7蛋白质或媒介对照(PBS),并且每隔一天用新鲜添加的Tat-Smad7更换培养基直至迁移细胞完全覆盖划痕的创口。

[0299] 细胞存活分析法.如先前所描述(Munshi,A.,Hobbs,M.以及Meyn,R.E.克隆源性细胞存活分析法(Clonogenic cell survival assay).Methods in molecular medicine 110,21-28(2005)),在稍作修改下进行细胞存活分析法。简单地说,将细胞以500个细胞/孔(对于未照射的细胞)涂于12孔板中,并且随着增加的辐射剂量而增加至1,500个细胞/孔。在将细胞涂板之后照射细胞24小时。将 $1\mu\text{g mL}^{-1}$ 的Tat-Smad7或相同体积的用于溶解Tat-Smad7的PBS(对照)添加至经过照射和未照射的细胞的培养基中。每隔一天用新鲜添加的Tat-Smad7或PBS更换培养基,持续10至14天。将集落固定于甲醇中,在0.5%结晶紫溶液(含有25%甲醇)中染色,进行计数,并且从每个实验中的4个孔计算平均值。对于每个细胞系进行两个至三个独立实验。如先前所描述计算相对存活分数,即,在每个辐射剂量下的绝对存活分数(集落数/总涂板细胞)除以未照射细胞的绝对存活分数。

[0300] 西方分析.如先前所描述(Li,A.G.,Lu,S.L.,Zhang,M.X.,Deng,C.以及Wang,X.J.Smad3敲除小鼠展现出对皮肤化学癌形成的抗性(Smad3knockout mice exhibit a resistance to skin chemical carcinogenesis).Cancer Res 64,7836-7845(2004))进行蛋白质提取和西方分析。这项研究中所使用的抗体包括针对Smad7的兔抗体(1:500)、针对Smad2的兔抗体(1:300,Zymed,51-1300)和针对Smad4的兔抗体(1:300,Epitomics,1676-1)、针对Smad3的兔抗体(1:300,Cell Signaling Technology,9513)、针对Rac1的小鼠抗体(1:500,BD Biosciences,610651)、针对CtBP1的兔抗体(1:500,Millipore,07-306)、针对微管蛋白的小鼠抗体(1:3000,Sigma,T5168)、针对GAPDH的小鼠抗体(1:5000,Abcam,Ab8245),以及针对肌动蛋白的小鼠抗体(1:1000,Santa Cruz Biotechnology,SC1616)。使用ODYSSEY[®] 1.2版软件(LI-COR Biosciences)获得灰度图像。

[0301] Rac1活化分析法.使用用于Rac1活化的BIOCHEM[™]试剂盒(Cytoskeleton Inc,BK035)检查活性GTP结合的Rac1。在15cm直径的组织培养板中培养野生型和Smad7转基因角质细胞并且使用所提供的裂解缓冲液制备蛋白质裂解物。为了分析Rac1活性,使用1mg细胞裂解物。为了检查总Rac1和Smad7蛋白质,使用50μg裂解物。为了测量小鼠舌头中GTP结合的Rac1,将一半舌头在液氮中磨成粉末并且用裂解缓冲液裂解以提取蛋白质,在每个样品2mg蛋白质裂解物中分析GTP结合的Rac1,并且加载50μg蛋白质裂解物用于总Rac1蛋白质西方墨点法。

[0302] ChIP分析法.如先前所描述(Hoot,K.E.等HGF上调促成具有角质细胞特异性Smad2缺失的小鼠中的血管生成(HGF upregulation contributes to angiogenesis in mice with keratinocyte-specific Smad2deletion).J Clin Invest 120,3606-3616(2010); Hoot,K.E.等角质细胞特异性Smad2消融使得在皮肤癌形成和进展期间上皮-间质转变增加(Keratinocyte-specific Smad2ablation results in increased epithelial-mesenchymal transition during skin cancer formation and progression).Owens等,J.Clin.Invest 118,2722-2732(2008);Smad4依赖性桥粒芯蛋白4表达促成毛囊完整性(Smad4-dependent desmoglein-4expression contributes to hair follicle integrity).Owens等,Dev.Biol.322:156-166(2008)),使用ChIP-IT表达试剂盒(Active Motive,53009)进行ChIP分析法。从初级小鼠角质细胞中分离DNA-蛋白质复合物。对于ChIP,将6.3μg剪切的染色质与蛋白质G磁珠和2μg下列每一者一起孵育:针对Smad2的兔抗体(Cell Signaling Technology,3122)、针对Smad3的兔抗体(Cell Signaling Technology,9523)、针对Smad4的兔抗体(Cell Signaling Technology,9515)、Smad7抗体(Santa Cruz Biotechnology,SC-11392)、CtBP1(Millipore)或阴性对照兔IgG(Santa Cruz Biotechnology,SC-2027)。从蛋白质-DNA复合物中洗脱的DNA用于PCR分析,并且通过凝胶图像上的ChIP色带强度或通过使用Power SYBR Green Master Mix(Applied Biosystems)的定量PCR,在野生型和Smad7转基因角质细胞中比较CtBP1结合至Rac1启动子。用于扩增Rac1SBE-1.5Kb启动子区的引物:

[0303] 5'-TGGAATTCCTGGTCTGGTTT-3'(有义)(SEQ ID NO:13)

[0304] 5'-GCCAAGCTGCTCTTCCAGTA-3'(反义)(SEQ ID NO:14)

[0305] 5'-TCTCAGGGGGCCAAAGGTGTT-3'(有义)(SEQ ID NO:15)

[0306] 5'-TCCCAGCACCTGAATCACATGG-3'(反义)(SEQ ID NO:16)

[0307] Rac1启动子荧光素酶报告基因构建体、定点诱变以及荧光素酶分析法.使用5' XhoI和3'HindIII标记的引物从野生型小鼠DNA扩增包涵SBE-1.5Kb位点的Rac1启动子的-1671bp至-789bp的883bp片段,并且将这个Rac1启动子片段克隆至pGL4.26载体(Promega)中以制造Rac1启动子-pGL4.26荧光素酶报告基因(Rac1-Luc)构建体.对于定点诱变,使SBE序列5'-TGTCTGTGCT-3'(SEQ ID NO:17)突变成5'-TGATAGAGCT-3'(SEQ ID NO:18).使用Lipofectamine 2000(Invitrogen)将Rac1-Luc和pGL4.74(1:20)与Smad7siRNA、CtBP1siRNA或乱序siRNA一起共转染至初级小鼠角质细胞或用Tat-Smad7($1\mu\text{g mL}^{-1}$)处理的初级小鼠角质细胞.在转染或Tat-Smad7处理之后48小时,使用DUAL-LUCIFERASE®报告基因分析试剂盒(Promega),遵循制造商的说明书,收集细胞裂解物并且进行荧光素酶分析法.用Glomax机器(Promega)测量Rac1-荧光素酶活性,并且由萤火虫活性与海肾活性的比率来表示.用于扩增Rac1启动子序列的引物是:

[0308] 5'-ATCCTCGAG-TATCCTCCAGGTCTGGG-3'(SEQ ID NO:19)

[0309] 5'-GCCAAGCTT-AGCGTCCAGCGTTAACCTG-3'(SEQ ID NO:20)

[0310] 统计分析.使用司徒登氏t检验(Student's t-test)分析分子分析和口腔粘膜炎症尺寸中的统计差异,并且除溃疡尺寸之外的所有数据都以平均值 \pm s.d.呈现,溃疡尺寸以平均值 \pm s.e.m呈现.通过费氏精确检验来分析口腔粘膜炎症发生率。

[0311] 实施例6:针对在大肠杆菌或酵母中的Smad7蛋白质产生进行密码子优化

[0312] 尽管许多哺乳动物蛋白质可以在不存在核苷酸序列修饰的情况下在细菌中产生,但分析指示,将需要修饰Smad7核苷酸序列以允许蛋白质在细菌中表达。

[0313] Smad7cDNA哺乳动物密码子使用的分析揭示由以下核苷酸编码的九个精氨酸氨基酸:7-9、43-45、169-171、403-405、490-492、526-528、526-528、823-825、1057-1059是稀有密码子(AGG,密码子利用率1.7%)。因为这些密码子是细菌中的稀有密码子,所以预期其可以停止或减少细菌中的蛋白质翻译和/或产生.由稀有精氨酸密码子编码的氨基酸在所说明的人类Smad7蛋白质中以下面的粗体大写字母来指示,包括在位置3、15、57、135、164、169、176、275以及353处的精氨酸.另外,以下精氨酸密码子还具有低频率使用.CGA(3.5%密码子利用率):核苷酸16-18、136-138、199-201、598-600,其编码位置6、46、67、200处的精氨酸;CGG(5.4%密码子利用率):核苷酸31-33、112-114、316-318、772-774、940-942、973-975、1135-1137、1276-1278,其编码位置11、38、106、258、314、325、379、426处的精氨酸;AGA(2.8%密码子利用率):核苷酸637-639、814-816,其编码位置213、272处的精氨酸.这些精氨酸残基以下面的粗体大写R突出显示并且其在密码子优化核酸序列中的至少一者中变成CGC(20.6%密码子利用率):

```

1      mfRtkRsalv RrlwRsrpag gedeeegagg gggggelRge gatdsRahga
51     ggggpgRagc clgkavrgak ghhhphppaa gagaaggaea dlkalthsvl
101    kklkeRqlle llqavesrgg trtacllpg rldcRlgpga pagaqpaqp
151    ssyslplllc kvfRwpdlRh ssevkRlccc esygkinpel vccnphhlsR
[0314] 201    lcelespppp ysRypmdflk ptadcpdvp ssaetggtny lapgglstdsq
251    llepgdRsh wcvvayweek tRvgRlycvq epsldifydl pqngfclgq
301    lnsdnksqlv qkvRskigcg iqltRevdgv wvynrssypi fiksatlndp
351    dsRtllvhkv fpgfsikafd yekayslqRp ndhefmqqpw tgftvqisfv
401    kgwgqcytrq fisscpwle vifnsR (SEQ ID NO: 12)

```

[0315] 基于这个分析,决定将Smad7核苷酸序列优化成据信允许在大肠杆菌或酵母中增加Tat-Smad7蛋白质产生的密码子。下面提供了由Genscript制造的优化核酸密码子序列。简单地说,序列具有以下组成:核苷酸1-6包括BamHI的限制识别位点;核苷酸7-36包括Tat序列;核苷酸37-1314包括密码子优化的人类Smad7cDNA;核苷酸1342-1383包括V5表位;核苷酸1384-1386是终止密码子;并且核苷酸1387-1392包括SalI的限制识别位点。在这个序列中,去除ATG以与GST一起使用。基于“密码子使用数据库”将整个经过设计的序列转变成大肠杆菌密码子。初始优化Smad7序列(SEQ ID NO:23)如下所示:

```

1      ggatccggcc gtaaaaaacg ccgtaacgc cgccgtggtt tccgtacgaa acgctcggcc
61     ctggtccgtc gcctgtggcg ctcccgtgct ccgggtggtg aagatgaaga agaaggtgct
[0316] 121    ggcggcggtg gcggtggcgg tgaactgcgt ggcgagggtg caaccgatag tcgtgcacac
181    ggtgcaggcg gtggcggtcc gggtcgtgct ggttgctgct tgggtaaagc tgtgcgcggc

```


241 gcgaaaggtc atcaccatcc gcacccgccg gcagcaggtg caggtgcagc tggcgggtgcg
 301 gaagccgatc tgaaagccct gaccatagtg gtcctgaaaa aactgaaaga acgtcagctg
 361 gagctgctgc tgcaagcagt agaatcccg ggcggtaccc gtacggcttg tctgctgctg
 421 ccgggtcgtc tggattgccg tctgggtccg ggtgcaccgg ctggtgcgca gccggcacia
 481 ccgccgagct cttacagcct gccgctgctg ctgtgtaaag tgtttcgttg gccggacctg
 541 cgccacagtt ccgaagttaa acgcctgtgc tgttgcgaga gctatggcaa aattaacccg
 601 gaactggttt gttgcaatcc gcaccatctg tctcgtctgt gtgaactgga gagccccgccg
 661 ccgccgtatt ctcgttacc gatggatttc ctgaaaccga ctgcagattg cccggacgca
 721 gtcccgtcat cggctgagac cggcggcacc aactatctgg caccgggcgg tctgagtgat
 781 tcccagctgc tgctggaacc gggcgaccgt tcacattggt gtgtggttgc ctattgggaa
 [0317] 841 gagaaaacgc gtgtcggctc cctgtactgc gtacaggaac cgtcgttga tatctttat
 901 gacctgccgc agggcaatgg tttctgtctg ggccaactga actcagataa taaatcgag
 961 ctggtgcaaa aagttcgtc aaaaattggc tgcggtatcc agctgaccgc tgaagttgac
 1021 ggtgtctggg tatataaccg cagctcttac ccgatttta taaaagtgc caccctggat
 1081 aatccggact ccgctacgt gctggccac aaagtatttc cgggcttctc aatcaaagcg
 1141 ttcgattacg agaaagccta ctcgtgcag cggccgaacg accatgaatt catgcagcaa
 1201 ccgtggacgg gttttactgt gcagatctct ttcgttaaag gctgggggtca atgctacacc
 1261 cgtcagttta tctcgtcctg tccgtgctgg ctggaagtga tttcaatag ccgcaagggc
 1321 gagtcaatt cgaagctga aggtaagcct atccctaacc ctctcctcgg tctcgattct
 1381 acgtgagtcg ac (SEQ ID NO: 23)

[0318] 下面提供了Tat-Smad7-V5 (SEQ ID NO:23) 与人类Smad7 (SEQ ID NO:22) cDNA之间的核苷酸序列比较。人类Smad7和密码子优化的Tat-Smad7-V5共享68%的密码子同源性。人类Smad7和密码子优化的Tat-Smad7共享71%的密码子同源性。人类Smad7和密码子优化的Smad7共享73%的密码子同源性。

[0319] 比对:针对参考分子的全局DNA比对

[0320] 参数:记分矩阵:线性(错配2,开放空位4,延伸空位1)

[0321] 参考分子:人类Smad7mRNA,区1-1281

[0322] 比对的序列数:2

[0323] 设置:相似显著性截止值:≥90%

[0324] 匹配百分比的概述:

[0325] 参考:人类Smad7mRNA 1-1281 (1281bp) --

[0326] 序列2:Tat-Smad7-V5 1-1392 (1392bp) 68%

[0331] 据信Smad7在体内具有数种活性,包括(但不限于)以下一者或多者:增强细胞增殖、增强细胞迁移、减少DNA损伤、减少细胞凋亡以及减轻发炎。Smad7对这些过程的作用归因于以下一者或多者:阻断TGF- β 信号传导、阻断NF- κ B信号传导、阻断CtBP1活性和/或增加Rac1表达和/或活性。据信Ptd-Smad7的较小功能域可能足以递送治疗作用(参看例如图15)。另外,预期所得较短蛋白质序列增强蛋白质产生。另外,据信含有部分Smad7序列的不同截短Tat-Smad7蛋白质可以适用于不同治疗。

[0332] 举例来说,据信Smad7的C端MH2域(约一半长度的Smad7蛋白质,例如,208-426aa)可以主要介导Smad7的消炎作用(Hong等,Nat Immunology,8,504-513,2007)。具有这种消炎功能的Smad7肽可能是足够的并且任选地是一种改善,用于治疗慢性发炎相关的病状,尤其诸如(但不限于)口腔粘膜炎、口炎以及牛皮癣。

[0333] Smad7的N端MH1域加上连接子区(约一半的蛋白质,例如,2-208aa)已知活化MAPK并且结合至泛素E3连接酶Smurf以使TGF- β 受体降解(Aragon等,Structure 20:1726-1736(2012))。据信其可以主要介导细胞迁移和/或阻断TGF- β 诱导的生长停滞和/或纤维化反应。具有这种细胞迁移和增殖功能的Smad7肽可能是足够的并且任选地是一种改善,用于增强与过度发炎不相关的愈合。可能得益于这种形式的治疗的创口类型尤其包括(但不限于)手术创口、纤维化疤痕以及糖尿病创口、缺陷性愈合和/或疤痕。

[0334] 设计了截短的Smad7N端和C端PTD-融合蛋白质。下面提供了Tat-Smad7-C端密码子优化的核苷酸和蛋白质序列的一个实例。在核酸序列中,核苷酸1-6包括BamHI的限制识别位点;核苷酸7-36包括Tat PTD序列;核苷酸37-810包括针对人类Smad7的C端氨基酸258至426进行的密码子优化;核苷酸568-609包括V5表位序列;核苷酸610-612包括终止序列;并且核苷酸613-618包括SalI的限制识别位点:

ggatccggccgtaaaaaacgccgtcaacgccgccgttcacattggtgtgtggtgcctat
G S G R K K R R Q R R R S H W C V V A Y
tgggaagagaaaaacgcgtgtcggctgcctgtactgcgtacaggaaccgtcgtggatatac
W E E K T R V G R L Y C V Q E P S L D I
ttttatgacctgccgcagggcaatggtttctgtctgggccaactgaactcagataataaa
F Y D L P Q G N G F C L G Q L N S D N K
tcgcagctggtgcaaaaagttcgtcaaaaattggctgcggtatccagctgacccgtgaa
S Q L V Q K V R S K I G C G I Q L T R E
gttgacgggtgtctgggtatataaccgcagctcttaccgatttttatcaaaagtgccacc
[0335] V D G V W V Y N R S S Y P I F I K S A T
ctggataatccggactcccgtacgtgctggtccacaaagtatttccgggcttctcaatc
L D N P D S R T L L V H K V F P G F S I
aaagcgttcgattacgagaaagcctactcgtgcagcgcccgaaacgaccatgaattcatg
K A F D Y E K A Y S L Q R P N D H E F M
cagcaaccgtggacgggttttactgtgcagatctcttctgtaaaaggctggggtcaatgc
Q Q P W T G F T V Q I S F V K G W G Q C
tacaccgctcagtttatctcgtcctgtccgtgctggctggaagtattttcaatagccgc
Y T R Q F I S S C P C W L E V I F N S R
aagggcgagctcaattcgaagcttgaaggttaagcctatccctaaccctctcctcggtctc
K G E L N S K L E G K P I P N P L L G L

[0336] gattctacgtgagtcgac (SEQ ID NO: 24)

D S T - (SEQ ID NO: 25)

[0337] 在核酸序列中,核苷酸1-6包括BamHI的限制识别位点;核苷酸7-36包括Tat PTD序列;核苷酸37-810包括针对人类Smad7的N端氨基酸1-258进行的密码子优化;核苷酸811-852包括V5表位序列(对应的氨基酸序列加粗);核苷酸853-855包括终止序列;并且核苷酸856-861包括SalI的限制识别位点。去除ATG以允许与GST融合:

[0338]

ggatccggccgtaaaaaacgccgtcaacgccgccgtggtttccgtacgaaacgctcggcc
G S G R K K R R Q R R R G F R T K R S A
ctggtcgctgcctgtggcgctcccggtgctccgggtggtgaagatgaagaagaagggtgct
L V R R L W R S R A P G G E D E E E G A
ggcggcggtggcggtggcggtgaactgcgtggcgagggtgcaaccgatagtcgtgcacac
G G G G G G G E L R G E G A T D S R A H
ggtgcaggcggtggcggtccgggtcgtgctggttgcgtgctgggtaaagctgtgcgcggc
G A G G G G P G R A G C C L G K A V R G
gcgaaaggatcatcaccatccgcacccgccggcagcaggtgcaggtgcagctggcggtgcg
A K G H H H P H P P A A G A G A A G G A
gaagccgatctgaaagccctgacctatagtgctctgaaaaactgaaagaacgtcagctg
E A D L K A L T H S V L K K L K E R Q L
gagctgctgctgcaagcagtagaatcccgtggcggtaccggtacggcttgcgtgctgctg
E L L L Q A V E S R G G T R T A C L L L
ccgggtcgtctggattgccgtctgggtccgggtgcaccggctggtgcgcagccggcacia
P G R L D C R L G P G A P A G A Q P A Q
ccgccgagctcttacagcctgccgtgctgctgtgtaaagtgttcgttggccggacctg
P P S S Y S L P L L L C K V F R W P D L
cgccacagttccgaagttaaacgcctgtgctgttgcgagagctatggcaaaattaaccg
R H S S E V K R L C C C E S Y G K I N P
gaactgggttgtgcaatccgcacccatctgtctcgtctgtgtaactggagagcccgcg
E L V C C N P H H L S R L C E L E S P P
ccgccgtattctcgttaccgatggatttctgaaaccgactgcagattgcccggacgca
P P Y S R Y P M D F L K P T A D C P D A
gtcccgtcatcggtgagaccggcgccaccaactatctggcaccggcggtctgagtgat
V P S S A E T G G T N Y L A P G G L S D
tcccagctgctgctggaaccggcgaccgtggtgaagcctatccctaaccctctcctcggt
S Q L L L E P G D R G K P I P N P L L G
ctcgattctacgtgagtcgac (SEQ ID NO: 26)
L D S T - (SEQ ID NO: 27)

[0339]

[0340] 实施例8:测试截短的Smad7蛋白质

[0341] 使用上文所描述的用于测试全长Smad7的体外和体内分析法以及其它分析法来测试截短的Smad7蛋白质的活性。这类分析法包括(但不限于)阻断Smad2的磷酸化和/或NF- κ B p50亚单元的核转位,增加细胞增殖,减少凋亡和/或辐射诱导的DNA损伤,减少发炎和/或血管生成,促进口腔粘膜炎、手术创口、糖尿病创口和/或与小鼠慢性发炎相关的创口的愈合的能力。

[0342] 在创口愈合分析法中,在野生型小鼠中进行6-mm钻取活组织检查,接下来每天局部施加C端或N端Tat-Smad7。通过测量大体创口闭合,发现上文所描述的两种截短的Smad7蛋白质(例如,Tat-Smad7C端和N端蛋白质)类似于全长Tat-Smad7促进创口愈合。

[0343] 实施例9:截短的Smad7蛋白质加速创口愈合

[0344] 在野生型小鼠中检查截短的Smad7蛋白质加速创口愈合的能力。图16A展示了C端截短(259-426aa)的Tat-C-Smad7对小鼠创口愈合模型的作用。使野生型C57BL/6小鼠麻醉,通过6-mm皮肤钻取活组织检查在背侧开创口,并且每隔一天通过局部施加PBS(对照,3只小鼠,4个创口/小鼠)、全长Tat-Smad7(0.4 μ g/10 μ L PBS/创口,3只小鼠,4个创口/小鼠)或Tat-C-Smad7(0.4 μ g/10 μ L PBS/创口,3只小鼠,4个创口/小鼠,处理总共12个创口)来处理。用佳能数码相机为创口区域拍照,并且通过在开创口之后第1天、第2天、第4天以及第5天使用Image J软件以照片内部的6-mm圆形用于标准化使分析成像来测量,并且对于每个处理组计算剩余的平均创口区域百分比。Tat-C-Smad7加速创口愈合的能力类似于全长Smad7(图16A)。

[0345] 图16B图解了N端截短(1-258aa)的Tat-N-Smad7对创口愈合的作用。使野生型C57BL/6小鼠麻醉,通过6-mm皮肤钻取活组织检查在背侧开创口,并且每隔一天通过局部施加PBS(对照,6只小鼠,4个创口/小鼠)、全长Tat-Smad7(0.4 μ g/10 μ L PBS,6只小鼠,4个创口/小鼠)或Tat-N-Smad7(0.4 μ g/10 μ L PBS,6只小鼠,4个创口/小鼠)来处理。通过在开创口之后第1天、第2天以及第4天使用Image J软件以照片内部的6-mm圆形用于标准化使分析成像来测量创口区域,并且对于每个处理组计算剩余的平均创口区域百分比。两天后在对照与Tat-Smad7处理的小鼠之间的创口愈合速率存在显著差异,后者更快愈合(图16B, $p < 0.05$)。尽管Tat-N-Smad7作为阴性对照来测试,但其意想不到地仅在一天后相对于对照小鼠促进显著加速的创口愈合,并且在两天后相对于全长Tat-Smad7处理的小鼠显著加速创口愈合(图16B, $p < 0.05$)。Tat-Cre相对于对照处理对创口闭合没有作用(数据未示出)。

[0346] 这些结果展示,全长Tat-Smad7与截短的Tat-Smad7蛋白质(Tat-C-Smad7和Tat-N-Smad7)都促进创口愈合。此外,这些结果展示,某些截短的Tat-Smad7蛋白质在加速创口愈合方面比全长Tat-Smad7更有效。

[0347] 因此,包含截短的Tat-Smad7蛋白质的组合物适用于治疗创口和加速创口愈合。

[0348] 实施例10:Smad7加速受损创口愈合模型中的创口愈合

[0349] 如上文实施例9中所描述通过6-mm皮肤钻取活组织检查将糖尿病(db/db)小鼠在背侧开创口,并且在8天前当创口并未完全被疮痂覆盖时每隔一天通过将PBS(对照,6只小鼠,4个创口/小鼠)、全长Tat-Smad7(0.4 μ g/10 μ L PBS,6只小鼠,4个创口/小鼠)或REGRANEX®乳膏(用棉签涂抹,6只小鼠,4个创口/小鼠)局部施加至创口来处理。10天后,对疮痂与创口外周之间的间隙局部施加处理以避免硬痂和未受创的角质层的阻挡。含

有重组人类血小板源性生长因子 (PDGF) 的 **REGRANEX®** 经批准用于局部施用于糖尿病性溃疡。

[0350] 在开创口之后第1天、第2天、第4天、第6天、第8天、第9天、第10天、第11天、第12天以及第13天在视觉上评估创口闭合 (图17A)。在每个时间点,用佳能数码相机为创口拍照。8天后,与对照相比在Tat-Smad7处理的小鼠中创口闭合有明显的视觉改善 (图17A)。

[0351] 通过在开创口之后第1天、第2天、第3天、第4天、第7天、第9天以及第11天使用 Image J 软件以照片内部的6-mm圆形用于标准化来测量创口尺寸,并且对于每个处理组计算剩余的平均创口区域百分比。Tat-Smad7加速创口愈合的能力类似于 **REGRANEX®** (图17B)。在第7天,图像分析结果指示,创口闭合在Tat-Smad7处理的小鼠中相对于对照显著加速 (图17B, $p < 0.05$)。这些结果类似于使用 **REGRANEX®** 所达成的那些。

[0352] 使用苏木精和伊红 (H&E) 对来自第8天的创口样品的福尔马林固定的石蜡背侧创口切片 (1mm) 进行染色。第8天的创口样品的组织学比较揭示在Smad7处理的db/db小鼠中相对于对照完全再上皮化和加速的创口闭合 (图17C)。

[0353] 总体来说,这些结果展示,Tat-Smad7适用于加速不良愈合的糖尿病性创口的创口闭合并且提供 **REGRANEX®** 治疗的替代方案。

[0354] 实施例11:针对Smad7蛋白质产生的额外密码子优化

[0355] Smad7核酸分子被设计成具有所选的额外核苷酸变化以增加蛋白质产生。举例来说,将操纵编码氨基酸Ser和His的密码子的利用。在上述实施例中的密码子优化的人类Smad7中,Ser密码子 (TCC或TCG) 具有约9%的密码子利用率的氨基酸频率。据信,使Ser的密码子变成AGC将增加Smad7蛋白质产生,这至少一部分是因为其可以任选地使密码子使用率增至15%。在Smad7蛋白质中存在33个Ser氨基酸 (在位置19-21、46-48、133-135、292-294、349-351、451-453、454-456、460-462、511-513、514-516、544-546、595-597、616-618、634-636、691-693、694-696、739-741、745-747、775-777、847-849、907-909、919-921、943-945、1006-1008、1009-1101、1030-1032、1054-1056、1093-1095、1126-1128、1192-1194、1237-1239、1240-1242、1273-1275处的核苷酸;对应的丝氨酸氨基酸位置7、16、45、98、117、151、152、154、171、172、182、199、206、212、231、232、247、249、259、283、303、307、315、336、337、344、352、365、376、398、413、414、425)。在这些当中,在不引入潜在的替代性开放阅读框架的情况下可以改变23个 (核苷酸19-21、292-294、349-351、451-453、454-456、460-462、511-513、514-516、544-546、616-618、634-636、691-693、694-696、739-741、745-747、775-777、847-849、907-909、919-921、1009-1101、1030-1032、1054-1056、1093-1095;对应的丝氨酸氨基酸位置7、98、117、151、152、154、171、172、182、206、212、231、232、247、249、259、283、303、307、337、344、352、365)。

[0356] 类似地,在上述实施例中的密码子优化的人类Smad7中,His密码子 (CAC) 具有9.6%的密码子使用率。据信,使His密码子变成CAT (任选地变至12.6%的使用率) 将增加Smad7蛋白质产生。在Smad7蛋白质中存在12个His (核苷酸142-144、214-216、217-219、220-222、226-228、289-291、589-591、778-780、1072-1074、1147-1149;对应于位置48、72、73、74、76、97、170、196、197、260、358、383处的组氨酸氨基酸)。在这些当中,在不引入潜在的替代性开放阅读框架的情况下可以改变4个 (核苷酸217-219、220-222、589-591、778-780,组

氨酸残基73、76、197、260)。

[0357] 另外,野生型人类Smad7包括Met氨基酸作为氨基酸216 (Met216)。这可以被例如细菌机构认为是替代性开放阅读框架,并且减少蛋白质产生。据信,使Met216变成Leu216 (ATG变成CTG),即具有最接近Met的生物化学特性并且因此预期不会改变蛋白质的3D结构的氨基酸,将增加蛋白质产生。

[0358] 下面提供了原始密码子优化的Tat-Smad7-V5与进一步改变之间的比较。上链:Tat-Smad7-V5 (SEQ ID NO:23):下链:在优化的Ser、His以及M216L突变之后 (SEQ ID No:30)。

[0359] 比对:局部DNA同源性

[0360] 参数:双链 方法:FastScan-Max Score

[0361] Mol 1 201203231113102B12-S54366-164160-1-PGEX-5.seq (1-1463) Mol 2TatSmad7 ser-His优化

[0362] 比对的序列数:2

[0363] 设置:相似显著性截止值:≥60%

[0364] 同源区块:匹配百分比94 得分1227 长度1392

[0365]

```

201203231131 30 ggatccggccgtaaaaaacgcgcgtcaacgcgcgcgtggtttccgtacgaaacgctcggcc
TatSmad7 Ser 1 .....agc...

201203231131 90 ctggtecgctcgctgtggcgctcccgtgctccgggtgggaagatgaagaagaaggtgct
TatSmad7 Ser 61 .....ag...

201203231131 150 ggcggcggtggcggtggcggtgaactgcgtggcgagggtgcaaccgatagtcgtgcacac
TatSmad7 Ser 121 .....C.....t...

201203231131 210 ggtgcaggcggtggcggtccgggtcgctgctggttgctgtctgggtaaagctgtgcgcggc
TatSmad7 Ser 181 .....

201203231131 270 gcgaaaggtcatcagcatccgcagccgcggcgagcaggtgcaggtgcagctggcggtgcg
TatSmad7 Ser 241 .....t...t...t...

201203231131 330 gaagccgatctgaaagccctgaccatagtgctcctgaaaaaactgaaagaacgtcagctg
TatSmad7 Ser 301 .....C.....

201203231131 390 gagctgctgctgcaagcagtagaatcccggtggcggtaccgctacggcttgctgctgctg
TatSmad7 Ser 361 .....ag...

201203231131 450 ccgggtcgctcggattgcgctcgggtccgggtgcaccggctggtgcgcagccggcacaa
TatSmad7 Ser 421 .....

201203231131 510 ccgcgcagctcttacagcctgcgcctgctgctgctgtaaagtgttccgttggccggacctg
TatSmad7 Ser 481 .....agc...

201203231131 570 cgccacagttccgaagttaaaccgcctgtgctgttgcgagagctatggcaaaattaaccgg
TatSmad7 Ser 541 .....t...cag...t...

201203231131 630 gaactggtttgttgcaatccgcaccatctgtctcgtctgtgtaactggagagccgcgcg
TatSmad7 Ser 601 .....t.....agc...

201203231131 690 ccgcgcgtattctcgttaaccgatggatttcctgaaaccgactgcagattgcccggaacga
TatSmad7 Ser 661 .....agc.....C.....

201203231131 750 gtcccgatcatcggtgagaccggcgccacaaactatctggcaccggcggtctgagtgat
TatSmad7 Ser 721 .....agcagc.....C...

201203231131 810 tcccagctgctgctggaaccggcgacccgttcacattgggtgtgtggttgccattgggaa
TatSmad7 Ser 781 ag.....agc...

201203231131 870 gagaaaaacgcgtgtcggtcgctgtactgcgtacaggaaccgctcgctggatatctttat
TatSmad7 Ser 841 .....agc...

201203231131 930 gacctgcgcagggaatggtttctgtctgggccaaactgaactcagataataaatccgag
TatSmad7 Ser 901 .....agc.....agc...

201203231131 990 ctggtgcaaaaagtctgcctcaaaaattggctgcggtatccagctgacccgtgaagttgac
TatSmad7 Ser 961 .....agc...

201203231131 1050 ggtgtctgggtatataaccgcagctcttacccgatttttatcaaaagtgcaccctggat
TatSmad7 Ser 1021 .....C.....

201203231131 1110 aatccggactcccgtaacgtgctggtccacaaagtatttcgggcttctcaatcaaaagcg
TatSmad7 Ser 1081 .....ag.....t.....agc...

201203231131 1170 ttcgattacgagaaagcctactcgtgcagcgcccgaaacgaccatgaattcatgcagaa
TatSmad7 Ser 1141 .....agc...

201203231131 1230 ccgtggacgggttttactgtgcagatctcttccgttaaaggctggggtcaatgctacacc
TatSmad7 Ser 1201 .....agc...

201203231131 1290 cgtcagtttatctcgtcctgctccgtgctggctggaagtgattttcaatagccgcaagggc
TatSmad7 Ser 1261 .....agcag...

201203231131 1350 gagctcaattcgaagcttgaaggtgaagcctatccctaaccctctcctcggtctcgattct
TatSmad7 Ser 1321 .....agc.....agc...

201203231131 1410 acgtgagtcgac
TatSmad7 Ser 1381 .....

```

[0366] 下面提供了将包括所有这些变化的核酸序列和其对应的氨基酸序列。氨基酸序列包括以粗体指示的V5表位,以及以斜体和下划线指示的pET101-Topo骨架。Tat-Smad7^{M216L}完全优化的全长核苷酸和蛋白质序列如下所示:

[0367]

ggatccggccgtaaaaaacgccgtcaacgccgccgtggtttccgtacgaaacgcagcgcc
G S G R K K R R Q R R R G F R T K R S A
ctggtccgtcgcctgtggcgcagccgtgctccgggtggtgaagatgaagaagaaggtgct
L V R R L W R S R A P G G E D E E E G A
ggcggcggtggcgggtggcgggtgaactgcgtggcgagggtgcaaccgatagccgtgcacat
G G G G G G G E L R G E G A T D S R A H
ggtgcaggcggtggcgggtccgggtcgtgctggttgcgtgctgggtaaagctgtgcgcggc
G A G G G G P G R A G C C L G K A V R G
gcgaaaggtcatcatcatccgcatccgccggcagcaggtgcaggtgcagctggcgggtgcg
A K G H H H P H P P A A G A G A A G G A
gaagccgatctgaaagccctgacctatagcgtcctgaaaaaactgaaagaacgtcagctg
E A D L K A L T H S V L K K L K E R Q L
gagctgctgctgcaagcagtagaaagccgtggcgggtaccggtacggcttgcgtgctgctg
E L L L Q A V E S R G G T R T A C L L L
ccgggtcgtctggattgccgtctgggtccgggtgcaccggctggtgcgcagccggcacia
P G R L D C R L G P G A P A G A Q P A Q
ccgccgagcagctacagcctgccgctgctgctgtgtaaagtgttcgttggccggacctg
P P S S Y S L P L L L C K V F R W P D L
cgccatagcagcgaagttaaagcgcctgtgctgttgcgagagctatggcaaaattaacccg
R H S S E V K R L C C C E S Y G K I N P
gaactggtttgttgaatccgcatcatctgagccgtctgtgtgaactggagagcccgcg
E L V C C N P H H L S R L C E L E S P P
ccgccgtatagccgttaccgctggatttctgaaaccgactgcagattgcccgagcgca
P P Y S R Y P L D F L K P T A D C P D A

gtcccgagcagcgctgagaccggcgccaccaactatctggcaccggcggtctgagcgat
V P S S A E T G G T N Y L A P G G L S D
agccagctgctgctggaaccggcgaccgtagccattggtgtgtggtgcctattgggaa
S Q L L L E P G D R S H W C V V A Y W E
gagaaaacgcgtgtcggtcgctgtactgcgtacaggaaccgagcctggatatctttat
E K T R V G R L Y C V Q E P S L D I F Y
gacctgccgcagggcaatggtttctgtctgggccaactgaacagcgataataaaagccag
D L P Q G N G F C L G Q L N S D N K S Q
ctggtgcaaaaagtctgcagcaaaattggctgcggtatccagctgacctgaagttgac
L V Q K V R S K I G C G I Q L T R E V D
ggtgtctgggtatataaccgcagcagctacccgattttatcaaaagcgccaccctggat
[0368] G V W V Y N R S S Y P I F I K S A T L D
aatccggacagccgtacgtgctggtccataaagtatttccgggcttcagcatcaaagcg
N P D S R T L L V H K V F P G F S I K A
ttcgattacgagaaagcctacagcctgcagcgcccgaacgacctgaattcatgcagcaa
F D Y E K A Y S L Q R P N D H E F M Q Q
ccgtggacgggtttactgtgcagatcagcttcgttaaaggctgggggtcaatgctacacc
P W T G F T V Q I S F V K G W G Q C Y T
cgctcagtttatcagcagctgtccgtgctggctggaagtattttcaatagccgcaagggc
R Q F I S S C P C W L E V I F N S R K G
gagctcaatagcaagcttgaaggttaagcctatccctaaccctctcctcggtctcgatagc
E L N S K L E G K P I P N P L L G L D S
acgtgagtcgac (SEQ ID NO: 30)
T - (SEQ ID NO: 31)

[0369] 优化的核苷酸和氨基酸序列还将用于制造多种N端和C端Tat-Smad7片段。下面提供了代表性实例。

[0370] 提供了Tat-N-Smad7-V5最优化的核苷酸和氨基酸序列。蛋白质序列包括以粗体大写字母指示的V5表位。

ggatccggccgtaaaaaacgccgtcaacgccgccgtggtttccgtacgaaacgcagcgcc

G S G R K K R R Q R R R G F R T K R S A

ctggtcgctgcctgtggcgagccgtgctccgggtggtgaagatgaagaagaaggtgct

L V R R L W R S R A P G G E D E E E G A

ggcggcggtggcggtggcggtgaactgcgtggcgagggtgcaaccgatagccgtgcacat

G G G G G G G E L R G E G A T D S R A H

ggtgcaggcggtggcggtccgggtcgtgctggttgcgtgctgggtaaagctgtgcgcggc

G A G G G G P G R A G C C L G K A V R G

gcgaaaggtcatcatcatccgcatccgccggcagcaggtgcaggtgcagctggcggtgcg

A K G H H H P H P P A A G A G A A G G A

gaagccgatctgaaagccctgacctatagcgtcctgaaaaaactgaaagaacgtcagctg

E A D L K A L T H S V L K K L K E R Q L

gagctgctgctgcaagcagtagaaagccgtggcggtaccggtacggcttgcgtgctgctg

[0371]

E L L L Q A V E S R G G T R T A C L L L

ccgggtcgtctggattgccgtctgggtccgggtgcaccggctggtgcgcagccggcacia

P G R L D C R L G P G A P A G A Q P A Q

ccgccgagcagctacagcctgccgtgctgctgtgtaaagtgttcgttggccggacctg

P P S S Y S L P L L L C K V F R W P D L

cgccatagcagcgaagttaaacgcctgtgctgttgcgagagctatggcaaaattaacccg

R H S S E V K R L C C C E S Y G K I N P

gaactggtttgttgaatccgcatcatctgagccgtctgtgtgaactggagagcccgcg

E L V C C N P H H L S R L C E L E S P P

ccgccgtatagccgttaccgatggatttctgaaaccgactgcagattgcccgagcgca

P P Y S R Y P M D F L K P T A D C P D A

gtcccagcagcgctgagaccggcggcaccaactatctggcaccggggcggtctgagcgat

V P S S A E T G G T N Y L A P G G L S D

agccagctgctgctggaaccgggcgaccgtggtaagcctatccctaaccctctcctcggt

S Q L L L E P G D R G K P I P N P L L G

ctcgattctacgtgagtcgac (SEQ ID NO: 32)

[0372]

L D S T - (SEQ ID NO: 27)

[0373] 提供了Tat-C-Smad7-V5最优化的核苷酸和氨基酸序列。蛋白质序列包括V5表位(以粗体大写字母指示),和pET101-Topo骨架(以加下划线的斜体指示)。

```
ggatccggccgtaaaaaacgccgtcaacgccgccgttcacattggtgtgtggtgcctat
G S G R K K R R Q R R R S H W C V V A Y
tggaagagaaaaacgcgtgtcggtcgctgtactgcgtacaggaaccgagcctggatc
W E E K T R V G R L Y C V Q E P S L D I
tttatgacctgccgcagggcaatggtttctgtctgggccaactgaacagcgataataa
F Y D L P Q G N G F C L G Q L N S D N K
agccagctggtgcaaaaagttcgagcaaaattggctgcggtatccagctgacctgaa
S Q L V Q K V R S K I G C G I Q L T R E
gttgacggtgtctgggtatataaccgcagcagctacccgattttatcaaaagcgccacc
V D G V W V Y N R S S Y P I F I K S A T
ctggataatccggacagccgtacgctgctgggtccataaagtatttccgggcttcagcatc
[0374] L D N P D S R T L L V H K V F P G F S I
aaagcgttcgattacgagaaagcctacagcctgcagcgcccgaacgaccatgaattcatg
K A F D Y E K A Y S L Q R P N D H E F M
cagcaaccgtggacgggttttactgtgcagatcagcttcgttaaaggctgggggtcaatgc
Q Q P W T G F T V Q I S F V K G W G Q C
tacaccgtcagtttatcagcagctgtccgtgctggctggaagtattttcaatagccgc
Y T R Q F I S S C P C W L E V I F N S R
aaggcgagctcaatagcaagcttgaaggtaagcctatccctaaccctctcctcggtctc
K G E L N S K L E G K P I P N P L L G L
gatagcacgtgagtcgac (SEQ ID NO: 34)
```

D S T - (SEQ ID NO: 25)

[0375] 下面提供了在上述优化之前与之后的比较。C端优化(上链)(比对以出现的顺序分别公开了SEQ ID NO 34和24)：

[0376] 比对：局部DNA同源性

[0377] 参数：双链 方法：FastScan-Max Score

[0378] Mol 1Tat-C-Smad7-ser-his优化(1-618) Mol 2Tat-C端Smad7-V5(1-618)

[0379] 比对的序列数：2

[0380] 设置：相似显著性截止值： $\geq 60\%$

[0381] 同源区块：匹配百分比93 得分541 长度618

```

Tat-C-Smad7-      1 ggatccggccgtaaaaaacgcgtcaacgcgcgcgttcacattggtgtgtggttgcctat
Tat-C-terminal    1 .....tgc.....

Tat-C-Smad7-      61 tgggaagagaaaaacgcgtgtcgggtcgctgtactgcgtacaggaaccgagcctggatgc
Tat-C-terminal    61 .....tgc.....

Tat-C-Smad7-     121 ttttatgacctgccgcagggaatggtttctgtctgggccaactgaacagcgataataaa
Tat-C-terminal   121 .....tca.....

Tat-C-Smad7-     181 agccagctggtgcaaaaagtctgcagcaaaattggctgcggtatccagctgaccggtgaa
Tat-C-terminal   181 tgc.....tca.....

Tat-C-Smad7-     241 gttgacggtgtctgggtatataaccgcagcagctacccgatttttatcaaaagcgccacc
Tat-C-terminal   241 .....tct.....

[0382] Tat-C-Smad7-     301 ctggataatccggacagccgtacgctgctggtccataaagtatttccgggcttcagcatc
Tat-C-terminal   301 .....tc.....c.....tca...

Tat-C-Smad7-     361 aaagcgttcgattacgagaaagcctacagcctgcagcgcccgaaacgaccatgaattcatg
Tat-C-terminal   361 .....tgc.....

Tat-C-Smad7-     421 cagcaaccgtggacgggttttactgtgcagatcagcttcgttaaaggctggggtcaatgc
Tat-C-terminal   421 .....tct.....

Tat-C-Smad7-     481 tacaccgctcagtttatcagcagctgtccgtgctggctggaagtgattttcaatagccgc
Tat-C-terminal   481 .....tcgtc.....

Tat-C-Smad7-     541 aagggcgagctcaatagcaagcttgaaggtaagcctatccctaaccctctcctcggtctc
Tat-C-terminal   541 .....tgc.....

Tat-C-Smad7-     601 gatagcacgtgagtcgac
Tat-C-terminal   601 ...tct.....

