(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 119506283 A
(43)申请公布日 2025.02.25

(21)申请号 202411642439.4
(22)申请日 2018.09.17
(30)优先权数据
2017903773 2017.09.15 AU
2018902840 2018.08.03 AU
2018902896 2018.08.08 AU
(62)分案原申请数据
201880072998.2 2018.09.17
(71)申请人 联邦科学技术研究组织
地址 澳大利亚澳大利亚首都领地
(72)发明人 N·A·史密斯 M-B·王 T·J·多兰 M·蒂泽德 A·D·阿鲁 I·K·格里夫斯 L·高 J·P·安德森 D·张 R·德菲特
(74)专利代理机构 永新专利商标代理有限公司 72002
专利代理师 左路
(51)Int.Cl.
C12N 15/113(2010.01)
C12N 15/82(2006.01)
A01H 5/00(2018.01)
A01H 6/20(2018.01)

权利要求书8页 说明书91页
序列表（电子公布） 附图31页

(54)发明名称

RNA分子

(57)摘要

本发明涉及新的双链RNA(dsRNA)结构及其在基因沉默中的用途。在一个实施方案中,dsRNA包含双发夹(哑铃)结构。在另一个实施方案中,dsRNA分子的有义链和反义链的一部分含有以非规范碱基对被碱基配对或者不被碱基配对的核糖核苷酸。

CN 119506283 A

1.一种RNA分子,所述RNA分子包含第一RNA组分、与所述第一RNA组分共价连接的第二RNA组分,以及任选地以下中的一种或多种或全部:(i)与所述第一和第二RNA组分共价连接的连接核糖核苷酸序列,(ii)5'前导序列和(iii)3'尾随序列,

其中所述第一RNA组分按5'至3'顺序由第一5'核糖核苷酸、第一RNA序列和第一3'核糖核苷酸组成,其中所述第一RNA组分中的所述第一5'和3'核糖核苷酸彼此碱基配对,其中所述第一RNA序列包含至少21个邻接核糖核苷酸的第一有义核糖核苷酸序列、至少4个核糖核苷酸的第一环序列和至少21个邻接核糖核苷酸的第一反义核糖核苷酸序列,其中所述RNA分子中的所述第一反义核糖核苷酸序列的所述至少21个邻接核糖核苷酸与所述第一有义核糖核苷酸序列的所述至少21个邻接核糖核苷酸碱基配对,其中所述第一反义核糖核苷酸序列能够与靶RNA分子的第一区杂交,

其中,如果存在连接核糖核苷酸序列,则经由所述连接核糖核苷酸序列,或如果不存在连接核糖核苷酸序列,则直接地,所述第二RNA组分与所述第一5'核糖核苷酸或所述第一3'核糖核苷酸共价连接,

其中所述第二RNA组分按5'至3'顺序由第二5'核糖核苷酸、第二RNA序列和第二3'核糖核苷酸组成,其中所述第二5'和3'核糖核苷酸在所述RNA分子中彼此碱基配对,其中所述第二RNA序列包含至少21个邻接核糖核苷酸的第二反义核糖核苷酸序列、至少4个核糖核苷酸的第二环序列和至少21个邻接核糖核苷酸的第二有义核糖核苷酸序列,其中在所述RNA分子中所述第二有义核糖核苷酸序列与所述第二反义核糖核苷酸序列杂交;

其中如果存在所述5'前导序列,则所述5'前导序列由下述核糖核苷酸序列组成,如果所述第二RNA组分与所述第一3'核糖核苷酸连接,则所述核糖核苷酸序列与所述第一5'核糖核苷酸共价连接,或者如果所述第二RNA组分与所述第一5'核糖核苷酸连接,则所述核糖核苷酸序列与所述第二5'核糖核苷酸共价连接,

其中如果存在所述3'尾随序列,则所述3'尾随序列由下述核糖核苷酸序列组成,如果所述第二RNA组分与所述第一3'核糖核苷酸连接,则所述核糖核苷酸序列与所述第二3'核糖核苷酸共价连接,或者如果所述第二RNA组分与所述第一5'核糖核苷酸连接,则所述核糖核苷酸序列与所述第一3'核糖核苷酸共价连接,

其中所述RNA分子由包含5'端和3'端的单个核糖核苷酸链组成,并且

其中(a)当第一有义核糖核苷酸序列与第一反义核糖核苷酸序列杂交时,所述第一有义核糖核苷酸序列和所述第一反义核糖核苷酸序列的总体上5%至40%的核糖核苷酸以非规范碱基对被碱基配对,或(b)当第二有义核糖核苷酸序列与第二反义核糖核苷酸序列杂交时,第二有义核糖核苷酸序列与第二反义核糖核苷酸序列的总体上5%至40%的核糖核苷酸以非规范碱基对被碱基配对,或(a)和(b)两者。

2.权利要求1的RNA分子,其中(a)第一有义核糖核苷酸序列和第一反义核糖核苷酸序列长度至少为40个邻接核苷酸,或(b)第二有义核糖核苷酸序列和第二反义核糖核苷酸序列长度至少为40个邻接核苷酸,或(a)和(b)两者。

3.权利要求1的RNA分子,其中第一反义核糖核苷酸序列的至少21个邻接核糖核苷酸均能够与靶RNA分子的第一区的核苷酸碱基配对。

4.权利要求3的RNA分子,其中第一反义核糖核苷酸序列的至少21个邻接核糖核苷酸与靶RNA分子的第一区的互补体相同。

5.权利要求1的RNA分子,其中(i)第一和第二反义核糖核苷酸序列位于第一和第二有义核糖核苷酸序列的侧翼,或(ii)第一和第二有义核糖核苷酸序列位于第一和第二反义核糖核苷酸序列的侧翼。

6.权利要求1的RNA分子,其特征为以下一项或多项:

(i)第一有义核糖核苷酸序列的至少21个邻接核糖核苷酸在序列上与靶RNA分子的第一区相同,

(ii)第一有义核糖核苷酸序列与第一5'核糖核苷酸共价连接而无任何间插核苷酸,或第一反义核糖核苷酸序列与第一3'核糖核苷酸共价连接而无任何间插核苷酸,或两者,

(iii)所述RNA分子包含连接核糖核苷酸序列,其中所述连接核糖核苷酸序列的长度小于20个核糖核苷酸,

(iv)所述连接核糖核苷酸序列与靶RNA分子杂交,

(v)所述连接核糖核苷酸序列与靶RNA分子的互补体的一部分相同,

(vi)所述连接核糖核苷酸序列长度在1至10个核糖核苷酸之间,

(vii)第一有义核糖核苷酸序列的至少21个核糖核苷酸和第二有义核糖核苷酸序列的至少21个核糖核苷酸在序列上与同一靶RNA分子的不同区域相同,

(viii)第一有义核糖核苷酸序列的至少21个核糖核苷酸和第二有义核糖核苷酸序列的至少21个核糖核苷酸在序列上与不同靶RNA分子的区域相同,

(ix)第二反义核糖核苷酸序列的至少21个核糖核苷酸和第二有义核糖核苷酸序列的至少21个核糖核苷酸完全碱基配对,

(x)第一反义核糖核苷酸序列的至少21个核糖核苷酸和第二反义核糖核苷酸序列的至少21个核糖核苷酸与同一靶RNA分子的不同区域互补,

(xi)第二反义核糖核苷酸序列与不同于第一反义核糖核苷酸序列的靶RNA分子的区域互补,

(xii)第一5'核糖核苷酸和第二3'核糖核苷酸在RNA分子中相邻,各自碱基配对且不直接共价结合彼此,或者第二5'核糖核苷酸和第一3'核糖核苷酸在RNA分子中相邻,各自碱基配对且不直接共价结合,

(xiii)第一和第二反义核糖核苷酸序列被连接核糖核苷酸序列隔开,

(xiv)第一有义核糖核苷酸序列与靶RNA分子的第一区至少60%相同,第二有义核糖核苷酸序列与靶RNA分子的第二区至少60%相同,靶RNA分子的第二区与靶RNA分子的第一区不同,其中第一和第二反义核糖核苷酸序列通过连接核糖核苷酸序列连接,

(xv)所述RNA分子具有5'前导序列或3'尾随序列,或两者,

(xvi)所述RNA分子中的至少一个环序列,或第一和第二环序列,长于20个核苷酸,

(xvii)所述RNA分子只有两个环,并且

(xviii)至少一个环,或所有环的长度在4到50个核糖核苷酸之间。

7.权利要求1所述的RNA分子,其中(a)第一反义核糖核苷酸序列与靶RNA分子的第一靶区的互补体在序列上具有80%至99.9%的同一性,或(b)第二反义核糖核苷酸序列独立地与靶RNA分子的第二靶区的互补体在序列上具有80%至99.9%的同一性,或(a)和(b)两者。

8.权利要求1所述的RNA分子,其中第一反义核糖核苷酸序列的至少21个核糖核苷酸和第二反义核糖核苷酸序列的至少21个核糖核苷酸与同一靶RNA分子的不同区域互补。

9.权利要求1所述的RNA分子,其中第一反义核糖核苷酸序列的至少21个核糖核苷酸和第二反义核糖核苷酸序列的至少21个核糖核苷酸与不同靶RNA分子的区域互补。

10.权利要求1所述的RNA分子,其中当所述有义和反义核糖核苷酸序列杂交时,(a)第一有义核糖核苷酸序列和第一反义核糖核苷酸序列总体上,以及(b)第二有义核糖核苷酸序列和第二反义核糖核苷酸序列总体上15%至30%之间的核糖核苷酸以非规范碱基对进行碱基配对。

11.权利要求10所述的RNA分子,其中当所述有义和反义核糖核苷酸序列杂交时,(a)第一有义核糖核苷酸序列和第一反义核糖核苷酸序列总体上,或(b)第二有义核糖核苷酸序列和第二反义核糖核苷酸序列总体上,或(a)和(b)两者的16%至25%之间的核糖核苷酸以非规范碱基对进行碱基配对。

12.权利要求1所述的RNA分子,其中(a)所述第一有义核糖核苷酸序列和所述第一反义核糖核苷酸序列的长度各自为40至1,000个核苷酸,并且(b)所述第二有义核糖核苷酸序列和所述第二反义核糖核苷酸序列的长度各自为40至1,000个核苷酸。

13.权利要求1所述的RNA分子,其中所述RNA分子中的所有非规范碱基对均为G:U碱基对。

14.权利要求1所述的RNA分子,其中所述RNA分子在细胞中产生。

15.权利要求14所述的RNA分子,其中所述RNA分子在植物细胞或真菌细胞中产生。

16.权利要求14所述的RNA分子,其中所述RNA分子在酵母细胞中产生。

17.权利要求1所述的RNA分子,其中所述RNA分子在无细胞表达系统中产生。

18.权利要求1所述的RNA分子,其中所述靶RNA分子在植物细胞、真菌细胞或动物细胞中。

19.权利要求1所述的RNA分子,其中所述靶RNA分子位于节肢动物细胞中。

20.权利要求1所述的RNA分子,其中所述靶RNA分子位于植物细胞中。

21.权利要求1所述的RNA分子,其中所述靶RNA分子位于真核细胞中,其中所述真核细胞与产生所述RNA分子的细胞不同。

22.权利要求1所述的RNA分子,其中所述RNA分子在第一有义核糖核苷酸序列或第一反义核糖核苷酸序列或两者中包含一个凸起、或两个或更多个凸起,每个凸起包含一个或多个未配对的核苷酸,其中所述dsRNA区域在与一个或多个未配对的核苷酸相对的互补序列中的相应位置处缺少非碱基配对的核苷酸。

23.权利要求1所述的RNA分子,其特征在于以下一项或多项:

(i)或者(a)第一有义核糖核苷酸序列中每四个到每六个核糖核苷酸中的一个与第一反义核糖核苷酸序列中的核糖核苷酸形成非规范碱基对,或者(b)第二有义核糖核苷酸序列中每四个到每六个核糖核苷酸中的一个与第二反义核糖核苷酸序列中的核糖核苷酸形成非规范碱基对,或者(a)和(b)两者,

(ii)或者(a)由第一有义核糖核苷酸序列和第一反义核糖核苷酸序列形成的dsRNA区域缺少8个邻接的规范碱基对,或者(b)由第二有义核糖核苷酸序列和第二反义核糖核苷酸序列形成的dsRNA区域缺少8个邻接的规范碱基对,或(a)和(b)两者,

(iii)第一或第二反义核糖核苷酸序列,或两者,分别独立地与靶RNA分子的第一或第二靶区的互补体具有小于100%的序列同一性,

(iv)第一或第二反义核糖核苷酸序列,或两者,分别与靶RNA分子的第一或第二靶区具有100%的序列同一性,

(v)或者(a)第一有义核糖核苷酸序列中的核糖核苷酸数目为第一反义核糖核苷酸序列中的核糖核苷酸数目的90%至110%,或(b)第二有义核糖核苷酸序列中的核糖核苷酸数目为第二反义核糖核苷酸序列中的核糖核苷酸数目的90%至110%,或(a)和(b)两者,以及

(vi)或者(a)第一有义核糖核苷酸序列中的核糖核苷酸数目与第一反义核糖核苷酸序列中的核糖核苷酸数目相同,或(b)第二有义核糖核苷酸序列中的核糖核苷酸数目与第二反义核糖核苷酸序列中的核糖核苷酸数目相同,或(a)和(b)两者。

24.一种RNA分子,所述RNA分子包含第一RNA组分、与所述第一RNA组分共价连接的第二RNA组分,以及任选地以下中的一种或多种或全部:(i)与所述第一和第二RNA组分共价连接的连接核糖核苷酸序列,(ii)5'前导序列和(iii)3'尾随序列,

其中所述第一RNA组分按5'至3'顺序由第一5'核糖核苷酸、第一RNA序列和第一3'核糖核苷酸组成,其中所述第一RNA组分中的所述第一5'和3'核糖核苷酸彼此碱基配对,其中所述第一RNA序列包含至少21个邻接核糖核苷酸的第一有义核糖核苷酸序列、至少4个核糖核苷酸的第一环序列和至少21个邻接核糖核苷酸的第一反义核糖核苷酸序列,其中第一反义核糖核苷酸序列与所述RNA分子中的第一有义核糖核苷酸序列杂交,其中第一反义核糖核苷酸序列能够与靶RNA分子的第一区杂交,

其中,如果存在所述连接核糖核苷酸序列,则经由所述连接核糖核苷酸序列,所述第二RNA组分与所述第一5'核糖核苷酸或所述第一3'核糖核苷酸共价连接,或如果不存在所述连接核糖核苷酸序列,则所述第二RNA组分与所述第一5'核糖核苷酸或所述第一3'核糖核苷酸直接共价连接,

其中所述第二RNA组分按5'至3'顺序由第二5'核糖核苷酸、第二RNA序列和第二3'核糖核苷酸组成,其中在所述RNA分子中的所述第二5'和3'核糖核苷酸彼此碱基配对,其中所述第二RNA序列包含至少21个邻接核糖核苷酸的第二反义核糖核苷酸序列、至少4个核糖核苷酸的第二环序列和至少21个邻接核糖核苷酸的第二有义核糖核苷酸序列,其中在所述RNA分子中所述第二有义核糖核苷酸序列与所述第二反义核糖核苷酸序列杂交,

其中如果存在所述5'前导序列,则所述5'前导序列由下述核糖核苷酸序列组成,如果所述第二RNA组分与所述第一3'核糖核苷酸连接,则所述核糖核苷酸序列与所述第一5'核糖核苷酸共价连接,或者如果所述第二RNA组分与所述第一5'核糖核苷酸连接,则所述核糖核苷酸序列与所述第二5'核糖核苷酸共价连接,

其中如果存在所述3'尾随序列,则所述3'尾随序列由下述核糖核苷酸序列组成,如果所述第二RNA组分与所述第一3'核糖核苷酸连接,则所述核糖核苷酸序列与所述第二3'核糖核苷酸共价连接,或者如果所述第二RNA组分与所述第一5'核糖核苷酸连接,则所述核糖核苷酸序列与所述第一3'核糖核苷酸共价连接,

其中,所述RNA分子由包含5'端和3'端的单个核糖核苷酸链组成,并且

其中,所述RNA分子在第一有义核糖核苷酸序列或第一反义核糖核苷酸序列或两者中包含一个凸起或两个或更多个凸起,每个凸起包含一个或多个未配对的核苷酸,其中所述dsRNA区在与所述一个或多个未配对核苷酸相对的互补序列中的相应位置处缺少非碱基配对的核苷酸。

25.权利要求24所述的RNA分子,其中第一反义核糖核苷酸序列包含一个凸起,或两个或更多个凸起。

26.权利要求24所述的RNA分子,其中第一反义核糖核苷酸序列的至少21个邻接核糖核苷酸均能够与靶RNA分子的第一区的核苷酸碱基配对。

27.权利要求26所述的RNA分子,其中第一反义核糖核苷酸序列的至少21个邻接核糖核苷酸与靶RNA分子的第一区的互补体相同。

28.权利要求24所述的RNA分子,其中(a)第一有义核糖核苷酸序列和第一反义核糖核苷酸序列的长度至少为40个核苷酸,或(b)第二有义核糖核苷酸序列和第二反义核糖核苷酸序列的长度各自至少为40个核苷酸,或(a)和(b)两者。

29.权利要求24所述的RNA分子,其中第一反义核糖核苷酸序列的至少21个核糖核苷酸和第二反义核糖核苷酸序列的至少21个核糖核苷酸与同一靶RNA分子的不同区域互补。

30.权利要求24所述的RNA分子,其中第一反义核糖核苷酸序列的至少21个核糖核苷酸和第二反义核糖核苷酸序列的至少21个核糖核苷酸与不同靶RNA分子的区域互补。

31.权利要求24所述的RNA分子,其中(a)当第一有义核糖核苷酸序列与第一反义核糖核苷酸序列杂交时,第一有义核糖核苷酸序列和第一反义核糖核苷酸序列中总体上5%至40%的核糖核苷酸以非规范碱基对进行碱基配对,并且(b)当第二有义核糖核苷酸序列与第二反义核糖核苷酸序列杂交时,第二有义核糖核苷酸序列和第二反义核糖核苷酸序列总体上5%至40%的核糖核苷酸以非规范碱基对进行碱基配对。

32.权利要求31所述的RNA分子,其中当所述有义和反义核糖核苷酸序列杂交时,(a)第一有义核糖核苷酸序列和第一反义核糖核苷酸序列中总体上15%至30%的核糖核苷酸,和(b)第二有义核糖核苷酸序列和第二反义核糖核苷酸序列中总体上15%至30%的核糖核苷酸以非规范碱基对进行碱基配对。

33.权利要求32所述的RNA分子,其中当所述有义和反义核糖核苷酸序列杂交时,(a)第一有义核糖核苷酸序列和第一反义核糖核苷酸序列总体上,或(b)第二有义核糖核苷酸序列和第二反义核糖核苷酸序列中总体上,或(a)和(b)两者中16%至25%的核糖核苷酸以非规范碱基对进行碱基配对。

34.权利要求24所述的RNA分子,其中(a)所述第一有义核糖核苷酸序列和所述第一反义核糖核苷酸序列的长度各自为40至1,000个核苷酸,并且(b)所述第二有义核糖核苷酸序列和所述第二反义核糖核苷酸序列的长度各自为40至1,000个核苷酸。

35.权利要求24所述的RNA分子,其中所述RNA分子中的所有非规范碱基对均为G:U碱基对。

36.权利要求24所述的RNA分子,其中所述RNA分子在植物细胞或真菌细胞中产生。

37.权利要求24所述的RNA分子,其中所述RNA分子在酵母细胞中产生。

38.权利要求24所述的RNA分子,其中所述靶RNA分子在植物细胞、真菌细胞或动物细胞中。

39.权利要求24所述的RNA分子,其中所述靶RNA分子位于节肢动物细胞中。

40.权利要求24所述的RNA分子,其中由第二有义核糖核苷酸序列和第二反义核糖核苷酸序列形成的dsRNA区中的所有核糖核苷酸都是碱基配对的。

41.权利要求24所述的RNA分子,其中(a)第一反义核糖核苷酸序列与靶RNA分子的第一

靶区域的互补体在序列上具有80％至99.9％的同一性,或(b)第二反义核糖核苷酸序列独立地与靶RNA分子的第二靶区域的互补体在序列上具有80％至99.9％的同一性,或(a)和(b)两者。

42.权利要求24所述的RNA分子,其特征在于以下一项或多项:

(i)或者(a)第一有义核糖核苷酸序列中每四个到每六个核糖核苷酸中的一个与第一反义核糖核苷酸序列中的核糖核苷酸形成非规范碱基对,或者(b)第二有义核糖核苷酸序列中每四个到每六个核糖核苷酸中的一个与第二反义核糖核苷酸序列中的核糖核苷酸形成非规范碱基对,或者(a)和(b)两者,

(ii)或者(a)由第一有义核糖核苷酸序列和第一反义核糖核苷酸序列形成的dsRNA区域缺少8个邻接的规范碱基对,或者(b)由第二有义核糖核苷酸序列和第二反义核糖核苷酸序列形成的dsRNA区域缺少8个邻接的规范碱基对,或(a)和(b)两者,

(iii)第一或第二反义核糖核苷酸序列,或两者,分别独立地与靶RNA分子的第一或第二靶区的互补体具有小于100％的序列同一性,

(iv)第一或第二反义核糖核苷酸序列,或两者,分别与靶RNA分子的第一或第二靶区具有100％的序列同一性,

(v)或者(a)第一有义核糖核苷酸序列中的核糖核苷酸数目为第一反义核糖核苷酸序列中的核糖核苷酸数目的90％至110％,或(b)第二有义核糖核苷酸序列中的核糖核苷酸数目为第二反义核糖核苷酸序列中的核糖核苷酸数目的90％至110％,或(a)和(b)两者,以及

(vi)或者(a)第一有义核糖核苷酸序列中的核糖核苷酸数目与第一反义核糖核苷酸序列中的核糖核苷酸数目相同,或(b)第二有义核糖核苷酸序列中的核糖核苷酸数目与第二反义核糖核苷酸序列中的核糖核苷酸数目相同,或(a)和(b)两者。

43.分离的和/或外源的多核苷酸,或包含该多核苷酸的载体,其编码权利要求1至42中任一项的RNA分子,其中所述多核苷酸可操作地连接到能够指导所述RNA分子在宿主细胞中表达的启动子。

44.宿主微生物,其包含(a)权利要求1至42中任一项的RNA分子,或(b)编码所述RNA分子的外源多核苷酸或载体,或(a)和(b)两者。

45.体外细胞,其包含(a)权利要求1至42中任一项的RNA分子,或(b)编码所述RNA分子的外源多核苷酸或载体,或(a)和(b)两者。

46.宿主细胞,其包含(a)权利要求1至42中任一项的RNA分子,或(b)编码所述RNA分子的外源多核苷酸或载体,或(a)和(b)两者,其中所述细胞是死亡的或不能繁殖。

47.权利要求46所述的宿主细胞,其中所述外源多核苷酸稳定地整合到所述细胞的基因组中。

48.一种产生权利要求1至42中任一项的RNA分子的方法,所述方法包括在宿主细胞或无细胞表达系统中表达编码所述RNA分子的多核苷酸。

49.一种生产转基因植物细胞的方法,该方法包括将编码权利要求1至42中任一项的RNA分子的外源多核苷酸引入植物细胞中,使得所述外源多核苷酸稳定地整合到所述转基因植物细胞的基因组中。

50.一种生产宿主细胞、体外细胞、或植物或其部分的方法,所述宿主细胞、体外细胞、或植物或其部分包含(a)权利要求1至42中任一项的RNA分子,或(b)编码所述RNA分子的外

源多核苷酸或载体,或(a)和(b)两者,所述方法包括使所述细胞、或植物或其部分与所述RNA分子、外源多核苷酸或载体接触。

51.一种组合物,其包含根据权利要求1至42中任一项的RNA分子和一种或多种合适的载体。

52.宿主细胞的提取物,其中所述提取物包含根据权利要求1至42中任一项的RNA分子,和/或编码所述RNA分子的多核苷酸或载体。

53.一种用于下调节肢动物或植物中靶RNA分子的水平和/或活性的方法,所述方法包括向所述节肢动物或植物递送根据权利要求1至42中任一项的RNA分子。

54.一种用于下调节肢动物中靶RNA分子的水平和/或活性的方法,所述方法包括向所述节肢动物递送根据权利要求44的宿主微生物或包含根据权利要求1至42中任一项的RNA分子的植物细胞。

55.一种用于下调节肢动物中靶RNA分子的水平和/或活性的方法,所述方法包括向所述节肢动物递送权利要求46的宿主细胞。

56.一种控制节肢动物或植物的方法,所述方法包括向所述节肢动物或植物递送根据权利要求1至42中任一项的RNA分子,其中所述RNA分子对所述节肢动物或植物具有有害作用。

57.一种控制节肢动物的方法,所述方法包括向所述节肢动物递送权利要求44的宿主微生物,其中所述RNA分子对所述节肢动物具有有害作用。

58.一种控制节肢动物的方法,所述方法包括向所述节肢动物递送权利要求46的宿主细胞,其中所述RNA分子对所述节肢动物具有有害作用。

59.权利要求57所述的方法,其用于控制昆虫。

60.权利要求59所述的方法,其中所述RNA分子在植物或其部分中产生,并且昆虫吃掉该植物或其部分,从而控制昆虫。

61.权利要求56所述的方法,其中所述RNA分子作为具有合适载体的制剂递送并且喷雾、撒粉或以其他方式施用至植物或其部分的表面。

62.根据权利要求1至42中任一项的RNA分子,用于预防或治疗受试者的疾病,其中所述RNA分子对所述疾病的至少一种症状或预防所述疾病具有有益效果。

63.根据权利要求1至42中任一项的RNA分子,用于制造用于预防或治疗受试者的疾病的药物,其中所述RNA分子对所述疾病的至少一种症状具有有益效果。

64.编码根据权利要求1至42中任一项的RNA分子的多核苷酸或载体,用于制造用于预防或治疗受试者的疾病的药物,其中所述RNA分子对所述疾病的至少一种症状具有有益效果。

65.权利要求44所述的宿主微生物,用于制造用于预防或治疗受试者的疾病的药物,其中所述RNA分子对所述疾病的至少一种症状具有有益效果。

66.权利要求46所述的宿主细胞,用于制造用于预防或治疗受试者的疾病的药物,其中所述RNA分子对所述疾病的至少一种症状具有有益效果。

67.权利要求63所述的RNA分子,其中所述RNA分子或所述药物是局部、口服或注射施用的。

68.将RNA分子引入植物的方法,所述方法包括将包含根据权利要求1至42中任一项的

RNA 分子的组合物局部施用到所述植物或其部分。

RNA分子

[0001] 本申请是申请日为2018年9月17日,申请号为201880072998.2(PCT申请号为PCT/AU2018/051015),发明名称为“RNA分子”的分案申请。

