



(12)发明专利

(10)授权公告号 CN 103849672 B

(45)授权公告日 2017.06.06

(21)申请号 201310658887.9

A01H 5/00(2006.01)

(22)申请日 2013.12.06

(56)对比文件

(65)同一申请的已公布的文献号  
申请公布号 CN 103849672 A

CN 101796182 A,2010.08.04,  
CN 102762738 A,2012.10.31,

(43)申请公布日 2014.06.11

Hongmei Luo et al..Analysis of the transcriptome of Panax notoginseng root uncovers putative triterpene saponin-biosynthetic genes and genetic markers.《BMC Genomics》.2011,第12卷1-15.

(66)本国优先权数据

201210520787.5 2012.12.06 CN  
201310227689.7 2013.06.07 CN

Hongmei Luo et al..Analysis of the transcriptome of Panax notoginseng root uncovers putative triterpene saponin-biosynthetic genes and genetic markers.《BMC Genomics》.2011,第12卷1-15.

(73)专利权人 中国科学院上海生命科学研究院  
地址 200031 上海市徐汇区岳阳路319号

(72)发明人 周志华 严兴 范云 王平平  
魏维 魏勇军 张珺 岳建民

(74)专利代理机构 上海一平知识产权代理有限公司 31266

代理人 祝莲君 刘真真

Luo et al..ACCESSION:GU997660,Panax notoginseng clone 13895n glycosyltransferase mRNA, complete cds.《Genbank》.2012,Features 和 Origin.

(51)Int.Cl.

C12P 33/02(2006.01)  
C12P 33/00(2006.01)  
C12N 9/10(2006.01)  
C12N 15/54(2006.01)  
C12N 15/63(2006.01)  
C12N 1/21(2006.01)  
C12N 1/19(2006.01)  
C12N 5/10(2006.01)

Luo et al..ACCESSION:GU997660,Panax notoginseng clone 13895n glycosyltransferase mRNA, complete cds.《Genbank》.2012,Features 和 Origin.

审查员 李翠莹

权利要求书3页 说明书44页  
序列表24页 附图20页

(54)发明名称

一组糖基转移酶及其应用

化合物的C-3位的第一个糖基上以延伸糖链。本发明糖基转移酶还可应用于构建人工合成稀有人参皂苷及多种新人参皂苷及其衍生物。

(57)摘要

本发明涉及一组糖基转移酶及其应用。具体地,提供了糖基转移酶gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7、3GT1、3GT2、3GT3、3GT4及其衍生多肽在萜类化合物糖基化催化及新皂苷合成中的应用,其中,糖基转移酶能特异和高效地催化四环三萜化合物底物的C-20位和/或C-6位和/或C-3的羟基糖基化,和/或将来自糖基供体的糖基转移到四环三萜类

CN 103849672 B

1. 一种体外糖基化方法,其特征在于,包括步骤:

在糖基转移酶存在下,将糖基供体的糖基转移到四环三萜类化合物的以下位点上:  
C-20位、C-6位;

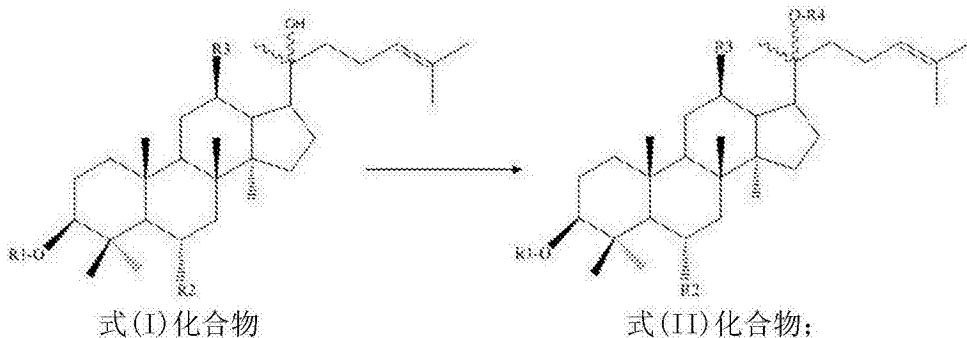
从而形成糖基化的四环三萜类化合物;

其中,所述的糖基转移酶选自:

如SEQ ID NOs.:2、16、18或20所示的糖基转移酶。

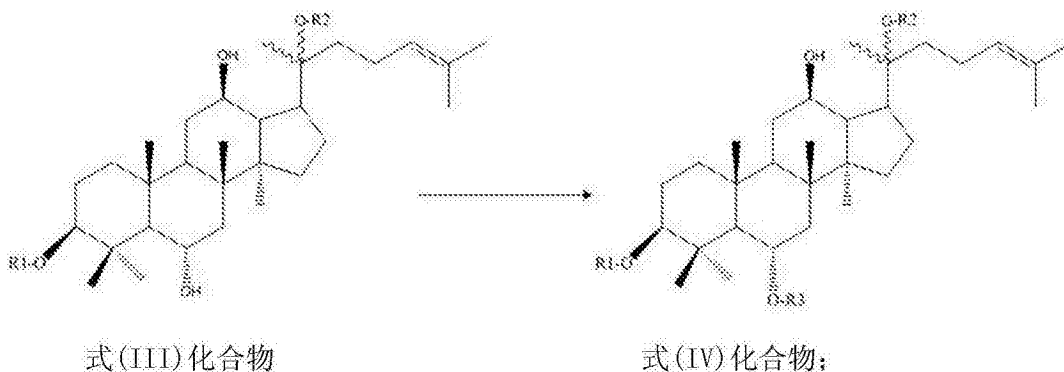
2. 如权利要求1所述的方法,其特征在于,所述糖基转移酶用于催化下述一种或多种反应或被用于制备催化下述一种或多种反应的催化制剂:

(A)



其中,R1为H,单糖糖基或多糖糖基;R2和R3为H或OH;R4为糖基;所述的糖基转移酶选自SEQ ID NOs.:2、16或18;

(B)



其中,R1为H或者糖基,R2为糖基,R3为糖基,所述的糖基转移酶选自SEQ ID NOs.:2、16、18或20;

或,R1为H或者糖基;R2为H或者糖基;R3为糖基,所述的糖基转移酶选自SEQ ID NO.:20。

3. 如权利要求2所述的方法,其特征在于,所述糖基转移酶催化的反应底物为式(I)或(III)化合物,且产物为(II)或(IV)化合物。

4. 如权利要求3所述的方法,其特征在于,所述的式(I)化合物为原人参二醇PPD,并且式(II)化合物为人参皂苷CK;

或,所述的式(I)化合物为人参皂苷Rh2,并且式(II)化合物为人参皂苷F2;

或,所述的式(I)化合物为人参皂苷Rg3,并且式(II)化合物为人参皂苷Rd;

或,所述的式(I)化合物为原人参三醇PPT,并且式(II)化合物为人参皂苷F1;

或,所述的式(I)化合物为达玛烯二醇DM,并且式(II)化合物为人参皂昔20-O-β-(D-吡喃葡萄糖基)-达玛烯二醇;

或,所述的式(III)化合物为原人参三醇PPT,并且式(IV)化合物为人参皂昔Rh1;

或,所述的式(III)化合物为人参皂昔F1,并且式(IV)化合物为人参皂昔Rg1。

5.一种分离的多肽,其特征在于,所述的多肽选自下组:

(a)如SEQ ID NO.:20所示氨基酸序列的多肽;(b)将SEQ ID NO.:20所示氨基酸序列的多肽添加信号肽序列后形成的、并具有糖基转移酶活性的多肽。

6.如权利要求5所述的多肽,其特征在于,所述的多肽还为选自下组的一种或多种额外多肽:

如SEQ ID NOs.:2、16或18所示氨基酸序列的多肽。

7.一种分离的多核苷酸,其特征在于,所述的多核苷酸选自下组:

(A) 编码权利要求5所述多肽的核苷酸序列;

(B) 编码如SEQ ID NO.:20所示多肽的核苷酸序列;

(C) 如SEQ ID NO.:19所示的核苷酸序列;

(F) 与(A)-(C)任一所述的核苷酸序列互补的核苷酸序列。

8.如权利要求7所述的多核苷酸,其特征在于,所述的多核苷酸还为选自下组的多核苷酸:

(A) 编码权利要求6所述多肽的核苷酸序列;

(B) 编码如SEQ ID NOs.:2、16或18所示多肽的核苷酸序列;

(C) 如SEQ ID NOs.:1、15或17所示的核苷酸序列。

9.一种载体,其特征在于,所述的载体含有权利要求7或8所述的多核苷酸。

10.权利要求5所述分离的多肽的用途,其特征在于,它被用于催化以下反应,或被用于制备催化以下反应的催化制剂:

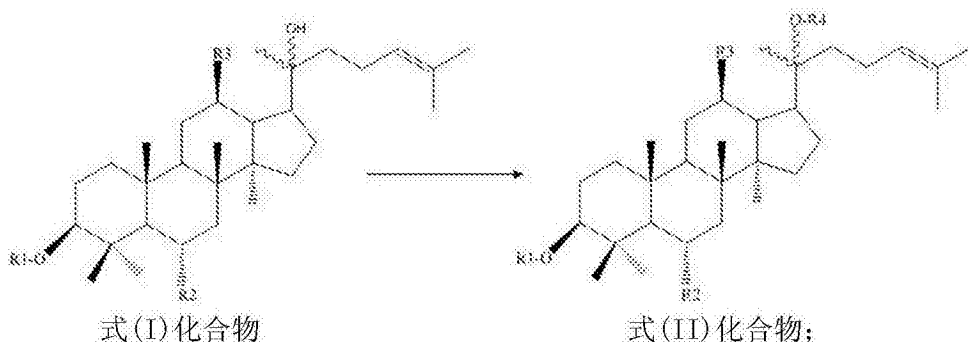
(ii) 将来自糖基供体的糖基转移到四环三萜类化合物的C-6位的羟基。

11.权利要求6所述分离的多肽的用途,其特征在于,它被用于催化以下反应,或被用于制备催化以下反应的催化制剂:

(ii) 将来自糖基供体的糖基转移到四环三萜类化合物的C-20位的羟基。

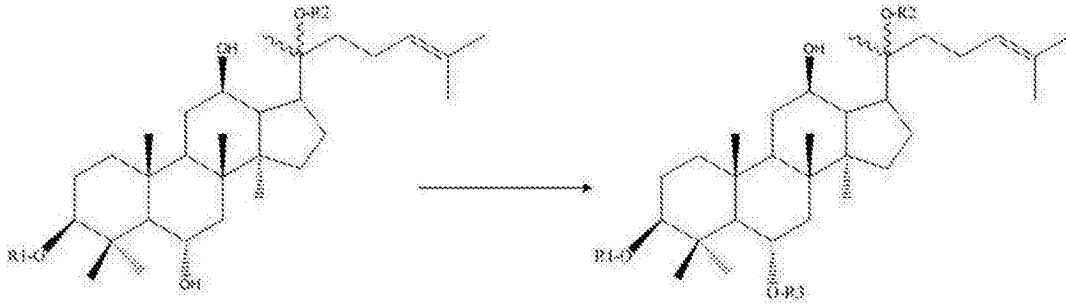
12.如权利要求10或11所述的用途,其特征在于,所述分离的多肽用于催化下述一种或多种反应或被用于制备催化下述一种或多种反应的催化制剂:

(A)



其中,R1为H,单糖糖基或多糖糖基;R2和R3为H或OH;R4为糖基;所述的多肽选自SEQ ID NOs.:2、16或18;

(B)



式(III)化合物

式(IV)化合物；

其中,R1为H或者糖基,R2为糖基,R3为糖基,所述的多肽选自SEQ ID NOs.:2、16、18或20;

或,R1为H或者糖基;R2为H或者糖基;R3为糖基,所述的多肽选自SEQ ID NO.:20。

13.一种进行糖基催化反应的方法,其特征在于,包括步骤:在权利要求5或6所述的多肽存在的条件下,进行糖基催化反应。

14.如权利要求13所述的方法,其特征在于,所述糖基催化反应的底物为式(I)或(III)化合物,且产物为(II)或(IV)化合物。

15.如权利要求14所述的方法,其特征在于,所述的式(I)化合物为原人参二醇PPD,并且式(II)化合物为人参皂苷CK;

或,所述的式(I)化合物为人参皂苷Rh2,并且式(II)化合物为人参皂苷F2;

或,所述的式(I)化合物为人参皂苷Rg3,并且式(II)化合物为人参皂苷Rd;

或,所述的式(I)化合物为原人参三醇PPT,并且式(II)化合物为人参皂苷F1;

或,所述的式(I)化合物为达玛烯二醇DM,并且式(II)化合物为人参皂苷20-O-β-(D-吡喃葡萄糖基)-达玛烯二醇;

或,所述的式(III)化合物为原人参三醇PPT,并且式(IV)化合物为人参皂苷Rh1;

或,所述的式(III)化合物为人参皂苷F1,并且式(IV)化合物为人参皂苷Rg1。

16.一种遗传工程化的宿主细胞,其特征在于,所述的宿主细胞含有权利要求9所述的载体,或其基因组中整合有权利要求7或8所述的多核苷酸。

17.权利要求16所述的宿主细胞的用途,其特征在于,用于制备酶催化试剂,或生产糖基转移酶、或作为催化细胞、或产生糖基化的四环三萜类化合物。

18.一种产生转基因植物的方法,其特征在于,包括步骤:将权利要求16所述的遗传工程化的宿主细胞再生为植物,并且所述的遗传工程化的宿主细胞为植物细胞。

19.一种分离的多肽,其特征在于,所述的多肽选自下组:

(a)如SEQ ID NOs.:2、16或18所示氨基酸序列的多肽。

20.一种分离的多核苷酸,其特征在于,所述的多核苷酸选自下组:

(A)编码如SEQ ID NOs.:2、16或18所示多肽的核苷酸序列;

(B)与(A)的核苷酸序列互补的核苷酸序列。

## 一组糖基转移酶及其应用

### 技术领域

[0001] 本发明涉及生物技术和植物生物学领域,具体地,本发明涉及一种糖基转移酶及其应用。

### 背景技术

[0002] 人参皂苷是从人参及其同属植物(如三七、西洋参等)中分离到的皂苷的总称,属于三萜类皂苷,是人参中的主要有效成份。目前,已经从人参中分离出了至少60种皂苷,其中一些人参皂苷被证实具有广泛的生理功能和药用价值:包括抗肿瘤、免疫调节、抗疲劳、护心、护肝等功能。

[0003] 从结构上来看,人参皂苷是皂苷元经过糖基化后形成的生物活性小分子。人参皂苷的皂苷元只有有限的几种,主要是达玛烷型的原人参二醇和原人参三醇,以及齐墩果烷酸。近年来,人们又从三七中分离到了新的皂苷元,25-OH-PPD和25-OCH<sub>3</sub>-PPD,这些新的皂苷元都具有非常好的抗肿瘤活性。

[0004] 皂苷元通过糖基化后,可以提高水溶性,并产生出不同的生理活性。原人参二醇型皂苷的糖链一般结合在皂苷元的C3(和)或C20的羟基上。原人参三醇皂苷与原人参二醇皂苷相比,在C6位多一个羟基,目前发现的原人参三醇型皂苷都是在C6(和)或C20的羟基上结合糖基化,在C3上结合糖基的原人参三醇型皂苷尚未见报道。糖基可以是葡萄糖、鼠李糖、木糖、和阿拉伯糖。

[0005] 不同的糖基结合位点,糖链组成和长度使人参皂苷在生理功能和药用价值上产生极大的差异。例如,人参皂苷Rb1,Rd和Rc都是以原人参二醇为皂苷元的皂苷,它们之间的差别只是糖基修饰上的差别,但它们之间的生理功能就有很多的差别。Rb1有稳定中心神经系统的功能,而Rc的功能却是抑制中心神经系统的功能,Rb1的生理功能非常广泛,而Rd却只有非常有限的几种功能。

[0006] 人参皂苷皂苷元和皂苷之间的结构多样性还体现在空间立体结构上,虽然四环三萜的骨架上存在很多的手性碳原子,但是能够产生空间立体结构的主要是在C20位上。几乎每种人参皂苷和皂苷元都存在C20位的差向异构体。在人参中,C20位S构型的人参皂苷和皂苷元含量要远远高于C20位R构型的人参皂苷和皂苷,所以一般情况下所说的人参皂苷和皂苷元都指的C20位S构型的人参皂苷和皂苷元。但是,人参皂苷和皂苷元C20位的差向异构体具有明显不同的生理活性。例如,S构型人参皂苷Rh2(3-O-β-(D-glucopyranosyl)-20(S)-protopanaxadiol)可以明显抑制前列腺癌细胞,但是R构型人参皂苷Rh2(3-O-β-(D-glucopyranosyl)-20(R)-protopanaxadiol)的抑制效果却很差。R构型人参皂苷Rh2可以选择性地抑制破骨细胞的生成且没有任何细胞毒性作用,但是S构型人参皂苷Rh2抑制破骨细胞生成的作用很弱,但是对破骨细胞却有很强的细胞毒性作用。并且S构型和R构型人参皂苷Rh2对于P-糖蛋白的调节作用也存在很大的差异。

[0007] 糖基转移酶的功能是将糖基供体(核苷二磷酸糖,例如UDP-葡萄糖)上的糖基转移到不同的糖基受体上。根据氨基酸序列的不同,目前糖基转移酶已有94个家族。目前已测序

的植物基因组中,发现了上百种以上不同的糖基转移酶。这些糖基转移酶的糖基受体包括糖、脂、蛋白、核酸、抗生素和其它的小分子。在人参中参与皂苷糖基化的糖基转移酶,其作用是把糖基供体上的糖基转移到皂苷元或者苷元的C-3、C-6或C-20的羟基上,从而形成具有不同药用价值的皂苷。

[0008] 目前,通过对人参,花旗参和三七的转录组分析,研究人员已经发现了大量的糖基转移酶基因,但是还不明确哪些糖基转移酶参与了人参皂苷的合成。由于人参中含有数量众多的糖基转移酶,且它们的含量都很低,所以它们的分离纯化的研究进展缓慢。

[0009] 稀有人参皂苷是指在人参中含量极低的皂苷。人参皂苷CK (20-O-β-(D-glucopyranosyl)-20(S)-protopanaxadiol) 属于原人参二醇类的皂苷,在皂苷元的C-20位羟基上连有一个葡萄糖基。人参皂苷CK在人参中含量极低,它是原人参二醇型皂苷在人体肠道中被微生物水解后产生的主要代谢产物。研究表明,多数原人参二醇型皂苷只有被代谢为CK后才能被人体吸收,所以人参皂苷CK是直接被体内吸收和发挥作用的真正实体,而其他皂苷只是药物前体。人参皂苷CK具有很好的抗肿瘤活性,可以诱导肿瘤细胞凋亡,抑制肿瘤细胞转移。它与放疗和化疗结合试验,可以增强放疗和化疗的效果。除此之外,人参皂苷CK还有具有抗过敏活性,抗炎活性,并且可以起到神经保护作用,抗糖尿病作用和抗皮肤衰老作用。人参皂苷CK的药理活性具有多靶点、高活性和低毒性的特点。

[0010] 人参皂苷F1 (20-O-β-D-glucopyranosyl-20(S)-protopanaxatriol) 属于原人参三醇型皂苷,它在人参中的含量也非常低,也属于稀有人参皂苷。人参皂苷F1的结构与CK非常接近,也是在皂苷元的C-20位羟基上连有一个葡萄糖基。人参皂苷F1也具有独特的药用价值。它具有抗衰老和抗氧化的功能。

[0011] 人参皂苷Rh1 (6-O-β-D-glucopyranosyl-20(S)-protopanaxatriol) 属于原人参三醇型皂苷,它在人参中的含量也非常低,也属于稀有人参皂苷。人参皂苷Rh1的结构与F1非常接近,但是它的糖基化位点是在C6位上的羟基。人参皂苷Rh1也具有特殊的生理功能,能够抗过敏和抗炎症。

[0012] 人参皂苷Rh2 (3-O-β-(D-glucopyranosyl)-20(S)-protopanaxadiol) 在人参中含量极低,它的含量大约只有人参干重的万分之一左右,也属于稀有人参皂苷。但是,人参皂苷Rh2具有良好的抗肿瘤活性,是人参中最主要的抗肿瘤活性成分之一,能够抑制肿瘤细胞生长、诱导肿瘤细胞凋亡、抗肿瘤转移。研究表明人参皂苷Rh2能够抑制lung cancer cells3LL(mice),Morris liver cancer cells(rats),B-16melanoma cells(mice),以及HeLa cells(human)的增值。在临床上,人参皂苷Rh2与放疗或化疗结合治疗,可以增强放疗和化疗的效果。此外,人参皂苷Rh2还具有抗过敏,提高机体免疫力的功能,抑制NO和PGE产生的炎症等作用。

[0013] 人参皂苷Rg3在人参中含量也很低,具有明显的抗肿瘤的作用,与人参皂苷Rh2在抗肿瘤的功能方面有互补性,临床使用证明,Rg3与Rh2联合使用其治疗肿瘤的增效作用进一步提高。

[0014] 由于稀有人参皂苷CK、F1、Rh1,Rh2和Rg3的含量非常低,所以目前生产的方法是从人参中的大量皂苷出发,通过选择性水解糖基的方法进行转化后再进行提取和纯化。以人参属植物的总皂苷或原人参二醇类皂苷为原料,转化、分离和提取20(S)-原人参皂苷-Rh2。该制备方法的优点是利用了大量的二醇类皂苷,但是需在高温和高压下反应(宋长春等.20

(S)-人参皂苷-Rh2的制备方法及其药物组合物及应用[P].中国专利:1225366,1999年)。韩国人参烟草研究所公开了2种从人参成分中制备20(R&S)-人参皂苷-Rh2的方法。其特征在于首先得到原人参二醇皂苷组分,再经酸水解处理得20(R&S)-人参皂苷-Rg3,然后将人参皂苷Rg3处理得到人参皂苷Rh2。以上这些方法主要缺陷在于产物的起始材料需要原人参二醇类系列单体皂苷,使得反应步骤较繁琐,原料损失较大,操作繁琐,从而导致成本增加,而且难以提高产率。由于CK和F1在C-20位的糖基容易在水解过程中被破坏,所以化学法不适用于CK和F1的生产。而酸法和碱法水解皂苷制备Rh1的产率很低,而且还存在很多副产物。

[0015] 由于酶转化法的条件温和,专一性强、产物易分离纯化的特点,是目前生产CK、F1和Rh1的主要方法。用于制备人参皂苷CK、F1、Rh1和Rh2的酶主要有柚苷酶、果胶酶、纤维素酶及乳糖酶等。人参皂苷CK也可以通过微生物转化的方法获得,主要是利用肠道来源的厌氧菌。虽然,生物转化法(酶法和微生物法)制备稀有人参皂苷CK、F1、Rh1和Rh2取得了很大的进步,但是由于其原料为人参皂苷,所以制备CK、F1、Rh1和Rh2的成本仍然非常高,并且其产量也相当有限(中国专利:CN1105781C;金东史等,大连轻工业学院学报,2001年)。

[0016] 由于人参皂苷Rh2的重要生物活性和巨大的经济价值,几十年来人们也一直尝试利用化学合成的方法来生产这种皂苷,其基本思路是由原人参二醇和相应糖基缩和而成,即半合成法(日本专利:特开平8-208688,1996年)。该方法以原人参二醇为原料半合成20(S)-原人参皂苷-Rh2,其合成步骤分为六步,且在糖苷化反应中使用了当量的碳酸银作为催化剂,价格贵重使得该方法的成本较高,同时该催化剂的立体选择性不高,产物得率低。另一种方法利用了芳香烃类的酰基或烷基取代原人参二醇C-12位羟基,再在有机溶剂和惰性气体的保护下,加入活化了C-1为羟基的葡萄糖基供体,在分子筛存在和路易斯酸的催化下生发生缩合反应,生成的产物经过柱层析或重结晶纯化后出去保护基团,得到20(S)-人参皂苷-Rh2(惠永正等,20(S)-人参皂苷-Rh2的制备方法,中国专利:CN1587273A,2005年)

[0017] 目前本领域尚缺乏一种有效的生产稀有人参皂苷CK、F1、Rh1、Rh2和Rg3的方法,因此迫切需要开发多种特异高效的糖基转移酶。

## 发明内容

[0018] 本发明的目的就是提供一组糖基转移酶及其应用。

[0019] 在本发明的第一方面,提供了一种体外糖基化方法,包括步骤:

[0020] 在糖基转移酶存在下,将糖基供体的糖基转移到四环三萜类化合物的以下位点上:

[0021] C-20位、C-6位、C-3位上或C-3位的第一个糖基上;

[0022] 从而形成糖基化的四环三萜类化合物;

[0023] 其中,所述的糖基转移酶选自:

[0024] 如SEQ ID NOs.:2、16、18、20、22、24、26、28、41、43、55、57、59或61所示的糖基转移酶。

[0025] 本发明第二方面,提供了一种分离的多肽,所述的多肽选自下组:

[0026] (a) 具有SEQ ID NOs.:2、16、18、20、26、28、41、43、55、57、59或61中任一条所示氨基酸序列的多肽;

[0027] (b) 将SEQ ID NOs.:2、16、18、20、26、28、41、43、55、57、59或61中任一条所示氨基

酸序列的多肽经过一个或多个氨基酸残基的取代、缺失或添加而形成的、或是添加信号肽序列后形成的、并具有糖基转移酶活性的衍生多肽；

[0028] (c) 序列中含有 (a) 或 (b) 中所述多肽序列的衍生多肽；

[0029] (d) 氨基酸序列与SEQ ID N0s.:2、16、18、20、26、28、41、43、55、57、59或61中任一条所示氨基酸序列的同源性 $\geq 85\%$ 或 $\geq 90\%$ (较佳地 $\geq 95\%$ )，并具有糖基转移酶活性的衍生多肽。

[0030] 在另一优选例中，所述的序列(c)为由(a)或(b)添加了标签序列、信号序列或分泌信号序列后所形成的融合蛋白。

[0031] 在另一优选例中，所述的多肽为SEQ ID N0s.:2、16、18、20、26、28、41、43、55、57、59或61所示氨基酸序列的多肽。

[0032] 本发明的第三方面，提供了一种分离的多肽，所述的多肽选自下组：

[0033] (a1) 具有SEQ ID N0s.:22、24中任一条所示氨基酸序列的多肽；

[0034] (b1) 序列中含有(a1)中所述多肽序列的多肽；和/或所述的多肽选自下组：

[0035] (a2) 具有SEQ ID N0s.:4或6中任一条所示氨基酸序列的多肽；

[0036] (b2) 将SEQ ID N0s.:4或6中任一条所示氨基酸序列的多肽经过一个或多个氨基酸残基的取代、缺失或添加而形成的、或是添加信号肽序列后形成的、并具有糖基转移酶活性的衍生多肽；

[0037] (c2) 序列中含有(b2)中所述多肽序列的衍生多肽；

[0038] (d2) 氨基酸序列与SEQ ID N0s.:4或6中任一条所示氨基酸序列的同源性 $\geq 85\%$ (较佳地 $\geq 95\%$ )，并具有糖基转移酶活性的衍生多肽。

[0039] 在另一优选例中，所述的序列(c2)为由(a2)或(b2)添加了标签序列、信号序列或分泌信号序列后所形成的融合蛋白。

[0040] 在本发明的第四方面，提供了一种分离的多核苷酸，所述的多核苷酸为选自下组的序列：

[0041] (A) 编码第一或第二方面所述多肽的核苷酸序列；

[0042] (B) 编码如SEQ ID N0s.:2、4、6、16、18、20、22、24、26、28、41、43、55、57、59或61所示多肽的核苷酸序列；

[0043] (C) 如SEQ ID N0s.:1、3、5、15、17、19、21、23、25、27、40、42、54、56、58或60所示的核苷酸序列；

[0044] (D) 与SEQ ID N0s.:1、3、5、15、17、19、21、27、40、42、54、56、58或60所示序列的同源性 $\geq 95\%$ (较佳地 $\geq 98\%$ )的核苷酸序列；

[0045] (E) 在SEQ ID N0s.:1、3、5、15、17、19、21、23、25、27、40、42、54、56、58或60所示核苷酸序列的5'端和/或3'端截短或添加1-60个(较佳地1-30，更佳地1-10个)核苷酸所形成的核苷酸序列；

[0046] (F) 与(A)-(E)任一所述的核苷酸序列互补(较佳地完全互补)的核苷酸序列。

[0047] 在另一优选例中，所述的核苷酸的序列如SEQ ID N0s.:1、3、5、15、17、19、21、23、25、27、40、42、54、56、58或60所示。

[0048] 在另一优选例中，序列如SEQ ID N0s.:1、3、5、15、17、19、21、23、25、27、40、42、54、56、58或60所示的多核苷酸编码氨基酸序列分别如SEQ ID N0s.:2、4、6、16、18、20、22、24、

26、28、41、43、55、57、59或61所示的多肽。

[0049] 在本发明的第五方面,提供了一种载体,所述的载体含有第三方面所述的多核苷酸。较佳地,所述载体包括表达载体、穿梭载体、整合载体。

[0050] 在本发明的第五方面,提供了本发明第一或第二方面所述分离的多肽的用途,它被用于催化以下一种或多种反应,或被用于制备催化以下一种或多种反应的催化制剂:将来自糖基供体的糖基转移到四环三萜类化合物C-20位和/或C-6位和/或C-3位的羟基以替换所述羟基的H,以及将来自糖基供体的糖基转移到四环三萜类化合物的C-3位的第一个糖基上以延伸糖链。

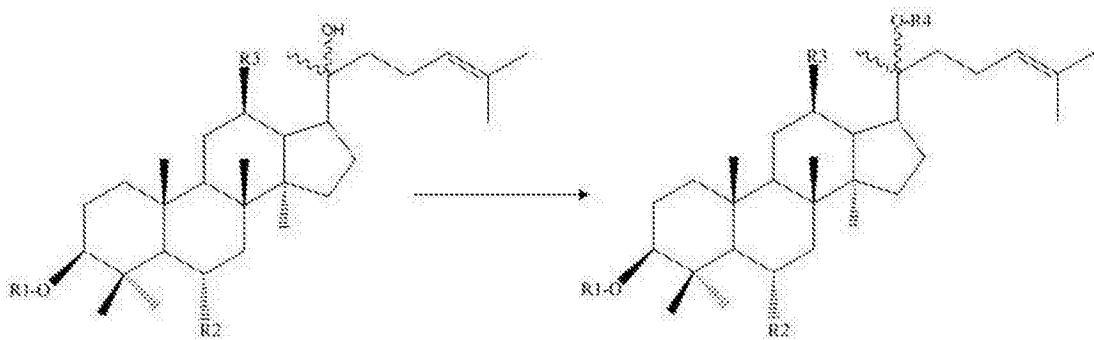
[0051] 在另一优选例中,所述的糖基供体包括选自下组的核苷二磷酸糖:UDP-葡萄糖,ADP-葡萄糖,TDP-葡萄糖,CDP-葡萄糖,GDP-葡萄糖,UDP-乙酰基葡萄糖,ADP-乙酰基葡萄糖,TDP-乙酰基葡萄糖,CDP-乙酰基葡萄糖,GDP-乙酰基葡萄糖,UDP-木糖,ADP-木糖,TDP-木糖,CDP-木糖,GDP-木糖,UDP-半乳糖醛酸,ADP-半乳糖醛酸,TDP-半乳糖醛酸,CDP-半乳糖醛酸,GDP-半乳糖醛酸,UDP-半乳糖,ADP-半乳糖,TDP-半乳糖,CDP-半乳糖,GDP-半乳糖,UDP-阿拉伯糖,ADP-阿拉伯糖,TDP-阿拉伯糖,CDP-阿拉伯糖,GDP-阿拉伯糖,UDP-鼠李糖,ADP-鼠李糖,TDP-鼠李糖,CDP-鼠李糖,GDP-鼠李糖,或其他核苷二磷酸己糖或核苷二磷酸戊糖,或其组合。

[0052] 在另一优选例中,所述的糖基供体包括选自下组的尿苷二磷酸(UDP)糖:UDP-葡萄糖,UDP-半乳糖醛酸,UDP-半乳糖,UDP-阿拉伯糖,UDP-鼠李糖,或其他尿苷二磷酸己糖或尿苷二磷酸戊糖,或其组合。

[0053] 在另一优选例中,所述分离的多肽用于催化下述一种或多种反应或被用于制备催化下述一种或多种反应的催化制剂:

[0054]

(A)



式(I)化合物

式(II)化合物

[0055] 其中,R1为H,单糖糖基或多糖糖基;R2和R3为H或OH;R4为糖基;所述的多肽选自SEQ ID NO:2、16或18或其衍生多肽。

[0056] 在另一优选例中,所述的单糖包括葡萄糖(Glc),鼠李糖(Rha),乙酰基葡萄糖(Glc(6)Ac),阿拉伯呋喃糖(Araf),阿拉伯吡喃糖(Arap),木糖(Xyl)等。

[0057] 在另一优选例中,所述的多糖包括Glc(2-1)Glc,Glc(6-1)Glc,Glc(6)Ac,Glc(2-1)Rha,Glc(6-1)Arap,Glc(6-1)Xyl,Glc(6-1)Araf,Glc(3-1)Glc(3-1),Glc(2-1)Glu(6)Ac,Glc(6-1)Arap(4-1)Xyl,Glc(6-1)Arap(2-1)Xyl,或Glc(6-1)Arap(3-1)Xyl等2-4个单

糖组成的多糖。

[0058] R1-R4经取代后的化合物如下表所示：

[0059]

底物	R1	R2	R3	R4	产物
PPD	H	H	OH	糖基	CK
Rh2	1个糖基	H	OH	糖基	F2
Rg3	2个糖基	H	OH	糖基	Rd
PPT	H	OH	OH	糖基	F1
DM	H	H	H	糖基	20-G-DM

[0060] 即当所述的R1,R2都为H,R3为OH时,所述的式(I)化合物为原人参二醇(PPD)

[0061] R1为一个葡萄糖基,R2为H,R3为OH时,所述的式(I)化合物为人参皂苷RH2

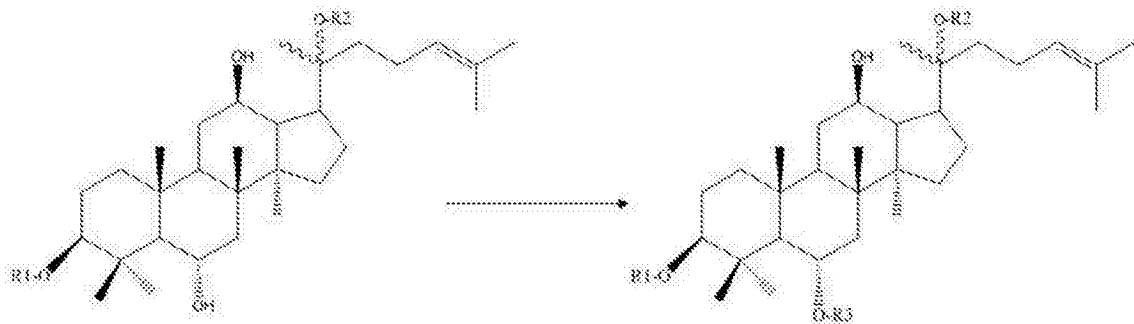
[0062] R1为两个葡萄糖基,R2为H,R3为OH时,所述的式(I)化合物为人参皂苷RG3

[0063] R1为H,R2为OH,R3为OH,所述的式(I)化合物为原人参三醇(PPT)

[0064] R1为H,R2为H,R3为H,所述的式(I)化合物达玛烯二醇(DM)

[0065]

(B)



式(III)化合物

式(IV)化合物

[0066] 其中,R1为H或者糖基,R2糖基,R3为糖基,所述的多肽选自SEQ ID NOs.:2、16、18或20或其衍生多肽；

[0067] 或,R1为H或者糖基;R2为H;R3为糖基,所述的多肽选自SEQ ID NO.:20或其衍生多肽。

[0068] R1-R3经取代后的化合物如下表所示：

[0069]

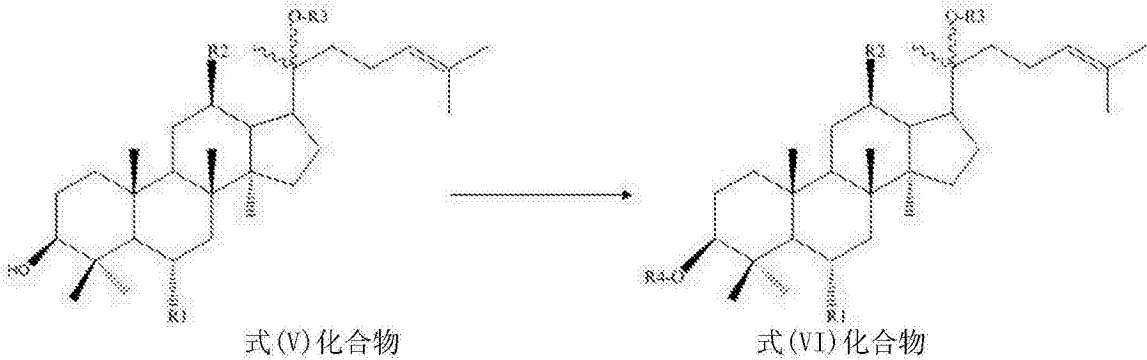
底物	R1	R2	R3	产物
F1	H	糖基	糖基	Rg1
PPT	H	H	糖基	Rh1

[0070] 即当R1,R2都为H时,所述的式(III)化合物为原人参三醇(PPT)。

[0071] R1为H,R2为葡萄糖基时,所述的式(III)化合物为人参皂苷F1。

[0072]

(C)



[0073] 其中,R1为H或者OH;R2为H或者OH;R3为H或者糖基;R4为糖基,所述的多肽选自SEQ ID NOs.:22、24、41或43或其衍生多肽。

[0074] R1-R4经取代后的化合物如下表所示:

[0075]

底物	R1	R2	R3	R4	产物
PPD	H	OH	H	糖基	Rh2
CK	H	OH	糖基	糖基	F2
PPT	OH	OH	H	糖基	3-G-PPT
F1	OH	OH	糖基	糖基	3-G-F1
DM	H	H	H	糖基	3-G-DM

[0076] 即当R1和R3都为H,R2为OH时,所述的式(V)化合物为原人参二醇(PPD);

[0077] R1为H,R2为OH,R3为葡萄糖基时,所述的式(V)化合物为人参皂苷CK;

[0078] R1为OH,R2为OH,R3为H时,所述的式(V)化合物为原人参三醇(PPT);

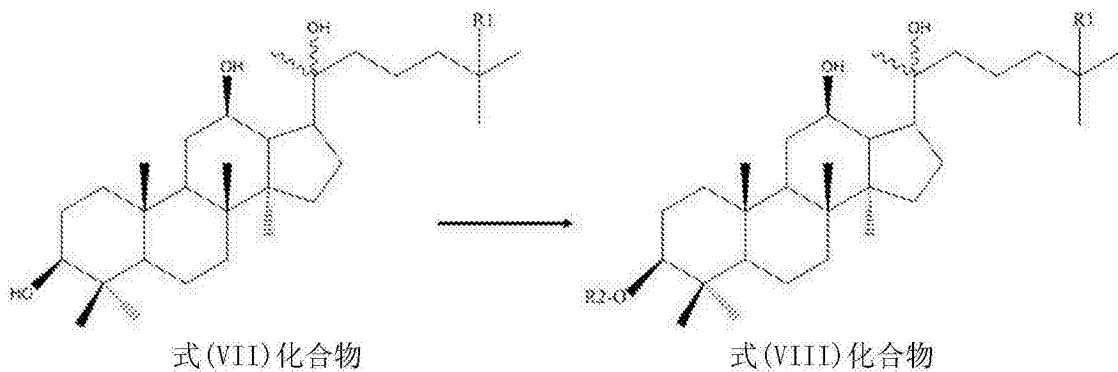
[0079] R1为OH,R2为OH,R3为葡萄糖基时,所述的式(V)化合物为人参皂苷F1;

[0080] R1为H,R2为OH,R3为H时,所述的式(V)化合物为达玛烯二醇(DM)。

[0081] 当底物为PPD时,所述的多肽选自SEQ ID NOs.:22、24、41或43或其衍生多肽;当底物为CK时,所述的多肽选自SEQ ID NOs.:22、24或43或其衍生多肽;当底物为PPT时,所述的多肽选自SEQ ID NOs.:22、24或41或其衍生多肽;当底物为F1和DM时,所述的多肽选自SEQ ID NOs.:22或24或其衍生多肽。

[0082]

(D)



[0083] 其中,R1为OH或者OCH<sub>3</sub>;R2为糖基,所述的多肽选自SEQ ID NOs.:22、24、41或43或其衍生多肽。

[0084] R1-R2经取代后的化合物如下表所示:

[0085]

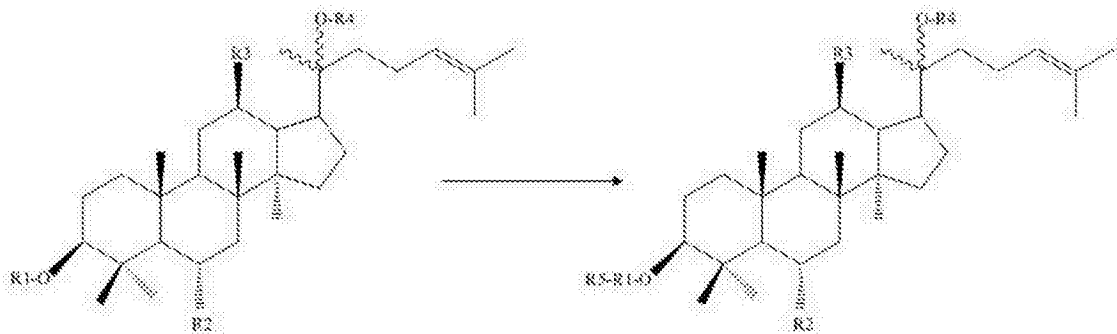
底物	R1	R2	产物
25-OH-PPD	OH	糖基	3-G-25-OH-PPD
25-OCH <sub>3</sub> -PPD	OCH <sub>3</sub>	糖基	3-G-25-OCH <sub>3</sub> -PPD

[0086] 即当R1为OH时,所述的式(VII)化合物为25-OH-PPD;

[0087] R1为OCH<sub>3</sub>时,所述的式(VII)化合物为25-OCH<sub>3</sub>-PPD。

[0088]

(E)



式(IX)化合物

式(X)化合物

[0089] 其中,R1为糖基;R2和R3为OH或者H;R4为糖基或者H;R5为糖基,R5-R1-O为C3第一个糖基衍生的糖基,所述的多肽选自SEQ ID NOs.:26、28、55、57、59或61或其衍生多肽。

[0090] R1-R4经取代后的化合物如下表所示:

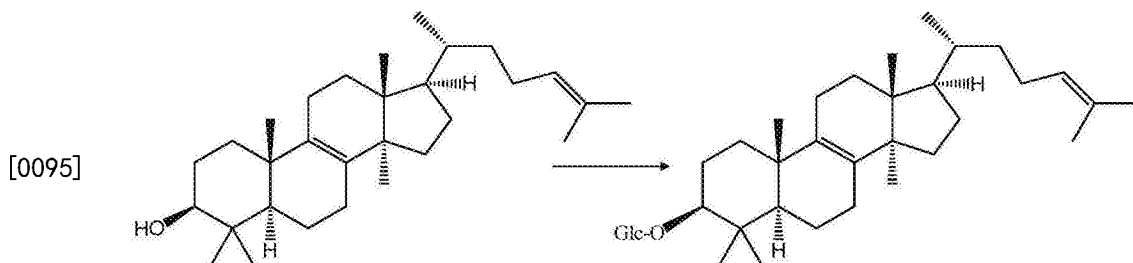
[0091]

底物	R1	R2	R3	R4	产物
Rh2	糖基	H	OH	H	Rg3
F2	糖基	H	OH	糖基	Rd

[0092] 即当R1为葡萄糖基;R2为H,R3为OH,R4为H,式(IX)化合物为Rh2。

[0093] R1为葡萄糖基;R2为H,R3为OH,R4为葡萄糖基,式(IX)化合物为F2。

[0094] (F)



[0095]

式(XI)化合物

式(XII)化合物

[0096] 所述的多肽选自SEQ ID NO:22或SEQ ID NO:24或其衍生多肽。式(XI)化合物为羊毛甾醇(lanosterol),并且式(XII)化合物为3-O-β-(D-glucopyranosyl)-lanosterol。

[0097] 在另一优选例中,所述的糖基选自:葡萄糖基、半乳糖醛酸基、半乳糖基、阿拉伯糖基、鼠李糖基,以及其他己糖基或戊糖基。

[0098] 在另一优选例中,所述反应式(I)、(III)、(V)、(VII), (IX) 或者(XI)化合物包括(但不限于):S构型或R构型的达玛烷型四环三萜类化合物、羊毛脂烷型四环三萜类化合物、甘遂烷型四环三萜类化合物、环阿屯烷(环阿尔廷烷)型四环三萜类化合物、葫芦烷四环三萜类化合物、或楝烷型四环三萜类化合物。

[0099] 在另一优选例中,所述的多肽选自下组:

[0100] (a) 具有SEQ ID NOs.:2、16、18、20、26、28、41、43、55、57、59或61中任一条所示氨基酸序列的多肽;

[0101] (b) 将SEQ ID NOs.:2、16、18、20、26、28、41、43、55、57、59或61中任一条所示氨基酸序列的多肽经过一个或多个氨基酸残基的取代、缺失或添加而形成的、或是添加信号肽序列后形成的、并具有糖基转移酶活性的衍生多肽;

[0102] (c) 序列中含有(a)或(b)中所述多肽序列的衍生多肽;

[0103] (d) 氨基酸序列与SEQ ID NO:2、16、18、20、26、28、41、43、55、57、59或61中任一条所示氨基酸序列的同源性 $\geq 85\%$ 或 $\geq 90\%$ (较佳地 $\geq 95\%$ ),并具有糖基转移酶活性的衍生多肽。

[0104] 在另一优选例中,所述的多肽选自下组:

[0105] (a1) 具有SEQ ID NOs.:22、24中任一条所示氨基酸序列的多肽;

[0106] (b1) 序列中含有(a1)中所述多肽序列的多肽;和/或

[0107] 所述的多肽选自下组:

[0108] (a2) 具有SEQ ID NOs.:4或6中任一条所示氨基酸序列的多肽;

[0109] (b2) 将SEQ ID NOs.:4或6中任一条所示氨基酸序列的多肽经过一个或多个氨基酸残基的取代、缺失或添加而形成的、或是添加信号肽序列后形成的、并具有糖基转移酶活性的衍生多肽;

[0110] (c2) 序列中含有(b2)中所述多肽序列的衍生多肽;

[0111] (d2) 氨基酸序列与SEQ ID NOs.:4或6中任一条所示氨基酸序列的同源性 $\geq 85\%$ (较佳地 $\geq 95\%$ ),并具有糖基转移酶活性的衍生多肽。

[0112] 在另一优选例中,编码所述多肽的核苷酸所述的多核苷酸为选自下组的序列:

[0113] (A) 编码本发明第一或第二方面所述多肽的核苷酸序列;

[0114] (B) 编码如SEQ ID NOs.:2、4、6、16、18、20、22、24、26、28、41、43、55、57、59或61所示多肽的核苷酸序列;

[0115] (C) 如SEQ ID NOs.:1、3、5、15、17、19、21、23、25、27、40、42、54、56、58或60所示的核苷酸序列;

[0116] (D) 与SEQ ID NOs.:1、3、5、15、17、19、21、23、25、27、40、42、54、56、58或60所示序列的同源性 $\geq 95\%$ (较佳地 $\geq 98\%$ )的核苷酸序列;

[0117] (E) 在SEQ ID NOs.:1、3、5、15、17、19、21、23、25、27、40、42、54、56、58或60所示核苷酸序列的5'端和/或3'端截短或添加1-60个(较佳地1-30,更佳地1-10个)核苷酸所形成的核苷酸序列;

[0118] (F) 与(A)-(E)任一所述的核苷酸序列互补(较佳地完全互补)的核苷酸序列。

[0119] 在另一优选例中,所述的核苷酸的序列如SEQ ID NOs.:1、3、5、15、17、19、21、23、25、27、40、42、54、56、58或60所示。

[0120] 在另一优选例中,序列如SEQ ID NOs.:1、3、5、15、17、19、21、23、25、27、40、42、54、56、58或60所示的多核苷酸编码氨基酸序列分别如SEQ ID NOs.:2、4、6、16、18、20、22、24、26、28、41、43、55、57、59或61所示的多肽。

[0121] 在本发明的第六方面,提供了一种进行糖基转移催化反应的方法,包括步骤:在本发明第二或第三方面所述的多肽或其衍生多肽存在的条件下,进行糖基转移催化反应。

[0122] 在另一优选例中,所述的方法还包括步骤:

[0123] 在糖基供体以及如本发明第二方面或第三方面所述多肽及其衍生多肽的存在下,将所述的式(I)化合物转化为所述的式(II)化合物,或式(III)化合物转化为所述的式(IV)化合物,或式(V)化合物转化为所述的式(VI)化合物,或式(VII)化合物转化为所述的式(VIII)化合物,或式(IX)化合物转化为所述的式(X)化合物,或式(XI)化合物转化为所述的式(XII)化合物;

[0124] 在另一优选例中,所述的方法还包括将所述的多肽及其衍生多肽分别加入催化反应;和/或

[0125] 将所述的多肽及其衍生多肽同时加入催化反应。

[0126] 在另一优选例中,所述的方法还包括在糖基供体以及如本发明第二方面和第三方面所述多肽及其衍生多肽中两种以上多肽共同存在的条件下,将式(I)化合物转化为式(IV),(VI),(VIII),(X)的化合物,或者式(III)化合物转化为式(II),(VI),(VIII),(X)的化合物;或者式(V)化合物转化式(II),(IV),(VIII),(X)的化合物,或者式(VII)化合物转化式(II),(IV),(VI),(X)的化合物,或者(IX)化合物转化式(II),(IV),(VI),(VIII)的化合物。

[0127] 在另一优选例中,所述的方法还包括将编码糖基转移酶的核苷酸序列与达玛稀二醇和/或原人参二醇和/或原人参三醇合成代谢途径中的关键基因在宿主细胞中共表达,从而获得所述的式(II),(IV),(VI),(VIII),(X)或(XII)化合物。

[0128] 在另一优选例中,所述的宿主细胞为酵母菌或大肠杆菌。

[0129] 在另一优选例中,所述的多肽为具有SEQ ID NOs.:2、4、6、16、18、20、22、24、26、28、41、43、55、57、59或61所示氨基酸序列的多肽及其衍生多肽。

[0130] 在另一优选例中,编码所述多肽的核苷酸序列如SEQ ID NOs.:1、3、5、15、17、19、21、23、25、27、40、42、54、56、58或60所示。

[0131] 在另一优选例中,所述方法还包括:向反应体系中提供用于调节酶活性的添加物。

[0132] 在另一优选例中,所述的用于调节酶活性的添加物是:提高酶活性或抑制酶活性的添加物。

[0133] 在另一优选例中,所述的用于调节酶活性的添加物选自下组:Ca<sup>2+</sup>、Co<sup>2+</sup>、Mn<sup>2+</sup>、Ba<sup>2+</sup>、Al<sup>3+</sup>、Ni<sup>2+</sup>、Zn<sup>2+</sup>、或Fe<sup>2+</sup>。

[0134] 在另一优选例中,所述的用于调节酶活性的添加物是:可以生成Ca<sup>2+</sup>、Co<sup>2+</sup>、Mn<sup>2+</sup>、Ba<sup>2+</sup>、Al<sup>3+</sup>、Ni<sup>2+</sup>、Zn<sup>2+</sup>、或Fe<sup>2+</sup>的物质。

[0135] 在另一优选例中,所述的糖基供体是核苷二磷酸糖,选自下组:UDP-葡萄糖,ADP-葡萄糖,TDP-葡萄糖,CDP-葡萄糖,GDP-葡萄糖,UDP-乙酰基葡萄糖,ADP-乙酰基葡萄糖,

TDP-乙酰基葡萄糖,CDP-乙酰基葡萄糖,GDP-乙酰基葡萄糖,UDP-木糖,ADP-木糖,TDP-木糖,CDP-木糖,GDP-木糖,UDP-半乳糖醛酸,ADP-半乳糖醛酸,TDP-半乳糖醛酸,CDP-半乳糖醛酸,GDP-半乳糖醛酸,UDP-半乳糖,ADP-半乳糖,TDP-半乳糖,CDP-半乳糖,GDP-半乳糖,UDP-阿拉伯糖,ADP-阿拉伯糖,TDP-阿拉伯糖,CDP-阿拉伯糖,GDP-阿拉伯糖,UDP-鼠李糖,ADP-鼠李糖,TDP-鼠李糖,CDP-鼠李糖,GDP-鼠李糖,或其他核苷二磷酸己糖或核苷二磷酸戊糖,或其组合。

[0136] 在另一优选例中,所述的糖基供体是尿苷二磷酸糖,选自下组:UDP-葡萄糖,UDP-半乳糖醛酸,UDP-半乳糖,UDP-阿拉伯糖,UDP-鼠李糖,或其他尿苷二磷酸己糖或尿苷二磷酸戊糖,或其组合。

[0137] 在另一优选例中,反应体系的pH为:pH4.0-10.0,优选pH为5.5-9.0。

[0138] 在另一优选例中,反应体系的温度为:10℃-105℃,优选20℃-50℃。

[0139] 在另一优选例中,所述的达玛烯二醇合成代谢途径中的关键基因包括(但不限于):达玛烯二醇合成酶基因。

[0140] 在另一优选例中,所述的原人参二醇合成代谢途径中的关键基因包括(但不限于):达玛烯二醇合成酶基因、细胞色素P450CYP716A47基因和P450CYP716A47的还原酶基因,或其组合。

[0141] 在另一优选例中,所述的原人参三醇合成代谢途径中的关键基因包括(但不限于):达玛烯二醇合成酶基因、细胞色素P450CYP716A47基因、P450CYP716A47的还原酶基因、及细胞色素P450CYP716A53V2基因及其还原酶基因,或其组合。

[0142] 在另一优选例中,所述糖基催化反应的底物为式(I)、(III)、(V)、(VII)、(IX)或(XI)化合物,且所述的底物为(II)、(IV)、(VI)、(VIII)、(X)或(XII)化合物;

[0143] 在另一优选例中,所述的式(I)化合物为原人参二醇PPD(Protopanaxadiol),并且式(II)化合物为人参皂苷CK(20-O-β-(D-吡喃葡萄糖基)-原人参二醇(20-O-β-(D-glucopyranosyl)-protopanaxadiol));

[0144] 或,所述的式(I)化合物为人参皂苷Rh2(3-O-β-(D-吡喃葡萄糖基)-原人参二醇(3-O-β-(D-glucopyranosyl)-protopanaxadiol)),并且式(II)化合物为人参皂苷F2(3-O-β-(D-吡喃葡萄糖基)-20-O-β-(D-吡喃葡萄糖基)-原人参二醇(3-O-β-(D-glucopyranosyl)-20-O-β-(D-glucopyranosyl)-protopanaxadiol));

[0145] 或,所述的式(I)化合物为人参皂苷Rg3,并且式(II)化合物为人参皂苷Rd;

[0146] 或,所述的式(I)化合物为原人参三醇PPT(Protopanaxatriol),并且式(II)化合物为人参皂苷F1(20-O-β-(D-吡喃葡萄糖基)-原人参三醇(20-O-β-(D-glucopyranosyl)-protopanaxatriol));

[0147] 或,所述的式(I)化合物为达玛烯二醇DM(Dammarenediol II),并且式(II)化合物为人参皂苷20-O-β-(D-吡喃葡萄糖基)-达玛烯二醇(20-O-β-(D-glucopyranosyl)-Dammarenediol II);

[0148] 或,所述的式(III)化合物为原人参三醇PPT,并且式(IV)化合物为人参皂苷Rh1(6-O-β-(D-吡喃葡萄糖基)-原人参三醇(6-O-β-(D-glucopyranosyl)-protopanaxatriol));

[0149] 或,所述的式(III)化合物为人参皂苷F1,并且式(IV)化合物为人参皂苷Rg1(6-O-

$\beta$ -(D-吡喃葡萄糖基)-20-O- $\beta$ -(D-吡喃葡萄糖基)-原人参二醇(6-O- $\beta$ -(D-glucopyranosyl)-20-O- $\beta$ -(D-glucopyranosyl)-protopanaxadiol));

[0150] 或,所述的式(V)化合物为原人参二醇,并且式(VI)化合物为人参皂苷Rh2(3-O- $\beta$ -(D-吡喃葡萄糖基)-原人参二醇(3-O- $\beta$ -(D-glucopyranosyl)-protopanaxadiol));

[0151] 或,所述的式(V)化合物为CK,并且式(VI)化合物为人参皂苷F2(3-O- $\beta$ -(D-吡喃葡萄糖基)-20-O- $\beta$ -(D-吡喃葡萄糖基)-原人参二醇(3-O- $\beta$ -(D-glucopyranosyl)-20-O- $\beta$ -(D-glucopyranosyl)-protopanaxadiol));

[0152] 或,所述的式(V)化合物为原人参三醇PPT,并且式(VI)化合物为人参皂苷3-O- $\beta$ -(D-吡喃葡萄糖基)-原人参三醇(3-O- $\beta$ -(D-glucopyranosyl)-protopanaxatriol);

[0153] 或,所述的式(V)化合物为人参皂苷F1,并且式(VI)化合物为人参皂苷3-O- $\beta$ -(D-吡喃葡萄糖基)-F1(3-O- $\beta$ -(D-glucopyranosyl)-F1);

[0154] 或,所述的式(V)化合物为达玛烯二醇DM,并且式(VI)化合物为人参皂苷3-O- $\beta$ -(D-吡喃葡萄糖基)-达玛烯二醇(3-O- $\beta$ -(D-glucopyranosyl)-Dammarendiol II);

[0155] 或,所述的式(VII)化合物为25-OH-原人参二醇(25-OH-protopanaxadiol),并且式(VIII)化合物为人参皂苷3-O- $\beta$ -(D-吡喃葡萄糖基)-25-OH-原人参二醇(3-O- $\beta$ -(D-glucopyranosyl)-25-OH-protopanaxadiol);

[0156] 或,所述的式(VII)化合物为25-OCH<sub>3</sub>-原人参二醇(25-OCH<sub>3</sub>-protopanaxadiol),并且式(VIII)化合物为人参皂苷3-O- $\beta$ -(D-吡喃葡萄糖基)-25-OCH<sub>3</sub>-原人参二醇(3-O- $\beta$ -(D-glucopyranosyl)-25-OCH<sub>3</sub>-protopanaxadiol);

[0157] 或,所述的式(IX)化合物为人参皂苷Rh2,并且式(X)化合物为人参皂苷Rg3;

[0158] 或,所述的式(IX)化合物为人参皂苷F2,并且式(X)化合物为人参皂苷Rd。

[0159] 或,所述的式(XI)化合物为羊毛甾醇(lanosterol),并且式(XII)化合物为3-O- $\beta$ -(D-吡喃葡萄糖基)-羊毛甾醇(3-O- $\beta$ -(D-glucopyranosyl)-lanosterol)。

[0160] 在本发明的第七方面,提供了一种遗传工程化的宿主细胞,所述的宿主细胞含有本发明第五方面所述的载体,或其基因组中整合本发明第四方面所述的多核苷酸。

[0161] 在另一优选例中,所述的糖基转移酶为如本发明第二或第三方面中所述的多肽或其衍生多肽。

[0162] 在另一优选例中,编码所述糖基转移酶的核苷酸序列如本发明第四方面所述。

[0163] 在另一优选例中,所述的细胞为原核细胞或真核细胞。

[0164] 在另一优选例中,所述的宿主细胞为真核细胞,如酵母细胞或植物细胞。

[0165] 在另一优选例中,所述的宿主细胞为酿酒酵母细胞。

[0166] 在另一优选例中,所述的宿主细胞原核细胞,如大肠杆菌。

[0167] 在另一优选例中,所述的宿主细胞为人参细胞。

[0168] 在另一优选例中,所述的宿主细胞不是天然产生式(II),(IV),(VI),(VIII),(X),(XII)化合物的细胞。

[0169] 在另一优选例中,所述的宿主细胞不是天然产生稀有人参皂苷CK和/或稀有人参皂苷F1和/或稀有人参皂苷Rh2和/或Rg3和/或Rh1,和/或新的人参皂苷20-O- $\beta$ -(D-glucopyranosyl)-dammarendiol II、3-O- $\beta$ -(D-glucopyranosyl)-PPT、3-O- $\beta$ -(D-glucopyranosyl)-F1、3-O- $\beta$ -(D-glucopyranosyl)-DM、3-O- $\beta$ -(D-glucopyranosyl)-25-OH-

PPD、3-O-β-(D-glucopyranosyl)-25-OCH<sub>3</sub>-PPD,以及和/或Rh1、F2、Rd和Rg1等的细胞。

[0170] 在另一优选例中,所述的达玛烯二醇合成代谢途径中的关键基因包括(但不限于):达玛烯二醇合成酶基因。

[0171] 在另一优选例中,所述的宿主细胞含有原人参二醇合成代谢途径中的关键基因包括(但不限于):达玛烯二醇合成酶基因、细胞色素P450CYP716A47基因、和P450CYP716A47的还原酶基因,或其组合。

[0172] 在另一优选例中,所述的宿主细胞含有原人参三醇合成代谢途径中的关键基因包括(但不限于):达玛烯二醇合成酶基因、细胞色素P450CYP716A47基因、P450CYP716A47的还原酶基因、及细胞色素P450CYP716A53V2基因,或其组合。

[0173] 在本发明的第八方面,提供了第七方面所述的宿主细胞的用途,用于制备酶催化试剂,或生产糖基转移酶、或作为催化细胞、或产生式(II),(IV),(VI),(VIII),(X)或(XII),化合物。

[0174] 在另一优选例中,所述的宿主细胞用于通过对达玛烯二醇(DM)和/或原人参二醇(PPD)和/或原人参三醇(PPT)的糖基化反应,生产新皂苷20-O-β-(D-glucopyranosyl)-dammarendiol III和/或3-O-β-(D-glucopyranosyl)-dammarendiol III、3-O-β-(D-glucopyranosyl)-protopanaxatriol、3-O-β-(D-glucopyranosyl)-F1和/或稀有人参皂苷CK和/或稀有人参皂苷F1和/或稀有人参皂苷Rh1和/或稀有人参皂苷Rh2和/或稀有人参皂苷Rg3。

[0175] 在本发明的第九方面,提供了一种产生转基因植物的方法,包括步骤:将第七方面所述的遗传工程化的宿主细胞再生为植物,并且所述的遗传工程化的宿主细胞为植物细胞。

[0176] 在另一优选例中,所述的遗传工程化的宿主细胞为人参细胞。

[0177] 应理解,在本发明范围内中,本发明的上述各技术特征和在下文(如实施例)中具体描述的各技术特征之间都可以互相组合,从而构成新的或优选的技术方案。限于篇幅,在此不再一一累述。

## 附图说明

[0178] 下列附图用于说明本发明的具体实施方案,而不用于限定由权利要求书所界定的本发明范围。

[0179] 图1为gGT25基因,gGT25-1基因,gGT25-3和gGT25-5基因PCR产物的琼脂糖凝胶电泳图谱。

[0180] 图2为SDS-PAGE检测gGT25,gGT25-1,gGT25-3和gGT25-5在酿酒酵母中的表达;泳道1,蛋白Marker的电泳结果(分子量由上到下依次为200、116、97.2、66.4、44.3kDa);泳道2,gt25-pYES2酵母重组子的裂解液上清;泳道3,gt25-1-pYES2酵母重组子的裂解液上清;泳道4,gt25-3-pYES2酵母重组子的裂解液上清;泳道5,gt25-5-pYES2酵母重组子的裂解液上清;泳道6,pYES2空载体重组子的裂解液上清。

[0181] 图3显示Western Blot检测gGT25,gGT25-1,gGT25-3和gGT25-5基因在酿酒酵母中的表达;泳道1,gt25-pYES2酵母重组子的裂解液上清;泳道2,gt25-1-pYES2酵母重组子的裂解液上清;泳道4,gt25-3-pYES2酵母重组子的裂解液上清;泳道5,gt25-5-pYES2酵母重

组子的裂解液上清;泳道3,pYES2空载体重组子的裂解液上清。

[0182] 图4为SDS-PAGE检测gGT13和gGT30在酿酒酵母中的表达;泳道1,gt30-pYES2酵母重组子的裂解液上清;泳道2,gt13-pYES2酵母重组子的裂解液上清;泳道3,空载体pYES2重组子的裂解液上清。

[0183] 图5显示Western Blot检测gGT13和gGT30在酿酒酵母中的表达;泳道1,gt30-pYES2酵母重组子的裂解液上清;泳道2,gt13-pYES2酵母重组子的裂解液上清;泳道3,空载体pYES2重组子的裂解液上清。

[0184] 图6显示糖基转移酶gGT25,gGT25-1和gGT25-3催化原人参二醇和原人参二醇型皂苷的产物TLC检测图谱。泳道25,gGT25粗酶液(gt25-pYES2酵母重组子裂解上清);泳道25-1,gGT25-1粗酶液(gt25-1-pYES2酵母重组子裂解上清);泳道25-3,gGT25-3粗酶液(gt25-3-pYES2酵母重组子裂解上清);泳道“-”,负对照,空载体酵母的裂解液上清代替酶液;泳道M,原人参二醇和原人参二醇型皂苷的混合标样。

[0185] 图7显示糖基转移酶gGT25,gGT25-1,gGT25-3和gGT25-5催化原人参三醇与原人参三醇型皂苷的产物TLC检测图谱;泳道M,原人参三醇(PPT)与原人参三醇型皂苷的混合标准样品;泳道25,gGT25粗酶液(gt25-pYES2酵母重组子裂解上清);泳道25-1,gGT25-1粗酶液(gt25-1-pYES2酵母重组子裂解上清);泳道25-3,gGT25-3粗酶液(gt25-3-pYES2酵母重组子裂解上清);泳道25-5,gGT25-5粗酶液(gt25-5-pYES2酵母重组子裂解上清);泳道“-”,负对照,空载体重组酵母裂解代替酶液。

[0186] 图8显示糖基转移酶gGT25,gGT25-1和gGT25-3催化达玛烯二醇的产物TLC检测图谱。泳道25,gGT25粗酶液(gt25-pYES2酵母重组子裂解上清);泳道25-1,gGT25-1粗酶液(gt25-1-pYES2酵母重组子裂解上清);泳道25-3,gGT25-3粗酶液(gt25-3-pYES2酵母重组子裂解上清);泳道“-”,负对照,空载体酵母裂解代替酶液;泳道M,标准样品达玛烯二醇(DM)。

[0187] 图9显示糖基转移酶gGT13和gGT30催化原人参二醇和三醇的产物TLC检测图谱;泳道M1,原人参二醇皂苷混合标准样品;泳道M2,原人参三醇皂苷混合标准样品;泳道1,gGT13粗酶液催化原人参二醇;泳道2,gGT30粗酶液催化原人参二醇;泳道3,负对照,ddH<sub>2</sub>O代替酶液;泳道4,gGT13粗酶液催化原人参三醇;泳道5,gGT30粗酶液催化原人参三醇;泳道6,负对照,ddH<sub>2</sub>O代替酶液。

[0188] 图10显示糖基转移酶gGT25催化原人参二醇的产物HPLC检测,第二行样品:原人参二醇(PPD)与多种皂苷(CK、Rh2、F2、Rg3)的混合标准样品;第一行样品:gGT25粗酶液催化PPD;第三行样品:负对照1,空载体重组酵母裂解液催化PPD;第四行样品:负对照2,dH<sub>2</sub>O。

[0189] 图11显示糖基转移酶gGT25催化原人参三醇的产物HPLC检测,第二行样品:原人参三醇(PPT)与多种三醇皂苷(F1、Rh1、Rg1)的混合标准样品;第一行样品:gGT25粗酶液催化PPT;第三行样品:负对照1,空载体重组酵母裂解液催化PPT;

[0190] 图12显示糖基转移酶gGT25催化原人参二醇产物的LC/MS检测,显示了图10中峰2(产物峰)和标准CK样品的质谱图。

[0191] 图13显示糖基转移酶gGT25催化原人参三醇产物的LC/MS检测,显示了图11中峰1(产物峰)和标准F1样品的质谱图。

[0192] 图14显示Western Blot检测gGT25-pET28a在E.coli BL21中的表达;泳道1-3分别

是50uMIPTG诱导后的总蛋白、上清和沉淀。

[0193] 图15显示gGT25-pET28a重组大肠杆菌细胞裂解液体外催化PPD的产物TLC检查图谱;泳道1,PPD和CK的标准样品混合物;泳道2,gGT25-pET28a重组大肠杆菌诱导后(50μM IPTG)细胞裂解液上清催化PPD。

[0194] 图16显示产CK酵母工程菌A细胞裂解液抽提物的HPLC检测,第一行样品:原人参二醇(PPD)、达玛烯二醇和CK的混合标准样品;第二行样品:产CK酵母工程菌A细胞裂解液;第三行样品:负对照1,酵母出发菌株的裂解液。

[0195] 图17显示糖基转移酶gGT25-5催化原人参三醇的产物HPLC检测,第一行样品:原人参三醇(PPT)与多种三醇皂苷(F1、Rh1、Rg1和Re)的混合标准样品;第二行样品:gGT25-5粗酶液催化PPT后的产物。

[0196] 图18显示糖基转移酶gGT25-5催化原人参三醇的产物LC/MS检测,显示了图17中P1峰(产物Rh1峰)和Rh1标准样品的质谱图。

[0197] 图19为(a) 3GT1和3GT2, (b) 3GT3和 (c) 3GT4基因PCR产物的琼脂糖凝胶电泳检测图谱。

[0198] 图20为SDS-PAGE检测 (a) 3GT1和3GT2, (b) 3GT3和 (c) 3GT4基因在大肠杆菌中的表达; (a) 泳道1, pet28a空载体大肠杆菌重组子的裂解液总蛋白;泳道2, 3GT1-pet28a大肠杆菌重组子的裂解液上清;泳道3, 3GT1-pet28a大肠杆菌重组子的裂解液沉淀;泳道4, 3GT1-pet28a大肠杆菌重组子的总蛋白; ;泳道5, 3GT2-pet28a大肠杆菌重组子的裂解液上清;泳道6, 3GT2-pet28a大肠杆菌重组子的裂解液沉淀;泳道7, 3GT2-pet28a大肠杆菌重组子的总蛋白;泳道8, 蛋白分子量Marker (b) 泳道1, 蛋白分子量Marker;泳道2, 3GT3-pET28a大肠杆菌重组子的裂解液上清;泳道3, 3GT3-pET28a大肠杆菌重组子的裂解液沉淀;泳道4, 3GT3-pET28a大肠杆菌重组子的裂解液总蛋白; (c) ,泳道1, 3GT4-pET28a大肠杆菌重组子的裂解液总蛋白;泳道2, 3GT4-pET28a大肠杆菌重组子的裂解液沉淀;泳道3, 3GT4-pET28a大肠杆菌重组子的裂解液上清;泳道4, pet28a空载体大肠杆菌重组子的裂解液;泳道5, 蛋白分子量Marker。箭头指向目的蛋白的位置。

[0199] 图21为Western Blot检测检测 (a) 3GT1和3GT2, (b) 3GT3和 (c) 3GT4基因在大肠杆菌中的表达; (a) 泳道1, pet28a空载体大肠杆菌重组子的裂解液总蛋白;泳道2, 3GT1-pet28a大肠杆菌重组子的裂解液上清;泳道3, 3GT1-pet28a大肠杆菌重组子的裂解液沉淀;泳道4, 3GT1-pet28a大肠杆菌重组子的总蛋白; ;泳道5, 3GT2-pet28a大肠杆菌重组子的裂解液上清;泳道6, 3GT2-pet28a大肠杆菌重组子的裂解液沉淀;泳道7, 3GT2-pet28a大肠杆菌重组子的总蛋白; (b) 泳道1, 3GT3-pET28a大肠杆菌重组子的裂解液上清;泳道2, 3GT3-pET28a大肠杆菌重组子的裂解液沉淀;泳道3, 3GT3-pET28a大肠杆菌重组子的裂解液总蛋白; (c) ,泳道1, 3GT4-pET28a大肠杆菌重组子的裂解液总蛋白;泳道2, 3GT4-pET28a大肠杆菌重组子的裂解液沉淀;泳道3, 3GT4-pET28a大肠杆菌重组子的裂解液上清;泳道4, pet28a空载体大肠杆菌重组子的裂解液。

[0200] 图22显示糖基转移酶3GT1和3GT2催化原人参二醇和CK的产物TLC检测图谱。泳道1,原人参二醇型皂苷标样,泳道2,糖基转移酶3GT1催化原人参二醇生成Rh2;泳道3,糖基转移酶3GT1催化人参皂苷CK生成F2;泳道4,糖基转移酶3GT2催化原人参二醇生成Rh2;泳道5,糖基转移酶3GT2催化人参皂苷CK生成F2。

[0201] 图23显示糖基转移酶3GT1和3GT2催化达玛烯二醇(DM)和25-OH-PPD的产物TLC检测图谱(A) 3GT1粗酶液粗酶液(3GT1-pet28a大肠杆菌重组子裂解上清)催化达玛烯二醇(DM)和25-OH-PPD。泳道1, 25-OH-PPD标样, 泳道2, 3GT1粗酶液催化25-OH-PPD生成3-O- $\beta$ -(D-glucopyranosyl)-25-OH-protopanaxadiol; 泳道3, 达玛烯二醇标样; 泳道4, 3GT1粗酶液催化达玛烯二醇生成3-O- $\beta$ -(D-glucopyranosyl)-dammarendiolIII; (B) 3GT2粗酶液(3GT2-pet28a大肠杆菌重组子裂解上清)催化达玛烯二醇(DM)和25-OH-PPD。泳道1, 25-OH-PPD标样, 泳道2, 3GT2粗酶液催化25-OH-PPD生成3-O- $\beta$ -(D-glucopyranosyl)-25-OH-protopanaxadiol; 泳道3, 达玛烯二醇标样; 泳道4, 3GT2粗酶液催化达玛烯二醇生成3-O- $\beta$ -(D-glucopyranosyl)-dammarendiolIII。

[0202] 图24显示糖基转移酶3GT1和3GT2催化PPT和F1的的产物TLC检测, 泳道1, 3GT1粗酶液(3GT1-pet28a大肠杆菌重组子裂解上清)催化PPT生成3-O- $\beta$ -(D-glucopyranosyl)-protopanaxatriol; 泳道2, 3GT1粗酶液催化F1生成3-O- $\beta$ -(D-glucopyranosyl)-F1; 泳道3, 3GT2粗酶液(3GT2-pet28a大肠杆菌重组子裂解上清)催化PPT生成3-O- $\beta$ -(D-glucopyranosyl)-protopanaxatriol; 泳道4, 3GT2粗酶液催化F1生成3-O- $\beta$ -(D-glucopyranosyl)-F1。

[0203] 图25显示糖基转移酶3GT1和3GT2催化20(R)-PPD的产物TLC检测。泳道1, 20(R)-PPD标样; 泳道2, 3GT1粗酶液(3GT1-pet28a大肠杆菌重组子裂解上清)催化20(R)-PPD生成20(R)-Rh2; 泳道3, 3GT2粗酶液(3GT2-pet28a大肠杆菌重组子裂解上清)催化20(R)-PPD生成20(R)-Rh2; 泳道4, 对照, pet28a空载体大肠杆菌重组子裂解上清取代酶液; 泳道5, 20(R)-Rh2标样。

[0204] 图26显示糖基转移酶3GT1催化羊毛甾醇(lanosterol)的产物TLC检测。泳道1, 3GT1粗酶液(3GT1-pet28a大肠杆菌重组子裂解上清)催化lanosterol; 泳道2, 3GT2粗酶液(3GT2-pet28a大肠杆菌重组子裂解上清)催化lanosterol; 泳道3: 对照, pet28a空载体大肠杆菌重组子裂解上清取代酶液。

[0205] 图27显示糖基转移酶3GT3催化原人参二醇, 原人参三醇和25-OH-PPD的产物TLC检测图谱。(a) 3GT3粗酶液(3GT3-pet28a大肠杆菌重组子裂解上清)催化原人参二醇生成Rh2, M为原人参二醇型皂苷混合标样, (b) 3GT3粗酶液催化原人参皂苷三醇生成3-O- $\beta$ -(D-glucopyranosyl)-PPT(3-G-PPT); (c) 3GT3粗酶液催化25-OH-PPD生成3-O- $\beta$ -(D-glucopyranosyl)-25-OH-PPD(3-G-25-OH-PPD)。

[0206] 图28显示糖基转移酶3GT4催化原人参二醇, CK和25-OH-PPD的产物TLC检测图谱。(a) 3GT4粗酶液(3GT4-pet28a大肠杆菌重组子裂解上清)催化原人参二醇生成Rh2, M为原人参二醇型皂苷混合标样, “+”代表加入3GT4粗酶液的样品, “-”为对照, 即用pet28a空载体大肠杆菌重组子裂解上清代替酶液 (b) 3GT4粗酶液催化CK生成F2; “+”代表加入3GT4粗酶液的样品, “-”为对照, 即用pet28a空载体大肠杆菌重组子裂解上清代替酶液; (c) 糖基转移酶3GT4催化25-OH-PPD生成3-O- $\beta$ -(D-glucopyranosyl)-25-OH-PPD(3-G-25-OH-PPD), “+”代表加入3GT4粗酶液的样品, “-”为对照, 即用pet28a空载体大肠杆菌重组子裂解上清代替酶液。

[0207] 图29显示糖基转移酶3GT1, 3GT3和3GT4催化原人参二醇生成Rh2的HPLC检测, 第一行样品: 人参皂苷CK、Rh2和F2的混合标准样品; 第二行样品: 糖基转移酶3GT1粗酶液(3GT1-

pet28a大肠杆菌重组子裂解上清)催化PPD后的产物;第三行;3GT3粗酶液(3GT3-pet28a大肠杆菌重组子裂解上清)催化PPD后的产物;第四行;3GT4粗酶液(3GT4-pet28a大肠杆菌重组子裂解上清)催化PPD后的产物。

[0208] 图30显示糖基转移酶3GT1,3GT3和3GT4催化原人参二醇产物的LC/MS检测。显示Rh2标准样品的质谱图和图29中P1峰(3GT1的产物峰),P2峰(3GT2的产物峰)以及P3峰(3GT4的产物峰)的质谱图。

[0209] 图31显示(a) gGT29/gGT29-3基因和(b) gGT29-4/gGT29-5/gGT29-6和gGT29-7基因PCR产物的琼脂糖凝胶电泳图。(b)泳道1,核酸Marker;泳道2,gGT29/gGT29-3基因PCR产物;(b)泳道1,gGT29-4/gGT29-5/gGT29-6基因PCR产物;泳道2,gGT29-7基因PCR产物;泳道3,核酸Marker。

[0210] 图32显示SDS-PAGE检测gGT29和gGT29-3在酿酒酵母中的表达;泳道1,空载体pYES2重组子的裂解液上清;泳道2,gGT29-pYES2酵母重组子的裂解液上清;泳道3,gGT29-3-pYES2酵母重组子的裂解液上清。

[0211] 图33显示Western Blot检测gGT29和gGT29-3在酿酒酵母中的表达;泳道1,空载体pYES2重组子的裂解液上清;泳道2,gGT29-pYES2酵母重组子的裂解液上清;泳道3,gGT29-3-pYES2酵母重组子的裂解液上清。

[0212] 图34显示糖基转移酶gGT29和gGT29-3催化人参皂苷Rh2和F2的产物的TLC检测图谱;泳道1,PPD及PPD类型皂苷混合标样,泳道2,gGT29粗酶液(gGT29-pYES2酵母重组子的裂解液上清)催化Rh2生成Rg3,泳道3,gGT29粗酶液催化Rh2对照,加入pYES2空质粒酵母重组子裂解液替代酶液;泳道4,gGT29催化F2生成Rd,泳道5,gGT29催化F2对照,加入pYES2空质粒酵母重组子裂解液替代酶液;泳道6,gGT29-3粗酶液(gGT29-3-pYES2酵母重组子的裂解液上清)催化Rh2生成Rg3;泳道7,gGT29-3粗酶液催化F2生成Rd。

[0213] 图35显示糖基转移酶gGT29和3GT1或gGT29和3GT4联合催化PPD产物的TLC检测;(a) gGT29和3GT1联合催化PPD,泳道1,PPD及PPD类型皂苷混合标样;泳道2,3GT1催化PPD生成Rh2;泳道3,gGT29催化Rh2生成Rg3;泳道4,3GT1和gGT29联合催化PPD生成Rg3;(b) gGT29和3GT4联合催化PPD,泳道1,PPD及PPD类型皂苷混合标样;泳道2,3GT4催化PPD生成Rh2;泳道3,PPD;泳道4,3GT4和gGT29联合催化PPD生成Rg3。

[0214] 图36显示糖基转移酶3GT1和gGT29分别催化20(R)-PPD和20(R)-PPD以及联合催化20(R)-PPD的产物TLC检测图谱;泳道1,3GT1催化20(R)-PPD生成20(R)-Rh2;泳道2,gGT29催化20(R)-Rh2生成20(R)-Rg3;泳道3,3GT1和gGT29联合催化20(R)-PPD生成20(R)-Rg3。

[0215] 图37显示糖基转移酶gGT29和3GT1或gGT29和3GT4联合催化PPD产物的HPLC检测结果。第一行,Rg3,Rh2和PPD混合标准样品;第二行,gGT29和3GT1联合催化PPD,第三行,gGT29和3GT4联合催化PPD。

[0216] 图38显示糖基转移酶gGT29和3GT1或gGT29和3GT4联合催化PPD的产物LC/MS检测结果。显示了标准样品Rg3的质谱图和图37中P1峰(gGT29和3GT1联合催化PPD的产物)和P2峰(gGT29和3GT4联合催化PPD的产物产物峰)的质谱图。

[0217] 图39显示产Rh2酵母工程菌A1细胞裂解液抽提物的HPLC检测结果,第一行样品:原人参二醇(PPD)、达玛烯二醇(DM),人参皂苷Rh2和Rg3的混合标准样品;第二行样品:产Rh2酵母工程菌A1细胞裂解液抽提物。

[0218] 图40显示产Rg3酵母工程菌A2细胞裂解液抽提物的HPLC检测结果,第一行样品:原人参二醇(PPD)、达玛烯二醇(DM),人参皂苷Rh2和Rg3的混合标准样品;第二行样品:产Rg3酵母工程菌A2细胞裂解液抽提物。

[0219] 图41显示产Rh1酵母工程菌A3细胞裂解液抽提物的HPLC检测结果,第一行样品:原人参三醇(PPT)和人参皂苷Rh1的标准样品;第二行样品:产Rh1酵母工程菌A3细胞裂解液抽提物;

[0220] 图42显示产F1酵母工程菌A4细胞裂解液抽提物的HPLC检测结果,第一行样品:原人参三醇(PPT)和人参皂苷F1的标准样品;第二行样品:产F1酵母工程菌A4细胞裂解液抽提物。

[0221] 图43显示产Rh2酵母工程菌A5细胞裂解液抽提物的HPLC检测结果,第一行样品:达玛烯二醇(DM),原人参二醇(PPD),人参皂苷Rh2和人参皂苷Rg3的标准样品;第二行样品:产Rh2酵母工程菌A5细胞裂解液抽提物。

[0222] 图44显示SDS-PAGE检测gGT29-4,gGT29-5,gGT29-6,gGT29-7在重组大肠杆菌中的表达。泳道1,gGT29-4-pET28a重组大肠杆菌体裂解总蛋白;泳道2,gGT29-4-pET28a重组大肠杆菌体裂解上清;泳道3,gGT29-5-pET28a重组大肠杆菌体裂解总蛋白;泳道4,gGT29-5-pET28a重组大肠杆菌体裂解上清;泳道5,gGT29-6-pET28a重组大肠杆菌体裂解总蛋白;泳道6,gGT29-6-pET28a重组大肠杆菌体裂解上清;泳道7,gGT29-7-pET28a重组大肠杆菌体裂解总蛋白;泳道8,gGT29-7-pET28a重组大肠杆菌体裂解上清;泳道9,蛋白分子量Marker。

[0223] 图45显示Western Blot检测gGT29-4,gGT29-5,gGT29-6,gGT29-7在重组大肠杆菌中的表达。泳道1,gGT29-4-pET28a重组大肠杆菌体裂解总蛋白;泳道2,gGT29-4-pET28a重组大肠杆菌体裂解上清;泳道3,gGT29-5-pET28a重组大肠杆菌体裂解总蛋白;泳道4,gGT29-5-pET28a重组大肠杆菌体裂解上清;泳道5,gGT29-6-pET28a重组大肠杆菌体裂解总蛋白;泳道6,gGT29-6-pET28a重组大肠杆菌体裂解上清;泳道7,gGT29-7-pET28a重组大肠杆菌体裂解总蛋白;泳道8,gGT29-7-pET28a重组大肠杆菌体裂解上清。

[0224] 图46显示糖基转移酶gGT29-4,gGT29-5,gGT29-6,gGT29-7分别催化Rh2和F2的产物TLC检测图谱。泳道Rh2表示用皂苷Rh2为底物;泳道F2表示用皂苷F2为底物。gGT29-4,gGT29-5,gGT29-6,gGT29-7表示用不同的酶液进行催化反应。

## 具体实施方式

[0225] 本发明人经过广泛而深入的研究,首次提供糖基转移酶gGT25(SEQ ID NO.:2)、gGT25-1(SEQ ID NO.:16)、gGT25-3(SEQ ID NO.:18)、gGT25-5(SEQ ID NO.:20)、gGT29(SEQ ID NO.:26)、gGT29-3(SEQ ID NO.:28)、gGT29-4(SEQ ID NO.:55)、gGT29-5(SEQ ID NO.:57)、gGT29-6(SEQ ID NO.:59)、gGT29-7(SEQ ID NO.:61)以及3GT1(SEQ ID NO.:22)、3GT2(SEQ ID NO.:24)、3GT3(SEQ ID NO.:41)、3GT4(SEQ ID NO.:43)、gGT13(SEQ ID NO.:4)、gGT30(SEQ ID NO.:6)在萜类化合物糖基化催化及新皂苷合成中的应用。具体地,本发明的糖基转移酶能特异和高效地催化四环三萜化合物底物的C-20位和/或C-6位和/或C-3的羟基糖基化,和/或将来自糖基供体的糖基转移到四环三萜类化合物的C-3位的第一个糖基上以延伸糖链。特别是能够将原人参二醇转化为具有抗癌活性的稀有人参皂苷CK和Rh2,将原人参三醇转化为具有抗衰老活性的稀有人参皂苷F1和抗过敏作用的稀有人参皂苷

Rh1,将Rh2转化为抗癌活性优良的稀有人参皂苷Rg3。本发明的糖基转移酶还能将达玛稀二醇、原人参三醇、F1、25-OH-PPD、25-OCH<sub>3</sub>-PPD合成之前未见报道的新皂苷20-0-β-(D-glucopyranosyl)-dammarendiolIII、3-0-β-(D-glucopyranosyl)-dammarendiolIII、3-0-β-(D-glucopyranosyl)-PPT、3-0-β-(D-glucopyranosyl)-F1、3-0-β-(D-glucopyranosyl)-25-OH-PPD、3-0-β-(D-glucopyranosyl)-25-OCH<sub>3</sub>-PPD。

[0226] 本发明的糖基转移酶还能将Rh2,CK,Rg3分别转化为人参皂苷人参皂苷F2,Rd和Rg1等。本发明还提供了转化和催化方法。本发明的糖基转移酶还可与达玛烯二醇和/或原人参二醇或原人参三醇合成代谢途径中的关键酶在宿主细胞中共表达,或者应用于制备达玛烯二醇(DM),原人参二醇(PPD)和原人参三醇(PPT)的遗传工程细胞中,应用于构建人工合成稀有人参皂苷CK,F1,Rh1,Rh2和Rg3,以及新的人参皂苷20-0-β-(D-glucopyranosyl)-dammarendiolIII、3-0-β-(D-glucopyranosyl)-PPT、3-0-β-(D-glucopyranosyl)-F1、3-0-β-(D-glucopyranosyl)-dammarendiolIII、3-0-β-(D-glucopyranosyl)-25-OH-PPD、3-0-β-(D-glucopyranosyl)-25-OCH<sub>3</sub>-PPD和F2,Rd和Rg1等的代谢途径中。在此基础上完成了本发明。

[0227] 定义

[0228] 如本文所用,术语“活性多肽”、“本发明的多肽及其衍生多肽”、“本发明的酶”、“糖基转移酶”、“本发明的gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、3GT1、3GT2、3GT3、3GT4蛋白”或“本发明的糖基转移酶”,均指糖基转移酶gGT25(SEQ ID NO.:2)、gGT13(SEQ ID NO.:4)、gGT30(SEQ ID NO.:6)、gGT25-1(SEQ ID NO.:16)、gGT25-3(SEQ ID NO.:18)、gGT25-5(SEQ ID NO.:20)、gGT29(SEQ ID NO.:26)、gGT29-3(SEQ ID NO.:28)、gGT29-4(SEQ ID NO.:55)、gGT29-5(SEQ ID NO.:57)、gGT29-6(SEQ ID NO.:59)、gGT29-7(SEQ ID NO.:61)以及3GT1(SEQ ID NO.:22)、3GT2(SEQ ID NO.:24)、3GT3(SEQ ID NO.:41)、3GT4(SEQ ID NO.:43)多肽及其衍生多肽。

[0229] 如非特别说明,本文所说的人参皂苷和皂苷元,是的C20位S构型的人参皂苷和皂苷元。

[0230] 如本文所用,“分离的多肽”是指所述多肽基本上不含天然与其相关的其它蛋白、脂类、糖类或其它物质。本领域的技术人员能用标准的蛋白质纯化技术纯化所述多肽。基本上纯的多肽在非还原聚丙烯酰胺凝胶上能产生单一的主带。所述多肽的纯度还可以用氨基酸序列进行进一步分析。

[0231] 本发明的活性多肽可以是重组多肽、天然多肽、合成多肽。本发明的多肽可以是天然纯化的产物,或是化学合成的产物,或使用重组技术从原核或真核宿主(例如,细菌、酵母、植物)中产生。根据重组生产方案所用的宿主,本发明的多肽可以是糖基化的,或可以是非糖基化的。本发明的多肽还可包括或不包括起始的甲硫氨酸残基。

[0232] 本发明还包括所述多肽的片段、衍生物和类似物。如本文所用,术语“片段”、“衍生物”和“类似物”是指基本上保持所述多肽相同的生物学功能或活性的多肽。

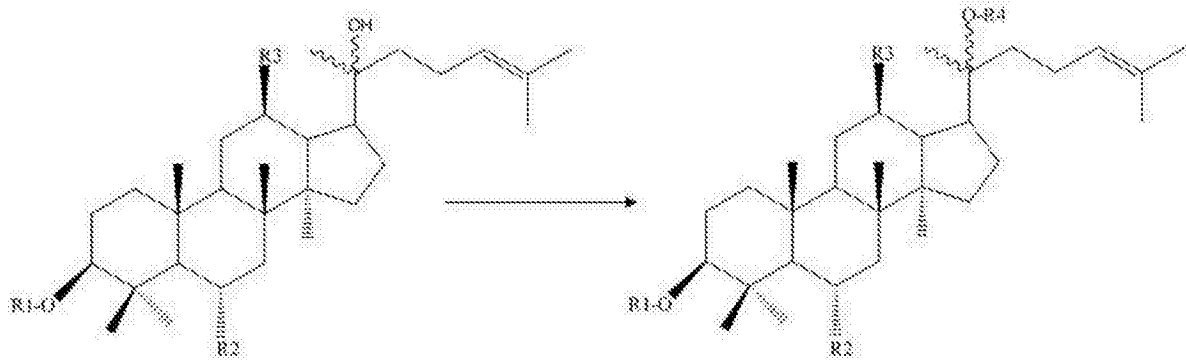
[0233] 本发明的多肽片段、衍生物或类似物可以是(i)有一个或多个保守或非保守性氨基酸残基(优选保守性氨基酸残基)被取代的多肽,而这样的取代的氨基酸残基可以是也可以不是由遗传密码编码的,或(ii)在一个或多个氨基酸残基中具有取代基团的多肽,或(iii)成熟多肽与另一个化合物(比如延长多肽半衰期的化合物,例如聚乙二醇)融合所形

成的多肽,或(iv)附加的氨基酸序列融合到此多肽序列而形成的多肽(如前导序列或分泌序列或用来纯化此多肽的序列或蛋白原序列,或与抗原IgG片段的形成的融合蛋白)。根据本文的教导,这些片段、衍生物和类似物属于本领域熟练技术人员公知的范围。

[0234] 在本发明的活性多肽具有糖基转移酶活性,并且能够催化以下一种或多种反应:

[0235]

(A)



式(I)化合物

式(II)化合物

[0236] 其中,R1为H,单糖糖基或多糖糖基;R2和R3为H或OH;R4为糖基;所述的多肽选自SEQ ID NO:2、16或18或其衍生多肽。

[0237] 在另一优选例中,所述的单糖包括葡萄糖(Glc),鼠李糖(Rha),乙酰基葡萄糖(Glc(6)Ac),阿拉伯呋喃糖(Araf),阿拉伯吡喃糖(Arap),木糖(Xyl)等。

[0238] 在另一优选例中,所述的多糖包括Glc(2-1)Glc,Glc(6-1)Glc,Glc(6)Ac,Glc(2-1)Rha,Glc(6-1)Arap,Glc(6-1)Xyl,Glc(6-1)Araf,Glc(3-1)Glc(3-1),Glc(2-1)Glu(6)Ac,Glc(6-1)Arap(4-1)Xyl,Glc(6-1)Arap(2-1)Xyl,Glc(6-1)Arap(3-1)Xyl等2-4个单糖组成的多糖。

[0239] R1-R4经取代后的化合物如下表所示:

[0240]

底物	R1	R2	R3	R4	产物
PPD	H	H	OH	糖基	CK

[0241]

Rh2	1个糖基	H	OH	糖基	F2
Rg3	2个糖基	H	OH	糖基	Rd
PPT	H	OH	OH	糖基	F1
DM	H	H	H	糖基	20-G-DM

[0242] 即当所述的R1,R2都为H,R3为OH时,所述的式(I)化合物为原人参二醇(PPD)

[0243] R1为一个葡萄糖基,R2为H,R3为OH时,所述的式(I)化合物为人参皂苷RH2.