```

[0383] N端优化(上链)(比对以出现的顺序分别公开了SEQ ID NO 32和26):

[0384] 比对:局部DNA同源性

[0385] 参数:双链 方法:FastScan-Max Score

[0386] Mol 1Tat-N-Smad7-V5-Ser-His优化(1861)Mol 2Tat-N-Smad7-V5(1-861)

[0387] 比对的序列数:2

[0388] 设置:相似显著性截止值:≥60%

[0389] 同源区块:匹配百分比95 得分781 长度861

Tat-N-Smad7- 1 ggatccggccgtataaaaaacgcgcgtcaacgcgcgcgtgggttccgtacgaaacgcagcgcc
 Tat-N-Smad7- 1tcg...
 Tat-N-Smad7- 61 ctgggtcgcgcgcgtgtggcgagccgtgctccgggtggtgaagatgaagaagaaggtgct
 Tat-N-Smad7- 61tc...
 Tat-N-Smad7- 121 ggcgccggtggcggtggcggtgaactgcgtggcgaggggtgcaaccgatagccgtgcacat
 Tat-N-Smad7- 121t...c
 Tat-N-Smad7- 181 ggtgcaggcggtggcggtccgggtcgtgctggttgcgtgtctgggtaaagctgtgcgcggc
 Tat-N-Smad7- 181
 Tat-N-Smad7- 241 gcgaaaggtcatcatcatccgcacccgcggcagcaggtgcaggtgcagctggcggtgcg
 Tat-N-Smad7- 241c.....c.....
 Tat-N-Smad7- 301 gaagccgatctgaaagccctgaccatagcgtcctgaaaaaactgaaagaacgtcagctg
 Tat-N-Smad7- 301t.....
 Tat-N-Smad7- 361 gagctgctgctgcaagcagtagaaagccgtggcggtaccggtacggcttgcctgctgctg
 Tat-N-Smad7- 361tc.....
 [0390] Tat-N-Smad7- 421 ccgggtcgtctggattgccgtctgggtccgggtgcaccggctggtgcgcagccggcaca
 Tat-N-Smad7- 421
 Tat-N-Smad7- 481 ccgcgagcagctacagcctgccgtgctgctgtgttaaagtgttccgttggccggacctg
 Tat-N-Smad7- 481tct.....
 Tat-N-Smad7- 541 cgccatagcagcgaagttaaaccgcctgtgctgttgcgagagctatggcaaaattaacccg
 Tat-N-Smad7- 541c...tct.....
 Tat-N-Smad7- 601 gaactggttgttgcaatccgcacatctgagccgtctgtgtgaactggagagcccgccg
 Tat-N-Smad7- 601c.....tct.....
 Tat-N-Smad7- 661 ccgcggtatagccggtaccgcgatggatttccctgaaaccgactgcagattgcccggaacga
 Tat-N-Smad7- 661tct.....
 Tat-N-Smad7- 721 gtcccgagcagcgtgagaccggcgccaccaactatctggcaccggggcgtctgagcgat
 Tat-N-Smad7- 721tcacg.....t...
 Tat-N-Smad7- 781 agccagctgctgctggaaccggcgaccgtggttaagcctatccctaaccctctcctcggt
 Tat-N-Smad7- 781 tc.....
 Tat-N-Smad7- 841 ctcgattctacgtgagtcgac
 Tat-N-Smad7- 841

[0391] 另外,还将评估其它密码子优化核酸在一个或多个表达系统中产生Smad7蛋白质的能力。下面提供了这样的序列的另一个实例。

[0392] 通过Optimizer程序优化的Tat-Smad7^{M7216L}-V5:

ggatccggtcgtaaaaaacgtcgtcagcgtcgtcgtggttccgtaccaaacgttctgcg

[0393] G S G R K K R R Q R R R G F R T K R S A

ctggttcgtcgtctgtggcgttctcgtgcgccgggtggtgaagacgaagaagaaggtgcg

L V R R L W R S R A P G G E D E E E G A

[0394]

gggtggtggtggtggtggtggtgaactgcgtggtgaaggtgcgaccgactctcgtagcgac
G G G G G G G E L R G E G A T D S R A H
ggtagcgggtggtggtggtccgggtcgtgcgggtgctgcctgggtaaagcgggtcgtggt
G A G G G G P G R A G C C L G K A V R G
gcgaaaggtcaccaccacccgcacccgccggcgcggtgcggtgcgcggggtggtgcg
A K G H H H P H P P A A G A G A A G G A
gaagcggacctgaaagcgtgacccactctgttctgaaaaactgaaagaacgtcagctg
E A D L K A L T H S V L K K L K E R Q L
gaactgctgctgcaggcgggtgaatctcgtggtggtaccggtaccggtgctgctgctg
E L L L Q A V E S R G G T R T A C L L L
ccgggtcgtctggactgccgtctgggtccgggtgcgccggcggtgcgagccggcgag
P G R L D C R L G P G A P A G A Q P A Q
ccgccgtcttctactctctgccgtgctgctgtgcaaagtttccgttgccggacctg
P P S S Y S L P L L L C K V F R W P D L
cgtcactcttctgaagttaaactgtgtgctgctgcgaatcttacggtaaaatcaaccgc
R H S S E V K R L C C C E S Y G K I N P
gaactggtttgctgcaaccgcaccacctgtctcgtctgtgcgaactggaatctccgccg
E L V C C N P H H L S R L C E L E S P P
ccgccgtactctcgttaccgctggacttctgaaaccgaccgcgactgcccggacgcg
P P Y S R Y P L D F L K P T A D C P D A
gttccgtcttctgcggaaccgggtggtaccaactacctggcgccgggtggtctgtctgac
V P S S A E T G G T N Y L A P G G L S D
tctcagctgctgctggaaccgggtgaccgttctcactggtgcgttggtgctactgggaa
S Q L L L E P G D R S H W C V V A Y W E
gaaaaaaccgtgttggtcgtctgtactgcgttcaggaaccgtctctggacatcttctac
E K T R V G R L Y C V Q E P S L D I F Y
gacctgccgcagggtaacggttctgcctgggtcagctgaactctgacaacaaatctcag
D L P Q G N G F C L G Q L N S D N K S Q

ctgggtcagaaagttcgttctaaaatcggttgcggtatccagctgacccgtgaagttgac
 L V Q K V R S K I G C G I Q L T R E V D
 ggtgtttgggtttacaaccgttcttaccgatcttcatcaaactcgcgaccctggac
 G V W V Y N R S S Y P I F I K S A T L D
 aaccggactctcgtaccctgctggtcacaagttttccgggtttctctatcaaagcg
 N P D S R T L L V H K V F P G F S I K A
 ttcgactacgaaaaagcgtactctctgcagcgtccgaacgaccacgaattcatgcagcag
 [0395] F D Y E K A Y S L Q R P N D H E F M Q Q
 ccgtggaccgggttcaccgttcagatctcttcgttaaaggtggggtcagtgtacacc
 P W T G F T V Q I S F V K G W G Q C Y T
 cgtcagttcatctcttcttgcctgctgggtggaagttatcttcaactctcgtggtaaa
 R Q F I S S C P C W L E V I F N S R G K
 ccgatcccgaaaccgctgctgggtctggactctacctgagtcgac (SEQ ID NO: 36)
 P I P N P L L G L D S T - - (SEQ ID NO: 37)

[0396] 核苷酸序列:

[0397] 1-6:BamHI;7-36:Tat;37-1314:密码子优化的人类Smad7;1315-1356:V5;137-1359:终止;1360-1365SalI

[0398] 去除ATG以与GST一起使用;682ATG变成CTG (M216变成L)

1 ggatccggtc gtaaaaaacg tcgtcagcgt cgtcgtgggt tccgtaccaa acgttctgcg
 61 ctggttcgtc gtctgtggcg ttctcgtgcg ccgggtgggtg aagacgaaga agaaggtgcg
 121 ggtggtgggtg gtggtgggtg tgaactgcgt ggtgaaggtg cgaccgactc tcgtgcgcac
 181 ggtgcgggtg gtggtgggtc gggtcgtgcg ggttgctgcc tgggtaaagc ggttcgtggt
 241 gcgaaaggtc accaccaccc gcaccgccc gcggcggggtg cgggtgcggc ggggtggtgcg
 [0399] 301 gaagcggacc tgaaagcgct gaccactct gttctgaaaa aactgaaaga acgtcagctg
 361 gaactgctgc tgcaggcgggt tgaatctcgt ggtggtaccc gtaccgctg cctgctgctg
 421 ccgggtcgtc tggactgccg tctgggtccg ggtgcgccgg cgggtgcgca gccggcgcag
 481 ccgccgtctt ctactctct gccgctgctg ctgtgcaaag tttccgttg gccggacctg
 541 cgtcactctt ctgaagttaa acgtctgtgc tgctgcgaat cttacggtaa aatcaaccgg

- 601 gaactggttt gctgcaacc gcaccacctg tctcgtctgt gcgaactgga atctccgccg
661 ccgccgtact ctcgttacc gctggacttc ctgaaaccga ccgcggactg cccggacgcg
721 gttccgtctt ctgcggaaac cgggtgtacc aactacctgg cgccgggtgg tctgtctgac
781 tctcagctgc tgctggaacc gggtgaccgt tctcactggg gcgttggtgc gtactgggaa
841 gaaaaaacc gtgttggtcg tctgtactgc gttcaggaac cgtctctgga catcttctac
901 gacctgccgc agggtaacgg tttctgcctg ggtcagctga actctgacaa caaatctcag
[0400] 961 ctgggtcaga aagttcgttc taaaatcggg tgcggtatcc agctgaccgc tgaagttgac
1021 ggtgtttggg ttacaaccg ttcttcttac ccgatcttca tcaaactctgc gaccctggac
1081 aaccgggact ctcgtaccct gctgggtcac aaagttttcc cgggtttctc tatcaaagcg
1141 ttcgactacg aaaaagcgta ctctctgcag cgtccgaacg accacgaatt catgcagcag
1201 ccgtggaccg gtttcaccgt tcagatctct ttcgttaaag gttgggggtca gtgctacacc
1261 cgtcagttca tctcttcttg cccgtgctgg ctggaagtta tcttcaactc tcgtggtaaa
1321 ccgateccga acccgctgct gggtctggac tctacctgag tcgac (SEQ ID NO: 36)
[0401] 与实施例4中所描述的Tat-Smad7M7216L-V5的序列比较(比对以出现的顺序分别公开了SEQ ID NO 36和30):
[0402] 比对:针对参考分子的全局DNA比对
[0403] 参数:记分矩阵:线性(错配2,开放空位4,延伸空位1)
[0404] 参考分子:Tat-Smad7-216L-V5-优化子,区1-1365
[0405] 比对的序列数:2
[0406] 设置:相似显著性截止值: $\geq 60\%$
[0407] 匹配百分比的概述;
[0408] 参考:Tat-Smad7-216L-V5-优化子1-1365(1365bp) --
[0409] 序列2:Tat-smad7Ser-His优化-682突变体 1-1392(1392bp) 79%

Tat-Smad7-21	1	ggatccggtcgtaaaaaacgtcgtcagcgtcgtcgtggtttccgtaccaaaccgttctgcg
TatSmad7 Ser	1	ggatccggccgtaaaaaacgcggtcaacgcgcgcgtggtttccgtacgaaacgcagcgc
Tat-Smad7-21	61	ctgggttcgtcgctctgtggcggttctcgtgcgcgggtgggtgaagacgaagaagaaggtgcg
TatSmad7 Ser	61	ctgggtccgtcgccctgtggcgccagccgtgctccgggtggtgaagatgaagaagaaggtgct
Tat-Smad7-21	121	ggtaggtggtggtggtggtggtgaactgcgtggtgaaggtgcgaccgactctcgtgcgcac
TatSmad7 Ser	121	ggcggcggtaggggtggcggtgaactgcgtggcgagggtgcaaccgatagccgtgcacat
Tat-Smad7-21	181	ggtagcgggtggtggtggtcgggtcggtgcgggttgcgtgctgggtaaagcggttcgtggt
TatSmad7 Ser	181	ggtagcaggcggtaggggtccgggtcggtgctgggtgctgctgggtaaagctgtgcgcggc
Tat-Smad7-21	241	gcgaaaggtcaccaccacccgcacccgcggcgggcggtgcgggtgcggcggggtggtgcg
TatSmad7 Ser	241	gcgaaaggtcatcatcatccgcacccgcggcagggtgcagggtgcagctggcggtgcg
Tat-Smad7-21	301	gaagcggacctgaaagcgtgacccactctgttctgaaaaaactgaaagaacgtcagctg
TatSmad7 Ser	301	gaagcggatctgaaagccctgacccatagcgtcctgaaaaaactgaaagaacgtcagctg
Tat-Smad7-21	361	gaactgctgctgcaggcggttgaatctcgtggtggtaccgcgtaccgcgtgctgctgctg
TatSmad7 Ser	361	gagctgctgctgcaggcagtagaaagccgtggcggtaccgcgtaccgcgttctgctgctgctg
Tat-Smad7-21	421	ccgggtcgctctggactgcgcgtctgggtccgggtgcgcggcggtgcgcagccggcgag
TatSmad7 Ser	421	ccgggtcgctctggattgcgcgtctgggtccgggtgcaccggctggtgcgcagccggcgaaa
Tat-Smad7-21	481	ccgcgcgtcttcttactctctgcgcgtgctgctggtgcaaaagtttccggttgccggacctg
TatSmad7 Ser	481	ccgcgcgagcagctacagccctgcgcgtgctgctggtgtaaaagtgttccgttgccggacctg
Tat-Smad7-21	541	cgtcactcttctgaagttaaacgtctgtgctgctgcgaatcttacgggtaaaatcaaccg
TatSmad7 Ser	541	cgccatagcagcgaagttaaacgcctgtgctggtgcgagagctatggcaaaatcaaccg
Tat-Smad7-21	601	gaactggtttgctgcgaaccgcacccactgtctcgtctgtgcgaactggaatctccgcg
TatSmad7 Ser	601	gaactggtttgttgcaatccgcacatctcgtgagccgtctgtgtgaactggagagccgcg
Tat-Smad7-21	661	ccgcgcgtactctcgttacccgcgtggacttccgtgaaaccgaccgcggactgcccggacgcg
TatSmad7 Ser	661	ccgcgcgtatagccgttacccgcgtggatttccgtgaaaccgactgcagattgccggacgcga
Tat-Smad7-21	721	gttcgcgtcttctgcggaaaccggtggtaccaaactacgtggcgccgggtggtctgtctgac
TatSmad7 Ser	721	gtcccagcagcgcgtgagaccggcgccaccaaactatctggcacccggcggtctgagcgat
Tat-Smad7-21	781	tctcagctgctgctggaaccgggtgaccgttctcactgggtgctggttcggtactgggaa
TatSmad7 Ser	781	agccagctgctgctggaaccggcgaccgttagccattgggtgtggttgcctattgggaa
Tat-Smad7-21	841	gaaaaaacccggtgttggtcgtctgtactggttcaggaaaccgtctctggacatcttctac
TatSmad7 Ser	841	gagaaaaacgcgtgtcgggtgcctgtactgcgtacaggaaccgagcctggatatcttctac
Tat-Smad7-21	901	gacctgcgcgagggtaacgggtttctgctgggtcagctgaactctgacacaaaatctcag
TatSmad7 Ser	901	gacctgcgcgagggtcaatgggtttctgctgggtcgaactgaacagcgataataaaagccag
Tat-Smad7-21	961	ctgggttcagaaaagtccgttctaaaatcgggtgcggtatccagctgaccgcgtgaagttgac
TatSmad7 Ser	961	ctgggtgcaaaaagttccgcagcaaaattggctgcgggtatccagctgaccgcgtgaagttgac
Tat-Smad7-21	1021	ggtggtttgggtttacaaccgttctcttaccgatcttcatcaaatctgcgaccctggac
TatSmad7 Ser	1021	ggtggtctgggtatatataaccgcagcagctaccgatcttcatcaaaagcgcaaccctggat
Tat-Smad7-21	1081	aaaccggactctcgtaccctgctgggttcacaaaagttttccgggtttctctatcaaaagcg
TatSmad7 Ser	1081	aatccggacagccgtacgctgctgggtccataaaagtatttccgggttcagcatcaaaagcg
Tat-Smad7-21	1141	ttcgactacgaaaaagcgtactctctgcagcgtccgaacgaccacgaattcatgcagcag
TatSmad7 Ser	1141	ttcgattacgagaaaagcctacagcctgcagcgcccgaacgaccatgaattcatgcagcaa
Tat-Smad7-21	1201	ccgtggaccgggtttcacggttcagatctcttccgttaaagggtgggggtcagtgctacacc
TatSmad7 Ser	1201	ccgtggaccgggttttactgtgcagatcagcttcgttaaagggtgggggtcaatgctacacc
Tat-Smad7-21	1261	cgtagttcatctcttcttgcccggtgctggctggaagtatatcttcaactctcgt-----
TatSmad7 Ser	1261	cgtagtttatcagcagctgtccgtgctggctggaagtgattttcaatagccgcaagggc
Tat-Smad7-21	1315	-----ggtaaacgatcccgaaaccgctgctgggtctggactct
TatSmad7 Ser	1321	gagctcaatagcaagcttgaaggtgaagcctatccctaaccctctcctcggtctcgatagc
Tat-Smad7-21	1354	acctgagtcgac
TatSmad7 Ser	1381	acgtgagtcgac

[0411] 通过将哺乳动物表达载体(例如,pCMV-6-Entry)转染至HEK293或CHO细胞中来使用下面的序列产生蛋白质。

[0412] 在pCMV6-Entry质粒中克隆的优化的GST-Tat-Smad7-myc-flag核苷酸序列:

1 gacgttgat acgactccta tagggcggcc gggaattcgt cgactggatc cgtaccgag
61 gagatctgcc gccgcgatcg ccatgtcccc cattctgggc tactggaaga ttaagggcct
121 ggtgcagcct actagactgc tgctggaata cctggaggaa aaatatgaag agcatctgta
181 tgaaagagac gagggggata aatggaggaa caagaaattc gaactgggac tggagtttcc
241 taatctgcca tactatattg acggcgatgt gaagctgact cagtctatgg ctatcattag
301 atacatcgca gacaaacaca acatgctggg cgggtgtcct aaggaaaggg cagagattag
361 tatgctggag ggagccgtgc tggatattag atacggcgtc tcacgcatcg cctatagcaa
421 agacttcgaa accctgaagg tggattttct gagcaaactg cctgaaatgc tgaagatgtt
481 cgaggacaga ctgtgccaca aaacctacct gaatggcgac catgtcacac acccagattt
541 tatgctgtac gacgccctgg atgtggctct gtatatggac cccatgtgtc tggatgcttt
601 ccctaagctg gtgtgcttta agaaaaggat cgaggcaatt cccagatcg ataagtacct
[0413] 661 gaaaagctcc aagtatatcg ctggcctct ccagggtgg caggcaacat tcggaggcgg
721 ggaccatccc cctaaaagcg acctggaggt gctgtttcag ggaccactgg gcagcggccg
781 gaagaagcgg cggcagaggc gccgaagtag gttccgcact aagcggtcag cactggtgag
841 gagactgtgg cgatctcggg ctctggagg agaggacgag gaagaggag caggcggcgg
901 cggaggagga ggagaactgc gcggggaggg agctacagat agccgagccc acggagctgg
961 aggaggagga ccagggcgag ccggatgtg tctgggcaaa gcagtgagag gcgccaaggg
1021 gcaccatcac ccccatccac ccgccgtgg cgcaggagca gccggcggag ctgaggcaga
1081 cctgaaagcc ctgactcaca gtgtgtgaa gaaactgaag gaaagacagc tggagctgct
1141 gctccaggca gtcgaatcac gcggaggcac ccgaacagct tgtctgtgc tgcccggccg
1201 gctggactgc cggctgggac ccggcgcccc tgctggggca cagccagccc agcctccatc
1261 tagttatagc ctgcccctgc tgctgtgtaa ggtgttcga tggcctgatc tgcggcattc
1321 aagcgaagtc aaaaggctgt gctgttgca gtctacggc aagattaacc cagaactggt

1381 gtgttgcaat ccccatcacc tgtctcgact gtgtgaactg gagtcccccc ctccacccta
1441 ctctaggtat cctatggact ttctgaagcc aaccgctgac tgcccagatg cagtgcctc
1501 ctctgccgag actgggggaa ccaactacct ggctcctggc ggactgagcg actcccagct
1561 gctgctggaa ccaggggatc gcagccactg gtgtgtggtc gcctactggg aagagaagac
1621 aagagtggga aggctgtatt gcgtccagga gccttcctg gacatcttct acgatctgcc
1681 acaggggaat ggattttgtc tgggccagct gaactctgac aataagagtc agctgggtgca
[0414] 1741 gaaagtcgg agcaagattg gctgcggcat ccagctgacc agggaggtgg acggcgtgtg
1801 ggtctacaac cgcagttcat atccaatctt catcaagagc gccactctgg acaatcccga
1861 ttcccgacc ctgctggtgc ataaggtctt ccccggttc agcatcaagg ccttcgacta
1921 cgagaaggct tatagtctcc agcggcccaa cgatcacgag ttcatgcagc agccttggac
1981 aggcttact gtgcagatca gttcgtcaa gggatggggc cagtgtaca caaggcagtt
2041 catctcaagc tgtccctgtt ggctggaagt catttcaat agtaggacgc gtac**gcggcc**
2101 **gctcgagcag** aaactcatct cagaagagga tctggcagca aatgatatcc tggattacaa
2161 ggatgacgac gataaggttt aa (SEQ ID NO: 87)

[0415] 使用来自GenScript的OPTIMUMGENE™程序对整个序列(直至NotI位点)针对哺乳动物表达进行密码子优化。

[0416] 1790-1792:使GAT变成GAC(保留Asp)以避免替代性ORF(突出显示)

[0417] 1-82:在pCMV6-Entry中的克隆位点+kozac序列

[0418] 83-777:GST+精确酶位点;776-805:Tat;

[0419] 806-2086:全长人类Smad7

[0420] 2087-2107:MluI、NotI、XhoI

[0421] 2108-2182:来自pCMV6-Entry的myc-flag

[0422] 粗体下划线:独特的NotI位点

[0423] 上述核苷酸序列与原生人类Smad7具有79%的同源性。

[0424] Mo1 1人类Smad7mRNA(1-3103) Mo1 2优化的GST-Tat-Smad7-myc-flag(1-2182)

[0425] 比对的序列数:2

[0426] 设置:由Score列出的分选结果

人类 Smad7	500	tcaccaccatccccacccgccagccgcggcgccggcgccggcgccggggcg
优化的 GS	1027c.....--.....---.t....a..a..a....c..a.
人类 Smad7	550	ccgaggcggatctgaaggcgctcacgcactcgggtgctcaagaaactgaag
优化的 GS	1071	.t....a..c....a..c..g..t...agt....g.....
人类 Smad7	600	gagcggcagctggagctgctgctccaggccgtggagtcccgcgccgggac
优化的 GS	1121	..aa.a.....a..c..a..a....a..c..
人类 Smad7	650	gcgcaccgcgtgcctctgctgcccggccgcctggactgcaggctgggcc
优化的 GS	1171	c..a..a..t..t..g.....g.....c.....a.
[0427] 人类 Smad7	700	cggggggcgcccgcggcgcgccagcctgcgcagccgccctcgtctactcg
优化的 GS	1221	.c..c..c..t..t..g..a....a..c.....t..a..tagt..tagc
人类 Smad7	750	ctccccctctgctgtgcaaagtgttcaggtggccggatctcaggcattc
优化的 GS	1271	..g.....g.....t..g.....c..a....t....gc.....
人类 Smad7	800	ctcggaagtcaagaggctgtgttgctgtgaatcttacgggaagatcaacc
优化的 GS	1321	aagc.....a.....c..t..c..g..c.....c....t....
人类 Smad7	850	ccgagctggtgtgctgcaacccccatcaccttagccgactctgcgaacta
优化的 GS	1371	.a..a.....t....t.....gtct....g..t....g
人类 Smad7	900	gagtctccccccctccttactccagatacccgatggattttctcaaacc

优化的 GS	1421c....t..a..c....t..g..t....c....g..g..
人类 Smad7	950	aactgcagactgtccagatgctgtgccttcctccgctgaaacagggggaa
优化的 GS	1471	...c..t....c.....a....c....t..c..g..t.....
人类 Smad7	1000	cgaattatctggcccctgggggctttcagattcccaacttcttctggag
优化的 GS	1521	.c..c..c....t....c..a..gagc..c....g..g..g.....a
人类 Smad7	1050	cctggggatcggtcacactgggtgcgtgggtggcatactgggaggagaagac
优化的 GS	1571	..a.....cagc.....t....c..c.....a.....
人类 Smad7	1100	gagagtggggaggctctactgtgtccaggagccctctctggatatcttct
优化的 GS	1621	a.....a....g..t..c.....t..c....c.....
[0428] 人类 Smad7	1150	atgatctacctcaggggaatggcttttgccctcgacagctcaattcggac
优化的 GS	1671	.c....g..a.....a....t..g..c....g..c..t...
人类 Smad7	1200	aacaagagtcagctgggtgcagaaggtgcggagcaaaatcggtgcggcat
优化的 GS	1721	..t.....a..c.....g..t.....
人类 Smad7	1250	ccagctgacgcgggaggtggatgggtgtgtgggtgtacaaccgcagcagtt
优化的 GS	1771ca.....c..c.....c.....ttca.
人类 Smad7	1300	accccatcttcatcaagtccgccacactggacaacccggactccaggacg
优化的 GS	1821	.t..a.....ag.....t.....t..c..t..c..c..c
人类 Smad7	1350	ctgttggtacacaaggtgttccccggtttctccatcaaggctttcgacta
优化的 GS	1871	...c....g..t....c.....c...ag.....c.....

人类 Smad7 1400 cgagaaggcgtacagcctgcagcgcccaatgaccacgagtttatgcagc
优化的 GS 1921t..t..c.....c..t.....c.....

人类 Smad7 1450 agccgtggacgggctttaccgtgcagatcagctttgtgaagggtgggt
优化的 GS 1971t....a....c..t.....c..c....a....c

[0429]

人类 Smad7 1500 cagtgtacacccgccagttcatcagcagctgcccgtgctgtagaggt
优化的 GS 2021aa.g.....tca.....t..c..t....g..a..

人类 Smad7 1550 catcttcaa (SEQ ID NO: 88)
优化的 GS 2071 ...t..... (SEQ ID NO: 89)

[0430] 来自上述核苷酸序列的优化的GST-Tat-Smad7-myc-flag蛋白质序列:

1 mspilgywki kglvqprrll leyleekyee hlyerdegdk wrnkkfelgl efpnlpyiid
61 gdvkltsma iiriyadkhn mlggcpkera eismlegavl dirygvria yskdfetlkv
121 dflsklpeml kmfedrlchk tylnghvth pdfmlydald vvlymdpmcl dafpklvckf
181 krieaipqid kylksskyia wplqgwqatf gggdhppksd levlfqgplg sgrkkrrqrr
241 rsrfrtkrsa lvrrlwrsra pggedeeega gggggggelr gegatdsrah gaggggpgra
301 gcclgkavrg akghhhphpp aagagaagga eadlkalths vlkkikerql ellqavesr
[0431] 361 ggtrtacill pgrldcrlgp gapagaqpaq ppssyslpil lckvfrwpdl rhssevkrlc
421 ccesygkinp elvccnphhl srlcelespp ppysrypmdf lkptadcpda vpssaetggt
481 nylapgglsd sqlllepgr shwcvvaywe ektrvgrylc vqepsldify dlpqngngfcl
541 gqlnsdnksq lvqkvrskig cgiqltrevd gvwwynrasy pifksatld npdsrtllvh
601 kvfpgfsika fdyekayslq rpndhefmqq pwtgftvqis fvkkgwgqcyt rqfisscpw
661 levifnsrtr trpleqklis eedlaandil dykdddkv (SEQ ID NO: 90)

[0432] 1-229:GST+精确位点

[0433] 232-668Tat-Smad7

[0434] 671-699:限制位点+myc-Flag

[0435] 在pCMV6-Entry质粒中克隆的优化的GST-Tat-N-Smad7-myc-flag核苷酸序列:

1 gaattcgtcg actggatccg gtaccgagga gatctgccgc cgcgatcgcc atgtcccca
61 ttctgggcta ctggaagatt aagggcctgg tgcagcctac tagactgctg ctggaatacc
121 tggaggaaaa atatgaagag catctgtatg aaagagacga gggggataaa tggaggaaca
181 agaaattcga actgggactg gagtttcta atctgccata ctatattgac ggcgatgtga
241 agctgactca gtctatggct atcattagat acatcgaga caaacacaac atgtggggcg
301 ggtgtcctaa ggaaaggga gagattagta tgctggaggg agccgtgctg gatattagat
361 acggcgtctc acgcatcgcc tatagcaaag acttcgaaac cctgaagggtg gattttctga
421 gcaaactgcc tgaaatgctg aagatgttcg aggacagact gtgccacaaa acctacctga
481 atggcgacca tgtcacacac ccagatttta tgctgtacga cgccctggat gtggtcctgt
541 atatggacc catgtgtctg gatgctttcc ctaagctggt gtgctttaag aaaaggatcg
601 aggcaattcc ccagatcgat aagtacctga aaagctcaa gtatatcgct tggcctctcc
661 agggctggca ggcaacattc ggaggcgggg accatcccc taaaagcgac ctggaggtgc
[0436] 721 tgtttcagg accactgggc agcggccgga agaagcggcg gcagaggcg cgaagtaggt
781 tccgcactaa gcggtcagca ctggtgcgga gactgtggcg atctcgggct cctggaggag
841 aggacgagga agaggagca ggcggcggcg gaggaggagg agaactgcgc ggggagggag
901 ctacagatag ccgagcccac ggagctggag gaggaggacc agggcgagcc ggatgtctgc
961 tgggcaaagc agtgagaggc gccaaggggc accatcacc ccatccacc gccgctggcg
1021 caggagcagc cggcggagct gaggcagacc tgaaagccct gactcacagt gtgtgaaga
1081 aactgaagga aagacagctg gagctgtgc tccaggcagt cgaatcacgc ggaggcacc
1141 gaacagcttg tctgtctg cccggccggc tggactgccg gctgggacc ggcgccctg
1201 ctggggcaca gccagcccag cctccatcta gttatagcct gcccctgctg ctgtgtaagg
1261 tgttccgatg gcctgatctg cggcattcaa gcgaagtcaa aaggctgtgc tgttgcagt
1321 cctacggcaa gattaacca gaactggtgt gttgcaatcc ccatcacctg tctcgactgt
1381 gtgaactgga gtccccct ccaccctact ctaggtatcc tatggacttt ctgaagccaa
1441 ccgctgactg ccagatgca gtgcctctct ctgccgagac tgggggaacc aactacctg
1501 ctctggcgg actgagcgac tccagctgc tgctggaacc aggggatcgg **cgccgctc**
1561 agcagaaact catctcagaa gaggatctgg cagcaaatga tatcctggat tacaaggatg
[0437] 1621 acgacgataa ggtttaa (SEQ ID NO: 91)

[0438] 使用来自GenScript的OPTIMUMGENETM程序对整个序列(直至NotI位点)针对哺乳动物表达进行密码子优化。

- [0439] 1-50: 在pCMV6-Entry中的克隆位点+kozac序列
- [0440] 51-743: GST+精确酶位点
- [0441] 744-773-: Tat
- [0442] 774-1549-: 编码N端1-258aa的人类Smad7
- [0443] 1550-1563: NotI、XhoI, 粗体下划线: 独特的NotI位点
- [0444] 1564-1637: 来自pCMV6-Entry的myc-flag
- [0445] 来自上述核苷酸序列的优化的GST-Tat-N-Smad7-myc-flag蛋白质序列:
- ```
1 mspilgywki kglvqprrll leyleekyee hlyerdegdk wrnkkfelgl efpnlpyyid
61 gdvkltsma iiriyadhkn mlggcpkera eismlegavl dirygvrsia yskdfetlkv
121 dflsklpeml kmfedrlchk tylnghdvth pdfmlydald vvlymdpmcl dafpklvcfk
181 krieaipqid kylksskyia wplqgwqatf gggdhppksd levlfqgplg sgrkrrqrr
```
- [0446] 241 rsrfrtkrsa lvrrlwrsra pggedeega gggggggelr gegatdsrah gaggggpgra
- ```
301  gcclgkavrg akghhhphpp aagagaagga eadlkalths vlklkerql ellqavesr
361  ggtrtaclll pgrldcrlgp gapagaqpaq ppssyslpil lckvfrwpdl rhssevkrlc
421  ccesygkinp elvccnphhl srlcelespp ppysrypmdf lkptadcpda vpssaetggg
481  nylapgglsd sqllepgdr rpleqklise edlaandild ykdddkv (SEQ ID NO: 92)
```
- [0447] 1-241: GST-Tat
- [0448] 244-500: N-Smad72-258
- [0449] 501-528: 限制位点+myc-Flag
- [0450] 在pCMV6-Entry质粒中克隆的优化的GST-Tat-C-Smad7-myc-flag核苷酸序列:
- ```
1 gaattcgctg actggatccg gtaccgagga gatctgccgc cgcgatcgcc atgtcccca
[0451] 61 ttctgggcta ctggaagatt aagggcctgg tgcagcctac tagactgctg ctggaatacc
```

- 121 tggaggaaaa atatgaagag catctgtatg aaagagacga gggggataaa tggaggaaca  
181 agaaattcga actgggactg gagtttcta atctgccata ctatattgac ggcgatgtga  
241 agctgactca gtctatggct atcattagat acatcgaga caaacacaac atgctgggag  
301 ggtgtcctaa ggaaagggca gagattagta tgctggaggg agccgtgctg gatattagat  
361 acggcgtctc acgcatcgcc tatagcaaag acttcgaaac cctgaagggtg gattttctga  
421 gcaaaactgcc tgaaatgctg aagatgttcg aggacagact gtgccacaaa acctacctga  
481 atggcgacca tgtcacacac ccagatttta tgctgtacga cgccctggat gtggctctgt  
541 atatggaccc catgtgtctg gatgctttcc ctaagctggt gtgctttaag aaaaggatcg  
601 aggcaattcc ccagatcgat aagtacctga aaagctcaa gtatatcgct tggcctctcc  
661 agggctggca ggcaacattc ggaggcgggg accatcccc taaaagcgac ctggagggtg  
[0452] 721 tgtttcaggg accactgggc agcggccgga agaagcggcg gcagaggcgc agccactggt  
781 gtgtggctgc ctactgggaa gagaagacaa gagtgggaag gctgtattgc gtccaggagc  
841 cttccctgga catcttctac gatctgccac aggggaatgg atttgtctg ggccagctga  
901 actctgacaa taagagtcag ctggtgcaga aagtcggag caagattggc tgcggcatcc  
961 agctgaccag ggagggtggac ggcgtgtggg tctacaaccg cagttcatat ccaattctca  
1021 tcaagagcgc cactctggac aatcccgatt cccgcaccct gctggtgcat aaggtcttcc  
1081 ccggcttcag catcaaggcc ttcgactacg agaaggctta tagtctccag cggcccaacg  
1141 atcacgagtt catgcagcag ccttgacag gcttactgt gcagatcagc ttcgtcaagg  
1201 gatggggcca gtgctacaca aggcagttca tctcaagctg tccctgttgg ctggaagtca  
1261 tttcaatag taggacgcgt acgcgccgc tcgagcagaa actcatctca gaagaggatc  
1321 tggcagcaaa tgatctctg gattacaagg atgacgacga taaggtttaa (SEQ ID NO: 93)  
[0453] 使用来自GenScript的OPTIMUMGENETM程序对整个序列(直至NotI位点)针对哺乳动物表达进行密码子优化。  
[0454] 1-50:在pCMV6-Entry中的克隆位点+kozac序列  
[0455] 51-743:GST+精确酶位点  
[0456] 744-773:Tat  
[0457] 774-1274:c端Smad7  
[0458] 1275:Mlu  
[0459] 1283:NotI  
[0460] 1275-1370:mluI、NotI、XhoI、myc+flag  
[0461] 来自上述核苷酸序列的优化的GST-Tat-C-Smad7-myc-flag蛋白质序列:

