



(12) 发明专利



(10) 授权公告号 CN 116250483 B

(45) 授权公告日 2024. 01. 30

(21) 申请号 202211393862.6

(22) 申请日 2016.06.24

(65) 同一申请的已公布的文献号
申请公布号 CN 116250483 A

(43) 申请公布日 2023.06.13

(30) 优先权数据
62/185,268 2015.06.26 US
62/186,854 2015.06.30 US
62/271,780 2015.12.28 US
62/335,772 2016.05.13 US

(62) 分案原申请数据
201680046625.9 2016.06.24

(73) 专利权人 奥驰亚客户服务公司
地址 美国弗吉尼亚州

(72) 发明人 A·C·亚当斯

M·F·德戈多伊鲁索

S·普拉莫德 许冬梅

(74) 专利代理机构 北京市金杜律师事务所
11256

专利代理师 孟凡宏 袁森

(51) Int.Cl.
A01H 5/12 (2018.01)
A01H 6/82 (2018.01)
C12N 15/82 (2006.01)
C12N 15/29 (2006.01)

(56) 对比文件
CN 102858983 A, 2013.01.02
CN 101611146 A, 2009.12.23

审查员 马彧博

权利要求书1页 说明书97页
序列表(电子公布) 附图7页

(54) 发明名称

用于生产烟草植物的组合物和方法以及具有改变的生物碱含量的制品

(57) 摘要

本公开提供烟草Nic1基因座的鉴定。还提供具有改变的总生物碱和烟碱含量以及商业上可接受的叶等级的烟草植物,其通过育种或转基因方法的开发以及从这些烟草植物生产烟草制品。还提供用于生产具有新的Nic1突变或等位基因以降低烟碱含量的烟草植物的组合物和方法。进一步提供用于育种具有降低的烟碱或无烟碱且同时保持烟叶等级和烟草制品质量的序列多态性和分子标记。

1. 来自烟草植物或其部分的熟化烟草材料,其包含选自下组的序列中的第一多态性:SEQ ID NO:132-135,以及选自下组的第二多态性:SEQ ID NO:138-141,其中当在相似生长条件下生长时,所述烟草植物包含的烟碱含量低于对照植物的烟碱含量的25%,其中当在相似的条件生长并熟化时,所述烟草植物能够生产熟化时USDA等级指数值为所述对照植物的叶的USDA等级指数值的至少65%的叶,其中所述对照植物与所述烟草植物具有除所述第一和第二多态性之外基本上相同的遗传背景,其中所述第一多态性包含选自下组的核苷酸:在SEQ ID NO:132的多态性位点处的T核苷酸、在SEQ ID NO:133的多态性位点处的G核苷酸、在SEQ ID NO:134的多态性位点处的C核苷酸和在SEQ ID NO:135的多态性位点处的T核苷酸,其中所述第二多态性包含选自下组的核苷酸:在SEQ ID NO:138的多态性位点处的A核苷酸、在SEQ ID NO:139的多态性位点处的A核苷酸、在SEQ ID NO:140的多态性位点处的G核苷酸和在SEQ ID NO:141的多态性位点处的G核苷酸,并且其中所述熟化烟草材料包含所述第一多态性和所述第二多态性。

2. 如权利要求1所述的熟化烟草材料,其中与来自不含所述突变的对照烟草植物的熟化烟草材料相比,所述熟化烟草材料包含更低含量的烟碱。

3. 如权利要求1所述的熟化烟草材料,其中所述烟草植物包含的烟碱含量为0.2% - 0.6%。

4. 如权利要求1所述的熟化烟草材料,其中所述熟化烟草材料通过选自下组的熟化过程制成:烤制熟化、晾制熟化、明火烤制熟化和晒制熟化。

5. 烟草制品,其包含如权利要求1所述的熟化烟草材料。

6. 如权利要求5所述的烟草制品,其中所述烟草制品选自下组:香烟、小雪茄、不透气过滤嘴香烟、透气过滤嘴香烟、雪茄、鼻烟、烟斗烟、雪茄烟、卷烟、嚼烟、叶烟、烟碎和烟丝。

7. 如权利要求5所述的烟草制品,其中所述烟草制品是无烟烟草制品。

8. 如权利要求1所述的熟化烟草材料,其中当在相似生长条件下生长时,所述烟草植物包含的烟碱含量比所述对照植物的烟碱含量低10%。

9. 如权利要求1所述的熟化烟草材料,其中所述烟草植物能够生产USDA等级指数值为50或更高的叶。

10. 如权利要求8所述的熟化烟草材料,其中当在相似生长条件下生长时,所述烟草植物包含的烟碱含量比所述对照植物的烟碱含量低8%。

11. 如权利要求1所述的熟化烟草材料,其中所述烟草植物是东方烟草。

12. 如权利要求1所述的熟化烟草材料,其中所述烟草植物包含低于2.0%的烟碱含量。

用于生产烟草植物的组合物和方法以及具有改变的生物碱含量的制品

[0001] 本申请是申请日为2016年06月24日、申请号为201680046625.9、发明名称为“用于生产烟草植物的组合物和方法以及具有改变的生物碱含量的制品”的发明专利申请的分案申请。

[0002] 申请号为201680046625.9的发明专利申请要求2015年6月26日提交的美国临时申请No.62/185,268(未决),2015年6月30日提交的美国临时申请No.62/186,854(未决),2015年12月28日提交的美国临时申请No.62/271,780以及2016年5月13日提交的美国临时申请No.62/335,772(未决)的权益,所有这些申请的全部内容通过引用并入本文。

[0003] 并入序列表

[0004] 序列表包含在3,579,889字节(在MS-Windows®中测量)的于2016年6月24日创建的命名为“P34319US04.txt”的文件中,该序列表包含127个核苷酸序列和37个氨基酸序列,其以电子方式提交并且其全部内容通过引用并入本文。

发明领域

[0005] 本公开提供烟草Nic1基因座的鉴定。还提供具有改变的总生物碱和烟碱含量以及商业可接受的叶等级的烟草植物,通过育种或转基因途径对它们进行开发,以及从这些烟草植物生产烟草制品。

[0006] 发明背景

[0007] 在烟草中发现了四种主要的生物碱:烟碱、降烟碱、假木贼碱和新烟草碱。烟碱是主要的生物碱,通常在商业烟草栽培种中占总生物碱的90%以上。烟碱的生物合成主要发生在烟草根部。然后,烟草植物通过维管束将烟碱转运到叶片,然后在那里将烟碱储存在液泡中。

[0008] 多种因素影响烟草的生物碱含量,包括基因型、环境、受精和农艺实践(例如,烟草生产受到打顶、外伤和食草动物伤害的刺激)。最初在古巴雪茄烟草品种株系中发现的低生物碱性状通过一系列回交被引入到香烟品种中。低生物碱烟草种质随后登记在栽培种白肋21的遗传背景中(Legg等,Crop Science,10:212(1970))。使用低生物碱白肋21(LA BU21)系的遗传研究表明,两个未连锁的基因座有助于烟叶中的烟碱含量。这两个基因座被称为Nic1和Nic2。LA BU21中的nic1和nic2突变是半显性的。它们对烟碱含量显示剂量依赖效应,nic1的作用比nic2强约2.4倍。已经报道了Nic2基因座的分子表征。显示nic2突变含有来自乙烯响应因子(ERF)家族的转录因子基因簇的缺失。

[0009] 降低烟草中的总生物碱含量可以有很多好处。它可以提高烟草作为生物质资源的价值。增加烟草植物中烟碱类生物碱的含量可以在保护植物抵御昆虫和食草动物方面发挥重要作用。

[0010] 与生物碱在昆虫防御中的作用一致,据报道LA BU21对虫害极为敏感(Legg等,Crop Science,10:212(1970))。比较总生物碱含量低(约0.20%)的烤制熟化烟草近等基因系和它们的“正常”轮回亲本(总生物碱含量为1.85-2.70%)的进一步研究表明,低生物碱

系中的产量、等级指数、总N和还原糖含量低于正常的烤制熟化栽培种(Chaplin和Weeks, Crop Science, 16(3):416-18(1976))。

[0011] 需要鉴定调控烟草烟碱含量的新基因,并且开发含有改变的烟碱含量(例如降低的烟碱)并同时保持烟叶质量(如果不是使质量更优的话)的烟草植物和制品。

[0012] 发明简述

[0013] 在一些方面,本公开提供烟草植物或其部分,其包含Nic1基因座中的突变、Nic2基因座中的突变或两者,其中该烟草植物能够生产USDA等级指数值为50或更高的叶。

[0014] 在其它方面,本公开提供烟草植物或其部分,其包含Nic1基因座中的突变、Nic2基因座中的突变或两者,其中当在相似的生长条件下生长时,该烟草植物能够生产USDA等级指数值与对照植物的USDA等级指数值相当的叶,其中对照植物与烟草植物具有除该突变之外基本上相同的遗传背景。

[0015] 在一些方面,本公开还提供非转基因烟草植物或其部分,其包含选自低于2.0%的烟碱含量,其中该烟草植物能够生产USDA等级指数值为50或更高的叶。

[0016] 在其它方面,本公开还提供烟草植物或其部分,其包含非转基因突变,其中当在相似的生长条件下生长时,该非转基因突变将烟草植物的烟碱含量降低至对照植物的烟碱含量的约20%或更低,其中该烟草植物能够生产USDA等级指数值与对照植物的USDA等级指数值相当的叶,并且其中对照植物与烟草植物具有除该非转基因突变之外基本上相同的遗传背景。

[0017] 在一些方面,本公开提供烟草植物或其部分,其包含Nic1基因座中的突变、Nic2基因座中的突变或两者,其中当在相似的生长条件下生长时,该烟草植物包含与对照烟草植物相比含量相似的选自下组的一种或多种烟草香气化合物:3-甲基戊酸、戊酸、异戊酸、labdenoid、西松烷、糖酯和还原糖。

[0018] 在其它方面,本公开提供烟草植物或其部分,其在Nic1基因座中包含突变,其中该突变在LA白肋21中不存在。在一些方面,与LA白肋21相比,本文提供的烟草植物在Nic1基因座处包含较短的染色体缺失。在其它方面,本文提供的烟草植物在Nic1基因座中不包含完整基因的缺失或完整基因编码序列的缺失。

[0019] 在一些方面,本公开提供烟草植物或其部分,其在一个或多个基因内包含一个或多个突变,该基因包含与选自下组的序列具有至少80%同一性的序列:SEQ ID NOs:9-28、75-82、86-100、145及它们的片段。在一些方面,本文提供的烟草植物在Nic1基因座处包含一个或多个非天然存在的突变等位基因,其降低或消除来自Nic1或Nic2基因座的一个或多个基因活性。在一些方面,这些突变等位基因导致更低的烟碱含量。

[0020] 在其它方面,本公开提供烟草植物或其部分,其在一个或多个基因内包含一个或多个突变,该基因包含与选自下组的序列具有至少80%同一性的序列:SEQ ID NOs:29-48、83、101-115、146及它们的片段。

[0021] 在一些方面,本公开提供烟草植物或其部分,其在一个或多个编码多肽的基因内包含一个或多个突变,该多肽与选自下组的序列具有至少80%同一性:SEQ ID NOs:49-68、84、116-130、147及它们的片段。

[0022] 在其它方面,本公开提供烟草植物或其部分,其包含异源表达盒,该异源表达盒包含基因的Nic1抑制序列,该基因包含与选自下组的序列具有至少80%同一性的序列:SEQ

ID N0s:9-28、75-82、86-100、145及它们的片段,其中该抑制序列与植物细胞中功能性的启动子可操作连接,并且其中该抑制序列与和选自下组的序列具有至少80%同一性的序列的至少21个核苷酸的片段具有至少90%序列同一性:SEQ ID N0s:9-28、75-82、86-100、145及它们的片段。

[0023] 在其它方面,本公开提供重组DNA构建体,其包含烟草细胞中功能性的并且与编码RNA分子的多核苷酸可操作连接的启动子,该RNA分子能够与编码多肽的RNA结合,该多肽的氨基酸序列与选自下组的氨基酸序列具有至少80%同一性:SEQ ID N0s:49-68、84、116-130、147及它们的片段,并且其中RNA分子抑制该多肽的表达。

[0024] 在一些方面,本公开提供重组DNA构建体,其包含烟草细胞中功能性的并且与编码多肽的多核苷酸可操作连接的启动子,该多肽的氨基酸序列与选自下组的氨基酸序列具有至少80%同一性:SEQ ID N0s:49-68、84、116-130、147及它们的片段。

[0025] 本公开进一步提供包含来自本文公开的烟草植物、系、品种或杂交种的植物材料的烤制烟草、烟草混合物、烟草制品。

[0026] 本公开还提供用于育种包含所需总生物碱或烟碱含量(例如低烟碱或无烟碱)的烟草系、栽培种或品种的方法。在一些方面,本公开提供将低烟碱性状导入烟草品种的方法,该方法包含:(a)将包含低烟碱性状的第一烟草品种与不具有低烟碱性状的第二烟草品种杂交以生产一个或多个后代烟草植物;(b)对所述一个或多个后代烟草植物进行基因分型检测与低烟碱性状连锁的多态性标记,其中该多态性标记位于侧翼为表3中列出的任何两个多态性位点或侧翼为表4中列出的任何两个多态性位点的染色体区间中;和(c)选择包含低烟碱性状的后代烟草植物。

[0027] 在一些方面,本公开提供将低烟碱性状导入烟草品种的方法,该方法包含:(a)将包含低烟碱性状的第一烟草品种与不具有低烟碱性状的第二烟草品种杂交以生产一个或多个后代烟草植物;(b)对所述一个或多个后代烟草植物进行基因分型检测与低烟碱性状连锁的多态性标记,其中该多态性标记位于表3、4、16和17中列出的任何一个多态性位点的20cM内;和(c)选择包含低烟碱性状的后代烟草植物。在一个方面,多态性标记是选自下组的SNP标记:SEQ ID Nos.131-144。

[0028] 在其它方面,本公开提供选择具有低烟碱性状的烟草植物的方法,该方法包含:(a)从烟草种质集合中分离核酸;(b)测定核酸检测与Nic1基因座紧密连锁的一个或多个标记;和(c)基于标记测定,选择具有低烟碱性状的烟草植物。

附图说明

[0029] 图1显示BU21、HI BU21、LI BU21和LA BU2的组内每个成对比较的差异性调控基因的数量。

[0030] 图2显示Nic1调控在烟碱生物合成中涉及的基因。红色箭头表示在Nic1存在下表达上调的基因,特别是在(i)BU21 vs LA BU21和(ii)HI BU21 vs LA BU21的比较中。

[0031] 图3显示通过其基因型(x轴)和每种植物的化学成分(y轴)作图的F₂群体的基因型数据和化学数据。

[0032] 图4与图3相同,除了排除可能由于采样错误而导致的8个异常值。

[0033] 图5显示2014年田地试验中的等基因BU21低生物碱系列(LA BU21、LI BU21、HI

BU21和BU21)的烟碱含量百分比。

[0034] 图6显示2015年田地试验中的等基因BU21低生物碱系列(LA BU21、LI BU21、HI BU21和BU21)的烟碱含量百分比。

[0035] 图7显示SEQ ID NO:85上的NDG1-NDG15基因的相对基因组位置,其显示在4个重叠片段中(从上到下,分别为核苷酸1至~100k、~100k至~200k、~200k至~300k和~300k至~425k)

[0036] 序列说明

[0037] SEQ ID No:1列出了从LA BU21鉴定的Nic1相关缺失的序列,Scaffold0002504。

[0038] SEQ ID No:2列出了从LA BU21鉴定的Nic2相关缺失的序列,Scaffold0000549。

[0039] SEQ ID Nos:3-8列出了用于PCR确认鉴定的Nic1和Nic2缺失的引物序列。

[0040] SEQ ID Nos:9-28列出了Nic1相关缺失Scaffold0002504中20个注释基因的基因组序列。

[0041] SEQ ID Nos:29-48列出了Nic1相关缺失Scaffold0002504中20个注释基因的cDNA序列。

[0042] SEQ ID Nos:49-68列出了由Nic1相关缺失Scaffold0002504中20个注释基因编码的氨基酸序列。

[0043] SEQ ID Nos:69-70列出了通过RNA干扰(RNAi)抑制g100614_Scaffold0002504和g100631_Scaffold0002504的示例性转化盒序列。

[0044] SEQ ID No:71和72列出了序列多态性的参考TN90等位基因。

[0045] SEQ ID No:73列出了烟草基因组序列组装NT2.0-Scaffold 4274的序列,其包含NT1.0-Scaffold0002504的再测序片段。具体地,NT2.0-Scaffold4272的核苷酸148796-282345对应于并且以负方向取代NT1.0-Scaffold0002504的核苷酸384701-542313。

[0046] SEQ ID No:74列出了烟草基因组序列组装NT2.0-Scaffold 14415的序列,其包含NT1.0-Scaffold0002504的再测序片段。具体地,NT2.0-Scaffold 14415的核苷酸1-59671对应于并且以负方向取代NT1.0-Scaffold0002504的核苷酸288601-363040。

[0047] SEQ ID Nos:75和82列出了Nic1相关缺失中8个注释基因的再测序和细化基因组序列。

[0048] 基于再测序,SEQ ID Nos:83和84进一步列出了g100623_Scaffold0002504中基因的注释cDNA和氨基酸序列。

[0049] SEQ ID No:85列出了LA BU21中完整Nic1缺失区的基因组序列。

[0050] SEQ ID Nos:86-100列出了SEQ ID No:85中15个注释基因的基因组序列。

[0051] SEQ ID Nos:101-115列出了SEQ ID No:85中15个注释基因的cDNA序列。

[0052] SEQ ID Nos:116-130列出了SEQ ID No:85中15个注释基因的蛋白序列。

[0053] SEQ ID Nos:131-142列出了nic1或nic2缺失侧翼的12个SNP标记序列。

[0054] SEQ ID Nos:143-144列出了与ERF-39样基因相关联的两个SNP标记序列。

[0055] SEQ ID Nos:145-147列出了ERF-39样基因的基因组、cDNA和蛋白序列。

[0056] SEQ ID Nos:148-164列出了靶向NDG1-NDG15的含有反向重复的RNAi盒的序列。

[0057] 本文公开的各种序列包括核苷酸序列中的“N”或氨基酸序列中的“X”。“N”可以是任何核苷酸,例如A、T、G、C或一个或多个核苷酸的缺失或插入。在某些情况下,显示一串

“N”。“N”的数量不一定与该位置上未确定的核苷酸的实际数量相关。实际的核苷酸序列可以比所显示的“N”片段更长或更短。类似地，“X”可以是任何氨基酸残基或一个或多个氨基酸的缺失或插入。此外，“X”的数量不一定与该位置上未确定的氨基酸的实际数量相关。实际的氨基酸序列可以比所显示的“X”片段更长或更短。

[0058] 发明详述

[0059] 除非另外定义,否则本文使用的技术术语和科学术语的含义与本领域普通技术人员通常理解的相同。本领域技术人员将理解,在本公开的实践中可以使用许多方法。实际上,本公开不限于所描述的方法和材料。为了本公开的目的,以下术语定义如下。

[0060] 本文引用的任何参考文献(例如,包括所有专利和出版物)的全部内容通过引用并入本文。

[0061] 如本文所用,除非上下文另有明确说明,否则单数形式“一”(a)、“一”(an)和“所述”(the)包括复数形式。例如,术语“一种化合物”或“至少一种化合物”可以包括多种化合物,包括它们的混合物。

[0062] 本文所用的术语“约”是指大约、大体上、左右或在其区间。当术语“约”与数值范围一起使用时,它通过将边界延伸至设定的数值之上或之下来改变所述范围。

[0063] 如本文所用,烟草植物可以来自烟草(*Nicotiana*)属的任何植物,其包括但不限于红花烟草(*Nicotiana tabacum*)、抱茎烟草(*Nicotiana amplexicaulis*)PI 271989;本塞姆氏烟草(*Nicotiana benthamiana*)PI555478;毕基劳氏烟草(*Nicotiana bigelovii*)PI 555485;迪勃纳氏烟草(*Nicotiana debneyi*);高烟草(*Nicotiana excelsior*)PI 224063;粘烟草(*Nicotiana glutinosa*)PI 555507;古特斯比氏烟草(*Nicotiana goodspeedii*)PI 241012;哥西氏烟草(*Nicotiana gossei*)PI 230953;西烟草(*Nicotiana hesperis*)PI 271991;奈特氏烟草(*Nicotiana knightiana*)PI 555527;*Nicotiana maritima* PI 555535;特大管烟草(*Nicotiana megalosiphon*)PI 555536;*Nicotiana nudicaulis* PI555540;圆锥烟草(*Nicotiana paniculata*)PI 555545;蓝茉莉叶烟草(*Nicotiana plumbaginifolia*)PI 555548;残波烟草(*Nicotiana repanda*)PI 555552;黄花烟草(*Nicotiana rustica*);*Nicotiana suaveolens* PI 230960;林烟草(*Nicotiana sylvestris*)PI 555569;绒毛烟草(*Nicotiana tomentosa*)PI 266379;绒毛状烟草(*Nicotiana tomentosiformis*);和三角叶烟草(*Nicotiana trigonophylla*)PI 555572。

[0064] 在一些方面,本公开提供烟草植物或其部分,其包含Nic1基因座中的突变,Nic2基因座中的突变或两者,其中该烟草植物能够生产USDA等级指数值为50或更高的叶。在一些方面,本文公开的烟草植物能够生产具有选自下组的USDA等级指数值的叶:55或更高、60或更高、65或更高、70或更高、75或更高、80或更高、85或更高、90或更高、95或更高。在其它方面,当在相似的生长条件下生长时,本文公开的烟草植物能够生产USDA等级指数值与对照植物的USDA等级指数值相当的叶,其中对照植物与烟草植物具有除该突变之外基本上相同的遗传背景。在进一步的方面,当在相似的生长条件下生长时,本文公开的烟草植物能够生产USDA等级指数值为对照植物的USDA等级指数值的至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、或至少约98%的叶,其中对照植物与烟草植物具有除该突变之外基本上相同的遗传背景。在进一步的方面,本文公开的烟草植物能够生产USDA等级指数值为对照植物的USDA等级指数值的65%-130%、70%-130%、75%-

130%、80%-130%、85%-130%、90%-130%、95%-130%、100%-130%、105%-130%、110%-130%、115%-130%或120%-130%的叶。在进一步的方面,本文公开的烟草植物能够生产USDA等级指数值为对照植物的USDA等级指数值的70%-125%、75%-120%、80%-115%、85%-110%或90%-100%的叶。在一些方面,当在相似的生长条件下生长时,本文公开的烟草植物包含的烟碱含量低于对照植物烟碱含量的1%、2%、5%、8%、10%、12%、15%、20%、25%、30%、40%、50%、60%、70%或80%,其中对照植物与烟草植物具有除该突变之外基本上相同的遗传背景。在其它方面,本文公开的烟草植物包含选自下组的烟碱含量:低于3%、低于2.75%、低于2.5%、低于2.25%、低于2.0%、低于1.75%、低于1.5%、低于1.25%、低于1%、低于0.9%、低于0.8%、低于0.7%、低于0.6%、低于0.5%、低于0.4%、低于0.3%、低于0.2%、低于0.1%和低于0.05%。在进一步的方面,本文公开的烟草植物进一步包含直接抑制编码选自下组产物的一个或多个基因的表达或活性的转基因或突变:PMT、MPO、QPT、BBL、A622和MATE转运蛋白。

[0065] 在一些方面,本公开提供烟草植物或其部分,其包含Nic1基因座中的突变、Nic2基因座中的突变或两者,其中当在相似的生长条件下生长时,该烟草植物能够生产USDA等级指数值与对照植物的USDA等级指数值相当的叶,其中对照植物与烟草植物具有除该突变之外基本上相同的遗传背景。在其它方面,本文公开的烟草植物能够生产具有选自下组的USDA等级指数值的叶:55或更高、60或更高、65或更高、70或更高、75或更高、80或更高、85或更高、90或更高、95或更高。在其它方面,本文公开的烟草植物能够生产具有选自下组的USDA等级指数值的叶:50-95、55-95、60-95、65-95、70-95、75-95、80-95、85-95、90-95、55-90、60-85、65-80、70-75、50-55、55-60、60-65、65-70、70-75、75-80、80-85、85-90和90-95。在进一步的方面,本文公开的烟草植物能够生产USDA等级指数值为对照植物的USDA等级指数值的至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%或至少约98%的叶。在进一步的方面,本文公开的烟草植物能够生产USDA等级指数值为对照植物的USDA等级指数值的65%-130%、70%-130%、75%-130%、80%-130%、85%-130%、90%-130%、95%-130%、100%-130%、105%-130%、110%-130%、115%-130%或120%-130%的叶。在进一步的方面,本文公开的烟草植物能够生产USDA等级指数值为对照植物的USDA等级指数值的70%-125%、75%-120%、80%-115%、85%-110%或90%-100%的叶。在其它方面,当在相似的生长条件下生长时,本文公开的烟草植物包含的烟碱含量低于对照植物的烟碱含量的1%、2%、5%、8%、10%、12%、15%、20%、25%、30%、40%、50%、60%、70%或80%。在其它方面,本文公开的烟草植物进一步包含直接抑制编码选自下组的产物的一个或多个基因的表达或活性的转基因或突变:PMT、MPO、QPT、BBL、A622和MATE转运蛋白。

[0066] 在一些方面,本公开还提供包含选自下组的突变的烟草品种、栽培种或系:nic1突变、nic2突变及它们的组合,其中当在相似的生长条件下生长时,该烟草品种、栽培种或系的叶等级与对照烟草品种、栽培种或系的叶等级相当,其中对照烟草品种与该烟草品种、栽培种或系具有除该突变之外基本上相同的遗传背景。

[0067] 在一些方面,本公开还提供非转基因烟草植物或其部分,其包含选自下组的烟碱含量:低于3%、低于2.75%、低于2.5%、低于2.25%、低于2.0%、低于1.75%、低于1.5%、低于1.25%、低于1%、低于0.9%、低于0.8%、低于0.7%、低于0.6%、低于0.5%、低于

0.4%、低于0.3%、低于0.2%、低于0.1%、和低于0.05%，其中该烟草植物能够生产USDA等级指数值为50或更高、55或更高、60或更高、65或更高、70或更高、75或更高、80或更高、85或更高、90或更高以及95或更高的叶。在其它方面，该非转基因烟草植物包含低于2.0%的烟碱含量并且能够生产USDA等级指数值为70或更高的叶。在进一步的方面，该非转基因烟草植物包含低于1.0%的烟碱含量并且能够生产USDA等级指数值为70或更高的叶。

[0068] 在一些方面，本公开还提供烟草植物或其部分，其包含非转基因突变，其中当在相似的生长条件下生长时，该非转基因突变将烟草植物的烟碱含量降低至低于对照植物的烟碱含量的1%、2%、5%、8%、10%、12%、15%、20%、25%、30%、40%、50%、60%、70%或80%，其中该烟草植物能够生产USDA等级指数值与对照植物的USDA等级指数值相当的叶，并且其中对照植物与该烟草植物具有除该非转基因突变之外基本上相同的遗传背景。

[0069] 在一些方面，本公开提供烟草植物或其部分，其在Nic1基因座中包含突变，其中该突变在LA白肋21中不存在。在一些方面，与LA白肋21相比，本文提供的烟草植物在Nic1基因座处包含较短的染色体缺失。在其它方面，本文提供的烟草植物在Nic1基因座中不包含完整基因的缺失或完整基因编码序列的缺失。在一些方面，本文提供的烟草植物在Nic1基因座处是纯合的。在其它方面，本文提供的烟草植物在Nic1基因座处是杂合的。在一些方面，本文提供的烟草植物包含选自下组的Nic1突变：点突变、缺失、插入、复制和逆转。在一些方面，本文提供的烟草植物中的Nic1突变通过选自下组的方法引入：随机诱变和定向诱变。在其它方面，本文提供的烟草植物中的Nic1突变通过选自下组的定向诱变方法引入：大范围核酸酶、锌指核酸酶、TALEN和CRISPR。

[0070] 在一些方面，本文提供的烟草植物在一个或多个基因内包含一个或多个突变，该基因包含与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%或100%同一性的序列：SEQ ID N0s:9-28、75-82、86-100、145及它们的片段。在一些方面，一个或多个突变降低一个或多个基因的表达或活性，该基因包含与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%或100%同一性的序列：SEQ ID N0s:9-28、75-82、86-100、145及它们的片段。

[0071] 在一些方面，本文提供的烟草植物在一个或多个基因内包含一个或多个突变，该基因包含与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%或100%同一性的编码序列：SEQ ID N0s:29-48、83、101-115、146及它们的片段。在一些方面，一个或多个突变降低一个或多个基因的表达或活性，该基因包含与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%或100%同一性的编码序列：SEQ ID N0s:29-48、83、101-115、146及它们的片段。

[0072] 在一些方面，本文提供的烟草植物在一个或多个编码多肽的基因内包含一个或多个突变，该多肽与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%或100%同一性：SEQ ID N0s:49-68、84、116-130、147及它们的片段。在一些方面，一个或多个突变降低一个或多个编码多肽的基因的表达或活性，该多肽与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%或100%同一性：SEQ ID N0s:49-68、84、116-130、147及它们的片段。

[0073] 在一些方面，本文提供的烟草植物在一个或多个基因内包含一个或多个突变，该基因包含与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少

98%、至少99%或100%同一性的序列:SEQ ID NOs:13、28、33、48、82、86、87、101、102、145、146及它们的片段。在一些方面,一个或多个突变降低一个或多个基因的表达或活性,该基因包含与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%或100%同一性的序列:SEQ ID NOs:13、28、33、48、82、86、87、101、102、145、146及它们的片段。

[0074] 在一些方面,本文提供的烟草植物在一个或多个基因内包含一个或多个突变,该基因包含与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%或100%同一性的编码序列:SEQ ID NOs:33、48、101、102、146及它们的片段。在一些方面,一个或多个突变降低一个或多个基因的表达或活性,该基因包含与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%或100%同一性的编码序列:SEQ ID NOs:33、48、101、102、146及它们的片段。

[0075] 在一些方面,本文提供的烟草植物在一个或多个编码多肽的基因内包含一个或多个突变,该多肽包含与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%或100%同一性的编码序列:SEQ ID NOs:53、68、116、117、147及它们的片段。在一些方面,一个或多个突变降低一个或多个编码多肽的基因的表达或活性,该多肽包含与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%或100%同一性的编码序列:SEQ ID NOs:53、68、116、117、147及它们的片段。

[0076] LA白肋21是低总生物碱烟草系,其通过几次回交将来自古巴雪茄品种的低生物碱基因引入白肋21中来生产(Legg等.1970)。与母本白肋21的约3.5%(干重)相比,它具有约0.2%的总生物碱(干重)。LA BU21的叶等级远低于商业上可接受的标准。

[0077] 除非另外指明,否则本文提到的烟草植物、品种、栽培种或系的生物碱或烟碱含量或叶等级指数值的测量值是指平均测量值,例如,包括单株植物的多个叶的平均值或来自单个品种、栽培种或系的烟草植物群体的平均测量值。用于确定平均测量值(例如,生物碱或烟碱含量或叶等级)的烟草植物群体或烟叶的集合可以具有任何大小,例如5、10、15、20、25、30、35、40或50。随后,用工业上接受的标准方案来确定平均测量值或等级指数值。

[0078] 如本文所用,“相似的生长条件”是指用于生长的相似环境条件和/或农艺实践,其能够在两种或更多种植物基因型之间进行有意义的比较,使得环境条件或农艺实践都不会有助于或解释在两种或更多种植物基因型之间观察到的任何差异。例如,环境条件包括光、温度、水(湿度)和营养(例如,氮和磷)。例如,农艺实践包括播种、修剪、根切、移栽、打顶和分株。参见Tobacco, Production, Chemistry and Technology, Davis&Nielsen, eds., Blackwell Publishing, Oxford(1999), pp 70-103的第4B章和第4C章。

[0079] “生物碱”是天然存在于植物中的复杂的含氮化合物,并且在人类和动物中具有药理作用。“烟碱”是商业化卷烟中的主要天然生物碱,其在红花烟草中约占生物碱含量的90%。烟草中的其他主要生物碱包括可替宁、降烟碱、麦斯明烟草碱、二烯烟碱、假木贼碱和新烟草碱。次要的烟草生物碱包括烟碱-n-氧化物、N-甲基新烟草碱、N-甲基假木贼碱、假氧烟碱、2,3-联吡啶等。

[0080] 在一些方面,当在相似生长条件下生长时,与没有Nic1突变的对照烟草植物相比,本文提供的烟草植物包含更低含量的总生物碱或单个生物碱。在其它方面,当在相似生长

条件下生长时,与对照烟草植物相比,本文提供的烟草植物包含更低含量的选自下组的一种或多种生物碱:可替宁、降烟碱、麦斯明烟草碱、二烯烟碱、假木贼碱和新烟草碱。在一些方面,更低的生物碱含量是指生物碱含量低于对照烟草植物的生物碱含量的1%、2%、5%、8%、10%、12%、15%、20%、25%、30%、40%、50%、60%、70%或80%。在其它方面,更低的生物碱含量是指生物碱含量为对照烟草植物的生物碱含量的约0.5%-1%、1%-2%、2%-3%、3%-4%、4%-5%、5%-6%、6%-7%、7%-8%、8%-9%、9%-10%、11%-12%、12%-13%、13%-14%、14%-15%、15%-16%、16%-17%、17%-18%、18%-19%、19%-20%、21%-22%、22%-23%、23%-24%、24%-25%、25%-26%、26%-27%、27%-28%、28%-29%或29%-30%。在进一步的方面,更低的生物碱含量是指生物碱含量为对照烟草植物的生物碱含量的约0.5%-5%、5%-10%、10%-20%、20%-30%。

[0081] 生物碱含量可以通过本领域已知的方法测定,例如基于气相-液相色谱、高效液相色谱、放射性免疫测定和酶联免疫吸附测定进行定量。例如,烟碱生物碱含量可以通过基于CORESTA推荐方法No.7 (1987)的GC-FID方法和ISO标准 (ISO TC 126N 394E,也可参见Hibi et al., Plant Physiology 100:826-35 (1992))的使用配备有毛细管柱和FID检测器的气相-液相色谱的方法来测量。

[0082] 或者,烟草总生物碱可以使用由Skalar Instrument Co (West Chester, PA)改良并且如文献Collins等, Tobacco Science 13:79-81 (1969)描述的用于分析烟草样品的分段流动比色法测量。简言之,在分析总生物碱和还原糖之前,将烟草样品干燥、研磨并提取。然后,该方法采用乙酸/甲醇/水提取并采用木炭进行脱色。总生物碱的测定是基于芳胺存在下氯化氰与烟碱生物碱反应形成在460nm处测量的有色复合物。除非另有说明,否则本文显示的总生物碱含量或烟碱含量是基于干重(例如,总生物碱百分比或烟碱百分比)。

[0083] 在一些方面,当在相似生长条件下生长时,与没有Nic1突变的对照烟草植物相比,本文提供的烟草植物包含更低含量的烟碱。在一些方面,更低的烟碱含量是指平均烟碱含量低于对照烟草植物的平均烟碱含量的1%、2%、5%、8%、10%、12%、15%、20%、25%、30%、40%、50%、60%、70%或80%。在其它方面,更低的烟碱含量是指平均烟碱含量为对照烟草植物的平均烟碱含量的约0.5%-1%、1%-2%、2%-3%、3%-4%、4%-5%、5%-6%、6%-7%、7%-8%、8%-9%、9%-10%、11%-12%、12%-13%、13%-14%、14%-15%、15%-16%、16%-17%、17%-18%、18%-19%、19%-20%、21%-22%、22%-23%、23%-24%、24%-25%、25%-26%、26%-27%、27%-28%、28%-29%或29%-30%。在进一步的方面,更低的烟碱含量是指平均烟碱含量为对照烟草植物的平均烟碱含量的约0.5%-5%、5%-10%、10%-20%、20%-30%。