技术领域

[0002] 本发明涉及新的双链RNA(dsRNA)结构及其在基因沉默中的用途。

背景技术

[0003] 用于在动物细胞中诱导RNA干扰(RNAi)的常规dsRNA分子是通过将单链有义和反义RNA或自互补RNA退火而形成的,所述自互补RNA具有允许形成茎-环结构的互补性区,也称为发夹RNA(hpRNA)。在植物中,已经证明表达长发夹RNA的转基因在诱导RNAi方面是高度有效的。一些对长hpRNA的研究表明,相对于hpRNA结构的开放端,靠近hpRNA环末端的序列所形成的短干扰RNA(siRNA)的加工RNA产物更加丰富,提示hpRNA的dsRNA茎的Dicer加工似乎是从环末端开始的。整个环可以从hpRNA结构上精确切除,这可以为RNA酶Dicer结合和启动siRNA加工提供dsRNA末端。

[0004] 虽然dsRNA诱导的基因沉默已被证明是改变生物体表型的有价值工具,但需要可用于RNAi的可选的、优选改善的dsRNA分子。

发明内容

[0005] 本发明人构思了RNAi诱导分子的新设计,在此称为环末端dsRNA(ledRNA),其具有一个或多个以下特征;容易合成,更容易形成dsRNA结构并诱导真核细胞中靶基因的有效沉默。当局部施用于植物叶时,ledRNA也是有效的。

[0006] 在第一方面,本发明提供核糖核酸(RNA)分子,其包含第一RNA组分、与第一RNA组分共价连接的第二RNA组分,以及任选的以下的一种或多种或全部:(i)与第一和第二RNA组分共价连接的连接核糖核苷酸序列,(ii)5’前导序列和(iii)3’尾随序列,

[0007] 其中所述第一RNA组分按5’至3’顺序由第一5’核糖核苷酸、第一RNA序列和第一3’核糖核苷酸组成,其中所述RNA分子中所述第一5’和3’核糖核苷酸彼此碱基配对,其中所述第一RNA序列包含至少20个邻接核糖核苷酸的第一有义核糖核苷酸序列、至少4个核糖核苷酸的第一环序列和至少20个邻接核糖核苷酸的第一反义核糖核苷酸序列,其中第一反义核糖核苷酸序列与RNA分子中的第一有义核糖核苷酸序列杂交,其中第一反义核糖核苷酸序列能够与靶RNA分子的第一区杂交,

[0008] 其中,如果存在连接核糖核苷酸序列,则经由所述连接核糖核苷酸序列,或如果不存在连接核糖核苷酸序列,则直接地,所述第二RNA组分与所述第一5’核糖核苷酸共价连接或所述第一3’核糖核苷酸,

[0009] 其中所述第二RNA组分按5’至3’顺序由第二5’核糖核苷酸、第二RNA序列和第二3’核糖核苷酸组成,其中所述第二5’和3’核糖核苷酸在所述RNA分子中彼此碱基配对,其中所述第二RNA序列包含第二有义核糖核苷酸序列、至少4个核糖核苷酸的第二环序列和第二反

义核糖核苷酸序列,其中所述第二有义核糖核苷酸序列与所述RNA分子中的所述第二反义核糖核苷酸序列杂交;

[0010] 其中如果存在5'前导序列,则所述5'前导序列由下述核糖核苷酸序列组成,如果所述第二RNA组分与所述第一3'核糖核苷酸连接,则所述核糖核苷酸序列与所述第一5'核糖核苷酸共价连接,或者如果所述第二RNA组分与所述第一5'核糖核苷酸连接,则所述核糖核苷酸序列与所述第二5'核糖核苷酸共价连接,并且

[0011] 其中如果存在3'尾随序列,则所述3'尾随序列由下述核糖核苷酸序列组成,如果所述第二RNA组分与所述第一3'核糖核苷酸连接,则所述核糖核苷酸序列与所述第二3'核糖核苷酸共价连接,或者如果所述第二RNA组分与所述第一5'核糖核苷酸连接,则所述核糖核苷酸序列与所述第一3'核糖核苷酸共价连接。

[0012] 在第二方面,本发明提供了一种RNA分子,其包含第一RNA组分、与第一RNA组分共价连接的第二RNA组分,以及任选的以下的一种或多种或全部:(i)与第一和第二RNA组分共价连接的连接核糖核苷酸序列,(ii)5'前导序列和(iii)3'尾随序列,

[0013] 其中所述第一RNA组分按5'至3'顺序由第一5'核糖核苷酸、第一RNA序列和第一3'核糖核苷酸组成,其中所述第一5'和3'核糖核苷酸碱基对,其中所述第一RNA序列包含第一有义核糖核苷酸序列、至少4个核糖核苷酸的第一环序列和第一反义核糖核苷酸序列,其中所述第一有义核糖核苷酸序列和第一反义核糖核苷酸序列各自由至少20个邻接的核糖核苷酸组成,其中所述第一有义核糖核苷酸序列的至少20个邻接的核糖核苷酸与所述第一反义核糖核苷酸序列的至少20个邻接的核糖核苷酸完全碱基配对,其中所述第一有义核糖核苷酸序列的至少20个邻接的核糖核苷酸或所述第一反义核糖核苷酸序列的至少20个邻接的核糖核苷酸在序列上分别与靶RNA分子或其互补体或两者的第一区相同,

[0014] 其中,如果存在所述连接核糖核苷酸序列,则经由所述连接核糖核苷酸序列,所述第二RNA组分与所述第一5'核糖核苷酸共价连接或所述第一3'核糖核苷酸,

[0015] 其中所述第二RNA组分按5'至3'顺序由第二5'核糖核苷酸、第二RNA序列和第二3'核糖核苷酸组成,其中所述第二5'和3'核糖核苷酸碱基配对,其中所述第二RNA序列包含第二有义核糖核苷酸序列、至少4个核糖核苷酸的第二环序列和第二反义核糖核苷酸序列,其中所述第二有义核糖核苷酸序列与所述第二反义核糖核苷酸序列碱基配对,

[0016] 其中如果存在5'前导序列,则所述5'前导序列由下述核糖核苷酸序列组成,如果所述第二RNA组分与所述第一3'核糖核苷酸连接,则所述核糖核苷酸序列与所述第一5'核糖核苷酸共价连接,或者如果所述第二RNA组分与所述第一5'核糖核苷酸连接,则所述核糖核苷酸序列与所述第二5'核糖核苷酸共价连接,并且

[0017] 其中如果存在3'尾随序列,则所述3'尾随序列由下述核糖核苷酸序列组成,如果所述第二RNA组分与所述第一3'核糖核苷酸连接,则所述核糖核苷酸序列与所述第二3'核糖核苷酸共价连接,或者如果所述第二RNA组分与所述第一5'核糖核苷酸连接,则所述核糖核苷酸序列与所述第一3'核糖核苷酸共价连接。

[0018] 在一个优选的实施方案中,本发明的RNA分子是嵌合RNA分子。

[0019] 在第三方面,本发明提供了嵌合核糖核酸(RNA)分子,其包含双链RNA(dsRNA)区,所述双链RNA区包含能够彼此杂交以形成dsRNA区的有义核糖核苷酸序列和反义核糖核苷酸序列,其中

[0020] i)所述有义核糖核苷酸序列由按5’至3’顺序共价连接的第一个5’核糖核苷酸、第一RNA序列和第一3’核糖核苷酸组成，

[0021] ii)所述反义核糖核苷酸序列由按5’至3’顺序共价连接的第二5’核糖核苷酸、第二RNA序列和第二3’核糖核苷酸组成，

[0022] iii)所述第一5’核糖核苷酸碱基与所述第二3’核糖核苷酸碱基配对，以形成所述dsRNA区的末端碱基对，

[0023] iv)所述第二5’核糖核苷酸碱基与所述第一3’核糖核苷酸碱基配对，以形成所述dsRNA区的末端碱基对，

[0024] v)所述有义核糖核苷酸序列和所述反义核糖核苷酸序列的总体上约5％至约40％的核糖核苷酸以非规范碱基对被碱基配对或者不被碱基配对，

[0025] vi)所述dsRNA区不包含20个邻接的规范碱基对，

[0026] vii)所述RNA分子能够在真核细胞中或在体外被加工，由此所述反义核糖核苷酸序列被切割以产生长度为20-24个核糖核苷酸的短反义RNA(asRNA)分子，

[0027] viii)所述RNA分子或至少一些asRNA分子，或两者，能够降低所述真核细胞中靶RNA分子的表达或活性，以及

[0028] ix)所述RNA分子能够通过在体外或在细胞中或两者中转录而酶促制备。

[0029] 在第四方面，本发明提供了嵌合RNA分子，其包含第一RNA组分和与第一RNA组分共价连接的第二RNA组分，

[0030] 其中所述第一RNA组分包含第一双链RNA(dsRNA)区和至少4个核苷酸的第一间插核糖核苷酸序列，所述第一双链RNA(dsRNA)区包含能够相互杂交形成第一dsRNA区的第一有义核糖核苷酸序列和第一反义核糖核苷酸序列，所述第一间插核糖核苷酸序列共价连接所述第一有义核糖核苷酸序列和所述第一反义核糖核苷酸序列，

[0031] 其中第二RNA组分包含第二有义核糖核苷酸序列、第二反义核糖核苷酸序列和至少4个核糖核苷酸的第二间插核糖核苷酸序列，所述第二间插核糖核苷酸序列共价连接第二有义核糖核苷酸序列和第二反义核糖核苷酸序列，其中RNA分子中的第二有义核糖核苷酸序列与第二反义核糖核苷酸序列杂交，

[0032] 其中在所述第一RNA组分中，

[0033] i)第一有义核糖核苷酸序列由至少20个按5’至3’顺序共价连接的邻接核糖核苷酸、第一5’核糖核苷酸、第一RNA序列和第一3’核糖核苷酸组成，

[0034] ii)第一反义核糖核苷酸序列由至少20个按5’至3’顺序共价连接的邻接核糖核苷酸、第二5’核糖核苷酸、第二RNA序列和第二3’核糖核苷酸组成，

[0035] iii)具有第二3’核糖核苷酸的第一5’核糖核苷酸碱基对，

[0036] iv)具有第一3’核糖核苷酸的第二5’核糖核苷酸碱基对，

[0037] v)第一有义核糖核苷酸序列和第一反义核糖核苷酸序列的总体上5％至40％的核糖核苷酸以非规范碱基对被碱基配对或者不被碱基配对，以及

[0038] vi)所述第一dsRNA区不包含20个邻接的规范碱基对，

[0039] 其中所述嵌合RNA分子能够在真核细胞中或在体外被加工，由此所述第一反义核糖核苷酸序列被切割以产生长度为20-24个核糖核苷酸的短反义RNA(asRNA)分子，并且其中

[0040] (a)所述嵌合RNA分子或至少一些asRNA分子,或两者,能够降低所述真核细胞中靶RNA分子的表达或活性,或

[0041] (b)所述第一反义核糖核苷酸序列包含至少20个邻接核糖核苷酸的序列,所述邻接核糖核苷酸在序列上与所述靶RNA分子的互补区至少50%相同,或

[0042] (c)(a)和(b)二者。

[0043] 在第一、第二和第四方面,第一RNA组分的第一5'核糖核苷酸和第一3'核糖核苷酸碱基配对以形成碱基对。该碱基对在本文中定义为通过第一RNA组分的自杂交形成的dsRNA区的末端碱基对。在第一有义核糖核苷酸序列与第一5'核糖核苷酸共价连接而没有任何间插核苷酸且第一反义核糖核苷酸序列与第一3'核糖核苷酸共价连接而没有任何间插核苷酸的实施方案中,第一5'核糖核苷酸直接与义序列和反义序列中的一个连接连接且第一3'核糖核苷酸与有义序列和反义序列中的另一个直接连接。

[0044] 在优选的实施方案中,第一反义核糖核苷酸序列的至少20个邻接的核糖核苷酸都能够与靶RNA分子的第一区的核苷酸碱基配对。在一个实施方案中,第一有义核糖核苷酸序列与第一5'核糖核苷酸共价连接而没有任何间插核苷酸,或第一反义核糖核苷酸序列与第一3'核糖核苷酸共价连接而没有任何间插核苷酸,或两者。

[0045] 在一个实施方案中,RNA分子包含连接核糖核苷酸序列,其中连接核糖核苷酸序列在序列上与靶基因相关,或者至少部分与靶RNA分子的区相同或者与其互补。在优选的实施方案中,连接核糖核苷酸序列与第一和第二RNA组分中的有义序列一起形成一个邻接有义序列的一部分,或与第一和第二RNA组分中的反义序列一起形成一个邻接反义序列的一部分。在一个实施方案中,RNA分子包含连接核糖核苷酸序列,其中连接核糖核苷酸序列小于20个核糖核苷酸。在一个实施方案中,连接核糖核苷酸序列与靶RNA分子杂交。在一个实施方案中,连接核糖核苷酸序列与靶RNA分子的互补体的一部分相同。在一个实施方案中,连接核糖核苷酸序列在1至10个核糖核苷酸之间。

[0046] 在第一、第二或第四方面的实施方案中,RNA分子包含以下的一个或多个或全部:(i)共价连接第一和第二RNA组分的连接核糖核苷酸序列,(ii)5'延伸序列和(iii)3'延伸序列,其中分别地,如果存在5'延伸序列,则其由与第一RNA组分或第二RNA组分共价连接的核糖核苷酸序列组成,并且其中如果存在3'延伸序列,则其由与第二RNA组分或第一RNA组分共价连接的核糖核苷酸序列组成。在一个实施方案中,第一RNA组分和第二RNA组分通过连接核糖核苷酸序列共价连接。在另一个实施方案中,第一RNA组分和第二RNA组分直接连接,没有任何连接核糖核苷酸序列。

[0047] 在第一至第四方面的实施方案中,RNA分子包含两个或更多个有义核糖核苷酸序列,每个有义核糖核苷酸序列在序列上与靶RNA分子的区相同,并且RNA分子包含基于与有义核糖核苷酸序列配对的一个或多个反义核糖核苷酸序列,其中一个或多个反义序列与靶分子的区互补,优选完全互补。在一个实施方案中,两个或更多个有义核糖核苷酸序列在序列上与同一靶RNA分子的不同区相同,其在靶RNA分子中可以是或可以不是邻接的。在一个实施方案中,两个或更多个有义核糖核苷酸序列在序列上与不同靶RNA分子的区相同。在一个实施方案中,两个或多个有义核糖核苷酸序列没有间插环序列,即它们相对于靶RNA分子是邻接的。

[0048] 在第一至第四方面的优选实施方案中,RNA分子包含两个或更多个反义核糖核苷

酸序列和与其碱基配对的有义核糖核苷酸序列,所述反义序列各自与靶RNA分子的区互补。靶RNA分子与其互补的区在靶RNA分子中可以是或不是邻接的。在一个实施方案中,所述两个或更多个反义核糖核苷酸序列与同一靶RNA分子的不同区互补。在一个实施方案中,所述两个或更多个反义核糖核苷酸序列中的第二个与不同于所述两个或更多个反义核糖核苷酸序列中的第一个的靶RNA分子的区互补。在一个优选的实施方案中,两个或多个反义核糖核苷酸序列没有间插环序列,即它们相对于靶RNA分子的互补序列是邻接的。在优选的实施方案中,两个或多个反义核糖核苷酸序列和有义核糖核苷酸序列中的一者或两者沿着它们的全长通过规范碱基对被碱基配对,或通过一些规范和一些非规范碱基对,优选G:U碱基对被碱基配对。

[0049] 在第一至第四方面的优选实施方案中,RNA分子是单链核糖核苷酸。例如,RNA分子可以包含具有5'端的单链核糖核苷酸、长度为至少21个核苷酸的至少一个有义核糖核苷酸序列、在至少21个邻接核苷酸上与每个有义核糖核苷酸序列完全碱基配对的反义核糖核苷酸序列、至少两个环序列和3'端。5'至3'顺序可以是有义核糖核苷酸序列,然后是反义核糖核苷酸序列,或反之亦然。在一个实施方案中,5'端的核糖核苷酸和3'端的核糖核苷酸是相邻的,每个碱基都配对并且不直接共价结合,参见例如图1。

[0050] 在第一至第四方面的另一个实施方案中,RNA分子包含与靶RNA的第一区杂交的第一反义核糖核苷酸序列;与靶RNA的第二区杂交的第二反义核糖核苷酸序列,所述靶RNA的第二区与所述靶RNA的第一区不同;并且所述RNA分子仅包含一个与所述靶RNA杂交的有义核糖核苷酸序列,其中所述两个反义序列在所述RNA分子中不邻接。在一个实施方案中,靶RNA的第一和第二区在靶RNA中是邻接的。或者,它们不是邻接的。

[0051] 在第一至第四方面的另一个实施方案中,RNA分子包含与靶RNA的第一区至少60%相同的第一有义核糖核苷酸序列;和与靶RNA的第二区至少60%相同的第二有义核糖核苷酸序列,所述靶RNA的第二区与所述靶RNA的第一区不同;并且所述RNA分子仅包含一个与所述靶RNA杂交的反义核糖核苷酸序列,其中所述两个有义序列在所述RNA分子中不邻接。在一个实施方案中,靶RNA的第一和第二区在靶RNA分子中是邻接的。或者,它们不是邻接的。在优选的实施方案中,第一和第二有义核糖核苷酸序列各自独立地与靶RNA的相应区至少70%、至少80%、至少90%、至少95%或至少99%相同,即第一有义序列可以与其靶区至少70%相同,第二序列与其靶序列至少80%相同,等等。

[0052] 在第一至第四方面的优选实施方案中,RNA分子是单链核糖核苷酸,其具有5'端、长度为至少21个核苷酸的至少一个有义核糖核苷酸序列、在至少21个邻接核苷酸上与每个有义核糖核苷酸序列完全碱基配对的反义核糖核苷酸序列、至少两个环序列和3'端。在更优选的实施方案中,RNA分子中的碱基配对包含在长度为至少21个邻接碱基对的双链区中,该双链区包括一些非规范碱基对,最优选一些G:U碱基对,该双链区包含至少一个长度为至少21个核苷酸的有义核糖核苷酸序列。

[0053] 在第一和第二方面的优选实施方案中,第二RNA组分的特征在于:

[0054] i)第二有义核糖核苷酸序列由至少20个邻接的按5'至3'顺序共价连接的第二5'核糖核苷酸、第三RNA序列和第三3'核糖核苷酸组成,

[0055] ii)第二反义核糖核苷酸序列由至少20个按5'至3'顺序共价连接的邻接核糖核苷酸、第三5'核糖核苷酸、第四RNA序列和第二3'核糖核苷酸组成,

[0056] iii)具有第二3'核糖核苷酸的第二5'核糖核苷酸碱基对,

[0057] iv)具有第三5'核糖核苷酸的第三3'核糖核苷酸碱基对,

[0058] 其中所述嵌合RNA分子能够在真核细胞中或在体外被加工,由此所述第二反义核糖核苷酸序列被切割以产生长度为20-24个核糖核苷酸的短反义RNA(asRNA)分子,并且其中最优选地,在没有或与从第一RNA组分的第一反义序列产生的asRNA组合的情况下,从第二反义序列产生的asRNA分子能够降低靶RNA的表达。

[0059] 在第四方面的优选实施方案中,第二RNA组分的特征在于:

[0060] i)第二有义核糖核苷酸序列由至少20个邻接的按5'至3'顺序共价连接的第三5'核糖核苷酸、第三RNA序列和第三3'核糖核苷酸组成,

[0061] ii)第二反义核糖核苷酸序列由至少20个按5'至3'顺序共价连接的邻接核糖核苷酸、第四5'核糖核苷酸、第四RNA序列和第四3'核糖核苷酸组成,

[0062] iii)具有第四3'核糖核苷酸的第三5'核糖核苷酸碱基对,

[0063] iv)具有第三5'核糖核苷酸的第三3'核糖核苷酸碱基对,

[0064] 其中所述嵌合RNA分子能够在真核细胞中或在体外被加工,由此所述第二反义核糖核苷酸序列被切割以产生长度为20-24个核糖核苷酸的短反义RNA(asRNA)分子,并且其中

[0065] 在这些优选的实施方案中,更优选第二有义核糖核苷酸序列和第二反义核糖核苷酸序列总体上5%至40%的核糖核苷酸以非规范碱基对被碱基配对或者不被碱基配对,和/或第二dsRNA区不包含20个邻接的规范碱基对。更优选地,第二有义核糖核苷酸序列和第二反义核糖核苷酸序列的核糖核苷酸中总体上约12%、约15%、约18%、约21%、约24%,或15%至30%,或甚至更优选16%至25%以非规范碱基对被碱基配对,或者不被碱基配对。在优选的实施方案中,所述第二dsRNA区的非规范碱基对的至少50%、至少60%、至少70%、至少80%、至少90%、至少95%、至少97%,或100%是G:U碱基对。最优选地,在这些实施方案中,

[0066] (a)所述嵌合RNA分子或至少一些asRNA分子,或两者,能够降低所述真核细胞中靶RNA分子的表达或活性,或

[0067] (b)所述第二反义核糖核苷酸序列包含至少20个邻接核糖核苷酸的序列,所述邻接核糖核苷酸在序列上与靶RNA分子的互补区至少50%相同、优选至少60%相同、更优选至少70%相同、甚至更优选至少80%相同,最优选与靶RNA分子的补体区至少90%相同或100%相同,或

[0068] (c)(a)和(b)二者。

[0069] 在第一至第四方面的实施方案中,RNA分子包含5'前导序列或5'延伸序列。在一个实施方案中,RNA分子包含3'尾随序列或3'延伸序列。在一个优选的实施方案中,RNA分子包含5'前导/延伸序列和3'尾随/延伸序列。

[0070] 在第一至第四方面的实施方案中,RNA分子的每个核糖核苷酸与另外两个核苷酸共价连接,即它是共价的闭合环。或者,RNA分子可表示为哑铃形(图1),但在双链结构的一部分中具有缺口或裂口。

[0071] 在第一至第四方面的实施方案中,RNA分子的至少一个或全部环序列长于20个核苷酸。在优选的实施方案中,RNA分子的至少一个环的长度为4至1,000个核糖核苷酸。在更

优选的实施方案中,所有环的长度为4至1,000个核糖核苷酸。在更优选的实施方案中,RNA分子的至少一个环的长度为4至200个核糖核苷酸。在甚至更优选的实施方案中,所有环的长度为4至200个核糖核苷酸。在甚至更优选的实施方案中,RNA分子的至少一个环的长度为4至50个核糖核苷酸。在最优选的实施方案中,所有环的长度为4-50个核糖核苷酸。在一个实施方案中,真核细胞是脊椎动物细胞,并且RNA分子的每个环的长度为20至50个,或20至30个核糖核苷酸。

[0072] 在第一至第四方面的实施方案中,RNA分子在双链区中没有,或有一个或两个或更多个隆起。在本文中,隆起是有义或反义核糖核苷酸序列中的核苷酸,或两个或更多个邻接核苷酸,其在dsRNA区中不被碱基配对并且在dsRNA区的互补序列中的相应位置不具有错配核苷酸。RNA分子的dsRNA区可以包含在有义或反义序列或两者中超过2个或3个核苷酸的序列,当dsRNA结构形成时,其从dsRNA区成环。环出的序列本身可以形成一些内部碱基配对,例如它本身可以形成茎-环结构。

[0073] 在第一至第四方面的实施方案中,RNA分子在双链区中没有,或有一个或两个或更多个隆起。在本文中,隆起是有义或反义核糖核苷酸序列中的核苷酸,或两个或更多个邻接核苷酸,其在dsRNA区中不被碱基配对并且在dsRNA区的互补序列中的相应位置不具有错配核苷酸。RNA分子的dsRNA区可以包含在有义或反义序列或两者中超过2个或3个核苷酸的序列,当dsRNA结构形成时,其从dsRNA区成环。环出的序列本身可以形成一些内部碱基配对,例如它本身可以形成茎-环结构。

[0074] 在一个实施方案中,RNA分子具有三个、四个或更多个环。在优选的实施方案中,RNA分子仅具有两个环。在一个实施方案中,RNA分子的第一双链区或第一dsRNA区和第二dsRNA区包含在双链区中不被碱基配对的一个或两个或更多个核苷酸,或在双链区中不被碱基配对的核苷酸的至多2%、3%、4%、5%、6%、7%、8%、9%或10%。

[0075] 在第一至第四方面的优选实施方案中,本发明的靶RNA分子或RNA分子或两者在真核细胞中。例如,真核细胞可以是植物细胞、动物细胞或真菌细胞。在一个实施方案中,真核细胞是真菌细胞。在一个实施方案中,本发明的RNA分子在细胞例如细菌细胞或其他微生物细胞中产生,所述细胞不同于包含靶RNA的细胞。类似地,在一个实施方案中,在生产本发明的RNA分子时,本发明的RNA分子是在不包含靶RNA的真核细胞中生产的,但是例如如果靶RNA是病毒RNA或其他引入的RNA,则包含本发明的RNA分子和/或其加工的RNA产物的真核细胞可以成为靶RNA的宿主。这类细胞可以针对病毒或其他引入的RNA被预防性地保护。

[0076] 在第一至第四方面的优选实施方案中,RNA分子能够通过体外或细胞中或两者中的转录酶促制备。在一个实施方案中,本发明的RNA分子在细胞中表达,即通过从编码RNA分子的一种或多种核酸转录而在细胞中产生。编码RNA分子的一种或多种核酸优选为DNA分子,其可存在于细胞中的载体上或整合入细胞的基因组,细胞的核基因组或细胞的质体DNA中。编码RNA分子的一种或多种核酸也可以是RNA分子,如病毒载体。

[0077] 因此,在一个实施方案中,本发明提供了包含本文所述RNA分子的细胞。在优选的实施方案中,本发明提供了在细胞中表达并且已经从细胞中分离和/或纯化的本文所述的RNA分子。因此,本发明提供了根据第一至第四方面中的一个或多个的分离的RNA分子的制剂,其适于施用至包含靶RNA或潜在地包含靶RNA的细胞。

[0078] 在一个实施方案中,一种或多种靶RNA编码蛋白质。或者,一种或多种靶RNA不编码

蛋白质,例如rRNA、tRNA、snoRNA或miRNA。

[0079] 在第一至第四方面的实施方案中,总体上形成dsRNA区的有义核糖核苷酸序列和反义核糖核苷酸序列的核糖核苷酸的约12%、约15%、约18%、约21%、约24%,或约15%至约30%,或优选约16%至约25%以非规范碱基对被碱基配对或不被碱基配对。在优选的实施方案中,dsRNA区或RNA分子中所有dsRNA区中至少50%、至少60%、至少70%、至少80%、至少90%、至少95%、至少97%或100%的非规范碱基对是G:U碱基对。每个G:U碱基对中的G核苷酸可以独立地在有义核糖核苷酸序列中或优选在反义核糖核苷酸序列中。关于dsRNA区的G:U碱基对中的G核苷酸,优选地至少50%在反义核糖核苷酸序列中,更优选地至少60%或70%,甚至更优选地至少80%或90%,并且最优选地至少95%在dsRNA区的反义核糖核苷酸序列中。该特征可应用于RNA分子中的所有dsRNA区。在一个实施方案中,RNA分子中总体上小于25%、小于20%、小于15%、小于10%、优选小于5%、更优选小于1%或最优选没有核糖核苷酸被碱基配对。在优选的实施方案中,dsRNA区中或dsRNA区中总体上,6个核糖核苷酸中的每4个至每1个形成非规范碱基对或在RNA分子内不被碱基配对。在优选的实施方案中,dsRNA区或dsRNA区中总体上不包含8个邻接的规范碱基对。在另一个实施方案中,dsRNA区包含至少8个邻接的规范碱基对,例如8-12或8-14个邻接的规范碱基对。在优选的实施方案中,dsRNA区中的所有核糖核苷酸以规范碱基对或非规范碱基对被碱基配对。在一个实施方案中,有义核糖核苷酸序列的一个或多个核糖核苷酸或反义核糖核苷酸序列的一个或多个核糖核苷酸,或两者,不被碱基配对。在一个实施方案中,在本发明的RNA分子中,每个有义核糖核苷酸序列的一个或多个核糖核苷酸和每个反义核糖核苷酸序列的一个或多个核糖核苷酸不被碱基配对。