- 1 mspilgywki kglvqprrll leyleekyee hlyerdegdk wrnkklfelgl efpnlpyyid  
 61 gdvkltsma iiriyadkhn mlggcpkera eismlegavl dirygsria yskdfetlkv  
 121 dflsklpeml kmfedrlchk tylngdhvth pdfmlydald vvlymdpmcl dafpklvckf  
 181 krieaipqid kylksskyia wplqgwqatf gggdhppksd levlfqgplg sgrkrrrr  
 [0462] 241 rshwcvvayw eektrvgrly cvqepsldif ydlpqngfc lgqlnsdnks qlvqkvrski  
 301 gcgiqltrev dgvvvynrss ypifiksatl dnpdsrtllv hkvfpghsik afdyekaysl  
 361 qrpndhefmq qpwtgftvqi sfvkgwgqcy trqfisscpc wlevifnsrt rtrpleqkli  
 421 seedlaandi ldykdddkv (SEQ ID NO: 94)
- [0463] 1-241:GST+精确+Tat
- [0464] 242-409:c-Smad7 (259-426aa)
- [0465] 410-440:Myu+Not+myc+flag:限制位点+myc-Flag
- [0466] 基于Smad7结构,预期含有PY域(203-217aa)的连接子(203-258aa)通过阻断TGF- $\beta$ 诱导的炎症或生长抑制将具有治疗作用(图15)。
- [0467] Tat-Smad7-连接子肽(203-258aa)的氨基酸序列是:
- [0468] GRKKRRQRRR-ELESPPPPYSRYPMDFLKPTADCPDAVPSSAETGGTNYLAPGGLSDSQLLLEPGDR  
 (SEQ ID NO:95)
- [0469] 1-10:Tat
- [0470] 在细菌产生中上述肽的密码子优化的核苷酸序列是:
- 1 ggccgtaaaa aacgccgtca acgccgccgt gaactggaga gcccgccgcc gccgtatagc  
 61 cgttaccga tggatttct gaaaccgact gcagattgcc cggacgcagt cccgagcagc  
 [0471] 121 gctgagaccg gcggcaccaa ctatctggca ccggcggtc tgagcgatag ccagctgctg  
 181 ctggaaccgg gcgaccgtgg taagcctatc cctaaccctc tctcgggtct cgattctacg  
 241 tgagtcgac (SEQ ID NO: 96)
- [0472] 1-30:Tat
- [0473] 31-198:Smad7连接子
- [0474] 199-249:V5
- [0475] 在哺乳动物细胞产生中上述肽的密码子优化的核苷酸序列是:
- 1 ggccggaaga agcgcgcgca gagcgccga gaactggagt cccccctcc accctactct  
 61 aggtatccta tggactttct gaagccaacc gctgactgcc cagatgcagt gccctctct  
 [0476] 121 gccgagactg ggggaaccaa ctacctggct cctggcgagc tgagcgactc ccagctgctg  
 181 ctggaaccag gggatcgcg gccgctcag cagaaactca tctcagaaga ggatctggca  
 241 gcaaatgata tcttgatta caaggatgac gacgataagg tttaa (SEQ ID NO: 97)
- [0477] 1-30:哺乳动物密码子优化的Tat

[0478] 31-198:Smad7连接子

[0479] 199-285:NotI、Myc标签、Flag标签

[0480] Tat-Smad7-PY肽 (203-217aa) 的氨基酸序列是:

[0481] GRKKRRQRRR-ELESPPPPYSRYPM (SEQ ID NO:98)

[0482] 1-10:Tat

[0483] 在细菌产生中上述肽的密码子优化的核苷酸序列是:

1 ggccgtaaaa aacgccgtca acgccgccgt gaactggaga gcccgccgcc gccgtatagc

[0484] 61 cgttaccgga tggatggtaa gcctatccct aaccctctcc tcggtctcga ttctacgtga

121 gtcgac (SEQ ID NO: 99)

[0485] 1-30:Tat

[0486] 31-75:Smad7-PY

[0487] 76-126:V5

[0488] 在哺乳动物细胞产生中上述肽的密码子优化的核苷酸序列是:

1 ggccggaaga agcggcgcca gagcgccga gaactggagt cccccctcc accctactct

[0489] 61 aggtatccta tggaccggcc gctcgagcag aaactcatct cagaagagga tctggcagca

121 aatgatatcc tggattacaa ggatgacgac gataaggttt aa (SEQ ID NO: 100)

[0490] 1-30:哺乳动物密码子优化的Tat

[0491] 31-75:Smad7PY

[0492] 76-162:NotI、Myc标签、Flag标签

[0493] 本发明技术的前述论述已经出于说明和描述的目的而呈现。前述内容并不意图将本发明技术限于本文所公开的形式。尽管本发明技术的描述已经包括一个或多个实施方案的描述以及某些变更和修改,但在理解本公开之后其它变更和修改处于本发明技术的范围内,例如,可以在本领域的技术人员的技能和知识范围内。意图获得在所允许的程度上包括替代性实施方案的权利,包括所要求的那些的替代的、可互换的和/或等效的结构、功能、范围或步骤,无论这类替代的、可互换的和/或等效的结构、功能、范围或步骤是否在本文中公开,并且不意图公开地奉献任何可取得专利的主题。

## 序列表

|        |                                                                   |    |
|--------|-------------------------------------------------------------------|----|
|        | <110> THE REGENTS OF THE UNIVERSITY OF COLORADO, A BODY CORPORATE |    |
|        | <120> PTD-SMAD7治疗剂                                                |    |
|        | <130> 089491-0303                                                 |    |
|        | <140> PCT/US2014/022052                                           |    |
|        | <141> 2014-03-07                                                  |    |
|        | <150> 61/775,252                                                  |    |
|        | <151> 2013-03-08                                                  |    |
|        | <160> 101                                                         |    |
|        | <170> PatentIn版本3.5                                               |    |
|        | <210> 1                                                           |    |
|        | <211> 30                                                          |    |
|        | <212> DNA                                                         |    |
|        | <213> 人类免疫缺陷病毒                                                    |    |
|        | <220>                                                             |    |
|        | <221> CDS                                                         |    |
|        | <222> (1).. (30)                                                  |    |
|        | <400> 1                                                           |    |
|        | ggc cgt aaa aaa cgc cgt caa cgc cgc cgt                           | 30 |
|        | Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg                           |    |
|        | 1 5 10                                                            |    |
| [0001] | <210> 2                                                           |    |
|        | <211> 10                                                          |    |
|        | <212> PRT                                                         |    |
|        | <213> 人类免疫缺陷病毒                                                    |    |
|        | <400> 2                                                           |    |
|        | Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg                           |    |
|        | 1 5 10                                                            |    |
|        | <210> 3                                                           |    |
|        | <211> 33                                                          |    |
|        | <212> DNA                                                         |    |
|        | <213> 人类免疫缺陷病毒                                                    |    |
|        | <400> 3                                                           |    |
|        | tatggccgta aaaaacgccg tcaacgccgc cgt                              | 33 |
|        | <210> 4                                                           |    |
|        | <211> 11                                                          |    |
|        | <212> PRT                                                         |    |
|        | <213> 人类免疫缺陷病毒                                                    |    |
|        | <400> 4                                                           |    |
|        | Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg                       |    |
|        | 1 5 10                                                            |    |
|        | <210> 5                                                           |    |
|        | <211> 21                                                          |    |
|        | <212> DNA                                                         |    |
|        | <213> 人类免疫缺陷病毒                                                    |    |
|        | <400> 5                                                           |    |
|        | ggccgtaaaa aacgccgtca a                                           | 21 |

|        |                                                                 |     |
|--------|-----------------------------------------------------------------|-----|
| [0002] | <210> 6                                                         |     |
|        | <211> 7                                                         |     |
|        | <212> PRT                                                       |     |
|        | <213> 人类免疫缺陷病毒                                                  |     |
|        | <400> 6                                                         |     |
|        | Gly Arg Lys Lys Arg Arg Gln                                     |     |
|        | 1 5                                                             |     |
|        | <210> 7                                                         |     |
|        | <211> 30                                                        |     |
|        | <212> DNA                                                       |     |
|        | <213> 人类免疫缺陷病毒                                                  |     |
|        | <400> 7                                                         |     |
|        | ggccgtaaaa aacgccgtca acgccgccgt                                | 30  |
|        | <210> 8                                                         |     |
|        | <211> 30                                                        |     |
|        | <212> DNA                                                       |     |
|        | <213> 人类免疫缺陷病毒                                                  |     |
|        | <400> 8                                                         |     |
|        | ggccgtaaaa aacgccgtca acgccgccgt                                | 30  |
|        | <210> 9                                                         |     |
|        | <211> 1497                                                      |     |
|        | <212> DNA                                                       |     |
|        | <213> 人工序列                                                      |     |
|        | <220>                                                           |     |
|        | <223> 人工序列的描述：合成多核苷酸                                            |     |
|        | <220>                                                           |     |
|        | <221> CDS                                                       |     |
|        | <222> (58).. (1461)                                             |     |
|        | <400> 9                                                         |     |
|        | ttccctcta gaaataattt tgtttaactt taagaaggaa ttcaggagcc cttcacc   | 57  |
|        | atg cgt aaa aaa cgc cgt caa cgc cgc cgt ggt ttc cgt acg aaa cgc | 105 |
|        | Met Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Phe Arg Thr Lys Arg |     |
|        | 1 5 10 15                                                       |     |
|        | tgc gcc ctg gtc cgt cgc ctg tgg cgc tcc cgt gct ccg ggt ggt gaa | 153 |
|        | Ser Ala Leu Val Arg Arg Leu Trp Arg Ser Arg Ala Pro Gly Gly Glu |     |
|        | 20 25 30                                                        |     |
|        | gat gaa gaa gaa ggt gct ggc ggc ggt ggc ggt ggc ggt gaa ctg cgt | 201 |
|        | Asp Glu Glu Glu Gly Ala Gly Gly Gly Gly Gly Gly Glu Leu Arg     |     |
|        | 35 40 45                                                        |     |
|        | ggc gag ggt gca acc gat agt cgt gca cac ggt gca ggc ggt ggc ggt | 249 |
|        | Gly Glu Gly Ala Thr Asp Ser Arg Ala His Gly Ala Gly Gly Gly Gly |     |
|        | 50 55 60                                                        |     |
|        | ccg ggt cgt gct ggt tgc tgt ctg ggt aaa gct gtg cgc ggc gcg aaa | 297 |
|        | Pro Gly Arg Ala Gly Cys Cys Leu Gly Lys Ala Val Arg Gly Ala Lys |     |
|        | 65 70 75 80                                                     |     |
|        | ggt cat cac cat ccg cac ccg ccg gca gca ggt gca ggt gca gct ggc | 345 |
|        | Gly His His His Pro His Pro Pro Ala Ala Gly Ala Gly Ala Ala Gly |     |
|        | 85 90 95                                                        |     |
|        | ggt gcg gaa gcc gat ctg aaa gcc ctg acc cat agt gtc ctg aaa aaa | 393 |
|        | Gly Ala Glu Ala Asp Leu Lys Ala Leu Thr His Ser Val Leu Lys Lys |     |

|        | 100                                                                                                                                | 105                                                                        | 110 |      |
|--------|------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------|-----|------|
|        | ctg aaa gaa cgt cag ctg gag<br>Leu Lys Glu Arg Gln Leu Glu                                                                         | ctg ctg ctg caa gca gta gaa tcc cgt<br>Leu Leu Leu Gln Ala Val Glu Ser Arg |     | 441  |
|        | 115                                                                                                                                | 120                                                                        | 125 |      |
|        | ggc ggt acc cgt acg gct tgt<br>Gly Gly Thr Arg Thr Ala Cys                                                                         | ctg ctg ctg ccg ggt cgt ctg gat tgc<br>Leu Leu Leu Pro Gly Arg Leu Asp Cys |     | 489  |
|        | 130                                                                                                                                | 135                                                                        | 140 |      |
|        | cgt ctg ggt ccg ggt gca ccg gct ggt gcg cag ccg gca caa ccg ccg<br>Arg Leu Gly Pro Gly Ala Pro Ala Gly Ala Gln Pro Ala Gln Pro Pro |                                                                            |     | 537  |
|        | 145                                                                                                                                | 150                                                                        | 155 | 160  |
|        | agc tct tac agc ctg ccg ctg ctg ctg tgt aaa gtg ttt cgt tgg ccg<br>Ser Ser Tyr Ser Leu Pro Leu Leu Leu Cys Lys Val Phe Arg Trp Pro |                                                                            |     | 585  |
|        | 165                                                                                                                                | 170                                                                        | 175 |      |
|        | gac ctg cgc cac agt tcc gaa gtt aaa cgc ctg tgc tgt tgc gag agc<br>Asp Leu Arg His Ser Ser Glu Val Lys Arg Leu Cys Cys Cys Glu Ser |                                                                            |     | 633  |
|        | 180                                                                                                                                | 185                                                                        | 190 |      |
|        | tat ggc aaa att aac ccg gaa ctg gtt tgt tgc aat ccg cac cat ctg<br>Tyr Gly Lys Ile Asn Pro Glu Leu Val Cys Cys Asn Pro His His Leu |                                                                            |     | 681  |
|        | 195                                                                                                                                | 200                                                                        | 205 |      |
|        | tct cgt ctg tgt gaa ctg gag agc ccg ccg ccg ccg tat tct cgt tac<br>Ser Arg Leu Cys Glu Leu Glu Ser Pro Pro Pro Pro Tyr Ser Arg Tyr |                                                                            |     | 729  |
|        | 210                                                                                                                                | 215                                                                        | 220 |      |
|        | ccg atg gat ttc ctg aaa ccg act gca gat tgc ccg gac gca gtc ccg<br>Pro Met Asp Phe Leu Lys Pro Thr Ala Asp Cys Pro Asp Ala Val Pro |                                                                            |     | 777  |
|        | 225                                                                                                                                | 230                                                                        | 235 | 240  |
| [0003] | tca tgc gct gag acc ggc ggc acc aac tat ctg gca ccg ggc ggt ctg<br>Ser Ser Ala Glu Thr Gly Gly Thr Asn Tyr Leu Ala Pro Gly Gly Leu |                                                                            |     | 825  |
|        | 245                                                                                                                                | 250                                                                        | 255 |      |
|        | agt gat tcc cag ctg ctg ctg gaa ccg ggc gac cgt tca cat tgg tgt<br>Ser Asp Ser Gln Leu Leu Leu Glu Pro Gly Asp Arg Ser His Trp Cys |                                                                            |     | 873  |
|        | 260                                                                                                                                | 265                                                                        | 270 |      |
|        | gtg gtt gcc tat tgg gaa gag aaa acg cgt gtc ggt cgc ctg tac tgc<br>Val Val Ala Tyr Trp Glu Glu Lys Thr Arg Val Gly Arg Leu Tyr Cys |                                                                            |     | 921  |
|        | 275                                                                                                                                | 280                                                                        | 285 |      |
|        | gta cag gaa ccg tgc ctg gat atc ttt tat gac ctg ccg cag ggc aat<br>Val Gln Glu Pro Ser Leu Asp Ile Phe Tyr Asp Leu Pro Gln Gly Asn |                                                                            |     | 969  |
|        | 290                                                                                                                                | 295                                                                        | 300 |      |
|        | ggc ttc tgt ctg ggc caa ctg aac tca gat aat aaa tgc cag ctg gtg<br>Gly Phe Cys Leu Gly Gln Leu Asn Ser Asp Asn Lys Ser Gln Leu Val |                                                                            |     | 1017 |
|        | 305                                                                                                                                | 310                                                                        | 315 | 320  |
|        | caa aaa gtt cgc tca aaa att ggc tgc ggt atc cag ctg acc cgt gaa<br>Gln Lys Val Arg Ser Lys Ile Gly Cys Gly Ile Gln Leu Thr Arg Glu |                                                                            |     | 1065 |
|        | 325                                                                                                                                | 330                                                                        | 335 |      |
|        | gtt gac ggt gtc tgg gta tat aac cgc agc tct tac ccg att ttt atc<br>Val Asp Gly Val Trp Val Tyr Asn Arg Ser Ser Tyr Pro Ile Phe Ile |                                                                            |     | 1113 |
|        | 340                                                                                                                                | 345                                                                        | 350 |      |
|        | aaa agt gcc acc ctg gat aat ccg gac tcc cgt acg ctg ctg gtc cac<br>Lys Ser Ala Thr Leu Asp Asn Pro Asp Ser Arg Thr Leu Leu Val His |                                                                            |     | 1161 |
|        | 355                                                                                                                                | 360                                                                        | 365 |      |
|        | aaa gta ttt ccg ggc ttc tca atc aaa gcg ttc gat tac gag aaa gcc<br>Lys Val Phe Pro Gly Phe Ser Ile Lys Ala Phe Asp Tyr Glu Lys Ala |                                                                            |     | 1209 |
|        | 370                                                                                                                                | 375                                                                        | 380 |      |
|        | tac tgc ctg cag cgc ccg aac gac cat gaa ttc atg cag caa ccg tgg                                                                    |                                                                            |     | 1257 |



Tyr Ser Leu Gln Arg Pro Asn Asp His Glu Phe Met Gln Gln Pro Trp  
 385 390 395 400  
 acg ggt ttt act gtg cag atc tct ttc gtt aaa ggc tgg ggt caa tgc 1305  
 Thr Gly Phe Thr Val Gln Ile Ser Phe Val Lys Gly Trp Gly Gln Cys  
 405 410 415  
 tac acc cgt cag ttt atc tcg tcc tgt ccg tgc tgg ctg gaa gtg att 1353  
 Tyr Thr Arg Gln Phe Ile Ser Ser Cys Pro Cys Trp Leu Glu Val Ile  
 420 425 430  
 ttc aat agc cgc aag ggc gag ctc aat tcg aag ctt gaa ggt aag cct 1401  
 Phe Asn Ser Arg Lys Gly Glu Leu Asn Ser Lys Leu Glu Gly Lys Pro  
 435 440 445  
 atc cct aac cct ctc ctc ggt ctc gat tct acg cgt acc ggt cat cat 1449  
 Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr Arg Thr Gly His His  
 450 455 460  
 cac cat cac cat tgagtttgat ccgctgcta acaaagcccg aaagga 1497  
 His His His His  
 465  
 <210> 10  
 <211> 468  
 <212> PRT  
 <213> 人工序列  
 <220>  
 <223> 人工序列的描述：合成多肽  
 <400> 10  
 Met Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Phe Arg Thr Lys Arg  
 1 5 10 15  
 Ser Ala Leu Val Arg Arg Leu Trp Arg Ser Arg Ala Pro Gly Gly Glu  
 20 25 30  
 Asp Glu Glu Glu Gly Ala Gly Gly Gly Gly Gly Gly Glu Leu Arg  
 35 40 45  
 Gly Glu Gly Ala Thr Asp Ser Arg Ala His Gly Ala Gly Gly Gly Gly  
 50 55 60  
 Pro Gly Arg Ala Gly Cys Cys Leu Gly Lys Ala Val Arg Gly Ala Lys  
 65 70 75 80  
 Gly His His His Pro His Pro Pro Ala Ala Gly Ala Gly Ala Ala Gly  
 85 90 95  
 Gly Ala Glu Ala Asp Leu Lys Ala Leu Thr His Ser Val Leu Lys Lys  
 100 105 110  
 Leu Lys Glu Arg Gln Leu Glu Leu Leu Leu Gln Ala Val Glu Ser Arg  
 115 120 125  
 Gly Gly Thr Arg Thr Ala Cys Leu Leu Leu Pro Gly Arg Leu Asp Cys  
 130 135 140  
 Arg Leu Gly Pro Gly Ala Pro Ala Gly Ala Gln Pro Ala Gln Pro Pro  
 145 150 155 160

[0004]

Ser Ser Tyr Ser Leu Pro Leu Leu Leu Cys Lys Val Phe Arg Trp Pro  
 165 170 175  
 Asp Leu Arg His Ser Ser Glu Val Lys Arg Leu Cys Cys Cys Glu Ser  
 180 185 190  
 Tyr Gly Lys Ile Asn Pro Glu Leu Val Cys Cys Asn Pro His His Leu  
 195 200 205  
 Ser Arg Leu Cys Glu Leu Glu Ser Pro Pro Pro Pro Tyr Ser Arg Tyr  
 210 215 220  
 Pro Met Asp Phe Leu Lys Pro Thr Ala Asp Cys Pro Asp Ala Val Pro  
 225 230 235 240  
 Ser Ser Ala Glu Thr Gly Gly Thr Asn Tyr Leu Ala Pro Gly Gly Leu  
 245 250 255  
 Ser Asp Ser Gln Leu Leu Leu Glu Pro Gly Asp Arg Ser His Trp Cys  
 260 265 270  
 Val Val Ala Tyr Trp Glu Glu Lys Thr Arg Val Gly Arg Leu Tyr Cys  
 275 280 285  
 Val Gln Glu Pro Ser Leu Asp Ile Phe Tyr Asp Leu Pro Gln Gly Asn  
 290 295 300  
 Gly Phe Cys Leu Gly Gln Leu Asn Ser Asp Asn Lys Ser Gln Leu Val  
 305 310 315 320  
 Gln Lys Val Arg Ser Lys Ile Gly Cys Gly Ile Gln Leu Thr Arg Glu  
 325 330 335  
 Val Asp Gly Val Trp Val Tyr Asn Arg Ser Ser Tyr Pro Ile Phe Ile  
 340 345 350  
 Lys Ser Ala Thr Leu Asp Asn Pro Asp Ser Arg Thr Leu Leu Val His  
 355 360 365  
 Lys Val Phe Pro Gly Phe Ser Ile Lys Ala Phe Asp Tyr Glu Lys Ala  
 370 375 380  
 Tyr Ser Leu Gln Arg Pro Asn Asp His Glu Phe Met Gln Gln Pro Trp  
 385 390 395 400  
 Thr Gly Phe Thr Val Gln Ile Ser Phe Val Lys Gly Trp Gly Gln Cys  
 405 410 415  
 Tyr Thr Arg Gln Phe Ile Ser Ser Cys Pro Cys Trp Leu Glu Val Ile  
 420 425 430  
 Phe Asn Ser Arg Lys Gly Glu Leu Asn Ser Lys Leu Glu Gly Lys Pro

[0005]

|        |                                                                 |     |         |
|--------|-----------------------------------------------------------------|-----|---------|
|        | 435                                                             | 440 | 445     |
|        | Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr Arg Thr Gly His His |     |         |
|        | 450                                                             | 455 | 460     |
|        | His His His His                                                 |     |         |
|        | 465                                                             |     |         |
|        | <210> 11                                                        |     |         |
|        | <211> 461                                                       |     |         |
|        | <212> PRT                                                       |     |         |
|        | <213> 人工序列                                                      |     |         |
|        | <220>                                                           |     |         |
|        | <223> 人工序列的描述：合成多肽                                              |     |         |
|        | <400> 11                                                        |     |         |
|        | Gly Ser Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Phe Arg Thr |     |         |
|        | 1                                                               | 5   | 10 15   |
|        | Lys Arg Ser Ala Leu Val Arg Arg Leu Trp Arg Ser Arg Ala Pro Gly |     |         |
|        |                                                                 | 20  | 25 30   |
|        | Gly Glu Asp Glu Glu Glu Gly Ala Gly Gly Gly Gly Gly Gly Glu     |     |         |
|        |                                                                 | 35  | 40 45   |
|        | Leu Arg Gly Glu Gly Ala Thr Asp Ser Arg Ala His Gly Ala Gly Gly |     |         |
|        |                                                                 | 50  | 55 60   |
| [0006] | Gly Gly Pro Gly Arg Ala Gly Cys Cys Leu Gly Lys Ala Val Arg Gly |     |         |
|        | 65                                                              | 70  | 75 80   |
|        | Ala Lys Gly His His His Pro His Pro Pro Ala Ala Gly Ala Gly Ala |     |         |
|        |                                                                 | 85  | 90 95   |
|        | Ala Gly Gly Ala Glu Ala Asp Leu Lys Ala Leu Thr His Ser Val Leu |     |         |
|        |                                                                 | 100 | 105 110 |
|        | Lys Lys Leu Lys Glu Arg Gln Leu Glu Leu Leu Leu Gln Ala Val Glu |     |         |
|        |                                                                 | 115 | 120 125 |
|        | Ser Arg Gly Gly Thr Arg Thr Ala Cys Leu Leu Leu Pro Gly Arg Leu |     |         |
|        |                                                                 | 130 | 135 140 |
|        | Asp Cys Arg Leu Gly Pro Gly Ala Pro Ala Gly Ala Gln Pro Ala Gln |     |         |
|        | 145                                                             | 150 | 155 160 |
|        | Pro Pro Ser Ser Tyr Ser Leu Pro Leu Leu Leu Cys Lys Val Phe Arg |     |         |
|        |                                                                 | 165 | 170 175 |
|        | Trp Pro Asp Leu Arg His Ser Ser Glu Val Lys Arg Leu Cys Cys Cys |     |         |
|        |                                                                 | 180 | 185 190 |
|        | Glu Ser Tyr Gly Lys Ile Asn Pro Glu Leu Val Cys Cys Asn Pro His |     |         |
|        |                                                                 | 195 | 200 205 |

His Leu Ser Arg Leu Cys Glu Leu Glu Ser Pro Pro Pro Tyr Ser  
 210 215 220  
 Arg Tyr Pro Met Asp Phe Leu Lys Pro Thr Ala Asp Cys Pro Asp Ala  
 225 230 235 240  
 Val Pro Ser Ser Ala Glu Thr Gly Gly Thr Asn Tyr Leu Ala Pro Gly  
 245 250 255  
 Gly Leu Ser Asp Ser Gln Leu Leu Leu Glu Pro Gly Asp Arg Ser His  
 260 265 270  
 Trp Cys Val Val Ala Tyr Trp Glu Glu Lys Thr Arg Val Gly Arg Leu  
 275 280 285  
 Tyr Cys Val Gln Glu Pro Ser Leu Asp Ile Phe Tyr Asp Leu Pro Gln  
 290 295 300  
 Gly Asn Gly Phe Cys Leu Gly Gln Leu Asn Ser Asp Asn Lys Ser Gln  
 305 310 315 320  
 Leu Val Gln Lys Val Arg Ser Lys Ile Gly Cys Gly Ile Gln Leu Thr  
 325 330 335  
 Arg Glu Val Asp Gly Val Trp Val Tyr Asn Arg Ser Ser Tyr Pro Ile  
 340 345 350  
 Phe Ile Lys Ser Ala Thr Leu Asp Asn Pro Asp Ser Arg Thr Leu Leu  
 355 360 365  
 Val His Lys Val Phe Pro Gly Phe Ser Ile Lys Ala Phe Asp Tyr Glu  
 370 375 380  
 Lys Ala Tyr Ser Leu Gln Arg Pro Asn Asp His Glu Phe Met Gln Gln  
 385 390 395 400  
 Pro Trp Thr Gly Phe Thr Val Gln Ile Ser Phe Val Lys Gly Trp Gly  
 405 410 415  
 Gln Cys Tyr Thr Arg Gln Phe Ile Ser Ser Cys Pro Cys Trp Leu Glu  
 420 425 430  
 Val Ile Phe Asn Ser Arg Lys Gly Glu Leu Asn Ser Lys Leu Glu Gly  
 435 440 445  
 Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr  
 450 455 460

<210> 12  
 <211> 426  
 <212> PRT  
 <213> 智人  
 <400> 12

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

[0008]

Asp Leu Pro Gln Gly Asn Gly Phe Cys Leu Gly Gln Leu Asn Ser Asp  
 290 295 300  
 Asn Lys Ser Gln Leu Val Gln Lys Val Arg Ser Lys Ile Gly Cys Gly  
 305 310 315 320  
 Ile Gln Leu Thr Arg Glu Val Asp Gly Val Trp Val Tyr Asn Arg Ser  
 325 330 335  
 Ser Tyr Pro Ile Phe Ile Lys Ser Ala Thr Leu Asp Asn Pro Asp Ser  
 340 345 350  
 Arg Thr Leu Leu Val His Lys Val Phe Pro Gly Phe Ser Ile Lys Ala  
 355 360 365  
 Phe Asp Tyr Glu Lys Ala Tyr Ser Leu Gln Arg Pro Asn Asp His Glu  
 370 375 380  
 Phe Met Gln Gln Pro Trp Thr Gly Phe Thr Val Gln Ile Ser Phe Val  
 385 390 395 400  
 Lys Gly Trp Gly Gln Cys Tyr Thr Arg Gln Phe Ile Ser Ser Cys Pro  
 405 410 415  
 Cys Trp Leu Glu Val Ile Phe Asn Ser Arg  
 420 425

[0009]

<210> 13  
 <211> 20  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成引物

<400> 13  
 tggaattcct ggtctggttt

20

<210> 14  
 <211> 20  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成引物

<400> 14  
 gccaaagctgc tcttcagta

20

<210> 15  
 <211> 21  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成引物

<400> 15  
 tctcaggggg ccaaggtgt t

21

|        |                                 |    |
|--------|---------------------------------|----|
| [0010] | <210> 16                        |    |
|        | <211> 22                        |    |
|        | <212> DNA                       |    |
|        | <213> 人工序列                      |    |
|        | <220>                           |    |
|        | <223> 人工序列的描述: 合成引物             |    |
|        | <400> 16                        |    |
|        | tcccagcacc tgaatcacat gg        | 22 |
|        | <210> 17                        |    |
|        | <211> 10                        |    |
|        | <212> DNA                       |    |
|        | <213> 未知                        |    |
|        | <220>                           |    |
|        | <223> 未知的描述: Smad结合元件寡核苷酸       |    |
|        | <400> 17                        |    |
|        | tgtctgtgct                      | 10 |
|        | <210> 18                        |    |
|        | <211> 10                        |    |
|        | <212> DNA                       |    |
|        | <213> 人工序列                      |    |
|        | <220>                           |    |
|        | <223> 人工序列的描述: 合成寡核苷酸           |    |
|        | <400> 18                        |    |
|        | tgatagagct                      | 10 |
|        | <210> 19                        |    |
|        | <211> 26                        |    |
|        | <212> DNA                       |    |
|        | <213> 人工序列                      |    |
|        | <220>                           |    |
|        | <223> 人工序列的描述: 合成引物             |    |
|        | <400> 19                        |    |
|        | atcctcgagt atcctccagg tctggg    | 26 |
|        | <210> 20                        |    |
|        | <211> 28                        |    |
|        | <212> DNA                       |    |
|        | <213> 人工序列                      |    |
|        | <220>                           |    |
|        | <223> 人工序列的描述: 合成引物             |    |
|        | <400> 20                        |    |
|        | gccaaagctta gcgtccagcg ttaacctg | 28 |
|        | <210> 21                        |    |
|        | <211> 1392                      |    |
|        | <212> DNA                       |    |
|        | <213> 人工序列                      |    |
|        | <220>                           |    |
|        | <223> 人工序列的描述: 合成多核苷酸           |    |
|        | <400> 21                        |    |

|        |                                                                    |      |
|--------|--------------------------------------------------------------------|------|
|        | ggatccggcc gtaaaaaacg ccgtcaacgc cgcctgggtt tccgtacgaa acgctcggcc  | 60   |
|        | ctggtccgtc gccgtggcgc ctcccgtgct cgggtgggtg aagatgaaga agaaggtgct  | 120  |
|        | ggcggcggtg gcggtggcgg tgaactgcgt ggcgagggtg caaccgatag tcgtgcacac  | 180  |
|        | ggtgcaggcg gtggcgggcc gggtcgtgct ggttgcctgc tgggtaaagc tgtgcgcggc  | 240  |
|        | gcgaaaggtc atcaccatcc gcacccgcgc gcagcaggtg cagggtgcgc tggcgggtgcg | 300  |
|        | gaagccgata tgaaagccct gaccatagtg gtctgaaaa aactgaaaga acgtcagctg   | 360  |
|        | gagctgctgc tgcaagcagt agaatccctg ggcggtaccc gtacggcttg tctgctgctg  | 420  |
|        | ccgggtcgtc tggattgccg tctgggtccg ggtgcaccgg ctggtgcgca gccggcacia  | 480  |
|        | ccgccgagct cttacagcct ccgcctgctg ctgtgtaaag tgtttcgttg gccggacctg  | 540  |
|        | cgcacagtt ccgaagttaa acgcctgtgc tgttgcgaga gctatggcaa aattaaccgc   | 600  |
|        | gaactgggtt gttgcaatcc gcaccatcgt tctcgtctgt gtgaactgga gagcccgccg  | 660  |
|        | ccgccgtatt ctcgttaccc gatggatttc ctgaaaccga ctgcagattg cccggacgca  | 720  |
|        | gtcccgtcat cggctgagac cggcggcacc aactatcttg caccgggcgg tctgagtgat  | 780  |
|        | tcccagctgc tgcctggaacc gggcgaccgt tcacattggt gtgtggttgc ctattgggaa | 840  |
|        | gagaaaacgc gtgtcggctc cctgtactgc gtacaggaac cgtcgtgga tatctttat    | 900  |
|        | gacctgcgcg agggcaatgg tttctgtctg ggccaactga actcagataa taaatcgag   | 960  |
|        | ctggtgcaaa aagttcgtc aaaaattggc tgcggtatcc agctgaccgc tgaagttgac   | 1020 |
| [0011] | ggtgtctggg tatataaccg cagctcttac ccgattttta tcaaaagtgc caccctggat  | 1080 |
|        | ggtgtctggg tatataaccg cagctcttac ccgattttta tcaaaagtgc caccctggat  | 1140 |
|        | aatccggact cccgtacgt gctggtccac aaagtatttc cgggcttctc aatcaaagcg   | 1200 |
|        | ccgtggacgg gttttactgt gcagatctct ttctgtaaag gctggggtca atgetacacc  | 1260 |
|        | cgtcagttta tctcgtcttg tccgtgctgg ctggaagtga tttcaatag ccgcaagggc   | 1320 |
|        | gagctcaatt cgaagcttga aggttaagct atccctaacc ctctctcgg tctcgattct   | 1380 |
|        | acgtgagtcg ac                                                      | 1392 |
|        | <210> 22                                                           |      |
|        | <211> 1281                                                         |      |
|        | <212> DNA                                                          |      |
|        | <213> 智人                                                           |      |
|        | <400> 22                                                           |      |
|        | atgttcagga ccaaacgata tgcctcgtc cggcgtctct ggaggagccg tgcgcccgcc   | 60   |
|        | ggcagaggacg aggaggaggc gcaggggga ggtggaggag gaggcgagct gcggggagaa  | 120  |
|        | ggggcgacgg acagccgagc gcatggggcc ggtggcggcg gcccgggcag ggctggatgc  | 180  |
|        | tgcttgggca aggcgggtgc aggtgccaaa ggtcaccacc atccccaccc gccagccgcg  | 240  |
|        | ggcgcggcgc cggccggggg gcgcgaggcg gatctgaagg cgtcacgca ctcggtgctc   | 300  |
|        | aagaaactga aggagcgga cgtggagctg ctgctccagg ccgtggagtc ccgcccggg    | 360  |
|        | acgcgcaccg cgtgcctcct gctgcccggc cgcctggact gcaggtggg cccgggggcg   | 420  |
|        | cccgccggcg cgcagcctgc gcagccgccc tctctctact cgtccccct cctgctgtgc   | 480  |



```

aaagtgttca ggtggccgga tctcaggcat tctcgggaag tcaagagget gtgttgcgtg 540
gaatcttiacg ggaagatcaa ccccgagctg gtgtgctgca acccccatca ccttagccga 600
ctctgcgaac tagagtctcc cccccctect tactccagat acccgatgga ttttctcaaa 660
ccaactgcag actgtccaga tgctgtgcct tcttccgctg aaacaggggg aacgaattat 720
ctggccccctg gggggcttct agattcccaa cttcttctgg agcctgggga tgggtcacac 780
tggtgctgtg tggcatactg ggaggagaag acgagagtgg ggaggctcta ctgtgtccag 840
gagccctctc tggatatctt ctatgatcta cctcagggga atggcttttg cctcggacag 900
ctcaattcgg acaacaagag tcagctgggtg cagaagggtc ggagcaaaat cggctgcggc 960
atccagctga cgggggaggt ggatgggtgtg tgggtgtaca accgcagcag ttaccccatc 1020
ttcatcaagt ccgccacact ggacaaccgc gactccagga cgtgtttggt acacaagggtg 1080
ttccccggtt tctccatcaa ggctttcgac tacgagaagg cgtacagcct gcagcggccc 1140
aatgaccacg agtttatgca gcagccgtgg acgggcttta ccgtgcagat cagctttgtg 1200
aagggtggg gtcagtgtca caccgccag ttcacagca gctgcccggtg ctggctagag 1260
gtcatcttca acagccggtg g 1281

```

&lt;210&gt; 23

&lt;211&gt; 1392

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

[0012]

&lt;223&gt; 人工序列的描述: 合成多核苷酸

&lt;400&gt; 23

```

ggatccggcc gtaaaaaacg ccgtcaacgc cggcgtggtt tccgtacgaa acgtccggcc 60
ctggtccgct gccgtggcg ctcctgtgct cgggtgtgtg aagatgaaga agaagggtct 120
ggcggcggtg gcggtggcgg tgaactgctg ggcgaggggt caaccgatag tctgtcacac 180
ggtgcaggcg gtggcggtcc gggctgtgct ggttgcgtgc tgggtaaage tgtgcgcggc 240
gcgaaaggtc atcaccatcc gcaccgccg gcagcagggt cagggtgcagc tggcggtgcg 300
gaagccgata tgaaagccct gaccatagt gtcctgaaaa aactgaaaga acgtcagctg 360
gagctgtctg tgcaagcagt agaattccgt ggccgttacc gtacggcttg tctgtctgtg 420
ccgggtcgtc tggattgccg tctgggtccg ggtgcaccgg ctggtgcgca gccggcaca 480
ccgccgagct cttacagcct gccgtgtctg ctgtgtaaag tgtttcgttg gccggacctg 540
cgccacagtt ccgaagttaa acgcctgtgc tgttgcgaga gctatggcaa aattaaccg 600
gaactggttt gttgcaatcc gcaccatctg tctcgtctgt gtgaactgga gagcccgccg 660
ccgccgtatt ctcgttaccg gatggatttc ctgaaaccga ctgcagattg cccggacgca 720
gtcccgctat cggctgagac cggcggcacc aactatctgg caccggcggt tctgagtgat 780
tcccagctgc tgciggaacc gggcgaccgt tcacattggt gtgtggttgc ctattgggaa 840
gagaaaaacg gtgtcggctc cctgtactgc gtacaggaac cgtcgtgga tatctttat 900
gacctgcgcg agggcaatgg tttctgtctg ggccaactga actcagataa taaatgcag 960
ctggtgcaaa aagttcgtc aaaaattggc tgcggtatcc agctgaccgc tgaagttgac 1020

