



## (12) 发明专利

(10) 授权公告号 CN 103958673 B

(45) 授权公告日 2021.10.26

(21) 申请号 201280053787.7

F·马丁

(22) 申请日 2012.08.31

(74) 专利代理机构 中国贸促会专利商标事务所  
有限公司 11038(65) 同一申请的已公布的文献号  
申请公布号 CN 103958673 A

代理人 刘晓东

(43) 申请公布日 2014.07.30

(51) Int.CI.

C12N 9/10 (2006.01)

(30) 优先权数据

C12N 15/82 (2006.01)

11179882.3 2011.09.02 EP

A01H 5/12 (2018.01)

(85) PCT国际申请进入国家阶段日  
2014.04.30

A01H 6/82 (2018.01)

(86) PCT国际申请的申请数据

(56) 对比文件

PCT/EP2012/003662 2012.08.31

M.Kroumova A.等.“PATHWAYS FOR  
SYNTHESIS, AND POSSIBILITIES FOR GENETIC  
MODIFICATION OF SUGAR ESTER ACYL GROUPS  
PRODUCED BY TRICHOMES OF SOLANACEOUS  
SPECIES”.《General and Applied Plant  
Physiology》.2009,第35卷(第3-4期),摘要.

(87) PCT国际申请的公布数据

审查员 薛娜

W02013/029799 EN 2013.03.07

权利要求书2页 说明书71页 附图7页

(73) 专利权人 菲利普莫里斯产品有限公司  
地址 瑞士,纳沙泰尔

(72) 发明人 N·巴卡尔 G·N·拜德勒

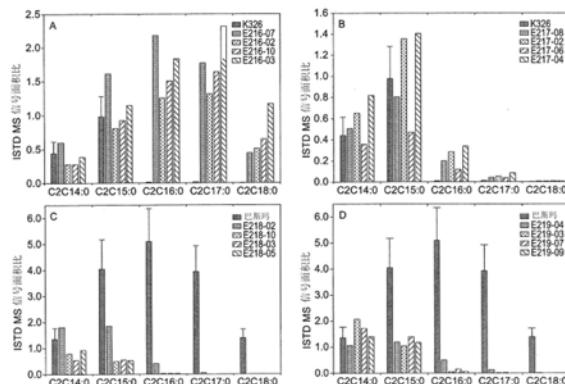
M·P·布朗克 S·格普费特

## (54) 发明名称

来自红花烟草的异丙基苹果酸合酶及其方  
法和用途

## (57) 摘要

B  
本发明涉及突变、非天然存在的或转基因植  
物细胞,包含:(i)至少一种多核苷酸,其包含、由  
或基本上由编码异丙基苹果酸合酶并且与SEQ  
ID NO:1或SEQ ID NO:10或SEQ ID NO:12或SEQ  
ID NO:14具有至少60%序列同一性的序列所组  
成;或者(ii)由所述多核苷酸编码的多肽;或者  
(iii)与SEQ ID NO:2或SEQ ID NO:11或SEQ ID  
NO:13或SEQ ID NO:15具有至少60%序列同一性的  
多肽;或者(iv)含有所述多核苷酸序列的构建  
体、载体或表达载体,任选地,其中所述构建体、  
载体或表达载体额外含有启动子,所述启动子包  
含、由或基本上由如SEQ ID NO:8所示的序列、或  
与其具有至少约60%同一性的其变体、或毛状体  
启动子所组成。



1. 一种用于在烟草植物的至少部分中调节蔗糖酯的量或类型的方法,所述方法包括以下步骤:

(i) 提高异丙基苹果酸合酶基因在烟草植物中的表达或活性,其中所述异丙基苹果酸合酶基因在重组构建体中与毛状体特异性启动子可操作连接,所述异丙基苹果酸合酶基因含有下述多核苷酸,所述多核苷酸由如SEQ ID NO:1所示的序列所组成;或者编码由所述多核苷酸编码的多肽;或者其中异丙基苹果酸合酶基因编码由SEQ ID NO:2所示的多肽;以及

(ii) 鉴定根据(i)所述异丙基苹果酸合酶基因的表达或活性已经被提高的植物,其中所述烟草植物中的一种或更多种蔗糖酯的量或类型与对照植物相比已经改变,所述对照植物中异丙基苹果酸合酶基因的表达或活性未得到调节,以及其中获自步骤(i)的烟草植物的视觉外观与对照植物基本相同。

2. 根据权利要求1所述的方法,还包括在获自步骤(i)的植物的至少部分中测量一种或更多种蔗糖酯的量或确定其类型。

3. 一种烟草制品,其包含:

(a) 包含构建体、载体或表达载体的烟草植物细胞,其中所述构建体、载体或表达载体包含:

(i) 与毛状体特异性启动子可操作性连接的至少一种多核苷酸,所述多核苷酸编码异丙基苹果酸合酶、由如SEQ ID NO:1所示的序列组成;或者

(ii) 由(i)所述多核苷酸编码的多肽;或

(iii) 与毛状体特异性启动子可操作性连接的至少一种多核苷酸,所述多核苷酸由编码SEQ ID NO:2所示的多肽的序列组成;或者

(b) 通过权利要求1或2的所述方法获得的或能够获得的植物的部分;或者

(c) 烟草植物的部分,其中序列为SEQ ID NO:2所示的异丙基苹果酸合酶蛋白的活性或编码所述异丙基苹果酸合酶蛋白的基因的表达是得到提高的,以及所述烟草植物的至少部分在一种或更多种蔗糖酯的组成上与对照烟草植物相比具有改变,所述对照烟草植物中异丙基苹果酸合酶的表达或活性未得到调节,以及其中所述烟草植物的视觉外观与对照烟草植物基本相同;或者

(d) 植物材料,所述植物材料包含来自(b)或(c)的植物的部分的细胞或组织。

4. 根据权利要求3所述的烟草制品,其中所述构建体、载体或表达载体还含有由SEQ ID NO:8所示的序列组成的启动子。

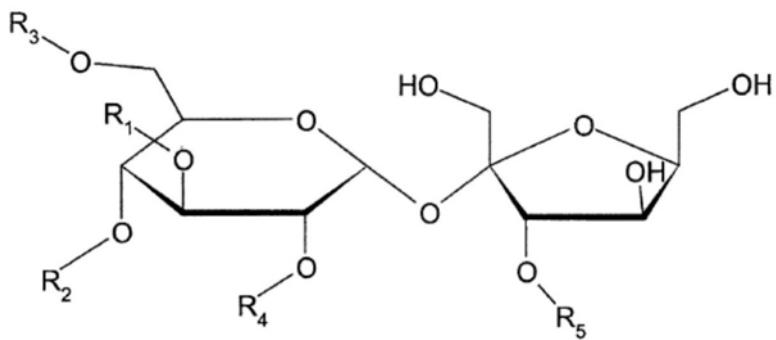
5. 一种用于产生含有一种或更多种蔗糖酯的组合物的方法,所述方法包括以下步骤:

(i) 提供权利要求3或4所限定的烟草植物的至少部分、植物材料、或者烟草制品;以及

(ii) 从其中提取蔗糖酯。

6. 根据权利要求5所述的方法,还包括分离或纯化提取出的蔗糖酯。

7. 根据权利要求1、2、5或6任一项所述的方法,其中蔗糖酯的一种或更多种具有下式所示的结构:



并且其中R3是乙酰基或氢;R1、R2和R4中的一个或更多个包含具有6个碳的酰基链;以及R5是乙酰基或氢。

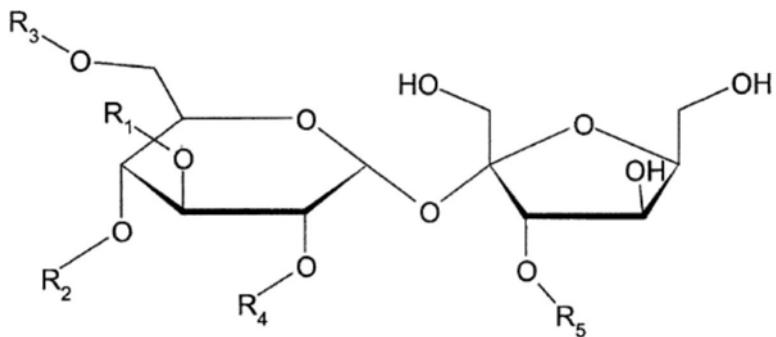
8.根据权利要求7所述的方法,其中,

R3是乙酰基;和/或

R1、R2和R4中的一个或更多个含有 $\beta$ -甲基戊酰基;和/或

R5是氢。

9.根据权利要求3或4所述的烟草制品,其中蔗糖酯的一种或更多种具有下式所示的结构:



并且其中R3是乙酰基或氢;R1、R2和R4中的一个或更多个包含具有6个碳的酰基链;以及R5是乙酰基或氢。

10.根据权利要求9所述的烟草制品,其中,

R3是乙酰基;和/或

R1、R2和R4中的一个或更多个含有 $\beta$ -甲基戊酰基;和/或

R5是氢。

11.一种用于调节烟草或烟草制品风味的方法,所述方法包括:

(i)将来自权利要求3或4所限定的烟草植物的植物部分,或者烟草植物材料加入至烟草或烟草制品中;或者

(ii)将含有通过权利要求5至8任一项所述的方法所获得的或能够获得的蔗糖酯的组合物加入至烟草或烟草制品中。

12.根据权利要求11所述的方法,其中所述植物的部分包含叶。

## 来自红花烟草的异丙基苹果酸合酶及其方法和用途

### 技术领域

[0001] 本发明公开来自红花烟草 (*Nicotiana tabacum*) 的异丙基苹果酸合酶基因以及其变体、同系物和片段。特别地，描述了为变更植物如烟草植物的蔗糖酯组合物而改变这种基因表达或由其编码的蛋白的活性。

### 背景技术

[0002] 蔗糖酯因在植物中具有杀虫性质而广为人知，其显示出针对软体昆虫的毒性以及抗生活性，所述软体昆虫包括蚜虫、螨、梨木虱、和粉虱。蔗糖酯也视作风味前体，在植物的生命期中积累在叶子表面上。其是可以在烤制材料和吸烟材料中鉴定出的稳定化合物。通过加热从蔗糖酯中释放出酯化的小羧酸。这些小羧酸是非常强效的风味分子，其被视为与烟草的东方风味部分地有关。蔗糖酯产生在植物的腺毛状体细胞中。所述腺毛状体细胞也是合成其它叶渗出物如蛋白质 (phylloplanin (无对应中译文)) 和双萜 (cembreneoid (无对应中译文) 和 labdenoid (无对应中译文)) 的地方。通过两种截然不同的代谢阻断产生所述蔗糖和小羧酸，其被酯化至蔗糖以生成蔗糖酯。

[0003] 烟草品种在其叶表面上存在的蔗糖酯的数量和质量上不同。作为一般规则，烤烟、白肋和马里兰品种积累低量的具有达5碳链长酰基的蔗糖酯，然而大多数东方品种和许多雪茄烟草型积累高量的具有达6碳链长酰基的蔗糖酯，以及在较小程度上可观察到具有7碳链长酰基的酯。在烟草品种中生产的蔗糖酯中观察到的二分 (dichotomy) 与单显性基因座相关，所述单显性基因座对其功能等位基因称为BMVSE (含有蔗糖酯的 $\beta$ -甲基戊酰基)、以及对其非功能等位基因称为bmvs。BMVSE基因组基因座定位在烟草的染色体A上。

[0004] Kroumova和Wagner (2009) General and Applied Plant Physiology 35, 3-4, p95-110描述了尝试使用双链干扰RNA通过反向遗传学方法来敲低异丙基苹果酸合酶的表达从而改变在各种植物中的蔗糖酯酰基含量。在这一研究中，所述来自潘那利番茄 (*Solanum pennellii*) 的异丙基苹果酸合酶基因用于从cDNA中扩增部分基因。继而将其以正义和反义方向引入至双链干扰RNA构建体中，再通过农杆菌 (*Agrobacterium*) 转化至红花烟草 T.I.1068、粘性烟草 (*Nicotiana glutinosa* cv.) 24a以及潘那利番茄。与未转化的对照相比，使用红花烟草植物获得的结果显示出通过减少的 $\beta$ -甲基戊酰基和增加的2-甲基丁酰化作用而在蔗糖酯的酰基丰度上的改变。然而，获得的红花烟草植物患有萎黄病并且一些具有卷曲叶所以其表型受损。同样地，转化的粘性烟草和潘那利番茄表型受损。作者推断异丙基苹果酸合酶是重要的酶以及其受损可能导致不健康的植物。

[0005] 在本领域中需要这样的植物，其中蔗糖酯组合物得到调节同时将对植物的不期望影响降至最低。本发明的一个目的是满足这种需要。

### 发明内容

[0006] 本发明至少在部分上基于出人意料的发现：调节来自红花烟草的异丙基苹果酸合酶基因的活性或表达导致在蔗糖酯组合物中的改变，其对所述植物的整体代谢具有更少的

不期望作用。所述作用的大多数集中在植物的毛状体和次级代谢上。因此，调节来自红花烟草的异丙基苹果酸合酶基因的活性或表达基本上不导致相对于对照植物的植物视觉外观改变。这是有利的，因为植物可用于各种产品的商业化生产，其中视觉外观上的改变将是产业上不可接受的、或者可导致产量不可接受的减少。有利地，从烟草中生成的烟雾的风味谱可被改变，因此可建立新的风味谱。此外，其中蔗糖酯谱已被改变的植物的抗虫性也可被改变。蔗糖酯的组合物可以从所述植物中提取以用于各种用途，例如在药物、食品添加剂、吸烟风味剂中、以及作为有机杀虫剂的组分等。

[0007] 本发明的方面和实施方案

[0008] 本发明的方面和实施方案如所附权利要求所示。

[0009] 在第一方面中，提供分离多核苷酸，其包含、由或基本上由编码异丙基苹果酸合酶、并且与SEQ ID NO:1、或SEQ ID NO:10、或SEQ ID NO:12、或SEQ ID NO:14具有至少60%序列同一性的序列所组成。

[0010] 在另一方面，提供由所述多核苷酸编码的分离多肽。

[0011] 在又一方面，提供与SEQ ID NO:2、或SEQ ID NO:11、或SEQ ID NO:13、或SEQ ID NO:15具有至少60%序列同一性的分离多肽。

[0012] 含有分离多核苷酸序列的构建体、载体或表达载体，任选地其中所述构建体、载体或表达载体另外包括启动子，所述启动子包含、由或基本上由如SEQ ID NO:8所示的序列、或者与所示的序列具有至少约60%同一性的其变体、或者毛状体启动子、或者天然异丙基苹果酸合酶启动子所组成。

[0013] 在另一方面，提供分离多核苷酸，其包含、由或基本上由SEQ ID NO:8、或与其具有至少约60%同一性的其变体所组成。

[0014] 在另一方面，提供突变、非天然存在的或转基因的植物细胞，其含有本发明的多核苷酸的至少一种、多肽或者构建体、载体或表达载体中的至少一种。合适地，调节异丙基苹果酸合酶的表达或由其编码的蛋白的活性，至少植物的一部分相较于对照植物具有蔗糖酯组合物的改变，对照植物中异丙基苹果酸合酶的表达或活性未经调节。合适地，植物的视觉外观基本上和对照植物相同。

[0015] 在另一方面，提供突变、非天然存在的或转基因的植物，其含有本发明的植物细胞。

[0016] 在另一方面，提供在植物的部分中用于调节蔗糖酯量的方法，其包括下列步骤：(i) 调节在植物中的异丙基苹果酸合酶的表达或活性，优选地，其中所述异丙基苹果酸合酶含有本文所述的多核苷酸序列或多肽序列；(ii) 测量在获自步骤(i)的突变、非天然存在的或转基因植物的至少部分中的蔗糖酯的量；以及(iii) 鉴定突变、非天然存在的或转基因植物，其中其蔗糖酯的量与对照植物相比已经改变，所述对照植物中异丙基苹果酸合酶的表达或活性没有被调节；以及优选地，其中所述突变、非天然存在的或转基因植物的视觉外观与对照植物基本相同。

[0017] 在另一方面，提供通过这种方法获得的或可获得的突变、非天然存在的或转基因植物。

[0018] 在另一方面，提供突变、非天然存在的或转基因植物，其中异丙基苹果酸合酶的表达、或由其编码的蛋白的活性是得到调节的，以及至少植物的部分与其中异丙基苹果酸合

酶的表达或活性没有得到调节的对照植物相比在蔗糖酯的组合物上具有改变，并且其中所述植物的视觉外观与对照植物基本相同。

[0019] 在另一方面，提供植物材料，包括含有来自所述植物的细胞或组织的生物质、种子或叶。

[0020] 在另一方面，提供含有根据本发明的植物的部分或植物材料的烟草制品。

[0021] 在另一方面，提供用于产生蔗糖酯的组合物的方法，包括下列步骤：(i) 提供根据本发明的突变、非天然存在的或转基因植物、植物材料或烟草制品的至少部分；(ii) 从其中提取（例如，一种或多种）蔗糖酯；以及 (iii) 任选地，纯化提取出的蔗糖酯。

[0022] 也提供通过所述方法获得的或可获得的蔗糖酯的组合物。

[0023] 适宜地，所述一种或多种蔗糖酯具有如图5所示的结构，其中R3是乙酰基或氢，优选乙酰基；R1、R2和R4中的一个或多个含有至少一种6碳的酰基链，优选 $\beta$ -甲基戊酰基；以及R5是乙酰基或氢，优选地是氢。

[0024] 适宜地，所述一种或多种蔗糖酯选自：如图5所示的蔗糖酯，其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=丙酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基；或者如图5所示的蔗糖酯，其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基；或者如图5所示的蔗糖酯，其中R3=乙酰基、R1=丁酰基、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基(C2C14:0)；如图5所示的蔗糖酯，其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基；或者如图5所示的蔗糖酯，其中R3=乙酰基、R1=戊酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基；或者如图5所示的蔗糖酯，其中R3=乙酰基、R1=丁酰基、R2=己酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基(C2C15:0)；如图5所示的通用结构，其中R3是乙酰基、R1是戊酰基或其同分异构体、R2是戊酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体(C2C16:0)；或者蔗糖酯具有如图5所示的通用结构，其中R3是乙酰基、R1是丙酰基或其同分异构体、R2是己酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体(C2C16:0)；蔗糖酯具有如图5所示的通用结构，其中R3是乙酰基、R1是戊酰基或其同分异构体、R2是己酰基、R4是己酰基、以及R5是氢原子或乙酰基半体(C2C17:0)；或蔗糖酯具有如图5所示的通用结构，其中R3是乙酰基、R1是己酰基、R2是己酰基、R4是己酰基、以及R5是氢原子或乙酰基半体(C2C18:0)或其组合。

[0025] 适宜地，所述一种或多种蔗糖酯选自：如图5所示的蔗糖酯，其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基；或者如图5所示的蔗糖酯，其中R3=乙酰基、R1=戊酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基；或者如图5所示的蔗糖酯，其中R3=乙酰基、R1=丁酰基、R2=己酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基(C2C15:0)；蔗糖酯具有如图5所示的通用结构，其中R3是乙酰基、R1是戊酰基或其同分异构体、R2是戊酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体(C2C16:0)；或者蔗糖酯具有如图5所示的通用结构，其中R3是乙酰基、R1是丙酰基或其同分异构体、R2是己酰基或其同分异构体、R4是己酰基或

其同分异构体、以及R5是氢原子或乙酰基半体(C2C16:0)；蔗糖酯具有如图5所示的通用结构，其中R3是乙酰基、R1是戊酰基或其同分异构体、R2是己酰基、R4是己酰基、以及R5是氢原子或乙酰基半体(C2C17:0)；或者蔗糖酯具有如图5所示的通用结构，其中R3是乙酰基、R1是己酰基、R2是己酰基、R4是己酰基、以及R5是氢原子或乙酰基半体(C2C18:0)或其组合。

[0026] 在另一方面，提供调节烟草或烟草制品风味的方法，包括：(i) 加入来自本文所述的突变、非天然存在的或转基因植物、或者植物材料的烟草或烟草制品、植物部分，优选叶，优选地其中所述突变、非天然存在的或转基因植物、或者植物材料属于烟草(Nicotinia)属；或者(ii) 将通过本文所述的方法获得的或可获得的包括含蔗糖酯的β-甲基戊酰基的组合物加入到烟草或烟草制品。

[0027] 在另一方面，提供用于产生β-甲基戊酸的方法，包括以下步骤：(i) 提供突变、非天然存在的或转基因植物、植物材料或烟草制品的至少部分；(ii) 水解步骤(i)提供的材料或其提取物；以及(iii) 任选地分离或纯化β-甲基戊酸。适宜地，所述植物是烟草植物，例如烟草属的植物或红花烟草种的植物。

[0028] 适宜地，所述植物的视觉外观与对照植物基本相同。适宜地，在田间移植后三个月、或者在去稍之后36天，所述植物的视觉外观与对照植物基本相同；优选其中，田间移植后三个月或者在去稍之后36天，突变、非天然存在的或转基因植物的秆高与对照植物的秆高基本相同；和/或在田间移植后三个月或者在去稍之后36天，所述突变、非天然存在的或转基因植物的叶绿素含量与对照植物的叶绿素含量基本相同。在其它实施方案中，在田间移植后三个月或者在去稍之后36天，突变、非天然存在的或转基因植物的下列特征的任何一个或多个：成熟度、每株叶片数、秆高、叶插入角度、叶尺寸(宽度和长度)、节间距离、叶片-中脉比、以及叶片着色与对照植物基本相同。

[0029] 再一方面涉及含有来自本文所述植物的细胞或组织的生物质、种子或叶，其是含有所述植物的部分或生物质、种子或叶的烟草制品。

[0030] 在本公开的另一方面中也提供通过这种方法鉴定的或可鉴定的烟草材料。

[0031] 所述植物的抗虫性可以被调节。

[0032] 本发明的其它方面如下所示。

[0033] 含有所述分离多核苷酸的嵌合基因可操作地连接到一个或多个调控序列。

[0034] 异丙基苹果酸合酶多核苷酸构建体包含、由或基本上由至少15-30个核苷酸、30-50个核苷酸、50-100个核苷酸、100-150个核苷酸、150-200个核苷酸、200-300个核苷酸、300-400个核苷酸、400-500个核苷酸、500-600个核苷酸、600-700个核苷酸、700-1000个核苷酸、1000-1300个核苷酸、1300-1500个核苷酸或1500-1900个核苷酸所组成。可消费的产物，其整合或利用根据本发明的植物材料、生物质、种子或叶。

[0035] 细胞或细胞系，其含有根据本发明的分离多核苷酸、嵌合基因、多核苷酸构建体、双链RNA、缀合物或表达载体等。

[0036] 用于在细胞中调节异丙基苹果酸合酶的表达、或由其编码的蛋白的活性的方法，所述方法包括施用本文所述的嵌合基因、多核苷酸构建体、双链RNA、缀合物或表达载体。

[0037] 用于检测、分离、扩增或分析异丙基苹果酸合酶多核苷酸的方法，所述方法包括以下步骤：提供含有多核苷酸的样本，使所述多核苷酸杂交至另一多核苷酸分子，所述另一多核苷酸分子含有来自根据本发明的分离核苷酸序列的至少10个连续核苷酸的核苷酸序列。

[0038] 调节异丙基苹果酸合酶的表达、或由其编码的蛋白的活性的试剂,用于调节在植物的至少部分如叶中的蔗糖酯含量的用途。

[0039] 在一个实施方案中,所述试剂是或源自异丙基苹果酸合酶核酸、嵌合异丙基苹果酸合酶基因、含有异丙基苹果酸合酶多核苷酸的多核苷酸构建体、反义RNA、双链RNA、cDNA、含有异丙基苹果酸合酶多核苷酸的缀合物以及与其共价附着的至少一种非核苷酸或非多核苷酸半体、核酶、诱变剂(mutagen)、锌指、小分子或大范围核酸酶。

[0040] 在另一个实施方案中,所述多核苷酸片段编码反义核酸、核酶、影响剪接体介导的反式剪接的RNA、干扰RNA (RNAi)、向导RNA、或其它非翻译RNA等等。在另一实施方案中,所述多核苷酸片段编码RNAi。

[0041] 另一方面涉及产生烟草制品的方法,包括以下步骤:(a)从本文所述的转基因、突变或非天然存在的植物中获得种子;(b)种植种子并使之生长成植物;(c)收获植物;以及(d)从收获的植物中制备烟草制品。

[0042] 上述实施方案作为上述各方面的实施方案而公开。

## 附图说明

[0043] 图1说明在叶渗出物上测量的蔗糖酯同分异构体的估计量。在来自bmvsse烟草品种希克斯阔叶 (HBL) 以及来自BMVSE烟草品种红色俄罗斯 (RR) 的绿叶盘的丙酮/甲醇洗涤中进行测量。从作为外标添加的蔗糖八乙酸酯中估计所述量。对每个品种而言,n=4。C2C12:0包括如图5所示的(饱和)蔗糖酯,其中R3=乙酰基、R1=丁酰基、R2=丙酰基或其同分异构体、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基;或者如图5所示的蔗糖酯,其中R3=乙酰基、R1=乙酰基、R2=戊酰基或其同分异构体、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基。C2C13:0包括如图5所示的(饱和)蔗糖酯其中R3=乙酰基、R1=丁酰基、R2=丁酰基、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基;或者如图5所示的蔗糖酯,其中R3=乙酰基、R1=丁酰基、R2=丙酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基;或者如图5所示的蔗糖酯,其中R3=乙酰基、R1=丁酰基、R2=丙酰基或其同分异构体、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基。C2C14:0包括如图5所示的(饱和)蔗糖酯其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=丙酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基;或者如图5所示的蔗糖酯,其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基;或者如图5所示的蔗糖酯,其中R3=乙酰基、R1=丁酰基、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基。C2C15:0包括如图5所示的(饱和)蔗糖酯,其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基;或者如图5所示的蔗糖酯,其中R3=乙酰基、R1=戊酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基;或者如图5所示的蔗糖酯,其中R3=乙酰基、R1=丁酰基、R2=己酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基;C2C16:0包括具有如图5所示的通用结构的(饱和)蔗糖酯,以及其中R3是乙酰基、R1是戊酰基或其同分异构体、R2是戊酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体;或者其中所述蔗糖酯具有图5所示的通用结构,其中R3是乙酰基、R1是

丙酰基或其同分异构体、R2是己酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体。C2C17:0包括如图5所示的(饱和)蔗糖酯,其中R3是乙酰基、R1是戊酰基或其同分异构体、R2是己酰基、R4是己酰基、以及R5是氢原子或乙酰基半体。C2C18:0包括如图5所示的(饱和)蔗糖酯,其中R3是乙酰基、R1是己酰基、R2是己酰基、R4是己酰基、以及R5是氢原子或乙酰基半体。

[0044] 图2显示主要的蔗糖酯同分异构体,其按照分子量以及其在生产或不产含 $\beta$ -甲基戊酯的蔗糖酯的烟草品种中的存在进行分类。在分别列为C5和C6的2和3甲基丁酰基或3和4甲基戊酰酯之间没有区别。下列术语用于定义酰基组合物。例如,提到C2C4C4C5这指分别在R3、R1、R2和R4上酰基中碳的数目。C2可以是乙酰基;C3是可以是丁酰基;C4可以是丙酰基或其同分异构体、或异丁酰基;C5可以是戊酰基(戊酰基(pentanoyl))或其同分异构体、或2-甲基-丁酰基(butyryl)、或异戊酰基、或异戊酰基、或戊酰基;C6可以是己酰基或其同分异构体、或2-甲基戊酰基、或 $\beta$ -甲基戊酰基、或4-甲基戊酰基。乙酰基或氢可存在于R5(未显示),氢也是适宜的。

[0045] 图3说明支链氨基酸合成途径、以及在蔗糖上酯化的酰基链的生成。框起的步骤代表核心支链氨基酸合成途径,其中通过每轮的缩合同分异构化作用(condensation isomerization)和脱羧作用将所述碳链延伸一个碳。1:乙酰乳酸合酶[2.2.1.6];2:乙酰羟酸还原异构酶[1.1.1.86];3:3,4-二羟酸脱水酶[4.2.1.9];4:支链氨基酸转氨酶[2.6.1.42];5:支链酮酸脱氢酶复合物(BCKD)([1.2.4.4][2.3.1.168][1.8.1.4]);6:异丙基苹果酸合酶[2.3.3.13];7:异丙基苹果酸异构酶[4.2.1.33];8:异丙基苹果酸脱氢酶[1.1.1.85];9:苏氨酸脱氨酶[4.3.1.19]。C4:异丁酰基酯(isobutanoyl ester);iC5:3-甲基-丁酰基酯(异戊酰基酯);aiC5:2-甲基-丁酰基酯(反异-戊酰基酯);iC6:4-甲基-戊酰基酯(异-己酰基酯);aiC6:3-甲基-戊酰基酯(反异-己酰基酯)、 $\beta$ -甲基戊酰基酯(BMV)。羧酸以中性形式存在,虽然名字反映出可溶的形式。建议的IPMS2催化步骤是圈中的。

[0046] 图4:显示转化植物的表面叶渗出物中的蔗糖四酯丰度的估计。从独立转化事件中所得的一月龄再生植物中收集叶。以2.5ng/mg鲜重的比添加内标(蔗糖八乙酸酯)。针对每种蔗糖酯特异的离子所记录的信号面积加以测量,并作为内标面积的比加以绘制。与内标信号面积的比允许将叶鲜重上的信号正态化。图2描述了此处所示的蔗糖酯的命名。A:亲本品种K326(烤烟,n=3)的叶渗出物的组合物和在毛状体特异启动子下表达SEQ ID NO:1的转化体;B:亲本品种K326的叶渗出物的组合物和在病毒启动子控制下表达SEQ ID NO:1的转化体;C:亲本品种巴斯玛(Basma)(东方,n=4)的叶渗出物的组合物和在毛状体特异启动子下表达SEQ ID NO:9的RNAi构建体的转化体;以及D:亲本品种巴斯玛的叶渗出物的组合物和在病毒启动子下表达SEQ ID NO:9的RNAi构建体的转化体。C2C14:0包括如图5所示的(饱和)蔗糖酯,其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=丙酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基;或如图5所示的蔗糖酯,其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基;或如图5所示的蔗糖酯,其中R3=乙酰基、R1=丁酰基、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基。C2C15:0包括如图5所示的(饱和)蔗糖酯,其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基;或如图5所示的蔗糖酯,其中R3

=乙酰基、R1=戊酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯，其中R3=乙酰基、R1=丁酰基、R2=己酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基；C2C16:0包括具有如图5所示的通用结构的(饱和)蔗糖酯，其中R3是乙酰基、R1是戊酰基或其同分异构体、R2是戊酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体；或者其中蔗糖酯具有如图5所示的通用结构以及其中R3是乙酰基、R1是丙酰基或其同分异构体、R2是己酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体。C2C17:0包括如图5所示的(饱和)蔗糖酯，其中R3是乙酰基、R1是戊酰基或其同分异构体、R2是己酰基、R4是己酰基、以及R5是氢原子或乙酰基半体。C2C18:0包括如图5所示的(饱和)蔗糖酯，其中R3是乙酰基、R1是己酰基、R2是己酰基、R4是己酰基以及R5是氢原子或乙酰基半体。

[0047] 图5说明蔗糖酯支架的结构。R1至R5是氢原子或通过酯键连接的酰基半体。

[0048] 图6说明在蔗糖酯上酯化的酰基链。R'代表蔗糖分子。1.乙酰基、2.丁酰基、3.丙酰基、4.异丁酰基、5.戊酰基(戊酰基)、6.2-甲基-丁酰基、7.异戊酰基、8.异戊酰基、9.戊酰基、10.己酰基或其同分异构体、11.2-甲基戊酰基、12.β-甲基戊酰基、13.4-甲基戊酰基。也考虑同分异构体及其衍生物。因此，例如，所述同分异构体可以是丙酰基的同分异构体，例如异丁酰基。再例如，所述同分异构体可以是戊酰基(戊酰基)的同分异构体，如2-甲基-丁酰基、或异戊酰基；或者具有双键的戊酰基(戊酰基)的同分异构体，例如异戊酰基或戊酰基。又例如，所述同分异构体可以是己酰基的同分异构体，例如2-甲基戊酰基、β-甲基戊酰基、或4-甲基戊酰基。所述同分异构体可包含或者可不包含一个或更多双键。

## 具体实施方案

[0049] 定义

[0050] 在本申请的范围内，所用的技术术语和表达通常被给予普遍适用于植物和分子生物学相关领域中的含义。所有下列术语定义适用于本申请中的全部内容。词“包括”不排除其它元素或步骤，并且不定冠词“一”或“一个”不排除复数。单个步骤可能实现权利要求中记载的几个特征的功能。一个给定计算值或范围的上下文中，术语“约”、“基本上”和“近似地”指给定值或范围的20%内、10%内、或5%、4%、3%、2%或1%内的值或范围。

[0051] 术语“分离的”是指取自其天然环境的任何实体，但是所述术语并不意味着任何程度的纯化。

[0052] 术语“载体”是指核酸载体，其包括能够转运核酸、核酸构建体和核酸缀合物等等的核酸组分的组合。合适的载体包括能够进行染色体外复制的附加体，如环形、双链核酸质粒；线性化的双链核酸质粒；以及任何来源的其它载体。

[0053] “表达载体”是指核酸载体，其包括能够表达核酸、核酸构建体和核酸缀合物等等的核酸组分的组合。合适的表达载体包括能够进行染色体外复制的附加体，如环形、双链核酸质粒；线性化的双链核酸质粒；以及任何来源的其它功能等同的表达载体。表达载体包括至少一个位于上游的并且可操作地连接至如下所限定的核酸、核酸构建体或核酸缀合物的启动子。

[0054] 术语“构建体”是指双链、重组核酸片段，其包括一个或多个多核苷酸。构建体包括

与互补的“有义或编码链”碱基配对的“模板链”。给定的构建体可以按两个可能的方向插入到载体中,相对于位于载体(如表达载体)中启动子的方向相同(或有义)方向或相反(或反义)方向。示例构建体如SEQ ID NO:9所示。

[0055] “启动子”是指通常位于上游并且可操作地连接至双链核酸片段的核酸元件/序列。启动子可完全源自靠近目标天然基因的区域,或可由源自不同天然启动子或合成核酸节段的不同元件所组成。示例启动子如SEQ ID NO:8所示。

[0056] 术语“同源性、同一性或相似性”是指通过序列比对比较的两个多肽之间或两个核酸分子之间序列相似性的程度。待比较两个离散核酸序列之间的同源性程度是在可比较的位置上相同的、或匹配的核苷酸数目的函数。百分比同一性可通过目测和数学计算来确定。或者,可以通过使用计算机程序如ClustalW、ClustalX、BLAST、FASTA或Smith-Waterman比较序列信息来确定两个核酸序列的百分比同一性。

[0057] 术语“植物”是指在其生命周期或发育的任何阶段的任何植物及其后代。在一个实施方案中,植物是烟草植物,其是指属于烟草属的植物。本文描述了烟草植物的优选物种、栽培品种、杂种和品种。

[0058] “植物细胞”指植物的结构和生理单元。植物细胞可以是没有细胞壁的原生质体的形式、分离的单细胞或培养细胞、或作为更高组织单元的部分,例如而不限于植物组织、植物器官或整个植物。

[0059] 术语“植物材料”是指可获自植物的任何固体、液体或气体组合物、或其组合,其包括生物质、叶、叶片、中脉、茎、根、花或花部分、果实、花粉、卵细胞、受精卵、种子、插条、分泌物、提取物、细胞或组织培养物、或植物的任何其它部分或产品。在一个实施方案中,植物材料包括或由生物质、种子或叶组成。在另一个实施方案中,植物材料包括或由叶组成。

[0060] 术语“品种”是指共享恒定特性的植物群体,所述特性将其从相同物种的其它植物中分离出来。当具有一个或多个独特性状时,品种还特征在于所述品种内个体间非常小的整体差异。品种往往是市售的。

[0061] 本文所用术语“系”或“育种系”表示植物育种期间所使用的一组植物。系不同于品种,因为系表现个体间一个或多个目标性状的小差异,尽管对于其它性状可能有一些个体间差异。

[0062] 术语“调节”可指减少、抑制、增加或其它影响多肽活性。所述术语也可指减少、抑制、增加或其它影响编码多肽的基因活性,其可包括而非限制于调节转录活性。

[0063] 此处所用术语“减少”或“减少的”是指从约10%至约99%的减少,或至少10%、至少20%、至少25%、至少30%、至少40%、至少50%、至少60%、至少70%、至少75%、至少80%、至少90%、至少95%、至少98%、至少99%、或至少100%、或更多的量或活性的减少,例如但不限于多肽活性、转录活性、和/或蛋白表达。

[0064] 如本文所用,术语“抑制”或“抑制的”是指从约98%至约100%的减少,或至少98%、至少99%、但尤其是100%的量或活性的减少,例如但不限于多肽活性、转录活性和/或蛋白表达。

[0065] 本文所用术语“增加”或“增加的”是指从约10%至约99%的增加,或至少10%、至少20%、至少25%、至少30%、至少40%、至少50%、至少60%、至少70%、至少75%、至少80%、至少90%、至少95%、至少98%、至少99%、或至少100%、至少150%、或至少200%或

更多的量或活性的增加,例如但不限于多肽活性、转录活性、和/或蛋白表达。

[0066] 术语“对照”在对照植物或对照植物细胞的背景中,是其中异丙基苹果酸合酶的表达或活性未经改变(例如增加或减少的)植物或植物细胞,因此其可以提供与异丙基苹果酸合酶的表达或活性已经过改变的植物作比较。对照植物可包括空载体。对照植物可与野生型植物相对应。

[0067] 发明详述

[0068] 在一方面,提供分离多核苷酸,其包含、由或基本上由多核苷酸序列所组成,并且与本文所述的任何序列具有至少60%序列同一性,包括在序列表中显示的任何多核苷酸。适宜地,所述分离多核苷酸包含、由或基本上由下列序列所组成:与其具有至少60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列。

[0069] 在一方面中,提供分离多核苷酸,其包含、由或基本上由编码异丙基苹果酸合酶且与SEQ ID N0:1具有至少60%序列同一性的多核苷酸序列所组成。合适地,所述分离多核苷酸由或基本上由与SEQ ID N0:1具有至少60%、61%、62%、63%、66%、67%、68%、69%、70%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列同一性的序列所组成。

[0070] 在另一方面,提供分离多核苷酸,其包含、由或基本上由下列多核苷酸序列所组成,所述多核苷酸序列编码异丙基苹果酸合酶且与SEQ ID N0:10具有至少60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性。

[0071] 在又一方面,提供分离多核苷酸,其包含、由或基本上由下列多核苷酸序列所组成,所述多核苷酸序列编码异丙基苹果酸合酶且与SEQ具有至少60%、61%、62%、63%、64%、65%、70%、75%、80%、85%、87%、88%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的序列同一性。

[0072] 在还一方面,提供分离多核苷酸,其包含、由或基本上由下列多核苷酸序列所组成,所述多核苷酸序列编码异丙基苹果酸合酶且与SEQ ID N0:14具有至少60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性。

[0073] 在再一方面,提供分离多核苷酸,其包含、由或基本上由下列多核苷酸序列所组成,所述多核苷酸序列与SEQ ID N0:8具有至少60%、61%、62%、65%、70%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、98%、99%或100%的序列同一性。SEQ ID N0:8编码启动子。

[0074] 术语“多核苷酸”是指核苷酸的聚合物,其可以是未修饰的脱氧核糖核酸(DNA)或核糖核酸(RNA)。因此,多核苷酸可以是,但不限于基因组DNA、互补DNA(cDNA)、mRNA、或RNA或其片段。此外,多核苷酸可以是单链、单链和双链区的混合物的核酸、cDNA的克隆片段、基因组DNA、寡核苷酸、或单个核苷酸、或上述的组合。虽然本文所述的多核苷酸序列以DNA序列显示,所述序列包括其对应的RNA序列、及其互补(例如完全互补)的DNA或RNA序列、包括其反向互补。

[0075] 术语“异丙基苹果酸合酶多核苷酸”包括编码来自红花烟草的异丙基苹果酸合酶

(IMPS)的多核苷酸,以及包括多核苷酸,多核苷酸含有、由或者基本由与SEQ ID NO:1、SEQ ID NO:10、SEQ ID NO:12或SEQ ID NO:14基本上同源(即序列相似)或基本同一的多核苷酸所组成;多核苷酸变体,其与SEQ ID NO:1、SEQ ID NO:10、SEQ ID NO:12或SEQ ID NO:14的序列具有至少约60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、或99%的序列同一性;多核苷酸的片段,其含有SEQ ID NO:1、SEQ ID NO:10、SEQ ID NO:12或SEQ ID NO:14的片段;以及SEQ ID NO:1、SEQ ID NO:10、SEQ ID NO:12或SEQ ID NO:14 的片段,其与如下的序列具有基本同源性(即序列相似)或基本同一性:所述序列与SEQ ID NO:1、SEQ ID NO:10、SEQ ID NO:12或SEQ ID NO:14的相应片段具有至少约60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的序列同一性。所述异丙基苹果酸合酶多核苷酸还包括这样的序列,序列含有与SEQ ID NO:1、SEQ ID NO:10、SEQ ID NO:12或SEQ ID NO:14具有足够或充分的同一性或相似性的程度以编码发挥异丙基苹果酸合酶功能的多肽。在一个实施方案中,术语“异丙基苹果酸合酶多核苷酸”指包含、由或者基本由在本文中指定为SEQ ID NO:1、SEQ ID NO:10、SEQ ID NO:12或SEQ ID NO:14的多核苷酸所组成的核苷酸的聚合物。