[0080] 在一个实施方案中,第一RNA组分或第二RNA组分或两者的反义RNA序列在序列上与靶RNA分子的一个区或两个这样的区的互补体与小于100%相同,或约80%至99.9%相同,或约90%至98%相同,或约95%至98%相同、优选98%至99.9%相同,所述两个区在靶RNA分子中可以是或可以不是邻接的。在优选的实施方案中,反义RNA序列在序列上与靶RNA分子的互补区100%相同,所述区为例如包含21、23、25、27、30或32个邻接核苷酸的区。在一个实施方案中,所述有义或反义核糖核苷酸序列或两者长度为至少40、至少50、至少约100、至少约200、至少约300、至少约400、至少约500、至少约600、至少约700、至少约800、至少约900、至少约1,000,或约100至约1,000个邻接核苷酸。当在植物细胞或真菌细胞或者对于非脊椎动物细胞中使用RNA分子时,优选至少100个核苷酸的长度。当在脊椎动物细胞中使用RNA分子时,dsRNA中有义和反义核糖核苷酸序列的长度为50个核苷酸或更短,例如31至50个核苷酸是优选的。在一个实施方案中,有义核糖核苷酸序列中的核糖核苷酸数目为反义核糖核苷酸序列中的核糖核苷酸数目的约90%至约110%、优选95%至105%、更优选98%至102%、甚至更优选99%至101%。在最优选的实施方案中,有义核糖核苷酸序列中的核糖核苷酸数目与反义核糖核苷酸序列中的核糖核苷酸数目相同。这些特征可应用于RNA分子中的每个dsRNA区。

[0081] 在第一至第四方面的实施方案中,RNA分子中的第一3'核糖核苷酸和第二5'核糖核苷酸通过由至少4个核糖核苷酸,或4至1,000个核糖核苷酸,或优选4至200个核糖核苷酸、更优选4至50个核糖核苷酸组成的环序列共价连接。在一个实施方案中,RNA分子还包含与第一5'核糖核苷酸共价连接的5'延伸序列或与第二3'核糖核苷酸共价连接的3'延伸序

列,或两者。在一个实施方案中,嵌合RNA分子还包含与第二5'核糖核苷酸共价连接的5'延伸序列或与第一3'核糖核苷酸共价连接的3'延伸序列,或两者。在该实施方案中,RNA分子包含杂交形成RNA分子的两条分开的RNA链,尽管它可能已经通过从核酸分子转录为单一RNA转录物并随后被加工以包含两条RNA链而产生。

[0082] 在从任何内含子剪接出来之后但在通过Dicer酶或其他RNA酶对RNA分子进行任何加工之前,作为单链RNA产生的本发明的RNA分子的总长度典型地在50至2000个核糖核苷酸、优选地60或70至2000个核糖核苷酸、更优选地在80或90至2000个核糖核苷酸、甚至更优选地100或110至2000个核糖核苷酸。在优选的实施方案中,RNA分子的最小长度是120、130、140、150、160、180或200个核苷酸,并且最大长度是400、500、600、700、800、900、1000、1200、1400、1500或2000个核糖核苷酸。可以考虑上述最小和最大长度的每种组合。通过体外转录或在诸如细菌的细胞或其他微生物细胞中或在其中靶基因被下调的真核细胞中转录,可容易地实现此类长度的RNA分子的产生。

[0083] 在第一至第四方面的实施方案中,所述嵌合RNA分子包含两个或更多个相同或优选不同的dsRNA区。

[0084] 在第一至第四方面的优选实施方案中,RNA分子在真核细胞中表达,即通过在细胞中转录产生。在这些实施方案中,与具有与规范碱基对被完全碱基配对的相应dsRNA区的类似RNA分子的加工相比,通过加工长度为22和/或20个核糖核苷酸的RNA分子形成更大比例的dsRNA分子。即,这些实施方案的RNA分子比其dsRNA区与规范碱基对被完全碱基配对的类似RNA分子更容易被加工,以提供22-和/或20-核糖核苷酸短反义RNA,作为由该RNA分子产生的总数20-24个核苷酸asRNA的一部分。

[0085] 在一个实施方案中,本发明的RNA分子包含本文所述RNA分子的两个或更多个特征的组合。

[0086] 另一方面,本发明提供了编码本文所述RNA分子,优选本文所述嵌合RNA分子的多核苷酸。在一个实施方案中,多核苷酸是可以整合到诸如染色体的较大DNA分子中的DNA构建体。在一个实施方案中,多核苷酸与能够指导RNA分子在宿主细胞中表达的启动子可操作地连接。所述宿主细胞可以是:细菌细胞,如大肠杆菌;真菌细胞,如酵母细胞;或真核细胞,如植物细胞或动物细胞。在一个实施方案中,启动子相对于多核苷酸是异源的。编码RNA分子的多核苷酸可以是嵌合或重组多核苷酸,或分离的和/或外源多核苷酸。在一个实施方案中,启动子可以在体外发挥功能,例如噬菌体启动子,诸如T7 RNA聚合酶启动子或SP6RNA聚合酶启动子。在一个实施方案中,启动子是RNA聚合酶III启动子,诸如U6启动子或H1启动子。在一个实施方案中,启动子是RNA聚合酶II启动子,其可以是组成型启动子、组织特异性启动子、发育调节的启动子或诱导型启动子。在一个实施方案中,多核苷酸编码在至少一个环序列中包含内含子的RNA前体分子,所述环序列能够在多核苷酸在宿主细胞中转录期间或之后剪接出来。在一个实施方案中,本发明提供了包含本文所述的多核苷酸的载体。在一个实施方案中,所述载体是病毒载体。在一个实施方案中,载体是质粒载体,例如适合与根癌农杆菌(Agrobacterium tumefaciens)一起使用的二元载体。

[0087] 在一个实施方案中,其中本发明的多核苷酸或载体在真核宿主细胞中,优选在植物中,多核苷酸或载体的启动子区与编码本发明RNA分子的区可操作地连接,与编码具有与规范碱基对完全碱基配对的相应dsRNA区的RNA分子的相应多核苷酸或载体的启动子相比,

具有较低水平的甲基化。在一个实施方案中,当与相应多核苷酸或载体的启动子相比时,较低水平的甲基化小于50%、小于40%、小于30%或小于20%。在一个实施方案中,宿主细胞包含编码本发明RNA分子的多核苷酸或载体的至少两个拷贝。在该实施例中:

[0088] i)相对于具有所述多核苷酸或载体的单拷贝的相应真核细胞,靶RNA分子在真核细胞中的表达和/或活性的降低水平至少相同,和/或

[0089] ii)当与包含具有与规范碱基对完全碱基配对的相应dsRNA区的RNA分子的相应细胞相比时,所述真核细胞中靶RNA分子的表达和/或活性降低的水平较低。

[0090] 在另一方面,本发明提供了包含本文所述RNA分子、本文所述多核苷酸或包含其的载体的宿主细胞。在一个实施方案中,宿主细胞是非人细胞,如细菌细胞、真菌细胞、植物细胞或非人动物细胞。在一个实施方案中,所述细胞为非人细胞或细胞培养物中的人细胞。在一个实施方案中,细胞是真核细胞,例如除了动物细胞之外的细胞。在一个实施方案中,细胞是微生物细胞如原核细胞。在一个实施方案中,宿主细胞是活的。在另一个实施方案中,宿主细胞是死亡的。

[0091] 在另一方面,本发明提供了一种包含本发明的RNA分子、优选本文所述的嵌合RNA分子非人生物体,包含其的本发明的多核苷酸或载体,或包含其的宿主细胞。在一个实施方案中,非人生物体是转基因的,只要它包含本发明的多核苷酸即可。在一个实施方案中,多核苷酸稳定地整合到非人生物体的基因组中。

[0092] 在另一方面,本发明提供了生产本发明的RNA分子的方法,所述方法包括在宿主细胞或无细胞表达系统中表达本发明的多核苷酸。在该实施方案中,该方法可进一步包括至少部分纯化或不纯化RNA分子。

[0093] 在另一方面,本发明提供了生产细胞或非人生物体、优选植物或真菌的方法,该方法包括将本发明的多核苷酸或载体导入细胞,优选植物细胞或真菌,优选使得编码RNA分子的多核苷酸或载体或其部分稳定地整合到细胞的基因组中。在一个实施方案中,非人生物体例如通过再生植物从细胞或子代细胞产生。在一个实施方案中,通过将细胞或一个或多个子代细胞导入非人生物体来产生非人生物体。作为将多核苷酸或载体稳定整合到细胞基因组中的替代方案,可以将多核苷酸或载体引入细胞而不将多核苷酸或载体整合到基因组中,例如以在细胞或生物体中瞬时表达RNA分子。

[0094] 在另一方面,本发明提供了本发明的宿主细胞或生物体或其部分的提取物,其中所述提取物包含本发明的RNA分子、通过加工所述RNA分子产生的小RNA分子(长度为20-24nt),或两者,和/或本发明的多核苷酸或载体。在一个实施方案中,本发明提供了一种组合物,其包含一种或多种本发明的RNA分子、通过加工RNA分子产生的小RNA分子(长度为20-24nt),或两者;本发明的多核苷酸;本发明的载体;本发明的宿主细胞;或通过本发明的方法产生的提取物;以及一种或多种合适的运载体。在一个实施方案中,所述组合物是药物组合物,例如适合施用于人或其他动物的组合物。该药物组合物可以适用于预防或治疗疾病,或适用于局部应用,如化妆品应用。在一个实施方案中,组合物适于施用至植物、优选地田地中的植物或植物群,或施用至昆虫或昆虫群。在一个实施方案中,所述组合物适于施用至作物,例如通过喷洒至田地中的作物植物。

[0095] 在一个实施方案中,包含本发明的RNA分子或通过加工RNA分子产生的小RNA分子(长度为20-24nt)或两者的提取物或组合物还包含至少一种增强RNA分子或多核苷酸和/或

载体的稳定性的化合物,其中所述至少一种化合物有助于RNA分子、多核苷酸或载体被例如生物体细胞的细胞摄取。在一个实施方案中,所述化合物是转染促进剂,例如含脂质化合物。

[0096] 在另一个方面,本发明提供用于降低或下调细胞或生物体中靶RNA分子的水平和/或活性的方法,所述方法包括向所述细胞或生物体递送一种或多种本发明的RNA分子或通过加工所述RNA分子产生的小RNA分子(长度为20-24nt),或两者;本发明的多核苷酸;本发明的载体或本发明的组合物。在一个实施方案中,靶RNA分子编码蛋白质。在一个实施方案中,所述方法降低多于一种靶RNA分子的水平和/或活性,例如降低两种或更多种靶RNA的水平和/或活性,所述靶RNA分子是不同的,所述靶RNA的水平和/或活性在例如来自基因家族的序列上相关。

[0097] 在另一方面,本发明提供了防治非人生物体的方法,所述方法包括向所述非人生物体递送一种或多种本发明的RNA分子或通过加工所述RNA分子产生的小RNA分子(长度为20-24nt),或两者;本发明的多核苷酸或载体;本发明的宿主细胞;通过本发明的方法产生的提取物;或本发明的组合物,其中所述RNA分子或小RNA分子对非人生物体具有有害作用。在一个实施方案中,所述非人生物体是节肢动物,例如昆虫;或植物,例如杂草。在一个实施方案中,所述非人类生物体是植物,并且节肢动物摄食该植物或其一部分,由此防治该节肢动物。

[0098] 在一个实施方案中,本发明提供了预防或治疗受试者的疾病的方法,所述方法包括向受试者施用一种或多种本发明的RNA分子或通过RNA分子的加工产生的小RNA分子(长度为20-24nt)或两者;本发明的多核苷酸或载体;本发明的宿主细胞;通过本发明的方法产生的提取物;或本发明的组合物,其中所述RNA分子或小RNA分子对所述疾病的至少一种症状具有有益效果。在一个实施方案中,RNA分子或小RNA分子、多核苷酸、载体或组合物是局部、口服或注射施用的。在一个实施方案中,受试者是脊椎动物。在一个实施方案中,脊椎动物是哺乳动物,如人、家畜如牛或绵羊;或鸟类,如鸡和其他家禽。

[0099] 在另一方面,本发明提供了本发明的RNA分子、本发明的多核苷酸或载体、本发明的宿主细胞、通过本发明的方法产生的提取物,或本发明的组合物用于治疗受试者的疾病的用途,其中所述RNA分子或小RNA分子对所述疾病的至少一种症状具有有益效果。在一个实施方案中,本发明提供了本发明的RNA分子或由其产生的小RNA分子、本发明的多核苷酸或载体、本发明的宿主细胞、通过本发明的方法产生的提取物,或本发明的组合物用于制造预防或治疗受试者的疾病的药物用途,其中所述RNA分子或由其产生的小RNA分子对所述疾病的至少一种症状具有有益效果。

[0100] 在另一方面,本发明提供试剂盒,其包含以下中的一种或多种:本发明的RNA分子或由其产生的小RNA分子、本发明的多核苷酸或载体、本发明的宿主细胞、通过本发明的方法产生的提取物,或本发明的组合物。

[0101] 除非另外具体说明,否则本文中的任何实施例应被理解为比照适用于任何其他实施例。

[0102] 本发明不限于这里描述的特定实施例的范围,这些实施例仅用于示例的目的。如本文所述,功能等同的产品、组合物和方法显然在本发明的范围内。

[0103] 在整个说明书中,除非另有特别说明或上下文另有要求,提及单个步骤、物质组

成、步骤组或物质组成组应被认为包括一个和多个(即一个或多个)那些步骤、物质组成、步骤组或物质组成组。

[0104] 以下通过以下非限制性实施例并参考附图描述本发明。

附图说明

[0105] 图1.两种ledRNA分子的示意图。(A)该ledRNA分子包含有义序列、反义序列和两个环,所述有义序列可被认为是两个相邻的有义序列,共价连接而没有间插间隔区序列,并且与靶RNA具有同一性,所述反义序列与所述有义序列互补并且被分成两个区,5'区和3'区,所述两个环将所述有义序列与所述反义序列分开。(B)该ledRNA分子包含反义序列、有义序列和两个环:所述反义序列可以被认为是两个相邻的反义序列,共价连接而没有间插间隔区序列,并且与靶RNA的互补体具有同一性,所述有义序列与所述反义序列互补并且被分成两个区,所述两个环将所述有义序列与所述反义序列分开。RNA分子的产生通过转录,例如通过从诸如T7或Sp6启动子的启动子进行体外转录,通过互补有义和反义序列之间的碱基配对自退火以形成双链区,所述双链区在每个末端具有环并且在反义或有义序列中具有"裂口"。额外的序列可以与5'和/或3'端连接作为5'或3'延伸序列。

[0106] 图2.与有义/反义退火或发夹RNA相比,ledRNA在形成dsRNA方面更有效。显示了双链RNA分子的三种形式的示意图:A,通过两条分开的链的退火形成的常规dsRNA;B,具有5'或3'延伸序列的发夹RNA;和C,ledRNA分子。下版面显示靶向GUS基因或GFP基因的三种类型的RNA分子的RNA转录物的凝胶电泳后的照片。

[0107] 图3.处理的(A和B)和未处理的远端(C和D)组织的Northern印迹杂交显示ledRNA比dsRNA更稳定并通过烟草叶组织传播。与强的ledRNA信号相反,在远端组织(C和D,上版面)中不能检测到dsRNA信号。

[0108] 图4.ledRNA处理在上述处理区(1)和未处理区(3)均诱导GUS下调。

[0109] 图5.ledRNA诱导本氏烟(N.benthamiana)叶中的FAD2.1基因沉默。

[0110] 图6.Northern印迹杂交证实通过在6和24小时用ledFAD2.1处理,FAD2.1 mRNA强烈下调。

[0111] 图7.将GUS靶基因区的核苷酸序列(SEQ ID NO:14)与hpGUS[G:U]构建体的有义序列(SEQ ID NO:11的核苷酸9至208)比对。52个胞嘧啶(C)核苷酸用胸腺嘧啶(T)核苷酸取代。保留的核苷酸用星号标记,取代的C没有用星号标记。

[0112] 图8.将GUS靶基因区的核苷酸序列(SEQ ID NO:14)与hpGUS[1:4]构建体的有义序列(SEQ ID NO:12的核苷酸9至208)比对。相对于相应的野生型有义序列,hpGUS[1:4]中的每第四个核苷酸被取代,由此对于每第四个核苷酸,C被改变为G,G被改变为C,A被改变为T,并且T被改变为A。保留的核苷酸用星号标记,取代的G和C没有用星号标记,取代的A和T以分号标记。

[0113] 图9.将GUS靶基因区的核苷酸序列(SEQ ID NO:14)与hpGUS[2:10]构建体的有义序列(SEQ ID NO:13的核苷酸9至208)比对。相对于相应的野生型有义序列,hpGUS[2:10]中10个核苷酸的各区段中的每第"9"和"10"个核苷酸被取代,由此对于每第"9"和"10"个核苷酸,C被改变为G,G被改变为C,A被改变为T,并且T被改变为A。保留的核苷酸用星号标记,取代的G和C没有用星号标记,取代的A和T以分号标记。

[0114] 图10.示意图显示编码靶向GUS mRNA的修饰的发夹RNA的遗传构建体的结构。

[0115] 图11.载体pWBPPGH的示意图,载体pWBPPGH用于转化烟草植物,提供GUS靶基因。T-DNA从载体的右边界(RB)延伸至左边界(LB)。T-DNA上的可选择标记基因是编码潮霉素抗性的35S-HPT-tml'基因。

[0116] 图12.用编码用于降低GUS靶基因表达的修饰的发夹RNA的构建体转化的植物中的GUS活性。无hp:对照PPGH11和PPGH24植物,无hpGUS构建体。在括号中给出了与相应的对照PPGH11或PPGH24植物相比显示出小于10％GUS活性的植物的数量以及此类植物相对于测试植物的数量的百分比。

[0117] 图13.(A)所有转基因植物的平均GUS活性:对于hpGUS[wt],59株,对于hpGUS[G:U]74株,对于hpGUS[1:4],33株,对于hpGUS[2:10],41株。(B)所有沉默植物的平均GUS活性(对于hpGUS[wt],32株,对于hpGUS[G:U],71株,对于hpGUS[1:4],33株,对于hpGUS[2:10],28株)。

[0118] 图14.含有hpGUS[wt]、hpGUS[G:U]或hpGUS[1:4]的转基因子代植物的GUS活性。

[0119] 图15.来自用hpGUS[G:U]构建体转化的16种植物的DNA的Southern印迹的放射自显影图。在凝胶电泳前用HindIII消化DNA并用OCS-T探针探测。泳道1:尺寸标记物(HindIII消化的λDNA);泳道2和3,亲代植物PPGHII和PPGH24的DNA;泳道4-19:16种不同转基因植物的DNA。

[0120] 图16.检测来自转基因烟草植物中表达的发夹RNA的有义(上版面)和反义(下版面)sRNA的Northern印迹杂交实验的放射自显影图。泳道1和2含有从缺乏hpGUS构建体的亲代植物PPGHII和PPGH24获得的RNA。泳道3-11含有来自hpGUS[wt]植物的RNA,且泳道12-20含有来自hpGUS[G:U]植物的RNA。

[0121] 图17.检测来自转基因植物的反义sRNA的Northern印迹杂交的放射自显影图。泳道1-10来自hpGUS[wt]植物,泳道11-19来自hpGUS[G:U]植物。反义sRNA具有对应于长度20-24nt的迁移率。用U6RNA的反义RNA作为泳道上样对照对印迹进行再探测。

[0122] 图18.检测来自转基因植物的反义sRNA的重复Northern印迹杂交的放射自显影图。

[0123] 图19.转基因植物中hpGUS构建体中35S启动子和有义GUS区的连接区的DNA甲基化分析。在用McrBC酶处理植物DNA之前,用(+)或不使用(-)对连接片段进行PCR扩增。

[0124] 图20.转基因植物中hpGUS构建体中的35S启动子区的DNA甲基化分析。在用McrBC酶处理植物DNA之前,用(+)或不使用(-)对35S片段进行PCR扩增。

[0125] 图21.经加工的RNA的尺寸分布和丰度。(A)EIN2构建体。(B)GUS构建体。

[0126] 图22.hpEIN2[G:U]构建体的有义序列(上部序列,SEQ ID NO:22的核苷酸17至216)和对应于拟南芥(A.thaliana)EIN2靶基因的cDNA区的核苷酸序列(下部序列,SEQ ID NO:27)的比对。用胸腺嘧啶(T)核苷酸取代野生型序列中的43个胞嘧啶(C)核苷酸,得到有义序列。保留的核苷酸用星号标记,取代的C没有用星号标记。

[0127] 图23.hpCHS[G:U]构建体的有义序列(上部序列,SEQ ID NO:24的核苷酸13至212)和对应于拟南芥(A.thaliana)CHS靶基因的cDNA区的核苷酸序列(SEQ ID NO:28,倒序)的比对。用胸腺嘧啶(T)核苷酸取代野生型序列中的65个胞嘧啶(C)核苷酸,得到有义序列。保留的核苷酸用星号标记,取代的C没有用星号标记。

[0128] 图24.hpEIN2[G:U/U:G]构建体的反义序列(上部序列,SEQ ID NO:25的核苷酸8至207)和对应于拟南芥(A.thaliana)EIN2靶基因和的互补体区的核苷酸序列(下部序列,SEQ ID NO:29)的比对。用胸腺嘧啶(T)核苷酸取代野生型序列中的49个胞嘧啶(C)核苷酸,得到反义序列。保留的核苷酸用星号标记,取代的C没有用星号标记。

[0129] 图25.hpCHS[G:U/U:G]构建体的反义序列(上部序列,SEQ ID NO:26的核苷酸13至212)和对应于拟南芥(A.thaliana)CHS靶基因的互补体区的核苷酸序列(下部序列,SEQ ID NO:30)的比对。用胸腺嘧啶(T)核苷酸取代野生型序列中的49个胞嘧啶(C)核苷酸,得到反义序列。保留的核苷酸用星号标记,取代的C没有用星号标记。

[0130] 图26.乙烯不敏感2(EIN2)和查耳酮合酶(CHS)hpRNA构建体的示意图。35S:CaMV 35S启动子;EIN2和CHS区显示为野生型序列(wt)或G:U修饰的序列(G:U)。箭头表示DNA片段的取向,从右到左的箭头表示反义序列。还显示了限制酶位点。

[0131] 图27.在EIN2测定中转基因拟南芥(A.thaliana)幼苗的下胚轴长度,含有hpEIN2[wt]或hpEIN2[G:U]。

[0132] 图28.对于转基因拟南芥(A.thaliana)中的CHS mRNA,qRT-PCR将hpCHS[wt]或hpCHS[G:U]构建体的转基因标准化至肌动蛋白(Actin)2RNA的水平。Col-0是野生型(非转基因)拟南芥(A.thaliana)。

[0133] 图29.用hpEIN2[wt]或hpEIN2[G:U]转化的植物的RNA的Northern印迹杂交的放射自显影图。上版面示出了品系的下胚轴长度。放射自显影图显示用EIN2有义探针探测以检测反义sRNA的Northern印迹。同样的印迹用U6 RNA探针作为上样对照(U6 RNA)进行再探测。

[0134] 图30.转基因拟南芥(A.thaliana)植物中的基因组DNA中35S启动子和35S-有义EIN2序列的DNA甲基化分析。

[0135] 图31.发夹RNA构建体的启动子和5'区中的DNA甲基化水平。

[0136] 图32.hpEIN2[wt]群体的甲基化最少的品系中的35S启动子仍然显示显著的甲基化。

[0137] 图33.G:U hpEIN2品系中的35S启动子仅显示弱甲基化(P<10%)。

[0138] 图34.72小时时在CHO和Vero细胞中具有G:U基因沉默的ledRNA和hpRNA。

[0139] 图35.48小时时在Hela细胞中测试哑铃型质粒。

[0140] 图36.dsRNA分子的可能修饰的实施例。

[0141] 图37.饲喂补充有ledRNA的人工饲料以下调绿色桃蚜中MpC002或MpRack-1基因的表达后,蚜虫的性能下降。上版面(A):用100μl的50ng/μl的ledRNA在十天周期之后每个成虫蚜虫的平均若虫数量。下版面(B):在喂食100μl含有MpC002、MpRack-1或对照ledGFP的200ng/μl ledRNA后,在五天的时间过程中存活的蚜虫的百分比。

[0142] 序列表的关键字

[0143] SEQ ID NO:1-GFP ledRNA的核糖核苷酸序列。

[0144] SEQ ID NO:2-GUS ledRNA的核糖核苷酸序列。

[0145] SEQ ID NO:3-本氏烟(N.benthamiana)FAD2.1 ledRNA的核糖核苷酸序列。

[0146] SEQ ID NO:4-编码GFP ledRNA的核苷酸序列。

[0147] SEQ ID NO:5-编码GUS ledRNA的核苷酸序列。

[0148] SEQ ID NO:6-编码本氏烟(N.benthamiana)FAD2.1 ledRNA的核苷酸序列。

[0149] SEQ ID NO:7-编码GFP的核苷酸序列。

[0150] SEQ ID NO:8-编码GUS的核苷酸序列。

[0151] SEQ ID NO:9-编码本氏烟(N.benthamiana)FAD2.1的核苷酸序列。

[0152] SEQ ID NO:10-用于为编码靶向GUS mRNA的发夹RNA分子的构建体提供GUS有义区的核苷酸序列。

[0153] SEQ ID NO:11-用于为编码发夹RNA分子hpGUS[G:U]的构建体提供GUS有义区的核苷酸序列。

[0154] SEQ ID NO:12-用于为编码发夹RNA分子hpGUS[1:4]的构建体提供GUS有义区的核苷酸序列。

[0155] SEQ ID NO:13-用于为编码发夹RNA分子hpGUS[2:10]的构建体提供GUS有义区的核苷酸序列。

[0156] SEQ ID NO:14-GUS基因的蛋白编码区的核苷酸781-1020的核苷酸序列。

[0157] SEQ ID NO:15-hpGUS[wt]RNA的发夹结构(包括其环)的核糖核苷酸序列。

[0158] SEQ ID NO:16-hpGUS[G:U]RNA的发夹结构(包括其环)的核糖核苷酸。

[0159] SEQ ID NO:17-hpGUS[1:4]RNA的发夹结构(包括其环)的核糖核苷酸。

[0160] SEQ ID NO:18-hpGUS[2:10]RNA的发夹结构(包括其环)的核糖核苷酸。

[0161] SEQ ID NO:19-对应于拟南芥(A.thaliana)EIN2基因,登录号NM_120406的cDNA的核苷酸序列。

[0162] SEQ ID NO:20-对应于拟南芥(A.thaliana)CHS基因,登录号NM_121396,1703nt的cDNA的核苷酸序列。

[0163] SEQ ID NO:21-包含来自对应于侧接限制酶位点的拟南芥(A.thaliana)EIN2基因的cDNA的200nt有义序列的DNA片段的核苷酸序列。