```

|  |                                                                    |      |
|--|--------------------------------------------------------------------|------|
|  | gggtgtctggg tatataaccg cagctcttac ccgattttta tcaaaagtgc caccctggat | 1080 |
|  | aatccggact cccgtacgct gctgggccac aaagtatttc egggtttctc aatcaaagcg  | 1140 |
|  | ttcgattiac agaaagccta ctgctgcag cgcccgaacg accatgaatt catgcagcaa   | 1200 |
|  | ccgtggacgg gttttactgt gcagatctct ttcgttaaag gctggggtea atgctacacc  | 1260 |
|  | cgtcagttaa tctctctctg tccgtgctgg ctggaagtga ttttcaatag ccgcaagggc  | 1320 |
|  | gagctcaatt cgaagcttga aggtaagcct atccctaacc ctctctcgg tctcgattct   | 1380 |
|  | acgtgagtcg ac                                                      | 1392 |
|  | <210> 24                                                           |      |
|  | <211> 618                                                          |      |
|  | <212> DNA                                                          |      |
|  | <213> 人工序列                                                         |      |
|  | <220>                                                              |      |
|  | <223> 人工序列的描述: 合成多核苷酸                                              |      |
|  | <220>                                                              |      |
|  | <221> CDS                                                          |      |
|  | <222> (1).. (609)                                                  |      |
|  | <400> 24                                                           |      |
|  | gga tcc ggc cgt aaa aaa cgc cgt caa cgc cgc cgt tca cat tgg tgt    | 48   |
|  | Gly Ser Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Ser His Trp Cys    |      |
|  | 1 5 10 15                                                          |      |
|  | gtg gtt gcc tat tgg gaa gag aaa acg cgt gtc ggt cgc ctg tac tgc    | 96   |
|  | Val Val Ala Tyr Trp Glu Glu Lys Thr Arg Val Gly Arg Leu Tyr Cys    |      |
|  | 20 25 30                                                           |      |
|  | gta cag gaa cgc tgc ctg gat atc ttt tat gac ctg ccg cag ggc aat    | 144  |
|  | Val Gln Glu Pro Ser Leu Asp Ile Phe Tyr Asp Leu Pro Gln Gly Asn    |      |
|  | 35 40 45                                                           |      |
|  | ggt ttc tgt ctg ggc caa ctg aac tca gat aat aaa tgc cag ctg gtg    | 192  |
|  | Gly Phe Cys Leu Gly Gln Leu Asn Ser Asp Asn Lys Ser Gln Leu Val    |      |
|  | 50 55 60                                                           |      |
|  | caa aaa gtt cgc tca aaa att ggc tgc ggt atc cag ctg acc cgt gaa    | 240  |
|  | Gln Lys Val Arg Ser Lys Ile Gly Cys Gly Ile Gln Leu Thr Arg Glu    |      |
|  | 65 70 75 80                                                        |      |
|  | gtt gac ggt gtc tgg gta tat aac cgc agc tct tac ccg att ttt atc    | 288  |
|  | Val Asp Gly Val Trp Val Tyr Asn Arg Ser Ser Tyr Pro Ile Phe Ile    |      |
|  | 85 90 95                                                           |      |
|  | aaa agt gcc acc ctg gat aat ccg gac tcc cgt acg ctg ctg gtc cac    | 336  |
|  | Lys Ser Ala Thr Leu Asp Asn Pro Asp Ser Arg Thr Leu Leu Val His    |      |
|  | 100 105 110                                                        |      |
|  | aaa gta ttt ccg ggc ttc tca atc aaa gcg ttc gat tac gag aaa gcc    | 384  |
|  | Lys Val Phe Pro Gly Phe Ser Ile Lys Ala Phe Asp Tyr Glu Lys Ala    |      |
|  | 115 120 125                                                        |      |
|  | tac tgc ctg cag cgc ccg aac gac cat gaa ttc atg cag caa ccg tgg    | 432  |
|  | Tyr Ser Leu Gln Arg Pro Asn Asp His Glu Phe Met Gln Gln Pro Trp    |      |
|  | 130 135 140                                                        |      |
|  | acg ggt ttt act gtg cag atc tct ttc gtt aaa ggc tgg ggt caa tgc    | 480  |
|  | Thr Gly Phe Thr Val Gln Ile Ser Phe Val Lys Gly Trp Gly Gln Cys    |      |
|  | 145 150 155 160                                                    |      |
|  | tac acc cgt cag ttt atc tgc tcc tgt ccg tgc tgg ctg gaa gtg att    | 528  |

[0013]

Tyr Thr Arg Gln Phe Ile Ser Ser Cys Pro Cys Trp Leu Glu Val Ile  
 165 170 175  
 ttc aat agc cgc aag ggc gag ctc aat tcg aag ctt gaa ggt aag cct 576  
 Phe Asn Ser Arg Lys Gly Glu Leu Asn Ser Lys Leu Glu Gly Lys Pro  
 180 185 190  
 atc cct aac cct ctc ctc ggt ctc gat tct acg tgagtcgac 618  
 Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr  
 195 200  
 <210> 25  
 <211> 203  
 <212> PRT  
 <213> 人工序列  
 <220>  
 <223> 人工序列的描述: 合成多肽  
 <400> 25  
 Gly Ser Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Ser His Trp Cys  
 1 5 10 15  
 Val Val Ala Tyr Trp Glu Glu Lys Thr Arg Val Gly Arg Leu Tyr Cys  
 20 25 30  
 Val Gln Glu Pro Ser Leu Asp Ile Phe Tyr Asp Leu Pro Gln Gly Asn  
 35 40 45  
 Gly Phe Cys Leu Gly Gln Leu Asn Ser Asp Asn Lys Ser Gln Leu Val  
 50 55 60  
 Gln Lys Val Arg Ser Lys Ile Gly Cys Gly Ile Gln Leu Thr Arg Glu  
 65 70 75 80  
 Val Asp Gly Val Trp Val Tyr Asn Arg Ser Ser Tyr Pro Ile Phe Ile  
 85 90 95  
 Lys Ser Ala Thr Leu Asp Asn Pro Asp Ser Arg Thr Leu Leu Val His  
 100 105 110  
 Lys Val Phe Pro Gly Phe Ser Ile Lys Ala Phe Asp Tyr Glu Lys Ala  
 115 120 125  
 Tyr Ser Leu Gln Arg Pro Asn Asp His Glu Phe Met Gln Gln Pro Trp  
 130 135 140  
 Thr Gly Phe Thr Val Gln Ile Ser Phe Val Lys Gly Trp Gly Gln Cys  
 145 150 155 160  
 Tyr Thr Arg Gln Phe Ile Ser Ser Cys Pro Cys Trp Leu Glu Val Ile  
 165 170 175  
 Phe Asn Ser Arg Lys Gly Glu Leu Asn Ser Lys Leu Glu Gly Lys Pro  
 180 185 190  
 Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr  
 195 200

[0014]

|        |                                                                 |     |
|--------|-----------------------------------------------------------------|-----|
| <210>  | 26                                                              |     |
| <211>  | 861                                                             |     |
| <212>  | DNA                                                             |     |
| <213>  | 人工序列                                                            |     |
| <220>  |                                                                 |     |
| <223>  | 人工序列的描述: 合成多核苷酸                                                 |     |
| <220>  |                                                                 |     |
| <221>  | CDS                                                             |     |
| <222>  | (1).. (852)                                                     |     |
| <400>  | 26                                                              |     |
|        | gga tcc ggc cgt aaa aaa cgc cgt caa cgc cgc cgt ggt ttc cgt acg | 48  |
|        | Gly Ser Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Phe Arg Thr |     |
|        | 1 5 10 15                                                       |     |
|        | aaa cgc tgc gcc ctg gtc cgt cgc ctg tgg cgc tcc cgt gct ccg ggt | 96  |
|        | Lys Arg Ser Ala Leu Val Arg Arg Leu Trp Arg Ser Arg Ala Pro Gly |     |
|        | 20 25 30                                                        |     |
|        | ggt gaa gat gaa gaa gaa ggt gct ggc ggc ggt ggc ggt ggc ggt gaa | 144 |
|        | Gly Glu Asp Glu Glu Glu Gly Ala Gly Gly Gly Gly Gly Gly Gly Glu |     |
|        | 35 40 45                                                        |     |
|        | ctg cgt ggc gag ggt gca acc gat agt cgt gca cac ggt gca ggc ggt | 192 |
|        | Leu Arg Gly Glu Gly Ala Thr Asp Ser Arg Ala His Gly Ala Gly Gly |     |
|        | 50 55 60                                                        |     |
|        | ggc ggt ccg ggt cgt gct ggt tgc tgt ctg ggt aaa gct gtg cgc ggc | 240 |
|        | Gly Gly Pro Gly Arg Ala Gly Cys Cys Leu Gly Lys Ala Val Arg Gly |     |
|        | 65 70 75 80                                                     |     |
| [0015] | gcg aaa ggt cat cac cat ccg cac ccg ccg gca gca ggt gca ggt gca | 288 |
|        | Ala Lys Gly His His His Pro His Pro Pro Ala Ala Gly Ala Gly Ala |     |
|        | 85 90 95                                                        |     |
|        | gct ggc ggt gcg gaa gcc gat ctg aaa gcc ctg acc cat agt gtc ctg | 336 |
|        | Ala Gly Gly Ala Glu Ala Asp Leu Lys Ala Leu Thr His Ser Val Leu |     |
|        | 100 105 110                                                     |     |
|        | aaa aaa ctg aaa gaa cgt cag ctg gag ctg ctg ctg caa gca gta gaa | 384 |
|        | Lys Lys Leu Lys Glu Arg Gln Leu Glu Leu Leu Gln Ala Val Glu     |     |
|        | 115 120 125                                                     |     |
|        | tcc cgt ggc ggt acc cgt acg gct tgt ctg ctg ctg ccg ggt cgt ctg | 432 |
|        | Ser Arg Gly Gly Thr Arg Thr Ala Cys Leu Leu Leu Pro Gly Arg Leu |     |
|        | 130 135 140                                                     |     |
|        | gat tgc cgt ctg ggt ccg ggt gca ccg gct ggt gcg cag ccg gca caa | 480 |
|        | Asp Cys Arg Leu Gly Pro Gly Ala Pro Ala Gly Ala Gln Pro Ala Gln |     |
|        | 145 150 155 160                                                 |     |
|        | ccg ccg agc tct tac agc ctg ccg ctg ctg ctg tgt aaa gtg ttt cgt | 528 |
|        | Pro Pro Ser Ser Tyr Ser Leu Pro Leu Leu Leu Cys Lys Val Phe Arg |     |
|        | 165 170 175                                                     |     |
|        | tgg ccg gac ctg cgc cac agt tcc gaa gtt aaa cgc ctg tgc tgt tgc | 576 |
|        | Trp Pro Asp Leu Arg His Ser Ser Glu Val Lys Arg Leu Cys Cys Cys |     |
|        | 180 185 190                                                     |     |
|        | gag agc tat ggc aaa att aac ccg gaa ctg gtt tgt tgc aat ccg cac | 624 |
|        | Glu Ser Tyr Gly Lys Ile Asn Pro Glu Leu Val Cys Cys Asn Pro His |     |
|        | 195 200 205                                                     |     |
|        | cat ctg tct cgt ctg tgt gaa ctg gag agc ccg ccg ccg ccg tat tct | 672 |
|        | His Leu Ser Arg Leu Cys Glu Leu Glu Ser Pro Pro Pro Pro Tyr Ser |     |
|        | 210 215 220                                                     |     |

cgt tac ccg atg gat ttc ctg aaa ccg act gca gat tgc ccg gac gca 720  
Arg Tyr Pro Met Asp Phe Leu Lys Pro Thr Ala Asp Cys Pro Asp Ala  
225 230 235 240

gtc ccg tca tgc gct gag acc ggc ggc acc aac tat ctg gca ccg ggc 768  
Val Pro Ser Ser Ala Glu Thr Gly Gly Thr Asn Tyr Leu Ala Pro Gly  
245 250 255

ggt ctg agt gat tcc cag ctg ctg ctg gaa ccg ggc gac cgt ggt aag 816  
Gly Leu Ser Asp Ser Gln Leu Leu Leu Glu Pro Gly Asp Arg Gly Lys  
260 265 270

cct atc cct aac cct ctc ctc ggt ctc gat tct acg tgagtegac 861  
Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr  
275 280

<210> 27  
<211> 284  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成多肽

<400> 27  
Gly Ser Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Phe Arg Thr  
1 5 10 15

Lys Arg Ser Ala Leu Val Arg Arg Leu Trp Arg Ser Arg Ala Pro Gly  
20 25 30

[0016]

Gly Glu Asp Glu Glu Glu Gly Ala Gly Gly Gly Gly Gly Gly Glu  
35 40 45

Leu Arg Gly Glu Gly Ala Thr Asp Ser Arg Ala His Gly Ala Gly Gly  
50 55 60

Gly Gly Pro Gly Arg Ala Gly Cys Cys Leu Gly Lys Ala Val Arg Gly  
65 70 75 80

Ala Lys Gly His His His Pro His Pro Pro Ala Ala Gly Ala Gly Ala  
85 90 95

Ala Gly Gly Ala Glu Ala Asp Leu Lys Ala Leu Thr His Ser Val Leu  
100 105 110

Lys Lys Leu Lys Glu Arg Gln Leu Glu Leu Leu Leu Gln Ala Val Glu  
115 120 125

Ser Arg Gly Gly Thr Arg Thr Ala Cys Leu Leu Leu Pro Gly Arg Leu  
130 135 140

Asp Cys Arg Leu Gly Pro Gly Ala Pro Ala Gly Ala Gln Pro Ala Gln  
145 150 155 160

Pro Pro Ser Ser Tyr Ser Leu Pro Leu Leu Leu Cys Lys Val Phe Arg  
165 170 175

Trp Pro Asp Leu Arg His Ser Ser Glu Val Lys Arg Leu Cys Cys Cys  
 180 185 190  
 Glu Ser Tyr Gly Lys Ile Asn Pro Glu Leu Val Cys Cys Asn Pro His  
 195 200 205  
 His Leu Ser Arg Leu Cys Glu Leu Glu Ser Pro Pro Pro Pro Tyr Ser  
 210 215 220  
 Arg Tyr Pro Met Asp Phe Leu Lys Pro Thr Ala Asp Cys Pro Asp Ala  
 225 230 235 240  
 Val Pro Ser Ser Ala Glu Thr Gly Gly Thr Asn Tyr Leu Ala Pro Gly  
 245 250 255  
 Gly Leu Ser Asp Ser Gln Leu Leu Leu Glu Pro Gly Asp Arg Gly Lys  
 260 265 270  
 Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr  
 275 280

<210> 28  
 <211> 1442  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成多核苷酸

[0017]

<400> 28  
 ggatccggcc gtaaaaaacg ccgtcaacgc cgcctggtt tccgtacgaa acgtcggcc 60  
 ctggtccgtc gcctgtggcg ctcccgctc cgggtggtg aagatgaaga agaaggtgct 120  
 ggcggcggtg cgggtggcgg tgaactcgt aagatgaaga agaaggtgct ggcggcggtg 180  
 gcggtggcgg tgaactcgt ggcgagggtg caaccgatag tcgtgcacac ggtgcaggcg 240  
 gtggcggtcc gggctcgtct ggttctgtc tgggtaaagc tgtgcgcgcgc gcgaaaggtc 300  
 atcaccatcc gcacccgcgc gcagcagggt cagggtgcgc tggcggtgcg gaagccgatc 360  
 tgaaagccct gaccatagtg gtctgaaaa aactgaaaga acgtcagctg gagctcgtgc 420  
 tgcaagcagt agaatccgt ggcggtaacc gtacggcttg tctgtcgtc cgggtcgtc 480  
 tggattgcgc tctgggtccg ggtgcaccgg ctggtgcgc gccggcaca ccccgagct 540  
 cttacagcct gccgtcgtc ctgtgtaaag tgttctgtt gccggacctg cgccacagtt 600  
 ccgaagttaa acgcctgtgc tgttcgaga gctatggcaa aattaaccgc gaactggttt 660  
 gttgcaatcc gcaccatctg tctcgtctgt gtgaactgga gagcccgccg ccgcctgatt 720  
 ctcgttaccc gatggatttc ctgaaacga ctgcagattg cccggacgca gtcctcgtcat 780  
 cggctgagac cggcggcacc aactatctgg caccggcgcg tctgagtgat tcccagctgc 840  
 tgcctggaacc gggcgaccgt tcacattggt gttgtgttc ctattgggaa gagaaaacgc 900  
 gtgtcggtcg cctgtactgc gtacaggaac cgtcgtgga tatcttttat gacctgccgc 960  
 agggcaatgg tttctgtctg ggccaactga actcagataa taaatgcag ctggtgcaaa 1020

```

aagttcgtc aaaaattggc tgcggtatcc agctgacccg tgaagttgac ggtgtctggg 1080
tatataaccg cagctcttac ccgattttta tcaaaagtgc caccctggat aatccggact 1140
cccgtacgct gctggteccac aaagtatttc cgggcttctc aatcaaagcg ttcgattacg 1200
agaaagccta ctgctgcag cgcccgaacg accatgaatt catgcagcaa ccgtggacgg 1260
gttttactgt gcagatctct ttcgttaaag gctggggtea atgctacacc cgtcagttta 1320
tctcgtcctg tccgtgctgg ctggaagtga ttttcaatag ccgcaagggc gagtcaatt 1380
cgaagcttga aggttaagcct atccctaacc ctctctcggg tctcgattct acgtgagtcg 1440
ac 1442

```

<210> 29  
 <211> 66  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成寡核苷酸

```

<400> 29
agcagctttc agtcagtagc agccagcagc cagagcagca gccagtagca gcagcagcag 60
agcage 66

```

<210> 30  
 <211> 1392  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成多核苷酸

<220>  
 <221> CDS  
 <222> (1)..(1383)

```

<400> 30
gga tcc ggc cgt aaa aaa cgc cgt caa cgc cgc cgt ggt ttc cgt acg 48
Gly Ser Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Phe Arg Thr
1 5 10 15

aaa cgc agc gcc ctg gtc cgt cgc ctg tgg cgc agc cgt gct ccg ggt 96
Lys Arg Ser Ala Leu Val Arg Arg Leu Trp Arg Ser Arg Ala Pro Gly
20 25 30

ggt gaa gat gaa gaa gaa ggt gct ggc ggc ggt ggc ggt ggc ggt gaa 144
Gly Glu Asp Glu Glu Glu Gly Ala Gly Gly Gly Gly Gly Gly Gly Glu
35 40 45

ctg cgt ggc gag ggt gca acc gat agc cgt gca cat ggt gca ggc ggt 192
Leu Arg Gly Glu Gly Ala Thr Asp Ser Arg Ala His Gly Ala Gly Gly
50 55 60

ggc ggt ccg ggt cgt gct ggt tgc tgt ctg ggt aaa gct gtg cgc ggc 240
Gly Gly Pro Gly Arg Ala Gly Cys Cys Leu Gly Lys Ala Val Arg Gly
65 70 75 80

gcg aaa ggt cat cat cat ccg cat ccg ccg gca gca ggt gca ggt gca 288
Ala Lys Gly His His His Pro His Pro Pro Ala Ala Gly Ala Gly Ala
85 90 95

gct ggc ggt gcg gaa gcc gat ctg aaa gcc ctg acc cat agc gtc ctg 336
Ala Gly Gly Ala Glu Ala Asp Leu Lys Ala Leu Thr His Ser Val Leu

```

[0018]

|        | 100                                                                                                                                                   | 105 | 110 |      |
|--------|-------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----|------|
|        | aaa aaa ctg aaa gaa cgt cag ctg gag ctg ctg ctg caa gca gta gaa<br>Lys Lys Leu Lys Glu Arg Gln Leu Glu Leu Leu Leu Gln Ala Val Glu<br>115 120 125     |     |     | 384  |
|        | agc cgt ggc ggt acc cgt acg gct tgt ctg ctg ctg ccg ggt cgt ctg<br>Ser Arg Gly Gly Thr Arg Thr Ala Cys Leu Leu Leu Pro Gly Arg Leu<br>130 135 140     |     |     | 432  |
|        | gat tgc cgt ctg ggt ccg ggt gca ccg gct ggt gcg cag ccg gca caa<br>Asp Cys Arg Leu Gly Pro Gly Ala Pro Ala Gly Ala Gln Pro Ala Gln<br>145 150 155 160 |     |     | 480  |
|        | ccg ccg agc agc tac agc ctg ccg ctg ctg ctg tgt aaa gtg ttt cgt<br>Pro Pro Ser Ser Tyr Ser Leu Pro Leu Leu Leu Cys Lys Val Phe Arg<br>165 170 175     |     |     | 528  |
|        | tgg ccg gac ctg cgc cat agc agc gaa gtt aaa cgc ctg tgc tgt tgc<br>Trp Pro Asp Leu Arg His Ser Ser Glu Val Lys Arg Leu Cys Cys Cys<br>180 185 190     |     |     | 576  |
|        | gag agc tat ggc aaa att aac ccg gaa ctg gtt tgt tgc aat ccg cat<br>Glu Ser Tyr Gly Lys Ile Asn Pro Glu Leu Val Cys Cys Asn Pro His<br>195 200 205     |     |     | 624  |
|        | cat ctg agc cgt ctg tgt gaa ctg gag agc ccg ccg ccg ccg tat agc<br>His Leu Ser Arg Leu Cys Glu Leu Glu Ser Pro Pro Pro Pro Tyr Ser<br>210 215 220     |     |     | 672  |
|        | cgt tac ccg ctg gat ttc ctg aaa ccg act gca gat tgc ccg gac gca<br>Arg Tyr Pro Leu Asp Phe Leu Lys Pro Thr Ala Asp Cys Pro Asp Ala<br>225 230 235 240 |     |     | 720  |
| [0019] | gtc ccg agc agc gct gag acc ggc ggc acc aac tat ctg gca ccg ggc<br>Val Pro Ser Ser Ala Glu Thr Gly Gly Thr Asn Tyr Leu Ala Pro Gly<br>245 250 255     |     |     | 768  |
|        | ggt ctg agc gat agc cag ctg ctg ctg gaa ccg ggc gac cgt agc cat<br>Gly Leu Ser Asp Ser Gln Leu Leu Leu Pro Gly Asp Arg Ser His<br>260 265 270         |     |     | 816  |
|        | tgg tgt gtg gtt gcc tat tgg gaa gag aaa acg cgt gtc ggt cgc ctg<br>Trp Cys Val Val Ala Tyr Trp Glu Glu Lys Thr Arg Val Gly Arg Leu<br>275 280 285     |     |     | 864  |
|        | tac tgc gta cag gaa ccg agc ctg gat atc ttt tat gac ctg ccg cag<br>Tyr Cys Val Gln Glu Pro Ser Leu Asp Ile Phe Tyr Asp Leu Pro Gln<br>290 295 300     |     |     | 912  |
|        | ggc aat ggt ttc tgt ctg ggc caa ctg aac agc gat aat aaa agc cag<br>Gly Asn Gly Phe Cys Leu Gly Gln Leu Asn Ser Asp Asn Lys Ser Gln<br>305 310 315 320 |     |     | 960  |
|        | ctg gtg caa aaa gtt cgc agc aaa att ggc tgc ggt atc cag ctg acc<br>Leu Val Gln Lys Val Arg Ser Lys Ile Gly Cys Gly Ile Gln Leu Thr<br>325 330 335     |     |     | 1008 |
|        | cgt gaa gtt gac ggt gtc tgg gta tat aac cgc agc agc tac ccg att<br>Arg Glu Val Asp Gly Val Trp Val Tyr Asn Arg Ser Ser Tyr Pro Ile<br>340 345 350     |     |     | 1056 |
|        | ttt atc aaa agc gcc acc ctg gat aat ccg gac agc cgt acg ctg ctg<br>Phe Ile Lys Ser Ala Thr Leu Asp Asn Pro Asp Ser Arg Thr Leu Leu<br>355 360 365     |     |     | 1104 |
|        | gtc cat aaa gta ttt ccg ggc ttc agc atc aaa gcg ttc gat tac gag<br>Val His Lys Val Phe Pro Gly Phe Ser Ile Lys Ala Phe Asp Tyr Glu<br>370 375 380     |     |     | 1152 |
|        | aaa gcc tac agc ctg cag cgc ccg aac gac cat gaa ttc atg cag caa                                                                                       |     |     | 1200 |



Lys Ala Tyr Ser Leu Gln Arg Pro Asn Asp His Glu Phe Met Gln Gln  
 385 390 395 400  
 ccg tgg acg ggt ttt act gtg cag atc agc ttc gtt aaa ggc tgg ggt 1248  
 Pro Trp Thr Gly Phe Thr Val Gln Ile Ser Phe Val Lys Gly Trp Gly  
 405 410 415  
 caa tgc tac acc cgt cag ttt atc agc agc tgt ccg tgc tgg ctg gaa 1296  
 Gln Cys Tyr Thr Arg Gln Phe Ile Ser Ser Cys Pro Cys Trp Leu Glu  
 420 425 430  
 gtg att ttc aat agc cgc aag ggc gag ctc aat agc aag ctt gaa ggt 1344  
 Val Ile Phe Asn Ser Arg Lys Gly Glu Leu Asn Ser Lys Leu Glu Gly  
 435 440 445  
 aag cct atc cct aac cct ctc ctc ggt ctc gat agc acg tgagtcgac 1392  
 Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr  
 450 455 460

&lt;210&gt; 31

&lt;211&gt; 461

&lt;212&gt; PRT

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 人工序列的描述: 合成多肽

&lt;400&gt; 31

Gly Ser Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Phe Arg Thr  
 1 5 10 15

Lys Arg Ser Ala Leu Val Arg Arg Leu Trp Arg Ser Arg Ala Pro Gly  
 20 25 30

Gly Glu Asp Glu Glu Glu Gly Ala Gly Gly Gly Gly Gly Gly Gly Glu  
 35 40 45

Leu Arg Gly Glu Gly Ala Thr Asp Ser Arg Ala His Gly Ala Gly Gly  
 50 55 60

Gly Gly Pro Gly Arg Ala Gly Cys Cys Leu Gly Lys Ala Val Arg Gly  
 65 70 75 80

Ala Lys Gly His His His Pro His Pro Pro Ala Ala Gly Ala Gly Ala  
 85 90 95

Ala Gly Gly Ala Glu Ala Asp Leu Lys Ala Leu Thr His Ser Val Leu  
 100 105 110

Lys Lys Leu Lys Glu Arg Gln Leu Glu Leu Leu Leu Gln Ala Val Glu  
 115 120 125

Ser Arg Gly Gly Thr Arg Thr Ala Cys Leu Leu Leu Pro Gly Arg Leu  
 130 135 140

Asp Cys Arg Leu Gly Pro Gly Ala Pro Ala Gly Ala Gln Pro Ala Gln  
 145 150 155 160

Pro Pro Ser Ser Tyr Ser Leu Pro Leu Leu Leu Cys Lys Val Phe Arg  
 165 170 175

[0020]

[0021]

Trp Pro Asp Leu Arg His Ser Ser Glu Val Lys Arg Leu Cys Cys Cys  
 180 185 190

Glu Ser Tyr Gly Lys Ile Asn Pro Glu Leu Val Cys Cys Asn Pro His  
 195 200 205

His Leu Ser Arg Leu Cys Glu Leu Glu Ser Pro Pro Pro Pro Tyr Ser  
 210 215 220

Arg Tyr Pro Leu Asp Phe Leu Lys Pro Thr Ala Asp Cys Pro Asp Ala  
 225 230 235 240

Val Pro Ser Ser Ala Glu Thr Gly Gly Thr Asn Tyr Leu Ala Pro Gly  
 245 250 255

Gly Leu Ser Asp Ser Gln Leu Leu Leu Glu Pro Gly Asp Arg Ser His  
 260 265 270

Trp Cys Val Val Ala Tyr Trp Glu Glu Lys Thr Arg Val Gly Arg Leu  
 275 280 285

Tyr Cys Val Gln Glu Pro Ser Leu Asp Ile Phe Tyr Asp Leu Pro Gln  
 290 295 300

Gly Asn Gly Phe Cys Leu Gly Gln Leu Asn Ser Asp Asn Lys Ser Gln  
 305 310 315 320

Leu Val Gln Lys Val Arg Ser Lys Ile Gly Cys Gly Ile Gln Leu Thr  
 325 330 335

Arg Glu Val Asp Gly Val Trp Val Tyr Asn Arg Ser Ser Tyr Pro Ile  
 340 345 350

Phe Ile Lys Ser Ala Thr Leu Asp Asn Pro Asp Ser Arg Thr Leu Leu  
 355 360 365

Val His Lys Val Phe Pro Gly Phe Ser Ile Lys Ala Phe Asp Tyr Glu  
 370 375 380

Lys Ala Tyr Ser Leu Gln Arg Pro Asn Asp His Glu Phe Met Gln Gln  
 385 390 395 400

Pro Trp Thr Gly Phe Thr Val Gln Ile Ser Phe Val Lys Gly Trp Gly  
 405 410 415

Gln Cys Tyr Thr Arg Gln Phe Ile Ser Ser Cys Pro Cys Trp Leu Glu  
 420 425 430

Val Ile Phe Asn Ser Arg Lys Gly Glu Leu Asn Ser Lys Leu Glu Gly  
 435 440 445

Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr

|        |                                                                                                                                                             |     |     |     |
|--------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----|-----|
|        | 450                                                                                                                                                         | 455 | 460 |     |
|        | <210> 32<br><211> 861<br><212> DNA<br><213> 人工序列                                                                                                            |     |     |     |
|        | <220><br><223> 人工序列的描述：合成多核苷酸                                                                                                                               |     |     |     |
|        | <220><br><221> CDS<br><222> (1).. (852)                                                                                                                     |     |     |     |
|        | <400> 32<br>gga tcc ggc cgt aaa aaa cgc cgt caa cgc cgc cgt ggt ttc cgt acg<br>Gly Ser Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Phe Arg Thr<br>1 5 10 15 |     |     | 48  |
|        | aaa cgc agc gcc ctg gtc cgt cgc ctg tgg cgc agc cgt gct ccg ggt<br>Lys Arg Ser Ala Leu Val Arg Arg Leu Trp Arg Ser Arg Ala Pro Gly<br>20 25 30              |     |     | 96  |
|        | ggt gaa gat gaa gaa gaa ggt gct ggc ggc ggt ggc ggt ggc ggt gaa<br>Gly Glu Asp Glu Glu Glu Gly Ala Gly Gly Gly Gly Gly Gly Gly Glu<br>35 40 45              |     |     | 144 |
|        | ctg cgt ggc gag ggt gca acc gat agc cgt gca cat ggt gca ggc ggt<br>Leu Arg Gly Glu Gly Ala Thr Asp Ser Arg Ala His Gly Ala Gly Gly<br>50 55 60              |     |     | 192 |
| [0022] | ggc ggt ccg ggt cgt gct ggt tgc tgt ctg ggt aaa gct gtg cgc ggc<br>Gly Gly Pro Gly Arg Ala Gly Cys Cys Leu Gly Lys Ala Val Arg Gly<br>65 70 75 80           |     |     | 240 |
|        | gcg aaa ggt cat cat cat ccg cat ccg ccg gca gca ggt gca ggt gca<br>Ala Lys Gly His His His Pro His Pro Pro Ala Ala Gly Ala Gly Ala<br>85 90 95              |     |     | 288 |
|        | gct ggc ggt gcg gaa gcc gat ctg aaa gcc ctg acc cat agc gtc ctg<br>Ala Gly Gly Ala Glu Ala Asp Leu Lys Ala Leu Thr His Ser Val Leu<br>100 105 110           |     |     | 336 |
|        | aaa aaa ctg aaa gaa cgt cag ctg gag ctg ctg ctg caa gca gta gaa<br>Lys Lys Leu Lys Glu Arg Gln Leu Glu Leu Leu Leu Gln Ala Val Glu<br>115 120 125           |     |     | 384 |
|        | agc cgt ggc ggt acc cgt acg gct tgt ctg ctg ctg ccg ggt cgt ctg<br>Ser Arg Gly Gly Thr Arg Thr Ala Cys Leu Leu Leu Pro Gly Arg Leu<br>130 135 140           |     |     | 432 |
|        | gat tgc cgt ctg ggt ccg ggt gca ccg gct ggt gcg cag ccg gca caa<br>Asp Cys Arg Leu Gly Pro Gly Ala Pro Ala Gly Ala Gln Pro Ala Gln<br>145 150 155 160       |     |     | 480 |
|        | ccg ccg agc agc tac agc ctg ccg ctg ctg ctg tgt aaa gtg ttt cgt<br>Pro Pro Ser Ser Tyr Ser Leu Pro Leu Leu Leu Cys Lys Val Phe Arg<br>165 170 175           |     |     | 528 |
|        | tgg ccg gac ctg cgc cat agc agc gaa gtt aaa cgc ctg tgc tgt tgc<br>Trp Pro Asp Leu Arg His Ser Ser Glu Val Lys Arg Leu Cys Cys Cys<br>180 185 190           |     |     | 576 |
|        | gag agc tat ggc aaa att aac ccg gaa ctg gtt tgt tgc aat ccg cat<br>Glu Ser Tyr Gly Lys Ile Asn Pro Glu Leu Val Cys Cys Asn Pro His<br>195 200 205           |     |     | 624 |
|        | cat ctg agc cgt ctg tgt gaa ctg gag agc ccg ccg ccg ccg tat agc<br>His Leu Ser Arg Leu Cys Glu Leu Glu Ser Pro Pro Pro Pro Tyr Ser                          |     |     | 672 |

|        |                                                                    |     |     |     |
|--------|--------------------------------------------------------------------|-----|-----|-----|
|        | 210                                                                | 215 | 220 |     |
|        | cgt tac ccg atg gat ttc ctg aaa ccg act gca gat tgc ccg gac gca    |     |     | 720 |
|        | Arg Tyr Pro Met Asp Phe Leu Lys Pro Thr Ala Asp Cys Pro Asp Ala    |     |     |     |
|        | 225                                                                | 230 | 235 | 240 |
|        | gtc ccg agc agc gct gag acc ggc ggc acc aac tat ctg gca ccg ggc    |     |     | 768 |
|        | Val Pro Ser Ser Ala Glu Thr Gly Gly Thr Asn Tyr Leu Ala Pro Gly    |     |     |     |
|        | 245                                                                | 250 | 255 |     |
|        | ggt ctg agc gat agc cag ctg ctg ctg gaa ccg ggc gac cgt ggt aag    |     |     | 816 |
|        | Gly Leu Ser Asp Ser Gln Leu Leu Leu Glu Pro Gly Asp Arg Gly Lys    |     |     |     |
|        | 260                                                                | 265 | 270 |     |
|        | cct atc cct aac cct ctc ctc ggt ctc gat tct acg tgagtcgac          |     |     | 861 |
|        | Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr                    |     |     |     |
|        | 275                                                                | 280 |     |     |
|        | <210> 33                                                           |     |     |     |
|        | <211> 558                                                          |     |     |     |
|        | <212> DNA                                                          |     |     |     |
|        | <213> 人工序列                                                         |     |     |     |
|        | <220>                                                              |     |     |     |
|        | <223> 人工序列的描述: 合成多核苷酸                                              |     |     |     |
|        | <400> 33                                                           |     |     |     |
|        | ggatccggcc gtaaaaaacg ccgtcaacgc cgccgttcac attggtgtgt ggttgcctat  |     |     | 60  |
|        | tggaagaga aaacgcgtgt cggctgcctg tactgcgtac aggaaccgag cctggatata   |     |     | 120 |
|        | ttttatgacc tgccgcaggg caatggtttc tgtctgggcc aactgaacag cgataataaa  |     |     | 180 |
| [0023] | agccagctgg tgcaaaaagt tcgcagcaaa attggctgctg gtatccagct gacccgtgaa |     |     | 240 |
|        | gttagcgggtg tctgggtata taaccgcagc agctaccga tttttatcaa aagcgcacc   |     |     | 300 |
|        | ctggataate cggacagccg tacgtctgtg gtccataaag tatttcgggg cttcagcatc  |     |     | 360 |
|        | aaagcgttcg attacgagaa agcctacagc ctgcagcgcc cgaacgacca tgaattcatg  |     |     | 420 |
|        | cagcaaccgt ggacgggttt tactgtgcag atcagcttcg ttaaaggctg gggccaatgc  |     |     | 480 |
|        | aagggcgagc tcaatagcaa gcttgaaggt aagcctatcc ctaacctct cctcggtctc   |     |     | 540 |
|        | gatagcacgt gagtcgac                                                |     |     | 558 |
|        | <210> 34                                                           |     |     |     |
|        | <211> 618                                                          |     |     |     |
|        | <212> DNA                                                          |     |     |     |
|        | <213> 人工序列                                                         |     |     |     |
|        | <220>                                                              |     |     |     |
|        | <223> 人工序列的描述: 合成多核苷酸                                              |     |     |     |
|        | <220>                                                              |     |     |     |
|        | <221> CDS                                                          |     |     |     |
|        | <222> (1).. (609)                                                  |     |     |     |
|        | <400> 34                                                           |     |     |     |
|        | gga tcc ggc cgt aaa aaa cgc cgt caa cgc cgc cgt tca cat tgg tgt    |     |     | 48  |
|        | Gly Ser Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Ser His Trp Cys    |     |     |     |
|        | 1                                                                  | 5   | 10  | 15  |
|        | gtg gtt gcc tat tgg gaa gag aaa acg cgt gtc ggt cgc ctg tac tgc    |     |     | 96  |
|        | Val Val Ala Tyr Trp Glu Glu Lys Thr Arg Val Gly Arg Leu Tyr Cys    |     |     |     |
|        | 20                                                                 | 25  | 30  |     |

|        |                                                                                                                                                       |     |
|--------|-------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
|        | gta cag gaa ccg agc ctg gat atc ttt tat gac ctg ccg cag ggc aat<br>Val Gln Glu Pro Ser Leu Asp Ile Phe Tyr Asp Leu Pro Gln Gly Asn<br>35 40 45        | 144 |
|        | ggg ttc tgt ctg ggc caa ctg aac agc gat aat aaa agc cag ctg gtg<br>Gly Phe Cys Leu Gly Gln Leu Asn Ser Asp Asn Lys Ser Gln Leu Val<br>50 55 60        | 192 |
|        | caa aaa gtt cgc agc aaa att ggc tgc ggt atc cag ctg acc cgt gaa<br>Gln Lys Val Arg Ser Lys Ile Gly Cys Gly Ile Gln Leu Thr Arg Glu<br>65 70 75 80     | 240 |
|        | gtt gac ggt gtc tgg gta tat aac cgc agc agc tac ccg att ttt atc<br>Val Asp Gly Val Trp Val Tyr Asn Arg Ser Ser Tyr Pro Ile Phe Ile<br>85 90 95        | 288 |
|        | aaa agc gcc acc ctg gat aat ccg gac agc cgt acg ctg ctg gtc cat<br>Lys Ser Ala Thr Leu Asp Asn Pro Asp Ser Arg Thr Leu Leu Val His<br>100 105 110     | 336 |
|        | aaa gta ttt ccg ggc ttc agc atc aaa gcg ttc gat tac gag aaa gcc<br>Lys Val Phe Pro Gly Phe Ser Ile Lys Ala Phe Asp Tyr Glu Lys Ala<br>115 120 125     | 384 |
|        | tac agc ctg cag cgc ccg aac gac cat gaa ttc atg cag caa ccg tgg<br>Tyr Ser Leu Gln Arg Pro Asn Asp His Glu Phe Met Gln Gln Pro Trp<br>130 135 140     | 432 |
|        | acg ggt ttt act gtg cag atc agc ttc gtt aaa ggc tgg ggt caa tgc<br>Thr Gly Phe Thr Val Gln Ile Ser Phe Val Lys Gly Trp Gly Gln Cys<br>145 150 155 160 | 480 |
|        | tac acc cgt cag ttt atc agc agc tgt ccg tgc tgg ctg gaa gtg att<br>Tyr Thr Arg Gln Phe Ile Ser Ser Cys Pro Cys Trp Leu Glu Val Ile<br>165 170 175     | 528 |
| [0024] | ttc aat agc cgc aag ggc gag ctc aat agc aag ctt gaa ggt aag cct<br>Phe Asn Ser Arg Lys Gly Glu Leu Asn Ser Lys Leu Glu Gly Lys Pro<br>180 185 190     | 576 |
|        | atc cct aac cct ctc ctc ggt ctc gat agc acg tgagtcgac<br>Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr<br>195 200                                       | 618 |
|        | <210> 35<br><211> 39<br><212> DNA<br><213> 人工序列                                                                                                       |     |
|        | <220><br><223> 人工序列的描述: 合成寡核苷酸                                                                                                                        |     |
|        | <400> 35<br>tcgtcagctgt catcttttcct catcgtcttc gtctcgtct                                                                                              | 39  |
|        | <210> 36<br><211> 1365<br><212> DNA<br><213> 人工序列                                                                                                     |     |
|        | <220><br><223> 人工序列的描述: 合成多核苷酸                                                                                                                        |     |
|        | <220><br><221> CDS<br><222> (1).. (1356)                                                                                                              |     |
|        | <400> 36                                                                                                                                              |     |