[0076] 本文所述的多核苷酸将普遍含有磷酸二酯键,然而在某些情况下包括多核苷酸类似物,类似物可具有包括例如氨基磷酸酯、硫代磷酸酯、二硫代磷酸酯或0-甲基氨基磷酸酯链接的替换主链;以及肽多核苷酸主链和链接。其它多核苷酸类似物包括具有阳性主链、非离子型主链和非核糖主链的那些。核糖-磷酸主链的修饰可以出于各种原因进行,例如为了增加这样的分子在生理环境中或在生物芯片上作为探针的稳定性和半衰期。可以制备天然存在的多核苷酸和类似物的混合物;或者,可以制备不同多核苷酸类似物的混合物,以及天然存在的多核苷酸和类似物的混合物。

[0077] 已知各种多核苷酸类似物,包括例如氨基磷酸酯、硫代磷酸酯、二硫代磷酸酯、0-甲基氨基磷酸酯链接以及肽多核苷酸主链和链接。其它多核苷酸类似物包括具有阳性主链、非离子型主链和非核糖主链的那些。也包括含有一个或更多个碳环糖的多核苷酸。

[0078] 其它类似物包括肽核酸,其是肽的多核苷酸类似物。与天然存在的多核苷酸的高电荷磷酸二酯主链不同,这些主链在中性条件下大体上是非离子型的。这可能会带来优势。首先,肽多核苷酸主链可能会表现出改善的杂交动力学。相对于完全匹配的碱基对,肽核酸对于错配的碱基对在解链温度上具有更大的变化。DNA和RNA对于内部错配在解链温度上通常表现出2-4°C的下降。对于非离子型的肽多核苷酸主链,下降接近7-9°C。相似地,由于其非离子性质,碱基附着至这些主链的杂交对盐浓度相对不敏感。此外,肽核酸可能不会被细胞酶降解或降解程度较小,因此可能会更稳定。

[0079] 在公开的多核苷酸、及其片段的组合的用途当中,是片段在核酸杂交分析中作为探针、或者在核酸扩增分析中用作引物的用途。这样的片段通常包括DNA序列的至少约10、11、12、13、14、15、16、17、18、19或20或更多个连续核苷酸。在其它实施方案中,DNA片段包括DNA序列的至少约10、15、20、30、40、50或60或更多个连续核苷酸。因为功能异丙基苹果酸合酶基因在本文所示是形成蔗糖酯所必须的,所以这些技术也可用于鉴定BMVSE和bmvs植物。邻近的遗传标志物与异丙基苹果酸合酶基因的等位基因可用作植物蔗糖酯产量的预

测物。这种测试可用于追踪植物品种或预测烟熏质量 (smoking quality) 或分批类型的同质性 (homogeneity)。因此在一方面,也提供用于检测如本文所述的异丙基苹果酸合酶多核苷酸的方法,包括使用能够特异性检测或特异性扩增所述多核苷酸的探针或引物、或者特异性检测且特异性扩增所述多核苷酸的探针或引物。在另一方面,提供用于鉴定能够产生蔗糖酯的植物的方法,所述方法包括步骤: (a) 提供含有来自目标植物核酸的样本; 以及 (b) 确定如本文所述的异丙基苹果酸合酶多核苷酸的存在, 在其中所述植物中所述多核苷酸的存在指示所述植物能够产生蔗糖酯。还提供用于检测至少一部分的异丙基苹果酸合酶多核苷酸的试剂盒, 所述试剂盒包括用于特异性检测至少一部分异丙基苹果酸合酶多核苷酸的一种或更多种 (one of more) 引物或探针。所述试剂盒可包括用于多核苷酸扩增如聚合酶链式反应 (PCR) 的试剂, 或用于核酸探针杂交检测技术如Southern印迹、Northern印迹、原位杂交或微阵列的试剂。所述试剂盒可包括用于抗体结合检测技术比如Western印迹、ELISA、SELDI质谱或测试条的试剂。所述试剂盒可包括用于DNA测序的试剂。所述试剂盒可包括用于确定蔗糖酯含量的试剂和/或说明书。在某些实施方案中, 试剂盒可包括所述一种或更多种方法的说明书。所述试剂盒可用于遗传同一性确定、系统发育研究、基因分型、单体型分析、谱系分析、或特别具有共显性评分 (co-dominant scoring) 的植物育种。所述方法和试剂盒可用于依赖BMVSE和bmvs品种之间杂交的育种方法、以及在后代中基于序列的植物选择。一般来说, 将选择显示有功能的异丙基苹果酸合酶遗传区的植物, 或者选择显示功能障碍的异丙基苹果酸合酶遗传区的植物, 所述有功能的异丙基苹果酸合酶遗传区存在于产生BMVSE的品种, 所述功能障碍的异丙基苹果酸合酶遗传区存在于不产生bmvs的品种。

[0080] 影响杂交条件选择的基本参数和制定合适条件的指南描述于Sambrook, J., E.F. Fritsch和T. Maniatis (1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.)。使用遗传密码子结合本文所述氨基酸序列的知识, 可以制备简并寡核苷酸的集合。如在PCR中, 这样的寡核苷酸用作引物, 由此分离并扩增DNA片段。在某些实施方案中, 可以用简并引物作为遗传文库的探针。这种文库包括但不限于cDNA文库、基因组文库、以及甚至电子表达序列标签或DNA文库。然后将通过此方法鉴定的同源序列用作探针, 来鉴定本文所鉴定序列的同系物。

[0081] 潜在用途还有多核苷酸和寡核苷酸(例如引物或探针), 其在降低的严谨条件下、通常中等严谨条件下、以及通常高度严谨条件下杂交至本文所述的多核苷酸。影响杂交条件选择的基本参数和制定合适条件的指南描述于Sambrook, J., E.F. Fritsch和T. Maniatis (1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.), 并且可以由本领域的普通技术人员基于如多核苷酸的长度或碱基组成很容易地确定。

[0082] 实现中等严谨条件的一种方法涉及使用含有 $5\times$ 标准柠檬酸钠、0.5%十二烷基硫酸钠、1.0mM乙二胺四乙酸 (pH8.0) 的预洗溶液, 约50%甲酰胺、 $6\times$ 标准柠檬酸钠的杂交缓冲液, 以及约55°C的杂交温度(或其它相似的杂交溶液, 如含有约50%甲酰胺的一种, 杂交温度约42°C), 以及 $0.5\times$ 标准柠檬酸钠、0.1%十二烷基硫酸钠中约60°C的洗涤条件。一般来说, 高度严谨条件定义为如上杂交条件, 但在约68°C,  $0.2\times$ 标准柠檬酸钠、0.1%十二烷基硫酸钠中洗涤。SSPE ( $1\times$  SSPE是0.15M氯化钠、10mM磷酸钠和1.25mM乙二胺四乙酸, pH值

7.4) 在杂交以及洗涤缓冲液中可替换为标准柠檬酸钠(1×标准柠檬酸钠是0.15M氯化钠和15mM柠檬酸钠);杂交完成后,进行15分钟的洗涤。应当理解,洗涤温度和洗涤盐浓度,通过应用控制杂交反应和双链体稳定性基本原则,可以根据需要进行调整以实现所需的严谨程度,如本领域技术人员所知的并在下面进一步描述(见,例如Sambrook,J.,E.F.Fritsch和T.Maniatis(1989,Molecular Cloning:A Laboratory Manual,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.)。当多核苷酸杂交至未知序列的靶多核苷酸时,杂交长度假定为杂交多核苷酸的长度。当已知序列的多核苷酸杂交时,杂交长度可通过比对多核苷酸的序列以及鉴定最佳序列互补性的区域或多个区域来确定。杂交期望小于50碱基对长度的杂交温度应该是低于杂交的解链温度5至10℃,其中根据下列等式确定解链温度。对于小于18碱基对长的杂交,解链温度(℃)=2(A+T碱基数目)+4(G+C碱基数目)。对于18碱基对长以上的杂交,解链温度(℃)=81.5+16.6(log<sub>10</sub>[Na<sup>+</sup>])+0.41(%G+C)-(600/N),其中N是杂交碱基数目,以及[Na<sup>+</sup>]是钠离子在杂交缓冲液中的钠离子的浓度(对于1×标准柠檬酸钠,[Na<sup>+</sup>]=0.165M)。一般来说,每个这种杂交多核苷酸具有这样的长度,是其所杂交的多核苷酸的长度的至少25%(通常至少50%、60%、或70%,并最常用为至少80%),并且与其所杂交的多核苷酸具有至少60%的序列同一性(例如,至少65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)。

[0083] 本领域技术人员将理解,线性DNA具有两个可能的方向:5'至3'方向和3'至5'方向。例如,如果参考序列处于5'至3'方向,并且如果第二序列在相同多核苷酸分子/链内处于5'至3'方向,那么参考序列和第二序列取向相同的方向,或具有相同的取向。一般来说,在给定启动子的调控下,启动子的序列和目标基因处于相同的取向。然而,相对于处于5'至3'方向的参考序列,如果第二序列在相同多核苷酸分子/链内处于3'至5'方向,那么参考序列和第二序列取向反义方向,或具有反义的取向。如果参考序列(5'至3'方向)和参考序列的反向互补序列(参考序列处于5'至3')处于相同多核苷酸分子/链内,相对于彼此具有反义取向的两条序列可以可选择地描述为具有相同的取向。本文所示序列以5'至3'方向显示。

[0084] 本文提供的重组构建体可用于转化植物或植物细胞以调节异丙基苹果酸合酶蛋白表达水平。重组多核苷酸构建体可包括编码如本文所述的异丙基苹果酸合酶多核苷酸的多核苷酸,其可操作地连接于适于在植物或细胞中表达异丙基苹果酸合酶多肽的调控区。因此,多核苷酸可包括编码本文所述的异丙基苹果酸合酶多肽的编码序列。由重组多核苷酸编码的异丙基苹果酸合酶多肽可以是天然的、也可以是对细胞异源的异丙基苹果酸合酶多肽。在某些情况下,重组构建体含有减少或抑制异丙基苹果酸合酶-调节多肽表达的多核苷酸,其可操作地连接于调控区。合适调控区的实例是本文所述的。

[0085] 还提供了如本文所述的那些含有重组多核苷酸构建体的载体。合适的载体骨架包括例如本领域常规使用的那些,比如质粒、病毒、人工染色体、BAC、YAC或PAC。合适的表达载体包括,但不限于,源自例如噬菌体、杆状病毒和逆转录病毒的质粒和病毒载体。许多载体和表达系统是商业上可获得的。

[0086] 载体还可以包括例如复制起点、支架附着区或标志物。标志物基因可以为植物细胞赋予可选择的表型。例如,标志物可以赋予杀生物剂抗性,比如对抗生素(例如卡那霉素、G418、博来霉素或潮霉素)或除草剂(例如草甘膦、氯磺隆或草铵膦)的抗性。此外,表达载体

可以包括设计用来促进操作或检测(例如,纯化或定位)所表达多肽的标签序列。标签序列,如荧光素酶、 $\beta$ 葡萄糖醛酸酶(GUS)、绿色荧光蛋白(GFP)、谷胱甘肽S-转移酶(GST)、多聚组氨酸、c-myc或血凝素序列,通常作为与所编码的多肽的融合体来表达。这样的标签可以插入多肽内的任何地方,包括在羧基末端或氨基末端。

[0087] 植物或植物细胞可以通过将重组多核苷酸整合至基因组进行转化来实现稳定的转化。稳定转化的细胞通常在各细胞分裂中保留所引入的多核苷酸。也可以瞬时转化植物或植物细胞,从而重组多核苷酸没有整合至其基因组中。瞬时转化的细胞通常在各细胞分裂中丢失引入的重组多核苷酸的所有或一些部分,从而引入的重组多核苷酸在足够次数的细胞分裂之后不能在子细胞中检测到。

[0088] 用于转化植物细胞一些方法在本领域中是可用的,这些方法全部包括在本文中,包括生物弹技术、基因枪技术、农杆菌介导的转化、病毒载体介导的转化和电穿孔法。用于向植物染色体中整合外源核酸的农杆菌系统已经针对植物遗传工程进行了广泛的研究、改进和探索。包括核酸序列的裸重组核酸分子通过常规方法被连接到适当的T-DNA序列上,其中核酸序列对应着目标纯化烟草蛋白,并按有义或反义方向可操作地连接至调控序列。通过聚乙二醇技术或通过电穿孔技术将这些引入烟草原生质体,两者都是标准的。或者,将这种含有编码目标纯化蛋白的重组核酸分子的载体引入活的农杆菌细胞,随后所述农杆菌细胞将核酸转移到植物细胞中。可以通过烟草原生质体与含有核酸的脂质体进行融合或者通过电穿孔可以实现不伴有T-DNA载体序列的裸DNA的转化。不伴有T-DNA载体序列的裸DNA也可通过惰性、高速微弹(microparticle)用于转化烟草细胞。

[0089] 如果细胞或培养的组织用作转化受体组织,如果需要的话通过本领域技术人员已知的技术,植物可以从转化的培养物中再生。

[0090] 待纳入重组构建体的调控区选择取决于若干因素,包括但不限于效率、可选择性、可诱导性、所需的表达水平和细胞的或组织的优先表达。对于本领域技术人员而言,通过适当选择和定位相对于编码序列的调控区来调节编码序列的表达是一个常规的事情。多核苷酸的转录可以以类似方式调节。一些合适的调控区仅仅或主要在某些细胞类型中起始转录。用于鉴定和表征植物基因组DNA中调控区的方法是本领域已知的。

[0091] 合适的启动子包括组织特异性因子所识别的组织特异性启动子,其存在于不同组织或细胞类型中(例如,根特异性启动子、茎特异性启动子、木质部特异性启动子)、或不同发育阶段出现的、或者响应不同环境条件而出现的。合适的启动子包括组成型启动子,其可以在大多数细胞类型中被激活而无需特定的诱导物。用于控制异丙基苹果酸合酶RNAi多肽生产的合适启动子的实例包括花椰菜花叶病毒35S(CaMV/35S)、SSU、OCS、1ib4、usp、STLS1、B33、nos或泛素启动子或菜豆素启动子。本领域技术人员能够产生重组启动子的多种变体。

[0092] 组织特异性启动子是转录控制元件,其仅在植物发育期间的特定时间中在特定细胞或组织中有活性,如在营养组织或繁殖组织中。例如,当优选在某些组织中表达多核苷酸时,组织特异性表达可以是有利的。发育控制下的组织特异性启动子实例包括可以仅在(或主要仅在)某些组织(如营养组织例如根或叶,或生殖组织如果实、胚珠、种子、花粉、pistols(无对应中译文)、花或任何胚组织)中起始转录的启动子。生殖组织特异性启动子可以是例如花药特异性、胚珠特异性、胚特异性、胚乳特异性、珠被特异性、种子和种皮特异性、花粉特异性、花瓣特异性、萼片特异性的或其组合。

[0093] 合适的叶特异性启动子包括来自C4植物(玉米)的丙酮酸、正磷酸二激酶(PPDK)启动子、来自玉米的cab-m1Ca+2启动子、拟南芥(*Arabidopsis thaliana*)myb相关基因的启动子(Atmyb5)、核酮糖二磷酸羧化酶(RBCS)启动子(例如,在叶和光照生长的幼苗中表达的番茄RBCS1、RBCS2和RBCS3A基因、在发育中的番茄果实中表达的RBCS1和RBCS2、或者几乎只在叶片和叶鞘的叶肉细胞中高水平表达的核酮糖二磷酸羧化酶启动子)。

[0094] 合适的衰老特异性启动子包括在果实成熟、衰老和叶脱落期间的番茄启动子,编码半胱氨酸蛋白酶的基因的玉米启动子。可以使用合适的花药特异性启动子。可以选择本领域技术人员已知合适的根优选启动子。合适的种子优选启动子包括种子特异性启动子(在种子发育期间有活性的那些启动子,比如种子贮藏蛋白的启动子)和种子萌发启动子(在种子萌发期间有活性的那些启动子)。这样的种子优选启动子包括但不限于Cim1(细胞分裂素诱导的信息);cZ19B1(玉米19kDa的玉米醇溶蛋白);milps(肌-肌醇-1-磷酸合酶);mZE40-2也称为Zm-40;nuclc;以及cel1A(纤维素合酶)。 $\gamma$ -玉米醇溶蛋白是胚乳特异性启动子。Glob-1是胚特 异性启动子。对于双子叶植物,种子特异性启动子包括但不限于豆 $\beta$ 菜豆素(bean. $\beta$ -phaseolin)、油菜籽蛋白、 $\beta$ -伴大豆球蛋白、大豆凝集素、十字花科蛋白(cruciferin)等。对于单子叶植物,种子特异性启动子包括但不限于玉米15kDa的玉米醇溶蛋白启动子、22kDa的玉米醇溶蛋白启动子、27kDa的玉米醇溶蛋白启动子、 $\gamma$ -玉米醇溶蛋白启动子、27kDa的 $\gamma$ -玉米醇溶蛋白启动子(如gzw64A启动子,参阅GenBank登录号S78780)、糯质基因(waxy)启动子、超甜基因(shrunken)1启动子、超甜基因2启动子、球蛋白1启动子(见GenBank登录号L22344)、Itp2启动子、cim1启动子、玉米end1和end2启动子、nuc1启动子、Zm40启动子、eep1和eep2;lec1、硫氧还蛋白H启动子;m1ip15启动子、PCNA2启动子;以及超甜基因2启动子。

[0095] 可诱导启动子的实例包括响应病原体攻击、条件、温度升高、光、干旱、低温或高盐浓度的启动子。病原体可诱导的启动子包括来自病理相关蛋白(PR)的那些,病理相关蛋白是病原体感染之后诱导的(例如PR蛋白、SAR、 $\beta$ -1,3-葡聚糖酶、壳多糖酶)。

[0096] 除了植物启动子,其它合适的启动子可源自细菌,例如,章鱼碱合酶启动子、胭脂碱合酶启动子和源于Ti质粒的其它),或可以源自病毒启动子(例如,花椰菜花叶病毒(CaMV)的35SRNA启动子、烟草病毒的组成型启动子、花椰菜花叶病毒(CaMV)19S和35S启动子、或玄参花叶启动子)。

[0097] 术语“异丙基苹果酸合酶多肽”指来自红花烟草的异丙基苹果酸合酶(IPMS)的多肽并包括多肽,所述多肽包含、由或基本上由多核苷酸变体编码的氨基酸序列所组成,所述多核苷酸变体与SEQ ID NO:2、SEQ ID NO:11、SEQ ID NO:13或S:15具有至少约60%、61%、62%、63%、64%、65%、66%、69%、70%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、98%或99%的序列同一性;异丙基苹果酸合酶多肽的片段;以及SEQ、SEQ ID NO:11、SEQ ID NO:13或SEQ ID NO:15的片段,其与SEQ ID NO:2、SEQ ID NO、ID NO:13或SEQ ID NO:15的相应片段具有至少约60%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98或100%的序列同一性。所述异丙基苹果酸合酶多肽也包括与SEQ、SEQ ID NO:11、SEQ ID NO:13或SEQ ID NO:15含有足够或很大程度的同一性或相似性以发挥异丙基苹果酸作用的序列。所述异丙基苹果酸合酶多肽的片段通常保留合酶活性。只要异丙基苹果酸合酶多肽仍作为异丙基苹果酸起作用,其也包括通

过引入任何类型的改变(例如,酸的插入、删除或取代;糖基化状态的变化;影响重新折叠或异构化、三维结构或自我结合状态的变化)的突变体,其可以有意地经过工程化改造天然分离的。异丙基苹果酸合酶多肽可以处于线性形式或使用已知的。术语“异丙基苹果酸合酶多肽”也可指多肽。

[0098] 在另一方面,提供含有、由或基本上由与本文所述的任何序列具有至少60% 序列同一性的多肽序列所组成的分离多肽,其包括在序列表中显示的任何多肽。适宜地,所述分离多肽含有、由或基本上由与其具有至少60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列同一性的序列所组成。

[0099] 多肽包括通过引入任何类型的改变所产生的变体(例如,氨基酸的插入、删除或取代;糖基化状态的变化;影响重新折叠或异构化、三维结构或自我结合状态的变化),其可以有意地经过工程化改造或是天然分离的。所述变体可具有产生沉默改变并导致功能等同蛋白的改变。只要保留底物的二级结合活性,基于在残基的极性、电荷、溶解度、疏水性、亲水性和两亲性上的相似性可进行有意的氨基酸取代。例如,带负电的氨基酸包括天冬氨酸和谷氨酸;带正电的氨基酸包括赖氨酸和精氨酸;以及具有相似亲水性值的具有不带电极性头部基团的氨基酸包括亮氨酸、异亮氨酸、缬氨酸、甘氨酸、丙氨酸、天冬酰胺、谷氨酰胺、丝氨酸、苏氨酸、苯丙氨酸和酪氨酸。可进行保守取代,例如根据下表。在第二列相同方框中的氨基酸以及优选在第三列相同行中的氨基酸可互相取代:

[0100]	脂肪族	非极性	Gly Ala Pro Ile Leu Val
		极性 - 不带电荷	Cys Ser Thr Met Asn Gly
		极性 - 带电荷	Asp Glu Lys Arg
	芳香族		His Phe TrpTyr

[0101] 所述多肽可以是成熟蛋白、或不成熟蛋白、或源自不成熟蛋白的蛋白。多肽可以处于线性形式或者使用已知方法环化。多肽通常包含至少10、至少20、至少30、或至少40个连续的氨基酸。

[0102] 在另一方面,提供分离的异丙基苹果酸合酶多肽,其包含、由或者基本由编码异丙基苹果酸合酶的并且与SEQ ID NO:2具有至少约60%、65%、70%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列同一性的序列所组成。

[0103] 在再一方面,提供分离的异丙基苹果酸合酶多肽,其包含、由或者基本由编码异丙基苹果酸合酶的并且与SEQ ID NO:11具有至少约60%、65%、70%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列同一性的序列所组成。

[0104] 在又一方面,提供分离的异丙基苹果酸合酶多肽,其包含、由或者基本由编码异丙基苹果酸合酶的并且与SEQ ID NO:13具有至少约60%、65%、70%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列同一性

的序列所组成。

[0105] 在还一方面,提供分离的异丙基苹果酸合酶多肽,其包含、由或者基本由编码异丙基苹果酸合酶的并且与SEQ ID NO:15具有至少约60%、65%、70%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列同一性的序列所组成。

[0106] 所述多肽序列的片段也在本文中公开,适宜地,这样的片段保留了全长序列的活性。

[0107] 突变多肽可用于建立含有突变多肽的突变植物、非天然存在的植物或转基因植物。适宜地,所述突变多肽保留未突变多肽的活性。突变多肽的活性可以比未突变多肽更高、更低或与之大约相同。

[0108] 通过在适合表达多肽的培养条件下培养转化的或重组的宿主细胞可以制备多肽。然后可使用已知的纯化方法从这种培养物中纯化出所产生的表达的多肽。多肽纯化可包括含有将与多肽结合的试剂的亲和柱;一个或多个穿过这种亲和树脂的柱步骤;一个或多个涉及疏水性相互作用色谱的步骤;或免疫亲和色谱。或者,所述多肽也可以表达为将促进纯化的形式。例如,其可以作为融合多肽表达,如麦芽糖结合多肽、谷胱甘肽-5-转移酶、或硫氧还蛋白中的那些。用于表达和纯化融合多肽的试剂盒是商业上可获得的。多肽也可以用表位标记,并随后通过使用针对这种表位的特异性抗体进行纯化。可以采用一个或多个液相色谱步骤,例如反相高效液相色谱法进一步纯化多肽。以不同的组合可采用部分或全部上述纯化步骤来提供基本上同质的重组多肽。因此,纯化的多肽可基本上不含其它多肽,在本文定义为“基本上纯化的多肽”;这样纯化的多肽包括多肽、片段、变体等。可以通过任何合适的技术包括但不限于本文所述的方法实现多肽和片段的表达、分离和纯化。

[0109] 也可能利用亲和柱,如产生的抗多肽的单克隆抗体来亲和纯化表达的多肽。可以用常规技术从亲和柱中除去这些多肽,例如在高盐洗脱缓冲液中,然后对较低盐的缓冲液进行透析备用、或取决于所用的亲和基质通过改变pH值或其它成分,或者使用天然存在的亲和半体的底物竞争性地除去。

[0110] 也可以通过已知的常规化学合成来产生多肽。通过合成方式来构建多肽或其片段的方法是本领域技术人员已知的。合成构建的多肽序列通过与天然多肽分享一级、二级或三级结构或构象特征,可拥有与其共同的生物学性质,包括生物活性。

[0111] 本文所用术语“非天然存在的”描述不由自然形成或者不存在于自然当中的实体(例如,多核苷酸、遗传突变、多肽、植物、植物细胞和植物材料)。可以通过本文所述或本领域中已知的方法制造、合成、起始、修饰、干预或操纵这样的非天然存在的实体或人工实体。因此,例如可以使用传统植物育种技术(如回交)或通过遗传操作技术(如反义RNA、干扰RNA、大范围核酸酶等)制成非天然存在的植物、非天然存在的植物细胞或非天然存在的植物材料。再例如,可以通过将来自第一植物或植物细胞一个或多个遗传突变(例如一个或多个多态性)基因渗入或通过转移至第二植物或植物细胞(其本身可以是天然存在的)来制备非天然存在的植物、非天然存在的植物细胞或非天然存在的植物材料,从而得到的植物、植物细胞或植物材料或其后代含有非天然形成或者非天然存在的遗传结构(例如基因组、染色体或其节段)。因此,所得植物、植物细胞或植物材料是人工或者非天然存在的。因此,如果第二植物或植物细胞含有不同于第一植物或植物细胞的遗传背景,可以通过在第一人工

或天然存在的植物或植物细胞中修饰遗传序列来制成人工的或非天然存在的植物或植物细胞,导致(i)人工的修饰的遗传序列、或(ii)天然存在于第二植物或植物细胞的修饰的遗传序列。可以通过表型差异或通过本领域已知的分子生物学技术如核酸测序、遗传标志物(例如微卫星RNA标志物)的存在或缺失,检测植物或植物细胞在遗传背景上的差异。

[0112]一些基于多核苷酸的方法可用于增加基因在植物中的表达。例如,可制备和待转化的植物兼容的构建体、载体或表达载体,其包括目标基因连同能够在植物中过表达所述基因的上游启动子。在本文描述示例性启动子。转化后当生长在合适条件下,所述启动子可驱动表达以调节(例如,增加)这种酶在植物或在其特定组织中的水平。在一个示例性实施方案中,生成携带异丙基苹果酸合酶的载体以在植物中过表达所述基因。所述载体携带位于异丙基苹果酸合酶基因上游的合适启动子,所述启动子驱动所述基因在植物的所有组织中表达。为转化的愈伤组织和细胞系赋予选择性,所述载体还可携带抗生素抗性基因。在一个实施方案中,所述启动子促进在植物毛状体中的转录活性。在一个实施方案中,所述启动子在植物毛状体中具有转录活性。在另一个实施方案中,所述启动子在植物毛状体中具有特异性转录活性、而在植物的任何其它部分没有或基本没有活性。适宜地,所述启动子包含、由或基本上由SEQ ID NO:8所示的序列、或毛状体启动子或天然异丙基苹果酸合酶启动子所组成。天然异丙基苹果酸合酶启动子将位于异丙基苹果酸合酶基因的5'端。例如在SEQ ID NO:6中,内源异丙基苹果酸合酶启动子将定位在核苷酸2826之前。因此,所述异丙基苹果酸合酶基因可在植物的毛状体中过表达。令人惊讶地,本发明人已发现,组成型表达并不导致与使用毛状体特异表达所达到的水平一样高的蔗糖酯积累(参见图4和其描述)。

[0113]根据一个实施方案,在已能够产生含 $\beta$ -甲基戊酰基的蔗糖酯的植物中,异丙基苹果酸合酶的表达或活性提高,从而导致其中异丙基苹果酸合酶的积累提高,并因此提高含 $\beta$ -甲基戊酰基的蔗糖酯的生产。根据另一个实施方案,在不能够产生含 $\beta$ -甲基戊酰基的蔗糖酯的植物中,异丙基苹果酸合酶的表达或活性提高,从而导致其中异丙基苹果酸合酶的积累以及因此含 $\beta$ -甲基戊酰基的蔗糖酯的生产。

[0114]一些基于多核苷酸的方法,包括反义RNA、核酶指导的RNA切割、转录后基因沉默如RNA干扰、以及转录基因沉默,可用于调节在植物中基因表达。虽然这样的方法典型地用于减少或抑制基因表达,其也可通过增加细胞生产力和表达出的蛋白的质量用于增加基因表达。这样的方法可包括细胞凋亡相关基因表达、蛋白的糖基化相关基因表达、参与细胞代谢的乳酸脱氢酶、以及用于基因扩增的二氢叶酸还原酶的沉默。其它方法可以扩展到减少或抑制涉及用于改变植物细胞中基因调节的不同细胞通路中的多个靶标。

[0115]在一些实施方案中,可以使用全长多核苷酸或其片段的互补物。通常,片段是至少10个连续核苷酸,例如至少10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、30、35、40、50、80、100、200、500个或更多连续核苷酸。在一个实施方案中,所述多核苷酸片段能够减少表达或活性,并可编码例如本文所述的反义核酸、核酶、影响剪接体介导的反式剪接的RNA、干扰RNA、向导RNA或其它非翻译RNA等。在另一个实施方案中,所述多核苷酸片段编码干扰RNA。

[0116]可减少异丙基苹果酸合酶表达或活性的组合物包括,但不限于可干扰一个或更多内源异丙基苹果酸合酶基因转录的序列特异性多核苷酸;可干扰异丙基苹果酸合酶RNA转录物(例如,双链RNA、siRNA、核酶)翻译的序列特异性多核苷酸;可干扰异丙基苹果酸合酶

蛋白稳定性的序列特异性多肽；可干扰异丙基苹果酸合酶蛋白的酶活性、或异丙基苹果酸合酶蛋白对于底物或调节蛋白结合活性的序列特异性多核苷酸；显示对异丙基苹果酸合酶蛋白特异性的抗体；可干扰异丙基苹果酸合酶蛋白的稳定性、或异丙基苹果酸合酶蛋白的酶活性、或者异丙基苹果酸合酶蛋白的结合活性的小分子化合物；结合异丙基苹果酸合酶多核苷酸的锌指蛋白；以及具有针对异丙基苹果酸合酶多核苷酸的活性的大范围核酸酶。

[0117] 反义技术是一种公知的方法，其可以用来调节（例如，减少或抑制）多肽的表达。待抑制的多核苷酸被克隆并被可操作地连接到调节区和转录终止序列，从而转录RNA的反义链。然后，将重组构建体转化到植物中并产生RNA的反义链。多核苷酸不需是待抑制基因的整个序列，但通常将基本上互补于待抑制基因的正义链至少部分。

[0118] 多核苷酸可被转录成核酶、或催化性RNA，其影响mRNA的表达。核酶可被设计成特异性地与几乎任何靶RNA配对，并在特定位置上切割磷酸二酯骨架，从而在功能上使靶RNA失活。异源多核苷酸可以编码设计用于切割特定mRNA转录物的核酶，从而阻止多肽表达。尽管可以使用在位点特异性识别序列上切割mRNA的各种核酶，锤头核酶用于破坏特定的mRNA。锤头核酶在侧翼区所规定的位置切割mRNA，侧翼区与靶mRNA形成互补碱基对。唯一要求是靶RNA包含5'-UG-3'核苷酸序列。锤头核酶的构建和生产是本领域已知的。锤头核酶序列可以被嵌入至稳定的RNA（如转移RNA）中来增加体内切割效率。

[0119] 在一个实施方案中，可干扰RNA转录物翻译的序列特异性多核苷酸是干扰RNA。RNA干扰或RNA沉默是进化上保守的过程，通过它特异性mRNA可以被 靶向酶促降解。必须引入或由细胞产生双链RNA（例如，双链RNA病毒、或干扰RNA多核苷酸）来启动干扰RNA途径。双链RNA可以被RNA酶III转换成多个21-23bp长的小干扰RNA双链体，RNA酶III是双链RNA特异性内切酶。小干扰RNA可以随后被RNA诱导的沉默复合物所识别，其通过ATP-依赖的过程促进小干扰RNA解旋。解旋的小干扰RNA的反义链将激活的RNA诱导的沉默复合物引导至含有互补于小干扰RNA反义链的序列的靶mRNA。靶mRNA和反义链可形成A-型螺旋，以及A-型螺旋的大沟可以被激活的RNA诱导的沉默复合物所识别。在小干扰RNA链的5'端结合位点所定义的单一位点处，靶mRNA可以被激活的RNA诱导的沉默复合物所切割。激活的RNA诱导的沉默复合物可以被回收以催化另一个切割事件。

[0120] 干扰RNA表达载体可包括编码干扰RNA多核苷酸的干扰RNA构建体，其通过降低mRNA、pre-mRNA或相关RNA变体的表达水平显示出RNA干扰活性。表达载体可包括位于上游并可操作地连接到干扰RNA构建体的启动子，如本文进一步描述的。干扰RNA表达载体可包括合适的最小核心启动子、目标干扰RNA构建体、上游（5'）调控区、下游（3'）调控区，包括转录终止和多腺苷酸化信号，以及本领域技术人员已知的其它序列，如各种选择标志物。

[0121] 多核苷酸可以以各种形式生产，包括双链结构（即，含有反义链和互补正义链的双链RNA分子）、双链发夹样结构、或单链结构（即，只含有反义链的单链RNA分子）。结构可以包括双链体、不对称双链体、发夹或不对称发夹二级结构、具有自身互补性的正义和反义链。双链干扰RNA可以被酶转化为双链小干扰RNA。小干扰RNA双链体的其中一条链可以退火至靶mRNA和相关RNA变体内的互补序列。小干扰RNA/mRNA双链体由RNA诱导的沉默复合物识别，其可以在多位点上以序列依赖性方式切割RNA，导致靶mRNA和相关RNA变体的降解。

[0122] 双链RNA分子可包括由单个寡核苷酸组装成茎环结构的小干扰RNA分子，其中小干扰RNA分子自身互补的正义和反义区通过基于多核苷酸或基于非多核苷酸接头的方式相连

接,以及环状单链RNA具有两个或多个环结构、以及含有自身互补的正义和反义链的茎,其中环状RNA可以在体内或在体外进行加工来产生能够介导干扰RNA的活性小干扰RNA分子。

[0123] 本文还关注了小发夹RNA分子的用途,包括除反向互补(正义)序列以外的特异性反义序列,其通常由间隔或环序列分开。间隔或环的切割提供单链RNA分子及其反向互补物,使得它们可以退火形成双链RNA分子(任选地具有附加的处理步骤,其可导致从一条或两条链的3'端或5'端添加或除去一个、两个、三个或更多个核苷酸)。间隔可以足够长以在切割间隔之前(和任选地随后的处理步骤,其可导致从一条或两条链的3'端或5'端添加或除去一个、两个、三个、四个或更多个核苷酸),允许反义和正义序列退火以形成双链结构(或茎)。间隔序列通常是无关的核苷酸序列,其位于退火成双链多核苷酸时的两个互补核苷酸序列区之间,包括小发夹RNA。间隔序列通常包括在约3至约100个核苷酸之间。

[0124] 通过选择适合产生发夹双链体的序列组合物、环尺寸和茎长度,可产生任何目标RNA多核苷酸。适合设计发夹双链体的茎长的范围,包括至少约10、11、12、13、14、15、16、17、18、19或20个核苷酸,如约14-30个核苷酸、约30-50个核苷酸、约50-100个核苷酸、约100-150个核苷酸、约150-200个核苷酸、约200-300个核苷酸、约300-400个核苷酸、约400-500个核苷酸、约500-600个核苷酸以及约600-700个核苷酸的茎长。设计发夹双链体的环长的合适范围,包括约4-25个核苷酸、约25-50个核苷酸的环长、或当发双链体的茎长相当长时则更长的环长。双链RNA或单链RNA分子是在约15和约40个核苷酸长之间。小干扰RNA分子可以是在约15和约35个核苷酸长之间的双链RNA或单链RNA分子。小干扰RNA分子可以是在约17和约30个核苷酸长之间的双链RNA或单链RNA分子。小干扰RNA分子可以是在约19和约25个核苷酸长之间的双链RNA或单链RNA分子。小干扰RNA分子是在约21至约23个核苷酸长之间的双链RNA或单链RNA分子。具有长于21个核苷酸的双链体区的发夹结构可促进有效的小干扰RNA指导的沉默,无论环序列和长度如何。

[0125] 靶mRNA序列可以是在约14至约50个核苷酸长之间。因此,可以扫描靶mRNA在约14和约50个核苷酸长之间的区域,其对于靶序列优选地满足以下标准的一个或多个:在约2:1和约1:2之间的A+T/G+C比;靶序列的5'端的AA二核苷酸或CA二核苷酸;靶mRNA独有的至少10个连续核苷酸的序列(即,在相同植物的其它mRNA序列上不存在的序列);以及没有三个以上连续鸟嘌呤(G)核苷酸或三个以上连续胞嘧啶(C)核苷酸的“串”。使用本领域已知的各种技术可以评估这些标准,例如,计算机程序如BLAST可用于搜索公共可用数据库以确定是否所选靶序列是靶mRNA独有的。或者,使用商业上可获得的计算机软件(例如,01igoEngine、Target Finder以及商业上可获得的小干扰RNA Design Tool可以选择靶序列(和设计小干扰RNA序列)。

[0126] 可选择靶mRNA序列,其满足上述标准中的一个或多个且在约14和约30个核苷酸长之间,或者可选择靶序列,其满足上述标准中的一个或多个且在约16和约30个核苷酸长之间。可选择靶序列,其满足上述标准中的一个或多个且在约19和约30个核苷酸长之间。选择靶序列,其满足上述标准中的一个或多个且在约19和约25个核苷酸长之间。

[0127] 小干扰RNA分子可包括特异性反义序列,其互补于本文所述的多核苷酸序列中任一个的至少12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30或更多个连续核苷酸。

[0128] 小干扰RNA分子所含有的特异性反义序列可以与靶序列的互补物一致或基本一

致。小干扰RNA分子所含有的特异性反义序列可与靶mRNA序列的互补物至少约75%、80%、85%、90%、95%、96%、97%、98%、99%或100%一致。确定序列同一性的方法是本领域已知的，并且能够通过，例如使用威斯康星大学计算机组(GCG)软件或NCBI网站上提供的BLASTN程序来确定。

[0129] 通过与靶mRNA序列在一个、两个、三个、四个或更多个核苷酸上有差别(例如，通过核苷酸取代，包括转换或颠换)，小干扰RNA分子的特异性反义序列可以表现出变异性。当这样的核苷酸取代存在于双链RNA分子的反义链中时，与取代核苷酸通常形成氢键碱基配对的正义链中的互补核苷酸可能或可能不会被相应地取代。也可以考虑双链RNA分子，其中一个或多个核苷酸取代发生在正义序列中，但不在反义链中。当小干扰RNA分子的反义序列包括在小干扰RNA的核苷酸序列和靶核苷酸序列之间的一个或多个错配时，如上所述，所述错配可能会发现在反义序列的3'末端、5'末端或中央部分。

[0130] 小干扰RNA分子可包括特异性反义序列，其能够在严谨条件下选择性地杂交至天然存在的靶基因或靶mRNA的部分。如本领域普通技术人员所知，可以通过改变杂交和洗涤步骤中所用时间、温度或溶液浓度来实现杂交条件的严谨程度的变化。适当的条件也可以部分依赖于所用的特定核苷酸序列，例如靶mRNA或基因的序列。

[0131] 一种用于在植物中诱导双链RNA沉默的方法是用产生发夹RNA的基因构建体进行转化(参见Smith等，(2000)Nature, 407, 319-320)。这样的构建体包含由合适间隔所分开的靶基因序列的反转区。由于生成内含子剪接的发夹RNA，插入作为间隔片段的功能植物内含子区，额外增加基因沉默诱导的效率(Wesley等，(2001)Plant J., 27, 581-590)。适宜地，茎长是约50核苷酸至约1千碱基长。用于产生内含子剪接发夹RNA的方法在本领域详述(参见例如，Bioscience, Biotechnology, and Biochemistry (2008) 72, 2, 615-617)。

[0132] 具有双链体或双链结构的干扰RNA分子，例如双链RNA或小发夹RNA，可具有平端，或者可具有3'或5'突出。如本文所用，“突出”是指未配对的核苷酸，或者当一条RNA链的3'-末端延伸超出另一条链的5'-末端时突出于双链体结构的核苷酸(3'突出)，或反之亦然(5'突出)。含有突出的核苷酸可以是核糖核苷酸、脱氧核糖核苷酸或其修饰的版本。在一个实施方案中，干扰RNA分子的至少一条链具有3'突出，其从约1至约6个核苷酸长。在其它实施方案中，所述3'突出是从约1至约5个核苷酸，从约1至约3个核苷酸以及从约2至约4个核苷酸长。

[0133] 当干扰RNA分子在分子的一端包括3'突出时，另一端可以是平端或也具有突出(5'或3')。当干扰RNA分子在分子的两端包括突出时，突出的长度可以相同或不同。在一个实施方案中，干扰RNA分子在分子两端包括约1至约3个核苷酸的3'突出。在另一个实施方案中，干扰RNA分子是在分子两端具有2个核苷酸的3'突出的双链RNA。在又一个实施方案中，含有干扰RNA突出的核苷酸是TT二核苷酸或UU二核苷酸。

[0134] 当确定含有一个或多个突出的干扰RNA分子与靶mRNA序列的百分比同一性时，可以考虑或不用考虑突出。例如，来自3'突出的核苷酸以及来自双链的5'-或3'-末端的至多2个核苷酸可以经过修饰而没有显著损失小干扰RNA分子的活性。