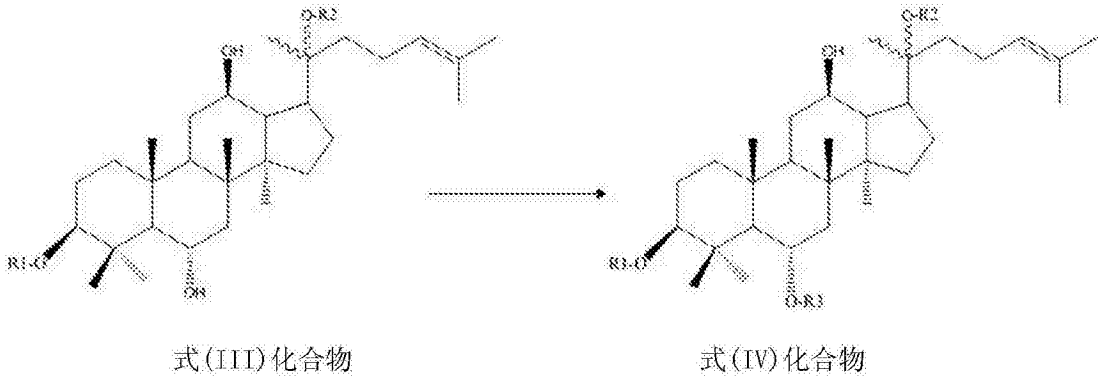
[0244] R1为两个葡萄糖基,R2为H,R3为OH时,所述的式(I)化合物为人参皂苷RG3.

[0245] R1为H,R2为OH,R3为OH,所述的式(I)化合物为原人参三醇(PPT)。

[0246] R1为H,R2为H,R3为H,所述的式(I)化合物达玛烯二醇(DM)。

[0247]

(B):



[0248] 其中,R1为H或者糖基,R2糖基,R3为糖基,所述的多肽选自SEQ ID NOs.:2、16、18或20或其衍生多肽;

[0249] 或,R1为H或者糖基;R2为H或者糖基;R3为糖基,所述的多肽选自SEQ ID NO.:20或其衍生多肽。

[0250] R1-R3经取代后的化合物如下表所示:

[0251]

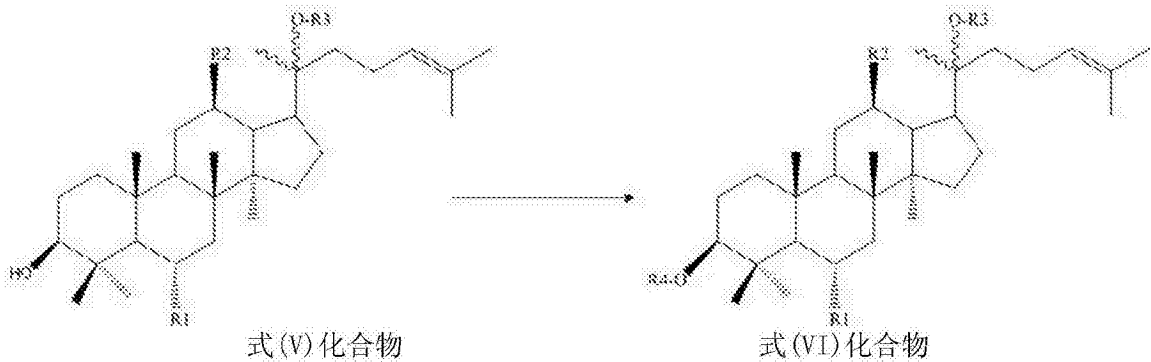
底物	R1	R2	R3	产物
F1	H	糖基	糖基	Rg1
PPT	H	H	糖基	Rh1

[0252] 即当R1,R2都为H时,所述的式(III)化合物为原人参三醇(PPT)。

[0253] R1为H,R2为葡萄糖基时,所述的式(III)化合物为人参皂苷F1。

[0254]

(C):



[0255] 其中,R1为H或者OH;R2为H或者OH;R3为H或者糖基;R4为糖基,所述的多肽选自SEQ ID NOs.:22、24、41或43或其衍生多肽。

[0256] R1-R4经取代后的化合物如下表所示:

[0257]

底物	R1	R2	R3	R4	产物
PPD	H	OH	H	糖基	Rh2
CK	H	OH	糖基	糖基	F2

[0258]

PPT	OH	OH	H	糖基	3-G-PPT
F1	OH	OH	糖基	糖基	3-G-F1
DM	H	H	H	糖基	3-G-DM

[0259] 当R1和R3都为H,R2为OH时,所述的式(V)化合物为原人参二醇(PPD),所述的多肽选自SEQ ID N0s.:22、24、41或43或其衍生多肽。

[0260] 当R1为H,R2为OH,R3为葡萄糖基时,所述的式(V)化合物为人参皂苷CK,所述的多肽选自SEQ ID N0s.:22、24或43或其衍生多肽;

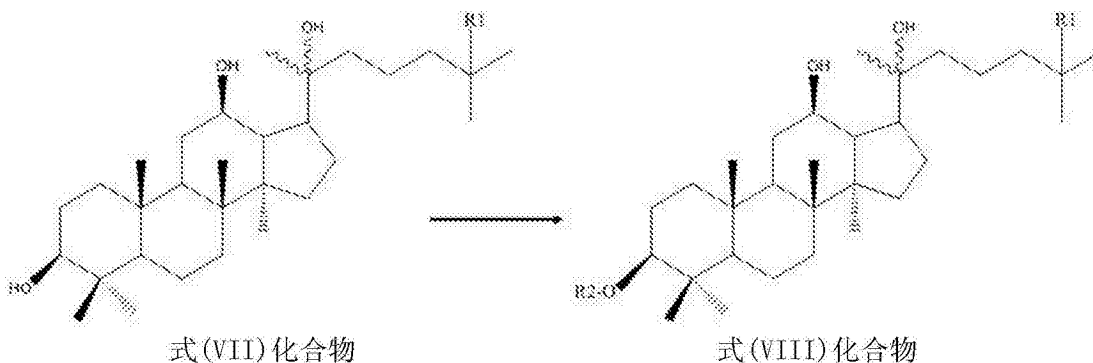
[0261] 当R1为OH,R2为OH,R3为H时,所述的式(V)化合物为原人参三醇(PPT),所述的多肽选自SEQ ID N0s.:22、24或41或其衍生多肽;

[0262] 当R1为OH,R2为OH,R3为葡萄糖基时,所述的式(V)化合物为人参皂苷F1,所述的多肽选自SEQ ID N0s.:22或24或其衍生多肽;

[0263] 当R1为H,R2为OH,R3为为H时,所述的式(V)化合物为达玛烯二醇(DM)所述的多肽选自SEQ ID N0s.:22或24或其衍生多肽。

[0264]

(D):



[0265] 其中,R1为OH或者OCH<sub>3</sub>;R2为糖基,所述的多肽选自SEQ ID N0s.:22、24、41或43或其衍生多肽。

[0266] R1-R2经取代后的化合物如下表所示:

[0267]

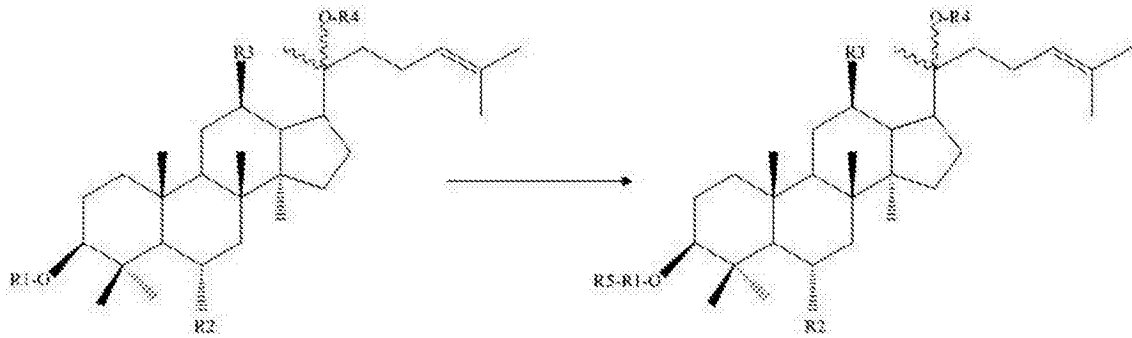
底物	R1	R2	产物
25-OH-PPD	OH	糖基	3-G-25-OH-PPD
25-OCH <sub>3</sub> -PPD	OCH <sub>3</sub>	糖基	3-G-25-OCH <sub>3</sub> -PPD

[0268] 即当R1为OH时,所述的式(VII)化合物为25-OH-PPD;

[0269] R1为OCH<sub>3</sub>时,所述的式(VII)化合物为25-OCH<sub>3</sub>-PPD。

[0270]

(E)



式(IX)化合物

式(X)化合物

[0271] 其中,R1为糖基;R2和R3为OH或者H;R4为糖基或者H;R5为糖基,所述的多肽选自SEQID NOs.:26、28、55、57、59或61或其衍生多肽。

[0272] R1-R4经取代后的化合物如下表所示:

[0273]

底物	R1	R2	R3	R4	产物
Rh2	糖基	H	OH	H	Rg3

[0274]

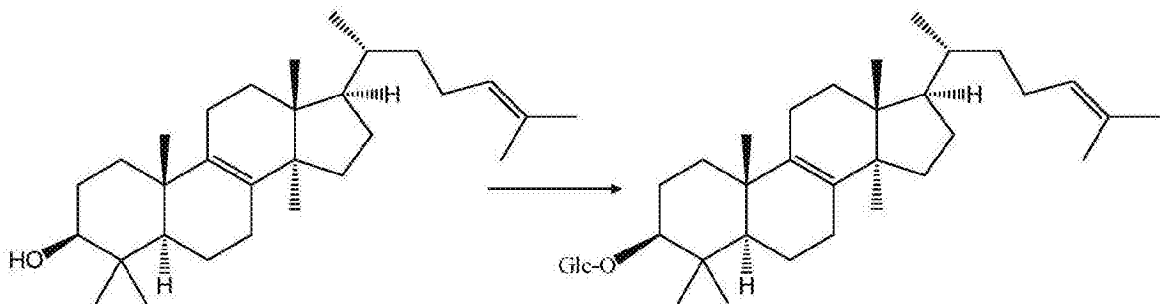
F2	糖基	H	OH	糖基	Rd
----	----	---	----	----	----

[0275] 即当R1为葡萄糖基;R2为H,R3为OH,R4为H,式(IX)化合物为Rh2。

[0276] R1为葡萄糖基;R2为H,R3为OH,R4为葡萄糖基,式(IX)化合物为F2。

[0277]

(F)



式(XI)化合物

式(XI1)化合物

[0278] 所述的多肽选自SEQ ID NOs.:22或24或其衍生多肽。

[0279] 所述的多肽优选序列为SEQ ID NOs.:2、16、18、20、22、24、41、26、28、43、55、57、59或61所示的多肽,该术语还包括具有与所示多肽具有相同功能的、SEQ ID NOs.:2、16、18、20、22、24、41、26、28、43、55、57、59或61序列的变异形式及衍生多肽。这些变异形式包括(但并不限于):一个或多个(通常为1-50个,较佳地1-30个,更佳地1-20个,最佳地1-10个)氨基酸的缺失、插入和/或取代,以及在C末端和/或N末端添加一个或数个(通常为20个以内,较佳地为10个以内,更佳地为5个以内)氨基酸。例如,在本领域中,用性能相近或相似的氨基酸进行取代时,通常不会改变蛋白质的功能。又比如,在C末端和/或N末端添加一个或数个

氨基酸通常也不会改变蛋白质的功能。该术语还包括人EGFRvA蛋白的活性片段和活性衍生物。本发明还提供所述多肽的类似物。这些类似物与天然人EGFRvA多肽的差别可以是氨基酸序列上的差异,也可以是不影响序列的修饰形式上的差异,或者兼而有之。这些多肽包括天然或诱导的遗传变异体。诱导变异体可以通过各种技术得到,如通过辐射或暴露于诱变剂而产生随机诱变,还可通过定点诱变法或其他已知分子生物学的技术。类似物还包括具有不同于天然L-氨基酸的残基(如D-氨基酸)的类似物,以及具有非天然存在的或合成的氨基酸(如 $\beta$ 、 $\gamma$ -氨基酸)的类似物。应理解,本发明的多肽并不限于上述例举的代表性的多肽。

[0280] 修饰(通常不改变一级结构)形式包括:体内或体外的多肽的化学衍生形式如乙酰化或羧基化。修饰还包括糖基化,如那些在多肽的合成和加工中或进一步加工步骤中进行糖基化修饰而产生的多肽。这种修饰可以通过将多肽暴露于进行糖基化的酶(如哺乳动物的糖基化酶或去糖基化酶)而完成。修饰形式还包括具有磷酸化氨基酸残基(如磷酸酪氨酸,磷酸丝氨酸,磷酸苏氨酸)的序列。还包括被修饰从而提高了其抗蛋白水解性能或优化了溶解性能的多肽。

[0281] 本发明的gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7以及3GT1、3GT2、3GT3、3GT4蛋白的氨基端或羧基端还可含有一个或多个多肽片段,作为蛋白标签。任何合适的标签都可以用于本发明。例如,所述的标签可以是FLAG、HA、HA1、c-Myc、Poly-His、Poly-Arg、Strep-TagII、AU1、EE、T7、4A6、 $\epsilon$ 、B、gE、以及Ty1。这些标签可用于对蛋白进行纯化。表1列出了其中的一些标签及其序列。

[0282] 表1

标签	残基数	序列
Poly-Arg	5-6 个(通常 5 个)	RRRRR
Poly-His	2-10 个(通常 6 个)	HHHHHH
FLAG	8 个	DYKDDDDK
Strep-TagII	8 个	WSHPQFEK
C-myc	10 个	WQKLISEEDL
GST	220 个	后面 6 个 LVPRGS

[0285] 为了使翻译的蛋白分泌表达(如分泌到细胞外),还可在所述gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7以及3GT1、3GT2、3GT3、3GT4的氨基酸氨基末端添加上信号肽序列,如pe1B信号肽等。信号肽在多肽从细胞内分泌出来的过程中可被切去。

[0286] 本发明的多核苷酸可以是DNA形式或RNA形式。DNA形式包括cDNA、基因组DNA或人工合成的DNA。DNA可以是单链的或是双链的。DNA可以是编码链或非编码链。编码成熟多肽的编码区序列可以与SEQ ID NOs.:1所示的编码区序列相同或者是简并的变异体。如本文所用,“简并的变异体”在本发明中是指编码具有SEQ ID NOs.:2、4、6、16、18、20、22、24、26、28、41、43、55、57、59或61的蛋白质,但与SEQ ID NOs.:1、3、5、15、17、19、21、23、25、27、40、42、54、56、58或60所示的编码区序列有差别的核酸序列。

[0287] 编码SEQ ID NOs.:2、4、6、16、18、20、22、24、26、28、41、43、55、57、59或61的成熟多肽的多核苷酸包括:只编码成熟多肽的编码序列;成熟多肽的编码序列和各种附加编码序列;成熟多肽的编码序列(和任选的附加编码序列)以及非编码序列。

[0288] 术语“编码多肽的多核苷酸”可以是包括编码此多肽的多核苷酸,也可以是还包括附加编码和/或非编码序列的多核苷酸。

[0289] 本发明还涉及上述多核苷酸的变异体,其编码与本发明有相同的氨基酸序列的多肽或多肽的片段、类似物和衍生物。此多核苷酸的变异体可以是天然发生的等位变异体或非天然发生的变异体。这些核苷酸变异体包括取代变异体、缺失变异体和插入变异体。如本领域所知的,等位变异体是一个多核苷酸的替换形式,它可能是一个或多个核苷酸的取代、缺失或插入,但不会从实质上改变其编码的多肽的功能。

[0290] 本发明还涉及与上述的序列杂交且两个序列之间具有至少50%,更佳地至少70%,更佳地至少80%相同性的多核苷酸。本发明特别涉及在严格条件(或严紧条件)下与本发明所述多核苷酸可杂交的多核苷酸。在本发明中,“严格条件”是指:(1)在较低离子强度和较高温度下的杂交和洗脱,如 $0.2 \times \text{SSC}$ , $0.1\% \text{SDS}$ , $60^\circ\text{C}$ ;或(2)杂交时加有变性剂,如50%(v/v)甲酰胺, $0.1\%$ 小牛血清/ $0.1\% \text{Ficol1}$ , $42^\circ\text{C}$ 等;或(3)仅在两条序列之间的相同性至少在90%以上,更好是95%以上时才发生杂交。并且,可杂交的多核苷酸编码的多肽与SEQ ID NOs.:2、4、6、16、18、20、22、24、26、28、41、43、55、57、59或61所示的成熟多肽有相同的生物学功能和活性。

[0291] 本发明还涉及与上述的序列杂交的核酸片段。如本文所用,“核酸片段”的长度至少含15个核苷酸,较好是至少30个核苷酸,更好是至少50个核苷酸,最好是至少100个核苷酸以上。核酸片段可用于核酸的扩增技术(如PCR)以确定和/或分离编码gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7以及3GT1、3GT2、3GT3、3GT4蛋白的多聚核苷酸。

[0292] 本发明中的多肽和多核苷酸优选以分离的形式提供,更佳地被纯化至均质。

[0293] 本发明的gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7以及3GT1、3GT2、3GT3、3GT4核苷酸全长序列或其片段通常可以用PCR扩增法、重组法或人工合成的方法获得。对于PCR扩增法,可根据本发明所公开的有关核苷酸序列,尤其是开放阅读框序列来设计引物,并用市售的cDNA库或按本领域技术人员已知的常规方法所制备的cDNA库作为模板,扩增而得有关序列。当序列较长时,常常需要进行两次或多次PCR扩增,然后再将各次扩增出的片段按正确次序拼接在一起。

[0294] 一旦获得了有关的序列,就可以用重组法来大批量地获得有关序列。这通常是将其克隆入载体,再转入细胞,然后通过常规方法从增殖后的宿主细胞中分离得到有关序列。

[0295] 此外,还可用人工合成的方法来合成有关序列,尤其是片段长度较短时。通常,通过先合成多个小片段,然后再进行连接可获得序列很长的片段。

[0296] 目前,已经可以完全通过化学合成来得到编码本发明蛋白(或其片段,或其衍生物)的DNA序列。然后可将该DNA序列引入本领域中已知的各种现有的DNA分子(或如载体)和细胞中。此外,还可通过化学合成将突变引入本发明蛋白序列中。

[0297] 应用PCR技术扩增DNA/RNA的方法被优选用于获得本发明的基因。特别是很难从文库中得到全长的cDNA时,可优选使用RACE法(RACE-cDNA末端快速扩增法),用于PCR的引物

可根据本文所公开的本发明的序列信息适当地选择,并可用常规方法合成。可用常规方法如通过凝胶电泳分离和纯化扩增的DNA/RNA片段。

[0298] 本发明也涉及包含本发明的多核苷酸的载体,以及用本发明的载体或gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7以及3GT1、3GT2、3GT3、3GT4蛋白编码序列经基因工程产生的宿主细胞,以及经重组技术产生本发明所述多肽的方法。

[0299] 通过常规的重组DNA技术,可利用本发明的多聚核苷酸序列可用来表达或生产重组的gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7以及3GT1、3GT2、3GT3、3GT4多肽。一般来说有以下步骤:

[0300] (1). 用本发明的编码gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7以及3GT1、3GT2、3GT3、3GT4多肽的多核苷酸(或变体),或用含有该多核苷酸的重组表达载体转化或转导合适的宿主细胞;

[0301] (2). 在合适的培养基中培养的宿主细胞;

[0302] (3). 从培养基或细胞中分离、纯化蛋白质。

[0303] 本发明中,gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7以及3GT1、3GT2、3GT3、3GT4多核苷酸序列可插入到重组表达载体中。术语“重组表达载体”指本领域熟知的细菌质粒、噬菌体、酵母质粒、植物细胞病毒、哺乳动物细胞病毒如腺病毒、逆转录病毒或其他载体。只要能在宿主体内复制和稳定,任何质粒和载体都可以用。表达载体的一个重要特征是通常含有复制起点、启动子、标记基因和翻译控制元件。

[0304] 本领域的技术人员熟知的方法能用于构建含gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7以及3GT1、3GT2、3GT3、3GT4编码DNA序列和合适的转录/翻译控制信号的表达载体。这些方法包括体外重组DNA技术、DNA合成技术、体内重组技术等。所述的DNA序列可有效连接到表达载体中的适当启动子上,以指导mRNA合成。这些启动子的代表性例子有:大肠杆菌的lac或trp启动子; $\lambda$ 噬菌体PL启动子;真核启动子包括CMV立即早期启动子、HSV胸苷激酶启动子、早期和晚期SV40启动子、反转录病毒的LTRs和其他一些已知的可控制基因在原核或真核细胞或其病毒中表达的启动子。表达载体还包括翻译起始用的核糖体结合位点和转录终止子。

[0305] 此外,表达载体优选地包含一个或多个选择性标记基因,以提供用于选择转化的宿主细胞的表型性状,如真核细胞培养用的二氢叶酸还原酶、新霉素抗性以及绿色荧光蛋白(GFP),或用于大肠杆菌的四环素或氨苄青霉素抗性。

[0306] 包含上述的适当DNA序列以及适当启动子或者控制序列的载体,可以用于转化适当的宿主细胞,以使其能够表达蛋白质。

[0307] 宿主细胞可以是原核细胞,如细菌细胞;或是低等真核细胞,如酵母细胞;或是高等真核细胞,如哺乳动物细胞。代表性例子有:大肠杆菌,链霉菌属;鼠伤寒沙门氏菌的细菌细胞;真菌细胞如酵母;植物细胞;果蝇S2或Sf9的昆虫细胞;CHO、COS、293细胞、或Bowes黑色素瘤细胞的动物细胞等。

[0308] 本发明的多核苷酸在高等真核细胞中表达时,如果在载体中插入增强子序列时将会使转录得到增强。增强子是DNA的顺式作用因子,通常大约有10到300个碱基对,作用于启

动子以增强基因的转录。可举的例子包括在复制起始点晚期一侧的100到270个碱基对的SV40增强子、在复制起始点晚期一侧的多瘤增强子以及腺病毒增强子等。

[0309] 本领域一般技术人员都清楚如何选择适当的载体、启动子、增强子和宿主细胞。

[0310] 用重组DNA转化宿主细胞可用本领域技术人员熟知的常规技术进行。当宿主为原核生物如大肠杆菌时,能吸收DNA的感受态细胞可在指数生长期后收获,用CaCl<sub>2</sub>法处理,所用的步骤在本领域众所周知。另一种方法是使用MgCl<sub>2</sub>。如果需要,转化也可用电穿孔的方法进行。当宿主是真核生物,可选用如下的DNA转染方法:磷酸钙共沉淀法,常规机械方法如显微注射、电穿孔、脂质体包装等。

[0311] 获得的转化子可以用常规方法培养,表达本发明的基因所编码的多肽。根据所用的宿主细胞,培养中所用的培养基可选自各种常规培养基。在适于宿主细胞生长的条件下进行培养。当宿主细胞生长到适当的细胞密度后,用合适的方法(如温度转换或化学诱导)诱导选择的启动子,将细胞再培养一段时间。

[0312] 在上面的方法中的重组多肽可在细胞内、或在细胞膜上表达、或分泌到细胞外。如果需要,可利用其物理的、化学的和其它特性通过各种分离方法分离和纯化重组的蛋白。这些方法是本领域技术人员所熟知的。这些方法的例子包括但不限于:常规的复性处理、用蛋白沉淀剂处理(盐析方法)、离心、渗透破菌、超处理、超离心、分子筛层析(凝胶过滤)、吸附层析、离子交换层析、高效液相层析(HPLC)和其它各种液相层析技术及这些方法的结合。

[0313] 应用

[0314] 本发明涉及的活性多肽或肽基转移酶gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7以及3GT1、3GT2、3GT3、3GT4的用途包括(但不限于):特异和高效地催化四环三萜化合物底物的C-20位和/或C-6位和/或C-3的羟基糖基化或将来自糖基供体的糖基转移到四环三萜类化合物的C-3位的第一个糖基上以延伸糖链。特别是能够将原人参二醇转化为具有抗癌活性的稀有人参皂苷CK和Rh2,将原人参三醇转化为具有抗衰老活性的稀有人参皂苷F1和抗过敏作用的稀有人参皂苷Rh1,将Rh2转化为抗癌活性更优良的稀有人参皂苷Rg3。本发明的糖基转移酶还能将达玛稀二醇、原人参三醇、F1、25-OH-PPD、25-OCH<sub>3</sub>-PPD合成之前未见报道的新皂苷20-O-β-(D-glucopyranosyl)-dammarendiolIII、3-O-β-(D-glucopyranosyl)-dammarendiolIII、3-O-β-(D-glucopyranosyl)-PPT、3-O-β-(D-glucopyranosyl)-F1、3-O-β-(D-glucopyranosyl)-25-OH-PPD、3-O-β-(D-glucopyranosyl)-25-OCH<sub>3</sub>-PPD。本发明的糖基转移酶还能将Rh2, CK, Rg3转化为人参皂苷人参皂苷F2, Rd和Rg1等。

[0315] 所述的四环三萜化合物包括(但不限于):S构型或R构型的达玛烷型、羊毛脂烷型、甘遂烷型、环阿屯烷(环阿尔廷烷)型、葫芦烷、楝烷型等四环三萜类化合物。

[0316] 本发明提供了一种工业催化方法,包括:在提供糖基供体的条件下,用本发明的gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7、3GT1、3GT2、3GT3和/或3GT4活性多肽或肽基转移酶获得(II),(IV),(VI),(VIII),(X)和(XII)化合物。具体是,所述的(a)反应中所用的多肽选自SEQ ID N0s.:2、16或18;所述的(b)反应中所用的多肽选自SEQ ID N0s.:20、2、16或18;所述的(c)和(d)反应所用的多肽选自SEQ ID N0s.:22、24、41和43;所述的(e)反应中所用的多肽选自SEQ ID N0s.:26、28、55、57、59或61所示氨基酸序列的活性多肽;所述的(F)反应中所用的多肽

选自SEQ ID NOs.:22或24所示氨基酸序列的活性多肽;。

[0317] 所述的糖基供体是核苷二磷酸糖,选自下组:UDP-葡萄糖,ADP-葡萄糖,TDP-葡萄糖,CDP-葡萄糖,GDP-葡萄糖,UDP-乙酰基葡萄糖,ADP-乙酰基葡萄糖,TDP-乙酰基葡萄糖,CDP-乙酰基葡萄糖,GDP-乙酰基葡萄糖,UDP-木糖,ADP-木糖,TDP-木糖,CDP-木糖,GDP-木糖,UDP-半乳糖醛酸,ADP-半乳糖醛酸,TDP-半乳糖醛酸,CDP-半乳糖醛酸,GDP-半乳糖醛酸,UDP-半乳糖,ADP-半乳糖,TDP-半乳糖,CDP-半乳糖,GDP-半乳糖,UDP-阿拉伯糖,ADP-阿拉伯糖,TDP-阿拉伯糖,CDP-阿拉伯糖,GDP-阿拉伯糖,UDP-鼠李糖,ADP-鼠李糖,TDP-鼠李糖,CDP-鼠李糖,GDP-鼠李糖,或其他核苷二磷酸己糖或核苷二磷酸戊糖,或其组合。

[0318] 所述的糖基供体优选是尿苷二磷酸糖,选自下组:UDP-葡萄糖,UDP-半乳糖醛酸,UDP-半乳糖,UDP-阿拉伯糖,UDP-鼠李糖,或其他尿苷二磷酸己糖或尿苷二磷酸戊糖,或其组合。

[0319] 在所述方法中,还可以添加酶活性添加物(提高酶活性或抑制酶活性的添加物)。所述酶活性的添加物可以选自下组:Ca<sup>2+</sup>、Co<sup>2+</sup>、Mn<sup>2+</sup>、Ba<sup>2+</sup>、Al<sup>3+</sup>、Ni<sup>2+</sup>、Zn<sup>2+</sup>、或Fe<sup>2+</sup>;或为可以生成Ca<sup>2+</sup>、Co<sup>2+</sup>、Mn<sup>2+</sup>、Ba<sup>2+</sup>、Al<sup>3+</sup>、Ni<sup>2+</sup>、Zn<sup>2+</sup>、或Fe<sup>2+</sup>的物质。

[0320] 所述方法的pH条件为:pH4.0-10.0,优选pH6.0-pH8.5,更优选8.5。

[0321] 所述方法的温度条件为:10℃-105℃,优选25℃-35℃,更优选35℃。

[0322] 本发明还提供了一种组合物,它含有有效量的本发明的活性多肽或肽基转移酶gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7、3GT1、3GT2、3GT3和3GT4,以及食品学上或工业上可接受的载体或赋形剂。这类载体包括(但并不限于):水、缓冲液、葡萄糖、水、甘油、乙醇、及其组合。

[0323] 所述的组合物中还可添加调节本发明的gGT25酶活性的物质。任何具有提高酶活性功能的物质均是可用的。较佳地,所述的提高本发明的gGT25酶活性的物质选自巯基乙醇。此外,很多物质可以降低酶活性,包括但不限于:Ca<sup>2+</sup>、Co<sup>2+</sup>、Mn<sup>2+</sup>、Ba<sup>2+</sup>、Al<sup>3+</sup>、Ni<sup>2+</sup>、Zn<sup>2+</sup>和Fe<sup>2+</sup>;或在添加至底物后可水解形成Ca<sup>2+</sup>、Co<sup>2+</sup>、Mn<sup>2+</sup>、Ba<sup>2+</sup>、Al<sup>3+</sup>、Ni<sup>2+</sup>、Zn<sup>2+</sup>和Fe<sup>2+</sup>的物质。

[0324] 在获得了本发明的gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7以及3GT1、3GT2、3GT3、3GT4后,本领域人员可以方便地应用该酶来发挥转糖基的作用,特别是对达玛稀二醇、原人参二醇和原人参三醇的转糖基作用。作为本发明的优选方式,还提供了二种形成稀有人参皂苷的方法,该方法之一包含:用本发明所述的gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7、3GT1和/或3GT2酶处理待转糖基的底物,所述的底物包括达玛稀二醇、原人参二醇和原人参三醇及其衍生物等四环三萜类化合物。较佳地,在pH3.5-10条件下,用所述的gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7以及3GT1、3GT2、3GT3、3GT4酶处理待转糖基的底物。较佳地,在温度30-105℃条件下,用所述的gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7以及3GT1、3GT2、3GT3、3GT4、gGT29-3、3GT1和/或3GT2酶处理待转糖基的底物。

[0325] 该方法之二包含:将本发明所述的gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7以及3GT1、3GT2、3GT3、3GT4基因转入可以合成达玛稀二醇、原人参二醇或原人参三醇的工程菌(例如,酵母或大肠杆菌工程

菌)中,或者,将gGT25、gGT13、gGT30、gGT25-1、gGT25-3、gGT25-5、gGT29、gGT29-3、gGT29-4、gGT29-5、gGT29-6、gGT29-7以及3GT1、3GT2、3GT3、3GT4基因与达玛烯二醇、原人参二醇和原人参三醇合成代谢途径中的关键基因于宿主细胞(例如酵母细胞或大肠杆菌)中共表达,获得直接生产稀有人参皂苷CK、Rh2、Rg3、Rh1或F1的重组菌。

[0326] 所述的达玛烯二醇合成代谢途径中的关键基因包括(但不限于):达玛烯二醇合成酶基因。

[0327] 在另一优选例中,所述的原人参二醇合成代谢途径中的关键基因包括(但不限于):达玛烯二醇合成酶基因、细胞色素P450CYP716A47基因、和P450CYP716A47的还原酶基因,或其组合。或者以上各种酶的同功酶及其组合。其中,达玛烯二醇合成酶将环氧角鲨烯(酿酒酵母自身合成)转化为达玛烯二醇,细胞色素P450CYP716A47及其还原酶再将达玛烯二醇转化为原人参二醇。(Han et al, plant & cell physiology, 2011, 52, 2062-73)

[0328] 在另一优选例中,所述的原人参三醇合成代谢途径中的关键基因包括(但不限于):达玛烯二醇合成酶基因、细胞色素P450CYP716A47基因、P450CYP716A47的还原酶基因、及细胞色素P450CYP716A53V2基因,或其组合。或者以上各种酶的同功酶及其组合。其中,达玛烯二醇合成酶将环氧角鲨烯(酿酒酵母自身合成)转化为达玛烯二醇,细胞色素P450CYP716A47及其还原酶再将达玛烯二醇转化为原人参二醇,细胞色素P450CYP716A53v2(JX036031)及P450CYP716A47的还原酶再进一步将原人参二醇转化为原人参三醇。(Han et al, plant & cell physiology, 2012, 53, 1535-45)

[0329] 本发明的主要优点:

[0330] (1) 本发明的糖基转移酶可以特异性和高效地将四环三萜化合物底物的C-20位和/或C-6位和/或C-3位羟基上转入葡萄糖;

[0331] (2) 本发明的糖基转移酶gGT29和gGT29-3可以将来自糖基供体的糖基转移到四环三萜类化合物的C-3位的第一个糖基上以延伸糖链

[0332] (3) 本发明的糖基转移酶特别能够分别将原人参二醇和原人参三醇转化为具有抗癌活性的稀有人参皂苷CK、Rh2或Rg3和具有抗衰老活性的稀有人参皂苷F1、具有抗过敏作用的稀有人参皂苷Rh1;

[0333] (4) 本发明的糖基转移酶还能将达玛烯二醇、原人参三醇、F1、25-OH-PPD、25-OCH<sub>3</sub>-PPD合成之前未见报道的新化合物20-O-β-(D-glucopyranosyl)-dammarendiolII、3-O-β-(D-glucopyranosyl)-dammarendiolII、3-O-β-(D-glucopyranosyl)-PPT、3-O-β-(D-glucopyranosyl)-F1、3-O-β-(D-glucopyranosyl)-25-OH-PPD、3-O-β-(D-glucopyranosyl)-25-OCH<sub>3</sub>-PPD、3-O-β-(D-glucopyranosyl)-lanosterol。

[0334] (5) 3GT1、3GT2、gGT29、gGT29-3和gGT25-5的催化活性并不受四环三萜化合物上20位羟基或者糖基空间构型的影响,既可以催化20(S)构型的人参皂苷(皂苷元)也可以催化20(R)构型的人参皂苷(皂苷元)。

[0335] (6) 在酵母中构建了人参皂苷元(达玛烯二醇、原人参二醇和原人参三醇)的合成途径,从而实现以葡萄糖等单糖为底物,用酵母来发酵生产新化合物20-O-β-(D-glucopyranosyl)-dammarendiolII、3-O-β-(D-glucopyranosyl)-dammarendiolII、3-O-β-(D-glucopyranosyl)-PPT、3-O-β-(D-glucopyranosyl)-F1、3-O-β-(D-glucopyranosyl)-lanosterol等以及稀有人参皂苷CK、F1、Rh1、Rh2和Rg3,这不但可以解决皂苷生产的原料来

源问题,而且可以大幅降低稀有皂苷CK、F1、Rh1、Rh2和Rg3的生产成本。

[0336] 下面结合具体实施例,进一步阐述本发明。应理解,这些实施例仅用于说明本发明而不适用于限制本发明的范围。下列实施例中未注明具体条件的实验方法,通常按照常规条件如Sambrook等人,分子克隆:实验室手册(New York:Cold Spring Harbor Laboratory Press,1989)中所述的条件,或按照制造厂商所建议的条件。

[0337] 实施例1

[0338] 糖基转移酶及其编码基因的分离

[0339] 从发表的人参属植物表达谱数据提取了100多条预测糖基转移酶的cDNA序列,从中克隆了60条cDNA全长序列并对它们进行了表达及转糖基反应分析,其中有11种表达产物对人参皂苷元及皂苷具有转糖基活性。

[0340] 提取人参RNA并进行反转录,获得人参的cDNA。以该cDNA为模板进行PCR扩增,使用其中引物对1(SEQ ID NOs.:7、8);引物对2(SEQ ID NOs.:9、10);引物对3(SEQ ID NOs.:11、12);引物对5(SEQ ID NOs.:34、35);引物对7(SEQ ID NOs.:46、47);引物对8(SEQ ID NOs.:62、63);引物对9(SEQ ID NOs.:64、65)全部获得扩增产物。DNA聚合酶选用宝生物工程有限公司的高保真的KOD DNA聚合酶。PCR产物经琼脂糖凝胶电泳检测(图1,19(c)和31)。在紫外下照射,切下目标DNA条带。然后采用Axygen Gel ExtractionKit(AEYGEN公司)从琼脂糖凝胶中回收DNA即为扩增出的DNA片段。将此DNA片段用宝生物工程有限公司的rTaq DNA聚合酶在末端加A后与市售的克隆载体pMD18-T Vector连接,连接产物转化市售的大肠杆菌EPI300感受态细胞,将转化后的大肠杆菌菌液涂布在添加氨苄青霉素50ug/mL、IPTG0.5mM、X-Gal25ug/mL的LB平板上,并进一步通过PCR和酶切验证重组克隆。分别选取其中一个克隆提取重组质粒后进行测序。用BESTORF软件寻找开放阅读框(ORF)。通过序列比对,ORF编码了糖基转移酶第1家族保守功能域,表明是糖基转移酶基因。

[0341] 用引物对1(SEQ ID NOs.:7、8)获得的基因具有SEQ ID NOs.:1、15、17和19所示的核苷酸序列,分别命名为gGT25,gGT25-1,gGT25-3和gGT25-5。自序列表中SEQ ID NO.:1的5'端第1至1425位核苷酸为gGT25的蛋白编码序列(CDS),自SEQ ID NO.:1的5'端的第1-3位核苷酸为gGT25基因的起始密码子ATG。序列表中SEQ ID NO.:15的5'端第1至1428位核苷酸为gGT25-1的开放阅读框(ORF),自SEQ ID NO.:15的5'端的第1-3位核苷酸为gGT25-1基因的起始密码子ATG,自SEQ ID NO.:15的5'端的第1426至1428位核苷酸为gGT25-1基因的终止密码子TAA。自序列表中SEQ ID NO.:17的5'端第1至1428位核苷酸为gGT25-3的开放阅读框(ORF),自SEQ ID NO.:17的5'端的第1-3位核苷酸为gGT25-3基因的起始密码子ATG,自SEQ ID NO.:17的5'端的第1426至1428位核苷酸为gGT25-3基因的终止密码子TAA。自序列表中SEQ ID NO.:19的5'端第1至1419位核苷酸为gGT25-5的开放阅读框(ORF),自SEQ ID NO.:19的5'端的第1-3位核苷酸为gGT25-5基因的起始密码子ATG,自SEQ ID NO.:19的5'端的第1426至1428位核苷酸为gGT25-5基因的终止密码子TAA。

[0342] 用引物对2(SEQ ID NOs.:9、10)获得的基因具有SEQ ID NO.:3所示的核苷酸序列,命名为gGT13。自序列表中SEQ ID NO.:3的5'端第1至1431位核苷酸为gGT13的开放阅读框(ORF),自SEQ ID NO.:3的5'端的第1-3位核苷酸为gGT13基因的起始密码子ATG,自SEQ ID NO.:1的5'端的第1429至1431位核苷酸为gGT13基因的终止密码子TAA。

[0343] 用引物对3(SEQ ID NOs.:11、12)获得的基因具有SEQ ID NO.:5所示的核苷酸序

列,命名为gGT30。自序列表中SEQ ID NO.:5的5'端第1至1353位核苷酸为gGT30的开放阅读框(ORF),自SEQ ID NO.:5的5'端的1-3位核苷酸为gGT30基因的起始密码子ATG,自SEQ ID NO.:5的5'端的第1351至1353位核苷酸为gG30基因的终止密码子TAA。

[0344] 用引物对5(SEQ ID NOs.:34,35)获得基因具有SEQ ID NOs.:25,27所示的核苷酸序列,分别命名为gGT29和gGT29-3。自序列表中SEQ ID NO.:25的5'端第1至1329位核苷酸为gGT29的开放阅读框(ORF),自SEQ ID NO.:25的5'端的1-3位核苷酸为gGT29基因的起始密码子ATG,自SEQ ID NO.:25的5'端的第1327至1329位核苷酸为gG29基因的终止密码子TAG。自SEQ ID NO.:27的5'端的1-3位核苷酸为gGT29-3基因的起始密码子ATG,自SEQ ID NO.:27的5'端的第1327至1329位核苷酸为gGT29-3基因的终止密码子TAG。