[0025]

|                                                                                                                                                           |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| gga tcc ggt cgt aaa aaa cgt cgt cag cgt cgt cgt ggt ttc cgt acc<br>Gly Ser Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Phe Arg Thr<br>1 5 10 15           | 48  |
| aaa cgt tct gcg ctg gtt cgt cgt ctg tgg cgt tct cgt gcg ccg ggt<br>Lys Arg Ser Ala Leu Val Arg Arg Leu Trp Arg Ser Arg Ala Pro Gly<br>20 25 30            | 96  |
| ggt gaa gac gaa gaa gaa ggt gcg ggt ggt ggt ggt ggt ggt gaa<br>Gly Glu Asp Glu Glu Glu Gly Ala Gly Gly Gly Gly Gly Gly Glu<br>35 40 45                    | 144 |
| ctg cgt ggt gaa ggt gcg acc gac tct cgt gcg cac ggt gcg ggt ggt<br>Leu Arg Gly Glu Gly Ala Thr Asp Ser Arg Ala His Gly Ala Gly Gly<br>50 55 60            | 192 |
| ggt ggt ccg ggt cgt gcg ggt tgc tgc ctg ggt aaa gcg gtt cgt ggt<br>Gly Gly Pro Gly Arg Ala Gly Cys Cys Leu Gly Lys Ala Val Arg Gly<br>65 70 75 80         | 240 |
| gcg aaa ggt cac cac cac ccg cac ccg ccg gcg gcg ggt gcg ggt gcg<br>Ala Lys Gly His His His Pro His Pro Pro Ala Ala Gly Ala Gly Ala<br>85 90 95            | 288 |
| gcg ggt ggt gcg gaa gcg gac ctg aaa gcg ctg acc cac tct gtt ctg<br>Ala Gly Gly Ala Glu Ala Asp Leu Lys Ala Leu Thr His Ser Val Leu<br>100 105 110         | 336 |
| aaa aaa ctg aaa gaa cgt cag ctg gaa ctg ctg ctg cag gcg gtt gaa<br>Lys Lys Leu Lys Glu Arg Gln Leu Glu Leu Leu Leu Gln Ala Val Glu<br>115 120 125         | 384 |
| tct cgt ggt ggt acc cgt acc gcg tgc ctg ctg ctg ccg ggt cgt ctg<br>Ser Arg Gly Gly Thr Arg Thr Ala Cys Leu Leu Leu Pro Gly Arg Leu<br>130 135 140         | 432 |
| gac tgc cgt ctg ggt ccg ggt gcg ccg gcg ggt gcg cag ccg gcg cag<br>Asp Cys Arg Leu Gly Tyr Pro Gly Ala Pro Ala Gly Ala Gln Pro Ala Gln<br>145 150 155 160 | 480 |
| ccg ccg tct tct tac tct ctg ccg ctg ctg ctg tgc aaa gtt ttc cgt<br>Pro Pro Ser Ser Tyr Ser Leu Pro Leu Leu Cys Lys Val Phe Arg<br>165 170 175             | 528 |
| tgg ccg gac ctg cgt cac tct tct gaa gtt aaa cgt ctg tgc tgc tgc<br>Trp Pro Asp Leu Arg His Ser Ser Glu Val Lys Arg Leu Cys Cys Cys<br>180 185 190         | 576 |
| gaa tct tac ggt aaa atc aac ccg gaa ctg gtt tgc tgc aac ccg cac<br>Glu Ser Tyr Gly Lys Ile Asn Pro Glu Leu Val Cys Cys Asn Pro His<br>195 200 205         | 624 |
| cac ctg tct cgt ctg tgc gaa ctg gaa tct ccg ccg ccg ccg tac tct<br>His Leu Ser Arg Leu Cys Glu Leu Glu Ser Pro Pro Pro Pro Tyr Ser<br>210 215 220         | 672 |
| cgt tac ccg ctg gac ttc ctg aaa ccg acc gcg gac tgc ccg gac gcg<br>Arg Tyr Pro Leu Asp Phe Leu Lys Pro Thr Ala Asp Cys Pro Asp Ala<br>225 230 235 240     | 720 |
| ggt ccg tct tct gcg gaa acc ggt ggt acc aac tac ctg gcg ccg ggt<br>Val Pro Ser Ser Ala Glu Thr Gly Gly Thr Asn Tyr Leu Ala Pro Gly<br>245 250 255         | 768 |
| ggt ctg tct gac tct cag ctg ctg ctg gaa ccg ggt gac cgt tct cac<br>Gly Leu Ser Asp Ser Gln Leu Leu Leu Glu Pro Gly Asp Arg Ser His<br>260 265 270         | 816 |
| tgg tgc gtt gtt gcg tac tgg gaa gaa aaa acc cgt gtt ggt cgt ctg<br>Trp Cys Val Val Ala Tyr Trp Glu Lys Thr Arg Val Gly Arg Leu<br>275 280 285             | 864 |

|        |                                                                                                                                                       |      |
|--------|-------------------------------------------------------------------------------------------------------------------------------------------------------|------|
|        | tac tgc gtt cag gaa ccg tct ctg gac atc ttc tac gac ctg ccg cag<br>Tyr Cys Val Gln Glu Pro Ser Leu Asp Ile Phe Tyr Asp Leu Pro Gln<br>290 295 300     | 912  |
|        | ggc aac ggt ttc tgc ctg ggt cag ctg aac tct gac aac aaa tct cag<br>Gly Asn Gly Phe Cys Leu Gly Gln Leu Asn Ser Asp Asn Lys Ser Gln<br>305 310 315 320 | 960  |
|        | ctg gtt cag aaa gtt cgt tct aaa atc ggt tgc ggt atc cag ctg acc<br>Leu Val Gln Lys Val Arg Ser Lys Ile Gly Cys Gly Ile Gln Leu Thr<br>325 330 335     | 1008 |
|        | cgt gaa gtt gac ggt gtt tgg gtt tac aac cgt tct tct tac ccg atc<br>Arg Glu Val Asp Gly Val Trp Val Tyr Asn Arg Ser Ser Tyr Pro Ile<br>340 345 350     | 1056 |
|        | ttc atc aaa tct gcg acc ctg gac aac ccg gac tct cgt acc ctg ctg<br>Phe Ile Lys Ser Ala Thr Leu Asp Asn Pro Asp Ser Arg Thr Leu Leu<br>355 360 365     | 1104 |
|        | gtt cac aaa gtt ttc ccg ggt ttc tct atc aaa gcg ttc gac tac gaa<br>Val His Lys Val Phe Pro Gly Phe Ser Ile Lys Ala Phe Asp Tyr Glu<br>370 375 380     | 1152 |
|        | aaa gcg tac tct ctg cag cgt ccg aac gac cac gaa ttc atg cag cag<br>Lys Ala Tyr Ser Leu Gln Arg Pro Asn Asp His Glu Phe Met Gln Gln<br>385 390 395 400 | 1200 |
|        | ccg tgg acc ggt ttc acc gtt cag atc tct ttc gtt aaa ggt tgg ggt<br>Pro Trp Thr Gly Phe Thr Val Gln Ile Ser Phe Val Lys Gly Trp Gly<br>405 410 415     | 1248 |
| [0026] | cag tgc tac acc cgt cag ttc atc tct tct tgc ccg tgc tgg ctg gaa<br>Gln Cys Tyr Thr Arg Gln Phe Ile Ser Ser Cys Pro Cys Trp Leu Glu<br>420 425 430     | 1296 |
|        | gtt atc ttc aac tct cgt ggt aaa ccg atc ccg aac ccg ctg ctg ggt<br>Val Ile Phe Asn Ser Arg Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly<br>435 440 445     | 1344 |
|        | ctg gac tct acc tgagtcgac<br>Leu Asp Ser Thr<br>450                                                                                                   | 1365 |
|        | <210> 37<br><211> 452<br><212> PRT<br><213> 人工序列                                                                                                      |      |
|        | <220><br><223> 人工序列的描述: 合成多肽                                                                                                                          |      |
|        | <400> 37<br>Gly Ser Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Phe Arg Thr<br>1 5 10 15                                                              |      |
|        | Lys Arg Ser Ala Leu Val Arg Arg Leu Trp Arg Ser Arg Ala Pro Gly<br>20 25 30                                                                           |      |
|        | Gly Glu Asp Glu Glu Glu Gly Ala Gly Gly Gly Gly Gly Gly Glu<br>35 40 45                                                                               |      |
|        | Leu Arg Gly Glu Gly Ala Thr Asp Ser Arg Ala His Gly Ala Gly Gly<br>50 55 60                                                                           |      |

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

[0027]



Phe Ile Lys Ser Ala Thr Leu Asp Asn Pro Asp Ser Arg Thr Leu Leu  
 355 360 365  
 Val His Lys Val Phe Pro Gly Phe Ser Ile Lys Ala Phe Asp Tyr Glu  
 370 375 380  
 Lys Ala Tyr Ser Leu Gln Arg Pro Asn Asp His Glu Phe Met Gln Gln  
 385 390 395 400  
 Pro Trp Thr Gly Phe Thr Val Gln Ile Ser Phe Val Lys Gly Trp Gly  
 405 410 415  
 Gln Cys Tyr Thr Arg Gln Phe Ile Ser Ser Cys Pro Cys Trp Leu Glu  
 420 425 430  
 Val Ile Phe Asn Ser Arg Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly  
 435 440 445  
 Leu Asp Ser Thr  
 450

&lt;210&gt; 38

&lt;211&gt; 1365

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 人工序列的描述: 合成多核苷酸

[0028]

&lt;400&gt; 38

```

ggatccggtc gtaaaaaacg tcgtcagcgt cgtcgtggtt tccgtacca acgttctgcg 60
ctggttcgtc gtcgtggcgc ttctcgtgcg ccgggtgggtg aagacgaaga agaaggtgcg 120
ggtgggtgggt gtgggtgggtg tgaactgcgt ggtgaagggtg cgaccgactc tcgtgcgcac 180
gggtcgggtg gtgggtgggtc gggtcgtgcg ggttgctgcc tgggtaaagc gggtcgtggt 240
gcgaaaggte accaccaccc gcaccgcgcg gcggcgggtg cgggtgcggc gggtgggtgcg 300
gaagcggacc tgaagcgct gaccactct gttctgaaaa aactgaaaga acgtcagctg 360
gaactgctgc tgcagcggtg tgaatctct ggtggtaccc gtaccgcgtg cctgctgctg 420
ccgggtcgtc tggactgccg tctgggtccg ggtgcgccgg cgggtgcgca gccggcgag 480
ccgccgtctt ctactctct gccgctgctg ctgtgcaaag ttttcggtg gccggacctg 540
cgtcactctt ctgaagttaa acgtctgtgc tgcgcgaat cttacggtaa aatcaaccg 600
gaactggttt gctgcaaccc gcaccacctg tctcgtctgt gcgaactgga atctccgccg 660
ccgccgtact ctggttacc gctggacttc ctgaaaccga ccgcggactg cccggacgcg 720
gttcctctt ctgcggaac cggtggtacc aactacctgg ccgggggtgg tctgtctgac 780
tctcagctgc tgcgtgaacc gggtagcgt tctcactggt gcgttgtgc gtactgggaa 840
gaaaaaaccc gtgttggtcg tctgtactgc gttcaggaac cgtctctgga catctctac 900
gacctgccgc agggtaacgg tttctgctg ggtcagctga actctgacaa caaatctcag 960
ctggttcaga aagttcgtc taaatcggt tgcggtatcc agctgaccgc tgaagttgac 1020

```

```

gggtgttggg tttaacaacg ttcttcttac cegatcttca tcaaatctgc gacctggac 1080
aaccggact ctctgacct gctggttcac aaagttttcc egggtttctc tatcaaagcg 1140
ttcgactiag aaaaagcgta ctctctgcag cgtccgaacg accacgaatt catgcagcag 1200
ccgtggaccg gtttcaccgt tcagatctct ttcgttaaag gttggggtea gtgtacacc 1260
cgtcagttea tctcttcttg cccgtgctgg ctggaagtta tcttcaactc tegtggtaaa 1320
ccgateccga acccgctgct gggctctggac tctacctgag tcgac 1365

```

<210> 39

<211> 1392

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述: 合成多核苷酸

<400> 39

```

ggatccggcc gtaaaaaacg ccgtcaacgc cgcctggtt tccgtacgaa acgcagcgcc 60
ctggtccgtc gcctgtggcg cagccgtgct cgggtgggtg aagatgaaga agaagggtct 120
ggcggcggtg gcggtggcgg tgaactgcgt ggcgaggggt caaccgatag ccgtgcacat 180
ggtgcaggcg gtggcggtcc gggctgtgct ggttgcgtgc tgggtaaagc tgtgcgcggc 240
gcgaaagtc atcatcatcc gcacccgcgc gcagcaggtg caggtgcagc tggcggtgcg 300
gaagccgacg tgaagccct gaccatagc gtcctgaaaa aactgaaaga acgtcagctg 360
[0029] gagctgtcgc tgcaagcagt agaaagccgt ggcggtaccc gtacggcttg tctgctgtc 420
ccgggtcgtc tggattgccg tctgggtccg ggtgcaccgg ctggtgcgca gccggcacia 480
ccgccgagca gctacagcct gccgtgtctg ctgtgtaaag tgttctgttg gccggacctg 540
cgccatagca gcgaagttaa acgctgtgct tgttgcgaga gctatggcaa aattaaccgc 600
gaactggttt gttgcaatcc gcacatctct agccgtctgt gtgaactgga gagccgcgcg 660
ccgccgtata gccgttaccc gctggatttc ctgaaaccga ctgcagattg cccggacgca 720
gtcccagca gcgtgagac cggcggcacc aactatcttg caccggcgcg tctgagcgat 780
agccagctgc tgcgtgaacc gggcgaccgt agccatttgt gtgtggttgc ctattgggaa 840
gagaaaaacg gtgtcggctg cctgtactgc gtacaggaac cgagcctgga tatctttat 900
gacctgccgc agggcaatgg tttctgtctg ggccaactga acagcgataa taaaagccag 960
ctggtgcaaa aagttcgag caaaattggc tgcggtatcc agctgaccgc tgaagttgac 1020
ggtgtctggg tatataaccg cagcagctac ccgattttta tcaaaagcgc caccctggat 1080
aatccggaca gccgtacgct gctgttccat aaagtatttc cgggcttcag catcaaagcg 1140
ttcgattacg agaaagccta cagcctgcag cgcgcgaacg accatgaatt catgcagcaa 1200
ccgtggaccg gttttactgt gcagatcagc ttcgttaaag gctggggtea atgtacacc 1260
cgtcagttaa tcagcagctg tccgtgctgg ctggaagtga ttttcaatag ccgcaagggc 1320
gagctcaata gcaagcttga aggtaaacct atccctaacc ctctctcgg tctcgatagc 1380
acgtgagtcg ac 1392

```

[0030]

<210> 40  
<211> 6  
<212> PRT  
<213> 人工序列  
  
<220>  
<223> 人工序列的描述: 合成6xHis标签  
  
<400> 40  
His His His His His His  
1 5  
  
<210> 41  
<211> 14  
<212> PRT  
<213> 人工序列  
  
<220>  
<223> 人工序列的描述: 合成肽  
  
<400> 41  
Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr  
1 5 10  
  
<210> 42  
<211> 8  
<212> PRT  
<213> 人工序列  
  
<220>  
<223> 人工序列的描述: 合成肽  
  
<400> 42  
Lys Tyr Lys Asp Asp Asp Asp Lys  
1 5  
  
<210> 43  
<211> 9  
<212> PRT  
<213> 人工序列  
  
<220>  
<223> 人工序列的描述: 合成肽  
  
<400> 43  
Tyr Pro Tyr Asp Val Pro Asp Tyr Ala  
1 5  
  
<210> 44  
<211> 23  
<212> PRT  
<213> 人工序列  
  
<220>  
<223> 人工序列的描述: 合成肽  
  
<400> 44  
Gly Ala Leu Phe Leu Gly Trp Leu Gly Ala Ala Gly Ser Thr Met Gly  
1 5 10 15  
  
Ala Lys Lys Lys Arg Lys Val  
20  
  
<210> 45

<211> 16  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成肽

<400> 45  
Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys  
1 5 10 15

<210> 46  
<211> 7  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成肽

<400> 46  
Arg Arg Met Lys Trp Lys Lys  
1 5

<210> 47  
<211> 16  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成肽

[0031] <400> 47  
Arg Arg Trp Arg Arg Trp Trp Arg Arg Trp Trp Arg Arg Trp Arg Arg  
1 5 10 15

<210> 48  
<211> 18  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成肽

<400> 48  
Arg Gly Gly Arg Leu Ser Tyr Ser Arg Arg Arg Phe Ser Thr Ser Thr  
1 5 10 15

Gly Arg

<210> 49  
<211> 9  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成肽

<400> 49  
Arg Lys Lys Arg Arg Gln Arg Arg Arg  
1 5

<210> 50  
<211> 11  
<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述: 合成肽

<400> 50

Tyr Ala Arg Ala Ala Ala Arg Gln Ala Arg Ala  
1 5 10

<210> 51

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述: 合成肽

<400> 51

Arg Arg Arg Arg Arg Arg Arg Arg  
1 5

<210> 52

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述: 合成肽

<400> 52

Lys Lys Lys Lys Lys Lys Lys Lys  
1 5

[0032]

<210> 53

<211> 27

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述: 合成肽

<220>

<221> MOD\_RES

<222> (25)..(25)

<223> 任何氨基酸

<400> 53

Gly Trp Thr Leu Asn Ser Ala Gly Tyr Leu Leu Gly Lys Ile Asn Leu  
1 5 10 15

Lys Ala Leu Ala Ala Leu Ala Lys Xaa Ile Leu  
20 25

<210> 54

<211> 18

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述: 合成肽

<400> 54

Leu Leu Ile Leu Leu Arg Arg Arg Ile Arg Lys Gln Ala Asn Ala His  
1 5 10 15

Ser Lys

<210> 55

<211> 16

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述：合成肽

<400> 55

Ser Arg Arg His His Cys Arg Ser Lys Ala Lys Arg Ser Arg His His  
1 5 10 15

<210> 56

<211> 11

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述：合成肽

<400> 56

Asn Arg Ala Arg Arg Asn Arg Arg Arg Val Arg  
1 5 10

<210> 57

<211> 15

<212> PRT

<213> 人工序列

[0033]

<220>

<223> 人工序列的描述：合成肽

<400> 57

Arg Gln Leu Arg Ile Ala Gly Arg Arg Leu Arg Gly Arg Ser Arg  
1 5 10 15

<210> 58

<211> 13

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述：合成肽

<400> 58

Lys Leu Ile Lys Gly Arg Thr Pro Ile Lys Phe Gly Lys  
1 5 10

<210> 59

<211> 10

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述：合成肽

<400> 59

Arg Arg Ile Pro Asn Arg Arg Pro Arg Arg  
1 5 10

<210> 60

<211> 18  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述：合成肽

<400> 60  
Lys Leu Ala Leu Lys Leu Ala Leu Lys Ala Leu Lys Ala Ala Leu Lys  
1 5 10 15

Leu Ala

<210> 61  
<211> 14  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述：合成肽

<400> 61  
Lys Leu Ala Lys Leu Ala Lys Lys Leu Ala Lys Leu Ala Lys  
1 5 10

<210> 62  
<211> 27  
<212> PRT  
<213> 人工序列

[0034]

<220>  
<223> 人工序列的描述：合成肽

<400> 62  
Gly Ala Leu Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Asn Gly  
1 5 10 15

Ala Trp Ser Gln Pro Lys Lys Lys Arg Lys Val  
20 25

<210> 63  
<211> 21  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述：合成肽

<400> 63  
Lys Glu Thr Trp Trp Glu Thr Trp Trp Thr Glu Trp Ser Gln Pro Lys  
1 5 10 15

Lys Lys Arg Lys Val  
20

<210> 64  
<211> 20  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述：合成肽

<400> 64  
Leu Lys Lys Leu Leu Lys Lys Leu Leu Lys Lys Leu Leu Lys Lys Leu  
1 5 10 15

Leu Lys Lys Leu  
20

<210> 65  
<211> 34  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述：合成多肽

<400> 65  
Gln Ala Ala Thr Ala Thr Arg Gly Arg Ser Ala Ala Ser Arg Pro Thr  
1 5 10 15

Glu Arg Pro Arg Ala Pro Ala Arg Ser Ala Ser Arg Pro Arg Arg Pro  
20 25 30

Val Glu

<210> 66  
<211> 23  
<212> PRT  
<213> 人工序列

[0035] <220>  
<223> 人工序列的描述：合成肽

<400> 66  
Met Gly Leu Gly Leu His Leu Leu Val Leu Ala Ala Ala Leu Gln Gly  
1 5 10 15

Ala Lys Ser Lys Arg Lys Val  
20

<210> 67  
<211> 26  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述：合成肽

<400> 67  
Ala Ala Val Ala Leu Leu Pro Ala Val Leu Leu Ala Leu Leu Ala Pro  
1 5 10 15

Ala Ala Ala Asn Tyr Lys Lys Pro Lys Leu  
20 25

<210> 68  
<211> 28  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述：合成肽



<400> 68  
Met Ala Asn Leu Gly Tyr Trp Leu Leu Ala Leu Phe Val Thr Met Trp  
1 5 10 15

Thr Asp Val Gly Leu Cys Lys Lys Arg Pro Lys Pro  
20 25

<210> 69  
<211> 24  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成肽

<400> 69  
Leu Gly Thr Tyr Thr Gln Asp Phe Asn Lys Phe His Thr Phe Pro Gln  
1 5 10 15

Thr Ala Ile Gly Val Gly Ala Pro  
20

<210> 70  
<211> 26  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成肽

[0036]

<220>  
<221> MOD\_RES  
<222> (24)..(24)  
<223> 任何氨基酸

<400> 70  
Asp Pro Lys Gly Asp Pro Lys Gly Val Thr Val Thr Val Thr Val Thr  
1 5 10 15

Val Thr Gly Lys Gly Asp Pro Xaa Pro Asp  
20 25

<210> 71  
<211> 14  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成肽

<400> 71  
Pro Pro Pro Pro Pro Pro Pro Pro Pro Pro Pro Pro Pro Pro  
1 5 10

<210> 72  
<211> 18  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成肽

<400> 72  
Val Arg Leu Pro Pro Pro Val Arg Leu Pro Pro Pro Val Arg Leu Pro  
1 5 10 15

Pro Pro

<210> 73  
<211> 10  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成肽

<400> 73  
Pro Arg Pro Leu Pro Pro Pro Arg Pro Gly  
1 5 10

<210> 74  
<211> 30  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成多肽

<400> 74  
Ser Val Arg Arg Arg Pro Arg Pro Pro Tyr Leu Pro Arg Pro Arg Pro  
1 5 10 15

[0037] Pro Pro Phe Phe Pro Pro Arg Leu Pro Pro Arg Ile Pro Pro  
20 25 30

<210> 75  
<211> 21  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成肽

<400> 75  
Thr Arg Ser Ser Arg Ala Gly Leu Gln Phe Pro Val Gly Arg Val His  
1 5 10 15

Arg Leu Leu Arg Lys  
20

<210> 76  
<211> 23  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成肽

<400> 76  
Gly Ile Gly Lys Phe Leu His Ser Ala Lys Lys Phe Gly Lys Ala Phe  
1 5 10 15

Val Gly Glu Ile Met Asn Ser  
20

<210> 77  
 <211> 37  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成多肽

<400> 77  
 Lys Trp Lys Leu Phe Lys Lys Ile Glu Lys Val Gly Gln Asn Ile Arg  
 1 5 10 15

Asp Gly Ile Ile Lys Ala Gly Pro Ala Val Ala Val Val Gly Gln Ala  
 20 25 30

Thr Gln Ile Ala Lys  
 35

<210> 78  
 <211> 28  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成肽

<400> 78  
 Ala Leu Trp Met Thr Leu Leu Lys Lys Val Leu Lys Ala Ala Lys  
 1 5 10 15

[0038] Ala Ala Leu Asn Ala Val Leu Val Gly Ala Asn Ala  
 20 25

<210> 79  
 <211> 26  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成肽

<400> 79  
 Gly Ile Gly Ala Val Leu Lys Val Leu Thr Thr Gly Leu Pro Ala Leu  
 1 5 10 15

Ile Ser Trp Ile Lys Arg Lys Arg Gln Gln  
 20 25

<210> 80  
 <211> 14  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成肽

<400> 80  
 Ile Asn Leu Lys Ala Leu Ala Ala Leu Ala Lys Lys Ile Leu  
 1 5 10

<210> 81

<211> 33  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成多肽

<400> 81  
Gly Phe Phe Ala Leu Ile Pro Lys Ile Ile Ser Ser Pro Leu Pro Lys  
1 5 10 15

Thr Leu Leu Ser Ala Val Gly Ser Ala Leu Gly Gly Ser Gly Gly Gln  
20 25 30

Glu

<210> 82  
<211> 15  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成肽

<400> 82  
Leu Ala Lys Trp Ala Leu Lys Gln Gly Phe Ala Lys Leu Lys Ser  
1 5 10 15

[0039]

<210> 83  
<211> 27  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成肽

<220>  
<221> MOD\_RES  
<222> (23)..(23)  
<223> 任何氨基酸

<400> 83  
Ser Met Ala Gln Asp Ile Ile Ser Thr Ile Gly Asp Leu Val Lys Trp  
1 5 10 15

Ile Ile Gln Thr Val Asn Xaa Phe Thr Lys Lys  
20 25

<210> 84  
<211> 41  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成多肽

<400> 84  
Leu Leu Gly Asp Phe Phe Arg Lys Ser Lys Glu Lys Ile Gly Lys Glu  
1 5 10 15

Phe Lys Arg Ile Val Gln Arg Ile Lys Gln Arg Ile Lys Asp Phe Leu  
20 25 30

Ala Asn Leu Val Pro Arg Thr Glu Ser  
35 40

<210> 85  
<211> 18  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成肽

<400> 85  
Pro Ala Trp Arg Lys Ala Phe Arg Trp Ala Trp Arg Met Leu Lys Lys  
1 5 10 15

Ala Ala

<210> 86  
<211> 18  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成肽

<400> 86  
Lys Leu Lys Leu Lys Leu Lys Leu Lys Leu Lys Leu Lys Leu  
1 5 10 15

[0040]

Lys Leu

<210> 87  
<211> 2182  
<212> DNA  
<213> 人工序列

<220>  
<223> 人工序列的描述: 合成多核苷酸

<400> 87  
gacgttgtat acgactceta tagggcggcc gggaattcgt cgactggatc cggtaccgag 60  
gagatctgcc gccgcgatcg ccatgtcccc cattctgggc tactggaaga ttaagggect 120  
ggtgcagcct actagactgc tgctggaata cctggaggaa aaatatgaag agcatctgta 180  
tgaaagagac gagggggata aatggaggaa caagaaattc gaactgggac tggagtttcc 240  
taatctgcca tactatattg acggcgatgt gaagctgact cagtctatgg ctatcattag 300  
atacatcgca gacaaacaca acatgctggg cgggtgtcct aaggaaaggg cagagattag 360  
tatgctggag ggagccgtgc tggatattag atacggcgtc tcacgcatcg cctatagcaa 420  
agacttcgaa accctgaagg tggattttct gagcaaactg cctgaaatgc tgaagatgtt 480  
cgaggacaga ctgtgccaca aaacctacct gaatggcgac catgtcacac acccagattt 540  
tatgtgttac gacgccctgg atgtggteet gtatatggac cccatgtgte tggatgcttt 600  
ccctaagctg gtgtgcttta agaaaaggat cgaggcaatt ccccagatcg ataagtacct 660

|        |                                                                   |      |
|--------|-------------------------------------------------------------------|------|
|        | gaaaagctcc aagtatatcg ctggcctct ccagggttg caggcaacat tcggaggcgg   | 720  |
|        | ggaccatccc cctaaaagcg acctggaggt gctgtttcag ggaccactgg gcagcgccg  | 780  |
|        | gaagaagcgg cggcagaggc gccgaagtag gtccgcact aagcggtcag cactgggtcg  | 840  |
|        | gagactgtgg cgatctcggg ctcttgagg agaggacgag gaagaggag caggcggcgg   | 900  |
|        | cggaggagga ggagaactgc gcggggagg agctacagat agccgagccc acggagctgg  | 960  |
|        | aggaggagga ccaggggcag ccggatgtg tctgggcaa gcagttagag gcgccaagg    | 1020 |
|        | gcaccatcac ccccatcac ccgccgtgg cgcaggagca gccggcggag ctgaggcaga   | 1080 |
|        | cctgaaagcc ctgactcaca gtgtgtgaa gaaactgaag gaaagacagc tggagctgct  | 1140 |
|        | gtccagga gtgaatcac gcggaggcac ccgaacagct tgtctgtgc tgcggccg       | 1200 |
|        | gctggactgc cggtgggac ccggcgcgc tctggggca cagccagccc agcctccatc    | 1260 |
|        | tagttatage ctgcccctgc tctgtgtaa ggtgttcga tggcctgate tgcggcattc   | 1320 |
|        | aagcgaagtc aaaaggtgt gctgttgcga gtctacggc aagattaacc cagaactggt   | 1380 |
|        | gtgttgcaat ccccatcac tgtctgact gtgtgaactg gattcccc ctcacccta      | 1440 |
|        | ctctaggtat cctatggact ttctgaagcc aaccgtgac tgcagatg cagtgcctc     | 1500 |
|        | ctctgccgag actgggggaa ccaactacct ggtcctggc ggactgagc actccagct    | 1560 |
|        | gctgtgga ccaggggatc gcagccactg gtgtgtggc gctactggg aagagaagac     | 1620 |
|        | aagagtggga aggtgtatt gcgtccagga gccttcctg gacatctct acgatctgc     | 1680 |
| [0041] | acaggggaat ggattttgtc tgggccagct gaactctgac aataagagtc agctgggtca | 1740 |
|        | gaaagtcgg agcaagattg gctcggcat ccagctgacc agggaggtgg acggcgtgtg   | 1800 |
|        | ggtctacaac cgcagttcat atccaattt catcaagagc gccactctgg acaatcccga  | 1860 |
|        | ttccgcacc ctgctgggtc ataagttct ccccgcttc agcatcaagg ccttcgacta    | 1920 |
|        | cgagaaggct tatagtctcc agcggcccaa cgtacagag ttcatgcagc agccttggac  | 1980 |
|        | aggttctact gtgcagatca gcttcgtcaa gggatggggc cagtgtaca caaggcagtt  | 2040 |
|        | catctcaagc tgtccctgtt ggctggaagt catcttcaat agtaggacgc gtacgggcc  | 2100 |
|        | gctcgagcag aaactcatct cagaagagga tctggcagca aatgatatcc tggattacaa | 2160 |
|        | ggatgacgac gataagttt aa                                           | 2182 |

&lt;210&gt; 88

&lt;211&gt; 1059

&lt;212&gt; DNA

&lt;213&gt; 智人

&lt;400&gt; 88

tcaccacat cccacccgc cagccgggg cgcggcgcg gccggggcg ccgaggcgga 60

tctgaaggcg ctacgcact cgggtctcaa gaaactgaag gagcggcagc tggagctgct 120

gtccaggcc gtggagtccc gcggcgggac gcgcaccgag tgcctctgc tgcggcgccg 180

cctggactgc aggtggggc cggggcgcc cgcggcgcg cagcctgcgc agccgcctc 240

gtcctactcg cteccctcc tctgtgcaa agtgttcagg tggccggatc tcaggcattc 300

ctcggaagtc aagaggtgt gtgtgtga atctacggg aagatcaacc ccgagctggt 360

```

gtgtctgaac ccccatcacc ttagecgact ctgcgaacta gagtctcccc cccctcetta 420
ctccagatac ccgatggatt ttctcaaacc aactgcagac tctccagatg ctgtgccttc 480
ctccgctgaa acagggggaa cgaattatct ggccccctggg gggctttcag attcccaact 540
tcttctggag cctgggggatc ggteacactg gtgcgtgggtg gcatactggg aggagaagac 600
gagagtgggg aggcctctact gtgtccagga gccctctctg gatattctct atgatctacc 660
tcaggggaat ggcttttgcc tcggacagct caattcggac aacaagagtc agctgggtgca 720
gaaggtgcgg agcaaaatcg gctgcggcat ccagctgacg cgggaggtgg atggtgtgtg 780
ggtgtacaac cgcagcagtt accccatctt catcaagtc gccacactgg acaaccgga 840
ctccaggacg ctgttggtag acaagggtgt ccccggtttc tccatcaagg ctttcgacta 900
cgagaaggcg tacagcctgc agcggcccaa tgaccacgag ttatgcagc agccgtggac 960
gggctttacc gtgcagatca gctttgtgaa gggctggggt cagtgtaca cccgccagtt 1020
catcagcagc tgeccgtgct ggctagaggt catcttcaa 1059

```

<210> 89

<211> 1053

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述: 合成多核苷酸

<400> 89

[0042]

```

tcaccccat ccacccgctg ctggcgagc agcagccggc ggagctgagg cagacctgaa 60
agccctgact cacagtgtgc tgaagaaact gaaggaaaga cagctggagc tgcgtctcca 120
ggcagtcgaa tcacgaggag gcacccgaac agcttgtctg ctgctgcccg gccggctgga 180
ctgcccgtgt ggacccggcg cccctgtctg ggcacagcca gccacgctc catctagtta 240
tagcctgccc ctgctgtgtg gtaagggttt ccgatggcct gatctgcggc attcaagcga 300
agtcaaaagg ctgtgtgtgt gcgagtccta cggcaagatt aaccagaac tgggtgtgtg 360
caatcccat caccgtgttc gactgtgtga actggagtc cccctccac cctactctag 420
gtatcctatg gactttctga agccaaccgc tgactgcca gatgcagtgc cctcctctgc 480
cgagactggg ggaaccaact acctggctcc tggcggactg agcgactccc agctgtgtct 540
ggaaccaggg gatcgagcc actggtgtgt ggtcgctac tgggaagaga agacaagagt 600
gggaaggtgt tattgcgtcc aggagcctc cctggacatc ttctacgate tgccacaggg 660
gaatggattt tgtctgggcc agctgaactc tgacaataag agtcagctgg tgcagaaagt 720
ccggagcaag attggctgcg gcattcagct gaccaggag gtggacggcg tgtgggtcta 780
caaccgcagt tcataccaa tcttcataa gagegccact ctggacaatc ccgattcccg 840
caccctgtgt gtgcataagg tcttccccg cttcagcatc aaggccttcg actacagaa 900
ggcttatagt ctccagcgcc ccaacgatca cgagttcatg cagcagcctt ggacaggctt 960
cactgtgcag atcagcttcg tcaagggatg gggccagtgc tacacaagge agttcatctc 1020
aagctgtccc tgttggctgg aagtcatttt caa 1053

```

<210> 90  
 <211> 699  
 <212> PRT  
 <213> 人工序列  
  
 <220>  
 <223> 人工序列的描述: 合成多肽  
  
 <400> 90  
 Met Ser Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly Leu Val Gln Pro  
 1 5 10 15  
  
 Thr Arg Leu Leu Leu Glu Tyr Leu Glu Glu Lys Tyr Glu Glu His Leu  
 20 25 30  
  
 Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys Lys Phe Glu Leu  
 35 40 45  
  
 Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp Gly Asp Val Lys  
 50 55 60  
  
 Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala Asp Lys His Asn  
 65 70 75 80  
  
 Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile Ser Met Leu Glu  
 85 90 95  
  
 Gly Ala Val Leu Asp Ile Arg Tyr Gly Val Ser Arg Ile Ala Tyr Ser  
 100 105 110  
  
 Lys Asp Phe Glu Thr Leu Lys Val Asp Phe Leu Ser Lys Leu Pro Glu  
 115 120 125  
  
 Met Leu Lys Met Phe Glu Asp Arg Leu Cys His Lys Thr Tyr Leu Asn  
 130 135 140  
  
 Gly Asp His Val Thr His Pro Asp Phe Met Leu Tyr Asp Ala Leu Asp  
 145 150 155 160  
  
 Val Val Leu Tyr Met Asp Pro Met Cys Leu Asp Ala Phe Pro Lys Leu  
 165 170 175  
  
 Val Cys Phe Lys Lys Arg Ile Glu Ala Ile Pro Gln Ile Asp Lys Tyr  
 180 185 190  
  
 Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln Gly Trp Gln Ala  
 195 200 205  
  
 Thr Phe Gly Gly Gly Asp His Pro Pro Lys Ser Asp Leu Glu Val Leu  
 210 215 220  
  
 Phe Gln Gly Pro Leu Gly Ser Gly Arg Lys Lys Arg Arg Gln Arg Arg  
 225 230 235 240  
  
 Arg Ser Arg Phe Arg Thr Lys Arg Ser Ala Leu Val Arg Arg Leu Trp  
 245 250 255

[0043]



Arg Ser Arg Ala Pro Gly Gly Glu Asp Glu Glu Glu Gly Ala Gly Gly  
 260 265 270  
 Gly Gly Gly Gly Gly Glu Leu Arg Gly Glu Gly Ala Thr Asp Ser Arg  
 275 280 285  
 Ala His Gly Ala Gly Gly Gly Gly Pro Gly Arg Ala Gly Cys Cys Leu  
 290 295 300  
 Gly Lys Ala Val Arg Gly Ala Lys Gly His His His Pro His Pro Pro  
 305 310 315 320  
 Ala Ala Gly Ala Gly Ala Ala Gly Gly Ala Glu Ala Asp Leu Lys Ala  
 325 330 335  
 Leu Thr His Ser Val Leu Lys Lys Leu Lys Glu Arg Gln Leu Glu Leu  
 340 345 350  
 Leu Leu Gln Ala Val Glu Ser Arg Gly Gly Thr Arg Thr Ala Cys Leu  
 355 360 365  
 Leu Leu Pro Gly Arg Leu Asp Cys Arg Leu Gly Pro Gly Ala Pro Ala  
 370 375 380  
 Gly Ala Gln Pro Ala Gln Pro Pro Ser Ser Tyr Ser Leu Pro Leu Leu  
 385 390 395 400  
 Leu Cys Lys Val Phe Arg Trp Pro Asp Leu Arg His Ser Ser Glu Val  
 405 410 415  
 Lys Arg Leu Cys Cys Cys Glu Ser Tyr Gly Lys Ile Asn Pro Glu Leu  
 420 425 430  
 Val Cys Cys Asn Pro His His Leu Ser Arg Leu Cys Glu Leu Glu Ser  
 435 440 445  
 Pro Pro Pro Pro Tyr Ser Arg Tyr Pro Met Asp Phe Leu Lys Pro Thr  
 450 455 460  
 Ala Asp Cys Pro Asp Ala Val Pro Ser Ser Ala Glu Thr Gly Gly Thr  
 465 470 475 480  
 Asn Tyr Leu Ala Pro Gly Gly Leu Ser Asp Ser Gln Leu Leu Leu Glu  
 485 490 495  
 Pro Gly Asp Arg Ser His Trp Cys Val Val Ala Tyr Trp Glu Glu Lys  
 500 505 510  
 Thr Arg Val Gly Arg Leu Tyr Cys Val Gln Glu Pro Ser Leu Asp Ile  
 515 520 525  
 Phe Tyr Asp Leu Pro Gln Gly Asn Gly Phe Cys Leu Gly Gln Leu Asn