[0164] SEQ ID NO:22-包含EIN2的200nt有义序列的DNA片段的核苷酸序列,与SEQ ID NO:21相同,只是用T'代替了43个C',用于构建hpEIN2[G:U]。

[0165] SEQ ID NO:23-包含来自对应于侧接限制酶位点的拟南芥(A.thaliana)CHS的cDNA的200nt有义序列的DNA片段的核苷酸序列。

[0166] SEQ ID NO:24-包含CHS的200nt有义序列的DNA片段的核苷酸序列,与SEQ ID NO:23相同,只是用T'代替了65个C',用于构建hpCHS[G:U]。

[0167] SEQ ID NO:25-包含EIN2的200nt反义序列的DNA片段的核苷酸序列,只是用50个C'代替了T',用于构建hpEIN2[G:U/U:G]。

[0168] SEQ ID NO:26-包含CHS的200nt反义序列的DNA片段的核苷酸序列,只是用49个C'代替了T',用于构建hpCHS[G:U/U:G]。

[0169] SEQ ID NO:27-对应于拟南芥(A.thaliana)EIN2基因(登录号NM_120406)的cDNA的核苷酸601-900的核苷酸序列。

[0170] SEQ ID NO:28-对应于拟南芥(A.thaliana)CHS基因(登录号NM_121396)的cDNA的核苷酸813-1112的核苷酸序列。

[0171] SEQ ID NO:29-对应于拟南芥(A.thaliana)EIN2基因(登录号NM_120406)的cDNA的核苷酸652-891的互补序列的核苷酸序列。

[0172] SEQ ID NO:30-对应于拟南芥(A.thaliana)CHS基因的cDNA的核苷酸804-1103的互补序列的核苷酸序列。

[0173] SEQ ID NO:31-拟南芥(Arabidopsis thaliana),登录号NM_001333162的cDNA的FANCM I蛋白编码区。靶区核苷酸675-1174(500个核苷酸)

[0174] SEQ ID NO:32-欧洲油菜(Brassica napus)的cDNA的FANCM I蛋白编码区。靶区核苷酸896-1395(500bp)

[0175] SEQ ID NO:33-编码hpFANCM-At[wt]的核苷酸序列,其靶向拟南芥(A.thaliana)的FANCM I蛋白编码区。FANCM有义序列,核苷酸38-537;环序列,核苷酸538-1306;FANCM反义序列,核苷酸1307-1806。

[0176] SEQ ID NO:34-编码hpFANCM-At[G:U]的核苷酸序列,其靶向拟南芥(A.thaliana)的FANCM I蛋白编码区。FANCM有义序列,核苷酸38-537;环序列,核苷酸538-1306;FANCM反义序列,核苷酸1307-1806。

[0177] SEQ ID NO:35-编码hpFANCM-Bn[wt]的核苷酸序列,其靶向欧洲油菜(B.napus)的FANCM I蛋白编码区。FANCM有义序列,核苷酸34-533;环序列,核苷酸534-1300;FANCM反义序列,核苷酸1301-1800。

[0178] SEQ ID NO:36-编码hpFANCM-Bn[G:U]的核苷酸序列,其靶向欧洲油菜(B.napus)的FANCM I蛋白编码区。FANCM有义序列,核苷酸34-533;环序列,核苷酸534-1300;FANCM反义序列,核苷酸1301-1800。

[0179] SEQ ID NO:37-对应于欧洲油菜(B.napus)DDM1基因;登录号XR_001278527的cDNA的蛋白编码区的核苷酸序列。

[0180] SEQ ID NO:38-编码hpDDM1-Bn[wt]的DNA的核苷酸序列,其靶向欧洲油菜(B.napus)的DDM1蛋白编码区。

[0181] SEQ ID NO:39-编码hpDDM1-Bn[G:U]的核苷酸序列,其靶向欧洲油菜(B.napus)的DDM1蛋白编码区。DDM1有义序列,核苷酸35-536;环序列,核苷酸537-1304;DDM1反义序列,核苷酸1305-1805。

[0182] SEQ ID NO:40-EGFPcDNA。

[0183] SEQ ID NO:41-hpEGFP[wt]的编码区的核苷酸序列,相对于启动子顺序为反义/环/有义。

[0184] SEQ ID NO:42-hpEGFP[G:U]的编码区的核苷酸序列,其在EGFP有义序列中具有157个C至T取代。

[0185] SEQ ID NO:43-ledEGFP[wt]的编码区的核苷酸序列,其在EGFP有义序列中没有C至T取代。

[0186] SEQ ID NO:44-ledEGFP[G:U]的编码区的核苷酸序列,其在EGFP有义序列中具有162个C至T取代。

[0187] SEQ ID NO:45-用于为编码不侧接限制酶位点的发夹RNA分子hpGUS[G:U]的构建体提供GUS有义区的核苷酸序列。

[0188] SEQ ID NO:46-用于为编码不侧接限制酶位点的发夹RNA分子hpGUS[1:4]的构建体提供GUS有义区的核苷酸序列。

[0189] SEQ ID NO:47-用于为编码不侧接限制酶位点的发夹RNA分子hpGUS[2:10]的构建

体提供GUS有义区的核苷酸序列。

[0190] SEQ ID NO:48-包含EIN2的200nt有义序列的DNA片段的核苷酸序列,与SEQ ID NO:21相同,只是用T'代替了43个C',用于构建没有侧接序列的hpEIN2[G:U]。

[0191] SEQ ID NO:49-包含CHS的200nt有义序列的DNA片段的核苷酸序列,与SEQ ID NO:23相同,只是用T'代替了65个C',用于构建没有侧接序列的hpCHS[G:U]。

[0192] SEQ ID NO:50-包含EIN2的200nt反义序列的DNA片段的核苷酸序列,只是用50个C'代替了T',用于构建没有侧接序列的hpEIN2[G:U/U:G]。

[0193] SEQ ID NO:51-包含CHS的200nt反义序列的DNA片段的核苷酸序列,只是用49个C'代替了T',用于构建没有侧接序列的hpCHS[G:U/U:G]序列。

[0194] SEQ ID No:52-用于扩增200bp GUS有义序列的寡核苷酸引物(GUS-WT-F)

[0195] SEQ ID No:53-用于扩增200bp GUS有义序列的寡核苷酸引物(GUS-WT-R)

[0196] SEQ ID NO:54-用于生产hpGUS[G:U]片段的寡核苷酸引物(正向),其中每个C被T取代(GUS-GU-F)

[0197] SEQ ID NO:55-用于生产hpGUS[G:U]片段的寡核苷酸引物(反向),其中每个C被T取代(GUS-GU-R)

[0198] SEQ ID NO:56-用于生产hpGUS[1:4]片段的寡核苷酸引物(正向),其中每第4个核苷酸被取代(GUS-4M-F)

[0199] SEQ ID NO:57-用于生产hpGUS[1:4]片段的寡核苷酸引物(反向),其中每第4个核苷酸被取代(GUS-4M-R)

[0200] SEQ ID NO:58-用于生产hpGUS[2:10]片段的寡核苷酸引物(正向),其中每第9和10位核苷酸被取代(GUS-10M-F)

[0201] SEQ ID NO:59-用于生产hpGUS[2:10]片段的寡核苷酸引物(反向),其中每第9和10位核苷酸被取代(GUS-10M-R)

[0202] SEQ ID No:60-编码正向引物(35S-F3)的核苷酸序列

[0203] SEQ ID NO:61-编码反向引物(GUSwt-R2)的核苷酸序列

[0204] SEQ ID NO:62-编码正向引物(GUSgu-R2)的核苷酸序列

[0205] SEQ ID NO:63-编码反向引物(GUS4m-R2)的核苷酸序列

[0206] SEQ ID No:64-编码正向引物(35S-F2)的核苷酸序列

[0207] SEQ ID NO:65-编码反向引物(35S-R1)的核苷酸序列

[0208] SEQ ID NO:66-用于扩增野生型200bp EIN2有义序列的寡核苷酸引物(EIN2wt-F)

[0209] SEQ ID NO:67-用于扩增野生型200bp EIN2有义序列的寡核苷酸引物(EIN2wt-R)

[0210] SEQ ID No:68-用于扩增野生型200bp CHS有义序列的寡核苷酸引物(CHSwt-F)

[0211] SEQ ID No:69-用于扩增野生型200bp CHS有义序列的寡核苷酸引物(CHSwt-R)

[0212] SEQ ID NO:70-用于生产hpEIN2[G:U]片段的寡核苷酸引物(正向),其中每个C被T取代(EIN2gu-F)

[0213] SEQ ID NO:71-用于生产hpEIN2[G:U]片段的寡核苷酸引物(反向),其中每个C被T取代(EIN2gu-R)

[0214] SEQ ID NO:72-用于生产hpCHS[G:U]片段的寡核苷酸引物(正向),其中每个C被T取代(CHSgu-F)

[0215] SEQ ID NO:73-用于生产hpCHS[G:U]片段的寡核苷酸引物(反向),其中每个C被T取代(CHSgu-R)

[0216] SEQ ID NO:74-用于产生hpEIN2[G:U/U:G]片段的寡核苷酸引物(正向),其中每个C被T取代(asEIN2gu-F)

[0217] SEQ ID NO:75-用于生产hpEIN2[G:U/U:G]片段的寡核苷酸引物(反向),其中每个C被T取代(asEIN2gu-R)

[0218] SEQ ID NO:76-用于生产hpCHS[G:U/U:G]片段的寡核苷酸引物(正向),其中每个C被T取代(asCHSgu-F)

[0219] SEQ ID NO:77-用于生产hpCHS[G:U/U:G]片段的寡核苷酸引物(反向),其中每个C被T取代(asCHSgu-R)

[0220] SEQ ID No:78-编码正向引物(CHS-200-F2)的核苷酸序列

[0221] SEQ ID NO:79-编码反向引物(CHS-200-R2)的核苷酸序列

[0222] SEQ ID No:80-编码正向引物(Actin2-For)的核苷酸序列

[0223] SEQ ID NO:81-编码反向引物(Actin2-Rev)的核苷酸序列

[0224] SEQ ID NO:82-编码正向引物(Top-35S-F2)的核苷酸序列

[0225] SEQ ID NO:83-编码反向引物(Top-35S-R2)的核苷酸序列

[0226] SEQ ID NO:84-编码正向引物(Link-35S-F2)的核苷酸序列

[0227] SEQ ID NO:85-编码反向引物(Link-EIN2-R2)的核苷酸序列

[0228] SEQ ID NO:86-有义si22-核糖核苷酸序列

[0229] SEQ ID NO:87-反义si22-核糖核苷酸序列

[0230] SEQ ID NO:88-正向引物的核糖核苷酸序列

[0231] SEQ ID NO:89-反向引物的核糖核苷酸序列

[0232] SEQ ID NO:90-正向引物的核糖核苷酸序列

[0233] SEQ ID NO:91-反向引物的核糖核苷酸序列

[0234] SEQ ID NO:92-dsRNA分子的可能修饰

[0235] SEQ ID NO:93-对应于欧洲油菜(B.napus)DDM1基因(登录号XR_001278527)的cDNA的核苷酸序列。

[0236] SEQ ID NO:94-编码靶向欧洲油菜(B.napus)DDM1基因的发夹RNAi(hpRNA)构建体的嵌合DNA的核苷酸序列。

[0237] SEQ ID NO:95-编码靶向欧洲油菜(B.napus)DDM1基因、具有G:U碱基对的发夹RNAi(hpRNA)构建体的嵌合DNA的核苷酸序列。

[0238] SEQ ID NO:96-编码靶向欧洲油菜(B.napus)DDM1基因的ledRNA构建体的嵌合DNA的核苷酸序列。

[0239] SEQ ID NO:97-对应于拟南芥(A.thaliana)FANCM基因(登录号NM_001333162)的cDNA的核苷酸序列。

[0240] SEQ ID NO:98-编码靶向拟南芥(A.thaliana)FANCM基因的发夹RNAi(hpRNA)构建体的嵌合DNA的核苷酸序列。

[0241] SEQ ID NO:99-编码靶向拟南芥(A.thaliana)FANCM基因、具有G:U碱基对的发夹RNAi(hpRNA)构建体的嵌合DNA的核苷酸序列。

[0242] SEQ ID NO:100-编码靶向拟南芥(A.thaliana)FANCM基因的ledRNA构建体的嵌合DNA的核苷酸序列。

[0243] SEQ ID NO:101-对应于欧洲油菜(B.napus)FANCM基因(登录号XM_022719486.1)的cDNA的核苷酸序列。

[0244] SEQ ID NO:102-编码靶向欧洲油菜(B.napus)FANCM基因的发夹RNAi(hpRNA)构建体的嵌合DNA的核苷酸序列。

[0245] SEQ ID NO:103-编码靶向欧洲油菜(B.napus)FANCM基因、具有G:U碱基对的发夹RNAi(hpRNA)构建体的嵌合DNA的核苷酸序列。

[0246] SEQ ID NO:104-编码靶向欧洲油菜(B.napus)FANCM基因的ledRNA构建体的嵌合DNA的核苷酸序列。

[0247] SEQ ID NO:105-对应于本氏烟(Nicotiana benthamiana)TOR基因的cDNA的蛋白质编码区的核苷酸序列。

[0248] SEQ ID NO:106-编码靶向本氏烟(Nicotiana benthamiana)TOR基因的ledRNA构建体的嵌合DNA的核苷酸序列。

[0249] SEQ ID NO:107:对应于大麦(Hordeum vulgare)(登录号LT601589)乙酰乳酸合酶(ALS)基因的cDNA的蛋白编码区的核苷酸序列。

[0250] SEQ ID NO:108-编码靶向大麦(H.vulgare)ALS基因的ledRNA构建体的嵌合DNA的核苷酸序列。

[0251] SEQ ID NO:109:对应于大麦(Hordeum vulgare)HvNCED1基因(登录号AK361999)的cDNA的蛋白编码区的核苷酸序列。

[0252] SEQ ID NO:110-对应于对应于大麦(Hordeum vulgare)HvNCEDD2基因(登录号DQ145931)的cDNA的蛋白质编码区的核苷酸序列。

[0253] SEQ ID NO:111-编码靶向大麦(Hordeum vulgare)和小麦(Triticum aestivum)的NCED1基因的ledRNA构建体的嵌合DNA的核苷酸序列。

[0254] SEQ ID NO:112-编码靶向大麦(Hordeum vulgare)和小麦(Triticum aestivum)的NCED2基因的ledRNA构建体的嵌合DNA的核苷酸序列。

[0255] SEQ ID NO:113:对应于编码ABA-OH-2的大麦基因(登录号DQ145933)的cDNA的蛋白质编码区的核苷酸序列。

[0256] SEQ ID NO:114-编码靶向大麦(Hordeum vulgare)和小麦(Triticum aestivum)的ABA-OH-2基因的ledRNA构建体的嵌合DNA的核苷酸序列。

[0257] SEQ ID NO:115-对应于应于编码EIN2的拟南芥(A.thaliana)基因(At5g03280)的cDNA蛋白质编码区的核苷酸序列。

[0258] SEQ ID NO:116-编码靶向拟南芥(A.thaliana)EIN2基因的ledRNA构建体的嵌合DNA的核苷酸序列。

[0259] SEQ ID NO:117:对应于编码CHS的拟南芥(A.thaliana)基因(登录号NM_121396)的cDNA的蛋白质编码区的核苷酸序列。

[0260] SEQ ID NO:118-编码靶向拟南芥(A.thaliana)CHS基因的ledRNA构建体的嵌合DNA的核苷酸序列。

[0261] SEQ ID NO:119:对应于编码狭叶羽扇豆(L.angustifolius)N样基因(登录号XM_

019604347)的cDNA的蛋白质编码区的核苷酸序列。

[0262] SEQ ID NO:120-编码靶向狭叶羽扇豆(L.angustifolius)N样基因的ledRNA构建体的嵌合DNA的核苷酸序列。

[0263] SEQ ID NO:121-对应于华东葡萄(Vitis pseudoreticulata)MLO基因(登录号KR362912)的cDNA的蛋白编码区的核苷酸序列。

[0264] SEQ ID NO:122-编码靶向华东葡萄(Vitis)MLO基因的第一ledRNA构建体的嵌合DNA的核苷酸序列。

[0265] SEQ ID NO:123-对应于桃蚜(Myzus persicae)MpCOO2基因的cDNA的蛋白编码区的核苷酸序列。

[0266] SEQ ID NO:124-对应于桃蚜(Myzus persicae)的MpRack-1基因的cDNA的蛋白编码区的核苷酸序列。

[0267] SEQ ID NO:125-编码靶向桃蚜(M.persicae)COO2基因的ledRNA的嵌合构建体的核苷酸序列。

[0268] SEQ ID NO:126-编码靶向桃蚜(M.persicae)Rack-1基因的ledRNA的嵌合构建体的核苷酸序列。

[0269] SEQ ID NO:127-对应于棉铃虫(Helicoverpa armigera)ABCwhite基因的cDNA的核苷酸序列。

[0270] SEQ ID NO:128-编码靶向棉铃虫(Helicoverpa armigera)的ABC转运蛋白白基因的ledRNA构建体的嵌合DNA的核苷酸序列。

[0271] SEQ ID NO:129-对应于阿根廷蚁(Linepithema humile)PBAN型神经肽样(登录号XM_012368710)的cDNA的核苷酸序列。

[0272] SEQ ID NO:130:编码靶向阿根廷蚁中PBAN基因(登录号XM_012368710)的ledRNA构建体的嵌合DNA的核苷酸序列。

[0273] SEQ ID NO:131:编码ledRNA构建体的嵌合DNA的核苷酸序列,所述ledRNA构建体靶向编码铜绿蝇(L.cuprina)的V型质子ATP酶催化亚基A的基因(登录号XM_023443547)。

[0274] SEQ ID NO:132-编码ledRNA构建体的嵌合DNA的核苷酸序列,所述ledRNA构建体靶向编码铜绿蝇(L.cuprina)的RNAse 1/2的基因。

[0275] SEQ ID NO:133-编码ledRNA构建体的嵌合DNA的核苷酸序列,所述ledRNA构建体靶向编码铜绿蝇(L.cuprina)的几丁质合酶的基因。

[0276] SEQ ID NO:134-编码ledRNA构建体的嵌合DNA的核苷酸序列,所述ledRNA构建体靶向编码铜绿蝇(L.cuprina)的蜕皮激素受体(EcR)的基因。

[0277] SEQ ID NO:135-编码ledRNA构建体的嵌合DNA的核苷酸序列,所述ledRNA构建体靶向编码铜绿蝇(L.cuprina)的γ-微管蛋白1/1的基因。

[0278] SEQ ID NO:136-TaMlo靶基因(AF384144)。

[0279] SEQ ID NO:137-编码ledRNA构建体的嵌合DNA的核苷酸序列,所述ledRNA构建体靶向编码TaMlo的基因。

[0280] SEQ ID NO:138-对应于华东葡萄(Vitis pseudoreticulata)MLO基因(登录号KR362912)的cDNA的蛋白编码区的核苷酸序列。

[0281] SEQ ID NO:139-编码靶向华东葡萄(Vitis)MLO基因的第一ledRNA构建体的嵌合

DNA的核苷酸序列。

具体实施方式

[0282] 一般技术和定义

[0283] 除非另外具体定义,否则本文使用的所有技术和科学术语应被认为具有与本领域普通技术人员通常理解的相同的含义(例如,细胞培养、分子遗传学、基因沉默、蛋白质化学和生物化学)。

[0284] 除非另有说明,用于本发明的重组蛋白、细胞培养物和免疫学技术是本领域技术人员熟知的标准方法。在以下来源的整个文献中描述了这种技术,诸如:J.Perbal,A Practical Guide to Molecular Cloning,John Wiley and Sons(1984),J.Sambrook et al.,Molecular Cloning:A Laboratory Manual,Cold Spring Harbour Laboratory Press(1989),T.A.Brown(editor),Essential Molecular Biology:A Practical Approach,Volumes 1and 2,IRL Press(1991),D.M.Glover and B.D.Hames(editors),DNA Cloning:A Practical Approach,Volumes 1-4,IRL Press(1995and 1996),and F.M.Ausubel et al.(editors),Current Protocols in Molecular Biology,Greene Pub.Associates and Wiley-Interscience(1988,including all updates until present),Ed Harlow and David Lane(editors)Antibodies:A Laboratory Manual,Cold Spring Harbour Laboratory,(1988),和J.E.Coligan et al.(editors)Current Protocols in Immunology,John Wiley&Sons(including all updates until present).

[0285] 本文所用的术语“反义调节元件”或“反义核糖核酸序列”或“反义RNA序列”是指与它所杂交的靶RNA分子的至少一部分至少部分互补的RNA序列。在某些实施方案中,例如通过减少靶RNA分子的翻译,反义RNA序列调节(增加或减少)靶RNA分子的表达或量或其活性。在某些实施方案中,反义RNA序列改变靶前体mRNA的剪接,导致不同的剪接变体。反义序列的示例性组分包括但不限于寡核苷酸、寡核苷、寡核苷酸类似物、寡核苷酸模拟物和这些的嵌合组合。

[0286] 术语“反义活性”在本发明的上下文中用于指归因于反义RNA序列与其靶RNA分子杂交的任何可检测和/或可测量的活性。这种检测和/或测量可以是直接的或间接的。在一个实施例中,通过检测和/或测量靶RNA分子转录物的量来评估反义活性。反义活性也可以检测为与靶RNA分子相关的表型变化。如本文所用,术语“靶RNA分子”是指由根据本发明的反义RNA序列调节的基因转录物。因此,“靶RNA分子”可以是其表达或活性能够通过反义RNA序列调节的任何RNA分子。示例性靶RNA分子包括但不限于从编码靶蛋白的DNA转录的RNA(包括但不限于前mRNA和mRNA或其部分)、rRNA、tRNA、小核RNA和miRNA,包括它们的前体形式。靶RNA可以是诸如病毒的病原体或害虫的基因组RNA,或由其衍生的RNA分子,诸如病毒病原体的复制形式,或其转录物。例如,靶RNA分子可以是来自内源基因(或从该基因转录的mRNA)的RNA,或被引入或可以被引入其表达与特定表型、性状、病症或疾病状态相关的真核细胞的基因,或来自感染性因子的核酸分子。在一个实施例中,靶RNA分子在真核细胞中。在另一个实施例中,靶RNA分子编码蛋白质。在本文中,可以通过检测和/或测量靶蛋白的量,例如通过其活性如酶活性,或不同于酶的功能,或通过与其功能相关的表型,来评估反义活性。如本文所用,术语“靶蛋白”是指由根据本发明的反义RNA序列调节的蛋白质。

[0287] 在某些实施方案中,通过检测和/或测量靶RNA分子和/或切割的靶RNA分子和/或选择性剪接的靶RNA分子的量来评估反义活性。

[0288] 可以使用各种方法检测或测量反义活性。例如,可以通过比较特定样品中的活性并将该活性与对照样品的活性进行比较来检测或评估反义活性。

[0289] 术语“靶向”在本发明的上下文中用于指反义RNA序列与特定靶RNA分子或靶RNA分子内的特定核苷酸区的缔合。在一个实施例中,根据本发明的反义RNA序列与靶RNA分子的至少一个区共享互补性。在本文中,术语“互补性”是指能够通过核糖核苷酸上碱基之间的氢键与靶RNA分子上的核糖核苷酸序列碱基配对的核糖核苷酸序列。例如,在RNA中,腺嘌呤(A)与尿嘧啶(U)互补,鸟嘌呤(G)与胞嘧啶(C)互补。

[0290] 在某些实施方案中,“互补碱基”是指能够与本发明的RNA分子或其靶RNA分子中的有义RNA序列的核糖核苷酸碱基配对的反义RNA序列的核糖核苷酸。例如,如果反义RNA序列的某个位置的核糖核苷酸能够与靶RNA分子的某个位置的核糖核苷酸氢键键合,则认为反义RNA序列与靶RNA分子之间的氢键键合的位置在该核糖核苷酸上是互补的。相反,术语“非互补”是指彼此不形成氢键或以其他方式支持杂交的一对核糖核苷酸。术语“互补的”也可用于指反义RNA序列通过互补性与另一核酸杂交的能力。在某些实施方案中,当每个分子中足够数目的相应位置被核糖核苷酸占据时,RNA序列和其靶是彼此互补的,所述核糖核苷酸可以彼此结合以允许本发明的RNA分子和/或靶RNA分子中的反义RNA序列和有义RNA序列之间的稳定缔合。本领域技术人员认识到可以包括错配而不消除反义RNA序列和靶保持缔合的能力。因此,本文描述了反义RNA序列,其可以包含至多约20%的错配的核苷酸(即,与靶序列的相应核苷酸不互补)。优选所述反义化合物含有不超过约15%、更优选不超过约10%、最优选不超过5%或没有错配。剩余的核糖核苷酸互补或不破坏反义RNA序列与有义RNA序列或靶RNA分子之间的杂交(例如,G:U或A:G对)。本领域普通技术人员将认识到本文所述的反义RNA序列与靶RNA分子的至少一个区至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%(完全)互补。

[0291] 如本文所用,“嵌合RNA分子”是指天然不存在的任何RNA分子。在一个实施例中,本文公开的嵌合RNA分子已经被修饰以在dsRNA的区中产生错配。例如,可以修饰嵌合RNA分子以将胞嘧啶转化为尿嘧啶。在一个实施例中,已经通过用亚硫酸氢盐在足以将未甲基化的胞嘧啶转化成尿嘧啶的时间和条件下处理来修饰嵌合RNA分子。

[0292] 本领域技术人员将理解各种核糖核苷酸组合可以碱基对。本发明考虑了规范和非规范碱基对。在一个实施例中,碱基配对可以在DNA分子中包含A:T或G:C或在RNA分子中包含U:A或G:C。在另一个实施例中,碱基配对可以包含A:G或G:T或U:G。

[0293] 本发明中使用的术语“规范碱基配对”意指两个核苷酸之间的碱基配对,对于脱氧核糖核苷酸为A:T或G:C,或者对于核糖核苷酸为A:U或G:C。

[0294] 如在本发明中使用的术语“非规范碱基配对”意指在两个DNA或两个RNA序列的上下文中除了规范碱基配对之外的两个核苷酸的碱基之间的相互作用。例如,非规范碱基配对包括G与U(G:U)之间或A与G(A:G)之间的配对。非规范碱基配对的实施例包括嘌呤-嘌呤或嘧啶-嘧啶。在本发明的上下文中,最常见的非规范碱基配对是G:U。非规范碱基对的其他实施例较不优选是A:C、G:T、G:G和A:A。