[0084] 在一些方面,基于干重,本文提供的烟草植物包含选自下组的平均烟碱含量:约0.01%、0.02%、0.05%、0.75%、0.1%、0.15%、0.2%、0.3%、0.35%、0.4%、0.5%、0.6%、0.7%、0.8%、0.9%、1%、1.1%、1.2%、1.3%、1.4%、1.5%、1.6%、1.7%、1.8%、1.9%、2%、2.1%、2.2%、2.3%、2.4%、2.5%、2.6%、2.7%、2.8%、2.9%、3%、3.1%、3.2%、3.3%、3.4%、3.5%、3.6%、3.7%、3.8%、3.9%、4%、5%、6%、7%、8%和9%。在其他方面,基于干重,本文提供的烟草植物包含选自下组的平均烟碱含量:约0.01%-0.02%、0.02%-0.05%、0.05%-0.75%、0.75%-0.1%、0.1%-0.15%、0.15%-0.2%、0.2%-0.3%、0.3%-0.35%、0.35%-0.4%、0.4%-0.5%、0.5%-0.6%、0.6%-0.7%、0.7%-

0.8%、0.8%-0.9%、0.9%-1%、1%-1.1%、1.1%-1.2%、1.2%-1.3%、1.3%-1.4%、1.4%-1.5%、1.5%-1.6%、1.6%-1.7%、1.7%-1.8%、1.8%-1.9%、1.9%-2%、2%-2.1%、2.1%-2.2%、2.2%-2.3%、2.3%-2.4%、2.4%-2.5%、2.5%-2.6%、2.6%-2.7%、2.7%-2.8%、2.8%-2.9%、2.9%-3%、3%-3.1%、3.1%-3.2%、3.2%-3.3%、3.3%-3.4%、3.4%-3.5%和3.5%-3.6%。在进一步的方面,基于干重,本文提供的烟草植物包含选自下组的平均烟碱含量:约0.01%-0.1%、0.02%-0.2%、0.03%-0.3%、0.04%-0.4%、0.05%-0.5%、0.75%-1%、0.1%-1.5%、0.15%-2%、0.2%-3%和0.3%-3.5%。

[0085] 本公开还提供对其他烟草性状(例如叶等级指数值)没有负面影响的具有改变的烟碱含量的烟草植物。在一个方面,本文公开的低烟碱或无烟碱烟草品种提供商业上可接受等级的熟化烟草。烟草等级的评估基于包括但不限于以下的因素:叶柄位置、叶大小、叶颜色、叶均匀性和完整性、成熟度、质地、弹性、光泽(与叶颜色的强度和深度以及光泽有关)、吸湿性(烟叶吸收和保持环境湿度的能力)以及绿色的细微差别或偏色。例如,可以使用由美国农业部农业市场服务部公布的官方标准等级(7U.S.C. §511)来确定叶等级。例如,参见1990年11月5日生效的白肋烟(Burley Tobacco)(美国31型和外国93型)官方标准等级(55F.R.40645);1989年3月27日生效的烤制熟化烟草(Flue-Cured Tobacco)(美国11、12、13、14型和外国92型)官方标准等级(54F.R.7925);1965年1月8日生效的宾夕法尼亚州子叶烟草(Seedleaf Tobacco)(美国41型)官方标准等级(29F.R.16854);1963年12月8日生效的俄亥俄州雪茄叶烟草(Cigar-Leaf Tobacco)(美国42、43和44型)官方标准等级(28F.R.11719和28F.R.11926);1969年11月20日生效的威斯康星州雪茄混合型烟草(美国54和55型)官方标准等级(34F.R.17061);1969年11月20日生效的威斯康星州雪茄混合型烟草官方标准等级(美国54和55型)(34F.R.17061);1971年4月生效的佐治亚州和佛罗里达州荫下种植雪茄包装烟草(美国62型)的官方标准等级。USDA等级指数值可以根据行业接受的等级指数来确定。例如,参见Bowman等,Tobacco Science,32:39-40(1988);传统烟草文献图书馆(Bates文件#523267826-523267833,1988年7月1日,拟议的白肋烟等级指数备忘录);和Miller等,1990,Tobacco Intern.,192:55-57(所有上述参考文献的全部内容通过引用并入本文)。在一个方面,USDA等级指数以0-100的数字代表接收的联邦等级,并且是所有叶柄位置的加权平均值。等级指数越高表示质量越高。或者,可以通过超光谱成像确定叶等级。例如,参见W0 2011/027315(2011年3月10日公开,并且全部内容通过引用并入本文)。

[0086] 在一些方面,当在相似生长条件下生长时,与对照烟草植物相比,本文提供的烟草植物包含相似含量的选自下组的一种或多种烟草香气化合物:3-甲基戊酸、戊酸、异戊酸、labdenoid、西松烷、糖酯和还原糖。在其它方面,本文提供的烟草植物包含Nic1突变、Nic2突变或它们的组合,其对选自下组的一种或多种烟草香气化合物的含量没有影响:3-甲基戊酸、戊酸、异戊酸、labdenoid、西松烷、糖酯和还原糖。

[0087] 如本文所用,烟草香气化合物是与烟草烟气的风味和香气相关的化合物。这些化合物包括但不限于3-甲基戊酸、戊酸、异戊酸、西松烷和labdenoid二萜、和糖酯。烟草香气化合物的浓度可以通过本领域中任何已知的代谢物谱分析方法来测量,包括但不限于气相色谱质谱(GC-MS)、核磁共振光谱、液相色谱-质谱联用。参见Weckwerth和Kahl编辑的“The Handbook of Plant Metabolomics”(Wiley-Blackwell)(2013年5月28日)。

[0088] 如本文所用,“还原糖(S)”是具有游离的或潜在游离的醛基或酮基的任何糖(单糖

或多糖)。通过降低烟气pH并且有效地减少“游离”未质子化烟碱的量,葡萄糖和果糖在香烟烟气中作为烟碱缓冲物。例如,通过改变烟碱和其他烟草生物碱的感官影响,还原糖平衡烟味。已经报道在各种烟草品种间、同一品种内,以及在因种植条件引起的同一株系内,糖含量和生物碱含量之间呈反比关系。还原糖含量可以使用由Skalar Instrument Co (West Chester, PA)改良并如文献Davis, Tobacco Science 20:139-144 (1976)描述的用于分析烟草样品的分段流动比色法测量。例如,样品用碳酸钠溶液透析。将铜新铜试剂 (Copper neocuproin) 添加到样品并且加热溶液。在糖存在下,铜新铜试剂螯合物被还原,形成在460nm处测量的有色复合物。

[0089] 在一些方面,本文提供的烟草植物在Nic1或Nic2基因座处包含一个或多个非天然存在的突变等位基因,其降低或消除来自Nic1或Nic2基因座的一个或多个基因活性。在一些方面,这些突变等位基因导致烟碱含量降低。可通过本领域已知的任何方法引入突变Nic1或Nic2等位基因,包括随机或定向诱变方法。

[0090] 这种诱变方法包括但不限于用甲基硫酸乙酯 (EMS) 处理种子 (Hildering和Verkerk, In, The use of induced mutations in plant breeding. Pergamon press, pp 317 - 320, 1965) 或UV - 辐射、X - 射线、和快中子辐照 (例如, 参见Verkerk, Neth. J. Agric. Sci. 19:197-203, 1971; 和Poehlman, Breeding Field Crops, Van Nostrand Reinhold, New York (3. sup. rd ed), 1987)、转座子标签 (Fedoroff等, 1984; 美国专利号4, 732, 856和美国专利号5, 013, 658) 以及T-DNA插入法 (Hoekema等, 1983; 美国专利号5, 149, 645)。EMS诱导的诱变在于在基因组长度上化学诱导随机点突变。快中子诱变在于将种子暴露于中子轰击,其通过双链DNA断裂引起大量缺失。转座子标签包含在内源基因内插入转座子以减少或消除基因的表达。例如,可能存在于烟草基因中的突变类型包括点突变、缺失、插入、复制和逆转。该突变期望存在于烟草基因的编码区;但是,烟草基因的启动子区、和内含子、或非翻译区中的突变也可能是期望的。

[0091] 此外,用于筛选化学诱导的突变的快速并且可自动化的方法,TILLING (在基因组中靶诱导局部病变) 也适用于本公开,其使用变性HPLC或选择性核酸内切酶消化选择的PCR产物。参见McCallum等 (2000) Nat. Biotechnol. 18:455-457。本文公开的影响基因表达或干扰基因功能的突变可以使用本领域公知的方法来确定。基因外显子中的插入突变通常导致无效突变。保守性残基中的突变可以特别有效地抑制蛋白的功能。在一些方面,本文公开的烟草植物包含本文所述的一个或多个Nic 1基因的无义 (例如终止密码子) 突变。

[0092] 在一些方面,本公开还提供具有改变的烟碱含量并同时保持商业上可接受的叶质量的烟草系。通过精确的基因组工程技术,例如转录激活因子样效应核酸酶 (TALEN)、大范围核酸酶、锌指核酸酶和CRISPR-cas9系统,可以在Nic1或Nic2基因座的一个或多个基因中引入突变来制备这些系。例如,参见Gaj等, Trends in Biotechnology, 31 (7):397-405 (2013)。

[0093] 诱变的烟草植物的筛选和选择可以通过本领域普通技术人员已知的任何方法进行。筛选和选择方法的实例包括但不限于Southern分析、用于检测多核苷酸的PCR扩增、Northern印迹、RNA酶保护、引物延伸、用于检测RNA转录物的RT-PCR扩增、用于检测酶或多肽和多核苷酸的核酶活性的酶测定、以及蛋白凝胶电泳、蛋白印迹法、免疫沉淀法以及用于检测多肽的酶联免疫测定法。诸如原位杂交、酶染色和免疫染色的其他技术也可用于检测

多肽和/或多核苷酸的存在或表达。实施所有引用技术的方法是已知的。

[0094] 本公开还提供用于在植物,特别是烟草属植物(包括各种商业化品种的烟草植物)中抑制来自Nic1基因座的一个或多个多肽的表达或功能的组合物和方法。

[0095] 在一些方面,本发明提供烟草植物或其部分,其包含异源表达盒,该异源表达盒包含基因的Nic1抑制序列,该基因包含与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%或100%同一性的序列:SEQ ID NOs:9-28、75-82、86-100、145及它们的片段,其中该抑制序列与植物细胞中功能性的启动子可操作连接,并且其中该抑制序列与和选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%或100%同一性的序列的至少21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、31个、32个、33个、34个、35个、36个、37个、38个、39个、40个、41个、42个、43个、44个、45个、46个、47个、48个、49个、50个、51个、52个、53个、54个、55个、56个、57个、58个、59个、60个、61个、62个、63个、64个、65个、66个、67个、68个、69个、70个、71个、72个、73个、74个、75个、76个、77个、78个、79个或80个核苷酸的片段具有至少90%序列同一性:SEQ ID NOs:9-28、75-82、86-100、145及它们的片段。在其它方面,本公开提供烟草植物或其部分,其包含异源表达盒,该异源表达盒包含基因的Nic1抑制序列,该基因包含与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%或100%同一性的序列:SEQ ID NOs:13、28、33、48、82、86、87、101、102、145、146及它们的片段,其中该抑制序列与植物细胞中功能性的启动子可操作连接,并且其中该抑制序列与和选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%或100%同一性的序列的至少21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、31个、32个、33个、34个、35个、36个、37个、38个、39个、40个、41个、42个、43个、44个、45个、46个、47个、48个、49个、50个、51个、52个、53个、54个、55个、56个、57个、58个、59个、60个、61个、62个、63个、64个、65个、66个、67个、68个、69个、70个、71个、72个、73个、74个、75个、76个、77个、78个、79个或80个核苷酸的片段具有至少90%序列同一性:SEQ ID NOs:13、28、33、48、82、86、87、101、102、145、146及它们的片段。在一些方面,Nic1抑制序列能够转录为选自下组的抑制多核苷酸:单链RNA多核苷酸、双链RNA多核苷酸及它们的组合。

[0096] 如本文所用,术语“抑制(inhibit)”、“抑制(inhibition)”和“抑制(inhibiting)”定义为本领域已知的或本文所述的任何降低目标基因产物(例如靶基因产物)的表达和功能的任何方法。“抑制”可以是在两种植物之间进行比较的情况下,例如,将遗传改变的植物与野生型植物进行比较。或者,抑制靶基因产物的表达或功能可以是在相同植物内或在不同植物之间的植物细胞、细胞器、器官、组织、或植物部分之间进行比较的情况下,以及包括在相同植物或在植物部分内或在植物或植物部分之间的发育阶段或时间阶段之间进行比较的情况下。“抑制”包括目标基因产物的功能或产量的任何相对减少,直至并且包括该基因产物的功能或产量的完全消除。术语“抑制”涵盖下调靶基因产物的翻译和/或转录或靶基因产物的功能活性的任何方法或组合物。在一些方面,本文公开的改造植物中来自Nic1基因座的一个或多个基因的mRNA或蛋白的含量低于不是突变体或未被基因修饰以抑制该基因表达的植物中的相同基因的蛋白含量的95%、90%、80%、70%、60%、50%、40%、30%、20%、10%、5%、4%、3%、2%或1%。

[0097] 术语“抑制序列”涵盖能够抑制来自植物中的Nic1基因座的参与烟碱生物合成调控的基因的表达或功能的任何多核苷酸或多肽序列,诸如全长多核苷酸或多肽序列、截短的多核苷酸或多肽序列、多核苷酸或多肽序列片段、多核苷酸或多肽序列变体、正义方向的核苷酸序列、反义方向的核苷酸序列、正义或反义方向的核苷酸序列的互补序列、核苷酸序列的反转区、核苷酸序列的发夹、双链核苷酸序列、单链核苷酸序列及它们的组合等。术语“多核苷酸序列”包括RNA、DNA、化学修饰的核酸、核酸类似物及它们的组合等的序列。

[0098] 本文中,抑制序列命名为靶基因产物的名称。因此,“Nic1抑制序列”是指能够抑制来自植物中的Nic1基因座的参与烟碱生物合成调控的基因的表达(例如在转录和/或翻译水平上),或能够抑制基因产物功能的抑制序列。当在多核苷酸抑制序列的上下文中使用短语“能够抑制”时,意欲是指抑制序列本身发挥抑制作用;或者,当抑制序列编码抑制核苷酸分子(例如发夹RNA、miRNA或双链RNA多核苷酸)或编码抑制多肽(例如,抑制靶基因产物的表达或功能的多肽时,在其转录之后(例如在编码发夹RNA、miRNA或双链RNA多核苷酸的抑制序列的情况下)或其转录和翻译之后(例如在编码抑制多肽的情况下),转录或翻译产物分别对靶基因产物发挥抑制作用(例如抑制靶基因产物的表达或功能)。

[0099] 本文公开的Nic1抑制序列可以是经由包括但不限于以下的本领域已知的任何沉默途径或机制触发基因沉默的序列:正义抑制/共抑制、反义抑制、双链RNA(dsRNA)干扰、发夹RNA干扰和含内含子的发夹RNA干扰、扩增子介导的干扰、核酶、小干扰性RNA、人造或合成微RNA、以及人造反式作用siRNA。取决于所需的结果,Nic1抑制序列可以是至少约20个核苷酸、约50个核苷酸、约70个核苷酸、约100个核苷酸、约150个核苷酸、约200个核苷酸、约250个核苷酸、约300个核苷酸、约350个核苷酸、约400个核苷酸,并且直至编码本公开的蛋白的全长多核苷酸。在一个方面,Nic1抑制序列可以是约50-约400个核苷酸、约70至-350个核苷酸、约90-约325个核苷酸、约90-约300个核苷酸、约90-约275个核苷酸、约100-约400个核苷酸、约100-约350个核苷酸、约100-约325个核苷酸、约100-约300个核苷酸、约125-约300个核苷酸、或约125-约275个核苷酸长度的片段。在一些实施方案中,细胞色素P450多核苷酸的片段是约50、约60、约70、约80、约90、约100、约125、约150、约175、约200、约225、约250、约275、约300、约325、约350、约400个核苷酸长度,以及约70和约400个核苷酸之前的其他此类值。

[0100] 术语“多核苷酸”的使用不旨在将本公开限制为包含DNA的多核苷酸。本领域普通技术人员将认识到,多核苷酸可以包含核糖核苷酸以及核糖核苷酸和脱氧核糖核苷酸的组合。该脱氧核糖核苷酸和核糖核苷酸包括天然存在的分子和合成的类似物。本公开的多核苷酸还涵盖所有形式的序列,包括但不限于单链形式、双链形式、发夹、茎环结构等。

[0101] 在一些方面,本公开提供重组DNA构建体,其包含在烟草细胞中是功能性的并且与编码RNA分子的多核苷酸可操作连接的启动子,该RNA分子能够与编码多肽的RNA结合,该多肽的氨基酸序列与选自下组的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%同一性:SEQ ID NOs:49-68、84、116-130、147及它们的片段,并且其中该RNA分子抑制该多肽的表达。在一些方面,RNA分子选自下组:微RNA、siRNA和反式作用siRNA。在其它方面,重组DNA构建体编码双链RNA。还提供包含这些重组DNA构建体的转基因烟草植物或其部分、熟化烟草材料或烟草制品。在一些方面,与不含重组DNA构建体的对照烟草植物相比,这些转基因植物、熟化烟草材料或烟草制品包含更低含量的烟碱。进

一步提供降低烟草植物的烟碱含量的方法,该方法包含用任何这些重组DNA构建体转化烟草植物。

[0102] 如本文所用,“可操作连接”是指两个或更多个元件之间的功能性连接。例如,目标多核苷酸和调控序列(例如启动子)之间的可操作连接是允许目标多核苷酸表达的功能性连接。可操作连接的元件可以是连续的或不连续的。

[0103] 如本文所用并且当用于序列时,“异源”是指源自外来物种的序列,或者如果来自同一物种,则通过有意的人为干预,基于其天然形式在组成和/或基因组位点方面进行实质性修饰的序列。该术语也适用于核酸构建体,其在本文中也称为“多核苷酸构建体”或“核苷酸构建体”。以这种方式,“异源”核酸构建体意在表示源自外来物种的构建体,或如果来自同一物种,则通过有意的人为干预,基于其天然形式在组成和/或基因组位点方面进行实质性修饰的构建体。异源核酸构建体包括但不限于已被引入植物或其部分的重组核苷酸构建体,例如通过转化方法或随后用另一种目标植物育种转基因植物。

[0104] 如本文所用,“基因表达”是指基因产物的生物合成或生产,包括基因产物的转录和/或翻译。

[0105] 本文还提供用于在植物,特别是烟草属植物(包括各种商业化品种的烟草植物)中过表达Nic1基因座的一个或多个多肽的组合物和方法。

[0106] 在一些方面,本公开提供重组DNA构建体,其包含在烟草细胞中功能性的并且与编码多肽的多核苷酸可操作连接的启动子,该多肽的氨基酸序列与选自下组的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%或100%同一性:SEQ ID NOs:49-68、84、116-130、147及它们的片段。还提供包含这些重组DNA构建体的转基因烟草植物或其部分、熟化烟草材料或烟草制品。在一些方面,与不含重组DNA构建体的对照烟草植物相比,这些转基因植物、熟化烟草材料或烟草制品包含增加含量的烟碱。进一步提供增加烟草植物的烟碱含量的方法,该方法包含用任何这些重组DNA构建体转化烟草植物。

[0107] 在一些方面,本文公开的重组DNA构建体或表达盒还可以包含用于选择转基因细胞的选择性标记基因。选择性标记基因包括但不限于编码抗生素抗性的基因,诸如编码新霉素磷酸转移酶II (NEO) 和潮霉素磷酸转移酶 (HPT) 的基因,以及赋予对诸如草铵膦、溴苯腈、咪唑啉酮和2,4-二氯苯氧基乙酸酯 (2,4-D) 的除草剂化合物的抗性的基因。额外的选择性标记包括诸如 β -半乳糖苷酶的表型标记和诸如绿色荧光蛋白 (GFP) 的荧光蛋白。

[0108] 在一些方面,本文公开的重组DNA构建体或表达盒包含选自下组的启动子:组成型启动子、诱导型启动子和组织偏好型启动子(例如,叶特异性或根特异性启动子)。示例性的组成型启动子包括美国专利号6,072,050公开的Rsyn7启动子的核心启动子;核心CaMV 35S启动子 (Odell等 (1985) Nature 313:810-812);泛素 (Christensen等 (1989) Plant Mol.Biol.12:619-632和Christensen等 (1992) Plant Mol.Biol.18:675-689);pEMU (Last等 (1991) Theor.Appl.Genet.81:581-588);MAS (Velten等 (1984) EMBO J 3:2723-2730);ALS启动子 (美国专利号5,659,026) 等。示例性的化学诱导型启动子包括通过水杨酸激活的烟草PR-1a启动子。其它目标化学诱导型启动子包括激素敏感型启动子(例如,参见Schna等 (1991) Proc.Natl.Acad.Sci.USA 88:10421-10425和McNellis等 (1998) 植物J.14(2):247-257中公开的糖皮质激素诱导型启动子)和四环素诱导型启动子(例如,参见Gatz等

(1991)Mol.Gen.Genet.227:229-237,以及美国专利号5,814,618和5,789,156)。可用于本文的额外的示例性启动子是负责以下的那些:热调控基因表达、光调控基因表达(例如豌豆rbcS-3A;玉米rbcS启动子;在豌豆中发现的叶绿素a1b结合蛋白基因;或Arabssu启动子)、激素调控的基因表达(例如来自小麦Em基因的脱落酸(ABA)响应序列;ABA诱导型HVA1和HVA22,以及大麦和拟南芥(Arabidopsis)的rd29A启动子)、伤口诱导的基因表达(例如,wun1)和器官特异性基因表达(例如,块茎特异性贮藏蛋白基因;描述的来自玉米的23kDa玉米醇溶蛋白基因;或法国菜豆(β -菜豆蛋白基因),或病原体诱导型启动子(例如,分别为PR-1、prp-1或 β -1,3葡聚糖酶启动子、小麦的真菌诱导型wirla启动子以及线虫诱导型启动子,烟草干燥西芹(tobacco arid parsley)的TobRB7-5A和Hmg-1)。

[0109] 在一些方面,本文提供的烟草植物进一步包含增加的或减少的参与烟碱生物合成或转运的基因的活性表达。参与烟碱生物合成的基因包括但不限于精氨酸脱羧酶(ADC)、甲基腐胺氧化酶(MPO)、NADH脱氢酶、鸟氨酸脱羧酶(ODC)、磷酸核糖基氨基苯甲酸酯异构酶(PRAI)、腐胺N-甲基转移酶(PMT)、喹啉磷酸核糖转移酶(QPT)和S-腺苷甲硫氨酸合成酶(SAMS)。虽然已经提出了两个候选基因(A622和NBB1),但是催化烟酸衍生物和甲基吡咯啉阳离子之间的缩合步骤的烟碱合成酶尚未阐明。参见US 2007/0240728 A1和US 2008/0120737A1。A622编码异黄酮还原酶样蛋白。另外,几种转运蛋白也可能参与烟碱的转运。已经克隆并表征了命名为MATE的转运蛋白基因(Morita等,PNAS 106:2447-52(2009))。

[0110] 在一些方面,与对照烟草植物相比,本文提供的烟草植物进一步包含增加或减少含量的mRNA、蛋白,或编码选自下组的产物的一个或多个基因:PMT、MPO、QPT、ADC、ODC、PRAI、SAMS、BBL、MATE、A622和NBB1。在其它方面,本文提供的烟草植物进一步包含直接抑制编码选自下组的产物的一个或多个基因的表达的转基因:PMT、MPO、QPT、ADC、ODC、PRAI、SAMS、BBL、MATE、A622和NBB1。在其它方面,本文提供的烟草植物进一步包含抑制编码选自下组的产物的一个或多个基因的表达或活性的转基因或突变:PMT、MPO、QPT、ADC、ODC、PRAI、SAMS、BBL、MATE、A622和NBB1。在其它方面,本文提供的烟草植物进一步包含过表达编码选自下组的产物的一个或多个基因的转基因:PMT、MPO、QPT、ADC、ODC、PRAI、SAMS、BBL、MATE、A622和NBB1。

[0111] 本文还公开了使用本领域已知的任何合适的转化方法用本文描述的重组构建体或表达盒转化烟草植物。用于将多核苷酸序列引入烟草植物的方法是本领域已知的,并且包括但不限于稳定转化方法、瞬时转化方法和病毒介导方法。“稳定转化”是指在其中引入植物的目标核苷酸构建体被整合到植物的基因组中并且能够被其后代遗传的转化。“瞬时转化”意指序列被引入植物中,并且仅在植物中暂时表达或仅短暂存在。

[0112] 将多核苷酸引入本公开的植物细胞的合适方法包括显微注射(Crossway等(1986)Biotechniques 4:320-334)、电穿孔(Shillito等(1987)Meth.Enzymol.153:313-336; Riggs等(1986)Proc.Natl.Acad.Sci.USA 83:5602-5606)、农杆菌介导的转化(美国专利号5,104,310、5,149,645、5,177,010、5,231,019、5,463,174、5,464,763、5,469,976、4,762,785、5,004,863、5,159,135、5,563,055和5,981,840)、直接基因转移(Paszkowski等.(1984)EMBO J.3:2717-2722)和弹道粒子加速(例如,参见美国专利号4,945,050、5,141,131、5,886,244、5,879,918和5,932,782;Tomes等.(1995)in Plant Cell,Tissue,和Organ Culture Fundamental Methods,ed.Gamborg和Phillips(Springer-Verlag,Berlin);

McCabe等.(1988) *Biotechnology* 6:923-926)。另参见Weissinger等.(1988) *Ann.Rev.Genet.*22:421-477;Christou等.(1988) *Plant Physiol.*87:671-674(大豆);McCabe等.(1988) *Bio/Technology* 6:923-926(大豆);Finer和McMullen(1991) *In Vitro Cell Dev.Biol.*27P:175-182(大豆);Singh等.(1998) *Theor.Appl.Genet.*96:319-324(大豆);De Wet等.(1985) in *The Experimental Manipulation of Ovule Tissues*, ed.Chapman等.(Longman,N.Y.), pp.197-209(花粉);Kaeppler等.(1990) *Plant Cell Reports* 9:415-418 and Kaeppler等.(1992) *Theor.Appl.Genet.*84:560-566(晶须介导的转化);D'Halluin等.(1992) *Plant Cell* 4:1495-1505(电穿孔)。

[0113] 在其它方面,可以通过使植物与病毒或病毒核酸接触而将本文公开的重组构建体或表达盒引入植物。通常,该方法涉及将本公开的表达盒整合入病毒DNA或RNA分子内。认识到用于本文公开的表达盒的启动子还涵盖通过病毒RNA聚合酶进行转录的启动子。将多核苷酸引入植物并表达其中(涉及病毒DNA或RNA分子)编码的蛋白的方法是本领域已知的。例如,参见美国专利号5,889,191、5,889,190、5,866,785、5,589,367、5,316,931和Porta等.(1996) *Molecular Biotechnology* 5:209-221。

[0114] 可以使用克隆方法随后繁殖的任何植物组织,无论是通过器官发生还是胚胎发生,都可以用本文公开的重组构建体或表达盒转化。“器官发生”意指芽和根从分生组织中心按顺序发育的过程。“胚胎发生”意指无论是来自体细胞还是配子,芽和根以协同方式(不是按顺序)一起发育的过程。适用于本文描述的各种转化方案的示例性组织包括但不限于愈伤组织、现有的分生组织(例如顶端分生组织、腋芽和根分生组织)和诱导的分生组织(例如子叶分生组织和下胚轴分生组织)、下胚轴、子叶、叶盘、花粉、胚等。

[0115] 在一些方面,本文提供的烟草植物来自选自下组的烟草类型:烤制熟化烟草、晾制熟化烟草、深色晾制熟化烟草、深色明火烤制熟化烟草、Galpao烟、和东方烟草。在其它方面,本文提供的烟草植物来自选自下组的烟草类型:白肋烟、马里兰烟和深色烟草。

[0116] 烤制熟化烟草(也称为弗吉尼亚烟或金黄色烟草)占世界烟草产量的约40%。因为在熟化期间其达到金黄色至深橘色,烤制熟化烟草也常常称为“金黄色烟草”。烤制熟化烟草具有淡淡的、明亮的香气和味道。烤制熟化烟草通常糖含量高,油含量低。主要的烤制熟化烟草种植国家有阿根廷、巴西、中国、印度、坦桑尼亚和美国。在一些方面,本文提供的低生物碱或低烟碱烟草植物或种子是选自下组的烤制熟化烟草背景:CC 13、CC 27、CC 33、CC 37、CC 65、CC 67、CC700、GF 318、GL 338、GL 368、GL 939、K 346、K 399、K326、NC102、NC 196、NC 291、NC 297、NC 299、NC 471、NC 55、NC 606、NC 71、NC 72、NC 92、PVH 1118、PVH 1452、PVH 2110、SPEIGHT168、SPEIGHT 220、SPEIGHT 225、SPEIGHT 227、SPEIGHT 236和基本上衍生自上述任一品种的任何品种。在其它方面,本文提供的低生物碱或低烟碱烟草植物或种子是选自下组的烤制熟化烟草背景:Coker 48、Coker 176、Coker 371-Gold、Coker 319、Coker 347、GL 939、K 149、K326、K 340、K 346、K 358、K 394、K 399、K 730、NC 27NF、NC 37NF、NC 55、NC 60、NC 71、NC 72、NC 82、NC 95、NC 297、NC 606、NC 729、NC 2326、McNair 373、McNair 944、Ox 207、Ox 414NF、Reams 126、Reams 713、Reams 744、RG 8、RG 11、RG 13、RG 17、RG 22、RG 81、RG H4、RG H51、Speight H-20、Speight G-28、Speight G-58、Speight G-70、Speight G-108、Speight G-111、Speight G-117、Speight 168、Speight 179、Speight NF-3、Va 116、Va 182和基本上衍生自上述任一品种的任何品种。参见WO

2004/041006 A1。在进一步的方面,本文公开的低生物碱或低烟碱烟草植物、种子、杂交种、品种或系是选自下组的任何烤制熟化烟草背景:K326、K346和NC196。

[0117] 晾制熟化烟草包括白肋烟、马里兰烟和深色烟草。共同因素是熟化主要没有人工来源的热和湿度。白肋烟的颜色是浅棕色至深棕色,油含量高且糖含量低。白肋烟在仓房里晾制熟化。主要的白肋烟种植国家有阿根廷、巴西、意大利、马拉维和美国。马里兰烟极度蓬松,具有良好的燃烧性能、低烟碱和中性的香气。主要的马里兰烟种植国家包括美国和意大利。在一些方面,本文提供的低生物碱或低烟碱烟草植物或种子是选自下组的白肋烟背景:Clay 402、Clay 403、Clay502、Ky 14、Ky 907、Ky 910、Ky 8959、NC 2、NC 3、NC 4、NC 5、NC 2000、Tn 86、Tn 90、Tn 97、R 610、R 630、R 711、R 712、NCBH129、Bu 21×Ky 10、HB04P、Ky 14×L 8、Kt 200、Newton 98、Pedigo561、Pf561和Va 509。在进一步的方面,本文提供的低生物碱或低烟碱烟草植物、种子、杂交种、品种或系是选自下组的白肋背景:TN90、KT 209、KT 206、KT212和HB 4488。在其它方面,本文提供的低生物碱或低烟碱烟草植物或种子是选自下组的马里兰烟背景:Md 10、Md 40、Md 201、Md 609、Md 872和Md 341。

[0118] 深色晾制熟化烟草主要因其熟化工艺而区别于其他类型,所述工艺为深色晾制熟化烟草提供中等棕色至深棕色的颜色和独特的香气与区分。深色晾制熟化烟草主要用于生产嚼烟和鼻烟。在一些方面,本文提供的低生物碱或低烟碱烟草植物或种子是选自下组的深色晾制熟化烟草背景:Sumatra、Jatim、Dominican Cubano、Besuki、One sucker、Green River、Virginia sun-cured和Paraguan Passado。

[0119] 深色明火烤制熟化烟草通常用封闭的熟化室地板上的低燃烧木火熟化。它们的叶糖含量低,但烟碱含量高。深色明火烤制熟化烟草用于制作斗烟、卷烟、嚼烟、鼻烟和味道强烈的雪茄。深色明火烤制熟化烟草的主要种植区域是美国的田纳西、肯塔基和弗吉尼亚。在一些方面,本文提供的低生物碱或低烟碱烟草植物或种子是选自下组的深色明火烤制熟化烟草背景:Narrow Leaf Madole、Improved Madole、Tom Rosson Madole、Newton's VH Madole、Little Crittenden、Green Wood、Little Wood、Small Stalk Black Mammoth、DT 508、DT 518、DT 592、KY 171、DF 911、DF 485、TN D94、TN D950、VA 309和VA 359。

[0120] 东方烟草也称为希腊、香气和土耳其烟草,由于它们通常种植在东地中海区域,诸如土耳其、希腊、保加利亚、马其顿、叙利亚、黎巴嫩、意大利和罗马尼亚。今天的东方品种的特征是植株和叶小,以及它独特的香气特性,这是植物在过去几个世纪中适应贫瘠土壤和恶劣的气候条件的结果。在一些方面,本文提供的低生物碱或低烟碱烟草植物或种子是选自下组的东方烟草背景:Izmir、Katerini、Samsun、Basma和Krumovgrad、Trabzon、Thesalian、Tasova、Sinop、Izmit、Hendek、Edirne、Semdinli、Adiyanman、Yayladag、Iskenderun、Duzce、Macedonian、Mavra、Prilep、Bafra、Bursa、Bucak、Bitlis、Balikesir和基本上衍生自上述任一品种的任何品种。

[0121] 在一些方面,本文提供的低生物碱或低烟碱烟草植物、种子、杂交种、品种或系基本上衍生自或是选自下组的基因背景:BU 64、CC 101、CC 200、CC 27、CC 301、CC 400、CC 500、CC 600、CC 700、CC 800、CC 900、Coker 176、Coker 319、Coker 371Gold、Coker 48、CU 263、DF911、Galpao烟、GL 26H、GL 350、GL 600、GL 737、GL 939、GL 973、HB 04P、K 149、K 326、K 346、K 358、K394、K 399、K 730、KDH 959、KT 200、KT204LC、KY 10、KY 14、KY 160、KY 17、KY 171、KY 907、KY907LC、KTY14 x L8 LC、Little Crittenden、McNair 373、McNair

944、msKY 14xL8、Narrow Leaf Madole、NC 100、NC 102、NC 2000、NC 291、NC 297、NC 299、NC 3、NC 4、NC 5、NC 6、NC7、NC 606、NC 71、NC 72、NC 810、NC BH 129、NC 2002、Neal Smith Madole、OXFORD 207、`Perique`tobacco、PVH03、PVH09、PVH19、PVH50、PVH51、R 610、R 630、R 7-11、R 7-12、RG 17、RG 81、RG H51、RGH 4、RGH 51、RS 1410、Speight 168、Speight 172、Speight 179、Speight 210、Speight 220、Speight 225、Speight 227、Speight234、Speight G-28、Speight G-70、Speight H-6、Speight H20、Speight NF3、TI 1406、TI 1269、TN 86、TN86LC、TN 90、TN 97、TN97LC、TN D94、TN D950、TR (Tom Rosson) Madole、VA 309或VA359、Maryland 609、HB3307PLC、HB4488PLC、KT206LC、KT209LC、KT210LC、KT212LC、R610LC、PVH2310、NC196、KTD14LC、KTD6LC、KTD8LC、PD7302LC、PD7305LC、PD7309LC、PD7318LC、PD7319LC、PD7312LC、ShireyLC或根据本领域已知的标准烟草育种技术的任何商业化烟草品种。

[0122] 所有上述提到的深色晾制熟化、白肋、马里兰、深色明火烤制熟化或东方型的特定品种仅仅是为了示范目的而列出。在本申请中还考虑任何额外的深色晾制熟化、白肋、马里兰、深色明火烤制熟化、东方品种。

[0123] 本文还提供本文所述烟草植物的群体。在一个方面，本文公开的烟草植物的群体的中值密度为每英亩约5,000-约8000、约5,000-约7,600、约5,000-约7,200、约5,000-约6,800、约5,000-约6,400、约5,000-约6,000、约5,000-约5,600、约5,000-约5,200、约5,200-约8,000、约5,600-约8,000、约6,000-约8,000、约6,400-约8,000、约6,800-约8,000、约7,200-约8,000或约7,600-约8,000株植物。另一方面，本文公开的烟草植物的群体种在低至中等肥力的土壤类型中。