[0044]

|        |                                                                    |     |     |     |
|--------|--------------------------------------------------------------------|-----|-----|-----|
|        | 530                                                                | 535 | 540 |     |
|        | Ser Asp Asn Lys Ser Gln Leu Val Gln Lys Val Arg Ser Lys Ile Gly    |     |     |     |
|        | 545                                                                | 550 | 555 | 560 |
|        | Cys Gly Ile Gln Leu Thr Arg Glu Val Asp Gly Val Trp Val Tyr Asn    |     |     |     |
|        |                                                                    | 565 | 570 | 575 |
|        | Arg Ser Ser Tyr Pro Ile Phe Ile Lys Ser Ala Thr Leu Asp Asn Pro    |     |     |     |
|        |                                                                    | 580 | 585 | 590 |
|        | Asp Ser Arg Thr Leu Leu Val His Lys Val Phe Pro Gly Phe Ser Ile    |     |     |     |
|        |                                                                    | 595 | 600 | 605 |
|        | Lys Ala Phe Asp Tyr Glu Lys Ala Tyr Ser Leu Gln Arg Pro Asn Asp    |     |     |     |
|        |                                                                    | 610 | 615 | 620 |
|        | His Glu Phe Met Gln Gln Pro Trp Thr Gly Phe Thr Val Gln Ile Ser    |     |     |     |
|        |                                                                    | 625 | 630 | 635 |
|        | Phe Val Lys Gly Trp Gly Gln Cys Tyr Thr Arg Gln Phe Ile Ser Ser    |     |     |     |
|        |                                                                    | 645 | 650 | 655 |
|        | Cys Pro Cys Trp Leu Glu Val Ile Phe Asn Ser Arg Thr Arg Thr Arg    |     |     |     |
|        |                                                                    | 660 | 665 | 670 |
| [0045] | Pro Leu Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Ala Ala Asn Asp    |     |     |     |
|        |                                                                    | 675 | 680 | 685 |
|        | Ile Leu Asp Tyr Lys Asp Asp Asp Asp Lys Val                        |     |     |     |
|        |                                                                    | 690 | 695 |     |
|        | <210> 91                                                           |     |     |     |
|        | <211> 1637                                                         |     |     |     |
|        | <212> DNA                                                          |     |     |     |
|        | <213> 人工序列                                                         |     |     |     |
|        | <220>                                                              |     |     |     |
|        | <223> 人工序列的描述: 合成多核苷酸                                              |     |     |     |
|        | <400> 91                                                           |     |     |     |
|        | gaattcgtcg actggatccg gtaccgagga gatctgccgc cgcgatcgcc atgtcccca   |     |     | 60  |
|        | ttctgggeta ctggaagatt aaggccctgg tgcagcctac tagactgctg ctggaatacc  |     |     | 120 |
|        | tggaggaaaa atatgaagag catctgtatg aaagagacga gggggataaa tggaggaaca  |     |     | 180 |
|        | agaaattcga actgggactg gagtttcta atctgccata ctatattgac ggcgatgtga   |     |     | 240 |
|        | agctgactca gtctatggct atcattagat acatgcaga caaacacaac atgtgggcg    |     |     | 300 |
|        | ggtgtcctaa ggaaggcca gagattagta tgctggagg agccgtgctg gatattagat    |     |     | 360 |
|        | acggcgtctc acgcategcc tatagcaaag acttcgaaac cctgaagggtg gattttctga |     |     | 420 |
|        | gcaaactgcc tgaaatgctg aagaigtctg aggacagact gtgccacaaa acctacctga  |     |     | 480 |
|        | atggcgacca tgtcacacac ccagatttta tgctgtacga cgcctggat gtggctctgt   |     |     | 540 |
|        | atatggacce catgtgtctg gatgctttcc ctaagctggt gtgctttaag aaaaggatcg  |     |     | 600 |

|                                                                   |      |
|-------------------------------------------------------------------|------|
| aggcaattcc ccagatcgat aagtacctga aaagctccaa gtatatcgct tggcctctcc | 660  |
| agggctggca ggcaacattc ggaggcgggg accatcccc taaaagcgac ctggagggtgc | 720  |
| tgtttcaggg accactgggc agcgccgga agaagcggcg gcagaggcgc cgaagtaggt  | 780  |
| tccgcactaa gcggtcagca ctggtgcgga gactgtggcg atctcgggct cctggaggag | 840  |
| aggacgagga agaggagca ggcggcggcg gaggaggagg agaactgcgc ggggagggag  | 900  |
| ctacagatag ccgagccac ggagctggag gaggaggacc agggcgagcc ggatgctgtc  | 960  |
| tgggcaaagc agtgagaggc gccaaagggc accatcacc ccatccacc gccgctggcg   | 1020 |
| caggagcagc cggcggagct gaggcagacc tgaaagccct gactcacagt gtgctgaaga | 1080 |
| aactgaagga aagacagctg gagctgctgc tccaggcagt cgaatcacgc ggaggcacc  | 1140 |
| gaacagcttg tctgctgctg ccggccggc tggactgccg gctgggaccc ggcgccccctg | 1200 |
| ctggggcaca gccagccag cctecateta gttatagcct gccctgctg ctgtgtaagg   | 1260 |
| tgttcgatg gcctgatctg cggcattcaa gcgaagtcaa aaggctgtgc tgttcgagt   | 1320 |
| cctacggcaa gattaacca gaactggtgt gttgcaatcc ccatcacctg tctcgactgt  | 1380 |
| gtgaactgga gtccccctt ccacctact ctaggtatcc tatggacttt ctgaagccaa   | 1440 |
| ccgtgactg cccagatgca gtgccctect ctgccgagac tgggggaacc aactacctgg  | 1500 |
| ctctggcgg actgagcag tcccagctgc tgetggaacc aggggategg cggccgctcg   | 1560 |
| agcagaaact catctcagaa gaggatctgg cagcaaatga tctctggat tacaaggatg  | 1620 |
| acgacgataa ggttttaa                                               | 1637 |

[0046]

&lt;210&gt; 92

&lt;211&gt; 528

&lt;212&gt; PRT

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 人工序列的描述: 合成多肽

&lt;400&gt; 92

Met Ser Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly Leu Val Gln Pro  
 1 5 10 15

Thr Arg Leu Leu Leu Glu Tyr Leu Glu Glu Lys Tyr Glu Glu His Leu  
 20 25 30

Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys Lys Phe Glu Leu  
 35 40 45

Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp Gly Asp Val Lys  
 50 55 60

Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala Asp Lys His Asn  
 65 70 75 80

Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile Ser Met Leu Glu  
 85 90 95

[0047]

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

Gly Ala Gln Pro Ala Gln Pro Pro Ser Ser Tyr Ser Leu Pro Leu Leu  
 385 390 395 400  
 Leu Cys Lys Val Phe Arg Trp Pro Asp Leu Arg His Ser Ser Glu Val  
 405 410 415  
 Lys Arg Leu Cys Cys Cys Glu Ser Tyr Gly Lys Ile Asn Pro Glu Leu  
 420 425 430  
 Val Cys Cys Asn Pro His His Leu Ser Arg Leu Cys Glu Leu Glu Ser  
 435 440 445  
 Pro Pro Pro Pro Tyr Ser Arg Tyr Pro Met Asp Phe Leu Lys Pro Thr  
 450 455 460  
 Ala Asp Cys Pro Asp Ala Val Pro Ser Ser Ala Glu Thr Gly Gly Thr  
 465 470 475 480  
 Asn Tyr Leu Ala Pro Gly Gly Leu Ser Asp Ser Gln Leu Leu Leu Glu  
 485 490 495  
 Pro Gly Asp Arg Arg Pro Leu Glu Gln Lys Leu Ile Ser Glu Glu Asp  
 500 505 510  
 Leu Ala Ala Asn Asp Ile Leu Asp Tyr Lys Asp Asp Asp Asp Lys Val  
 515 520 525

[0048]

<210> 93  
 <211> 1370  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成多核苷酸

<400> 93  
 gaattcgctg actggatccg gtaccgagga gatctgccgc cgcgatcgcc atgtccccc 60  
 ttctgggcta ctggaagatt aaggccctgg tgcagcctac tagactgctg ctggaatacc 120  
 tggaggaaaa atatgaagag catctgtatg aaagagacga gggggataaa tggaggaaca 180  
 agaaattcga actgggactg gagtttccta atctgccata ctatattgac ggcgatgtga 240  
 agctgactca gtctatggct atcattagat acatcgcaga caaacacaac atgctgggcg 300  
 ggtgtcctaa ggaaaggcca gagattagta tcttgagggg agccgtgctg gatattagat 360  
 acggcgtctc acgcategcc tatagcaaag acttcgaaac cctgaagggtg gattttctga 420  
 gcaaactgcc tgaaatgctg aagatgttcg aggacagact gtgccacaaa acctacctga 480  
 atggcgacca tgtcacacac ccagatttta tgctgtacga cgccctggat gtggctcctgt 540  
 atatggacce catgtgtctg gatgctttcc ctaagctggg gtgctttaag aaaaggatcg 600  
 aggcaattcc ccagatcgat aagtacctga aaagctccaa gtatatcgct tggcctctcc 660  
 agggctggca ggcaacattc ggaggcgggg accatcccc taaaagcgac ctggaggtgc 720  
 tgtttcaggg accactgggc agcgcccgga agaagcggcg gcagaggcgc agccactggt 780

```

gtgtggtcgc ctactgggaa gagaagacaa gagtgggaag gctgtattgc gtccaggagc 840
cttccctgga catctttctac gatctgccac aggggaatgg attttgtctg ggccagctga 900
actctgacaa taagagtcag ctggigcaga aagtccggag caagattggc tgcggcatcc 960
agctgaccag ggaggtggac ggcgtgtggg tctacaaccg cagttcatat ccaatcttca 1020
tcaagagcgc cactctggac aatcccgatt cccgcaccct gctggtgcat aaggtcttcc 1080
ccggetttag catcaaggcc ttgactacg agaaggetta tagtctccag cggcccaacg 1140
atcacgagtt catgcagcag ccttgacag gcttactgt gcagatcagc ttctcaagg 1200
gatggggcca gtgctacaca aggcagttca tctcaagctg tcctgttgg ctggaagtca 1260
ttttcaatag taggacgcgt acgcggccgc tcgacagaa actcatctca gaagaggatc 1320
tggcagcaaa tgatatctg gattacaagg atgacgacga taaggtttaa 1370

```

&lt;210&gt; 94

&lt;211&gt; 440

&lt;212&gt; PRT

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 人工序列的描述: 合成多肽

&lt;400&gt; 94

```

Met Ser Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly Leu Val Gln Pro
1 5 10 15

```

[0049]

```

Thr Arg Leu Leu Leu Glu Tyr Leu Glu Lys Tyr Glu Glu His Leu
 20 25 30

```

```

Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys Lys Phe Glu Leu
 35 40 45

```

```

Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp Gly Asp Val Lys
 50 55 60

```

```

Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala Asp Lys His Asn
65 70 75 80

```

```

Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile Ser Met Leu Glu
 85 90 95

```

```

Gly Ala Val Leu Asp Ile Arg Tyr Gly Val Ser Arg Ile Ala Tyr Ser
 100 105 110

```

```

Lys Asp Phe Glu Thr Leu Lys Val Asp Phe Leu Ser Lys Leu Pro Glu
 115 120 125

```

```

Met Leu Lys Met Phe Glu Asp Arg Leu Cys His Lys Thr Tyr Leu Asn
 130 135 140

```

```

Gly Asp His Val Thr His Pro Asp Phe Met Leu Tyr Asp Ala Leu Asp
 145 150 155 160

```

Val Val Leu Tyr Met Asp Pro Met Cys Leu Asp Ala Phe Pro Lys Leu  
 165 170 175  
 Val Cys Phe Lys Lys Arg Ile Glu Ala Ile Pro Gln Ile Asp Lys Tyr  
 180 185 190  
 Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln Gly Trp Gln Ala  
 195 200 205  
 Thr Phe Gly Gly Gly Asp His Pro Pro Lys Ser Asp Leu Glu Val Leu  
 210 215 220  
 Phe Gln Gly Pro Leu Gly Ser Gly Arg Lys Lys Arg Arg Gln Arg Arg  
 225 230 235 240  
 Arg Ser His Trp Cys Val Val Ala Tyr Trp Glu Glu Lys Thr Arg Val  
 245 250 255  
 Gly Arg Leu Tyr Cys Val Gln Glu Pro Ser Leu Asp Ile Phe Tyr Asp  
 260 265 270  
 Leu Pro Gln Gly Asn Gly Phe Cys Leu Gly Gln Leu Asn Ser Asp Asn  
 275 280 285  
 Lys Ser Gln Leu Val Gln Lys Val Arg Ser Lys Ile Gly Cys Gly Ile  
 290 295 300  
 Gln Leu Thr Arg Glu Val Asp Gly Val Trp Val Tyr Asn Arg Ser Ser  
 305 310 315 320  
 Tyr Pro Ile Phe Ile Lys Ser Ala Thr Leu Asp Asn Pro Asp Ser Arg  
 325 330 335  
 Thr Leu Leu Val His Lys Val Phe Pro Gly Phe Ser Ile Lys Ala Phe  
 340 345 350  
 Asp Tyr Glu Lys Ala Tyr Ser Leu Gln Arg Pro Asn Asp His Glu Phe  
 355 360 365  
 Met Gln Gln Pro Trp Thr Gly Phe Thr Val Gln Ile Ser Phe Val Lys  
 370 375 380  
 Gly Trp Gly Gln Cys Tyr Thr Arg Gln Phe Ile Ser Ser Cys Pro Cys  
 385 390 395 400  
 Trp Leu Glu Val Ile Phe Asn Ser Arg Thr Arg Thr Arg Pro Leu Glu  
 405 410 415  
 Gln Lys Leu Ile Ser Glu Glu Asp Leu Ala Ala Asn Asp Ile Leu Asp  
 420 425 430  
 Tyr Lys Asp Asp Asp Asp Lys Val  
 435 440

[0050]

<210> 95  
 <211> 66  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成多肽

<400> 95  
 Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Glu Leu Glu Ser Pro Pro  
 1 5 10 15

Pro Pro Tyr Ser Arg Tyr Pro Met Asp Phe Leu Lys Pro Thr Ala Asp  
 20 25 30

Cys Pro Asp Ala Val Pro Ser Ser Ala Glu Thr Gly Gly Thr Asn Tyr  
 35 40 45

Leu Ala Pro Gly Gly Leu Ser Asp Ser Gln Leu Leu Leu Glu Pro Gly  
 50 55 60

Asp Arg  
 65

<210> 96  
 <211> 249  
 <212> DNA  
 <213> 人工序列

[0051] <220>  
 <223> 人工序列的描述: 合成多核苷酸

<400> 96  
 ggccgtaaaa aacgccgtca acgccgccgt gaactggaga gcccgccgcc gccgtatagc 60  
 cgttacccga tggatttcct gaaaccgact gcagattgcc cggacgcagt cccgagcagc 120  
 gctgagaccg gcggcaccaa ctatctggca ccgggcggtc tgagcgatag ccagctgctg 180  
 ctggaaccgg gcgaccgtgg taagcctatc cctaaccctc tctcgggtct cgattctacg 240  
 tgagtcgac 249

<210> 97  
 <211> 285  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成多核苷酸

<400> 97  
 ggccggaaga agcggcgga gagcgccga gaactggagt cccccctcc accctactct 60  
 aggtatccta tggactttct gaagccaacc gctgactgcc cagatgcagt gccctctctt 120  
 gccgagactg ggggaaccaa ctacctggct cctggcggac tgagcgactc ccagctgctg 180  
 ctggaaccag gggatcggcg gccgctcgag cagaaactca tctcagaaga ggatctggca 240  
 gcaaatgata tctggatta caaggatgac gacgataagg tttaa 285

<210> 98



<211> 25  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成肽

<400> 98  
 Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Glu Leu Glu Ser Pro Pro  
 1 5 10 15

Pro Pro Tyr Ser Arg Tyr Pro Met Asp  
 20 25

<210> 99  
 <211> 126  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成多核苷酸

<400> 99  
 ggccgtaaaa aacgccgtca acgccgccgt gaactggaga gcccgccgcc gccgtatagc 60  
 cgttaccgga tggatggtaa gcctatccct aacctctctc tcggtctcga ttctacgtga 120  
 gtcgac 126

[0052]

<210> 100  
 <211> 162  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成多核苷酸

<400> 100  
 ggccggaaga agcgccggca gagcgccga gaactggagt cccccctcc accctactct 60  
 aggtatccta tggaccggcc gctcgagcag aaactcatct cagaagagga tctggcagca 120  
 aatgatatcc tggattacaa ggatgacgac gataaggttt aa 162

<210> 101  
 <211> 13  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 人工序列的描述: 合成肽

<400> 101  
 Gly Leu Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser  
 1 5 10