[0345] 用引物对6(SEQ ID NOs.:46,47)获得基因具有SEQ ID NO.:42所示的核苷酸序列,命名为3GT4。自序列表中SEQ ID NO.:42的5'端第1至1374位核苷酸为3GT4的开放阅读框(ORF),自SEQ ID NO.:42的5'端的1-3位核苷酸为3GT4基因的起始密码子ATG,自SEQ ID NO.:42的5'端的第1372至1374位核苷酸为3GT4基因的终止密码子TAG。

[0346] 用引物对7(SEQ ID NOs.:62,63)获得基因具有SEQ ID NO.54、56、58所示的核苷酸序列,命名为gGT29-4、gGT29-5和gGT29-6。自序列表中SEQ ID NO.:54的5'端第1至1341位核苷酸为gGT29-4的开放阅读框(ORF),自SEQ ID NO.:54的5'端的1-3位核苷酸为gGT29-4基因的起始密码子ATG,自SEQ ID NO.:54的5'端的第1339至1341位核苷酸为gGT29-4基因的终止密码子TAG。自序列表中SEQ ID NO.:56的5'端第1至1341位核苷酸为gGT29-5的开放阅读框(ORF),自SEQ ID NO.:56的5'端的1-3位核苷酸为gGT29-5基因的起始密码子ATG,自SEQ ID NO.:56的5'端的第1339至1341位核苷酸为gGT29-5基因的终止密码子TAG。自序列表中SEQ ID NO.:58的5'端第1至1341位核苷酸为gGT29-6的开放阅读框(ORF),自SEQ ID NO.:58的5'端的1-3位核苷酸为gGT29-6基因的起始密码子ATG,自SEQ ID NO.:58的5'端的第1339至1341位核苷酸为gGT29-6基因的终止密码子TAG。

[0347] 用引物对8(SEQ ID NOs.:64,65)获得基因具有SEQ ID NO.60所示的核苷酸序列,命名为gGT29-7。自序列表中SEQ ID NO.:60的5'端第1至1341位核苷酸为gGT29-7的开放阅读框(ORF),自SEQ ID NO.:60的5'端的1-3位核苷酸为gGT29-7基因的起始密码子ATG,自SEQ ID NO.:60的5'端的第1339至1341位核苷酸为gGT29-7基因的终止密码子TAG。人工合成具有SEQ ID NOs.:21、23和40所示的核苷酸序列,分别命名为3GT1、3GT2和3GT3。自序列表中SEQ ID NO.:21的5'端第1至1488位核苷酸为3GT1的开放阅读框(ORF),自SEQ ID NO.:21的5'端的第1-3位核苷酸为3GT1基因的起始密码子ATG,自SEQ ID NO.:21的5'端的第1486至1488位核苷酸为3GT1基因的终止密码子TAA。自序列表中SEQ ID NO.:23的5'端第1至1488位核苷酸为3GT2的开放阅读框(ORF),自SEQ ID NO.:23的5'端的第1-3位核苷酸为3GT2基因的起始密码子ATG,自SEQ ID NO.:23的5'端的第1486至1488位核苷酸为3GT2基因的终止密码子TAA。自序列表中SEQ ID NO.:40的5'端第1至1494位核苷酸为3GT3的开放阅读框(ORF),自SEQ ID NO.:40的5'端的第1-3位核苷酸为3GT3基因的起始密码子ATG,自SEQ ID NO.:40的5'端的第1492至1494位核苷酸为3GT3基因的终止密码子TAA。用引物对4(SEQ ID NOs.:29,30)对其中两条人工合成的基因(SEQ ID NO.:21和SEQ ID NO.:23)进行PCR扩增,获得的PCR产物具有SEQ ID NO.:21和SEQ ID NO.:23所示的核苷酸序列(图19(a));用引物对6(SEQ ID NOs.:44,45)对另外一条人工合成的基因(SEQ ID NO.:40)进行PCR扩增,

获得的PCR产物具有SEQ ID NO.:40所示的核苷酸序列(图19(b))。

[0348] 糖基转移酶基因gGT25编码一个含有475个氨基酸的蛋白质gGT25,具有序列表中SEQ ID NO.:2所示的的氨基酸序列。用软件预测该蛋白质的理论分子量大小为53kDa,等电点pI为5.14。自SEQ ID NO.:2的氨基端的第344-387位为糖基转移酶第1家族保守功能域。该糖基转移酶与人参转录组中预测的皂苷糖基转移酶基因的氨基酸序列一致性低于52%。

[0349] 糖基转移酶基因gGT25-1编码一个含有475个氨基酸的蛋白质gGT25-1,具有序列表中SEQ ID NO.:16所示的的氨基酸序列。用软件预测该蛋白质的理论分子量为53kDa,等电点pI为4.91。自SEQ ID NO.:16的氨基端的第344-387位为糖基转移酶第1家族保守功能域。该糖基转移酶与人参转录组中预测的皂苷糖基转移酶基因的氨基酸序列一致性低于52%。

[0350] 糖基转移酶基因gGT25-3编码一个含有475个氨基酸的蛋白质gGT25-3,具有序列表中SEQ ID NO.:18所示的的氨基酸序列。用软件预测该蛋白质的理论分子量为53kDa,等电点pI为5.05。自SEQ ID NO.:18的氨基端的第344-387位为糖基转移酶第1家族保守功能域。该糖基转移酶与人参转录组中预测的皂苷糖基转移酶基因的氨基酸序列一致性低于52%。

[0351] 糖基转移酶基因gGT25-5编码一个含有472个氨基酸的蛋白质gGT25-5,具有序列表中SEQ ID NO.:20所示的的氨基酸序列。用软件预测该蛋白质的理论分子量为53kDa,等电点pI为4.98。自SEQ ID NO.:20的氨基端的第343-386位为糖基转移酶第1家族保守功能域。该糖基转移酶与人参转录组中预测的皂苷糖基转移酶基因的氨基酸序列一致性低于52%。

[0352] 糖基转移酶基因gGT13编码一个含有476个氨基酸的蛋白质gGT13,具有序列表中序列SEQ ID NO.:4所示的的氨基酸序列。用软件预测该蛋白质的理论分子量为53kDa,等电点pI为4.91。自SEQ ID NO.:4的氨基端的第343-386位为糖基转移酶第1家族保守功能域。该糖基转移酶与人参转录组中预测的皂苷糖基转移酶基因的氨基酸序列一致性最高为99.5%。

[0353] 糖基转移酶基因gGT30编码一个含有451个氨基酸的蛋白质gGT30,具有序列表中SEQ ID NO.:6所示的的氨基酸序列。用软件预测该蛋白质的理论分子量为51kDa,等电点pI为6.79。自SEQ ID NO.:6的氨基端的第318-361位为糖基转移酶第1家族保守功能域。该糖基转移酶同酿酒葡萄(*Vitis vinifera*)的糖基转移酶(XP\_002271587)相似性最高(53%),表明该糖基转移酶为新酶。

[0354] 糖基转移酶基因3GT1编码一个含有495个氨基酸的蛋白质3GT1,具有序列表中SEQ ID NO.:22所示的的氨基酸序列。用软件预测该蛋白质的理论分子量为56kDa,等电点pI为5.52。自SEQ ID NO.:22的氨基端的第355-398位为糖基转移酶第1家族保守功能域。该糖基转移酶与欧洲山芥(*Barbarea vulgaris*)来源的糖基转移酶UGT73C10同源性>99%

[0355] 糖基转移酶基因3GT2编码一个含有495个氨基酸的蛋白质3GT2,具有序列表中SEQ ID NO.:24所示的的氨基酸序列。用软件预测该蛋白质的理论分子量为56kDa,等电点pI为5.62。自SEQ ID NO.:24的氨基端的第355-398位为糖基转移酶第1家族保守功能域。该糖基转移酶与欧洲山芥(*Barbarea vulgaris*)来源的糖基转移酶UGT73C12同源性>99%

[0356] 糖基转移酶基因gGT29编码一个含有442个氨基酸的蛋白质gGT29,具有序列表中

SEQ ID NO.:26所示的的氨基酸序列。用软件预测该蛋白质的理论分子量为49kDa,等电点pI为5.93。自SEQ ID NO.:26的氨基端的第317-360位为糖基转移酶第1家族保守功能域。该糖基转移酶与酿酒葡萄 (*Vitis vinifera*) 来源的糖基转移酶的序列相似性低于56%。

[0357] 糖基转移酶基因gGT29-3编码一个含有442个氨基酸的蛋白质gGT29-3,具有序列表中SEQ ID NO.:28所示的的氨基酸序列。用软件预测该蛋白质的理论分子量为49kDa,等电点pI为5.48。自SEQ ID NO.:28的氨基端的第317-360位为糖基转移酶第1家族保守功能域。该糖基转移酶与酿酒葡萄 (*Vitis vinifera*) 来源的糖基转移酶的序列相似性低于56%。

[0358] 糖基转移酶基因3GT3编码一个含有个497氨基酸的蛋白质3GT3,具有序列表中SEQ ID NO.:41所示的的氨基酸序列。用软件预测该蛋白质的理论分子量为55kDa,等电点pI为5.50。自SEQ ID NO.:41的氨基端的第350-393位为糖基转移酶第1家族保守功能域。该糖基转移酶与蒺藜苜蓿 (*Medicago truncatula*) 来源的糖基转移酶同源性>99%。

[0359] 糖基转移酶基因3GT4编码一个含有458个氨基酸的蛋白质3GT4,具有序列表中SEQ ID NO.:43所示的的氨基酸序列。用软件预测该蛋白质的理论分子量为51kDa,等电点pI为5.10。自SEQ ID NO.:43的氨基端的第333-376位为糖基转移酶第1家族保守功能域。该糖基转移酶与酿酒葡萄 (*Vitis vinifera*) 来源的糖基转移酶的序列同源性低于50%。

[0360] 糖基转移酶基因gGT29-4编码一个含有446个氨基酸的蛋白质gGT29-4,具有序列表中SEQ ID NO.:55所示的的氨基酸序列。用软件预测该蛋白质的理论分子量为50kDa,等电点pI为5.78。自SEQ ID NO.:55的氨基端的第321-364位为糖基转移酶第1家族保守功能域。该糖基转移酶与北柴胡 (*Bupleurum chinense*) 来源的糖基转移酶的序列相似性低于57%。

[0361] 糖基转移酶基因gGT29-5编码一个含有446个氨基酸的蛋白质gGT29-5,具有序列表中SEQ ID NO.:57所示的的氨基酸序列。用软件预测该蛋白质的理论分子量为50kDa,等电点pI为5.93。自SEQ ID NO.:57的氨基端的第321-364位为糖基转移酶第1家族保守功能域。该糖基转移酶与北柴胡 (*Bupleurum chinense*) 来源的糖基转移酶的序列相似性低于58%。

[0362] 糖基转移酶基因gGT29-6编码一个含有446个氨基酸的蛋白质gGT29-6,具有序列表中SEQ ID NO.:59所示的的氨基酸序列。用软件预测该蛋白质的理论分子量为50kDa,等电点pI为6.03。自SEQ ID NO.:59的氨基端的第321-364位为糖基转移酶第1家族保守功能域。该糖基转移酶与北柴胡 (*Bupleurum chinense*) 来源的糖基转移酶的序列相似性低于59%。

[0363] 糖基转移酶基因gGT29-7编码一个含有446个氨基酸的蛋白质gGT29-7,具有序列表中SEQ ID NO.:61所示的的氨基酸序列。用软件预测该蛋白质的理论分子量为50kDa,等电点pI为5.80。自SEQ ID NO.:61的氨基端的第321-364位为糖基转移酶第1家族保守功能域。该糖基转移酶与北柴胡 (*Bupleurum chinense*) 来源的糖基转移酶的序列相似性低于57%。

[0364] 表2

[0365]

糖基转移酶	C-3	C-3 第一个糖基	C6	C20
gGT25 (SEQ ID NO. : 2)			✓	✓
gGT25-1 (SEQ ID NO. : 16)			✓	✓
gGT25-3 (SEQ ID NO. : 18)			✓	✓
gGT25-5 (SEQ ID NO. : 20)			✓	
gGT29 (SEQ ID NO. : 26)		✓		
gGT29-3 (SEQ ID NO. : 28)		✓		
gGT29-4 (SEQ ID NO. : 55)		✓		
gGT29-5 (SEQ ID NO. : 57)		✓		
gGT29-6 (SEQ ID NO. : 59)		✓		
gGT29-7 (SEQ ID NO. : 61)		✓		
3GT1 (SEQ ID NO. : 22)	✓			
3GT2 (SEQ ID NO. : 24)	✓			
3GT3 (SEQ ID NO. : 40)	✓			
3GT4 (SEQ ID NO. : 41)	✓			
gGT13 (SEQ ID NO. : 4)	ND	ND	ND	ND
gGT30 (SEQ ID NO. : 6)	ND	ND	ND	ND

[0366] 实施例2

[0367] 糖基转移酶基因gGT25, gGT25-1, gGT25-3和gGT25-5的酵母重组表达载体的构建

[0368] 以实施例1构建的含有gGT25, gGT25-1, gGT25-3和gGT25-5基因的质粒gGT25-pMD18T, gGT25-1-pMD18T, gGT25-3-pMD18T和gGT25-5-pMD18T为模板扩增目标基因。

[0369] 所用正向引物均为:

[0370] 5'-GCCGGAGCTCATGAAGTCAGAATTGATATTC-3' (SEQ ID NO.:13), 其5'端添加SacI识别位点:GAGCTC;

[0371] 所用反向引物均为:

[0372] 5'-GCCGCTCGAGTTAATGATGATGATGATGATGCATAATTTCTCAAATAGCTTC-3' (SEQ ID NO.:14), 其5'端添加XhoI识别位点:CTCGAG, 反向引物引入6×His Tag以便进行Western Blot检测表达和纯化。

[0373] 利用上述引物和模板通过PCR方法扩增gGT25, gGT25-1, gGT25-3和gGT25-5基因。DNA聚合酶选用Toyobo公司的高保真DNA聚合酶kod, 参考其说明书设定PCR程序:94℃2min; 94℃15s, 58℃30s, 68℃1.5min, 共30个循环;68℃10min;10℃保温。PCR产物经琼脂糖凝胶电泳检测, 在紫外光下, 切下与目标DNA大小一致的条带。然后采用AXYGEN公司的AxyPrep DNA Gel Extraction Kit从琼脂糖凝胶中回收DNA片段。用Takara公司的QuickCut限制性内切酶Kpn I和Xba I双酶切回收的DNA片段30min, 用AXYGEN公司的AxyPrepPCR Cleanup Kit对酶切产物进行清洁回收。利用NEB公司的T4DNA连接酶将酶切产物与酿酒酵母表达质粒pYES2 (同样以Kpn I和Xba I酶切并割胶回收) 25℃连接2h。连接产物转化E.coli TOP10感受态细胞, 并涂布于添加100μg/mL氨苄青霉素的LB平板上。通过菌落PCR验证阳性转化

子,并测序进一步验证,结果表明表达质粒gt25-pYES2,gt25-1-pYES2,gt25-3-pYES2和gt25-5-pYES2构建成功。

[0374] 实施例3

[0375] 糖基转移酶基因gGT25,gGT25-1,gGT25-3和gGT25-5在酿酒酵母中的表达

[0376] 通过电转化方法将构建好的表达质粒gt25-pYES2转化到酿酒酵母(*Saccharomyces cerevisiae*)中,涂布于筛选平板SC-Ura(0.67%酵母无氨基酸基本氮源,2%葡萄糖)。通过菌落PCR验证酵母重组子。挑酵母重组子菌落于10mL SC-Ura(2%葡萄糖)培养基中,30℃200rpm培养20h。4℃3500g离心收集菌体,用无菌去离子水清洗菌体两次,用诱导培养基SC-Ura(2%半乳糖)重悬菌体,并接种到50mL诱导培养基中,使OD<sub>600</sub>在0.4左右,30℃200rpm开始诱导表达。4℃3500g离心收集诱导表达12h的菌体,用无菌去离子水清洗菌体两次,重悬于酵母裂解缓冲液中,使OD<sub>600</sub>在50-100之间。用Fastprep细胞破碎仪震荡破碎酵母细胞,4℃12000g离心10min去除细胞碎片,收集细胞裂解液上清。取适量裂解液上清进行SDS-PAGE电泳检测,与pYES2空载体的重组子相比,gt25-pYES2,gt25-1-pYES2,gt25-3-pYES2,gt25-5-pYES2重组子没有明显的条带特征,如图2。采用anti-6×His Tag Western Blot检测表达情况,如图3所示,表达gGT25,gGT25-1,gGT25-3和gGT25-5的酿酒酵母重组子显示很强的Western Blot信号,表明gGT25,gGT25-1,gGT25-3和gGT25-5在酵母中可溶表达,而转pYES2空载体的重组子则没有anti-6×His Tag Western Blot信号。

[0377] 实施例4

[0378] 酵母表达产物gGT25,gGT25-1,gGT25-3和gGT25-5转糖基反应和产物鉴定

[0379] 以表达gGT25,gGT25-1,gGT25-3和gGT25-5的重组酵母裂解上清做为酶液来催化底物原人参二醇(Protopanoxadiol PPD),原人参三醇(Protopanaxatriol PPT)和达玛烯二醇(Dammarenediol II,DM)的转糖基反应,表达空载体的重组酵母裂解上清为对照。100μL反应体系如表3:

[0380] 表3

[0381]

9%吐温20	11.1μL
50mM UDP-葡萄糖	10μL
1M Tris-HCl pH8.5	5μL
100mM底物(乙醇溶解)	0.5μL
酶液	73.4μL

[0382] 反应在35℃下进行12h,然后加入100μL丁醇终止反应,并抽提产物。产物真空干燥后,用甲醇溶解。

[0383] 反应产物先用薄层层析(TLC)进行检测,表达gGT25或gGT25-1或gGT25-3的重组酵母裂解上清酶液可以在原人参二醇和原人参三醇的20位羟基糖基化,分别转化为稀有人参皂苷CK和F1(图6和图7);3位羟基已糖基化的原人参二醇型皂苷(Rh2和Rg3),可以在gGT25,gGT25-1和gGT25-3的催化下,将20位羟基继续糖基化,分别生成F2和Rd(图6);gGT25,gGT25-1和gGT25-3不但可以将PPT的20位羟基糖基化生成F1,还能进一步在6位羟基继续糖基化形成Rg1(图7);gGT25,gGT25-1和gGT25-3还能将PPD的前体物达玛烯二醇的20位羟基糖基化,得到一种未见报道的皂苷20-O-β-(D-glucopyranosyl)-dammarendiolIII(图8)。但

是,6位羟基已经糖基化的原人参三醇型皂苷(Rh1,Rg2和Rf)则不能在gGT25,gGT25-1和gGT25-3的催化下将20位羟基糖基化。同时,gGT25,gGT25-1和gGT25-3也不能催化糖链的延伸。gGT25-5与gGT25,gGT25-1和gGT25-3的催化活性存在差别,它不能像gGT25,gGT25-1和gGT25-3一样,在PPD,PPT或者达玛烯二醇的20位羟基糖基化,它只能将PPT的6位羟基糖基化,转化为稀有人参皂苷Rh1(图7)。

[0384] 进一步对gGT25的转化产物用HPLC进行鉴定(图10和图11)。在图10中,有三个峰出现,其中峰2与标样样品中CK的保留时间一致,峰3与原人参二醇的峰一致。峰3已经很小,这说明原人参二醇基本都被转换为了CK。峰1,在负对照的图谱中也出现,所以应该与原人参二醇的转化无关。在图11中,也有三个峰出现,峰1与标样样品中F1的保留时间一致,峰3与原人参三醇的峰一致。峰3已经很小,这说明原人参三醇基本都被转换为了F1。峰2,在负对照的图谱中也出现,所以应该与原人参三醇的转化无关。

[0385] 最后用LC/MS对产物做更进一步的鉴定(图12和图13)。图12是对原人参二醇转化产物中CK峰(图10中峰2)的质谱图,它与标准CK样品的质谱图完全一致。图13是对原人参三醇转化产物中F1峰(图11中峰1)的质谱图,它与标准样品F1的质谱图完全一致。这些结果进一步证实了原人参二醇和三醇的gGT25转化产物分别为CK和F1。

#### [0386] 实施例5

[0387] 糖基转移酶基因gGT13和gGT30的克隆、表达及其表达产物的转糖基反应

[0388] 用与实施例2同样的方法,获得了gGT13和gGT30的克隆,构建了它们的酵母重组表达载体并转入酿酒酵母。按照实施例3相同的步骤诱导表达糖基转移酶的表达,虽然在SDS-PAGE胶上没有明显的目的蛋白的条带(图4),但是Western Blot检测到明显的杂交信息,这表明gGT13和gGT30均已在酵母中表达(图5)。

[0389] 用与实施例4同样的方法,利用表达gGT13和gGT30的重组酵母细胞裂解液分别催化原人参二醇(PPD)和原人参三醇(PPT)。

[0390] 结果发现,gGT13和gGT30的蛋白表达产物都无法转化PPD或PPT(图9);并且gGT13和gGT30也都无法转化原人参二醇型皂苷Rh2、CK、F2和Rg3以及原人参三醇型F1、Rh1和Rg1。

[0391] 以上结果表明,尽管gGT13与人参转录组中预测的人参皂苷糖基转移酶氨基酸序列的一致性很高(99.5%),但是gGT13和gGT30对上述的底物都没有转糖基作用。

#### [0392] 实施例6

[0393] 糖基转移酶基因gGT25在大肠杆菌中的表达及其表达产物的转糖基反应

[0394] 以实施例1构建的含有gGT25基因的质粒gGT25-pMD18T为模板扩增目标基因gGT25,并将其克隆到大肠杆菌表达载体pet28a(购自Merck公司)中,构建大肠杆菌表达载体gt25-pet28a,转化到市售的E.coli BL21中。接种一个重组子到LB培养基中,30℃200rpm培养至OD<sub>600</sub>约0.6-0.8,使菌液降温至4℃,加入终浓度为50μM的IPTG,18℃200rpm诱导表达15h。4℃离心收集菌体,超声破碎细胞,4℃12000g离心收集细胞裂解液上清,取样品进行SDS-PAGE电泳。

[0395] Western Blot(图14)表明,在50μM IPTG诱导条件下,糖基转移酶gGT25在大肠杆菌中也可以表达。以该重组大肠杆菌的细胞裂解液上清为粗酶液来进行转糖基反应,反应的条件与实施例4中相同。

[0396] 反应在35℃下进行12h,然后加入100μL丁醇终止反应,并抽提产物。产物真空干燥

后,用甲醇溶解。反应产物先用薄层层析(TLC)进行检测,图15中可见gGT25粗酶液可将PPD转化为CK。

[0397] 实施例7

[0398] 产CK酵母工程菌的构建与产物鉴定

[0399] 在pESC-HIS质粒((Stratagene,Agilent)上,同时组装达玛烯二醇合成酶(Dammarenediol synthase)(ACZ71036.1)(GAL1/GAL10GAL10侧启动子,ADH1终止子)、细胞色素P450CYP716A47(AEY75213.1)(FBA1启动子,CYC1终止子)以及糖基转移酶GT25(GAL1/GAL10GAL1侧启动子,TDH2终止子),构成游离型质粒,转化酿酒酵母BY4742,并将拟南芥来源的细胞色素P450还原酶ATR2-1(NP\_849472.2)整合于酿酒酵母BY4742染色体中染色体trp1基因位点(GAL1启动子,利用trp1原有终止子),构建了重组酵母A。同样方法,构建了重组酵母B,区别在于将拟南芥来源还原酶ATR2-1整合于含有达玛烯二醇合成酶、细胞色素P450CYP716A47及糖基转移酶GT25的重组质粒上,ATR2-1有启动子与终止子分别TEF2启动子和TPI1终止子,其他3个基因的启动子与终止子与重组菌A的对应基因相同。

[0400] 用与重组酵母菌B相同的方法构建了重组酵母菌C,但重新规划了各个基因的启动子与终止子,如表4。

[0401] 表4主要酶启动子与终止子构成:

[0402]

主要酶	启动子	终止子
达玛烯二醇合成酶	GAL1/GAL10GAL10侧	ADH1
CYP716A47	GAL1/GAL10GAL1侧	TDH2
ATR2-1	TEF2	TPI1
GT25	FBA1	CYC1

[0403] 重组酵母菌株A,B,C在SC-Ura(0.67%酵母无氨基酸基本氮源,2%半乳糖)培养基中发酵,各重组菌需要额外添加的氨基酸或尿嘧啶见表5,取50mL重组酵母的发酵液,离心后沉淀的菌体用5mL酵母裂解缓冲液(50mM Tris-HCl,1mM EDTA,1mM PMSF,5%甘油,pH7.5)重悬后,用Fastprep震荡裂解酵母,设置功率6M/S,震荡7~8次使酵母充分裂解。将裂解液转移到2mL EP管中,每个管装1mL,加入等体积(1mL)的正丁醇抽提约30min后12000g离心10min。吸取上清至一新的EP管中。45℃并真空条件下使正丁醇蒸干。用100μL甲醇溶解后用于HPLC检测。

[0404] 通过HPLC分析,重组酵母A的细胞裂解液中含有达玛烯二醇、原人参二醇(PPD)及人参皂苷活性代谢物CK(图16),酵母A合成CK的产量达到0.6mg/L。HPLC分析同样发现重组酵母B和C的细胞裂解液含有微量CK。

[0405] 表5重组酵母菌需补加的相应氨基酸或尿嘧啶

[0406]

重组酵母菌株	补加氨基酸或尿嘧啶
A	0.01%色氨酸(tryptophan),亮氨酸(leucine),赖氨酸(lysine)
B	0.01%尿嘧啶(uracil),亮氨酸(leucine),赖氨酸(lysine)
C	0.01%尿嘧啶(uracil),亮氨酸(leucine),赖氨酸(lysine)

[0407] 实施例8

[0408] 产Rh1酵母工程菌的构建与产物鉴定

[0409] 在pESC-HIS质粒((Stratagene,Agilent)上,同时组装达玛烯二醇合成酶(Dammareniol synthase)(ACZ71036.1)(GAL1/GAL10GAL10侧启动子,ADH1终止子)、细胞色素P450CYP716A47(AEY75213.1)(FBA1启动子,CYC1终止子)、细胞色素P450CYP716A53V2基因(EN02启动子,CYC1终止子)以及糖基转移酶gGT25-5(GAL1/GAL10GAL1侧启动子,TDH2终止子),构成游离型质粒,转化酿酒酵母BY4742,并将拟南芥来源的细胞色素P450还原酶ATR2-1(NP\_849472.2)整合于酿酒酵母BY4742染色体中染色体trp1基因位点(GAL1启动子,利用trp1原有终止子),构建了重组酵母A3。重组酵母菌需补加的相应氨基酸或尿嘧啶见表5。

[0410] 将重组酵母A3裂解液转移到2mL EP管中,每个管装1mL,加入等体积(1mL)的正丁醇抽提约30min后12000g离心10min。吸取上清至一新的EP管中。45℃并真空条件下使正丁醇蒸干。用100μL甲醇溶解后用于HPLC检测。

[0411] 通过HPLC分析,重组酵母A3的细胞裂解液中含有原人参三醇(PPT)及人参皂苷活性代谢物Rh1(图41)。

[0412] 实施例9

[0413] 糖基转移酶基因3GT1,3GT2,3GT3和3GT4的大肠杆菌重组表达载体的构建

[0414] 以实施例1构建的含有3GT1和3GT2基因的质粒3GT1-pMD18T,3GT2-pMD18T,为模板扩增目标基因。

[0415] 3GT1和3GT2所用正向引物均为SEQ ID NO.:31,其5'端添加BamH I识别位点:GGATCC;3GT1所用反向引物为SEQ ID NO.:32,其5'端添加Sal I识别位点:CTCGAG;3GT2所用反向引物为SEQ ID NO.:33,其5'端添加Sal I识别位点:CTCGAG。

[0416] 利用上述引物和模板通过PCR方法扩增3GT1和3GT2基因。DNA聚合酶选用Toyobo公司的高保真DNA聚合酶KOD,参考其说明书设定PCR程序:94℃2min;94℃15s,58℃30s,68℃1.5min,共35个循环;68℃10min;10℃保温。PCR产物经琼脂糖凝胶电泳检测,在紫外光下,切下与目标DNA大小一致的条带。然后采用AXYGEN公司的AxyPrep DNA Gel Extraction Kit从琼脂糖凝胶中回收DNA片段。用Takara公司的QuickCut限制性内切酶Kpn I和Xba I双酶切回收的DNA片段30min,用AXYGEN公司的AxyPrep PCR Cleanup Kit对酶切产物进行清洁回收。利用NEB公司的T4DNA连接酶将酶切产物与大肠杆菌表达质粒pET28a(同样以BamH I和Sal I酶切并割胶回收)16℃连接4h。连接产物转化E.coli EPI300感受态细胞,并涂布于添加50μg/mL卡那霉素的LB平板上。通过菌落PCR验证阳性转化子,并测序进一步验证表达质粒3GT1-pET28a和3GT2-pET28a构建成功。

[0417] 以实施例1构建的含有3GT3和3GT4基因的质粒3GT3-pMD18T,3GT4-pMD18T,为模板扩增目标基因。

[0418] 3GT3所用的正向引物如SEQ ID NO.:48所示,其5'端添加了与载体pET28a同源的序列:ACTTTAAGAAGGAGATATACC;3GT3所用反向引物为如SEQ ID NO.:49所示,其5'端添加了与载体pET28a同源的序列:CTCGAGTGC GGCCGCAAGCTT。

[0419] 3GT4所用正向引物为SEQ ID NO.:50,其5'端添加了与载体pET28a同源的序列:ACTTTAAGAAGGAGATATACC;3GT4所用反向引物为SEQ ID NO.:51,其5'端添加了与载体pET28a同源的18个碱基片段:CTCGAGTGC GGCCGCAAGCTT。

[0420] 利用上述引物通过PCR方法扩增3GT3和3GT4的基因。扩增基因选用NEB公司的Q5高保真DNA聚合酶,参考其说明书设定PCR程序:98℃30s;98℃15s,58℃30s,72℃1min,共35个循环;72℃2min;10℃保温。

[0421] 同时,使用SEQ ID NO.:52和SEQ ID NO.:53分别作为正向和反向引物扩增载体pET28a,获得线性化的载体pET28a。扩增pET28a线性化载体也选用NEB公司的Q5高保真DNA聚合酶,参考其说明书设定PCR程序:98℃30s;98℃15s,58℃30s,72℃3min,共35个循环;72℃2min;10℃保温。

[0422] 上述3GT3和3GT4基因PCR产物以及线性化的载体pET28a,经琼脂糖凝胶电泳检测后,在紫外光下切下与目标DNA大小一致的条带。然后采用AXYGEN公司的AxyPrep DNA Gel Extraction Kit从琼脂糖凝胶中回收DNA片段。参考诺晶生物科技有限公司的BGclonart无缝克隆试剂盒说明书,将回收的线性化的pET28a载体片段,回收的3GT3或者3GT4基因片段及诺晶生物科技有限公司的BGclonart无缝克隆反应液以适当比例进行混合,共20 $\mu$ l。混匀后在50℃孵育30分钟,然后将混合反应液转移到冰上。使用5 $\mu$ l反应液转化E.coli EPI300感受态细胞,并涂布于添加50 $\mu$ g/mL卡那霉素的LB平板上。通过菌落PCR验证阳性转化子,并测序进一步验证表达质粒3GT3-pET28a和3GT4-pET28a构建成功。

[0423] 实施例10

[0424] 糖基转移酶基因3GT1,3GT2,3GT3和3GT4在大肠杆菌中的表达

[0425] 以实施例9构建的大肠杆菌表达载体3GT1-pET28a,3GT2-pET28a,3GT3-pET28a和3GT4-pET28a,转化到市售的E.coli BL21中。接种一个重组子到LB培养基中,30℃200rpm培养至OD<sub>600</sub>约0.6-0.8,使菌液降温至4℃,加入终浓度为50 $\mu$ M的IPTG,18℃200rpm诱导表达15h。4℃离心收集菌体,超声破碎细胞,4℃12000g离心收集细胞裂解液上清,取样品进行SDS-PAGE电泳(图20)。与pet28a空载体的重组子相比,3GT1-pet28a,3GT2-pet28a,3GT3-pet28a和3GT4-pet28a重组子有明显的条带(大约55KD)表征3GT1,3GT2,3GT3和3GT4。从Western Blot的结果看(图21),也证明目标蛋白3GT1,3GT2,3GT3和3GT4在宿主中实现了可溶表达。

[0426] 实施例11

[0427] 大肠杆菌表达产物3GT1,3GT2,3GT3和3GT4转糖基反应和产物的鉴定

[0428] 以表达3GT1,3GT2,3GT3和3GT4的重组大肠杆菌裂解上清为粗酶液来催化人参皂苷和皂苷元的转糖基反应,表达空载体的重组大肠杆菌裂解上清为对照。100 $\mu$ L反应体系如表3所示。反应在35℃下进行12h,然后加入100 $\mu$ L丁醇终止反应,并抽提产物。产物真空干燥后,用甲醇溶解。

[0429] 反应产物先用薄层层析(TLC)进行检测(图22-28),用3GT1,3GT2,3GT3和3GT4的粗酶液可以将原人参二醇(PPD)的C3位羟基糖基化,分别转化为稀有人参皂苷Rh2(图22,27(a)和28(a));20位羟基己糖基化的原人参二醇型皂苷(CK),可以在3GT1,3GT2和3GT4粗酶液的催化下,将3位羟基继续糖基化,分别生成F2(图22和28(b));糖基转移酶3GT1和3GT2还可以分别将达玛烯二醇DM的C3位羟基糖基化,转化为新化合物3-O- $\beta$ -(D-glucopyranosyl)-Dammareniol II(图23);糖基转移酶3GT1,3GT2,3GT3和3GT4可以将25-OH-PPD的C3位羟基糖基化,转化为新化合物3-O- $\beta$ -(D-glucopyranosyl)-25-OH-PPD。(图23,图27(c)和图28(c));3GT1,3GT2和3GT3还可以将原人参三醇(PPT)的C3位羟基糖基

化,转化为之前没有报道过的新皂苷3-O- $\beta$ -(D-glucopyranosyl)-PPT(图24和27(b));3GT1和3GT2还可以将F1的C3位羟基糖基化,转化为之前没有报道过的新皂苷3-O- $\beta$ -(D-glucopyranosyl)-F1(图24);3GT1和3GT2还可以将羊毛甾醇(lanosterol)的C3位羟基糖基化转化为新化合物3-O- $\beta$ -(D-glucopyranosyl)-lanosterol(图26)。同时,3GT1和3GT2的催化活性并不受20位羟基或者糖基空间构型的影响,例如既可以催化20(S)-PPD,也可以催化20(R)-PPD生成稀有人参皂苷20(R)-人参皂苷Rh2(图25)。虽然3GT1,3GT2,3GT3和3GT4四种糖基转移酶都可以在四环三萜皂苷元的3位加入糖基,但是它们可以催化的底物谱有很大的差别。如表6所示,3GT1和3GT2可以催化的底物最多,3GT3可催化的底物最少,但是3GT4的专一性最好,它只能催化原人参二醇及原人参二醇型(CK)皂苷。

[0430] 进一步用HPLC对3GT1,3GT3和3GT4催化PPD的产物进行检测(图29)。在图29中,糖基转移酶3GT1,3GT3和3GT4催化PPD的产物中都出现了保留时间相同的峰(P1,P2和P3),它们与标样中人参皂苷Rh2的保留时间一致,说明糖基转移酶3GT1,3GT3和3GT4催化PPD生成了人参皂苷Rh2。最后,用LC/MS对图29中P1,P2和P3三个样品峰作了质谱鉴定(图30),它们与标准样品人参皂苷Rh2的质谱完全一致,这进一步说明了糖基转移酶3GT1,3GT3和3GT4催化PPD生成的产物为Rh2。

[0431] 糖基转移酶3GT1,3GT2,3GT3和3GT4可催化底物的比较如表6所示:

[0432] 表6

[0433]

	PPD	(20R)-PPD	CK	PPT	F1	DM	25-OH-PPD	羊毛甾醇
3GT1/3GT2	√	√	√	√	√	√	√	√
3GT3	√	×	×	√	×	×	√	×
3GT4	√	×	√	×	×	×	√	×

[0434] 实施例12

[0435] 产Rh2酵母工程菌的构建与产物鉴定

[0436] 12.1在pESC-HIS质粒((Stratagene,Agilent)上,同时组装达玛烯二醇合成酶(Dammarenediol synthase)(ACZ71036.1)(GAL1/GAL10GAL10侧启动子,ADH1终止子)、细胞色素P450CYP716A47(AEY75213.1)(FBA1启动子,CYC1终止子)以及糖基转移酶3GT4(GAL1/GAL10GAL1侧启动子,TDH2终止子),构成游离型质粒,转化酿酒酵母BY4742,并将拟南芥来源的细胞色素P450还原酶ATR2-1(NP\_849472.2)整合于酿酒酵母BY4742染色体中染色体trp1基因位点(GAL1启动子,利用trp1原有终止子),构建了重组酵母A1。重组酵母菌需补加的相应氨基酸或尿嘧啶见表5。

[0437] 将重组酵母A1裂解液转移到2mL EP管中,每个管装1mL,加入等体积(1mL)的正丁醇抽提约30min后12000g离心10min。吸取上清至一新的EP管中。45℃并真空条件下使正丁醇蒸干。用100 $\mu$ L甲醇溶解后用于HPLC检测。

[0438] 通过HPLC分析(图39),重组酵母A1的细胞裂解液中含有达玛烯二醇、原人参二醇(PPD)及人参皂苷活性代谢物Rh2。

[0439] 12.2方法同12.1,区别在于用糖基转移酶3GT1代替3GT4,得到重组酵母A5。

[0440] 结果见图43,通过HPLC分析,重组酵母A5的细胞裂解液中含有达玛烯二醇、原人参二醇(PPD)及人参皂苷活性代谢物Rh2。

[0441] 实施例13

[0442] 糖基转移酶基因gGT29和gGT29-3的酵母重组表达载体的构建

[0443] 分别以实施例1构建的含有gGT29和gGT29-3基因的质粒gGT29-pMD18T和gGT29-3-pMD18T为模板扩增目标基因。

[0444] gGT29所用正向引物均为(SEQ ID NO.:36),其5'端添加Kpn I识别位点:GGATCC;所用反向引物均为(SEQ ID NO.:37),其5'端添加XhoI识别位点:CTCGAG,反向引物引入6×HisTag以便进行Western Blot检测表达和纯化。

[0445] gGT29-3所用正向引物均为(SEQ ID NO.:38),其5'端添加Kpn I识别位点:GGATCC;所用反向引物均为(SEQ ID NO.:39),其5'端添加XhoI识别位点:CTCGAG,反向引物引入6×His Tag以便进行Western Blot检测表达和纯化。

[0446] 以质粒gGT29-pMD18T和gGT29-3-pMD18T为模板,利用上述引物通过PCR方法扩增gGT29和gGT29-3的基因。DNA聚合酶选用Toyobo公司的高保真DNA聚合酶kod,参考其说明书设定PCR程序:94℃2min;94℃15s,58℃30s,68℃1.5min,共30个循环;68℃10min;10℃保温。PCR产物经琼脂糖凝胶电泳检测,在紫外光下,切下与目标DNA大小一致的条带。然后采用AXYGEN公司的AxyPrep DNA Gel Extraction Kit从琼脂糖凝胶中回收DNA片段。用Takara公司的QuickCut限制性内切酶Kpn I和Xba I双酶切回收的DNA片段30min,用AXYGEN公司的AxyPrep PCR Cleanup Kit对酶切产物进行清洁回收。利用NEB公司的T4DNA连接酶将酶切产物与酿酒酵母表达质粒pYES2(同样以Kpn I和Xba I酶切并割胶回收)25℃连接2h。连接产物转化E.coli TOP10感受态细胞,并涂布于添加100μg/mL氨苄青霉素的LB平板上。通过菌落PCR验证阳性转化子,并测序进一步验证表达质粒gGT29-pYES2和gGT29-3-pYES2构建成功。

[0447] 实施例14

[0448] 糖基转移酶基因gGT29和gGT29-3在酿酒酵母中的表达

[0449] 通过电转化方法将构建好的表达质粒gGT29-pYES2和gGT29-3-pYES2转化到酿酒酵母(*Saccharomyces cerevisiae*)中,涂布于筛选平板SC-Ura(0.67%酵母无氨基酸基本氮源,2%葡萄糖)。通过菌落PCR验证酵母重组子。挑酵母重组子菌落于10mL SC-Ura(2%葡萄糖)培养基中,30℃200rpm培养20h。4℃3500g离心收集菌体,用无菌去离子水清洗菌体两次,用诱导培养基SC-Ura(2%半乳糖)重悬菌体,并接种到50mL诱导培养基中,使OD<sub>600</sub>在0.4左右,30℃200rpm开始诱导表达。4℃3500g离心收集诱导表达12h的菌体,用无菌去离子水清洗菌体两次,重悬于酵母裂解缓冲液中,使OD<sub>600</sub>在50-100之间。用Fastprep细胞破碎仪震荡破碎酵母细胞,4℃12000g离心10min去除细胞碎片,收集细胞裂解液上清。取适量裂解液上清进行SDS-PAGE电泳检测,与pYES2空载体的重组子相比,gGT29-pYES2和gGT29-3-pYES2重组子没有明显的条带表征(图32)。采用anti-6×His Tag Western Blot检测表达情况,表达gGT29和gGT29-3的酿酒酵母显示很强的Western Blot信号,表明gGT29和gGT29-3在酵母中均可溶表达,而转pYES2空载体的重组子则没有anti-6×His Tag Western Blot信号(图33)。

[0450] 实施例15

[0451] 酵母表达产物gGT29和gGT29-3转糖基反应和产物鉴定

[0452] 以表达gGT29和gGT29-3的重组酵母裂解上清做为酶液来催化人参皂苷Rh2和F2的

转糖基反应,表达空载体的重组酵母裂解上清为对照。100 $\mu$ L反应体系如表3所示。反应在35 $^{\circ}$ C下进行12h,然后加入100 $\mu$ L丁醇终止反应,并抽提产物。产物真空干燥后,用甲醇溶解。

[0453] 反应产物先用薄层层析(TLC)进行检测,表达gGT29和gGT29-3的酵母宿主裂解上清酶液可以在人参皂苷Rh2和F2的3位糖基再延伸一个糖基,转化为人参皂苷Rg3和Rd(图34)。gGT29和gGT29-3的催化活性不受人参皂苷20位糖基或者羟基构型的影响,可以将20(R)-Rh2转化为20(R)-Rg3(图36)。

[0454] 实施例16

[0455] 糖基转移酶3GT1/3GT4和gGT29的联合转糖基反应和产物鉴定

[0456] 以表达3GT1或3GT4的大肠杆菌宿主裂解上清和以表达gGT29的酵母宿主裂解上清做为酶液来共同催化原人参二醇(PPD)。100 $\mu$ L反应体系如表3所示。73.4 $\mu$ L的酶液中40 $\mu$ L为3GT1的大肠宿主裂解上清,剩余33.4 $\mu$ L为表达gGT29的酵母宿主裂解上清。反应在35 $^{\circ}$ C下进行12h,然后加入100 $\mu$ L丁醇终止反应,并抽提产物。产物真空干燥后,用甲醇溶解。反应产物先用薄层层析(TLC)进行检测(图35),可以看到糖基转移酶3GT1和gGT29或者3GT4和gGT29联合使用都可将PPD转化为Rg3。

[0457] 糖基转移酶3GT1和gGT29或者3GT2和gGT29联合使用都可以催化20(R)-PPD,生成20(R)-Rg3(图36)。

[0458] 实施例17

[0459] 产Rg3酵母工程菌的构建与产物鉴定

[0460] 17.1在pESC-HIS质粒((Stratagene,Agilent)上,同时组装达玛烯二醇合成酶(Dammarenediol synthase)(ACZ71036.1)(GAL1/GAL10GAL10侧启动子,ADH1终止子)、细胞色素P450CYP716A47(AEY75213.1)(FBA1启动子,CYC1终止子)以及糖基转移酶3GT4和gGT29(GAL1/GAL10GAL1侧启动子,TDH2终止子),构成游离型质粒,转化酿酒酵母BY4742,并将拟南芥来源的细胞色素P450还原酶ATR2-1(NP\_849472.2)整合于酿酒酵母BY4742染色体中染色体trp1基因位点(GAL1启动子,利用trp1原有终止子),构建了重组酵母A2。重组酵母菌需补加的相应氨基酸或尿嘧啶见表5。

[0461] 将重组酵母A2裂解液转移到2mL EP管中,每个管装1mL,加入等体积(1mL)的正丁醇抽提约30min后12000g离心10min。吸取上清至一新的EP管中。45 $^{\circ}$ C并真空条件下使正丁醇蒸干。用100 $\mu$ L甲醇溶解后用于HPLC检测。

[0462] 通过HPLC分析,重组酵母A2的细胞裂解液中含有达玛烯二醇、原人参二醇(PPD)和人参皂苷活性代谢物Rg3(图40)。

[0463] 17.2方法同17.1,区别在于用糖基转移酶3GT1代替3GT4,得到重组酵母A6。通过HPLC分析,重组酵母A6的细胞裂解液中也含有达玛烯二醇、原人参二醇(PPD)及人参皂苷活性代谢物Rg3。