[0124] 本文还提供来自本文所述烟草植物的种子的容器。本公开的烟草种子的容器可以容纳任何数量、重量或体积的种子。例如，容器可以容纳至少或超过约100粒、200粒、300粒、400粒、500粒、600粒、700粒、800粒、900粒、1000粒、1500粒、2000粒、2500粒、3000粒、3500粒、4000或更多粒种子。或者，容器可以容纳至少或超过约1盎司、5盎司、10盎司、1磅、2磅、3磅、4磅、5磅或更多的种子。烟草种子的容器可以是本领域可得任何容器。非限制性实施例的方式中，容器可以是盒、袋、小包、小袋、带卷、管或瓶子。

[0125] 本文还提供由本文所述的低生物碱或低烟碱烟草植物制作的熟化烟草材料。进一步提供由本文所述的烟草植物制作的具有高生物碱或烟碱含量的熟化烟草材料。

[0126] “熟化”是陈化过程，其降低水分并导致叶绿素的破坏，从而使烟叶变成金黄色，并且通过它将淀粉转化成糖。因此，与收获的绿叶相比，熟化烟草具有更高的还原糖含量和更低的淀粉含量。在一些方面，本文提供的绿叶烟草可以使用常规手段来熟化，例如烤制熟化、阴干熟化、明火烤制熟化、晾制熟化或晒制熟化。例如，参见Tso (1999, Tobacco, Production, Chemistry and Technology, Davis&Nielsen, eds., Blackwell Publishing, Oxford第1章)，其描述了不同类型的熟化方法。熟化烟草通常在压缩条件下在木桶（例如大桶）或纸箱中陈化几年（例如2-5年），含水量为10%-约25%。参见美国专利号4,516,590和5,372,149。然后，可以对熟化的并且陈化的烟草进行进一步加工。进一步加工包括在真空中引入或不引入各种温度的蒸汽调节烟草、巴氏灭菌和发酵。通常，发酵的特征是初始含水量高、发热、以及干重减少10-20%。例如，参见美国专利号4,528,993、4,660,577、4,848,373、5,372,149；美国公开号2005/0178398；和Tso (1999, Tobacco, Production, Chemistry

and Technology, Davis&Nielsen, eds., Blackwell Publishing, Oxford第1章)。可以对熟化、陈化并发酵的烟草进行进一步加工(例如切丝、切碎、膨化或混合)。例如,参见美国专利号4,528,993;4,660,577;和4,987,907。在一个方面,本公开的熟化烟草材料是晒制熟化的。在另一方面,本公开的熟化烟草材料是烤制熟化、晾制熟化或明火烤制熟化。

[0127] 从本公开的烟草系、品种或杂交种获得的烟草材料可用于制作烟草制品。如本文所用,“烟草制品”定义为旨在供人类使用或消费的由烟草制作或衍生自烟草的任何制品。

[0128] 本文提供的烟草制品包括但不限于香烟制品(例如香烟和比迪烟)、雪茄制品(例如雪茄包装烟和小雪茄)、斗烟制品、衍生自烟草的制品、来自烟草的烟碱制品、无烟烟草制品(例如湿鼻烟、干鼻烟和嚼烟)、膜、咀嚼片、标签、成形部分、凝胶、消费单元、不溶性基质、中空形状、再造烟草、膨胀烟草等。例如,参见美国专利公开号US 2006/0191548。

[0129] 如本文所用,“香烟”是指具有“棒”和“填料”的烟草制品。香烟“棒”包括香烟纸、过滤嘴、包装纸(用于容纳过滤材料)、将香烟纸(包括填料)固定在过滤嘴上的接装纸以及将这些部件固定在一起的所有胶水。“填料”包括(1)所有烟草,包括但不限于再造的并膨胀的烟草;(2)非烟草替代物(包括但不限于草药、非烟草植物材料和可以卷入香烟纸内伴随烟草的其他香料);(3)加料;(4)调味剂;(5)所有其他添加剂(其混入烟草和替代物中并卷成香烟)。

[0130] 如本文所用,“再造烟草”是指由烟草粉末和其它烟草废料制成的烟草填料的一部分,其加工成片状并切成条状以与烟草相像。除了节省成本之外,再造烟草有助于使用氨和糖之间的反应来加工风味开发的香烟味道,因此非常重要。

[0131] 如本文所用,“膨胀烟草”是指经过合适气体的膨胀加工的烟草填料的一部分,以便使烟草“膨化”,从而导致密度降低和填充容量更大。它减少了香烟中使用的烟草重量。

[0132] 衍生自本公开植物的烟草制品还包括香烟和其它吸烟用品,特别是那些包括滤嘴元件的吸烟用品,其中抽烟材料的棒包括在烟草混合物内的熟化烟草。在一个方面,本公开的烟草制品选自下组:小雪茄、不通气过滤嘴香烟、通气过滤嘴香烟、雪茄、鼻烟、烟斗烟、雪茄烟、卷烟、嚼烟、叶烟、水烟、烟碎和烟丝。在另一方面,本公开的烟草制品是无烟烟草制品。无烟烟草制品不燃烧,并且包括但不限于嚼烟、湿无烟烟草、湿鼻烟和干鼻烟。嚼烟是粗分烟叶,其通常包装在大的袋状包装中,并用在塞或烟草卷中。湿无烟烟草是一种潮湿的更细分的烟草,其以散装形式或袋装形式提供并通常包装在圆形罐中,用作夹或用在置于成年烟草消费者的脸颊和牙龈之间的小袋中。湿鼻烟是一种热处理的无烟烟草。干鼻烟是放在口中或用于鼻腔的磨细的烟草。在进一步的方面,本公开的烟草制品选自下组:散叶嚼烟、塞嚼烟(plug chewing tobacco)、湿鼻烟和鼻用鼻烟(nasal snuff)。在另一方面,本公开的烟草制品选自下组:电子加热的香烟、电子香烟、电子蒸发装置。

[0133] 在一个方面,本公开的烟草制品可以是混合烟草制品。在另一个方面,本公开的烟草制品可以是低烟碱烟草制品。在进一步的方面,本公开的烟草制品可以包含含量少于约3mg/g的降烟碱。例如,该制品中的降烟碱含量为3.0mg/g、2.5mg/g、2.0mg/g、1.5mg/g、1.0mg/g、750μg/g、500pg/g、250pg/g、100pg/g、75pg/g、50pg/g、25pg/g、10pg/g、7.0pg/g、5.0pg/g、4.0pg/g、2.0pg/g、1.0pg/g、0.5pg/g、0.4pg/g、0.2pg/g、0.1pg/g、0.05pg/g、0.01pg/g或不可检测。

[0134] 在一些方面,基于干重,本文提供的熟化烟草材料或烟草制品包含选自下组的平

均烟碱含量:约0.01%、0.02%、0.05%、0.75%、0.1%、0.15%、0.2%、0.3%、0.35%、0.4%、0.5%、0.6%、0.7%、0.8%、0.9%、1%、1.1%、1.2%、1.3%、1.4%、1.5%、1.6%、1.7%、1.8%、1.9%、2%、2.1%、2.2%、2.3%、2.4%、2.5%、2.6%、2.7%、2.8%、2.9%、3%、3.1%、3.2%、3.3%、3.4%、3.5%、3.6%、3.7%、3.8%、3.9%、4%、5%、6%、7%、8%和9%。在其它方面,基于干重,本文提供的熟化烟草材料或烟草制品包含选自下组的平均烟碱含量:约0.01%-0.02%、0.02%-0.05%、0.05%-0.75%、0.75%-0.1%、0.1%-0.15%、0.15%-0.2%、0.2%-0.3%、0.3%-0.35%、0.35%-0.4%、0.4%-0.5%、0.5%-0.6%、0.6%-0.7%、0.7%-0.8%、0.8%-0.9%、0.9%-1%、1%-1.1%、1.1%-1.2%、1.2%-1.3%、1.3%-1.4%、1.4%-1.5%、1.5%-1.6%、1.6%-1.7%、1.7%-1.8%、1.8%-1.9%、1.9%-2%、2%-2.1%、2.1%-2.2%、2.2%-2.3%、2.3%-2.4%、2.4%-2.5%、2.5%-2.6%、2.6%-2.7%、2.7%-2.8%、2.8%-2.9%、2.9%-3%、3%-3.1%、3.1%-3.2%、3.2%-3.3%、3.3%-3.4%、3.4%-3.5%和3.5%-3.6%。在进一步的方面,基于干重,本文提供的熟化烟草材料或烟草制品包含选自下组的平均烟碱含量:约0.01%-0.1%、0.02%-0.2%、0.03%-0.3%、0.04%-0.4%、0.05%-0.5%、0.75%-1%、0.1%-1.5%、0.15%-2%、0.2%-3%和0.3%-3.5%。

[0135] 本公开还提供用于育种包含所需含量的总生物碱或烟碱(例如低烟碱或无烟碱)的烟草系、栽培种或品种的方法。育种可以通过任何已知的程序进行。可以在标记-辅助选择(MAS)育种项目中使用DNA指纹图谱、SNP作图、单倍型作图或相似技术以将所需的性状或等位基因转移或育种到烟草植物中。例如,育种者可以使用本文公开的F1杂交植物在F2代或回交世代中生成分离群体,或者进一步将F1杂交种植物与其他具有农艺上所需基因型的供体植物杂交。可以使用本领域已知的技术或本文列出的技术之一筛选F2或回交世代的植物的所需农艺性状或所需化学特征。取决于预期的遗传模式或使用的MAS技术,可以在每个回交周期之前对所选植物进行自花授粉,以帮助鉴定所需的单个植株。回交或其他育种程序可重复进行,直到轮回亲本的所需表型恢复。本公开中的轮回亲本可以是烤制熟化品种、白肋品种、深色晾制熟化品种、深色明火烤制熟化品种或东方品种。例如,其它育种技术可以在Wernsman, E.A., and Rufty, R.C. 1987. Chapter Seventeen. Tobacco. Pages 669-698 In: Cultivar Development. Crop Species. W.H. Fehr (ed.), MacMillan Publishing Co., Inc., New York, N.Y. 中找到,其全部内容通过引用并入本文。

[0136] 使用本文所述的烟草植物的植物育种程序的结果包括本公开的有用的系、栽培种、品种、后代、近交系和杂交种。如本文所用,术语“品种”是指具有将其与相同物种的其他植物分离的恒定特征的植物群体。品种通常是,尽管并不总是,商业化销售的。尽管具有一个或多个区别性状,品种的进一步特征是在该品种内的个体之间的总体差异非常小。可以通过数代自花授粉和选择,或使用组织或细胞培养技术从单个亲本无性繁殖生成“纯系”品种。品种可以基本上衍生自另一个系或品种。根据国际公约对植物新品种保护所定义的(1961年12月2日,于1972年11月10日、1978年10月23日和1991年3月19日在日内瓦修订),品种“基本上衍生”自初始品种,如果:a)它主要衍生自初始品种,或衍生自主要衍生自初始品种的品种,同时保留由所述初始品种的基因型或基因型组合产生的必要特征的表达;b)它与初始品种有明显区别;c)除了由衍生行为产生的差异,它在由所述初始品种的基因型或基因型组合导致的必要特征的表达中与初始品种相符合。例如,可通过选择天然或诱导的

突变体、体细胞无性变异、来自初始品种植物的变异个体、回交或转化获得基本上衍生的品种。认为第一烟草品种和第二烟草品种(第一烟草品种基本衍生自第二烟草品种)具有基本相同的遗传背景。与品种不同,“系”最通常是指非商业化使用,例如用于植物研究的一组植物。系通常在个体间的一个或多个目标性状显示非常小的总体差异,尽管在个体间的其他性状可能存在一些差异。

[0137] 在一些方面,本公开提供将低烟碱性状导入烟草品种的方法,该方法包括:(a) 将包含低烟碱性状的第一烟草品种与不具有低烟碱性状的第二烟草品种杂交以生产一个或多个后代烟草植物;(b) 对所述一个或多个后代烟草植物进行基因分型检测与低烟碱性状连锁的多态性标记,其中该多态性标记位于侧翼为表3列出的任何两个多态性位点或侧翼为表4列出的任何两个多态性位点的染色体区间中;和(c) 选择包含低烟碱性状的后代烟草植物。在其它方面,这些方法进一步包含将选择的后代烟草植物与第二烟草品种回交。在进一步的方面,这些方法进一步包含(d) 将选择的后代植物与其自身或与第二烟草品种杂交以产生一个或多个进一步的后代烟草植物;和(e) 选择包含低烟碱性状的进一步的后代烟草植物。在一些方面,选择步骤(e) 包含标记-辅助选择。在一些方面,这些方法产生包含低烟碱性状的单基因转化。在一些方面,这些方法产生包含Nic1导入的单基因转化。在某些方面,第二烟草品种是优良品种。在其它方面,这些方法的基因分型步骤涉及一个或多个分子标记测定。在其它方面,该方法使用的多态性标记包含选下组的多态性:单核苷酸多态性(SNP)、DNA序列中的插入或缺失(Indel)、DNA序列的简单序列重复(SSR)、限制性片段长度多态性(RFLP) 和标签SNP。在其它方面,与LA白肋21相比,所选后代烟草植物在Nic1基因座处包含更短的染色体缺失。

[0138] 在其它方面,本公开提供将低烟碱性状导入烟草品种的方法,该方法包括:(a) 将包含低烟碱性状的第一烟草品种与不具有低烟碱性状的第二烟草品种杂交以生产一个或多个后代烟草植物;(b) 对所述一个或多个后代烟草植物进行基因分型检测与低烟碱性状连锁的多态性标记,其中所述多态性标记在表3和4中列出的任何一个多态性位点的20cM内;和(c) 选择包含低烟碱性状的后代烟草植物。在一些方面,该方法包含同时或一并选择与Nic1基因座相关联或紧密连锁的一种或多种分子标记以及与Nic2基因座相关联或紧密连锁的一种或多种分子标记。

[0139] 在一些方面,本公开提供选择具有低烟碱性状的烟草植物的方法,该方法包含:(a) 从烟草种质集合中分离核酸;(b) 测定核酸检测与Nic1基因座紧密连锁的一个或多个标记;和(c) 基于标记测定,选择具有低烟碱性状的烟草植物。在一些方面,测定的与Nic1基因座紧密连锁的一个或多个标记在表3列出的任何一个多态性位点的约20cM、10cM、5cM、4cM、3cM、2cM、1cM、0.5cM或小于0.5cM内。在其它方面,该方法进一步包含测定与Nic2基因座紧密连锁的一个或多个标记。在一些方面,测定的与Nic2基因座紧密连锁的一个或多个标记在表4列出的任何一个多态性位点的约20cM、10cM、5cM、4cM、3cM、2cM、1cM、0.5cM或小于0.5cM内。在一些方面,该方法进一步包含确定所选植物的烟碱含量以确认低烟碱性状。

[0140] 本公开还提供将低烟碱性状导入烟草品种的方法,该方法包含:(a) 将包含低烟碱性状的第一烟草品种与不具有低烟碱性状的第二烟草品种杂交以生产一个或多个后代烟草植物;(b) 对所述一个或多个后代烟草植物进行基因分型检测与低烟碱性状连锁的多态性标记,其中所述多态性标记位于侧翼为表4列出的任何两个多态性位点的染色体区间中;

和(c)选择包含低烟碱性状的后代烟草植物。在一些方面,这些方法产生包含低烟碱性状的单基因转化。在一些方面,这些方法产生包含Nic2基因导入的单基因转化。在一些方面,第二烟草品种是优良品种。在其它方面,这些方法的基因分型步骤涉及一个或多个分子标记测定。在其它方面,该方法使用的多态性标记包含选自下组的多态性:单核苷酸多态性(SNP)、DNA序列中的插入或缺失(Indel)、DNA序列的简单序列重复(SSR)、限制性片段长度多态性(RFLP)和标签SNP。在其它方面,与LA白肋21相比,所选后代烟草植物在Nic2基因座处包含更短的染色体缺失。

[0141] 如本文所用,“基因座”是多态性核酸、性状决定因子、基因或标记所在的染色体区域。本公开的基因座包含群体中的一种或多种多态性;例如,在一些个体中存在替代等位基因。如本文所用,“等位基因”是指特定基因座处的替代核酸序列。等位基因的长度可以小至1个核苷酸碱基,但通常更大。例如,第一等位基因可以出现在一条染色体上,而第二等位基因出现在第二同源染色体上,例如,如出现在杂合子个体的不同染色体上,或者在群体中不同的纯合子或杂合子个体间出现。如本文所用,术语“染色体区间”是指位于单一染色体上的基因组DNA的连续线性跨度。

[0142] 如本文所用,厘摩(“cM”)是重组频率的测量单位。一cM等于1%的概率,由于只跨越一代,一个遗传基因座上的标记会与第二基因座上的标记分离。可以使用Kosambi函数(Kosambi, The estimation of map distances from recombination values. *Annals of Eugenics*, 12:172-75 (1944))从重组值计算本文涉及的遗传距离。

[0143] 如本文所用,“与...紧密连锁”或“与...相关联”是指标记或基因座在另一个标记或基因座的约20cM、10cM、5cM、1cM、0.5cM或小于0.5cM内。例如,20cM意味着标记与基因座之间的重组频率等于或小于约20%。

[0144] 如本文所用,“导入(introgression)”或“导入(introgress)”是指将遗传基因座的所需等位基因从一个遗传背景传递到另一个遗传背景。

[0145] 如本文所用,“杂交(crossed)”或“杂交(cross)”是指通过受精(例如细胞、种子或植物)产生后代,并且包括植物(有性)和自我受精(自交)之间的杂交。

[0146] 如本文所用,“回交(backcross)”和“回交(backcrossing)”是指后代植物被反复杂交回其亲本之一的过程。在回交方案中,“供体”亲本指具有要导入的所需基因或基因座的亲本植物。“受体”亲本(使用一次或多次)或“轮回”亲本(使用两次或更多次)是指基因或基因座被导入其中的亲本植物。初始杂交产生F1代。术语“BC1”是指轮回亲本的第二次使用,“BC2”是指轮回亲本的第三次使用,等等。在一些方面,反复进行回交,每个连续回交世代的后代个体本身回交至相同的亲本基因型。

[0147] 如本文所用,“单基因转化(single gene converted)”或“单基因转化(single gene conversion)”是指使用已知为回交的植物育种技术或通过基因工程开发的植物,其中除了通过回交技术或通过遗传工程转移到品种的单基因之外,恢复了品种基本上全部的所需形态和生理特征。

[0148] 如本文所用,“优良品种”意指由育种和选择得到的农艺学表现优异的任何品种。

[0149] 如本文所用,在标记-辅助选择或育种的上下文中的“选择(selecting)”或“选择(selection)”是指基于某些预定标准,一般从群体中挑选或选取所需个体的行为。

[0150] 如本文所用,术语“性状”是指可被基因型影响的细胞或生物体的一种或多种可检

测特征。表型可以通过肉眼,或通过本领域已知的任何其他评估手段来观察,例如显微镜、生化分析、遗传分析、特定疾病耐受性测定等。在一些情况下,表型直接由单个基因或遗传基因座控制,例如“单基因性状”。在其他情况下,表型是几个基因的结果。

[0151] 如本文所用,“标记测定”意指使用特定方法检测特定基因座的多态性的方法,例如测量至少一种表型(诸如种子颜色、花色或其它视觉可检测的性状)、限制性片段长度多态性(RFLP)、单碱基延伸、电泳、序列比对、等位基因特异性寡核苷酸杂交(ASO)、随机扩增多态性DNA(RAPD)、基于微阵列的技术和核酸测序技术等。

[0152] 如本文所用,“标记-辅助选择”(MAS)是基于标记基因型选择表型的过程。“标记-辅助选择育种”是指通过检测植物中的一种或多种核酸选择一株植物或多株植物中所需的一个或多个性状,其中核酸与所需性状连锁,然后选择具有那些一种或多种核酸的植物或种质。

[0153] 如本文所用,“多态性”是指群体中一种或多种变异的存在。多态性可以表现为核酸的核苷酸序列的变异或蛋白的氨基酸序列的变异。多态性包括在一个或多个个体的群体中的一个或多个位点处存在的核酸序列或核酸特征的一个或多个变异。该变异可以包括但不限于一个或多个核苷酸碱基的改变、一个或多个核苷酸的插入或一个或多个核苷酸的缺失。多态性可以通过突变从核酸复制的随机过程产生作为移动基因组元件的结果、从拷贝数变异产生以及在减数分裂过程期间(诸如不等位交杂、基因组复制以及染色体断裂和融合)产生。这种变异可以在群体内普遍发现或可能以低频率存在,前者在一般的植物育种中具有更大的效用,而后者可能与罕见但重要的表型变异相关。有用的多态性可以包括单核苷酸多态性(SNP)、DNA序列中的插入或缺失(Indel)、DNA序列的简单序列重复(SSR)、限制性片段长度多态性(RFLP)和标签SNP。基因标记、基因、DNA衍生序列、RNA衍生序列、启动子、基因的5'非翻译区、基因的3'非翻译区、微RNA、siRNA、耐受基因座、卫星标记、转基因、mRNA、ds mRNA、转录谱和甲基化模式也可以包含多态性。另外,前述拷贝数的存在、不存在或变异可以包含多态性。

[0154] 如本文所用,“SNP”或“单核苷酸多态性”意指当基因组序列中的单个核苷酸(A、T、C或G)被改变或可变时发生的序列变异。当SNP被映射到基因组上的位点时,存在“SNP标记”。

[0155] 如本文所用,“标记”或“分子标记”或“标记基因座”是用于表示足够独特以表征基因组上的特定基因座的核酸或氨基酸序列的术语。任何可检测的多态性性状都可以用作标记,只要它是差异性遗传的并且与目标表型性状表现出连锁不平衡。因而,每个标记都指示DNA特定片段,具有独特的核苷酸序列。地图位置提供特定标记彼此之间相对位置的测量。当性状描述为与给定的标记连锁时,应当理解,其序列影响性状的DNA片段通常与标记共分离。如果在性状的两侧都鉴定了标记,则可以获得更精确且更确切的性状定位。通过测量杂交后代中标记的出现,可以通过相对简单的分子测试来检测性状的存在,而不需要实际评估性状本身的出现(其可能是困难的且耗时的),因为实际评估性状需要植物生长到可表达性状的阶段和/或环境条件下。在一些方面,使用本文公开的Nic1或Nic2基因座,本文所用的标记的LOD得分为2或更高、3或更高、4或更高、5或更高、6或更高、7或更高、8或更高、或9或更高,其使用本领域已知的方法(诸如Qgene Version 2.23(1996))和默认参数进行测量。

[0156] 在一个方面,本公开提供烟草植物或其部分,其包含侧翼为并且不包含Nic1标记Nos.1-207中的任何两个的第一染色体缺失,侧翼为并且不包含Nic2标记Nos.1-340中的任何两个的第二染色体缺失,以及所述第一和所述第二染色体缺失两者,其中当在相似的生长条件下生长时,所述烟草植物能够生产USDA等级指数值为50或更高、55或更高、60或更高、65或更高、70或更高、75或更高、80或更高、85或更高、90或更高、95或更高的叶。在另一方面,烟草植物包含选自下组的烟碱含量:低于3%、低于2.75%、低于2.5%、低于2.25%、低于2.0%、低于1.75%、低于1.5%、低于1.25%、低于1%、低于0.9%、低于0.8%、低于0.7%、低于0.6%、低于0.5%、低于0.4%、低于0.3%、低于0.2%、低于0.1%、低于0.05%。在一个方面,进一步提供本段落中的烟草植物群体、由其制成的熟化烟草材料、包含所述熟化烟草材料的烟草混合物和包含所述熟化烟草材料的烟草制品。

[0157] 在另一方面,本公开提供烟草植物或其部分,其包含选自下组的突变:侧翼为并且不包含Nic1标记Nos.1-207中的任何两个的第一染色体缺失,侧翼为并且不包含Nic2标记Nos.1-340中的任何两个的第二染色体缺失,以及所述第一和所述第二染色体缺失两者,其中当在相似的生长条件下生长时,所述烟草植物能够生产USDA等级指数值与对照植物的USDA等级指数值相当的叶,其中所述对照植物与所述烟草植物具有除所述突变之外基本上相同的遗传背景。在一个方面,当在相似的生长条件下生长时,烟草植物包含的烟碱含量低于所述对照植物的烟碱含量的1%、2%、5%、8%、10%、12%、15%、20%、25%、30%、40%、50%、60%、70%或80%。在一个方面,进一步提供本段落中的烟草植物群体、由其制成的熟化烟草材料、包含所述熟化烟草材料的烟草混合物和包含所述熟化烟草材料的烟草制品。

[0158] 在一个方面,第一染色体缺失的侧翼为并且不包含Nic1标记Nos.1-20、21-40、41-60、61-80、81-100、101-120、121-140、141-160、161-180、181-200或201-207中的任何两个。在另一个方面,第一染色体缺失的侧翼为并且不包含Nic1标记Nos.1-10、11-20、21-30、31-40、41-50、51-60、61-70、71-80、81-90、91-100、101-110、111-120、121-130、131-140、141-150、151-160、161-170、171-180、181-190、191-200或201-207中的任何两个。

[0159] 在一个方面,第二染色体缺失的侧翼为并且不包含Nic2标记Nos.1-20、21-40、41-60、61-80、81-100、101-120、121-140、141-160、161-180、181-200、201-220、221-240、241-260、261-280、281-300、301-320或321-340中的任何两个。在另一个方面,第二染色体缺失的侧翼为并且不包含Nic2标记Nos.1-10、11-20、21-30、31-40、41-50、51-60、61-70、71-80、81-90、91-100、101-110、111-120、121-130、131-140、141-150、151-160、161-170、171-180、181-190、191-200、201-210、211-220、221-230、231-240、241-250、251-260、261-270、271-280、281-290、291-300、301-310、311-320、321-330或331-340中的任何两个。

[0160] 应当理解,本公开的任何烟草植物可以进一步包含额外的农艺上所需的性状,例如,通过使用本领域已知的技术用遗传构建体或转基因进行转化。非限制性地,所需性状的实例是除草剂抗性、抗虫性、抗病性;高产;高等级指数值;可熟化性;熟化质量;机械收获性能;耐熟能力;叶质量;高度、植物成熟(例如早熟、早熟到中熟、中熟、中晚熟或晚熟);茎的大小(例如小的、中等的或大的茎);或每株植物的叶数(例如少的(例如5-10叶)、中等的(例如11-15叶)或多的(例如16-21)叶数)或任意组合。在一些方面,本文公开的低烟碱或无烟碱的烟草植物或种子包含表达一种或多种杀虫蛋白的一种或多种转基因,所述杀虫蛋白例如苏云金杆菌(*Bacillus thuringiensis*)的晶体蛋白或来自蜡样芽孢杆菌(*Bacillus*

cereus)的营养型杀虫蛋白,诸如VIP3(例如,参见Estruch等.(1997)Nat.Biotechnol.15:137)。在其它方面,本文公开的烟草植物进一步包含赋予对棕色茎腐病的抗性(美国专利号5,689,035)或对抗囊胞线虫的抗性(美国专利号5,491,081)的导入性状。

[0161] 本公开还提供烟草植物,其包含改变的烟碱含量,但其产量与没有该改变的烟碱含量的相应的初始烟草植物的产量相当。在一个方面,本文公开的低烟碱或无烟碱烟草品种提供选自下组的产量:约1200-3500、1300-3400、1400-3300、1500-3200、1600-3100、1700-3000、1800-2900、1900-2800、2000-2700、2100-2600、2200-2500和2300-2400lbs/英亩。在另一个方面,本文公开的低烟碱或无烟碱烟草品种提供选自下组的产量:约1200-3500、1300-3500、1400-3500、1500-3500、1600-3500、1700-3500、1800-3500、1900-3500、2000-3500、2100-3500、2200-3500、2300-3500、2400-3500、2500-3500、2600-3500、2700-3500、2800-3500、2900-3500、3000-3500、3100-3500lbs/英亩。在进一步的方面,本文公开的低烟碱或无烟碱的烟草植物提供的产量为对照植物的产量的65%-130%、70%-130%、75%-130%、80%-130%、85%-130%、90%-130%、95%-130%、100%-130%、105%-130%、110%-130%、115%-130%、120%-130%,所述对照植物除Nic1突变、Nic2突变、Nic1转基因、Nic2转基因或其组合之外具有基本相同的遗传背景。在进一步的方面,本文公开的低烟碱或无烟碱的烟草植物提供的产量为对照植物的产量的70%-125%、75%-120%、80%-115%、85%-110%或90%-100%,所述对照植物除Nic1突变、Nic2突变、Nic1转基因、Nic2转基因或其组合之外具有基本相同的遗传背景。

[0162] 在一个方面,本文公开的低烟碱或无烟碱烟草品种适用于机器收获。在另一个方面,本文公开的低烟碱或无烟碱烟草品种通过机械收获。

[0163] 在一些方面,本文提供的烟草植物是杂交种植物。可以通过以下来产生杂交种:防止第一品种的母本植物(例如种子亲本)的自花授粉,允许来自第二品种的父本植物的花粉使母本植物受精,并且允许在雌株上形成F1杂交种子。可以通过在花发育的早期阶段将花去雄防止雌株的自花授粉。或者,可以使用雄性不育的形式防止母本植物的花粉形成。例如,雄性不育可由雄性不育(MS)、或转基因雄性不育(其中转基因抑制小孢子发生和/或花粉形成)或自交不亲和性产生。含有MS的母本植物是特别有用的。当母本植物是MS时,可以从雄性可育植物收获花粉并人工施用到MS母本植物的柱头上,并收获所得的F1种子。

[0164] 植物可以用于形成单交烟草F1杂交种。将来自父本植物的花粉人工转移到去雄的母本植物或雄性不育的母本植物以形成F1种子。或者,可以进行三交,其中将单交F1杂交种用作母本,并与不同的父本杂交。作为另一种选择,可以生成双交杂交种,其中两个不同单交的F1后代本身杂交。特别有利地,自交不亲和性可在形成双交杂交种时用于防止母本的自花授粉。

[0165] 在一个方面,本文公开的低烟碱或无烟碱烟草品种是雄性不育的。在另一个方面,本文公开的低烟碱或无烟碱烟草品种是细胞质雄性不育的。雄性不育的烟草植物可以通过本领域已知的任何方法产生。产生雄性不育烟草的方法描述于Wernsman,E.A.,and Rufty,R.C.1987.Chapter Seventeen.Tobacco.Pages 669-698In:Cultivar Development.Crop Species.W.H.Fehr(ed.),MacMillan Publishing Co.,Inc.,New York,N.Y.761pp。

[0166] 在进一步的方面,本文提供的烟草部分包括但不限于叶、茎、根、种子、花、花粉、花药、胚珠、花梗、果实、分生组织、子叶、下胚轴、荚果、胚、胚乳、外植体、愈伤组织、组织培养

物、嫩芽、细胞和原生质体。在一个方面,本文提供的烟草部分不包括种子。在一个方面,本公开提供不是繁殖材料并且不介导植物的天然繁殖的烟草植物细胞、组织和器官。另一方面,本公开还提供作为繁殖材料并且介导植物的天然繁殖的烟草植物细胞、组织和器官。另一方面,本公开提供不能通过光合作用维持自身的烟草植物细胞、组织和器官。另一方面,本公开提供体细胞烟草植物细胞。与繁殖细胞相反,体细胞不介导植物繁殖。

[0167] 所提供的细胞、组织和器官可以来自种子、果实、叶、子叶、下胚轴、分生组织、胚、胚乳、根、嫩芽、茎、荚果、花、花序、杆、花梗、花柱、柱头、花托、花瓣、花萼、花粉、花药、花丝、子房、胚珠、果皮、韧皮部、维管组织。在另一个方面,本公开提供烟草植物叶绿体。在进一步的方面,本公开提供表皮细胞、气孔细胞、叶毛或根毛、储藏根或块茎。在另一个方面,本公开提供烟草原生质体。

[0168] 本领域技术人员理解,烟草植物通过种子自然繁殖,而不是通过无性繁殖或营养繁殖。在一个方面,本公开提供烟草胚乳。在另一个方面,本公开提供烟草胚乳细胞。在进一步的方面,本公开提供雄性或雌性不育烟草植物,其在无人工干预的情况下不能繁殖。

[0169] 在一些方面,本公开提供包含与选自下组的序列具有至少约40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的核酸分子:SEQ ID NOs:9-48、75-82、86-15和131-146及它们的片段。在其它方面,本公开提供包含与选自下组的序列具有至少约40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的核酸分子:SEQ ID NOs:13、28、33、48、82、86、87、101、102、145、146及它们的片段。在一些方面,本公开提供包含与选自下组的氨基酸序列具有至少约40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的多肽或蛋白:SEQ ID NOs:49-68、116-130和147。在一些方面,本公开提供包含与选自下组的氨基酸序列具有至少约40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列相似性的多肽或蛋白:SEQ ID NOs:49-68、116-130和147。在其它方面,本公开提供具有选自下组的氨基酸序列的蛋白的生物活性变体:SEQ ID NOs:49-68、116-130和147。本公开的蛋白的生物活性变体可与该蛋白只相差1-15个氨基酸残基、只相差10个、9个、8个、7个、6个、5个、4个、3个、2个或1个氨基酸残基。本文还提供来自Nic1基因座的基因或蛋白的直系同源基因或蛋白。“直系同源”是衍生自共同祖先基因的基因,并且由于物种形成而在不同物种中发现。在核苷酸序列和/或蛋白序列水平上,直系同源可以具有至少60%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的序列同一性或相似性。直系同源的功能通常在物种间高度保守。

[0170] 如本文所用,在两个多核苷酸或多肽序列的上下文中的术语“序列同一性”或“同一性”是指当在特定的比较窗口上比对最大一致性时,两个序列中相同的残基。当序列同一性百分比用于指蛋白时,公认不相同的残基位置经常因保守的氨基酸取代而不同,其中氨基酸残基被具有相似化学性能(例如电荷或疏水性)的其他氨基酸残基取代,因而不会改变分子的功能特性。当序列在保守取代中不同时,可以上调序列同一性百分比以校正取代的保守性质。通过这种保守取代而不同的序列被认为具有“序列相似性”或“相似性”。

[0171] 本文提供的核酸分子、多肽或蛋白可以是分离的或实质上纯化的。“分离的”或“纯

化的”核酸分子、多肽、蛋白或其生物活性部分实质上或基本上不含一般伴随在其天然存在的环境中发现的多核苷酸或蛋白或与之相互作用的组分。例如,当通过重组技术产生时,分离的或纯化的多核苷酸或蛋白实质上不含其他细胞材料或培养基,或当化学合成时,实质上不含化学前体或其他化学品。

[0172] 本公开进一步提供从本文公开的烟草植物制造包含烟草材料的烟草制品的方法。在一些方面,本文公开的方法包含调节由本文公开的烟草植物制备的陈化烟草材料,以将其含水量从约12.5%-约13.5%增加至约21%,混合经调节的烟草材料以生产所需混合物。在一个方面,本文公开的制造烟草制品的方法进一步包含对混合物进行加料或调味。通常,在加料过程中,向混合物添加加料或调味材料,以通过平衡化学组成来提高质量,并开发某些所需的风味特性。关于加料过程的进一步细节可以在Tobacco Production, Chemistry and Technology, Edited by L.Davis and M.Nielsen, Blackwell Science, 1999中找到。

[0173] 本文提供的烟草材料还可以使用包括但不限于热处理(例如烹饪、烘烤)、调味、酶处理、膨胀和/或熟化的方法加工。可以使用这些技术来加工发酵的和未发酵的烟草。合适的加工烟草的实例包括深色晾制熟化、深色明火烤制熟化、白肋、烤制熟化和雪茄填料或包装物,以及来自全叶填塞操作的制品。在一些方面,以鲜重计,烟草纤维包括高达70%的深色烟草。例如,如美国公开号2004/0118422或2005/0178398中所述,烟草可通过加热、发汗和/或巴氏消毒步骤进行调节。

