(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 119604615 A
(43)申请公布日 2025.03.11

(21)申请号 202380056688.2
(22)申请日 2023.08.01
(30)优先权数据
2022-123600 2022.08.02 JP
(85)PCT国际申请进入国家阶段日
2025.01.24
(86)PCT国际申请的申请数据
PCT/JP2023/028120 2023.08.01
(87)PCT国际申请的公布数据
WO2024/029529 JA 2024.02.08
(71)申请人 株式会社UniBio
地址 日本新泻县
(72)发明人 吉中乔慈 中林修 大川昭一郎 结城雅之
(74)专利代理机构 北京市柳沈律师事务所 11105
专利代理师 王珂
(51)Int.Cl.
C12N 15/12(2006.01)
C12N 5/14(2006.01)
C12N 15/82(2006.01)
C12P 21/02(2006.01)

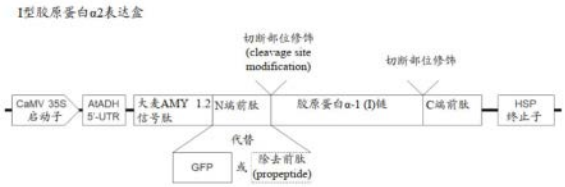
权利要求书1页 说明书30页
序列表（电子公布） 附图4页

(54)发明名称

胶原蛋白产生方法

(57)摘要

本发明涉及制造胶原蛋白的植物或分离得到的植物细胞的制造方法和其中使用的核酸构建物,以及使用了所述植物或分离得到的植物细胞的胶原蛋白的制造方法。



CN 119604615 A

1.一种从植物体或分离得到的植物细胞产生原胶原蛋白或胶原蛋白的方法,其包含:

A)准备包含编码原胶原蛋白的核酸构建物的植物体或分离得到的植物细胞的工序;

B)在植物体或分离得到的植物细胞中,使原胶原蛋白瞬时性地表达的工序;

C)将通过工序B分泌至细胞外的原胶原蛋白或胶原蛋白进行分离的工序。

2.一种从植物体或分离得到的植物细胞产生原胶原蛋白或胶原蛋白的方法,其包含:

在工序B之后,切断分泌至细胞外的原胶原蛋白的N末端或C末端的工序。

3.根据权利要求1或2所述的方法,其中,

所述核酸构建物包含编码原胶原蛋白的核酸序列,所述原胶原蛋白相对于天然型原胶原蛋白具有80%以上的氨基酸序列相同性。

4.根据权利要求1或2所述的方法,其中,

所述核酸构建物包含编码原胶原蛋白全长的核酸序列,并具有使原胶原蛋白的N末端在植物细胞内不被切断的突变。

5.一种原胶原蛋白或胶原蛋白,其通过权利要求1或2所述的方法而产生得到。

6.一种表达原胶原蛋白或胶原蛋白的核酸构建物,其用于权利要求1所述的方法。

7.一种表达载体,其包含权利要求6所述的核酸构建物。

8.一种植物体或分离得到的植物细胞,其包含权利要求6所述的核酸构建物或权利要求7所述的表达载体。

9.一种原胶原蛋白或胶原蛋白,其为权利要求6所述的核酸构建物编码的原胶原蛋白或胶原蛋白,其中,

所述原胶原蛋白或胶原蛋白具有非天然型氨基酸序列。

10.根据权利要求9所述的原胶原蛋白或胶原蛋白,其中,

在N末端前区融合有异源肽。

11.一种作为药物组合物、化妆品组合物或食品组合物的组合物,其包含:

通过权利要求1或2所述的方法而产生得到的原胶原蛋白或胶原蛋白,或者

权利要求9或10所述的原胶原蛋白或胶原蛋白。

胶原蛋白产生方法

技术领域

[0001] 本发明涉及制造胶原蛋白的植物或分离得到的植物细胞的制造方法和其中使用的核酸构建物,以及使用了所述植物或分离得到的植物细胞的胶原蛋白的制造方法。

背景技术

[0002] 胶原蛋白是在所有哺乳类动物结缔组织中发现的细胞外基质蛋白家族,胶原蛋白分子包含由三条被称为α链的多肽而形成三螺旋区域的结构。I型胶原蛋白等胶原蛋白参与动物的早期发育、器官形成、细胞粘附、细胞增殖、造血、伤口愈合、组织重建等各种生物反应。着眼于其结构的特性、较低免疫原性等,胶原蛋白广泛应用于医药领域、治疗领域、化妆品领域、食品领域等(非专利文献1:Ruggiero et al., 2000,FEBS Letter, 2000, Vol.469, pp.132-136. 非专利文献2:Merle et al., FEBS Letter, 2002, Vol.515, pp.114-118)。

[0003] 然而,通常用于这些用途的胶原蛋白是从动物组织中提取得到的,存在被病毒、其他感染性物质污染的可能性。因此,所述各领域的行业中,期待确保从动物组织以外得到胶原蛋白的方法。作为为了安全地大量生产动物性质蛋白质的优异性价比的大量生产方法,尝试了利用使用了基因修饰植物的表达系统,在基因修饰烟草等植物体中产生胶原蛋白(非专利文献2:Merle et al., FEBS Letter, 2002, Vol.515, pp.114-118,专利文献1:国际公开2006/035442号)。

[0004] 然而,胶原蛋白是具有复杂的结构的蛋白质,所述复杂的结构具有由3条肽链形成的三股螺旋结构,为了在植物体内制造可以用于医药领域、治疗领域、化妆品领域、食品领域等的胶原蛋白,需要进一步进行技术改善。例如,就如烟草植物这样的用于蛋白质产生的植物细胞而言,具有切断偏离人本来的前区(proregion)的切断部位的位置的酶,即使单一地在植物细胞内表达,也很难完全地制造与人胶原蛋白结构一致的高品质的胶原蛋白。因此,即使单一地在植物细胞内表达胶原蛋白,也很难制造具有可以用于医药领域、治疗领域、化妆品领域、食品领域等的品质的胶原蛋白。

[0005] 现有技术文献

[0006] 专利文献

[0007] 专利文献1:国际公开2006/035442号

[0008] 非专利文献

[0009] 非专利文献1:Ruggiero et al., 2000, FEBS Letter, 2000, Vol.469, pp.132-136.

[0010] 非专利文献2:Merle et al., FEBS Letter, 2002, Vol.515, pp.114-118.

发明内容

[0011] 发明所解决的技术问题

[0012] 本发明的目的在于提供一种用于在植物或分离得到的植物细胞中制造胶原蛋白

的进一步改善的方法,所述胶原蛋白可以用于医药领域、治疗领域、化妆品领域、食品领域等。

[0013] 解决问题的技术手段

[0014] 本发明的发明人对于用于在植物或分离得到的植物细胞中制造所述胶原蛋白的方法进行了深入试验、研究,其结果发现了,通过在植物细胞内瞬时性地表达原胶原蛋白(Procollagen),抑制植物体内的原胶原蛋白的N末端切断等,从而成功在植物或分离得到的植物细胞中效率良好地产生胶原蛋白,而完成了本申请发明。

[0015] 在本发明的一种方式中,提供一种从植物体或分离得到的植物细胞产生原胶原蛋白或胶原蛋白的方法,其包含:在植物体或分离得到的植物细胞中,使原胶原蛋白瞬时性地表达的工序。

[0016] 在本发明的一种方式中,提供一种胶原蛋白或原胶原蛋白,其在植物体或分离得到的植物细胞中产生得到。

[0017] 在本发明的一种方式中,提供一种核酸构建物,其用于在植物体或分离得到的植物细胞中产生胶原蛋白或原胶原蛋白。

[0018] 在本发明的一种方式中,提供一种载体和含有所述载体的植物体或分离得到的植物细胞,所述载体含有用于在植物体或分离得到的植物细胞中产生胶原蛋白或原胶原蛋白的核酸构建物。

[0019] 在本发明的一种方式中,提供一种具有非天然序列的原胶原蛋白或胶原蛋白,所述非天然序列是用于在植物体或分离得到的植物细胞中产生胶原蛋白或者原胶原蛋白的核酸构建物编码的。

[0020] 在本发明的一种方式中,提供一种包含原胶原蛋白或胶原蛋白的组合物。

[0021] 更具体而言,本发明提供以下内容。

[0022] [项1]一种从植物体或分离得到的植物细胞产生原胶原蛋白或胶原蛋白的方法,其包含:

[0023] A)准备包含编码原胶原蛋白的核酸构建物的植物体或分离得到的植物细胞的工序;

[0024] B)在植物体或分离得到的植物细胞中,使原胶原蛋白瞬时性地表达的工序;

[0025] C)将通过工序B分泌至细胞外的原胶原蛋白或胶原蛋白进行分离的工序。

[0026] [项2]一种从植物体或分离得到的植物细胞产生原胶原蛋白或胶原蛋白的方法,其包含:

[0027] 在工序B之后,切断分泌至细胞外的原胶原蛋白的N末端或C末端的工序。

[0028] [项3]根据项1或2所述的方法,其中,

[0029] 所述核酸构建物包含编码原胶原蛋白的核酸序列,所述原胶原蛋白相对于天然型原胶原蛋白具有80%以上的氨基酸序列相同性。

[0030] [项4]根据项1或2所述的方法,其中,

[0031] 所述核酸构建物包含编码原胶原蛋白全长的核酸序列,并具有使原胶原蛋白的N末端在植物细胞内不被切断的突变。

[0032] [项5]一种原胶原蛋白或胶原蛋白,其通过项1~4中任一项所述的方法而产生得到。

[0033] [项6]一种表达胶原蛋白的核酸构建物,其用于项1~4中任一项所述的方法。

[0034] [项7]一种表达载体,其包含项6所述的核酸构建物。

[0035] [项8]一种植物体或分离得到的植物细胞,其包含项5所述的核酸构建物或项7所述的表达载体。

[0036] [项9]一种原胶原蛋白或胶原蛋白,其为项6所述的核酸构建物编码的原胶原蛋白或胶原蛋白,其中,

[0037] 所述原胶原蛋白或胶原蛋白具有非天然型氨基酸序列。

[0038] [项10]根据项9所述的原胶原蛋白或胶原蛋白,其中,

[0039] 在N末端前区融合有异源肽。

[0040] [项11]一种作为药物组合物、化妆品组合物或食品组合物的组合物,其包含:

[0041] 通过项1~4中任一项所述的方法而产生得到的原胶原蛋白或胶原蛋白,或者

[0042] 项9或10所述的原胶原蛋白或胶原蛋白。

[0043] 发明的效果

[0044] 本发明具有以下效果:提供一种用于在植物或分离得到的植物细胞中制造胶原蛋白的进一步改善的方法,所述胶原蛋白可以用于医药领域、治疗领域、化妆品领域、食品领域等。

附图说明

[0045] [图1a] 图1a是表示I型胶原蛋白α1表达盒的构成的示意图。

[0046] [图1b] 图1b是表示I型胶原蛋白α2表达盒的构成的示意图。

[0047] [图1c] 图1c是表示脯氨酰4-羟化酶α亚基(P4Hα)、脯氨酰4-羟化酶β亚基(P4Hβ)表达盒的构成的示意图。

[0048] [图1d] 图1d是表示赖氨酰羟化酶3 (LH3)表达盒的构成的示意图。

[0049] [图1e] 图1e是表示骨形态发生蛋白1(BMP1)表达盒的构成的示意图。

[0050] [图1f] 图1f是表示番茄丛矮病毒(Tomato bushy stunt virus)来源RNA沉默抑制子(silencing suppressor)p19表达盒的构成的示意图。

[0051] [图2] 图2是表示通过使I型胶原蛋白α1表达盒在本氏烟草(Nicotiana benthamiana)植物表达系统瞬时性地表达,确认了在植物体内积蓄有具有N末端前区和C末端前区的全长胶原蛋白的图。泳道3表示提取了由I型胶原蛋白α1表达盒表达得到的全长胶原蛋白。泳道4~18表示全长胶原蛋白通过对应的蛋白酶进行了切断,其中,所述全长胶原蛋白由具有与各种蛋白酶对应的切断部位突变的I型胶原蛋白α1表达盒表达得到。

[0052] [图3] 图3是表示通过针对绿色荧光蛋白质(GFP)融合胶原蛋白α1提取液的提取液照射激发光,而确认了荧光的图。

[0053] [图4] 图4是表示通过使I型胶原蛋白α1表达盒和I型胶原蛋白α2表达盒在植物表达系统共表达,而确认α1和α2胶原蛋白同时在细胞体内积蓄的蛋白质印迹实验的结果的图。

[0054] [图5] 图5是表示通过质量分析仪器而确认了表达了α1和α2胶原蛋白的试验的结果的图。

具体实施方式

[0055] 胶原蛋白由3条被称为α链的肽链构成,并且通过与细胞质的内质网(endoplasmic reticulum)膜结合的核糖体从mRNA合成得到的各肽在合成的同时转移至内质网内,通过各个C末端的S-S键合反应等而形成包含3条肽的三链(三螺旋)原胶原蛋白分子。就从内质网经过高尔基体而分泌至细胞外的三链原胶原蛋白分子而言,通过切断N末端和C末端而得到成熟三链胶原蛋白(Ricard-Blum, Cold Spring Harb Perspect Biol. 2011 Jan 1;3(1):a004978)。