[0295] 本发明涉及跨越一系列核糖核苷酸“杂交”的RNA组分。本领域技术人员将理解术

语例如“杂交”和“杂交化”用于描述基于互补核酸序列退火的分子。这样的分子不需要100％互补以便杂交(即,它们不需要“完全碱基对”。例如,序列互补性可能存在一个或多个错配。在一个实施例中,本文定义的RNA组分在严格杂交条件下杂交。术语“严格杂交条件”是指本领域所熟悉的参数,包括杂交温度随RNA分子长度的变化。核糖核苷酸杂交参数可以在编辑这些方法的参考文献中找到,Sambrook等(见上文),和Ausubel等(见上文)。例如,本文所用的严格杂交条件可以是指在65℃下在杂交缓冲液(3.5×SSC,0.02％聚蔗糖,0.02％聚乙烯吡咯烷酮,0.02％牛血清白蛋白(BSA),2.5mM NaH_2PO_4(pH7),0.5％SDS,2mM EDTA)中的杂交,然后在50℃下在0.2×SSC,0.01％BSA中洗涤一次或多次。较短的RNA组分,如长度为20-24个核苷酸的RNA序列在较低严格条件下杂交。术语“低严格性杂交条件”是指本领域所熟悉的参数,包括杂交温度随RNA分子长度的变化。例如,本文所用的低严格杂交条件可以是指在42℃下在杂交缓冲液(3.5×SSC,0.02％聚蔗糖,0.02％聚乙烯吡咯烷酮,0.02％牛血清白蛋白(BSA),2.5mM NaH_2PO_4(pH7),0.5％SDS,2mM EDTA)中的杂交,然后在30℃下在0.2×SSC,0.01％BSA中洗涤一次或多次。

[0296] 本发明还涵盖跨越邻接核糖核苷酸的“完全碱基对”的RNA组分。术语“完全碱基对”在本发明的上下文中用于指一系列邻接的核糖核苷酸碱基对。邻接核糖核苷酸的完全碱基配对系列在该系列内不包含缺口或非碱基配对核苷酸。术语“邻接”用于指一系列核糖核苷酸。包含邻接系列的核糖核苷酸将通过连续系列的磷酸二酯键连接,每个核糖核苷酸直接与下一个键合。

[0297] 本发明的RNA分子包含有义序列和相应的反义序列。在此定义了这些序列之间的关系。本文还定义了反义序列与靶RNA分子的序列关系和活性。

[0298] 术语“共价连接”在本发明的上下文中用于指第一和第二RNA组分或任何RNA序列或核糖核苷酸之间的连接。如本领域技术人员将理解的,共价连接或键是涉及原子之间共享电子对的化学键。在一个实施例中,该第一和第二RNA组分或该有义RNA序列和该反义RNA序列是作为单个RNA链的一部分共价连接的,该单一RNA链可以通过自身互补性折回到其自身上。在该实施例中,所述组分通过磷酸二酯键跨越一个或多个核糖核苷酸共价连接。

[0299] 在本发明的上下文中,术语“杂交”意指通过互补碱基的碱基配对来对互补多核苷酸进行配对。虽然不限于特定机制,但最常见的配对机制涉及互补核糖核苷酸之间的氢键,其可以是沃森-克里克氢键。

[0300] 如本文所用,短语“RNA分子对非人生物体具有有害作用”或类似短语意指分子的靶RNA分子存在于非人生物体中,并且当与缺少RNA分子的相同细胞相比时,将表达靶RNA分子的细胞暴露于靶RNA分子导致靶RNA分子的水平和/或活性降低。在一个实施方案中,靶RNA分子编码对生长、繁殖或存活重要的蛋白质。作为一个实施例,如果该非人类生物体是作物害虫或病原体,或动物的害虫或病原体,则该RNA分子可以对该害虫或病原体的摄食、细胞凋亡、细胞分化和发育、性生殖的能力或欲望、肌肉形成、肌肉抽搐、肌肉收缩、保幼激素形成、保幼激素调节、离子调节和转运、细胞膜电位的维持、氨基酸生物合成、氨基酸降解、精子形成、信息素合成、信息素感测、触角形成、翅膀形成、腿形成、卵形成、幼虫成熟、消化酶形成、血淋巴合成、血淋巴维持、神经传递、幼虫期转变、化蛹、从化蛹出现、细胞分裂、能量代谢、呼吸、几丁质代谢、细胞骨架结构的形成有有害作用。在另一个实施例中,所述非人类生物体是杂草,并且所述RNA分子对氨基酸生物合成、光合作用、脂肪酸合成、细胞膜完

整性、色素合成或生长具有有害作用。

[0301] 如本文所用,短语“RNA分子对疾病的至少一种症状具有有益作用”或类似短语意指分子的靶RNA存在于受试者中,并且当与缺少RNA分子的相同细胞相比时,表达靶RNA的细胞暴露于RNA分子导致靶RNA的水平和/或活性降低。在一个实施方案中,靶RNA编码在疾病存在下发挥作用的蛋白质。在一个实施方案中,所述疾病为癌症或癌性疾病、感染性疾病、心血管疾病、神经疾病、朊病毒病、炎性疾病、自身免疫性疾病、肺病、肾病、肝病、线粒体疾病、内分泌疾病、生殖相关疾病和病症,以及可响应于细胞或生物体中表达的基因产物的水平的任何其他适应症。

[0302] 可以将根据本发明的RNA分子以及包含该RNA分子的组合物施用至受试者。术语例如“受试者”、“患者”或“个体”是在上下文中可以在本发明中互换使用的术语。在一个实施例中,受试者是哺乳动物。该哺乳动物可以是宠物,如狗或猫,或家畜,如马或母牛。在一个实施例中,受试者是人。例如,受试者可以是成人。在另一个实施例中,受试者可以是孩子。在另一个实施例中,受试者可以是青少年。在另一个实施例中,可以将根据本发明的RNA分子以及包含它们的组合物施用至昆虫。在另一个实施例中,可将根据本发明的RNA分子和包含所述RNA分子的组合物施用至植物。在另一个实施例中,可以将根据本发明的RNA分子以及包含它们的组合物施用至真菌细胞或群体。

[0303] 术语“和/或”,例如“X和/或Y”应理解为意指“X和Y”或“X或Y”,并且应理解为提供对两种含义或两种含义的明确支持。

[0304] 如本文所用,除非另有相反说明,术语“约”是指指定值的+/-20%、更优选+/-10%。

[0305] 在整个说明书中,词语“包含(comprise)”或诸如“包含(comprises)”或“包含(comprising)”的变化将被理解为暗示包含所述元件、整体或步骤,或元件、整体或步骤的组,但不排除任何其他元件、整体或步骤,或元件、整体或步骤的组。

[0306] RNA分子

[0307] 在某些实施方案中,本发明的RNA分子包含与第二RNA组分共价连接的第一RNA组分。在优选的实施方案中,RNA分子自身杂交或折叠形成“哑铃”或ledRNA结构,例如参见图1。在一个实施方案中,所述分子还包含以下中的一种或多种:

[0308] -连接核糖核苷酸序列,其共价连接第一和第二RNA组分;

[0309] -5’前导序列;和,

[0310] -3’尾随序列。

[0311] 在一个实施方案中,所述第一RNA组分按5’至3’顺序由第一5’核糖核苷酸、第一RNA序列和第一3’核糖核苷酸组成,其中所述RNA分子中所述第一5’和3’核糖核苷酸彼此碱基配对,其中所述第一RNA序列包含至少20个邻接核糖核苷酸的第一有义核糖核苷酸序列、至少4个核糖核苷酸的第一环序列和至少20个邻接核糖核苷酸的第一反义核糖核苷酸序列,其中第一反义核糖核苷酸序列与RNA分子中的第一有义核糖核苷酸序列杂交,其中第一反义核糖核苷酸序列能够与靶RNA分子的第一区杂交。

[0312] 在另一个实施方案中,所述第一RNA组分按5’至3’顺序由第一5’核糖核苷酸、第一RNA序列和第一3’核糖核苷酸组成,其中所述RNA分子中所述第一5’和3’核糖核苷酸彼此碱基配对,其中所述第一RNA序列包含至少20个邻接核糖核苷酸的第一有义核糖核苷酸序列、

至少4个核糖核苷酸的第一环序列和至少20个邻接核糖核苷酸的第一反义核糖核苷酸序列,其中所述第一反义核糖核苷酸序列与所述RNA分子中的所述第一有义核糖核苷酸序列完全碱基配对,其中所述第一反义核糖核苷酸序列与靶RNA分子的第一区的互补体的序列相同。这两个实施方案的该第一RNA组分的实施例示意性地显示在图1A的左半部分或图1B的右半部分中。

[0313] 在另一个实施方案中,所述第一RNA组分由第一5'核糖核苷酸、第一RNA序列和第一3'核糖核苷酸组成,其中所述第一RNA组分中的第一5'和3'核糖核苷酸彼此碱基配对,其中第一RNA序列包含第一有义核糖核苷酸序列、至少4个核糖核苷酸的第一环序列和第一反义核糖核苷酸序列,其中所述第一有义核糖核苷酸序列和第一反义核糖核苷酸序列各自至少有20个连续核糖核苷酸,其中所述第一有义核糖核苷酸序列的至少20个连续核糖核苷酸与所述第一反义核糖核苷酸序列的至少20个连续核糖核苷酸完全碱基配对,其中所述第一有义核糖核苷酸序列的至少20个连续核糖核苷酸在序列上与靶RNA分子的第一区域基本上相同。

[0314] 在这些实施方案中,第一5'核糖核苷酸和第一3'核糖核苷酸之间形成的碱基对被认为是通过第一个RNA组分的自身杂交形成的dsRNA区的末端碱基对,即它限定了dsRNA区的末端。

[0315] 在一个实施方案中,第一有义序列与靶RNA的区具有基本上的序列同一性,该同一性可以是长度小于20个核苷酸的序列。在一个实施方案中,第一有义核糖核苷酸序列的至少15个、至少16个、至少17个、至少18个或至少19个邻接核糖核苷酸、优选至少20个邻接核糖核苷酸与靶RNA分子的第一区在序列上至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%,或99%相同。在另一个实施方案中,第一有义核糖核苷酸序列的至少15个、至少16个、至少17个、至少18个、至少19个邻接核糖核苷酸与靶RNA分子的第一区100%相同。在一个实施方案中,来自第一有义核糖核苷酸序列的5'端的前3、前4、前5、前6或前7个核糖核苷酸与靶RNA分子的区100%相同,其余的核糖核苷酸与该靶RNA分子至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%相同。

[0316] 在一个实施方案中,第一有义核糖核苷酸序列的至少20个邻接核糖核苷酸与靶RNA分子的第一区至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%相同。同样,在该实施方案中,前3、前4、前5、前6或前7个核糖核苷酸可以与靶RNA分子的区100%相同,其余的核糖核苷酸与靶RNA分子至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%相同。在另一个实施方案中,第一有义核糖核苷酸序列的至少20个邻接核糖核苷酸与靶RNA分子的第一区100%相同。

[0317] 在一个实施方案中,所述第一反义序列与所述靶RNA的区的互补体具有基本上的序列同一性,该同一性可以是与所述互补体的长度小于20个核苷酸的序列的同一性。在一个实施方案中,第一反义核糖核苷酸序列的至少15个、至少16个、至少17个、至少18个或至少19个邻接核糖核苷酸、优选至少20个邻接核糖核苷酸与靶RNA分子的第一区的互补体在序列上至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%,或99%相同。在另一个实施方案中,第一反义核糖核

苷酸序列的至少15个、至少16个、至少17个、至少18个、至少19个邻接核糖核苷酸与靶RNA分子的第一区的互补体100％相同。在一个实施方案中,来自第一反义核糖核苷酸序列的5'端的前3、前4、前5、前6或前7核糖核苷酸与靶RNA分子区的互补序列100％相同,其余的核糖核苷酸与该靶RNA分子的互补体至少60％、至少65％、至少70％、至少75％、至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％相同。

[0318] 在一个实施方案中,第一反义核糖核苷酸序列的至少20个邻接核糖核苷酸与靶RNA分子的第一区的互补体至少60％、至少65％、至少70％、至少75％、至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％相同。同样,在该实施方案中,前3、前4、前5、前6或前7个核糖核苷酸与靶RNA分子区的互补体100％相同,其余的核糖核苷酸与靶RNA分子的互补体至少60％、至少65％、至少70％、至少75％、至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％相同。在另一个实施方案中,第一反义核糖核苷酸序列的至少20个邻接核糖核苷酸与靶RNA分子的第一区100％相同。

[0319] 在另一个实施方案中,所述第二RNA组分按5'至3'顺序由第二5'核糖核苷酸、第二RNA序列和第二3'核糖核苷酸组成,其中所述第二5'和3'核糖核苷酸碱基配对,其中所述第二RNA序列包含第二有义核糖核苷酸序列、至少4个核糖核苷酸的第二环序列和第二反义核糖核苷酸序列,其中所述第二有义核糖核苷酸序列与所述第二反义核糖核苷酸序列碱基配对。在该实施方案中,第二5'核糖核苷酸和第二3'核糖核苷酸之间形成的碱基对被认为是通过第二RNA组分的自身杂交形成的dsRNA区的末端碱基对。

[0320] 在一个实施方案中,RNA分子包含5'前导序列或5'延伸序列,其可以作为从遗传构建体中的启动子,从转录的起始位点到编码RNA分子的剩余部分的多核苷酸的起点转录的结果而产生。优选该5'前导序列或5'延伸序列与分子的剩余部分相比相对较短,并且可以在转录后从RNA分子中除去,例如通过RNA酶处理。5'前导序列或5'延伸序列可以大部分是非碱基配对的,或者可以含有一个或多个茎-环结构。在该实施方案中,5'前导序列可以由核糖核苷酸序列组成,如果所述第二RNA组分与所述第一3' 核糖核苷酸连接,则所述核糖核苷酸序列与所述第一5' 核糖核苷酸共价连接,或者如果所述第二RNA组分与所述第一5' 核糖核苷酸连接,则所述核糖核苷酸序列与所述第二5' 核糖核苷酸共价连接。在一个实施方案中,5'前导序列为至少10、至少20、至少30、至少100、至少200个核糖核苷酸长,优选至250个核糖核苷酸的最大长度。在另一个实施方案中,5'前导序列为至少50个核糖核苷酸长。在一个实施方案中,5'前导序列可以作为延伸序列用于通过合适的扩增反应扩增RNA分子。对于实施方案,延伸序列可以促进通过聚合酶的扩增。

[0321] 在另一个实施方案中,RNA分子包含3'尾随序列或3'延伸序列,其可能由于连续转录直到编码RNA分子的构建体中的转录终止或聚腺苷酸化信号而产生。3'尾随序列或3'延伸序列可以包含多聚A(polyA)尾。优选该3'尾随序列或3'延伸序列与分子的剩余部分相比相对较短,并且可以在转录后从RNA分子中除去,例如通过RNA酶处理。3'尾随序列或3'延伸序列可以大部分是非碱基配对的,或者可以含有一个或多个茎-环结构。在该实施方案中,3'尾随序列可以由核糖核苷酸序列组成,如果所述第二RNA组分与所述第一3' 核糖核苷酸连接,则所述核糖核苷酸序列与所述第二3' 核糖核苷酸共价连接,或者如果所述第二RNA组分与所述第一5' 核糖核苷酸连接,则所述核糖核苷酸序列与所述第一3' 核糖核苷酸共价连

接。在一个实施方案中,3'前导序列为至少10、至少20、至少30、至少100、至少200个核糖核苷酸长,优选至250个核糖核苷酸的最大长度。在另一个实施方案中,3'前导序列为至少50个核糖核苷酸长。在一个实施方案中,3'尾随序列可以作为延伸序列用于通过合适的扩增反应扩增RNA分子。对于实施方案,延伸序列可以促进通过聚合酶的扩增。

[0322] 在一个实施方案中,除了两个核糖核苷酸之外的所有核苷酸与两个其他核苷酸共价连接,即RNA分子仅由一条具有自身互补区的RNA链组成,且因此仅具有一个5'末端核苷酸和一个3'末端核苷酸。在另一个实施方案中,除了4个核糖核苷酸之外的所有核苷酸与两个其他核苷酸共价连接,即RNA分子由具有杂交的互补区的2条RNA链组成,因此仅具有两个5'末端核苷酸和两个3'末端核苷酸。在另一个实施方案中,每个核糖核苷酸与另外两个核苷酸共价连接,即RNA分子是环状的并且具有自身互补区,因此没有5'末端核苷酸和3'末端核苷酸。

[0323] 在一个实施方案中,RNA分子的双链区可以包含一个或多个由有义RNA序列或反义RNA序列或两者中的未配对核苷酸产生的隆起。在一个实施方案中,RNA分子包含一系列突起。对于实施方案,RNA分子的双链区可以具有2、3、4、5、6、7、8、9、10或更多个隆起。每个隆起部分可以独立地为一个、两个或更多个未配对的核苷酸,至多达10个核苷酸。较长的序列可以环出dsRNA区中的有义或反义序列,其可以在内部碱基配对或保持不配对。在另一个实施方案中,RNA分子的双链区不包含隆起,即沿着dsRNA区的全长完全碱基配对。

[0324] 在另一个实施方案中,第一有义核糖核苷酸序列与第一5'核糖核苷酸共价连接而没有任何间插核苷酸,或第一反义核糖核苷酸序列与第一3'核糖核苷酸共价连接而没有任何间插核苷酸,或两者。在另一个实施方案中,存在至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个间插核苷酸。应理解这样的间插核苷酸在序列上与靶RNA分子无关,但可有助于稳定相邻有义和反义序列的碱基配对。

[0325] 在另一个实施方案中,第一有义核糖核苷酸序列的20个邻接核苷酸与第一5'核糖核苷酸共价连接,没有任何间插核苷酸,并且第一反义核糖核苷酸序列的20个邻接核苷酸与第一3'核糖核苷酸共价连接,没有任何间插核苷酸。在另一个实施方案中,存在至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个间插核苷酸。间插核苷酸可作为RNA分子的双链区的一部分进行碱基配对,但在序列上与靶RNA无关。它们可以帮助为双链区提供增加的稳定性或将RNA分子的两端结合在一起而不留下未配对的5'或3'端,或两者。

[0326] 在一个实施方案中,上述第一和第二RNA组分包含连接核糖核苷酸序列。在一个实施方案中,连接核糖核苷酸序列作为第一有义核糖核苷酸序列和分子的其他组分之间的间隔物,所述第一有义核糖核苷酸序列在序列上与靶RNA分子的第一区基本相同。例如,连接核糖核苷酸序列可以作为该区和环之间的间隔物。在另一个实施方案中,RNA分子包含多个有义核糖核苷酸序列和连接核糖核苷酸序列,所述有义核糖核苷酸序列在序列上与靶RNA分子的第一区基本相同,所述连接核糖核苷酸序列充当这些序列之间的间隔物。在一个实施方案中,在RNA分子中提供了在序列上与靶RNA分子的第一区基本相同的至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个核糖核苷酸序列,每个核糖核苷酸序列通过连接核糖核苷酸序列彼此分开。

[0327] 在一个实施方案中,上述RNA分子包含5'前导序列。在一个实施方案中,5'前导序

列由核糖核苷酸序列组成,如果所述第二RNA组分与所述第一3’核糖核苷酸连接,则所述核糖核苷酸序列与所述第一5’核糖核苷酸共价连接,或者如果所述第二RNA组分与所述第一5’核糖核苷酸连接,则所述核糖核苷酸序列与所述第二5’核糖核苷酸共价连接。在一个实施方案中,RNA分子具有修饰的5’或3’端,例如通过诸如胆固醇的连接脂质基团,或诸如生物素的维生素,或多肽。这样的修饰可以帮助RNA分子摄入其中RNA将起作用的真核细胞。

[0328] 在一个实施方案中,连接核糖核苷酸序列的长度小于100个核糖核苷酸。在一个实施方案中,连接核糖核苷酸序列的长度小于50个核糖核苷酸。在一个实施方案中,连接核糖核苷酸序列的长度小于20个核糖核苷酸。在一个实施方案中,连接核糖核苷酸序列的长度小于10个核糖核苷酸。在一个实施方案中,连接核糖核苷酸序列的长度小于5个核糖核苷酸。在一个实施方案中,连接核糖核苷酸序列的长度为1至100个核糖核苷酸。在一个实施方案中,连接核糖核苷酸序列的长度为1至50个核糖核苷酸。在一个实施方案中,连接核糖核苷酸序列的长度为1至20个核糖核苷酸。在一个实施方案中,连接核糖核苷酸序列的长度为1至10个核糖核苷酸。在一个实施方案中,连接核糖核苷酸序列的长度为1至5个核糖核苷酸。在一个实施方案中,连接核糖核苷酸序列的核糖核苷酸不是碱基配对的。在优选的实施方案中,连接核糖核苷酸序列的核糖核苷酸全部被碱基配对,或者除了1、2或3个核糖核苷酸之外全部被碱基配对。

[0329] 在一个实施方案中,第一或第二RNA组分包含发夹结构。在一个优选的实施方案中,第一和第二RNA组分各自包含发夹结构。在这些实施方案中,发夹结构可以是茎-环。因此,在一个实施方案中,RNA分子可包含各自包含发夹结构的第一和第二RNA组分,其中发夹通过接头序列共价结合。例如,参见图1。在一个实施方案中,所述接头序列是一个或多个未配对核糖核酸(RNA)。在一个实施方案中,接头序列为1至10个未配对核糖核苷酸。

[0330] 在一个实施方案中,RNA分子具有双发夹结构,即“1edRNA结构”或“哑铃结构”。在该实施方案中,第一发夹是第一RNA组分并且第二发夹是第二RNA组分。在这些实施方案中,第一个3’核糖核苷酸和第二个5’核糖核苷酸,或第二个3’核糖核苷酸和第一个5’核糖核苷酸,但并非两者,共价连接。在该实施方案中,其他5’/3’核糖核苷酸可以通过裂口(即dsRNA分子中的不连续性,其中在5’/3’核糖核苷酸之间没有磷酸二酯键)分开。图1B示出了该类型布置的实施例。在另一个实施方案中,相应的5’/3’核糖核苷酸可以通过环分开。5’前导序列和3’尾随序列的长度可以相同或不同。对于实施方案,5’前导可以比3’尾随序列长大约5、10、15、20、25、50、100、200、500个核糖核苷酸,反之亦然。

[0331] 在RNA分子具有双发夹结构的实施方案中,第二发夹(除了第一发夹结构之外)包含有义RNA序列和反义RNA序列,其序列分别与靶RNA分子或其互补体的区基本相同。在一个实施方案中,每个发夹具有与同一靶RNA分子的区在序列上基本相同的一系列核糖核苷酸。在一个实施方案中,每个发夹具有与同一靶RNA分子的不同区在序列上基本上相同的一系列核糖核苷酸。在一个实施方案中,每个发夹具有与不同靶RNA分子的区在序列上基本相同的一系列核糖核苷酸,即RNA分子可用于降低在序列上不相关的两个靶RNA分子的表达和/或活性。

[0332] 在RNA分子的双发夹结构的每个发夹结构中,按5’至3’顺序,每个发夹结构中有义和反义RNA序列中的顺序可以独立地为有义然后反义,或反义然后有义。在优选的实施方案中,RNA分子的双发夹结构中的有义和反义序列的顺序是其中两个有义序列是邻接的反义-

有义-有义-反义(图1A),或者是其中两个反义序列是邻接的有义-反义-反义-有义(图1B)。

[0333] 在一个实施方案中,按5'至3'顺序,所述RNA分子可以包含5'前导序列,第一环,有义RNA序列,第二环和3'尾随序列,其中所述5'和3'前导序列共价结合至所述有义链以形成dsRNA序列。在一个实施方案中,5'前导序列和3'尾随序列彼此不共价结合。在一个实施方案中,5'前导序列和3'尾随序列被裂口分开。在一个实施方案中,将5'前导序列和3'尾随序列连接在一起以提供具有闭合结构的RNA分子。在另一个实施方案中,5'前导序列和3'尾随序列由环分开。

[0334] 在本发明的上下文中使用的术语"环"是指在此公开的RNA分子中由一系列非互补核糖核苷酸形成的环结构。环通常跟随第一和第二RNA组分之间的一系列碱基对或连接第一和第二RNA组分之一或两者中的有义RNA序列和反义RNA序列。在一个实施方案中,通常对于4-10个核糖核苷酸的较短的环,所有环核糖核苷酸是非互补的。在其他实施方案中,一个或多个环中的一些核糖核苷酸是互补的并且能够在环序列内碱基配对,只要这些碱基配对能够形成环结构。例如,至少5%、至少10%或至少15%的环核糖核苷酸是互补的。环的实施方案包括茎环或发夹、假结和四环。

[0335] 在一个实施方案中,RNA分子仅包含两个环。在另一个实施方案中,RNA分子包含至少两个、至少三个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个或至少10个环,优选至多10个环。例如,RNA分子可包含4个环。

[0336] 本发明考虑了各种大小的环。例如,环可以包含4、5、6、7、8、9、10、11或12个核糖核苷酸。在其他实施方案中,环包含15、20、25或30个核苷酸。在一个实施方案中,环序列中的一个或全部长于20个核苷酸。在其他实施方案中,环较大,例如包含50、100、150、200或300个核糖核苷酸。在一个实施方案中,环包含160个核糖核苷酸。在另一个较不优选的实施方案中,环包含200、500、700或1,000个核糖核苷酸,条件是环不干扰有义和反义RNA序列的杂交。在一个实施方案中,每个环具有相同数目的核糖核苷酸。例如,环可以具有100至1,000个核糖核苷酸的长度。例如,环可以具有600至1,000个核糖核苷酸的长度。例如,环可以具有4至1,000个核糖核苷酸。例如,环优选具有4-50个核糖核苷酸。在另一个实施方案中,环包含不同数目的核糖核苷酸。

[0337] 在另一个实施方案中,一个或多个环包含可以从RNA分子剪接出来的内含子。在一个实施方案中,内含子来自植物基因。示例性的内含子包括玉米醇脱氢酶1(Adh1)(GenBank:AF044293)的内含子3,大豆β-伴大豆球蛋白α亚基(GenBank:AB051865)的内含子4;核酮糖-1,5-二磷酸羧化酶(RBC)小亚基的豌豆rbcS-3A基因(GenBank:X04333)的内含子之一。在(McCullough和Schuler,1997;Smith等,2000)中讨论了合适的内含子的其他实施方案。

[0338] 在各种实施方案中,环可以在至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个邻接碱基对的末端,其可以是规范碱基对或可以包括一个或多个非规范碱基对。在其他实施方案中,特别是脊椎动物细胞而言,不太优选地,环可以在至少20、30、50、100、200、500或更多邻接碱基对的末端。

[0339] 在另一个实施方案中,RNA分子包含两个或更多个有义核糖核苷酸序列和完全基于其碱基配对的反义核糖核苷酸序列,它们各自在序列上与靶RNA分子的区相同。例如,RNA分子可包含2、3、4、5、6、7、8、9、10、11、12个或更多个有义核糖核苷酸序列,和与其完全配对

的反义核糖核苷酸序列,所述有义核糖核苷酸序列各自独立地在序列上与靶RNA分子的区相同。在该实施方案中,任何一个或多个或所有序列可由连接核糖核苷酸序列分开。在该实施方案中,任何一个或多个或所有序列可以由环分开。

[0340] 在一个实施方案中,两个或多个有义核糖核苷酸序列在序列上与同一靶RNA分子的不同区相同。例如,所述序列可以与同一靶分子的至少2个、至少3个、至少4个、至少5个、至少6个区相同。在另一个实施方案中,两个或更多个有义核糖核苷酸序列在序列上是相同的。在一个实施方案中,两个或多个有义核糖核苷酸序列在序列上与同一靶RNA分子的相同区相同。在另一个实施方案中,两个或多个有义核糖核苷酸序列在序列上与不同的靶RNA分子相同。对于实施方案,所述序列可以与不同靶分子的至少2个、至少3个、至少4个、至少5个、至少6个区相同。