[0174] 本文提供的烟草材料可以发酵。通常,发酵的特征在于初始含水量高、发热、干重减少10-20%。例如,参见美国专利号4,528,993、4,660,577、4,848,373和5,372,149。除了改变叶的香气外,发酵还可以改变叶的颜色、质地或两者。在发酵过程期间,可以产生逸出气体,可以吸收氧气,可以改变pH,并且可以改变保留的水量。例如,参见美国公开号2005/0178398和Tso (1999, Chapter 1 in Tobacco, Production, Chemistry and Technology, Davis&Nielsen, eds., Blackwell Publishing, Oxford)。熟化的、或熟化并且发酵的烟草可以在引入口腔制品之前进一步加工(例如切丝、膨化、混合、研磨或粉碎)。在一些情况下,与共聚物和任选的调味剂和其它添加剂混合之前,烟草是烘箱挥发物含量为48-50重量%的长切丝发酵熟化湿烟草。

[0175] 在一些方面,可以将本文提供的烟草材料加工成所需的大小。在某些方面,可以将烟草纤维加工成平均纤维大小小于200微米。在一些方面,烟草纤维为75-125微米。在其它方面,将烟草纤维加工成具有75微米或更小的大小。在一些方面,烟草纤维包括长切烟丝,其可以被切丝或切碎成约10切丝刀数/英寸-约110切丝刀数/英寸的宽度和约0.1英寸-约1英寸的长度。双切丝烟纤维可以具有一定范围的颗粒大小,使得双切丝烟纤维的约70%落入-20目至80目的网目大小中。

[0176] 可以将本文提供的烟草材料加工成具有约10重量%或更大;约20重量%或更大;约40重量%或更大;约15重量%-约25重量%;约20重量%-约30重量%;约30重量%-约50重量%;约45重量%-约65重量%或约50重量%-约60重量%的总烘箱挥发物含量。本领域技术人员将会理解,“湿”烟草通常是指烘箱挥发物含量为约40重量%-约60重量%(例如,约45重量%-约55重量%、或约50重量%)的烟草。如本文所用,通过计算样品在预热的强力通风烘箱中、110℃下干燥3.25小时之后的重量损失百分比来确定“烘箱挥发物”。与用于制作口腔制品的烟草纤维的烘箱挥发物含量相比,口腔制品可具有不同的总体烘箱挥发物含

量。本文所述的加工步骤可以减少或增加烘箱挥发物含量。

[0177] 以下段落提供示例性实施方案的列表。

[0178] 实施方案1:烟草植物或其部分,其包含Nic1基因座中的突变、Nic2基因座中的突变或两者,其中所述烟草植物能够生产USDA等级指数值为50或更高的叶。

[0179] 实施方案2:实施方案1的烟草植物或其部分,其中所述烟草植物能够生产具有选自下组的USDA等级指数值的叶:55或更高、60或更高、65或更高、70或更高、75或更高、80或更高、85或更高、90或更高、95或更高。

[0180] 实施方案3:实施方案1的烟草植物或其部分,其中当在相似的生长条件下生长时,所述烟草植物能够生产USDA等级指数值与对照植物的USDA等级指数值相当的叶,其中所述对照植物与所述烟草植物具有除所述突变之外基本上相同的遗传背景。

[0181] 实施方案4:实施方案1的烟草植物或其部分,其中当在相似的生长条件下生长时,所述烟草植物能够生产USDA等级指数值为对照植物的USDA等级指数值的至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%或至少约98%的叶,其中所述对照植物与所述烟草植物具有除所述突变之外基本上相同的遗传背景。

[0182] 实施方案6:实施方案1的烟草植物或其部分,其中所述烟草植物包含选自下组的烟碱含量:低于3%、低于2.75%、低于2.5%、低于2.25%、低于2.0%、低于1.75%、低于1.5%、低于1.25%、低于1%、低于0.9%、低于0.8%、低于0.7%、低于0.6%、低于0.5%、低于0.4%、低于0.3%、低于0.2%、低于0.1%、低于0.05%。

[0183] 实施方案7:实施方案1的烟草植物或其部分,其中所述烟草植物进一步包含直接抑制编码选自下组的产物的一个或多个基因的表达或活性的转基因或突变:PMT、MPO、QPT、BBL、A622和MATE转运蛋白。

[0184] 实施方案8:烟草植物或其部分,其包含Nic1基因座中的突变、Nic2基因座中的突变或两者,其中当在相似的生长条件下生长时,所述烟草植物能够生产USDA等级指数值与对照植物的USDA等级指数值相当的叶,其中所述对照植物与所述烟草植物具有除所述突变之外基本上相同的遗传背景。

[0185] 实施方案9:实施方案8的烟草植物或其部分,其中所述烟草植物能够生产具有选自下组的USDA等级指数值的叶:55或更高、60或更高、65或更高、70或更高、75或更高、80或更高、85或更高、90或更高、95或更高。

[0186] 实施方案10:实施方案8的烟草植物或其部分,其中所述烟草植物能够生产USDA等级指数值为所述对照植物的等级指数值的至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%或至少约98%的叶。

[0187] 实施方案11:实施方案8的烟草植物或其部分,其中当在相似的生长条件下生长时,所述烟草植物包含的烟碱含量低于所述对照植物的烟碱含量的1%、2%、5%、8%、10%、12%、15%、20%、25%、30%、40%、50%、60%、70%或80%。

[0188] 实施方案12:实施方案8的烟草植物或其部分,其中所述烟草植物进一步包含直接抑制编码选自下组的产物的一个或多个基因的表达或活性的转基因或突变:PMT、MPO、QPT、BBL、A622和MATE转运蛋白。

[0189] 实施方案13:烟草品种植物,其包含Nic1基因座中的突变、Nic2基因座中的突变或两者,其中当在相似的生长条件下生长时,所述烟草品种的叶等级指数与对照烟草品种的

叶等级指数相当,其中所述对照烟草品种与所述烟草品种具有除所述突变之外基本上相同的遗传背景。

[0190] 实施方案14:非转基因烟草植物或其部分,其包含选自下组的烟碱含量:低于3%、低于2.75%、低于2.5%、低于2.25%、低于2.0%、低于1.75%、低于1.5%、低于1.25%、低于1%、低于0.9%、低于0.8%、低于0.7%、低于0.6%、低于0.5%、低于0.4%、低于0.3%、低于0.2%、低于0.1%以及低于0.05%,其中所述烟草植物能够生产USDA等级指数值为50或更高、55或更高、60或更高、65或更高、70或更高、75或更高、80或更高、85或更高、90或更高以及95或更高的叶。

[0191] 实施方案15:权利要求14的非转基因烟草植物或其部分,其中所述非转基因烟草植物包含低于2.0%的烟碱含量并且能够生产USDA等级指数值为70或更高的叶。

[0192] 实施方案16:权利要求14的非转基因烟草植物或其部分,其中所述所述非转基因烟草植物包含低于1.0%的烟碱含量并且能够生产USDA等级指数值为70或更高的叶。

[0193] 实施方案17:烟草植物或其部分,其包含非转基因突变,其中当在相似的生长条件下生长时,所述非转基因突变将所述烟草植物的烟碱含量降低至低于对照植物的烟碱含量的1%、2%、5%、8%、10%、12%、15%、20%、25%、30%、40%、50%、60%、70%或80%,其中所述烟草植物能够生产USDA等级指数值与所述对照植物的USDA等级指数值相当的叶,并且其中所述对照植物与所述烟草植物具有除所述非转基因突变之外基本上相同的遗传背景。

[0194] 实施方案18:实施方案1-17中任一项的烟草植物的群体。

[0195] 实施方案19:来自权利要求1-17中任一项的烟草植物的熟化烟草材料。

[0196] 实施方案20:实施方案19的熟化烟草材料,其中所述熟化烟草材料通过选自下组的熟化过程制成:烤制熟化、晾制熟化、明火烤制熟化和晒制熟化。

[0197] 实施方案21:烟草混合物,其包含权利要求19的熟化烟草材料。

[0198] 实施方案22:实施方案21的烟草混合物,其中按重量计,所述熟化烟草材料在所述烟草混合物中占熟化烟草的至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%或至少95%。

[0199] 实施方案23:实施方案21的烟草混合物,其中按体积计,该熟化烟草材料在所述烟草混合物中占熟化烟草的至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%或至少95%。

[0200] 实施方案24:烟草制品,其包含权利要求19的熟化烟草材料。

[0201] 实施方案25:实施方案24的烟草制品,其中该烟草制品选自下组:香烟、小雪茄、不通气过滤嘴香烟、通气过滤嘴香烟、雪茄、鼻烟、烟斗烟、雪茄烟、卷烟、嚼烟、叶烟、烟碎和烟丝。

[0202] 实施方案26:实施方案24的烟草制品,其中该烟草制品是无烟烟草制品。

[0203] 实施方案27:实施方案26的烟草制品,其中该无烟烟草制品选自下组:散叶嚼烟、塞嚼烟、湿鼻烟和鼻用鼻烟。

[0204] 实施方案28:再造烟草,其包含权利要求19的熟化烟草材料。

[0205] 实施方案29:烟草植物或其部分,其包含Nic1基因座中的突变,其中所述突变在LA白肋21品种中不存在。

[0206] 实施方案30:实施方案29的烟草植物或其部分,其中与所述LA白肋21品种相比,所述烟草植物在Nic1基因座处包含较短的染色体缺失。

[0207] 实施方案31:实施方案29的烟草植物或其部分,其中当在相似的生长条件下生长时,与不含所述突变的对照烟草植物相比,所述烟草植物包含更低的烟碱含量。

[0208] 实施方案32:实施方案29的烟草植物或其部分,其中当在相似的生长条件下生长时,所述烟草植物包含的烟碱含量低于不含所述突变的对照烟草植物的烟碱含量的1%、2%、5%、8%、10%、12%、15%、20%、25%、30%、40%、50%、60%、70%或80%。

[0209] 实施方案33:实施方案29的烟草植物或其部分,其中当在相似的生长条件下生长时,与不含所述突变的对照烟草植物相比,所述烟草植物包含更低的总生物碱含量。

[0210] 实施方案34:实施方案29的烟草植物或其部分,其中当在相似的生长条件下生长时,与不含所述突变的对照烟草植物相比,所述烟草植物包含更低含量的选自下组的一个或多个生物碱:烟碱、降烟碱、假木贼碱和新烟草碱。

[0211] 实施方案35:实施方案29的烟草植物或其部分,其中当在相似的生长条件下生长时,与不含所述突变的对照烟草植物相比,所述烟草植物包含相似含量的选自下组的一个或多个化合物:3-甲基戊酸、戊酸、异戊酸、labdenoid、西松烷、糖酯和还原糖。

[0212] 实施方案36:实施方案29的烟草植物或其部分,其中所述突变是纯合的。

[0213] 实施方案37:实施方案29的烟草植物或其部分,其中所述突变是杂合的。

[0214] 实施方案38:实施方案29的烟草植物或其部分,其中所述突变选自下组:点突变、缺失、插入、复制和逆转。

[0215] 实施方案39:实施方案29的烟草植物或其部分,其中所述突变通过选自下组的方法引入:随机诱变和定点诱变。

[0216] 实施方案40:实施方案39的烟草植物或其部分,其中所述定点诱变通过大范围核酸酶、锌指核酸酶、TALEN或CRISPR介导。

[0217] 实施方案41:实施方案29的烟草植物或其部分,其中所述烟草植物进一步包含Nic2基因座中的突变。

[0218] 实施方案42:实施方案29的烟草植物或其部分,其中所述突变位于一个基因内,该基因包含与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%同一性的序列:SEQ ID NOs:9-28、75-82、86-100、145及它们的片段。

[0219] 实施方案43:实施方案29的烟草植物或其部分,其中所述突变降低所述基因的表达或活性。

[0220] 实施方案44:实施方案29的烟草植物或其部分,其中所述突变位于一个基因内,该基因包含与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%同一性的序列:SEQ ID NOs:13、28、33、48、82、86、87、101、102、145、146及它们的片段。

[0221] 实施方案45:实施方案29的烟草植物或其部分,其中所述突变位于与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%同一性的序列内:SEQ ID NOs:13、28、33、48、82、86、87、101、102、145、146及它们的片段。

[0222] 实施方案46:实施方案29的烟草植物或其部分,其中所述突变位于一个基因内,该基因包含与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%同一性的编码序列:SEQ ID NOs:33、48、101、102、146及它们的片段。

[0223] 实施方案47:实施方案29的烟草植物或其部分,其中所述突变位于与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%同一性的序列内:SEQ ID NOs:33、48、101、102、146及它们的片段。

[0224] 实施方案48:实施方案29的烟草植物或其部分,其中所述突变位于编码多肽的基因内,该多肽与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%同一性:SEQ ID NOs:53、68、116、117、147及它们的片段。

[0225] 实施方案49:实施方案29的烟草植物或其部分,其中当在相似的生长条件下生长时,与不含所述突变的烟草植物相比,所述植物进一步包含减少含量的编码选自下组的产物的一种或多种基因的mRNA、蛋白、或两者:PMT、MPO、QPT、BBL、MATE和A622。

[0226] 实施方案50:实施方案29的烟草植物或其部分,其中所述植物进一步包含抑制编码选自下组的产物的一种或多种基因的表达或活性的转基因或突变:PMT、MPO、QPT、BBL、A622和MATE转运蛋白。

[0227] 实施方案51:实施方案29的烟草植物或其部分,其中所述烟草植物是杂交种。

[0228] 实施方案52:实施方案29的烟草植物或其部分,其中所述部分选自下组:叶、茎、根、种子、花、花粉、花药、胚珠、花梗、果实、分生组织、子叶、下胚轴、荚果、胚、胚乳、外植体、愈伤组织、组织培养物、嫩芽、细胞和原生质体。

[0229] 实施方案53:实施方案29的烟草植物或其部分,其中所述烟草植物来自选自下组的品种:烤制熟化烟草、晾制熟化烟草、深色明火烤制熟化烟草、Galpao烟和东方烟草。

[0230] 实施方案54:实施方案29的烟草植物或其部分,其中所述烟草植物来自选自下组的品种:白肋烟、马里兰烟和深色晾制熟化烟草。

[0231] 实施方案55:实施方案29的烟草植物的群体。

[0232] 实施方案56:来自实施方案29的烟草植物的熟化烟草材料。

[0233] 实施方案57:实施方案56的熟化烟草材料,其中与来自不含所述突变的对照烟草植物的熟化烟草材料相比,所述熟化烟草材料包含更低含量的烟碱。

[0234] 实施方案58:实施方案56的熟化烟草材料,其中所述烟草植物包含的烟碱含量为0.2%-0.6%。

[0235] 实施方案59:实施方案56的熟化烟草材料,其中所述烟草植物包含的烟碱含量为1.0%-3.0%。

[0236] 实施方案60:实施方案56的熟化烟草材料,其中所述熟化烟草材料通过选自下组的熟化过程制成:烤制熟化、晾制熟化、明火烤制熟化和晒制熟化。

[0237] 实施方案61:烟草混合物,其包含权利要求56的熟化烟草材料。

[0238] 实施方案62:烟草制品,其包含权利要求56的熟化烟草材料。

[0239] 实施方案63:实施方案62的烟草制品,其中该烟草制品选自下组:香烟、小雪茄、不通气过滤嘴香烟、通气过滤嘴香烟、雪茄、鼻烟、烟斗烟、雪茄烟、卷烟、嚼烟、叶烟、烟碎和烟丝。

[0240] 实施方案64:实施方案62的烟草制品,其中该烟草制品是无烟烟草制品。

[0241] 实施方案65:实施方案64的烟草制品,其中该无烟烟草制品选自下组:散叶嚼烟、塞嚼烟、湿鼻烟和鼻用鼻烟。

[0242] 实施方案66:再造烟草,其包含权利要求56的熟化烟草材料。

[0243] 实施方案67:重组DNA构建体,其包含在烟草细胞中功能性的并且与编码多肽的多核苷酸可操作连接的启动子,该多肽的氨基酸序列与选自下组的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%同一性:SEQ ID NOs:49-68、84、116-130、147及它们的片段。

[0244] 实施方案68:烟草植物或其部分,其包含权利要求67的重组DNA构建体。

[0245] 实施方案69:实施方案68的烟草植物或其部分,其中与不含所述重组DNA构建体的对照烟草植物相比,所述烟草植物包含更高含量的烟碱。

[0246] 实施方案70:来自实施方案68的烟草植物的熟化烟草材料。

[0247] 实施方案71:烟草制品,其包含权利要求70的熟化烟草材料。

[0248] 实施方案72:增加烟草植物的烟碱含量的方法,所述方法包含用权利要求67的重组DNA构建体转化烟草植物。

[0249] 实施方案73:重组DNA构建体,其包含在烟草细胞中功能性的并且与编码RNA分子的多核苷酸可操作连接的启动子,该RNA分子能够与编码多肽的RNA结合,该多肽的氨基酸序列与选自下组的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%同一性:SEQ ID NOs:49-68、84、116-130、147及它们的片段,并且其中所述RNA分子抑制所述多肽的表达。

[0250] 实施方案74:烟草植物或其部分,其包含权利要求73的重组DNA构建体。

[0251] 实施方案75:实施方案74的烟草植物或其部分,其中所述RNA分子选自下组:微RNA、siRNA和反式作用siRNA。

[0252] 实施方案76:实施方案74的烟草植物或其部分,其中所述多核苷酸编码双链RNA。

[0253] 实施方案77:实施方案74的烟草植物或其部分,其中与不含所述重组DNA构建体的对照烟草植物相比,所述烟草植物包含更低含量的烟碱。

[0254] 实施方案78:来自实施方案74的烟草植物的熟化烟草材料。

[0255] 实施方案79:烟草制品,其包含权利要求78的熟化烟草材料。

[0256] 实施方案80:降低烟草植物的烟碱含量的方法,所述方法包含用权利要求73的重组DNA构建体转化烟草植物。

[0257] 实施方案81:烟草植物或其部分,其包含异源表达盒,该异源表达盒包含基因的Nic1抑制序列,该基因包含选自下组的序列:SEQ ID NOs:9-28、75-82、86-100、145及它们的片段,其中所述抑制序列与在植物细胞中功能性的启动子可操作连接,并且其中所述抑制序列与选自以下序列的至少21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、31个、32个、33个、34个、35个、36个、37个、38个、39个、40个、41个、42个、43个、44个、45个、46个、47个、48个、49个、50个、51个、52个、53个、54个、55个、56个、57个、58个、59个、60个、61个、62个、63个、64个、65个、66个、67个、68个、69个、70个、71个、72个、73个、74个、75个、76个、77个、78个、79个或80个核苷酸的片段具有至少90%序列同一性:SEQ ID NOs:9-28、75-82、86-100、145及它们的片段。

[0258] 实施方案82:实施方案81的烟草植物或其部分,其中所述Nic1抑制序列能够转录

为选自下组的抑制多核苷酸：单链RNA多核苷酸、双链RNA多核苷酸及它们的组合。

[0259] 实施方案83：实施方案81的烟草植物或其部分，其中所述启动子选自下组：组成型启动子、诱导型启动子和组织偏好型启动子。

[0260] 实施方案84：实施方案81的烟草植物或其部分，其中所述启动子是根特异性启动子。

[0261] 实施方案85：烟草植物或其部分，其包含异源表达盒，该异源表达盒包含基因的Nic1抑制序列，该基因包含选自下组的序列：SEQ ID NOs：13、28、33、48、82、86、87、101、102、145、146及它们的片段，其中所述抑制序列与在植物细胞中功能性的启动子可操作连接，并且其中所述抑制序列与选自以下序列的至少21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、31个、32个、33个、34个、35个、36个、37个、38个、39个、40个、41个、42个、43个、44个、45个、46个、47个、48个、49个、50个、51个、52个、53个、54个、55个、56个、57个、58个、59个、60个、61个、62个、63个、64个、65个、66个、67个、68个、69个、70个、71个、72个、73个、74个、75个、76个、77个、78个、79个或80个核苷酸的片段具有至少90%序列同一性：SEQ ID NOs：13、28、33、48、82、86、87、101、102、145、146及它们的片段。

[0262] 实施方案86：将低烟碱性状导入烟草品种的方法，所述方法包含：

[0263] (a) 将包含低烟碱性状的第一烟草品种与不具有所述低烟碱性状的第二烟草品种杂交以生产一个或多个后代烟草植物；

[0264] (b) 对所述一个或多个后代烟草植物进行基因分型检测与低烟碱性状连锁的多态性标记，其中该多态性标记位于侧翼为表3列出的任何两个多态性位点或侧翼为表4列出的任何两个多态性位点的染色体区间中；和

[0265] (c) 选择包含所述低烟碱性状的后代烟草植物。

[0266] 实施方案87：实施方案86的方法，其中该方法进一步包含将所述选择的后代烟草植物与所述第二烟草品种回交。

[0267] 实施方案88：实施方案86的方法，其中该方法进一步包含：

[0268] (d) 将选择的后代植物与其自身或与第二烟草品种杂交以产生一个或多个进一步的后代烟草植物；和

[0269] (e) 选择包含所述低烟碱性状的进一步的后代烟草植物。

[0270] 实施方案89：实施方案88的方法，其中选择步骤(e)包含标记-辅助选择。

[0271] 实施方案90：实施方案86的方法，其中该方法产生包含所述低烟碱性状的单基因转化。

[0272] 实施方案91：实施方案86的方法，其中该第二烟草品种是优良品种。

[0273] 实施方案92：实施方案86的方法，其中该基因分型涉及一个或多个分子标记测定。

[0274] 实施方案93：实施方案86的方法，其中该多态性标记包含选自下组的多态性：单核苷酸多态性(SNP)、DNA序列中的插入或缺失(Indel)、DNA序列的简单序列重复(SSR)、限制性片段长度多态性(RFLP)和标签SNP。

[0275] 实施方案94：实施方案86的方法，其中该基因分型包含测定位于与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%同一性的序列内的核酸序列的存在或不存在：SEQ ID NOs：9-28、75-82、86-100、145及它们的片段。

[0276] 实施方案95：实施方案86的方法，其中该基因分型包含测定位于与选自下组的序

列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%同一性的序列内的核酸序列的存在或不存在:SEQ ID Nos:13、28、33、48、82、86、87、101、102、145、146及它们的片段。

[0277] 实施方案96:实施方案86的方法,其中该第一烟草品种是LA白肋21。

[0278] 实施方案97:实施方案86的方法,其中与LA白肋21相比,选择的后代烟草植物在Nic1基因座处包含较短的染色体缺失。

[0279] 实施方案98:将低烟碱性状导入烟草品种的方法,所述方法包含:

[0280] a.将包含低烟碱性状的第一烟草品种与不具有低烟碱性状的第二烟草品种杂交以生产一个或多个后代烟草植物;

[0281] b.对所述一个或多个后代烟草植物进行基因分型检测与低烟碱性状连锁的多态性标记,其中所述多态性标记位于表3和表4列出的任何一个多态性位点的20cM内,或为SEQ ID Nos.131-144中的任何一个;和

[0282] c.选择包含所述低烟碱性状的后代烟草植物。

[0283] 实施方案99:实施方案98的方法,其中该基因分型包含测定位于与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%同一性的序列内的核酸序列的存在或不存在:SEQ ID Nos:9-28、75-82、86-100、145及它们的片段。

[0284] 实施方案100:实施方案98的方法,其中该基因分型包含测定位于与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%同一性的序列内的核酸序列的存在或不存在:SEQ ID Nos:13、28、33、48、82、86、87、101、102、145、146及它们的片段。

[0285] 实施方案101:选择具有低烟碱性状的烟草植物的方法,所述方法包含:

[0286] a.从烟草种质集合中分离核酸;

[0287] b.测定核酸检测与Nic1基因座或Nic2基因座紧密连锁的一个或多个标记;和

[0288] c.基于标记测定,选择具有低烟碱性状的烟草植物。

[0289] 实施方案102:实施方案101的方法,其中所述一个或多个标记在表3和表4列出的任何一个多态性位点的约20cM、10cM、5cM、4cM、3cM、2cM、1cM、0.5cM或小于0.5cM内,或为SEQ ID Nos.131-144中的任何一个。

[0290] 实施方案103:实施方案101的方法,其中该测定包含测定位于与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%同一性的序列内的核酸序列的存在或不存在:SEQ ID Nos:9-28、75-82、86-100、145及它们的片段。

[0291] 实施方案104:实施方案101的方法,其中该测定包含测定位于与选自下组的序列具有至少80%、至少85%、至少90%、至少95%、至少97%、至少98%、至少99%同一性的序列内的核酸序列的存在或不存在:SEQ ID Nos:13、28、33、48、82、86、87、101、102、145、146及它们的片段。

[0292] 实施方案105:实施方案101的方法,其中该方法进一步包括确定所述选择的植物的烟碱含量以确认所述低烟碱性状。

[0293] 实施方案106:实施方案101的方法,其中烟草种质集合是单倍体育种群体。

[0294] 实施方案107:烟草植物或其部分,其包含侧翼为并且不包含Nic1标记No.1-207中的任何两个的第一染色体缺失,侧翼为并且不包含Nic2标记No.1-340中的任何两个的第二

染色体缺失,以及所述第一和所述第二染色体缺失两者,其中所述烟草植物能够生产USDA等级指数值为50或更高、55或更高、60或更高、65或更高、70或更高、75或更高、80或更高、85或更高、90或更高、或95或更高的叶。

[0295] 实施方案108:实施方案107的烟草植物或其部分,其中所述烟草植物包含选自下组的烟碱含量:低于3%、低于2.75%、低于2.5%、低于2.25%、低于2%、低于1.75%、低于1.5%、低于1.25%、低于1%、低于0.9%、低于0.8%、低于0.7%、低于0.6%、低于0.5%、低于0.4%、低于0.3%、低于0.2%、低于0.1%、低于0.05%。

[0296] 实施方案109:实施方案107的烟草植物或其部分,其中所述第一染色体缺失的侧翼为并且不包含Nic1标记No.1-20、21-40、41-60、61-80、81-100、101-120、121-140、141-160、161-180、181-200、或201-207中的任何两个。

[0297] 实施方案110:实施方案107的烟草植物或其部分,其中所述第一染色体缺失的侧翼为并且不包含Nic1标记Nos.1-10、11-20、21-30、31-40、41-50、51-60、61-70、71-80、81-90、91-100、101-110、111-120、121-130、131-140、141-150、151-160、161-170、171-180、181-190、191-200、或201-207中的任何两个。

[0298] 实施方案111:实施方案107的烟草植物或其部分,其中所述第二染色体缺失的侧翼为并且不包含Nic2标记Nos.1-20、21-40、41-60、61-80、81-100、101-120、121-140、141-160、161-180、181-200、201-220、221-240、241-260、261-280、281-300、301-320或321-340中的任何两个。

[0299] 实施方案112:实施方案107的烟草植物或其部分,其中所述第二染色体缺失的侧翼为并且不包含Nic2标记Nos.1-10、11-20、21-30、31-40、41-50、51-60、61-70、71-80、81-90、91-100、101-110、111-120、121-130、131-140、141-150、151-160、161-170、171-180、181-190、191-200、201-210、211-220、221-230、231-240、241-250、251-260、261-270、271-280、281-290、291-300、301-310、311-320、321-330或331-340中的任何两个。

[0300] 实施方案113:烟草植物或其部分,其包含选自下组的突变:侧翼为并且不包含Nic1标记Nos.1-207中的任何两个的第一染色体缺失、侧翼为并且不包含Nic2标记Nos.1-340中的任何两个的第二染色体缺失、以及所述第一和所述第二染色体缺失两者,其中当在相似的生长条件下生长时,所述烟草植物能够生产USDA等级指数值与对照植物的USDA等级指数值相当的叶,其中所述对照植物与所述烟草植物具有除所述突变之外基本上相同的遗传背景。

[0301] 实施方案114:实施方案113的烟草植物或其部分,其中当在相似的生长条件下生长时,所述烟草植物包含的烟碱含量低于对照植物的烟碱含量的1%、2%、5%、8%、10%、12%、15%、20%、25%、30%、40%、50%、60%、70%或80%。

[0302] 实施方案115:实施方案107-114中任一项的烟草植物的群体。

[0303] 实施方案116:来自权利要求107-114中任一项的烟草植物的熟化烟草材料。

[0304] 实施方案117:实施方案116的熟化烟草材料,其中所述熟化烟草材料通过选自下组的熟化过程制成:烤制熟化、晾制熟化、明火烤制熟化和晒制熟化。

[0305] 实施方案118:烟草混合物,其包含权利要求116的熟化烟草材料。

[0306] 实施方案119:实施方案118的烟草混合物,其中按重量计,该熟化烟草材料在所述烟草混合物中占熟化烟草的约至少10%、至少15%、至少20%、至少25%、至少30%、至少

35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%或至少95%。

[0307] 实施方案120:实施方案118的烟草混合物,其中按体积计,该熟化烟草材料在所述烟草混合物中占熟化烟草的约至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%或至少95%。

[0308] 实施方式121:烟草制品,其包含权利要求116的熟化烟草材料。

[0309] 实施方案122:实施方案121的烟草制品,其中该烟草制品选自下组:香烟、小雪茄、不透气过滤嘴香烟、透气过滤嘴香烟、雪茄、鼻烟、烟斗烟、雪茄烟、卷烟、嚼烟、叶烟、烟碎、和烟丝。

[0310] 实施方案123:实施方案121的烟草制品,其中该无烟烟草制品选自下组:散叶嚼烟、塞嚼烟、湿鼻烟和鼻用鼻烟。

[0311] 现在已经概括地描述了本公开,其通过参照以下实施例将更容易理解,除非特别说明,所述实施例以示例的方式提供,而不旨在限制本公开。

实施例

[0312] 实施例1:具有低生物碱的烟草系的全基因组测序

[0313] 使用全基因组测序来确定低生物碱(LA)白肋21中的nic1突变的潜在基因损伤。对四个烟草系进行测序。它们是LA白肋21(nic1nic2,基于干重,平均烟碱~0.3%(范围~0.2-0.6%))、低中级(LI)白肋21(nic1 Nic2,平均烟碱~2.3%(范围~1.5-3.0%))、高中级(HI)白肋21(Nic1 nic2,平均烟碱~3.7%(范围~2.5-5.0%))和野生型白肋21(也称为“BU21”)(Nic1 Nic2,平均烟碱~4.7%(范围~4.0-6.0%))。LA白肋21(也称为“LA BU21”)是一个低总生物碱系,其通过将古巴雪茄品种的低生物碱基因通过几次回交整合入白肋21中产生(Legg等.1970)。

[0314] 从烟叶制备基因组DNA样品。从各自属于BU21和LA BU21,以及高中级和低中级(分别为HI BU21和LI BU21)的三株植物收集绿色叶组织。将组织在液氮中快速冷冻并储存在-80℃。将储存的组织在液氮中研磨,并将来自四个系中的每一个的三株植物的等量组织合并在一起(~10g总最终重量)用于测序。

[0315] 按照标准工业方案进行DNA文库制备和测序,其包括例如剪切DNA以生成平均插入大小为500bp的成对末端文库。将原始序列映射到专有的TN90基因组,并且使用全基因组变体检测流水线来鉴定单核苷酸多态性(SNP)以及插入和缺失(InDel)。变体检测流水线基于序列质量对原始测序数据进行修整和过滤,然后选择最小长度为75bp和未知碱基最大数量为2的读长。然后使用软件gaMap v2.0.0(BETA)将这些读长映射到Altria的专有烟草(TN90)基因组。使用软件gaVariant v.2.0.0BETA将只有质量得分至少为40的映射读长进一步用于SNP和InDel检测。只有最小变异质量为37、基因型质量为85并且覆盖度为7x的SNP和InDel才被进一步分析。表1提供四个系中每一个的双端测序读长、过滤读长、映射读长以及映射读长百分比。4个品种的平均覆盖度为22-40x,其超过用于变体检测的推荐覆盖度(大约20x)。表2提供在四个系中的每一个中检测到的SNP和InDel的数量的细节。

[0316] 表1:四个测序的白肋系的全基因组测序统计。

[0317]	样品	原始读长	过滤读长	映射读长	%过滤	%映射
	BU21	1,030,266,578	987,988,031	815,146,282	95.9	82.51
	HI BU21	928,747,952	889,038,519	738,942,909	95.72	83.12
	LI BU21	838,709,526	800,400,171	663,625,600	95.43	82.91
	LA BU21	792,837,196	755,706,369	621,019,698	95.32	82.18
	总计	3,590,561,252	3,433,133,090	2,838,734,489	95.59	82.68

[0318] 表2:四个测序的白肋系的基因分型统计。

[0319]	样品	SNPs	插入	缺失	%SNP	%插入	%缺失
	BU21	2,349,123	304,814	397,514	76.984	9.98915	13.02705
	HI BU21	2,253,395	292,635	386,493	76.842	9.97895	13.17954
	LI BU21	2,138,321	276,921	371,838	76.723	9.93588	13.34149
	LA BU21	2,083,729	265,561	358,655	76.949	9.80674	13.24455
	总计	8,824,568	1,139,931	1,514,500	76.8745	9.92768	13.19816

[0320] 实施例2:分析基因组序列以鉴定Nic1和Nic2基因座。

[0321] 分析来自实施例1的基因组序列以鉴定LA BU21中的nic1和nic2突变。首先分析Nic2基因座的序列。Shoji等(2010)先前报道了Nic 2基因座包含7个乙烯反应因子(ERF)基因的缺失。这些缺失的ERF基因被映射到TN90基因组的单个连续区域。我们对BU21、HI BU21、LI BU21和LA BU21的测序数据进行了扫描,并鉴定出该基因组区域中的变体图谱由BU21和LI BU21的纯合参考(TN90)等位基因型,以及HI BU21和LA BU21的丢失数据组成。这表明Nic2缺失在变体基因型中表现为丢失数据。

[0322] 先前的文献表明,Nic1和Nic2基因座可能含有各自源自祖先物种(绒毛状烟草或林烟草)之一的重复基因(Hibi等,1994)。此外,Shoji等(2010)报道,Nic2基因座衍生自绒毛状烟草,并且nic2是缺失的。它们起源的相似性表明nic1也可能是缺失的。编写定制的perl脚本来检测BU21和HI BU21中纯合或杂合的参考等位基因的模式,以及LI BU21和LA BU21中的丢失数据。定制的perl脚本共检测到总计14035个拼接序列(scaffold),其中总计近37916个变异位点与LA BU21和LI BU21中的丢失数据的模式以及HI BU21和BU21中的纯合基因型相匹配。每个拼接序列观察到的变体的平均数量接近3(2.701)。一个连续区或拼接序列(Scaffold0002504,SEQ ID No:1)含有具有该模式的207个变体位点(表3)。由此可见,该异常拼接序列中具有该特定模式的变体数量是剩余拼接序列的变体的平均数量的76.67倍。Scaffold0002504为Nic1基因座的位置提供第一个指示。

[0323] 基于拼接序列之间的读长覆盖,没有观察到测序误差来解释LI BU21和LA BU21中读长的不存在。此外,在已知的Nic2基因座(Scaffold0000549,SEQ ID No:2)处,观察到相似的模式,其中LA BU21和HI BU21显示丢失数据,并且BU21和LI BU21显示纯合的参考等位基因呼叫(与呈现该变异位点模式的所有拼接序列中观察到的2.49个位点相比,Scaffold0000549上有340个位点)(表4)。

[0324] 表3:Nic1缺失片段(NT1.0-Scaffold0002504)中的多态性位点。POS表示Scaffold0002504上每个多态性位点的核苷酸位置。REF是指参考TN90等位基因的序列。ALT是指K326、Narrow Leaf Madole或东方型中发现的序列多态性。“0/0”是指参考TN90等位基因纯合,而“./.”是指丢失数据(例如缺失)。

[0325]

Nic1 标 记号	POS	REF	ALT	BU21	HI	LI	LA
1	240	T	A	0/0	0/0	/.	/.
2	1750	CA	C	0/0	0/0	/.	/.
3	2064	G	A	0/0	0/0	/.	/.
4	2671	G	A	0/0	0/0	/.	/.
5	2678	G	T	0/0	0/0	/.	/.
6	5174	G	T	0/0	0/0	/.	/.
7	5176	A	C	0/0	0/0	/.	/.
8	5781	C	T	0/0	0/0	/.	/.
9	5812	A	G	0/0	0/0	/.	/.
10	6312	G	T	0/0	0/0	/.	/.
11	6349	T	G	0/0	0/0	/.	/.
12	6373	T	C	0/0	0/0	/.	/.
13	8028	T	G	0/0	0/0	/.	/.
14	12004	AT	A	0/0	0/0	/.	/.
15	12668	T	C	0/0	0/0	/.	/.
16	12939	A	G	0/0	0/0	/.	/.
17	13171	G	T	0/0	0/0	/.	/.
18	13644	G	T	0/0	0/0	/.	/.
19	14078	G	A	0/0	0/0	/.	/.
20	14085	T	C	0/0	0/0	/.	/.
21	29126	G	T	0/0	0/0	/.	/.
22	46969	G	A	0/0	0/0	/.	/.
23	54856	C	T	0/0	0/0	/.	/.
24	54923	T	C	0/0	0/0	/.	/.
25	63743	T	C	0/0	0/0	/.	/.
26	69020	G	T	0/0	0/0	/.	/.
27	78215	A	G	0/0	0/0	/.	/.
28	87817	T	G	0/0	0/0	/.	/.
29	91989	A	G	0/0	0/0	/.	/.
30	92936	C	G	0/0	0/0	/.	/.
31	104894	G	T	0/0	0/0	/.	/.
32	108471	A	G	0/0	0/0	/.	/.