[0464] 实施例18

[0465] 产F1酵母工程菌的构建与产物鉴定

[0466] 在pESC-HIS质粒((Stratagene,Agilent)上,同时组装达玛烯二醇合成酶(Dammarenediol synthase)(ACZ71036.1)(GAL1/GAL10GAL10侧启动子,ADH1终止子)、糖基转移酶gGT25(GAL1/GAL10GAL1侧启动子,TDH2终止子),细胞色素P450CYP716A47(AEY75213.1)(FBA1启动子,FBA1终止子),细胞色素P450CYP716A53V2(EN02启动子,CYC1终

止子)。构成游离型质粒,转化酿酒酵母BY4742,并将拟南芥来源的细胞色素P450还原酶ATR2-1(NP\_849472.2)整合于酿酒酵母BY4742染色体中染色体trp1基因位点(GAL1启动子,利用trp1原有终止子),构建了重组酵母A4。重组酵母菌需补加的相应氨基酸或尿嘧啶见表5。

[0467] 将重组酵母A4裂解液转移到2mL EP管中,每个管装1mL,加入等体积(1mL)的正丁醇抽提约30min后12000g离心10min。吸取上清至一新的EP管中。45℃并真空条件下使正丁醇蒸干。用100μL甲醇溶解后用于HPLC检测。

[0468] 通过HPLC分析,重组酵母A4的细胞裂解液中含有原人参三醇(PPT)和人参皂苷活性代谢物F1(图42)。

[0469] 实施例19

[0470] 糖基转移酶基因gGT29-4, gGT29-5, gGT29-6和gGT29-7的大肠杆菌重组表达载体的构建

[0471] 以实施例1构建的含有gGT29-4, gGT29-5, gGT29-6和gGT29-7基因的质粒gGT29-4-pMD18T, gGT29-5-pMD18T, gGT29-6-pMD18T和gGT29-7-pMD18T为模板扩增目标基因。

[0472] gGT29-5和gGT29-6基因所用的正向引物如SEQ ID NO.:66所示,其5'端添加了与载体pET28a同源的序列:CTGGTGCCGCGCGGCAGC;所用反向引物为如SEQ ID NO.:68所示,其5'端添加了与载体pET28a同源的序列:TGCGGCCGCAAGCTTGTC。

[0473] gGT29-4和gGT29-7基因所用正向引物为SEQ ID NO.:67,其5'端添加了与载体pET28a同源的序列:CTGGTGCCGCGCGGCAGC;所用反向引物为SEQ ID NO.:68,其5'端添加了与载体pET28a同源的18个碱基片段:TGCGGCCGCAAGCTTGTC。

[0474] 利用上述引物通过PCR方法扩增gGT29-4, gGT29-5, gGT29-6和gGT29-7基因。扩增基因选用NEB公司的Q5高保真DNA聚合酶,参考其说明书设定PCR程序:98℃30s;98℃15s,58℃30s,72℃1min,共35个循环;72℃2min;10℃保温。

[0475] 同时,使用SEQ ID NO.:69和SEQ ID NO.:70分别作为正向和反向引物扩增载体pET28a,获得线性化的载体pET28a。扩增pET28a线性化载体也选用NEB公司的Q5高保真DNA聚合酶,参考其说明书设定PCR程序:98℃30s;98℃15s,58℃30s,72℃3min,共35个循环;72℃2min;10℃保温。

[0476] 上述gGT29-4, gGT29-5, gGT29-6和gGT29-7基因PCR产物以及线性化的载体pET28a,经琼脂糖凝胶电泳检测后,在紫外光下切下与目标DNA大小一致的条带。然后采用AXYGEN公司的AxyPrep DNA Gel Extraction Kit从琼脂糖凝胶中回收DNA片段。参考诺晶生物科技有限公司的BGclonart无缝克隆试剂盒说明书,将回收的线性化的pET28a载体片段,回收的gGT29-4, gGT29-5, gGT29-6和gGT29-7基因片段及诺晶生物科技有限公司的BGclonart无缝克隆反应液以适当比例进行混合,共20μl。混匀后在50℃孵育30分钟,然后将混合反应液转移到冰上。使用5μl反应液转化E.coli EPI300感受态细胞,并涂布于添加50μg/mL卡那霉素的LB平板上。通过菌落PCR验证阳性转化子,并测序进一步验证表达质粒gGT29-4-pET28a, gGT29-5-pET28a, gGT29-6-pET28a和gGT29-7-pET28a构建成功。

[0477] 实施例20

[0478] 糖基转移酶基因gGT29-4, gGT29-5, gGT29-6和gGT29-7在大肠杆菌中的表达

[0479] 以实施例19构建的大肠杆菌表达载体gGT29-4-pET28a, gGT29-5-pET28a, gGT29-

6-pET28a和gGT29-7-pET28a转化到市售的E.coli BL21中。接种一个重组子到LB培养基中, 30℃ 200rpm培养至OD<sub>600</sub>约0.6-0.8, 使菌液降温至4℃, 加入终浓度为50μM的IPTG, 18℃ 200rpm诱导表达15h。4℃离心收集菌体, 超声破碎细胞, 4℃12000g离心收集细胞裂解液上清, 取样品进行SDS-PAGE电泳(图44)。gGT29-4-pET28a, gGT29-5-pET28a, gGT29-6-pET28a和gGT29-7-pET28a重组子裂解液和总蛋白和上清中都有明显的目的蛋白条带(大约50KD), 分别表征糖基转移酶gGT29-4, gGT29-5, gGT29-6和gGT29-7。从Western Blot的结果看(图45), 也证明目标蛋白gGT29-4, gGT29-5, gGT29-6和gGT29-7在宿主中实现了可溶表达。

[0480] 实施例21

[0481] 大肠杆菌表达产物gGT29-4, gGT29-5, gGT29-6和gGT29-7转糖基反应和产物的鉴定

[0482] 以表达gGT29-4, gGT29-5, gGT29-6和gGT29-7的重组酵母裂解上清做为酶液来催化人参皂苷Rh2和F2的转糖基反应。100μL反应体系如表3所示。反应在35℃下进行12h, 然后加入100μL丁醇终止反应, 并抽提产物。产物真空干燥后, 用甲醇溶解。

[0483] 反应产物用薄层层析(TLC)进行检测, gGT29-6的粗酶液可以在人参皂苷Rh2和F2的3位糖基上再延伸一个糖基, 分别生成为人参皂苷Rg3和Rd(图46); gGT29-4, gGT29-5和gGT29-7的粗酶液可以在人参皂苷F2的3位糖基再延伸一个糖基生成皂苷Rd, 但是它们不能催化皂苷Rh2(图46)。

[0484] 在本发明提及的所有文献都在本申请中引用作为参考, 就如同每一篇文献被单独引用作为参考那样。此外应理解, 在阅读了本发明的上述讲授内容之后, 本领域技术人员可以对本发明作各种改动或修改, 这些等价形式同样落于本申请所附权利要求书所限定的范围。

序列表

<110> 中国科学院上海生命科学研究院  
 <120> 一维糖基转移酶及其应用  
 <130> P2013-1222  
 <160> 70  
 <170> PatentIn version 3.5

<210> 1  
 <211> 1425  
 <212> DNA  
 <213> 人参(Panax ginseng)

<220>  
 <221> misc feature  
 <223> gGT25  
 <400> 1

```

atgaagtcag aattgatatt ettgcecgcc cgggeccatcg gacacctcgt gggaaatggtg      60
gagatggceta aactcttcat cagtcgacat gaaaacctcf cggtcaccgt cctcatcgcg      120
aaattcttaca tggatagcggg ggtagacaac tacaataaat cactcttaac aaaccctacc      180
cecggtctca caattgtaaa tctcccggaa accgaccccc aaaactatat gctcaaacca      240
cgccatgcca tctttcctag cgtcatcgag actcagaaga cacacgtgag agacataata      300
tcaggcatga ctcagtcgga gtcgaactgg gtcgttggit tgcctggcga ccttttgctc      360
atcaacatta tggacattgc caatgagttc aatgttccaa cttatgtata ctcccctgcc      420
ggagccggtc atcttggcct cgcgttccat ctccagacac tcaacgacaa aaagcaagat      480
gtgaccggagt tcaggaaact ggacaactgag ttattggtag cgagttttgc aaaccgggtt      540
cccgccggagg tcttggcgtc gatgtatgtg gataaagaag gtgggtatga ttatttgttt      600
tcattgttcc ggagtlgcag agagtc aaag gcaatlatta ttaacagtl ttaggagctg      660
gaacctatg cgataaatte cctccggatg gatagtatga tccctccgat ctaccgggtg      720
ggaccctaac taaatctcaa egggtatggc caaaaactcgc atgaggctgc tgtgatcctt      780
ggttggttag atgatcaacc accttcatct gtggtgtttt tgtgclttgg tagctatgga      840
agctttcaag aaaaccaggt gaaggagatt gcaatgggtc tagagcgcag tggccatcgc      900
ttcttgggtt ccttggctcc gtcctatcct aaaggcgaga caaagcttea gcttaaatc      960
tcaaatlga aagaattctt cccagttaga ttcltggaca ggacatcatg cgtcggaaaa      1020
gtgattggat gggcccgcga agtggccgtg ctgggacatg agtcagtcgg agggttccctg      1080
tctcattgag gttggaattc gacattggag agtgtttggt gtgggggtgc cgttgcaaca      1140
tggccaatgt atggf gagea acaactcaat gcttttgaga tggtttaagga gtttaggtatt      1200
gcggtgga aa ttgaggtgga ctataagaaa gattatttta acatgaagaa tgattttatt      1260
gttaggcgag aagaatega gacaaaaata aagaagtgtg tgatggatga aaataatagt      1320
gaaataagaa agaaggtaaa ggaatgaaa gaaaagagta ggctgcaat gtctgagaat      1380
ggatcatctt ataattcatt ggccaagcta tttgaggaaa ttatg      1425

```

[0001]

<210> 2  
 <211> 475  
 <212> PRT  
 <213> 人参(Panax ginseng)

<220>  
 <221> misc feature  
 <223> gGT25  
 <400> 2

```

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

```

Ile Asn Ser Leu Arg Met Asp Ser Met Ile Pro Pro Ile Tyr Pro Val  
 225 230 235 240  
 Gly Pro Ile Leu Asn Leu Asn Gly Asp Gly Gln Asn Ser Asp Glu Ala  
 245 250 255  
 Ala Val Ile Leu Gly Trp Leu Asp Asp Gln Pro Pro Ser Ser Val Val  
 260 265 270  
 Phe Leu Cys Phe Gly Ser Tyr Gly Ser Phe Gln Glu Asn Gln Val Lys  
 275 280 285  
 Glu Ile Ala Met Gly Leu Glu Arg Ser Gly His Arg Phe Leu Trp Ser  
 290 295 300  
 Leu Arg Pro Ser Ile Pro Lys Gly Glu Thr Lys Leu Gln Leu Lys Tyr  
 305 310 315 320  
 Ser Asn Leu Lys Glu Ile Leu Pro Val Gly Phe Leu Asp Arg Thr Ser  
 325 330 335  
 Cys Val Gly Lys Val Ile Gly Trp Ala Pro Gln Val Ala Val Leu Gly  
 340 345 350  
 His Glu Ser Val Gly Gly Phe Leu Ser His Cys Gly Trp Asn Ser Thr  
 355 360 365  
 Leu Glu Ser Val Trp Cys Gly Val Pro Val Ala Thr Trp Pro Met Tyr  
 370 375 380  
 Gly Glu Gln Gln Leu Asn Ala Phe Glu Met Val Lys Glu Leu Gly Ile  
 385 390 395 400  
 Ala Val Glu Ile Glu Val Asp Tyr Lys Lys Asp Tyr Phe Asn Met Lys  
 405 410 415  
 Asn Asp Phe Ile Val Arg Ala Glu Glu Ile Glu Thr Lys Ile Lys Lys  
 420 425 430  
 Leu Met Met Asp Glu Asn Asn Ser Glu Ile Arg Lys Lys Val Lys Glu  
 435 440 445 460  
 Met Lys Glu Lys Ser Arg Ala Met Ser Glu Asn Gly Ser Ser Tyr  
 450 455 465  
 Asn Ser Leu Ala Lys Leu Phe Glu Glu Ile Met  
 470 475

<210> 3  
 <211> 1431  
 <212> DNA  
 <213> 人参(Panax ginseng)

<220>  
 <221> misc feature  
 <223> gGT13  
 <400> 3

[0002]

```

atgggagcgg agctcatctt aatccctct cccgggagtt gtcactcgtt atctactgtt      60
gagattgcaa agctcctcat cagtcagat gaaccgcttt ceateacagt ccttgtaatg      120
aagttctcgc atgacactgg tctgactgcc tacactcggg eattgcagaa agatgctctt      180
aacccgatag tctttgtgga ccttctcag aaccgagtc tttatctcctt gcccagttct      240
ttctttacta gcttcatega gactengac agtcgggta gagattctgt cagacaaatt      300
gtgagtcggg ctgattctaa taagctcgtt ggcttcgtea tgcacatggt etgcacccca      360
atgatagacg tggcaaatga atttggagtc ccaacctatg tttcttcac ttcaggtgct      420
gcatttcttg gcttcagtt ttaccatctg agtctcagtg atgaacataa ccaggacctt      480
accgagtata aggacacgga tttttagtta tctatcccga gtttcatcaa cccagtgcct      540
gctaaagttt tgccttcggt gattctgaac aaggaaggat cgacctgctt ccaatctatt      600
tcccgaaggt ttaaagaagc caaggccatt ctagtcaaca cgttcgggga gctggaacca      660
catgccatta aggccttgg tgataactgc aagatcccct ctatctatcc cgtgggacct      720
ataatcaacc tcaagaacaa ggagggaaca acccaaaacc atagttctga agatgggtatc      780
attagctggt tggacaatca gccacctct tcggtagtgt tttgtgctt tgggagcttt      840
gggagctttg atgaaggcca agtcaggag atagcaaacg gattggagca gagtggacag      900
cgattcttgt ggtctctac cggcggccca gaaaaaatgg aattgcctaa agactatgag      960
aatcccgaag aagtgttgc agaaggatc atagaacgaa catcagggat ggggaaggtg      1020
atcggatggg cgcacaaac ggcgattctt tcccacctg ctgtggagg attcgtgtct      1080
catgtggat ggaattctac attggagagt atatggtgt gggttccaat ggtacttgg      1140
ectatataatg cagagcagca aatcaatcgg tttgagttgg tgaaggagt gggaatggct      1200
gtggagatca aatggatta cagagaagat tataattttg cactgaaaa taatttagtt      1260
gtgaccgag accagataga gaaaggaatg cgatgtctga tgatggatgg agagagtga      1320
atgaggaaga aggtggaaga gatgaaagag aagagcagaa tggccatggt gaaggcggg      1380
tctcttaca tttcactcgg gcattttatt gaggatgtca tgcgtaatta a      1431

```

<210> 4  
 <211> 476  
 <212> PRT  
 <213> 人参(Panax ginseng)

<220>  
 <221> misc feature  
 <223> gGT13  
 <400> 4

Met Gly Ala Glu Ile Leu Ile Pro Ser Pro Gly Val Gly His Leu  
 1 5 10 15  
 Val Ser Thr Val Glu Ile Ala Lys Leu Leu Ile Ser Arg Asp Glu Arg  
 20 25 30

Leu Ser Ile Thr Val Leu Val Met Lys Phe Ser His Asp Thr Gly Val  
 35 40 45  
 Thr Ala Tyr Thr Arg Ser Leu Gln Lys Asp Ala Pro Asn Arg Ile Val  
 50 55 60  
 Phe Val Asp Leu Pro Gln Asn Glu Ser Leu Ile Ser Ser Pro Lys Ser  
 65 70 75 80  
 Phe Phe Thr Ser Phe Ile Glu Ser Gln Thr Ser Pro Val Arg Asp Ser  
 85 90 95  
 Val Arg Gln Ile Val Ser Arg Ser Asp Ser Asn Lys Leu Ala Gly Phe  
 100 105 110  
 Val Ile Asp Met Phe Cys Thr Pro Met Ile Asp Val Ala Asn Glu Phe  
 115 120 125  
 Gly Val Pro Thr Tyr Val Phe Phe Thr Ser Gly Ala Ala Phe Leu Gly  
 130 135 140  
 Leu Gln Phe Tyr His Leu Ser Leu Ser Asp Glu His Asn Gln Asp Leu  
 145 150 155 160  
 Thr Glu Tyr Lys Asp Thr Asp Val Glu Leu Ser Ile Pro Ser Phe Ile  
 165 170 175  
 Asn Pro Val Pro Ala Lys Val Leu Pro Ser Val Ile Leu Asn Lys Glu  
 180 185 190  
 Gly Ser Thr Met Leu Gln Ser Ile Ser Arg Arg Phe Lys Glu Ala Lys  
 195 200 205  
 Ala Ile Leu Val Asn Thr Phe Ala Glu Leu Glu Pro His Ala Ile Lys  
 210 215 220  
 Ala Leu Gly Asp Asn Cys Lys Ile Pro Pro Ile Tyr Pro Val Gly Pro  
 225 230 235 240  
 Ile Ile Asn Leu Lys Asn Lys Glu Gly Thr Thr Gln Asn His Ser Ser  
 245 250 255  
 Glu Asp Gly Ile Ile Ser Trp Leu Asp Asn Gln Pro Pro Ser Ser Val  
 260 265 270  
 Val Phe Leu Cys Phe Gly Ser Phe Gly Ser Phe Asp Glu Gly Gln Val  
 275 280 285  
 Arg Glu Ile Ala Asn Gly Leu Glu Gln Ser Gly Gln Arg Phe Leu Trp  
 290 295 300  
 Ser Leu Arg Arg Arg Pro Glu Lys Met Glu Leu Pro Lys Asp Tyr Glu  
 305 310 315 320  
 Asn Pro Glu Glu Val Leu Pro Glu Gly Phe Ile Glu Arg Thr Ser Gly  
 325 330 335  
 Met Gly Lys Val Ile Gly Trp Ala Pro Gln Thr Ala Ile Leu Ser His  
 340 345 350  
 Pro Ala Val Gly Gly Phe Val Ser His Cys Gly Trp Asn Ser Thr Leu  
 355 360 365  
 Glu Ser Ile Trp Cys Gly Val Pro Met Ala Thr Trp Pro Ile Tyr Ala  
 370 375 380  
 Glu Gln Gln Ile Asn Ala Phe Glu Leu Val Lys Glu Leu Gly Met Ala  
 385 390 395 400  
 Val Glu Ile Lys Met Asp Tyr Arg Glu Asp Tyr Ile Phe Ala Pro Glu  
 405 410 415  
 Asn Asn Leu Val Val Thr Ala Asp Gln Ile Glu Lys Gly Met Arg Cys  
 420 425 430  
 Leu Met Met Asp Gly Glu Ser Glu Met Arg Lys Lys Val Glu Glu Met  
 435 440 445  
 Lys Glu Lys Ser Arg Met Ala Met Val Lys Gly Gly Ser Ser Tyr Ile  
 450 455 460  
 Ser Leu Gly His Phe Ile Glu Asp Val Met Arg Asn  
 465 470 475

[0003]

<210> 5  
 <211> 1353  
 <212> DNA  
 <213> 人参(Panax ginseng)  
 <220>  
 <221> misc feature  
 <223> gGT30  
 <400> 5  
 atggcaacc aaaaatgttt cegtgtctc gtgttcccat gtttagctca tggtcacatc 60  
 tccccttcc tagaactagc cagaaaaact atagaaaaag gaaacttttt catttacttc 120  
 tgetccactc ctataaatct catctccata aagaaaaaac ttagtggtga tgateatcaa 180  
 aattacacca agtcaataca actagtggaa cacaatctac cgaccttacc teaacttctc 240  
 cctcaactacc aaaccaccga eggcctccca ccaaatctca atcctaccct tggaaaaggca 300  
 tttgaaatgt caaaactate ctcccacaac accetaaaca ctttgaace agatctacte 360  
 atttfgtatg acttatttca atggccagaa atagtagctt catcacatgg tgttccggtt 420  
 gttcggttcc aaacatgcag cgtgacagcc gctagttttc tagctcatat ttttaagaac 480  
 ccagacggtta cataaccttt tccagccatt tatcttcag agtatgaaac tgatcagatc 540  
 aggcgttgcg ttgacggcgt ttttgaage ggcagagaag aatctegcaa cctgttggtt 600  
 gteaacnct ctaaagcgat egaggaaaaa tafttttatt attattctct actteggggt 660  
 aacacaaaa ttatgccagt tggctcgcct attcageaac ccccaaatgg cgacgaggat 720  
 atgaaggtea tcgaatggct tgacaagaag gatccgtgtt caaccgtgtt cgtgtccttt 780  
 gggagcgagt attttatgca aaaagaagag gttgaagaga tgctcatgg ttttagagctt 840  
 agcaacgtea atttcatctg ggtttttagg gctccggtgg gagcggaana ggttaagctg 900  
 ccgttagggt ttgttgagag ggttgggggg aggggaattg ttatggaggg gtgggcccc 960

```

caggcaagga ttttgggaca ttcaagtatt ggtgggtttg tgagtcattg tgggtggaat 1020
tctgtgttgg aaactafaaa ttttgggttt ccaataatag ggaigccaat gaaatttgag 1080
cagcctatga atgctagget tctatctgaa ctttgggttt gtgttgaaat tgtgggggac 1140
gaaactagga ggtttggaag agaagaggta ggaatgtaa taagagaagt ggttgggtggg 1200
aagatagggg acgatttgag aaggaaagtg aaagaacttg gagcaaatg aaaggaaaaa 1260
caagaggaag agatggatga tbtgcttgat gaattagtac aaatttgtaa caagaaaaaa 1320
eggactgttg tacatcatca tcatcatcat tga. 1353
    
```

```

<210> 6
<211> 450
<212> PRT
<213> 人参(Panax ginseng)
    
```

```

<220>
<221> misc feature
<223> gGT30
<400> 6
    
```

[0004]

```

Met Ala Thr Gln Lys Cys Phe Arg Val Leu Val Phe Pro Trp Leu Ala
1 5 10 15
His Gly His Ile Ser Pro Phe Leu Glu Leu Ala Arg Lys Leu Ile Glu
20 25 30
Lys Gly Asn Phe Phe Ile Tyr Phe Cys Ser Thr Pro Ile Asn Leu Ile
35 40 45
Ser Ile Lys Lys Lys Leu Ser Gly Asp Asp His Gln Asn Tyr Thr Lys
50 55 60
Ser Ile Gln Leu Val Glu His Asn Leu Pro Thr Leu Pro Gln Leu Pro
65 70 75 80
Pro His Tyr His Thr Thr Asp Gly Leu Pro Pro Asn Leu Asn Pro Thr
85 90 95
Leu Arg Lys Ala Phe Glu Met Ser Lys Leu Ser Phe Pro Asn Thr Leu
100 105 110
Asn Thr Leu Lys Pro Asp Leu Leu Ile Cys Asp Asp Leu Phe Gln Trp
115 120 125
Pro Glu Ile Val Ala Ser Ser His Gly Val Pro Val Val Arg Phe Gln
130 135 140
Thr Cys Ser Val Thr Ala Ala Ser Phe Leu Ala His Ile Phe Thr Asn
145 150 155 160
Pro Asp Val Thr Tyr Pro Phe Pro Ala Ile Tyr Leu His Glu Tyr Glu
165 170 175
Thr Asp Gln Ile Arg Arg Cys Val Asp Ala Val Phe Glu Ser Gly Arg
180 185 190
Glu Glu Ser Arg Asn Leu Leu Val Val Asn Thr Ser Lys Ala Ile Glu
195 200 205
Glu Lys Tyr Phe Tyr Tyr Tyr Ser Leu Leu Arg Gly Asn Thr Lys Ile
210 215 220
Met Pro Val Gly Pro Leu Ile Gln Gln Ala Pro Asn Gly Asp Glu Asp
225 230 235 240
Met Lys Val Ile Glu Trp Leu Asp Lys Lys Asp Pro Cys Ser Thr Val
245 250 255
Phe Val Ser Phe Gly Ser Glu Tyr Phe Met Gln Lys Glu Glu Val Glu
260 265 270
Glu Met Ala His Gly Leu Glu Leu Ser Asn Val Asn Phe Ile Trp Val
275 280 285
Phe Arg Ala Pro Val Gly Ala Glu Lys Val Lys Leu Pro Leu Gly Phe
290 295 300
Val Glu Arg Val Gly Gly Arg Gly Ile Val Met Glu Gly Trp Ala Pro
305 310 315 320
Gln Ala Arg Ile Leu Gly His Ser Ser Ile Gly Gly Phe Val Ser His
325 330 335
Cys Gly Trp Asn Ser Val Leu Glu Thr Ile Asn Phe Gly Val Pro Ile
340 345 350
Ile Gly Met Pro Met Lys Phe Glu Gln Pro Met Asn Ala Arg Leu Leu
355 360 365
Ser Glu Leu Gly Val Cys Val Glu Ile Val Gly Asp Glu Thr Arg Arg
370 375 380
Phe Gly Arg Glu Glu Val Gly Asn Val Ile Arg Glu Val Val Gly Gly
385 390 395 400
Lys Ile Gly Asp Asp Leu Arg Arg Lys Val Lys Glu Leu Gly Ala Lys
405 410 415
Ile Lys Glu Lys Gln Glu Glu Glu Met Asp Asp Val Leu Asp Glu Leu
420 425 430
Val Gln Ile Cys Asn Lys Lys Lys Arg Thr Val Val His His His His
435 440 445
His His
450
    
```

```

<210> 7
<211> 26
<212> DNA
<213> 人参(Panax ginseng)
<400> 7
    
```

	ccaagttcat tcaagatgaa gtcaga	26
	<210> 8	
	<211> 28	
	<212> DNA	
	<213> 人工序列	
	<400> 8	
	cataatttcc tcaaataget tegccaat	28
	<210> 9	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<400> 9	
	actggagcag ttgcaacata eaaa	24
	<210> 10	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<400> 10	
	tcattttatgc gettttctgg atgt	24
	<210> 11	
	<211> 21	
	<212> DNA	
	<213> 人工序列	
	<400> 11	
	acgaggggtgg ttttaatggc a	21
	<210> 12	
	<211> 27	
	<212> DNA	
	<213> 人工序列	
	<400> 12	
	acaccattca taateataca acagtcc	27
[0005]	<210> 13	
	<211> 29	
	<212> DNA	
	<213> 人工序列	
	<400> 13	
	atatggtacc atggcagtag ccggegetg	29
	<210> 14	
	<211> 57	
	<212> DNA	
	<213> 人工序列	
	<400> 14	
	gcgetctaga ttaatgatga tgaatgatgat gttttattat atgttttttg gggtcgc	57
	<210> 15	
	<211> 1428	
	<212> DNA	
	<213> 人参(Panax ginseng)	
	<220>	
	<221> misc_feature	
	<223> gGT25-1	
	<400> 15	
	atgaagtcag aattgatatt ctlgeccgcc ccgcccateg gacacctcgt gggaaatggtg	60
	gagatggeta aactttcat eagtegacat gaaaacctct cggteaccgt cctcctcggc	120
	aaattctaca tggatcggg ggtagacaac tacaataaat cactcttaac aaacctace	180
	ccgcgtctca caattgtaaa tctcccggaa accgaccccc aaaactatat gctcaaacca	240
	cgccatgoca ttttcttag cgtcctcgag actcagaaga cacacgtgcg agacataata	300
	tcaggcatga ctcagtcaga gtgcactcgg gtctgttggt tgcctgctga ccttttcttc	360
	atcaaacatta tggacattgc caatgagtic aatgttccaa cttatgtata ctccccgcc	420
	ggagccggtc atcttggect cgcgttccat ctccagacac tcaacgacaa aaageaagat	480
	gtgaccgagt tcaggaaactc ggcactgag ttattggtae cgagttttgc aaacctgggt	540
	ccgcccaggg tcttgccgtc gatgtatgtg gataaagaag gtgggtatga ttatttgttt	600
	teattgttcc ggaggtgcag agagtcaaaag gcaattatta ttaacacggt tgaggagctg	660
	gaacctatg cgatcaattc cctccggatg gatagtatga tccctccgat ctacctgggt	720
	ggacctatc taatctcaa cggatgagc caaaactcgc atgaggetgc tgtgatcctt	780
	ggttggttag atgatcaacc acctteatct gtggtgtttt tgtgcttttg tagctatgga	840
	acctttcaag aaaaccaggt gaaggagatt gcaatgggtc tagagegcag tgggcatcgc	900

```

ttcttgtggt ccttgcgtcc gtctatccct aaagcggaga caaagcttca gcttaaatac 960
tcaaatttgg aagaaattct cccagtcgga ttcttggaca ggacatcatg cgtcggaaaa 1020
gttattggat gggccccgea agtggcggtg ctccggacacg aggeagtcgg agggttctctg 1080
tctcattgtg gttggaattc gacattagag agtgtgtggt gtggcgtgcc cgtcggcaaca 1140
tgccaatgt aeggcgagca acaactcaat gcttttgaga tggtaagga gttaggattt 1200
gcggtggaaa ttgaggtgga ctataagaat gaattttta acatgaataa tgattttatt 1260
gttagggcag aagaaatcga gacgaaaata aagaagtga tgatggatga aaagaatagt 1320
gnaataaggn agaaggtaaa ggaantgaaa gaaaagagta ggcttgcnaat gtcctgagaat 1380
ggatcatctt ataattcatt ggcgaageta tttgaggaaa tcatgtaa 1428
    
```

<210> 16  
 <211> 475  
 <212> PRT  
 <213> 人参(Panax ginseng)

<220>  
 <221> misc feature  
 <223> gGT25-1  
 <400> 16

[0006]

```

Met Lys Ser Glu Leu Ile Phe Leu Pro Ala Pro Ala Ile Gly His Leu
1 5 10 15
Val Gly Met Val Glu Met Ala Lys Leu Phe Ile Ser Arg His Glu Asn
20 25 30
Leu Ser Val Thr Val Leu Ile Ala Lys Phe Tyr Met Asp Thr Gly Val
35 40 45
Asp Asn Tyr Asn Lys Ser Leu Leu Thr Asn Pro Thr Pro Arg Leu Thr
50 55 60
Ile Val Asn Leu Pro Glu Thr Asp Pro Gln Asn Tyr Met Leu Lys Pro
65 70 75 80
Arg His Ala Ile Phe Pro Ser Val Ile Glu Thr Gln Lys Thr His Val
85 90 95
Arg Asp Ile Ile Ser Gly Met Thr Gln Ser Glu Ser Thr Arg Val Val
100 105 110
Gly Leu Leu Ala Asp Leu Leu Phe Ile Asn Ile Met Asp Ile Ala Asn
115 120 125
Glu Phe Asn Val Pro Thr Tyr Val Tyr Ser Pro Ala Gly Ala Gly His
130 135 140
Leu Gly Leu Ala Phe His Leu Gln Thr Leu Asn Asp Lys Lys Gln Asp
145 150 155 160
Val Thr Glu Phe Arg Asn Ser Asp Thr Glu Leu Leu Val Pro Ser Phe
165 170 175
Ala Asn Pro Val Pro Ala Glu Val Leu Pro Ser Met Tyr Val Asp Lys
180 185 190
Glu Gly Gly Tyr Asp Tyr Leu Phe Ser Leu Phe Arg Arg Cys Arg Glu
195 200 205
Ser Lys Ala Ile Ile Ile Asn Thr Phe Glu Glu Leu Glu Pro Tyr Ala
210 215 220
Ile Asn Ser Leu Arg Met Asp Ser Met Ile Pro Pro Ile Tyr Pro Val
225 230 235 240
Gly Pro Ile Leu Asn Leu Asn Gly Asp Gly Gln Asn Ser Asp Glu Ala
245 250 255
Ala Val Ile Leu Gly Trp Leu Asp Asp Gln Pro Pro Ser Ser Val Val
260 265 270
Phe Leu Cys Phe Gly Ser Tyr Gly Thr Phe Gln Glu Asn Gln Val Lys
275 280 285
Glu Ile Ala Met Gly Leu Glu Arg Ser Gly His Arg Phe Leu Trp Ser
290 295 300
Leu Arg Pro Ser Ile Pro Lys Gly Glu Thr Lys Leu Gln Leu Lys Tyr
305 310 315 320
Ser Asn Leu Glu Glu Ile Leu Pro Val Gly Phe Leu Asp Arg Thr Ser
325 330 335
Cys Val Gly Lys Val Ile Gly Trp Ala Pro Gln Val Ala Val Leu Gly
340 345 350
His Glu Ala Val Gly Gly Phe Leu Ser His Cys Gly Trp Asn Ser Thr
355 360 365
Leu Glu Ser Val Trp Cys Gly Val Pro Val Ala Thr Trp Pro Met Tyr
370 375 380
Gly Glu Gln Gln Leu Asn Ala Phe Glu Met Val Lys Glu Leu Gly Ile
385 390 395
Ala Val Glu Ile Glu Val Asp Tyr Lys Asn Glu Tyr Phe Asn Met Asn
400 405 410 415
Asn Asp Phe Ile Val Arg Ala Glu Glu Ile Glu Thr Lys Ile Lys Lys
420 425 430
Leu Met Met Asp Glu Lys Asn Ser Glu Ile Arg Lys Lys Val Lys Glu
435 440 445
Met Lys Glu Lys Ser Arg Leu Ala Met Ser Glu Asn Gly Ser Ser Tyr
450 455 460
Asn Ser Leu Ala Lys Leu Phe Glu Glu Ile Met
465 470 475
    
```

<210> 17  
 <211> 1428

<212> DNA  
 <213> 人参(Panax ginseng)

<220>  
 <221> misc feature  
 <223> gGT25-3  
 <400> 17

```

atgaagtcaag aattgatatt cttgcecgcc cgggceatcg gacacctcgt gggaatgggtg 60
gagatggceta aactcttcat cagtegacat gaaaatctct cggtcaccgt cctcatcgcg 120
aaattctaca lggatcgggg gglagacaac tacaataaat cactcttaac aaaccclacc 180
ecggcgtetea caattgtaaa teteceggaa accgaccccc aaaactatat gctcaaaeca 240
egeeacgceca tcttctctag cgtcategag actcagaaga cacacgtcgg agacataata 300
tccggatga ctcagtccega gtegaicag gtegttggtt tgetggciga ccttttgttc 360
atcaacatca iggacattgc caatgagttc aatgttccaa ctatgtata ctccctcgcc 420
ggagccggtc atcttggcct cgcgttccat ctccagaenc teaacgacaa aaaacaagat 480
glgaccgagt tcaggaacte ggatactgag ttatttggtae cgagttttgc aaaccgggtt 540
cccgcggagg tcttgcgctc gatgtatgtg gataaangaag gtgggtatga ttatctgttt 600
tcattgttcc ggaggtgcag agagteaag gcaattatta ttaacacggt tgaggagctg 660
gaacctatg cgateaatte cctccggatg gatagatga tccctccgat ctaccgggtg 720
ggaccctatac taattctcaa cggatgagcc caaaaactcgc atgaggctgc tgtgatcctt 780
ggttgggttag atgatacaacc accctcatct gtgggtgttt tgtgcttgg tagctatgga 840
agctttcaag aaaaccaggt gaaggagatt gcaatgggta tagagecgag tgggcategc 900
tcttgttggc ccttgcgtcc gctctatccc aaaggcgaga caaagcttca gcttaaaatac 960
tcaaatttga aagaaattct cccagtagga ttcttggaca ggacatcatg cgtcggaaaa 1020
gtgattggat gggecccgca agtggccgtg ctccgacatg agtcagtcgg aggttctctg 1080
tctcattgag gttggaatte gacattggag agtgtttggt gtgggttgc cgttgcacaa 1140
tggecaatgt atggtgagca acaactcaat gcttttgaga tggttaagga gtttaggtatt 1200
gcggtgaaaa ttgaggtgga ctataagaaa gattatttta acatgaagaa tgattttatt 1260
ggttagggcg aagaatcga gacaaaaata aagaagttag tgatggatga aaataatagt 1320
gaaataagaa aagaagttaa gaaaatgaaa gaaaagagta ggcttgcact gtctgagaat 1380
ggatcatctt ataattcatt ggcgaagccta tttaggaaaa ttatgtaa 1428
    
```

<210> 18  
 <211> 475  
 <212> PRT  
 <213> 人参(Panax ginseng)

[0007]

<220>  
 <221> misc feature  
 <223> gGT25-3  
 <400> 18

```

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

```

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

```

```

<210> 19
<211> 1419
<212> DNA
<213> 人参(Panax ginseng)

```

```

<220>
<221> misc feature
<223> gGT25-5
<400> 19

```

[0008]

```

atgaagtcag aattgatatt cgtgcccgtc cggccttcg ggcacctcgt ggggatggg      60
gagatggcta aacttttcat cagtcgacac gaaaacctat cagtcaccgt ccctatttcg      120
aaatttttca ttgataccgg gatagacaac tacaataaat cactctttagc gaaacctacc      180
cecgctctca caattataaa tctcccggaa atcgatcccc aaaaaatatt gctcaaacca      240
cgttgcgeca tctttccctc cctcatcgag aateagaaga cacacgtgeg agacgtaatg      300
tccgcgatga ctcagtcoga gtcgactcgg gtcgttggtt tgetggcaga cattttgttc      360
gtcgacaatc tggacattgc cgatgagttc aatgttccaa cttatgtata ctcctccfgec      420
ggagccggtt ttcttggect cgcgttccat ctccagacac tcaacgacga caaaaagcaa      480
gatgtgaccg agttcaggaa ttcggacact gagttatlgg taccgagttl tgcaaaaccg      540
gtcccgcceg agttcttggc gtcgatattt ttggaanaag atggtaggca tgatgttttg      600
ttatcatgtt actggaggtg cagggaggca aagggaafta ttgttaacac gtttgaggag      660
ctggaaccct atcgatcaaa ttcctctccg atggatagta tgatccctcc gatctaccgg      720
gtgggaccce tactaaatct caacggtgag ggacaaaact ccgatgaggc tgctgtgac      780
cttggtttgt tagatgatca accaccttca tctgtggtgt ttttgtgett tggtagcttt      840
ggaagctttc cagaaaacca ggtgaaggag attgcaatgg gtttagagcg cagcgggcat      900
egeltcttgt ggtccttgeg tccgtgtatc tctgaagggt agacaacgct tcaactlaaa      960
tactcaaatf fggaacttcc gcccggatcc ttggatagga catcatcgct cggaaaagtg      1020
attgatggg ccccaaaaat ggccatctca gcacacgagg cagtcggagg gttcgtgtct      1080
catgtgtggt ggaattcggc actagagagt gttgtgtatg gcatgcctgt cgcacaatgg      1140
ccaatgtacc gtgagcaaca actcaacgct tttgagatgg ttaaggagtt gggtcttggc      1200
gtgaaattg aggtggaacta taggaatgaa tataacaagt ctgattttat tgtaaggct      1260
gacgaaattg agacaaaaat aaagaagttg atgatggatg gaaagaatag taaaataagg      1320
aagaaggtaa aggaaatgaa agaaaagagt agggttgcca tctcggagaa tgggtcatct      1380
tatacttcat tggegaaget atttgaggaa attatgtaa

```

```

<210> 20
<211> 472
<212> PRT
<213> 人参(Panax ginseng)

```

```

<220>
<221> misc feature
<223> gGT25-5
<400> 20

```

```

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

```

115 120 125  
 Glu Phe Asn Val Pro Thr Tyr Val Tyr Ser Pro Ala Gly Ala Gly Phe  
 130 135 140  
 Leu Gly Leu Ala Phe His Leu Gln Thr Leu Asn Asp Asp Lys Lys Gln  
 145 150 155 160  
 Asp Val Thr Glu Phe Arg Asn Ser Asp Thr Glu Leu Leu Val Pro Ser  
 165 170 175  
 Phe Ala Asn Pro Val Pro Ala Glu Phe Leu Pro Ser Ile Phe Leu Glu  
 180 185 190  
 Lys Asp Gly Arg His Asp Val Leu Leu Ser Leu Tyr Trp Arg Cys Arg  
 195 200 205  
 Glu Ala Lys Gly Ile Ile Val Asn Thr Phe Glu Glu Leu Glu Pro Tyr  
 210 215 220  
 Ala Ile Asn Ser Leu Arg Met Asp Ser Met Ile Pro Pro Ile Tyr Pro  
 225 230 235 240  
 Val Gly Pro Ile Leu Asn Leu Asn Gly Glu Gly Gln Asn Ser Asp Glu  
 245 250 255  
 Ala Ala Val Ile Leu Gly Trp Leu Asp Asp Gln Pro Pro Ser Ser Val  
 260 265 270  
 Val Phe Leu Cys Phe Gly Ser Phe Gly Ser Phe Pro Glu Asn Gln Val  
 275 280 285  
 Lys Glu Ile Ala Met Gly Leu Glu Arg Ser Gly His Arg Phe Leu Trp  
 290 295 300  
 Ser Leu Arg Pro Cys Ile Ser Glu Gly Glu Thr Thr Leu Gln Leu Lys  
 305 310 315 320  
 Tyr Ser Asn Leu Glu Leu Pro Ala Gly Phe Leu Asp Arg Thr Ser Cys  
 325 330 335  
 Val Gly Lys Val Ile Gly Trp Ala Pro Gln Met Ala Ile Leu Ala His  
 340 345 350  
 Glu Ala Val Gly Gly Phe Val Ser His Cys Gly Trp Asn Ser Val Leu  
 355 360 365  
 Glu Ser Val Trp Tyr Gly Met Pro Val Ala Thr Trp Pro Met Tyr Gly  
 370 375 380  
 Glu Gln Gln Leu Asn Ala Phe Glu Met Val Lys Glu Leu Gly Leu Ala  
 385 390 395 400  
 Val Glu Ile Glu Val Asp Tyr Arg Asn Glu Tyr Asn Lys Ser Asp Phe  
 405 410 415  
 Ile Val Lys Ala Asp Glu Ile Glu Thr Lys Ile Lys Lys Leu Met Met  
 420 425 430  
 Asp Gly Lys Asn Ser Lys Ile Arg Lys Lys Val Lys Glu Met Lys Glu  
 435 440 445  
 Lys Ser Arg Val Ala Met Ser Glu Asn Gly Ser Ser Tyr Thr Ser Leu  
 450 455 460  
 Ala Lys Leu Phe Glu Glu Ile Met  
 465 470

[0009]

<210> 21  
 <211> 1488  
 <212> DNA  
 <213> 欧洲山芥(Barbarea vulgaris)  
  
 <220>  
 <221> misc\_feature  
 <223> 3GT1  
 <400> 21  
 atggtttccg aaatcacc caaataattat ccttttca ct ttgtttctt ccctttcaatg 60  
 gctcaagcc acatgatcc catggtt gat attgcaaggo tcttggctca ggcgggtgtg 120  
 aaaataacaa ttgtcacaac gccgcacaat gcagcaggt tegagaatgt cctaagccgt 180  
 gccattgagt ctggcttgc catcagcata gtgcaagtea agcttccate tcaagaagct 240  
 gctttaccag aaggaaatga gactttgat tcacttgtct caacaaagt t gctggtacct 300  
 ttctttaaag cggttaecat gcttgaagaa ccggtccaga agctctttga agagatgagc 360  
 cctcaacc aa gctgtataat ttctgatttt tgtttgcctt atacaagcaa aatcgccaag 420  
 aagtccaata tcccaagat ccttttccat ggcattgttt gcttttgtct tetgtgatg 480  
 catgttttac gcaagaaccg tgagatcttg gaaaacttaa agtctgaca ggagcatttc 540  
 gttgttctt attttctga tctgattgaa ttcacaagac ctcaagttcc attggcaaca 600  
 tatgttctg ggaatgga egagatcaag gaggatagg tagaagegga taagacttcc 660  
 tatggttga tagtcaacac atatcaagag ctgagaccg ct tttatgcaa cgctacaag 720  
 gaggcaaggt ctggtaaagc atggaccatt ggacctgtt cctttgtgca caaggtggga 780  
 gccgacaag cagagaggg aaacaagca gacattgat aagatgagtg tcttaaatgg 840  
 cttgatteta aagaagaag ttctggtteta tatgtttgct ttggaagtat ctgcagtctt 900  
 cctctgtct agctcaagga gctgggcta ggccttgagg aateccaaag accttcaat 960  
 tggctcgtaa gaggttggga gaagaacaa gattacttg agtggttctc ggagagcgga 1020  
 ttgaagaaa gagtaaaaga cagaggctt ctcatcaag gatgtcacc tcaaatgctt 1080  
 atccttgac atcattccgt tggaggctt ttaacacact gtggatgga ctgcaccctc 1140  
 gaaggaatca ettcaggegt tccattgctc acttggccac tgtttggaga ccaattctgc 1200  
 aaccaaaaac ttgtctgca ggtgctaaaa gtgggtgtaa gtgcgggggt tgaagaggtt 1260  
 acgaattggg gagaagagga gaaaatagga gtattagtg ataaagagg agtgaagaag 1320  
 gcagtggag aattaatggg tgagagtgat gatctaaag aaataagaaa aagagtcaaa 1380  
 gagcttggac aattagctca caaggctgtg gaggaaggag gctcatctca ttctaatac 1440  
 acatccttgc tagaagacat aatgcaacta gcacaaccta ataattaa 1488

<210> 22

<211> 495  
 <212> PRT  
 <213> 欧洲山芥(*Barbarea vulgaris*)

<220>  
 <221> misc\_feature  
 <223> 3GT1  
 <400> 22

Met Val Ser Glu Ile Thr His Lys Ser Tyr Pro Leu His Phe Val Leu  
 1 5 10 15  
 Phe Pro Phe Met Ala Gln Gly His Met Ile Pro Met Val Asp Ile Ala  
 20 25 30  
 Arg Leu Leu Ala Gln Arg Gly Val Lys Ile Thr Ile Val Thr Thr Pro  
 35 40 45  
 His Asn Ala Ala Arg Phe Glu Asn Val Leu Ser Arg Ala Ile Glu Ser  
 50 55 60  
 Gly Leu Pro Ile Ser Ile Val Gln Val Lys Leu Pro Ser Gln Glu Ala  
 65 70 75  
 Gly Leu Pro Glu Gly Asn Glu Thr Phe Asp Ser Leu Val Ser Thr Lys  
 80 85 90  
 Leu Leu Val Pro Phe Phe Lys Ala Val Asn Met Leu Glu Glu Pro Val  
 100 105 110  
 Gln Lys Leu Phe Glu Glu Met Ser Pro Gln Pro Ser Cys Ile Ile Ser  
 115 120 125  
 Asp Phe Cys Leu Pro Tyr Thr Ser Lys Ile Ala Lys Lys Phe Asn Ile  
 130 135 140  
 Pro Lys Ile Leu Phe His Gly Met Cys Cys Phe Cys Leu Leu Cys Met  
 145 150 155 160  
 His Val Leu Arg Lys Asn Arg Glu Ile Leu Glu Asn Leu Lys Ser Asp  
 165 170 175  
 Lys Glu His Phe Val Val Pro Tyr Phe Pro Asp Arg Val Glu Phe Thr  
 180 185 190  
 Arg Pro Gln Val Pro Leu Ala Thr Tyr Val Pro Gly Glu Trp His Glu  
 195 200 205  
 Ile Lys Glu Asp Met Val Glu Ala Asp Lys Thr Ser Tyr Gly Val Ile  
 210 215 220  
 Val Asn Thr Tyr Gln Glu Leu Glu Pro Ala Tyr Ala Asn Gly Tyr Lys  
 225 230 235 240  
 Glu Ala Arg Ser Gly Lys Ala Trp Thr Ile Gly Pro Val Ser Leu Cys  
 245 250 255  
 Asn Lys Val Gly Ala Asp Lys Ala Glu Arg Gly Asn Lys Ala Asp Ile  
 260 265 270  
 Asp Gln Asp Glu Cys Leu Lys Trp Leu Asp Ser Lys Glu Glu Gly Ser  
 275 280 285  
 Val Leu Tyr Val Cys Leu Gly Ser Ile Cys Ser Leu Pro Leu Ser Gln  
 290 295 300  
 Leu Lys Glu Leu Gly Leu Gly Leu Glu Glu Ser Gln Arg Pro Phe Ile  
 305 310 315 320  
 Trp Val Val Arg Gly Trp Glu Lys Asn Lys Glu Leu Leu Glu Trp Phe  
 325 330 335  
 Ser Glu Ser Gly Phe Glu Glu Arg Val Lys Asp Arg Gly Leu Leu Ile  
 340 345 350  
 Lys Gly Trp Ser Pro Gln Met Leu Ile Leu Ala His His Ser Val Gly  
 355 360 365  
 Gly Phe Leu Thr His Cys Gly Trp Asn Ser Thr Leu Glu Gly Ile Thr  
 370 375 380  
 Ser Gly Val Pro Leu Leu Thr Trp Pro Leu Phe Gly Asp Gln Phe Cys  
 385 390 395 400  
 Asn Gln Lys Leu Val Val Gln Val Leu Lys Val Gly Val Ser Ala Gly  
 405 410 415  
 Val Glu Glu Val Thr Asn Trp Gly Glu Glu Glu Lys Ile Gly Val Leu  
 420 425 430  
 Val Asp Lys Glu Gly Val Lys Lys Ala Val Glu Glu Leu Met Gly Glu  
 435 440 445  
 Ser Asp Asp Ala Lys Glu Ile Arg Lys Arg Val Lys Glu Leu Gly Gln  
 450 455 460  
 Leu Ala His Lys Ala Val Glu Glu Gly Gly Ser Ser His Ser Asn Ile  
 465 470 475 480  
 Thr Ser Leu Leu Glu Asp Ile Met Gln Leu Ala Gln Pro Asn Asn  
 485 490 495

[0010]

<210> 23  
 <211> 1488  
 <212> DNA  
 <213> 欧洲山芥(*Barbarea vulgaris*)

<220>  
 <221> misc\_feature  
 <223> 3GT2  
 <400> 23