[0341] 在另一个实施方案中,两个或多个有义核糖核苷酸序列没有间插环(间隔物)序列。

[0342] 在一个实施方案中,RNA分子可以包含具有5'端的单链核糖核苷酸、长度为至少21个核苷酸的至少一个有义核糖核苷酸序列、在至少21个邻接核苷酸上与每个有义核糖核苷酸序列完全碱基配对的反义核糖核苷酸序列、至少两个环序列和3'端。在该实施方案中,5'端的核糖核苷酸和3'端的核糖核苷酸不直接共价结合,而是与每个碱基配对的位置相邻。

[0343] 在另一个实施方案中,RNA组分的邻接碱基对间隔至少一个缺口。在一个实施方案中,"缺口"由未配对的核糖核苷酸提供。在另一个实施方案中,"缺口"由未连接的5'前导序列和/或3'尾随序列提供。在该实施例中,缺口可称为"未连接缺口"。错配和未连接缺口可位于RNA分子的不同位置。对于实施方案,未连接缺口可以紧跟在反义序列之后。在另一个实施方案中,未连接的缺口可以靠近RNA分子的环。在另一个实施例中,未连接的缺口被定位成在至少两个环之间大约等距。

[0344] 在一个实施方案中,RNA分子由单链RNA产生。在一个实施方案中,单链不是环状闭合的,例如包含未连接的缺口。在另一个实施方案中,RNA分子是环状闭合分子。可通过例如用RNA连接酶连接包含未连接缺口的上述RNA分子来产生闭合分子。

[0345] 在另一个实施方案中,RNA分子包含5'-或3'-延伸序列或两者。例如,RNA分子可以包含与第一5'核糖核苷酸共价连接的5'延伸序列。在另一个实施方案中,RNA分子包含与第二3'核糖核苷酸共价连接的3'延伸序列。在另一个实施方案中,RNA分子包含与第一5'核糖核苷酸共价连接的5'延伸序列和与第二3'核糖核苷酸共价连接的3'延伸序列。

[0346] 在另一个实施方案中,RNA分子包含与第二5'核糖核苷酸共价连接的5'延伸序列。在另一个实施方案中,RNA分子包含与第一3'核糖核苷酸共价连接的3'延伸序列。在另一个实施方案中,RNA分子包含与第二5'核糖核苷酸共价连接的5'延伸序列和与第一3'核糖核苷酸共价连接的3'延伸序列。

[0347] 在另一个实施方案中,RNA分子可以包含以下中的一种或多种:

[0348] -与第一5'核糖核苷酸共价连接的5'延伸序列;

[0349] -与第二3'核糖核苷酸共价连接的3'延伸序列;

[0350] -与第一5'核糖核苷酸共价连接的5'延伸序列和与第二3'核糖核苷酸共价连接的3'延伸序列;

[0351] -与第二5'核糖核苷酸共价连接的5'延伸序列;

[0352] -与第一3'核糖核苷酸共价连接的3'延伸序列;

[0353] -与第二5'核糖核苷酸共价连接的5'延伸序列和与第一3'核糖核苷酸共价连接的3'延伸序列。

[0354] 非规范碱基配对

[0355] 在一个实施方案中,本发明的RNA分子包含有义核糖核苷酸序列和反义核糖核苷酸序列,它们能够彼此杂交,以形成具有一些非规范碱基配对即具有规范和非规范碱基配对的组合的双链(ds)RNA区。在一个实施方案中,本发明的RNA分子包含两个或更多个有义核糖核苷酸序列,其各自能够与一个(邻接)反义核糖核苷酸序列的区杂交,以形成具有一些非规范碱基配对的dsRNA区。参见例如图1B。在一个实施方案中,本发明的RNA分子包含两个或更多个反义有义核糖核苷酸序列,其各自能够与一个(邻接)有义核糖核苷酸序列的区杂交,以形成具有一些非规范碱基配对的dsRNA区。参见例如图1A。在一个实施方案中,本发明的RNA分子包含两个或更多个反义有义核糖核苷酸序列和两个或更多个有义核糖核苷酸序列,其中每个反义核糖核苷酸序列能够与反义核糖核苷酸序列杂交以形成两个或更多个dsRNA区,一个或两个dsRNA区包含一些非规范碱基配对。

[0356] 在以下实施方案中,如果仅存在一个(邻接的)dsRNA区,则认为本发明的RNA分子的dsRNA区的全长(即整个dsRNA区)是特征的上下文,或者如果在RNA分子中存在两个或更多个dsRNA区,则认为是RNA分子的每个dsRNA区的上下文。在一个实施方案中,dsRNA区中至少5%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中至少6%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中至少7%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中至少8%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中至少9%或10%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中至少11%或12%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中至少15%或约15%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中至少20%或约20%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中至少25%或约25%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中至少30%或约30%的碱基对是非规范碱基对。在这些实施方案的每一个中,优选dsRNA区中最多40%的碱基对是非规范碱基对,更优选dsRNA区中最多35%的碱基对是非规范碱基对,还更优选dsRNA区中最多30%的碱基对是非规范碱基对。在一个较不优选的实施方案中,dsRNA区中约35%的碱基对是非规范碱基对。在一个实施方案中,甚至更较不优选地,dsRNA区中约40%的碱基对是非规范碱基对。在上述实施方案的每一个中,dsRNA区可以在有义序列或反义序列或两者中包含或不包含一个或多个非碱基配对的核糖核苷酸。

[0357] 在一个实施方案中,本发明的RNA分子的dsRNA区中10%-40%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中10%-35%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中10%-30%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中10%-25%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中10%-20%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中10%-15%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中15%-30%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中15%-25%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中15%-20%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中5%-30%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中5%-25%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中

5%-20%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中5%-15%的碱基对是非规范碱基对。在一个实施方案中,dsRNA区中5%-10%的碱基对是非规范碱基对。在上述实施方案的每一个中,dsRNA区可以在有义序列或反义序列或两者中包含或不包含一个或多个非碱基配对的核糖核苷酸。

[0358] 在一个实施方案中,本发明的RNA分子的dsRNA区包含20个邻接碱基对,其中所述20个邻接碱基对中的至少一个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含20个邻接碱基对,其中20个邻接碱基对中的至少2个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含20个邻接碱基对,其中20个邻接碱基对中的至少3个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含20个邻接碱基对,其中20个邻接碱基对中的至少4个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含20个邻接碱基对,其中20个邻接碱基对中的至少5个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含20个邻接碱基对,其中20个邻接碱基对中的至少6个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含20个邻接碱基对,其中20个邻接碱基对中的至少7个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含20个邻接碱基对,其中20个邻接碱基对中的至少8个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含20个邻接碱基对,其中20个邻接碱基对中的至少9个碱基对是非规范碱基对。在这些实施方案的每一个中,优选dsRNA区中20个邻接碱基对中最多10个是非规范碱基对,更优选dsRNA区中最多9个是非规范碱基对,还更优选dsRNA区中最多8个碱基对是非规范碱基对,甚至更优选dsRNA区中最多7个碱基对是非规范碱基对,最优选dsRNA区中最多6个碱基对是非规范碱基对。优选地,在上述实施方案中,非规范碱基对包含至少一个G:U碱基对,更优选所有的非规范碱基对为G:U碱基对。优选地,上述实施方案的特征适用于本发明RNA分子中存在的20个邻接碱基对中的每一个。

[0359] 在一个实施方案中,本发明的RNA分子的dsRNA区包含21个邻接碱基对,其中所述21个邻接碱基对中的至少一个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含21个邻接碱基对,其中21个邻接碱基对中的至少2个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含21个邻接碱基对,其中21个邻接碱基对中的至少3个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含21个邻接碱基对,其中21个邻接碱基对中的至少4个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含21个邻接碱基对,其中21个邻接碱基对中的至少5个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含21个邻接碱基对,其中21个邻接碱基对中的至少6个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含21个邻接碱基对,其中21个邻接碱基对中的至少7个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含21个邻接碱基对,其中21个邻接碱基对中的至少8个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含21个邻接碱基对,其中21个邻接碱基对中的至少9个碱基对是非规范碱基对。在这些实施方案的每一个中,优选dsRNA区中21个邻接碱基对中最多10个是非规范碱基对,更优选dsRNA区中最多9个是非规范碱基对,还更优选dsRNA区中最多8个碱基对是非规范碱基对,甚至更优选dsRNA区中最多7个碱基对是非规范碱基对,最优选dsRNA区中最多6个碱基对是非规范碱基对。优选地,在上述实施方案中,非规范碱基对包含至少一个G:U碱基对,更优选所有的非规范碱基对为G:U碱基对。优选地,上述实施方案的特征适用于本发明RNA分子中存在的21个邻接碱基对中的每一个。

[0360] 在一个实施方案中,本发明的RNA分子的dsRNA区包含22个邻接碱基对,其中所述

22个邻接碱基对中的至少一个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含22个邻接碱基对,其中22个邻接碱基对中的至少2个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含22个邻接碱基对,其中22个邻接碱基对中的至少3个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含22个邻接碱基对,其中22个邻接碱基对中的至少4个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含22个邻接碱基对,其中22个邻接碱基对中的至少5个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含22个邻接碱基对,其中22个邻接碱基对中的至少6个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含22个邻接碱基对,其中22个邻接碱基对中的至少7个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含22个邻接碱基对,其中22个邻接碱基对中的至少8个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含22个邻接碱基对,其中22个邻接碱基对中的至少9个碱基对是非规范碱基对。在这些实施方案的每一个中,优选dsRNA区中22个邻接碱基对中最多10个是非规范碱基对,更优选dsRNA区中最多9个是非规范碱基对,还更优选dsRNA区中最多8个碱基对是非规范碱基对,甚至更优选dsRNA区中最多7个碱基对是非规范碱基对,最优选dsRNA区中最多6个碱基对是非规范碱基对。优选地,在上述实施方案中,非规范碱基对包含至少一个G:U碱基对,更优选所有的非规范碱基对为G:U碱基对。优选地,上述实施方案的特征适用于本发明RNA分子中存在的22个邻接碱基对中的每一个。

[0361] 在一个实施方案中,本发明的RNA分子的dsRNA区包含23个邻接碱基对,其中所述23个邻接碱基对中的至少一个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含23个邻接碱基对,其中23个邻接碱基对中的至少2个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含23个邻接碱基对,其中23个邻接碱基对中的至少3个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含23个邻接碱基对,其中23个邻接碱基对中的至少4个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含23个邻接碱基对,其中23个邻接碱基对中的至少5个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含23个邻接碱基对,其中23个邻接碱基对中的至少6个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含23个邻接碱基对,其中23个邻接碱基对中的至少7个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含23个邻接碱基对,其中23个邻接碱基对中的至少8个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含23个邻接碱基对,其中23个邻接碱基对中的至少9个碱基对是非规范碱基对。在这些实施方案的每一个中,优选dsRNA区中23个邻接碱基对中最多10个是非规范碱基对,更优选dsRNA区中最多9个是非规范碱基对,还更优选dsRNA区中最多8个碱基对是非规范碱基对,甚至更优选dsRNA区中最多7个碱基对是非规范碱基对,最优选dsRNA区中最多6个碱基对是非规范碱基对。优选地,在上述实施方案中,非规范碱基对包含至少一个G:U碱基对,更优选所有的非规范碱基对为G:U碱基对。优选地,上述实施方案的特征适用于本发明RNA分子中存在的23个邻接碱基对中的每一个。

[0362] 在一个实施方案中,本发明的RNA分子的dsRNA区包含24个邻接碱基对,其中所述24个邻接碱基对中的至少一个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含24个邻接碱基对,其中24个邻接碱基对中的至少2个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含24个邻接碱基对,其中24个邻接碱基对中的至少3个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含24个邻接碱基对,其中24个邻接碱基对中的至少4个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含24个邻接碱基对,其中24个邻接碱

基对中的至少5个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含24个邻接碱基对,其中24个邻接碱基对中的至少6个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含24个邻接碱基对,其中24个邻接碱基对中的至少7个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含24个邻接碱基对,其中24个邻接碱基对中的至少8个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含24个邻接碱基对,其中24个邻接碱基对中的至少9个碱基对是非规范碱基对。在这些实施方案的每一个中,优选dsRNA区中24个邻接碱基对中最多10个是非规范碱基对,更优选dsRNA区中最多9个是非规范碱基对,还更优选dsRNA区中最多8个碱基对是非规范碱基对,甚至更优选dsRNA区中最多7个碱基对是非规范碱基对,最优选dsRNA区中最多6个碱基对是非规范碱基对。优选地,在上述实施方案中,非规范碱基对包含至少一个G:U碱基对,更优选所有的非规范碱基对为G:U碱基对。优选地,上述实施方案的特征适用于本发明RNA分子中存在的24个邻接碱基对中的每一个。

[0363] 在以下实施方案中,如果仅存在一个(邻接的)dsRNA区,则认为本发明的RNA分子的dsRNA区的全长(即整个dsRNA区)是特征的上下文,或者如果在RNA分子中存在两个或更多个dsRNA区,则认为是RNA分子的每个dsRNA区的上下文。在一个实施方案中,dsRNA区不包含20个邻接的规范碱基对,即20个邻接碱基对的每个子区包含至少一个非规范碱基对,优选至少一个G:U碱基对。在一个实施方案中,dsRNA区不包含19个邻接的规范碱基对。在一个实施方案中,dsRNA区不包含18个邻接的规范碱基对。在一个实施方案中,dsRNA区不包含17个邻接的规范碱基对。在一个实施方案中,dsRNA区不包含16个邻接的规范碱基对。在一个实施方案中,dsRNA区不包含15个邻接的规范碱基对。在一个实施方案中,dsRNA区不包含14个邻接的规范碱基对。在一个实施方案中,dsRNA区不包含13个邻接的规范碱基对。在一个实施方案中,dsRNA区不包含12个邻接的规范碱基对。在一个实施方案中,dsRNA区不包含11个邻接的规范碱基对。在一个实施方案中,dsRNA区不包含10个邻接的规范碱基对。在一个实施方案中,dsRNA区不包含9个邻接的规范碱基对。在一个实施方案中,dsRNA区不包含8个邻接的规范碱基对。在一个实施方案中,dsRNA区不包含7个邻接的规范碱基对。在上述实施方案中,优选RNA分子的dsRNA区或RNA分子中每个dsRNA区的邻接规范碱基配对的最长子区是5、6或7个邻接规范碱基配对,即朝向所述较短的长度。上述实施方案的每个特征优选与以下特征组合在RNA分子中。在一个实施方案中,dsRNA区包含10-19或20个邻接碱基对。在一个优选的实施方案中,dsRNA区包含12-19或20个邻接碱基对。在一个实施方案中,dsRNA区包含14-19或20个邻接碱基对。在这些实施方案中,dsRNA区包含15个邻接碱基对。在一个实施方案中,dsRNA区包含16、17、18或19个邻接碱基对。在一个实施方案中,dsRNA区包含20个邻接碱基对。优选地,在上述实施方案中,邻接碱基对包含包括至少一个G:U碱基对的至少一个非规范碱基对,更优选地,邻接碱基对区中的所有非规范碱基对为G:U碱基对。

[0364] 在一个实施方案中,dsRNA区包含侧接非规范碱基对的4个规范碱基对的子区,即邻近4个规范碱基对的每一端的至少1个、优选1个或2个(不超过2个)非规范碱基对。在一个实施方案中,dsRNA区包含2个子区,每个子区为侧接非规范碱基对的4个标准碱基对。在一个实施方案中,dsRNA区包含3个子区,每个子区为侧接非规范碱基对的4个标准碱基对。在一个实施方案中,dsRNA区包含4个或5个子区,每个子区为侧接非规范碱基对的4个标准碱基对。在一个实施方案中,dsRNA区包含6个或7个子区,每个子区为侧接非规范碱基对的4个标准碱基对。在一个实施方案中,dsRNA区包含8至10个子区,每个子区为侧接非规范碱基对

的4个标准碱基对。在一个实施方案中,dsRNA区包含11至15个子区,每个子区为侧接非规范碱基对的4个标准碱基对。在一个实施方案中,dsRNA区包含2至50个子区,每个子区为侧接非规范碱基对的4个标准碱基对。在一个实施方案中,dsRNA区包含2至40个子区,每个子区为侧接非规范碱基对的4个标准碱基对。在一个实施方案中,dsRNA区包含2至30个子区,每个子区为侧接非规范碱基对的4个标准碱基对。在一个实施方案中,dsRNA区包含2至20个子区,每个子区为侧接非规范碱基对的4个标准碱基对。优选地,在上述实施方案中,非规范碱基对包含至少一个G:U碱基对,更优选地,在子区中邻接的规范碱基对侧接的所有非规范碱基对是G:U碱基对。在上述实施方案的变体中,对于一些或所有的子区,侧接非规范碱基对中的一个或两个在有义序列、反义序列或在两个序列中,被非碱基配对的核糖核苷酸替代。容易理解,在上述实施方案中,子区的最大数目由RNA分子中dsRNA区的长度决定。

[0365] 在一个实施方案中,dsRNA区包含侧接非规范碱基对的5个规范碱基对的子区。在一个实施方案中,dsRNA区包含2个子区,每个子区为侧接非规范碱基对的5个标准碱基对。在一个实施方案中,dsRNA区包含3个子区,每个子区为侧接非规范碱基对的5个标准碱基对。在一个实施方案中,dsRNA区包含4个或5个子区,每个子区为侧接非规范碱基对的5个标准碱基对。在一个实施方案中,dsRNA区包含6个或7个子区,每个子区为侧接非规范碱基对的5个标准碱基对。在一个实施方案中,dsRNA区包含8至10个子区,每个子区为侧接非规范碱基对的5个标准碱基对。在一个实施方案中,dsRNA区包含11至15个子区,每个子区为侧接非规范碱基对的5个标准碱基对。在一个实施方案中,dsRNA区包含2至50个子区,每个子区为侧接非规范碱基对的5个标准碱基对。在一个实施方案中,dsRNA区包含2至50个子区,每个子区为侧接非规范碱基对的5个标准碱基对。在一个实施方案中,dsRNA区包含2至30个子区,每个子区为侧接非规范碱基对的5个标准碱基对。在一个实施方案中,dsRNA区包含2至20个子区,每个子区为侧接非规范碱基对的5个标准碱基对。优选地,在上述实施方案中,非规范碱基对包含至少一个G:U碱基对,更优选地,在子区中邻接的规范碱基对侧接的所有非规范碱基对是G:U碱基对。在上述实施方案的变体中,对于一些或所有的子区,侧接非规范碱基对中的一个或两个在有义序列、反义序列或在两个序列中,被非碱基配对的核糖核苷酸替代。容易理解,在上述实施方案中,子区的最大数目由RNA分子中dsRNA区的长度决定。

[0366] 在一个实施方案中,dsRNA区包含侧接非规范碱基对的6个规范碱基对的子区。在一个实施方案中,dsRNA区包含2个子区,每个子区为侧接非规范碱基对的6个标准碱基对。在一个实施方案中,dsRNA区包含3个子区,每个子区为侧接非规范碱基对的6个标准碱基对。在一个实施方案中,dsRNA区包含4个或5个子区,每个子区为侧接非规范碱基对的6个标准碱基对。在一个实施方案中,dsRNA区包含6个或7个子区,每个子区为侧接非规范碱基对的6个标准碱基对。在一个实施方案中,dsRNA区包含8至10个子区,每个子区为侧接非规范碱基对的6个标准碱基对。在一个实施方案中,dsRNA区包含11至16个子区,每个子区为侧接非规范碱基对的6个标准碱基对。在一个实施方案中,dsRNA区包含2至60个子区,每个子区为侧接非规范碱基对的6个标准碱基对。在一个实施方案中,dsRNA区包含2至60个子区,每个子区为侧接非规范碱基对的6个标准碱基对。在一个实施方案中,dsRNA区包含2至30个子区,每个子区为侧接非规范碱基对的6个标准碱基对。在一个实施方案中,dsRNA区包含2至20个子区,每个子区为侧接非规范碱基对的6个标准碱基对。优选地,在上述实施方案中,非规范碱基对包含至少一个G:U碱基对,更优选地,在子区中邻接的规范碱基对侧接的所有非

规范碱基对是G:U碱基对。在上述实施方案的变体中,对于一些或所有的子区,侧接非规范碱基对中的一个或两个在有义序列、反义序列或在两个序列中,被非碱基配对的核糖核苷酸替代。容易理解,在上述实施方案中,子区的最大数目由RNA分子中dsRNA区的长度决定。

[0367] 在一个实施方案中,dsRNA区包含10个邻接碱基对的子区,其中2-4个碱基对是非规范碱基对。在一个实施方案中,dsRNA区包含2个子区,每个子区为10个邻接碱基对,其中10个邻接碱基对中的2-4个是非规范碱基对。在一个实施方案中,dsRNA区包含3个子区,每个都是10个邻接碱基对,其中10个邻接碱基对中的2-4个是非规范碱基对。在一个实施方案中,dsRNA区包含4个子区,每个都是10个邻接碱基对,其中10个邻接碱基对中的2-4个是非规范碱基对。在一个实施方案中,dsRNA区包含5个子区,每个都是10个邻接碱基对,其中10个邻接碱基对中的2-4个是非规范碱基对。在一个实施方案中,dsRNA区包含10个子区,每个都是10个邻接碱基对,其中10个邻接碱基对中的2-4个是非规范碱基对。在一个实施方案中,dsRNA区包含4个子区,每个都是15个邻接碱基对,其中15个邻接碱基对中的2-6个是非规范碱基对。在一个实施方案中,dsRNA区包含2-50个子区,每个子区包含10个邻接碱基对,其中10个邻接碱基对中的2-4个是非规范碱基对。在一个实施方案中,dsRNA区包含2-40个子区,每个子区包含10个邻接碱基对,其中10个邻接碱基对中的2-4个是非规范碱基对。在一个实施方案中,dsRNA区包含2-30个子区,每个子区包含10个邻接碱基对,其中10个邻接碱基对中的2-4个是非规范碱基对。在一个实施方案中,dsRNA区包含2-20个子区,每个子区包含10个邻接碱基对,其中10个邻接碱基对中的2-4个是非规范碱基对。优选地,在上述实施方案中,非规范碱基对包含至少一个G:U碱基对,更优选子区中的所有非规范碱基对为G:U碱基对。在上述实施方案的变体中,对于一些或所有的子区,2-4个或2-6个非规范碱基对中的一个或多个在有义序列、反义序列或在两个序列中,被非碱基配对的核糖核苷酸替代。容易理解,在上述实施方案中,子区的最大数目由RNA分子中dsRNA区的长度决定。

[0368] 在一个实施方案中,dsRNA区中规范与非规范碱基对的比率在2.5:1和3.5:1之间,例如约3:1。在一个实施方案中,dsRNA区中规范与非规范碱基对的比率在3.5:1和4.5:1之间,例如约4:1。在一个实施方案中,dsRNA区中规范与非规范碱基对的比率在4.5:1和5.5:1之间,例如约5:1。在一个实施方案中,dsRNA区中规范与非规范碱基对的比率在5.5:1和6.5:1之间,例如约6:1。RNA分子中不同的dsRNA区可以具有不同的比例。

[0369] 在上述实施方案中,RNA分子的dsRNA区中的非规范碱基对优选都是G:U碱基对。在一个实施方案中,至少99%的非规范碱基对是G:U碱基对。在一个实施方案中,至少98%的非规范碱基对是G:U碱基对。在一个实施方案中,至少97%的非规范碱基对是G:U碱基对。在一个实施方案中,至少95%的非规范碱基对是G:U碱基对。在一个实施方案中,至少90%的非规范碱基对是G:U碱基对。在一个实施方案中,90-95%的非规范碱基对是G:U碱基对。例如,如果存在10个非规范碱基对,则至少9个(90%)为G:U碱基对。

[0370] 包含非规范碱基配对的dsRNA区包含20个邻接核苷酸的反义序列,其充当反义调节元件。在一个实施方案中,所述反义调节元件与真核细胞中的靶RNA分子至少80%、优选至少90%、更优选至少95%或最优选100%互补。在一个实施方案中,dsRNA区包含2、3、4或5个与同一靶RNA分子互补(即与同一靶RNA分子的不同区互补)或与不同靶RNA分子互补的反义调节元件。

[0371] 在一个实施方案中,当有义和反义序列杂交时,所述有义核糖核苷酸序列的一个

或多个核糖核苷酸或所述反义核糖核苷酸序列的一个或多个核糖核苷酸或两者在dsRNA区中不被碱基配对。在该实施方案中,dsRNA区不包括共价连接有义和反义序列的任何环序列。dsRNA区或子区的一个或多个核糖核苷酸可以不被碱基配对。因此,在该实施方案中,dsRNA区的有义链与其相应的反义链不完全碱基配对。

[0372] 此外,在一个实施方案中并且任选地与上述实施方案的任何特征组合,有义序列中的核糖核苷酸总数和反义序列中的核糖核苷酸总数可以不相同,尽管优选它们是相同的。在一个实施方案中,dsRNA区的有义核糖核苷酸序列中的核糖核苷酸总数为反义核糖核苷酸序列中的核糖核苷酸总数的90%-110%。在一个实施方案中,有义核糖核苷酸序列中的核糖核苷酸总数为反义核糖核苷酸序列中的核糖核苷酸总数的95%-105%。在一个实施方案中,本发明的嵌合RNA分子可以包含一个或多个结构元件,例如内部或末端隆起部分或环。以上讨论了隆起和环的各种实施例。在一个实施方案中,dsRNA区由诸如隆起或环的结构元件分隔开。在一个实施方案中,dsRNA区由间插(间隔)序列分开。间隔序列的一些核糖核苷酸可以与RNA分子中的其他核糖核苷酸碱基配对,例如与间隔序列中的其他核糖核苷酸碱基配对,或者它们在RNA分子中可以不碱基配对,或者它们各自中的一些不碱基配对。在一个实施方案中,dsRNA区与末端环连接。在一个实施方案中,dsRNA区侧接末端环。

[0373] 在一个实施方案中,当本发明RNA分子的dsRNA区在5个邻接碱基对的任何子区中具有至少3个非规范碱基对时,非规范碱基对不邻接,而是由一个或多个规范碱基对隔开,即dsRNA区不具有3个或更多个邻接的非规范碱基对。在一个实施方案中,dsRNA区不具有4个或更多个邻接的非规范碱基对。例如,在一个实施方案中,dsRNA区在10个碱基对的子区中包含至少3个非规范碱基对,其中每个非规范碱基对被4个规范碱基对分开。

[0374] 在一个实施方案中,本发明的RNA分子包含多于一个dsRNA区。例如,所述RNA分子包含2、3、4、5、6、7、8、9、10或更多个dsRNA区。在该实施例中,dsRNA区中的一个或多个或全部可以包含以上例示的性质,例如非规范碱基配对和/或若干反义调节元件。