[0326]

Nic1 标 记号	POS	REF	ALT	BU21	HI	LI	LA
33	108492	G	A	0/0	0/0	./.	./.
34	108503	A	C	0/0	0/0	./.	./.
35	109907	A	C	0/0	0/0	./.	./.
36	112193	A	T	0/0	0/0	./.	./.
37	114778	C	T	0/0	0/0	./.	./.
38	114852	C	T	0/0	0/0	./.	./.
39	117178	A	G	0/0	0/0	./.	./.
40	123634	G	A	0/0	0/0	./.	./.
41	130508	A	G	0/0	0/0	./.	./.
42	130966	T	C	0/0	0/0	./.	./.
43	131000	C	T	0/0	0/0	./.	./.
44	131045	C	A	0/0	0/0	./.	./.
45	134007	A	G	0/0	0/0	./.	./.
46	134046	A	G	0/0	0/0	./.	./.
47	136601	C	T	0/0	0/0	./.	./.
48	136884	GT	G	0/0	0/0	./.	./.
49	150080	G	A	0/0	0/0	./.	./.
50	150585	T	A	0/0	0/0	./.	./.
51	153975	T	C	0/0	0/0	./.	./.
52	159146	A	G	0/0	0/0	./.	./.
53	162586	C	T	0/0	0/0	./.	./.
54	163446	G	A	0/0	0/0	./.	./.
55	163641	T	C	0/0	0/0	./.	./.
56	167356	T	C	0/0	0/0	./.	./.
57	171095	A	T	0/0	0/0	./.	./.
58	181539	A	G	0/0	0/0	./.	./.
59	182327	C	T	0/0	0/0	./.	./.
60	190959	C	T	0/0	0/0	./.	./.
61	198368	G	A	0/0	0/0	./.	./.
62	198393	A	G	0/0	0/0	./.	./.
63	204573	G	C	0/0	0/0	./.	./.
64	205220	A	G	0/0	0/0	./.	./.
65	205252	C	T	0/0	0/0	./.	./.
66	206301	C	A	0/0	0/0	./.	./.
67	206500	G	A	0/0	0/0	./.	./.
68	206634	A	C	0/0	0/0	./.	./.
69	207061	A	G	0/0	0/0	./.	./.
70	207101	G	A	0/0	0/0	./.	./.
71	207131	T	G	0/0	0/0	./.	./.
72	207181	T	A	0/0	0/0	./.	./.

[0327]

Nic1 标 记号	POS	REF	ALT	BU21	HI	LI	LA
73	207204	C	G	0/0	0/0	./.	./.
74	207512	C	A	0/0	0/0	./.	./.
75	208518	A	C	0/0	0/0	./.	./.
76	208522	C	T	0/0	0/0	./.	./.
77	208556	T	C	0/0	0/0	./.	./.
78	208572	T	C	0/0	0/0	./.	./.
79	209297	T	C	0/0	0/0	./.	./.
80	209315	C	T	0/0	0/0	./.	./.
81	209661	C	T	0/0	0/0	./.	./.
82	209674	T	C	0/0	0/0	./.	./.
83	210485	T	G	0/0	0/0	./.	./.
84	211269	T	C	0/0	0/0	./.	./.
85	216116	C	T	0/0	0/0	./.	./.
86	222130	G	A	0/0	0/0	./.	./.
87	224309	G	T	0/0	0/0	./.	./.
88	224568	C	T	0/0	0/0	./.	./.
89	232343	C	T	0/0	0/0	./.	./.
90	238580	C	T	0/0	0/0	./.	./.
91	239069	T	A	0/0	0/0	./.	./.
92	242812	A	T	0/0	0/0	./.	./.
93	243675	A	G	0/0	0/0	./.	./.
94	244449	A	G	0/0	0/0	./.	./.
95	244950	T	C	0/0	0/0	./.	./.
96	245176	A	G	0/0	0/0	./.	./.
97	254755	G	A	0/0	0/0	./.	./.
98	255169	C	T	0/0	0/0	./.	./.
99	259166	G	A	0/0	0/0	./.	./.
100	271206	G	A	0/0	0/0	./.	./.
101	272201	T	C	0/0	0/0	./.	./.
102	273944	T	C	0/0	0/0	./.	./.
103	276518	C	T	0/0	0/0	./.	./.
104	276838	T	A	0/0	0/0	./.	./.
105	281675	G	T	0/0	0/0	./.	./.
106	284726	G	T	0/0	0/0	./.	./.
107	286609	C	T	0/0	0/0	./.	./.
108	286915	A	G	0/0	0/0	./.	./.
109	286966	T	A	0/0	0/0	./.	./.
110	286987	A	G	0/0	0/0	./.	./.
111	293036	A	G	0/0	0/0	./.	./.
112	300478	G	T	0/0	0/0	./.	./.

[0328]

Nic1 标 记号	POS	REF	ALT	BU21	HI	LI	LA
113	300631	C	T	0/0	0/0	./.	./.
114	300759	A	T	0/0	0/0	./.	./.
115	301119	G	A	0/0	0/0	./.	./.
116	303538	A	C	0/0	0/0	./.	./.
117	304744	A	G	0/0	0/0	./.	./.
118	306236	C	T	0/0	0/0	./.	./.
119	308259	C	A	0/0	0/0	./.	./.
120	313448	A	G	0/0	0/0	./.	./.
121	317046	C	G	0/0	0/0	./.	./.
122	318876	GGGG GGT	G	0/0	0/0	./.	./.
123	320878	G	T	0/0	0/0	./.	./.
124	321339	T	G	0/0	0/0	./.	./.
125	321361	A	G	0/0	0/0	./.	./.
126	321363	C	T	0/0	0/0	./.	./.
127	321532	A	C	0/0	0/0	./.	./.
128	324556	A	G	0/0	0/0	./.	./.
129	324990	G	A	0/0	0/0	./.	./.
130	329545	T	A	0/0	0/0	./.	./.
131	349886	T	C	0/0	0/0	./.	./.
132	350146	A	C	0/0	0/0	./.	./.
133	350163	A	G	0/0	0/0	./.	./.
134	350515	T	G	0/0	0/0	./.	./.
135	351620	A	C	0/0	0/0	./.	./.
136	353656	A	G	0/0	0/0	./.	./.
137	354639	CT	C	0/0	0/0	./.	./.
138	366150	AG	A	0/0	0/0	./.	./.
139	388388	T	G	0/0	0/0	./.	./.
140	390555	G	A	0/0	0/0	./.	./.
141	391708	C	T	0/0	0/0	./.	./.
142	392358	G	C	0/0	0/0	./.	./.
143	392712	T	C	0/0	0/0	./.	./.
144	401096	G	T	0/0	0/0	./.	./.
145	401445	G	C	0/0	0/0	./.	./.
146	404189	A	C	0/0	0/0	./.	./.
147	406610	T	G	0/0	0/0	./.	./.
148	407131	C	T	0/0	0/0	./.	./.
149	411268	A	C	0/0	0/0	./.	./.
150	412538	T	A	0/0	0/0	./.	./.
151	412550	G	T	0/0	0/0	./.	./.

[0329]

Nic1 标 记号	POS	REF	ALT	BU21	HI	LI	LA
152	413003	A	C	0/0	0/0	./.	./.
153	413373	A	T	0/0	0/0	./.	./.
154	413437	C	T	0/0	0/0	./.	./.
155	415425	A	G	0/0	0/0	./.	./.
156	415449	A	C	0/0	0/0	./.	./.
157	415458	T	G	0/0	0/0	./.	./.
158	417103	A	C	0/0	0/0	./.	./.
159	418099	G	A	0/0	0/0	./.	./.
160	418558	A	G	0/0	0/0	./.	./.
161	423778	T	C	0/0	0/0	./.	./.
162	430030	G	T	0/0	0/0	./.	./.
163	437935	T	C	0/0	0/0	./.	./.
164	440647	G	A	0/0	0/0	./.	./.
165	440664	C	T	0/0	0/0	./.	./.
166	442849	T	A	0/0	0/0	./.	./.
167	445778	A	G	0/0	0/0	./.	./.
168	446871	T	C	0/0	0/0	./.	./.
169	447998	A	T	0/0	0/0	./.	./.
170	450145	A	G	0/0	0/0	./.	./.
171	452523	A	G	0/0	0/0	./.	./.
172	452968	G	A	0/0	0/0	./.	./.
173	452987	T	C	0/0	0/0	./.	./.
174	453033	C	T	0/0	0/0	./.	./.
175	453112	G	A	0/0	0/0	./.	./.
176	453174	T	C	0/0	0/0	./.	./.
177	453188	G	A	0/0	0/0	./.	./.
178	454399	T	C	0/0	0/0	./.	./.
179	456245	A	C	0/0	0/0	./.	./.
180	466990	T	C	0/0	0/0	./.	./.
181	473415	C	T	0/0	0/0	./.	./.
182	479185	G	A	0/0	0/0	./.	./.
183	481321	T	C	0/0	0/0	./.	./.
184	481329	T	C	0/0	0/0	./.	./.
185	481922	T	C	0/0	0/0	./.	./.
186	485778	G	A	0/0	0/0	./.	./.
187	486466	A	G	0/0	0/0	./.	./.
188	487632	G	A	0/0	0/0	./.	./.
189	495057	TA	T	0/0	0/0	./.	./.
190	496107	G	A	0/0	0/0	./.	./.
191	496129	G	T	0/0	0/0	./.	./.

[0330]

Nic1 标记号	POS	REF	ALT	BU21	HI	LI	LA
192	498389	C	T	0/0	0/0	./.	./.
193	502623	G	T	0/0	0/0	./.	./.
194	515727	T	C	0/0	0/0	./.	./.
195	517529	C	T	0/0	0/0	./.	./.
196	521866	T	C	0/0	0/0	./.	./.
197	525885	AAAA AAAA AAAA AAAA AAAA AAAA AAAA AAAA AAAA AAAA AAAA AAAA AAAA AAAA AAAA AAAA AAAA CTTT (SEQ ID NO: 71)	A	0/0	0/0	./.	./.
198	528989	A	G	0/0	0/0	./.	./.
199	542757	A	G	0/0	0/0	./.	./.
200	542859	C	A	0/0	0/0	./.	./.
201	543285	T	C	0/0	0/0	./.	./.
202	543440	G	C	0/0	0/0	./.	./.
203	543565	G	A	0/0	0/0	./.	./.
204	543749	C	T	0/0	0/0	./.	./.
205	543771	T	C	0/0	0/0	./.	./.
206	543941	C	A	0/0	0/0	./.	./.
207	543958	G	C	0/0	0/0	./.	./.

[0331] 表4:Nic2缺失片段 (NT1.0-Scaffold0000549) 中的多态性位点。REF是指参考TN90等位基因的序列。ALT是指在K326、Narrow Leaf Madole或东方型中发现的序列多态性。“0/0”是指参考TN90等位基因纯合,而“./.”是指丢失数据 (例如缺失)。

[0332]

Nic2 标 记号	POS	REF	ALT	BU21	HI	LI	LA
1	8514	T	A	0/0	/.	0/0	/.
2	12287	A	C	0/0	/.	0/0	/.
3	15020	C	T	0/0	/.	0/0	/.
4	20105	C	T	0/0	/.	0/0	/.
5	21133	G	T	0/0	/.	0/0	/.
6	21182	A	C	0/0	/.	0/0	/.
7	25267	T	C	0/0	/.	0/0	/.
8	27211	A	G	0/0	/.	0/0	/.
9	27965	T	A	0/0	/.	0/0	/.
10	28780	A	G	0/0	/.	0/0	/.
11	28800	G	A	0/0	/.	0/0	/.
12	28976	T	A	0/0	/.	0/0	/.
13	29013	G	A	0/0	/.	0/0	/.
14	31000	C	T	0/0	/.	0/0	/.
15	31074	T	C	0/0	/.	0/0	/.
16	31118	A	G	0/0	/.	0/0	/.
17	31136	G	A	0/0	/.	0/0	/.
18	31930	T	G	0/0	/.	0/0	/.
19	32523	A	G	0/0	/.	0/0	/.
20	33617	A	C	0/0	/.	0/0	/.
21	33661	A	T	0/0	/.	0/0	/.
22	35656	T	C	0/0	/.	0/0	/.
23	35674	C	T	0/0	/.	0/0	/.
24	44008	C	T	0/0	/.	0/0	/.
25	62895	T	C	0/0	/.	0/0	/.
26	62916	G	A	0/0	/.	0/0	/.
27	64381	A	G	0/0	/.	0/0	/.
28	77284	A	T	0/0	/.	0/0	/.
29	77363	A	G	0/0	/.	0/0	/.
30	77909	G	C	0/0	/.	0/0	/.
31	77975	A	T	0/0	/.	0/0	/.
32	77985	A	G	0/0	/.	0/0	/.
33	80196	C	T	0/0	/.	0/0	/.
34	87731	C	T	0/0	/.	0/0	/.
35	87799	G	A	0/0	/.	0/0	/.
36	89358	G	T	0/0	/.	0/0	/.
37	92022	T	C	0/0	/.	0/0	/.
38	92542	C	A	0/0	/.	0/0	/.
39	92675	C	T	0/0	/.	0/0	/.
40	92695	G	A	0/0	/.	0/0	/.

[0333]

Nic2 标 记号	POS	REF	ALT	BU21	HI	LI	LA
41	94612	C	G	0/0	/.	0/0	/.
42	94683	C	T	0/0	/.	0/0	/.
43	103131	G	T	0/0	/.	0/0	/.
44	108577	A	G	0/0	/.	0/0	/.
45	108967	G	A	0/0	/.	0/0	/.
46	113914	T	C	0/0	/.	0/0	/.
47	118142	A	T	0/0	/.	0/0	/.
48	118151	T	C	0/0	/.	0/0	/.
49	119780	A	T	0/0	/.	0/0	/.
50	121195	G	A	0/0	/.	0/0	/.
51	135236	G	T	0/0	/.	0/0	/.
52	135239	C	T	0/0	/.	0/0	/.
53	135401	T	C	0/0	/.	0/0	/.
54	135683	T	A	0/0	/.	0/0	/.
55	136546	C	T	0/0	/.	0/0	/.
56	136553	T	G	0/0	/.	0/0	/.
57	137241	G	A	0/0	/.	0/0	/.
58	137643	CT	C	0/0	/.	0/0	/.
59	138384	C	G	0/0	/.	0/0	/.
60	138450	T	C	0/0	/.	0/0	/.
61	138457	G	A	0/0	/.	0/0	/.
62	138663	T	C	0/0	/.	0/0	/.
63	138791	G	A	0/0	/.	0/0	/.
64	152981	C	T	0/0	/.	0/0	/.
65	152994	G	A	0/0	/.	0/0	/.
66	158500	C	T	0/0	/.	0/0	/.
67	159059	G	A	0/0	/.	0/0	/.
68	159855	A	C	0/0	/.	0/0	/.
69	162242	C	A	0/0	/.	0/0	/.
70	162272	T	C	0/0	/.	0/0	/.
71	162616	T	C	0/0	/.	0/0	/.
72	162698	A	T	0/0	/.	0/0	/.
73	163071	T	C	0/0	/.	0/0	/.
74	163389	T	A	0/0	/.	0/0	/.
75	163549	C	T	0/0	/.	0/0	/.
76	169105	A	C	0/0	/.	0/0	/.
77	169578	A	G	0/0	/.	0/0	/.
78	170180	G	A	0/0	/.	0/0	/.
79	170221	G	T	0/0	/.	0/0	/.
80	170292	A	G	0/0	/.	0/0	/.

[0334]

Nic2 标 记号	POS	REF	ALT	BU21	HI	LI	LA
81	176741	TG	T	0/0	/.	0/0	/.
82	177610	G	A	0/0	/.	0/0	/.
83	189973	C	A	0/0	/.	0/0	/.
84	190622	T	C	0/0	/.	0/0	/.
85	192179	T	C	0/0	/.	0/0	/.
86	192203	A	G	0/0	/.	0/0	/.
87	192520	G	A	0/0	/.	0/0	/.
88	198596	A	T	0/0	/.	0/0	/.
89	210148	C	T	0/0	/.	0/0	/.
90	211238	A	G	0/0	/.	0/0	/.
91	211276	G	A	0/0	/.	0/0	/.
92	211298	C	T	0/0	/.	0/0	/.
93	213281	C	G	0/0	/.	0/0	/.
94	221865	T	G	0/0	/.	0/0	/.
95	224164	T	A	0/0	/.	0/0	/.
96	226470	A	G	0/0	/.	0/0	/.
97	228965	C	A	0/0	/.	0/0	/.
98	230571	G	A	0/0	/.	0/0	/.
99	232100	C	G	0/0	/.	0/0	/.
100	232459	G	A	0/0	/.	0/0	/.
101	234537	C	T	0/0	/.	0/0	/.
102	241822	G	A	0/0	/.	0/0	/.
103	244741	C	T	0/0	/.	0/0	/.
104	246403	G	A	0/0	/.	0/0	/.
105	246519	C	T	0/0	/.	0/0	/.
106	247174	G	A	0/0	/.	0/0	/.
107	252132	G	A	0/0	/.	0/0	/.
108	252717	T	C	0/0	/.	0/0	/.
109	252728	T	C	0/0	/.	0/0	/.
110	255099	A	G	0/0	/.	0/0	/.
111	257486	C	T	0/0	/.	0/0	/.
112	258008	C	A	0/0	/.	0/0	/.
113	258101	G	A	0/0	/.	0/0	/.
114	260634	C	T	0/0	/.	0/0	/.
115	260944	C	T	0/0	/.	0/0	/.
116	261816	T	C	0/0	/.	0/0	/.
117	264558	T	C	0/0	/.	0/0	/.
118	265660	G	A	0/0	/.	0/0	/.
119	275884	A	G	0/0	/.	0/0	/.
120	277012	A	G	0/0	/.	0/0	/.

[0335]

Nic2 标 记号	POS	REF	ALT	BU21	HI	LI	LA
121	278708	A	G	0/0	/.	0/0	/.
122	278762	T	C	0/0	/.	0/0	/.
123	278781	G	T	0/0	/.	0/0	/.
124	281156	A	C	0/0	/.	0/0	/.
125	282192	G	T	0/0	/.	0/0	/.
126	285956	A	G	0/0	/.	0/0	/.
127	286075	A	G	0/0	/.	0/0	/.
128	290957	G	T	0/0	/.	0/0	/.
129	295851	C	A	0/0	/.	0/0	/.
130	297396	G	C	0/0	/.	0/0	/.
131	298942	G	T	0/0	/.	0/0	/.
132	299095	A	G	0/0	/.	0/0	/.
133	300592	C	T	0/0	/.	0/0	/.
134	301121	C	G	0/0	/.	0/0	/.
135	301147	C	T	0/0	/.	0/0	/.
136	306986	G	A	0/0	/.	0/0	/.
137	310002	C	T	0/0	/.	0/0	/.
138	314892	C	T	0/0	/.	0/0	/.
139	324211	G	A	0/0	/.	0/0	/.
140	327796	T	G	0/0	/.	0/0	/.
141	330726	A	G	0/0	/.	0/0	/.
142	331046	G	A	0/0	/.	0/0	/.
143	332636	T	C	0/0	/.	0/0	/.
144	336046	G	C	0/0	/.	0/0	/.
145	338147	C	A	0/0	/.	0/0	/.
146	340868	C	T	0/0	/.	0/0	/.
147	341170	T	A	0/0	/.	0/0	/.
148	359439	T	A	0/0	/.	0/0	/.
149	359447	C	A	0/0	/.	0/0	/.
150	362131	A	G	0/0	/.	0/0	/.
151	363467	G	A	0/0	/.	0/0	/.
152	365467	T	A	0/0	/.	0/0	/.
153	367311	A	G	0/0	/.	0/0	/.
154	370267	G	A	0/0	/.	0/0	/.
155	384941	A	G	0/0	/.	0/0	/.
156	393244	G	T	0/0	/.	0/0	/.
157	394169	T	C	0/0	/.	0/0	/.
158	394200	A	G	0/0	/.	0/0	/.
159	394213	A	C	0/0	/.	0/0	/.
160	394228	G	A	0/0	/.	0/0	/.

[0336]

Nic2 标 记号	POS	REF	ALT	BU21	HI	LI	LA
161	396927	C	T	0/0	/.	0/0	/.
162	404142	T	C	0/0	/.	0/0	/.
163	404761	A	G	0/0	/.	0/0	/.
164	406475	G	A	0/0	/.	0/0	/.
165	406481	T	A	0/0	/.	0/0	/.
166	410940	G	A	0/0	/.	0/0	/.
167	411032	G	A	0/0	/.	0/0	/.
168	411069	T	A	0/0	/.	0/0	/.
169	411317	GA	G	0/0	/.	0/0	/.
170	413027	C	A	0/0	/.	0/0	/.
171	413058	C	T	0/0	/.	0/0	/.
172	414268	C	G	0/0	/.	0/0	/.
173	416798	T	C	0/0	/.	0/0	/.
174	417540	G	C	0/0	/.	0/0	/.
175	420742	A	G	0/0	/.	0/0	/.
176	421259	T	C	0/0	/.	0/0	/.
177	426709	A	G	0/0	/.	0/0	/.
178	427690	G	C	0/0	/.	0/0	/.
179	430705	G	A	0/0	/.	0/0	/.
180	431773	G	A	0/0	/.	0/0	/.
181	433900	A	G	0/0	/.	0/0	/.
182	442372	T	C	0/0	/.	0/0	/.
183	451131	G	A	0/0	/.	0/0	/.
184	454572	C	A	0/0	/.	0/0	/.
185	460976	G	A	0/0	/.	0/0	/.
186	461460	C	T	0/0	/.	0/0	/.
187	464177	A	G	0/0	/.	0/0	/.
188	465089	T	C	0/0	/.	0/0	/.
189	465495	T	G	0/0	/.	0/0	/.
190	468266	C	A	0/0	/.	0/0	/.
191	474980	C	T	0/0	/.	0/0	/.
192	480131	A	T	0/0	/.	0/0	/.
193	483164	T	C	0/0	/.	0/0	/.
194	484777	A	G	0/0	/.	0/0	/.
195	485948	G	A	0/0	/.	0/0	/.
196	486626	G	C	0/0	/.	0/0	/.
197	494004	C	T	0/0	/.	0/0	/.
198	494393	T	C	0/0	/.	0/0	/.
199	495953	C	T	0/0	/.	0/0	/.
200	496552	A	G	0/0	/.	0/0	/.

[0337]

Nic2 标 记号	POS	REF	ALT	BU21	HI	LI	LA
201	499298	A	G	0/0	/.	0/0	/.
202	499947	T	C	0/0	/.	0/0	/.
203	501352	G	A	0/0	/.	0/0	/.
204	505470	A	G	0/0	/.	0/0	/.
205	509926	G	A	0/0	/.	0/0	/.
206	511336	A	G	0/0	/.	0/0	/.
207	513808	TC	T	0/0	/.	0/0	/.
208	515055	C	T	0/0	/.	0/0	/.
209	516444	A	G	0/0	/.	0/0	/.
210	519420	A	T	0/0	/.	0/0	/.
211	521028	G	C	0/0	/.	0/0	/.
212	521834	T	C	0/0	/.	0/0	/.
213	524240	A	C	0/0	/.	0/0	/.
214	524256	T	C	0/0	/.	0/0	/.
215	524317	A	G	0/0	/.	0/0	/.
216	524986	G	A	0/0	/.	0/0	/.
217	526266	A	C	0/0	/.	0/0	/.
218	526905	C	T	0/0	/.	0/0	/.
219	526987	C	A	0/0	/.	0/0	/.
220	530341	A	G	0/0	/.	0/0	/.
221	531695	A	T	0/0	/.	0/0	/.
222	541587	T	C	0/0	/.	0/0	/.
223	541648	G	A	0/0	/.	0/0	/.
224	544386	A	G	0/0	/.	0/0	/.
225	545716	T	C	0/0	/.	0/0	/.
226	546047	C	A	0/0	/.	0/0	/.
227	546372	A	G	0/0	/.	0/0	/.
228	546416	A	G	0/0	/.	0/0	/.
229	546434	T	A	0/0	/.	0/0	/.
230	546775	C	T	0/0	/.	0/0	/.
231	547015	A	G	0/0	/.	0/0	/.
232	554248	A	G	0/0	/.	0/0	/.
233	554496	C	T	0/0	/.	0/0	/.
234	556239	A	G	0/0	/.	0/0	/.
235	558480	C	T	0/0	/.	0/0	/.
236	562524	C	T	0/0	/.	0/0	/.
237	563642	A	T	0/0	/.	0/0	/.
238	563802	T	A	0/0	/.	0/0	/.
239	563862	G	T	0/0	/.	0/0	/.
240	573521	C	T	0/0	/.	0/0	/.

[0338]

Nic2 标 记号	POS	REF	ALT	BU21	HI	LI	LA
241	573552	T	C	0/0	/.	0/0	/.
242	574370	C	A	0/0	/.	0/0	/.
243	576866	C	T	0/0	/.	0/0	/.
244	577137	C	T	0/0	/.	0/0	/.
245	582336	T	A	0/0	/.	0/0	/.
246	585822	G	A	0/0	/.	0/0	/.
247	586808	A	G	0/0	/.	0/0	/.
248	592379	A	C	0/0	/.	0/0	/.
249	594088	T	C	0/0	/.	0/0	/.
250	594122	T	A	0/0	/.	0/0	/.
251	594128	C	A	0/0	/.	0/0	/.
252	596697	T	C	0/0	/.	0/0	/.
253	599034	G	C	0/0	/.	0/0	/.
254	609636	T	C	0/0	/.	0/0	/.
255	610500	C	T	0/0	/.	0/0	/.
256	610533	T	A	0/0	/.	0/0	/.
257	610938	A	G	0/0	/.	0/0	/.
258	614216	A	G	0/0	/.	0/0	/.
259	614885	TTA	T	0/0	/.	0/0	/.
260	617065	G	A	0/0	/.	0/0	/.
261	621974	G	A	0/0	/.	0/0	/.
262	623585	T	G	0/0	/.	0/0	/.
263	645256	C	G	0/0	/.	0/0	/.
264	645263	T	C	0/0	/.	0/0	/.
265	646476	C	T	0/0	/.	0/0	/.
266	652446	AT	A	0/0	/.	0/0	/.
267	660505	C	T	0/0	/.	0/0	/.
268	667339	G	A	0/0	/.	0/0	/.
269	668081	A	G	0/0	/.	0/0	/.
270	669231	T	G	0/0	/.	0/0	/.
271	680482	G	A	0/0	/.	0/0	/.
272	680899	T	C	0/0	/.	0/0	/.
273	681503	A	G	0/0	/.	0/0	/.
274	681579	TA	T	0/0	/.	0/0	/.
275	685669	G	A	0/0	/.	0/0	/.
276	687781	T	C	0/0	/.	0/0	/.
277	690142	C	T	0/0	/.	0/0	/.
278	696619	G	A	0/0	/.	0/0	/.
279	696973	C	T	0/0	/.	0/0	/.
280	699014	T	C	0/0	/.	0/0	/.

[0339]

Nic2 标 记号	POS	REF	ALT	BU21	HI	LI	LA
281	699059	A	G	0/0	/.	0/0	/.
282	699441	A	G	0/0	/.	0/0	/.
283	699482	A	G	0/0	/.	0/0	/.
284	699507	A	C	0/0	/.	0/0	/.
285	699568	AC	A	0/0	/.	0/0	/.
286	699991	C	T	0/0	/.	0/0	/.
287	701221	G	A	0/0	/.	0/0	/.
288	701338	C	T	0/0	/.	0/0	/.
289	704968	A	G	0/0	/.	0/0	/.
290	705347	C	T	0/0	/.	0/0	/.
291	708907	T	C	0/0	/.	0/0	/.
292	710281	C	T	0/0	/.	0/0	/.
293	711996	T	C	0/0	/.	0/0	/.
294	714019	C	A	0/0	/.	0/0	/.
295	714445	T	C	0/0	/.	0/0	/.
296	714483	T	C	0/0	/.	0/0	/.
297	714491	T	C	0/0	/.	0/0	/.
298	716492	A	G	0/0	/.	0/0	/.
299	716551	A	G	0/0	/.	0/0	/.
300	721681	G	A	0/0	/.	0/0	/.
301	722271	T	A	0/0	/.	0/0	/.
302	723002	G	C	0/0	/.	0/0	/.
303	732955	T	C	0/0	/.	0/0	/.
304	734117	T	G	0/0	/.	0/0	/.
305	743601	TA	T	0/0	/.	0/0	/.
306	745383	C	T	0/0	/.	0/0	/.
307	745938	C	T	0/0	/.	0/0	/.
308	747238	T	C	0/0	/.	0/0	/.
309	750673	A	G	0/0	/.	0/0	/.
310	751683	G	T	0/0	/.	0/0	/.
311	751702	C	T	0/0	/.	0/0	/.
312	762212	A	G	0/0	/.	0/0	/.
313	762934	C	T	0/0	/.	0/0	/.
314	765864	C	T	0/0	/.	0/0	/.
315	768978	G	A	0/0	/.	0/0	/.
316	770377	T	A	0/0	/.	0/0	/.
317	773111	G	A	0/0	/.	0/0	/.
318	773112	T	A	0/0	/.	0/0	/.
319	773155	A	T	0/0	/.	0/0	/.
320	778101	C	A	0/0	/.	0/0	/.

[0340]

Nic2 标 记号	POS	REF	ALT	BU21	HI	LI	LA
321	778231	T	A	0/0	./.	0/0	./.
322	779756	C	T	0/0	./.	0/0	./.
323	780381	A	G	0/0	./.	0/0	./.
324	780398	A	C	0/0	./.	0/0	./.
325	782275	A	C	0/0	./.	0/0	./.
326	782502	T	G	0/0	./.	0/0	./.
327	788379	A	G	0/0	./.	0/0	./.
328	788416	T	A	0/0	./.	0/0	./.
329	789721	C	T	0/0	./.	0/0	./.
330	795120	G	A	0/0	./.	0/0	./.
331	797627	G	A	0/0	./.	0/0	./.
332	799891	A	C	0/0	./.	0/0	./.
333	800525	A	G	0/0	./.	0/0	./.
334	801525	T	C	0/0	./.	0/0	./.
335	802686	ATGC CATG TGGT TTAT ATTA TTGG CACG TGAG TTGT CCGT GCGA GTCC AGAT ATTT ATAC TATA GC (SEQ ID NO: 72)	A	0/0	./.	0/0	./.
336	802777	A	T	0/0	./.	0/0	./.
337	807008	T	C	0/0	./.	0/0	./.
338	808634	T	G	0/0	./.	0/0	./.
339	813994	G	A	0/0	./.	0/0	./.
340	817957	A	G	0/0	./.	0/0	./.