[0135] 干扰RNA分子可以包括一个或多个5'或3'-帽结构。干扰RNA分子可以包括帽结构，所述帽结构在正义链、反义链、或正义和反义链两者的3'-端；或在干扰RNA分子的正义链、反义链、或正义和反义链两者的5'-端。或者，干扰RNA分子可以在干扰RNA分子的3'-端和

5'-端两者含有帽结构。术语“帽结构”指在寡核苷酸的任一末端并入的化学修饰，其保护分子免于外切核酸酶降解，并且还可有助于细胞内的递送或定位。

[0136] 适于干扰RNA分子的另一种修饰是将一个或多个半体或缀合物化学链接至干扰RNA分子，所述半体或缀合物增强干扰RNA分子的活性、细胞分布、细胞吸收、生物利用度或稳定性。所述多核苷酸可以通过本领域公认的方法进行合成或修饰。化学修饰可包括但不限于2'修饰、引入非天然碱基、与配体的共价附着、以及用硫代磷酯键替换磷酯键。在本实施方案中，双链体结构的完整性通过至少一个、以及典型地两个化学键得到加强。化学链接可以通过各种公知技术的任意者来实现，例如通过引入共价、离子或氢键；疏水相互作用、范德华力或堆积相互作用；通过金属离子配位的方式、或通过使用嘌呤类似物。

[0137] 在另一个实施方案中，在两条单链的一个或两个上的核苷酸可被修饰以减少或抑制细胞酶的激活，如例如但不限于某些核酸酶。用于减少或抑制细胞酶激活的技术是本领域已知的，包括但不限于2'-氨基修饰、2'-氟修饰、2'-烷基修饰、不带电荷的骨架修饰、吗啉基修饰、2'-0-甲基修饰和氨基磷酸酯。因此，在双链RNA上的核苷酸的至少一个2'-羟基基团被化学基团所取代。此外，至少一个核苷酸可被修饰以形成锁核苷酸。这种锁核苷酸包含连接核糖的2'-氧与核糖的4'-碳的亚甲基或亚乙基桥。将锁核苷酸引入到寡核苷酸改善了对互补序列的亲和性并将解链温度提高若干度。

[0138] 配体可缀合至干扰RNA分子，例如，以便增强其细胞吸收。在某些实施方案中，疏水性配体缀合至分子以利于细胞膜的直接渗透。使用这些方法以促进反义寡核苷酸的细胞渗透。在某些情况下，阳离子配体和寡核苷酸的缀合通常导致改善的对核酸酶的抗性。阳离子配体的代表性实例包括丙基铵和二甲基丙基铵。当阳离子配体分散在寡核苷酸时，反义寡核苷酸可以保留其对mRNA的高结合亲和性。

[0139] 根据一个实施方案，在已经能够产生含β-甲基戊酰基的蔗糖酯的植物中，异丙基苹果酸合酶的表达或活性是减少的，因此导致其内异丙基苹果酸合酶的积累减少，以及因此减少含β-甲基戊酰基的蔗糖酯的生产。

[0140] 本文所述的分子和核苷酸可使用公知的固相合成技术来制备。本领域已知用于这样合成的任何其它方法可以附加地或可选择地采用。

[0141] 各种实施方案针对含有一个或多个本文所述多核苷酸的表达载体、或含有一个或多个多核苷酸的干扰RNA构建体。

[0142] 各种实施方案针对含有一个或多个多核苷酸的表达载体、或一个或多个干扰RNA构建体。

[0143] 各种实施方案针对含有一个或多个多核苷酸的表达载体、或一个或多个干扰RNA构建体，所述构建体编码一个或多个能够自退火以形成发夹结构的干扰RNA多核苷酸，其中所述构建体包括：(a)一个或多个本文所述的多核苷酸；(b)编码间隔元件的第二序列，其包括剪接时形成发夹结构环的内含子；以及(c)含有第一序列的反向互补序列的第三序列，其定位在与第一序列相同的方向，其中第二序列定位在第一序列和第三序列之间，并且第二序列可操作地连接于第一序列和第三序列。

[0144] 可利用所述公开的序列来构建各种不形成发夹结构的多核苷酸。例如，通过(1)通过可操作地连接到第一启动子来转录所述DNA的第一链、和(2)通过可操作地连接到第二启动子来转录所述DNA片段的第一链的反向互补序列，可形成双链RNA。可以从同一表达载体，

或从不同表达载体转录出所述多核苷酸的每条链。具有RNA干扰活性的RNA双链体可以通过酶转化为小干扰RNA来降低RNA水平。

[0145] 因此,各种实施方案针对包含多核苷酸的表达载体、或编码能够自退火的干扰RNA多核苷酸的干扰RNA构建体,其中所述构建体包括(a)一个或多个本文所述的多核苷酸;和(b)含有第一序列的互补(例如,反向互补)序列的第二序列,其定位在与第一序列相同的方向。

[0146] 通过促进基因表达的共抑制来提供各种组合物和方法用于降低异丙基苹果酸合酶的内源表达水平。作为向植物细胞宿主中引入多拷贝转基因的结果而发生共抑制现象。多拷贝转基因的整合可导致转基因和靶向内源性基因的表达减少。共抑制程度依赖于转基因和靶向内源性基因之间序列同一性程度。内源性基因和转基因两者的沉默可以通过沉默的基因座(即,目标内源性启动子和内源性基因)的广泛甲基化而发生,其可以阻止转录。或者在某些情况下,内源性基因和转基因的共抑制可以通过转录后基因沉默发生,其中可以产生转录物,但增强的降解速度阻止转录物的积累。转录后基因沉默的共抑制机制被认为类似于RNA干扰,其中RNA似乎是这些过程中的重要的起始子(initiator)和靶,并且可以通过相同的分子机制、可能通过RNA引导的mRNA降解至少部分地被介导。

[0147] 可以通过将多个拷贝的核酸或其片段作为转基因整合到目标植物的基因组中来实现核酸的共抑制。使用含有可操作地连接到核酸或其片段的启动子的表达载体可转化宿主植物。各种实施方案针对用于促进内源性基因共抑制的表达载体,其包括可操作地连接到多核苷酸的启动子。

[0148] 也描述了用于获得保守突变多核苷酸和多肽的方法。任何目标植物包括植物细胞或植物材料,可以通过各种已知用于诱导诱变的方法进行遗传修饰,包括位点定向诱变、寡核苷酸定向诱变、化学诱导的诱变、辐射诱导的诱变、利用修饰碱基的诱变、利用缺口双链体DNA的诱变、双链断裂诱变、利用修复缺陷的宿主株的诱变、通过全基因合成的诱变、DNA改组和其它等同的方法。

[0149] 或者,可以通过向目标植物基因组中引入转座子(例如,IS元件)使基因靶向失活。可以通过有性杂交受精引入这些可移动的遗传元件,通过蛋白活性丢失筛选插入突变体。通过将亲本植物与未经转座子诱导诱变的植物进行杂交,例如通过有性杂交受精,可以将亲本植物中破坏的基因引入到其它植物中。可以利用本领域技术人员已知的任何标准育种技术。在一个实施方案中,可以通过插入一个或多个转座子使一个或多个基因失活。突变可导致一个或多个基因的纯合子破坏,导致一个或多个基因的杂合子破坏,或者如果破坏一个以上基因则导致纯合子破坏和杂合子破坏的组合。合适的可转座的元件包括反转录转座子、反转录子(retroposon)和SINE样元件。这样的方法是本领域的技术人员已知的。

[0150] 或者,可以通过引入源自一些能够自我切割并在植物中复制的小环状RNA的核酶靶向基因进行失活。这些RNA可以单独复制(类病毒RNA)或与辅助病毒(卫星RNA)一起复制。合适的RNA实例包括源自鳄梨日斑类病毒(avocado sunblotch viroid)的那些、以及源自烟草环斑病毒(tobacco ringspot virus)、紫花苜蓿短暂条纹病毒(lucerne transient streak virus)、绒毛烟斑驳病毒(velvet tobacco mottle virus)、茄属植物斑驳病毒(solanum nodiflorum mottle virus)和地三叶草斑驳病毒(subterranean clover mottle virus)的卫星RNA。各种靶RNA特异性核酶是本领域技术人员已知的。

[0151] 在一些实施方案中,通过非转基因方式(例如在基因中建立突变)来调节异丙基苹果酸合酶多肽的表达。在基因序列中随机引入突变的方法可以包括化学诱变、EMS诱变和辐射诱变。向细胞中引入一个或多个靶向突变的方法包括但不限于基因组编辑技术,特别是锌指核酸酶介导的诱变、tilling(靶向诱导的基因组局部损伤)、同源重组、寡核苷酸定向诱变以及大范围核酸酶介导的诱变。

[0152] 突变的一些非限制性实例是至少一个核苷酸的删除、插入和错义突变、单核苷酸多态性和简单序列重复。突变后,可以进行筛选以鉴定产生早期终止密码子或者其它非功能基因的突变。突变后,可以进行筛选以鉴定产生能够以升高的水平表达的功能基因的突变。通过测序、或通过使用对基因或蛋白特异的探针或引物可以进行突变体的筛选。也可以产生多核苷酸中的特异性突变,其导致减少的或增加的基因表达、减少或增加的mRNA稳定性、或者减少或增加的蛋白稳定性。这样的植物本文称为突变的植物。

[0153] 突变的植物可具有一个或多个突变的任意组合,突变导致调节的多肽水平。例如,突变植物可以具有在单基因中的单突变;在单基因中的多突变;在两个或更多或者三个或更多基因中的单突变;在两个或更多或者三个或更多基因中的多突变。因此,公开了含有突变多肽变体的突变植物。

[0154] 在一个实施方案中,来自植物的种子被诱变,然后生长成第一代突变植物。然后,允许第一代植物自花传粉,来自第一代植物的种子生长成第二代植物,然后筛选第二代植物基因座中的突变。虽然可以筛选诱变植物材料的突变,筛选第二代植物的优点是所有的体细胞突变对应于种系突变。本领域技术人员理解,为产生突变植物可诱变各种植物材料,包括但不限于种子、花粉、植物组织或植物细胞。然而,当筛选植物核酸的突变时,诱变的植物材料类型可能产生影响。例如,当在非诱变植物授粉前对花粉进行诱变时,所述授粉所产生的种子生长成第一代植物。第一代植物的每一个细胞将包含花粉中产生的突变;因此,随后可以筛选这些第一代植物的突变,而不是等到第二代。

[0155] 产生点突变、短的删除、插入、颠换和或转换的诱变剂,包括化学诱变剂或辐射,可用于产生突变。诱变剂包括但不限于甲磺酸乙酯、甲磺酸甲酯、N-乙基-N-亚硝基脲、三乙基三聚氰胺、N-甲基-N-亚硝基脲、甲基苄肼、苯丁酸氮芥、环磷酰胺、硫酸二乙酯、丙烯酰胺单体、美法仑、氮芥、长春新碱、二甲基亚硝胺、N-甲基-N'-硝基-亚硝基胍、亚硝基脲、2-氨基嘌呤、7,12二甲基-苯并蒽、环氧乙烷、六甲基磷酰胺、白消安(bisulfan)、二环氧烷烃(二环氧辛烷、二环氧丁烷等)、2-甲氧基-6-氯-9[3-(乙基-2-氯-乙基)氨丙基氨基]吖啶二盐酸盐和甲醛。只要它们产生期望的表型,也可以考虑不是由诱变剂直接造成的基因座中的自发突变。合适的诱变试剂还包括,例如电离辐射,如X-射线、 $\gamma$ 射线、快中子辐射和UV辐射。

[0156] 本领域技术人员已知的植物核酸制备的任何方法可用来制备用于突变筛选的植物核酸。

[0157] 来自个体植物、植物细胞、或植物材料的制备核酸可以任选地进行合并,以便加快筛选源自诱变的植物组织、细胞或材料的植物群体的异丙基苹果酸合酶基因中的突变。可以筛选植物、植物细胞或植物材料的一代或随后更多代。任选合并组的尺寸取决于所用筛选方法的灵敏度。

[0158] 任选地合并核酸样本之后,可对其进行多核苷酸特异性扩增技术,如聚合酶链式反应(PCR)。对异丙基苹果酸合酶基因特异的、或对异丙基苹果酸合酶基因紧邻的序列特异

的任何一个或多个引物或探针,可用于扩增任选合并的核酸样本中的序列。优选地,设计一个或多个引物用于扩增最有可能产生有用突变的异丙基苹果酸合酶基因座区域。最优选的是,设计引物用于检测异丙基苹果酸合酶多核苷酸区域内的突变。此外,优选引物以避免已知多态性位点以便于点突变的筛选。为方便扩增产物的检测,可以使用任何常规标记方法来标记一个或多个引物或探针。可以基于本文所述序列使用本领域公知的方法设计引物或探针。通过本领域已知的方法可鉴定多态性。在其它方面,提供制备突变植物的方法。方法涉及提供含有编码功能性异丙基苹果酸合酶多肽基因的植物的至少一个细胞。接着,在有效调节异丙基苹果酸合酶基因活性的条件下,处理所述植物的至少一个细胞。然后,将所述至少一个突变植物细胞繁殖成突变植物,与对照植物相比突变植物具有调节的异丙基苹果酸合酶多肽水平。在制成突变植物的这种方法的一个实施方案中,处理步骤涉及在有效产生至少一个突变植物细胞的条件下,让至少一个细胞经受如上所述的化学诱变试剂。在本方法的另一个实施方案中,处理步骤涉及在有效产生至少一个突变植物细胞的条件下,让至少一个细胞经受辐射源。术语“突变植物”包括突变植物,其中与对照植物相比,适当地通过遗传工程或遗传修饰以外的方式修饰突变植物的基因型。

[0159] 在某些实施方案中,突变植物、突变植物细胞或突变植物材料可包括一个或多个突变并赋予期望的性状,所述突变天然存在于另一植物、植物细胞或植物材料中。这种突变可以并入(例如,基因渗入)另一植物、植物细胞或植物材料(例如,具有与突变所源自植物的遗传背景不同的植物、植物细胞或植物材料)来赋予其性状。因此,举例来说,天然存在于第一植物中的突变可被引入第二植物,如具有与第一植物遗传背景不同的第二植物。因此,技术人员能够搜索和鉴定在其基因组中天然携带异丙基苹果酸合酶基因的一个或多个赋予期望性状的突变等位基因的植物。可以通过各种方法(包括育种、回交和基因渗入)将天然存在的突变等位基因转移到第二植物,以产生在异丙基苹果酸合酶基因中具有一个或多个突变的系、品种或杂种。可从突变植物的库中筛选出显示期望性状的植物。合适地,利用如本文所述的异丙基苹果酸合酶核苷酸序列的知识进行选择。因而,可能筛选指示蔗糖酯组合物与对照相比的改变的遗传性状。这种筛选方法可涉及如本文所述的常规核酸扩增和/或杂交技术的应用。因此,本发明的其它方面涉及用于鉴定突变植物的方法,所述方法包括步骤:(a)提供含有来自植物的异丙基苹果酸合酶多核苷酸的样本;以及(b)确定异丙基苹果酸合酶多核苷酸的核酸序列,其中所述异丙基苹果酸合酶多核苷酸与对照植物的异丙基苹果酸合酶多核苷酸相比在序列上的差异,指示所述植物是异丙基苹果酸合酶突变植物。另一个方面,提供鉴定突变植物的方法,所述突变植物与对照植物相比在蔗糖酯的组合物上有改变,所述方法包括步骤:(a)提供来自待筛选植物的样本;(b)确定所述样本是否含有在异丙基苹果酸合酶多核苷酸中的一个或多个突变;以及(c)确定在所述植物中的蔗糖酯的组合物;其中如果所述样本在异丙基苹果酸合酶多核苷酸上含有一个或多个突变(与对照植物相比,其调节编码蛋白的表达或活性),那么与对照植物(在其中异丙基苹果酸合酶的表达或活性并未减少)相比所述植物的部分在蔗糖酯的组合物上有改变则指示着在蔗糖酯组合物上有改变的突变植物。在另一方面中,提供用于制备突变植物的方法,所述突变植物和对照植物相比在蔗糖酯的组合物上有改变,所述方法包括步骤:(a)提供来自第一植物的样本;(b)确定所述样本是否在异丙基苹果酸合酶多核苷酸中含有一个或多个突变,所述突变导致样本中蔗糖酯的组合物的改变;以及(c)将一个或多个突变转移到第二植物

内。可以使用本领域已知的各种方法(如通过遗传工程、遗传操作、基因渗入、植物育种、回交等)将突变转移到第二植物中。在一个实施方案中,第一植物是天然存在的植物。在一个实施方案中,第二植物具有不同于第一植物的遗传背景。在另一方面,提供制备突变植物的方法,所述突变植物与对照植物相比在蔗糖酯的组合物上有改变,所述方法包括步骤:(a)提供来自第一植物的样本;(b)确定所述样本是否包含异丙基苹果酸合酶多核苷酸中的一个或多个突变,所述突变导致样本中蔗糖酯的组合物的改变;以及(c)将一个或多个突变从第一植物中基因渗入到第二植物。在一个实施方案中,基因渗入的步骤包括植物育种,任选地包括回交等。在一个实施方案中,第一植物是天然存在的植物。在一个实施方案中,第二植物具有不同于第一植物的遗传背景。在一个实施方案中,第一植物不是栽培品种或优良栽培品种。在一个实施方案中,第二植物是栽培品种或优良栽培品种。再一方面涉及通过本文所述方法获得或可获得的突变植物(包括栽培品种或优良栽培品种突变植物)。在某些实施方案中,突变植物可具有仅位于植物特定区域的一个或多个突变,如在异丙基苹果酸合酶多核苷酸序列内。根据本实施方案,突变植物的其余基因组序列和诱变之前的植物相同或基本上相同。

[0160] 在某些实施方案中,突变植物可以具有位于植物一个以上的区域(如异丙基苹果酸合酶多核苷酸的序列内)以及基因组的一个或多个其它区域中的一个或多个突变。根据本实施方案,突变植物的其余基因组序列和诱变之前的植物不相同或基本上不相同。在某些实施方案中,突变植物可以不在异丙基苹果酸合酶多核苷酸的一个或多个、两个或多个、三个或多个、四个或多个或五个或多个外显子中具有一个或多个突变;或者可以不在异丙基苹果酸合酶多核苷酸的一个或多个、两个或多个、三个或多个、四个或多个或五个或多个内含子中具有一个或多个突变;或者可以不在异丙基苹果酸合酶多核苷酸的启动子中具有一个或多个突变;或者可以不在异丙基苹果酸合酶3'非翻译区具有一个或多个突变;可以不在异丙基苹果酸合酶5'非翻译区具有一个或多个突变;可以不在异丙基苹果酸合酶编码区具有一个或多个突变;或可以不在异丙基苹果酸合酶非编码区具有一个或多个突变;或两个或多个、三个或多个、四个或多个、五个或多个;或六个或多个其部分的任意组合。在另一方面中,提供了鉴定在编码异丙基苹果酸合酶基因中含有突变的植物、植物细胞或植物材料的方法,包括:(a)使植物、植物细胞或植物材料经诱变;(b)从所述植物、植物细胞或植物材料或其后代获得核酸样本;以及(c)确定编码异丙基苹果酸合酶、或其变体或片段的基因的核酸序列,其中所述序列中的差异指示着其中一个或多个突变。

[0161] 锌指蛋白可用于调节异丙基苹果酸合酶的表达或活性。在各种实施方案中,通过锌指核酸酶介导的诱变来修饰含有部分或全部多核苷酸编码序列的基因组DNA序列。搜索基因组DNA序列以找寻锌指蛋白结合的独特位点。或者,搜索基因组DNA序列以找寻锌指蛋白结合的两个独特位点,其中两个位点在相对的链上并彼此靠近,例如,相隔1、2、3、4、5、6或更多个碱基对。因此,提供结合本文所述多核苷酸的锌指蛋白。

[0162] 可将锌指蛋白工程化以识别在异丙基苹果酸合酶基因上的选择的靶向位点。通过截短、或扩展、或定点诱变的方法结合选择方法(例如而非限制于噬菌体展示选择、细菌双杂交选择或细菌单杂交选择),锌指蛋白可含有源自天然锌指DNA结合结构域和非天然锌指DNA结合结构域的基序的任何组合。术语“非天然锌指DNA结合结构域”是指锌指DNA结合结构域,其结合靶DNA中的3碱基对序列并且不存在于含有待修饰DNA的细胞或生物中。锌指蛋

白的设计方法是本领域已知的，所述锌指蛋白结合对靶基因独特的特异核苷酸序列。

[0163] 可以通过制备第一多核苷酸和第二多核苷酸的融合体来构建锌指核酸酶，所述第一多核苷酸编码结合多核苷酸的锌指蛋白，以及第二多核苷酸编码非特异性核酸内切酶，例如但不限于IIS型核酸内切酶的那些。在锌指蛋白和所述核酸酶之间的融合蛋白可包括由两个碱基对组成的间隔区(spacer)、或者可替代地，所述间隔区可由3、4、5、6或7或更多碱基对所组成。在各种实施方案中，锌指核酸酶在多核苷酸的基因组DNA序列的调控区、编码区或非编码区引入一个双链断裂、并导致多核苷酸表达水平的降低、或者多核苷酸编码的蛋白活性的降低。锌指核酸酶的切割频繁导致在切割位点上的DNA删除，随后是通过非同源端连接的DNA修复。

[0164] 在其它实施方案中，可选择锌指蛋白来结合异丙基苹果酸合酶多核苷酸的调控序列。更具体地，所述调控序列可包含转录起始位点、起始密码子、外显子区、外显子-内含子边界、终止子或终止密码子。因此，本公开提供在异丙基苹果酸合酶基因附近或内部通过锌指核酸酶介导的诱变所产生的突变的、非天然存在的或转基因的植物或植物细胞，以及通过锌指核酸酶介导的诱变制成这种植物或植物细胞的方法。用于将锌指蛋白和锌指核酸酶递送至植物的方法与下述用于递送大范围核酸酶的那些相似。

[0165] 在另一方面，提供了使用大范围核酸酶(比如大范围核酸酶I-CreI)产生重组的、突变的、非天然存在的或转基因的或者其它遗传修饰的植物的方法。天然存在的大范围核酸酶以及重组大范围核酸酶可用于在植物基因组DNA中的单个位点或在相对较少位点上特异性造成双链断裂，从而允许异丙基苹果酸合酶多核苷酸的破坏。大范围核酸酶可以是具有改变的DNA识别性能的工程化的大范围核酸酶。可以通过本领域已知的各种不同机制将大范围核酸酶蛋白递送到植物细胞中。

[0166] 本公开包括大范围核酸酶在植物细胞或植物中失活一个或多个本文所述的多核苷酸的用途。方面还涉及使用大范围核酸酶失活植物中多核苷酸的方法，包括：(a) 提供含有一个或多个本文所述多核苷酸的植物细胞；(b) 将大范围核酸酶或编码大范围核酸酶的构建体引入所述植物细胞中；以及(c) 允许大范围核酸酶使所述多核苷酸基本上失活。

[0167] 大范围核酸酶可用于切割多核苷酸编码区内的大范围核酸酶识别位点。这种切割频繁导致大范围核酸酶识别位点的DNA删除，随后是通过非同源端连接的致突变的DNA修复。这种在基因编码序列中的突变通常足以使所述基因失活。首先，这种改变植物细胞的方法涉及使用合适的转化方法将大范围核酸酶表达盒递送至植物细胞。为获得最高效率，需要将大范围核酸酶表达盒连接至可选择的标志物，并在选择试剂存在下选择成功转化的细胞。这种方法将导致大范围核酸酶表达盒整合到基因组中，然而如果植物有可能需要监管批准，这种过方法则是不可取的。在这种情况下，可以使用常规育种技术将大范围核酸酶表达盒(以及相连的选择性标志物基因)在后代植物中分离(segregate)出来。或者，可以最初用缺乏选择性标志物的大范围核酸酶表达盒转化植物细胞，所述植物细胞可在缺乏选择试剂的介质中生长。在这种条件下，部分处理后的细胞将获得大范围核酸酶表达盒并瞬时表达工程化的大范围核酸酶而没有将大范围核酸酶表达盒整合到基因组中。因为没有考虑转化效率，所述后者转化过程需要筛选更多处理的细胞以获得所需的基因组修饰。当使用锌指蛋白或锌指核酸酶时，上述方法也可用于修饰植物细胞。

[0168] 递送大范围核酸酶表达盒之后，植物细胞最初在对于所用特定转化过程典型的条

件下生长。这可能意味着,通常在黑暗中在低于26度C的温度下的介质上,使转化细胞生长。这样的标准条件可使用一段时间,优选1-4天以允许植物细胞从转化过程中恢复。此初始恢复期后的任何点上,可提高生长温度以刺激工程化的大范围核酸酶的活性来切割并突变大范围核酸酶识别位点。

[0169] 对于某些应用,可能需要从植物基因组中精确除去多核苷酸。通过使用一对工程化的大范围核酸酶这种应用是可行的,其中每种大范围核酸酶切割位于所预期的删除各侧的大范围核酸酶识别位点。

[0170] 大范围核酸酶和其它DNA修复策略可用于触发在植物细胞所含的DNA和靶DNA之间的异源重组。靶DNA通常含有与植物细胞所含的DNA同源的至少一个、优选两个区。细胞修复DNA之后,同源端结合导致靶DNA被引入细胞内。

[0171] 在其它实施方案中,因为异丙基苹果酸合酶是在植物细胞核中编码的叶绿体蛋白,所以通过将所述基因直接插入到叶绿体基因组中可增加植物中的异丙基苹果酸合酶的表达水平。同样,提高异丙基苹果酸合酶蛋白转运或输入至叶绿体也可调节活性模式。

[0172] 一个目的是提供突变、转基因或非天然存在的植物,其显示出调节的蔗糖酯水平同时基本保持了与对照植物相同的视觉外观。因此,本文描述了突变、转基因或非天然存在的植物或细胞,其与对照细胞或对照植物相比具有调节的蔗糖酯水平。通过调节编码本文所述多核苷酸序列一个或多个多肽的表达,所述突变、转基因或非天然存在的植物或细胞已经被改变来调节本文所述的异丙基苹果酸合酶的合成或活性。

[0173] 另一方面,涉及突变、非天然存在的或转基因的植物或细胞,其中异丙基苹果酸合酶的表达或活性被调节,并且所述植物的部分与对照植物相比在蔗糖酯水平上具有至少5%的增加或减少,所述对照植物中所述酶的表达或活性未被调节。在又一方面,涉及突变、非天然存在的或转基因的植物或细胞,其中异丙基苹果酸合酶的表达或其所编码的蛋白的活性被调节,以及在其中烟雾(aerosol)中的蔗糖酯水平与对照植物的烟雾相比增加或减少了至少5%。

[0174] 与对照植物相比,在蔗糖酯含量上的改变可以是至少约5%、至少约10%、至少约20%、至少约25%、至少约30%、至少约40%、至少约50%、至少约60%、至少约70%、至少约75%、至少约80%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或约100%或更多。

[0175] 可加热植物至100°C或以上,如至少125°C、至少150°C、至少175°C或至少200°C,来释放烟雾。

[0176] 适用于遗传修饰的植物包括单子叶和双子叶植物以及植物细胞系统,包括来自下列科之一的种:爵床科(Acanthaceae)、葱科(Alliaceae)、六出花科(Alstroemeriaceae)、石蒜科(Amaryllidaceae)、夹竹桃科(Apocynaceae)、槟榔科(Arecaceae)、菊科(Asteraceae)、小檗科(Berberidaceae)、红木科(Bixaceae)、十字花科(Brassicaceae)、凤梨科(Bromeliaceae)、大麻科(Cannabaceae)、石竹科(Caryophyllaceae)、三尖杉科(Cephalotaxaceae)、藜科(Chenopodiaceae)、秋水仙科(Colchicaceae)、葫芦科(Cucurbitaceae)、薯蓣科(Dioscoreaceae)、麻黄科(Ephedraceae)、古柯科(Erythroxylaceae)、大戟科(Euphorbiaceae)、蝶形花科(Fabaceae)、唇形科(Lamiaceae)、亚麻科(Linaceae)、石松科(Lycopodiaceae)、锦葵科(Malvaceae)、黑药花科

(Melanthiaceae)、芭蕉科(Musaceae)、桃金娘科(Myrtaceae)、蓝果树科(Nyssaceae)、罂粟科(Papaveraceae)、松科(Pinaceae)、车前草科(Plantaginaceae)、禾本科(Poaceae)、蔷薇科(Rosaceae)、茜草科(Rubiaceae)、杨柳科(Salicaceae)、无患子科(Sapindaceae)、茄科(Solanaceae)、紫杉科(Taxaceae)、山茶科(Theaceae)或葡萄科(Vitaceae)。

[0177] 合适的种可包括黄秋葵(*Abelmoschus*)、冷杉(*Abies*)、槭(*Acer*)、翦股颖(*Agrostis*)、葱(*Allium*)、六出花(*Alstroemeria*)、凤梨(*Ananas*)、穿心莲(*Andrographis*)、须芒草(*Andropogon*)、蒿(*Artemisia*)、芦竹(*Arundo*)、颠茄(*Atropa*)、小檗(*Berberis*)、甜菜(*Beta*)、红木(*Bixa*)、芸苔(*Brassica*)、金盏菊(*Calendula*)、山茶花(*Camellia*)、喜树(*Camptotheca*)、大麻(*Cannabis*)、辣椒(*Capsicum*)、红花(*Carthamus*)、长春花(*Catharanthus*)、三尖杉(*Cephalotaxus*)、菊花(*Chrysanthemum*)、金鸡纳(*Cinchona*)、西瓜(*Citrullus*)、咖啡(*Coffea*)、秋水仙(*Colchicum*)、彩叶草(*Coleus*)、黄瓜(*Cucumis*)、南瓜(*Cucurbita*)、狗牙根(*Cynodon*)、曼陀罗(*Datura*)、石竹(*Dianthus*)、洋地黄(*Digitalis*)、薯蓣(*Dioscorea*)、油棕(*Elaeis*)、麻黄草(*Ephedra*)、斑茅(*Erianthus*)、吉柯(*Erythroxylum*)、桉树(*Eucalyptus*)、羊茅(*Festuca*)、草莓(*Fragaria*)、雪花莲(*Galanthus*)、大豆(*Glycine*)、棉花(*Gossypium*)、向日葵(*Helianthus*)、橡胶树(*Hevea*)、大麦(*Hordeum*)、天仙子(*Hyoscyamus*)、麻疯树(*Jatropha*)、莴苣(*Lactuca*)、亚麻(*Linum*)、黑麦草(*Lolium*)、羽扇豆(*Lupinus*)、番茄(*Lycopersicon*)、石松(*Lycopodium*)、木薯(*Manihot*)、苜蓿(*Medicago*)、薄荷(*Mentha*)、芒草(*Misanthus*)、香蕉(*Musa*)、烟草(*Nicotiana*)、水稻(*Oryza*)、黍(*Panicum*)、罂粟(*Papaver*)、银胶菊(*Parthenium*)、狼尾草(*Pennisetum*)、矮牵牛(*Petunia*)、𬟁草(*Phalaris*)、猫尾(*Phleum*)、松(*Pinus*)、早熟禾(*Poa*)、一品红(*Poinsettia*)、胡杨(*Populus*)、萝芙木(*Rauwolfia*)、蓖麻(*Ricinus*)、蔷薇(*Rosa*)、甘蔗(*Saccharum*)、柳(*Salix*)、血根草(*Sanguinaria*)、莨菪(*Scopolia*)、黑麦(*Secale*)、茄(*Solanum*)、高粱(*Sorghum*)、米草(*Spartina*)、菠菜(*Spinacea*)、艾菊(*Tanacetum*)、紫杉(*Taxus*)、可可(*Theobroma*)、小黑麦(*Triticosecale*)、小麦(*Triticum*)、北美穗草(*Uniola*)、藜芦(*Veratrum*)、长春花(*Vinca*)、葡萄(*Vitis*)和玉米(*Zea*)属的成员。

[0178] 合适的种可包括黍(*Panicum spp.*)、高粱(*Sorghum spp.*)、芒(*Misanthus spp.*)、甘蔗(*Saccharum spp.*)、蔗茅(*Erianthus spp.*)、杨(*Populus spp.*)、大须芒草(*Andropogon gerardii*) (大须芒草(big bluestem))、象草(*Pennisetum purpureum*) (象草(elephant grass))、𬟁草(*Phalaris arundinacea*) (𬟁草(reed canarygrass))、狗牙根(*Cynodon dactylon*) (狗牙根(bermudagrass))、高羊茅(*Festuca arundinacea*) (高羊茅(tall fescue))、草原索草(*Spartina pectinata*) (草原绳草(prairie cord-grass))、紫花苜蓿(*Medicago sativa*) (紫花苜蓿(alfalfa))、芦竹(*Arundo donax*) (芦竹(giant reed))、黑麦(*Secale cereale*) (黑麦(rye))、柳(*Salix spp.*) (柳(willow))、桉树(*Eucalyptus spp.*) (桉树(eucalyptus))、小黑麦(*Triticosecale*)、竹(bamboo)、向日葵(*Helianthus annuus*) (向日葵(sunflower))、红花(*Carthamus tinctorius*) (红花(safflower))、麻风(*Jatropha curcas*) (麻风(jatropha))、蓖麻(*Ricinus communis*) (蓖麻(castor))、油棕(*Elaeis guineensis*) (油棕(palm))、亚麻(*Linum usitatissimum*) (亚麻(flax))、芥菜(*Brassica juncea*)、甜菜(*Beta vulgaris*) (甜菜(sugarbeet))、木薯(*Manihot esculenta*) (木薯(cassaya))、番茄(*Lycopersicon esculentum*) (西红柿)

(tomato))、莴苣 (*Lactuca sativa*) (莴苣 (lettuce))、香蕉 (*Musa paradisiaca*) (香蕉 (banana))、马铃薯 (*Solanum tuberosum*) (马铃薯 (potato))、甘蓝 (*Brassica oleracea*) (西兰花 (broccoli))、花椰菜 (cauliflower)、芽甘蓝 (*Brussels sprouts*)、茶 (*Camellia sinensis*) (茶 (tea))、草莓 (*Fragaria ananassa*) (草莓 (strawberry))、可可 (*Theobroma cacao*) (可可 (cocoa))、咖啡 (*Coffea arabica*) (咖啡 (coffee))、葡萄 (*Vitis vinifera*) (葡萄 (grape))、菠萝 (*Ananas comosus*) (菠萝 (pineapple))、辣椒 (*Capsicum annum*) (辣甜椒 (hot&sweet pepper))、洋葱 (*Allium cepa*) (洋葱 (onion))、甜瓜 (*Cucumis melo*) (甜瓜 (melon))、黄瓜 (*Cucumis sativus*) (黄瓜 (cucumber))、南瓜 (*Cucurbita maxima*) (南瓜 (squash))、南瓜 (*Cucurbita moschata*) (南瓜 (squash))、菠菜 (*Spinacea oleracea*) (菠菜 (Spinacea))、西瓜 (*Citrullus lanatus*) (西瓜 (watermelon))、秋葵 (*Abelmoschus esculentus*) (秋葵 (okra))、茄子 (*Solanum melongena*) (茄子 (eggplant))、蔷薇 (*Rosa spp.*) (蔷薇 (rose))、康乃馨 (*Dianthus caryophyllus*) (康乃馨 (carnation))、矮牵牛 (*Petunia spp.*) (矮牵牛 (petunia))、一品红 (*Poinsettia pulcherrima*) (一品红 (poinsettia))、羽扇豆 (*Lupinus albus*) (羽扇豆 (lupin))、燕麦 (*Uniola paniculata*) (燕麦 (oats))、翦股颖 (bentgrass) (翦股颖 (*Agrostis spp.*))、山杨 (*Populus tremuloides*) (山杨 (aspen))、松 (*Pinus spp.*) (松 (pine))、冷杉 (*Abies spp.*) (冷杉 (fir))、槭 (*Acer spp.*) (枫 (maple))、大麦 (*Hordeum vulgare*) (大麦 (barley))、早熟禾 (*Poa pratensis*) (早熟禾 (bluegrass))、黑麦草 (*Lolium spp.*) (黑麦草 (ryegrass)) 和猫尾草 (*Phleum pratense*) (猫尾草 (timothy))、柳 (*Panicum virgatum*) (柳 (switchgrass))、高粱 (*Sorghum bicolor*) (高粱 (sorghum))、苏丹草 (sudangrass)、芒草 (*Miscanthus giganteus*) (芒草 (miscanthus))、甘蔗 (*Saccharum sp.*) (甘蔗 (energycane))、白杨 (*Populus balsamifera*) (白杨 (poplar))、玉米 (*Zea mays*) (玉米 (corn))、大豆 (*Glycine max*) (大豆 (soybean))、油菜 (*Brassica napus*) (油菜 (canola))、小麦 (*Triticum aestivum*) (小麦 (wheat))、棉花 (*Gossypium hirsutum*) (棉花 (cotton))、水稻 (*Oryza sativa*) (水稻 (rice))、向日葵 (*Helianthus annuus*) (向日葵 (sunflower))、紫花苜蓿 (*Medicago sativa*) (紫花苜蓿 (alfalfa))、甜菜 (*Beta vulgaris*) (甜菜 (sugarbeet)) 或者珍珠粟 (*Pennisetum glaucum*) (珍珠粟 (pearl millet))。

[0179] 各种实施方案针对突变植物、非天然存在植物或转基因植物，其经过修饰以减少异丙基苹果酸合酶基因表达水平，从而产生这样的植物如烟草植物，在其中目标植物组织内的异丙基苹果酸合酶表达水平和对照植物相比是减少的。公开的组合物和方法可应用于烟草属的任何种，包括黄花烟草 (*N.rustica*) 和红花烟草 (*N.tabacum*) (例如，LA B21、LN KY171、TI1406、巴斯玛 (Basma)、Galpao (无对应译文)、珀里克 (Perique)、Beinhart1000-1、K326，希克斯阔叶 (Hicks Broadleaf) 和Petico (无对应译文))。其它物种包括 *N.acaulis* (无对应译文)、智利尖叶烟草 (*N.acuminata*)、*N.acuminata* var. *multiflora* (无对应译文)、非洲烟草 (*N.africana*)、花烟草 (*N.alata*)、抱茎烟草 (*N.amplexicaulis*)、*Narentsii* (无对应译文)、*N.attenuata* (无对应译文)、*N.benavidesii* (无对应译文)、本塞姆氏烟草 (*N.benthamiana*)、*N.bigelovii* (无对应译文)、*N.bonariensis* (无对应译文)、*N.cavicola* (无对应译文)、克氏烟草 (*N.clevelandii*)、心叶烟草 (*N.cordifolia*)、*N.corymbosa* (无对应译文)、*N.debneyi* (无对应译文)、*N.excelsior* (无对应译文)、福氏烟草 (*N.forgetiana*)、

*N.fragrans* (无对应译文)、粉蓝烟草 (*N.glauca*)、粘性烟草 (*N.glutinosa*)、*N.goodspeedii* (无对应译文)、野生烟草 (*N.gossei*)、*N.hybrid* (无对应译文)、*N.ingulba* (无对应译文)、*N.kawakamii* (无对应译文)、*N.knightiana* (无对应译文)、郎氏烟草 (*N.langsdorffii*)、*N.linearis* (无对应译文)、长花烟草 (*N.longiflora*)、海滨烟草 (*N.maritima*)、麦格隆熄丰烟草 (*N.megalosiphon*)、*N.miersii* (无对应译文)、夜花烟草 (*N.noctiflora*)、*N.nudicaulis* (无对应译文)、*N.obtusifolia* (无对应译文)、西方烟草 (*N.occidentalis*)、*N.occidentalis* subsp.*hesperis* (无对应译文)、耳状烟草 (*N.otophora*)、圆锥烟草 (*N.paniculata*)、*N.pauciflora* (无对应译文)、矮牵牛状烟草 (*N.petunioides*)、蓝茉莉叶烟草 (*N.plumbaginifolia*)、*N.quadrivalvis* (无对应译文)、*N.raimondii* (无对应译文)、波缘烟草 (*N.repanda*)、莲座叶烟草 (*N.rosulata*)、*N.rosulata* subsp.*ingulba* (无对应译文)、*N.rotundifolia* (无对应译文)、*N.setchellii* (无对应译文)、*N.simulans* (无对应译文)、*N.solanifolia* (无对应译文)、*N.spegazzinii* (无对应译文)、斯托克通氏烟草 (*N.stocktonii*)、*N.suaveolens* (无对应译文)、种林烟草 (*N.sylvestris*)、*N.thyrsiflora* (无对应译文)、绒毛烟草 (*N.tomentosa*)、拟茸毛烟草 (*N.tomentosiformis*)、*N.trigonophylla* (无对应译文)、*N.umbratica* (无对应译文)、波缘烟草 (*N.undulata*)、*N.velutina* (无对应译文)、*N.wigandiodoides* (无对应译文) 和 *N.× sanderae* (无对应译文)。在高度优选的实施方案中,所述植物是烟草植物,例如烟草属的植物或红花烟草种。