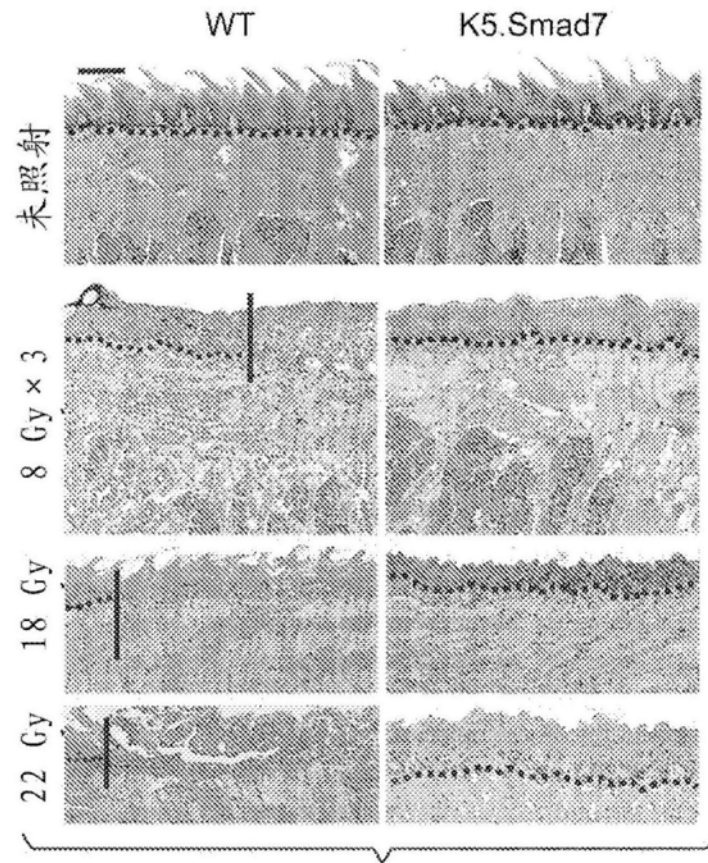


图1A

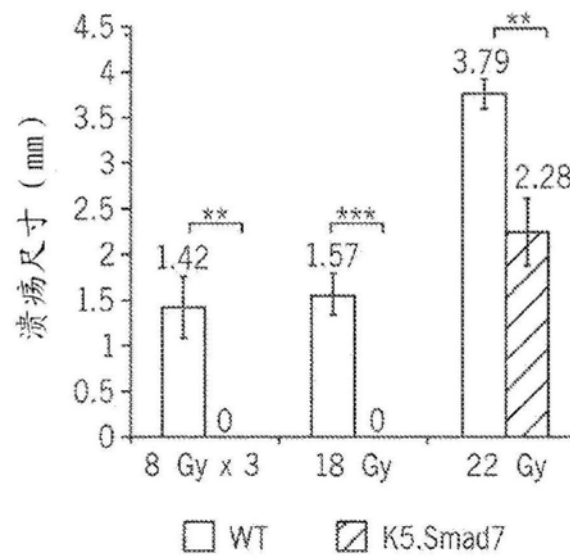


图1B

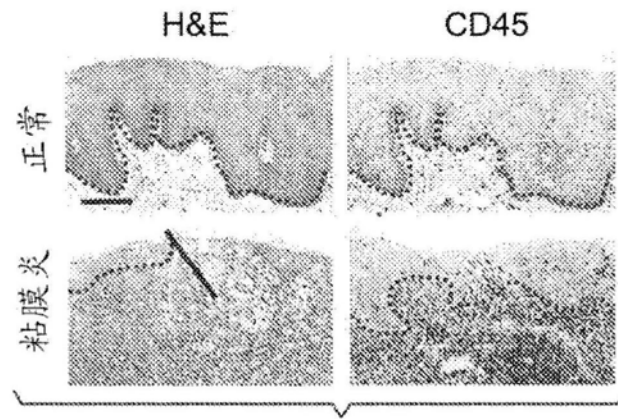


图1C

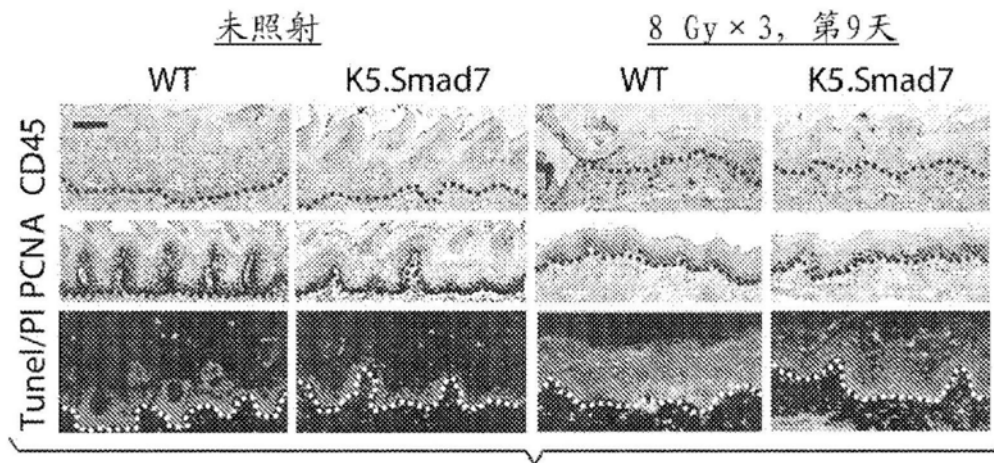


图1D

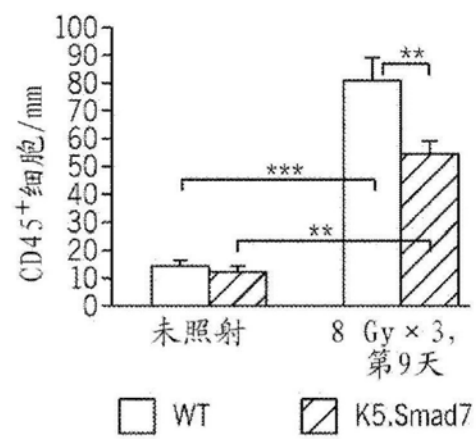


图1E

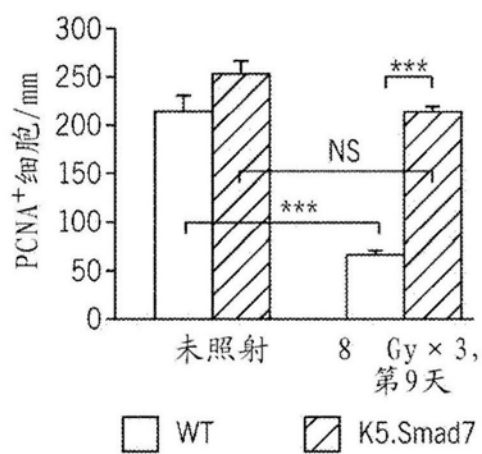


图1F

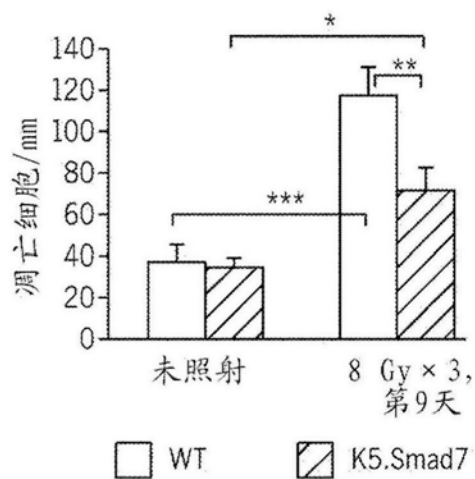


图1G

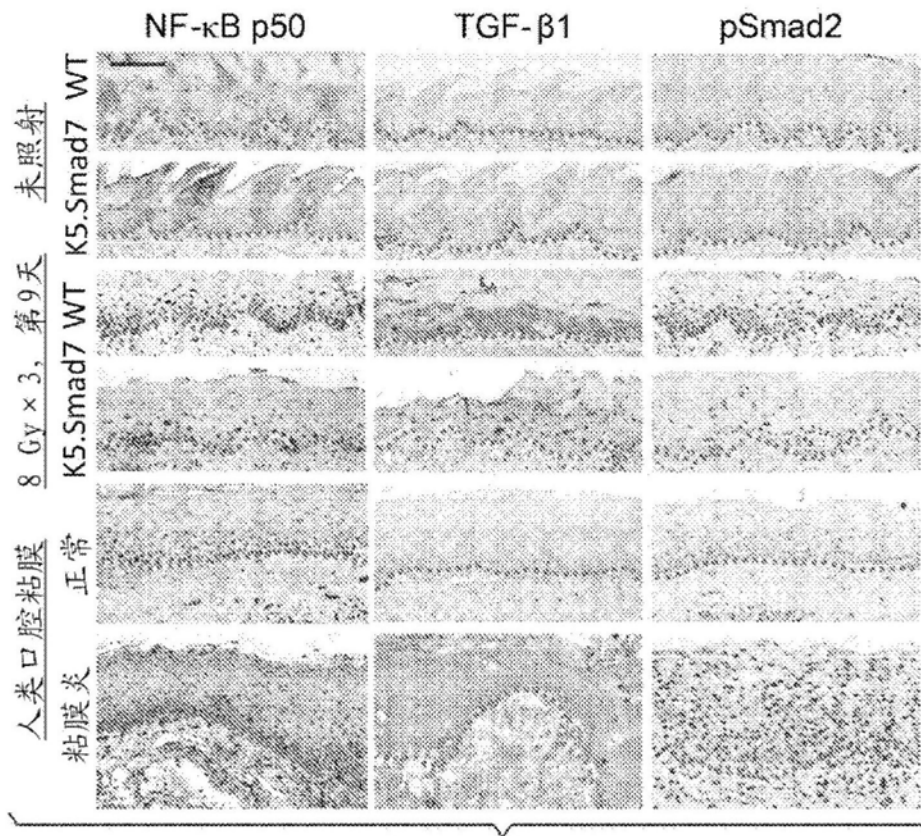


图2A

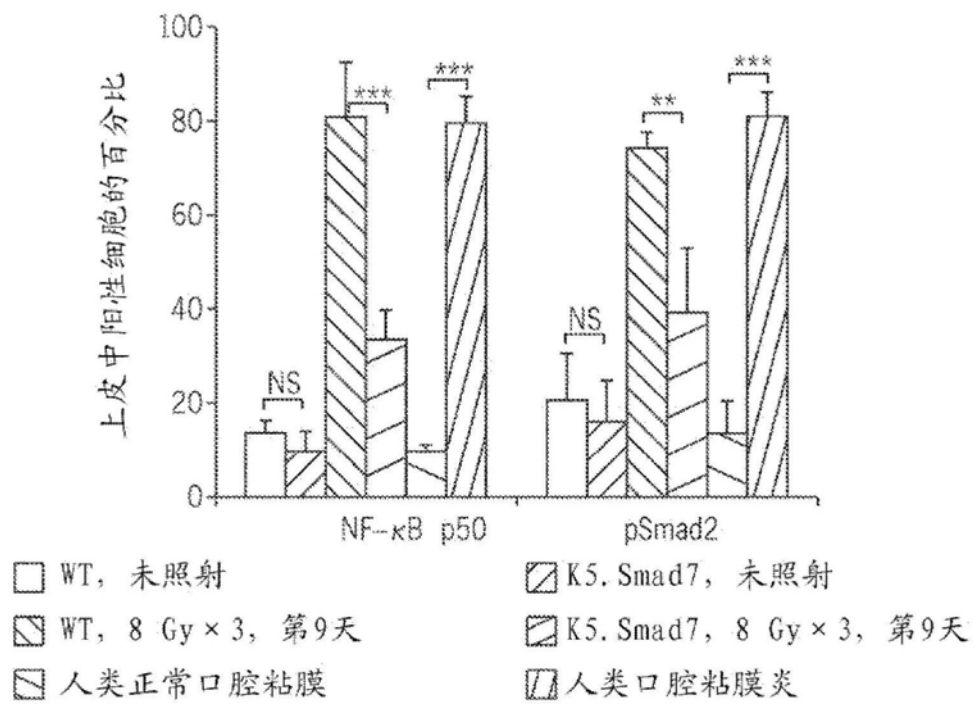


图2B

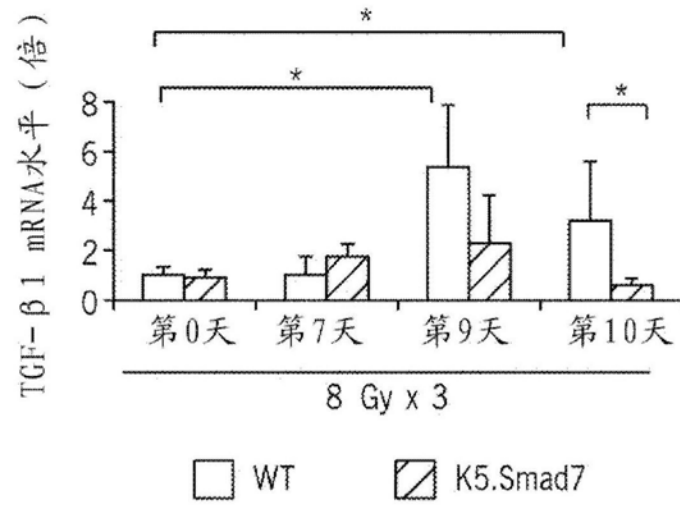


图2C

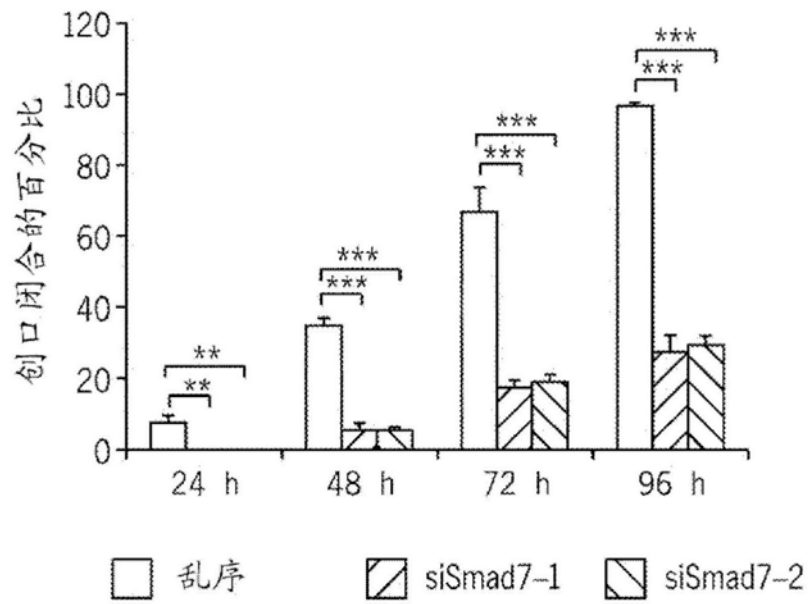


图2D

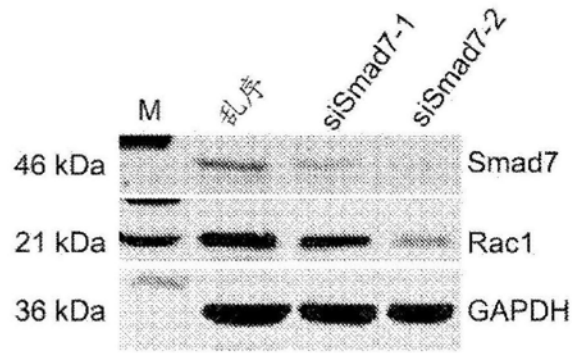


图2E

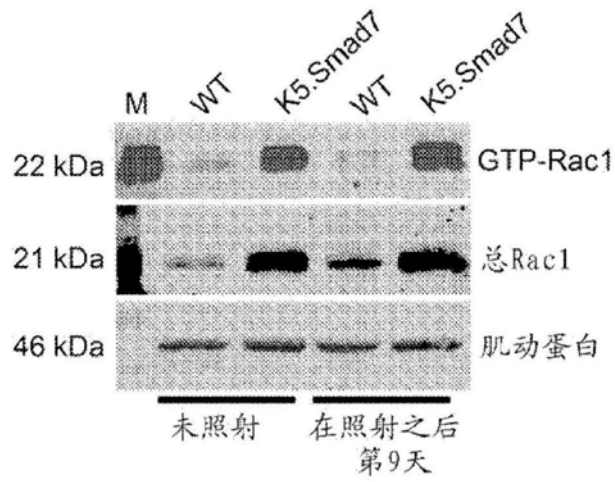


图2F

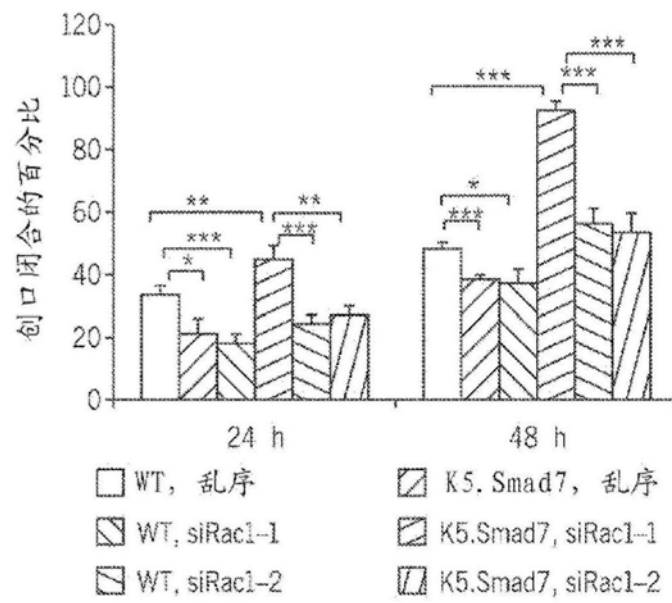


图2G

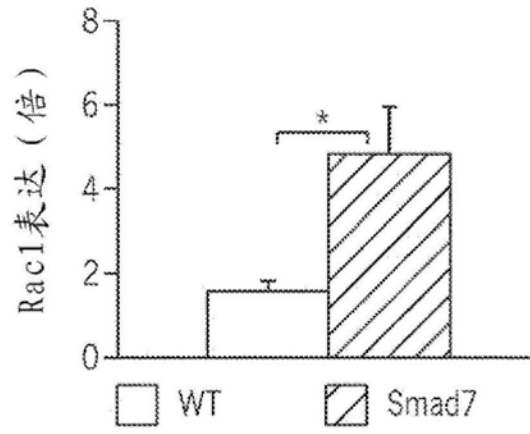


图3A

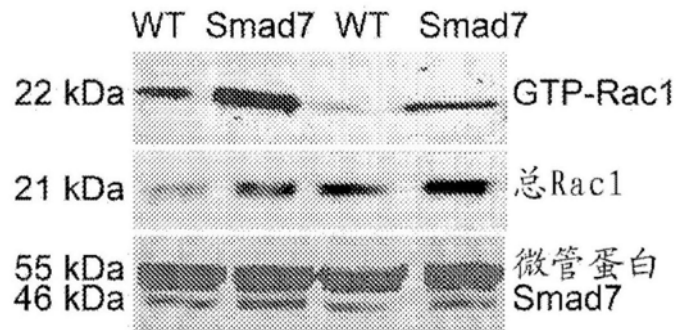


图3B

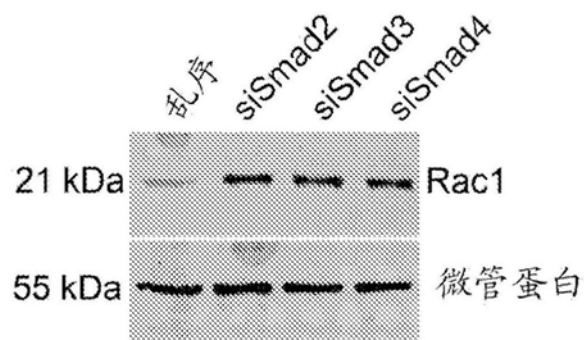


图3C



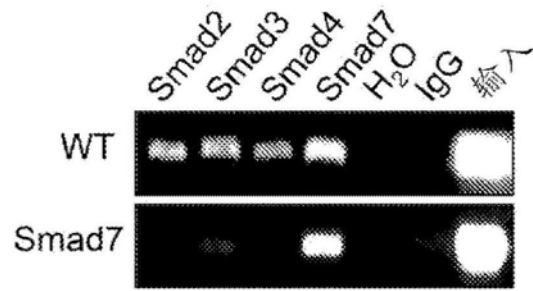


图3D

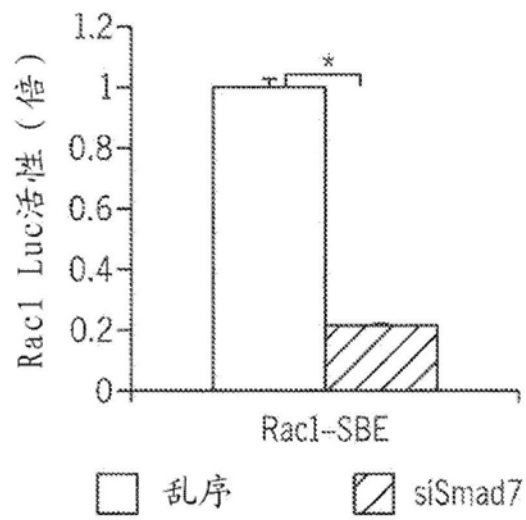


图3E

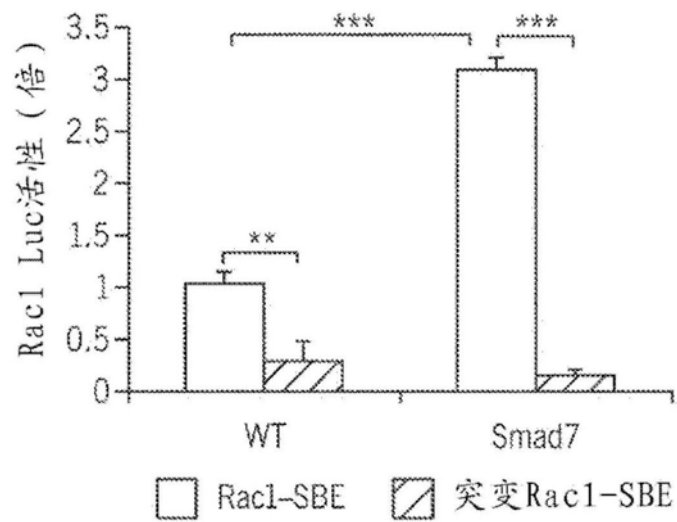


图3F

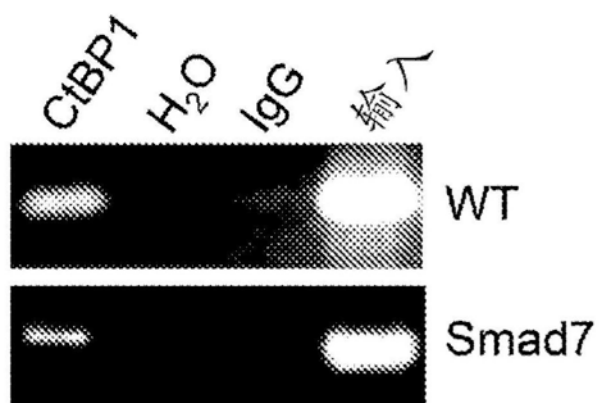


图3G

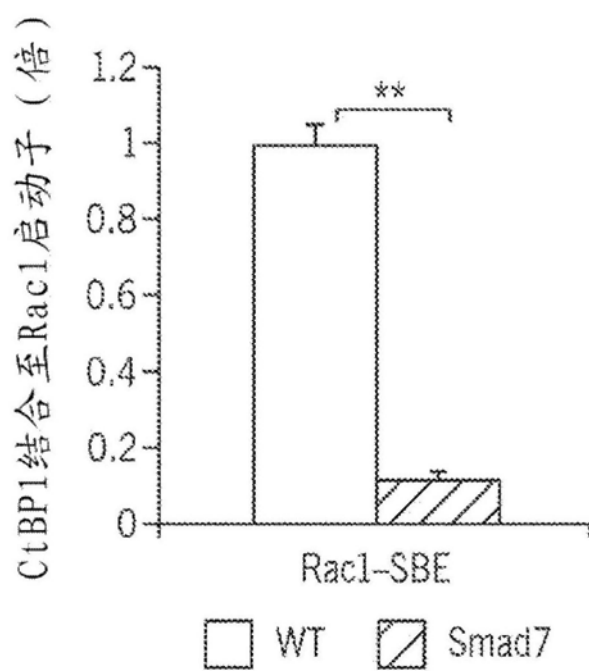


图3H

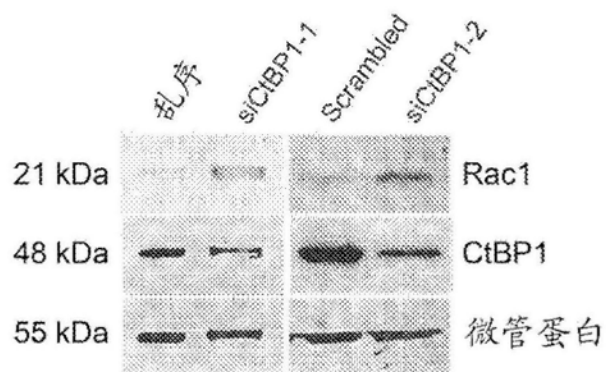


图4A

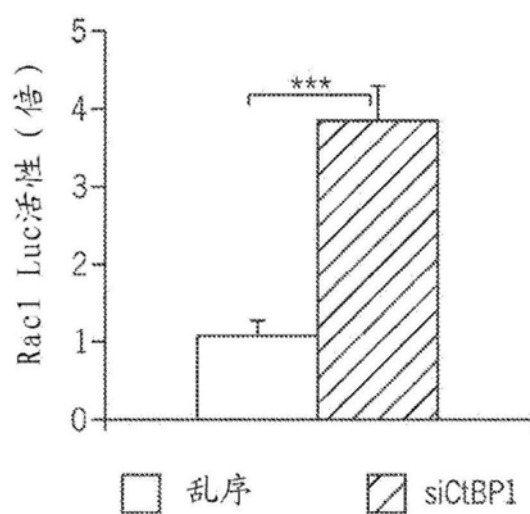


图4B

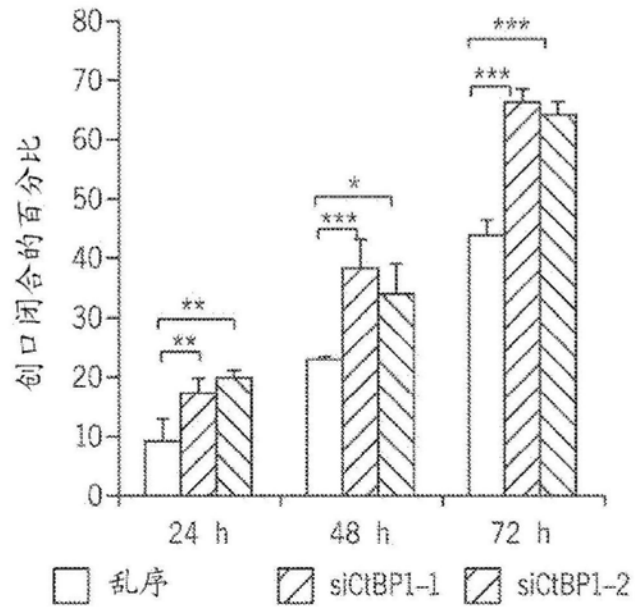


图4C

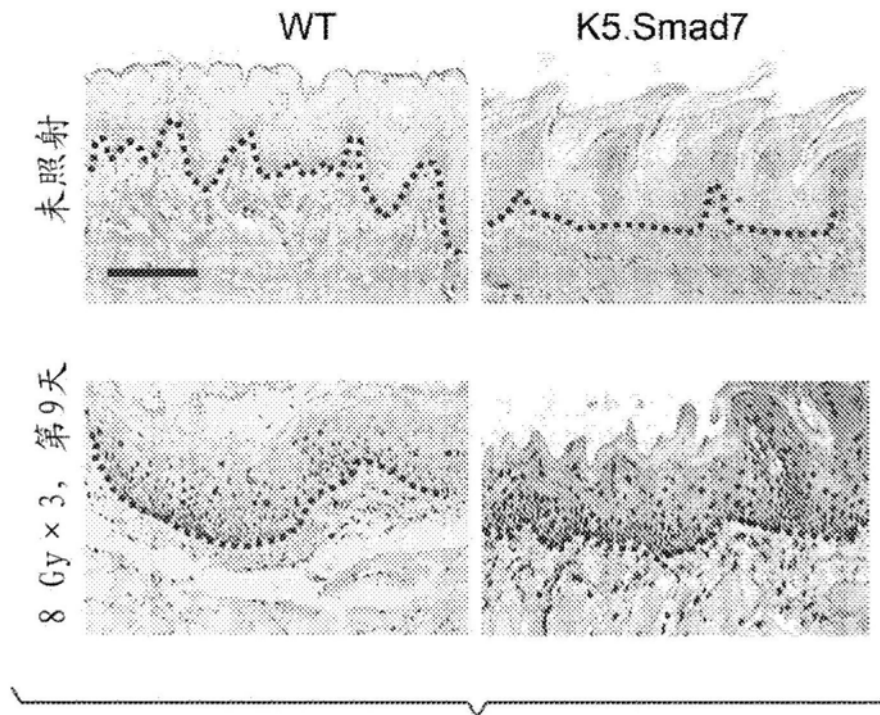


图4D

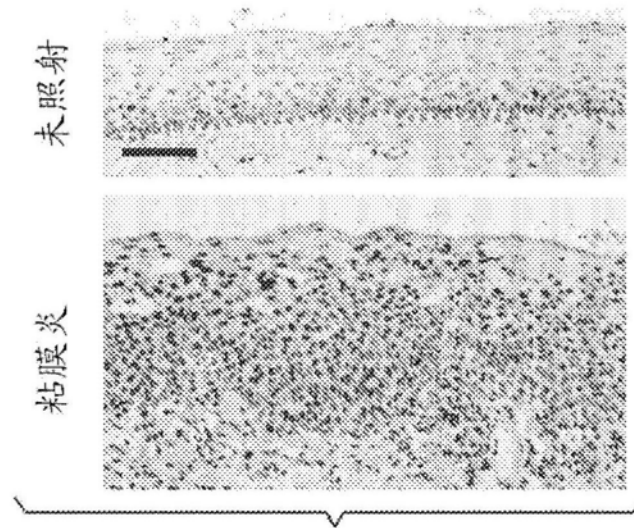


图4E

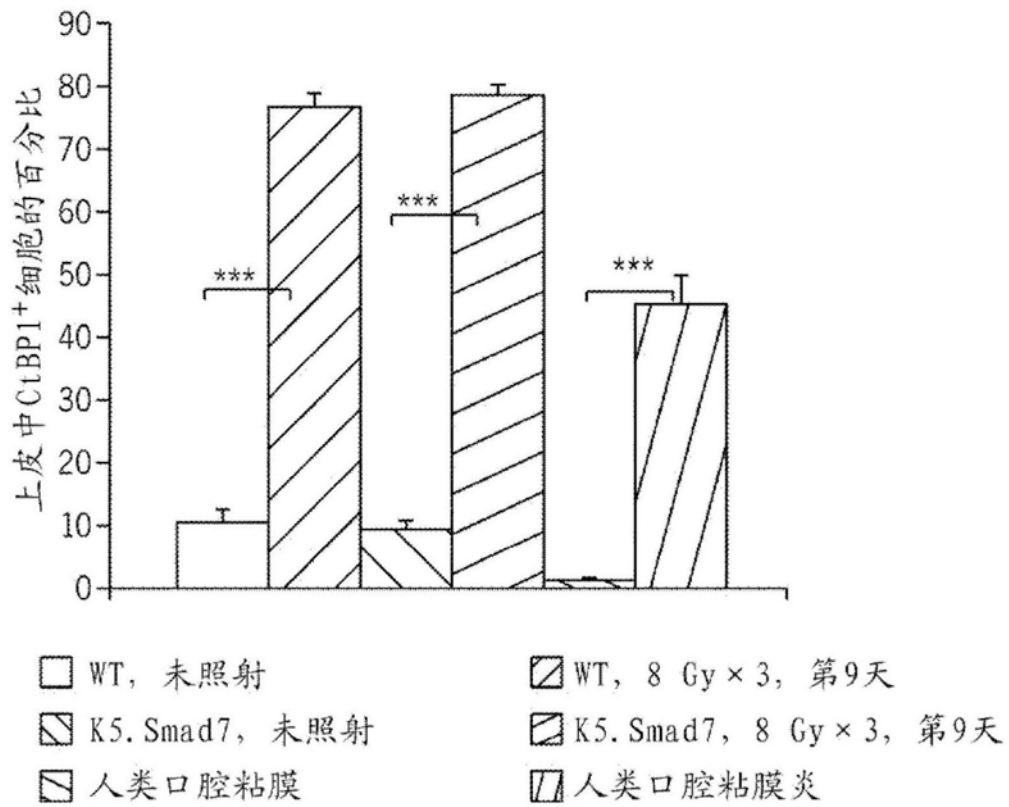


图4F

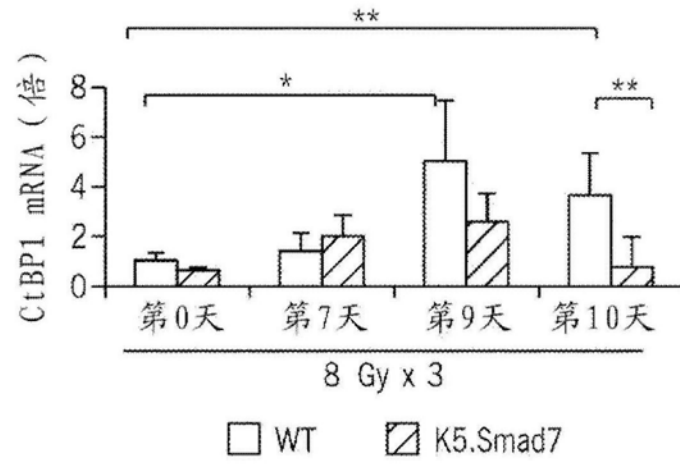


图4G

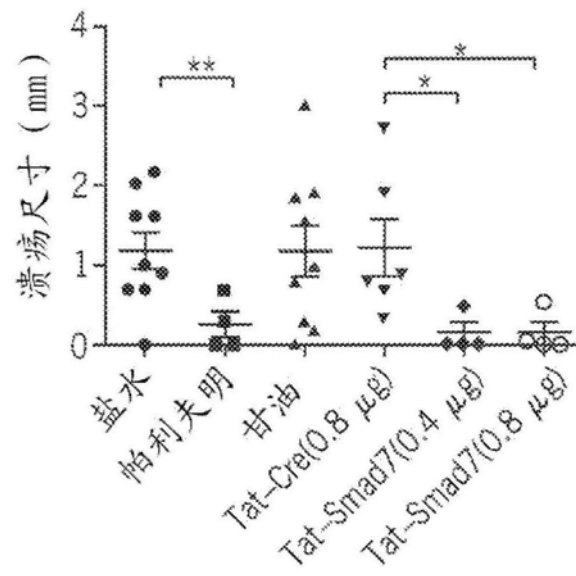


图5A

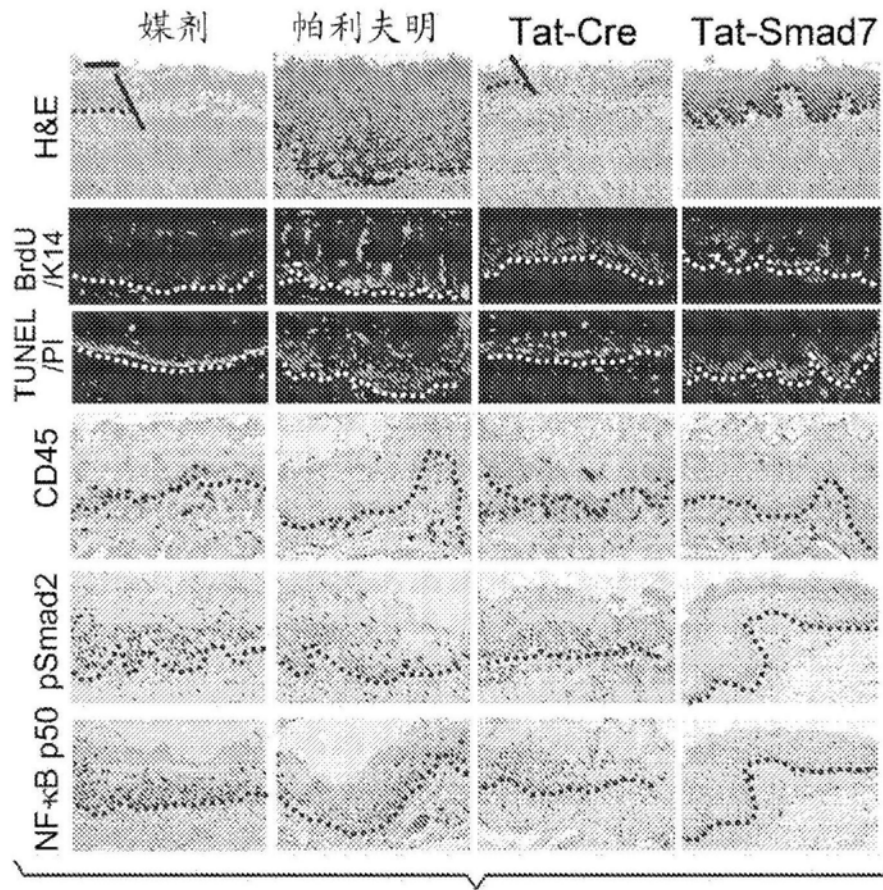


图5B

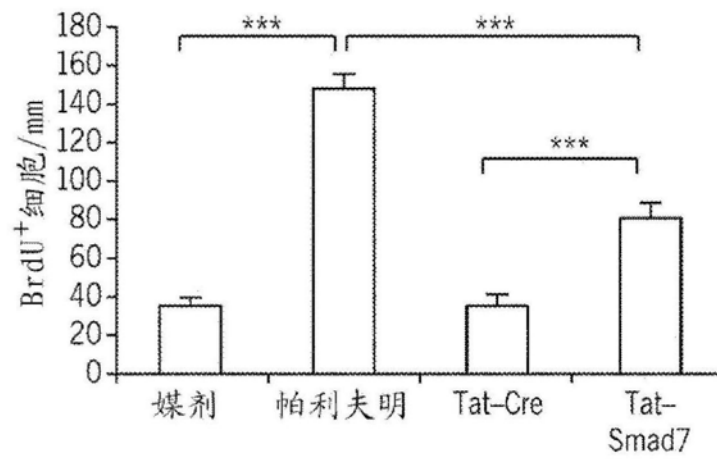


图5C

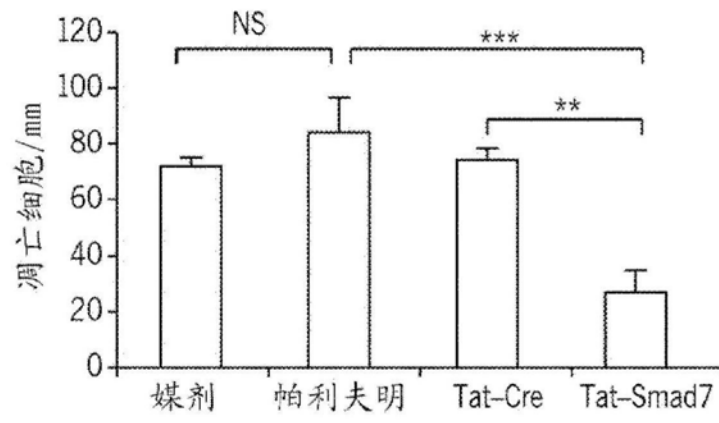


图5D

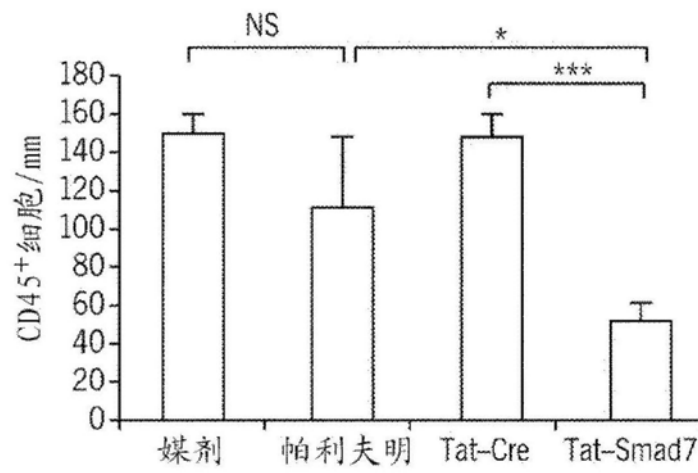


图5E

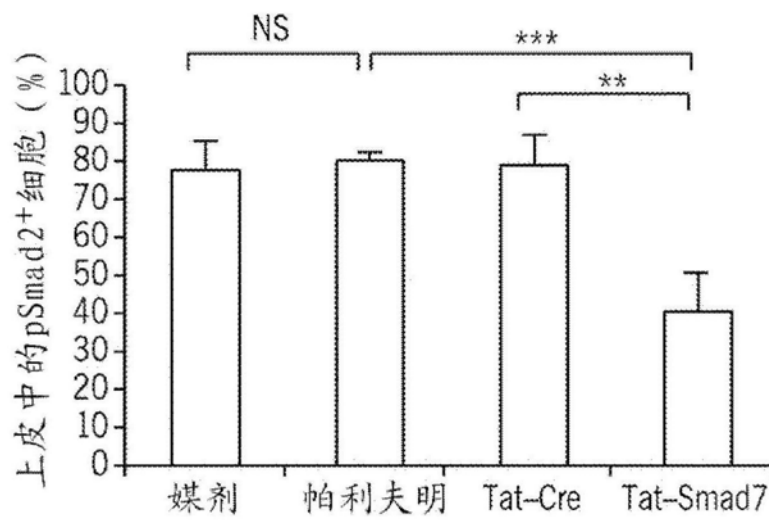


图5F



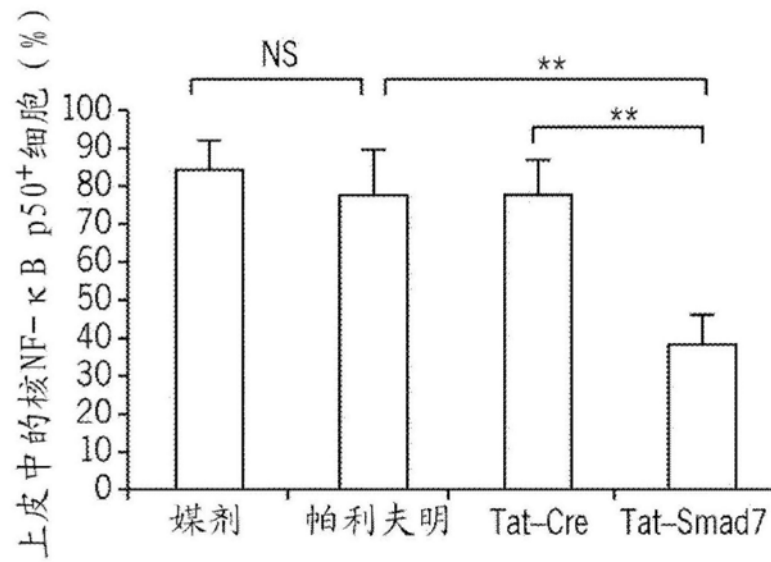


图5G

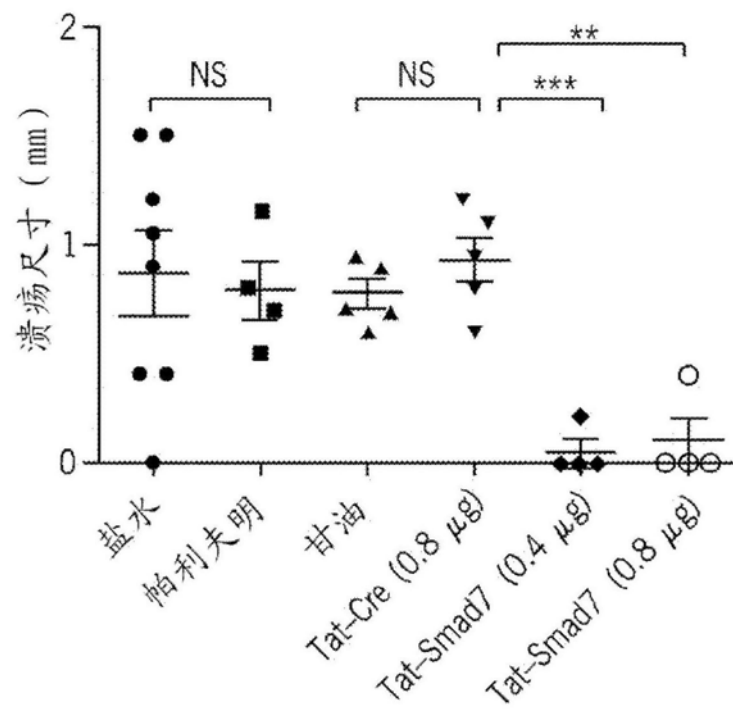


图6A

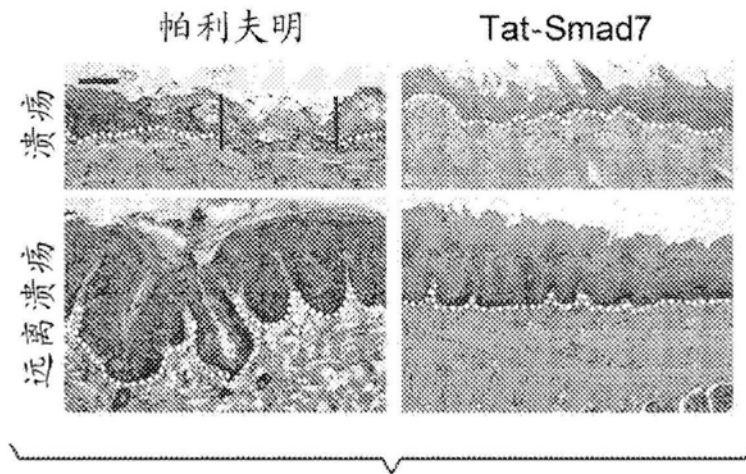


图6B

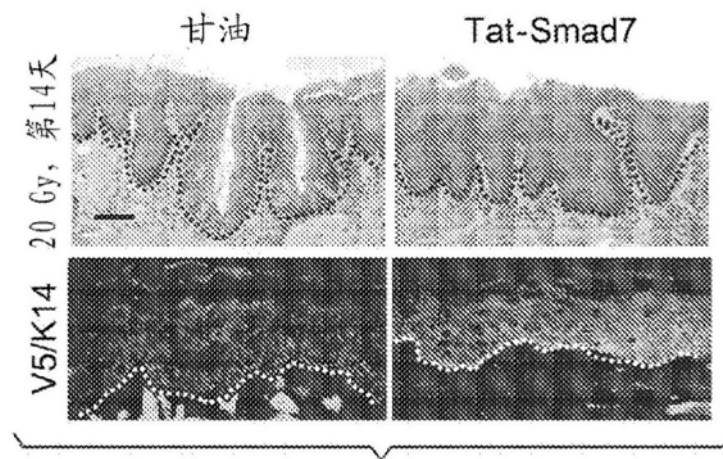


图6C

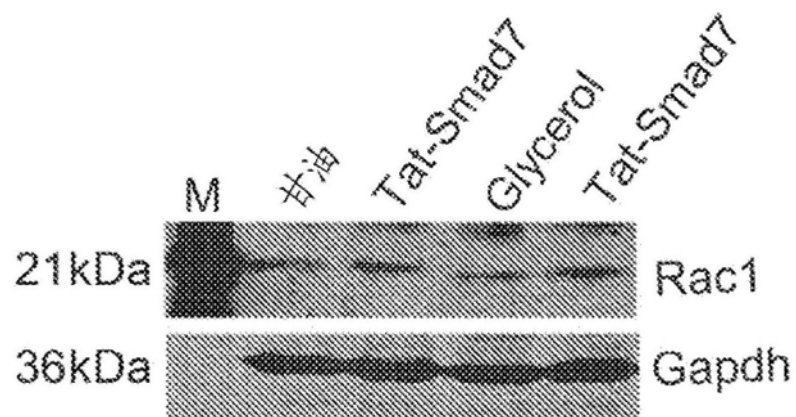


图6D

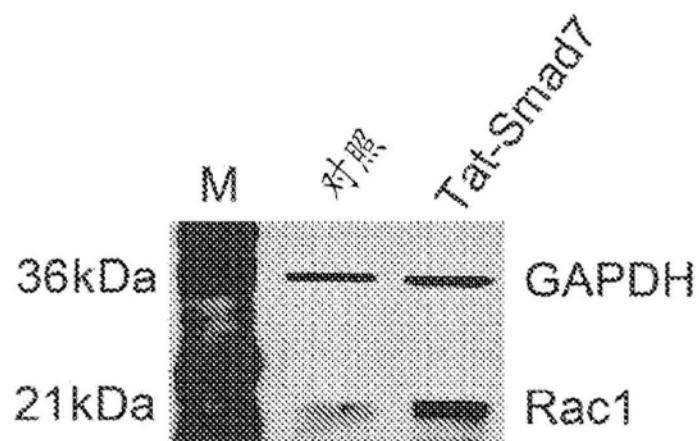


图6E

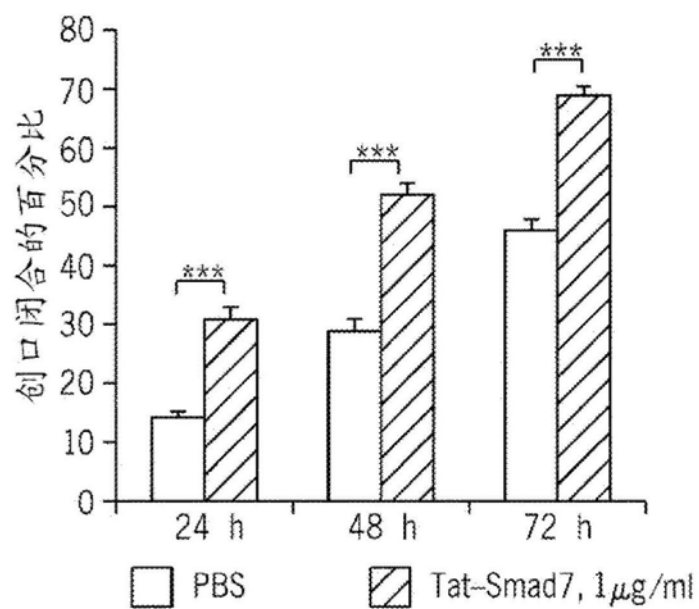


图6F

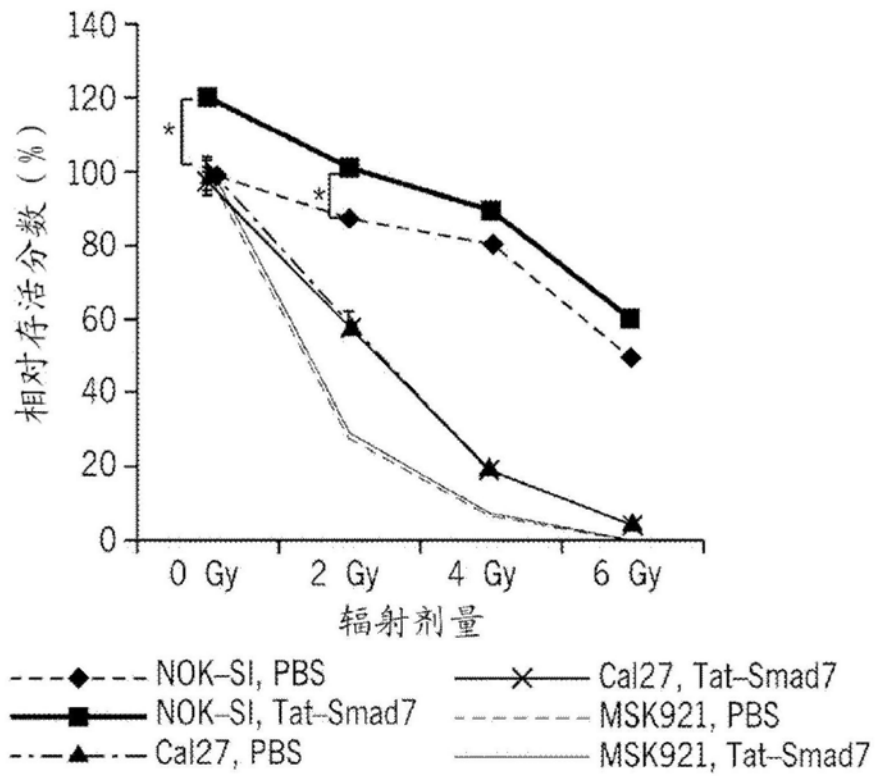


图6G

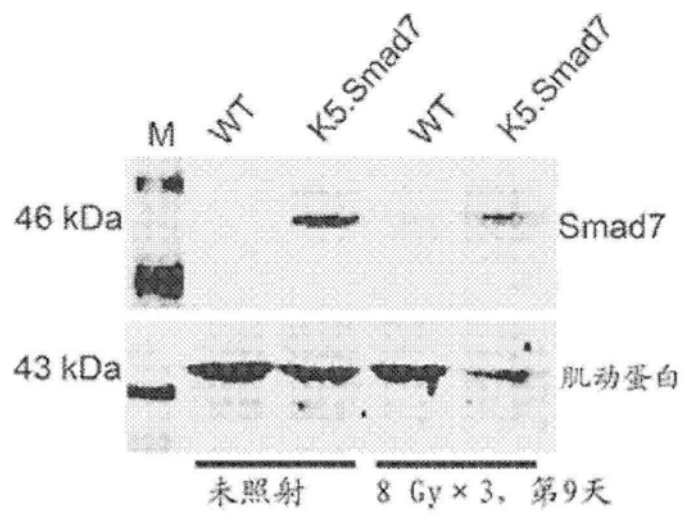


图7A

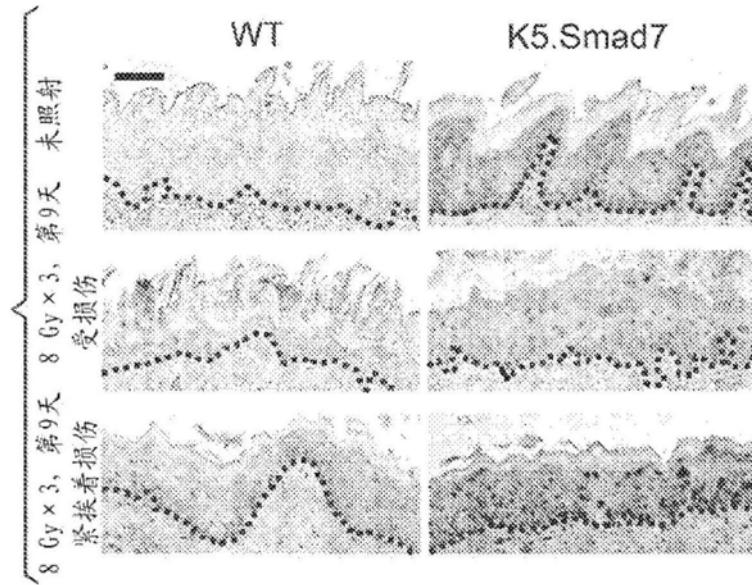


图7B

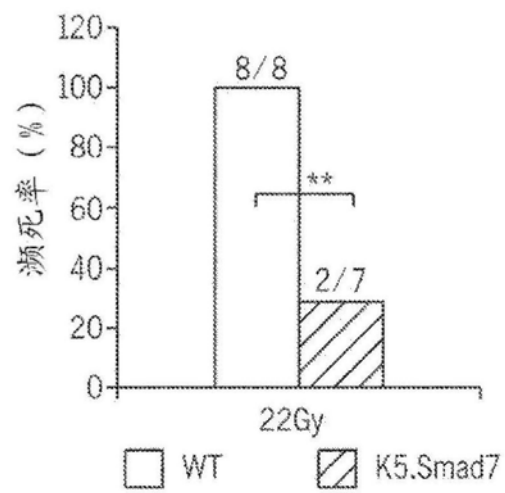


图7C

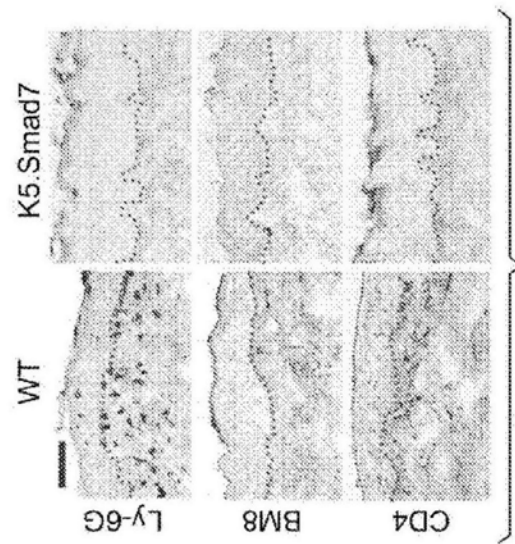


图7D