[0056] 成熟三链胶原蛋白根据构成的α链的种类、组合、整体的结构或特性而分为I型胶原蛋白、II型胶原蛋白、III型胶原蛋白~XXVIII型胶原蛋白等各个种类。例如,I型胶原蛋白主要由2条α1肽和1条α2肽构成,II型胶原蛋白由3条α1肽构成。各α链根据生物合成时的剪接方式的不同而进一步产生变异,III型胶原蛋白与II型胶原蛋白同样地由3条α1肽构成,但构成这些的α链的剪接方式不同(Ricard-Blum, Cold Spring Harb Perspect Biol. 2011 Jan 1;3(1):a004978)。

[0057] 一些种类的成熟三链胶原蛋白通过进一步彼此交联,而形成纤维结构,一些种类的成熟三链胶原蛋白形成网络结构。这些各种胶原蛋白在生物体内的各种位置生成,负责各种生物体功能(Ricard-Blum, Cold Spring Harb Perspect Biol. 2011 Jan 1;3(1):a004978)。

[0058] 为了大量生产这样的胶原蛋白而用于医药领域、治疗领域、化妆品领域,对于用于在植物或分离得到的植物细胞中制造胶原蛋白的方法,本发明的发明人进行了深入试验、研究的结果发现了,通过在植物细胞内使原胶原蛋白瞬时性地表达,从而可以大幅抑制由植物细胞的内在性酶引起的表达胶原蛋白的分解。此外,发现了通过向胶原蛋白基因导入突变,而可以从植物体回收三链原胶原蛋白分子或成熟三链胶原蛋白分子,可以效率良好地产生胶原蛋白。

[0059] 更具体而言,在哺乳类生物体内,天然胶原蛋白作为在N末端和C末端具有延长区域的原胶原蛋白而分泌至细胞外之后,N末端和C末端通过酶切断而得到成熟胶原蛋白分子,与之相对的,在植物细胞内表达原胶原蛋白时,在植物细胞内切断N末端和C末端,其结果使水溶性大幅降低。所述植物细胞内切断的N末端切断部位与本来原胶原蛋白在成熟为胶原蛋白时N末端被切断的部位相比,位于更靠近C末端的位置,从植物细胞内表达的原胶原蛋白得到的成熟胶原蛋白与天然成熟胶原蛋白相比,N末端被缩短。此外,植物细胞内切断的N末端切断部位的切断位置是不清楚的。通过这些植物细胞内的非天然位置的切断,在植物细胞内产生得到的胶原蛋白可以具有与天然胶原蛋白不同的特性。

[0060] 本发明的发明人等发现了,通过使植物细胞内的胶原蛋白瞬时性地表达,可以抑制植物细胞内的原胶原蛋白的N末端切断。此外,发现了可以抑制植物细胞内或产生工序中的原胶原蛋白的N末端或C末端的切断的氨基酸突变,通过使具有这些突变的原胶原蛋白表达,从而能够效率良好地回收分泌至细胞外的原胶原蛋白或胶原蛋白。此外,通过期望的酶切断原胶原蛋白的N末端和C末端延长区域,并且通过将酶识别序列导入胶原蛋白基因,能够由期望的工序将原胶原蛋白转换为成熟胶原蛋白分子,所述酶识别序列用于使切断后的成熟胶原蛋白分子的N末端和C末端的氨基酸序列与天然成熟胶原蛋白分子的氨基酸序列相同。

[0061] (定义)

[0062] 只要没有其他定义,本说明书中使用的技术用语和科学用语全部具有与由本发明所属技术领域的本领域技术人员通常所理解的相同意义。

[0063] 本说明书中,在表示多个数值的范围的情况下,也同样地意指包含这些多个范围的任意的下限值和上限值的组合的范围。

[0064] 本说明书中"实质上"这样的用语,具有与由本发明所属技术领域的本领域技术人员通常所理解的相同意义,例如,考虑到存在生物学上或化学上的现象无法完全实现目的状态的情况,使用该表达旨在包含目的状态和由于生物学上或化学上的特性而不可避免地无法达成的状态。

[0065] 本说明书中,在涉及数值,使用"约"这样的用语的情况下,所述值是指例如,可以在10%、9%、8%、7%、6%、5%、4%、3%、2%、1%、0.5%、0.1%、0.05%或0.01%的范围内变动。

[0066] 本说明书中,"包含"这样的用语具有与由本发明所属技术领域的本领域技术人员通常所理解的相同意义,例如,包括"包含"和"由…组成",具体而言,"包含" A的组合物不仅包含仅A的情况,也包含其他构成要素B。

[0067] 本说明书中,对于组合物所使用的"由…组成"或"由…构成"这样的用语具有与由本发明所属技术领域的本领域技术人员通常所理解的相同意义,并且用于表示排他性地构成所述组合物的构成要素。例如,由A"构成"的组合物排他性地仅包含A。然而在一种方式中,由A"构成"的组合物包含以下方式:包含基于生物学和化学的性质而在制造上无法避免A以外的混入物。

[0068] 本发明中,氨基酸序列或碱基序列的"相同性"是指,以使得待比较的两个氨基酸序列的氨基酸或碱基序列的碱基尽可能多地一致的方式排列这两个序列,将一致的氨基酸数或碱基数除以全氨基酸数或碱基数而得到的数以百分数来表示。在所述排列时,根据需要,向待比较的两个序列的一方或双方插入适宜间隙。这样的序列的排列化可以使用例如BLAST、FASTA、CLUSTAL W等周知的程序进行。在插入间隙的情况下,所述全氨基酸数或碱基数是将一个间隙计数为一个氨基酸或碱基而得到的数。由此计数得到的全氨基酸数或碱基数在待比较的两个序列间不同的情况下,相同性(%)是将一致的氨基酸数或碱基数除以较长一方的序列的全氨基酸数或碱基数而进行计算。但是,待比较的序列与其他任意的序列连接的状态的情况下,仅将相应的区域取出来对比序列,计算相同性。

[0069] 本发明中,原胶原蛋白或胶原蛋白的"瞬时性地"表达是指,具有与由本发明所属技术领域的本领域技术人员通常所理解的相同意义,典型性而言,在不导入宿主细胞的基因组的状态下,使编码原胶原蛋白或胶原蛋白的核酸序列在宿主细胞内表达。这样的瞬时性地表达可以通过例如使用病毒载体而实施。

[0070] (在植物或分离得到的植物细胞中产生胶原蛋白的方法)

[0071] 在一种方式中,本发明提供一种在植物或分离得到的植物细胞中产生胶原蛋白的方法。

[0072] 本发明的胶原蛋白产生方法中,包含:通过使包含编码原胶原蛋白的核酸的核酸构建物在植物体或分离得到的植物细胞中瞬时性地表达,而回收分泌至细胞外的原胶原蛋白或胶原蛋白分子的工序,最终产生原胶原蛋白或胶原蛋白分子。

[0073] 典型性而言,本发明的原胶原蛋白或胶原蛋白产生方法包含以下工序:

[0074] A)准备包含编码原胶原蛋白的核酸构建物的植物体或分离得到的植物细胞的工序;

[0075] B)在植物体或分离得到的植物细胞中,使原胶原蛋白瞬时性地表达的工序;

[0076] C)将通过工序B分泌至细胞外的原胶原蛋白或胶原蛋白进行分离的工序。

[0077] 就所述工序A)“准备包含编码原胶原蛋白的核酸构建物的植物体或分离得到的植物细胞的工序”中使用的核酸构建物而言,典型性而言,是具有下文(用于植物或分离得到的植物细胞中的胶原蛋白产生的核酸构建物)的构成的核酸构建物。

[0078] 本发明的原胶原蛋白或胶原蛋白的产生方法中产生的胶原蛋白的种类没有特别限定,通过组合适当的编码α链的核酸构建物,能够产生任意的胶原蛋白,例如,I型胶原蛋白、II型胶原蛋白、III型胶原蛋白等。例如,如下文(核酸构建物编码的原胶原蛋白或胶原蛋白)中所记载的,可以产生具有非天然型氨基酸序列的胶原蛋白或将异源肽进行融合的胶原蛋白。

[0079] 作为本发明的原胶原蛋白或胶原蛋白产生方法中使用的植物,例如可举出烟草属植物、苔藓植物(小立碗藓(P. patens)等)、土豆(S. tuberosum等)、禾本科植物(O. sativa等)、十字花科植物、生菜,优选为烟草属植物。作为烟草属植物,例如,可举出本氏烟草(N. benthamiana)、烟草(N. tabacum)、高烟草(Nicotiana excelsior)等,但不限于此。

[0080] 所述工序B)“在植物体或分离得到的植物细胞中,使原胶原蛋白瞬时性地表达的工序”中,典型性而言,在不导入宿主细胞的基因组的状态下,使包含编码原胶原蛋白的核酸序列的病毒载体等表达载体在植物体或分离得到的植物细胞中表达。

[0081] 通过瞬时性地表达原胶原蛋白,能够抑制植物细胞内的N末端切断。就植物体或分离得到的植物细胞内的原胶原蛋白瞬时性地表达而言,可以通过本发明所属技术领域的本领域技术人员所公知的方法进行,就原胶原蛋白的表达量、表达期间等而言,通过对使用的载体、启动子、培养基的构成、培养条件等进行适宜调节,本领域技术人员可以在不脱离本发明的主旨的范围内适宜调节。

[0082] 例如,在将本发明的核酸构建物作为重组病毒使用的情况下,在感染了病毒的植物体或分离得到的植物细胞内病毒进行增殖,在植物细胞内生产大量的胶原蛋白。例如,在使用植物体而产生胶原蛋白的情况下,病毒接种后培育植物约3~15天后,可以从叶等植物组织中提取回收胶原蛋白或原胶原蛋白。

[0083] 所述工序C)“将通过工序B分泌至细胞外的原胶原蛋白或胶原蛋白进行分离的工序”中,可以利用本发明所属技术领域的本领域技术人员所公知的任意的方法。例如,可以使用C. D. Mount et al., Archives of Biochemistry and Biophysics, 240 33 (1985)、U. H. Gregory and I. R. Willshire, Hoppe-Sayler's Z. Physiol. Chem., 356 1765 (1975)中所记载的方法,但不限于此。更具体而言,例如,在烟草植物体中使胶原蛋白表达而产生胶原蛋白或原胶原蛋白的情况下,可以将经过冷冻保存的烟草植物的叶在适当的缓冲液中磨碎,通过盐析从缓冲液提取、纯化胶原蛋白或原胶原蛋白。

[0084] 在一种方式中,就本发明的原胶原蛋白或胶原蛋白的产生方法而言,可以包含:在工序B之后,切断分泌至细胞外的原胶原蛋白的N末端或C末端的工序。所述切断可以通过对原胶原蛋白用化学处理或酶处理或这些的组合进行处理而实施。所述化学或酶处理的至少一部分在植物细胞内以外的位置进行,可以在培养工序中、提取纯化工序中、提取纯化工序

后等工序中进行。例如,已知可以通过将完全长原胶原蛋白在酸性环境进行提取纯化,从而切除天然型原胶原蛋白的C末端。在一种方式中,可以在提取纯化工序后通过蛋白酶(蛋白质分解酶)而切除原胶原蛋白的N末端或C末端。可以通过在提取纯化工序后由蛋白酶切除原胶原蛋白的N末端或C末端,从而抑制向切除反应体系中混入夹杂物,使蛋白酶在适当的反应条件下进行切除反应,从而提高切除反应的特异性,获得生成物的高稳定性。如此,不仅能够制造稳定的胶原蛋白作为最终销售形态,也能够有以下应用方法:将完全长原胶原蛋白作为生物3D打印机的生物油墨的原料,在通过生物3D打印机构建立体结构后追加切断N末端或C末端的酶(例:BMP1),从而使三螺旋结构更牢固。这是相当于数码相机普及前的照片胶片的"显影"的工序。

[0085] 切断所述原胶原蛋白的N末端或C末端延长区域的酶没有特别限定。在一种方式中,可以通过使用切断哺乳类细胞中的天然型原胶原蛋白的N末端和C末端的酶,由天然型原胶原蛋白生成天然型成熟胶原蛋白。作为这样的切断天然型原胶原蛋白的N末端和C末端的酶的非限定性的实例,可以使用Procollagen N-Endopeptidase(原胶原N内肽酶)、BMP1等。此外,可以通过利用氨基酸突变,向原胶原蛋白的N末端和C末端导入期望的蛋白酶识别序列,从而利用对应的蛋白酶。作为这样的酶,例如,可以使用HRV3C、弗林蛋白酶(Furin)、肠激酶(Enterokinase),但不限于此。需要说明的是,HRV3C、弗林蛋白酶和肠激酶的识别序列和切断部位如下。

[0086]

蛋白酶	识别序列和切断部位(用/表示)
HRV3C	Leu Glu Val Leu Phe Gln / Gly Pro
弗林蛋白酶	Arg Xaa Arg Arg /
	Arg Xaa Lys Arg /
肠激酶	Asp Asp Asp Asp Lys /

[0087] 在一种方式中,本发明的原胶原蛋白或胶原蛋白的产生方法可以包含对原胶原蛋白或胶原蛋白进行翻译后修饰的工序。例如,可以通过P4Hα、P4Hβ或LH3等,对原胶原蛋白或胶原蛋白进行翻译后修饰,从而提高原胶原蛋白或胶原蛋白的稳定性。在本发明的原胶原蛋白或胶原蛋白的产生方法中,就进行翻译后修饰的工序而言,可以在原胶原蛋白或胶原蛋白的翻译后的任意时间点进行,可以在植物细胞内或从植物细胞中分泌后的阶段进行。例如,通过将P4Hα和P4Hβ与原胶原蛋白在同一植物细胞进行表达,可以在植物细胞内例如内质网内进行翻译后修饰。优选所述翻译后修饰的至少一部分在植物细胞的液胞以外的位置进行。