[0341] 实施例3:通过遗传分离分析确认鉴定的Nic1基因座。

[0342] 为了帮助确认鉴定为Nic1基因座的区域的信息,开发了分离Nic1和Nic2基因座的

F2群体,目的是观察与表型分离的nic1和nic2标记。F2群体来自TN90与LA BU21之间的杂交,并且群体大小为173。在自动化DNA提取仪Oktopure (LGC genomics,LLC,Beverly,MA)上使用定制的sbeadex maxi植物试剂盒 (LGC genomics,LLC,Beverly,MA),从单个F2植株以及亲本系和四个已知的白肋系 (BU21、HI BU21、LI BU21和LA BU21)的绿叶组织提取DNA。从这个群体随机选择173个个体中的九十 (90) 个进行Nic1和Nic2基因座以及对照基因座的基因分型。通过PCR扩增所鉴定的缺失的选定区域来确认所鉴定的Nic1缺失。表5示出了使用Primer3网站V.4.0.0为Nic1基因座设计的PCR引物。表5也提供了从Shoji等 (2010) 获得的用于Nic2基因座和对照基因座的引物序列。

[0343] 对包含以下的30 μ L反应进行的PCR:5 μ L的~10ng/ μ L DNA、15 μ L的2xAmpliTaq Gold 360MasterMix (Applied Biosystems,Foster City,CA)、1.2 μ L的10 μ M正向引物、1.2 μ L的10 μ M反向引物和7.6 μ L的ddH₂O。PCR反应开始在95 $^{\circ}$ C初始变性5min,然后是具有以下温度分布的35个循环:94 $^{\circ}$ C、20秒,最佳退火温度下20秒 (表5),和72 $^{\circ}$ C、1min。然后在72 $^{\circ}$ C下进行7分钟的最后延伸阶段,并无限期保持在4 $^{\circ}$ C。

[0344] 表5:用于PCR的引物序列和退火温度列表

	基因座	特定基因	引物	引物序列 (SEQ ID)	PCR T _m
[0345]	Nic1	g100614_Scaffold 0002504	F	CAACCTAGCCACTGGTCCAT (SEQ ID No: 3)	62
			R	TCAAACCAAGAGCGAGGAGT (SEQ ID No: 4)	
	Nic2	ERF 189	F	GGGCAATGGAAATGAATCTAGC (SEQ ID No: 5)	55
			R	CTTCCTTCCTTTCACATAG (SEQ ID No: 6)	
[0346]	对照	ERF 199	F	CCATTCATTTTCATCCAAACCC (SEQ ID No: 7)	55
			R	CGGAGTACTTTTCATGGGATTC (SEQ ID No: 8)	

[0347] 然后在eGene (Qiagen N.V.,Venlo,荷兰)上观察PCR产物,所需大小的PCR产物带存在评为基因座存在,其不存在表示基因座缺失。用于基因ERF199的引物 (表5) 用作对照,其不存在表明DNA质量有问题,如果不存在,将样品从进一步的分析中去除。使用基于CORESTA推荐方法 (CRM NO.7) 和ISO标准 (ISO TC 126N 394E.)的GC-FID方法,还对来自每个个体植株打顶后2周收集的叶样品进行烟碱、降烟碱、新烟草碱和假木贼碱的生物碱测量 (表6)。生物碱测量作为nic1和nic2基因型的额外确认。

[0348] 所有90个F2个体以及TN90、BU21、HI BU21、LI BU21和LA BU21显示成功扩增对照ERF 199引物。在90个F2个体的19个中以及在HI BU21和LA BU21中观察到Nic2基因座的缺失。在90个F2个体的21个中以及在LI BU21和LA BU21中观察到Nic1基因座的缺失。同样,Nic1和Nic2基因座在90个F2植物的5个中以及在LA BU21中一起缺失。表6提供173个F2个体中的90个的基因型的信息。表7提供对于每个基因型观察到的和预期的植物数量,其给出的两个基因座的分离比为9:3:3:1。表8显示了F2分离群体 ((LA白肋21 $\times$ TN 90LC) (X)) 中观察

到的四种基因型的每一种的总生物碱含量的平均百分比。

[0349] 表6:用于分离nic1和nic2的F2群体((LA白肋21×TN 90LC)(X))的单个植株的Nic1和Nic2基因型和生物碱测量。Nic1和Ni2基因型信息不存在表明那些植物没有基因分型。BLQ:低于定量水平。

[0350]	植物	Nic1 基因型 (存在=1, 不存在=0)	Nic2 基因型 (存在=1, 不存在=0)	总生物碱百分比	烟碱百分比
	植物 1	1	0	3.0011	2.8628
	植物 2	1	1	2.6729	2.5628
	植物 3	0	1	2.2933	2.2051
	植物 4	0	0	BLQ	BLQ

[0351]

植物	Nic1 基因型 (存在=1, 不存在=0)	Nic2 基因型 (存在=1, 不存在=0)	总生物碱百分比	烟碱百分比
植物 5	1	1	3.1597	3.0139
植物 6	1	1	2.9731	2.8266
植物 7	0	0	BLQ	BLQ
植物 8	1	1	2.8809	2.763
植物 9	1	1	4.1221	3.9034
植物 10	1	1	3.994	3.8198
植物 11	1	1	3.7119	3.5467
植物 12	1	1	3.6687	3.5089
植物 13	1	1	4.2203	4.0304
植物 14	1	0	3.2027	3.0498
植物 15	1	1	3.7481	3.5971
植物 16	0	1	BLQ	BLQ
植物 17	1	1	4.422	4.1744
植物 18	1	1	2.8292	2.7127
植物 19	1	0	2.7538	2.6094
植物 20			1.9167	1.8179
植物 21	1	1	4.3081	4.0381
植物 22	1	1	3.6128	3.4443
植物 23	1	1	BLQ	BLQ
植物 24	1	1	BLQ	BLQ
植物 25	0	1	1.9377	1.8346
植物 26	1	1	2.8042	2.6369
植物 27	1	1	3.2544	3.1033
植物 28	1	1	3.4886	3.3306
植物 29	1	1	2.7115	2.5954
植物 30	1	1	BLQ	BLQ
植物 31	1	1	4.1091	3.9219
植物 32	1	1	3.2426	3.1062
植物 33	0	1	1.9005	1.7845
植物 34			BLQ	BLQ
植物 35	0	1	3.0363	2.8753
植物 36	1	1	3.6491	3.4489
植物 37	0	1	1.7179	1.6199
植物 38	1	0	2.8954	2.7684
植物 39	1	1	3.5936	3.4294

[0352]

植物	Nic1 基因型 (存在=1, 不存在=0)	Nic2 基因型 (存在=1, 不存在=0)	总生物碱百分比	烟碱百分比
植物 40	0	1	2.8582	2.7421
植物 41	1	0	BLQ	BLQ
植物 42			4.4196	4.2055
植物 43			BLQ	BLQ
植物 44	0	1	1.6122	1.5324
植物 45	1	1	2.4573	2.3326
植物 46	1	1	3.4145	3.1655
植物 47	1	1	3.5314	3.3588
植物 48	1	1	3.7216	3.5305
植物 49	1	0	3.9067	3.7233
植物 50	1	0	3.0601	2.9073
植物 51	1	1	3.8504	3.6463
植物 52	0	1	BLQ	BLQ
植物 53	1	1	4.759	4.47
植物 54			3.6204	3.4389
植物 55	0	0	0.5823	0.5489
植物 56	1	1	3.5092	3.3467
植物 57			4.107	3.9257
植物 58	0	1	2.4405	2.3477
植物 59	0	0	0.4188	0.3927
植物 60	1	1	3.8854	3.7403
植物 61	1	1	5.0457	4.8093
植物 62	0	1	2.1401	2.0358
植物 63	1	1	3.5212	3.3729
植物 64	1	1	5.1868	4.9157
植物 65			2.1012	1.9855
植物 66	1	0	0.3436	0.3282
植物 67			2.3158	2.2265
植物 68	1	1	4.5611	4.3195
植物 69	1	0	3.5517	3.4176
植物 70	1	1	3.945	3.7651
植物 71	0	1	3.8097	3.6456
植物 72	1	0	4.1787	3.9818
植物 73	1	0	5.8965	5.5549
植物 74			3.5723	3.393

[0353]

植物	Nic1 基因型 (存在=1, 不存在=0)	Nic2 基因型 (存在=1, 不存在=0)	总生物碱百分比	烟碱百分比
植物 75	1	1	5.6703	5.3808
植物 76	1	1	3.1131	2.9747
植物 77	1	1	4.2168	4.0215
植物 78	0	1	2.9419	2.79
植物 79	1	1	4.2109	3.9834
植物 80			4.3812	4.1685
植物 81	1	0	0.7258	0.6868
植物 82	1	0	4.0284	3.8092
植物 83	0	1	2.076	1.9724
植物 84	1	1	5.1348	4.8334
植物 85	1	1	5.6781	5.3311
植物 86	1	1	4.3462	4.1069
植物 87	1	1	4.0997	3.8974
植物 88	1	1	3.47	3.3158
植物 89	0	1	2.7871	2.6465
植物 90			3.5846	3.4166
植物 91			2.3103	2.187
植物 92			2.8975	2.7746
植物 93			3.5385	3.3521
植物 94			5.2231	4.9531
植物 95			0.5855	0.5538
植物 96			2.6133	2.4813
植物 97			4.3772	4.1254
植物 98			5.453	5.0786
植物 99			3.5055	3.2963
植物 100			2.0902	1.9902
植物 101			4.5207	4.2845
植物 102			2.0429	1.943
植物 103			BLQ	BLQ
植物 104			3.448	3.3
植物 105			4.3806	4.1605
植物 106			4.3696	4.1293
植物 107			4.0754	3.8637
植物 108			4.1516	3.8525
植物 109			3.6425	3.4421

[0354]

植物	Nic1 基因型 (存在=1, 不存在=0)	Nic2 基因型 (存在=1, 不存在=0)	总生物碱百分比	烟碱百分比
植物 110			4.5431	4.3192
植物 111			4.4435	4.1513
植物 112			3.9523	3.737
植物 113			3.4997	3.3269
植物 114			3.8171	3.638
植物 115			3.4582	3.3031
植物 116	1	1	3.5845	3.4243
植物 117			3.945	3.7507
植物 118	0	1	2.5889	2.4632
植物 119	1	1	3.5472	3.3747
植物 120	0	1	2.4497	2.3085
植物 121	1	1	4.1965	3.9772
植物 122	1	0	3.8719	3.6731
植物 123	1	1	4.5964	4.3886
植物 124	1	1	3.8232	3.667
植物 125			4.7023	4.4466
植物 126			3.4026	3.2343
植物 127			3.2645	3.1016
植物 128			2.842	2.7461
植物 129			3.8524	3.6496
植物 130			4.7094	4.4891
植物 131			0.4548	0.4254
植物 132			3.3652	3.211
植物 133			3.2598	3.1137
植物 134			BLQ	BLQ
植物 135			3.3438	3.2074
植物 136			3.7641	3.5279
植物 137			3.6367	3.4453
植物 138	0	0	0.4207	0.3995
植物 139			2.2626	2.0909
植物 140			4.102	3.8806
植物 141			4.4366	4.2298
植物 142			4.386	4.1759
植物 143			5.7992	5.4971
植物 144			BLQ	BLQ

[0355]

植物	Nic1 基因型 (存在=1, 不存在=0)	Nic2 基因型 (存在=1, 不存在=0)	总生物碱百分比	烟碱百分比
植物 145			2.8282	2.6992
植物 146			2.7978	2.6543
植物 147			3.6949	3.5188
植物 148			2.3355	2.2257
植物 149			1.7436	1.6735
植物 150			BLQ	BLQ
植物 151			1.2908	1.2512
植物 152			2.09	1.999
植物 153			3.2661	3.1309
植物 154			BLQ	BLQ
植物 155			2.9794	2.8095
植物 156			5.3031	5.0703
植物 157			3.6654	3.4826
植物 158			3.7191	3.5198
植物 159			3.5148	3.359
植物 160			3.0729	2.9446
植物 161			4.0171	3.8311
植物 162			2.4201	2.3083
植物 163			4.4837	4.2994
植物 164			3.5759	3.4229
植物 165			3.8303	3.6354
植物 166			4.0415	3.8397
植物 167			2.3684	2.2563
植物 168			3.5194	3.3464
植物 169			4.3816	4.1756
植物 170			3.7741	3.5736
植物 171			2.1715	2.0453
植物 172			6.5454	6.1323
植物 173			6.9723	6.6948

[0356] 表7:在显著性阈值0.05上,被研究的F2群体中观察到的nic1和nic2的分离比与预期分离比9:3:3:1没有明显差异($\chi^2=0.9827$,df=3,P值=0.1944)。

[0357]

基因型	预期比率	预期的植物数量	观察到的植物数量
-----	------	---------	----------

[0358]	Nic1 Nic2	9/16	50.625	55
	Nic1 <i>nic2</i>	3/16	16.875	14
	<i>nic1</i> Nic2	3/16	16.875	16
	<i>nic1 nic2</i>	1/16	5.625	5

[0359] 表8:在F2分离群体((LA白肋21×TN 90LC)(X))中观察到的四种基因型的每一种的总生物碱含量的平均百分比。显示平均值和标准偏差。数量表示测量了生物碱含量的每种基因型的植物数量。每种基因型的植物数量不同于表7中列出的数量,因为这里排除了生物碱含量低于定量水平(BLQ)的植物。

[0360]	基因型	数量	平均值	标准偏差	标准误差平均值	低于95%	高于95%
	Nic1 Nic2	49	3.8419	0.74094	0.10585	3.6291	4.0547
	Nic1 <i>nic2</i>	13	3.18588	1.43233	0.39726	2.3203	4.0514
	<i>nic1</i> Nic2	15	2.43933	0.58525	0.15111	2.1152	2.7634
	<i>nic1 nic2</i>	3	0.47393	0.09385	0.05419	0.2408	0.7071

[0361] 实施例4:基于鉴定的Nic1缺失片段中的基因表达模式,鉴定Nic1缺失中的基因。

[0362] 所鉴定的Nic1缺失片段NT1.0-Scaffold0002504具有至少20个注释基因。基于根RNASeq数据集,分析它们在烟草根中的表达模式。根RNASeq数据集是基于来自四个品种(BU21、HI BU21、LI BU21和LA BU21)的根组织的RNA测序和转录图谱构建的。

[0363] 在温室中种植这四个品种的烟草植物,直到大部分系开始出芽,然后收获根和叶组织样品。收集来自这四个品种中的每一个的15株植物(未打顶植物和打顶植物)的根部组织(~100mg)。表9提供从这些植物收集样品的样品数量、时间点和条件的细节。在不同的时间点收获各基因型的根。从根提取RNA并合并在一起以形成每个基因型的两个单独的合并样品;一个在打顶之前,一个在打顶之后。具体而言,将烟草根在液氮中快速冷冻,然后使用2000Geno/Grinder Spex样品制备物进行浸渍。使用组织RNA提取试剂盒(Promega, Madison, WI),将浸软的组织进一步在自动化Maxwell[®] 16systems (Promega, Madison, WI)上提取RNA。使用Nanodrop1000对分离的RNA进行定量以确认所有样品满足HiSeq2000的2x100bp双端测序所需的最低产量。

[0364] 测序后,整理并进一步过滤原始RNASeq读长。然后使用CLC基因组学工作台v.7.1 (Qiagen, N.V., Velno, The Netherlands)中的转录组映射应用,将过滤的RNASeq读长映射到ALCS的专有烟草基因组。基因的原始表达测量为烟草基因组内204,695个注释基因的每一个的映射读长的数量。然后,使用每百万读长中每千碱基转录本的读长数(RPKM)值将每个基因的基因表达值标准化,以获得基因表达的相对水平。

[0365] 在鉴定的Nic1缺失片段(NT1.0-Scaffold0002504)中注释的20个基因中,18个基因在任何条件下、任何变种中均不表达。只有2个基因在根中表达(表10)。它们是g100614_Scaffold0002504和g100631_Scaffold0002504(分别具有SEQ ID NO:28和13的基因组序列,SEQ ID NO:48和33的cDNA序列,SEQ ID NO:68和53的蛋白序列)。两个表达的基因都注释为“晚疫病抗性蛋白同源物”。对于Nic2基因座处的基因也观察到相似的表达模式(表11)。

[0366] 表9:基于RNAseq的烟草根转录谱的组织取样设计。

基因型	条件	取样植物的数量				
		时间点				
		30 mins	2 hrs	6 hrs	24 hrs	72 hrs
[0367]	白肋 21	打顶前	3	3	3	3
		打顶后	3	3	3	3
	HI 白肋 21	打顶前	3	3	3	3
		打顶后	3	3	3	3
	LI 白肋 21	打顶前	3	3	3	3
		打顶后	3	3	3	3
	LA 白肋 21	打顶前	3	3	3	3
		打顶后	3	3	3	3

[0368] 实施例5:鉴定由Nic1和Nic2调控的基因。

[0369] 来自实施例4的根RNASeq数据集也允许鉴定其表达由Nic1或Nic2调控的基因。在四个品种:BU21、HI BU21、LI BU21和LA BU21中进行基因表达的成对比较。图1显示在错误发现率(FDR)校正的P值为0.05时,在每个可能的成对组合品种之间鉴定为上调或下调基因的数量。

[0370] BU21 vs LA BU21的比较以及HI BU21vs LA BU21的比较是最具信息性的,其在腐胺形成后的烟碱生物合成途径中鉴定所有基因。Nic1基因座(在BU21和HI BU21中存在,但在LA BU21中不存在)的作用也被确认。图2表示Nic1存在下被上调的一些基因。

[0371] 实施例6:在Nic1和Nic2缺失片段处或其附近开发分子标记。

[0372] 所鉴定的Nic1缺失片段(NT1.0-Scaffold0002504)的长度为至少544,860bps。在该片段内,在参考TN90基因组序列和代表缺失模式的BU21系之间检测到至少207个变异位点(表3)

[0373] 在scaffoldNT1.0-Scaffold0000549(NT1.0-Scaffold0000549的总长度=1,142,469bps)内鉴定的Nic2缺失片段的长度为至少820,000bps。在该片段内,在参考TN90基因组序列和代表缺失模式的白肋21系之间检测到至少340个多态性位点(表4)。

[0374] 实施例7:育种含有低烟碱的烟草品种

[0375] 所鉴定的nic1缺失片段,其内部的基因以及与其相关的分子标记用于育种并生产包含nic1缺失或部分缺失的低烟碱烟草杂交种、品种和系。这些基因和标记也用于筛选来自各种烟草种质(例如,不同的烟草种或红花烟草系)的额外的nic1和nic2等位基因。美国专利号7,700,834的表8中提供可以筛选的四十三三个烟草种、四十九个黄花烟草系和大约六百个红花烟草系的集合。

[0376] 鉴定为具有新的nic1或nic2等位基因的种质用作与栽培烟草育种的源材料。可以使用与标准育种方法(诸如回交或谱系方法)相结合的种间或种内杂交方法将所需的nic1或nic2突变等位基因从供体来源转移到栽培烟草中。例如,将包含nic1、nic2或两个突变体等位基因的低烟碱品种(例如供体亲本,诸如LA白肋21)与具有所需遗传背景和农艺优良性状的优良高烟碱品种杂交。任选地,在来自该杂交的F1后代植物中测定表9和10示例的一种或多种分子标记。然后将F1后代植物与母本优良高烟碱品种(轮回亲本)回交。对来自BC1代的植物使用表9和10示例的分子标记进行基因分型,以选择具有较小nic1或nic2缺失片段的烟草植物。经过多轮回交(例如5-7代),获得包含低烟碱性状和来自轮回亲本优良系的其

他所需性状的新的优良烟草品种。由于Nic1和Nic2基因座周围的遗传重组事件,这种新的优良烟草品种也没有与低烟碱性状相关的任何遗传阻力。这些重组事件将nic1和nic2突变从任何相关的有害突破解开,从而减少或避免遗传阻力。使用上述育种和标记辅助选择策略,还可以实现低烟碱性状与其他降低烟碱或降烟碱含量的转基因或天然等位基因的聚合或叠加。

[0377] 低烟碱烟草杂交种、品种或系可以是白肋型、深色型、烤制熟化型、马里兰型或东方型烟草,或可以基本上衍生自:BU 64、CC 101、CC 200、CC 27、CC 301、CC 400、CC 500、CC 600、CC 700、CC 800、CC 900、Coker 176、Coker 319、Coker 371Gold、Coker 48、CU 263、DF911、Galpao烟、GL 26H、GL 350、GL 600、GL 737、GL 939、GL 973、HB 04P、K 149、K 326、K 346、K 358、K394、K 399、K 730、KDH 959、KT 200、KT204LC、KY 10、KY 14、KY 160、KY 17、KY 171、KY 907、KY907LC、KTY14 x L8 LC、Little Crittenden、McNair 373、McNair 944、msKY 14xL8、Narrow Leaf Madole、NC 100、NC 102、NC 2000、NC 291、NC 297、NC 299、NC 3、NC 4、NC 5、NC 6、NC7、NC 606、NC 71、NC 72、NC 810、NC BH 129、NC 2002、Neal Smith Madole、OXFORD 207、`Perique`烟、PVH03、PVH09、PVH19、PVH50、PVH51、R 610、R 630、R 7-11、R 7-12、RG 17、RG 81、RG H51、RGH 4、RGH 51、RS 1410、Speight 168、Speight 172、Speight 179、Speight 210、Speight 220、Speight 225、Speight 227、Speight234、Speight G-28、Speight G-70、Speight H-6、Speight H20、Speight NF3、TI 1406、TI 1269、TN 86、TN86LC、TN 90、TN 97、TN97LC、TN D94、TN D950、TR (Tom Rosson) Madole、VA 309或VA359、Maryland 609、HB3307PLC、HB4488PLC、KT206LC、KT209LC、KT210LC、KT212LC、R610LC、PVH2310、NC196、KTD14LC、KTD6LC、KTD8LC、PD7302LC、PD7305LC、PD7309LC、PD7318LC、PD7319LC、PD7312LC、ShireyLC、或根据本领域已知的标准烟草育种技术的任何商业化烟草品种。

[0378] 表10:在Nic1缺失拼接序列(NT1.0-Scaffold0002504(SEQ ID No.1))中注释的基因。“起始”和“终止”表示注释基因在拼接序列NT1.0-Scaffold0002504上的起始和终止核苷酸的位置。在打顶之前和之后,仅观察到两个粗体表示的基因(g100631_Scaffold0002504和g100614_Scaffold0002504)在BU21和HI BU21的根中表达。在LI BU21或LA BU21根中未检测到这两个基因的表达。基于根RNASeq数据集,所有其他列出的基因在BU21、HI BU21、LI BU21或LA BU21中均未显示根表达。星号(*)表示SEQ ID Nos.83和84分别对应于SEQ ID NO.80的细化cDNA和氨基酸序列。

[0379]

基因	Blast 命中 描述	起始	终止	基因 组序 列 (SEQ ID)	细化的基 因组序 列 (SEQ ID)	cDNA 序列 (SEQ ID)	氨基 酸序 列 (SEQ ID)
g100624_Scaffold0002504	gi 113205363 gb AAT66771.2 推 定多聚蛋白,相同 [<i>Solanum demissum</i>]	23574	44762	9		29	49

[0380]

基因	Blast 命中 描述	起始	终止	基因组序列 (SEQ ID)	细化的基因组序列 (SEQ ID)	cDNA 序列 (SEQ ID)	氨基酸序列 (SEQ ID)
g100616_Scaffold0002504	gi 113205363 gb AAT66771.2 推定多聚蛋白, 相同 [<i>Solanum demissum</i>]	49007	62258	10		30	50
g100619_Scaffold0002504	gi 460410342 ref XP_004250589.1 预测: 未表征蛋白 LOC101263780 [番茄 (<i>Solanum lycopersicum</i>)]	63561	67079	11		31	51
g100629_Scaffold0002504	gi 113205363 gb AAT66771.2 推定多聚蛋白, 相同 [<i>Solanum demissum</i>]	72790	84448	12		32	52
g100631_Scaffold0002504	gi 460370553 ref XP_004231117.1 预测: 推定晚疫抗性蛋白同源物 R1A-10-样 [番茄]	124722	127361	13		33	53
g100627_Scaffold0002504	gi 113205316 gb ABI34339.1 多聚蛋白, 3'-部分, 推定的[<i>Solanum demissum</i>]	130345	131552	14		34	54

[0381]

基因	Blast 命中 描述	起始	终止	基因组序列 (SEQ ID)	细化的基因组序列 (SEQ ID)	cDNA 序列 (SEQ ID)	氨基酸序列 (SEQ ID)
g100630_Scaffold0002504	gi 113205316 gb ABI34339.1 多聚蛋白, 3'-部分, 推定的[<i>Solanum demissum</i>]	146435	147770	15		35	55
g100633_Scaffold0002504	gi 113205316 gb ABI34339.1 多聚蛋白, 3'-部分, 推定的[<i>Solanum demissum</i>]	174793	179215	16		36	56
g100620_Scaffold0002504	NA	186534	189120	17		37	57
g100625_Scaffold0002504	gi 460395064 ref XP_004243109.1 预测: 未表征蛋白 LOC101263429 [番茄]	247160	250636	18		38	58
g100615_Scaffold0002504	gi 113205363 gb AAT66771.2 推定多聚蛋白, 相同[<i>Solanum demissum</i>]	263965	268932	19		39	59
g100618_Scaffold0002504	gi 460410504 ref XP_004250667.1 预测: 未表征蛋白 LOC101267192 [番茄]	282902	284495	20		40	60
g100622_Scaffold0002504	NA	305097	306605	21	75	41	61

[0382]

基因	Blast 命中 描述	起始	终止	基因组序列 (SEQ ID)	细化的基因组序列 (SEQ ID)	cDNA 序列 (SEQ ID)	氨基酸序列 (SEQ ID)
g100617_Scaffold0002504	gi 113205345 gb AAT38783.2 假定蛋白 SDM1_46t00006 [<i>Solanum demissum</i>]	351391	354310	22	76	42	62
g100621_Scaffold0002504	gi 113205363 gb AAT66771.2 推定多聚蛋白, 相同[<i>Solanum demissum</i>]	385927	388324	23	77	43	63
g100632_Scaffold0002504	NA	421150	422045	24	78	44	64
g100628_Scaffold0002504	NA	427647	432434	25	79	45	65
g100623_Scaffold0002504	gi 89179421 gb ABD63156.1 反转录转座子 gag 蛋白[芦笋 (<i>Asparagus officinalis</i>)]	471868	472786	26	80*	46	66
g100626_Scaffold0002504	gi 147845547 emb CAN78493.1 假定蛋白 VITISV_037041 [葡萄 (<i>Vitis vinifera</i>)]	477222	483825	27	81	47	67
g100614_Scaffold0002504	gi 460406698 ref XP_004248798.1 预测: 推定晚疫抗性蛋白同源物 R1B-14-样[番茄]	530011	535390	28	82	48	68

[0383] 表11:在Nic2缺失拼接序列(NT1.0-Scaffold0000549(SEQ ID No.2))中注释的基因。“起始”和“终止”表示注释基因在Scaffold NT1.0-Scaffold0002504上的起始和终止核苷酸的位置。在打顶之前和之后,仅观察到三个粗体表示的基因(g38885_Scaffold0000549、g38878_Scaffold0000549和g38864_Scaffold0000549)在BU21和LI BU21的根中表达。在HI BU21或LA BU21根中没有检测到这三个基因的表达。基于根RNASeq数据集,所有其他列出的基因在BU21、HI BU21、LI BU21或LA BU21中均未显示根表达。

[0384]

基因	Blast 命中 描述	起始	终止
g38875_Scaffold000054 9	gi 47824950 gb AAT38724.1 推定反转录转座子蛋白, identical [<i>Solanum demissum</i>]	1064	2604
g38854_Scaffold000054 9	gi 460414233 ref XP_004252477.1 预测: 未表征蛋 白 LOC101245629 [番茄]	12488	12827
g38847_Scaffold000054 9	gi 460407027 ref XP_004248959.1 预测: 未表征蛋 白 LOC101266468 [番茄]	40706	42338
g38857_Scaffold000054 9	gi 460410342 ref XP_004250589.1 预测: 未表征蛋 白 LOC101263780 [番茄]	63740	66314
g38861_Scaffold000054 9	gi 460395064 ref XP_004243109.1 预测: 未表征蛋 白 LOC101263429 [番茄]	69289	71053
g38873_Scaffold000054 9	gi 460407027 ref XP_004248959.1 预测: 未表征蛋 白 LOC101266468 [番茄]	71359	78187
g38871_Scaffold000054 9	gi 460415745 ref XP_004253217.1 预测: 未表征蛋 白 LOC101263890 [番茄]	99417	112568
g38850_Scaffold000054 9	NA	136925	137758
g38855_Scaffold000054 9	NA	141858	144628
g38872_Scaffold000054 9	NA	147982	150752
g38849_Scaffold000054 9	gi 147865536 emb CAN81563.1 假定蛋白 VITISV_019697 [葡萄]	151800	153027
g38866_Scaffold000054 9	gi 460407027 ref XP_004248959.1 预测: 未表征蛋 白 LOC101266468 [番茄]	159812	161152
g38853_Scaffold000054 9	gi 460407027 ref XP_004248959.1 预测: 未表征蛋 白 LOC101266468 [番茄]	165374	166727

[0385]

基因	Blast 命中 描述	起始	终止
g38868_Scaffold000054 9	NA	205238	206813
g38884_Scaffold000054 9	gi 460406842 ref XP_004248867.1 预测: 推定核糖核 酸酶 H 蛋白 Atlg65750-样 [番茄]	223785	228636
g38876_Scaffold000054 9	NA	246679	246970
g38856_Scaffold000054 9	NA	263028	264922
g38859_Scaffold000054 9	NA	267551	269920
g38848_Scaffold000054 9	gi 460410377 ref XP_004250606.1 预测: 未表征蛋 白 LOC101247390 [番茄]	274617	278325
g38885_Scaffold000054 9	gi 296278604 gb ADH04266.1 ERF1 [本塞姆氏烟 草]	279007	279583
g38862_Scaffold000054 9	NA	281949	283475
g38877_Scaffold000054 9	gi 147773804 emb CAN60970.1 假定蛋白 VITISV_026408 [葡萄]	313840	316312
g38869_Scaffold000054 9	gi 147775355 emb CAN65719.1 假定蛋白 VITISV_020846 [葡萄]	320254	322003
g38881_Scaffold000054 9	gi 460410342 ref XP_004250589.1 预测: 未表征蛋 白 LOC101263780 [番茄]	336602	338552
g38852_Scaffold000054 9	gi 113205363 gb AAT66771.2 推定多聚蛋白, 相同 [<i>Solanum demissum</i>]	352315	358329
g38886_Scaffold000054 9	gi 460387720 ref XP_004239522.1 预测: 未表征蛋 白 LOC101244956 [番茄]	359255	372066
g38878_Scaffold000054 9	gi 296278604 gb ADH04266.1 ERF1 [本塞姆氏烟 草]	372621	375461
g38860_Scaffold000054 9	gi 460366233 ref XP_004228993.1 预测: 未表征蛋 白 LOC101255727 [番茄]	386045	387443
g38864_Scaffold000054 9	gi 296278604 gb ADH04266.1 ERF1 [本塞姆氏烟 草]	416812	419577
g38863_Scaffold000054 9	NA	449315	449560
g38865_Scaffold000054 9	NA	449684	449943

基因	Blast 命中 描述	起始	终止
[0386] g38879_Scaffold000054 9	gi 470132088 ref XP_004301918.1 预测: 未表征蛋白 LOC101298139 [野草莓 (<i>Fragaria vesca</i>) 亚种. <i>vesca</i>]	456673	458830
	g38867_Scaffold000054 9	gi 460395064 ref XP_004243109.1 预测: 未表征蛋白 LOC101263429 [番茄]	550298 551726
	g38883_Scaffold000054 9	NA	666517 668921
	g38846_Scaffold000054 9	NA	679387 681160
	g38843_Scaffold000054 9	gi 113205316 gb ABI34339.1 多聚蛋白, 3'-部分, 推定的 [<i>Solanum demissum</i>]	696597 697503
	g38880_Scaffold000054 9	gi 4406792 gb AAD20101.1 推定逆转录因子 pol 多聚蛋白[拟南芥 (<i>Arabidopsis thaliana</i>)]	748082 752804
	g38870_Scaffold000054 9	gi 156603850 ref XP_001618917.1 假定蛋白 NEMVEDRAFT_v1g68789 [<i>Nematostella vectensis</i>] gi 156200895 gb EDO26817.1 预测蛋白 [<i>Nematostella vectensis</i>]	766567 768613
	g38874_Scaffold000054 9	gi 460415871 ref XP_004253277.1 预测: 未表征蛋白 LOC101244169 [番茄]	769094 773060
g38844_Scaffold000054 9	gi 460415745 ref XP_004253217.1 预测: 未表征蛋白 LOC101263890 [番茄]	781866	785313

[0387] 实施例8:通过转基因方法开发具有所需烟碱含量的烟草品种。

[0388] 采用过表达和抑制表达两种方法来研究Nic1基因的功能。产生两组转基因植物, 一组使用全长编码序列, 另一组使用RNAi序列。为了表达全长编码序列或RNAi序列, 可以构建表达载体以具有CsVMV启动子和NOS终止子, 以及具有在肌动蛋白2启动子的指导下的卡那霉素选择标记 (NPT II) 并且具有NOS终止子的盒。靶向Nic缺失段中的基因的RNAi构建体的示例性转化盒序列可以在SEQ ID Nos:69和70中找到。本领域普通技术人员理解, 其他靶序列可用于构建RNAi构建体或用于基因沉默的其它转基因方法中 (例如人造微RNA、反式作用siRNA等)。

[0389] 使用DNA轰击或基因枪方法, 将携带目标转基因的核酸构建体引入烟草叶盘。例如, 参见Sanford等, 1993, Methods Enzymol., 217:483-510; 和Okuzaki and Tabei, 2012, Plant Biotechnology, 29:307-310。简而言之, 按如下方法, 将含有转化盒的质粒DNA涂在1 μ m金颗粒 (DNA/金) 上。将1 μ m的金颗粒在180°C下烘烤12小时, 并制备母液 (40mg/ml)。为了制作10次的混合物, 将100 μ l母液与40 μ l表达载体DNA (1 μ g/ μ l)、100 μ l 2.5M CaCl₂和40 μ l 0.1M亚精胺在1.5-ml管中混合。该混合物在13,000 $\times$ g下离心30秒, 并用500 μ l 100%乙醇洗涤沉淀。将DNA/金混合物悬浮于100 μ l水中, 并将10 μ l施加到巨载体上, 干燥, 然后轰击。在PDS-1000/He系统 (Bio-Rad Laboratories, Hercules, CA, USA) 中, 在部分真空 (711mmHg) 下, 使用1,100psi破裂圆盘轰击每个平板两次。Narrow Leaf Madole (NLM) 和Tennessee 90 (TN90) 烟草叶盘用于RNAi构建体和全长基因构建体的转化。将整个烟叶 (长约45 $\times$ 30mm) 置于MS培养基上过夜, 并在第二天用构建体轰击叶盘。然后将叶切成小块 (约5 $\times$ 5mm), 并放回

在TOM培养基(具有20g蔗糖/L、1mg/L IAA和2.5mg/L BAP的MS培养基)上,于27℃生长3-5天,然后转移到含有300mg/l卡那霉素(TOM-Kan)的TOM培养基中生长。每2-3周将组织转移到新的TOM-Kan平板上,持续4-6周(27℃、16小时光照)。轰击4-6周后,将卡那霉素抗性主芽再生。将芽转移到MS-卡那霉素平板上以生长根。然后评估T1植物(和随后世代)的叶和/或根,以确定一种或多种生物碱和/或一种或多种TSNA的量。

[0390] 实施例9:通过随机诱变开发新的nic1突变

[0391] 使用乙磺酸乙酯(EMS)诱变或快速中子轰击对烟草植物进行随机诱变。EMS诱变是在基因组长度上化学诱导的随机点突变。快速中子诱变是将种子暴露于中子轰击,其通过双链DNA断裂引起大量缺失。

[0392] 对于EMS诱变,将1克(大约10,000粒种子)Tennessee 90烟草(TN90)种子在0.1%吐温中洗涤15分钟,然后在30ml ddH₂O中浸泡2小时。然后将一百五十(150)μl的0.5%EMS(Sigma,目录号M-0880)混合到种子/ddH₂O溶液中并在室温下(RT;大约20℃)在工作台下孵育8-12小时(以30rpm旋转)。然后从种子中除去液体,并混入1M NaOH中过夜以进行净化和清理。然后将种子用100ml ddH₂O洗涤2-4小时。然后将洗过的种子悬浮在0.1%琼脂溶液中。

[0393] 将琼脂溶液中的经EMS处理的种子以~2000粒种子/层为单位均匀地分散在浸水的Carolina's Choice烟草混合物(Carolina Soil Company, Kinston, NC)上。然后盖上保鲜膜并置于生长室中。一旦幼苗从土壤中萌发,保鲜膜被刺破以允许湿度逐渐下降。两周后完全除去保鲜膜。将该单位转移到温室,并用NPK肥料施肥。将幼苗重新插入浮盘并生长至移栽大小。随后将植物移植到田地里。在生长期,植物自花授粉以形成M1种子。在成熟阶段,从每个植株收获五个荚果,并为来自每个植物的一组种子单独命名。这形成了M1群体。生长来自每个M0植物的M1种子的复合物,并收集来自M1植物的叶用于DNA提取。扩增并测序靶基因以进行突变鉴定。

[0394] 实施例10:通过定向诱变开发新的nic1突变

[0395] 通过精确的基因组工程技术(例如转录激活子样效应物核酸酶(TALEN)、大范围核酸酶、锌指核酸酶和CRISPR),通过将突变引入Nic1基因座来生产具有低烟碱并同时保持高叶质量的烟草系,。在诸如TN90、K326和Narrow Leaf Madole的商业化烟草品种中进行基因组修饰。

[0396] 例如,来自Nic1基因的特定靶序列可以充当TALEN靶结合区。TALEN位点对于来自Nic1基因座的单基因或DNA片段是特异性的。TALEN区1和2将用于破坏Nic1的关键区域并部分融合在一起以生成延伸缺失,但小于整个Nic1缺失区。预期生成的植物将具有低烟碱含量,并具有高叶质量。

[0397] 基于靶DNA序列,合成用于转录激活因子样(TAL)效应蛋白的序列并将其克隆到植物表达载体中以充当进入载体。取决于目的,使用不同的方案来产生诱变烟草系:1) 将一个或多个进入载体(含有靶TAL的pALCS1)直接转化到烟草原生质体中以产生随机序列缺失或插入诱变烟草系;2) 将左侧和右侧侧接有与靶插入序列同源的序列的供体序列(例如报告基因,例如GUS基因)与一个或多个进入载体(含有靶TALs的pALCS1)共转化到烟草原生质体中以产生含有报告基因的诱变烟草系;和3) 将含有具有点突变的靶TAL的供体序列与一个或多个进入载体(含有靶TAL的pALCS1)共转化到烟草原生质体中以产生具有点突变的诱变

烟草系;4) 含有组织特异性启动子序列的供体序列,以产生以组织特异性方式表达内源基因的突变烟草系;和5) 含有前述供体序列与报告基因构建体的组合的供体序列,以促进突变烟草筛选。

[0398] 从在生长室中的Magenta盒中生长的TN90烟叶分离烟草原生质体。将3-4周龄植物中扩张良好的叶(5cm)从叶中部切成0.5-1mm的叶小条。通过浸渍小条的两侧,将小条转移到制备的酶溶液(1%纤维素酶R10、0.25%离析酶R10、0.4M甘露醇、20mM KCl、20mM MES(pH 5.7)、10mM CaCl_2 、0.1%BSA)中。在黑暗中使用干燥器真空浸渗叶片30min,在室温下、黑暗中持续消化4小时至过夜,而不摇动。将原生质体在100 μm 尼龙过滤器中过滤并用3ml淋巴细胞分离液(Lymphoprep)纯化。将原生质体离心并用W5n溶液(154mM NaCl、125mM CaCl_2 、5mM KCl、2mM MES、991mg/l葡萄糖pH5.7)洗涤,并以 $5 \times 10^5/\text{ml}$ 的浓度悬浮于W5n溶液中。原生质体在冰上保持30min以通过重力沉降在管的底部。去除W5n溶液并在室温下将原生质体重新悬浮在P2溶液中。将50 μl DNA(10-20 μg 质粒)、500 μl 原生质体(2×10^5 原生质体)和550 μl PEG溶液(40%、v/v 10ml 4g PEG4000、0.2M甘露醇、0.1M CaCl_2) 在15-ml微量离心管中缓慢混合,将混合物在室温下孵育5min。