[0180] 本文也考虑使用烟草栽培品种和优良烟草栽培品种。因此转基因、非天然存在的或突变的植物可以是烟草品种或优良烟草栽培品种,其包括一个或多个转基因或一个或多个遗传突变或其组合。遗传突变(例如,一个或多个多态性)可以是不天然存在于个体烟草品种或烟草栽培品种(例如,优良烟草栽培品种)的突变,或只要突变不天然存在于个体烟草品种或烟草栽培品种(例如,优良烟草栽培品种)中,所述遗传突变可以是天然存在的遗传突变。特别有用的红花烟草品种包括白肋烟型、深型、烤烟型和东方型烟草。品种或栽培品种的非限制性实例是:BD64、CC101、CC200、CC27、CC301、CC400、CC500、CC600、CC700、CC800、CC900、Coker176 (无对应译文)、Coker319 (无对应译文)、Coker371Gold (无对应译文)、Coker48 (无对应译文)、CD263、DF911、DT538LC Galpao烟草、GL26H、GL350、GL600、GL737、GL939、GL973、HB04P、HB04P LC、HB3307PLC、杂交403LC、杂交404LC、杂交501LC、K149、K326、K346、K358、K394、K399、K730、KDH959、KT200、KT204LC、KY10、KY14、KY160、KY17、KY171、KY907、KY907LC、KTY14xL8LC、Little Crittenden (无对应译文)、McNair373、McNair944、msKY14xL8、窄叶Madole (无对应译文)、窄叶Madole LC、NBH98、N-126、N-777LC、N-7371LC、NC100、NC102、NC2000、NC291、NC297、NC299、NC3、NC4、NC5、NC6、NC7、NC606、NC71、NC72、NC810、NC BH129、NC2002、Neal Smith Madole (无对应译文)、OXFORD207、PD7302LC、PD7309LC、PD7312LC、“珀里克”烟草、PVH03、PVH09、PVH19、PVH50、PVH51、R610、R630、R7-11、R7-12、RG17、RG81、RG H51、RGH4、RGH51、RS1410、Speight168 (无对应译文)、Speight172 (无对应译文)、Speight179 (无对应译文)、Speight210 (无对应译文)、Speight220 (无对应译文)、Speight225 (无对应译文)、Speight227 (无对应译文)、Speight234 (无对应译文)、Speight G-28 (无对应译文)、Speight G-70 (无对应译文)、Speight H-6 (无对应译文)、Speight H20 (无对应译文)、Speight NF3 (无对应译文)、TI1406、TI1269、TN86、TN86LC、

TN90、TN97、TN97LC、TN D94、TN D950、TR (Tom Rosson) Madole (无对应译文)、VA309、VA359、AA37-1、B13P、Xanthi (Mitchell-Mor (无对应译文))、Bel-W3、79-615、沙姆逊福尔摩斯NN、KTRDC2号杂交49、白肋21、KY8959、KY9、MD609、PG01、PG04、P01、P02、P03、RG11、RG8、VA509、AS44、Banket A1 (无对应译文)、巴斯玛Drama B84/31 (无对应译文)、巴斯玛I Zichna ZP4/B (无对应译文)、巴斯玛Xanthi BX2A (无对应译文)、Batek (无对应译文)、Besuki Jember (无对应译文)、C104、Coker347、Criollo Misionero (无对应译文)、Delcrest (无对应译文)、Djebel181 (无对应译文)、DVH405、Galpão Comum (无对应译文)、HB04P、希克斯阔叶、Kabakulak Elassona (无对应译文)、Kutsage E1 (无对应译文)、LA BU21、NC2326、NC297、PVH2110、红色俄罗斯 (Red Russian)、Samsun (无对应译文)、Saplak (无对应译文)、Simmaba (无对应译文)、Talgar28 (无对应译文)、Wislica (无对应译文)、Yayaldag (无对应译文)、Prilep HC-72 (无对应译文)、Prilep P23 (无对应译文)、Prilep PB156/1 (无对应译文)、Prilep P12-2/1 (无对应译文)、Yaka JK-48 (无对应译文)、Yaka JB125/3 (无对应译文)、TI-1068、KDH-960、TI-1070、TW136、巴斯玛、TKF4028、L8、TKF2002、GR141、巴斯玛xanthi (无对应译文)、GR149、GR153、Petit Havana (无对应译文)。即使本文未特别指明，也考虑上述低转化亚变种 (low converter subvariety)。在某些实施方案中，栽培品种和优良栽培品种是优选的。所述红花烟草种可用在某些其它实施方案中。可使用红花烟草白肋型。

[0181] 实施方案还针对用于产生突变、非天然存在的、或转基因植物的组合物和方法，其中植物已被修饰以调节异丙基苹果酸合酶表达或活性，从而导致植物或植物成分与对照相比具有蔗糖酯的调节水平。在植物中调节蔗糖酯水平可用于生成从其中生成的烟雾风味谱被改变的植物。因此，本发明可提供新的混合机会以生成期望的烟草风味。因此，例如通过增加异丙基苹果酸合酶的表达或活性可增加植物的蔗糖酯含量，其可导致植物能够释放升高水平的与东方烟草更相似的风味分子。因此，再例如，通过降低异丙基苹果酸合酶的表达或活性可降低植物的蔗糖酯含量，其可导致植物能够释放较低水平的与烤烟更相似的风味分子。

[0182] 在再一方面，提供调配烟草的方法，包括步骤：在烟草调配中替换或添加一种或更多类型的烟草，所述烟草 (例如，烟叶、或者源自或可源自烟叶的组合物) 源自转基因、突变或非天然存在的植物，所述植物中的植物蔗糖酯含量如本文所述通过调节 (优选，增加) 异丙基苹果酸合酶的表达或活性已经得到调节 (例如增加或减少)。在另一方面，提供调配烟草的方法，包括步骤：在烟草调配中替换、减少或省略东方类型的烟草；而添加源自转基因、突变或非天然存在植物的烟草 (例如，烟叶、或者源自或可源自烟叶的组合物)，所述植物中的植物蔗糖酯含量如本文所述通过调节 (优选，增加) 异丙基苹果酸合酶的表达或活性已经得到调节 (优选，增加)。烟草如东方烟草的替换、减少或添加可对应于被替换、减少或添加的烟草的约10%、20%、30%、40%、50%、60%、70%、80%、90%或100%。

[0183] 在又一方面涉及调节烟草或烟草制品风味的方法，包括：加入烟草 (例如，烟叶、或者源自或可源自烟叶的组合物) 或烟草制品，所述烟草源自或可源自转基因、突变或非天然存在的植物，如本文所述植物中的植物蔗糖酯含量已经得到调节。

[0184] 在还一方面涉及使用调节异丙基苹果酸合酶的表达或活性的试剂来调节烟草或烟草制品的风味。

[0185] 调节在植物中的蔗糖酯水平还可用于生成其中植物抗虫性被改变的植物。通过如

本文所述地增加蔗糖酯生产水平可增加植物的抗虫性。通过如本文所述地减少蔗糖酯生产水平可减少植物的抗虫性。

[0186] 如本文所述的在植物中产生的蔗糖酯的组合物可从其中提取和可选地纯化以用于各种其它用途,例如在药物、食品添加剂、吸烟风味剂(smoking flavourant)以及作为有机杀虫剂的成分等等。在一个实施方案中,从叶或叶表面提取蔗糖酯。因此,本公开的另一方面是产品,例如含有蔗糖酯的药物、食品添加剂、吸烟风味剂以及有机杀虫剂。

[0187] 根据本公开的蔗糖酯具有图5所示的通用结构,其中R1、R2、R3、R4和R5是氢原子或酰基链。在蔗糖酯上可被酯化的酰基链如图6所示。一些具体蔗糖酯的实例如表1和图2所示。

[0188] 在一个或多个蔗糖酯中每一个的R1、R2和R4上的酰基链中的碳总数可以是,例如12、13、14、15、16、17或18;适宜地,14、15、16、17或18;适宜地,16、17或18;或适宜地,14或15。如果这个总数的碳存在于每个蔗糖酯上,那么组成的酰基链可选自如图6所示那些中的任意者。适宜地,所述蔗糖酯在R3上具有乙酰基、以及在R5上具有乙酰基或氢。适宜地,所述蔗糖酯在R3上具有乙酰基、以及在R5上具有氢。适宜地,所述酰基链的一个或更多、两个或更多、或者三个或更多将是 $\beta$ -甲基戊酰基,特别是当6碳酰基链存在时。

[0189] 在一个实施方案中,所述蔗糖酯是这样的蔗糖酯,其在R3和R5上具有乙酰基或氢、以及在R1、R2和R4位的任意者上具有3、4、5或6碳的酰基链。适宜地,所述蔗糖酯在R3上具有乙酰基、以及在R5上具有乙酰基或氢。适宜地,所述蔗糖酯在R3上具有乙酰基、以及在R5上具有氢。在另一个实施方案中,所述蔗糖酯在R3和R5上具有乙酰基或氢、以及在R1、R2和R4位的任意者上具有4、5或6碳的酰基链。适宜地,所述蔗糖酯具有在R3上的乙酰基、以及在R5上具有乙酰基或氢。在又一个实施方案中,所述蔗糖酯在R3和R5上具有乙酰基或氢、以及在R1、R2和R4位的任意者上具有5或6碳的酰基链。适宜地,所述蔗糖酯在R3上的乙酰基、以及在R5上具有乙酰基或氢。适宜地,所述蔗糖酯在R3上具有乙酰基、以及在R5上具有氢。在还有一个实施方案中,所述蔗糖酯在R3和R5上具有乙酰基或氢、以及在R1、R2和R4位上具有6碳的酰基链。适宜地,所述蔗糖酯在R3上具有乙酰基、以及在R5上具有乙酰基或氢。适宜地,所述蔗糖酯在R3上具有乙酰基、以及在R5上具有氢。

[0190] 在一个实施方案中,所述蔗糖酯在R3和R5上具有乙酰基或氢、以及在R1位上具有3、4、5或6碳的酰基链、在R2位上具有3、4、5或6碳的酰基链和在R4位上具有3、4、5或6碳的酰基链。适宜地,所述蔗糖酯在R3上具有乙酰基、以及R5上具有乙酰基或氢。适宜地,所述蔗糖酯在R3上具有乙酰基、以及在R5上具有氢。在另一个实施方案中,所述蔗糖酯在R3和R5上具有乙酰基或氢、以及在R1位上具有4、5或6碳的酰基链、在R2位上具有4、5或6碳的酰基链和在R4位上具有4、5或6碳的酰基链。适宜地,所述蔗糖酯在R3上具有乙酰基、以及R5上具有乙酰基或氢。适宜地,所述蔗糖酯在R3上具有乙酰基、以及在R5上具有氢。在又一个实施方案中,所述蔗糖酯在R3和R5上具有乙酰基或氢、以及在R1位上具有5或6碳的酰基链、在R2位上具有5或6碳的酰基链和在R4位上具有5或6碳的酰基链。适宜地,所述蔗糖酯在R3上具有乙酰基、以及在R5上具有乙酰基或氢。适宜地,所述蔗糖酯在R3上具有乙酰基、以及在R5上具有氢。

[0191] 在一个实施方案中,所述蔗糖酯在R3和R5上具有乙酰基或氢、以及至少在R1、R2和

R4位之一上具有6碳的酰基链。适宜地，所述蔗糖酯在R3上具有乙酰基、以及R5上具有乙酰基或氢。适宜地，所述蔗糖酯在R3上具有乙酰基、以及在R5上具有氢。

[0192] 在一个实施方案中，所述蔗糖酯在R3和R5上具有乙酰基或氢、以及在R1位上具有6碳的酰基链、在R2位上具有6碳的酰基链和在R4位上具有6碳的酰基链。适宜地，所述蔗糖酯在R3上具有乙酰基、以及R5上具有乙酰基或氢。适宜地，所述蔗糖酯在R3上具有乙酰基、以及在R5上具有氢。在一个实施方案中，所述蔗糖酯是含有(支链) $\beta$ -甲基戊酰基的蔗糖酯。

[0193] 在一个实施方案中，所述蔗糖酯具有图5所示的通用结构，其中R1至R5的每一个是这样的酰基链中的任一个，所述酰基链选自乙酰基、丁酰基、丙酰基或其同分异构体、异丁酰基、戊酰基(戊酰基)、2-甲基-丁酰基、异戊酰基、异戊烯酸酰基、戊烯酸酰基、己酰基或其同分异构体、2-甲基戊酰基、 $\beta$ -甲基戊酰基和4-甲基戊酰基。在另一个实施方案中，蔗糖酯具有图5所示的通用结构，并且其中R3是乙酰基以及R1、R2、R4和R5中的每一个是这样的酰基链中的任一个，所述酰基链选自乙酰基、丁酰基、丙酰基或其同分异构体、异丁酰基、戊酰基(戊酰基)、2-甲基-丁酰基、异戊酰基、异戊烯酸酰基、戊烯酸酰基、己酰基或其同分异构体、2-甲基戊酰基、 $\beta$ -甲基戊酰基和4-甲基戊酰基。在一个实施方案中，优选6碳的酰基链是支链。在另一个实施方案中，优选6碳的酰基链是 $\beta$ -甲基戊酰基。

[0194] 在另一个实施方案中，蔗糖酯具有图5所示的通用结构，并且其中R3是氢或乙酰基、R5是氢或乙酰基、以及R1、R2和R4中的每一个是这样的酰基链中的任一个，所述酰基链选自乙酰基、丁酰基、丙酰基或其同分异构体、异丁酰基、戊酰基(戊酰基)或其同分异构体、2-甲基-丁酰基、异戊酰基、异戊烯酸酰基、戊烯酸酰基、己酰基或其同分异构体、2-甲基戊酰基、 $\beta$ -甲基戊酰基和4-甲基戊酰基。个体蔗糖酯公开在表1和图2中。

[0195] 在另一个实施方案中，蔗糖酯具有图5所示的通用结构，并且其中R3是乙酰基、R5是氢或乙酰基、以及R1、R2和R4中的每一个是这样的酰基链中的任一个，所述酰基链选自乙酰基、丁酰基、丙酰基或其同分异构体、异丁酰基、戊酰基(戊酰基)或其同分异构体、2-甲基-丁酰基、异戊酰基、异戊烯酸酰基、戊烯酸酰基、己酰基或其同分异构体、2-甲基戊酰基、 $\beta$ -甲基戊酰基和4-甲基戊酰基。

[0196] 在另一个实施方案中，蔗糖酯具有图5所示的通用结构，并且其中R3是乙酰基、R5是氢或乙酰基；R1是丁酰基、丙酰基或其同分异构体、戊酰基(戊酰基)或其同分异构体、或者己酰基或其同分异构体；R2是丙酰基或其同分异构体、戊酰基(戊酰基)或其同分异构体、或者己酰基或其同分异构体；以及R4是戊酰基(戊酰基)或其同分异构体、或者己酰基或其同分异构体。

[0197] 在另一个实施方案中，蔗糖酯具有图5所示的通用结构，并且其中R3是乙酰基、R5是氢或乙酰基；R1是丙酰基或其同分异构体、戊酰基(戊酰基)或其同分异构体、或者己酰基或其同分异构体；R2是戊酰基(戊酰基)或其同分异构体、或者己酰基或其同分异构体；以及R4是己酰基或其同分异构体。

[0198] 在一个实施方案中，得到调节的蔗糖酯具有图5所示的通用结构，其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=丙酰基或其同分异构体、R4=己酰基或其同分异构体以及R5是氢或乙酰基；或者蔗糖酯如图5所示，其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=戊酰基或其同分异构体以及R5是氢或乙酰基；或者蔗糖酯如图5所示，其中R3=乙酰基、R1=丁酰基、R2=戊酰基或其同分异构体、R4=己酰基或

其同分异构体以及R5是氢或乙酰基(C<sub>2</sub>C<sub>14</sub>:0)。

[0199] 在一个实施方案中,蔗糖酯具有图5所示的通用结构,其中R<sub>3</sub>=乙酰基、R<sub>1</sub>=丙酰基或其同分异构体、R<sub>2</sub>=戊酰基或其同分异构体、R<sub>4</sub>=己酰基或其同分异构体以及R<sub>5</sub>是氢或乙酰基;或者蔗糖酯如图5所示,其中R<sub>3</sub>=乙酰基、R<sub>1</sub>=戊酰基或其同分异构体、R<sub>2</sub>=戊酰基或其同分异构体、R<sub>4</sub>=戊酰基或其同分异构体以及R<sub>5</sub>是氢或乙酰基;或者蔗糖酯如图5所示,其中R<sub>3</sub>=乙酰基、R<sub>1</sub>=丁酰基、R<sub>2</sub>=己酰基或其同分异构体、R<sub>4</sub>=己酰基或其同分异构体以及R<sub>5</sub>是氢或乙酰基(C<sub>12</sub>C<sub>15</sub>:0)。

[0200] 在一个实施方案中,蔗糖酯具有图5所示的通用结构,并且其中R<sub>3</sub>是乙酰基、R<sub>1</sub>是戊酰基或其同分异构体、R<sub>2</sub>是戊酰基或其同分异构体、R<sub>4</sub>是己酰基或其同分异构体以及R<sub>5</sub>是氢原子或乙酰基半体(C<sub>2</sub>C<sub>16</sub>:0)。

[0201] 在一个实施方案中,被调节的蔗糖酯具有图5所示的通用结构,并且其中R<sub>3</sub>是乙酰基、R<sub>1</sub>是丙酰基或其同分异构体、R<sub>2</sub>是己酰基或其同分异构体、R<sub>4</sub>是己酰基或其同分异构体以及R<sub>5</sub>是氢原子或乙酰基半体(C<sub>2</sub>C<sub>16</sub>:0)。

[0202] 在一个实施方案中,蔗糖酯具有图5所示的通用结构,并且其中R<sub>3</sub>是乙酰基、R<sub>1</sub>是戊酰基或其同分异构体、R<sub>2</sub>是己酰基或其同分异构体、R<sub>4</sub>是己酰基或其同分异构体、以及R<sub>5</sub>是氢原子或乙酰基半体(C<sub>2</sub>C<sub>17</sub>:0)。

[0203] 在一个实施方案中,被调节的蔗糖酯具有图5所示的通用结构,并且其中R<sub>3</sub>是乙酰基、R<sub>1</sub>是己酰基或其同分异构体、R<sub>2</sub>是己酰基或其同分异构体、R<sub>4</sub>是己酰基或其同分异构体、以及R<sub>5</sub>是氢原子或乙酰基半体(C<sub>2</sub>C<sub>18</sub>:0)。

[0204] 在另一个实施方案中,所述蔗糖酯具有从约594道尔顿至692道尔顿的分子量。在一个实施方案中,所述蔗糖酯具有从约608道尔顿至692道尔顿的分子量。在一个实施方案中,所述蔗糖酯具有从约622道尔顿至692道尔顿的分子量。在一个实施方案中,所述蔗糖酯具有从约636道尔顿至692道尔顿的分子量。在一个实施方案中,所述蔗糖酯具有从约650道尔顿至692道尔顿的分子量。在另一个实施方案中,所述蔗糖酯具有从约594道尔顿至678道尔顿的分子量。在一个实施方案中,所述蔗糖酯具有从约608道尔顿至678道尔顿的分子量。在一个实施方案中,所述蔗糖酯具有从约622道尔顿至678道尔顿的分子量。在一个实施方案中,所述蔗糖酯具有从约636道尔顿至678道尔顿的分子量。在一个实施方案中,所述蔗糖酯具有从约650道尔顿至678道尔顿的分子量。在一个实施方案中,所述蔗糖酯具有从约664道尔顿至678道尔顿的分子量。

[0205] 另一方面涉及调节烟草或烟草制品风味的方法,包括向其中添加此处所述的蔗糖酯组合物。

[0206] 在另一个实施方案中,在植物中所述蔗糖酯的浓度、水平、或产量增加了。在另一个实施方案中,在植物中所述蔗糖酯的浓度、水平、或产量减少了。

[0207] 在一个具体实施方案中,提供用于产生突变、非天然存在的或转基因植物的组合物和方法,其中所述植物已被修饰以增加异丙基苹果酸合酶表达或活性,这可导致其中蔗糖酯的浓度、水平、或产量增加了的植物或植物材料。具体地,本文所述的蔗糖酯的水平或产量是增加的或诱导的。

[0208] 在另一个具体实施方案中,提供用于产生突变、非天然存在的或转基因植物的组

合物和方法,其中所述植物已被修饰以减少或抑制异丙基苹果酸合酶表达或活性,这可导致其中蔗糖酯的浓度、水平、或产量减少了的植物或植物材料。具体地,本文所述的蔗糖酯的水平是减少的或废除的。

[0209] 有利地,根据本文所述方法获得的突变、非天然存在的或转基因植物与对照植物在视觉外观上相似或基本相同。在一个实施方案中,突变、非天然存在的或转基因植物的秆高(stalk height)与对照植物基本相同,例如在田间移植后一、二、或三或更多个月、或者在去稍后10、20、30或36或更多天。例如,突变、非天然存在的或转基因植物的秆高与不低于对照植物基本的秆高。在另一个实施方案中,突变、非天然存在的或转基因植物的叶绿素含量与对照植物基本相同。在另一个实施方案中,突变、非天然存在的或转基因植物的秆高与对照植物基本相同、并且突变、非天然存在的或转基因植物的叶绿素含量与对照植物基本相同。在其它实施方案中,突变、非天然存在的或转基因植物的叶子的尺寸、或形状、或数目、或颜色与对照植物基本相同。适宜地,所述植物是烟草植物。

[0210] 在另一方面,提供至少在植物的部分中用于调节蔗糖酯含量(如 $\beta$ -甲基戊酰基蔗糖酯含量)的方法,其包括下列步骤:(i)调节在植物中的异丙基苹果酸合酶的表达或活性,优选地,其中所述异丙基苹果酸合酶含有本文所述的多核苷酸序列或多肽序列;(ii)任选地测量在获自步骤(i)的突变、非天然存在的或转基因植物的至少部分中的蔗糖酯含量;以及(iii)识别其中蔗糖酯含量与对照植物相比已得到调节的突变、非天然存在的或转基因植物。适宜地,所述突变、非天然存在的或转基因植物的视觉外观与对照植物基本相同。适宜地,所述植物是烟草植物。

[0211] 在另一方面,提供至少在植物的部分(例如,叶子)中用于增加蔗糖酯含量(如 $\beta$ -甲基戊酰基蔗糖酯含量)的方法,其包括下列步骤:(i)增加在植物中的异丙基苹果酸合酶的表达或活性,优选地,其中所述异丙基苹果酸合酶含有本文所述的多核苷酸序列或多肽序列;(ii)任选地测量在获自步骤(i)的突变、非天然存在的或转基因植物的至少部分中的蔗糖酯含量;以及(iii)鉴定其中蔗糖酯含量与对照植物相比已增加的突变、非天然存在的或转基因植物。

[0212] 在另一方面,提供至少在植物的部分中用于减少蔗糖酯含量(如 $\beta$ -甲基戊酰基蔗糖酯含量)的方法,其包括下列步骤:(i)减少在植物中的异丙基苹果酸合酶的表达或活性,优选地,其中所述异丙基苹果酸合酶含有本文所述的多核苷酸序列或多肽序列;(ii)任选地测量在获自步骤(i)的突变、非天然存在的或转基因植物的至少部分中的蔗糖酯含量;以及(iii)鉴定其中蔗糖酯含量与对照植物相比已减少的突变、非天然存在的或转基因植物。

[0213] 有利地,本文所述的方法也可用于通过增加异丙基苹果酸合酶的表达或活性而生成东方性状的植物。根据这种实施方案,非东方植物品种,例如烤烟品种植物可被工程化以使得其中异丙基苹果酸合酶的表达或活性增加。这可赋予植物东方性状,因为其可改变(例如,增加或提高)其东方所以是有利地,从而产生新的芳香植物系、品种或杂种。

[0214] 本文所述的方法也可用于生成植物,通过减少异丙基苹果酸合酶的表达和活性减少、抑制或废除所述植物中的东方。因此,根据这种实施方案,东方植物品种可以从而异丙基苹果酸合酶的表达或活性减少。植物内东方性状的减少或废除可以是有利的,因为改变或减少了其东方风味,从而产生新的芳香品种或杂种。

[0215] 异丙基苹果酸合酶表达与对照相比的增加从约5%至约100%或更多、或至少

10%、至少25%、至少30%、至少40%、至少50%、至少60%、至少70%、%、至少80%、至少90%、至少95%、至少98%、100%、150%、或200的增加,其包括在转录活性或蛋白表达或两者上的增加。

[0216] 异丙基苹果酸合酶活性与对照植物相比的增加从约5%至约100%或更多、或至少10%、至少20、25%、至少30%、至少40%、至少50%、至少60%、至少70%、至少80%、至少90%、至少95%、至少98%、100%、150%、或200%或更多的增加。

[0217] 异丙基苹果酸合酶表达与对照相比的减少从约5%至约100%、或至少10%、至少20%、至少30%、至少40%、至少50%、至少60%、至少70%、至少75、80%、至少90%、至少95%、至少98%、或100%的减少,其包括转录活性或蛋白表达或两者。

[0218] 异丙基苹果酸合酶活性与对照植物相比的减少从约5%至约100%、或至少10%、至少20%、至少30%、至少40%、至少50%、至少60%、至少70%、至少75%、%、至少90%、至少95%、至少98%、或100%或更多的减少。

[0219] 蔗糖酯含量与对照植物相比的减少可从约5%至约100%、或至少10%、至少20%、至少25%、至少30%、至少40%、至少50%、至少60%、至少70%、至少75%、至少80%、至少90%、至少95%、至少98%、或多达100%的减少。

[0220] 本文所述的多核苷酸和重组构建体可用于在目标植物种(适宜地烟草)中调节本文所述的异丙基苹果酸合酶的表达。

[0221] 一方面是本公开的突变植物、非天然存在植物、杂交植物、转基因植物的种子。优选地,所述种子是烟草种子。另一方面是本公开的突变植物、非天然存在植物、杂交植物、转基因植物的花粉或胚珠。此外,本发明提供所述的突变植物、非天然存在植物、杂交植物、转基因植物,其进一步包括赋予雄性不育的核酸。

[0222] 本公开还提供本发明突变植物、非天然存在植物、杂交植物、或转基因植物或其部分的可再生细胞的组织培养,其培养能够表达亲代的所有形态和生理特征的可再生植物。可再生细胞包括,但不限于来自叶、花粉、胚、子叶、下胚轴、根、根尖、花药、花和其部分、胚珠、芽、茎(stem)、秆(stalk)、髓和荚膜、或源自其的愈伤组织或原生质体的细胞。

[0223] 根据本公开,携带突变异丙基苹果酸合酶等位基因的植物可用于植物育种程序中以建立有用的系、品种和杂种。特别地,可使所述突变异丙基苹果酸合酶等位基因进行基因渗入至以上所述的商业上重要的品种。因此,提供植物育种的方法,其包括将本文所述的突变植物、非天然存在植物或转基因植物与另一种植物进行杂交、育种或基因渗入,所述另一种植物例如具有不同遗传同一性、不同遗传背景、不同品种、不同系或者不同杂种。一种方法包括将本文所述的突变植物、非天然存在植物或转基因植物与栽培品种、或优良栽培品种或烟草属植物如红花烟草(例如,白肋型)进行杂交、育种或基因渗入。方法可以进一步包括后代植物与另一植物杂交,和任选地重复杂交直到获得具有期望的遗传性状或遗传背景的后代。所述方法可进一步包括鉴定其中异丙基苹果酸合酶的表达或活性被调节的植物的步骤。所述方法可进一步包括选择其中异丙基苹果酸合酶的表达或活性与对照植物相比被调节的植物的步骤。这样的育种方法的一个目的是向其它品种、育种系、杂种或栽培品种,尤其是具有商业利益的那些,引入期望的遗传性状。另一个目的是便于在单一植物品种、系、杂种或栽培品种中叠加不同基因的遗传修饰。考虑种内和种间交配。这样的杂交所产生的后代植物也称为育种系,是本公开的非天然存在植物的实例。

[0224] 在一个实施方案中,提供用于生产非天然存在植物的方法,包括:(a)突变或转基因植物与第二植物杂交以产生后代烟草种子;(b)在植物生长条件下使后代烟草种子生长以产生非天然存在植物。所述方法可进一步包括鉴定其中异丙基苹果酸合酶的表达或活性被调节的植物的任选步骤。所述方法可进一步包括选择其中异丙基苹果酸合酶的表达或活性与对照植物相比被调节的植物的任选步骤。所述方法可进一步包括:(c)非天然存在植物的上一代与自身或另一植物杂交以产生后代烟草种子;(d)在植物生长条件下使步骤(c)的后代烟草种子生长,以产生额外的非天然存在植物;以及(e)重复多次(c)和(d)的杂交和生长步骤,以产生非天然存在植物的其它世代。方法在步骤(a)之前可任选地包括提供亲本植物的步骤,亲本植物包括表征的遗传同一性且不同于突变或转基因植物。在一些实施方案中,根据育种程序,杂交和生长步骤重复从0到2次、从0到3次、从0到4次、从0到5次、从0到6次、从0到7次、从0到8次,从0到9次或从0到10次,以产生非天然存在植物的世代。回交是这种方法的实例,其中后代与其亲本之一或与其亲本遗传相似的另一植物杂交,以便获得具有接近亲本之一的遗传同一性的下一代中的后代植物。用于植物育种,尤其是烟草植物育种的技术,是公知的并可以用于所述的方法。本公开还提供了通过这些方法产生的、获得的或可获得的非天然存在的植物。

[0225] 在本文所述方法的一些实施方案中,使用标准田间程序在田间评价获自育种和筛选的系的变体异丙基苹果酸合酶基因。包括原始未诱变亲本的对照基因型包括在内,并且按随机的完全区组设计或其它适当的田间设计将入选者(entries)排布于田间。对于烟草,使用标准的农学操作,例如在焙烘(curing)之前和期间,收割、称量和取样烟草用于化学和其它常见测试。进行数据的统计分析以确认所选择系与亲本系的相似性。任选进行所选植物的细胞遗传学分析来确认染色体互补性和染色体配对关系。

[0226] 在一个实施方案中,通过上述育种、杂交或基因渗入方法所获得的或可获得的植物将不产生顺式冷杉醇、将基本上不产生顺式冷杉醇、或将不产生可检测水平的顺式冷杉醇。在另一个实施方案中,蔗糖酯组合物和通过本文所述方法所获得或可获得的蔗糖酯组合物将不包含或含有顺式冷杉醇、将不包含或含有实质水平的顺式冷杉醇、或者将不包含或含有可检测水平的顺式冷杉醇。因此,顺式冷杉醇性状可以存在或者其不存在。

[0227] DNA指纹、单核苷酸多态性、微卫星标志物或类似技术可用在标志物辅助的选择(MAS)育种程序中,以便向其它如本文所述的烟草转移或培育异丙基苹果酸合酶基因的突变等位基因。例如,育种人员可以采用含有突变等位基因的基因型与农学期望的基因型进行杂交而建立分离的群体。可使用本文所列出的技术之一,使用从异丙基苹果酸合酶基因组序列或其片段所开发的标志物来筛选F2中的植物或回交世代。鉴定为具有突变等位基因的植物可以回交或自花传粉以建立待筛选的第二群体。取决于预期遗传样式或所用MAS技术,有必要在每轮回交之前对所选择的植物进行自花授粉,以帮助鉴定期望的个体植物。可以重复进行回交或其它育种过程直到恢复轮回亲本的所需表型。

[0228] 根据本公开,在育种程序中,成功的杂交产生可育的F1植物。所选择的F1植物可以与亲本之一杂交,并且第一回交世代植物进行自花传粉以产生再次筛选变体异丙基苹果酸合酶基因表达的群体(例如,异丙基苹果酸合酶基因的无效版本)。重复回交、自花授粉和筛选的过程,例如至少4次,直到最终筛选产生可育并且与轮回亲本相当相似的植物。如果需要的话,这种植物自花授粉,并且随后再次筛选后代以确认植物显示出变体异丙基苹果酸

合酶基因表达。在一些实施方案中，筛选F2代中植物群体的变体异丙基苹果酸合酶基因表达，例如，根据标准方法，例如通过使用PCR方法采用基于本文所述异丙基苹果酸合酶的核苷酸序列信息的引物，鉴定由于缺失异丙基苹果酸合酶基因而没有表达异丙基苹果酸合酶的植物。

[0229] 杂交烟草品种可通过以下产生：阻止第一品种的雌性亲本植物（即，种子亲本）的自花传粉、允许来自第二品种的雄性亲本植物的花粉使雌性亲本植物受精、并允许在雌性植物上形成F1杂种种子。可以通过在花发育早期将花朵去雄防止雌性植物的自花授粉。或者，可以使用雄性不育的形式防止雌性亲本植物上的花粉形成。例如，通过细胞质雄性不育（CMS）或转基因雄性不育可以产生雄性不育，在转基因雄性不育中转基因抑制孢子和/或花粉形成或自交不兼容。含有CMS的雌性亲本植物特别有用。在实施方案中，其中雌性亲本植物是CMS，从雄性可育植物收获花粉并人工施加到CMS雌性亲本植物的柱头，从而收获得到的F1种子。

[0230] 本文所述品种和系可用于形成单杂交烟草F1杂种。在这样的实施方案中，亲本品种的植物可以生长为基本均匀的相邻种群，以便雄性亲本植物对雌性亲本植物的天然交叉授粉。通过常规方式选择性收获雌性亲本植物上形成的F1种子。还可以大批培养两个亲本植物品种，并收获作为自花授粉结果的雌性亲本上形成的F1杂种种子和雄性亲本上形成的种子的混合。或者，可以进行三元杂交，其中单杂交F1杂种用作雌性亲本，并与不同的雄性亲本杂交。作为另一种选择，可以建立双杂交杂种，其中两个不同单杂交的F1后代进行自交。

[0231] 可以筛选或选择突变的、非天然存在的或转基因的植物群体中具有期望性状或表型的那些群体成员。例如，可以筛选单一转化事件的后代群体中具有期望多肽或多核苷酸表达水平的那些植物。物理和生物化学的方法可用于鉴定表达水平。这些包括用于检测多核苷酸的Southern分析或PCR扩增；用于检测RNA转录物的Northern印迹、S1RNA酶保护、引物延伸或RT-PCR扩增；用于检测多肽和多核苷酸的酶或核酶活性的酶法分析；以及检测多肽的蛋白凝胶电泳、Western印迹、免疫沉淀和酶联免疫分析。其它技术如原位杂交、酶染色和免疫染色也可用于检测多肽或多核苷酸的存在或表达。

[0232] 如本文所述的突变的、非天然存在的或转基因植物细胞和植物包括一个或多个重组多核苷酸，例如本文所述的一个或多个分离多核苷酸、一个或多个多核苷酸构建体、一个或多个双链RNA、一个或多个缀合物或者一个或多个载体/表达载体。

[0233] 可以使用包括例如RT-PCR、Northern印迹、RNA酶保护、引物延伸、Western印迹、蛋白凝胶电泳、免疫沉淀、酶联免疫分析、芯片分析和质谱法的方法评价基因表达。应当注意，如果多肽在组织优选的或广泛表达启动子的控制下表达，可以在整个植物或在选定的组织中来评价表达。同样，如果多肽在特定时间例如在特定的发育时间或经诱导表达，可在期望的时间段内选择性地评价表达。

[0234] 没有限制，本文所述的植物出于其它目的在表达或活性被调节之前或之后可经过修饰。一个或多个如下遗传修饰可以存在于本公开的突变的、非天然存在的或转基因植物中。在一个实施方案中，修饰参与重金属吸收或重金属转运的一个或多个基因，产生比未经修饰的对照植物或其部分具有更低重金属含量的植物或植物部分（如叶）。非限制性实例包括助阳离子扩散蛋白（CDF）家族、Zrt-，Irt样蛋白（ZIP）家族、阳离子交换蛋白（CAX）家族、

铜转运蛋白(COPT)家族、重金属P型ATP酶家族(HMA,如W02009074325中所述)、天然抗性相关的巨噬细胞蛋白(NRAMP)同系物家族以及参与重金属(如镉)转运的ATP结合盒(ABC)转运蛋白家族中的基因。如本文所用,术语重金属包括过渡金属。在另一个实施方案中,参与含氮代谢中间体转化的一个或多个基因被修饰,产生这样的植物或植物部分(如叶),当加热时产生和对照植物或其部分相比更低水平的至少一种烟草特异性亚硝胺(例如4-(甲基亚硝胺)-1-(3-吡啶基)-1-丁酮、N-亚硝基去甲基烟碱、N-亚硝基新烟草碱和N-亚硝基假木贼碱)。可以被修饰的基因的非限制性实例包括编码烟碱去甲基化酶的基因,如CYP82E4、CYP82E5和CYP82E10,其参与烟碱而去甲基烟碱的转化,并在W02006091194、W02008070274、W02009064771和PCT/US2011/021088中描述。

[0235] 其它修饰的实例包括除草剂耐受性,例如,草甘膦是许多广谱除草剂的活性成分。已经通过转移aroA基因(来自鼠伤寒沙门氏菌(*Salmonella typhimurium*)和大肠杆菌(*E.coli*)的草甘膦EPSP合成酶)开发了草甘膦抗性转基因植物。已经通过转化来自拟南芥的突变ALS(乙酰乳酸合成酶)基因制备了磺脲抗性植物。来自突变绿穗苋(*Amaranthus hybridus*)的光系统II的OB蛋白已经转移至植物中产生阿特拉津抗性转基因植物;以及通过并入来自细菌肺炎克雷伯菌(*Klebsiella pneumoniae*)的bxn基因产生溴苯腈抗性转基因植物。另一示例性的修饰产生昆虫抗性植物。苏云金芽孢杆菌(*Bacillus thuringiensis*)Bt毒素可以提供延迟Bt抗性害虫出现的有效途径,如最近在西兰花中所示,其中渐增的(pyramided)cry1Ac和cry1C Bt基因控制小菜蛾对单一蛋白的抗性,并显著延迟抗性昆虫的进化。另一个示例性修饰产生抵抗病原体(如病毒、细菌、真菌)引起的疾病的植物。工程改造了表达Xa21基因的植物(细菌性叶斑病抗性)以及表达Bt融合基因和几丁质酶基因的植物(三化螟抗性和鞘耐受)。另一个示例性的修饰产生改变的生殖能力,如雄性不育。另一个示例性的修饰产生耐受非生物胁迫(例如,干旱、温度、盐度)的植物,通过转移来自拟南芥的酰基甘油磷酸酶产生耐受转基因植物;编码参与甘露糖醇和山梨糖醇合成的甘露糖醇脱氢酶和山梨醇脱氢酶的基因改善抗旱性。另一个示例性的修饰产生这样的植物,其产生具有人体使用中良好的免疫原性特性的蛋白质。例如,能够产生在其N-聚糖中基本上缺失 $\alpha$ -1,3-连接的岩藻糖残基、 $\beta$ -1,2-连接的木糖残基、或两者的蛋白质的植物可以是有用的。其它示例性修饰可以产生具有改善的贮藏蛋白和油的植物、具有增强的光合效率的植物、具有延长的货架期的植物、具有增强的碳水化合物含量的植物以及抗真菌的植物;编码参与生物碱生物合成的酶的植物。也可考虑这样的转基因植物,其中S-腺苷-L-甲硫氨酸(SAM)和/或胱硫醚 $\gamma$ -合酶(CGS)的表达已被调节。

[0236] 一个或更多这样的性状可基因渗入到来自另一个植物或烟草栽培品种的突变的、非天然存在的或转基因的植物,或可以直接转化至其中。可以通过本领域已知的任何植物育种方法实现将性状基因渗入到突变的、非天然存在的或转基因植物中,所述植物育种方法例如谱系育种、回交、双单倍体育种等(参见,Wernsman,E.A,和Rufty,R.C.1987.第17章Tobacco.669-698页,在Cultivar Development.Crop Species.W.H.Fehr(编),MacMillan Publishing Co,Inc.,New York,N.Y761页中)。如上所述的基于分子生物学的技术,尤其是RFLP和微卫星标志物,可用于这样的回交中,来鉴定与轮回亲本具有最高遗传同一程度的后代。这允许加快生产与轮回亲本具有至少90%、优选至少95%、更优选至少99%遗传同一性的烟草品种,以及更优选地与轮回亲本遗传上一致的烟草品种,并还包括来自供体亲

本的基因渗入性状。这种遗传同一性的确定可以基于本领域已知的分子标志物。

[0237] 最后的回交世代可进行自交以便为所转移的核酸提供纯的育种后代。除了转移的性状(例如,一个或多个单基因性状)以外,所得植物一般还具有基本上本公开的所有突变的、非天然存在的或转基因的植物的形态和生理特性。精确的回交操作规程将取决于待改变的性状来确定合适的测试操作规程。虽然当被转移的性状是显性等位基因时回交方法简化了,也可以转移隐性等位基因。在这种情况下,可能有必要引入对后代的测试以确定是否已成功转移期望性状。

[0238] 可以通过使重组多核苷酸整合至其基因组中方式转化植物或植物细胞而使其成为稳定转化的。稳定转化的细胞一般在每次细胞分裂时保留所引入的多核苷酸。也可以瞬时转化植物或植物细胞,使得重组多核苷酸没有整合到其基因组中。瞬时转化的细胞通常在每次细胞分裂时失去所引入的重组多核苷酸的所有或某些部分,从而在足够次数的细胞分裂之后所引入的重组多核苷酸在子细胞中检测不到。

[0239] 公开了缀合半体,其包括大分子化合物,例如蛋白质(例如抗体)、脂肪酸链、糖残基、糖蛋白、聚合物(例如,聚乙二醇),或其组合。寡核苷酸可以缀合至增加寡核苷酸细胞吸收的半体上。半体的非限制性实例包括但不限于抗体、多肽、脂质半体如胆固醇半体、胆酸、硫醚例如己基-s-三苯甲硫醇、巯基胆固醇、脂肪族链例如十二烷二醇或十一烷基残基、磷脂、聚胺或聚乙二醇链、金刚烷乙酸、棕榈基半体、十八烷基胺或己基氨基-羧基-羟胆固醇半体。半体可以是带正电荷的聚合物,如带正电荷的肽,例如其长度是约1至50个氨基酸残基,或聚亚烷基氧化物如聚乙二醇(PEG)或聚丙二醇。适当地,带正电荷的聚合物,如聚亚烷基氧化物可以通过接头例如可释放的接头附着至低聚物。