```

atggtttccg aaatcaccca taatcttat cctcttcact ttgtttcttt ccttttcatt 60
gcctcaagcc acatgatccc catggttga attgcaagge tcttggecca gcgcgggtgt 120
aaaaatacaa ttgtcacaca cccgcacaat gcagcgaggt tcaagaatgt cctaagtcgt 180
gccattgagt ctgggttggc catcagcata gtgcaagtca agcttccate tcaagaagct 240
ggcttaccag aaggaatga gactctgat tcacttgtct cgatggagtt gatgatacat 300
tctttaaaga cggttaacat getggaagaa cccgtcccaga agctcttga agagatgagc 360
cctcaaccaa gctgtataat tctgatctt ttgttgcctt atacaagcaa aatcgccaag 420
aagttcaata tcccaagat cctcttccat ggcatgtgt gettttgtet tctgtgtatg 480
catatlltac gcaagaaccg tgagatcgtg gaaaacttaa agtctgacaa ggagcatttc 540
gttgttccct attttcctga tegagttgaa ttcacaagac ctcaagttee agtggcaaca 600
tatgttccct gagactggca cgagatcaag gaggatattg tagaaggcga taagacttec 660
tatggttfga tagtcaacac atatcaagag ctogagcctg cttatgceaa cgactacaag 720
gaggaaggt ctggtaaaag atggaaccat ggacctgtt ccttgtgcaa caaggtggga 780
gggacaaga cagagagggg aaacaaagca gacattgac agatgatgtg tctttaaattg 840
cttaattcta aagaagaagg ttcggttcta tatgtttgce ttggaagtat ctgcaattct 900
cctctgtctc agctcaagga gctcgggcta ggctttgagg aatccccaaag acctttcatt 960
tgggtcataa gaggttggga gaagaacaaa gagttacatg agtggttctc ggagagegga 1020
ttcgaagaaa gaatcaaaaga cagaggactt ctcatcaaaag gatgggctcc tcaaatgctt 1080
atactttcac atacttccgt tggagggttc tiaacacact gtgatggaa ctogactctt 1140
gaggggctaa ccgctggctt accactgctg acatggccc ttlttcgaga ccagttctgc 1200
aacgagaaac ttgccgtgca ggtattaaaa gccggtgtaa gcgccgggtg tgaccagcct 1260
atgaaatggg gagaagagga gaaaatagga gtgttgggtg ataaagaagg agtgaagaag 1320
gcagtggaa aattaaatgg tgagagtgat gatgctaaag agataagaag aagagccaaa 1380
gagctggag aattagctca caagctgtg gaggaaggga gctcatctca tctaatatc 1440
acatcccttc tagaagacat aatgcaacta gcacaatcca ataattaa 1488

```

- <210> 24
- <211> 495
- <212> PRT
- <213> 欧洲山芥 (*Barbarea vulgaris*)

- <220>
- <221> misc\_feature
- <223> 3GT2
- <400> 24

[0011]

```

Met Val Ser Glu Ile Thr His Lys Ser Tyr Pro Leu His Phe Val Leu
1
Phe Pro Phe Met Ala Gln Gly His Met Ile Pro Met Val Asp Ile Ala
20 25
Arg Leu Leu Ala Gln Arg Gly Val Lys Ile Thr Ile Val Thr Thr Pro
35 40 45
His Asn Ala Ala Arg Phe Lys Asn Val Leu Ser Arg Ala Ile Glu Ser
50 55 60
Gly Leu Pro Ile Ser Ile Val Gln Val Lys Leu Pro Ser Gln Glu Ala
65 70 75 80
Gly Leu Pro Glu Gly Asn Glu Thr Leu Asp Ser Leu Val Ser Met Glu
85 90 95
Leu Met Ile His Phe Leu Lys Ala Val Asn Met Leu Glu Glu Pro Val
100 105 110
Gln Lys Leu Phe Glu Glu Met Ser Pro Gln Pro Ser Cys Ile Ile Ser
115 120 125
Asp Phe Cys Leu Pro Tyr Thr Ser Lys Ile Ala Lys Lys Phe Asn Ile
130 135 140
Pro Lys Ile Leu Phe His Gly Met Cys Cys Phe Cys Leu Leu Cys Met
145 150 155 160
His Ile Leu Arg Lys Asn Arg Glu Ile Val Glu Asn Leu Lys Ser Asp
165 170 175
Lys Glu His Phe Val Val Pro Tyr Phe Pro Asp Arg Val Glu Phe Thr
180 185 190
Arg Pro Gln Val Pro Val Ala Thr Tyr Val Pro Gly Asp Trp His Glu
195 200 205
Ile Thr Glu Asp Met Val Glu Ala Asp Lys Thr Ser Tyr Gly Val Ile
210 215 220
Val Asn Thr Tyr Gln Glu Leu Glu Pro Ala Tyr Ala Asn Asp Tyr Lys
225 230 235 240
Glu Ala Arg Ser Gly Lys Ala Trp Thr Ile Gly Pro Val Ser Leu Cys
245 250 255
Asn Lys Val Gly Ala Asp Lys Ala Glu Arg Gly Asn Lys Ala Asp Ile
260 265 270
Asp Gln Asp Glu Cys Leu Lys Trp Leu Asn Ser Lys Glu Glu Gly Ser
275 280 285
Val Leu Tyr Val Cys Leu Gly Ser Ile Cys Asn Leu Pro Leu Ser Gln
290 295 300
Leu Lys Glu Leu Gly Leu Gly Leu Glu Glu Ser Gln Arg Pro Phe Ile
305 310 315 320
Trp Val Ile Arg Gly Trp Glu Lys Asn Lys Glu Leu His Glu Trp Phe
325 330 335
Ser Glu Ser Gly Phe Glu Glu Arg Ile Lys Asp Arg Gly Leu Leu Ile
340 345 350
Lys Gly Trp Ala Pro Gln Met Leu Ile Leu Ser His His Ser Val Gly
355 360 365
Gly Phe Leu Thr His Cys Gly Trp Asn Ser Thr Leu Glu Gly Leu Thr

```

序 列 表

```

370          375          380
Ala Gly Leu Pro Leu Leu Thr Trp Pro Leu Phe Ala Asp Gln Phe Cys
385          390          395
Asn Glu Lys Leu Ala Val Gln Val Leu Lys Ala Gly Val Ser Ala Gly
405          410          415
Val Asp Gln Pro Met Lys Trp Gly Glu Glu Lys Ile Gly Val Leu
420          425          430
Val Asp Lys Glu Gly Val Lys Lys Ala Val Glu Glu Leu Met Gly Glu
435          440          445
Ser Asp Asp Ala Lys Glu Ile Arg Arg Arg Ala Lys Glu Leu Gly Glu
450          455          460
Leu Ala His Lys Ala Val Glu Glu Gly Gly Ser Ser His Ser Asn Ile
465          470          475
Thr Ser Leu Leu Glu Asp Ile Met Gln Leu Ala Gln Ser Asn Asn
485          490          495

```

```

<210> 25
<211> 1329
<212> DNA
<213> 人参(Panax ginseng)

```

```

<220>
<221> misc feature
<223> gGT29
<400> 25

```

```

atggataacc aaaaaggtag aatcagtata gcggtgetac cattttttagc ccattggtcac      60
atatctecct tctttgagct agccaaacaa ctgcggaana gaaattgcaa tgttttctctc      120
tgttctacc caatcaalct laggctccatc aaggataaagg atlccctctgc ttctataaaa      180
ctagttgagc ttcattctcc atcttccctc gatcttccctc ctcactatca caccacaaat      240
ggctccctt cccatctcat geteccactc agaaaagcct ttgaaactgc aggcccccacc      300
ttctctgaaa tcttlaaaac cttaaacccc gatttgccta tttatgattt caatccctca      360
tgggcaccgg agatcgettc gtctcacaat attccggcag tttatttccct aaccacggca      420
gcagccagct ctccattgg cctacatgct ttcaaaaacc caggtgaaaa ataccattt      480
ccagatttt atgataacag taatattacc cctgaaccac ctctgcaga taacatgaag      540
ctacttcatg attttatcgc ttgtttcgaa cgalcttgcg acattatltt gallaagagt      600
tttagagaac tagangggna atatattgat ttgctttcca ctttatctga taaaactttg      660
gttctgttg gtccactcgt tcaagatcct atgggceata atgaagatcc aaaaacagag      720
cagattataa actggttga caaaaaggct gaatctacag tsgtgtttgt ctgctttgga      780
agtgagtatt ttctctccaa tgaggaattg gaagaagtag caattgggct agagattagc      840
acggttaatt tcaatgggc tgtgagatta attgaaggag agaaaaaagg gatlttacca      900
gaggggtttg ttcaaagggt aggagacaga ggattgggtg tggaggggtg ggtcccacag      960
gcaagaattt taggacattc aagcaccggt gggtttgtga gccattgtgg gtggagtct      1020
attggggaga gtatgaagtt tggggttcca gtaattgcca tggccaggca tcttgatcag      1080
cctttgaatg gtaagctggc ggcggaggtt ggtgtgggca tggaggttgt gagagatgag      1140
aatgggaagt ataagagaga agggattgca gaggtaataa gaaaagtgtt tgtggagaaa      1200
agtggggagg ttatcaggag gaaagcaagg gagttagtg agaaaatgaa agagaaagga      1260
gagcaagaga ttgatagggc attggaggag ctatlacaaa ttltgaagaa gaagaaagat      1320
gnacnafag

```

[0012]

```

<210> 26
<211> 442
<212> PRT
<213> 人参(Panax ginseng)

```

```

<220>
<221> misc feature
<223> gGT29
<400> 26

```

```

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

```

Cys Asp Ile Ile Leu Ile Lys Ser Phe Arg Glu Leu Glu Gly Lys Tyr  
 Ile Asp Leu Leu Ser Thr Leu Ser Asp Lys Thr Leu Val Pro Val Gly  
 Pro Leu Val Gln Asp Pro Met Gly His Asn Glu Asp Pro Lys Thr Glu  
 Gln Ile Ile Asn Trp Leu Asp Lys Arg Ala Glu Ser Thr Val Val Phe  
 Val Cys Phe Gly Ser Glu Tyr Phe Leu Ser Asn Glu Glu Leu Glu Glu  
 Val Ala Ile Gly Leu Glu Ile Ser Thr Val Asn Phe Ile Trp Ala Val  
 Arg Leu Ile Glu Gly Glu Lys Lys Gly Ile Leu Pro Glu Gly Phe Val  
 Gln Arg Val Gly Asp Arg Gly Leu Val Val Glu Gly Trp Ala Pro Gln  
 Ala Arg Ile Leu Gly His Ser Ser Thr Gly Gly Phe Val Ser His Cys  
 Gly Trp Ser Ser Ile Ala Glu Ser Met Lys Phe Gly Val Pro Val Ile  
 Ala Met Ala Arg His Leu Asp Gln Pro Leu Asn Gly Lys Leu Ala Ala  
 Glu Val Gly Val Gly Met Glu Val Val Arg Asp Glu Asn Gly Lys Tyr  
 Lys Arg Glu Gly Ile Ala Glu Val Ile Arg Lys Val Val Val Glu Lys  
 Ser Gly Glu Val Ile Arg Arg Lys Ala Arg Glu Leu Ser Glu Lys Met  
 Lys Glu Lys Gly Glu Gln Glu Ile Asp Arg Ala Leu Glu Glu Leu Val  
 Gln Ile Cys Lys Lys Lys Lys Asp Glu Gln

<210> 27  
 <211> 1329  
 <212> DNA  
 <213> 人參(Panax ginseng)

[0013]

<220>  
 <221> misc\_feature  
 <223> gGT29-3  
 <400> 27

atggataacc	aagaaggtag	aatcagtata	gcgttgetac	catttttagc	ccatggctac	60
atatctccct	tccttgagct	agccaaacaa	ctcgcaaaaa	gaaattgcaa	tgttttcctc	120
tgttctaccc	caatcaatct	tagctccate	aagaataagg	attcctctgc	ttctataaaa	180
ctagttgagc	ttcatcttcc	atcttccctc	gatcttcttc	ctcactatca	caccacaaat	240
ggctccctt	cccatctcat	ggteccacte	ataaacgctt	tfgaacagc	aggecccaac	300
ttctctgaaa	tccttaaaac	cttaaacccc	gatttgctta	tttatgattt	caatccctca	360
tgggcacegg	agategcttc	gtctcacaat	attccggcag	tttatttctt	aaccacggca	420
gcagccagct	cttccattgg	cttacaatgt	ttcaaaaacc	caggtgaaaa	atacceattt	480
ccagattttt	atgataacag	taataatacc	cttgaaccac	cttctgcaga	taacatgaag	540
ctacttcatg	atfttatcgc	ttgtttcgaa	cgatcttggc	acattatttt	gattaagagt	600
tttatagaac	tagaagggaa	atatatcgat	ttgctttcca	ctttatctga	taaaaactttg	660
gttctctgtt	gtccaactgt	tcaagatcct	atgggccata	atgaagatcc	aaaaacagag	720
cagattataa	acttgcttga	caaaaaggct	gaatctacag	tgggttttgt	ctgctttgga	780
agtgagtatt	ttctctccaa	tgaggaattg	gaagaagtag	caattgggct	agagattagc	840
atggttaatt	tcatatgggc	tgtfngatta	attgaaggag	ngnaaanagg	ggttttaccn	900
gagggatttg	ttcaaagggt	aggagacaga	ggatttggtg	tggaggggtg	ggctccacag	960
gcaagaattt	taggacattc	aagcaccggg	gggtttgtga	gccattgttg	gtggagtctt	1020
attgeggaga	glatgaagtt	tggggltcca	gtaattgcca	tggccaggca	tcttgatcag	1080
ctttgaaatg	gtaagctggc	ggcggaggtt	gggtgaggca	tggaggttgt	gagagatgaa	1140
aatgggnagt	ataagagaga	agggattgca	gaggtaatat	gaanaagctg	tgtggagaa	1200
agtggggagg	ttatgaggag	gaaagcaagg	gaattgagtg	agaaaatgaa	agagaaagga	1260
gaggaagaga	ttgatagggc	agtggaggag	ctagtacaaa	tttgaagaa	gaagaaagat	1320
gcacaatag						1329

<210> 28  
 <211> 442  
 <212> PRT  
 <213> 人參(Panax ginseng)

<220>  
 <221> misc\_feature  
 <223> gGT29-3  
 <400> 28

Met Asp Asn Gln Glu Gly Arg Ile Ser Ile Ala Leu Leu Pro Phe Leu  
 1 5 10 15  
 Ala His Gly His Ile Ser Pro Phe Phe Glu Leu Ala Lys Gln Leu Ala  
 20 25 30  
 Lys Arg Asn Cys Asn Val Phe Leu Cys Ser Thr Pro Ile Asn Leu Ser

Ser Ile Lys Asn Lys Asp Ser Ser Ala Ser Ile Lys Leu Val Glu Leu  
 35 40 45  
 50 55 60  
 His Leu Pro Ser Ser Pro Asp Leu Pro Pro His Tyr His Thr Thr Asn  
 65 70 75 80  
 Gly Leu Pro Ser His Leu Met Val Pro Leu Ile Asn Ala Phe Glu Thr  
 85 90 95  
 Ala Gly Pro Thr Phe Ser Glu Ile Leu Lys Thr Leu Asn Pro Asp Leu  
 100 105 110  
 Leu Ile Tyr Asp Phe Asn Pro Ser Trp Ala Pro Glu Ile Ala Ser Ser  
 115 120 125  
 His Asn Ile Pro Ala Val Tyr Phe Leu Thr Thr Ala Ala Ala Ser Ser  
 130 135 140  
 Ser Ile Gly Leu His Ala Phe Lys Asn Pro Gly Glu Lys Tyr Pro Phe  
 145 150 155 160  
 Pro Asp Phe Tyr Asp Asn Ser Asn Asn Thr Pro Glu Pro Pro Ser Ala  
 165 170 175  
 Asp Asn Met Lys Leu Leu His Asp Phe Ile Ala Cys Phe Glu Arg Ser  
 180 185 190  
 Cys Asp Ile Ile Leu Ile Lys Ser Phe Ile Glu Leu Glu Gly Lys Tyr  
 195 200 205  
 Ile Asp Leu Leu Ser Thr Leu Ser Asp Lys Thr Leu Val Pro Val Gly  
 210 215 220  
 Pro Leu Val Gln Asp Pro Met Gly His Asn Glu Asp Pro Lys Thr Glu  
 225 230 235 240  
 Gln Ile Ile Asn Trp Leu Asp Lys Arg Ala Glu Ser Thr Val Val Phe  
 245 250 255  
 Val Cys Phe Gly Ser Glu Tyr Phe Leu Ser Asn Glu Glu Leu Glu Glu  
 260 265 270  
 Val Ala Ile Gly Leu Glu Ile Ser Met Val Asn Phe Ile Trp Ala Val  
 275 280 285  
 Arg Leu Ile Glu Gly Glu Lys Lys Gly Val Leu Pro Glu Gly Phe Val  
 290 295 300  
 Gln Arg Val Gly Asp Arg Gly Leu Val Val Glu Gly Trp Ala Pro Gln  
 305 310 315 320  
 Ala Arg Ile Leu Gly His Ser Ser Thr Gly Gly Phe Val Ser His Cys  
 325 330 335  
 Gly Trp Ser Ser Ile Ala Glu Ser Met Lys Phe Gly Val Pro Val Ile  
 340 345 350  
 Ala Met Ala Arg His Leu Asp Gln Pro Leu Asn Gly Lys Leu Ala Ala  
 355 360 365  
 Glu Val Gly Val Gly Met Glu Val Val Arg Asp Glu Asn Gly Lys Tyr  
 370 375 380  
 Lys Arg Glu Gly Ile Ala Glu Val Ile Arg Lys Val Val Val Glu Lys  
 385 390 395 400  
 Ser Gly Glu Val Met Arg Arg Lys Ala Arg Glu Leu Ser Glu Lys Met  
 405 410 415  
 Lys Glu Lys Gly Glu Glu Glu Ile Asp Arg Ala Val Glu Glu Leu Val  
 420 425 430 435  
 Gln Ile Cys Lys Lys Lys Lys Asp Ala Gln 440

[0014]

<210> 29  
 <211> 20  
 <212> DNA  
 <213> 人工序列  
 <400> 29  
 atggtttccg aaatcaccca 20

<210> 30  
 <211> 24  
 <212> DNA  
 <213> 人工序列  
 <400> 30  
 ttaattatir grttgtgcta gttg 24

<210> 31  
 <211> 35  
 <212> DNA  
 <213> 人工序列  
 <400> 31  
 gactggatcc atggtttccg aaatcaccca taaat 35

<210> 32  
 <211> 30  
 <212> DNA  
 <213> 人工序列  
 <400> 32  
 gcgcgtcgac attattaggt tgtgctagtt 30

	<210> 33		
	<211> 30		
	<212> DNA		
	<213> 人工序列		
	<400> 33		
	ggeggtgac attattggat lgtgctagtt		30
	<210> 34		
	<211> 26		
	<212> DNA		
	<213> 人工序列		
	<400> 34		
	agtaagaaaa acagagttca tcatgg		26
	<210> 35		
	<211> 20		
	<212> DNA		
	<213> 人工序列		
	<400> 35		
	gectcggta ggctagctgt		20
	<210> 36		
	<211> 35		
	<212> DNA		
	<213> 人工序列		
	<400> 36		
	ggeggtacc atggataacc aaaaaggtag aatca		35
	<210> 37		
	<211> 56		
	<212> DNA		
	<213> 人工序列		
	<400> 37		
	gcegcctgag tcaatgatga tgatgatgat gttgttcate tttcttcttc ttacaa		56
[0015]	<210> 38		
	<211> 35		
	<212> DNA		
	<213> 人工序列		
	<400> 38		
	ggeggtacc atggataacc aagaaggtag aatca		35
	<210> 39		
	<211> 56		
	<212> DNA		
	<213> 人工序列		
	<400> 39		
	gcegcctgag tcaatgatga tgatgatgat gttgttcate tttcttcttc ttacaa		56
	<210> 40		
	<211> 1494		
	<212> DNA		
	<213> 蒺藜苜蓿 (Medicago truncatula)		
	<220>		
	<221> misc_feature		
	<223> 3GT3		
	<400> 40		
	atggaagtg ttgaagttga acaaccattg aaagtttact tcattccatt tcttgcattc		60
	ggacatatga tccctctttt tgacatagca actatgtttg cateccgttg ccagcaagta		120
	acagtcata ccactcccc caacgcaaaa tcccttacca aatctctctc atccgacgtc		180
	cttccattcc ttctgttctta caccgttgac tteccctccc aacaagtctg cctccctgaa		240
	ggtattgaat ctatgtcttc aactacggac cccaccacca cttggaagat ccatactggc		300
	ggatgctcc ttaaagaacc tattggggat ttcattgaga atgatccacc ggattgtatc		360
	atctccgaet ccacgtacce atgggttaat gacttggcgc ataagtctca gatcccaaac		420
	atcacatca atfgattgfg cctttttgct gctccctcgc tggaaacctt caaaacaaag		480
	aatttactta agtctcagac agattctgat teggattcaa gttcctttgt tgttccaaat		540
	tttctcacc atatcacctt gtttggaaaa ccgccaaaagg taatcgggat attcatggga		600
	atgatgcttg agacgggtct taaaagtata gcaactaata teaacaactt cagtgaactt		660
	gatggagaag agtgcataca acactacgag aaagccacgg gtccacaaggt ttggcalett		720
	ggtccaaact ctcttattcg caaaactgct caagagaaat cagagagggg aatgagggt		780
	gctgtgaatg tgcacgagag cctgagttgg ctgatccag agagagttaa ctcagtgttg		840
	tacatagtt ttggaagcat caactatfff tctgataaac aactatacga gatggcatgt		900
	ggatagaag cctccggtea cccattcata tgggttgttc ctgagaagaa agggaaaaga		960
	gatgagagcg aagaagagaa agaaaagttg ttaccgaagg gatattgaaga gagaaatate		1020

```

gggaagaagg gtttgatcat taggggttgg gcccacagg ttaagatatt gageccacct 1080
gcagtggggg gattfatgac gcattgctgg gggaactcaa ccgtagagge tgttagtgcg 1140
ggagttccaa tgataacgtg gceggttcat ggagatcaat tctacaatga gaaactgata 1200
acacaattcc gaggaattgg agttgaagtc ggtgcaacag aatgggtgac gagtgggtgc 1260
gcggagagaa agaagttagt gaggcagat agcatagaga aggcgtgaag gagattgatg 1320
gacgggtgtg atgaagctga aaatatcagg ctacgtgctc gagagtttgg agaaaaagct 1380
atacaagcta ttcagaaggg tggetcgtct tataataatt tgttggett gattgacgaa 1440
cttaaaagat cgagagacct taaaagattg agagacctca agetggatga ttaa 1494
    
```

```

<210> 41
<211> 497
<212> PRT
<213> 蒺藜苜蓿 (Medicago truncatula)
    
```

```

<220>
<221> misc feature
<223> 3GT3
<400> 41
    
```

[0016]

```

Met Glu Gly Val Glu Val Glu Gln Pro Leu Lys Val Tyr Phe Ile Pro
1 5 10 15
Phe Leu Ala Ser Gly His Met Ile Pro Leu Phe Asp Ile Ala Thr Met
20 25 30
Phe Ala Ser Arg Gly Gln Gln Val Thr Val Ile Thr Thr Pro Ala Asn
35 40 45
Ala Lys Ser Leu Thr Lys Ser Leu Ser Ser Asp Ala Pro Ser Phe Leu
50 55 60
Arg Leu His Thr Val Asp Phe Pro Ser Gln Gln Val Gly Leu Pro Glu
65 70 75 80
Gly Ile Glu Ser Met Ser Ser Thr Thr Asp Pro Thr Thr Thr Trp Lys
85 90 95
Ile His Thr Gly Ala Met Leu Leu Lys Glu Pro Ile Gly Asp Phe Ile
100 105 110
Glu Asn Asp Pro Pro Asp Cys Ile Ile Ser Asp Ser Thr Tyr Pro Trp
115 120 125
Val Asn Asp Leu Ala Asp Lys Phe Gln Ile Pro Asn Ile Thr Phe Asn
130 135 140
Gly Leu Cys Leu Phe Ala Val Ser Leu Val Glu Thr Leu Lys Thr Asn
145 150 155 160
Asn Leu Leu Lys Ser Gln Thr Asp Ser Asp Ser Asp Ser Ser Ser Phe
165 170 175
Val Val Pro Asn Phe Pro His His Ile Thr Leu Cys Gly Lys Pro Pro
180 185 190
Lys Val Ile Gly Ile Phe Met Gly Met Met Leu Glu Thr Val Leu Lys
195 200 205
Ser Lys Ala Leu Ile Ile Asn Asn Phe Ser Glu Leu Asp Gly Glu Glu
210 215 220
Cys Ile Gln His Tyr Glu Lys Ala Thr Gly His Lys Val Trp His Leu
225 230 235 240
Gly Pro Thr Ser Leu Ile Arg Lys Thr Ala Gln Glu Lys Ser Glu Arg
245 250 255
Gly Asn Glu Gly Ala Val Asn Val His Glu Ser Leu Ser Trp Leu Asp
260 265 270
Ser Glu Arg Val Asn Ser Val Leu Tyr Ile Cys Phe Gly Ser Ile Asn
275 280 285
Tyr Phe Ser Asp Lys Gln Leu Tyr Glu Met Ala Cys Ala Ile Glu Ala
290 295 300
Ser Gly His Pro Phe Ile Trp Val Val Pro Glu Lys Lys Gly Lys Glu
305 310 315 320
Asp Glu Ser Glu Glu Lys Glu Lys Trp Leu Pro Lys Gly Phe Glu
325 330 335
Glu Arg Asn Ile Gly Lys Lys Gly Leu Ile Ile Arg Gly Trp Ala Pro
340 345 350
Gln Val Lys Ile Leu Ser His Pro Ala Val Gly Gly Phe Met Thr His
355 360 365
Cys Gly Asn Ser Thr Val Glu Ala Val Ser Ala Gly Val Pro Met
370 375 380
Ile Thr Trp Pro Val His Gly Asp Gln Phe Tyr Asn Glu Lys Leu Ile
385 390 395 400
Thr Gln Phe Arg Gly Ile Gly Val Glu Val Gly Ala Thr Glu Trp Cys
405 410 415
Thr Ser Gly Val Ala Glu Arg Lys Lys Leu Val Ser Arg Asp Ser Ile
420 425 430
Glu Lys Ala Val Arg Arg Leu Met Asp Gly Gly Asp Glu Ala Glu Asn
435 440 445
Ile Arg Leu Arg Ala Arg Glu Phe Gly Glu Lys Ala Ile Gln Ala Ile
450 455 460
Gln Glu Gly Gly Ser Ser Tyr Asn Asn Leu Leu Ala Leu Ile Asp Glu
465 470 475 480
Leu Lys Arg Ser Arg Asp Leu Lys Arg Leu Arg Asp Leu Lys Leu Asp
485 490 495
Asp
    
```

```

<210> 42
<211> 1374
<212> DNA
<213> 人参(Panax ginseng)

<220>
<221> misc_feature
<223> 3GT4
<400> 42
atggagagag aantgttgag caaaaetcaac attatgitten tcccatlccc agctcaagge      60
cacatgagcc caatgatgca attcgccaag cgtttagcct ggaaggcct gcgaatcacg      120
atagttcttc eggcetcaaat tggagatttc atgcaataaa ccaaccatt gatcaaacct      180
gagtgcctct cctttgattt tgataaagac gatgggatgc cataagecat gcaggettat      240
atgggagttg taaaactcaa ggtcacaaat aaactgagtg acctactega gaagcaaaga      300
acaaatggcf accctggtta tttgctagt gttgattcaf tatatccatc tgggtagaa      360
atgtgccacc aacttggggg aaaaggaget ccatttttca ctactcttg tgcgttggt      420
gceatttatt ataatgctcg cttagggaaa ttgaagatac ctctgagga agggitgaet      480
tcgttttcat tgccttcaat tccattgttg gggagagatg atttgcceat tattaggact      540
ggcaactttc ctfatctctt tgagcatttg gggatcagtt tttcagatct tgataaagcg      600
atgttgatct tttcaatac ttttgataag ctfgaaaatg aggaagcaaa atggtctct      660
agccaatggc caattacatc catcgacca ttaatccctt caatgtaett agacaacaaa      720
ttaccaaaag acaaagacaa tggcattaat ttctacaagg cagacgtcgg atcglgcatc      780
aagtggctag acgccaaaaga ccttggetcg gtatctctac cctcattcgg gagcgtgaag      840
cacaacctcg ccatgacta catggacgaa gtacatggg gcttgttaca tagcaaatat      900
cacttcatat gggttgttat agaatecgaa cgtacaaaag tctctagcga tttcttggea      960
gagccagagg cagaggaaaa aggcetaata gtgagttggt gcccctcaact ceaagttttg      1020
tcacataaat ctatagggag ttttatgaet cattgtggtt ggaactcgac ggttaggca      1080
ttgagtttgg gcgtgccaat ggtgacctg ccacacagtt ttgatcagcc tgetaatgce      1140
aagtatatcg tggatgtatg gcaaatggg gttcgggttc cgatttggta agagggggtt      1200
gttttgaggg gagaagttgc taactgtata aaggatgtta tggaggggga aataggggat      1260
gagcttagag gaaatgcltt gaaatggaag gggttggcig tggaggcaat ggagaagggg      1320
ggtagctctg ataagaatat tgatgagttc atttcaagc ttgtttctc ctga      1374

```

```

<210> 43
<211> 457
<212> PRT
<213> 人参(Panax ginseng)

```

[0017]

```

<220>
<221> misc_feature
<223> 3GT4
<400> 43
Met Glu Arg Glu Met Leu Ser Lys Thr His Ile Met Phe Ile Pro Phe
1      5      10      15
Pro Ala Gln Gly His Met Ser Pro Met Met Gln Phe Ala Lys Arg Leu
20     25     30
Ala Trp Lys Gly Leu Arg Ile Thr Ile Val Leu Pro Ala Gln Ile Arg
35     40     45
Asp Phe Met Gln Ile Thr Asn Pro Leu Ile Asn Thr Gln Cys Ile Ser
50     55     60
Phe Asp Phe Asp Lys Asp Asp Gly Met Pro Tyr Ser Met Gln Ala Tyr
65     70     75     80
Met Gly Val Val Lys Leu Lys Val Thr Asn Lys Leu Ser Asp Leu Leu
85     90     95
Glu Lys Gln Arg Thr Asn Gly Tyr Pro Val Asn Leu Leu Val Val Asp
100    105    110
Ser Leu Tyr Pro Ser Arg Val Glu Met Cys His Gln Leu Gly Val Lys
115    120    125
Gly Ala Pro Phe Phe Thr His Ser Cys Ala Val Gly Ala Ile Tyr Tyr
130    135    140
Asn Ala Arg Leu Gly Lys Leu Lys Ile Pro Pro Glu Glu Gly Leu Thr
145    150    155    160
Ser Val Ser Leu Pro Ser Ile Pro Leu Leu Gly Arg Asp Asp Leu Pro
165    170    175
Ile Ile Arg Thr Gly Thr Phe Pro Asp Leu Phe Glu His Leu Gly Asn
180    185    190
Gln Phe Ser Asp Leu Asp Lys Ala Asp Trp Ile Phe Phe Asn Thr Phe
195    200    205
Asp Lys Leu Glu Asn Glu Glu Ala Lys Trp Leu Ser Ser Gln Trp Pro
210    215    220
Ile Thr Ser Ile Gly Pro Leu Ile Pro Ser Met Tyr Leu Asp Lys Gln
225    230    235
Leu Pro Asn Asp Lys Asp Asn Gly Ile Asn Phe Tyr Lys Ala Asp Val
240    245    250    255
Gly Ser Cys Ile Lys Trp Leu Asp Ala Lys Asp Pro Gly Ser Val Val
260    265    270
Tyr Ala Ser Phe Gly Ser Val Lys His Asn Leu Gly Asp Asp Tyr Met
275    280    285
Asp Glu Val Ala Trp Gly Leu Leu His Ser Lys Tyr His Phe Ile Trp
290    295    300
Val Val Ile Glu Ser Glu Arg Thr Lys Leu Ser Ser Asp Phe Leu Ala

```

305  
 Glu Ala Glu Ala Glu Glu Lys Gly Leu Ile Val Ser Trp Cys Pro Gln  
 310  
 325  
 Leu Gln Val Leu Ser His Lys Ser Ile Gly Ser Phe Met Thr His Cys  
 330  
 340  
 Gly Trp Asn Ser Thr Val Glu Ala Leu Ser Leu Gly Val Pro Met Val  
 345  
 355  
 Ala Leu Pro Gln Gln Phe Asp Gln Pro Ala Asn Ala Lys Tyr Ile Val  
 360  
 370  
 Asp Val Trp Gln Ile Gly Val Arg Val Pro Ile Gly Glu Glu Gly Val  
 375  
 385  
 Val Leu Arg Gly Glu Val Ala Asn Cys Ile Lys Asp Val Met Glu Gly  
 390  
 400  
 Glu Ile Gly Asp Glu Leu Arg Gly Asn Ala Leu Lys Trp Lys Gly Leu  
 405  
 415  
 Ala Val Glu Ala Met Glu Lys Gly Gly Ser Ser Asp Lys Asn Ile Asp  
 420  
 430  
 Glu Phe Ile Ser Lys Leu Val Ser Ser  
 435  
 445  
 450  
 455

<210> 44  
 <211> 24  
 <212> DNA  
 <213> 人工序列  
 <400> 44  
 atggaaggtg ttgaagttga acaa 24  
  
 <210> 45  
 <211> 24  
 <212> DNA  
 <213> 人工序列  
 <400> 45  
 ttaatcatcc agettgaggt ctct 24  
  
 <210> 46  
 <211> 25  
 <212> DNA  
 <213> 人工序列  
 <400> 46  
 tacataaata tataatacat aggca 25  
  
 <210> 47  
 <211> 25  
 <212> DNA  
 <213> 人工序列  
 <400> 47  
 aaaatacata caaaatcttg aaata 25  
  
 <210> 48  
 <211> 45  
 <212> DNA  
 <213> 人工序列  
 <400> 48  
 actttaagaa ggagatatac catggaaggt gttgaagttg acaaa 45  
  
 <210> 49  
 <211> 45  
 <212> DNA  
 <213> 人工序列  
 <400> 49  
 ctcgagtgcg gccgcaagct tatcatccag cttgaggtct ctcaa 45  
  
 <210> 50  
 <211> 46  
 <212> DNA  
 <213> 人工序列  
 <400> 50  
 actttaagaa ggagatatac catggagaga gaaatgttga gcaaaa 46  
  
 <210> 51  
 <211> 46  
 <212> DNA  
 <213> 人工序列  
 <400> 51  
 ctcgagtgcg gccgcaagct tggaggaaac aagctttgaa atgaac 46

[0018]

<210> 52  
 <211> 27  
 <212> DNA  
 <213> 人工序列  
 <400> 52  
 ggtatatctc ettettaag ttaaaca 27

<210> 53  
 <211> 27  
 <212> DNA  
 <213> 人工序列  
 <400> 53  
 aagettggcg ccgcactega gcaccac 27

<210> 54  
 <211> 1341  
 <212> DNA  
 <213> 人参(Panax ginseng)

<220>  
 <221> misc feature  
 <223> gGT29-4  
 <400> 54  
 atggataacc aaaaaggtag aatcagtata gggttgetac catttttagc ccatggccac 60  
 atttctccat tctttgagct agccaagcat ctctcaaaaa gaaattgtaa tatattcctc 120  
 tgttctaccc caatcaatct tagctccatc aagaacagaa tatctgataa ggattcctct 180  
 gcttctataa aactagtaga gcttcatctt ccatcttccc ctgactctcc tccctactac 240  
 cacaccacaa atggctctcc tcccatctc atggctccac tcagaacgc ctttgaacaa 300  
 gcagccccca ccttctctga aatccttaaa accttaaac ctgatttgc tttttatgat 360  
 ttcaatecct catgggcacc ggagatcgct tcgtctcaca atattccggc agtttgttcc 420  
 ataattgggg gageagcctc ctttccatg agcctacata gtttcaaaaa cccaggtgaa 480  
 aaatacccat ttctagattt tgatgataac agtaaatata cccctgaacc accttcagca 540  
 gataaacatga agttattact tgattttatg acttgtttcg aacgatcttg cgacattatt 600  
 ttgattaaga gttttagaga actagaaggg aaatatattg atttttatct tactttatct 660  
 gataaaaact tggttcctgt tggtecacic gttcaagatc ctatgggeca taatgaagat 720  
 ccaaaaaacag agcagtttat aaactggcctt gacaaaaagg ctgaatctac agtgggtgtt 780  
 gctctgcttg gaagttagta tttctctccc aatgaggaa tggagaagat agcaattggg 840  
 ctagagatta gcattggttaa ttteatatgg gctgtgagat taattgaagg agagaaaaaa 900  
 ggggttttac cagaggggtt tgttcaaagg gtaggagaca gaggattggt tgtggagggg 960  
 tggctctcac aggeaagaat tttaggacat tcaagcacc gttgggtttgt gagccattgt 1020  
 ggggggagtt ctattacgga gagtatgaag tttggggttc cagtaattgc catggccagg 1080  
 catcttgatc agcctttgaa tggtaagctg gcgcccaggg ttgggtgtgg catggaggtt 1140  
 gtgagagatg aaaaatgggaa gtataagaga gaagggatg cagagtaat aagaaaaagc 1200  
 gttgtggaga aaagtgggga ggttatgagg aggaaagcaa ggggaattgag tgagaaaatg 1260  
 aaagagaaag gagaggaaga gattgatagg gcagtgagg agctagtaca aatttctaag 1320  
 aagaagaag atgaacaata g 1341

[0019]

<210> 55  
 <211> 446  
 <212> PRT  
 <213> 人参(Panax ginseng)

<220>  
 <221> misc feature  
 <223> gGT29-4  
 <400> 55  
 Met. Asp Asn Gln Lys Gly Arg Ile Ser Ile Ala Leu Leu Pro Phe Leu  
 1 5 10 15  
 Ala His Gly His Ile Ser Pro Phe Phe Glu Leu Ala Lys His Leu Ser  
 20 25 30  
 Lys Arg Asn Cys Asn Ile Phe Leu Cys Ser Thr Pro Ile Asn Leu Ser  
 35 40 45  
 Ser Ile Lys Asn Arg Ile Ser Asp Lys Asp Ser Ser Ala Ser Ile Lys  
 50 55 60  
 Leu Val Glu Leu His Leu Pro Ser Ser Pro Asp Leu Pro Pro His Tyr  
 65 70 75 80  
 His Thr Thr Asn Gly Leu Pro Ser His Leu Met Val Pro Leu Arg Asn  
 85 90 95  
 Ala Phe Glu Thr Ala Ala Pro Thr Phe Ser Glu Ile Leu Lys Thr Leu  
 100 105 110  
 Asn Pro Asp Leu Leu Ile Tyr Asp Phe Asn Pro Ser Trp Ala Pro Glu  
 115 120 125  
 Ile Ala Ser Ser His Asn Ile Pro Ala Val Cys Phe Ile Ile Gly Gly  
 130 135 140  
 Ala Ala Ser Phe Ser Met Ser Leu His Ser Phe Lys Asn Pro Gly Glu  
 145 150 155 160  
 Lys Tyr Pro Phe Leu Asp Phe Asp Asp Asn Ser Asn Ile Thr Pro Glu  
 165 170 175  
 Pro Pro Ser Ala Asp Asn Met Lys Leu Leu Leu Asp Phe Met Thr Cys