[0399] 将原生质体沉淀,并重悬于1mM M 2X 8EN1(8EN1:不含 NH_4NO_3 的MS盐、MS维生素、0.2%肌肉肌醇、4mM MES、1mg/l NAA、1mg/l IAA、0.5M甘露醇、0.5mg/l BAP、1.5%蔗糖)。用等量低熔琼脂糖(LMA)将转化的原生质体胶化,并且将0.2ml的原生质体-LAM滴下以形成珠。向该珠中添加10ml 8EN1,7天后取出5ml 8EN1并添加5ml 8EN2(含0.25M甘露醇的8EN1);再过7天(14天),取出10ml 8EN2并添加10ml 8EN2;再过7天(21天)中,取出5ml 8EN2并添加5ml 8EN3(含3%蔗糖并且不含甘露醇的8EN1);再过7天(28天)后,取出10ml 8EN3并添加10ml 8EN3。保持原生质体两周直至微小愈伤组织生长。将愈伤组织转移到NCM固体培养基中,直到达到约5mm(通常约两周)。将愈伤组织转移到TOM-Kan固体培养基中以生长芽,并使用本文所述的方法再生转化的烟草植物。

[0400] 实施例11:对鉴定的Nic1损伤进行进一步遗传确认。

[0401] 对来自TN90与LA BU21之间杂交的F2群体5的22株植物进行遗传分离分析(表12)。如实施例3所述,对该F2群体((LA白肋21x TN 90LC)(X))进行Nic1和Nic2基因座的基于PCR的基因分型。还如实施例3一样,测量了每株植物的总生物碱百分比和烟碱含量百分比。简而言之,在中间生长阶段(分支阶段)收集烟草样品用于基因分型。在伸长芽阶段,将植物打顶。将植物打顶后两周,取样用于化学分析的样品。

[0402] 基因型数据再次确认了F2群体中Nic1和Nic2基因座的分离比为9:3:3:1。四种基因型中每一种的平均总生物碱百分比和平均烟碱含量百分比也与分离数据和先前的观察结果(nic1比nic2对总生物碱和烟碱含量具有更强的作用)一致(表13和14)。根据其基因型(x轴)和每种植物的化学物(y轴)作图,图3和图4进一步示出了该F2群体的基因型和化学数据。图4与图3的不同之处在于排除了由于采样或处理错误而可能交换的8株植物。

[0403] 此外,在两年内测试了等基因BU21低生物碱系列种质组(LA BU21、LI BU21、HI BU21和BU21)。低生物碱性状不受叶生物碱含量每年变化的影响。如图5和图6所示,烟碱含量的降低百分比在两年内一致,表明低生物碱性状不受环境因素或田地条件的影响。

[0404] 表12:用于分离nic1和nic2((LA白肋21x TN 90LC)(X))的较大F2群体的单个植物的Nic1和Nic2基因型和生物碱测量。N/S表示没有样品。星号表示在取样、加工或分析阶段

DNA或叶样品可能交换的四对相邻植物(共8个,植物编号分别为163和164、174和175、361和362、481和482,也用黑体标出)。四对植物中的每一对在田地中是并排的。

[0405]

植物编号	Nic1 (存在=1, 缺失=0)	Nic2 (存在=1, 缺失=0)	烟碱百分比	总生物碱百 分比
1	1	1	1.33	1.389
2	1	1	1.59	1.66
3	0	1	0.72	0.7511
4	1	1	2.59	2.753
5	1	0	0.81	0.8422
6	1	1	1.04	1.096
7	1	1	1.76	1.843
8	1	1	1.07	1.121
9	1	0	1.85	1.938
10	1	1	1.58	1.662
11	1	0	1.1	1.147
12	1	0	0.68	0.7045
13	1	1	1.33	1.39
14	1	0	1.77	1.828
15	1	1	1.83	1.93
16	1	1	3.34	3.513
17	0	0	0.059	0.0643
18	1	1	3.36	3.543
19	0	0	0.062	0.0661
20	1	1	2.31	2.437
21	1	0	1.23	1.286
22	1	0	1.61	1.701
23	1	1	1.18	1.241
24	0	0	0.106	0.1099
25	0	1	1.41	1.487
26	1	1	2.47	2.6
27	1	1	2.16	2.245
28	1	1	2.22	2.334
29	1	0	2.54	2.667
30	1	1	1.37	1.429
31	1	1	1.86	1.956

[0406]

植物编号	Nic1 (存在=1, 缺失=0)	Nic2 (存在=1, 缺失=0)	烟碱百分比	总生物碱百 分比
32	1	1	1.68	1.762
33	1	0	1.91	1.996
34	1	1	2.89	3.051
35	0	1	1.15	1.205
36	1	1	2.92	3.064
37	1	0	2.84	3.005
38	0	1	1.49	1.566
39	0	0	0.17	0.1796
40	1	0	3.02	3.211
41	1	0	1.29	1.35
42	1	1	1.73	1.805
43	1	0	1.37	1.469
44	1	1	2.18	2.279
45	0	0	0.16	0.1719
46	1	0	2.38	2.495
47	1	1	2.34	2.455
48	1	0	1.39	1.455
49	1	0	2.9	3.044
50	0	1	2.18	2.307
51	0	1	1.34	1.393
52	1	0	1.9	1.98
53	1	1	1.49	1.566
54	1	1	2.22	2.347
55	1	1	3.21	3.359
56	1	1	3.67	3.862
57	1	1	2.07	2.17
58	0	1	0.83	0.8704
59	1	1	2.37	2.492
60	1	1	2.25	2.395
61	1	1	2.53	2.678
62	1	1	2.19	2.326
63	1	1	1.92	1.997
64	1	1	3.06	3.229
65	1	1	3.2	3.369
66	1	1	2.83	2.997
67	1	1	4.07	4.257

[0407]

植物编号	Nic1 (存在=1, 缺失=0)	Nic2 (存在=1, 缺失=0)	烟碱百分比	总生物碱百 分比
68	1	1	2.4	2.498
69	0	1	1.66	1.734
70	1	1	1.93	2.045
71	1	1	3.79	3.99
72	0	0	0.079	0.0849
73	1	1	2.17	2.289
74	1	0	3.58	3.778
75	1	1	2.63	2.76
76	0	1	1.52	1.603
77	1	1	3.04	3.211
78	1	1	2.82	2.979
79	1	1	2.6	2.74
80	1	1	2.48	2.597
81	1	1	1.88	1.977
82	1	1	2.98	3.106
83	0	1	0.99	1.0255
84	0	0	0.3	0.3194
85	1	1	1.79	1.873
86	1	1	2.37	2.477
87	1	1	2.79	2.932
88	1	1	2.14	2.301
89	1	1	4.35	4.588
90	0	0	0.09	0.0972
91	0	1	1.59	1.67
92	1	0	1.71	1.789
93	1	1	3.94	4.114
94	1	1	2.65	2.755
95	0	1	1.14	1.211
96	1	0	2.14	2.248
97	1	1	2.62	2.743
98	1	0	0.95	0.9865
99	1	1	2.61	2.729
100	1	1	3.33	3.485
101	1	0	2.77	2.878
102	1	1	2.31	2.424
103	0	1	0.49	0.5099

[0408]

植物编号	Nic1 (存在=1, 缺失=0)	Nic2 (存在=1, 缺失=0)	烟碱百分比	总生物碱百 分比
104	1	1	2.5	2.611
105	0	1	0.97	1.012
106	1	1	2.29	2.404
107	1	1	1.11	1.16
108	1	0	1.87	1.947
109	0	1	0.95	1
110	1	1	3.05	3.209
111	1	0	3.28	3.437
112	1	0	3.46	3.637
113	1	1	4.33	4.594
114	1	1	4.35	4.552
115	1	1	3.56	3.746
116	0	1	1.85	1.93
117	1	1	5.94	6.305
118	0	1	1.47	1.559
119	1	1	3.8	4.028
120	0	1	1.9	2.007
121	1	0	3.38	3.573
122	0	0	0.3	0.3191
123	1	1	4.04	4.245
124	1	1	3.4	3.622
125	0	1	1.09	1.141
126	1	0	2.71	2.828
127	1	1	4.93	5.178
128	1	0	2.65	2.801
129	1	1	1.62	1.693
130	1	0	2.13	2.232
131	0	1	2.64	2.789
132	1	1	3.17	3.329
133	0	1	1.99	2.108
134	1	1	2.28	2.377
135	1	1	3.85	4.019
136	1	1	3.62	3.792
137	1	0	3.19	3.355
138	0	1	1.49	1.565
139	1	1	4.33	4.516

[0409]

植物编号	Nic1 (存在=1, 缺失=0)	Nic2 (存在=1, 缺失=0)	烟碱百分比	总生物碱百 分比
140	1	1	3.88	4.097
141	1	1	5.68	6.003
142	1	1	3.15	3.292
143	1	0	2.35	2.456
144	1	0	3.63	3.786
145	1	1	3.32	3.532
146	1	1	3.26	3.407
147	1	1	3.37	3.507
148	1	0	3.21	3.388
149	1	1	3.7	3.873
150	0	0	0.4	0.4279
151	1	1	1.41	1.482
152	1	1	2.18	2.273
153	1	1	2.58	2.725
154	0	1	2.89	3.088
155	1	1	1.51	1.593
156	0	1	2.03	2.143
157	0	1	2.03	2.151
158	1	0	1.27	1.323
159	1	1	1.62	1.734
160	1	1	4.48	4.67
161	1	1	1	1.045
162	1	1	3.17	3.309
163*	1	1	0.18	0.1906
164*	0	1	2.38	2.493
165	1	1	3.62	3.745
166	1	1	N/S	N/S
167	1	0	2.62	2.717
168	1	1	2.31	2.428
169	1	0	2.52	2.645
170	1	1	4.35	4.527
171	1	0	1.66	1.741
172	1	0	3.03	3.159
173	1	0	3.2	3.343
174*	1	1	0.059	0.0653
175*	0	0	4.03	4.197

[0410]

植物编号	Nic1 (存在=1, 缺失=0)	Nic2 (存在=1, 缺失=0)	烟碱百分比	总生物碱百 分比
176	1	1	2.85	3.01
177	1	1	2.91	3.086
178	1	1	4.21	4.395
179	1	1	1.09	1.154
180	0	1	1.15	1.206
181	1	0	3.2	3.345
182	1	1	3.28	3.406
183	1	1	3.17	3.324
184	1	0	3.24	3.392
185	1	1	0.94	0.987
186	1	1	3.95	4.123
187	1	0	0.68	0.7094
188	0	1	1.94	2.007
189	1	1	0.46	0.4809
190	0	1	2.28	2.356
191	1	1	2.25	2.336
192	1	0	2.74	2.842
193	1	1	3.9	4.091
194	1	1	3.39	3.578
195	1	0	2.59	2.698
196	0	0	0.131	0.1383
197	1	1	2.28	2.365
198	1	0	2.27	2.395
199	0	0	0.1	0.1063
200	1	1	2.39	2.496
201	0	1	1.02	1.09
202	1	1	1.99	2.09
203	1	1	3.44	3.594
204	1	1	3.05	3.167
205	0	1	0.85	0.8839
206	0	1	2.18	2.331
207	1	1	1.89	1.971
208	1	1	1.37	1.445
209	1	1	3.26	3.432
210	1	1	1.28	1.3224
211	0	1	1.86	1.966

[0411]

植物编号	Nic1 (存在=1, 缺失=0)	Nic2 (存在=1, 缺失=0)	烟碱百分比	总生物碱百 分比
212	1	1	1.53	1.59
213	0	0	0.142	0.1513
214	1	1	2.22	2.331
215	1	1	2.94	3.075
216	0	1	1.72	1.801
217	1	1	3.89	4.044
218	1	1	4.47	4.691
219	0	1	1.63	1.696
220	1	0	3.28	3.432
221	1	0	1.14	1.193
222	1	1	3.91	4.089
223	1	0	0.87	0.9002
224	1	1	1.8	1.891
225	1	1	2.25	2.369
226	1	0	2.59	2.738
227	0	1	0.89	0.933
228	0	1	0.64	0.669
227B	1	1	1.38	1.439
228B	1	1	3.25	3.421
229	1	0	3.57	3.71
230	0	1	2.43	2.583
231	1	1	3.23	3.398
232	1	1	4.85	5.187
233	0	1	1.09	1.138
234	0	0	0.24	0.2521
235	1	1	2.49	2.597
236	1	1	4.07	4.327
237	0	1	1.17	1.235
238	0	0	0.16	0.1683
239	1	1	3.18	3.323
240	1	0	2.39	2.5
241	1	1	1.72	1.805
242	0	1	0.95	0.992
243	0	1	0.75	0.787
244	1	1	2.28	2.398
245	1	1	1.89	1.989

[0412]

植物编号	Nic1 (存在=1, 缺失=0)	Nic2 (存在=1, 缺失=0)	烟碱百分比	总生物碱百 分比
246	1	0	3.91	4.084
247	1	0	2.53	2.648
248	0	1	0.68	0.7146
249	1	1	2.82	2.966
250	1	1	1.76	1.841
251	0	0	0.18	0.1936
252	1	1	1.39	1.452
253	1	1	3.54	3.722
254	1	1	4.05	4.247
255	1	1	2.09	2.192
256	1	1	1.47	1.551
257	1	0	2.54	2.677
258	1	1	3.83	4.002
259	1	0	1.23	1.301
260	1	1	1.92	2.017
261	1	1	2.17	2.323
262	1	0	2.41	2.552
263	1	0	2.3	2.435
262B	1	0	1.34	1.408
263B	1	1	3.39	3.577
264	0	1	1.98	2.093
265	0	1	0.8	0.854
266	1	1	1.76	1.863
267	1	1	1.87	1.989
268	0	0	0.19	0.2031
269	0	0	0.12	0.1293
270	0	1	0.57	0.607
271	0	0	0.139	0.1481
272	1	0	1.81	1.893
273	1	0	2.13	2.244
274	0	1	0.52	0.5523
275	0	1	1.7	1.805
276	1	0	1.83	1.922
277	1	1	2	2.101
278	0	1	0.91	0.9535
279	1	1	3.25	3.439

[0413]

植物编号	Nic1 (存在=1, 缺失=0)	Nic2 (存在=1, 缺失=0)	烟碱百分比	总生物碱百 分比
280	1	1	3.45	3.586
281	1	1	4.62	4.807
282	1	1	3.41	3.566
283	0	1	1.34	1.407
284	1	1	2.43	2.556
285	1	1	1.44	1.514
286	1	1	2.49	2.614
287	0	0	0.137	0.1445
288	1	1	1.78	1.863
289	0	1	0.53	0.5681
290	1	1	N/S	N/S
291	0	1	0.73	0.775
292	1	1	1.43	1.504
293	1	1	1.92	2.012
294	0	1	0.51	0.5388
295	1	0	2.07	2.181
296	1	1	1.4	1.463
297	1	1	4.21	4.548
298	1	0	3.11	3.225
299	1	1	3.93	4.158
300	1	1	1.99	2.062
301	0	1	0.56	0.5858
302	1	1	1.93	2.018
303	1	1	1.06	1.114
304	1	1	6.89	7.302
305	1	0	2.25	2.366
306	1	0	1.42	1.479
307	1	0	2.39	2.591
308	1	1	2.52	2.635
309	1	0	1.84	1.933
310	1	1	2.29	2.412
311	1	1	1.62	1.707
312	0	1	1.89	1.988
313	1	1	2.48	2.577
314	0	0	1.01	1.056
315	1	1	2.21	2.334

[0414]

植物编号	Nic1 (存在=1, 缺失=0)	Nic2 (存在=1, 缺失=0)	烟碱百分比	总生物碱百 分比
316	1	1	1.54	1.617
317	1	1	N/S	N/S
318	0	1	0.65	0.6821
319	1	1	3.32	3.493
320	1	0	2.68	2.8
321	1	1	2.71	2.835
322	1	0	1.97	2.068
323	1	1	2.76	2.905
324	1	0	3.92	4.122
325	1	0	2.56	2.724
326	1	1	3.18	3.355
327	0	1	2.08	2.165
328	1	0	2.77	2.89
329	1	1	2.11	2.216
330	1	0	1.81	1.878
331	0	0	0.2	0.2096
332	1	1	1.8	1.894
333	0	1	0.75	0.7875
334	1	0	2.78	2.92
335	0	0	0.19	0.2044
336	1	1	2.2	2.31
337	1	1	1.54	1.606
338	0	0	0.19	0.1986
339	1	1	2.05	2.153
340	1	0	1.57	1.644
341	1	1	2.64	2.768
342	1	1	2.4	2.522
343	1	1	2.1	2.213
344	1	1	2.83	2.989
345	1	0	2.77	2.899
346	1	1	1.44	1.509
347			2.89	3.042
348	1	1	1.76	1.841
349	0	1	0.59	0.6235
350	1	1	3.62	3.804
351			0.67	0.707

[0415]

植物编号	Nic1 (存在=1, 缺失=0)	Nic2 (存在=1, 缺失=0)	烟碱百分比	总生物碱百 分比
352	1	0	2.25	2.352
353	1	1	1.7	1.786
354	1	1	2.36	2.463
355	1	1	N/S	N/S
356	1	1	3.36	3.549
357	1	1	1.63	1.706
358	0	1	0.84	0.887
359	0	1	0.5	0.5223
360	1	1	2.44	2.568
361*	0	1	1.86	1.931
362*	0	0	2.98	3.122
363	0	0	0.072	0.0754
364	0	1	1.64	1.739
365	1	1	3.4	3.602
366	1	0	2.72	2.872
367	0	1	1.12	1.205
368	1	1	3	3.207
369	1	1	1.45	1.542
370	1	1	4.65	4.975
371	1	1	3.1	3.271
372	1	1	1.2	1.252
373	1	1	1.94	2.049
374	0	0	0.24	0.2523
375	1	1	3.4	3.528
376	1	0	2.73	2.863
377	0	1	0.57	0.5979
378	1	1	2.34	2.447
379	1	1	2.9	3.073
380	1	1	2.03	2.156
381	1	1	2.07	2.177
382	0	1	1.23	1.301
383	1	1	2.42	2.554
384	1	1	3.32	3.518
385	1	1	2.18	2.293
386	1	1	1.58	1.668
387	0	1	1.21	1.285

[0416]

植物编号	Nic1 (存在=1, 缺失=0)	Nic2 (存在=1, 缺失=0)	烟碱百分比	总生物碱百 分比
388	0	0	0.074	0.0789
389	1	1	1.65	1.725
390	1	1	2.36	2.507
391	1	1	1.94	2.055
392	1	1	2	2.102
393	1	1	2.61	2.705
394	0	1	1	1.06
395	1	1	2.53	2.641
396	1	1	1.03	1.081
397	0	0	0.136	0.1417
398	1	1	0.99	1.033
399	1	1	1.96	2.062
400	0	0	0.113	0.1193
401	1	1	1.75	1.841
402	1	1	2.34	2.442
403	1	1	1.92	2.002
404	1	1	3.87	4.06
405	1	1	2.94	3.068
406	1	1	1.15	1.204
407	1	1	3.15	3.359
408	1	1	1.83	1.923
409	1	1	2.28	2.402
410	1	1	2.76	2.906
411	1	1	1.88	1.994
412	1	1	1.28	1.342
413	1	1	1.12	1.175
414	0	1	1.28	1.341
415	1	1	2.22	2.334
416	0	1	1.05	1.099
417	1	1	3.01	3.148
418	1	1	3.27	3.444
419	1	1	1.77	1.86
420	1	1	1.44	1.497
421	0	1	0.55	0.5777
422	1	0	2.7	2.825
423	1	1	1.24	1.315

[0417]

植物编号	Nic1 (存在=1, 缺失=0)	Nic2 (存在=1, 缺失=0)	烟碱百分比	总生物碱百 分比
424	1	1	2.95	3.106
425	0	1	0.63	0.6607
426	1	1	2.88	3.014
427	1	1	2.31	2.455
428	1	0	1.92	1.994
429	1	1	3.05	3.232
430	1	1	1.76	1.88
431	0	0	0.11	0.1164
432	0	1	0.78	0.8129
433	0	1	2.01	2.134
434	1	1	1.98	2.095
435			1.71	1.801
436	1	0	3.27	3.447
437	0	1	0.89	0.944
438	1	1	2.9	3.049
439	0	0	0.18	0.1905
440	1	1	4.12	4.367
441	0	1	1.53	1.623
442	1	1	4.07	4.315
443	0	1	1.35	1.429
444	1	0	2.98	3.12
445	1	1	4.98	5.272
446	0	1	1.22	1.288
447	1	1	4.34	4.578
448	1	1	2.89	3.028
449	1	1	5.8	6.135
450	1	1	4.34	4.553
451	1	0	3.6	3.791
452	1	1	4.67	4.947
453	0	1	1.83	1.918
454	1	1	2.6	2.692
455	1	1	5.03	5.301
456	1	1	4.41	4.695
457	1	0	2.62	2.741
458	0	0	0.133	0.1425
459	0	1	0.53	0.5534

[0418]

植物编号	Nic1 (存在=1, 缺失=0)	Nic2 (存在=1, 缺失=0)	烟碱百分比	总生物碱百 分比
460	0	1	0.63	0.6663
461	0	1	2.78	2.943
462	1	1	2.64	2.785
463	0	1	1.23	1.296
464	0	1	1.55	1.629
465	1	1	2.49	2.626
466	1	1	2.8	2.936
467	0	0	0.19	0.2026
468	0	1	1.91	2.029
469	1	1	1.37	1.442
470	0	1	1.29	1.375
471	1	1	3.74	3.919
472	0	1	1.67	1.78
473	1	1	1.39	1.479
474	1	1	0.88	0.926
475			2.53	2.679
476	1	1	3.59	3.756
477	1	1	3.96	4.196
478	1	1	1.59	1.679
479	0	1	2.27	2.385
480	0	0	0.108	0.1147
481*	0	1	2.21	2.363
482*	1	1	0.54	0.5726
483	1	1	1.76	1.854
484	1	1	3.34	3.512
485	1	1	1.4	1.468
486	1	0	0.87	0.9085
487	0	1	0.38	0.3976
488	1	1	2.53	2.662
489	0	1	1.34	1.412
490	0	1	1.17	1.246
491	0	1	0.43	0.4505
492	1	0	3.55	3.741
493	0	1	2.16	2.294
494	1	1	2.6	2.736
495	1	1	2.73	2.905

[0419]

植物编号	Nic1 (存在=1, 缺失=0)	Nic2 (存在=1, 缺失=0)	烟碱百分比	总生物碱百分比
496	1	1	2.46	2.591
497	1	1	1.8	1.906
498	0	1	0.75	0.7913
499	1	1	1.37	1.461
500	1	0	1.63	1.731
501	1	1	1.34	1.417
502	1	0	0.98	1.0156
503	1	1	0.79	0.8276
504	1	0	1.75	1.812
505	1	1	1.49	1.571
506	0	0	0.084	0.0892
507	1	1	1.2	1.247
508	1	0	1.98	2.081
509	1	0	1.03	1.0704
510	1	0	1.4	1.475
511	1	0	0.76	0.791
512	1	1	1.4	1.491
513	1	1	2.29	2.402
514	1	1	1.14	1.19
515	1	1	1	1.044
516	1	1	1.26	1.356
517	0	1	0.52	0.5535
518	1	1	0.45	0.4786
519	1	0	1.59	1.651
520	1	1	1.59	1.667
521	0	1	0.46	0.4803
522	1	0	1.1	1.162

[0420] 表13:F2分离群体 ((LA白肋21x TN 90LC) (X)) 的遗传分离以及四种分离基因型中的每一种的总生物碱含量平均百分比和烟碱含量平均百分比。显示了平均值和标准偏差。三株丢失样品的植物被排除,使得分析的植物总数为518株。

[0421]

基因型	预期比	预期的 植物数	观察到的 植物数量	烟碱平均 百分比	标准偏差 (烟碱)	总生物 碱平均	标准偏差 (总生物
-----	-----	------------	--------------	-------------	--------------	------------	--------------

[0422]

		量				百分比	碱)
Nic1 Nic2	9/16	291.9	288	2.55	1.08	2.68	1.14
Nic1 <i>nic2</i>	3/16	97.3	96	2.25	0.83	2.36	0.87
<i>nic1</i> Nic2	3/16	97.3	97	1.30	0.61	1.37	0.65
<i>nic1 nic2</i>	1/16	32.4	37	0.36	0.78	0.38	0.81

[0423] 表14:表12中的分离群体的再分析进一步排除了可能交换的四对植物(植物编号163和164、174和175、361和362、481和482),使得总植物数量为510。

	基因型	预期比	预期的植物数量	观察到的植物数量	烟碱平均百分比	标准偏差(烟碱)	总生物碱平均百分比	标准偏差(总生物碱)
[0424]	Nic1 Nic2	9/16	286.88	285	2.57	1.06	2.71	1.11
	Nic1 <i>nic2</i>	3/16	95.63	96	2.25	0.83	2.36	0.87
	<i>nic1</i> Nic2	3/16	95.63	95	1.28	0.60	1.35	0.64
	<i>nic1 nic2</i>	1/16	31.88	34	0.18	0.16	0.19	0.17

[0425] 实施例12:基因组重测序以验证和细化所鉴定的Nic1病变中的序列。

[0426] 使用PacBio的SMRT测序技术(NT2.0)对鉴定的Nic1拼接序列和注释基因的基因组序列进行重测序。Nic1基因座区(NT1.0-Scaffold0002504)的序列质量得到改善。简而言之,NT2.0-Scaffold4274(SEQ ID No:73)的碱基对(bp)1-133,550对应于并以负方向取代NT1.0-Scaffold0002504的bp384,701-542,313。NT2.0-Scaffold14415(SEQ ID No:74)的碱基对1-59,671对应于并以负方向取代NT1.0-Scaffold0002504的bp 288,601-363,040。

[0427] 对于之前在SEQ ID Nos:21-28中鉴定的8个基因,SEQ ID Nos:75-82分别提供额外的基因组序列数据。还鉴定了一组新的“g100623_Scaffold0002504”的cDNA和蛋白序列(新序列SEQ ID No:83和84分别相对于之前的序列SEQ ID No:46和66)。

[0428] 本申请中提及的每一个美国或者外国专利、公开的专利申请、非专利文献或者任何其他参考文献的全部内容通过引用并入本文。

[0429] 实施例13:使用细菌人工染色体测序对Nic1损伤区完整测序。

[0430] 使用细菌人工染色体(BAC)步行方法来制备Nic1基因座周围的完整序列,并进一步鉴定基因组缺失的边界序列。简而言之,鉴定与全基因组谱(WGP)重叠群相关的21个BAC克隆,并将其合并以形成3个各具有7个BAC的池。每个池在单独的SMRT细胞(PacBio)上分开测序。将序列组装成14个重叠群,基于通过在每个重叠群任一端检索(blast)1000bp获得的重叠序列信息,将其进一步组装成一个重叠群。所得的参考重叠群用于重新映射来自BU21、HI、LI和LA BU21系的重测序数据。鉴定了跨越整个Nic1缺失的连续425,001bp的区域(SEQ ID NO:85)。表15提供基于最佳BLAST命中的十五个鉴定的具有编码蛋白的功能信息的基因模型的细节。这15个基因在SEQ ID NO:85上的相对位置可以在图7中找到。提供十五个Nic1缺失基因(NDG1-NDG15)的基因组序列(SEQ ID NO:86-100)、编码序列(CDS;SEQ ID NOs:101-115)和蛋白序列(SEQ ID NO:116-130)。在这些中,SEQ ID NOs:86、87和90-100中提供的基因组序列数据分别对应于之前在SEQ ID NOs:28、13、20、22、11、17、23、26、21、27、14、25、和24中鉴定的13个基因模型。设计RNAi构建体以沉默NDG1-NDG15,并且还用于制备具有更低生物碱含量的转基因烟草植物。

[0431] 实施例14:开发用于选择*nic1*和*nic2*的额外的SNP标记。

[0432] 为了提供对包含*nic1*和*nic2*突变的植物的更有效的选择,鉴定了*nic1*和*nic2*缺失两侧的12个SNP(每个基因座6个SNP)以开发共显性标记。使用每个缺失区上游的3个SNP和下游的3个SNP。相应的缺失区的每个SNP多态性小于100kb。表16提供每个SNP的相对位置和序列多态性的更多细节,并且还提供每个SNP的上游60bp和下游60bp的序列信息。基于已鉴

定的SNP,设计高通量Kompetitive等位基因特异性PCRTM或KASPTM测定(LGC Genomics, Beverly, Massachusetts)进行基因分型。这些SNP标记可用来检测nic1或nic2缺失的杂合植物,并且也可以区分来自不良PCR反应的真正的纯合缺失。表18示出了用于对实施例11(参见表12的植物编号1-72)的F2群体((LA白肋21x TN 90LC)(X))进行基因分型的两个SNP标记(Seq ID Nos.135和137)的有效性。

[0433] 实施例15:鉴定与nic1相关的ERF基因。

[0434] 鉴定了Nic1缺失区的下游307,823bp处的名为“ERF-39样”的ERF基因。鉴于该基因紧密靠近,鉴定了两个相邻的SNP标记(SEQ ID Nos.143和144)用于基因分型和追踪目的。表17提供这两个SNP的细节。SEQ ID Nos:145-147分别提供了ERF-39样基因的基因组、cDNA和蛋白序列。制备ERF-39基因过表达或下调的植物(例如通过RNAi、人工miRNA、化学诱变或目标基因组编辑),并评估该基因对植物烟碱含量的影响。

[0435] 实施例16:通过基因沉默和基因组编辑开发低生物碱烟草植物

[0436] 通过目标基因组编辑技术,通过向Nic1基因座或Nic2基因座引入功能缺失或无效突变,生产具有低烟碱并同时保持高叶质量的烟草系。如实施例10所述,基因组编辑技术的实例是转录激活因子样效应物核酸酶(TALEN)、大范围核酸酶、锌指核酸酶和CRISPR。在诸如TN90、K326和Narrow Leaf Madole的商业化烟草品种中进行基因组修饰。通过基因沉默(例如,RNAi、人造/合成微RNA或其他小RNA介导的技术),产生以下具有抑制表达的其他烟草系:一种或多种NDG、Nic2基因座上的一个或多个ERF基因,或其组合。表15列出了靶向NDG1-NDG15的示例性RNAi盒序列。

[0437] 例如,按实施例10所述的方法制备原生质体。合成含有用于转录激活因子样效应蛋白,对Nic1(例如NDG1或NDG2)、ERF-39样和Nic2(例如ERF189或来自表11的另一ERF基因)的靶序列特异性序列的一个或多个进入载体,并将其转化入分离的原生质体。将原生质体培养成愈伤组织。或者,通过CRISPR实现Nic1和Nic2基因座突变,而不使用质粒。在这种情况下,用Cas9、引导RNA和聚乙二醇(PEG)转化原生质体。

[0438] 生成的植物在Nic1、Nic2或两个基因座中含有基因突变(例如,无效突变或缺失),并构成低烟碱品种。或者,将含有Nic1突变等位基因的品种与含有Nic2突变等位基因的另一品种杂交。这种杂交将导致具有所需遗传背景的品种具有甚至更低的烟碱含量,但高叶质量。作为替代选择,可以如实施例10所述利用基因沉默和基因组编辑技术来同时引入Nic1和Nic2突变或缺失。

[0439] 表15:该表提供Nic1基因座内的编码序列(CDS)的位置,以及具有相应蛋白的最低E值的BLAST命中的细节。含反向重复序列的RNAi盒的序列以SEQ ID Nos列出。RNAi载体的T-DNA盒序列可以在SEQ ID Nos.69和70中找到。

[0440]

名称	SEQ ID NO: 85 上的位置 坐标	BLAST 命中 描述	Blast 命中 ID	基因组 序列 (SEQ ID)	cDNA 序列 (SEQ ID)	氨基酸 序列 (SEQ ID)	表 10 的 相应基 因	RNAi 盒 序列 (SEQ ID)
NDG1	连接 (385015..385 449,387135.. 387299,3873 83..387795,3 87857..3892 64)	预测的: 推定 晚疫抗性蛋 白同系物 R1A-3 [绒毛 状烟草]	XP_009624787.1	86	101	116	g100614 _Scaffold 0002504	148 和 149
NDG2	互补(连接 (73886..7431 1,74491..746 80,74794..74 934,75026..7 6149,76285.. 76524))	预测的: 推定 晚疫抗性蛋 白同系物 R1B-16 [绒 毛状烟草]	XP_009589795.1	87	102	117	g100631 _Scaffold 0002504	150 和 151

[0441]

名称	SEQ ID NO: 85 上的位置 坐标	BLAST 命中 描述	Blast 命中 ID	基因组 序列 (SEQ ID)	cDNA 序列 (SEQ ID)	氨基酸 序列 (SEQ ID)	表 10 的 相应基 因	RNAi 盒 序列 (SEQ ID)
NDG3	连接 (323943..323 949,324085.. 324464)	预测的: 未表 征蛋白 LOC1040907 58 [绒毛状烟 草]	XP_009594228.1	88	103	118		152
NDG4	315634..316 095	预测的: 未表 征蛋白 LOC1041065 16 [绒毛状烟 草]	XP_009613372.1	89	104	119		153
NDG5	互补(连接 (179863..179 982,180074.. 180152,1808 39..181455))	预测的: 未表 征蛋白 LOC1040917 89 [绒毛状烟 草]	XP_009595504.1	90	105	120	g100618 _Scaffold 0002504	154
NDG6	互补(连接 (233956..234 093,234729.. 235030,2351 66..235378,2 35496..2360 68,236374..2 36438,23663 3..236874))	预测的: 未表 征蛋白 LOC1040933 28 [绒毛状烟 草]	XP_009597352.1	91	106	121	g100617 _Scaffold 0002504	155
NDG7	连接 (36266..3662 9,36753..369 11,37083..37 384,37655..3 7736,37857.. 38347,39697 ..39783)	预测的: 未表 征蛋白 LOC1070012 34 [<i>Solanum pennellii</i>]	XP_015054850.1	92	107	122	g100619 _Scaffold 0002504	156

[0442]

名称	SEQ ID NO: 85 上的位置 坐标	BLAST 命中 描述	Blast 命中 ID	基因组 序列 (SEQ ID)	cDNA 序列 (SEQ ID)	氨基酸 序列 (SEQ ID)	表 10 的 相应基 因	RNAi 盒 序列 (SEQ ID)
NDG8	连接 (114418..114 612,115436.. 115956,1162 68..116427,1 16558..1167 49,116833..1 17003)	预测的: 未表 征蛋白 LOC1040865 61 [绒毛状烟 草]	XP_009589152.1	93	108	123	g100620 _Scaffold 0002504	157
NDG9	互补(连接 (262599..262 791,264178.. 264595,2646 77..264995))	预测的: 未表 征蛋白 LOC1070299 76 [<i>Solanum pennellii</i>]	XP_015086886.1	94	109	124	g100621 _Scaffold 0002504	158
NDG10	335270..336 187	预测的: 未表 征蛋白 LOC1042175 32 [林烟草]	XP_009766114.1	95	110	125	g100623 _Scaffold 0002504	159
NDG11	互补(连接 (196787..197 330,197459.. 197772,1978 60..198026,1 98131..1982 93))	预测的: 未表 征蛋白 LOC1041078 87 [绒毛状烟 草]	XP_009615097.1	96	111	126	g100622 _Scaffold 0002504	160
NDG12	互补(连接 (340626..340 850,340943.. 341242,3413 77..342443,3 42827..3428 91,345833..3 46446,34678 2..347228))	预测的: 未表 征蛋白 LOC1042255 35 [林烟草]	XP_009775660.1	97	112	127	g100626 _Scaffold 0002504	161

[0443]

名称	SEQ ID NO: 85 上的位置 坐标	BLAST 命中 描述	Blast 命中 ID	基因组 序列 (SEQ ID)	cDNA 序列 (SEQ ID)	氨基酸 序列 (SEQ ID)	表 10 的 相应基 因	RNAi 盒 序列 (SEQ ID)
NDG13	连接 (79509..7974 4,79863..801 41,80301..80 715)	预测的: 未表 征蛋白 LOC1040955 12 [绒毛状烟 草]	XP_009599950.1	98	113	128	g100627 _Scaffold 0002504	162
NDG14	互补(连接 (297606..297 727,299439.. 299551,3002 49..300331,3 00629..3008 02,301088..3 01258,30132 6..301450,30 1535..30163 6,301731..30 1859,301942 ..302392))	预测的: 未表 征蛋白 LOC1042159 65 [林烟草]	XP_009764210.1	99	114	129	g100628 _Scaffold 0002504	163
NDG15	连接 (291110..291 242,291376.. 291556,2916 65..292004)	预测的: DNA 连接酶 1-样 [林烟草]	XP_009785055.1	100	115	130	g100632 _Scaffold 0002504	164