[0375] 沉默活性

[0376] 本发明的RNA分子具有反义活性,因为它们包含与靶RNA分子的区基本上互补的有义核糖核苷酸序列。例如,核糖核苷酸序列与真核细胞中靶RNA分子的区基本上互补。在一个实施例中,靶RNA分子可以在细菌细胞、真菌细胞、植物细胞、昆虫细胞或动物细胞中。本文定义的RNA分子的此类组分可称为“反义调节元件”。“基本上互补”意指有义核糖核苷酸序列与真核细胞中靶RNA分子的互补序列相比可以具有插入、缺失和单个点突变。优选地,具有反义活性的有义核糖核苷酸序列与靶RNA分子之间的同源性为至少80%、优选至少90%、优选至少95%、最优选100%。例如,所述有义核糖核苷酸序列可包含序列与真核细胞中靶RNA分子的第一区相同的约15、约16、约17、约18、约19或更多个邻接核苷酸。在另一个实施例中,有义核糖核苷酸序列可包含序列与真核细胞中靶RNA分子的第一区相同的约20个邻接核苷酸。

[0377] “反义活性”在本发明的上下文中用于指来自本文定义的RNA分子的反义调节元件,其调节(增加或减少)靶RNA分子的表达。

[0378] 在各种实施例中,根据本发明的反义调节元件可包含通过连接基团连接在一起的多个单体亚基。实施例包括引物、探针、反义化合物、反义寡核苷酸、外部指导序列(EGS)寡核苷酸、交替剪接器、gapmers、siRNA和微RNA。因此,根据本发明的RNA分子可包含具有单

链、双链、环状、分支或发夹结构的反义调节元件。在一个实施例中,反义序列可以含有结构元件,例如内部或末端隆起部分或环。

[0379] 在一个实施例中,本发明的RNA分子包含嵌合寡聚组分,如嵌合寡核苷酸。例如,RNA分子可以包含不同修饰的核苷酸、混合骨架反义寡核苷酸或其组合。在一个实施例中,嵌合寡聚化合物可包含至少一个经修饰的区,以赋予对核酸酶降解的抗性增加,细胞摄取增加和/或对靶RNA分子的结合亲和力增加。

[0380] 反义调节元件可以具有多种长度。在各种实施例中,本发明提供了由X-Y连接的碱基组成的反义调节元件,其中X和Y各自独立地选自8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49和50(条件是X<Y)。例如,在某些实施方案中,本发明提供了反义调节元件,其包含8-9、8-10、8-11、8-12、8-13、8-14、8-15、8-16、8-17、8-18、8-19、8-20、8-21、8-22、8-23、8-24、8-25、8-26、8-27、8-28、8-29、8-30、9-10、9-11、9-12、9-13、9-14、9-15、9-16、9-17、9-18、9-19、9-20、9-21、9-22、9-23、9-24、9-25、9-26、9-27、9-28、9-29、9-30、10-11、10-12、10-13、10-14、10-15、10-16、10-17、10-18、10-19、10-20、10-21、10-22、10-23、10-24、10-25、10-26、10-27、10-28、10-29、10-30、11-12、11-13、11-14、11-15、11-16、11-17、11-18、11-19、11-20、11-21、11-22、11-23、11-24、11-25、11-26、11-27、11-28、11-29、11-30、12-13、12-14、12-15、12-16、12-17、12-18、12-19、12-20、12-21、12-22、12-23、12-24、12-25、12-26、12-27、12-28、12-29、12-30、13-14、13-15、13-16、13-17、13-18、13-19、13-20、13-21、13-22、13-23、13-24、13-25、13-26、13-27、13-28、13-29、13-30、14-15、14-16、14-17、14-18、14-19、14-20、14-21、14-22、14-23、14-24、14-25、14-26、14-27、14-28、14-29、14-30、15-16、15-17、15-18、15-19、15-20、15-21、15-22、15-23、15-24、15-25、15-26、15-27、15-28、15-29、15-30、16-17、16-18、16-19、16-20、16-21、16-22、16-23、16-24、16-25、16-26、16-27、16-28、16-29、16-30、17-18、17-19、17-20、17-21、17-22、17-23、17-24、17-25、17-26、17-27、17-28、17-29、17-30、18-19、18-20、18-21、18-22、18-23、18-24、18-25、18-26、18-27、18-28、18-29、18-30、19-20、19-21、19-22、19-23、19-24、19-25、19-26、19-29、19-28、19-29、19-30、20-21、20-22、20-23、20-24、20-25、20-26、20-27、20-28、20-29、20-30、21-22、21-23、21-24、21-25、21-26、21-27、21-28、21-29、21-30、22-23、22-24、22-25、22-26、22-27、22-28、22-29、22-30、23-24、23-25、23-26、23-27、23-28、23-29、23-30、24-25、24-26、24-27、24-28、24-29、24-30、25-26、25-27、25-28、25-29、25-30、26-27、26-28、26-29、26-30、27-28、27-29、27-30、28-29、28-30,或29-30个连接的碱基。

[0381] 根据本发明的RNA分子可以包含多个反义调节元件。例如,RNA分子可包含至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个反义调节元件。在一个实施例中,反义调节元件是相同的。在该实施例中,RNA分子可包含反义调节元件的至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个拷贝。在另一个实施例中,根据本发明的RNA分子可以包含不同的反义调节元件。例如,可以提供反义调节元件以靶向诸如脂质生物合成的途径中的多个基因。在该实施例中,RNA分子可包含至少2个,至少3个,至少4个,至少5个,至少6个,至少7个,至少8个,至少9个,至少10个不同的反义调节元件。

[0382] 根据本发明的反义调节元件可以调节(增加或减少)各种靶RNA分子的表达或量。

在一个实施例中,靶RNA分子是脂肪酸生物合成基因。这类基因的实施例包括编码下述的基因:乙酰转酰酶;酰基转运蛋白(例如“酰基载体蛋白”);去饱和酶,诸如硬脂酰基去饱和酶或微粒体D12-去饱和酶,特别是Fad2-1基因;丙二酰转酰酶;-酮脂酰-ACP合酶;3-酮-ACP还原酶;烯酰基-ACP水合酶;硫酯酶,如酰基-ACP硫酯酶;烯酰基-ACP还原酶。在一个实施例中,靶RNA分子是FAD2基因(例如由Genbank登录号:描述的那些基因)。未发布:AF124360(埃塞俄比亚芥(Brassica carinata))、AF042841(白菜型油菜(Brassica rapa))、L26296(拟南芥(Arabidopsis thaliana))、A65102(欧洲油菜)。例如,靶RNA分子可以是FAD2.1基因。在另一个实施例中,靶RNA分子可以是FAD2.2基因。在另一个实施例中,靶RNA分子可以是FAD2.1和FAD2.2基因。参与修饰可以是靶RNA分子的脂质合物的其他基因的实施例是本领域已知的(Shure et al.(1983)Cell35:225-233;Preiss et al.(1987)Tailoring Genes for Crop Improvement(Bruening et al.,eds.),Plenum Press,S.133-152;Gupta et al.(1988)Plant Mol.Biol.10:215-224;Olive et al.(1989)Plant Mol Biol 12:525-538;Bhattacharyya et al.(1990)Cell 60:155-122;Dunwell J M(2000)J Exp Botany 51Spec No:487-96;Brar D S et al.(1996)Biotech Genet.Eng Rev 13:167-79;Kishore G M and Somerville C R(1993)Curr Opin Biotech4(2):152-8;US 5,530,192和WO 94/18337).

[0383] 在另一个实施例中,靶RNA分子是节肢动物基因,如昆虫基因转录物。这样的基因的实施例包括几丁质合酶基因,例如CHS1和/或CHS2或控制昆虫活性、行为、繁殖、生长和/或发育的其他基因。各种病原体的各种必需基因是本领域技术人员已知的(例如,在WO 93/10251、WO 94/17194中总结了线虫抗性基因)。

[0384] 在另一个实施例中,靶RNA分子与疾病相关。例如,靶RNA分子可以是癌基因或肿瘤抑制基因转录物。示例性癌基因包括ABL1、BCL1、BCL2、BCL6、CBFA2、CBL、CSF1R、ERBA、ERBB、EBRB2、FGR、FOS、FYN、HRAS、JUN、LCK、LYN、MYB、MYC、NRAS、RET或SRC。示例性肿瘤抑制基因包括BRCA1或BRCA2;粘附分子;细胞周期蛋白激酶及其抑制剂。

[0385] 在另一个实施例中,靶RNA分子与果实成熟的延迟相关。延迟果实成熟可以例如通过降低选自下述基因的基因表达来实现,所述基因选自由以下组成的组:聚半乳糖醛酸酶、果胶酯酶,β-(1-4)-葡聚糖酶(纤维素酶)、半乳糖苷酶(β-半乳糖苷酶);乙烯生物合成基因,诸如1-氨基环丙烷-1-羧酸合酶;类胡萝卜素生物合成的基因,例如举例而言,前八氢番茄红素或八氢番茄红素的生物合成的基因。

[0386] 在另一个实施例中,靶RNA分子与衰老症状的延迟相关。合适的靶RNA分子包括肉桂酰基-CoA:NADPH还原酶或肉桂酰基醇脱氢酶。在WO1995/07993中进一步描述了靶RNA分子。

[0387] 在另一个实施例中,靶RNA分子与食品、优选种子中的纤维含量的修饰相关。例如,所述RNA分子可以降低咖啡酸O-甲基转移酶或肉桂酰基醇脱氢酶的表达。

[0388] 编码RNA分子的核酸

[0389] 本领域技术人员从前面的描述将理解,本发明还提供了编码本文公开的RNA分子的分离的核酸及其组成部分。例如,包含SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9中任何一个或多个所示序列的核酸。所述核酸可以在宿主细胞中表达后被部分纯化。术语“部分纯化的”

用于指通常已与脂质、核酸、其他肽和其他污染分子分离的RNA分子,所述RNA分子在宿主细胞中与所述脂质、核酸、其他肽和其他污染分子缔合。优选地,部分纯化的多核苷酸至少60%不含、更优选至少75%不含、更优选至少90%不含与它相关的其他组分。

[0390] 在另一个实施例中,根据本发明的多核苷酸是异源多核苷酸。术语"异源多核苷酸"是本领域熟知的,并且是指不是细胞内源的多核苷酸,或者是天然序列已经改变的天然多核苷酸,或者是由于通过重组DNA技术操纵细胞而定量改变其表达的天然多肽。

[0391] 在另一个实施例中,根据本发明的多核苷酸是合成多核苷酸。例如,可以使用不需要预先存在的核酸序列的技术,如DNA印刷和寡核苷酸合成,来生产多核苷酸。在另一个实施例中,多核苷酸由异种核酸产生。

[0392] 在一个实施例中,本文公开的多核苷酸编码RNA前体分子,所述RNA前体分子在至少一个环序列中包含内含子,所述内含子能够在多核苷酸在宿主细胞中转录期间剪接出来。在另一个实施例中,环序列包含两个、三个、四个、五个或更多个内含子。本发明还提供了表达构建体,例如包含与启动子可操作地连接的本发明的分离的核酸的DNA构建体。在一个实施例中,此类分离的核酸和/或表达构建体在细胞或非人类生物体中提供。在一个实施例中,分离的核酸稳定地整合到细胞或非人生物体的基因组中。下面讨论合适的表达构建体、启动子和包含它们的细胞的各种实施例。

[0393] 根据本发明的RNA分子的合成可以使用本领域已知的各种方法实现。实施例部分提供了体外合成的实施例。在该实施例中,包含本文公开的RNA分子的构建体经受在3'端限制,沉淀,纯化和定量。在转化HT115电感受态细胞并使用T7、IPTG系统诱导RNA合成后,可以在细菌培养物中实现RNA合成。

[0394] 重组载体

[0395] 本发明的一个实施方案包括重组载体,其包含至少一种本文定义的RNA分子并且能够将所述RNA分子递送至宿主细胞中。重组载体包括表达载体。重组载体含有异源多核苷酸序列,即与本文定义的RNA分子天然不相邻的多核苷酸序列,其优选衍生自不同物种。载体可以是RNA或DNA,并且通常是衍生自病毒的病毒载体或质粒。

[0396] 各种病毒载体可用于递送和介导根据本发明的RNA分子的表达。病毒载体的选择通常取决于各种参数,例如,用于递送的细胞或组织,载体的转导效率和致病性。在一个实施例中,病毒载体整合到宿主细胞染色质中(例如慢病毒)。在另一个实施例中,病毒载体主要作为染色体外附加体(例如腺病毒)保留在细胞核中。这些类型的病毒载体的实施例包括肿瘤逆转录病毒、慢病毒、腺相关病毒、腺病毒、疱疹病毒和逆转录病毒。

[0397] 质粒载体通常包括提供表达盒在原核细胞中容易选择、扩增和转化的另外的核酸序列,例如pUC衍生的载体、pGEM衍生的载体或含有一个或多个T-DNA区的二元载体。另外的核酸序列包括提供载体自主复制的复制起点,优选编码抗生素或除草剂抗性的选择标记基因,提供多个位点以插入核酸序列或在核酸构建体中编码的基因的独特多克隆位点,以及增强原核和真核(特别是植物)细胞的转化的序列。

[0398] 如本文所用,"可操作连接"是指两个或更多个核酸(例如,DNA)片段之间的功能关系。通常,它是指转录调节元件(启动子)与转录序列的功能关系。例如,如果启动子在合适的细胞中刺激或调节编码序列的转录,则启动子与本文定义的RNA分子的编码序列可操作地连接。通常,与转录序列可操作连接的启动子转录调控元件与转录序列物理邻接,即它们

是顺式作用的。然而,一些诸如增强子的转录调节元件不需要物理邻接或位于其转录增强的编码序列附近。

[0399] 当存在多个启动子时,每个启动子可以独立地相同或不同。

[0400] 为了便于鉴定转化体,重组载体期望地包含可选择的或可筛选的标记基因。“标记基因”是指赋予表达该标记基因的细胞不同表型的基因,因此允许将这种转化的细胞与不具有该标记的细胞区分开。可选择标记基因赋予一种性状,人们可以基于对选择剂(例如,除草剂、抗生素等)的抗性来“选择”该性状。可筛选标记基因(或报道基因)赋予可通过观察或测试鉴定的性状,即,通过“筛选”(例如,β-葡糖醛酸糖苷酶、荧光素酶、GFP或未转化细胞中不存在的其他酶活性)。用于选择植物转化体的示例性选择标记包括但不限于编码潮霉素B抗性的hyg基因;赋予卡那霉素、巴龙霉素抗性的新霉素磷酸转移酶(nptII)基因;来自大鼠肝脏的谷胱甘肽-S-转移酶基因,其赋予对谷胱甘肽衍生的除草剂的抗性,如例如EP256223中所述;谷氨酰胺合酶基因,其在过表达时赋予对诸如膦丝菌素的谷氨酰胺合酶抑制剂的抗性,如WO 87/05327中所述;来自产绿色链霉菌(Streptomyces viridopromogenes)的乙酰转移酶基因,其赋予对选择性试剂膦丝菌素的抗性,如例如EP275957中所述;编码5-烯醇异硫氰酸酯-3-磷酸合酶(EPSPS)的基因,其赋予对N-膦酰甲基甘氨酸的耐受性,如例如Hinchee等人所述(1988);bar基因,其赋予对双丙氨膦的抗性,例如WO91/02071中所述;腈水解酶基因,例如来自克雷白氏杆菌(Klebsiella ozaenae)的bxn,其赋予对溴苯腈的抗性(Stalker等人,1988);二氢叶酸还原酶(DHFR)基因,其赋予对甲氨蝶呤的抗性(Thillet等人,1988);突变乙酰乳酸合酶基因(ALS),其赋予对咪唑啉酮,磺酰脲或其他ALS抑制化学品的抗性(参见EP154,204);突变的邻氨基苯甲酸盐(酯)合酶基因,其赋予对5-甲基色氨酸的抗性;或茅草枯(dalapon)脱卤酶基因,其赋予除草剂抗性。

[0401] 优选地,所述重组载体稳定整合到细胞诸如植物细胞的基因组中。因此,重组载体可包含允许载体掺入基因组或细胞染色体的适当元件。

[0402] 表达载体

[0403] 如本文所用,“表达载体”是能够转化宿主细胞并影响本文定义的RNA分子表达的DNA载体。本发明的表达载体包含调节序列,例如转录控制序列、翻译控制序列、复制起点,以及与宿主细胞相容并且控制根据本发明的RNA分子的表达的其他调节序列。特别地,本发明的表达载体包括转录控制序列。转录控制序列是控制转录的起始、延伸和终止的序列。特别重要的转录控制序列是控制转录起始的那些序列,例如启动子、增强子、操纵子和阻遏子序列。所使用的调节序列的选择取决于靶生物体,诸如所关注的植物和/或靶器官或组织。这样的调节序列可以获自任何真核生物,例如植物或植物病毒,或者可以是化学合成的。

[0404] 适用于稳定转染植物细胞或建立转基因植物的示例性载体已描述于例如,Pouwels等,Cloning Vectors(克隆载体):A Laboratory Manual(实验室手册),1985,supp.1987,Weissbach和Weissbach,Methods for Plant Molecular Biology(植物分子生物的方法),Academic Press(美国学术出版社),1989,以及Gelvin等.,Plant Molecular Biology Manual(植物分子生物学手册),Kluwer Academic Publishers(克吕韦尔学术出版社),1990。典型地,植物表达载体包括例如在5’和3’调节序列和显性选择标记的转录控制下的一个或多个克隆的植物基因。这样的植物表达载体还可以包含启动子调节区(例如,调控可诱导的或组成型的、环境或发育调控的,或细胞或组织特异性表达的调节区)、转录

起始开始位点、核糖体结合位点、转录终止位点,和/或聚腺苷酸化信号。

[0405] 本发明的载体还可用于在无细胞表达系统中产生本文定义的RNA分子,这种系统是本领域熟知的。

[0406] 在一个实施例中,编码根据本发明的RNA分子的多核苷酸可操作地连接至能够在宿主细胞中直接表达RNA分子的启动子。在一个实施例中,启动子在体外起作用。在一个实施例中,启动子是RNA聚合酶启动子。例如,启动子可以是RNA聚合酶III启动子。在另一个实施例中,启动子可以是RNA聚合酶II启动子。然而,启动子的选择可取决于靶生物体,诸如所关注的植物、昆虫和/或组织。示例性的哺乳动物启动子包括CMV,EF1α,SV40,PGK1,Ubc,人β肌动蛋白,CAG,TRE,UAS,CaMKIIa,CAL1,10,TEF1,GDS,ADH1,CaMV35S,Ubi,H1和U6。示例性昆虫启动子包括Ac5和多角体蛋白。已经描述了许多在植物细胞中有活性的组成型启动子。用于在植物中组成型表达的合适启动子包括但不限于花椰菜花叶病毒(CaMV)35S启动子、玄参花叶病毒(FMV)35S、来自核酮糖-1,5-二磷酸羧化酶的小亚基(SSU)的光诱导启动子、水稻细胞质三糖磷酸异构酶启动子、拟南芥(Arabidopsis)腺嘌呤磷酸核糖转移酶启动子、水稻肌动蛋白1基因启动子、甘露氨酸合酶和章鱼氨酸合酶启动子、Adh启动子、蔗糖合酶启动子、R基因复合物启动子、叶绿素α/α结合蛋白基因启动子。这些启动子已经用于产生已经在植物中表达的DNA载体,参见例如WO 84/02913。所有这些启动子已经用于产生各种类型的植物可表达的重组DNA载体。

[0407] 为了在植物的源组织如叶、种子、根或茎中表达,优选用于本发明的启动子在这些特定组织中具有相对高的表达。为此,可以从许多启动子中选择具有组织或细胞特异性或增强表达的基因。文献中报道的此类启动子的实施例包括来自豌豆的叶绿体谷氨酰胺合酶GS2启动子、来自小麦的叶绿体果糖-1,6-二磷酸酶启动子、来自马铃薯的核光合ST-LSI启动子、来自拟南芥(Arabidopsis thaliana)的丝氨酸/苏氨酸激酶启动子和葡糖淀粉酶(CHS)启动子。还报道了来自东方落叶松(Larix laricina)的核酮糖-1,5-二磷酸羧化酶启动子,来自松树的Cab基因、Cab6的启动子,来自小麦的Cab-1基因的启动子,来自菠菜的Cab-1基因的启动子,来自水稻的Cab 1R基因的启动子,来自玉蜀黍(Zea mays)的丙酮酸磷酸双激酶(PPDK)的启动子,烟草Lhcb1*2基因的启动子,拟南芥(Arabidopsis thaliana)Suc2蔗糖-H^{30}同向转运蛋白启动子,菠菜的类囊体膜蛋白基因(PsaD、PsaF、PsaE、PC、FNR、AtpC、AtpD、Cab、RbcS)的启动子。在本发明中还可以使用叶绿素α/β-结合蛋白的其他启动子,例如来自白芥(Sinapis alba)的LhcB基因和PsbP基因的启动子。

[0408] 响应于环境、激素、化学和/或发育信号调节的多种植物基因启动子也可用于在植物细胞中表达RNA结合蛋白基因,包括通过(1)热;(2)光(例如,豌豆RbcS-3A启动子,玉米RbcS启动子);(3)激素,例如脱落酸;(4)损伤(例如WunI);或(5)化学品,例如茉莉酮酸甲酯、水杨酸、类固醇激素、醇、

[0409] (Safeners)(WO 97/06269);或其也可有利地使用(6)器官特异性启动子。

[0410] 如本文所用,术语“植物贮藏器官特异性启动子”是指当与其他植物组织相比时优先指导植物贮藏器官中的基因转录的启动子。为了在植物的库组织,如马铃薯植物的块茎,番茄的果实,或大豆、油菜、棉花、玉蜀黍、小麦、水稻和大麦的种子中表达,优选本发明中使用的启动子在这些特定组织中具有相对高的表达。可使用β-伴大豆球蛋白启动子,或其他种子特异性启动子,诸如油菜籽蛋白、玉米醇溶蛋白、linin和菜豆蛋白启动子。也可以使用

根特异性启动子。这种启动子的实施例是酸性壳多糖酶基因的启动子。在根组织中的表达也可以通过利用已经鉴定的CaMV 35S启动子的根特异性亚结构域来实现。

[0411] 在一个特别优选的实施方案中,启动子指导发生脂质生物合成的组织和器官中的表达。这类启动子可在合适的时间在种子发育中起作用以修饰种子中的脂质组成。用于种子特异性表达的优选启动子包括:1)来自编码参与脂质生物合成和在种子中积累的酶(诸如去饱和酶和延伸酶)的基因的启动子,2)来自编码种子贮藏蛋白的基因的启动子,和3)来自编码参与碳水化合物生物合成和在种子中积累的酶的基因的启动子。合适的种子特异性启动子是,油菜的油菜籽蛋白基因启动子(US 5,608,152),蚕豆(Vicia faba)USP启动子(Baumlein等人,1991),拟南芥(Arabidopsis)油质蛋白启动子(WO 98/45461),菜豆(WO 98/45461)的菜豆蛋白启动子(US 5,504,200),芸苔属(Brassica)Bce4启动子(WO 91/13980)或豆球蛋白B4启动子(Baumlein等人,1992)和导致种子在单子叶植物如玉米、大麦、小麦、黑麦、水稻等中特异性表达的启动子。合适的值得注意的启动子是大麦lpt2或lpt1基因启动子(WO 95/15389和WO 95/23230),或WO 99/16890中描述的启动子(来自大麦醇溶蛋白基因、水稻谷蛋白基因、水稻oryzin基因、水稻醇溶蛋白基因、小麦醇溶蛋白基因、小麦谷蛋白基因、玉米醇溶蛋白基因、燕麦谷蛋白基因、高粱kasirin基因,黑麦的黑麦碱(secalin)基因的启动子)。其他启动子包括Broun等,(1998),Potenza等,(2004),US20070192902和US20030159173中描述的那些。在一个实施方案中,种子特异性启动子优选在种子的限定部分如子叶或胚乳中表达。子叶特异性启动子的实施例包括但不限于FPI启动子(Ellerstrom et al.,1996),豌豆豆球蛋白启动子(Perrin等,2000年);和菜豆植物凝集素启动子(Perrin等,2000年)。胚乳特异性启动子的实施例包括但不限于玉米玉米醇溶蛋白-1启动子(Chikwamba等,2003),水稻谷蛋白-1启动子(Yang等,2003);大麦D-大麦醇溶蛋白启动子(Horvath等,2000);和小麦HMW麦谷蛋白启动子(Alvarez等,2000)。在另一个实施方案中,种子特异性启动子在胚芽和/或种子发芽后不表达或仅以低水平表达。

[0412] 在另一个实施方案中,所述植物贮藏器官特异性启动子是果实特异性启动子。实施例包括但不限于番茄多聚半乳糖醛酸酶E8和Pds启动子,以及苹果ACC氧化酶启动子(综述参见Potenza等,2004)。在一个优选的实施方案中,相对于果实的果皮或果实内的种子,启动子优先指导在例如果实的果髓的果实可食用部分中的表达。

[0413] 在一个实施方案中,诱导型启动子是构巢曲霉(Aspergillus nidulans)alc系统。在Padidam(2003)和Corrado和Karali(2009)的综述中描述了可用于代替构巢曲霉alc系统的诱导型表达系统的实施例。在另一个实施方案中,诱导型启动子是安全剂诱导型启动子,例如玉米ln2-1或ln2-2启动子(Hershey和Stoner,1991),安全剂诱导型启动子是玉米GST-27启动子(Jepson等,1994)或大豆GH2/4启动子(Ulmasov等,1995)。

[0414] 在另一个实施方案中,诱导型启动子是衰老诱导型启动子,例如来自拟南芥(Arabidopsis)的衰老诱导型启动子SAG(衰老相关基因)12和SAG13(Gan,1995;Gan和Amasino,1995)和来自欧洲油菜(Brassica napus)的LSC54(Buchanan-Wollaston,1994)。这样的启动子在大约植物组织、特别是叶子衰老开始时显示增加的表达。

[0415] 为了在营养组织中表达,可以使用叶特异性启动子,例如核酮糖二磷酸羧化酶(RBCS)启动子。例如,在叶片和轻生长的幼苗中表达番茄RBCS1、RBCS2和RBCS3A基因(Meier等,1997)。可以使用Matsuoka等(1994)描述的核酮糖二磷酸羧化酶启动子,其几乎只在叶

片和叶鞘的叶肉细胞中高水平表达。另一种叶特异性启动子是集光叶绿素a/b结合蛋白基因启动子(参见Shiina等,1997)。Li等(1996)描述的拟南芥(Arabidopsis thaliana)myb相关基因启动子(Atmby5),是叶特异性的。Atmyb5启动子在幼龄莲座和花椰菜叶片边缘发育的叶毛、托叶和表皮细胞以及未成熟种子中表达。由Busk等(1997)在玉米中鉴定的叶启动子,也可以使用。

[0416] 在一些情况下,例如当LEC2或BBM重组表达时,可能需要转基因不以高水平表达。可以在这种情况下使用的启动子的实施例是保留种子特异性表达模式但具有降低的表达水平的截短的油菜籽蛋白A启动子(Tan等,2011)。

[0417] 该5'非翻译前导序列可以衍生自经选择以表达本发明的RNA分子的异源基因序列的启动子,或可以相对于有待产生的酶的编码区是异源的,并且如果希望的话可以进行特异性修饰以增加mRNA的翻译。关于优化转基因表达的综述,参见Koziel等,(1996)。5'非翻译区也可从合适的真核基因、植物基因(小麦和玉米叶绿素a/b结合蛋白基因前导序列)或合成基因序列的植物病毒RNA(烟草花叶病毒、烟草蚀纹病毒、玉米矮花叶病毒、苜蓿花叶病毒等)获得。本发明不限于其中非翻译区衍生自伴随启动子序列的5'非翻译序列的构建体。前导序列也可以衍生自不相关的启动子或编码序列。在本发明上下文中有用的前导序列包括玉米Hsp70前导序列(US 5,362,865和US 5,859,347)和TMV ω元件。