[0240] 当多肽表达被测量时,对编码多肽的mRNA在细胞中量的检测可以通过例如PCR或Northern印迹来进行定量。当测量到样本中多肽量的变化时,通过使用抗体检测它,可用于使用已知技术来定量多肽在细胞中的量。或者,使用本领域公知的方法来测量酶的生物活性。

[0241] 在另一个实施方案中,本文提供了对多肽具有免疫反应性的抗体。如本文所示的多肽、片段、变体、融合多肽等可以在生产对其具有免疫反应性的抗体中用作“免疫原”。这样的抗体可通过抗体的抗原结合位点特异性结合多肽。特异性结合的抗体是特异性识别并结合多肽、同系物和变体、而不识别和结合其它分子的那些。在一个实施方案中,所述抗体特异于具有本文所示氨基酸序列的多肽,并且不与其它多肽交叉反应。

[0242] 更具体而言,所述多肽、片段、变体、融合多肽等含有激发抗体形成的抗原决定簇或表位。这些抗原决定簇或表位可以是线性的或构象的(不连续的)。线性表位的组成有单一部分的多肽氨基酸,而构象的或不连续表位的组成有来自多肽链不同区域的氨基酸部分,其通过多肽折叠而极为靠近。可以通过本领域已知的任何方法鉴定表位。此外,来自多肽的表位可以用作分析中的研究试剂,以便用于从物质(如多克隆血清或来自培养杂交瘤的上清)中纯化特异性结合的抗体。可以使用本领域已知的技术如固相合成、多肽的化学或酶切割、或使用重组DNA技术来生产这样的表位或其变体。

[0243] 可以通过常规技术制备多肽的多克隆抗体和单克隆抗体两者。本文也考虑产生对多肽特异的单克隆抗体的杂交瘤细胞系。可以通过常规技术生产并鉴定这样的杂交瘤。对于抗体的生产而言,可以通过注射多肽、片段、变体或其突变体来免疫各种宿主动物。举几

个来说,这样的宿主动物可以包括,但不限于,兔、小鼠和大鼠。可以使用各种佐剂来增加免疫应答。根据宿主种类,这样的佐剂包括,但不限于,弗氏(完全和不完全)、矿物凝胶如氢氧化铝、表面活性物质如溶血卵磷脂、复合多元醇、聚阴离子、肽、油乳液、钥孔血蓝蛋白、二硝基苯酚和潜在有用的人佐剂如BCG(卡介苗)和短小棒状杆菌(*Corynebacterium parvum*)。可以通过常规技术回收单克隆抗体。这样的单克隆抗体可以是任何免疫球蛋白类,包括IgG、IgM、IgE、IgA、IgD及其任何亚类。

[0244] 抗体也可用在分析中来检测多肽或片段在体内或体外的存在。抗体也可以用于通过免疫亲和层析来纯化多肽或片段。

[0245] 各种实施方案提供了突变的植物、非天然存在的植物或转基因的植物、以及生物质,其中异丙基苹果酸合酶多核苷酸的表达水平是调节的。

[0246] 这样的植物,特别是烟草植物的部分,以及更特别的是烟草植物的叶片和中脉,可以并入或用于制造各种消费品,包括但不限于烟雾形成材料、烟雾形成装置、吸烟物品、可抽吸物品、无烟制品和烟草制品。烟雾形成材料的实例包括但不限于烟草组合物、烟草、烟草提取物、烟丝、切丝填料、烤烟、膨胀烟草、均质烟草、再造烟草和烟斗烟草。吸烟物品和可抽吸物品是烟雾形成装置的类型。吸烟物品或可抽吸物品的实例包括但不限于香烟、小雪茄和雪茄。无烟制品的实例包括咀嚼烟草和鼻烟。在某些烟雾形成装置而不是燃烧装置中,烟草组合物或其它烟雾形成材料被一个或多个电加热组件进行加热而产生烟雾。在加热烟雾形成装置的另一种类型中,通过将热从可燃性燃料组件或热源传递至物理上分开的可位于热源之内、周围或下游的烟雾形成材料来产生烟雾。无烟烟草制品和各种含烟草的烟雾形成材料可包含任何形式的烟草,包括干燥的颗粒、片、小粒、粉或匀浆,沉积在、混合于、包围有或组合有任何形式比如片、膜、卡(tab)、泡沫、或珠的其它成分。如本文所用,术语“烟”用来描述一种类型的烟雾,其由吸烟物品如香烟、或通过烟雾形成材料的燃烧而产生。

[0247] 在一个实施方案中,还提供了来自本文所述突变的、转基因的和非天然存在的植物的干燥(cured)材料。干燥新鲜烟叶的工艺是本领域技术人员已知的,并包括但不限于空气干燥、火干燥、烤干和太阳干燥。干燥新鲜烟叶的工艺取决于收获的烟草的类型。例如,弗吉尼亚烤烟(轻型(bright))烟草通常是烤干的,白肋烟和某些深色株通常是空气干燥的,而烟斗烟草、咀嚼烟草和鼻烟通常是火干燥的。

[0248] 在另一个实施方案中,描述了包括含烟草的烟雾形成材料的烟草制品,所述烟雾形成材料包括来自本文所述的突变烟草植物、转基因烟草植物或者非天然存在的烟草植物的叶子,优选干燥的叶。本文所述的烟草制品可以是混合的烟草制品,其中还可以包括未修饰的烟草。

[0249] 当与源自非突变、非天然存在的或非转基因的对应植物的烟雾形成材料或烟草组合物相比时,在这些烟雾形成材料或烟草组合物中的%蔗糖酯可以是至少约5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%、和100%、150%、或200%或更高。

[0250] 当与源自非突变、非天然存在的或非转基因的对应植物的烟雾形成材料或烟草组合物相比时,在这些烟雾形成材料或烟草组合物中的%蔗糖酯可以是至少约5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%、以及100%更低。

[0251] 突变的、非天然存在的或转基因的植物可以具有其它用途,例如在农业中。例如,本文所述的突变的、非天然存在的或转基因的植物可用于制造动物饲料和人类食品。

[0252] 本公开还提供了用于生产种子的方法,所述方法包括:培养本文所述的突变植物、非天然存在的植物或转基因植物,并从培养的植物收集种子。来自本文所述植物的种子可以通过本领域中已知的方式进行调整处理并包装在包装材料中而形成制品。包装材料如纸和布是本领域众所周知的。种子的包装可以带有描述其中种子性质的标记,例如固定到包装材料的标签或标记、印刷在包装材料上的标记或插入到包装内的标记。

[0253] 另一方面涉及产生含有蔗糖酯的组合物的方法,包括以下步骤:(a)提供本文所述的突变、非天然存在的或转基因植物的部分;生物质、种子或叶;或烟草制品;和(b)从其中提取蔗糖酯;以及(c)任选地纯化蔗糖酯。

[0254] 再一方面涉及产生 $\beta$ -甲基戊酸的方法,包括以下步骤:(a)提供本文所述的突变、非天然存在的或转基因植物的部分;生物质、或叶;或烟雾形成材料;(b)水解获自步骤(a)的材料或其提取物;以及(c)任选地收集 $\beta$ -甲基戊酸。 $\beta$ -甲基戊酸可用于生产聚羟基脂肪酸酯等,其可赋予包括塑料在内的材料以有益的性能。可以生产的塑料可以是可生物降解的,并因此可用于生产生物塑料。细菌,如大肠杆菌或假单胞菌属(*Pseudomonas*),可以工程化以表达本文所述的异丙基苹果酸合酶。适宜的,所用的细菌也可被工程化用来产生聚羟基脂肪酸酯。使用微生物系统的用于生产聚羟基脂肪酸酯的方法描述在,例如US5,750,848和US6,492,134中。因此,一方面涉及含有异丙基苹果酸合酶的细菌以及任选的一种或多种用于生产聚羟基脂肪酸酯的基因。另一方面涉及用于生产聚羟基脂肪酸酯的方法,其包括使用所述细菌。

[0255] 在一个实施方案中,提供来自本文所述的突变、转基因以及非天然存在植物的干燥材料。例如,干燥新鲜烟叶的工艺是本领域技术人员已知的,并包括而不限于空气干燥、火干燥、烤干和太阳干燥。干燥新鲜烟叶的工艺取决于收获的烟草的类型。例如,弗吉尼亚烤烟(轻型(bright))烟草通常是烤干的,白肋烟和某些深色株通常是空气干燥的,而烟斗烟草、咀嚼烟草和鼻烟通常是火干燥的。

[0256] 在另一实施方案中,描述了烟草制品,其包括含有来自本文所述的突变、转基因以及非天然存在植物的叶(如干燥叶)的烟草制品、或者通过本文所述的方法生产的烟草制品。本文所述的烟草制品还可以包括未修饰的烟草。

[0257] 在又一实施方案中,描述了烟草制品,其含有来自本文所述的突变、转基因以及非天然存在植物的植物材料,优选叶如干燥叶。例如,可添加植物材料到该烟草制品的内部或外部,因此在燃烧时所期望的芳香得以释放。根据这种实施方案的烟草制品甚至可以是未修饰的烟草或改良的烟草。根据这种实施方案的烟草制品甚至可以源自突变、转基因或非天然存在的植物,其除了本文所公开的基因还具有在一个或多个基因上的修饰。

[0258] 另一方面涉及(分离的)蔗糖酯组合物,其包含、由或基本上由本文所述的蔗糖酯的一种或多种组成。所述蔗糖酯可以是在其中R1、R2和R4上酰基链中碳的总数是14的蔗糖酯;在其中R1、R2和R4上酰基链中碳的总数是15的蔗糖酯;在其中R1、R2和R4上酰基链中碳的总数是16的蔗糖酯;在其中R1、R2和R4上酰基链中碳的总数是17的蔗糖酯;以及在其中R1、R2和R4上酰基链中碳的总数是18的蔗糖酯。如果碳的这个总数存在,则组成的酰基链可选自图6所示那些中的任意者。适宜地,蔗糖酯在R3上具有乙酰基、以及在R5上具有乙酰基

或氢。适宜地，蔗糖酯在R3上具有乙酰基以及在R5上具有氢。适宜地，所述酰基链的一个或更多、两个或更多、或者三个或更多将是 $\beta$ -甲基戊酰基。因此，例如，蔗糖酯可以是如图5所示的蔗糖酯，其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=丙酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯，其中R3=乙酰基，R1=丙酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯，其中R3=乙酰基、R1=丁酰基、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基(C2C14:0)；如图5所示的蔗糖酯，其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯，其中R3=乙酰基、R1=戊酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯，其中R3=乙酰基、R1=丁酰基、R2=己酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基(C2C15:0)；蔗糖酯具有图5中所示的通用结构，并且其中R3为乙酰基、R1为戊酰基或其同分异构体、R2是戊酰基或其同分异构体、R4为己酰基或其同分异构体、以及R5是氢原子或乙酰基半体(C2C16:0)；或者其中所述蔗糖酯具有图5所示的通用结构，并且其中R3为乙酰基、R1为丙酰基或其同分异构体、R2是己酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体(C2C16:0)；如图5所示的蔗糖酯，其中R3为乙酰基、R1是戊酰基或其同分异构体、R2是己酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体(C2C17:0)；并且，如图5所示的蔗糖酯，其中R3是乙酰基、R1是己酰基或其同分异构体，R2是己酰基或其同分异构体，R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体(C2C18:0)。

[0259] 或者，所述蔗糖酯可以是在其中R1、R2和R4上酰基链中碳的总数是14的蔗糖酯；在其中R1、R2和R4上酰基链中碳的总数是15的蔗糖酯；在其中R1、R2和R4上酰基链中碳的总数是16的蔗糖酯；在其中R1、R2和R4上酰基链中碳的总数是17的蔗糖酯；以及在其中R1、R2和R4上酰基链中碳的总数是18的蔗糖酯不存在、基本不存在、或者在可检测的水平不存在(C2C18:0)。如果存在碳的这个总数，则组成的酰基链可选自图6所示那些中的任意者。适宜地，蔗糖酯在R3上具有乙酰基、以及在R5上具有乙酰基或氢。适宜地，蔗糖酯在R3上具有乙酰基、以及在R5上具有氢。适宜地，所述酰基链的一个或更多、两个或更多、或者三个或更多将是 $\beta$ -甲基戊酰基。因此，例如，蔗糖酯可以是如图5所示的蔗糖酯，其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=丙酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯，其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯，其中R3=乙酰基、R1=丁酰基、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基(C2C14:0)；如图5所示的蔗糖酯，其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯，其中R3=乙酰基、R1=戊酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯，其中R3=乙酰基、R1=丁酰基、R2=己酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基(C2C15:0)；蔗糖酯具有图5中所示的通用结构，并且其中R3是

乙酰基、R1是戊酰基或其同分异构体、R2是戊酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体(C2C16:0)；或其中所述蔗糖酯具有图5所示的通用结构，并且其中R3是乙酰基、R1是丙酰基或其同分异构体、R2是己酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体(C2C16:0)；如图5所示的蔗糖酯，其中R3是乙酰基、R1是戊酰基或其同分异构体、R2是己酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体(C2C17:0)；并且其中如图5所示的蔗糖酯，其中R3是乙酰基、R1是己酰基或其同分异构体、R2是己酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体，不存在、基本不存在或者在可检测水平不存在(C2C18:0)。

[0260] 或者，所述蔗糖酯可以是在其中R1、R2和R4上酰基链中碳的总数是14的蔗糖酯；在其中R1、R2和R4上酰基链中碳的总数是15的蔗糖酯；在其中R1、R2和R4上酰基链中碳的总数是16的蔗糖酯；在其中R1、R2和R4上酰基链中碳的总数是17的蔗糖酯不存在、基本不存在、或者在可检测的水平不存在；以及在其中R1、R2和R4上酰基链中碳的总数是18的蔗糖酯不存在、基本不存在、或者在可检测的水平不存在。如果碳的这个总数存在，则组成的酰基链可选自图6所示那些中的任意者。适宜地，蔗糖酯在R3上具有乙酰基、以及在R5上具有乙酰基或氢。适宜地，蔗糖酯在R3上具有乙酰基、以及在R5上具有氢。适宜地，所述酰基链的一个或更多、两个或更多、或者三个或更多将是 $\beta$ -甲基戊酰基。因此，例如，蔗糖酯可以是如图5所示的蔗糖酯，其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=丙酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯，其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯，其中R3=乙酰基、R1=丁酰基、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基(C2C14:0)；如图5所示的蔗糖酯，其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯，其中R3=乙酰基、R1=戊酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯其中R3=乙酰基、R1=丁酰基、R2=己酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基(C2C15:0)；蔗糖酯具有图5所示的通用结构，并且其中R3是乙酰基、R1是戊酰基或其同分异构体、R2是戊酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体(C2C16:0)；或其中所述蔗糖酯具有图5所示的通用结构，并且其中R3是乙酰基、R1是丙酰基或其同分异构体、R2是己酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体(C2C16:0)；如图5所示的蔗糖酯，其中R3是乙酰基、R1是戊酰基或其同分异构体、R2是己酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体，不存在、基本不存在或者在可检测的水平不存在(C2C17:0)；并且其中如图5所示的蔗糖酯，其中R3是乙酰基、R1是己酰基或其同分异构体、R2是己酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体，不存在、基本上不存在或者在可检测的水平不存在(C2C18:0)。

[0261] 或者，所述蔗糖酯可以是在其中R1、R2和R4上酰基链中碳的总数是14的蔗糖酯；在其中R1、R2和R4上酰基链中碳的总数是15的蔗糖酯；在其中R1、R2和R4上酰基链中碳的总数

是16的蔗糖酯不存在、基本不存在、或者在可检测的水平不存在；在其中R1、R2和R4上酰基链中碳的总数是17的蔗糖酯不存在、基本不存在、或者在可检测的水平不存在；以及在其中R1、R2和R4上酰基链中碳的总数是18的蔗糖酯不存在、基本不存在、或者在可检测的水平不存在。如果碳的这个总数存在，则组成的酰基链可选自图6所示那些中的任意者。适宜地，蔗糖酯在R3上具有乙酰基、以及在R5上具有乙酰基或氢。适宜地，蔗糖酯在R3上具有乙酰基、以及在R5上具有氢。适宜地，所述酰基链的一个或更多、两个或更多、或者三个或更多将是 $\beta$ -甲基戊酰基。因此，例如，蔗糖酯可以是如图5所示的蔗糖酯，其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=丙酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯，其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯，其中R3=乙酰基、R1=丁酰基、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基(C2C14:0)；如图5所示的蔗糖酯其中R3=乙酰基、R1=丙酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯，其中R3=乙酰基、R1=戊酰基或其同分异构体、R2=戊酰基或其同分异构体、R4=戊酰基或其同分异构体、以及R5是氢或乙酰基；或如图5所示的蔗糖酯，其中R3=乙酰基、R1=丁酰基、R2=己酰基或其同分异构体、R4=己酰基或其同分异构体、以及R5是氢或乙酰基(C2C15:0)；蔗糖酯具有图5所示的通用结构，并且其中R3是乙酰基、R1是戊酰基或其同分异构体、R2是戊酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体，不存在、基本不存在、或者在可检测的水平不存在(C2C16:0)；或其中所述蔗糖酯具有图5所示的通用结构，并且其中R3是乙酰基、R1是丙酰基或其同分异构体、R2是己酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是一个氢原子或乙酰基团，不存在、基本不存在、或者在可检测的水平不存在(C2C16:0)；如图5所示的蔗糖酯，其中R3是乙酰基、R1是戊酰基或其同分异构体、R2是己酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体，不存在、基本不存在、或者在可检测的水平不存在(C2C17:0)；并且其中如图5所示的蔗糖酯，其中R3是乙酰基、R1是己酰基或其同分异构体、R2是己酰基或其同分异构体、R4是己酰基或其同分异构体、以及R5是氢原子或乙酰基半体，不存在、基本不存在、或者在可检测的水平不存在(C2C18:0)。

[0262] 另一方面涉及突变、非天然存在的或转基因植物(适宜地，烟草植物)，其中本文所述的异丙基苹果酸合酶的表达、或者由其所编码的蛋白的活性与(其中异丙基苹果酸合酶的表达、或者由其所编码的蛋白的活性没有增加的)对照植物相比是增加的，并且其中所述植物与所述对照植物相比以更高水平生产在其中含 $\beta$ -甲基戊酰基的蔗糖酯。再一方面涉及突变、非天然存在的或转基因植物(适宜地，烟草植物)，其中本文所述的异丙基苹果酸合酶的表达、或者由其所编码的蛋白的活性与(其中异丙基苹果酸合酶的表达、或者由其所编码的蛋白的活性没有减少的)对照植物相比是减少的，并且其中含 $\beta$ -甲基戊酰基的蔗糖酯的生产比所述对照植物更低。

[0263] 还一方面涉及突变、非天然存在的或转基因烤烟烟草植物，其中本文所述的异丙基苹果酸合酶的表达、或者由其所编码的蛋白的活性与(其中异丙基苹果酸合酶的表达、或者由其所编码的蛋白的活性没有增加的)对照植物相比是增加的，并且其中含 $\beta$ -甲基戊酰基的蔗糖酯的生产比所述对照植物更高。

[0264] 又一方面涉及突变、非天然存在的或转基因东方烟草植物,其中本文所述的异丙基苹果酸合酶的表达、或者由其所编码的蛋白的活性与(其中异丙基苹果酸合酶的表达、或者由其所编码的蛋白的活性没有减少的)对照植物相比是减少的,并且其中含 $\beta$ -甲基戊酰基的蔗糖酯的生产比所述对照植物更低。

[0265] 可收获本文所述突变、非天然存在的或转基因烟草植物,源自其的烟草通过使用对所选烟草类型已知的操作进行处理或干燥。使用例如水或有机溶剂洗涤以及随后的干燥,可以收集叶表面化合物。然后,所得到的粉末可以用来提高吸烟材料的风味性能。然后,在洗涤后残留的烟草材料可用作例如膨胀烟草的材料。

[0266] 在另一方面,提供用于至少在植物的部分中调节抗虫性的方法,包括以下步骤:(i)调节在植物中的异丙基苹果酸合酶的表达或活性,优选地,其中所述异丙基苹果酸合酶含有本文所述的多核苷酸序列或多肽序列;(ii)任选地测量在获自步骤(i)的突变、非天然存在的或转基因植物的至少部分中的 $\beta$ -甲基戊酰基蔗糖酯的量;(iii)鉴定突变、非天然存在的或转基因植物,其中其含 $\beta$ -甲基戊酰基的蔗糖酯的量与(其中异丙基苹果酸合酶的表达或活性没有被调节的)对照植物相比已改变,以及优选地,其中所述突变、非天然存在的或转基因植物的视觉外观与对照植物基本相同;以及(iv)获得其中抗虫性已经得到调节的植物。

[0267] 合适地,提供用于至少在植物的部分中增加抗虫性的方法,包括以下步骤:(i)增加在植物中的异丙基苹果酸合酶的表达或活性,优选地,其中所述异丙基苹果酸合酶含有本文所述的多核苷酸序列或多肽序列;(ii)任选地测量在获自步骤(i)的突变、非天然存在的或转基因植物的至少部分中的含 $\beta$ -甲基戊酰基的蔗糖酯的量;(iii)识别突变、非天然存在的或转基因植物,其中其含 $\beta$ -甲基戊酰基的蔗糖酯的量与(其中异丙基苹果酸合酶的表达或活性没有增加的)对照植物相比已增加,以及优选地,其中所述突变、非天然存在的或转基因植物的视觉外观与对照植物基本相同;以及(iv)获得其中抗虫性已经增加的植物。

[0268] 还提供通过所述方法获得的或可获得的植物或植物材料。

[0269] 另一方面涉及突变、非天然存在的或转基因植物,其中异丙基苹果酸合酶的表达、或者由其所编码的蛋白的活性是得到调节的;以及与(其中异丙基苹果酸合酶的表达或活性没有得到调节)对照植物相比,至少所述植物的部分在含 $\beta$ -甲基戊酰基的蔗糖酯的组合物上具有改变,以及其中所述植物的视觉外观与对照植物基本相同,并且其中其抗虫性也是得到调节的。

[0270] 合适地,提供突变、非天然存在的或转基因植物,其中异丙基苹果酸合酶的表达、或者由其所编码的蛋白的活性是增加的以及与(其中异丙基苹果酸合酶的表达或活性没有增加的)对照植物相比,至少所述植物的部分在含 $\beta$ -甲基戊酰基的蔗糖酯的组合物上具有改变,以及其中所述植物的视觉外观与对照植物基本相同,并且其中其抗虫性也是增加的。

[0271] 包括用于对植物基因分型以鉴定、选择或育种的组合物、方法和试剂盒,并可包括检测样本中异丙基苹果酸合酶多核苷酸存在的方法。因此,描述了组合物,其包括一个或多个用于特异性扩增异丙基苹果酸合酶多核苷酸至少一部分的引物,以及用于进行扩增或检测的任选的一个或多个探针和任选的一种或多种试剂。公开了基因特异性寡核苷酸引物或探针,其包括对应于异丙基苹果酸合酶多核苷酸的约10个或更多的连续多核苷酸。所述引物或探针可包括或由约15、20、25、30、40、45或50或更多连续多核苷酸组成,其杂交(例如,

在例如严谨的杂交条件下特异性地杂交)至异丙基苹果酸合酶多核苷酸。在一些实施方案中,引物或探针可包括或由约10至50个连续核苷酸、约10至40个连续核苷酸、约10至30个连续核苷酸、或约15至30个连续核苷酸组成,其可用于基因鉴定的序列依赖性方法(例如,Southern杂交)或分离(例如,细菌克隆或噬菌体噬菌斑的原位杂交)或基因检测(例如,作为核酸扩增或检测中的一个或多个扩增引物)。可设计一个或多个特异性引物或探针并用于扩增或检测异丙基苹果酸合酶多核苷酸的部分或全部。通过具体的实例,两个引物可用于聚合酶链式反应规程中来扩增编码异丙基苹果酸合酶核酸的核酸片段,如DNA或RNA。也可以使用源自异丙基苹果酸合酶核酸序列的一个引物和杂交至异丙基苹果酸合酶核酸序列(如异丙基苹果酸合酶启动子序列、mRNA前体3'端或源自载体的序列等)的上游或下游序列的第二引物进行聚合酶链式反应。可用于多核苷酸体外扩增的热和等温技术的实例是本领域中公知的。样本可以是或可源自植物、植物细胞或植物材料、或者从本文所述的植物、植物细胞或植物材料所制得的或来源的烟草制品。

[0272] 因此,在另一方面中,还提供了检测样本中异丙基苹果酸合酶多核苷酸的方法,包括步骤:(a)提供包括、或疑似包括多核苷酸的样本;(b)使所述样本接触多个引物之一、或一个或多个探针以便特异性检测异丙基苹果酸合酶多核苷酸的至少部分;以及(c)检测扩增产物的存在,其中扩增产物的存在指示着样本中存在异丙基苹果酸合酶多核苷酸。在另一方面中,还提供用于特异性检测异丙基苹果酸合酶多核苷酸的至少部分的一个或多个引物或探针的用途。还提供了用于检测异丙基苹果酸合酶多核苷酸的至少部分的试剂盒,其包括用于特异性检测异丙基苹果酸合酶多核苷酸的至少部分的一个或多个引物或探针。试剂盒可以包括用于多核苷酸扩增如聚合酶链反应(PCR)的试剂,或用于核酸探针杂交检测技术如Southern印迹、Northern印迹、原位杂交或微阵列的试剂。试剂盒可以包括用于抗体结合检测技术,如Western印迹、ELISA、SELDI质谱法或测试条的试剂。试剂盒可以包括用于核酸测序的试剂。试剂盒可以包括用于确定蔗糖酯含量的试剂和/或说明书。在一些实施方案中,试剂盒可以包括一个或多个所述方法的说明书。所述试剂盒可用于遗传同一性确定、系统发育研究、基因分型、单倍体分型、谱系分析或植物育种,特别是共显性评分。

[0273] 本公开还提供了对含有异丙基苹果酸合酶多核苷酸的植物、植物细胞或植物材料进行基因分型方法。基因分型提供了区分染色体对的同系物的方式,并可以用来区分植物群体中的分离子(segregant)。分子标志物方法可用于系统发育研究、表征作物品种之间的遗传关系、鉴定杂交或体细胞杂种、定位影响单基因性状的染色体片段、图位克隆(map based cloning)和定量遗传研究。基因分型的具体方法可以采用任意数目的分子标志物分析技术,包括扩增片段长度多态性(AFLP)。AFLP是由核苷酸序列变异性引起的扩增片段之间等位基因差异的产物。因此,本发明还提供了使用像AFLP分析这样的技术来跟踪异丙基苹果酸合酶基因或核酸的分离、以及遗传连锁至这些基因或核酸的染色体序列的方式。

[0274] 将于如下实施例中进一步描述本公开,提供所述实施例以更详细地描述本公开。这些实施例,其阐述用于实施本发明的当前设想的优选方式,是意欲阐明而非限制本公开。

#### [0275] 实施例

[0276] 实施例1:在红花烟草品种烤烟希克斯阔叶(HBL)和原始的红色俄罗斯(Red Russian,RR)上的蔗糖酯分析。

[0277] 为研究在烟草中的蔗糖酯合成,选择两种模式品种:烤烟希克斯阔叶(HBL)和原始

的红色俄罗斯(RR)。从国家遗传资源计划(National Genetic Resources Program)(种质资源信息网(Germplasm Resources Information Network)-(GRIN),国家种质资源实验室(National Germplasm Resources Laboratory),贝茨维尔,马里兰)可获得的初始数据的目录显示在这两种品系之间含β-甲基戊酰基的蔗糖酯的不同含量。

[0278] 分析HBL和RR叶渗出物在蔗糖酯上的组合物,结果显示在图1中。RR蔗糖酯组合物与HBL蔗糖酯组合物不同。在HBL中,蔗糖酯主要由C2C15:0组成,然而在RR中,观察到存在更高分子量的种,如C2C16:0、C2C17:0和C2C18:0。如图2所示,RR和HBL之间的差异是因为在分子中甲基戊酰基酯的存在或缺失。因此,结论是RR是BMVSE品种而HBL是bmvs品种。

[0279] 实施例2:数量性状基因座分析

[0280] 在分别源自136个自交的F2基因型植物的8F3植物的136个群体中分析叶渗出物中蔗糖酯的组合物,所述F2基因型植物由HBL和RR亲本杂交所得。进行数量性状基因座研究,以使微卫星标志物和含β-甲基戊酰基的蔗糖酯表型相关联。使用数量性状基因座制图法分析数据。含β-甲基戊酰基的蔗糖酯的性状与在烟草基因组上的一个单一基因座相关联。这个基因座定位在烟草的连锁群15(染色体A)上。

[0281] 实施例3:含β-甲基戊酰基的蔗糖酯的相关基因的鉴定

[0282] 研究烟草外显子在毛状体中的转录水平,以在编码图3所述函数(function)的全部可能基因中选择优选的基因靶标。为研究含β-甲基戊酰基的蔗糖酯的生产曲线,使用微阵列在几种选择的品种中研究表达。

[0283] 所述结果鉴定出异丙基苹果酸合酶相关外显子探针,因为其表达与每个品种的每个提取物中测量的含β-甲基戊酰基的蔗糖酯的估计量相关( $R>0.95$ )。

[0284] 鉴定出四种烟草序列,其编码番茄异丙基苹果酸合酶IPMSA和IPMSB的同系物。其命名为NtIPMS1A(SEQ ID NO:10)、NtIPMS1Bv1(SEQ ID NO:12)、NtIPMS1Bv2(SEQ ID NO:14)和NtIPMS2(SEQ ID NO:1)。根据EST数据库,因为大多SEQ ID NO:1的表达可以在毛状体中观察到,所以其是选择的候选者。

[0285] 用NtIPMS2cDNA作为PCR诱饵筛选从红花烟草品种HBL构建的BAC文库,鉴定出命名为PISOGE的BAC。用鸟枪测序这个BAC。组装产生几个重叠群,其中pisoge1和pisoge2含有与NtIPMS2相关的序列。

[0286] 内含子-外显子结构在烟草和番茄的异丙基苹果酸合酶中是保守的。对NtIPMS1Bv1、NtIPMS1Bv2、S1IPMSB和S1IPMSA而言,观察到所述基因分离在12个外显子中。还预期NtIPMS2由12个外显子组成,但由于S1IPMSA第一外显子对应于pisoge1上的2个外显子,所以假定NtIPMS2含有13个外显子。

[0287] 通过PCR来测试基因组序列预测的可靠性。普遍地,使用纯化的PISOGE BAC或从希克斯阔叶或红色俄罗斯植物中分离出的基因组DNA,来验证来自pisoge组装的预测序列。鉴定出一个与NtIPMS2对应的全长基因组结构。此外,观察到两个其它基因样结构,并有可能代表假基因重复。由PISOGE BAC提出的结构部分得到验证,但观察到所述BAC结构和HBL基因组DNA之间的一些差异。

[0288] 用PCR作遗传标志物来区分NtIPMS2的RR和HBL等位基因。在数量性状基因座作图中所用的全部F2植物的基因组DNA的PCR扩增上,使用所述IPMS2标志物。使用这个额外的标志物重新计算蔗糖酯数量性状基因座。数量性状基因座峰鉴定出性状和标志物之间关联

的概率最大的位置。所述IPMS2标志物与简单序列重复标志物PT30172一起,是对含 $\beta$ -甲基戊酰基的蔗糖酯的性状的影响概率最高的标志物。

[0289] 实施例4:证明NtIPMS2的活性作为在含 $\beta$ -甲基戊酰基的蔗糖酯的合成中的关键因素

[0290] 从毛状体EST文库中分离出与NtIPMS2对应的表达序列标签,所述毛状体EST文库以如下形式:GT09L09(来源:Galpao(无对应中译文)[TI1068]毛状体)和OT05C20(来源:Orinocco(无对应中译文)[TI81]毛状体)。两个分子在核苷酸序列上一致,因此只有OT05C20cDNA用于表达实验。使用基于共有序列的NtIPMS2的第一外显子和第一内含子来生成RNAi构建体。将两个构建体转移至双元载体内,所述双元载体含有对毛状体特异的CPS2p启动子(CPS2指copalyl合酶2,其参与仅在毛状体中发生的顺式冷杉醇合成)、或MMVp病毒启动子(MMV指紫茉莉花叶病毒)。在东方烟草品种巴斯玛xanthi(无对应中译文)和烤烟品种K326中进行使用分别携带双元载体的根瘤农杆菌(*Agrobacterium tumefaciens*)AGL1的稳定转化。

[0291] 从独立转化事件所得的一月龄再生植物中收集叶子。以2.5ng/mg鲜重的比添加内标(蔗糖八乙酸酯)。测量针对特异于每种蔗糖酯的离子所记录的信号面积,并作为内标面积的比进行绘制。从分子到分子的信号反应不相同,所以仅在每两个分子之间进行的比较是无效的。与内标信号面积的比允许将叶鲜重上的信号正态化。图2描述了此处所示的蔗糖酯的命名。参见图4:(A)亲本品种K326(烤烟,n=3)的叶渗出物的组合物和在毛状体特异启动子下表达NtIPMS2的转化体;(B)亲本品种K326的叶渗出物的组合物和在病毒启动子下表达NtIPMS2的转化体;(C)亲本品种巴斯玛(东方,n=4)的叶渗出物的组合物和在毛状体特异启动子下表达NtIPMS2-RNAi构建体的转化体;(D):亲本品种巴斯玛的叶渗出物的组合物和在病毒启动子下表达NtIPMS2-RNAi构建体的转化体。

[0292] 呈现的数据足以证明NtIPMS2是在含 $\beta$ -甲基戊酸的蔗糖酯的生产中的关键基因。在含 $\beta$ -甲基戊酸的蔗糖酯的品种(东方)中抑制IPMS2,减少或废除了含 $\beta$ -甲基戊酸的蔗糖酯的生产(参见C和D),从而证明了IPMS2是生物合成途径的关键成分。RNA干扰构建体的组成型和毛状体特异性表达导致蔗糖酯C2C16:0、C2C17:0和C2C18:0积累的损失,所述蔗糖酯是可含 $\beta$ -甲基戊酸的酯。

[0293] 在bmvs品种(烤烟)引入NtIPMS2可以诱导产生蔗糖酯(参见A和B)。这证明,仅NtIPMS2cDNA表达足以恢复蔗糖酯的生产。这里的NtIPMS2cDNA的组成型和毛状体特异性表达导致蔗糖酯C2C16:0、C2C17:0和C2C18:0的重新积累,蔗糖酯是可以含 $\beta$ 甲基戊酸的酯。有趣的是,所述cDNA的组成型表达(B)并未导致蔗糖酯的高度积累,然而毛状体特异性表达导致蔗糖酯积累到这样的水平,所述水平比得上例如在含 $\beta$ -甲基戊酸的蔗糖酯的品种中可观察到的。

[0294] 实施例5:用于选择锌指核酸酶靶位点的搜索规程

[0295] 本实施例说明如何搜索异丙基苹果酸合酶基因来筛选相较于给定基因组数据库在给定基因序列内独特靶位点的存在,以便开发用于改变基因表达的工具。在本公开内包括:通过本公开的方法所鉴定的靶位点、序列基序、以及任何位点或基序在改变植物(如烟草)中对应的基因序列中的用途。

[0296] 搜索算法.开发了计算机程序,其允许在DNA数据库中使用后缀阵列筛选输入查询

(目标)核苷酸序列中被给定间隔尺寸所隔开的两个固定长度子字符串DNA基序的存在。后缀阵列的构建和搜索使用开放资源libdivsufsort文库2.0.0 (<http://code.google.com/p/libdivsufsort/>) ,其将任何输入的字符串直接转换成Burrows-Wheeler转化字符串。程序扫描整个输入(目标)核苷酸序列并返回在选定DNA数据库中发生次数少于所选次数的所有子字符串组合。

[0297] 选择查询序列的锌指核酸酶介导的诱变的靶位点。锌指DNA结合结构域识别3碱基对的核苷酸序列。锌指核酸酶包括含有一个、两个、三个、四个、五个、六个或更多个锌指DNA结合结构域的锌指蛋白,以及IIS型限制性酶的非特异性核酸酶。锌指核酸酶可用于向靶序列中引入双链断裂。为引入双链断裂,需要一对锌指核酸酶,其中一个结合靶序列的正(上)链,另一个结合同一靶序列的相隔0、1、2、3、4、5、6或更多个核苷酸的负(下)链。通过对两个固定长度子字符串DNA基序的每一个使用3的复数,程序可用来鉴定相隔给定间隔长度的两个锌指蛋白靶位点。

[0298] 程序输入:

[0299] 1、靶查询DNA序列

[0300] 2、待搜索的DNA数据库

[0301] 3、第一子字符串DNA基序的固定尺寸

[0302] 4、间隔序列的固定尺寸

[0303] 5、第二子字符串DNA基序的固定尺寸

[0304] 6、在程序输入2的所选DNA数据库中,由程序输入4隔开的程序输入3和5的组合发生数目的阈值。

[0305] 程序输出:核苷酸序列列表,其中对于各序列,序列在DNA数据库中发生的次数具有程序输入6阈值的最大值。

[0306] 本文引用或描述的任何出版物提供本申请申请日之前公开的有关信息。本文的陈述不能被解释为承认本发明人无权早于这种公开。在以上说明书中提到的所有出版物通过参考并入本文。本公开的各种改变和变化,在不脱离本公开的范围 和精神时,对于本领域技术人员是显而易见的。虽然本公开已结合具体的优选实施方案进行了描述,应当理解,所要求的本发明不应该被不适当当地限制于这种具体的实施方案。实际上,用于实施本发明的所述形式的各种改变,其对于细胞、分子和植物生物学或相关领域的技术人员是显而易见的,其意欲在如下权利要求的范围内。

[0307] 序列

[0308] SEQ ID NO:1 (NtIPMS2核苷酸序列)

[0309] ATGGCTTCTCTCTGTAAATTCTATAATTCCCTGAGCACTCCCTTCATTACATTCTAAAAACCCA  
CTTATTACAGTGTCTTCAGTTCACGCCTCAACCACAAGGCATTCAGCTATATGCTGCTCAAATATTGCGCGC  
GCCGGAGTATAACATGGAAAATTCTCTGACCCCTGATTATGTTGGTATTTGACACCAGTCTCGCGATGGCGAAC  
AGGCCGCTGGCTACCATGACTAGTAAAGAAAACTGGACATTGCACGTCAAGTGGCTAAGCTTGGTGTGATGTT  
ATTGAGGCCGGTTTCCTTGCCTCTGAAGCTGAGTCGAGCTGTAAAGTTGATAGCACAGGAAATTGGTAATAA  
CGTAGACAAAGAAGGATACGTGCCGATGATATGTGCCTAGCTAGGTCTAGTAAGAAGGATATTGAAAGAGCTGGG  
ATGCTTAAAGTATGCAAAGAACCAATGCTTCATATGTTATTGCGACGAGTGTATACATATGAAGTACAAGTTA  
AAGATGAGTAGAGAAGAAATTGTGGAGACAGCTAGGAGTACGGTGGCTATGCAAAACCCATTGAGGATGTTCG

GTTCAGCGCTGAAGATGCTGCAAGATCTGATAGGGAGTTCCCTTATCATATTATTGGAGAAGTTACAAAGCTGGTG  
 CAACAGTGATTGCCCTCCCTGATACAGTGGATGCAATTGCCAGTGAATATGCACAACACTGATTCTGATATAAAA  
 GCCAATACCCAGGAATACAAGATGCAAACACATTCAACACACTGTCACAACGATCTGGGCTGCTACTGCCAACTC  
 CTTAGCTGGAATTGCGCAGGCAGACTAGTAGATGTTACCATCAATGGAATTGGTAAAGAGCTGGAAATGCTT  
 CTCTGGAGGAGATTGTAATGGCCTTAAATATCGTGGAGAGCAAGTACTAGGTGTATCTATACTGGGATTAATACA  
 AAGCATATATTGACGAGCAAATGGTAGAAGAGTACAGTGGCTTAAGCTGCAGCACATAAGGCCATTGTTGG  
 AGCTAATGCATTTCTCATGAGAGTGGCATCCATCAGGATGGAGTGTAAAGAACAGAGATACATATGAGTTGTAT  
 CTCATGAAGATGTTGGGTATCGTCGTGCTAATGAAAACGGTATTAGTCTGGAAAGCTCAGTGGCCCATGCATTG  
 AAAGCCAAAATGGCTGAGCTGGATATGACTTGATGGAAAAGAACTTGATGACCTCTTCGTCGATTCAAGTCACT  
 AGCTGAGAGGAAAAAGAAAATTACAGATGATGACTTGAGAGCACTGTATCAGATGACGTTTCCAGCCTCAAGTT  
 CCTGGCAACTTGGAGATGTACAGATTACTGTGGAAATGTTGGCCGCTCACAGCAAATGTTAAGCTTATTGACAGC  
 GATGGTCAAGAGCACACTGCCTTCTGTTGGAACAGGACCTGTTGATGCAGCTACAAGGCAGTTGACCTCATTGT  
 AAAGGTACCTGTAACACTCGTTAATTCGGTTAATGCAATCACAAACGTATAAATTCCACAGCTCAACCAGAG  
 TGTAGTTCGTGGGAATGACTATGCATCGTTAATACTCAAACGGGCAAATGTTAATCGTACAGTTAGTGGA  
 ACAGGAGCGCATATGGACATTGTCGTTCAAGTGTCCAAGCCTATGTTGAGGCAGTGAACAAATATTCAAGTACAA  
 AAAAACAGGTCTCGTGAACAAATTGAAGGCAGTGCAGCAATCGTAA