[0088] (在植物或分离得到的植物细胞中产生得到的胶原蛋白)

[0089] 在一种方式中,本发明提供一种在植物或分离得到的植物细胞中产生得到的原胶原蛋白或胶原蛋白。就通常在医药、化妆品或食品领域等中以往所提供的胶原蛋白而言,大多将存在于生物体组织的胶原蛋白分子进一步聚合得到的胶原蛋白纤维等进行分解而提供,这样的胶原蛋白通过分解处理而失去胶原蛋白的结构和功能的一部分。另一方面,就在本发明的植物或分离得到的植物细胞中产生得到的原胶原蛋白或胶原蛋白而言,可以不经过这样的分解处理而产生,因此可以在更完整的状态下保持胶原蛋白的结构和功能。

[0090] 在一种方式中,就通过本发明的产生方法而产生的原胶原蛋白而言,包含具有N末

端和C末端的原胶原蛋白。这样的原胶原蛋白可以通过使包含编码天然型原胶原蛋白的核酸的核酸构建物在植物细胞内进行瞬时性地表达,提取、纯化分泌至细胞外的原胶原蛋白,而产生。

[0091] 在一种方式中,本发明中产生的原胶原蛋白包含具有天然型氨基酸序列的原胶原蛋白。就所述具有天然型序列的原胶原蛋白而言,氨基酸序列与天然原胶原蛋白相同,但是根据其产生工序中的各种条件,在翻译后修饰状态等方面可以不同。如上所述,在本发明的植物或分离得到的植物细胞中产生胶原蛋白的方法中,可以通过各种方法对翻译后修饰状态进行控制,由此可以具备各种生理活性上的优点。

[0092] 在一种方式中,本发明中产生的原胶原蛋白包含:在N末端切断部位或C末端切断部位或这两者的部位具有氨基酸突变的非天然型原胶原蛋白。在一种方式中,所述突变抑制植物细胞内或产生工序中的原胶原蛋白的切断。在一种方式中,所述突变提供对应于期望的蛋白酶的识别序列。在一种方式中,所述突变提供通过产生工序的化学处理而诱导切断的序列。就非天然型完全长原胶原蛋白的N末端切断部位或C末端切断部位的氨基酸突变而言,可以是最终生成天然型成熟胶原蛋白的突变,也可以是生成在N末端或C末端或这两者的末端具有非天然型序列的成熟胶原蛋白的突变。

[0093] 所述氨基酸突变可以是相对于天然型氨基酸序列的氨基酸添加、缺失、置换或这些的组合。

[0094] 在一种方式中,就通过本发明的产生方法产生的原胶原蛋白而言,包含仅N末端或C末端中的一个被切除的N末端缺失原胶原蛋白和N末端缺失原胶原蛋白。所述N末端缺失原胶原蛋白具有与通过植物细胞内切除得到的不同的N末端缺失。在一种方式中,所述N末端或C末端缺失原胶原蛋白可以通过将分泌至细胞外的完全长或不完全长原胶原蛋白进行化学处理或酶处理或这些的组合处理,而生成。

[0095] 在一种方式中,本发明中产生的原胶原蛋白是包含在N末端或C末端或这两者融合了具有特定功能的肽的原胶原蛋白。融合的肽没有特别限定,可以为异源蛋白质来源的肽、新合成的非天然肽。例如,可以利用促进向细胞外的分泌的肽、可以用于原胶原蛋白的纯化的亲和性肽、产生荧光的标记肽等。在一种方式中所述进行了融合的肽可以使用由本发明的所属的技术领域的本领域技术人员已知的接头序列而与原胶原蛋白进行融合。在一种方式中所述进行了融合的肽的一部分区域或全区域可以通过化学或酶反应,而从原胶原蛋白中切除。

[0096] 在一种方式中,本发明中产生的成熟胶原蛋白在N末端和C末端具有天然型氨基酸序列。在另一种方式中,本发明中产生的胶原蛋白在N末端和C末端的至少一者中具有非天然型氨基酸序列。N末端的非天然型氨基酸序列与通过植物细胞内切除而生成的不同。

[0097] 在一种方式中,本发明中产生的原胶原蛋白或胶原蛋白包含具有上述以外的任意的氨基酸突变的原胶原蛋白或胶原蛋白。在一种方式中,就本发明中产生的原胶原蛋白或胶原蛋白而言,包含上述以外的原胶原蛋白或胶原蛋白的全区域或一部分区域与其他肽融合而成的原胶原蛋白或胶原蛋白。

[0098] (用于植物或分离得到的植物细胞中的胶原蛋白产生的核酸构建物)

[0099] 在一种方式中,本发明提供一种用于植物或分离得到的植物细胞中的胶原蛋白产生的核酸构建物。

[0100] 本发明的核酸构建物具有编码至少一种构成胶原蛋白的α链的核酸序列。在一种方式中本发明的核酸构建物具有表达原胶原蛋白的序列。就所述原胶原蛋白而言,包含具有所述(在植物或分离得到的植物细胞中产生得到的胶原蛋白)提及的各种突变的原胶原蛋白以及融合有其他肽的原胶原蛋白。本发明的核酸构建物编码的α链的种类没有特别限定,可以编码α1、α2和这些的剪接异构体中的任一种。

[0101] 本发明的核酸构建物可以是DNA、RNA或由本发明的所属技术领域中本领域技术人员已知的任意的这些的类似物。例如,本发明的核酸构建物可以包含非天然碱基,并且其中一部分可以包含烃链、肽链。

[0102] 本发明的核酸构建物可以通过由本发明的所属技术领域的本领域技术人员已知的常规方法的化学合成方法或基因工程方法而制备。例如,可以通过合成得到具有天然存在的序列的cDNA后,以所述cDNA为模板,使用适当的引物而导入适宜突变而合成。或者,可以将待合成的核酸构建物分成一些较短区域而对多个DNA片段进行化学合成,通过公知的fusion PCR法而连接各片段从而合成DNA构建物。RNA构建物可以通过来自DNA构建物的转录反应而制备。

[0103] 作为向核酸构建物的序列导入突变的方法,可以利用由本发明的所属技术领域的本领域技术人员已知的任意的方法。例如,本领域技术人员可以使用Kunkel法、Gapped duplex法等定点诱变法、重叠延伸PCR(overlap extension PCR)法、QuikChange法等方法,但不限于此。

[0104] 在一种方式中,本发明的核酸构建物可以包含用于在植物体内瞬时性地表达的核酸片段。可以包含例如编码在通过转录反应而合成重组病毒RNA时所需的适当的RNA聚合酶启动子(例如,T7启动子等噬菌体聚合酶启动子)、病毒复制酶、迁移蛋白、外壳蛋白等的核酸片段。

[0105] 在一种方式中,本发明的核酸构建物可以包含编码用于在植物体内使胶原蛋白稳定表达的有益的酶等的核酸片段。例如,可以包含编码增强胶原蛋白的稳定性的P4Hα、P4Hβ或LH3等的核酸片段或这些的组合。作为这些酶,可以利用各种生物来源的酶,例如,可以使用来源于线虫、小鼠、人等的酶。

[0106] 在一种方式中,就编码这些酶等的核酸片段的一部分或全部而言,可以包含于与包含编码原胶原蛋白的α链的核酸的核酸构建物不同的第2核酸构建物中,包含编码原胶原蛋白的α链的核酸的第1核酸构建物可以与所述第2核酸构建物组合使用。

[0107] 在将包含编码原胶原蛋白的α链的核酸的第1核酸构建物与包含编码酶等的核酸片段的第2核酸构建物进行组合使用的情况下,可以将这些导入相同的植物体或分离得到的植物细胞,也可以导入不同的植物体或分离得到的植物细胞,并使得分泌至细胞外的原胶原蛋白与酶进行反应。

[0108] (包含核酸构建物的表达载体)

[0109] 在一种方式中,本发明提供一种包含核酸构建物的载体,所述核酸构建物用于植物或分离得到的植物细胞中的胶原蛋白产生。所述载体可用于在植物宿主内瞬时性地表达胶原蛋白,典型性而言,作为植物病毒载体而提供。在一种方式中,本发明的表达载体可以以“重组病毒”的方式提供。所述病毒为RNA病毒时本发明的核酸构建物为RNA,病毒为DNA病毒时所述核酸构建物为DNA。在一种方式中,可以使载体质粒转化至根癌农杆菌

(Agrobacterium tuminiferens),使该根癌农杆菌感染植物而使靶蛋白瞬时性地表达。

[0110] 作为本发明的重组病毒提供的表达载体可以通过常规方法对由本发明的所属技术领域的本领域技术人员已知的任意的植物病毒载体(参照例如,“蛋白质 核酸 酶”,vol.45,p.607-613等)进行修饰而制备。在一种方式中,优选修饰番茄丛矮病毒而作为重组病毒。

[0111] (包含核酸构建物或表达载体的植物体或分离得到的植物细胞)

[0112] 在本发明的一种方式中,提供一种包含核酸构建物或表达载体的植物体或分离得到的植物细胞。所述植物体或分离得到的植物细胞除了是经过核酸构建物或表达载体的导入处理而得到的之外,也可以是这些被处理植物体或植物细胞经过生长发育或增殖而得到的。

[0113] (核酸构建物编码的原胶原蛋白或胶原蛋白)

[0114] 在本发明的一种方式中,包含所述本发明的核酸构建物编码的原胶原蛋白或胶原蛋白。所述原胶原蛋白或胶原蛋白不限于从植物体或分离得到的植物细胞产生得到的。就所述原胶原蛋白或胶原蛋白而言,典型性而言,具有非天然型氨基酸序列。在一种方式中,本发明的原胶原蛋白或胶原蛋白相对于天然型原胶原蛋白或胶原蛋白,具有70%以上、75%以上、80%以上、85%以上、90%以上或95%以上的氨基酸序列相同性。本发明的原胶原蛋白或胶原蛋白相对于天然型原胶原蛋白或胶原蛋白,可以包含一个以上的氨基酸的置换、插入、缺失。

[0115] 在一种方式中,本发明的原胶原蛋白或胶原蛋白可以代替N末端前区,而融合异源蛋白质,由此添加各种功能。例如,可以通过代替N末端前区而融合GFP,产生荧光胶原蛋白。所述原胶原蛋白或胶原蛋白可以通过由本发明的所属技术领域的本领域技术人员所知的任意的方法而进行表达和纯化,可以用于分析化学特性、生理活性等目的。

[0116] (包含原胶原蛋白或胶原蛋白的组合物)

[0117] 在本发明的一种方式中,提供一种组合物,其包含:由植物体或者分离得到的植物细胞产生的原胶原蛋白或胶原蛋白,或者通过其他方法而产生的具有非天然型氨基酸序列的原胶原蛋白或胶原蛋白。

[0118] 在一种方式中,本发明的组合物的方式不限,例如以药物组合物、化妆品组合物或食品组合物的方式提供,除了包含原胶原蛋白或胶原蛋白之外,还可以包含这些各领域的组合物中允许的载体、赋形剂等。在本发明的组合物中,原胶原蛋白或胶原蛋白可以为进一步通过分子间的交联处理而具有包含多个原胶原蛋白或胶原蛋白分子的纤维状或网络状结构的多聚体的状态。本发明的组合物含有的胶原蛋白的量没有特别限定,可以根据剂型、目的等而适宜调节。

[0119] 就本发明的包含原胶原蛋白或胶原蛋白的药物组合物而言,包含所述组合物本身为药物的情况和作为制造药物时的原料或中间制品的组合物。本发明的药物组合物的形态没有特别限定,可以是固体状、粉状、颗粒状、液体状。本发明的药物组合物针对对象给药的方式的情况下,其剂型没有特别限定,例如,可以作为片剂、细粒剂、丸剂、锭剂、胶囊剂或这些的原材料或中间制品而提供。

[0120] 本发明的药物组合物包含再生医疗用药物或作为制造再生医疗用药物时的原料或中间制品的组合物。例如,本发明的药物组合物作为片状的组合物而提供,可以作为用于

促进对象的体内或体表组织再生的组合物或用于培养片状的组织的支架材料而提供。在一种方式中,本发明的药物组合物可以作为用于构成用3D打印技术等而具有3维结构的再生医疗制品的材料组合物,例如作为生物油墨的原料而提供。

[0121] 在一种方式中,本发明的药物组合物包含作为可植入的医疗设备或制造所述医疗设备时的原料或中间制品的组合物。这些组合物可以用于疾病或障碍的预防或治疗、美容目的等。所述组合物的非限定性的实例包含:创伤敷料、人造皮肤、人造血管、用于除去流体或向患者递送流体的导管和其他设备、人造心脏、人造肾脏、用于整形外科的针、板和植入物。例如本发明的药物组合物可以用作3D生物打印得到的乳房切除后植入物、美容整形用填充剂。

[0122] 就本发明的包含原胶原蛋白或胶原蛋白的化妆品组合物而言,包含所述组合物本身为化妆品的情况和作为制造化妆品时的原料或中间制品的组合物。本发明的化妆品组合物的方式没有特别限定,可以是固体状、粉状、颗粒状、液体状。本发明的化妆品可以局部适用或口服给药,例如,可以作为软膏、洗剂、霜剂、微乳液、凝胶、油、溶液等局部给药剂、片剂、细颗粒剂、丸剂、锭剂、胶囊剂等口服给药剂或这些的原材料或中间制品而提供。