[0444] 表16: 位于Nic1或Nic2缺失区侧翼的SNP标记。REF是指参考TN90等位基因的序列。ALT是指与相应的缺失区相关的替代等位基因。显示了4个品种 (BU21、HI、LI和LA) 的基因型, 其中“00”是指参考TN90等位基因纯合, 而“11”是指ALT等位基因纯合。序列中的多态性位点用“X”表示。

[0445]

SNP 标记 (SEQ ID No.)	Asse'd 基因 座	相对 于缺 失的 位置	REF	ALT	BU21	HI	L I	LA	序列
131	Nic1	上游	C	A	00	00	11	11	ATCATGTCTAATTGATTTAATTGCT GTATTTGCTCAAACCTGCCTTATTTG GACTATGTGAXACATGCTAGGTTA GAAATATATGTTTTAACTTGGTGT GAAATTTAATTTAATTGAGTATT (SEQ ID NO. 131)
132	Nic1	上游	C	T	00	00	11	11	GCGGCTGTATACCATTTTGTACGG ACCGCAGTGGGCTCACCGTGGCCT CAATCGAGTTTGXGCGGTTTCATAA AGATGGGGGTTTCAGAGAGTTGGGA GTTTAGAGATTAAGACCAATACGG T(SEQ ID NO. 132)
133	Nic1	上游	T	G	00	00	11	11	GGAAGCACTCAAGCATCCACTCTT GGAGGTGGTGGGGGAGGGTCTGG AATATTATCATTGXCCTGGTGGCTT CTCCTTTGTCTTTGAGGTACAATAG GAACCTCATCATCAATATTGTCAT(SEQ ID NO. 133)
134	Nic1	下游	T	C	00	00	11	11	ACGAAGGATAAAGTGTTTGGGTAG CAGAACAAAATGCCTTCGTCATTC CAGTCTTTAACAXATGCCAAGTGC AAACAATACAATTAAATTTGTAG TCTCTTCTGATGGTGTGGACTTAC (SEQ ID NO. 134)
135	Nic1	下游	C	T	00	00	11	11	ACACCTTCTTCCGGGTAAACAGAA TTCCATACTCGGATTTCTGGTTCGC AGACTGTAATAXGGAGTCAATCTT TTCCTCGATTTGGGATTTGAACCGG TGATTTGGGAAATCATAATTATC(S EQ ID NO. 135)

[0446]

SNP 标记 (SEQ ID No.)	Asse'd 基因 座	相对 于缺 失的 位置	REF	ALT	BU21	HI	L I	LA	序列
136	Nic1	下游	A	G	00	00	11	11	TAGCTAACAAGGAATTGGATCAAT TGAGAGATTGATTAACCCAATTAA AGAGTTTAACCTXGAGATAGTAAC AACATGACTTGAGCTCTTATCAAC AGTTTTGGTTGATACCTTTTGGTCT(SEQ ID NO. 136)
137	Nic2	上游	T	C	00	11	00	11	TTACATAAATATAAAGGTTTAATT GAAAGTTATACTTTTGGTCAAAC ACAAATACCGTAXCAAAATAGTTC GATACGGTTAGGTATTTCTTGTTT GGTTCGGTACGGCTTCGATATTAT(SEQ ID NO. 137)
138	Nic2	上游	G	A	00	11	00	11	GAAAATACCGACCGAAAACGGTC GAAAATAACATATTTTTAATTATT CCAACCGACTTCXGTCGGTTTGTTA AGATTTAAAAAGAAATGCAAAATT ATCACATAATTATATTTCCGACCG(SEQ ID NO. 138)
139	Nic2	上游	G	A	00	11	00	11	GTCTGCACTTTTCTTATTGCTTATT GTTTACCCGAAAAATGGATAGAGT TGAATTTATACXTAGTTTAAAGGG TATGTGGTATAATTTAATACAAAT CGTAAGAATAAGTAGAAATATCAA (SEQ ID NO. 139)
140	Nic2	下游	G	A	00	11	00	11	TAGGATTTTGACCGTGGTTGGGAA CTATGTGAAGACAACTCCGGAATA GAGTTATGTCGXTTCTGTTAGCTCC ATTAGATGATTTTGGACTTAGGGG CGTGTCCGATTATGTTTTGGAG(S EQ ID NO. 140)

[0447]

SNP 标记 (SEQ ID No.)	Asse'd 基因座	相对于缺失的位置	REF	ALT	BU21	HI	L I	LA	序列
141	Nic2	下游	A	G	00	11	00	11	GTCCCTTACGAATTTGTCTTCAAAT TGATACTTCTCCTTGCTAAAACACC ACGATCCTTAXCCACAACCTCACCC CACGAACCCCTAGCATAGAACCACA ACACCCTACGGCCCTTAAGAAAC(S EQ ID NO. 141)
142	Nic2	下游	T	C	00	11	00	11	GCCTCCCAGCTTAGCAAGAGTTCA TCCGGCATTTCATTTTCGATGCTA CAAAGGTATCAXGGCAATCCGTCT CATGTGTTGGATTTCAGCTCAGTCC AGTTGGACAAGGATCTATCTTAT(S EQ ID NO. 142)

[0448] 表17:与Nic1缺失区相关的ERF-39样基因侧翼的SNP标记。此表列出了相对位置(位置,上游/下游)、TN90中的参考等位基因(REF)、候选等位基因(ALT),4个品种的基因型(BU21、HI、LI和LA BU21)以及具有侧翼位置的SNP的序列。“00”是指参考TN90等位基因纯合,而“11”是指ALT等位基因纯合,并且“01”是指杂合。序列中的多态性位点用“X”表示。

[0449]

SNP 标记 (SEQ ID No.)	Asse'd 基因座	相对于 ERF-39 的位置	REF	ALT	BU21	HI	L I	LA	序列
143	ERF-39	上游	A	G	00	00	01	01	CCAAATTATTTTGA CTGTTTGACCAAAAT AGCGACCAACGTTG GTCGCCATTTTGTAT CXTTGTGACCAAAAT GGCGACCAACTTGT GTCGCTATATTTGAA AATAAATAAATAAA ATAAT (SEQ ID NO. 143)

[0450]

144	ERF-3 9	下游	TG	T	00	00	11	11	TTGTTTCTCAAGAAA ATCAGCATCGATCTT TTTCTTTATTATAAA GGTAATATAGATGC TXXGGGTATTAAAA GAGGACAAAATATT CTGGCATCATTTGAG ATTTGCTGAGTGCAA TAACA (SEQ ID NO. 144)
-----	------------	----	----	---	----	----	----	----	--

[0451] 表18:基因特异性引物与用于对实施例11的F2群体(分离nic1和nic2((LA白肋21x TN 90LC) (X)), F2_1-F2_72植物是指表12中的植物编号1-72)进行基因分型的侧翼SNP标记的比较。*基因型未解决;^a nic1或nic2纯合;^b nic1或nic2杂合。阴影格子显示其中基因特异性引物和侧翼SNP之间的基因分型结果相匹配的植物。

[0452]

使用如表 5 所示的 g100614_Scaffold0002504 的基因特异性引物对 Nic 1 基因座进行基因分型。									
	1	2	3	4	5	6	7	8	9
A	F2_1	F2_9	F2_17 ^a	F2_25 ^a	F2_33	F2_41	F2_49	F2_57	F2_65
B	F2_2	F2_10	F2_18	F2_26	F2_34	F2_42	F2_50 ^a	F2_58 ^a	F2_66
C	F2_3 ^a	F2_11	F2_19 ^a	F2_27	F2_35 ^a	F2_43	F2_51 ^a	F2_59	F2_67
D	F2_4	F2_12	F2_20	F2_28	F2_36	F2_44	F2_52	F2_60	F2_68
E	F2_5	F2_13	F2_21	F2_29	F2_37	F2_45 ^a	F2_53	F2_61	F2_69 ^a
F	F2_6	F2_14	F2_22	F2_30	F2_38 ^a	F2_46	F2_54	F2_62	F2_70
G	F2_7	F2_15	F2_23	F2_31	F2_39 ^a	F2_47	F2_55	F2_63	F2_71
H	F2_8	F2_16	F2_24 ^a	F2_32	F2_40	F2_48	F2_56	F2_64	F2_72 ^a
使用 SNP 标记 SEQ ID No.135 对 Nic 1 进行 KASP 基因分型									
	1	2	3	4	5	6	7	8	9
A	F2_1*	F2_9 ^b	F2_17 ^a	F2_25 ^a	F2_33	F2_41 ^b	F2_49	F2_57 ^b	F2_65*
B	F2_2 ^b	F2_10 ^b	F2_18	F2_26	F2_34	F2_42 ^b	F2_50 ^a	F2_58 ^a	F2_66 ^b
C	F2_3 ^a	F2_11 ^b	F2_19 ^a	F2_27	F2_35 ^a	F2_43 ^b	F2_51 ^a	F2_59 ^b	F2_67 ^b
D	F2_4	F2_12 ^b	F2_20	F2_28 ^b	F2_36 ^b	F2_44 ^b	F2_52 ^b	F2_60 ^b	F2_68 ^b
E	F2_5 ^b	F2_13 ^b	F2_21 ^b	F2_29 ^b	F2_37	F2_45 ^a	F2_53 ^b	F2_61 ^b	F2_69 ^a
F	F2_6 ^b	F2_14	F2_22 ^b	F2_30	F2_38 ^a	F2_46 ^b	F2_54 ^b	F2_62	F2_70 ^b
G	F2_7*	F2_15 ^b	F2_23 ^b	F2_31 ^b	F2_39 ^a	F2_47 ^b	F2_55 ^b	F2_63 ^b	F2_71 ^b
H	F2_8 ^b	F2_16*	F2_24*	F2_32 ^b	F2_40 ^b	F2_48 ^b	F2_56*	F2_64 ^b	F2_72 ^a

[0453]

[0454]

使用如表 5 所示的 ERF 189 的基因特异性引物对 Nic 2 基因座进行基因分型。									
	1	2	3	4	5	6	7	8	9
A	F2_1	F2_9 ^a	F2_17 ^a	F2_25	F2_33 ^a	F2_41 ^a	F2_49 ^a	F2_57	F2_65
B	F2_2	F2_10	F2_18	F2_26	F2_34	F2_42	F2_50	F2_58	F2_66
C	F2_3	F2_11 ^a	F2_19 ^a	F2_27	F2_35	F2_43 ^a	F2_51	F2_59	F2_67
D	F2_4	F2_12 ^a	F2_20	F2_28	F2_36	F2_44	F2_52 ^a	F2_60	F2_68
E	F2_5 ^a	F2_13	F2_21 ^a	F2_29 ^a	F2_37 ^a	F2_45 ^a	F2_53	F2_61	F2_69
F	F2_6	F2_14 ^a	F2_22 ^a	F2_30	F2_38	F2_46 ^a	F2_54	F2_62	F2_70
G	F2_7	F2_15	F2_23	F2_31	F2_39 ^a	F2_47	F2_55	F2_63	F2_71
H	F2_8	F2_16	F2_24 ^a	F2_32	F2_40 ^a	F2_48 ^a	F2_56	F2_64	F2_72 ^a
使用 SNP 标记 SEQ ID No. 137 对 Nic 2 进行 KASP 基因分型									
	1	2	3	4	5	6	7	8	9
A	F2_1*	F2_9 ^a	F2_17 ^a	F2_25	F2_33 ^b	F2_41 ^a	F2_49 ^a	F2_57 ^b	F2_65
B	F2_2 ^b	F2_10 ^b	F2_18 ^b	F2_26 ^b	F2_34 ^a	F2_42 ^b	F2_50 ^b	F2_58 ^b	F2_66 ^b
C	F2_3 ^b	F2_11 ^a	F2_19 ^a	F2_27	F2_35	F2_43 ^a	F2_51 ^b	F2_59 ^b	F2_67 ^b
D	F2_4	F2_12 ^a	F2_20 ^b	F2_28	F2_36 ^b	F2_44 ^b	F2_52 ^a	F2_60	F2_68 ^b
E	F2_5 ^a	F2_13 ^b	F2_21 ^a	F2_29 ^a	F2_37 ^b	F2_45 ^a	F2_53 ^b	F2_61	F2_69
F	F2_6*	F2_14 ^a	F2_22 ^a	F2_30 ^b	F2_38	F2_46 ^a	F2_54 ^b	F2_62 ^b	F2_70
G	F2_7	F2_15 ^b	F2_23 ^b	F2_31	F2_39 ^a	F2_47 ^b	F2_55 ^b	F2_63	F2_71 ^b
H	F2_8 ^b	F2_16 ^b	F2_24*	F2_32 ^b	F2_40 ^b	F2_48 ^a	F2_56*	F2_64 ^b	F2_72*

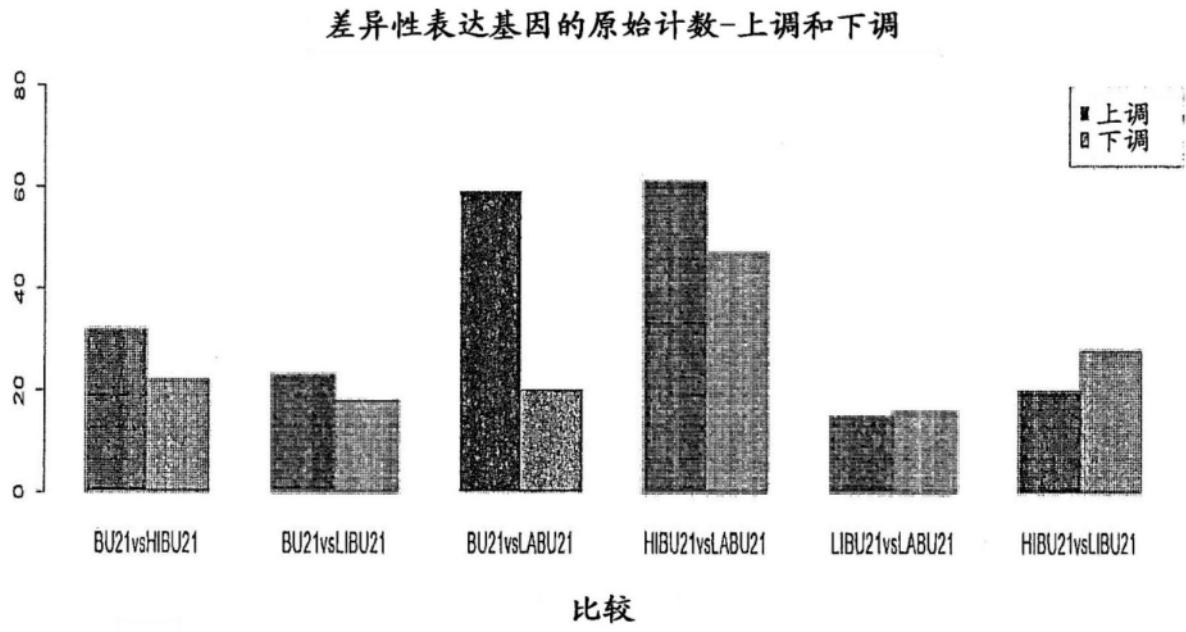


图1

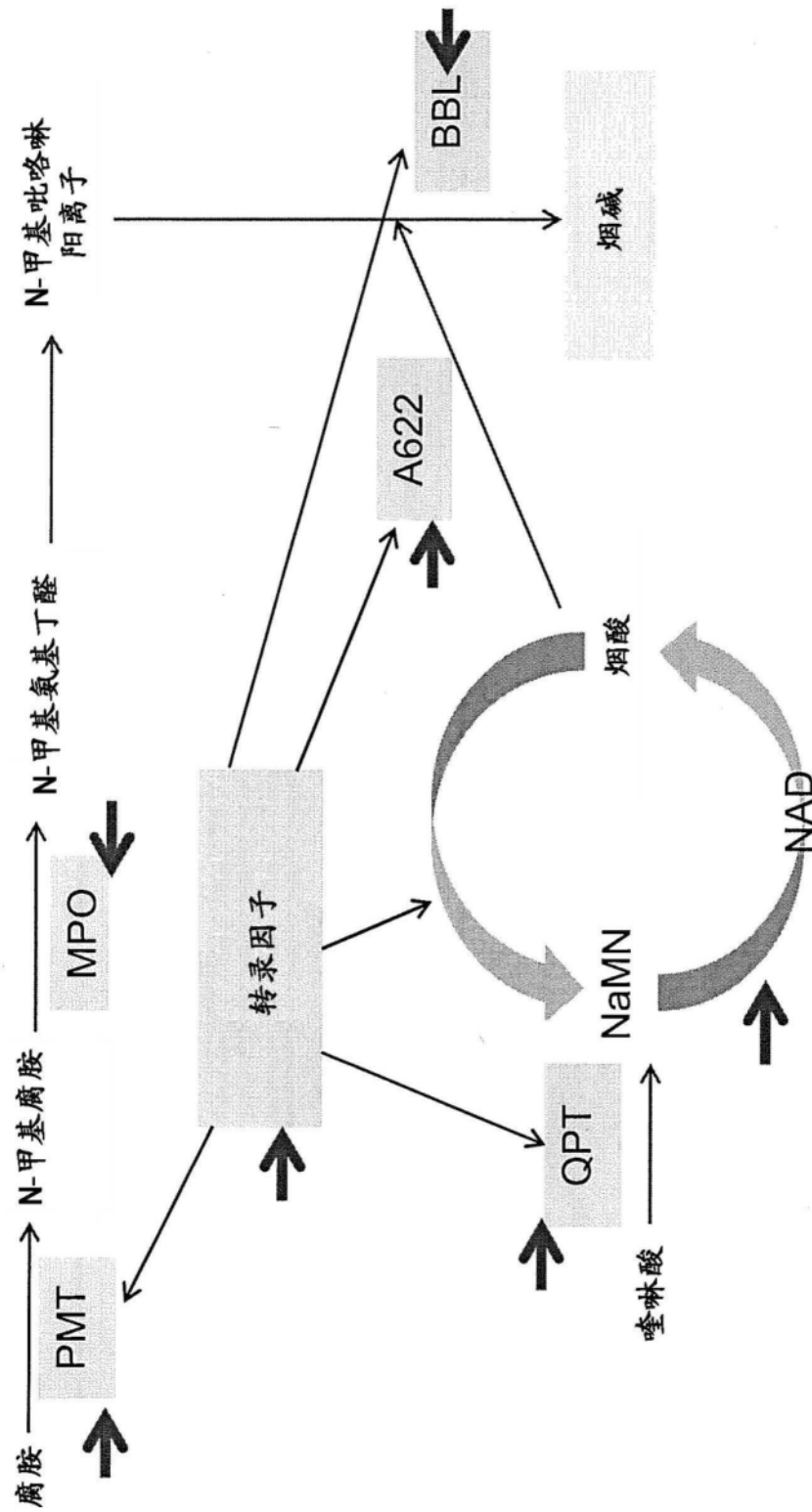


图2

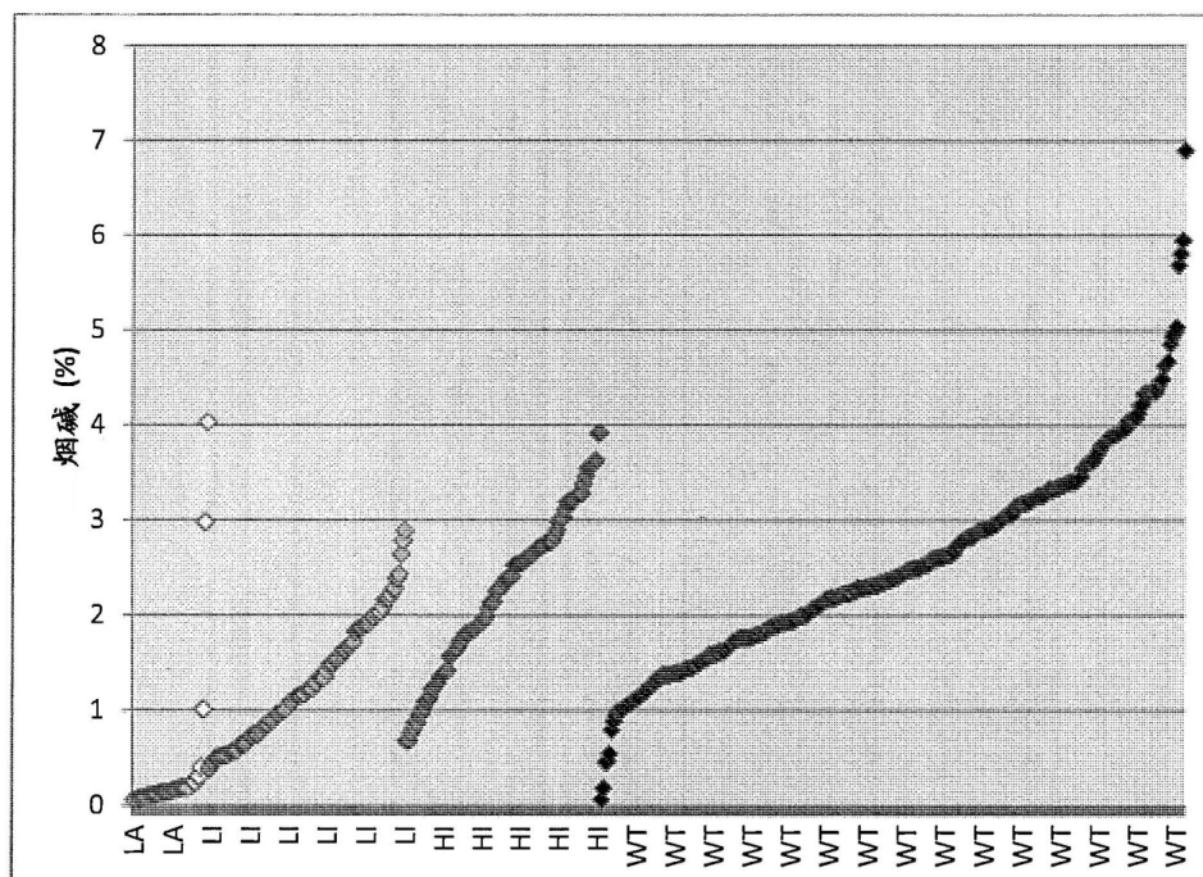


图3

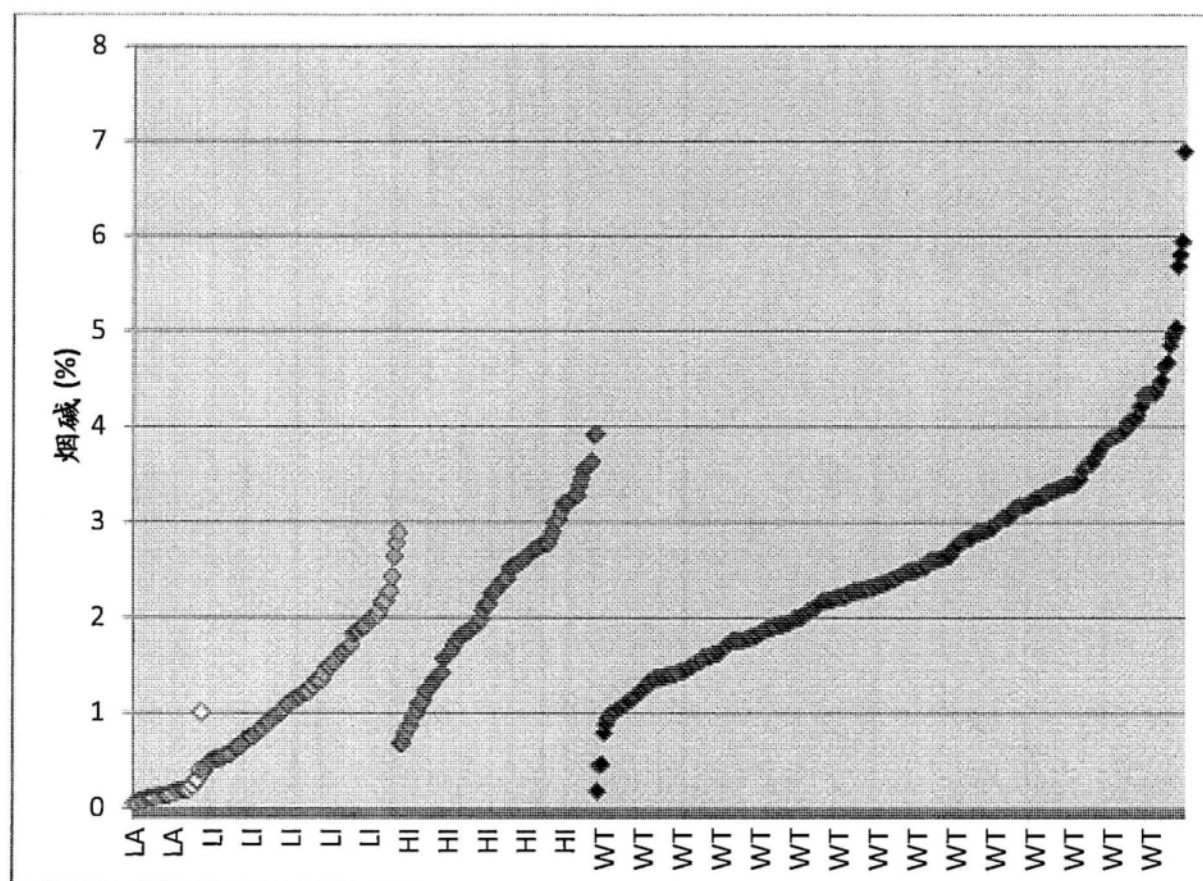


图4

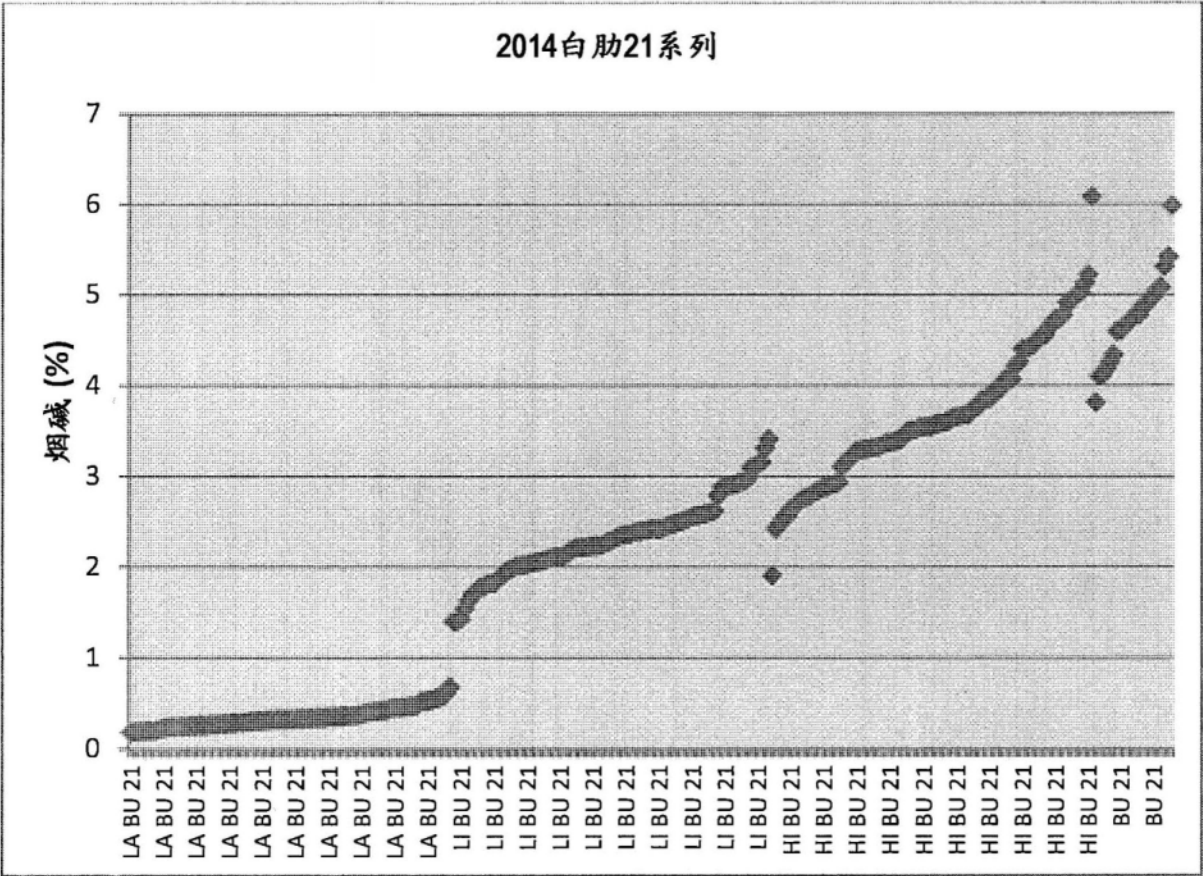


图5

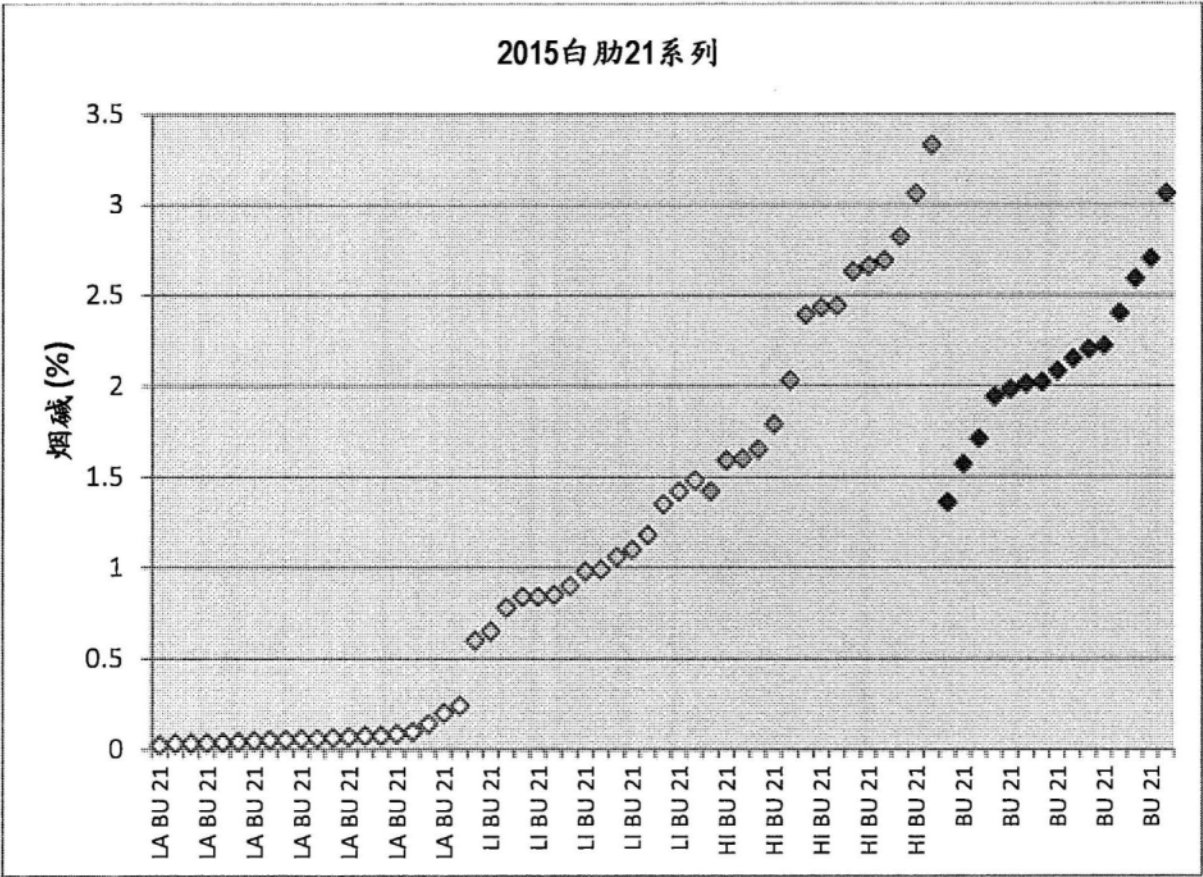


图6

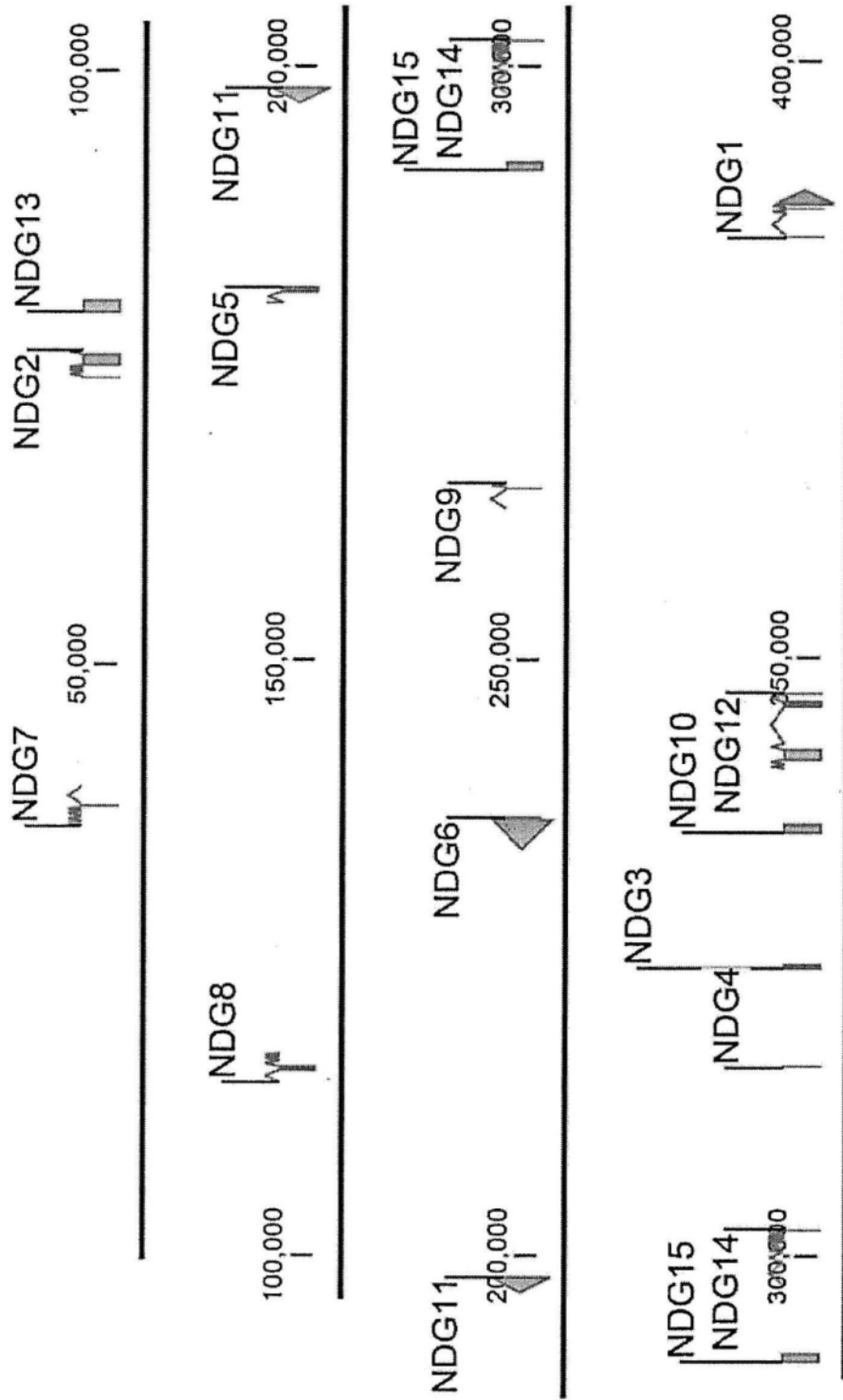


图7