[0310] SEQ ID NO:2 (NtIPMS2氨基酸序列)

[0311] MASLSVNSIISLSTSLSLHSKNPLIHSVFSFTPSTTRHSAICCSNIRRPEYKHGKFSDPDYVGIFDTS  
 LRDGEQAAGATMTSKEKLDIARQLAKLGVDVIEAGFPFASEAEFELVKLIAQEIGNVDKEGYVPMICALRSSKKD  
 IERAWDALKYAKKPMILHMFIAITSDIHKYKLKMSREEIVETARSTVAYAKTLFEDVRFAEDAARSDFLYHIIGE  
 VIKAGATVIGLPDTVGCLNPSEYAQLISDIKANTPGIQDANISTHCHNDLGLATANSLAGICAGARLVDVTINGIGE  
 RAGNASLEEIVMALKYRGEQVLGGIYTGINTRHIFMTSKMVEEYSGLKLQPHKAVGANAFSHESGIHQDGVLKNRD  
 TYEFVSHEDVGYRRANENGISLGKLSGRHALKAKMAELGYDFDGKELDDLFRRFKSLAERKKITDDDLRALVSDDV  
 FQPQVSWQLGDVQITCGNVRSTANVKLIDSDGQEHTAFSGTGPVDAAYKAVDLIVKVPVTLVEYSVNAITKRINS  
 TASTRVLVRGNDYASFNTSNGQTVNRTVSGTGAHMDIVVSSVQAYVEALNKIFSYYKKTGLVNKFEGSAQS

[0312] SEQ ID NO:3 (用于互补的NtIPMS2cDNA序列) ctctttcttagggaaaagaataatggct  
 tctctctgtaaattctataattccctgagcacccatccattacattctaaaaaccacttattcacagtg  
 tcttcagttcacgcctcaaccacaaggcattcagctatgctcaaatattcgccggccggagataaa  
 acatggaaaattctgtaccctgattatgttgttgcacccatgtggcttgcacgttggctaaagtttatggatgttgc  
 gtttcctttgcctctgaagctgagttcgagcttgcacccatgtggcttgcacccatgtggatgttgc  
 agaaggatacgccatgtgccttagcttaggtcttagtaagaaggatattgaaagagcttggatgttgc  
 aagtatgcataagaaaccatgttcataatgtttattgcgcacgagtatgcatacatatgaagttaaagatga  
 gtagagaagaaattgtggagacagctaggagttacgggtttatgcacccatattggagaagtttatcaagctggatgt  
 cgctgaagatgtcaagatctgtatggatgcatttgcacccatgtggatgttgcacccatgtggatgttgc  
 gtgattggcccccgtatacgatgtggatgcatttgcacccatgtggatgttgcacccatgtggatgttgc  
 atacccccaggaatacaagatgcacacattcaacacactgtcacaacgcatttgcacccatgtggatgttgc  
 agctggatattgcgcaggcgcaagacttagtagatgttaccatcaatggatgtggaaagagctggaaatgttgc  
 ctggaggagattgtatggcttaaaatatcgatgtggagagcaagtaggtgttatctatactggatataacaa

agcatatattcatgacgagcaaaatggtagaagagttacagtggcttaagctgcagccacataaggcattgttgg  
 agctaattgcatttcattcatgagagtggcatccatcaggatggagtgttaaaaagaacagagatacatatgagttgt  
 tctcatgaagatgtggatcgctgtgctaatgaaaacggtattagtctggaaagctcagtgccgcatgcatt  
 tgaaagccaaaatggctgagctggatatgactttgtatggaaaagaacttgcacccctttcgatcaagtc  
 actagctgagaggaaaaagaaaattacagatgatgacttgagagcacttgtatcagatgacgcccccttgc  
 gttcctggcaacttggagatgtacagattacttgtgaaatgtggccgtcacagcaaattgtaagcttatt  
 acagcgatggcaagagcacactgcctttgttggaaacaggacctgttatgcagcttacaaggcagttgac  
 cattgttaaggtaacctgttaacactcggttatgcattcaatcacaacgtataattccacagcttca  
 accagagtgttagtgcgtggaaatgtgactatgcattcaatcacaacggccaaactgttatactgtac  
 ttagtggaaacaggagcgcataatggacattgtcggttatgcattcaatcacaacggccaaactgttata  
 cagttacaaaaaaacaggctcgtaacaaattgtgaaaggcgtgcgcaatcgtaaaagtgtatgggtgt  
 aaacaaaagcttaatgctttagtgctgatgatcataattatgcttattcatgcattactgaaatt  
 aattaagaagtaatcttattgtataagatgtaaacatgtttattcatgcattactatgaaatt  
 tctgagccaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

[0313] SEQ ID NO:4(覆盖外显子XI、XII和XIII的NtIPMS2区的基因组DNA序列,如在红色俄罗斯中观察到的。下划线表示cDNA匹配区。粗体显示红色俄罗斯和希克斯阔叶DNA序列之间的保守DNA区)

**AGCAAGATGTCAAGAACGACACACTTGCCTTTCCGTGGAACAGGACCTGTTGATGCAGCTTACAAGGCAGTGACC**  
**TCATTGTAAGGTTATAATGAAACTGAAAACCTCTCATATGTTGTATCTTGTCAACTAGTAGGATAGGTAAATT**  
**AGGTTTATATACGCTGAAGGTATAAAAATTTACTCAATCAGAGCGCTTACAAGACAATTGCAAGGTAACTCTATG**  
**AGTAGCTTAGACTAAAAGATAAACGTAGTAATTCAATGTAATAGTTAGGAGACATGCGAATCTTCCTTGACAA**  
 AGAAAATGAAACTTGCAGCACCTCAATCACCCCGTCTGTGAATTGAACGAACTAAAAGTACAACC  
 AGCTTAAACTGATTTTACCTATTTTAGTGTTCAGACGGATCTGCAGCAACCCTGATTGATTCAACCGGTGCA  
 CAAATATCCTCACCGTGCATTGAACTCCCTTCTCCACATCCTAAAAATCCTCTTAAGATGTCTTGTGTTGCA  
 AACTCTAGCCGTCATCTTGGGAAACATGTCTCTTAATATGTGTATTATGTGTTATAATAGTATTAG  
 GAGAATGGCAAGTTAACAGACTAACACCTATTATGAATGGATCCAACCGAATAGAAATATTCAACATTAG  
 CTAAGTAAGCTGCCAAGTCAAATAAGACTGATAGAGTAAAACCTTCTATAGCAATGAATATAACCTAAATCC  
 TTTAGTTACTGCTAGCAATTGTAGGTATAATTGCTTGTGTTGAAACATCTATGTAACCTTGTGTTATAA  
 GTAATGTTCTTGTCAATTGCAAGGTACCTGTAACACTCGTTGAATATCGTTAATGCAATCACAAACGTATAAATTCCACA  
GCTTCAACCAAGAGTGTAGTCTCGTGGGAATGACTATGCCATCCCTTAATACTTCAAACGGCAAACGTAAATC  
GTACAGTTAGGTATGAAATATGACAGAGAAGAATTAAACGATTCTAAAGGAGAAAATAATATGTGATAGGAATGA  
 AATGTTGTCAAGTTATATAGTTCTCAAATTGTTACTACGGGTATCTAGAGCCAACCTGAGAATAGAAAGTA  
 TCATATAAGAAATATACACAGCAGTTAGATAGAGTACATTTTTAATTATTTAATTATTTATGCAATGCC  
TCAACACGTTACTCACCTCTAGCTAATTGTTACATGGGATCGTTCTACATTTGAGTAGCTGAGAGATTTG  
AACTTAGGACCTTTATTTCTCTAACACTATGTTATATTGCCAGTGGAACAGGAGCGCATATGGACATTGTCGTTTC  
AAAGTCCAACGCTATGTTGAGGCGTTGAACAAAATATTCAAGTTACAAAAAAACAGNTCTCGTGAACAAATTGAAAG  
GCAGTGCAGCAATCGTAAAAGTGTAGGGTGTGCTNNCTANAAG

[0315] SEQ ID NO:5(覆盖外显子XI和XIII的NtIPMS2区的基因组DNA序列,如在希克斯阔叶中观察到的。下划线表示cDNA匹配区。粗体显示红色俄罗斯和希克斯阔叶DNA序列之间的保守DNA区)

[0316]

GTCAAGACCACTGCCTTCTGTTGGAACAGGACCTGTTGATGCAGCTACAAGCAGTTGACCTCATTGTAAAG  
GTTTATAATGAAACTGAAAACCTCTCATATAATGTTGTATCTTGTACACTAGTAGGATAGGTAAATTAAAGTTTATAT  
ACGCTGAAGGTATAAAAATTTACTCAATCAGAGCGTTACAAGACAATTGCAGGTAACTCTATGAGTAGCTTAGA  
CTAAAAGTTTAATTATTTAATTATTTATGATGCATGCCTAACACGTTACTCACCTCTTAGCCTAATTTTGTTCAC  
ATGGGATCGTCTACATTGAGTAGCTGAGAGATTGAACTTAGGACCTTATTCTCTAAACACTATGTTATATT  
GCCAGTGGAACAGGAGCGCATATGGACATTGTCGTTCAAGTGTCAGGCTATGTTGAGGCGTGAACAAAATATT  
CAGTTACAAAAAAACAGGTCTCGTGAACAAATTGAAGGCAGTGCAGTGCCTAAAGTGATGGTGTGCTATCTG

**CNAAAAA**CTCGA

[0317] SEQ ID NO:6 (NtIPMS2对应BAC重叠群序列(总BAC长度是76832bp。仅显示14232bp)

[0318] TGATGATTCTCATTATAATTGTAAGTAGGGTGGCAAATGGTAAAAGAAAAGTTATCCATCCA  
 TATTATTCACTAAAAATGGGTGTATAATGAACCTTTAAAAAGGATCAATTATGGATAAGAACCATATTATCC  
 GCTTAGAAAATGGATAACCAATAAGTTAACCTACACTGTAAAGCTTCAAATTGGGGTTCTCAAGTTGTGA  
 GAAGAAGAATTCTCCCAAAGTGATCATATTCAAGAAGTCTTGGATAATGAATATCCATTATCCGCCATTATC  
 TCGTTTAATCCGTATTAATATGAGTCGGTCAATAATTATCTGTTGTATTATCTACTTGGACATGTCC  
 ATACCCGACCCGACCCCGCTCGTTGTCACCCCTAGTTGTAACATATACTGTTGCTTACCTAACGGTTGAGTTA  
 AGTACCATCACGACTTGGTGGATTGGCTCTACAAAACATAATTGAGCCTGATTATAATTATTATTATAAA  
 ATTGAGGTTGGTATAATTATTTGCTGATCAAATCTGGTATTCAAAGTTCAACTGAAATATGCAATACTAAA  
 TTATTTCTTATTAAATAAATTGATGAGCTAATTATAATTGAATAAAAAAAATTCAAATAACTAGTA  
 TTTTGAGTTGAATAGAAATTGTCAGTACAGTTGAATATAACAACTGTAATTATTCAAATAATTAGTTGAT  
 ATAATAATAATCTTCAAAAACATTCTTACAGAACTGACTGAAACCATAATAAGTATCTCACTACTTTGTGC  
 GACAAGGTTCTCCTAAATGAGATCAACAACTTCTTAAGGATTTTATTATTACAAATAT  
 TACTCAATATCTAAGTAATCTTAAAATCTGTATACTCATTACTATAAGCATACTGCAAACCGTACAATAA  
 AAACTAATTAAAGGGGAGGAGAAGGGATGAACAGAACCCCTATTAAAGCAAAATTAAAT  
 CCAACAAATTGAACTGTTAAGATTAAAAAGAAAAAGAAACTTCATTGTCAGTAAGTCAAAGTATCACT  
 AGCCGTAAAATTGGCACCTTAAAGAATGATAATATACTAGGGAAAGCCTAAACCTCTGGTATTAAATA  
 ACTATAGTTAGAAATTAGGCAAAGCCGCTACTAGGGTAGTTAATTATTGGTCTGTTATTATTCATTCTCGT  
 TTTGTTGATAATGATGAGCAACGTTCCATTACATGAAGAAACCACGTCAATGTTATTCCCTATTGTTCTAA  
 ATTATTCGAGATGTGAGTAAACCTCCCCCACACACCCCCCTTTTTGCTGTTAAAGTTGACAA  
 TAAGAACGTTAGAGTTCTGAACCTGTTATTCTGCAGAAATTATAATTGGAGTCAACCACTGCATCCAA  
 ATTGATGCAGACTTATCTGAATAATTACAGTATATAATTATAATCTATTCTATGTTCTAAATTATGTG  
 ACACATTCTTATCGTTAGTTAAAATGATAAATTATCAGTATATAATTATAATCTATTCTATGTTCTAAATT  
 ATTATCCATTTCACCTTAATGAGAAACTTTGTAACCATAAAATGTTATAGAATATTTAGAACACAAGT  
 TTCAAATATTTCAGTCACATAATTAAACACACAAATTAGAGCAAATTCAAACACTTCACATAATTGAA  
 ACAGAGGGAAAATAACTCCCTCCGTTCCCTTACTGGCACGTTGACTTTACGCCCTTAAGAAATAA  
 TAAATGAAGTGCATAATTACCATGATAACCCATTAAATTGATGTTATTGAGTTGAAAAATGATTGAA  
 ATGAGTAATAATATCGTGGGTATAACAAAAAAATGATAAATTATTCAGAATAATTAACTTTAAATT  
 TTTTATTTCATTCCATTTCACCTTAATGAGAAACTTTGTAACCATAAAATGTTATAGAATATTTAGAACAA  
 CAAGTTCAAATATTTCAGTCACATAATTAAACACACAAATTAGAGCAAATTCAAACACTTCACATAAA  
 TTGAAACAGGGAAAATAACTCCCTCCGTTCCCTTACTGGCACGTTGACTTTACGCCCTTAAGA  
 AATAATAATGAAGTGCATAATTACCATGATAACCCATTAAATTGATGTTATTGAGTTGAAAAATGAT  
 TTGAAATGAGTAATAATATCGTGGGTATAACAAAAAAATTATCTCTTGATGCGTAAAGTGACAAGCA

AAAATGAAATCTATTTAGTATACATGCCAATAAAAGTAACGGTAAAATGAACGGAGGGAGTATATATTAT  
TACTACTAAAACGGAAACTGGCATGCTAAGTTTCATTGCTATTCAAGTCCGATTAATAGCCATAATATAA  
TGTTCTAGAAACAAAGAACACGCCAAGATAAATTTCCAAAAAGTAGCACAAATGTAACAAATTCTTACCTC  
AAATTTAATCTAAACACATAAATTTCTTATAATTGTTCTAAAAACTAAACTAAATGCAACAAG  
TTAAATGAGCAAGGAAGTATTCAGTTCTATTCTGAATTAAAGAATGAAATATTCGTTGTTAAGTTACAAACCGTAAGGGTGGTAATT  
CTGGTTCTGTCTAGCTAGCTTAAGACGAAAATTGCTTGTAAAGTTACAAACCGTAAGGGTGGTAATT  
AATTTCAATTGAGTATGAGCAATTAACTTCAACACAAAATCTACGTGAAATAATATACTCTAGTGTTGTT  
GCGTACGAAAAGAAAGTCTACATAATTAAATGCAGCGTCCATAGTT GATATTAAAATAACCCGTAATAAATT  
ATATAAAACTTATGTTATTCTTGTAAACAGCAACCAGACACACCAGGGCTACTTGGTCATTCT  
CCAATAAAACCCCTGTTCACCGCTTGTAGGTGTTGAATTCAACAGCGAACATTCCATTAACTCCGATCC  
ACTCTCTTCTTAGGAAAAGAATAATGGCTCTCTGTAAATTCTATAATTCCCTGAGCACTCCCTTC  
ATTACATTCTAAAAACCCACTTACAGTGTCTCAGTTCACGCCCTCAACCACAAGGCATTAGCTATATGC  
TGCTCAAATATTGTCGGCGCCGGAGTATAACATGAAAATTCTCTGACCTGATTATGTTGGTATTTGACA  
CCAGTCTCGCGATGGCGAACAGGCCGCTGGTGTACCATGACTAGTAAAGAAAAACTGGACATTGCACGTCAGTT  
GGCTAAGCTTGGTGTGATGTTATTGAGGCCGTTTCCCTTGCCTCTGAAGCTGAGTCAGCTGAAAGTTG  
ATAGCACAGGAAATTGTAACCTTAATGTTAACCGTTCACATTCTAATATTACTTATTGTAACATGTC  
ACGTGTTAGTTCTTCTTATGAAACAAACATGCATGCAAAGATATTAGTATTTGGACGGCGAGTGAGA  
TTGAAACTAGGACCGTTGCCTGATAACATTAAATGTAACCATTGTAAGTTAAACTGTTGATA  
GTAGCATATTCTTACTTTATTGTAAGTACTATATTCAACAGGTAATAACGTAGACAAAGAAGGATACGTGCC  
GATGATATGTCCTTAGCTAGGTCTAGTAAGAAGGATATTGAAAGAGCTGGGATGTTAAAGTATGCAAAGAAA  
CCAATGCTCATAGTAAGAAATATTCAAGATCGAATTAGGTCAAAGTGTAACTGAAAAGTAAAGAGAGAA  
TCTAAAGGTACTTATTGTTCTGGCTTTCTTTAATATAAGCGCCCTGAGATTAGGAGGTTGATATA  
GACCTCTACCATGTAATTCTTATTGCTTTGTTAGCTTAAACCTTCCCTGACTGAAATAGATT  
TAATGGCCTTAAATACGTGGAGAGCAAGTACTAGGTGGTATCTACTGGGATTAATACAAAGCATATTCT  
GACGAGCAAATGGTACTCTCATCTTAATCTTAGTAAAGGATATTACAGTTAAAGCTGCTTAATT  
TTGTCCTTCTGTTGACACAATTGTTGCTATGTAGGTAGAAGAGTACAGTGGCTTAAGCTGCAGCCACATAAGG  
CCATTGTCGGAGCTAATGCATTCTCATGAGAGTGGCATCCATCAGGTTTGCCTTAAAGCTTCTAAAGTT  
ACTTCTACATACACTAGATTGTAAGGATTACTGAATTGGCACCTCTACACTAGGTAAATTGCAATAAAC  
TTCATACCAACTACCGGTTGATATAAAATAACCTTCTCGTTAAGATAAGATGGTAAAGAATTCAAAG  
TTGAATTATTCAATTAGAAATGTATCATTGTTGTAACAGACTAATAAGGAAATTGTCATCTAAATA  
CACATGAAGTATTGGATTCAAGCCCATTGCCCTCCCTTCAATTGCTTAATCAGGATTCTACTCCAT  
AGTATCATTCAAGTTCTTCAATTCTCTTTAAATTGTTAATGTGGTTAGGATGGAGTGTAAAGAACAGAGA  
TACATATGAATTGATCTCTGAAGATGTTGGGTATCGTGTGCTAATGAAAACGGTATTAGTCTGGAAAGCTC  
AGGTATGATTGATTTCAAGGTGATCGATTGTTATGTTGATGGTTGATAGAGATTATTGTTCAAATTCTTG  
GTCTTCCCCATATTGATGAGTCATAATTGACTAATATAGTCTACCAATTGGTCAACATAT  
GTACGAATTGAATATAGTCATCAACTCAACTATGCTTCAACTATCCATACTGTAAGGTATACTCCCTTACATTAAAA  
AATACAAGTGAACAGAAAAAGACATTAAAGTGTGCTATCTCAACCGTTAAATTGTTAGATT  
AGATTATCATAACAGTTCAAACACTTAAACTCGTACTATCCCCTAGAATTCAAGTACCTCCGTTGATTGAATCA  
ACCCCAACCCCTCATTTCATTACATTAGGCACATTGATGTACCTCTTCAAGGAATACATTGTAA

CATGGTAGAATTTGTGAGTGGCCGCATGCATTGAAAGCCAAATGGCTGAGGTAAGATAACATAATCTATAG  
TACGATATACTACAAATATTAGTATGGAAATGGATCAAATCATTGTTACAGCTGGATATGACT  
TTGATGAAAAGAACCTGATGACCTTCGTCATTCAAGTCAGTGCTGAGAGGAAAAGGTGATTACGTTGCT  
GATATTCACTAGTCAACAGTCAAATTGTGATCGCACCTTCGTTCTAGTGTGTCATATTGAGTTA  
ATTTCAAATATGACTAACTATTCTACTATAACAATAAAATTAAATTGAGTCAAGTGA  
ACAATAAAATATTGAAAATATGAAAATTACGCCAAGAAAAACTCTTGACTCTGACATCCGAGC  
ATCACCACATAGTAATAAATGAGAAAATTATTGAGTCAAAATTAAAGTA  
ACTCAATTACTATTACTAAATGAGGAAACTACGCACTATAGTGGCAAAACCTATCCATATTCAAGGTGACTTTAAGTACCGGTCT  
AGTAGGTAGATTGTTACTGTTAGTGGATAAAATTAAAGTCAGTTGATAATTGGAAGAATA  
GCCTGACAGACATCTCACTTATTGGCTTGACATCCGGCCCCACTCCATATAATTCAACCAAAACTTAA  
ACTTGATAAAACATTAGTCTAACACCTCTGATACTGTTGTTCACTCGTGTGATATAAGGACACAT  
CAGATCCACCTCAGATATAATGCCACATCATTCTATTGAGAAAACACTCA  
CTCATCCCTTCTCCAAAAACAAACCTTATTCTCTCCACCATTCCATCTGCCTCTCCACCC  
GTCACCGCCACTACCAAACACTGCACCCCTCTCCATGAGAAAACACTCA  
CACTCCAAACCAACTATCTCCTCTCATATTGACTCTTCAAAACCTAGACATTGAACGAG  
TTAATAAAATCTATCCTAAATTCTTAGCCAAATACCCAAAGCAAAACCAAAAAATTGAAAAGAAGA  
GAAGAAATTCTACTCAAAATTCTATTGACCCCTCAGCAACGATAAAATTATAAAACTAACAAATCAAGCTT  
CGCGTTCAACAACCGATGACGATCCAATTCTGGCTTAAACACTATTCAAGCTTCAATGTC  
CTGCTAAACTCTATTAGTTGGTCTGTTACTGCTCTTTATTATCATAGTCCAGTTGATTGCTAAA  
ATTGTAAGAGAAATAATGTAATGCTGTTGAGAAAACACTACTTGAGAAAACCAAAATTAA  
ATAGAGGGTGTGGA  
GCGTGTCAATTCAATGGCGTTTAAGAGATTGGGGTAAGAAGAAGCGGCTGGATAGAGAGGGCGAGATAA  
TGGCTAGCATTGTTTAATTCTTCTTCAATTCTCTTTATTATTTCTAATCAAATTGTA  
ACTCGCTTCTAGGAGCGTCAACTTATTGTTGACTCACAAACAAAGGCTACATAAGTTGGTCAATTGGT  
TAGATGAGTTAAACTAATGCTTACAGTCAGTGCTGATTGAAACTATGTTGAGTAGGGTGTG  
GTCAAAGCCGAAAGTCGAGGCAGTGGCGAAGCCAAGAAATTCAACAAGGGTATTCAAACCTTGCCAGTGGCT  
ATGTAAGGGTGTCAAGCCTATTAAATCAATAACAAGTAATATTACCTTACGGAGTATAATTCTGGC  
GAAGGGTAGTCAGTGACCACCCCTGATTGACGTAGCTCGCCCTGAGTCGAGGTGTCTTAGGCTAAATAG  
CGTATAATTAAATTTATAAAATTATTGAGTAAAGTAATAAAATTATTGAGTAAAGTTAAG  
TAAATTATAAAACTTAAAGAAAATTGAAAGCTTACATATGAGGTCGGCTCGACCCACTTACGTTGATCCC  
CCTCTTCTTCTTCAAAATATGCGTCTCTGAATAGCTCATGACTCATTGGACGCCACTTCTTCT  
CTGAAAATAATTAAATTAGTTAATAAAATAA GATATACTTAGAAAAGTAAATAATTAGCAATTACGTA  
CCAATCTTCTTTATTGTCCTAGGCTGATCGCACCTAGTGTGCAAGGAGCCTCCCTCGGATGCGCAGC  
TTCTTCTTCTTCTGCTCTCAAGAACCTTGCAGGTAAGCCATTGCTTGCATACTCCACAAATTCTCA  
AGTAACCCAGCCAGGCCTCTGTTCTTCTAGCATCCGAGAAAGCATCCGCAATCTTACTAGCTTGTGAT  
GAAAATTGAGCCACTTCCGCGTATAGCGGTCTTCCATACACACTGCTGTATTAAAATTAAATT  
GATGGAAACATCATAAAATATAAATTAAACTGTTAAAAAAATTACCTTAAATTGATTGAAATTGCT  
CTTCACTGAGAATGGACAATCAGTCCAAGTCGATAAGGCCATCATAAGCTTCTGATGGCATTAGT  
GATTAT

CTTCGTAGTCTTATTACCCGGCCTGAACCTACAATTAAATAATAAAAACACATTAATATCTTAGTAAAATGTTAAC  
AAATCAATAGACGGAAAAATAATGAATAACACTTACCCATCACCTTAGGGACTATGATGATCCTGCCATATCGAT  
CATATGCACTACCTCGTCATCACTATTGAAGCATGTGTATCAGAGGCATGTGAAGACGGTGTAGGCAGGTCAA  
GCTACTGCCTCGTAGCGAAGGCCGCAATAGATGGAGTCGATAATGAAGATGGTTGTGATCCCTGTGACTACAA  
ATAGCTGGATGTGACCCAAGTGGATGTGATAATGCTGGCTGTGACCCGAGTGGCTGTGACATGGATGGATGCGATC  
CATGTGGCTGTGATATGTAGGGATGTGATCCATATGGATGTGATGTAGATGGCTGCGATGCAGATGGCTGTGAGGC  
AGCTGGACTGTATGTCGAACATATAGTCCTATGCCCTGTGATGGAAGACCTGGTGTGACAGAAGGTGTATGAT  
TCGTGATGCTGTGAAACTCAGTATAGCCGTGGTGAGGATAAGACATAGGCATCTCTGAAAAGGGTGGACAAG  
AGGGACTATTATTTCTACCCCTCTTCCCTACCCCTTCTGACCCCTGAGAACTAGTAGGGTCATTGTT  
ACCTTGACCCCTGCCATCTGTATGATATAATGCACATTAGATTGAAACATATAAGAAACTAGCTATAAAC  
AAGAGTGCTTAAAAAACCAACTTTAATCCTCATCGACGTATTGTTCTCGTCCGAGAAATTCTCGCCTCATT  
TGTTGAGCTTCATCAGTTGATTCTCGCCTCATTTCTTAATTGTTACTTCATTATACACCTCTTCAAT  
ATGCATTCAAGGATGTTCAAATCATTCTAACTGATCGTCCACTATTGGTGAATATTAGAAATATCATTGAT  
ATGCAATATTAAACACATTCTGACTTCCACCCCTACCTACATGCTTAGTTTATTACAACCCACCAATCGGACTT  
ATTCCGCCGCAATGGATAAGGAGCATAATACACTTGCTAACGTTATGTGCAATTATGAAAGGATCATAGCGATCA  
TACTCCCTCGATGATTAACCTCAATTATGTTGATTGGTGTACACTGTACCTCTGTTGGATTGGTCAA  
ACCACTTGATATAAAGAGTATCAATTCTATATGCCAACCTGTATATTCTAGTTGAATATTCTTGACCAC  
ACCATAATAATCAATATCTCAACTGGTGCATCACCACCTGAACCCACACCCCGCTGTTACTATT  
TTTTAGAGCCATCCTTGTATGAAACTTAAACCAATTCACTACGTACTTAGACATTGTTGACCTGAAGCCCAG  
GTCCCCAAGATATCTTCAAAATTGATTACACCATTATTGGATTATTACCTACATAGTGTAAAGGTT  
CATAAAGTTAGCATAACTTATAATCAATTCTATACATTCACTACATATATATATATATATAACTAGTGT  
ATATATATATACAAACTGTTGAACACGTATCAAATCTGTATATACAACATCATGCCAAATTGACCTACGAAGT  
GACTGTGACACATCAAATTGTTAGTAGTTGATCAATATGTAGTCACAGTAAATTACATTGATAACTAAT  
AAATCAATACTTACTTGAGAAATGGTACAACCTCGGAACAATTAGCAACACATGAAGTGTAGCTGACTTGACTC  
CATATCACTCAAACCTCTTCTAACATCCTAGAACATCGGCCTGGTTGATTGAATATAGATATTGGTGGATAT  
AATGGATCATTCACTATTGACCGTGTGCCTATTGGCCTATTCTAGAACATGGCACGTTATTCTCAAATAAT  
AAGAACAAAAATATGCAGTTCTTGCAAGATAGGCTCGATATAGATCCTAACATCCCTGCTTAACA  
AATTGTTGCATTGCCAATTGCTTACATAATGTATGTTAGCCAAGAACATTCAAATAAAATAGAATATTACAT  
CAAACCTATAATATTACCTCTCAAAGGGATACATCCATCTGCATTGAACATGCCCTCCAAGTCGCTCGTAC  
AAGGTGTATTGGAAGGTGTTCCATTACATCAAAGAACCATGGGAATTCTTCCATCTTACTAGAAATTACA  
CGAATGTTCTGGTCCATCCGAAGTAGGTTCTTCCATTGAACACAGCCTTGAAAACAAATTAA  
TCTTATGATGGTTCCAGATTCTTCACTGCAAGTATAATTCACTGTTGAGCAACATGGTACTAAGGTCTCCATGAAAACATG  
GCAGTCATGACTTCAAATGGCTCAACTTCCCTACCTCCATATCTATTGTTCAAGATTGACGCATAACCCCTC  
AGGCATCTCAATTGTAACCCAAATCACAAATTGTCGTCTTCCCAAATGAATGTGTAACCTGCTTGGAC  
TTGAACACCTTACATTGTTGCTGCAAGTATAATTCACTGTTGAGCAACATGGTACTAAGGTCTCCATGAAAACATG  
CCTCGGGGTATCTTGTCTTACCTTAACATCCATCACTATGTTGAACAAATTGTCAAATAATTCTCTTCTAT  
ATGCATGACATCAAGGTTGTCGGAGAAGATTATCCTCCAATAAGGCAACTCCAAAATATACTTGTGTC  
CAATTATGATTAACACCATACTCGAGGAATCTATAAGGTGGAGCCTCAGTAACCTTACTGAAGTTCTGGACCCCTC  
CCCAAATTCTCACCTGAAAGTATCGGAGGTGGAGAATCATATTCACTTATTCTTGAATGCATTTCAT

CCTCTAAACTCATGATCATCAGGCAAGAATTGACGGTGACAATCAAACCATGATTACTTCGGCCATGTTCAAA  
GTGAACGCTTACTATTTCTACAGTAAGGATAAGCTAGCTTCCAGCAGTCATCCACCCAGACAACATCCAT  
ACCGAGGAAAATCGTAATAGTCCACATTAAATTAGGACGCAAATTGAAATTCTGCTGGTTGATATGTATATGT  
TTCAACACCACATACCCACAATTGTTAGCTCATCAATCAAAGGTTACAAATATAAATCAAACCTTTGGA  
TTACGTGGACTGAGGATAATATAATTTCAGAATATATGGACTAGTCATACAACTAAGGTGGTAGATTATAAG  
GTATAAGAAAGACAGGCTAACATGAGTATGGTGTGCAGATATAGAAAAAGGCGTAAGCCATCCGCACACAGACC  
TAACCGAATGTTCTGATTCACTAGCAAATCTGGATATGTCCTATCAAAGTGTCCAAGCTTCCATCTGAA  
AGATGACACATAACACCGGGTGGTCTTCTATTTCAAAGTGCCATCTCATATGAGGAGCAGAACTCATGACGCAT  
ATAACCTTTAACCTAGGTATAAGAAGTAATAATGCATGCCATTGACAACGCCATTCCCACGGAAAGCCT  
CTTGAACGAGGATTTCGAAAATTACAACGTCTAACTAGAAACCAATCTCTTGCCTCATAGAAATCACAGGTA  
AGTTGATATTAGGGTCAACTAGTCACTCATAAGGTCAATGAAAGAGTCCATGGCTGCTGAGAAATATTCAAC  
AGATTGATACTTAGTAATCTAACAGACAGCTCAGAGTGCAGACTCCTCACGTAGTAGATGACTAGCT  
TCCTCTAGTGTCTAAACATTTGCGTCTCATTAGGAGTTGTTCAATATTTCATTGGCTCACCCCCGAA  
GTGCATCCAAAAGCATCCGAACCATATCCTGAATTCTAGAATCAAGATTGTATTCTCCACCGACCTACTACTT  
TCACCAACAACCATGTTACGAAATATCCCACGGCTACCATCGATCTCCATGATTAGTCCACACAAAGTAATTCT  
CTATAACCCCTTCATAAA GATGAAGCTTAATTCCCTCGATTGGTGAAGTGCATTGATGCTAATAAGTCATCAACCCCTCT  
ACAAAATCCTCCCGCAATCCCCGCCATTAGGATAATTCTATTGTACATCCAAGTACATGTTCCATCTACAAAT  
AAAACAAGAACAAATTAAATTCTACAATTATAAATTATCTTTTAGTTAATTCAAGATAATTATTGT  
TTCTAATTATACCAATTATATCCTAAAAGTTCAATTCACTACACCCAAAGTCCATTCTATCTAAAGTTAAT  
TCATATCCTAAAAGTTAATTCAAGAACAAATCCTAAAACCTAAAGTTGATTCAAGAACAAATTAAATTCTACAA  
TAAACGAACTAAACTAAAGAAATTCAACCCATACCCAAAAGTTGATTCAAGAACAAATTAAATTCTACAA  
TTATAAATTAAATTATTTTTAGTTAATTCAAGATAATTATTTCTAATTACACCAATTATATCCTAAAA  
GTTCAATTCATCCAAAAGGTCCAATTCAATTCTAAAAGTTCAATTCAATTCTAAAGTTCAATTCAACTAGAAC  
AAATCCTAAAACCTAAATTCAATTCAATTCTACGAGTTAAACATAACTAAACTAAACCTAAAGTTCAACCC  
ATACCCCTAAAAGTTGATTCAAGAACAAATCATAACCCCTAGATTCTAACTAAACTATTCAAGAACAAATTCAAAGT  
TAATAATTCAATTCTAAGACTATTCAAGACAATACAAACTCAAATTGAAACTCATCAATTAAACTATTCA  
ATAACTAATTGACTAACTAACAAACTATTAGTTAATAAAACTAATTAAACAAACTATTAAAGACACCTAAC  
CCTAATTCTCAATAAAATGCAATTGTTCAAATAATTGAAACTATTCAATTGAAACTATTCTACAAATTACAA  
AACCTAGAAATAATAAAATGGGTTGTAATTCAAACGTAGAATTAGGAGATGGAGAAGGGAGGGGGCAG  
TGGCAGTGGCAGTGGCAGTGGCAGCGCGATGGCGATGGCAGCGCGACGGTATGGCTGGCAGTGGCAG  
GCAGGGGTGGGAATGAGATTATGTGAGGGAAATATAGAAGGAGAGGGGAAGGTTTGAGTGAGAAAAATAGAGAAG  
GAGAGGGGAAGATAAGAGAGAAATGGGGGTTCCCCCGTTTCAATTCTGATTTCAGAATTACCGACCAAAGTTGGT  
CAGTAATTGGTCCGTACATTAAAGTGTGACCGTTGACCAAAACCGACCAACTTCGGTCCGTTTTTTAA  
AATTAAACTTTTTGGTGTGTTTCTAAATATTATAAACTATAAAACTATTAAAGTATAATTACAAATAATTATTAAC  
TTTATTAACTTAATAAAATTATAATAGCATCTATCTAAGTAAATTCTAAGTTAATTAAAGTATGTGAGTAATTTC  
TCACATACACACATACACTATTGTAATTGATGCTAATTACTCTTTCTATTGTTCATCAACTAAATGCATG

ACATAGATTAATCGATATAGTCGAGACGTTGATGAATCATCGATTAATATTGATACATCTAAGTAAGT  
TATAAGTCGATGAATCAATTACAAATAGATATTTGATGATAAGTATATACTAATTAGTATCTTCCCAAGTC  
TATCCCTAAAAGAACCGACCCTATGGCCTAGCGCTCGTCTTATATACGGCTACCTATATATAT  
ACACACACACACACACACAAATACACTCACTGTAATACTTAACATATATATATATATATATATATATA  
TATATATATA

[0319] SEQ ID NO:7 (NtIPMS2对应BAC重叠群序列(总BAC长度10082bp))

[0320] TAAAACACGCTTCGTTGACTACTTAACATACATATAGCATATTATATAGTATATGTTAGTGT  
TTCATTATTGTATTACAAAAGTGTAAACGAATTACATTATATTAGTAAATATAAAAGTAATTGAA  
AGTTGACCAATGTTGACGTAATAACCGACCAACTTGGTCGGTATTTACAGAAAATATCAATTACCGACCAA  
GTTGGTCGGTAAATACCTTCCCTGCATTAACCGACCAACGTTGGTCGATAATTAAATATAATTCTTAAATATT  
ATAAAATATTAAATAAAATAATTATTAAAATTGGTCCAGATTACCGACCAACGTTGGTCGG  
TAATCCAAAATTTCTGCTGACCAGCTAGGTCAATTGTTGACCAATTGCCGACCAACGTTGGTCGGTATT  
AGTCAAAAAAATGTAATTTGTTCCAGTTCTGTCAGCTGGACGGTTGGTCGTTTTTT  
CCGACCAACGTTGGTCGGAAATATTGGTCAGTTTATCAGATTAGTAGTGTCAAACACTTAAACTCGTT  
TTATCCCTAGAATTCAGTACCTCCATTGTATTGAATCAACCCCAACCCTCCATTTCCTACATTAGGCG  
CATTGATGTACCTCTTCCTTTAAGGAATATATTGTAACATGGTGAATATTGTCAGTGCCACCATGCAT  
TGAAAGCCAAAATGGCTGAGGTAAGATAACATAATCTATAGTACGATATACTACAAATATTAGTATATGGACTTG  
GATCAAATCATTATTGTTACAGCTGGATATGACTTGTGATGGAAAGAACTTGATGACCTTCTCGAT  
TCAAGTCACTGGCTGAGAGGAAAAGGTGATTAAGTTGCTGATCTTCATAGTGAATGGTCAAACCTGTGATCGC  
ACCTTTCTGCCTAGTGTGTTCATATTGAGTTAAGTATGACTCAATTCTATTATAATAAT  
AAAATTAAATTATTACATTAAATGACTGAACTAACAATTAAAATATTGAAAAACATATGAAAAATT  
ACGGCCAAGAAAAATTCTTGAACATCTCAAACATCCAAGCATCACCATAGTAATAATGAAAAATT  
TATCACATTAAAGTAACTGAATTACTTACTATTGATAATAAAAGAAAACCACTAAACTTACGCACATAGTG  
GAAAACCTATCCATTCAAGGTGACTTAAAGTACTGGCTAAAGTAGATTGATTTACTGTTATAGTG  
GATAAAATTAGTAACTTTTTGATAATTAGCAGAATAGCCTGAGTGACATCTCACTAATTGGCTTGACATC  
TCGGTCCCCTACTCCACAGTTCAACCAACACTGAACTGTATAAAATATTAGTTAAACACCTCTGATACT  
GTGTGTTGTCCTCGCGCTGATATAATGACACATCAGAGTCACCTCAGATACATTAAATGCCACGTCTT  
ATTTTACAAAAATTTTTAATAAAATCAAACATCATTCTTTCTCCAAAAAGAAAACAAACCTTA  
GTTCTCTCCACCGTCCATCTGCCTCTCCACCCATCAGGCCACTACAAACACTGCACCCCTCTTCC  
ATGAGAAAACACTCAATCAAACCACTTTATCTTTTGCTCAAATTCTCATTACTACAATCTATTCAACC  
ACCCATGGAAGTTGTCGCCACCATGCCAGCAACCACTCCAAAACCACCACTATCTCCTCTTCCATAT  
TCATGTCACCCATTCAAAACCTAGACATTGAACGAGTTAAGAAAATTATCCTCAATTCTTAGCCAAAC  
ACCCAAAGAAAACCCAAAATTGAAAAAGAGAAGAAATTCTTAATATTGGTACCTCAAATTCTTA  
TTAACCCCTCTAGCTACGGCAAATTATAAAACTAACAAATCAAGCTCGACGTCTCAACAAACGATGACGAC  
CCAATTGGCTTAAACACTATTCAAGCTCAATGTCTTAAATTCTCGACGTCTTATTGTCTCTCAGT  
TATTTGAGAAAAACAAAAATTAAACAACCTTGTTC CTTGACTAGCTGCTGCTAAGCTCTATTAGTTGATG  
TTGTTATGCTACACCTTTGTTTATCATAGTCCAGTTGATGCTAAGATTGTAAGAGAAATAATGTAATGTT  
ATTGAGAAAACACTACTGAGAAACAAAATTAAATATTAGAGGGTGTGGAGCGTGCCAATTCAATGGCGTTT  
TAAGGGATTGGGGTAAGAAGAAGCGGATGGATAGAGGGCGAGATAATGGCTAGCATTGTTTAATTTTA