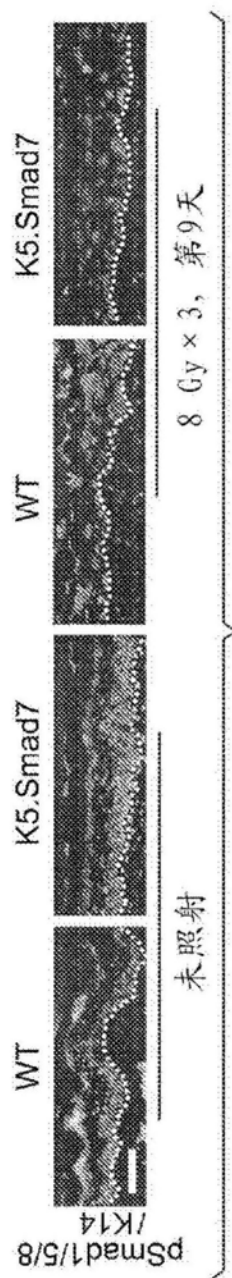


图7E

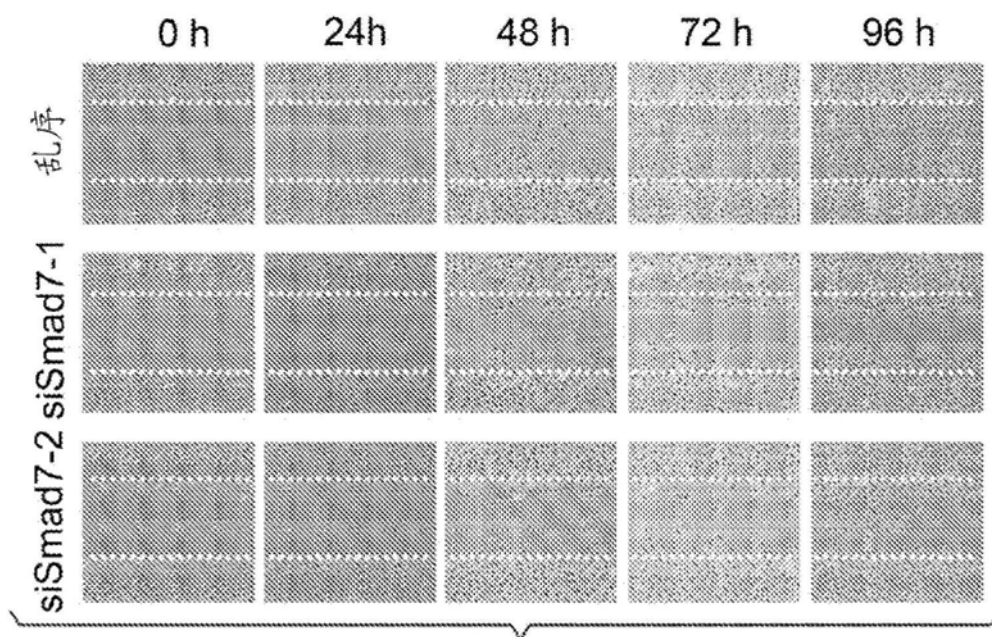


图8A

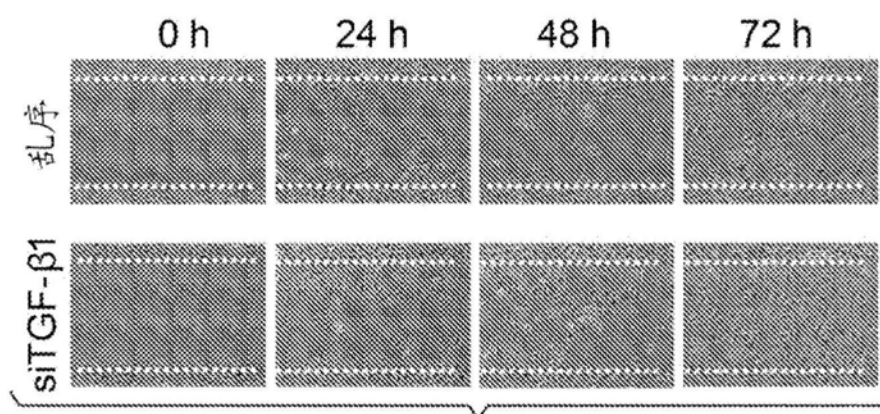


图8B



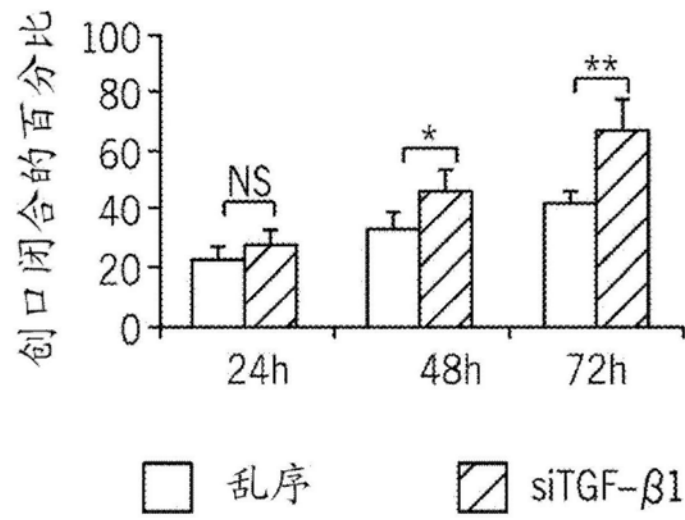


图8C

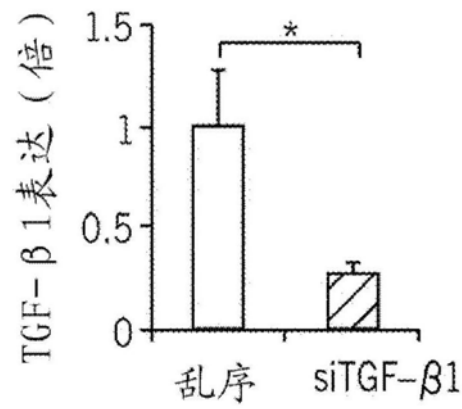


图8D

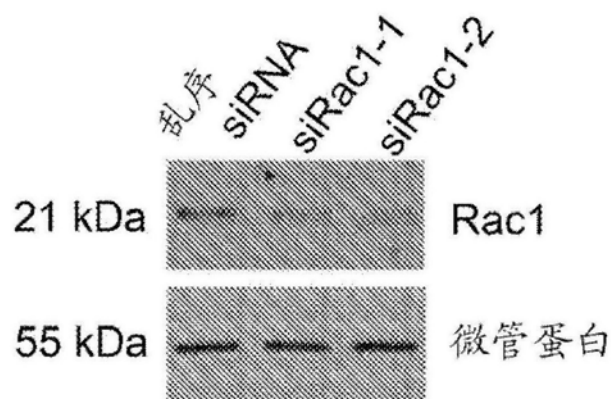


图9A

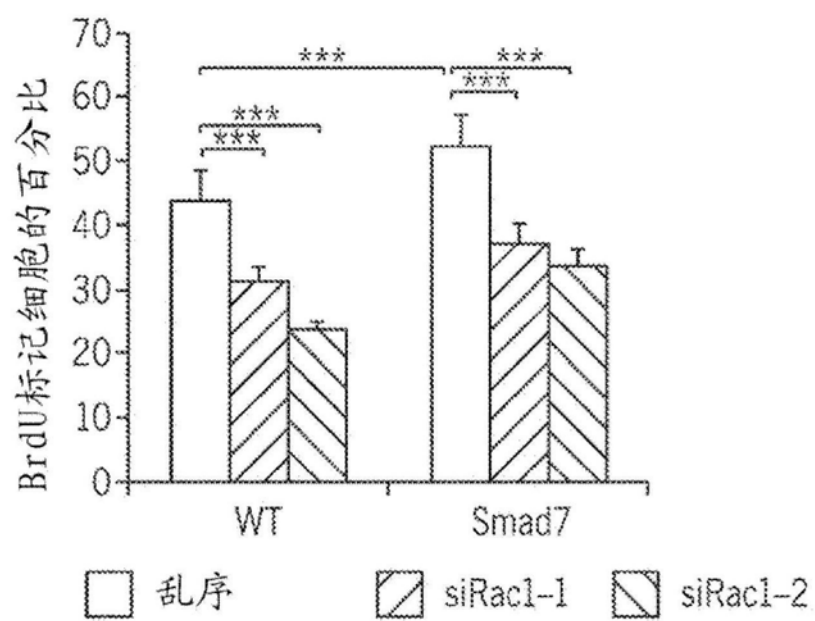


图9B

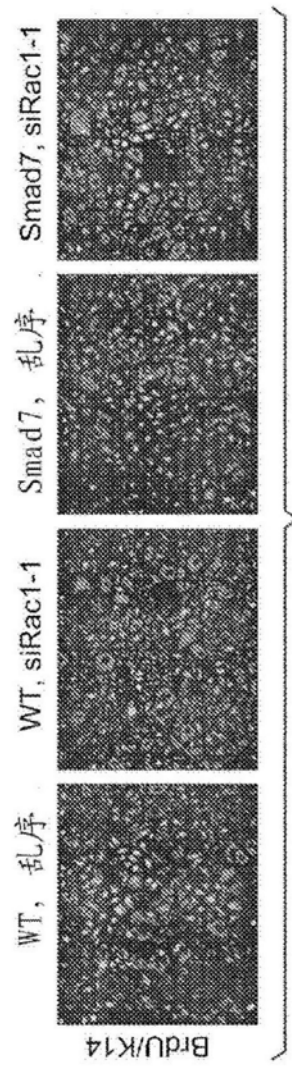


图9C

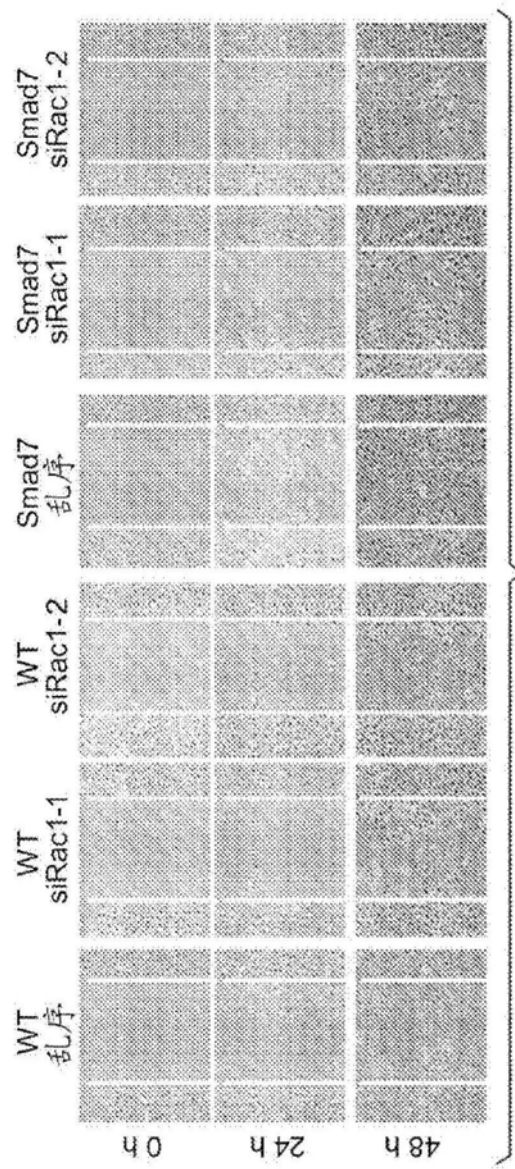


图9D

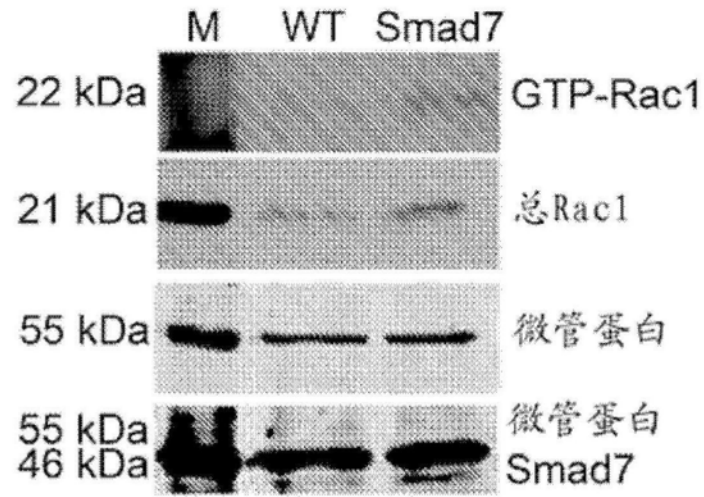


图10A

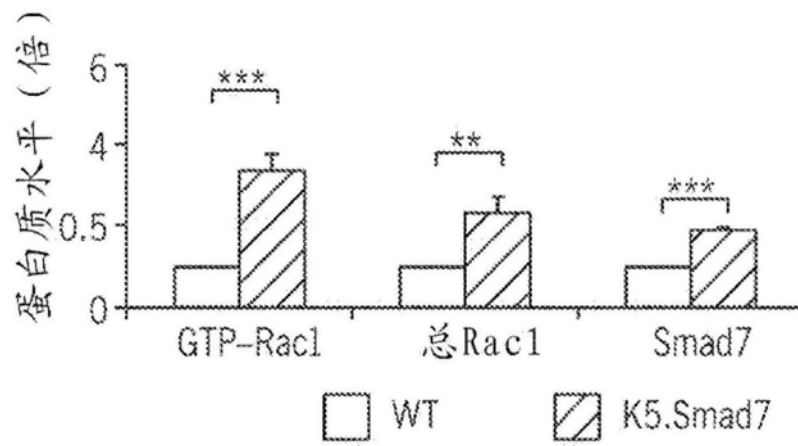


图10B

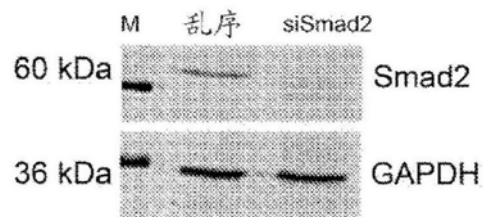


图10C

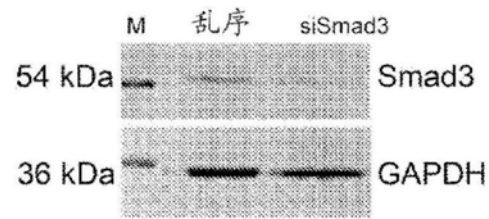


图10D



图10E

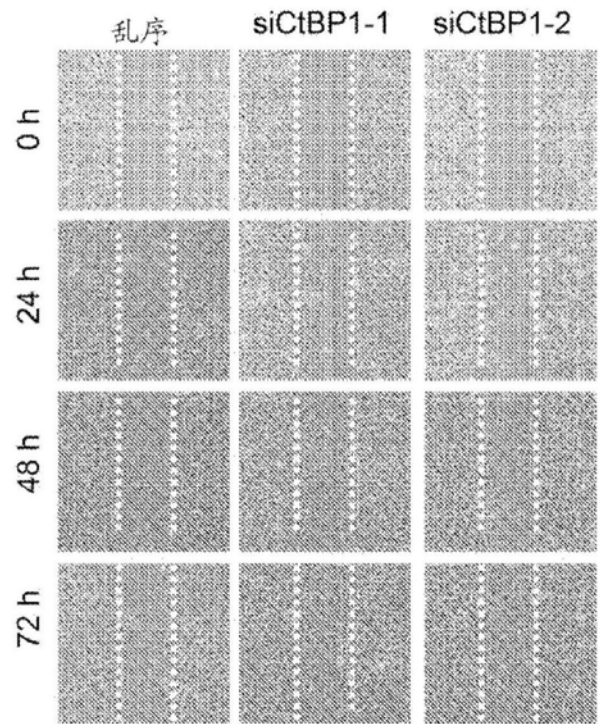


图10F

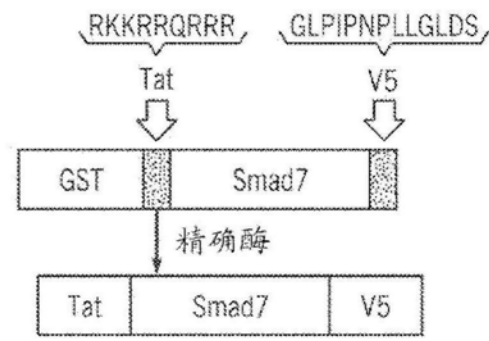


图11A

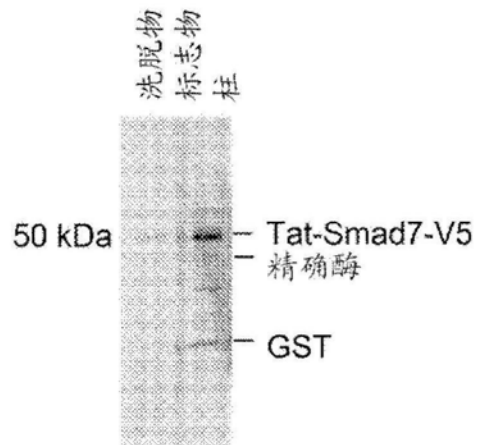


图11B

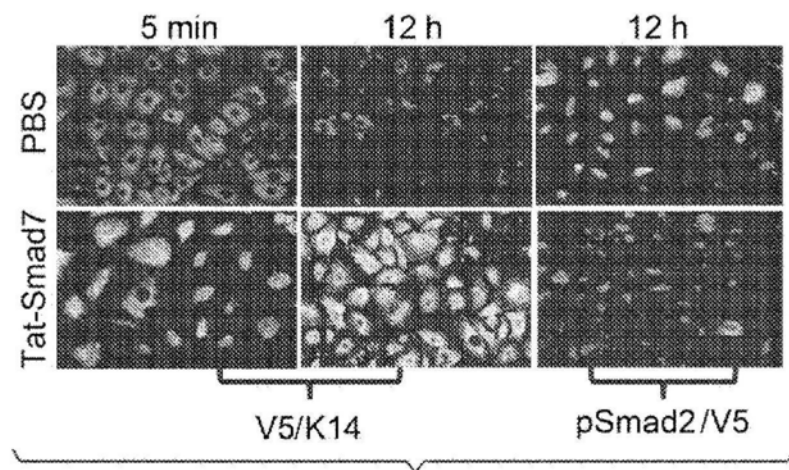


图11C

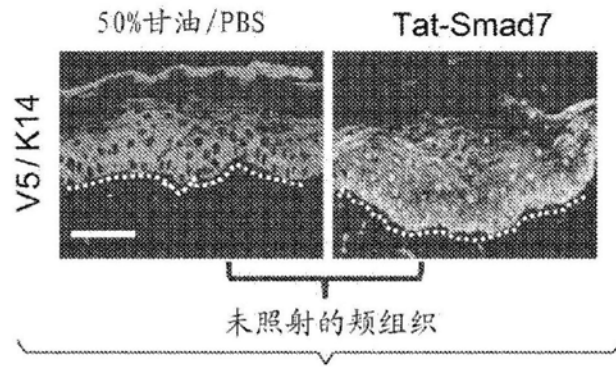


图11D

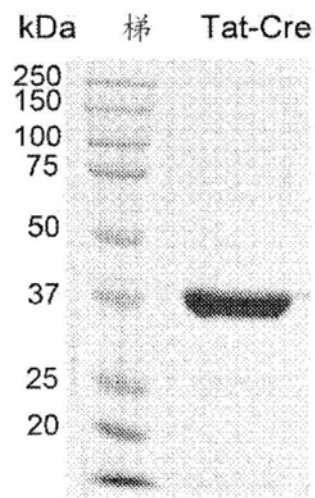


图11E

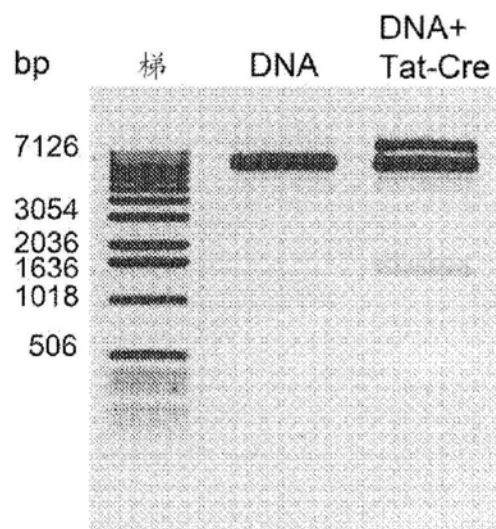


图11F



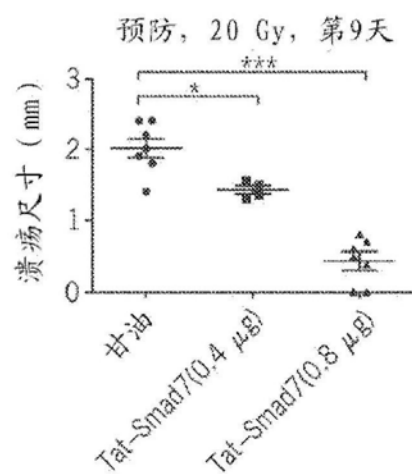


图11G

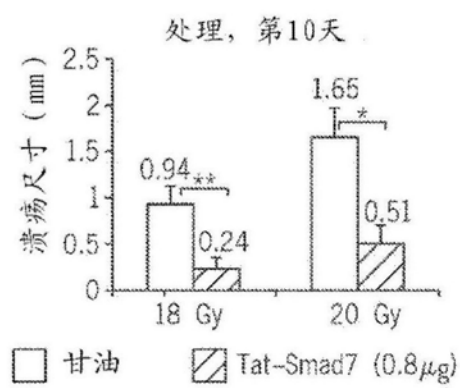


图12A

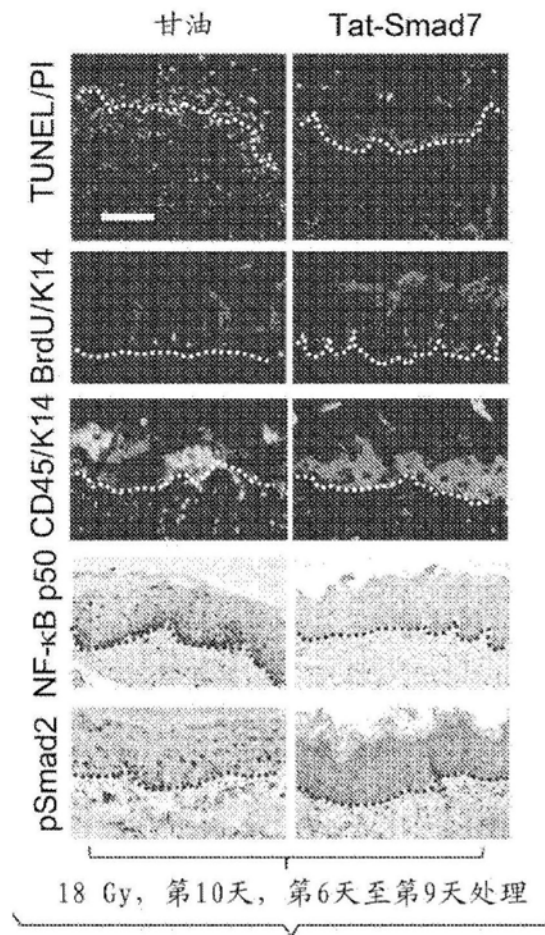


图12B

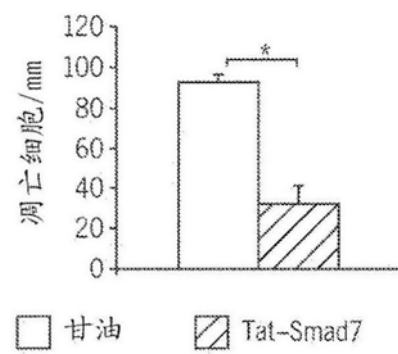


图12C

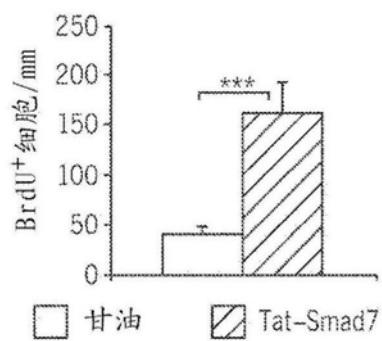


图12D

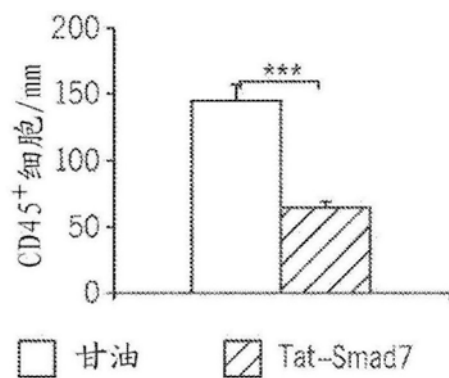


图12E

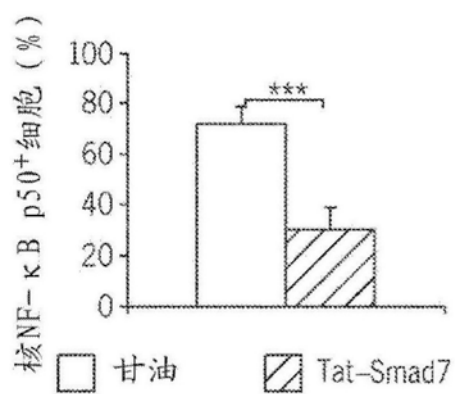


图12F

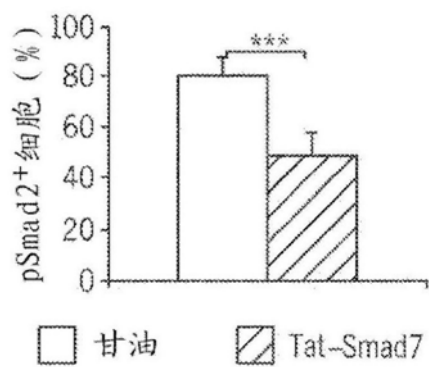


图12G

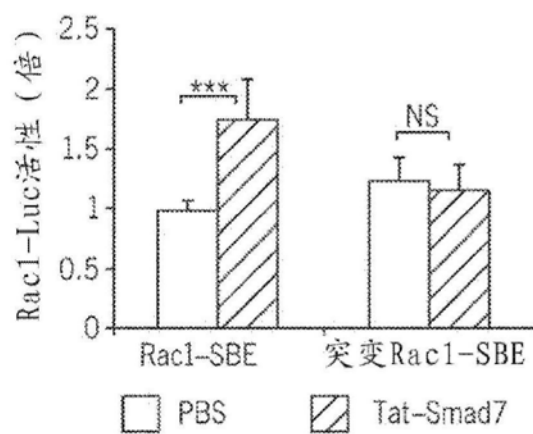


图12H

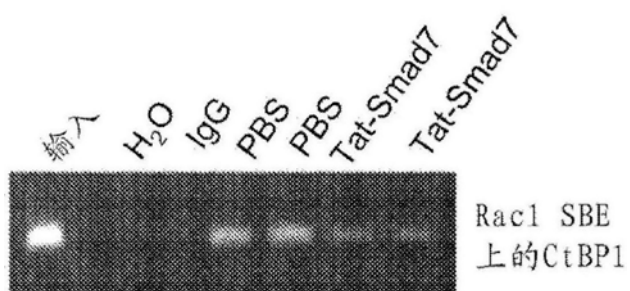


图12I

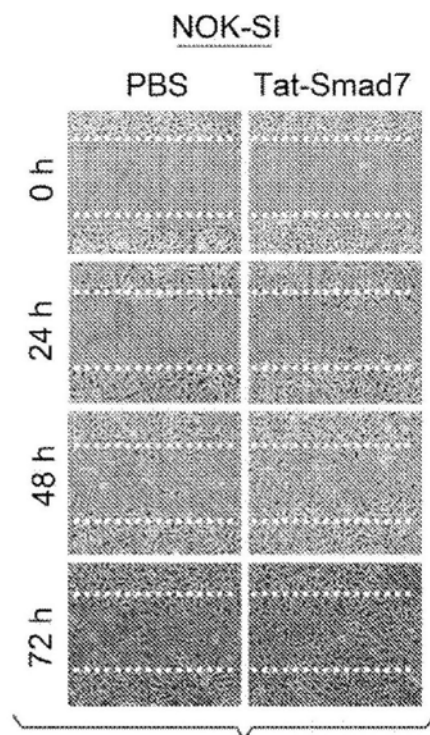


图13A

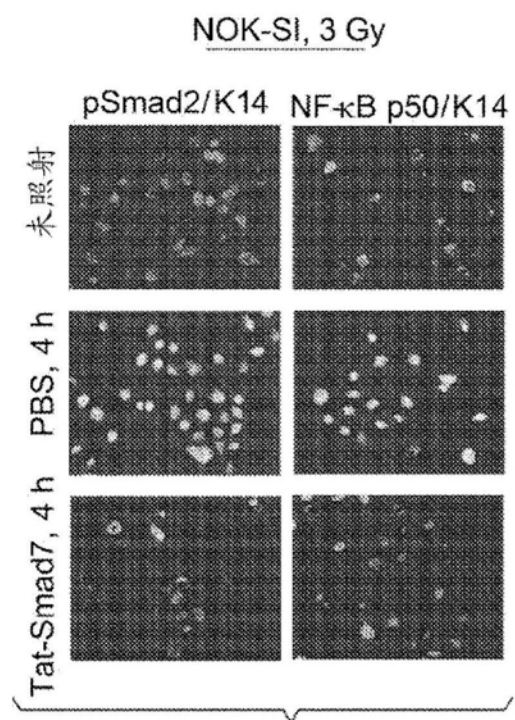


图13B

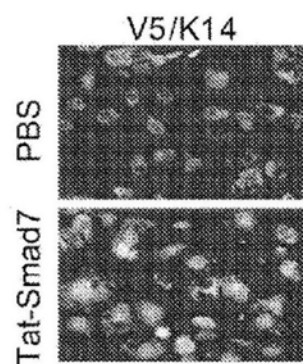


图13C

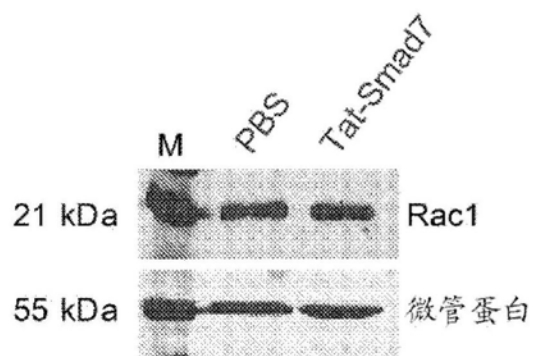


图13D

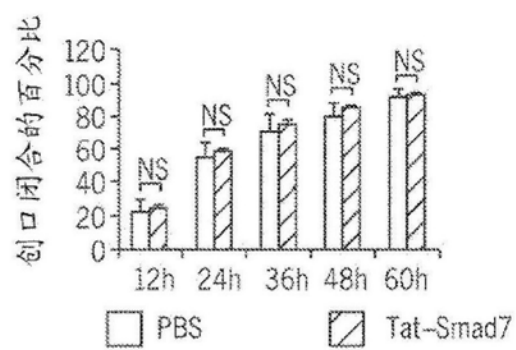


图13E

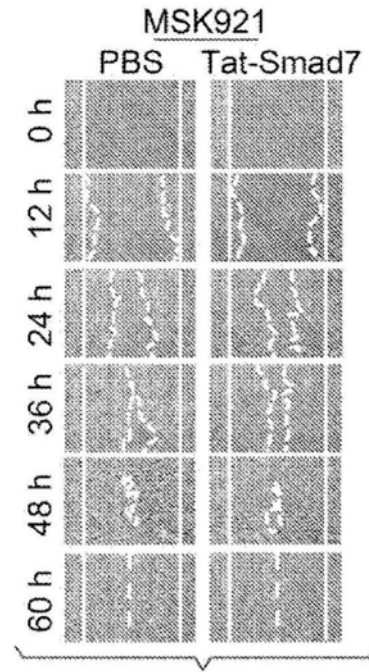


图13F

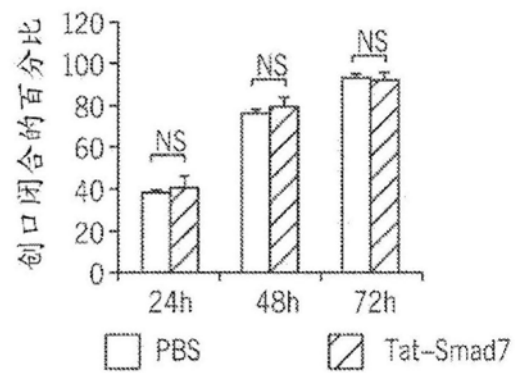


图13G

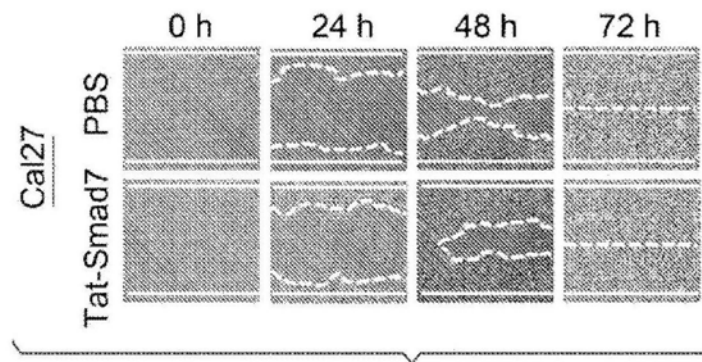


图13H

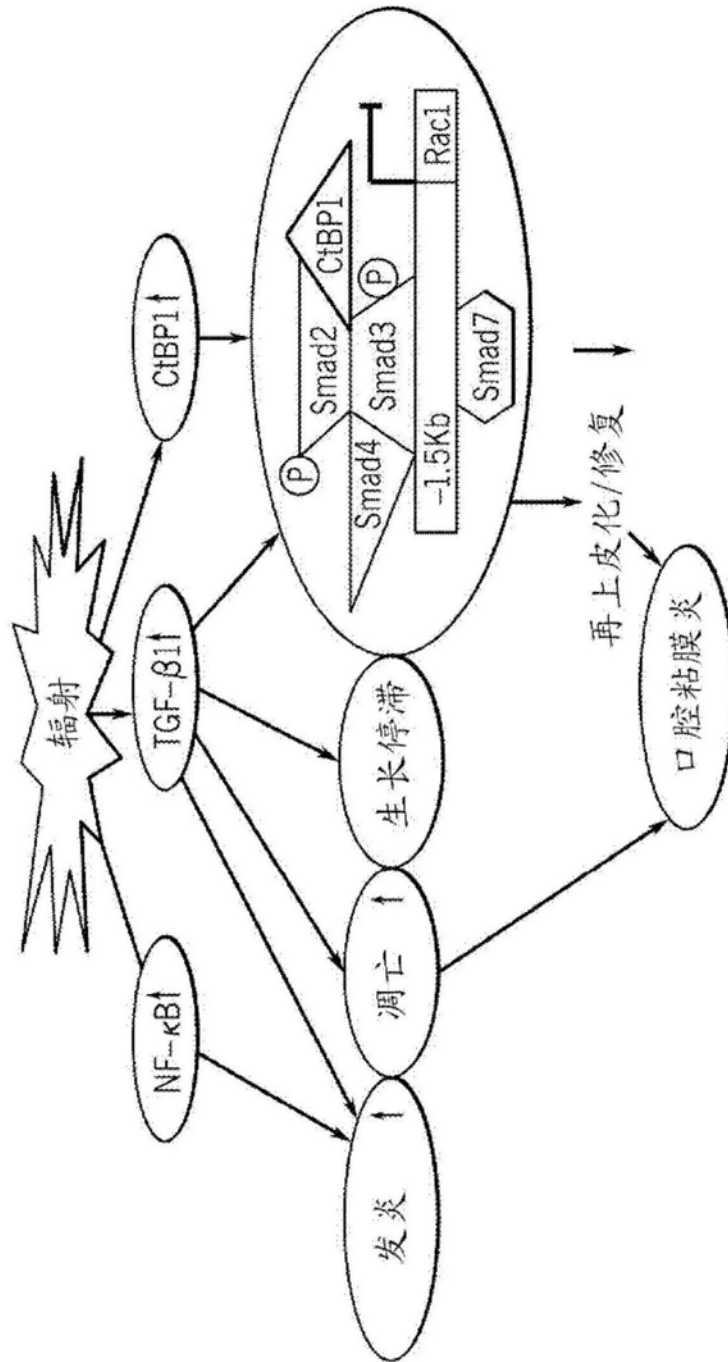


图14A



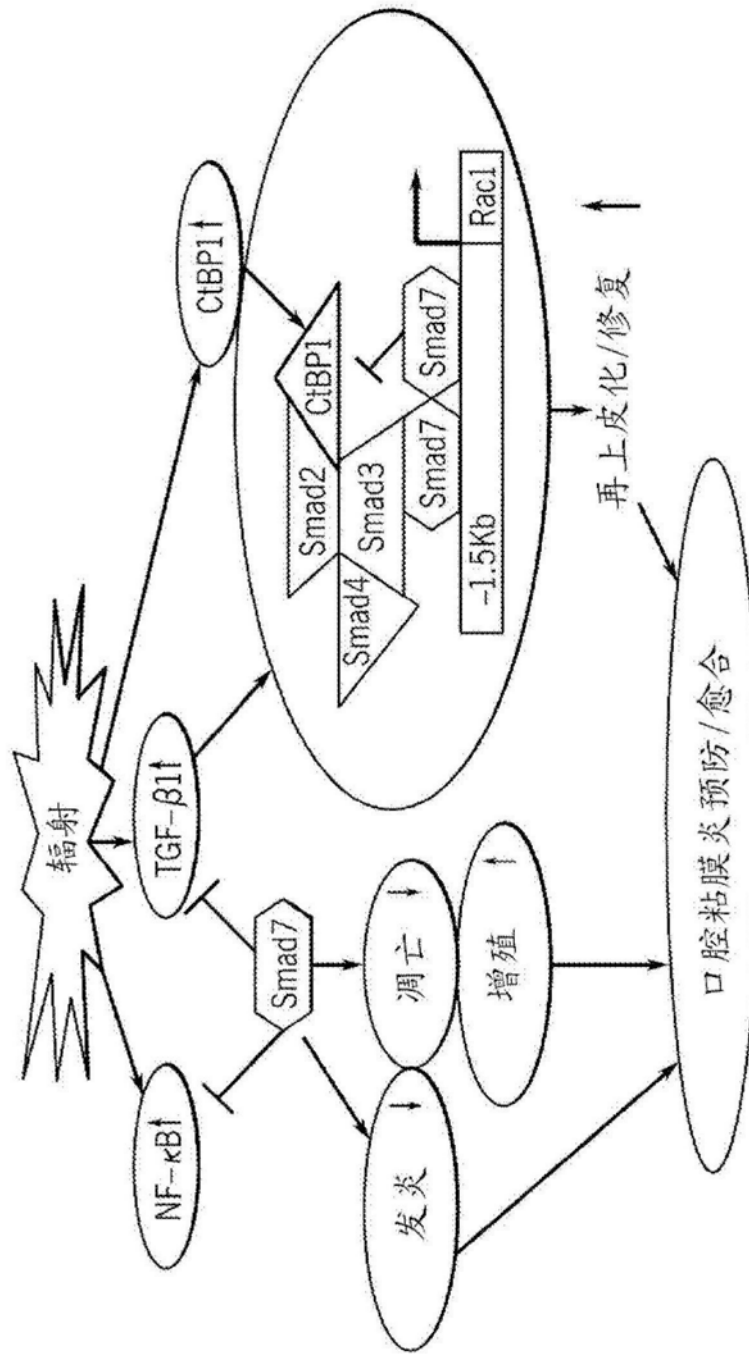


图14B

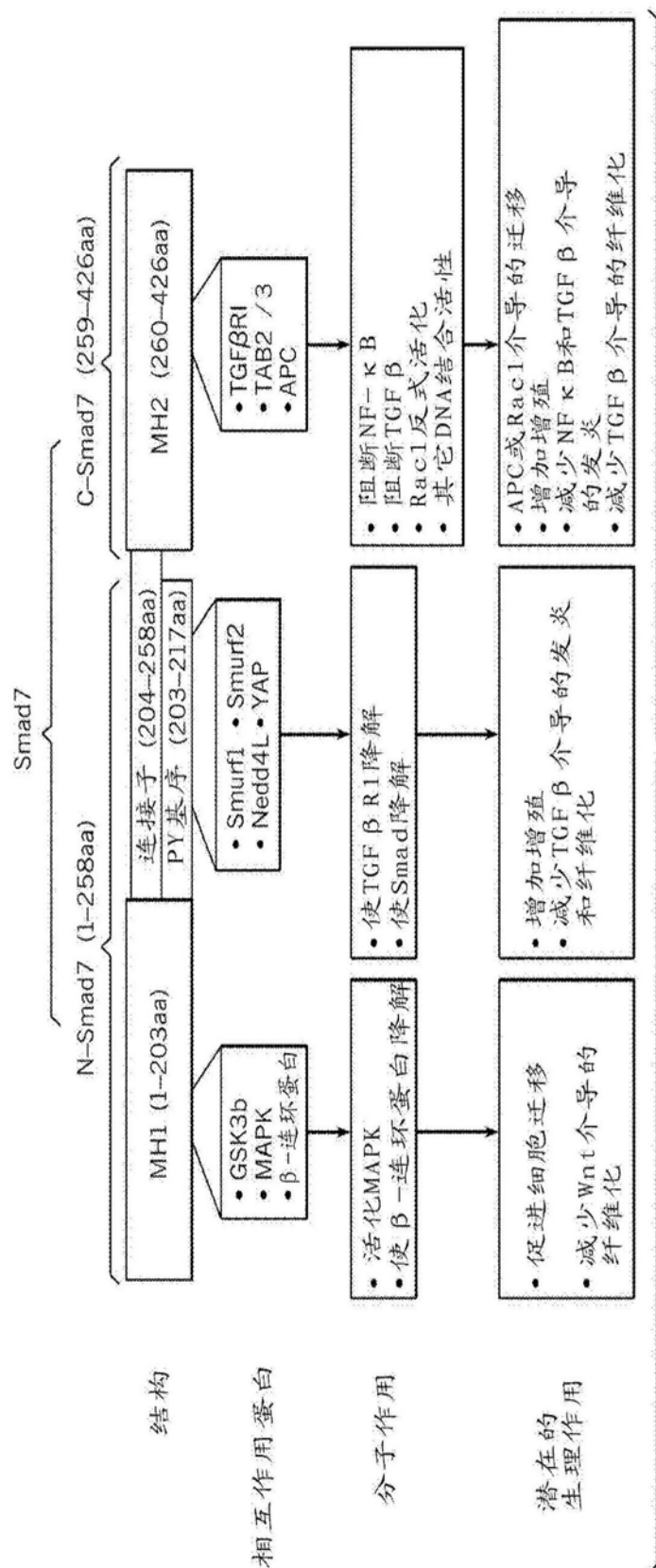


图15

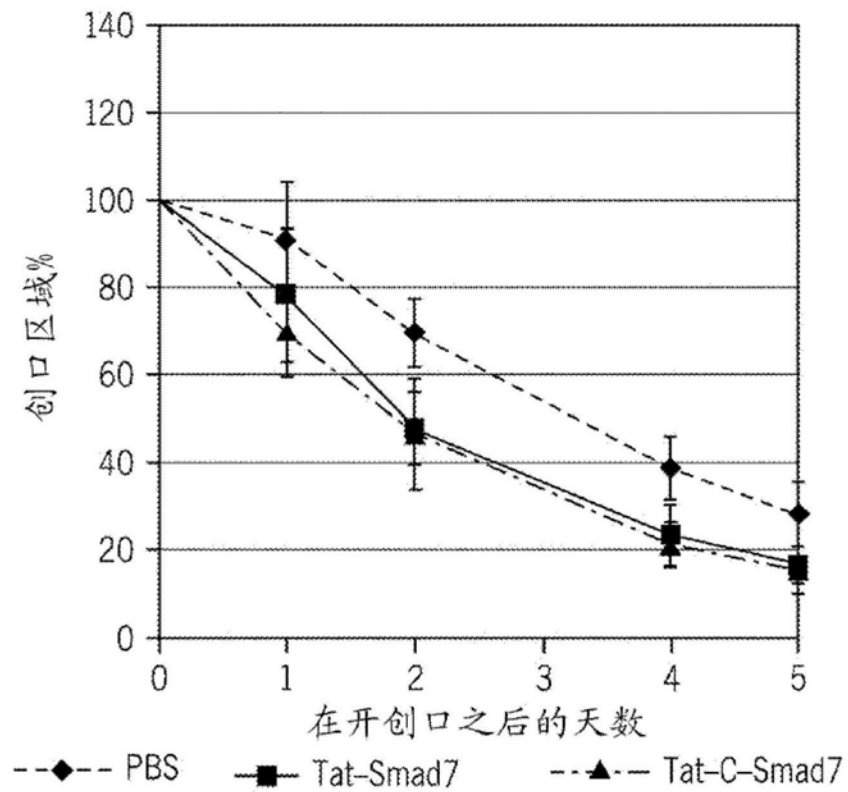


图16A

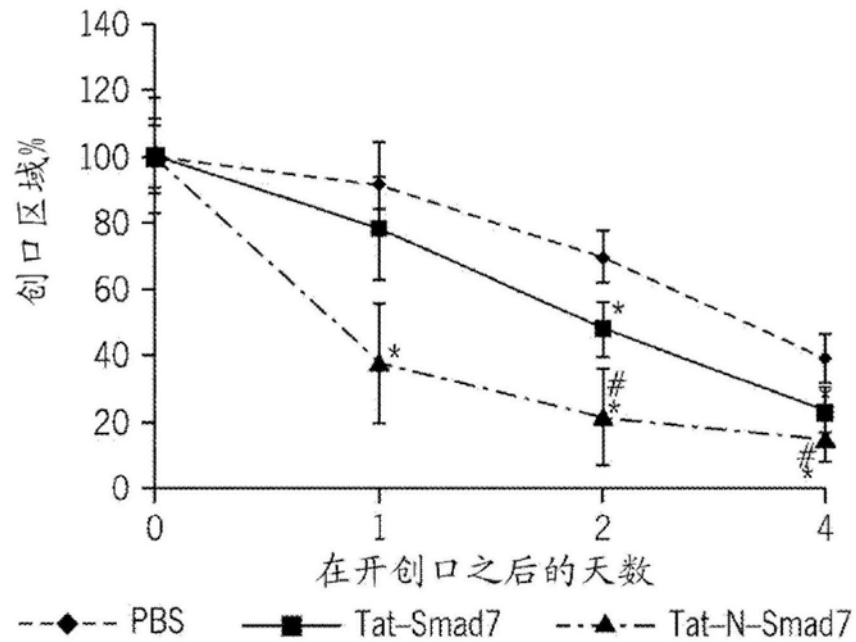


图16B

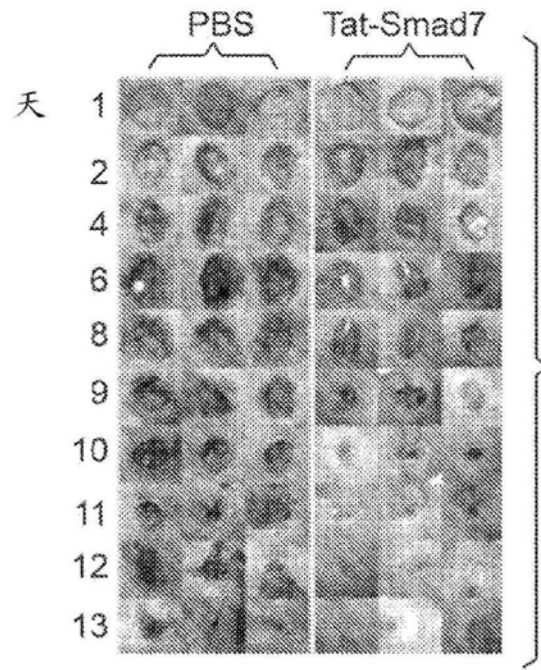


图17A

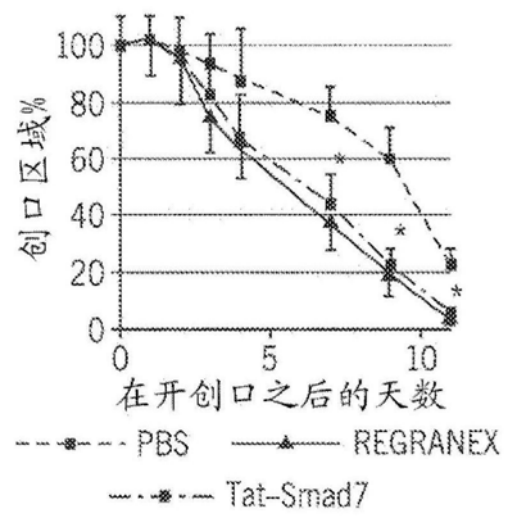


图17B

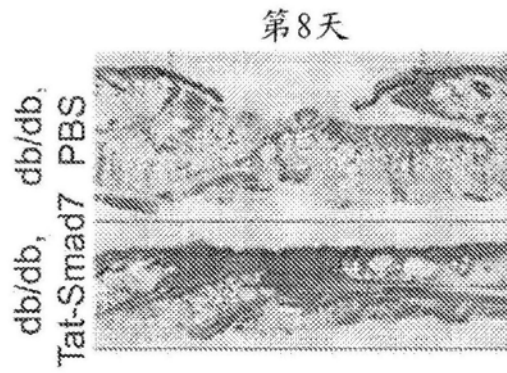


图17C