[0123] 就本发明的包含原胶原蛋白或胶原蛋白的食品组合物而言,包含所述组合物为本身为食品的情况和作为制造食品时的原料或中间制品的组合物。所述食品包含健康食品、功能性食品、特定保健用食品、患者用食品,在以人以外的动物作为对象来使用的情况下包含饲料。所述食品的方式没有特别限定,可以是固体状或液态。作为人以外的动物没有特别限定,例如可以以哺乳类、鸟类、爬虫类、两栖类、鱼类、昆虫等各种动物作为对象来使用。作为具体的食品的种类,例如包含:清凉饮料、碳酸饮料、营养饮料、果汁饮料、乳饮料等饮料或这些饮料的浓缩原液和调制用粉末;冰淇淋、冰果子露、刨冰等冷冻甜点;荞麦面、乌冬面、粉条、饺子皮、烧麦皮、中式面条、方便面等面条类;糖、口香糖、糖果、软糖、透明果味糖、焦糖、巧克力、硬糖、零食、饼干等烘焙点心、果冻、果酱、奶油等点心类;鱼板、汉堡包、火腿、香肠等水产/畜产加工食品;加工乳、发酵乳、酸奶、黄油、奶酪等乳制品;色拉油、天妇罗油、人造黄油、蛋黄酱、起酥油、生奶油、沙拉酱等油脂和油脂加工食品;酱料、蘸料等调味料;汤、炖菜、咖喱、面包、果酱、沙拉、配菜、泡菜等,但不限于此。

[0124] 在下文中使用实施例对本发明进一步详细地说明,这些实施例为例举,实施例的范围不限定本发明的范围。

[0125] 实施例

[0126] (材料和方法)

[0127] 本研究中使用的表达盒的构建如图1a~f所示。本研究中的编码序列根据需要为了烟草中的表达而进行最优化,经过化学合成使得具有所需的侧翼区域(flanking region)。

[0128] 使用的序列如下。

[0129] [表1]

[0130]

SEQ ID NO.	序列名	序列
1	CaMV 35S 启动子 (DNA序列)	TGAGACTTTTCAACAAAGGGTAATATCCGGAAAC CTCCTCGGATTCCATTGCCCAGCTATCTGTCACT TTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTC CTACAAATGCCATCATTGCGATAAAGGAAAGGCC ATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCA AAGATGGACCCCCACCCACGAGGAGCATCGTGGA AAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAG TGGATTGATGTGATATCTCCACTGACGTAAGGGA TGACGCACAATCCCACTATCCTTCGCAAGACCCTT CCTCTATATAAGGAAGTTCATTTCATTTGGAGAG AACA
2	TMV omega (DNA序列)	ACAATTACCAACAACAACAAACAACAAACAACATT ACAATTACTATTTACAATTAC
3	AtADH 5'-UTR (DNA序列)	TACATCACAATCACACAAAACTAACAAAAGATCAA AAGCAAGTTCTTCACTGTTGATA
4	HSP-终止子 (DNA序列)	TATGAAGATGAAGATGAAATATTTGGTGTGTCAA ATAAAAAGCTAGCTTGTGTGCTTAAGTTTGTGTT TTTTTCTTGGCTTGTTGTGTTATGAATTTGTGGC TTTTTCTAATATTAAATGAATGTAAGATCTCATTA TAATGAATAAACAAATGTTTCTATAATCCATTGTG AATGTTTTGTTGGATCTCTTCGCATATAACTACT GTATGTGCTATGGTATGGACTATGGAATATGATT AAAGATAAG
5	大麦AMY 1.2 信号	MANKHLSLSLFLVLLGLSASLASG

[0131]

	肽 (氨基酸序列)	
6	巯基蛋白酶aleurain(大麦)液泡信号(氨基酸序列)	MAHARVLLLALAVLATAAVAVASSSSFADSNPIRPVTDRAASTL
7	胶原蛋白α-1 N端前肽(氨基酸序列)	QEEGQVEGQDEDIPPITCVQNGLRYHDRDVWKPEPCRICVCDNGKVLCDDVICDETKNCPGAEVPEGECCPVCPDGSESPTDQETTGVEGPKGDTGPRGPRGPAGPPGRDGIPGQPGLPGPPGPPGPPGPPGLGGNFAP
8	胶原蛋白α-1(I)链(氨基酸序列)	QEEGQVEGQDEDIPPITCVQNGLRYHDRDVWKPEPCRICVCDNGKVLCDDVICDETKNCPGAEVPEGECCPVCPDGSESPTDQETTGVEGPKGDTGPRGPRGPAGPPGRDGIPGQPGLPGPPGPPGPPGPPGLGGNFAPQLSYGYDEKSTGGISVPGPMGPSGPRGLPGPPGAPGPQGFQGPPGEPGEPGASGPMGPRGPPGPPGKNGDDGEAGKPGRPGERGPPGPQGARGLPGTAGLPGMKGHRGFSGLDGAKGDAGPAGPKGEPGSPGENGAPGQMGPRGLPGERGRPGAPGPAGARGNDGATGAAGPPGPTGPAGPPGFPGAVGAKGEAGPQGPRGSEGPQGVRGEPGPPGPAGAAGPAGNPGADGQPGAKGANGAPGIAGAPGFPGARGPSGPQGPGGPPGPKGNSGEPGAPGSKGDTGAKGEPGPVGVQGPPGPAGEEGKRGARGEPGPTGLPGPPGERGGPGSRGFPGADGVAGPKGPAGERGSPGPAGPKGSPGEAGRPGEAGLPGAKGLTGSPGS

[0132]

		PGPDGKTGPPGPAGQDGRPGPPGPPGARGQAGVM GFPGPKGAAGEPGKAGERGVPGPPGAVGPAGKDGE AGAQGPPGPAGPAGERGEQGPAGSPGFQGLPGPAG PPGEAGKPGEQGVPGDLGAPGPSGARGERGFPGER GVQGPPGPAGPRGANGAPGNDGAKGDAGAPGAPG SQGAPGLQGMPGERGAAGLPGPKGDRGDAGPKGA DGSPGKDGVRGLTGPIGPPGPAGAPGDKGESGPSG PAGPTGARGAPGDRGEPGPPGPAGFAGPPGADGQP GAKGEPGDAGAKGDAGPPGPAGPAGPPGPIGNVGA PGAKGARGSAGPPGATGFPGAAGRVGPPGPSGNAG PPGPPGPAGKEGGKGPRGETGPAGRPGEVGPPGPP GPAGEKGSPGADGPAGAPGTPGPQGIAGQRGVVGL PGQRGERGFPGLPGPSGEPGKQGPSGASGERGPPG PMGPPGLAGPPGESGREGAPGAEGSPGRDGSPGAK GDRGETGPAGPPGAPGAPGAPGPVGPAGKSGDRGE TGPAGPTGPVGPVGARGPAGPQGPRGDKGETGEQ GDRGIKGHRGFSGLQGPPGPPGSPGEQGPSGASGP AGPRGPPGSAGAPGKDGLNGLPGPIGPPGPRGRTG DAGPVGPPGPPGPPGPPGPPSAGFDFSFLPQPPQEK AHDGGRYYRADDANVVRDRDLEVDTTLKSLSQQIE NIRSPEGSRKNPARTCRDLKMCHSDWKSGEYWIDPN QGCNLDAIKVFCNMETGETCVYPTQPSVAQKNWYI SKNPKDKRHVWFGESMTDGFQFEYGGQGSDPADV AIQLTFLRLMSTEASQNITYHCKNSVAYMDQQTGNL KKALLLQGSNEIEIRAEGNSRFTYSVTVDGCTSHTGA WGKTVIEYKTTKTSRLPIIDVAPLDVGAPDQEFGFD VGPVCFL
9	胶原蛋白α-1 C端前肽(氨基酸序列)	DDANVVRDRDLEVDTTLKSLSQQIENIRSPEGSRKN PARTCRDLKMCHSDWKSGEYWIDPNQGCNLDAIKV FCNMETGETCVYPTQPSVAQKNWYISKNPKDKRHV WFGESMTDGFQFEYGGQGSDPADVAIQLTFLRLMS TEASQNITYHCKNSVAYMDQQTGNLKKALLLQGSN

[0133]

		EIEIRAEGNSRFTYSVTVDGCTSHTGAWGKTVIEYKT TKTSRLPIIDVAPLDVGAPDQEFGFDVGPVCFL
10	胶原蛋白α-2 N端前肽(氨基酸序列)	QSLQEETVRKGPAGDRGPRGERGPPGPPGRDGEDG PTGPPGPPGPPGPPGLGGNFAA
11	胶原蛋白α-2(I)链(氨基酸序列)	QSLQEETVRKGPAGDRGPRGERGPPGPPGRDGEDG PTGPPGPPGPPGPPGLGGNFAAQYDGKGVGLGPGP MGLMGPRGPPGAAGAPGPQGFQGPAGEPGEPGQT GPAGARGPAGPPGKAGEDGHPGKPGRPGERGVVG PQGARGFPGTPGLPGFKGIRGHNGLDGLKGQPGAP GVKGEPGAPGENGTPGQTGARGLPGERGRVGAPG PAGARGSDGSVGPVGPAGPIGSAGPPGFPGAPGPKG EIGAVGNAGPAGPAGPRGEVGLPGLSGPVGPPGNP GANGLTGAKGAAGLPGVAGAPGLPGPRGIPGPVGA AGATGARGLVGEPGPAGSKGESGNKGEPGSAGPQG PPGPSGEEGKRGPNGEAGSAGPPGPPGLRGSPGSRG LPGADGRAGVMGPPGSRGASGPAGVRGPNGDAGR PGEPGLMGPRGLPGSPGNIGPAGKEGPVGLPGIDGR PGPIGPAGARGEPGNIGFPGPKGPTGDPGKNGDKG HAGLAGARGAPGPDGNNGAQGPPGPQGVQGGKG EQGPPGPPGFQGLPGPSGPAGEVGKPGERGLHGEF GLPGPAGPRGERGPPGESGAAGPTGPIGSRGPSGPP GPDGNKGEPGVVGAVGTAGPSGPSGLPGERGAAGI PGGKGEKGEPGLRGEIGNPGRDGARGAPGAVGAPG PAGATGDRGEAGAAGPAGPAGPRGSPGERGEVGPA GPNGFAGPAGAAGQPGAKGERGAKGPKGENGVVG PTGPVGAAGPAGPNGPPGPAGSRGDGGPPGMTGF PGAAGRTGPPGPSGISGPPGPPGPAGKEGLRGPRGD QGPVGRTGEVGAVGPPGFAGEKGPSGEAGTAGPPG

[0134]

		TPGPQGLLGAPGILGLPGSRGERGLPGVAGAVGEPG PLGIAGPPGARGPPGAVGSPGVNGAPGEAGRDGNP GNDGPPGRDGQPGHKGERGYPGNIGPVGAAGAPG PHGPVGPAGKHGNRGETGPSGPVGPAGAVGPRGPS GPQGIRGDKGEPGEKGPRGLPGLKGHNGLQGLPGI AGHHGDQGAPGSVGPAGPRGPAGPSGPAGKDGRT GHPGTVGPAGIRGPQGHQGPAGPPGPPGPPGPPGV SGGGYDFGYDGDFYRADQPRSAPSLRPKDYEVDATL KSLNNQIETLLTPEGSRKNPARTCRDLRLSHPEWSSG YYWIDPNQGCTMDAIKVYCDFSTGETCIRAQPENIP AKNWYRSSKDKKHVWLGETINAGSQFEYNVEGVTS KEMATQLAFMRLLANYASQNITYHCKNSIAYMDEET GNLKKAVILQGSNDVELVAEGNSRFTYTVLVDGCSK KTNEWGKTIIEYKTNKPSRLPFLDIAPLDIGGADQEF FVDIGPVCFK
12	胶原蛋白α-2 C端前肽 (氨基酸序列)	DQPRSAPSLRPKDYEVDATLKSLNNQIETLLTPEGSR KNPARTCRDLRLSHPEWSSGYYWIDPNQGCTMDAI KVYCDFSTGETCIRAQPENIPAKNWYRSSKDKKHVW LGETINAGSQFEYNVEGVTSKEMATQLAFMRLLANY ASQNITYHCKNSIAYMDEETGNLKKAVILQGSNDVE LVAEGNSRFTYTVLVDGCSKKTNEWGKTIIEYKTNK PSRLPFLDIAPLDIGGADQEFFVDIGPVCFK
13	GFP (氨基酸序列)	SKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDA TYGKLTLKFICTTGKLPVPWPTLVTTFSYGVQCFSRY PDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTR AEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYN SHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHY QQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDH MVLLEFVTAAGITHGMDELYK
14	GST (氨基酸序列)	SPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG DKWRNKKFELGLEFPNLPYYIDGDVKLTQSMAIIRYI ADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYS

[0135]

		KDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDH VTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPK
15	MBP(氨基酸序列)	KIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVT VEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYA QSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPI AVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKS ALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDV GVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAF NKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKG QPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEG LEAVNKDKPLGAVALKSYEEELVKDPRIAATMENAQ KGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEAL KDAQTNSSSNNNNNNNNNNLG
16	HN标签(氨基酸序列)	HNHNHNHNHNHN
17	人P4H α亚基(氨基酸序列)	HPGFFTSIGQMTDLIHTEKDLVTSLKDYIKAEEDKLE QIKKWAEKLDRLTSTATKDPEGFVGHPVNAFKLMK RLNTEWSELENLVLKDMSDGFISNLTIQRQYFPNDE DQVGAAKALLRLQDTYNLDTDTISKGNLPGVKHKS FLTAEDCFELGKVAYTEADYYHTELWMEQALRQLD EGEISTIDKVSVLDYLSYAVYQQGDLDKALLLTKKLL ELDPEHQRANGNLKYFEYIMAKEKDVNKSASDDQS DQKTTPKKKGVAVDYLPERQKYEMLCRGEGIKMTP RRQKKLFCRYHDGNRNPKFILAPAKQEDEWDKPRII RFHDIISDAEIEIVKDLAKPRLRRATISNPITGDLETVH YRISKSAWLSGYENPVVSRINMRIQDLTGLDVSTAEE LQVANYGVGGQYEPHFDFARKDEPDAFKELGTGNR IATWLFYMSDVSAGGATVFPEVGASVWPKKGTAVF WYNLFASGEGDYSTRHAACPVLVGNKWVSNKWLHE RGQEFRRPCTLSELE