ATTTTTATTTCTTATATTCTAATCAAATTTGTATAGTTACTGCTTTAGGAGTGTGCAATCACT  
TATTTGCTGACTCAACAAACAAAGGCCGTATAAGCTGGTCATGGCATATGGTTAAAACATAATGTCTTA  
TACAAGTTCAAGTACCTGATTGAAACTATGTGGAGTAGGCCGATCGTAATAGCAAAGCCGAATAAGTCGAGGTGTC  
TCTTAGTCTAAAGTAGCCTATTTATAGTGAGTAAAGTTATGTCATTATATTATATGAAATTAAACCTATAG  
GTTTCCAGTTCTTGGTAAGCTTACCTAATTAAAATTTACCTGTGTCGGTATAATGAATATGTATGTCCGCCT  
ATTATCTCCCTACATCTATTTCAGAAAATTACAGATGATGACTTGAGAGCACTGTATCAGATGACGTTTC  
CAGCCTCAAGTTCCGGCAACTGGAGATGTACAGGTACATGTAATACTATTCATGTGTTTGCCTAAGTTG  
CTCTGACACGGCAGTTAGGTGACACGCCGTGCGCAAGACACTAGTATGAGTGTGGATGTGAGATCCGTACCG  
GATCTGGTCAAATAATTGGGTACTTGACCGCAGCGTACGAAAAATTAGAGATAAGATAACAATTGATCCCG  
AAATCAGAACCAAAGCTAGGGTAATTGAACAAATAGCATACTATCTAGGAAATCAATCCTTACTTATCTA  
CAACTCGAGAATAAAAAATAACACATTACAAGTTACGTAAGTATTCCACAAAATTCTCGTAATTAGA  
TATATTATTTATTAAATTGAATTATTTAGTCGGATCCCCGACCCGTATCCGTATTAGGATCAGTAT  
CCAAAATCTAGAATTACATCTGAAGGATCCGACCTCTAGATCTGCACCTGTATCGGATACCGTACCGTAT  
CTGAGCAACTTAGGTTTGCTAGTCTCACTTAGTGTGGATTGAAAAGGAATTATGTTCACCTAAATCTAATGA  
TTGCTGTTATTGTACAGATTACTGTGGAAATGTTGGCGCTCTACAGCAAATGTTAAGCTTATTGACAGCGAT  
GGTCAAGAGCACACTGCCCTTCTGTTGAAACAGGACCTGTTGATGCGCTTACAAGGCAGTTGACCTCATGTA  
AGGTTATAATGAAAAGCTCTACATAATGTTGATCTGTCACACTAGTAGGATAGGAAATTAAAGTTTA  
TATACGCTGAAGGTATAAAAATTTACTCAATCAGAGCGCTTACAAGACAATTGCAAGGTAACTCTATGAGTAGCT  
TAGACTAAAAGTTAATTATTAAATTATTTATGCATGCCTCAACACGTTACTCACCTTACGCCTAATT  
GTTACATGGGATCGTTCTACATTTGAGTAGCTGAGAGATTGAACTTAGGACCTTATTTCTCTAACACTATGT  
TATATTGCCAGTGGAACAGGAGCGCATATGACATTGCTTCAAGTGTCCAAGCCTATGTTGAGGCCTGTAACA  
AAATATTCAAGTTACAAAAAACAGGTCTCGTAACAAATTGAAGGCAGTGCAGTAAAGTGTGGTGT  
GCTCCAGAAACAAAAGCTTAAATGCTTACTGCTGAGTACGATCATAAATTGCTTATTGATGCATTAC  
TGAAGTTAATTAGAAGTAATCTTAAATTGTATAAGATGTAACATGTTCTATTGATGCATTACTATGAAATTAA  
TACAATTCTGAGCCAATTTTATTGGTCTAAAATTACAAATCTTATTGACTAACAGA  
TTTCGGGAATTAGGGCAAGTAGTCCAATCAGAACATCACATTAGCTGGAGAGTAAAATCAAATCTGCACTCCTG  
AACAAACCTGACTCAAAGTACCTTAAACATATTAAAACATGAAAAGGAGCAGCCATATGAGAAGCAGCCATAT  
GAGAAGCAGCCATGTGGCCTCACATTGCTTGAGCCTGCAGGTATTCTGAGCTGCACAATTGATCTGCAGCT  
GCACCTCAGCTCTACGATCGCACAATTGATGCGACTCCGACTCTGGCAAGGCTTCAGTCTGTTCTATTG  
ACGCTCTGAACCTTGTATTCTTCACTGCAACTATGTTCTGCGGCCACACAAATTGTCGGTCTGCAAT  
TCTCTCAATCACTCTAGGCTTACATTAGATCTATGCCGTAGAGAGAATTATGTCAGACTTTATTGTTGG  
CGGCCACAGAAACTTTGCGGACCGCACTCTGCTCTACTGCGCTCCTTGCCTTGTGAGCCGAAACACTCC  
TTTGAGTTGAATTCTCATTAGAGTACAAATTCCAACATTGCTTAATTGAGACTTTATTGTTGG  
AACATAAAATCAATACTTTGGACTAAAACATGAAAGCAAGAAGGTACTAATAAGTGGTCAAACACTTACACC  
AAATCCATACCCAAAATAACATCAAACATCTACCATATTAGTAACAAAGATTAACCTAGTCTCGTAGCCCCAATA  
GTGACAAAACAAGAACAAATAGACACAGTCTACCAATAGAACATCCCACGGGTAGACACATAAGTAGGAGTAT  
TGAGAGAACATGAGACATATCCAAATATGAAGAAAAGTAGGATGACACATACGAATATGTAGAACCCAGATCAA  
TCAAACAAATGCCTCTATGACAGATCAGAATGATACTGTGATAACTGCATCTGAAGCAACTGCCTCAGTCCTA  
CCAAGGAAAGAATAACAACGAGCTGGCCTCCTCTATGATGACCTCTACCCGCTGTCCTCACCTCCAGCTA

GCTGTGCAGGTGAAGTAGCAACTGGAGCGGGAACTGTAGCCTAGTACCTGCTGAGGTCCATCCATCTTAAGTCT  
GGGACAATCTCACCTGTCCTAGTATCATCACACTCTAAATAACCTCTATGCGAGCGTGCTGCTGGTACTGA  
GTTGCCCGTGACGATTGGAATGACCACTCTAGGAACCTCATGCAGAAGGCACACTGAAATATGTCCGGTACGAGT  
GCTCTAAGATCCATGGCTAGCTGGAGCACCACAGGAATGTGAGGTGCATACTGAACACTAGTCAGTACAAAGCCT  
CTGCCATAAAGGGCTAACGCTATAGCTAGTAAACCAACTAAATCCTCCAGAACCGAGGCTTATCCTCCCTAT  
CCTCCCTCTCCTGACAGCGATACGCTAACATCCTAGCAATCTCCACAAACCTGATGAAATGCAATCTCAGTCTC  
TAACCTCCGTGCCATACCATATCTGAGGCTAAAGGCAGTCCTCAATAATCTACGAACACTCTCTATCAGT  
GAAACCAAGGAAGGTGCATGACGAGATAACTCCGTAACCTCATAGCATACTCGGACACAGTCATGCTCCCCTGGC  
GCAACTGCTCAAACACTATTGACAACACTCATCCCTGCAGGTCTAACGCTCCATCAAGAAAAGATCAGAGAATAGGGAC  
CATGTAATGGTGTGCGCCAATATCATACTCGCCTCATAAATCTACCACCATATGTATGCCGAACCCGTCAACT  
GGAAAAGTAGTAAGGTTGACCCCCCTGACTCCACAAACCTATGGTACGTAGAATGCGATGATAATCCAATA  
AACTTTGTGCATCCTCAGAAGTCCCTCACTAAACTAGGGAGGATGAAACTTCAGTAACCTCTCCAAACACCTCTG  
TCCCTCAGCCGAGACAACACTGGCCCTACCTCGGGCTGAAC TGCGTTATTGGCTGCACCAACACAACACCCAGTGTC  
TGGTAAGCATGAGTCGGCTACTCTGGTGTGCGGGCTGCGTGAGTTGAGCTCCCTCCCCAGTATATGAAGTAGCTG  
GTGCAACCAGAATCAACTCTCCTCGGCCAAACTCACAAACATACTCATGAAGTATGCCAAGGTCTCCTAGAGCGC  
CAGTGTAGGAATAGGCTGCTCAGGTGCACGCTAACCTGCTCGGCTCTAACCTCAACTGGATCTGTTGAGAT  
CCACCGTTGCTACTCTAGTATAGCATGTGCCCTCCTCGGCCCTACCTCGGCCCTGGCCTTTGCGACTCTAAT  
AGCGGGCACTTGTGCCTGCTAGCTAATTGGTAGAGCGTGTCTCGCCATCTATGGGAGAATAGAAGAGTAAAGG  
TTCAAACTTGGAATAACAAAT CGACATGATAAAGAATGAAGGAAGGTGATGTTCTAACAGTTGGTAGCCTC  
TCAAAGATAAGTACAAATGTCTTGTACCGATCGGAAGATTATACTAAACCTGCTCATGACAACCTGAGACTTA  
TGAACCTAGAGCTATAACCAACTTGTATGATCCAATCTCACCATAGGGTCTTGATGAAACCTAGTCTATAAA  
ACTAGGTAAGTCGACCACTAACAAATTAAATCAGCAAAACATGATATAAGGCAGAGTTAATATGAAAGCA  
GAAAATAGAGTTAACATCAGC AAAATGGCAACTCTAAATATTCCCCAGAATTAGGTAGTACAAAGTCATGAGCT  
CTGATAGAATACATAAGGATGTCTTAAATTACAATACTGTTGAATATGAAAATAACAATATAATAGGGAGGG  
ACTCCAAGGGAATGCGCCGGTCATGCAGGTCTACCTGAATCCTGCGGTCATGTAAGCCTCACTGCCTAGGTCCG  
CTGCCCTCAACACCTAAATCTACACAAAATATGTAGAATTATAGTATGAGTACACTACTATCGATACCAGTAAGT  
CAAGACTAACCTCGGTGAAGTAGTGTAGGTTCAAGTCATTGATACTCACTAGTCAAATAAGCTGTACAATATAT  
CAAAGACTAGAAATAGAATGTATAATAGAAAATAATCAACAACTCTTTGGAACATATGATAACAAACTCAATGCA  
AGTAAAGGTTCATCTGACAAAAAGTTATCAAATAGAAACACAAACGTATGATACTCGTAACACAATAACAATATC  
CATCTGCATCACTCAGCCCTGACAAAATAACAATACCTAAACTTATCTCTCGTCTGACACAATAACAATATCCAT  
CTGTATCGCTCTGCCCTGATATAACAATATCCATCTGTATTACTCCGCCCTGACATAACAATAATCCACCCG  
TGTGTTCCACCCGTACACAATAACAATATCCATCTGTTGCTCGGCCCTGACACAATAACAATAATCCATATGTAT  
CGCTCAGAACTGTCACAATAACAATCACAATCGTACGACAGAAACCTCGTGCCAACACCCAAACAATCCGCCAACAT  
GTCCATATGTGTCACAATTATAACAAAACAATAATGTCAAGTTCATCAAATACTAGCTCACAATTATCAACAAGG  
TGCACAAGGACATGATAACAATAAAATGTGGATGGTATTCAACATATAAACCATGACTACGACTAATTCAATTGA  
AAAGATTATGTCATGAGTCACAAAGTGATTAAGTATGTTCATGTTAGAGGACATTCCAAATTAGGCATATTAA  
TAGCCTAAGGTCTAATCTAGTCATAAGCCGTATATAACGCCATGTACACGCTCGTCACCTCATGTACACGTCGCTT  
TTCACATAACACGAATAATAAGGACAAATCATAAGGGTATTCCCCACCCACAAGATTAGGCAAGATACTT  
ACCTCAAATAAGGCAAATCAACTTTAAAAGCCCTTGCCTATCGAATCGGCCTCCGAAAGGGTCAAATATAGCCT

AAAAACAACTCAATAATCAAATACGACTATAGAAATTGATTCAAGTAAGAAGCTCTGATCTTATCAAATATC  
 AAAAATTCACTCCGAACCTCGCCTCCCGAACCCGATAAAATTACAATCTCGAGTACCCATTGAATACGAGTTCA  
 AATGTGCAAGTTTATCCAATTGGACTCCGAATCGGGTTCAAATCTCTATTTTATTTAGAATAGTTTTCC  
 AAAATTCCAATTTCTCACTTGATTAACCAATAAAAGCCAAATTCAAGATGGAATCATGAACAATAATCAAATC  
 CGAGTAAAGAACACTTACCCAATCCAAGTGGGTGAAAATGCCCTAAAGGATTCCCTCAATTGAGCTTAAATCTCA  
 AAATATGATAAAATAAGTCAAACCTCAATTAAAATACTCTACTCAGTGATTTGCTCTGCGTACCATCCATCG  
 CATCTGCAATGCCGCTTGTGAGCCCCTCTCGCATATGCAAGTTCACTGCAGTGGCTCCCCGATCCTGCTCTG  
 TACTATTCTCGCTTCTACGGAGTCGCTCTGCACCCAAACTAGCATACTGCCCAAAGTCCGTACCTACGGACA  
 ATGCTGCCTTCCTCTCCCTCGCATCTGTAGCCCTTTGTCATCTGCAGACTCAAACACTCGTAGTGCAGAAC  
 ACTAGAATCTGAAGTGTACCATATAATCAAACACACCCGAGGCCTCGGGACCCATGTAATCATA  
 CCAATCAATCCAAAATATAACACGGACCTACTCGAGTCCGAAATCACATAGAATAACATCAAACCATAATCAC  
 ACATCAATTCAAGTTAATGAACTAACTAACCTCAAACACTCACGCCAAACATATCTAAACTACTTAG  
 AATGACCTTAAATTTCACACACAAGTCTCAAATTACATAATTACCTATTCAAACATTTCAGAATAACAATCCAGACC  
 CGACAGCCTCAATGTCAACTCCCGTAAACCTATGAACCTCCAAACCTTAAATTCAAACATTGCTACATCGCAAATAGA  
 GCCAAATAAACCTAGGAACCTCCAAATCTAAATCCGGACATACGCCATAAGTCTAAATCACAATACAATCTATTGG  
 AACCATCAAATACCATTCCAGAACATCATTATATAAAAGTCAAACATCGGTCAACTCCTATAACTTAAGCTGCAT  
 GCCTGCAGGTGACTCTAGAGGATCCCCGGTACCGAGCTCGAATTGCCCTATAGTGAGTCGTATTACAATTCACT  
 GGCGTCGTTTACAACGTCGTACTGGAAAACCCCTGGCGTTACCCAACTTAATGCCCTGCAGCACATCCCCCTT  
 TCGCCAGCTGGCGTAATGCAAGAGGCCGCACCGATGCCCTCCAAACAGTTGCGCAGCTGAATGGCAATGGC  
 GCCTGATGCGTATTTCTCCTACGCATCTGTGTTATTACACCGCATATGGCACTCTCAGTACAATCTGC  
 TCTGATGCCCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGAACCCCTGCGGCCATCGA  
 ATATAACTTCGTATAATGT

[0321] SEQ ID NO:8 (用于驱动毛状体特异性表达的启动子序列。此启动子分离自红花烟草，并位于编码参与顺式冷杉醇合成的copaly1合酶的基因的1.5kb上游)

[0322] agttctgcaaattcccaacattatcaccttgcgcctgtaaaggcaagtagctaaaggcatatttt  
 aagaatgaccacacaaaataatagtagtagtaaaaggactacatataattagacccatcatataggaaaa  
 ctagccaatactttccaattaagattaggttttttttttttttttttttttttttttttttttttttttttttttt  
 tgcaaaaacacagaaaatataatccaaactaccccttcctccttaaacattttttttttttttttttttttttt  
 ttgaaacaatcagggatattgttaaggtaataaatgactgggttttttttttttttttttttttttttttttttt  
 ttcaatttgaaaaacgactaataattaaaggactggtaacgctacgtttaactacactaagaaaatataatgtgt  
 tgacaatataatgaagtcaagaataagaactggcatttatatatgtttcaagtaggttttaggttatggcaaaat  
 taataaggacttaatttgcgttacaaaataatgttttttttttttttttttttttttttttttttttttttttttt  
 aatagagagttgcctatagtttgcattcgctcgcccttttttttttttttttttttttttttttttttttttttt  
 aaattttgtgtgaattgttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt  
 taagaaaaataaaaaatttgagtataaaataaaaggatgacattgcacttcttttttttttttttttttttttt  
 acttcttgcgcacgccattgattcaatctatcatagttctgttaatgttttttttttttttttttttttttttt  
 atttttaataaagtcaatcatttttaacatctagatttgcgttttttttttttttttttttttttttttttttt  
 atagcactaaggactaattgtgacttgcaagttaaaatacaatgcaatgctgctgaaacaacag

agcaactcgttccagtaaaaatctaaagttaactactttcacaaaaactaataaaaagtttagagtgcgttgac  
aatttatttcatgcaaaggatcgagaacaatcaacacagaattaggctgaagtgtctaagagaatttatattgc  
ccttcattcaagaggcaatataaataaaggcctgcgccattgcaacaactcaaaccattccaatattgcctcaca  
gtcagtagtgccttcctctcaaacgttcattgtcttatctccctccccattctcattggaagaataaaac  
aaaaattaaatttagaaaatg

[0323] SEQ ID NO:9(用于通过RNAi机制触发NtIPMS2基因表达抑制的合成构建体。由NtIPMS2第一内含子的片段加上第一外显子组成,第一内含子的片段在外显子之后是反向重复的。理论上这种构建体的翻译和剪接导致呈发夹结构的双链RNA)

[0324] gctaagcttgcaccatggaaattctctgaccctgattatgttgtattttgcacaccagtccgcatggcaacaggcccgttgtctaccatgactagtaaagaaaaactggacattgcacgtcagttgtcaagcttgtgtatgt  
tattgaggcccggtttcetttgcctctgaagctgagttcgagcttgtaaaagttagatagcacaggaaatttgtaaCC  
ttaaatgtttaacccgttcacattctaataatttacttatttgtaacatgtcgtacgtttagtttcatctttta  
tgaacccaacatgcattgtcaaaagatattttagatattttggacccgtcgagttagatgtaaactaggaccgttgcctg  
atacaatattaaaatatgttaaccattttatgttacaagttaaacttgtatagtagcatatttttacttttattta  
agtatactatattccaacagtaaGTTAACcaatttcgtctatcaactttacaagctcgactcagcttcaagag  
gcaaaaggaaaaccggcctcaataacatcaacaccaagcttagccaaactgacgtcaatgtccagtttttcttact  
agtcatgttagccccagccggcctttgcctatgcqaaactgtgtcaaaaataccacataatcagggtcagagaatttg  
g<sup>+</sup>  
gcgcgcacattgacc

[0325] 粗体和下划线:外显子1、21-229

[0326] 粗体:剪接边界

[0327] 斜体: 剪接位点

### [0328] 斜体和下划线:内含子1

[0329] SEQ ID NO:10 (NtIPMS1A核苷酸序列)

[0330] ATGTCTTCTCTGTTCAAACCTGCAACTTCTTAGTTGAATAACTCCTCCAATCCAAAATCC  
TCTTCTTCACACCCTTAACCTCTTCCCTCAATCAAACACACTCTGTTCCCATATACAGTTATCCGGTGC  
TCAATTCAAAGCGACCTGAATATACCGAGTAAATCTCGACCACAAATACGTACGCATTTGACACGACTC  
TCCGCGATGGAGAGCAATCCCCGGCGCTACGATGACTACGAAAGAAAAATTAGATGTTGCCGTAAATTAGCGAA  
ACTCGGAGTTGACATAATCGAAGCTGGTTCCAGCTCATCTGAAGCTGATTTGAAGCTGTGAGATTAATAGCA  
GAGGAATTGGTAATAATAGCGACGGTGATTATGTGCCGGTGAAGTGTGCGAAGAGACCTAGGGTTCATACGTTATAGCGACGAGTGAGATACA  
TTGATAAAGCGTGGAGCTGTGAAGTGTGCGAAGAGACCTAGGGTTCATACGTTATAGCGACGAGTGAGATACA  
TATGAAGTATAAGTTGAAGATGAGTAAAGAAGAAGTAGTGGAGAAAGCGCGAGTATGGTTGCTTATGCGAGGAGT  
TTGGGATGTGAGGATGTTGAAGTTAGCCCTGAAGATGCTGGAAGGTCTGAGCGTGAGTTCTTACCATATCCTG  
GAGAAGTTATCAAAGCTGGTGCAACAAACCTAACATACCTGATACTGTTGGATACACTGTGCCACTGAATTGG  
ACAATTAAATTGCtgcacataaaagccaatccccaggaattgaaaatgtgatcattctacacactgccagaatgat  
cttggactttctactgccaacacttagctggagcttgccatggggcaagacaagttagtgcatttcacatggca  
tttgtgaaagagctggaaatgctctggaggagttgtaatggcctaaaatgtcgtagagagcaagttactgg  
tggcctgtatacaggaattaatacacaacatatactcatgtcaagcaagatggtagaggatcacccggcttcat  
gtgcagccacacaaggccatgtggagctaatgctttgctcatgaaagtggcatccatcagGATGGAATGTTAA  
AACACAAAGATAATGAGATTATCTCCTGAAGATATTGGCTAGTCGTGCTAATGAAGCCGGTATTGTCCTT  
GGGAAGCTCAGTGGCGCCATGCATTGAAATCCAAAGCTGAGCTGGATATGACATTGAGGGAAAAGAACTGG  
GGACCTTCTGGCGATTTAAGTCGGTGGCTGAGAAGAAAAAGAAAATTACAGATGATGACATAATAGCACTGATGT

CAGATGAAGTTTCCAGCCTCAAGTTGGCAACTGCAGATGTACAGATTGCTGTGGAAGTCTTGGCCTCT  
ACAGCAACTGTTAAGCTTATTGACAGTGATGGTCAAGAGCATGTTGCTGTTGGAACCGGACCAGTGTGATGC  
AGCTTATAAGGCAGTTGACCTCATTGTAAGGTACCTATAACACTCCTCGAGTATTCCATGAATGCAGTCACAGAAG  
GTATAGATGCCATAGCCTAACCGAGTATTAAATCCGGGGAGGATGACCATGCTATAACCAATGGTCAATTGGA  
CCGACTCATCACCGTATATTAGTGGAACTGGAGCTGATATGGACGTTGTCATCTTAGTGTCCGAGCCTATATTGG  
TGCATTGAACAAAATGTTGAGTTCGGGAGCTGGTTGAGGTACAAGAACCTGAAGGTAGTGTGGTAGTATAA

[0331] SEQ ID NO:11 (NtIPMS1A氨基酸序列)

[0332] MSSLCNSSATSLSCNNSFQSKNPLLHTIFNFFPSIKPHSCFPYTVIRCSIQKRPEYIPSKIISDHKYVRI  
FDTTLRDGEQSPGATMTTKEKLDVARKLAKLGVDIEAGFPASSEADFEAVRLIAEEIGNNSDGDYVPVICGLARCN  
KRDIDKAWEAVKCAKRPRVHTFIATSEIHMKYKLKMSKEEVVEKARSMVAYARSLGCEDVEFSPEDAGRSEREFLYH  
ILGEVIKAGATTLNIPDTVGTVPTEFQQLIADIKANTPGIENVIISTHCQNDLGLSTANTLAGACAGARQVEVTIN  
GIGEPAGNASLEEVVMALKCRGEQVLGGLYTGINTQKILMSSKMVEEYTGLHVQPHKAIVGANAFAHESGIHQDGML  
KHKDTYEIISPEDIGLSRANEAGIVLGKLSGRHALSKMELGYDIEGKELEDLFWRFKSVAEKKKITDDDIIALM  
SDEVFQPQVVWQLADVQIACGSLGLSTATVKLIDSQGEHVACSGVTGPVDAAYKAVDLIVKVPITLLEYSMNAVTE  
GIDAIASTRVLIRGEDDHAITNGSIGPTHHRIFSGTGADMDVVISSVRAYIGALNKMLSGKLVSRYKKPEGSVVV

[0333] SEQ ID NO:12 (NtIPMS1Bv1核苷酸序列) ATGGCGTCTATCACCATAAACCATTCATTTC

CGTAACCCCTAACATCTCATTCCATCCCCAAAATCCTCTCATTCAAACCCAAGCTCTTCAATTCAAACCATCAAT  
CTCCAAATGTTCCCTATTATCCACTGCGCAATCCGCCGTGACCCGAATATACCCGAGCCACATTCCGACCCGA  
ACTACATTGCATCTCGACACCCTCTCCGCACGGCGAACAAATCCCCAGGCGCCACAATGACCACAAAGAAAAAA  
CTCGACGTTGCGCGTCAGTTAGCTAACGCTGGTGTGACATAATTGAAGCCGTTTCTGCTTCTGAAGCTGA  
TCTCGAAGCTGTGAAATTAAAGCGAAGGAAGTTGAAATGGTGTGAATGAAGAGGGACATGTTCCGTAATTGTG  
GACTTGCAGGGTGTAAATAAGAGGGATTGATAAGGCTTGGGAGGCTGTGAAGTATGCGAAAAAACCGAGGATTCA  
ACGTTTATTGCGACTAGTGAGATACATATGAAGTTAACGTTAACGATGAGTAGAGATGAAGTTGTGGAGAAAGCTAG  
GAGTATGGTTGCTTATGCTAGGAGTATTGGTTGTGAGGATGTTGAATTAGCCCAGAAGATGCTGGAAGATCCGATC  
CAGAGTTCCCTATCATATCCTGGAGAGGTCAAAAGCTGGGCAACAACCCCTAACATCCCTGATACTGTTGGA  
TACACTGTTCCAGCGAATTGGAAAATTGATTGCTGATATAAAGGCCAATACCCAGGAATTGGAGATGTGATCAT  
CTCAACACACTGCCAGAACGATCTGGGCTTCACTGCCAACACCTAGCTGGAGCATGCGCAGGTGCAAGACAAG  
TAGAAGTGACCATCAACCGAACCGTGAAAGAGCTGGAAATGCTCTTGGAGGAGGTGTAATGGCTTAAAGT  
CGTGGAGAGCAAGTACTAGGTGGCCTGTATACAGGAATTAAACACATATACTCATGTCAAGCAAGATGGT  
GGAGTACACCGGGCTTCATGTGCAGCCACACAAGGCCATTGTTGGAGCTAATGCGTTGCTCATGAAAGTGGCATCC  
ATCAGGATGGAATGTTAAACACAAAGATAATGAGATTATCTCCTGAAGATATTGGCTTAACCGAGTTAAT  
GAATCTGGCATCGTCCTGGAAACTCAGTGGCGTCATGCTTGCAAGCCAAATGCTCGAGCTGGATACGATAT  
TGAGGGAAAAGAACTTGAGGACCTTTGGCGATTCAAATCTGTTGGCCGAGAAGAAAAGAAAATTACAGATGATG  
ACCTGATAGCATTAAATGTCAGATGAAGTTCCAGCCTCAATTGTTGGCAACTGAAAATGTACAGGTTACATGT  
GGAAGTCTTGGCCTTCTACGGCAACTGTTAACGCTATTGACGCTGATGGTCAAGAGCATGTTCTGTTG  
AACGGGGCCAGTTGATGCGGCTTATAAGGCAGTTGATCTCATTGTAAGGTACCTGAGCAGTCCTGAATATTCT  
TGAATGCAGTCACGGAAAGGTATAGATGCCATAGCTCAACCAGAGTTAACGTTGAGGAGAATGGCATA  
ACCCATGCTTAACTGGAGAGACTGTACACCGTTTTAGTGGAAACCGGAGCAGATATGGATATTGTTATCTCCAG  
TGTCCGAGCCTATATTGGTGCATTGAATAAGATGTTGAGTTGAGAAAGCTGGTATCGAAACACAGCAAACCTGAAG

GCAGTGCAGTCGTATAG

[0334] SEQ ID NO:13 (NtIPMS1Bv1氨基酸序列)

[0335] MASITINHSFSRNPNIKFHPQNPLIQTQALFNFKPSISKCSPIIHCAIRRPEYTPSHIPDPNYIRIFDTTLRDGEQSPGATMTTKEKLDVARQLAKLGVDIIIEAGFPASSEADLEAVKLIAKEVGNGVNEEGHVPVICGLARCNKRDIDKAWEAVKYAKKPRIHTFIATSEIHMKFKLMSRDEVVEKARSMVAYARSIGCEDVEFSPEDAGRSDPEFLYHILGEVIKAGATTNLIPDTVGYTVPSEFGKLIADIKANTPGIGDVIISTHCQNDLGLSTANTLAGACAGARQVEVTINGIGERAGNASLEEVVMALKCRGEQVLGGLYTGINTQHILMSSKMVEEYTGLHVQPHKAIVGANAFAHESGIHQDGMLKHKDTYEIISPEDIGLNRVNESGIVLGKLSGRHALQAKMLELGYDIEGKELEDFWRFKSVAEKKKITDDDLIALMSDEVFQPQFWQLENVQVTCGSLGLSTATVKLIDADGQEHVSCSGTGPVDAAYKAVDLIVKVPVALLEYSLNAVTEGIDAIASTRVLIRGENGHTSTHALTGETVHRSFSGTGADMDIVISSVRAYIGALNKMLSRKLVSKHSKPEGSAVV

[0336] SEQ ID NO:14 (NtIPMS1Bv2核苷酸序列)

[0337] ATGGCGTCTATCACCGCAAACCATACTTTCCGTAACCTAACATCTCATTGCATCCCCAAAATCCTCTCATTCAAACCCAAGCTCTTCAACTCAAATCATCAATCCCCAAATGTTCCCTATTATCTGCTGCGCAATCCGCGTCGACCCACTCTCCGAGCGAACAATCTCCAGGCGCCACAATGACCACAAAAGAAAAACTCGACGTTGCGCGTCAGTTAGCTAAGCTTGGTGTGACATAATTGAAGCCGGTTTCCTGCTTCTGAAGCTGATCTGAAGCTGTGAAATTAAATAGCGAAGGAAGTTGGAAATGGTGTGATGAAGAGGGACATGTTCCGTAATCTGTGGACTTGCAGGGTGAATAAGAGGGATATTGATAAGGCTTGGGAGGCTGTGAAGTATGCGAAAAAACCGAGGATTACATGTTATTGCGACTAGTGAGATAATATGAAGTTAAAGTTGAAGATGAGTAGAGATGAAGTTGGAGAAAGCTAGAAGTATGGTTGCTTATGCTAGGAGTATTGGTTGTGAGATGTTGAATTAGCCTGAAGATGCTGGAAGATCTGATCCTGAGTTCCTCTATCATATCCTGGAGAGGTCATCAAAAGCTGGGCAACAACCTTAACATCCCTGATACTGTTGGATACACTGTTCCCAGTGAATTGGAAAATTGATCGCTGATATAAAGGCCAATACCCAGGAATTGGAGATGTGATCATCTCAACGCAGTGCAGAACGATCTGGCTTCTAC TGCCAACACCTTAGCTGGAGCATGTGCAGGTGCAAGACAAGTAGAAGTGACCATCAATGGAATCGGTGAAAGAGCTGAAATGCTTGGAGGAGGTTGTAATGGCCTTAAATGTCGTGGAGAGCAAGTACTAGGTGGCCTGTATACAGGAATTAATACACAACATATACTCATGTCAAGCAAGATGGTAGAGGAGTACACCGGGCTTACATGTGCAGGCCACACAAGGCATTGTTGGAGCTAATGCTTTGCTCATGAAAGTGGCATCCATCAGGATGGAATGTTAAACACAAAGATACTATAGAGATTATATCTCCTGAAGATATTGGCCTTAACCGAGCTAATGAATCTGGTATCGTCCTCGGGAAACTCAGTGGCGCTCATGCTTGGCAAGCCAAATGCTCGAGCTTGGATACGATATTGAGGGAAAAGAACTTGAGGACCTTTGGCGATTCAAATCTGTGGCTGAGAAGAAAAAGAAAATTACAGATGATGACCTGATAGCATTGATGTCAGATGAAGTTCCAGCCTCAATTGTTGGCAACTCGAAAATGTACAGGTTACATGTGGAGTCTTGCCTTACGGCAACTGTTAAGCTATTGACGCTGATGGTCAAGAGCATGTTCTGTTGGAACGGGCCAGTTGATGCGGCTTACAAGGCAGTTGATCTCATTGTAATGCAGTCACGGAAAGGTATAGATGCCATAGCTTCAACCAGAGTTAATCGTGGGAGAATGGACATACATCAACCCATGTTAACTGGAGAGACTGTACACCGTTCTTTAGTGGAACCGGAGCAGATATGGATATTGTTATCTCTAGTGTCCGAGCCTATATTGGAGCATTGAATAAGATGCTGAGTTCAGAAAAGCTGGTGTGAAACACACAGCAGACCTGAAGGCAGTGCAGTCGTATAG

[0338] SEQ ID NO:15 (NtIPMS1Bv2氨基酸序列) MASITANHTFSRNPNIKSFHPQNPLIQTQALFNFKSIPKCSPIICCAIRRDPYTPSHIPDPKYIRIFDTTLRDGEQSPGATMTTKEKLDVARQLAKLGVDIIIEAGFPASEADLEAVKLIAKEVGNGVYEEGHVPVICGLARCNKRDIDKAWEAVKYAKKPRIHTFIATSEIHMKFKLMSRDEVVEKARSMVAYARSIGCEDVEFSPEDAGRSDPEFLYHILGEVIKAGATTNLIPDTVGYTVPSEFGKLIADIKANTPGIG

DVIISTHCQNDLGLSTANTLAGACAGARQVEVTINGIGERAGNASLEEVVMALKCRGEVQLGGLYTGINTQHILMSS  
 KMVEEYTGLHVQPHKAIVGANAFAHESGIHQDGMLKHKDTYEIISPEDIGLNRANESGIVLGKLSGRHALQAKMEL  
 GYDIEGKELEDLFWRFKSVAEKKKITDDDLIALMSDEVFQPQFWQLENVQVTCGSLGLSTATVKLIDADGQEHSV  
 CSVGTGPVDAAYKAVDLIVKVPVALLEYSLNAVTEGIDAIASTRVLIRGENGHTSTHALTGETVHRSFSGTGADM  
 VISSVRAYIGALNKMMLSFRKLVSKHRSPEGSAVV

[0339] 表1

[0340] 蔗糖酯的结构。蔗糖酯的通用结构在图5中显示，并描述了R1、R2、R3、R4

[0341] 和R5基的各种组合。

R5	R3	R1	R2	R4	名称
-	C2	C4i	C4i	C4i	C2C12
-	C2	C4i	C4i	C5ai	C2C13
-	C2	C4i	C4i	C5i	C2C13
-	C2	C4i	C4i	C6ai	C2C14
-	C2	C4i	C4i	C6i	C2C14
-	C2	C4i	C5ai	C4i	C2C13
-	C2	C4i	C5ai	C5ai	C2C14
-	C2	C4i	C5ai	C5i	C2C14
-	C2	C4i	C5ai	C6ai	C2C15
-	C2	C4i	C5ai	C6i	C2C15
-	C2	C4i	C5i	C4i	C2C13
-	C2	C4i	C5i	C5ai	C2C14
-	C2	C4i	C5i	C5i	C2C14
-	C2	C4i	C5i	C6ai	C2C15
-	C2	C4i	C5i	C6i	C2C15
-	C2	C4i	C6ai	C4i	C2C14
-	C2	C4i	C6ai	C5ai	C2C15
-	C2	C4i	C6ai	C5i	C2C15
-	C2	C4i	C6ai	C6ai	C2C16
-	C2	C4i	C6ai	C6i	C2C16
-	C2	C4i	C6i	C6ai	C2C16
-	-	C4i	C4i	C4i	C12
[0342]	-	C4i	C4i	C5ai	C13
-	-	C4i	C4i	C5i	C13
-	-	C4i	C4i	C6ai	C14
-	-	C4i	C4i	C6i	C14
-	-	C4i	C5ai	C4i	C13
-	-	C4i	C5ai	C5ai	C14
-	-	C4i	C5ai	C5i	C14
-	-	C4i	C5ai	C6ai	C15
-	-	C4i	C5ai	C6i	C15
-	-	C4i	C5i	C4i	C13
-	-	C4i	C5i	C5ai	C14
-	-	C4i	C5i	C5i	C14
-	-	C4i	C5i	C6ai	C15
-	-	C4i	C5i	C6i	C15
-	-	C4i	C5i	C6i	C15
-	-	C4i	C6ai	C4i	C14
-	-	C4i	C6ai	C5ai	C15
-	-	C4i	C6ai	C5i	C15
-	-	C4i	C6ai	C6ai	C16
-	-	C4i	C6ai	C6i	C16
-	-	C4i	C6i	C4i	C14
-	-	C4i	C6i	C5ai	C15
-	-	C4i	C6i	C5i	C15
-	-	C4i	C6i	C6ai	C16
-	-	C4i	C6i	C6i	C16
C2	-	C4i	C4i	C4i	C12C2
C2	-	C4i	C4i	C5ai	C13C2
C2	-	C4i	C4i	C5i	C13C2
C2	-	C4i	C4i	C6ai	C14C2
C2	-	C4i	C4i	C6i	C14C2

C2 - C4i C5ai C4i C13C2  
 C2 - C4i C5ai C5ai C14C2  
 C2 - C4i C5ai C5i C14C2  
 C2 - C4i C5ai C6ai C15C2  
 C2 - C4i C5ai C6i C15C2  
 C2 - C4i C5i C4i C13C2  
 C2 - C4i C5i C5ai C14C2  
 C2 - C4i C5i C5i C14C2  
 C2 - C4i C5i C6ai C15C2  
 C2 - C4i C6ai C4i C14C2  
 C2 - C4i C6ai C5ai C15C2  
 C2 - C4i C6ai C5i C15C2  
 C2 - C4i C6ai C6ai C16C2  
 C2 - C4i C6ai C6i C16C2  
 C2 - C4i C6i C4i C14C2  
 C2 - C4i C6i C5ai C15C2  
 C2 - C4i C6i C5i C15C2  
 C2 - C4i C6i C6ai C16C2  
 C2 - C4i C6i C6i C16C2  
 C2 C2 C4i C4i C4i C2C12C2  
 C2 C2 C4i C4i C5ai C2C13C2  
 C2 C2 C4i C4i C5i C2C13C2  
 C2 C2 C4i C4i C6ai C2C14C2  
 C2 C2 C4i C4i C6i C2C14C2  
 C2 C2 C4i C5ai C4i C2C13C2  
 C2 C2 C4i C5ai C5ai C2C14C2  
 C2 C2 C4i C5ai C5i C2C14C2  
 C2 C2 C4i C5ai C6ai C2C15C2  
 C2 C2 C4i C5ai C6i C2C15C2  
 C2 C2 C4i C5i C4i C2C13C2  
 C2 C2 C4i C5i C5ai C2C14C2  
 C2 C2 C4i C5i C5i C2C14C2  
 C2 C2 C4i C5i C6ai C2C15C2  
 C2 C2 C4i C5i C6i C2C15C2  
 C2 C2 C4i C6ai C4i C2C14C2  
 C2 C2 C4i C6ai C5ai C2C15C2  
 C2 C2 C4i C6ai C5i C2C15C2  
 C2 C2 C4i C6ai C6ai C2C16C2  
 C2 C2 C4i C6ai C6i C2C16C2  
 C2 C2 C4i C6ai C6i C2C14C2  
 C2 C2 C4i C6i C4i C2C14C2  
 C2 C2 C4i C6i C5ai C2C15C2  
 C2 C2 C4i C6i C5i C2C15C2  
 C2 C2 C4i C6i C6ai C2C16C2  
 C2 C2 C4i C6i C6i C2C16C2  
 - C2 C5ai C4i C4i C2C13  
 - C2 C5ai C4i C5ai C2C14  
 - C2 C5ai C4i C5i C2C14  
 - C2 C5ai C4i C6ai C2C15  
 - C2 C5ai C4i C6i C2C15  
 - C2 C5ai C5ai C4i C2C14  
 - C2 C5ai C5ai C5ai C2C15  
 - C2 C5ai C5ai C5i C2C15  
 - C2 C5ai C5ai C6ai C2C16  
 - C2 C5ai C5ai C6i C2C16  
 - C2 C5ai C5i C4i C2C14  
 - C2 C5ai C5i C5ai C2C15  
 - C2 C5ai C5i C5i C2C15  
 - C2 C5ai C5i C6ai C2C16  
 - C2 C5ai C5i C6i C2C16  
 - C2 C5ai C6ai C4i C2C15  
 - C2 C5ai C6ai C5ai C2C16  
 - C2 C5ai C6ai C5i C2C16  
 - C2 C5ai C6ai C6ai C2C17  
 - C2 C5ai C6ai C6i C2C17  
 - C2 C5ai C6ai C4i C2C15  
 - C2 C5ai C6i C5ai C2C16  
 - C2 C5ai C6i C5i C2C16  
 - C2 C5ai C6i C6ai C2C17  
 - C2 C5ai C6i C6i C2C17  
 - - C5ai C4i C4i C13  
 - - C5ai C4i C5ai C14  
 - - C5ai C4i C5i C14  
 - - C5ai C4i C6ai C15  
 - - C5ai C4i C6i C15  
 - - C5ai C5ai C4i C14

[0344]