```

180          185          190
Phe Glu Arg Ser Cys Asp Ile Ile Leu Ile Lys Ser Phe Arg Glu Leu
195          200          205
Glu Gly Lys Tyr Phe Asp Phe Tyr Ser Thr Leu Ser Asp Lys Thr Leu
210          215          220
Val Pro Val Gly Pro Leu Val Gln Asp Pro Met Gly His Asn Glu Asp
225          230          235
Pro Lys Thr Glu Gln Phe Ile Asn Trp Leu Asp Lys Arg Ala Glu Ser
240          245          250
Thr Val Val Phe Val Cys Phe Gly Ser Glu Tyr Phe Leu Ser Asn Glu
255          260          265
Glu Leu Glu Glu Val Ala Ile Gly Leu Glu Ile Ser Met Val Asn Phe
270          275          280
Ile Trp Ala Val Arg Leu Ile Glu Gly Glu Lys Lys Gly Val Leu Pro
285          290          295
Glu Gly Phe Val Gln Arg Val Gly Asp Arg Gly Leu Val Val Glu Gly
300          305          310
Trp Ala Pro Gln Ala Arg Ile Leu Gly His Ser Ser Thr Gly Gly Phe
315          320          325
Val Ser His Cys Gly Trp Ser Ser Ile Thr Glu Ser Met Lys Phe Gly
325          330          335
Val Pro Val Ile Ala Met Ala Arg His Leu Asp Gln Pro Leu Asn Gly
340          345          350
Lys Leu Ala Ala Glu Val Gly Val Gly Met Glu Val Val Arg Asp Glu
355          360          365
Asn Gly Lys Tyr Lys Arg Glu Gly Ile Ala Glu Val Ile Arg Lys Val
370          375          380
Val Val Glu Lys Ser Gly Glu Val Met Arg Arg Lys Ala Arg Glu Leu
385          390          395
Ser Glu Lys Met Lys Glu Lys Gly Glu Glu Glu Ile Asp Arg Ala Val
400          405          410
Glu Glu Leu Val Gln Ile Cys Lys Lys Lys Lys Asp Glu Glu
415          420          425
435          440          445

```

```

<210> 56
<211> 1341
<212> DNA
<213> 人参(Panax ginseng)

```

[0020]

```

<220>
<221> misc_feature
<223> gGT29-5
<400> 56

```

```

atggataacc aaaagggtag aatcngtata gttatgetgc cattttttagc ceatggccac 60
atttctccat tctttgagct agccaaagcat ctctcaaaaa gaaattgtaa tatattctct 120
tgttctacec caatcaatct tagctccate aagaacagaa tatctgataa ggattctctt 180
gcttctataa aactagtaga gcttccatctt ccactcttccc ctgatcttcc tcctactctac 240
cacaccacaa atggctctcc tteccatctc atggctccac tcagaaaagc ctttgaacaa 300
gcagcccceca ccttctctga aatccttaaa accttaaac ctgatttget tatttatgat 360
ttcaatccct catgggcacc ggagatgect tctctcaca atattccggc agtttgttte 420
ataattgggg gageagcctc cltttccatg agcctacata gtttcaaaaa ccaggtgaa 480
aaatacceat tcttagattt tgatgataac agtaatatta cccctgaacc accttcagca 540
gataaacatga agttattact tgattttatg acttgtttcg aacgatcttg cgacattatt 600
ttgattaaga gttttagaga actagaaggg aaatatactg atttgctttc cactttatct 660
gafaaaaact ttggttctgt tggtcactc gttcaagate ctatgggcca taatgaagat 720
ccaaaaacag agcagattat aaactggcct gacaaaaagg ctgaatctac agtgggtgtt 780
gtctgctttg gaagttagta tttctctccc aatgaggnaat tggagaagat agcaattggg 840
ctagagatta gcatgtttaa ttcatatgg gctgtgagat taattgaagg agagaaaaaa 900
gggtttttac cagagggatt tgttcaaagg gtagggagaca gaggatttgt tgtggagggg 960
tgggctccac aggcagaagt ttaggacat tcaagcaccg gtgggtttgt gagccattgt 1020
gggtggagtt ctattgcgga gatatgaag tttggggttc cagtaattgc catggccagg 1080
cafcttgate agccttfgaa tggtaagctg gcgcggagg ttggtgtggg catggaggtt 1140
gtgagagatg aaaatgggaa gtataagaga gaagggattg cagaggtaat aagaaaagtc 1200
gttctggaga aaagtgggga gttatcagg aggaagcaa gggaattgag tgagaaaatg 1260
aaagagatag gagagcaatt gattgatagg gcagtggagg agctagtaca aatttghtaag 1320
aagaagaag atgaacaata g
1341

```

```

<210> 57
<211> 446
<212> PRT
<213> 人参(Panax ginseng)

```

```

<220>
<221> misc_feature
<223> gGT29-5
<400> 57

```

```

Met Asp Asn Gln Lys Gly Arg Ile Ser Ile Val Met Leu Pro Phe Leu
1          5          10          15
Ala His Gly His Ile Ser Pro Phe Phe Glu Leu Ala Lys His Leu Ser
20          25          30
Lys Arg Asn Cys Asn Ile Phe Leu Cys Ser Thr Pro Ile Asn Leu Ser

```

Ser Ile Lys Asn Arg Ile Ser Asp Lys Asp Ser Ser Ala Ser Ile Lys  
 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445

Leu Val Glu Leu His Leu Pro Ser Ser Pro Asp Leu Pro Pro His Tyr  
 His Thr Thr Asn Gly Leu Pro Ser His Leu Met Val Pro Leu Arg Asn  
 Ala Phe Glu Thr Ala Ala Pro Thr Phe Ser Glu Ile Leu Lys Thr Leu  
 Asn Pro Asp Leu Leu Ile Tyr Asp Phe Asn Pro Ser Trp Ala Pro Glu  
 Ile Ala Ser Ser His Asn Ile Pro Ala Val Cys Phe Ile Ile Gly Gly  
 Ala Ala Ser Phe Ser Met Ser Leu His Ser Phe Lys Asn Pro Gly Glu  
 Lys Tyr Pro Phe Leu Asp Phe Asp Asp Asn Ser Asn Ile Thr Pro Glu  
 Pro Pro Ser Ala Asp Asn Met Lys Leu Leu Leu Asp Phe Met Thr Cys  
 Phe Glu Arg Ser Cys Asp Ile Ile Leu Ile Lys Ser Phe Arg Glu Leu  
 Glu Gly Lys Tyr Ile Asp Leu Leu Ser Thr Leu Ser Asp Lys Thr Leu  
 Val Pro Val Gly Pro Leu Val Gln Asp Pro Met Gly His Asn Glu Asp  
 Pro Lys Thr Glu Gln Ile Ile Asn Trp Leu Asp Lys Arg Ala Glu Ser  
 Thr Val Val Phe Val Cys Phe Gly Ser Glu Tyr Phe Leu Ser Asn Glu  
 Glu Leu Glu Glu Val Ala Ile Gly Leu Glu Ile Ser Met Val Asn Phe  
 Ile Trp Ala Val Arg Leu Ile Glu Gly Glu Lys Lys Gly Val Leu Pro  
 Glu Gly Phe Val Gln Arg Val Gly Asp Arg Gly Leu Val Val Glu Gly  
 Trp Ala Pro Gln Ala Arg Ile Leu Gly His Ser Ser Thr Gly Gly Phe  
 Val Ser His Cys Gly Trp Ser Ser Ile Ala Glu Ser Met Lys Phe Gly  
 Val Pro Val Ile Ala Met Ala Arg His Leu Asp Gln Pro Leu Asn Gly  
 Lys Leu Ala Ala Glu Val Gly Val Gly Met Glu Val Val Arg Asp Glu  
 Asn Gly Lys Tyr Lys Arg Glu Gly Ile Ala Glu Val Ile Arg Lys Val  
 Val Val Glu Lys Ser Gly Glu Val Ile Arg Arg Lys Ala Arg Glu Leu  
 Ser Glu Lys Met Lys Glu Ile Gly Glu Gln Leu Ile Asp Arg Ala Val  
 Glu Glu Leu Val Gln Ile Cys Lys Lys Lys Lys Asp Glu Gln

[0021]

- <210> 58
- <211> 1341
- <212> DNA
- <213> 人参(Panax ginseng)

- <220>
- <221> misc\_feature
- <223> gGT29-6
- <400> 58

```

atggataacc aaaagggtag aatcagtata gttatgetgc catittttagc ccatggccac 60
atttcccat tctttgaget agccaagcat ctctcaaaaa gaaattgtaa tatattctc 120
tgttctacce caatcaatct tagctccatc aagaacagaa tatctgataa ggattcctct 180
gcttctataa aactagtaga gcttcatctt ccattctccc ctgatcttcc tctcactac 240
cacaccacaa atggcctccc ttcccatctc atggctccc tcagaaacgc ctttgaaca 300
gcagccccca ccttctctga aatccttaaa acctaaace ctgatttget tatttatgat 360
ttcaatccct catgggcacc ggagatcget tegtctcaca atattccggc agtttatttc 420
ctaaccacgg cagcagccag ctcttccatt ggcctacatg ctttcaaaaa cceaggtgaa 480
aaatacccat ttccngattt ttatgataac agtaataata cccctgaacc acctcttgea 540
gataaacatga agctacttca tgattttatc gcttgtttcg aacgatcttg cgacattatt 600
ttgattaaga gttttataga actagaaggg aaatatatcg atttgettte cactttatct 660
gataaaactt tggttcctgt tggccacac gtccaagatc ctatgggcca taatgaagat 720
ccaaaaacag agcagattat aaactggctt gacaaaaggg ctgaatctac agtgggtgtt 780
gctctgctttg gaagttagta tttctctcc aatgaggaat tggaaagaag agcaattggg 840
ctagagatta gcatgggtaa tttcatatgg gctgtgagat taattgaagg agagaaaaaa 900
ggggttttac cagagggatt ttctcaaagg gtaggagaca gaggattggt tgtggagggg 960
tggactccac agcaagaat tttaggacat tcaagcaecg tggggtttgt gagccattgt 1020
gggtggagtt ctattggcga gagtatgaag tttggggttc cagtaattgc catggccagg 1080
catcttgatc agcctttgaa tggtaagctg gggccggagg ttgggtgggg catggaggtt 1140
gtgagagatg aaatgggaa gtataagaga gaagatattg caggggtaat aagaaaagtc 1200
gtgtgtggaga aaagtgggga ggttatcagg aggaaagcaa gggaaattgag tgagaaaatg 1260
    
```

aaagagatag gagagcaatt gattgatagg gcagtgagg agctagtaca aatttgaag 1320  
aagaagaag atgaacaata g 1341

<210> 59  
<211> 446  
<212> PRT  
<213> 人参(Panax ginseng)

<220>  
<221> misc\_feature  
<223> gGT29-6  
<400> 59

Met Asp Asn Gln Lys Gly Arg Ile Ser Ile Val Met Leu Pro Phe Leu  
1 5 10 15  
Ala His Gly His Ile Ser Pro Phe Phe Glu Leu Ala Lys His Leu Ser  
20 25 30  
Lys Arg Asn Cys Asn Ile Phe Leu Cys Ser Thr Pro Ile Asn Leu Ser  
35 40 45  
Ser Ile Lys Asn Arg Ile Ser Asp Lys Asp Ser Ser Ala Ser Ile Lys  
50 55 60  
Leu Val Glu Leu His Leu Pro Ser Ser Pro Asp Leu Pro Pro His Tyr  
65 70 75 80  
His Thr Thr Asn Gly Leu Pro Ser His Leu Met Val Pro Leu Arg Asn  
85 90 95  
Ala Phe Glu Thr Ala Ala Pro Thr Phe Ser Glu Ile Leu Lys Thr Leu  
100 105 110  
Asn Pro Asp Leu Leu Ile Tyr Asp Phe Asn Pro Ser Trp Ala Pro Glu  
115 120 125  
Ile Ala Ser Ser His Asn Ile Pro Ala Val Tyr Phe Leu Thr Thr Ala  
130 135 140  
Ala Ala Ser Ser Ser Ile Gly Leu His Ala Phe Lys Asn Pro Gly Glu  
145 150 155 160  
Lys Tyr Pro Phe Pro Asp Phe Tyr Asp Asn Ser Asn Asn Thr Pro Glu  
165 170 175  
Pro Pro Ser Ala Asp Asn Met Lys Leu Leu His Asp Phe Ile Ala Cys  
180 185 190  
Phe Glu Arg Ser Cys Asp Ile Ile Leu Ile Lys Ser Phe Ile Glu Leu  
195 200 205  
Glu Gly Lys Tyr Ile Asp Leu Leu Ser Thr Leu Ser Asp Lys Thr Leu  
210 215 220  
Val Pro Val Gly Pro Leu Val Gln Asp Pro Met Gly His Asn Glu Asp  
225 230 235 240  
Pro Lys Thr Glu Gln Ile Ile Asn Trp Leu Asp Lys Arg Ala Glu Ser  
245 250 255  
Thr Val Val Phe Val Cys Phe Gly Ser Glu Tyr Phe Leu Ser Asn Glu  
260 265 270  
Glu Leu Glu Glu Val Ala Ile Gly Leu Glu Ile Ser Met Val Asn Phe  
275 280 285  
Ile Trp Ala Val Arg Leu Ile Glu Gly Glu Lys Lys Gly Val Leu Pro  
290 295 300  
Glu Gly Phe Val Gln Arg Val Gly Asp Arg Gly Leu Val Val Glu Gly  
305 310 315 320  
Trp Ala Pro Gln Ala Arg Ile Leu Gly His Ser Ser Thr Gly Gly Phe  
325 330 335  
Val Ser His Cys Gly Trp Ser Ser Ile Ala Glu Ser Met Lys Phe Gly  
340 345 350  
Val Pro Val Ile Ala Met Ala Arg His Leu Asp Gln Pro Leu Asn Gly  
355 360 365  
Lys Leu Ala Ala Glu Val Gly Val Gly Met Glu Val Val Arg Asp Glu  
370 375 380  
Asn Gly Lys Tyr Lys Arg Glu Asp Ile Ala Gly Val Ile Arg Lys Val  
385 390 395 400  
Val Val Glu Lys Ser Gly Glu Val Ile Arg Arg Lys Ala Arg Glu Leu  
405 410 415  
Ser Glu Lys Met Lys Glu Ile Gly Glu Gln Leu Ile Asp Arg Ala Val  
420 425 430  
Glu Glu Leu Val Gln Ile Cys Lys Lys Lys Asp Glu Gln  
435 440 445

[0022]

<210> 60  
<211> 1341  
<212> DNA  
<213> 人参(Panax ginseng)

<220>  
<221> misc\_feature  
<223> gGT29-7  
<400> 60

atggataacc aanaaggtag aatcagtata gcgttgetac catttttagc ceatggtcac 60  
atatctccct tctttgagct agccaacaa ctgcacaaa gaaattgeaa tgttttctc 120  
tgtttacc ccattcaatct tagctccatc aagaacagag tatctgataa ggattctct 180

```

gcttcataaa aactagtaga gcttcattt ccatcttccc ctgatcttcc tctcactac 240
cacaccacaa atggcctccc tcccactctc atgatccccc tcagaaaagc cttfgataca 300
gcaggcccaca ccttctctga aatccttaaa accttaaaccc ctgatttget tatttatgat 360
ttcaatccct catgggcacc ggagatcget tegtctcaca atattccggc agtttgtttc 420
ataattggtg gagcagcctc ctcttccatg agcctacata gtttcaaaaa cccaggtgaa 480
aaatacccat ttctagattt tgatgataac agtaatatta cccctgaacc acctcageca 540
gataaacatga agctattaat taattttatg acttgtttcc aacgatcttg cgacattatt 600
ttgalttaaga gttttlagaga actagaaggg aaataltttg atlllllltcc cactllatct 660
gataaaaactt lggttcctgt lggtcacctc gttcaagatc ctatgggcca taatgaagal 720
ccaaaaacag agcagtttat aaactggctt gacaaaaggg ctgaatctac agtgggtgtt 780
gtctgctttg gaagtgagtg tttctctccc aatgaggaat tggagaagt agcgattggg 840
ctagagatta gcatggtfaa tttcatatgg gctgtgagat taattgaagg agagaaaaaa 900
ggggttttac cagaggggtt tgttcaagg gtaggagaca gaggattggt tgtggaggag 960
tgggctccac aggcagaat tttaggacat tcaagcaccg gtgggtttgt gagecattgt 1020
gggtggaaatt ctattacgga gagtatgaag tttgggttcc cagtaattgc catggccagg 1080
cattttgate agcctttgaa tggtaagctg cgggcccagg ttgggtgtgg catggaggtt 1140
gtgagagatg aaaaatgggaa gtataagaga gaagggattg cagaggtaat aagaaaagtc 1200
gttggtgaga aaagtgggga ggttatcagg aggaaagcaa ggaattgag tgagaaaatg 1260
aaagagaaag gagagcaaga gattgatagg gtagtggagg agctagtaca aatttgaag 1320
aagaagaaag atgaacaata g

```

```

<210> 61
<211> 446
<212> PRT
<213> 人参(Panax ginseng)

```

```

<220>
<221> misc feature
<223> gGT29-7
<400> 61

```

[0023]