[0136]

18	秀丽隐杆线虫P4H α亚基(氨基酸序列)	DLFTSIADMQNLLETERNIPKILDKYIHDEEERLVQL KKLSEEYSKKNEISIENGLKDITNPINAFLLIKRKIFDW KEIESKMNANKAGNVVSSITDDSYGVRYPTADDLSG AAIGLLRLQDTYRLDTKDLADGKIYADQGNYTFSAK DCFEIARAAYNEHDFYHTVMWMEEAQRRLGDEVEP TVEVEDILEYLAFALYKQNNLKHALKLTEELYKMNP THPRAKGNVKWYEDLLEQEGVRRSDMRKNLPPIQN RRPDSVLGNTERTMYEALCRNEVPVSQKDISRLYCYY KRDRPFLVYAPIKVEIKRFNPLAVLFKDVISDDEVAAI QELAKPKLARATVHDSVTGKLVTATYRISKSAWLKE WEGDVVETVNKRIGYMTNLEMETAEELQIANYGIG GHYDPHFDHAKKEESKSFESLGTGNRIATVLFYMSQ PSHGGGTVFTEAKSTILPTKNDALFWYNLYKQGDG NPDTRHAACPVLVGIKWVSNKWIHEKGNEFRRPCG LKSSDYERFVGDLGYGPEPRNAPNVSPNLAKDVWET L
19	人P4H β亚基(氨基酸序列)	DAPEEEDHVLVLRKSNFAEALAAHKYLLVEFYAPWC GHCKALAPEYAKAAGKLKAEGSEIRLAKVDATEESDL AQQYGVRGYPTIKFFRNGDTASPKEYTAGREADDIV NWLKKRTGPAATTLPDGAAAESLVESSEVAVIGFFK DVESDSAKQFLQAAEAIDDIPFGITSNSDVFSKYQLD KDGVVLFKKFDEGRNNFEGEVTKENLLDFIKHNQLP LVIEFTEQTAPKIFGGEIKTHILLFLPKSVSDYDGKLS NFKTAAESFKGKILFIFIDSDHTDNQRILEFFGLKKEE CPAVRLITLEEEMTKYKPESEELTAERITEFCHRFLEG KIKPHLMSQELPEDWDKQPVKVLVGKNFEDVAFDE KKNVFVEFYAPWCGHCKQLAPIWDKLGETYKDHEN IVIAKMDSTANEVEAVKVHSFPTLKFFPASADRTVID YNGERTLDGFKKFLESGGQDGAGDDDDLEDLEEAE EPDMEEDDDQKAV
20	小鼠P4H β	DALEEEDNVLVLKKSNFEEALAAHKYLLVEFYAPWC GHCKALAPEYAKAAAKLKAEGSEIRLAKVDATEESDL

[0137]

	亚基(氨基酸序列)	AQQYGVRGYPTIKFFKNGDTASPKEYTAGREADDIV NWLKKRTGPAATTLSDTAAAESLVDSSEVTVIGFFK DVESDSAKQFLLAAEAIDDIPFGITSNSGVFSKYQLDK DGVVLFKKFDEGRNNFEGEITKEKLLDFIKHNQLPL VIEFTEQTAPKIFGGEIKTHILLFLPKSVSDYDGKLSSF KRAAEGFKGKILFIFIDSDHTDNQRILEFFGLKKEECP AVRLITLEEEMTKYKPESDELTAEKITEFCHRFLEGKI KPHLMSQEVPEDWDKQPVKVLVGANFEEVAFDEKK NVFVEFYAPWCGHCKQLAPIWDKLGETYKDHENIII AKMDSTANEVEAVKVHSFPTLKFFPASADRTVIDYN GERTLDGFKKFLESGGQDGAGDDEDLDLEEALEPD MEEDDDQKAVKDEL
21	LH3(氨基酸序列)	SDRPRGRDPVNPEKLLVITVATAETEGYLRFLRSAEF FNYTVRTLGLGEEWRGGDVARTVGGGQKVRWLKK EMEKYADREDMIIMFVDSYDVILAGSPTELLKKFVQS GSRLLFSAESFCWPEWGLAEQYPEVGTGKRFLNSGG FIGFATTIHQIVRQWKYKDDDDDQLFYTRLYLDPGL REKLSLNLDHKSRIFQNLNGALDEVVLKFDRNRVRIR NVAYDTLPIVVHGNGPTKLQLNYLGNYVPNGWTPE GGCGFCNQDRRTLPGGQPPPRVFLAVFVEQPTPFL PRFLQRLLLLDYPPDRVTLFLHNNEVFHEPHIADSW PQLQDHFSAVKLVGPEEALSPGEARDMAMDLCRQD PECEFYFSLDADAVLTNLQTLRILIEENRKVIAPMLSR HGKLWSNFWGALSPDEYYARSEDYVELVQRKRVGV WNVPYISQAYVIRGDTLRMELPQRDVFSGSDTDPD MAFCKSFRDKGIFLHLSNQHEFGRLLATSRYDTEHL HPDLWQIFDNPVDWKEQYIHENYSRALEGEGIVEQP CPDVYWFPLLSEQMCDELVAEMEHYGQWSGGRHE DSRLAGGYENVPTVDIHMKQVGYEDQWLQLLRTYV GPMTESLFPGYHTKARAVMNFVVRYRPDEQPSLRP HHDSSTFTLNVALNHKGLDYEGGGCRFLRYDCVISS PRKGWALLHPGRLTHYHEGLPTTWGTRYIMVSFVD

[0138]

		P
22	BMP1(氨基酸序列)	AATSRPERVWPDGVIPFVIGGNFTGSQRAVFRQAMR HWEKHTCVTFLERTDEDSYIVFTYRPCGCCSYVGRR GGGPQAISIGKNCDKFGIVVHELGHVVGFWHEHTRP DRDRHVSIVRENIQPGQEYNFLKMEPQEVESLGETY DFDSIMHYARNTFSRGIFLDTIVPKYEVNGVKPPIGQ RTRLSKGDIAQARKLYKCPACGETLQDSTGNFSSPEY PNGYSAHMHCVWRISVTPGEKIILNFTSLDLYRSRLC WYDYVEVRDGFWRKAPLRGRFCGSKLPEPIVSTDSR LWVEFRSSSNWVGKGFFAVYEAICGGDVKKDYGHIQ SPNYPDDYRPSKVCIWRIQVSEGFHVGLTFQSFEIER HDSCAYDYLEVRDGHSESSTLIGRYCGYEKPDDIKST SSRLWLKFVSDGSINKAGFAVNFFKEVDECSRPNRGG CEQRCLNTLGSYKCSCDPGYELAPDKRRCEAACGGF LTKLNGSITSPGWPKEYPPNKNCIWQLVAPTQYRISL QFDFFETEGNDVCKYDFVEVRSGLTADSKLHGKFCG SEKPEVITSQYNNMRVEFKSDNTVSKKGFKAHFFSEK RPALQPPRGRPHQLKFRVQKRNRTPQ
23	RNA沉默抑制子p19(氨基酸序列)	MERAIQGNDAREQANSERWDGGSGSSTSPFQLPDES PSWTEWRLHNDETNSNQDNPLGFKESWGFGKVVF KRYLRYERTETSLHRVLGSWTGDSVNYAASRFFGVN QIGCTYSIRFRGVSVTISGGSRTLQHLCEMAIRSKQEL LQLTPVEVESNVSRGCPEGVETFEEESE

[0139] 1.胶原蛋白基因表达盒的构建

[0140] 将合成基因克隆至包含CaMV 35S启动子(SEQ ID NO.1)和拟南芥ADH来源5'-UTR(SEQ ID NO.3)以及HSP终止子(SEQ ID NO.4)的表达盒中(图1a),所述合成基因编码融合于大麦淀粉酶信号肽(Barley AMY 1.2 Signal Peptide(大麦AMY1.2信号肽),SEQ ID NO.5)的胶原蛋白α-1 (I)链(Collagen alpha-1 (I) chain) (Col1,SEQ ID NO.8)。向Col1导入突变,构建了添加如下修饰的Col1的表达盒:N末端或C末端前区切断部位的修饰(基于HRV3C、弗林蛋白酶或肠激酶的切断部位的导入)、N末端前区的除去、基于GFP(SEQ ID NO.13)的N末端前区的置换等修饰。

[0141] 具体而言,N末端前区切断部位的修饰通过将N末端前区的C末端区域分别置换为HRV3C、弗林蛋白酶或肠激酶的识别序列而进行。此时,弗林蛋白酶和肠激酶切断识别序列的C末端,因此切断处理后的原胶原蛋白为具有天然N末端端肽序列的胶原蛋白。另一方面,HRV3C切断识别序列的内部,因此切断处理后的原胶原蛋白为具有添加了HRV3C的识别序列的C末端侧序列GlyPro的N末端端肽序列的胶原蛋白。在植物细胞内离天然N末端切断部位

的2氨基酸残基(GlnLeu)C末端侧的位置进行切断。作为再现所述植物细胞内的切断的修饰,通过向跨越N末端前区和N末端端肽区域的区域导入肠激酶的识别序列,从而构建得到Col1的表达盒,所述Col1的表达盒产生具有缺失了N末端的2氨基酸残基(GlnLeu)的N末端端肽序列的胶原蛋白。

[0142] 此外,就C末端前区的修饰而言,为了除去在植物细胞内切断C末端前区的未知的蛋白酶的识别序列,而构建了多个表达盒,所述表达盒在从C末端端肽区域到C末端前区的不同位置对肠激酶的识别序列进行氨基酸置换。

[0143] 将合成基因克隆至包含CaMV 35S启动子(SEQ ID NO.1)和拟南芥ADH来源5'-UTR(SEQ ID NO.3)以及HSP终止子(SEQ ID NO.4)的表达盒中(图1b),所述合成基因编码融合于大麦淀粉酶信号肽(SEQ ID NO.5)的胶原蛋白α-2 (I)链(Collagen alpha-2 (I) chain) (Col2,SEQ ID NO.11)。向Col2导入突变,构建了添加如下修饰的Col2的表达盒:N末端或C末端前区切断部位的修饰、N末端前区的除去、向N末端的绿色荧光蛋白质的融合、HN标签(SEQ ID NO.16)、GST标签(SEQ ID NO.14)、MBP标签(SEQ ID NO.15)的添加等修饰。各修饰的内容与Col1的表达盒修饰是相同的。

[0144] 3.根癌农杆菌的转化

[0145] 使用通过上述2制备得到的pRI 201-AN植物转化载体来源的质粒,通过冷冻融解法而转化根癌农杆菌GV3101株。将用各质粒进行了转化的根癌农杆菌在单独或混合2个菌株的条件下用于浸润(infiltration)。

[0146] 将植物本氏烟草(Nicotiana benthamiana)在装有市售的椰壳基质的穴盘中从种子开始培育。使所述植物在温室下以光照16小时/避光8小时的光照周期和全天25℃的温度日程的条件下生长发育。对于从接种起4周后的植栽实施了农杆菌浸润。在即将实施之前,通过从植物摘取芽,而除去了顶芽。

[0147] 2.修饰酶表达载体的构建

[0148] 将合成基因克隆至包含CaMV 35S启动子(SEQ ID NO.1)以及拟南芥ADH来源5'-UTR(SEQ ID NO.3)或者TMV omega序列(SEQ ID NO.2)、HSP来源终止子(SEQ ID NO.4)的表达盒中(图1c),所述合成基因编码融合于液胞信号(Thiol protease aleurain (Barley) vacuolar signal(巯基蛋白酶aleurain (大麦)液泡信号),SEQ ID NO.6)或者大麦淀粉酶信号肽(SEQ ID NO.5)的人P4Hα(SEQ ID NO.17)、人P4Hβ(SEQ ID NO.19)、线虫P4Hα(SEQ ID NO.18)和小鼠P4Hβ(SEQ ID NO.20)。除了在内质网中积蓄的情况,从P4Hβ基因的3'末端除去KDEL序列。

[0149] 将合成基因克隆至包含CaMV 35S启动子(SEQ ID NO.1)以及拟南芥ADH来源5'-UTR(SEQ ID NO.3)、HSP来源终止子(SEQ ID NO.4)的表达盒中(图1d),所述合成基因编码融合于液胞信号(SEQ ID NO.6)或者大麦淀粉酶信号肽(SEQ ID NO.5)的LH3(SEQ ID NO.21)。

[0150] 将编码融合有HN标签(SEQ ID NO.16)的BMP1(SEQ ID NO.22)的合成基因克隆至包含CaMV 35S启动子(SEQ ID NO.1)以及拟南芥ADH来源5'-UTR(SEQ ID NO.3)、HSP来源终止子(SEQ ID NO.4)的表达盒中(图1f),所述HN标签(SEQ ID NO.16)融合于液胞信号(SEQ ID NO.6)或者大麦淀粉酶信号肽(SEQ ID NO.5)。只有在内质网中积蓄的情况下,向BMP1基因的3'末端添加了KDEL序列。