- - C5ai C5ai C5ai C15  
 - - C5ai C5ai C5i C15  
 - - C5ai C5ai C6ai C16  
 - - C5ai C5ai C6i C16  
 - - C5ai C5i C4i C14  
 - - C5ai C5i C5ai C15  
 - - C5ai C5i C5i C15  
 - - C5ai C5i C6ai C16  
 - - C5ai C5i C6i C16  
 - - C5ai C6ai C4i C15  
 - - C5ai C6ai C5ai C16  
 - - C5ai C6ai C5i C16  
 - - C5ai C6ai C6ai C17  
 - - C5ai C6ai C6i C17  
 - - C5ai C6i C4i C15  
 - - C5ai C6i C5ai C16  
 - - C5ai C6i C5i C16  
 - - C5ai C6i C6ai C17  
 - - C5ai C6i C6i C17  
 C2 - C5ai C4i C4i C13C2  
 C2 - C5ai C4i C5ai C14C2  
 C2 - C5ai C4i C5i C14C2  
 C2 - C5ai C4i C6ai C15C2  
 C2 - C5ai C4i C6i C15C2  
 C2 - C5ai C5ai C4i C14C2  
 C2 - C5ai C5ai C5ai C15C2  
 C2 - C5ai C5ai C5i C15C2  
 C2 - C5ai C5ai C6ai C16C2  
 C2 - C5ai C5ai C6i C16C2  
 C2 - C5ai C5i C4i C14C2  
 C2 - C5ai C5i C5ai C15C2  
 C2 - C5ai C5i C5i C15C2  
 C2 - C5ai C5i C6ai C16C2  
 C2 - C5ai C5i C6i C16C2  
 C2 - C5ai C6ai C4i C15C2  
 C2 - C5ai C6ai C5ai C16C2  
 C2 - C5ai C6ai C5i C16C2  
 C2 - C5ai C6ai C6ai C17C2  
 C2 - C5ai C6ai C6i C17C2  
 C2 - C5ai C6i C4i C15C2  
 C2 - C5ai C6i C5ai C16C2  
 C2 - C5ai C6i C5i C16C2  
 C2 - C5ai C6i C17C2  
 C2 - C5ai C6i C6i C17C2  
 C2 C2 C5ai C4i C4i C2C13C2  
 C2 C2 C5ai C4i C5ai C2C14C2  
 C2 C2 C5ai C4i C5i C2C14C2  
 C2 C2 C5ai C4i C6ai C2C15C2  
 C2 C2 C5ai C4i C6i C2C15C2  
 C2 C2 C5ai C5ai C4i C2C14C2  
 C2 C2 C5ai C5ai C5ai C2C15C2  
 C2 C2 C5ai C5ai C5i C2C15C2  
 C2 C2 C5ai C5ai C6ai C2C16C2  
 C2 C2 C5ai C5ai C6i C2C16C2  
 C2 C2 C5ai C5i C4i C2C14C2  
 C2 C2 C5ai C5i C5ai C2C15C2  
 C2 C2 C5ai C5i C5i C2C15C2  
 C2 C2 C5ai C5i C6ai C2C16C2  
 C2 C2 C5ai C5i C6i C2C16C2  
 C2 C2 C5ai C5i C6i C2C17C2  
 C2 C2 C5ai C6ai C4i C2C15C2  
 C2 C2 C5ai C6ai C5ai C2C16C2  
 C2 C2 C5ai C6i C5i C2C16C2  
 C2 C2 C5ai C6i C6ai C2C17C2  
 C2 C2 C5ai C6i C6i C2C17C2  
 - C2 C5i C4i C4i C2C13  
 - C2 C5i C4i C5ai C2C14  
 - C2 C5i C4i C5i C2C14  
 - C2 C5i C4i C6ai C2C15  
 - C2 C5i C4i C6i C2C15  
 - C2 C5i C5ai C4i C2C14  
 - C2 C5i C5ai C5ai C2C15

- C2 C5i C5ai C5i C2C15  
 - C2 C5i C5ai C6ai C2C16  
 - C2 C5i C5ai C6i C2C16  
 - C2 C5i C5i C4i C2C14  
 - C2 C5i C5i C5ai C2C15  
 - C2 C5i C5i C5i C2C15  
 - C2 C5i C5i C6ai C2C16  
 - C2 C5i C5i C6i C2C16  
 - C2 C5i C6ai C4i C2C15  
 - C2 C5i C6ai C5ai C2C16  
 - C2 C5i C6ai C5i C2C16  
 - C2 C5i C6ai C6ai C2C16  
 - C2 C5i C6ai C6i C2C17  
 - C2 C5i C6ai C6i C2C17  
 - C2 C5i C6i C4i C2C15  
 - C2 C5i C6i C5ai C2C16  
 - C2 C5i C6i C5i C2C16  
 - C2 C5i C6i C6ai C2C17  
 - C2 C5i C6i C6i C2C17  
 - - C5i C4i C4i C13  
 - - C5i C4i C5ai C14  
 - - C5i C4i C5i C14  
 - - C5i C4i C6ai C15  
 - - C5i C4i C6i C15  
 - - C5i C5ai C4i C14  
 - - C5i C5ai C5ai C15  
 - - C5i C5ai C5i C15  
 - - C5i C5ai C6ai C16  
 - - C5i C5ai C6i C16  
 - - C5i C5i C4i C14  
 - - C5i C5i C5ai C15  
 - - C5i C5i C5i C15  
 - - C5i C5i C6ai C16  
 - - C5i C5i C6i C16  
 - - C5i C6ai C4i C15  
 - - C5i C6ai C5ai C16  
 - - C5i C6ai C5i C16  
 - - C5i C6ai C6ai C17  
 - - C5i C6ai C6i C17  
 [0345] - - C5i C6i C4i C15  
 - - C5i C6i C5ai C16  
 - - C5i C6i C5i C16  
 - - C5i C6i C6ai C17  
 - - C5i C6i C6i C17  
 C2 - C5i C4i C4i C13C2  
 C2 - C5i C4i C5ai C14C2  
 C2 - C5i C4i C5i C14C2  
 C2 - C5i C4i C6ai C15C2  
 C2 - C5i C4i C6i C15C2  
 C2 - C5i C5ai C4i C14C2  
 C2 - C5i C5ai C5ai C15C2  
 C2 - C5i C5ai C5i C15C2  
 C2 - C5i C5ai C6ai C16C2  
 C2 - C5i C5ai C6i C16C2  
 C2 - C5i C5i C4i C14C2  
 C2 - C5i C5i C5ai C15C2  
 C2 - C5i C5i C5i C15C2  
 C2 - C5i C5i C6ai C16C2  
 C2 - C5i C5i C6i C16C2  
 C2 - C5i C6ai C4i C15C2  
 C2 - C5i C6ai C5ai C16C2  
 C2 - C5i C6ai C5i C16C2  
 C2 - C5i C6ai C6ai C17C2  
 C2 - C5i C6ai C6i C17C2  
 C2 - C5i C6i C4i C15C2  
 C2 - C5i C6i C5ai C16C2  
 C2 - C5i C6i C5i C16C2  
 C2 - C5i C6i C6ai C17C2  
 C2 - C5i C6i C6i C17C2  
 C2 C2 C5i C4i C4i C2C13C2  
 C2 C2 C5i C4i C5ai C2C14C2  
 C2 C2 C5i C4i C5i C2C14C2  
 C2 C2 C5i C4i C6ai C2C15C2  
 C2 C2 C5i C4i C6i C2C15C2  
 C2 C2 C5i C5ai C4i C2C14C2  
 C2 C2 C5i C5ai C5ai C2C15C2  
 C2 C2 C5i C5ai C5i C2C15C2

C2 C2 C5i C5ai C6ai C2C16C2  
 C2 C2 C5i C5ai C6i C2C16C2  
 C2 C2 C5i C5i C4i C2C14C2  
 C2 C2 C5i C5i C5ai C2C15C2  
 C2 C2 C5i C5i C5i C2C15C2  
 C2 C2 C5i C5i C6ai C2C16C2  
 C2 C2 C5i C5i C6i C2C16C2  
 C2 C2 C5i C6ai C4i C2C15C2  
 C2 C2 C5i C6ai C5ai C2C16C2  
 C2 C2 C5i C6ai C5i C2C16C2  
 C2 C2 C5i C6ai C6ai C2C17C2  
 C2 C2 C5i C6ai C6i C2C17C2  
 C2 C2 C5i C6i C4i C2C15C2  
 C2 C2 C5i C6i C5ai C2C16C2  
 C2 C2 C5i C6i C5i C2C16C2  
 C2 C2 C5i C6i C6ai C2C17C2  
 C2 C2 C5i C6i C6i C2C17C2  
 - C2 C6ai C4i C4i C2C14  
 - C2 C6ai C4i C5ai C2C15  
 - C2 C6ai C4i C5i C2C15  
 - C2 C6ai C4i C6ai C2C16  
 - C2 C6ai C4i C6i C2C16  
 - C2 C6ai C5ai C4i C2C15  
 - C2 C6ai C5ai C5ai C2C16  
 - C2 C6ai C5ai C5i C2C16  
 - C2 C6ai C5ai C6ai C2C17  
 - C2 C6ai C5ai C6i C2C17  
 - C2 C6ai C5i C4i C2C15  
 - C2 C6ai C5i C5ai C2C16  
 - C2 C6ai C5i C5i C2C16  
 - C2 C6ai C5i C6ai C2C17  
 - C2 C6ai C5i C6i C2C17  
 - C2 C6ai C6ai C4i C2C16  
 - C2 C6ai C6ai C5ai C2C17  
 - C2 C6ai C6ai C5i C2C17  
 - C2 C6ai C6ai C6ai C2C18  
 - C2 C6ai C6ai C6i C2C18  
 - C2 C6ai C6i C4i C2C16  
 - C2 C6ai C6i C5ai C2C17  
 - C2 C6ai C6i C5i C2C17  
 - C2 C6ai C6i C6ai C2C18  
 - C2 C6ai C6i C6i C2C18  
 - - C6ai C4i C4i C14  
 - - C6ai C4i C5ai C15  
 - - C6ai C4i C5i C15  
 - - C6ai C4i C6ai C16  
 - - C6ai C4i C6i C16  
 - - C6ai C5ai C4i C15  
 - - C6ai C5ai C5ai C16  
 - - C6ai C5ai C5i C16  
 - - C6ai C5ai C6ai C17  
 - - C6ai C5ai C6i C17  
 - - C6ai C5i C4i C15  
 - - C6ai C5i C5ai C16  
 - - C6ai C5i C5i C16  
 - - C6ai C5i C6ai C17  
 - - C6ai C5i C6i C17  
 - - C6ai C6ai C4i C16  
 - - C6ai C6ai C5ai C17  
 - - C6ai C6ai C5i C17  
 - - C6ai C6ai C6ai C18  
 - - C6ai C6ai C6i C18  
 C2 - C6ai C4i C4i C14C2  
 C2 - C6ai C4i C5ai C15C2  
 C2 - C6ai C4i C5i C15C2  
 C2 - C6ai C4i C6ai C16C2  
 C2 - C6ai C4i C6i C16C2  
 C2 - C6ai C5ai C4i C15C2  
 C2 - C6ai C5ai C5ai C16C2  
 C2 - C6ai C5ai C5i C16C2  
 C2 - C6ai C5ai C6ai C17C2

[0346]

[0347]

C2 - C6ai C5ai C6i C17C2  
C2 - C6ai C5i C4i C15C2  
C2 - C6ai C5i C5ai C16C2  
C2 - C6ai C5i C5i C16C2  
C2 - C6ai C5i C6ai C17C2  
C2 - C6ai C5i C6i C17C2  
C2 - C6ai C6ai C4i C16C2  
C2 - C6ai C6ai C5ai C17C2  
C2 - C6ai C6ai C5i C17C2  
C2 - C6ai C6ai C6ai C18C2  
C2 - C6ai C6ai C6i C18C2  
C2 - C6ai C6i C4i C16C2  
C2 - C6ai C6i C5ai C17C2  
C2 - C6ai C6i C5i C17C2  
C2 - C6ai C6i C6ai C18C2  
C2 - C6ai C6i C6i C18C2  
C2 C2 C6ai C4i C4i C2C14C2  
C2 C2 C6ai C4i C5ai C2C15C2  
C2 C2 C6ai C4i C5i C2C15C2  
C2 C2 C6ai C4i C6ai C2C16C2  
C2 C2 C6ai C4i C6i C2C16C2  
C2 C2 C6ai C5ai C4i C2C15C2  
C2 C2 C6ai C5ai C5ai C2C16C2  
C2 C2 C6ai C5ai C6ai C2C17C2  
C2 C2 C6ai C5ai C6i C2C17C2  
C2 C2 C6ai C5i C4i C2C15C2  
C2 C2 C6ai C5i C5ai C2C16C2  
C2 C2 C6ai C5i C5i C2C16C2  
C2 C2 C6ai C5i C6ai C2C17C2  
C2 C2 C6ai C5i C6i C2C17C2  
C2 C2 C6ai C6ai C4i C2C16C2  
C2 C2 C6ai C6i C5ai C2C17C2  
C2 C2 C6ai C6i C5i C2C15C2  
C2 C2 C6ai C6i C6ai C2C18C2  
C2 C2 C6ai C6i C6i C2C18C2  
C2 C2 C6ai C6i C4i C2C16C2  
C2 C2 C6ai C6i C5ai C2C17C2  
C2 C2 C6ai C6i C5i C2C17C2  
C2 C2 C6ai C6ai C6ai C2C18C2  
C2 C2 C6ai C6ai C6i C2C18C2  
C2 C2 C6ai C6i C4i C2C16C2  
C2 C2 C6ai C6i C5ai C2C17C2  
C2 C2 C6ai C6i C5i C2C15C2  
C2 C2 C6ai C6i C6ai C2C16C2  
C2 C2 C6ai C6i C6i C2C17C2  
C2 C2 C6ai C6i C5ai C2C15C2  
C2 C2 C6ai C6i C5i C2C16C2  
C2 C2 C6ai C6i C6ai C2C16C2  
C2 C2 C6ai C6i C6i C2C17C2  
C2 C2 C6ai C6i C5ai C2C15C2  
C2 C2 C6ai C6i C5i C2C16C2  
C2 C2 C6ai C6i C6ai C2C16C2  
C2 C2 C6ai C6i C6i C2C17C2  
C2 C2 C6ai C6i C5ai C2C15C2  
C2 C2 C6ai C6i C5i C2C16C2  
C2 C2 C6ai C6i C6ai C2C16C2  
C2 C2 C6ai C6i C6i C2C17C2  
C2 C2 C6ai C6i C4i C2C16C2  
C2 C2 C6ai C6i C5ai C2C17C2  
C2 C2 C6ai C6i C5i C2C17C2  
C2 C2 C6ai C6ai C6ai C2C18C2  
C2 C2 C6ai C6ai C6i C2C18C2  
C2 C2 C6ai C6i C4i C2C16C2  
C2 C2 C6ai C6i C5ai C2C17C2  
C2 C2 C6ai C6i C5i C2C17C2  
C2 C2 C6ai C6ai C6ai C2C18C2  
C2 C2 C6ai C6ai C6i C2C18C2  
C2 C2 C6ai C6i C4i C2C16C2  
C2 C2 C6ai C6i C5ai C2C17C2  
C2 C2 C6ai C6i C5i C2C17C2  
C2 C2 C6ai C6i C6ai C2C18C2  
C2 C2 C6ai C6i C6i C2C18C2  
C2 C2 C6ai C4i C4i C14  
C2 C2 C6ai C4i C5ai C15  
C2 C2 C6i C4i C5i C15  
C2 C2 C6i C4i C6ai C16  
C2 C2 C6i C4i C6i C16  
C2 C2 C6i C5ai C4i C15  
C2 C2 C6i C5ai C5ai C16  
C2 C2 C6i C5ai C5i C16  
C2 C2 C6i C5ai C6ai C17  
C2 C2 C6i C5ai C6i C17

- - C6i C5i C4i C15  
 - - C6i C5i C5ai C16  
 - - C6i C5i C5i C16  
 - - C6i C5i C6ai C17  
 - - C6i C5i C6i C17  
 - - C6i C6ai C4i C16  
 - - C6i C6ai C5ai C17  
 - - C6i C6ai C5i C17  
 - - C6i C6ai C6ai C18  
 - - C6i C6ai C6i C18  
 - - C6i C6i C4i C16  
 - - C6i C6i C5ai C17  
 - - C6i C6i C5i C17  
 - - C6i C6i C6ai C18  
 - - C6i C6i C6i C18  
 C2 - C6i C4i C4i C14C2  
 C2 - C6i C4i C5ai C15C2  
 C2 - C6i C4i C5i C15C2  
 C2 - C6i C4i C6ai C16C2  
 C2 - C6i C4i C6i C16C2  
 C2 - C6i C5ai C4i C15C2  
 C2 - C6i C5ai C5ai C16C2  
 C2 - C6i C5ai C5i C16C2  
 C2 - C6i C5ai C6ai C17C2  
 C2 - C6i C5ai C6i C17C2  
 C2 - C6i C5i C4i C15C2  
 C2 - C6i C5i C5ai C16C2  
 C2 - C6i C5i C6ai C17C2  
 C2 - C6i C5i C6i C17C2  
 C2 - C6i C6ai C4i C16C2  
 C2 - C6i C6ai C5ai C17C2  
 C2 - C6i C6ai C5i C17C2  
 C2 - C6i C6ai C6ai C18C2  
 C2 - C6i C6ai C6ai C18C2  
 C2 - C6i C6ai C6i C18C2  
 C2 - C6i C6ai C4i C2C14C2  
 C2 - C6i C4i C5ai C2C15C2  
 C2 - C6i C4i C5i C2C15C2  
 C2 - C6i C4i C6ai C2C16C2  
 C2 - C6i C4i C6i C2C16C2  
 C2 - C6i C5ai C4i C2C15C2  
 C2 - C6i C5ai C5ai C2C16C2  
 C2 - C6i C5ai C5i C2C16C2  
 C2 - C6i C5ai C6ai C2C17C2  
 C2 - C6i C5ai C6i C2C17C2  
 C2 - C6i C5i C4i C2C15C2  
 C2 - C6i C5i C5ai C2C16C2  
 C2 - C6i C5i C5i C2C16C2  
 C2 - C6i C5i C6ai C2C17C2  
 C2 - C6i C5i C6i C2C17C2  
 C2 - C6i C6ai C4i C2C16C2  
 C2 - C6i C6ai C5ai C2C17C2  
 C2 - C6i C6ai C5i C2C17C2  
 C2 - C6i C6ai C6ai C2C18C2  
 C2 - C6i C6ai C6i C2C18C2  
 C2 - C6i C6i C4i C2C16C2  
 C2 - C6i C6i C5ai C2C17C2  
 C2 - C6i C6i C5i C2C17C2  
 C2 - C6i C6i C6ai C2C18C2  
 C2 - C6i C6i C6i C2C18C2

[0348] C2表示乙酰基;C3表示丁酰基;C4表示丙酰基;C4i表示异丁酰基;C5表示戊酰基(戊酰基);C5ai表示2-甲基-丁酰基;C5i表示异戊酰基;C5:1i表示异戊酰基;C5:1表示戊酰基;C6表示己酰基;C6 (2) 表示2-甲基戊酰基;C6ai表示β-甲基戊酰基;C6i表示4-甲基戊酰基;-通常是氢。

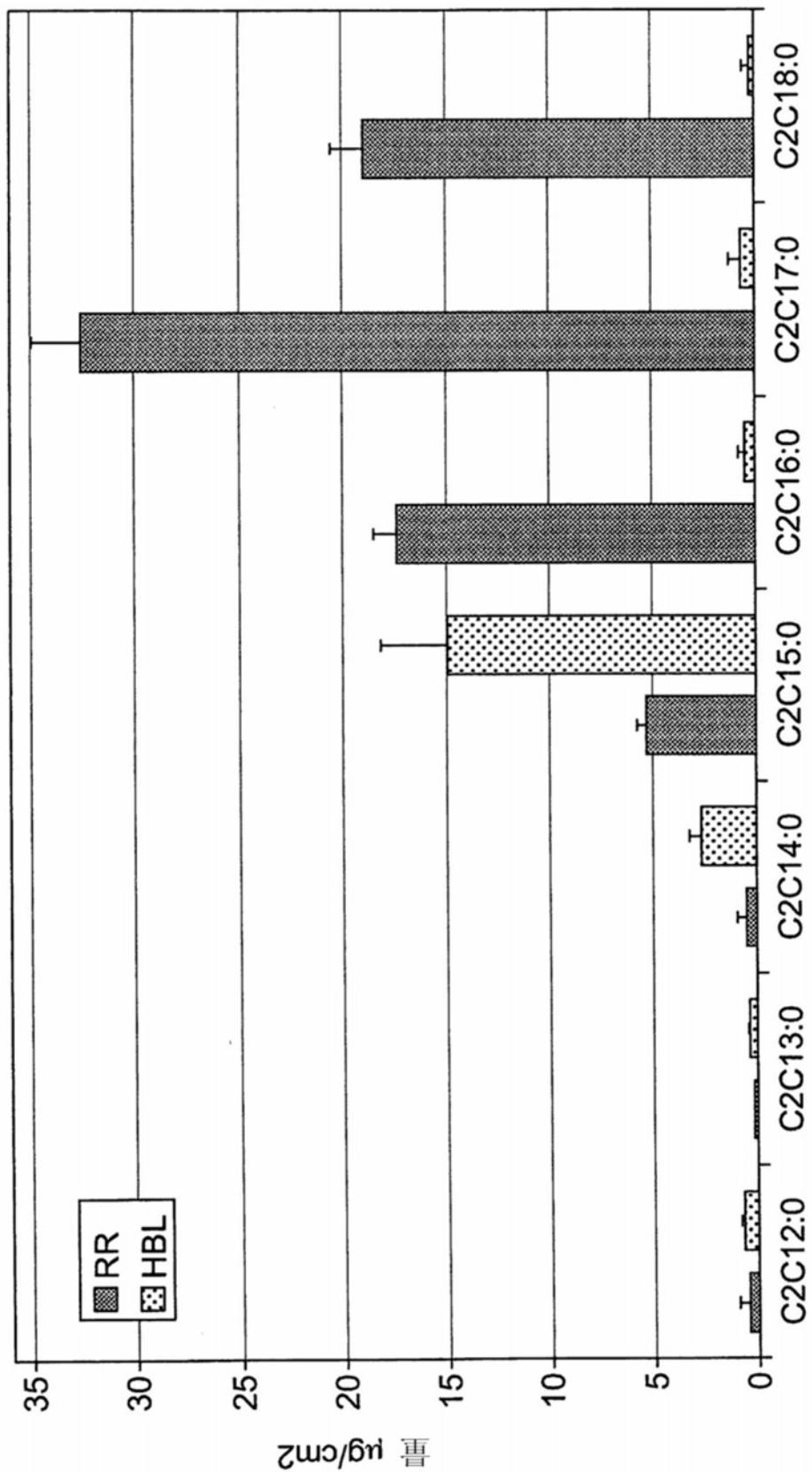


图1

MW(Da)	描述	酰基组合物	在bmwse烟草品种中存在		在BMVSE烟草品种中存在	
			在bmwse烟草品种中存在	在BMVSE烟草品种中存在	在bmwse烟草品种中存在	在BMVSE烟草品种中存在
594	C15:0	C5C5C5 C6C5C4	+	+	+	+
	C2C12:0	C2C3C4C5 C2C2C5C5	+	+	+	+
608	C16:0	C6C5C5	+	+	+	+
	C2C13:0	C2C4C4C5 C2C3C4C6	+	+	+	+
622	C17:0	C2C3C5C5	+	+	+	+
	C2C14:0	C6C6C5 C2C4C4C6 C2C4C5C5	+	+	+	+
636	C18:0	C2C3C5C6	+	+	+	+
	C2C15:0	C6C6C6 C2C4C5C6 C2C5C5C5	+	+	+	+
650	C2C16:0	C2C3C6C6	+	+	+	+
		C2C5C5C6 C2C4C6C6	+	+	+	+
664	C2C17:0	C2C5C6C6	+	+	+	+
678	C2C18:0	C2C6C6C6	+	+	+	+
692	C2C19:0	C2C6C6C7	+	+	+	+

图2

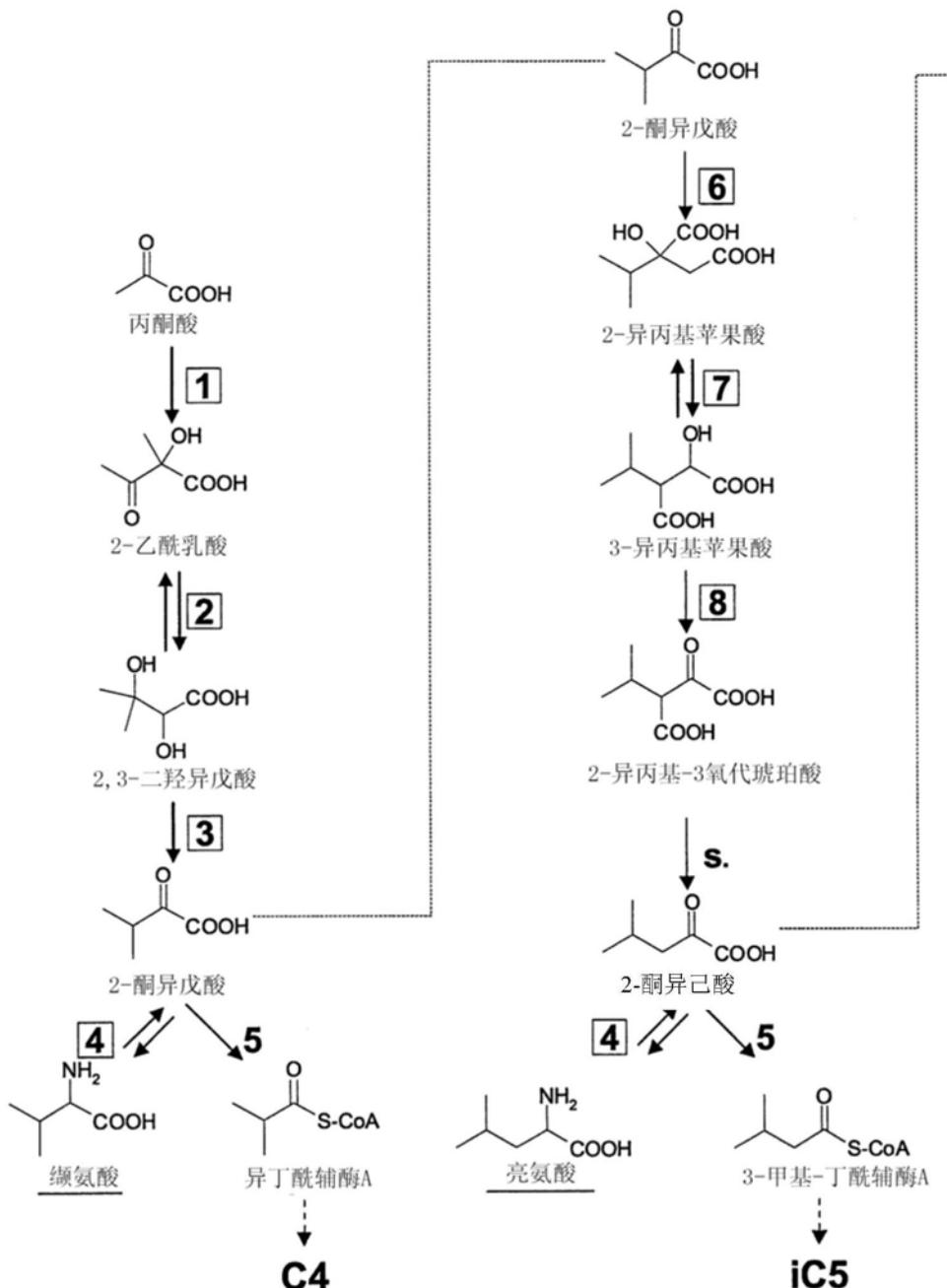
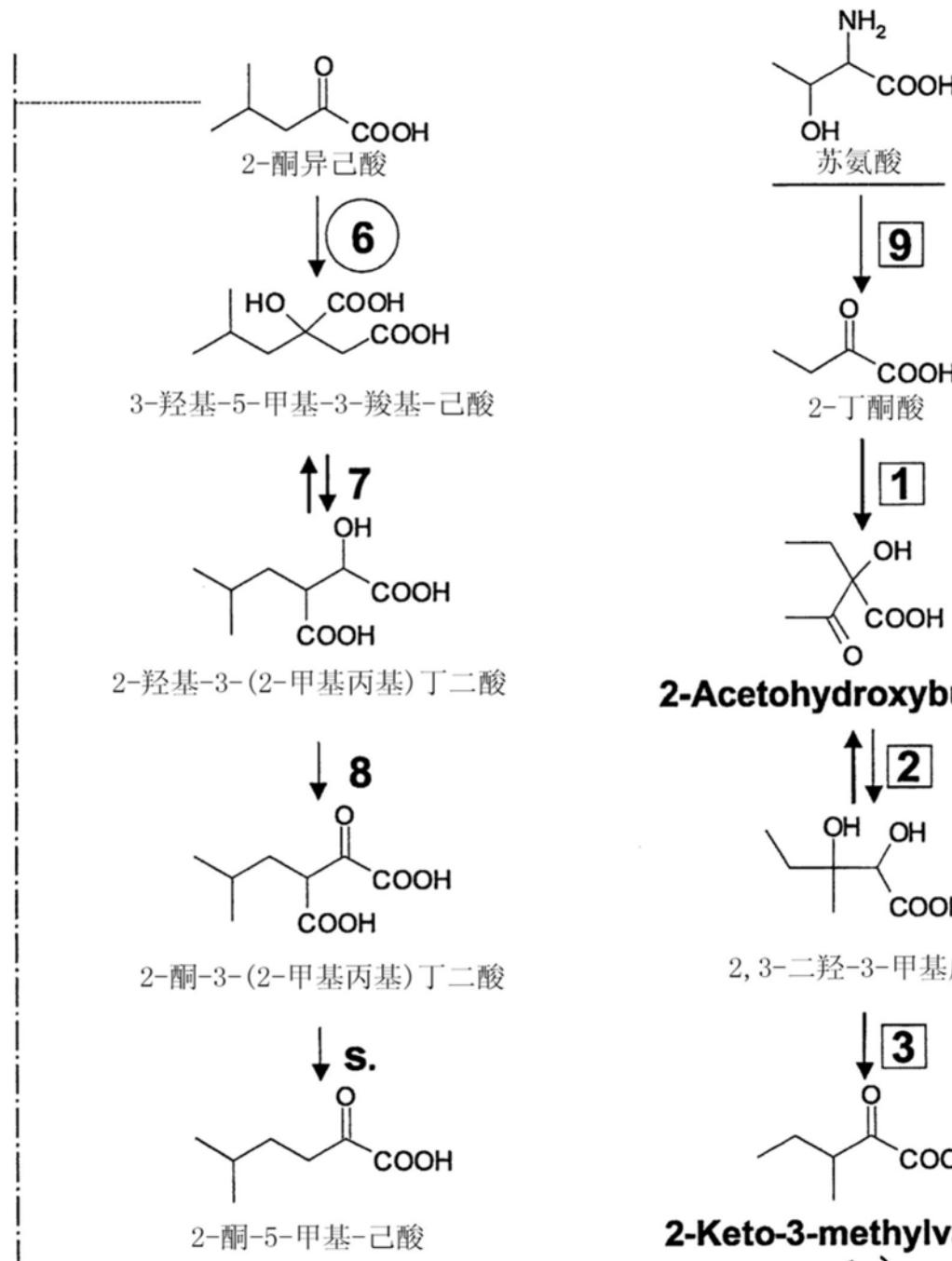


图3



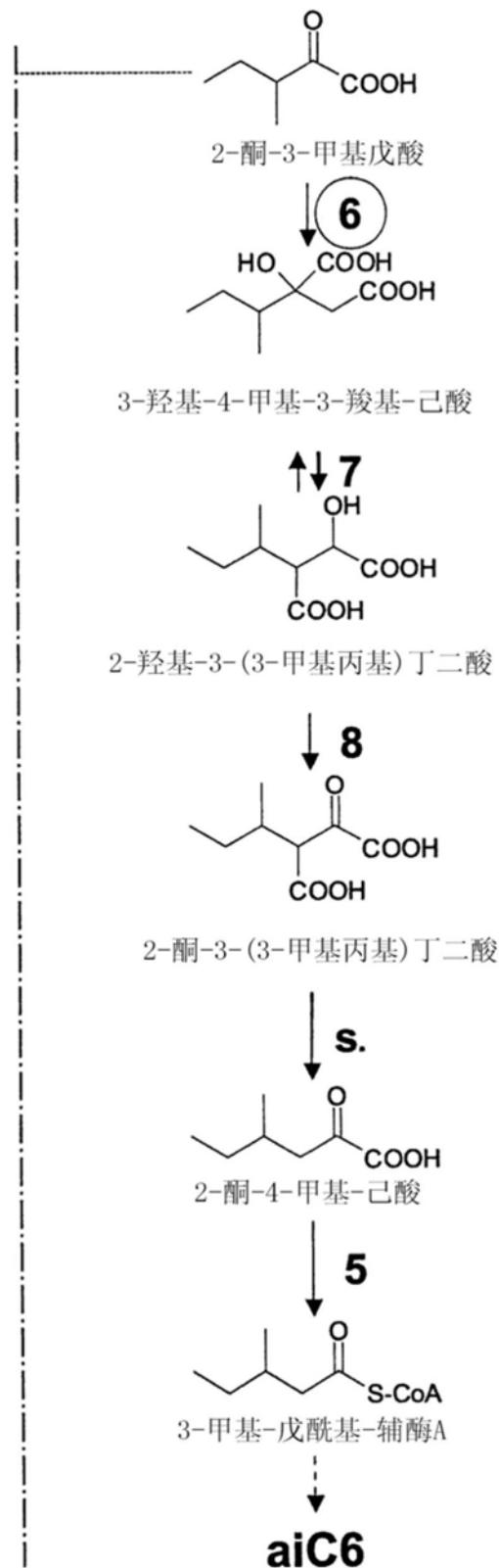


图3续

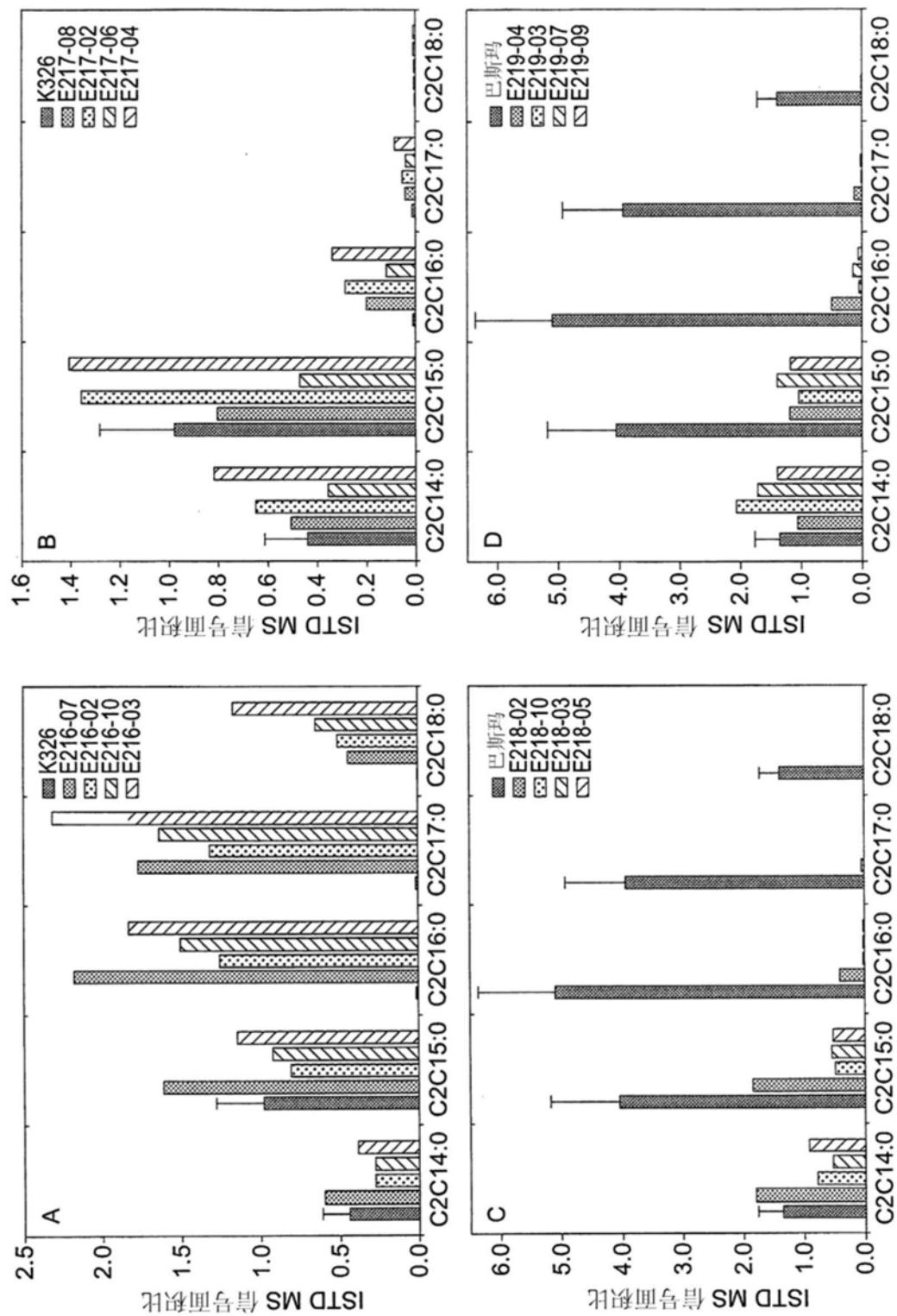


图4

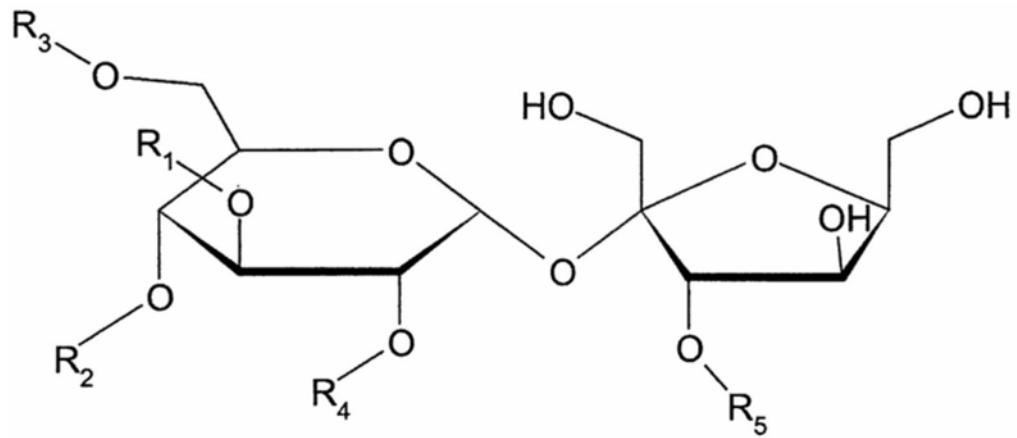


图5

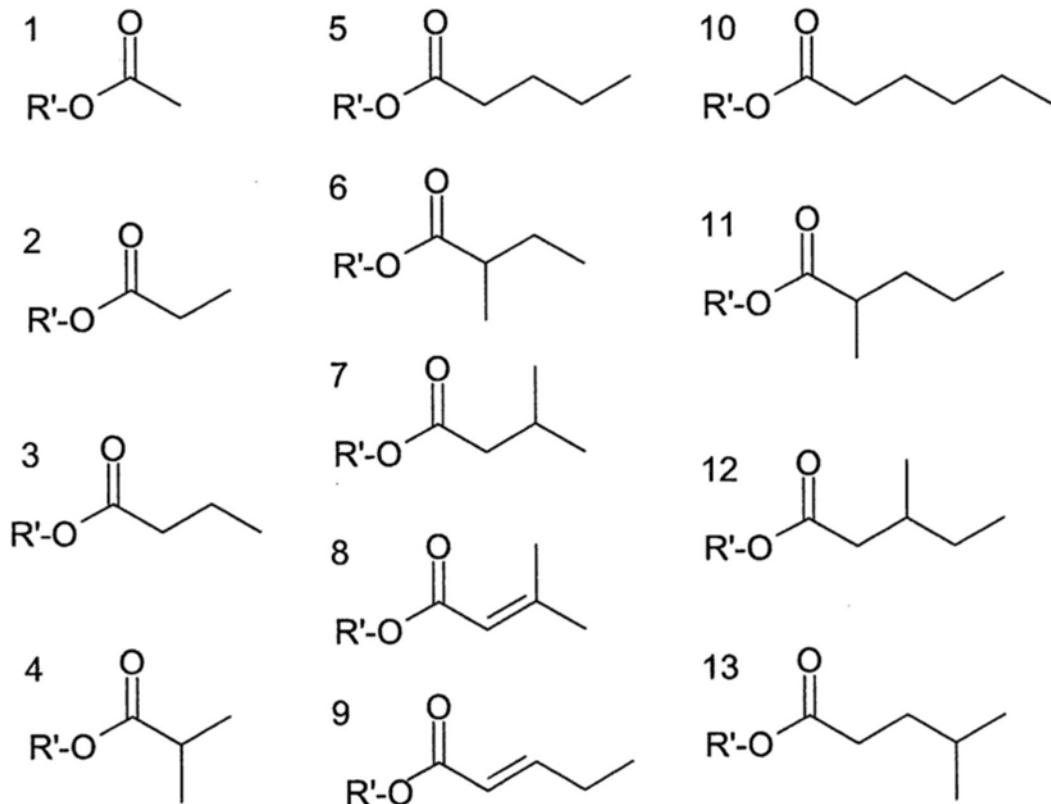


图6