```

Met Asp Asn Gln Lys Gly Arg Ile Ser Ile Ala Leu Leu Pro Phe Leu
1 5 10 15
Ala His Gly His Ile Ser Pro Phe Phe Glu Leu Ala Lys Gln Leu Ala
20 25 30
Lys Arg Asn Cys Asn Val Phe Leu Cys Ser Thr Pro Ile Asn Leu Ser
35 40 45
Ser Ile Lys Asn Arg Val Ser Asp Lys Asp Ser Ser Ala Ser Ile Lys
50 55 60
Leu Val Glu Leu His Leu Pro Ser Ser Pro Asp Leu Pro Pro His Tyr
65 70 75 80
His Thr Thr Asn Gly Leu Pro Ser His Leu Met Ile Pro Leu Arg Asn
85 90 95
Ala Phe Asp Thr Ala Gly Pro Thr Phe Ser Glu Ile Leu Lys Thr Leu
100 105 110
Asn Pro Asp Leu Leu Ile Tyr Asp Phe Asn Pro Ser Trp Ala Pro Glu
115 120 125
Ile Ala Ser Ser His Asn Ile Pro Ala Val Cys Phe Ile Ile Gly Gly
130 135 140
Ala Ala Ser Ser Ser Met Ser Leu His Ser Phe Lys Asn Pro Gly Glu
145 150 155 160
Lys Tyr Pro Phe Leu Asp Phe Asp Asp Asn Ser Asn Ile Thr Pro Glu
165 170 175
Pro Pro Ser Ala Asp Asn Met Lys Leu Leu Ile Asn Phe Met Thr Cys
180 185 190
Phe Glu Arg Ser Cys Asp Ile Ile Leu Ile Lys Ser Phe Arg Glu Leu
195 200 205
Glu Gly Lys Tyr Phe Asp Phe Ser Thr Leu Ser Asp Lys Thr Leu
210 215 220
Val Pro Val Gly Pro Leu Val Gln Asp Pro Met Gly His Asn Glu Asp
225 230 235 240
Pro Lys Thr Glu Gln Phe Ile Asn Trp Leu Asp Lys Arg Ala Glu Ser
245 250 255
Thr Val Val Phe Val Cys Phe Gly Ser Glu Cys Phe Leu Ser Asn Glu
260 265 270
Glu Leu Glu Glu Val Ala Ile Gly Leu Glu Ile Ser Met Val Asn Phe
275 280 285
Ile Trp Ala Val Arg Leu Ile Glu Gly Glu Lys Lys Gly Val Leu Pro
290 295 300
Glu Gly Phe Val Gln Arg Val Gly Asp Arg Gly Leu Val Val Glu Glu
305 310 315 320
Trp Ala Pro Gln Ala Arg Ile Leu Gly His Ser Ser Thr Gly Gly Phe
325 330 335
Val Ser His Cys Gly Trp Asn Ser Ile Thr Glu Ser Met Lys Phe Gly
340 345 350
Val Pro Val Ile Ala Met Ala Arg His Phe Asp Gln Pro Leu Asn Gly
355 360 365
Lys Leu Ala Ala Glu Val Gly Val Gly Met Glu Val Val Arg Asp Glu
370 375 380
Asn Gly Lys Tyr Lys Arg Glu Gly Ile Ala Glu Val Ile Arg Lys Val
385 390 395 400
Val Val Glu Lys Ser Gly Glu Val Ile Arg Arg Lys Ala Arg Glu Leu
405 410 415

```

	Ser Glu Lys Met Lys Glu Lys Gly Glu Gln Glu Ile Asp Arg Val Val	
	420 425 430	
	Glu Glu Leu Val Gln Ile Cys Lys Lys Lys Lys Asp Glu Gln	
	435 440 445	
	<210> 62	
	<211> 26	
	<212> DNA	
	<213> 人工序列	
	<400> 62	
	acageaagag agagacacag agttca	26
	<210> 63	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<400> 63	
	gttcaaagcc caacctaacg gca	23
	<210> 64	
	<211> 25	
	<212> DNA	
	<213> 人工序列	
	<400> 64	
	atgaaattat acagagaggg agaga	25
	<210> 65	
	<211> 25	
	<212> DNA	
	<213> 人工序列	
	<400> 65	
	tggttttctt cacaacacaa agtac	25
[0024]	<210> 66	
	<211> 43	
	<212> DNA	
	<213> 人工序列	
	<400> 66	
	ctgggtgccgc gcggcagcat ggataaccaa aaggtagaa tea	43
	<210> 67	
	<211> 43	
	<212> DNA	
	<213> 人工序列	
	<400> 67	
	ctgggtgccgc gcggcagcat ggataaccaa aaaggtagaa tea	43
	<210> 68	
	<211> 43	
	<212> DNA	
	<213> 人工序列	
	<400> 68	
	tgcggccgca agettgcttt gttcatcttt cttcttctta caa	43
	<210> 69	
	<211> 30	
	<212> DNA	
	<213> 人工序列	
	<400> 69	
	gctgccgcgc ggcaccagge cgtgtgtgtg	30
	<210> 70	
	<211> 30	
	<212> DNA	
	<213> 人工序列	
	<400> 70	
	tccgtcgaca agettgccgc egcaactcgag	30

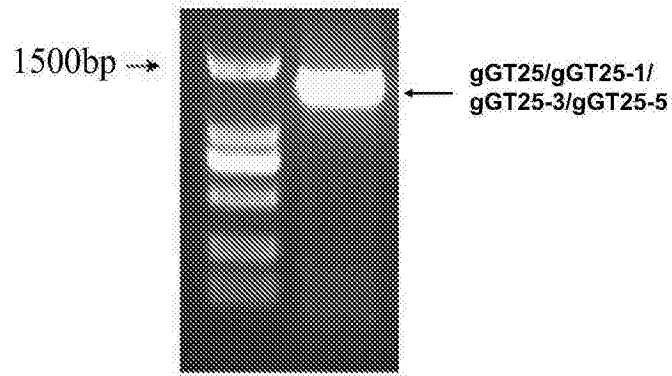


图1

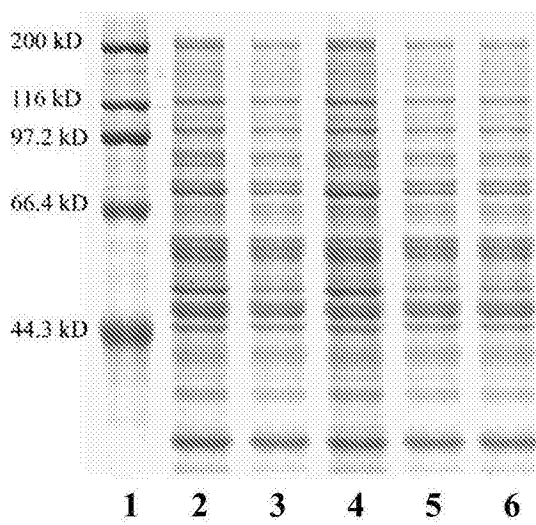


图2

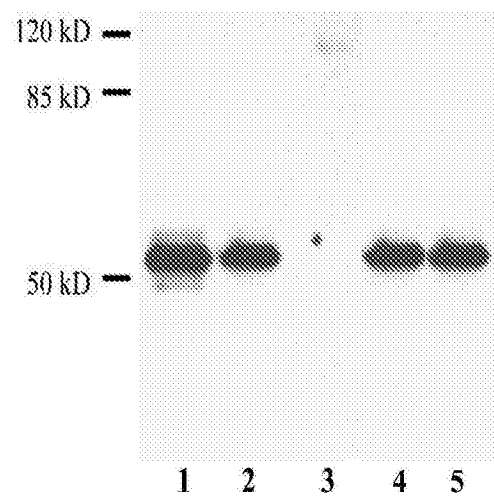


图3

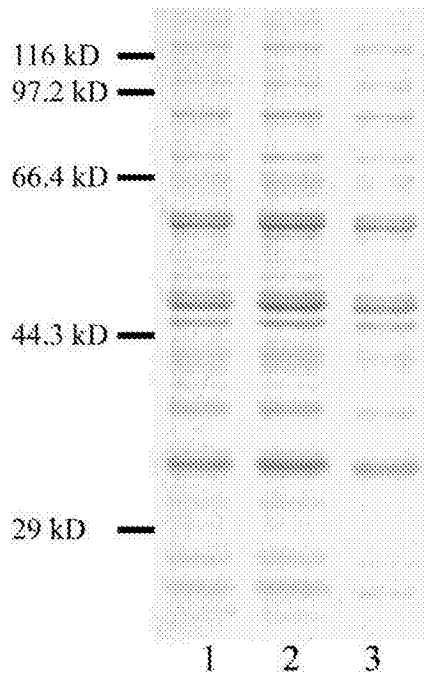


图4



图5

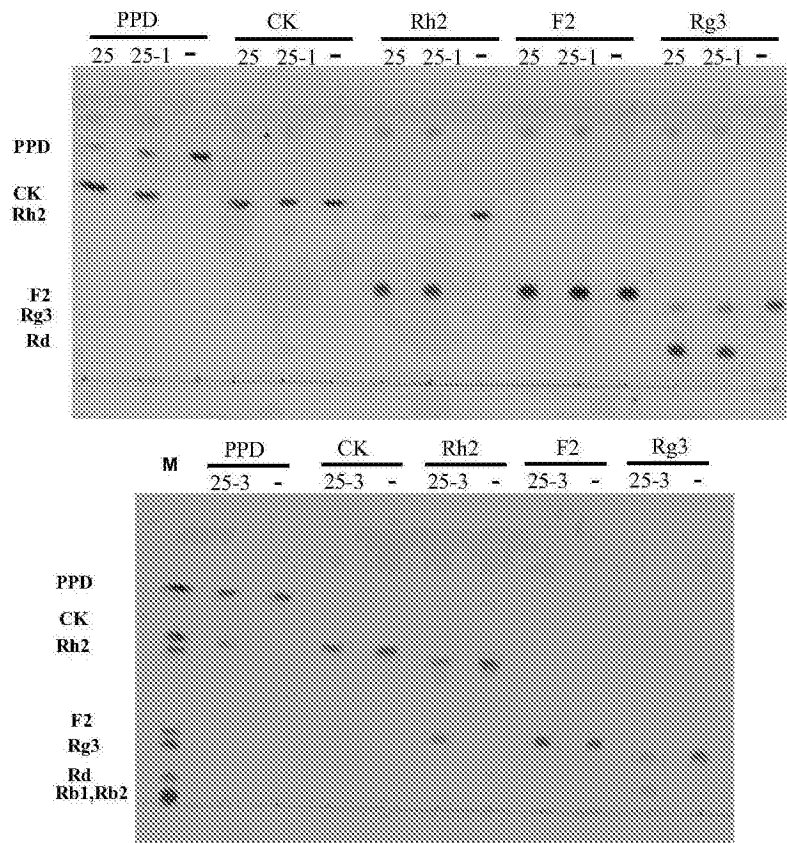


图6

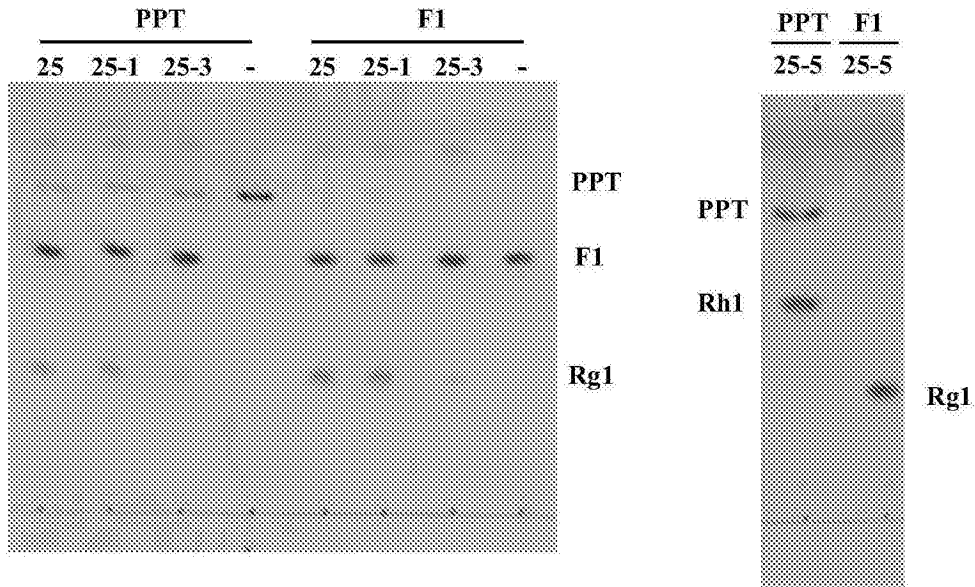


图7

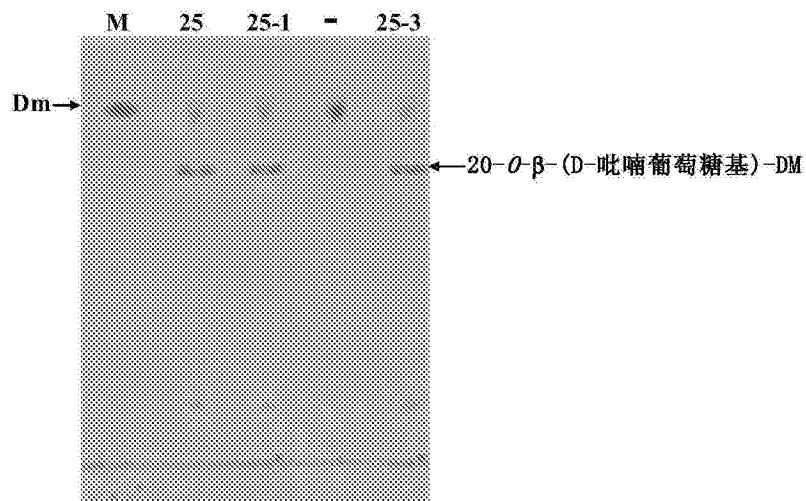


图8

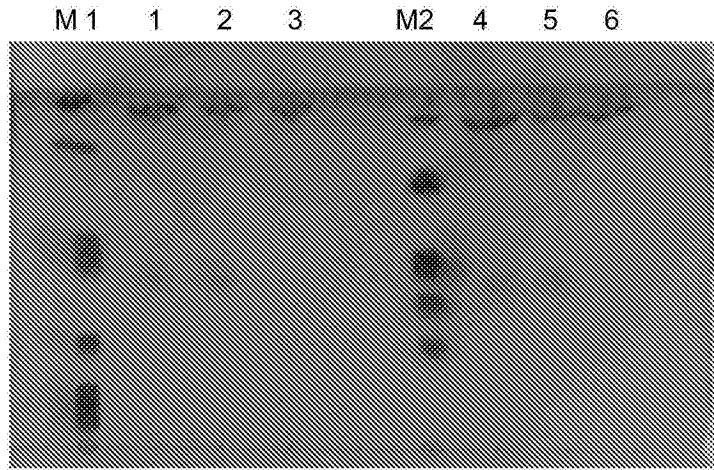


图9

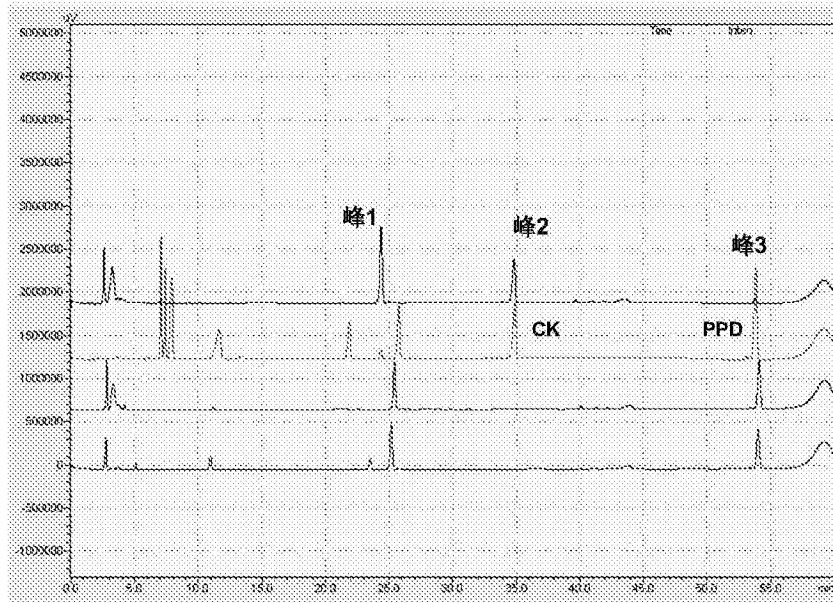


图10

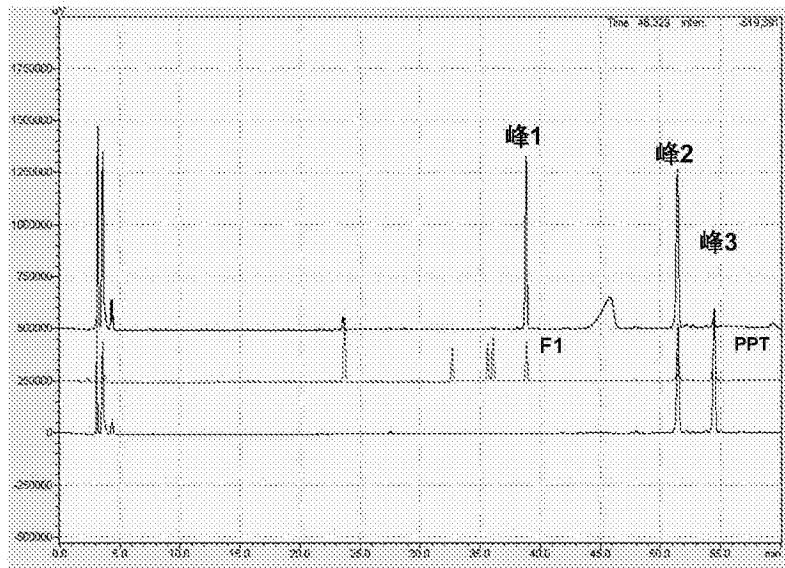


图11

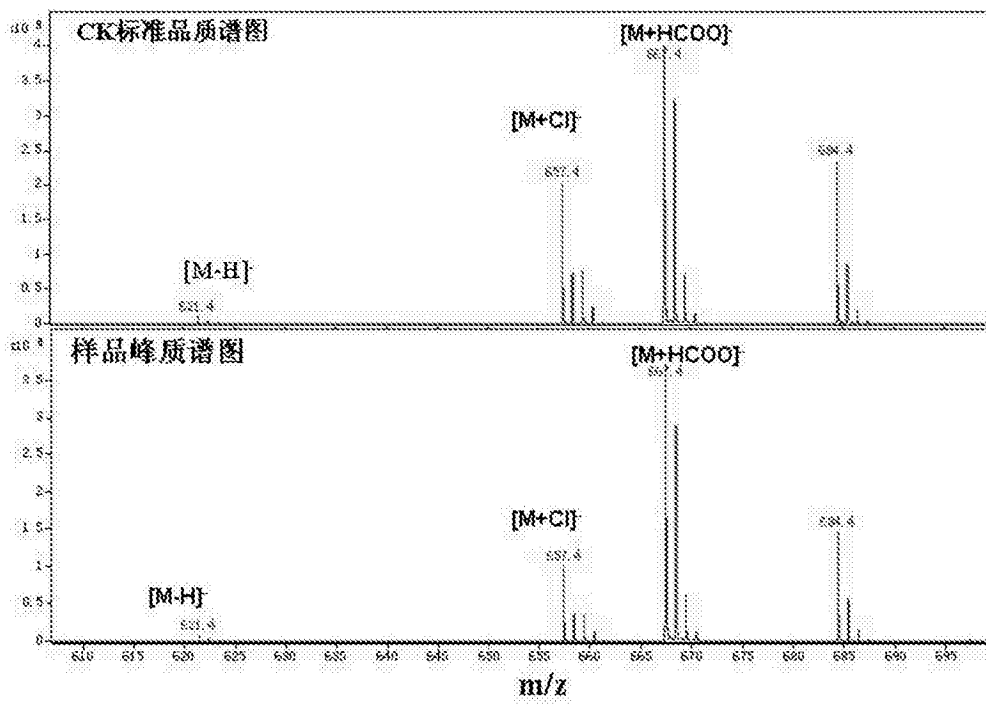


图12

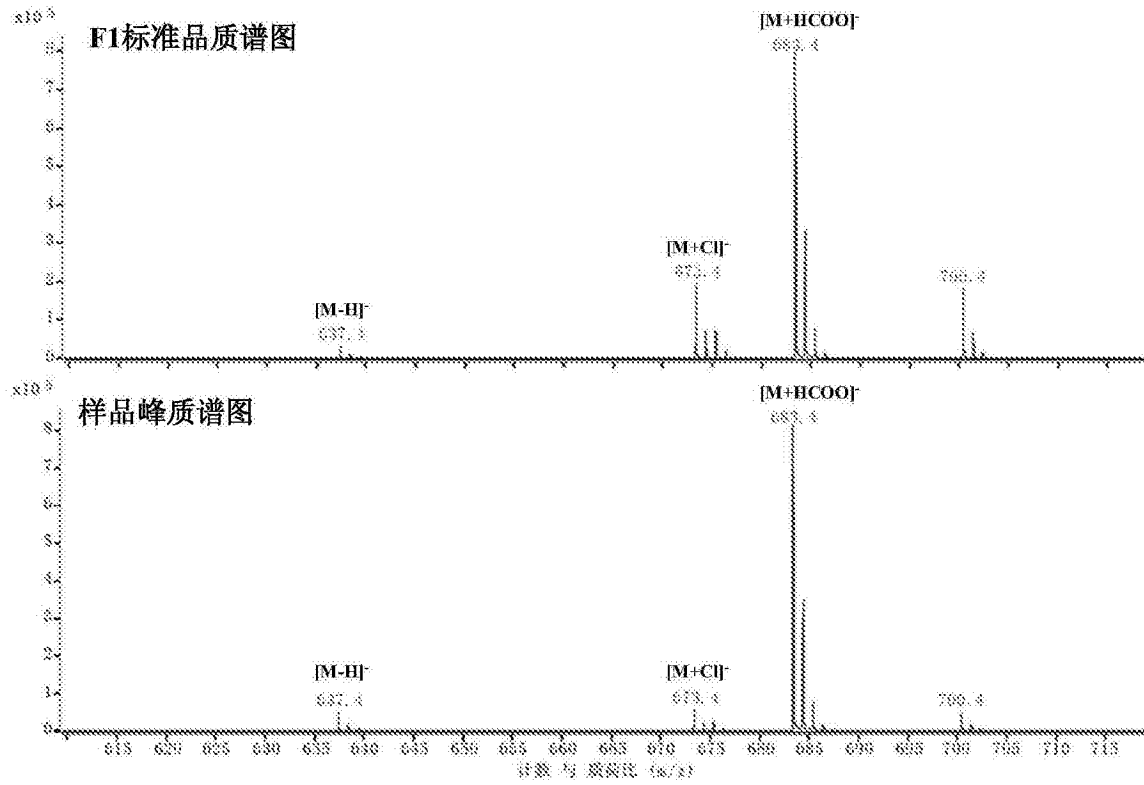


图13

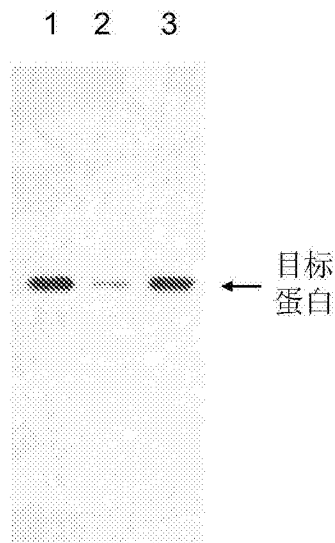


图14

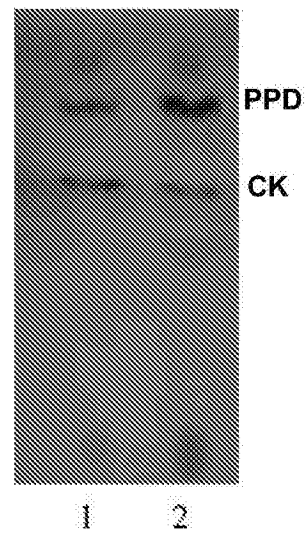


图15

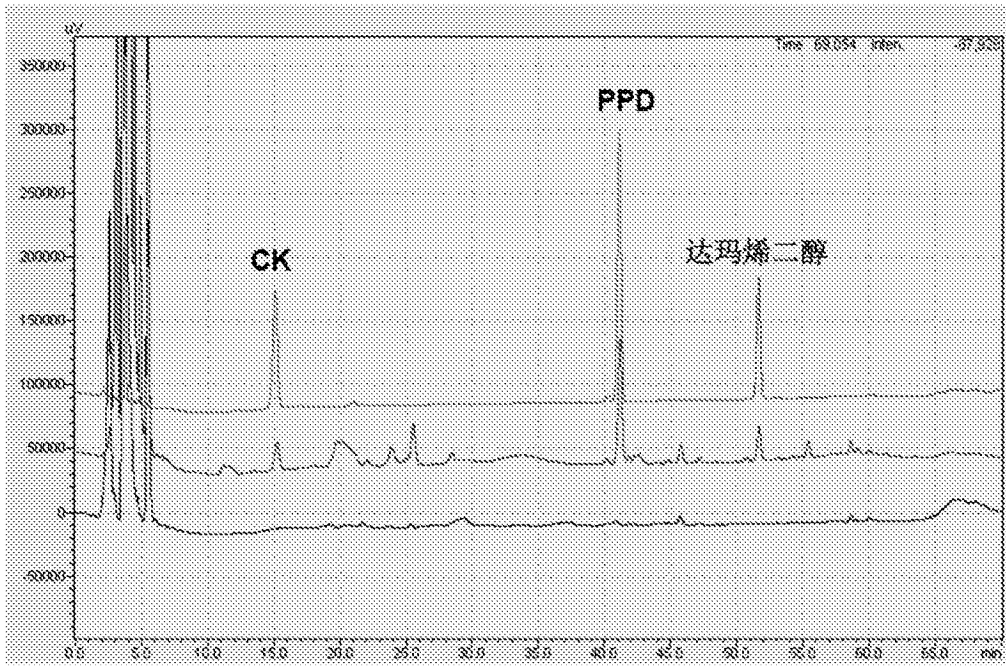


图16

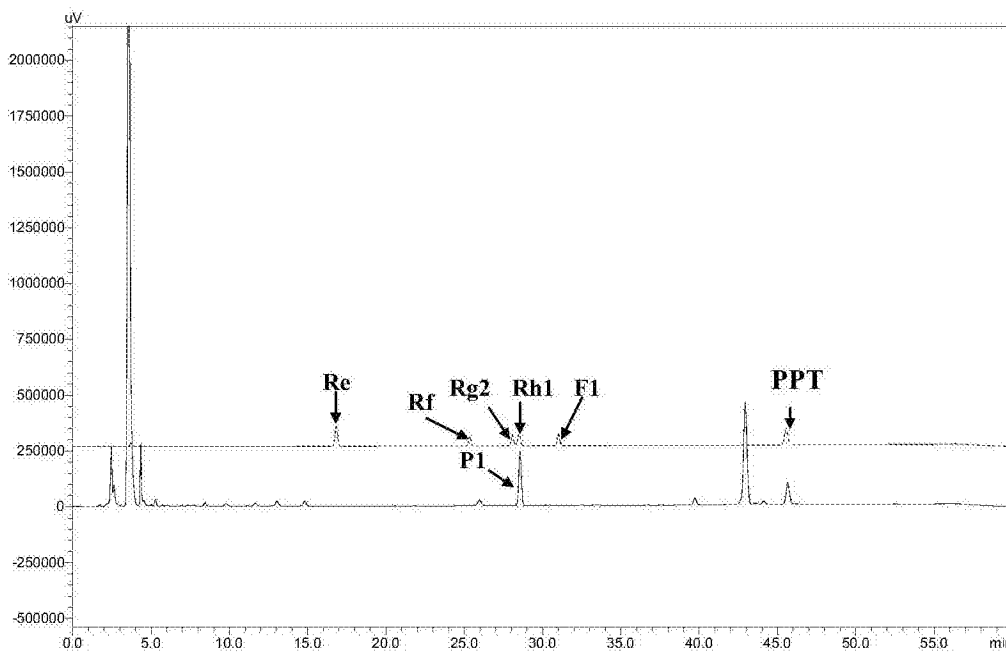


图17

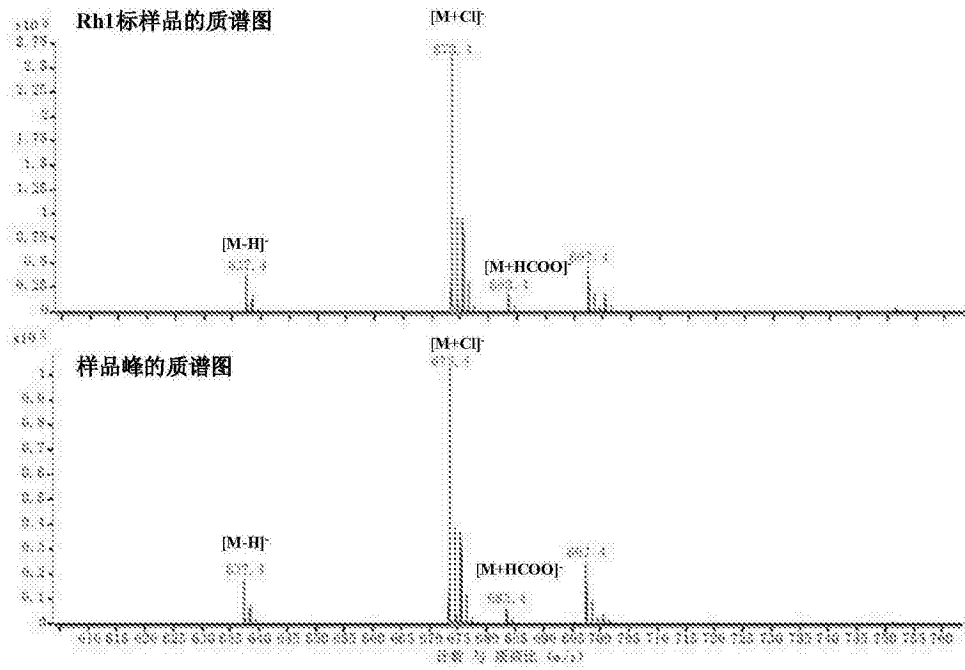


图18

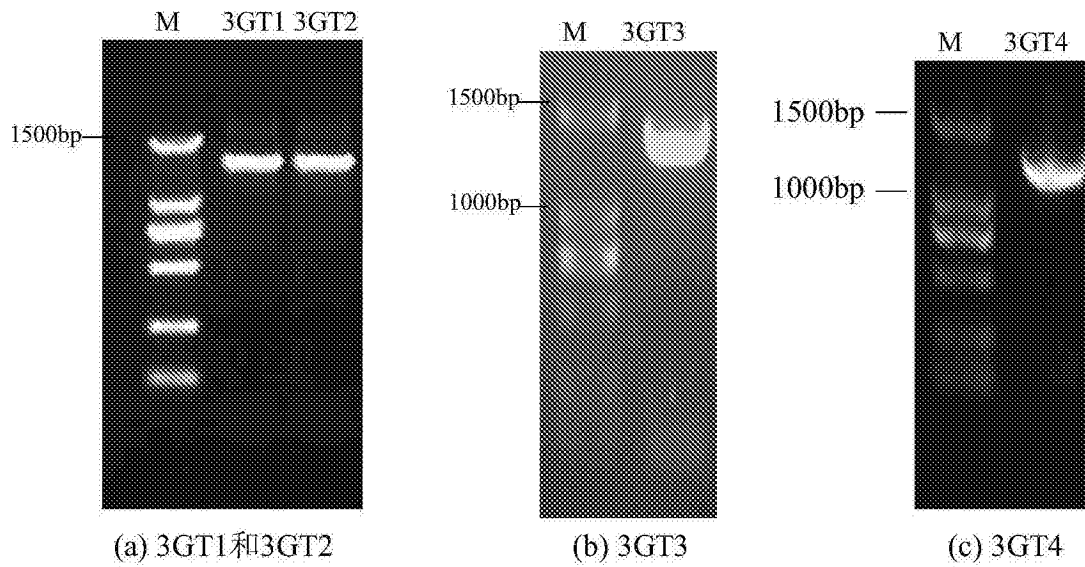


图19

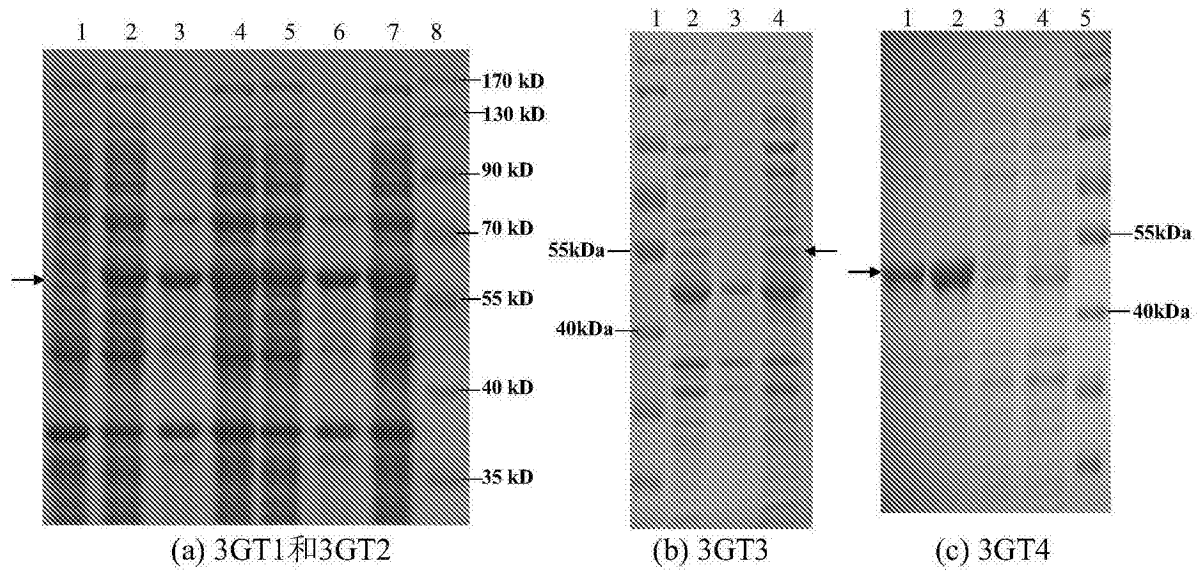


图20

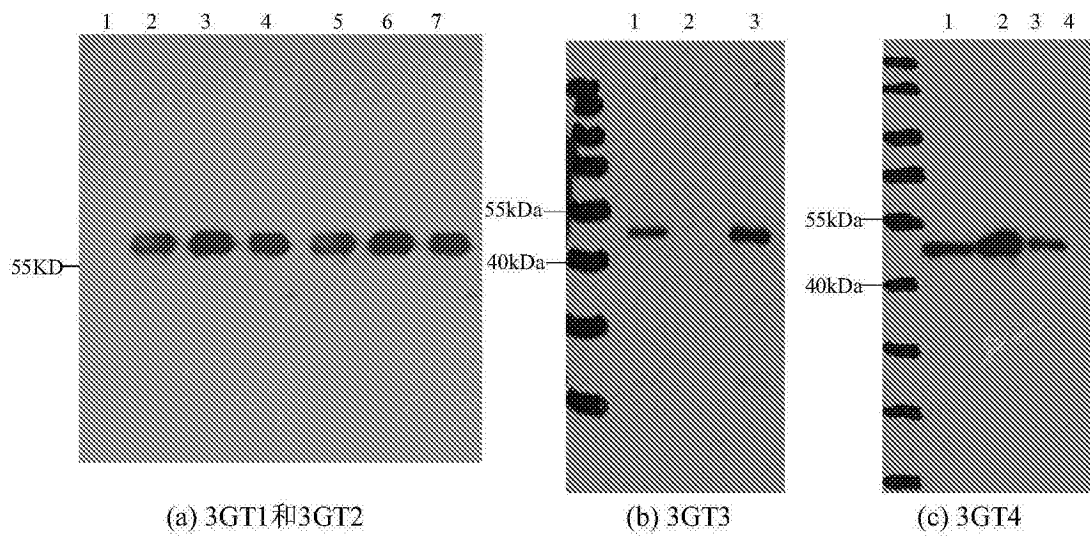


图21

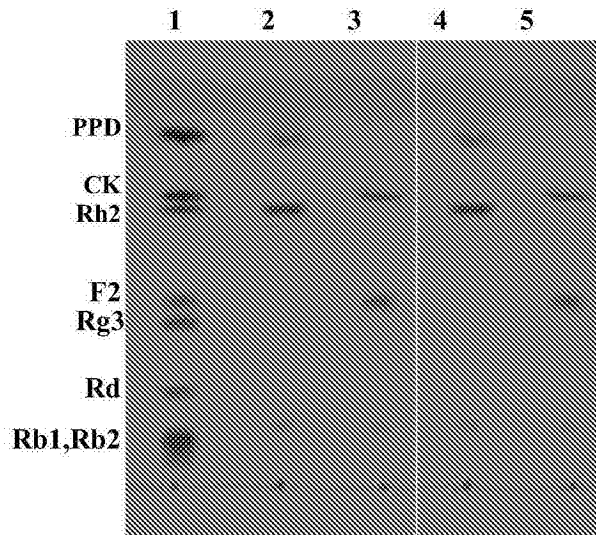


图22

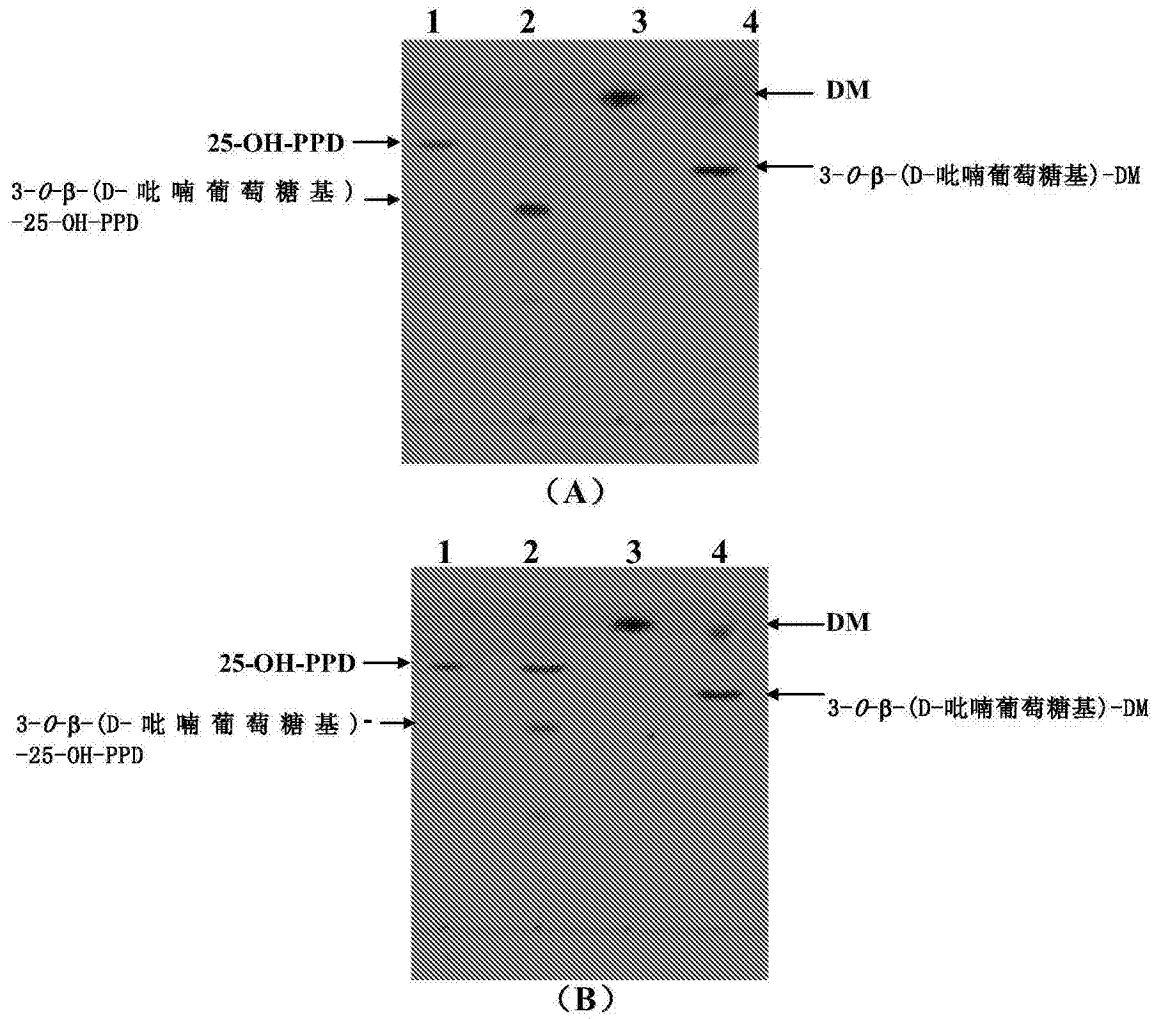


图23

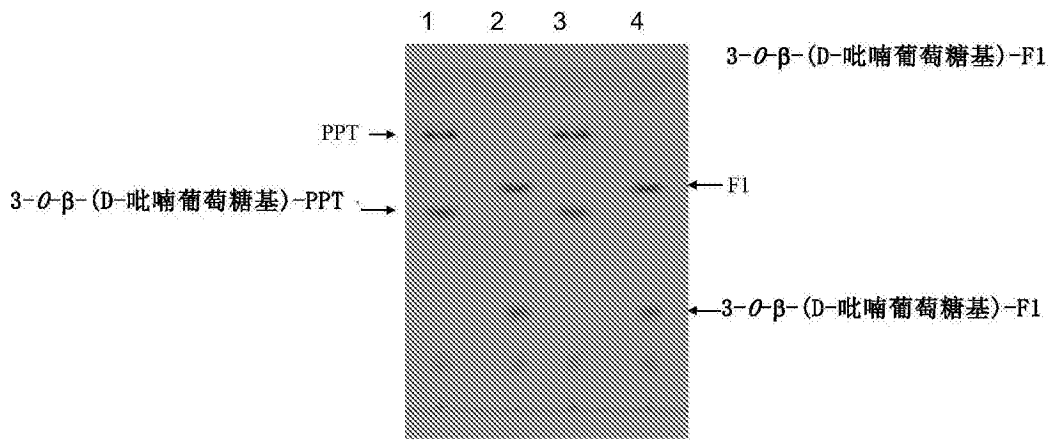


图24

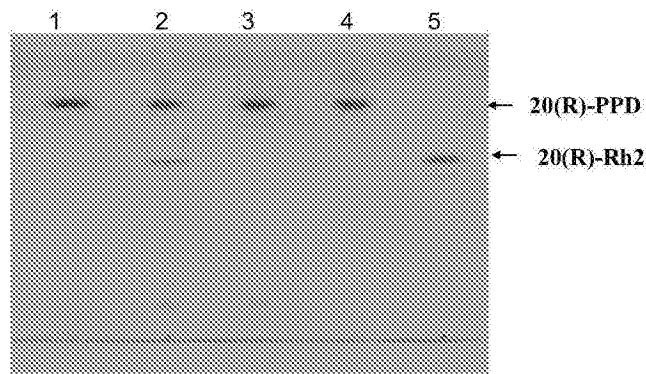


图25

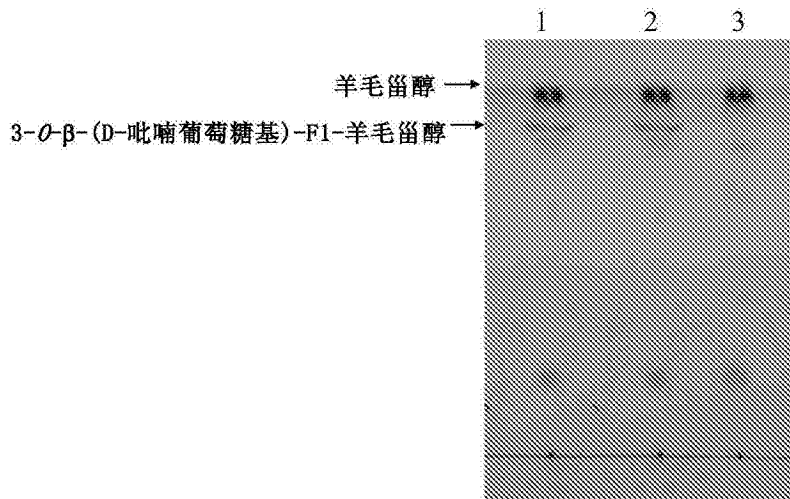


图26

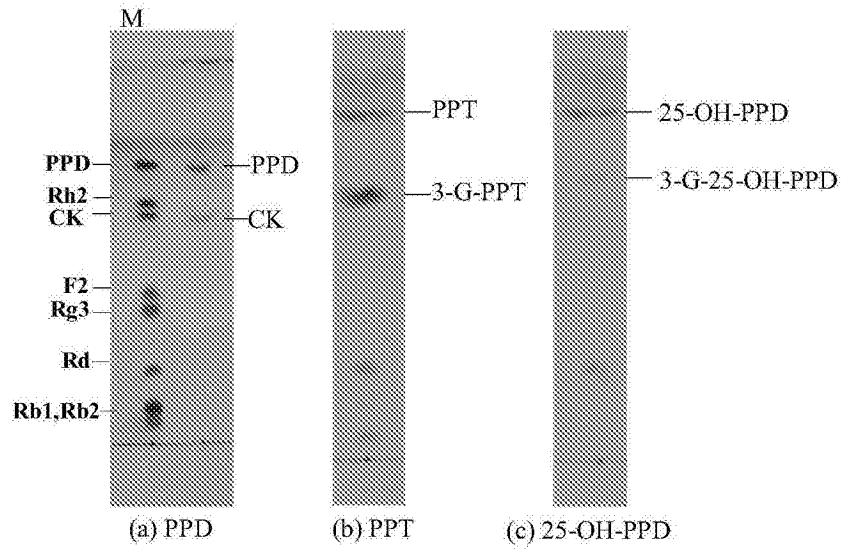


图27

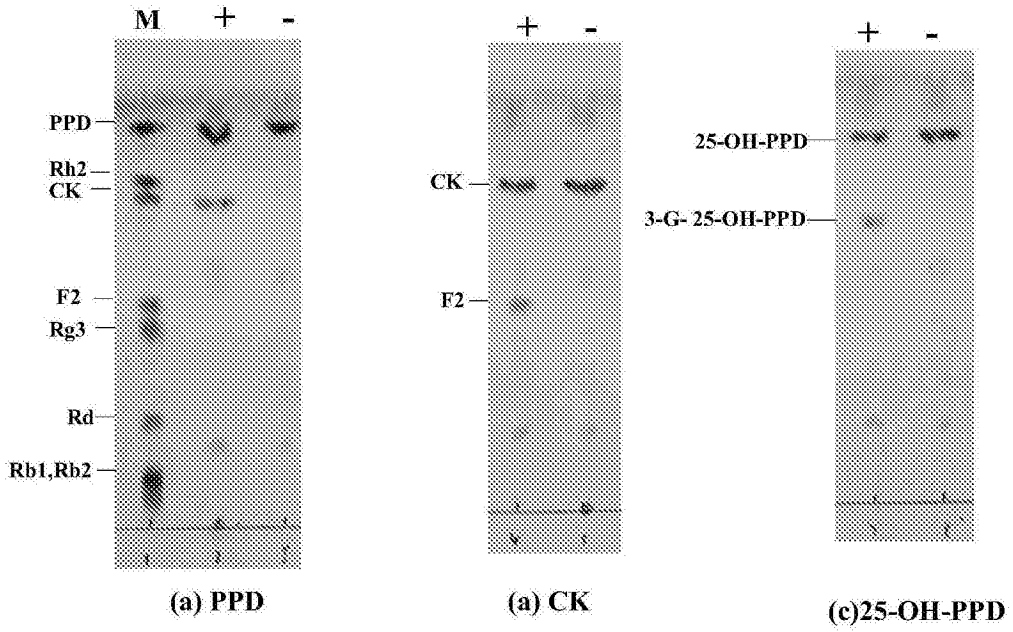


图28

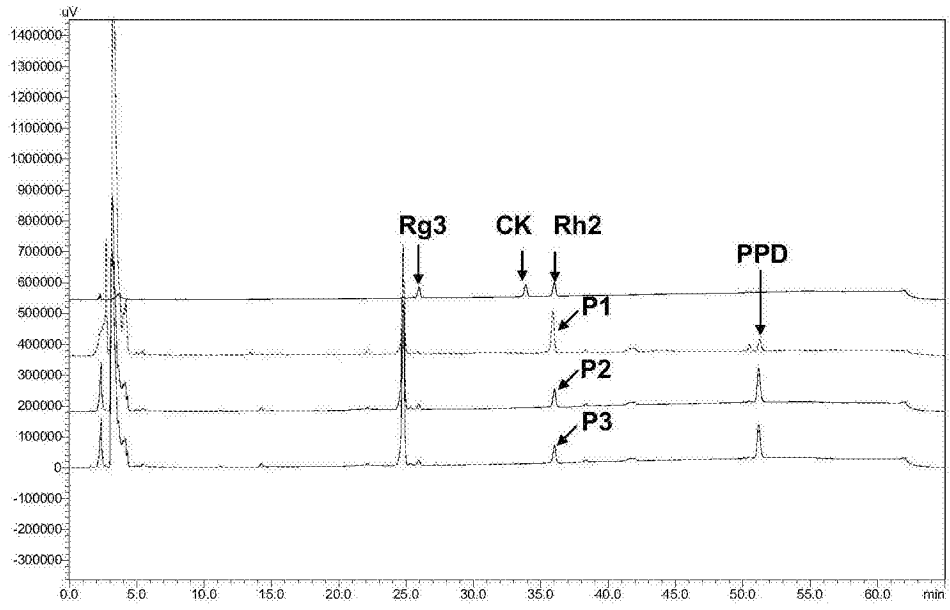


图29

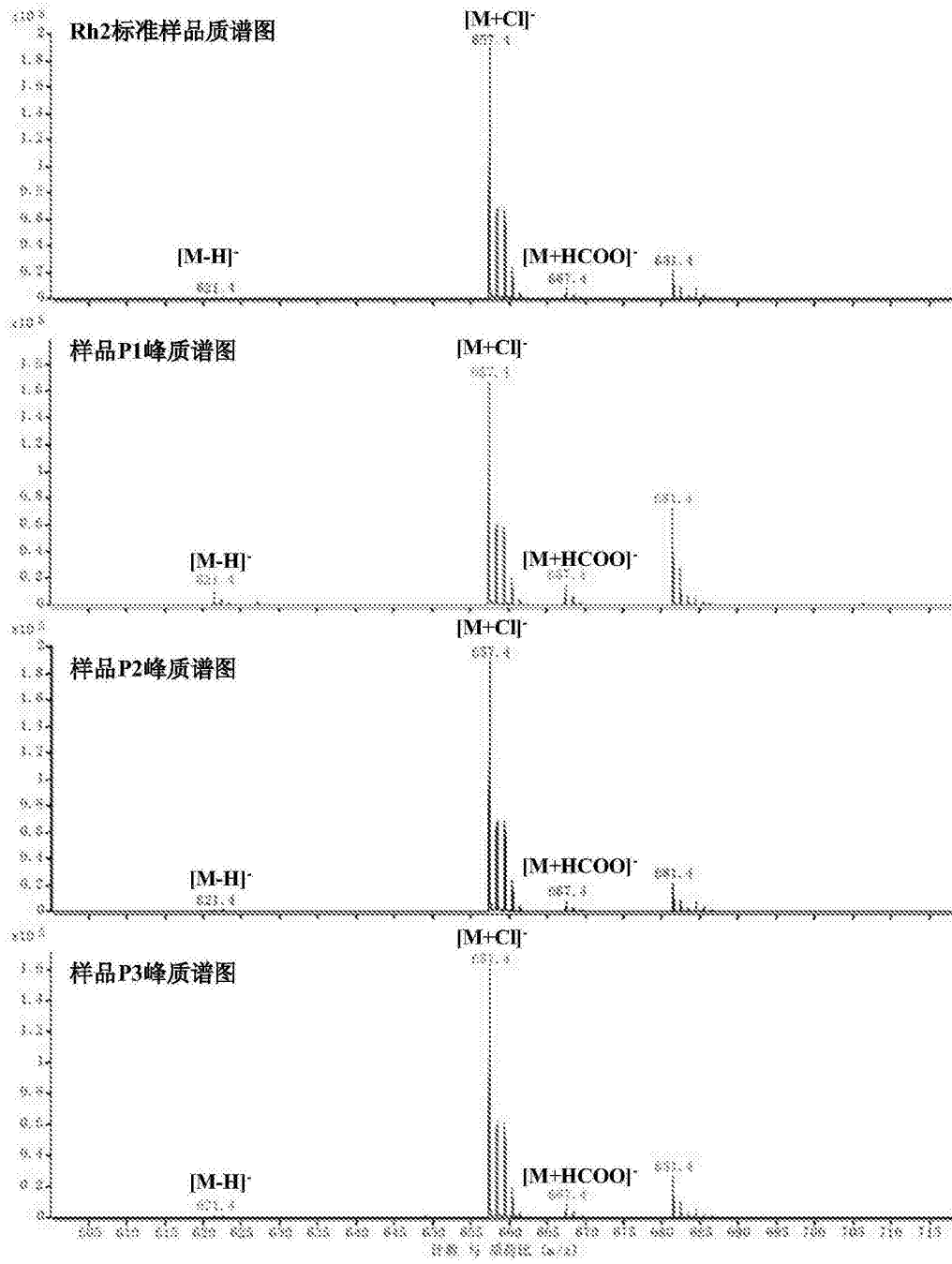


图30

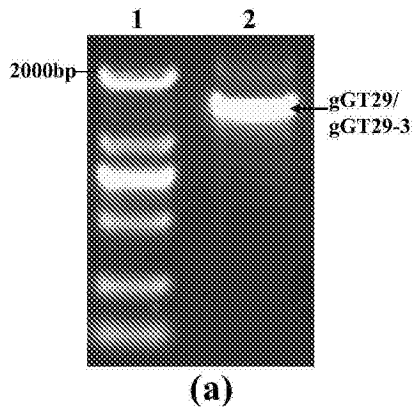


图31

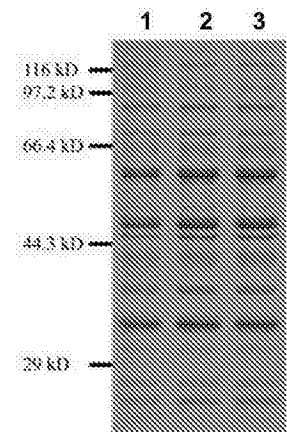
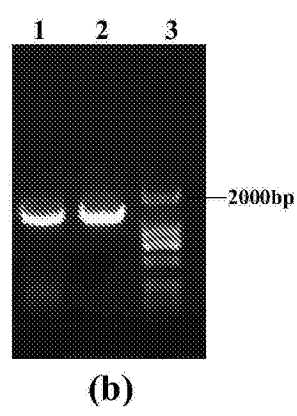


图32

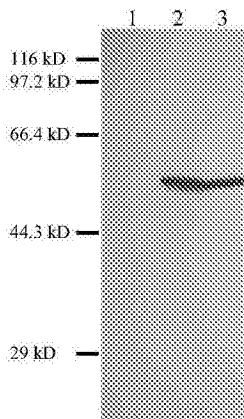


图33

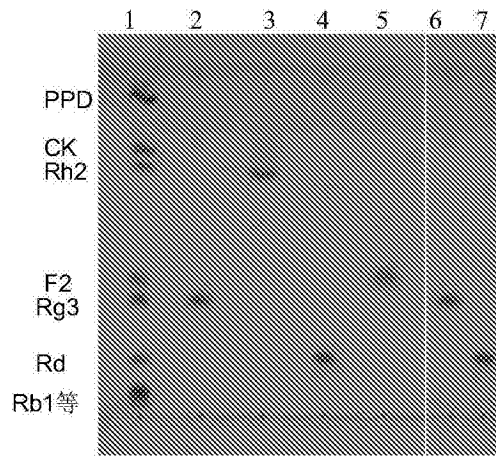


图34

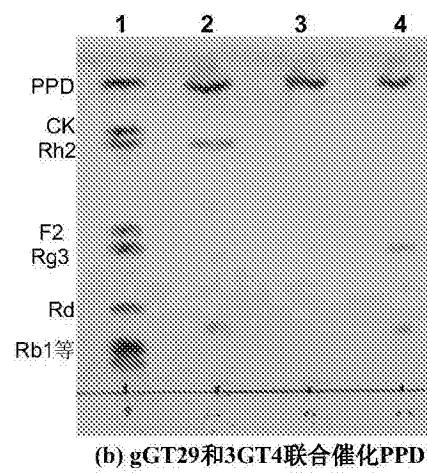
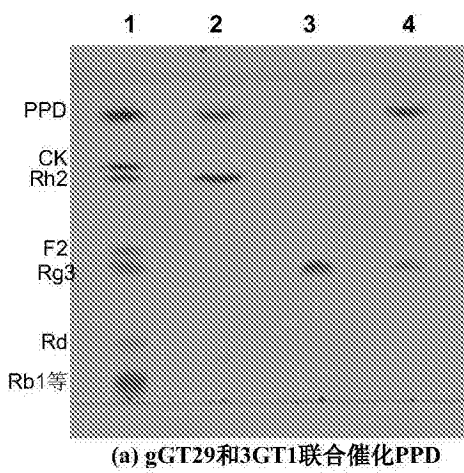


图35

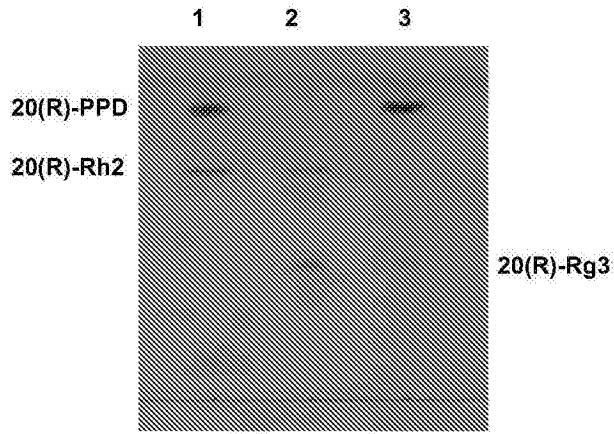


图36

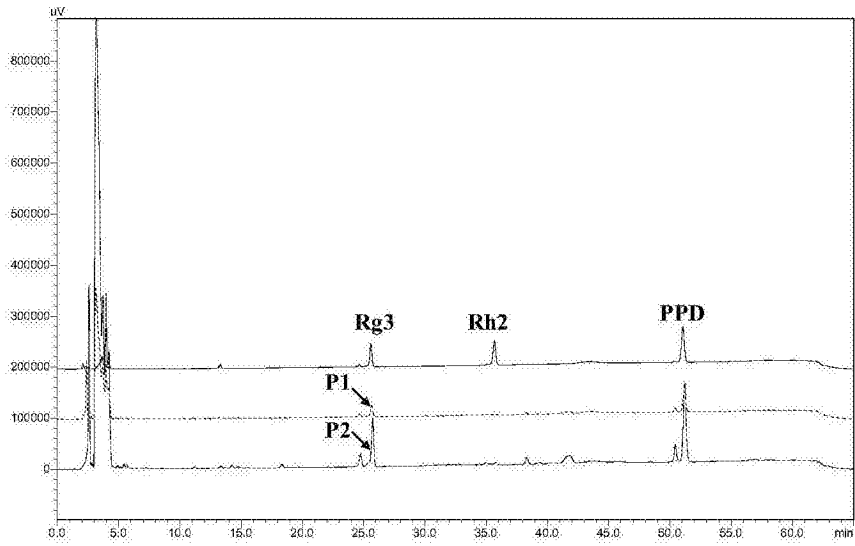


图37

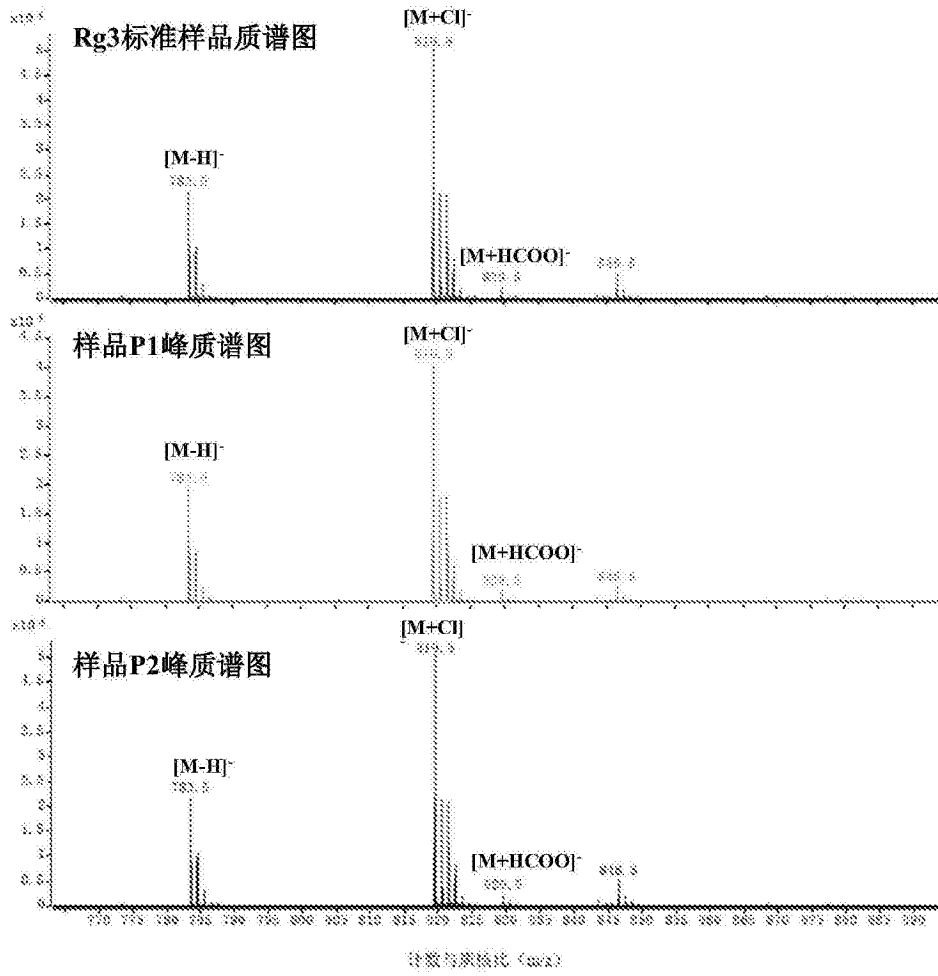


图38

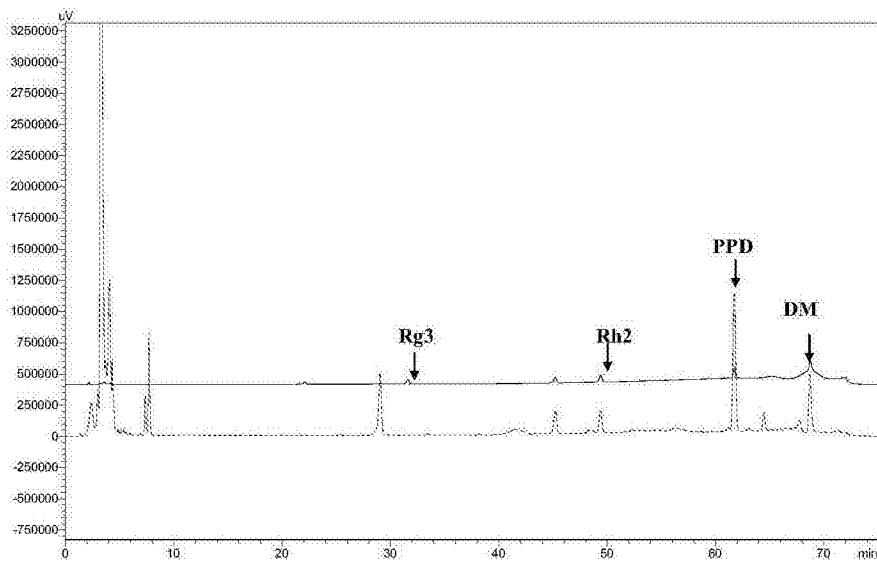


图39

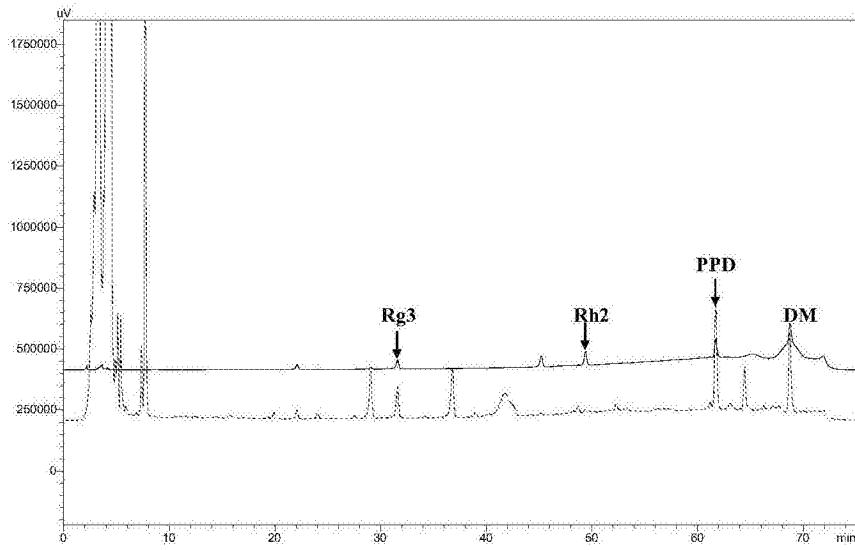


图40

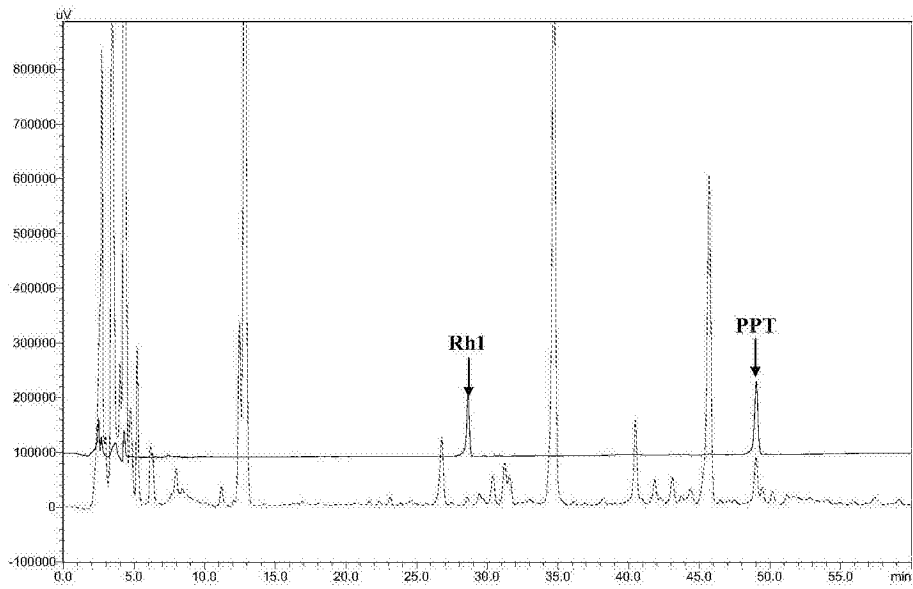


图41

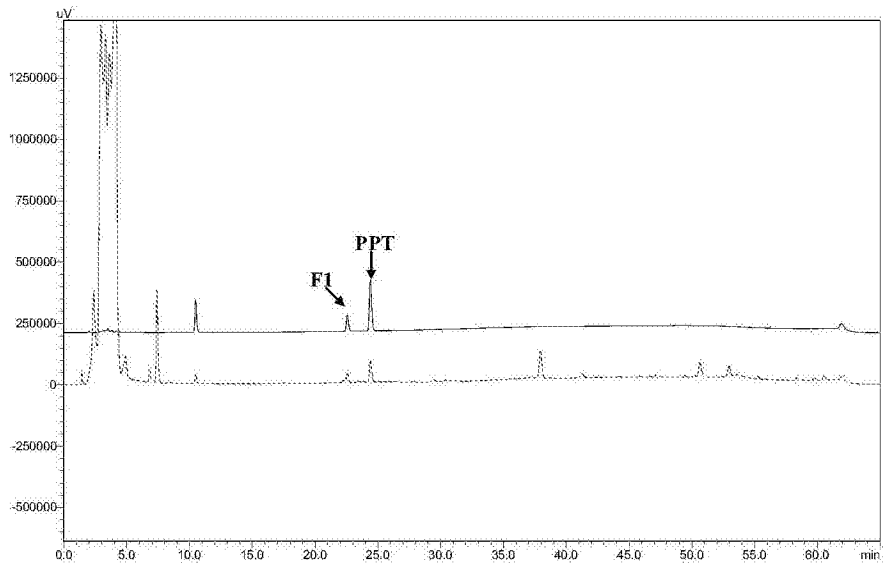


图42

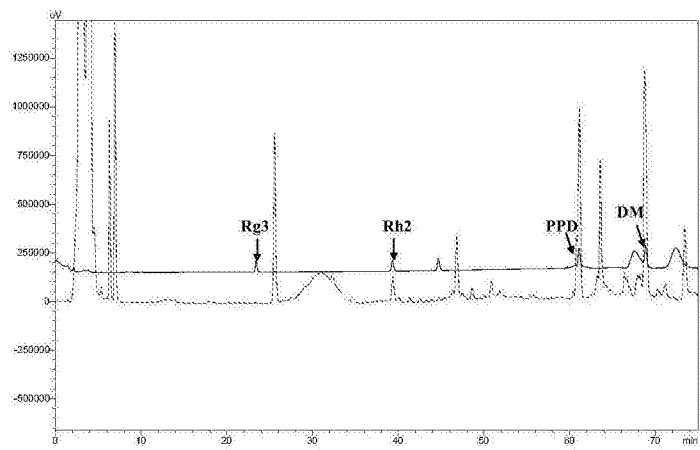


图43

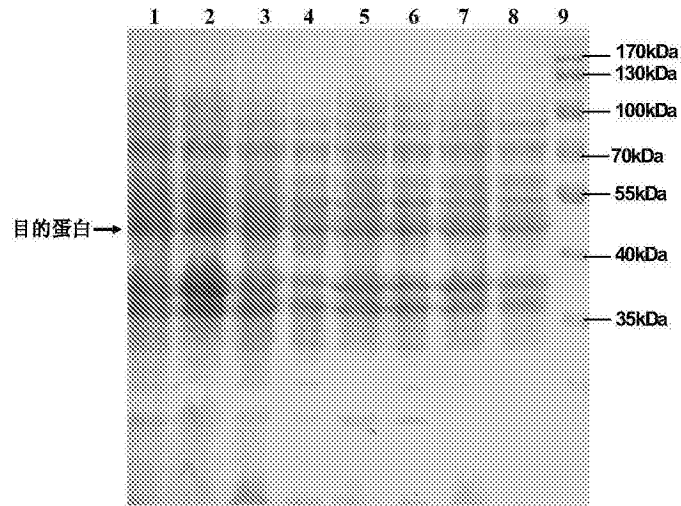


图44

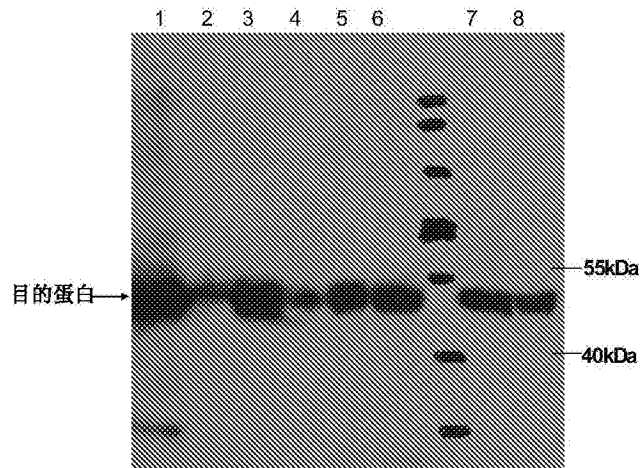


图45

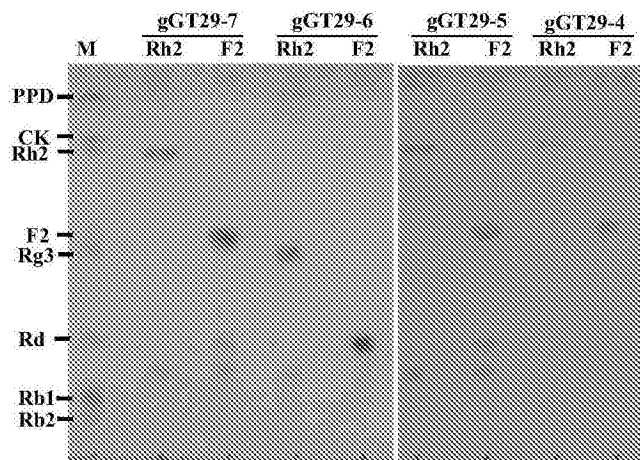


图46