[0151] 将合成基因(SEQ ID NO.23)克隆至包含CaMV 35S启动子(SEQ ID NO.1)以及拟南芥ADH来源5'-UTR(SEQ ID NO.3)、HSP来源终止子(SEQ ID NO.4)表达盒中(图1d),所述合成基因编码番茄丛矮病毒来源RNA沉默抑制子p19(UniProtKB - P50628)。

[0152] 将构建的各基因的表达盒以以下组合克隆至pRI 201-AN植物转化载体的多克隆位点中。

[0153] (1)Col1+p19

[0154] (2)Col2+p19

[0155] (3)Col1+Col2+P4Hα+P4Hβ+LH3+p19

[0156] (4)Col1+Col2+p19

[0157] (5)P4Hα+P4Hβ+LH3+p19

[0158] (6)Col1+P4Hα+P4Hβ+p19

[0159] (7)Col1+LH3+p19

[0160] (8)Col1+BMP1+p19

[0161] 4.从根癌农杆菌获得胶原蛋白

[0162] 将用各质粒进行了转化的根癌农杆菌在包含50 μg/ml 卡那霉素(kanamycin)的选择替代蛋白质来源(Select Alternative Protein Source(APS))培养基中在28℃下振荡培养一晚,进行增殖使其OD_{600}从4.5达到5.5。将根癌农杆菌培养液进行离心,将沉淀重悬于浸润缓冲液(10mM $MgCl_2$和10mM MES(pH5.6))中。菌液的浓度以使得OD_{600}为0.4(同时接种2个菌株的情况下为0.8)的方式进行调节而实施浸润。作为浸润后的植栽而孵化6天,仅收获浸润了根癌农杆菌的叶。收获的叶在-80℃下冷冻,粉碎为小片。

[0163] 就全可溶性蛋白质而言,对于约0.1 g的冷冻粉碎的叶,添加3体积份的50 mMTris(pH7.4)、0.15 M NaCl并磨碎,在4℃下5分钟,20000 g的条件下离心而提取得到。此外,作为另一提取条件,实施了使用3体积份的0.5 M 乙酸、0.2 M NaCl的提取。

[0164] 5.获得的胶原蛋白的分析

[0165] 对在以下条件下提取得到的胶原蛋白进行处理,通过SDS-PAGE而确认了是否切断。

[0166] 胶原蛋白提取液的蛋白酶处理:4℃ 16小时反应

[0167] HRV3C:提取液20μL+HRV3C (Millipore#71493) 1μL

[0168] 弗林蛋白酶:提取液20μL+6mM $CaCl_2$ 4μL +10倍稀释弗林蛋白酶 (R&D #1503-SE-010) 2μL

[0169] 肠激酶:提取液20μL+6mM $CaCl_2$ 10μL+10倍稀释肠激酶 (NEB#P8070S) 2μL

[0170] 6.质量分析

[0171] 通过DTT将全可溶性蛋白质还原后,通过进行使用了MOPS缓冲液的SDS-PAGE电泳从而分离质量分析样品。泳动后,通过CBB染色,进行分离蛋白质的检测时,其结果,确认了在表达了胶原蛋白的全可溶性蛋白质的约130 kDa附近有预想的胶原蛋白的条带,切出所述部分的凝胶,进行了包含的蛋白质的质量分析。质量分析向日本PROTEOMICS(株式会社)委托解析。将凝胶片段脱色后,进行胰肮酶消化,通过nanoLC-MS/MS进行质量的测定后,将数据用MascotSearch进行蛋白质鉴定(https://www.jproteomics.com/)。

[0172] 7.使用蛋白质印迹法确认烟草叶中的胶原蛋白表达

[0173] 通过使用Tris甘氨酸缓冲液的SDS-PAGE电泳而分离从烟草叶(本氏烟草)中提取的全可溶性蛋白质。泳动后,转印至PVDF膜,使用针对胶原蛋白α1和标签序列的抗体而进行蛋白质印迹法,确认了胶原蛋白α1和胶原蛋白α2的表达。

[0174] (结果)

[0175] 实施例1:植物表达系统中的分泌型的I型胶原蛋白α1的瞬时性地表达

[0176] 通过上述方法,使Col1的表达盒在本氏烟草植物表达系统瞬时性地表达,通过SDS-PAGE确认了是否切断(图2)。各泳道的样品如下所示。

[0177]

泳道	样品
1	标记物
2	阴性对照(negative control)提取液
3	原胶原蛋白 α1 提取液
4	除去 N 末端前区的胶原蛋白 α1 提取液
5	N 末端切断部位修饰(HRV3C)原胶原蛋白 α1 提取液
6	N 末端切断部位修饰(HRV3C)原胶原蛋白 α1 提取液 +HRV3C 处理
7	N 末端切断部位修饰(弗林蛋白酶)原胶原蛋白 α1 提取液
8	N 末端切断部位修饰(弗林蛋白酶)原胶原蛋白 α1 提取液 + 弗林蛋白酶处理
9	N 末端切断部位修饰(肠激酶)原胶原蛋白 α1 提取液
10	N 末端切断部位修饰(肠激酶)原胶原蛋白 α1 提取液 + 肠激酶处理
11	N 末端切断部位修饰(植物内切断位置)(肠激酶)原胶原蛋白 α1 提取液
12	N 末端切断部位修饰(植物内切断位置)(肠激酶)原胶原蛋白 α1 提取液 + 肠激酶处理
13	N 末端 GFP 结合切断部位修饰(HRV3C)原胶原蛋白 α1 提取液
14	N 末端 GFP 结合切断部位修饰(HRV3C)原胶原蛋白 α1 提取液 +HRV3C 处理
15	N 末端 GFP 结合切断部位修饰(肠激酶)原胶原蛋白 α1 提取液
16	N 末端 GFP 结合切断部位修饰(肠激酶)原胶原蛋白 α1 提取液 +HRV3C 处理
17	C 末端切断部位修饰(肠激酶)原胶原蛋白 α1 提取液
18	C 末端切断部位修饰(肠激酶)原胶原蛋白 α1 提取液 +HRV3C 处理

[0178] 其结果,确认了植物体内积蓄有具有N末端前区和C末端前区的全长胶原蛋白(图2,泳道3)。就所述结果而言,与F Ruggiero等的2000年的研究所示的植物稳定表达系统中分泌型的I型胶原蛋白α1的N末端前区和C末端前区被切断的实验结果相比,是意料之外的结果。此外,显示了可以通过将具有切断部位修饰的I型胶原蛋白α1在提取液中用针对切断部位修饰的酶进行处理,而切断I型胶原蛋白α1 N末端和C末端的前区(图2,泳道4~18)。

[0179] 实施例2:GFP融合胶原蛋白α1的获得

[0180] 通过对GFP融合胶原蛋白α1提取液的提取液照射激发光,确认了荧光。N末端和C末端的前区中,C末端前区对于胶原蛋白形成三重链是必要的,N末端前区不是必须的,实质上除去了N末端前区的序列也确认了表达。通过本实验,显示了可以通过代替N末端前区而融合GFP这样的异源肽序列,而产生赋予了新功能的胶原蛋白(图3)。

[0181] 实施例3:胶原蛋白α1和胶原蛋白α2的同时表达的确认

[0182] 通过上述方法,在植物表达系统使Col1的表达盒和Col2的表达盒进行瞬时性地共表达时,其结果,由蛋白质印迹确认了细胞体内同时积蓄α1和α2胶原蛋白(图4)。各泳道的样品如下所示。

[0183]

泳道	样品
1	阴性对照(negative control)提取液
2	原胶原蛋白 α1 提取液
3	HisTrap 纯化液 除去 N 末端前区的胶原蛋白 α1 以及 添加 HN 标签且除去 N 末端前区的胶原蛋白 α2
4	HisTrap 纯化液 原胶原蛋白 α1 以及 添加 HN 标签的原胶原蛋白 α2
5	HisTrap 纯化液 除去 N 末端前区的胶原蛋白 α1 以及 添加 MBP 标签且除去 N 末端前区的胶原蛋白 α2

[0184] 此外,通过质量分析,确认了α1和α2胶原蛋白的表达(图5)。各泳道的样品与图4相同。

[0185] 从以上的实施例意外地发现了,可以通过在植物表达系统使胶原蛋白瞬时性地表达,而稳定地获得分泌型的胶原蛋白。此外,确认了可以通过使代替N末端前区而融合了具有各种功能的肽的胶原蛋白表达,而获得高功能型胶原蛋白。

[0186]

<Seq ID No:1 CaMV 35S Promoter DNA Seq; DNA; Cauliflower mosaic virus>
TGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCAC
TTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGG
CCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGT
GGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTA
AGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCATT
TGGAGAGAACA

<Seq ID No:2 TMV omega DNA Seq; DNA; Tobacco mosaic virus>
ACAATTACCAACAACAACAAACAACAAACAACATTACAATTACTATTTACAATTAC

<Seq ID No:3 AtADH 5'-UTR DNASeq; DNA; Arabidopsis thaliana>
TACATCACAATCACACAAAACTAACAAAAGATCAAAAGCAAGTTCTTCACTGTTGATA

<Seq ID No:4 HSP-terminator DNA Seq; DNA; Arabidopsis thaliana>
TATGAAGATGAAGATGAAATATTTGGTGTGTCAAATAAAAAGCTAGCTTGTGTGCTTAAGTTTGTGT
TTTTTTCTTGGCTTGTTGTGTTATGAATTTGTGGCTTTTTCTAATATTAAATGAATGTAAGATCTCA
TTATAATGAATAAACAAATGTTTCTATAATCCATTGTGAATGTTTTGTTGGATCTCTTCGCATATAA
CTACTGTATGTGCTATGGTATGGACTATGGAATATGATTAAAGATAAG

<Seq ID No:5 Barley AMY 1.2 Signal Peptide AA Seq; AA; Hordeum vulgare>
MANKHLSLSLFLVLLGLSASLASG

<Seq ID No:6 Thiol protease aleurain (Barley) vacuolar signal AA Seq; AA; Hordeum vulgare>
MAHARVLLLALAVLATAAVAVASSSSFADSNPIRPVTDRAASTL

<Seq ID No:7 Collagen alpha-1 N-terminal propeptide AA Seq; AA; Homo sapiens>
QEEGQVEGQDEDIPPITCVQNGLRYHDRDVWKPEPCRICVCDNGKVLCDDVICDETKNCPGAEVPEG
ECCPVCPDGSESPTDQETTGVEGPKGDTGPRGPRGPAGPPGRDGIPGQPGLPGPPGPPGPPGPPGLG
GNFAP

<Seq ID No:8 Collagen alpha-1 (I) chain AA Seq; AA; Homo sapiens>

QEEGQVEGQDEDIPPITCVQNGLRYEDRDVWKPEPCRICVCDNGKVLCDDVICDETKNCPGAEVPEG
ECCPVCPDGSESPTDQETTGVEGPKGDTGPRGPRGPAGPPGRDGIPCQPGLPGPPGPPGPPGPPGLC
GNFAPQLSYGYDEKSTGGISVPGPMGPSGPRGLPGPPGAPGPQGFQGPPGEPGEPGASGPMGPRGPP
GPPGKNGDDGEAGKPGRPGERGPPGPQGARGLPGTAGLPGMKGHRGFSGLDGAKGDAGPAGPKGEPG
SPGENGAPGQMGPRGLPGERGRPGAPGPAGARGNDGATGAAGPPGPTGPAGPPGPFPGAVGAKGEAGP
QGPRGSEGPQGVRGEPGPPGPAGAAGPAGNPGADGQPGAKGANGAPGIAGAPGFPGARGPSGPQGPG
GPPGPKGNSGEPGAPGSKGDTGAKGEPGPVGVQGPPGPAGEEGKRGARGEPGPTGLPGPPGERGGPG
SRGFPGADGVAGPKGPAGERGSPGPAGPKGSPGEAGRPGEAGLPGAKGLTGSPGSPGPDGKTGPPGP
AGQDGRPGPPGPPGARGQAGVMGFPGPKGAAGEPGKAGERGVPGPPGAVGPAGKDGEAGAQGPPGPA
GPAGERGEQGPAGSPGPQGLPGPAGPPGEAGKPGEQGVPGDLGAPGPSGARGERGFPGERGVQGPPG
PAGPRGANGAPGNDGAKGDAGAPGAPGSQGAPGLQGMPGERGAAGLPGPKGDRGDAGPKGADGSPGK
DGVRGLTGPIGPPGPAGAPGDKGESGPSGPAGPTGARGAPGDRGEPGPPGPAGFAGPPGADGQPGAK
GEPGDAGAKGDAGPPGPAGPAGPPGPIGNVGAPGAKGARGSAGPPGATGFPGAAGRVGPPGPSGNAG
PPGPPGPAGKEGGKGPRGETGPAGRPGEVGPPGPPGPAGEKGSPGADGPAGAPGTPGPQGIAGQRGV
VGLPGQRGERGFPGLPGPSGEPGKQGPSGASGERGPPGPMGPPGLAGPPGESGREGAPGAEGSPGRD
GSPGAKGDRGETGPAGPPGAPGAPGAPGPVGPAGKSGDRGETGPAGPTGPVGPVGARGPAGPQGPRG
DKGETGEQGDRGIKGHRGFSGLQGPPGPPGSPGEQGPSGASGPAGPRGPPGSAGAPGKDGLNGLPGP
IGPPGPRGRTGDAGPVGPPGPPGPPGPPSAGFDFSFLPQPPQEKAHDGGRYYRADDANVVRDRD
LEVDTTLKSLSQQIENIRSPEGSRKNPARTCRDLKMCHSDWKSGEYWIDPNQGCNLDAIKVFCNMET
GETCVYPTQPSVAQKNWYISKNPKDKRHVWFGESMTDGFQFEYGGQGSDPADVAIQLTFLRLMSTEA
SQNITYHCKNSVAYMDQQTGNLKKALLLQGSNEIEIRAEGNSRFTYSVTVDGCTSHTGAWGKTVIEY
KTTKTSRLPIIDVAPLDVGAPDQEFGFDVGPVCFL

[0187]

<Seq ID No:9 Collagen alpha-1 C-terminal propeptide AA Seq; AA; synthetic construct>

DDANVVRDRDLEVDTTLKSLSQQIENIRSPEGSRKNPARTCRDLKMCHSDWKSGEYWIDPNQGCNLD
AIKVFCNMETGETCVYPTQPSVAQKNWYISKNPKDKRHVWFGESMTDGFQFEYGGQGSDPADVAIQL
TFLRLMSTEASQNITYHCKNSVAYMDQQTGNLKKALLLQGSNEIEIRAEGNSRFTYSVTVDGCTSHT
GAWGKTVIEYKTTKTSRLPIIDVAPLDVGAPDQEFGFDVGPVCFL

<Seq ID No:10 Collagen alpha-2 N-terminal propeptide AA Seq; AA; synthetic construct>

QSLQEETVRKGPAGDRGPRGERGPPGPPGRDGEDGPTGPPGPPGPPGPPGLGGNFAA

<Seq ID No:11 Collagen alpha-2 (I) chain AA seq; AA; Homo sapiens>

QSLQEETVRKGPAGDRGPRGERGPPGPPGRDGEDGPTGPPGPPGPPGPPGLGGNFAAQYDGKGVGLG

PGPMGLMGPRGPPGAAGAPGPQGFQGPAGEPGEPGQTGPAGARGPAGPPGKAGEDGHPGKPGRPGER

GVVGPQGARGFPGTPGLPGFKGIRGHNGLDGLKGQPGAPGVKGEPGAPGENGTPGQTGARGLPGERG

RVGAPGPAGARGSDGSVGPVGPAGPIGSAGPPGFPGAPGPKGEICAVGNAGPAGPAGPRGEVGLPGL

SGPVGPPGNPGANGLTGAKGAAGLPGVAGAPGLPGPRGIPGPVGAAGATGARGLVGEPGPAGSKGES

GNKGEPGSAGPQGPPGPSGEEGKRGPNGEAGSAGPPGPPGLRGSPGSRGLPGADGRAGVMGPPGSRG

ASGPAGVRGPNGDAGRPGEPGLMGPRGLPGSPGNIGPAGKEGPVGLPGIDGRPGPIGPAGARGEPGN

IGFPGPKGPTGDPGKNGDKGEAGLAGARGAPGPDGNNGAQGPPGPQGVQGGKGEQGPPGPPGFQGLP

GPSGPAGEVGKPGERGLHGEFGLPGPAGPRGERGPPGESGAAGPTGPIGSRGPSGPPGPDGNKGEPG

VVGAVGTAGPSGPSGLPGERGAAGIPGGKGEKGEPGLRGEIGNPGRDGARGAPGAVGAPGPAGATGD

RGEAGAAGPAGPAGPRGSPGERGEVGPAGPNGFAGPAGAAGQPGAKGERGAKGPKGENGVVGPTGPV

GAAGPAGPNGPPGPAGSRGDGGPPGMTGFPGAAGRTGPPGPSGISGPPGPPGPAGKEGLRGPRGDQG

PVGRTGEVGAVGPPGFAGEKGPSGEAGTAGPPGTPGPQGLLGAPGILGLPGSRGERGLPGVAGAVGE

PGPLGIAGPPGARGPPGAVGSPGVNGAPGEAGRDGNPGNDGPPGRDGQPGHKGERGYPGNIGPVGAA

GAPGPHGPVGPAGKHGNRGETGPSGPVGPAGAVGPRGPSGPQGIRGDKGEPGEKGPRGLPGLKGHNG

LQGLPGIAGHHGDQGAPGSVGPAGPRGPAGPSGPAGKDGRTGHPGTVGPAGIRGPQGHQGPAGPPGP

PGPPGPPGVSGGGYDFGYDGDFYRADQPRSAPSLRPKDYEVDATLKSLNNQIETLLTPEGSRKNPAR

TCRDLRLSHPEWSSGYYWIDPNQGCTMDAIKVYCDFSTGETCIRAQPENIPAKNWYRSSKDKKHVWL

GETINAGSQFEYNVEGVTSKEMATQLAFMRLLANYASQNITYHCKNSIAYMDEETGNLKKAVILQGS

NDVELVAEGNSRFTYTVLVDGCSKKTNEWGKTIIEYKTNKPSRLPFLDIAPLDIGGADQEFFVDIGP

VCFK

[0188]

<Seq ID No:12 Collagen alpha-2 C-terminal propeptide AA Seq;AA; synthetic construct>

DQPRSAPSLRPKDYEVDATLKSLNNQIETLLTPEGSRKNPARTCRDLRLSHPEWSSGYYWIDPNQGC

TMDAIKVYCDFSTGETCIRAQPENIPAKNWYRSSKDKKHVWLGETINAGSQFEYNVEGVTSKEMATQ

LAFMRLLANYASQNITYHCKNSIAYMDEETGNLKKAVILQGSNDVELVAEGNSRFTYTVLVDGCSKK

TNEWGKTIIEYKTNKPSRLPFLDIAPLDIGGADQEFFVDIGPVCFK

<Seq ID No:13 GFP AA Seq; AA; synthetic construct>

SKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFSYGV

QCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGN

ILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLS

TQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK

<Seq ID No:14 GST AA Seq;AA; synthetic construct>

SPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQS

MAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFED
RLCHKTYLNGDEVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWP
LQGWQATFGGGDHPPK

<Seq ID No:15 MBP AA Seq; AA; Escherichia coli>
KIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRF
GGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPA
LDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKN
KHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINA
ASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELVKDPRIAATMENAQKGEIMPNIPQ
MSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNNNNNLG

<Seq ID No:16 HN Tag AA Seq; AA; synthetic construct>
HNHNHNHNHNHN

<Seq ID No:17 Human P4H alpha subunit AA Seq; AA; Homo sapiens>
HPGFFTSIGQMTDLIHTEKDLVTSLKDYIKAEEDKLEQIKKWAEKLDRLTSTATKDPEGFVGHPVNA
FKLMKRLNTEWSELENLVLKDMSDGFISNLTLQRQYFPNDEDQVGAAKALLRLQDTYNLDTDTISKG
[0189] NLPGVKHKSFLTAEDCFELGKVAYTEADYYHTELWMEQALRQLDEGEISTIDKVSVLDYLSYAVYQQ
GDLDKALLLTKKLLELDPEHQRANGNLKYFEYIMAKEKDVNKSASDDQSDQKTTPKKKGVAVDYLPE
RQKYEMLCRGEGIKMTPRRQKKLFCRYHDGNRNPKFILAPAKQEDEWDKPRIIRFHDIISDAEIEIV
KDLAKPRLRRATISNPITGDLETVHYRISKSAWLSGYENPVVSRINMRIQDLTGLDVSTAEELQVAN
YGVGGQYEPHFDFARKDEPDAFKELGTGNRIATWLFYMSDVSAGGATVFPEVGASVWPKKGTAVFWY
NLFASGEGDYSTRHAACPVLVGNKWVSNKWLHERGQEFRRPCTLSELE

<Seq ID No:18 C. elegans P4H alpha subunit AA Seq; AA; Caenorhabditis
elegans>
DLFTSIADMQNLLETRNIPKILDKYIHDEEERLVQLKKLSEEYSKKNEISIENGLKDITNPINAFL
LIKRKTFDWKEIESKMNANKAGNVVSSITDDSYGVRYPTADDISGAATGLLRLQDTYRLDTKDLADG
KIYADQGNYCFSAKDCFEIARAAYNEHDFYHCVMWMEEAQRRLGDEVEPTVEVEDILEYLAFALYKQ
NNLKHALKLTEELYKMNPIHPRAKGNVKWYEDLLEQEGVRRSDMRKNLPPIQNRRPDSVLGNTERTM
YEALCRNEVPVSQKDISRLYCYYKRDRPFLVYAPIKVEIKRFNPLAVLFKDVISDDEVAAIQELAKP
KLARATVEDSVTGKLVTATYRISKSAWLKEWEGDVVECVNKRTGYMCNLEMETAEELQIANYGIGGH
YDPHFDHAKKEESKSFESLGTGNRIATVLFYMSQPSHGGGTVFTEAKSTILPTKNDALFWYNLYKQG
DGNPDTRFAACPVLVGIKWVSNKWIHEKGNEFRRPCGLKSSDYEREVGDLGYGPEPRNAPNVSPNLA
KDVWETL

<Seq ID No:19 Human P4H beta subunit AA Seq; AA; Homo sapiens>
DAPEEEDHVLVLRKSNFAEALAAHKYLLVEFYAPWCGHCKALAPEYAKAAGKLKAEGSEIRLAKVDA
TEESDLAQQYGVRGYPTIKFFRNGDTASPKEYTAGREADDIVNWLKKRTGPAATTLPDGAAAESLVE
SSEVAVIGFFKDVESDSAKQFLQAAEAIDDIPFGITSNSDVFSKYQLDKDGVVLFKKFDEGRNNFEG
EVTKENLLDFIKHNQLPLVIEFTEQTAPKIFGGEIKTHILLFLPKSVSDYDGKLSNFKTAAESFKGK
ILFIFIDSDHTDNQRILEFFGLKKEECPAVRLITLEEEMTKYKPESEELTAERITEFCHRFLEGKIK
PHLMSQELPEDWDKQPVKVLVGKNFEDVAFDEKKNVFVEFYAPWCGHCKQLAPIWDKLGETYKDHEN
IVIAKMDSTANEVEAVKVHSFPTLKFFPASADRTVIDYNGERTLDGFKKFLESGGQDGAGDDDDLED
LEEAEEPDMEEDDDQKAV

<Seq ID No:20 Mouse P4H beta subunit AA Seq; AA; Mus musculus>
DALEEEDNVLVLKKSNFEEALAAHKYLLVEFYAPWCGHCKALAPEYAKAAAKLKAEGSEIRLAKVDA
TEESDLAQQYGVRGYPTIKFFKNGDTASPKEYTAGREADDIVNWLKKRTGPAATTLSDTAAAESLVD
SSEVTVIGFFKDVESDSAKQFLLAAEAIDDIPFGITSNSGVFSKYQLDKDGVVLFKKFDEGRNNFEG
EITKEKLLDFIKHNQLPLVIEFTEQTAPKIFGGEIKTHILLFLPKSVSDYDGKLSSFKRAAEGFKGK
ILFIFIDSDHTDNQRILEFFGLKKEECPAVRLITLEEEMTKYKPESDELTAEKITEFCHRFLEGKIK
PHLMSQEVPEDWDKQPVKVLVGANFEEVAFDEKKNVFVEFYAPWCGHCKQLAPIWDKLGETYKDHEN
[0190] IIIAKMDSTANEVEAVKVHSFPTLKFFPASADRTVIDYNGERTLDGFKKFLESGGQDGAGDDEDLDL
EEALEPDMEEDDDQKAVKDEL

<Seq ID No:21 LH3 AA Seq; AA; Homo sapiens>
SDRPRGRDPVNPEKLLVITVATAETEGYLRFLRSAEFFNYTVRTLGLGEEWRGGDVARTVGGGQKVR
WLKKEMEKYADREDMIIMFVDSYDVILAGSPTELLKKFVQSGSRLLFSAESFCWPEWGLAEQYPEVG
TGKRFLNSGGFIGFATTIHQIVRQWKYKDDDDDQLFYTRLYLDPGLREKLSLNLDHKSRIFQNLNGA
LDEVVLKFDRNRVRIRNVAYDTLPIVVHGNGPTKLQLNYLGNYVPNGWTPEGGCGFCNQDRRTLPGG
QPPPRVFLAVFVEQPTPFLPRFLQRLLLLDYPPDRVTLFLHNNEVFHEPHIADSWPQLQDHFSAVKL
VGPEEALSPGEARDMAMDLCRQDPECEFYFSLDADAVLTNLQTLRILIEENRKVIAPMLSRHGKLWS
NFWGALSPDEYYARSEDYVELVQRKRVGVWNVPYISQAYVIRGDTLRMELPQRDVFSGSDTDPDMAF
CKSFRDKGIFLHLSNQHEFGRLLATSRYDIEHLHPDLWQIFDNPVDWKEQYIHENYSRALEGEGIVE
QPCPDVYWFPLLSEQMCDELVAEMEHYGQWSGGRHEDSRLAGGYENVPTVDIHMKQVGYEDQWLQLL
RTYVGPMTESLFPGYHTKARAVMNFVVRYRPDEQPSLRPHHDSSTFTLNVALNHKGLDYEGGGCRFL
RYDCVISSPRKGWALLHPGRLTHYHEGLPITWGTRYIMVSFVDP

<Seq ID No:22 BMP1 AA seq; AA; Homo sapiens>
AATSRPERVWPDGVIPFVIGGNFTGSQRAVFRQAMRHWEKHTCVTFLERTDEDSYIVFTYRPCGCCS

YVGRRGGPQAISIGKNCDKFGIVVHELGHVVGFWEEHTRPDRDRHVSIVRENIQPGQEYNFLKMEP
QEVESLGETYDFDSIMHYARNTFSRGIFLDTIVPKYEVNGVKPPIGQRTRLSKGDIAQARKLYKCPA
CGETLQDSTGNFSSPEYPNGYSAEMHCVWRISVTPGEKIILNFTSLDLYRSRLCWYDYVEVRDGFWR
KAPLRGRFCGSKLPEPIVSTDSRLWVEFRSSSNWVGKGFFAVYEAICGGDVKKDYGHIQSPNYPDDY
RPSKVCIWRIQVSEGFHVGLTFQSFEIERHDSCAYDYLEVRDGHSESSTLIGRYCGYEKPDDIKSTS
SRLWLKFVSDGSINKAGFAVKFFKEVDECSRPNRGGCEQRCLNTLGSYKCSCDPGYELAPDKRRCEA
ACGGFLTKLNGSITSPGWPKEYPPNKNCIWQLVAPTQYRISLQFDFFETEGNDVCKYDFVEVRSGLT
ADSKLHGKFCGSEKPEVITSQYNNMRVEFKSDNTVSKKGFKAHFFSEKRPALQPPRGRPEQLKFRVQ
KRNRTPQ

<Seq ID No:23 RNA silencing suppressor p19 AA Seq; AA; Tomato bushy stunt virus>

[0191] MERATQGNDAREQANSERWDGGSGSSTSPFQLPDFSPSWTFWRLHNDFTNSNQDNPLGFKFSWGFGK
VVFKRYLRYERTFTSLHRVLGSWTGDSVNYAASRFFGVNQTGCTYSIRFRGVSVTTSGGSRTLQHLC
EMAIRSKQFLLQLTPVFVFSNVSRGCPFGVFTFEEFSR

<Seq ID No:24 HRV3C recognition sequence ; AA; synthetic construct>
LEVLFQGP

<Seq ID No:25 Furin recognition sequence 1; AA; synthetic construct>
RXRR

<Seq ID No:26 Furin recognition sequence 2; AA; synthetic construct>
RXKR

<Seq ID No:27 Enterokinase recognition sequence; AA; synthetic construct>
DDDDK

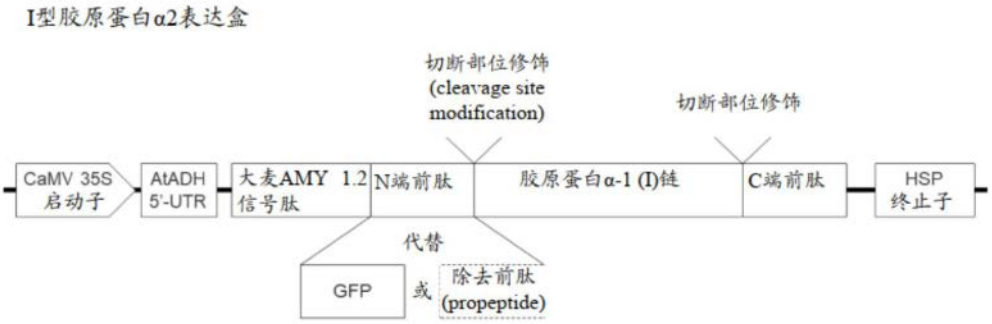

I型胶原蛋白α2表达盒
切断部位修饰
(cleavage site
modification)
切断部位修饰
CaMV 35S
启动子
AtADH
5'-UTR
大麦AMY 1.2
信号肽
N端前肽
胶原蛋白α-1 (I)链
C端前肽
HSP
终止子
代替
GFP
或
除去前肽
(propeptide)

图1a

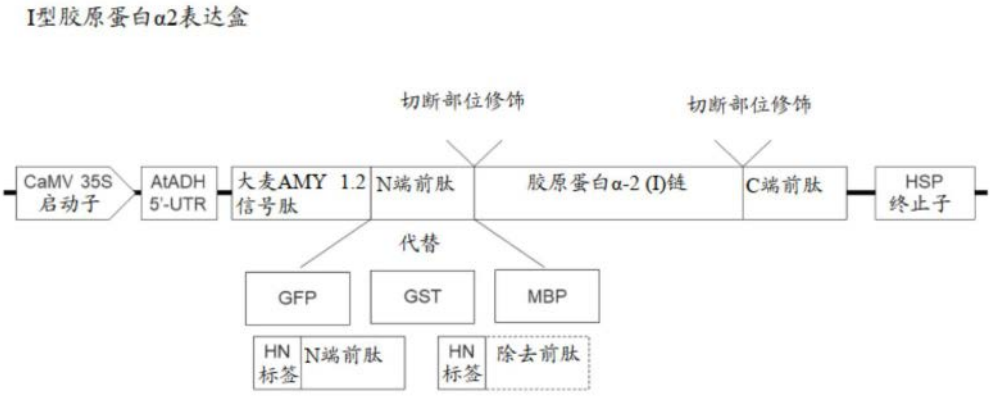

I型胶原蛋白α2表达盒
切断部位修饰
切断部位修饰
CaMV 35S
启动子
AtADH
5'-UTR
大麦AMY 1.2
信号肽
N端前肽
胶原蛋白α-2 (I)链
C端前肽
HSP
终止子
代替
GFP
GST
MBP
HN
标签
N端前肽
HN
标签
除去前肽

图1b

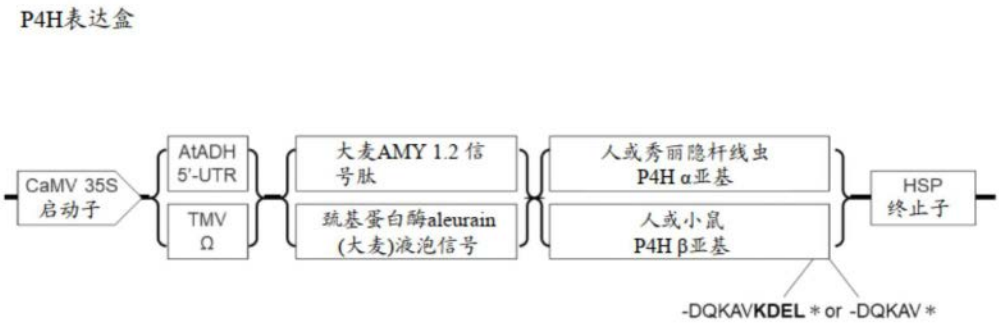

P4H表达盒
CaMV 35S
启动子
AtADH
5'-UTR
TMV
Ω
大麦AMY 1.2 信
号肽
疏基蛋白酶aleurain
(大麦)液泡信号
人或秀丽隐杆线虫
P4H α亚基
人或小鼠
P4H β亚基
HSP
终止子
-DQKAVKDEL＊or -DQKAV＊

图1c

LH表达盒

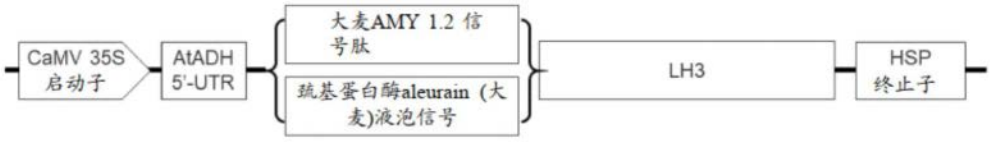


图1d

BMP-1表达盒

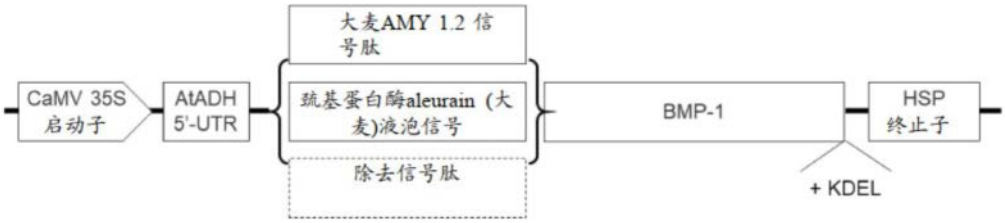


图1e

P19表达盒

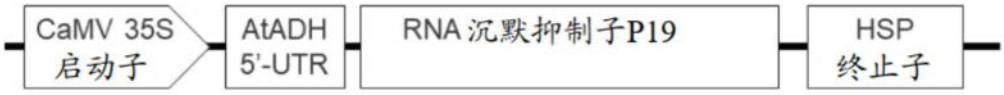
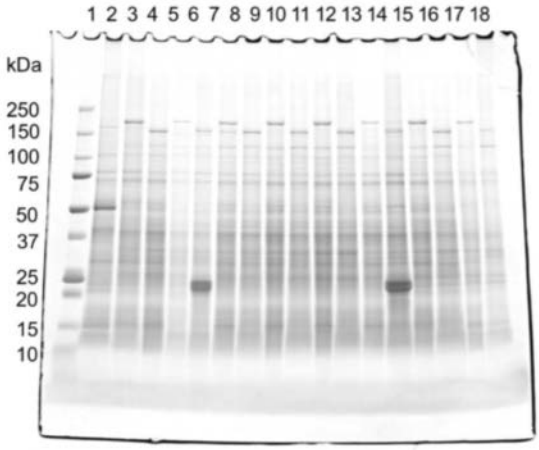


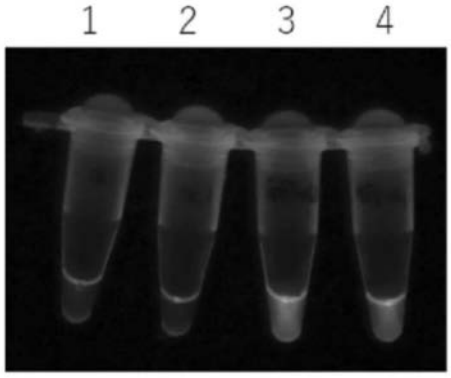
图1f



泳道	样品
1	标记物
2	阴性对照提取液
3	原胶原蛋白α1提取液
4	除去N末端前区的胶原蛋白α1提取液
5	N末端切断部位修饰(HRV3C)原胶原蛋白α1提取液
6	泳道5提取液+ HRV3C处理
7	N末端切断部位修饰(弗林蛋白酶)原胶原蛋白α1提取液
8	泳道7提取液+ 弗林蛋白酶处理
9	N末端切断部位修饰(肠激酶)原胶原蛋白α1提取液
10	泳道9提取液+ 肠激酶处理
11	N末端切断部位修饰(肠激酶)原胶原蛋白α1提取液
12	泳道11提取液+ 肠激酶处理
13	N末端GFP融合(HRV3C)胶原蛋白α1提取液
14	泳道13提取液+ HRV3C处理
15	N末端GFP融合(肠激酶)胶原蛋白α1提取液
16	泳道15提取液+ 肠激酶处理
17	C末端切断部位修饰原胶原蛋白α1提取液
18	泳道17提取液+ 肠激酶处理

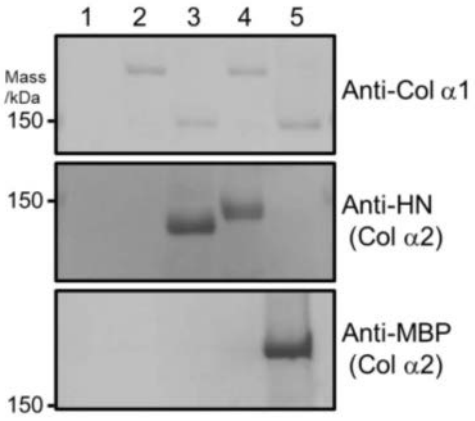
确认一系列的修饰胶原蛋白的基于蛋白酶的切断
各修饰胶原蛋白的构造参照图1a

图2



1,2：胶原蛋白α1提取液
3,4：GFP融合胶原蛋白α1提取液

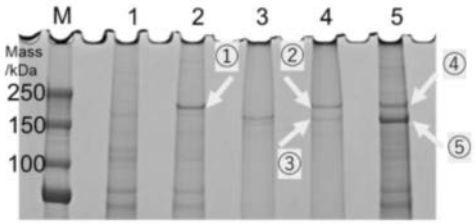
图3



泳道	样品
1	阴性对照提取液
2	原胶原蛋白α1提取液
3	除去N末端前区的胶原蛋白α1 以及 添加HN标签并除去N末端前区的胶原蛋白α2提取液
4	原胶原蛋白α1 以及 添加HN标签的原胶原蛋白α2提取液
5	除去N末端前区的胶原蛋白α1 以及 添加MBP标签并除去N末端前区的胶原蛋白α2提取液

通过蛋白质印记法确认了修饰胶原蛋白α1和α2的共表达。
各修饰胶原蛋白的构造参照图1a和b

图4



泳道	样品
1	阴性对照提取液
2	原胶原蛋白α1提取液
3	HisTrap纯化液 除去N末端前区的胶原蛋白α1以及 添加HN标签并除去N末端前区的胶原蛋白α2
4	HisTrap纯化液 原胶原蛋白α1以及 添加HN标签的原胶原蛋白α2
5	HisTrap纯化液 除去N末端前区的胶原蛋白α1以及 添加MBP标签并除去N末端前区的胶原蛋白α2

	MS检测出的肽
①	胶原蛋白α-1 (I)链
②	胶原蛋白α-1 (I)链
③	细胞质肌动蛋白1 胶原蛋白α-2 (I)链 热休克蛋白β-1 组蛋白H4 组蛋白H2B 2-F型
④	胶原蛋白α-2 (I)链 胶原蛋白α-1 (I)链
⑤	胶原蛋白α-1 (I)链

通过质量分析确认了修饰胶原蛋白α1和α2的共表达。各修饰胶原蛋白的构造参照图1a和b

图5