[0418] 转录的终止通过在表达载体中可操作地连接至所关注RNA分子的3'非翻译DNA序列实现。重组DNA分子的3'非翻译区含有聚腺苷酸化信号,其在植物中起作用以引起将腺苷酸核苷酸添加至RNA的3'端。3'非翻译区可以从在植物细胞中表达的各种基因获得。胭脂碱合酶3'非翻译区、豌豆小亚基Rubisco基因的3'非翻译区、大豆7S种子贮藏蛋白基因的3'非翻译区通常以这种能力使用。含有农杆菌(Agrobacterium)肿瘤诱导(Ti)质粒基因的聚腺苷酸信号的3'转录、非翻译区也是合适的。

[0419] 转移核酸

[0420] 转移核酸可用于将外源多核苷酸递送至细胞并包含一个、优选两个边缘序列和一个或多个所关注的RNA分子。转移核酸可以编码或不编码可选择标记。优选地,转移核酸形成细菌中二元载体的一部分,其中二元载体进一步包含允许载体在细菌中复制,选择或维持含有二元载体的细菌细胞的元件。在转移到真核细胞后,二元载体的转移核酸组分能够整合到真核细胞的基因组中,或者对于瞬时表达实验,仅能够在细胞中表达。

[0421] 如本文所用,术语"染色体外转移核酸"是指能够从细菌(诸如农杆菌属(Agrobacterium)物种)转移至诸如植物叶细胞的真核细胞的核酸分子。染色体外转移核酸是熟知为能够被转移的元件的遗传元件,随后将包含在其边缘内的核苷酸序列整合到受体细胞的基因组中。在这方面,转移核酸通常侧接两个"边缘"序列,尽管在一些情况下可以使用在一端的单个边缘并且在转移过程中随机产生转移核酸的第二端。所关注的RNA分子通常位于转移核酸的左边缘样序列和右边缘样序列之间。包含在转移核酸内的RNA分子可以与多种不同的启动子和终止子调节元件可操作地连接,所述调节元件促进其表达,即RNA分子的转录和/或RNA分子的翻译。转移来自农杆菌属(Agrobacterium)物种的DNA(T-DNAs),诸如根癌农杆菌(Agrobacterium tumefaciens)或发根农杆菌(Agrobacterium rhizogenes)及其人工变体/突变体可能是转移核酸的最佳特征实施例。另一个实施例是P-DNA("植物-DNA"),其包含来自植物的T-DNA边缘样序列。

[0422] 如本文所用,“T-DNA”是指根癌农杆菌(Agrobacterium tumefaciens)Ti质粒或发根农杆菌(Agrobacterium rhizogenes)Ri质粒的T-DNA,或其用于将DNA转移到植物细胞中的变体。T-DNA可以包含完整的T-DNA,包括右和左边界序列,但只需要包含顺式转移所需的最小序列,即右T-DNA边界序列。本发明的T-DNA已插入到它们中,在右和左边界序列(如果存在的话)之间的任何地方插入所关注的RNA分子。编码将T-DNA转移到植物细胞中所需的反式因子(如vir基因)的序列可以插入到T-DNA中,或者可以存在于与T-DNA相同的复制子上,或者优选地在农杆菌(Agrobacterium)宿主中的相容复制子上是反式的。这样的“二元载体系统”是本领域熟知的。如本文所用,“P-DNA”是指分离自植物基因组或其人造变体/突变体的转移核酸,并且在每个末端或仅在一个末端包含T-DNA边缘样序列。

[0423] 如本文所用,转移核酸的“边缘”序列可以分离自所选生物体,例如植物或细菌,或者是其人造变体/突变体。所述边缘序列促进并利于与其连接的RNA分子的转移,并且可以促进其在受体细胞基因组中的整合。在一个实施方案中,边界序列的长度为10-80bp。农杆菌属(Agrobacterium)物种T-DNA边缘序列是本领域熟知的并且包括描述于Lacroix等(2008)的那些。

[0424] 而传统上仅仅是农杆菌属(Agrobacterium)物种已经用于将基因转移到植物细胞中,现在已经鉴定/开发了大量的系统,它们以与农杆菌属(Agrobacterium)物种相似的方式起作用。最近已经遗传修饰了几种非农杆菌属(Agrobacterium)物种以具有基因转移的能力(Chung等,2006;Broothaerts等,2005)。这些包括根瘤菌属(Rhizobium)物种NGR234,苜蓿中华根瘤菌(Sinorhizobium meliloti)和百脉根中慢生根瘤菌(Mezorhizobium loti)。

[0425] 通过哺乳动物细胞和携带质粒的大肠杆菌的原生质体的融合,在数十年前首次实现了真核表达质粒从细菌到真核宿主的直接转移(Schaffner,1980)。自此,能够将基因递送到哺乳动物细胞中的细菌的数量已经稳定地增加(Weiss,2003),由四组独立地发现(Sizemore等,1995;Courvalin等,1995;Powell等,1996;Darji等,1997)。

[0426] 如本文所用,术语“转染”、“转化”及其变化形式通常可互换使用。“转染的”或“转化的”细胞可以已经被操作以导入所关注的RNA分子,或者可以是由其衍生的子代细胞。

[0427] 重组细胞

[0428] 本发明还提供了重组细胞,例如重组细菌细胞、真菌细胞、植物细胞、昆虫细胞或动物细胞,其为用本文定义的一种或多种RNA分子或载体或其组合转化的宿主细胞。本发明的合适细胞包括可以用根据本发明的RNA分子或重组载体转化的任何细胞。在一个实施例中,转化的宿主细胞是死亡的。

[0429] 所述重组细胞可以是培养中的细胞,体外细胞,或例如植物的生物体中的细胞,或例如种子或叶的器官中的细胞。优选地,所述细胞在植物中,更优选在植物的种子中。在一个实施方案中,重组细胞是非人细胞。因此,在一个实施例中,本发明涉及包含在此公开的RNA分子的一个或多个或全部的非人生物体。

[0430] 在一个实施例中,所述细胞是昆虫细胞。在一个实施例中,昆虫细胞来源于粉纹夜蛾(Trichoplusia)。

[0431] 合适的宿主细胞的另一个实施例是电感受态HT115细胞。

[0432] 导入RNA分子的宿主细胞可以是未转化的细胞或已经用至少一种核酸转化的细

胞。此类核酸可与脂质合成相关,或不相关。本发明的宿主细胞可以内源性地(即,天然地)能够表达本文定义的RNA分子,在这种情况下,由其衍生的重组细胞具有增强的产生RNA分子的能力,或者仅在用至少一种本文定义的RNA分子转化后才能够产生所述RNA分子的能力。在一个实施例中,所述细胞是能够用于产生脂质的细胞。在一个实施方案中,本发明的重组细胞具有增强的产生诸如TAG的非极性脂质的能力。

[0433] 本发明的宿主细胞可以是能够表达本文所述的至少一种RNA分子的任何细胞,并且包括细菌、真菌(包括酵母)、寄生虫、节肢动物、动物和植物细胞。宿主细胞的实施例包括沙门氏菌属(Salmonella)、埃希氏菌属(Escherichia)、芽孢杆菌属(Bacillus)、李斯特菌属(Listeria)、酵母属(Saccharomyces)、灰翅夜蛾属(Spodoptera)、分枝杆菌属(Mycobacteria)、粉纹夜蛾属(Trichoplusia)、农杆菌属(Agrobacterium)、BHK(幼仓鼠肾)细胞、MDCK细胞、CRFK细胞、CV-1细胞、COS(例如COS-7)细胞和Vero细胞。宿主细胞的进一步实施例是大肠杆菌,包括大肠杆菌K-12衍生物;伤寒沙门氏菌(Salmonella typhi);鼠伤寒沙门氏菌(Salmonella typhimurium),包括减毒株;草地贪夜蛾(Spodoptera frugiperda);粉纹夜蛾(Spodoptera frugiperda);和非肿瘤发生性小鼠成肌细胞G8细胞(例如,ATCC CRL 1246)。其他合适的哺乳动物细胞宿主包括其他肾细胞系、其他成纤维细胞系(例如,人、鼠或鸡胚胎成纤维细胞系)、骨髓瘤细胞系、中国仓鼠卵巢细胞、小鼠NIH/3T3细胞、LMTK细胞和/或HeLa细胞。

[0434] 在一个优选的实施方案中,植物细胞是种子细胞,特别是种子的子叶或胚乳中的细胞。在一个实施方案中,细胞是动物细胞。动物细胞可以是任何类型的动物,诸如举例而言,非人动物细胞、非人脊椎动物细胞、非人哺乳动物细胞,或诸如鱼或甲壳动物的水生动物、无脊椎动物、昆虫等的细胞。用作本发明宿主细胞的藻类细胞的实施例包括例如衣藻属(Chlamydomonas)物种(例如,莱茵衣藻(C lamydomonas reinhardtii))、杜氏藻属(Dunaliella)物种、红球藻属(Haematococcus)物种、小球藻属(Chlorella)物种、破囊壶菌属(Thraustochytrium)物种、裂殖壶菌属(Schizochytrium)物种,和团藻属(Volvox)物种。

[0435] 转基因植物

[0436] 本发明还提供了包含一种或多种本文定义的外源RNA分子的植物、根据本发明的细胞、根据本发明的载体或其组合。当用作名词时,术语“植物”是指完整植物,而术语“其部分”是指植物器官(例如,叶、茎、根、花、果实、种子);单细胞(例如,花粉);种子;种子部分,如胚、胚乳、胚盘或种皮;植物组织,如维管组织;植物细胞及其后代。如本文所用,植物部分包含植物细胞。

[0437] 如本文所用,在对植物进行修饰的上下文中,术语“在植物中”和“在该植物中”意指修饰已经在植物的至少一部分中发生,包括修饰已经在整个植物中发生的情况,并且不排除修饰仅在植物的一个或多个而不是所有部分中发生的情况。例如,组织特异性启动子被称为“在植物中”表达,即使它可能仅在植物的某些部分中表达。类似地,“增加植物中一种或多种糖酵解和/或脂肪酸生物合成基因的表达的转录因子多肽”意指增加的表达发生在植物的至少一部分中。

[0438] 如本文所用,术语“植物”以其最广泛的含义使用,包括植物界中的任何生物体。还包括红藻、褐藻以及绿藻。其包括但不限于任何种类的开花植物、草、作物或谷类(例如,油籽、玉米、大豆)、饲料或草料、水果或蔬菜植物、草本植物、木本植物或树。这并不意味着将

植物限制于任何特定结构。它还指单细胞植物(例如,微藻)。关于植物的术语“其部分”是指植物细胞及其后代,多个植物细胞,存在于植物发育的任何阶段的结构,或植物组织。这样的结构包括但不限于叶、茎、花、果实、坚果、根、种子、种皮、胚。术语“植物组织”包括植物的分化的和未分化的组织,包括存在于叶、茎、花、果实、坚果、根、种子中的那些,例如,胚组织、胚乳、真皮组织(例如,表皮、周皮)、维管组织(例如,木质部、韧皮部),或基本组织(包含薄壁组织、厚角组织,和/或间充质细胞),以及培养物中的细胞(例如,单细胞、原生质体、愈伤组织、胚胎等)。植物组织可以是植物、器官培养物、组织培养物或细胞培养物。

[0439] 在不同植物的糖脂中发现不同量的18:3和16:3脂肪酸。这用于区分具有3个双键的脂肪酸通常总是C_{18}原子长的18:3植物和同时包含C_{16}和C_{18}脂肪酸的16:3植物。在18:3叶绿体中,催化磷脂酸转化为二酰基甘油和二酰基甘油转化为单半乳糖基二酰基甘油(MGD)的酶活性显著低于16:3叶绿体。在18:3的叶片中,叶绿体在基质中合成硬脂酰ACP2,在饱和烃链中引入第一个双键,然后水解硫酯。释放的油酸酯输出穿过叶绿体包膜进入细胞真核部分的膜,可能是内质网,在那里它被掺入到PC中。PC连接的油酰基在这些膜中去饱和,随后移回到叶绿体中。MGD连接的酰基是用于引入第三双键以产生具有两个亚麻酸残基的MGD的底物。该半乳糖脂类是诸如菊科(Asteraceae)和豆科(Fabaceae)的18:3植物的特征。在例如由伞形科(Apiaceae)和十字花科(Brassicaceae)的成员代表的16:3植物的光合活性细胞中,两个途径并行运行以提供具有MGD的类囊体。协同的“真核”序列通过“原核”途径得到不同程度的补充。其反应限于叶绿体并导致酰基的典型排列以及它们一旦酯化成MGD就完全去饱和。原核DAG骨架携带C16:0及其在C18:脂肪酸被排除C-2位置的去饱和产物。C-1位置被C18脂肪酸占据并且在小程度上被C16基团占据。16:3植物中蓝藻脂质DAG骨架与叶绿体转运途径合成的DAG骨架的相似性说明了这一系统发生关系,并证明了原核细胞的存在。

[0440] 如本文所用,术语“营养组织”或“营养植物部分”是除了用于植物有性繁殖的器官之外的任何植物组织、器官或部分。植物有性繁殖的器官具体为带种器官、花、花粉、果实和种子。营养组织和部分至少包括植物叶,茎(包括茎球和分蘖,但不包括头端),块茎和根,但不包括花,花粉,包括种皮的种子,胚和胚乳,包括中果皮组织的果实,带有种子的荚和带有种子的头端。在一个实施方案中,植物的营养部分是地上植物部分。在另一个或进一步的实施方案中,营养植物部分是绿色部分,如叶或茎。

[0441] “转基因植物”或其变体是指含有在相同物种、品种或栽培种的野生型植物中未发现的转基因的植物。如本发明上下文中所定义的转基因植物包括已经使用重组技术遗传修饰以在所需植物或其部分中产生本文所定义的至少一种多肽的植物及其后代。转基因植物部分具有相应的含义。

[0442] 术语“种子”和“谷粒”在本文中可互换使用。“谷粒”是指成熟的谷粒,例如收获的谷粒或仍在植物上但准备收获的谷粒,但根据上下文,也可以指吸胀或发芽后的谷粒。成熟谷粒通常具有小于约18%的水分含量。在一个优选的实施例中,该谷粒的水分含量处于通常被认为是储存安全的水平,优选地5%-15%、6%-8%、8%-10%,或10%-15%。如本文所用,“发育中的种子”是指成熟前的种子,通常在受精或开花后在植物的生殖结构中发现,但也可指从植物分离的成熟前的此类种子。成熟种子通常具有小于约12%的水分含量。

[0443] 如本文所用,术语“植物储存器官”是指专门以例如蛋白质、碳水化合物、脂质的形

式储存能量的植物的一部分。植物储存器官的实施例是种子、果实、块茎根和块茎。本发明优选的植物贮藏器官是种子。

[0444] 如本文所用,术语“表型正常”是指遗传修饰的植物或其部分,例如转基因植物,或与未修饰的植物或其部分相比不具有显著降低的生长和繁殖能力的本发明的贮藏器官,如种子、块茎或果实。优选地,在相同条件下生长时,所产生的生物量、生长速率、发芽率、贮藏器官大小、种子大小和/或存活种子的数目不低于缺乏所述重组多核苷酸的植物的生物量、生长速率、发芽率、贮藏器官大小、种子大小和/或存活种子的数目的90%。该术语不包括可能与野生型植物不同但不影响植物用于商业目的的有用性的植物特征,例如幼苗叶片的无内里(ballerina)表型。在一个实施方案中,表型正常的所述经遗传修饰的植物或其部分包含重组多核苷酸,所述重组多核苷酸编码沉默抑制因子,所述沉默抑制因子可操作地连接至植物贮藏器官特异性启动子并且具有与不包含所述多核苷酸的相应植物或其部分基本相同的生长或繁殖能力。

[0445] 本发明提供的或预期用于实施本发明的植物包括单子叶植物和双子叶植物。在优选的实施方案中,本发明的植物是农作物(例如,谷类和干豆、玉米、小麦、马铃薯、水稻、高粱、粟、木薯、大麦)或豆类如大豆、菜豆或豌豆。所述植物可以生长用于生产可食用的根、块茎、叶、茎、花或果实。植物可以是其营养部分用作食物的蔬菜植物。本发明的植物可以是:马库巴棕榈(Acrocomia aculeata),拟南芥(Arabidopsis thaliana),花生(Aracinis hypogaea),棕榈乳脂(Astrocaryum murumuru),星果棕(Astrocaryum vulgare),天竺葵(Attalea geraensis),美洲油棕(Attalea humilis),油茶(Attalea oleifera)(andaiá),Attalea phalerata(uricuri),巴西棕榈(Attalea speciosa)、燕麦(Avena sativa)、甜菜(Beta vulgaris),芥属(Brassica sp),诸如埃塞俄比亚芥(Brassica carinata)、印度芥(Brassica juncea)、芜菁甘兰(Brassica napobrassica)、欧洲油菜(Brassica napus)、亚麻荠(Camelina sativa)、红花(Carthamus tinctorius),巴西油桃木果(Caryocar brasiliense),椰子(Cocos nucifera),埃塞俄比亚两节荠(Crambe abyssinica)、甜瓜(Cucumis melo)、非洲棕榈(Elaeis guineensis),大豆(Glycine max),棉(Gossypium hirsutum),向日葵属(Helianthus)物种,诸如向日葵(Helianthus annuus),大麦(Hordeum vulgare),麻疯树(Jatropha curcas),安达树(Joannesia princeps),浮萍属(Lemna)物种,诸如细脉浮萍(Lemna aequinoctialis),Lemna disperma,Lemna ecuadoriensis,溶胀浮萍(Lemna gibba),日本浮萍(Lemna japonica),小浮萍(Lemna minor),单脉萍(Lemna minuta),Lemna obscura,稀脉萍(Lemna paucicostata),稀脉浮萍(Lemna perpusilla),Lemna Lemnaminuta tenera,品萍(Lemna trisulca),鳞根萍(Lemna turionifera),Lemna valdiviana,Lemna yungensis,坚硬利堪蔷薇木(Licania rigida),亚麻(Linum usitatissimum),羽扇豆(Lupinus angustifolius),毛瑞榈(Mauritia flexuosa),伊纳雅榈(Maximiliana maripa),芒属(Miscanthus)物种,诸如大芒(Miscanthusx giganteus)和中华芒(Miscanthus sinensis),烟草属(Nicotiana)物种,诸如烟草(Nicotiana tabacum)或本氏烟(Nicotiana benthamiana),Oenocarpus bacaba,巴套阿酒实棕(Oenocarpus bataua),Oenocarpus distichus,稻属(Oryza),诸如水稻(Oryza sativa)和稃稻(Oryza glaberrima),柳枝稷(Panicum virgatum),马里(Paraqueiba paraensis),牛油果(Persea amencana),水黄皮(水黄皮),毛果杨(Populus trichocarpa),蓖麻(Ricinus communis),

甘蔗属(Saccharum)物种,芝麻(Sesamumindicum),马铃薯(Solanum tuberosum),高粱属(Sorghum)物种,诸如双色高粱(Sorghum bicolor),高粱(Sorghum vulgare),大叶罗汉松(Theobroma grandiforum),三叶草属(Trifolium)物种,巴西针叶棕榈(Trithrinax brasiliensis),小麦属(Triticum)物种,诸如小麦(Triticum aestivum),玉蜀黍(Zea mays),苜蓿(Medicago sativa),黑麦(Secale cerale),甘薯(Lopmoea batatus),木薯(Manihot esculenta),咖啡属(Cofea)物种,菠萝(Anana comosus),柑橘属(Citrus)物种,可可(Theobroma cacao),茶(Camellia senensis),香蕉属(Musa)物种,鳄梨(Persea americana)无花果(Ficus casica),番石榴(Psidium guajava),芒果(Mangifer indica),橄榄(Olea europaea),番木瓜(Carica papaya),腰果(Anacardium occidentale),澳洲坚果(Macadamia intergrifolia)和杏仁(Prunus amygdalus)。例如,本发明的植物可以是本氏烟(Nicotiana benthamiana)。

[0446] 其他优选的植物包括C4草,例如,除了上述那些之外,须芒草(Andropogon gerardi)、垂穗草(Bouteloua curtipendula)、细野麻(B.gracilis)、野牛草(Buchloe dactyloides)、帚裂稃草(Schizachyrium scoparium)、黄假高粱(Sorghastrum nutans)、沙鼠尾粟(Sporoboluscreptandrus);C3草,例如加拿大披碱草(Elymus canadensis),豆科头状胡枝子(Lespedeza capitata)和簇毛半枝莲(Petalostemum villosum),杂类草(Aster azureus);和木本植物,如椭圆栎(Quercus ellipsoidalis)和大果栎(Q.macrocarpa)。其他优选的植物包括C3草。

[0447] 在优选的实施方案中,植物是被子植物。

[0448] 在一个实施方案中,植物是油籽植物,优选油籽作物植物。如本文所用,“油籽植物”是用于从植物的种子商业生产脂质的植物物种。油籽植物可以是例如油籽油菜(例如油菜)、玉米、向日葵、红花、大豆、高粱、亚麻(亚麻籽)或甜菜。此外,油籽植物可以是其他芸苔、棉花、花生、芸苔、芥菜、蓖麻、芝麻、红花、麻疯树(Jatropha curcas)或产生坚果的植物。所述植物可在其果实施例如橄榄、油棕或椰子中产生高水平的脂质。可以应用本发明的园艺植物是莴苣,甘蓝或蔬菜芸苔,包括卷心菜、椰菜或花椰菜。本发明可应用于烟草、瓜类、胡萝卜、草莓、番茄或胡椒。

[0449] 在一个优选的实施方案中,所述转基因植物对于已经导入的每个基因(转基因)是纯合的,使得其后代不分离所需表型。转基因植物对于导入的转基因也可以是杂合的,优选对于转基因是一致杂合的,例如在已经从杂交种子生长的F1子代中。这类植物可提供本领域熟知的优点,如杂种优势。

[0450] 转化

[0451] 本文公开的RNA分子可以稳定地导入上述宿主细胞和/或非人生物体,如植物。为了避免疑问,本发明的实施例涵盖用在此公开的RNA分子稳定转化的以上提及的植物。如本文所用,术语“稳定转化”、“稳定转化的”及其变体是指RNA分子或编码RNA分子的核酸整合到细胞的基因组中,使得它们在细胞分裂期间被转移至子代细胞而不需要阳性选择它们的存在。稳定的转化体或其后代可以通过本领域已知的任何方法鉴定,例如在染色体DNA上的Southern印迹,或基因组DNA的原位杂交,使得它们能够被选择。

[0452] 转基因植物可以使用本领域已知的技术产生,例如在Slater等人,植物生物技术-植物的遗传操作(Plant Biotechnology-The Genetic Manipulation of Plants),牛津大

学出版社(2003),和Christou和Klee,植物生物技术手册(Handbook of Plant Biotechnology),John Wiley and Sons(2004年)中一般描述的那些技术。

[0453] 在一个实施方案中,可以通过将根据本发明的RNA分子局部施用于植物或其部分来转化植物。例如,RNA分子可以作为具有合适载体的制剂提供并且喷雾、撒粉或以其他方式施用至植物或其部分的表面。因此,在一个实施例中,本发明的方法涵盖将在此公开的RNA分子引入植物,该方法包括将包含该RNA分子的组合物局部施用于该植物或其部分。

[0454] 农杆菌介导的转移是用于将基因导入植物细胞的广泛适用的系统,因为可以将DNA导入组织培养中的整个植物组织、植物器官或外植体中的细胞中,用于瞬时表达或用于DNA在植物细胞基因组中的稳定整合。例如,可以使用花卉浸渍(in planta)法。使用农杆菌(Agrobacterium)介导的植物整合载体将DNA导入植物细胞是本领域熟知的。待转移的DNA区由边缘序列限定,间插DNA(T-DNA)通常插入到植物基因组中。由于基因转移的简便和明确的性质,因此选择该方法。

[0455] 可以使用的加速方法包括例如微粒轰击等。用于将转化核酸分子递送至植物细胞的方法的一个实施例是微粒轰击。该方法已被综述于文献Yang等,用于基因转移的粒子轰击技术(Particle Bombardment Technology for Gene Transfer),牛津出版社,牛津,英国(1994)。非生物颗粒(微粒),其可以用核酸包被并通过推进力递送到例如未成熟胚的细胞中。示例性颗粒包括由钨、金、铂等组成的那些。

[0456] 在另一种方法中,可以稳定地转化质粒。所公开的用于高等植物中的质粒转化的方法包括含有选择标记的DNA的粒子枪递送和通过同源重组将DNA靶向至质粒基因组(US 5,451,513、US 5,545,818、US 5,877,402、US 5,932479,和WO 99/05265)。也可以使用细胞转化的其他方法,包括但不限于通过将DNA直接转移到花粉中,通过将DNA直接注射到植物的生殖器官中,或通过将DNA直接注射到未成熟胚的细胞中,然后再水化干燥的胚,从而将DNA导入植物中。

[0457] 来自单一植物原生质体转化体或来自各种转化外植体的植物的再生、发育和培养是本领域熟知的(Weissbach等,在:植物分子生物学方法(植物分子生物学方法),学术出版社,圣地亚哥,加利福尼亚州,(1988))。这种再生和生长过程通常包括以下步骤:选择转化的细胞,在从生根的小植株阶段到胚胎发育的通常阶段培养那些个体化的细胞。转基因胚胎和种子类似地再生。然后将得到的转基因生根的芽种植在合适的植物生长培养基,如土壤中。

[0458] 含有外源基因的植物的发育或再生是本领域熟知的。优选地,再生的植物自花授粉以提供纯合的转基因植物。否则,将再生植株获得的花粉与农艺上重要的品系的种子生长植株杂交。相反,来自这些重要品系的植物的花粉用于给再生的植物授粉。使用本领域技术人员熟知的方法培养含有所需多核苷酸的本发明的转基因植物。

[0459] 为了确认转基因在转基因细胞和植物中的存在,可以使用本领域技术人员已知的方法进行聚合酶链反应(PCR)扩增或Southern印迹分析。转基因的表达产物可以以多种方式中的任一种检测,这取决于产物的性质,并且包括Northern印迹杂交、蛋白质印迹和酶测定。一旦获得转基因植物,它们可以生长以产生具有所需表型的植物组织或部分。可以收获植物组织或植物部分,和/或收集种子。种子可用作生长具有具有所需特征的组织或部分的额外植物的来源。优选地,在非极性脂质的产量最高时收获营养植物部分。在一个实施方案

中,营养植物部分大约在开花时或开花开始后收获。优选地,所述植物部分在大约衰老开始时收获,通常表现为叶片变黄和干燥。

[0460] 使用农杆菌(Agrobacterium)或其他转化方法形成的转基因植物通常在一个染色体上含有单个基因座。这样的转基因植物可以被认为是所添加基因的半合子。更优选的是对于添加的基因是纯合的转基因植物,即,含有两个添加的基因的转基因植物,一个基因在染色体对的每个染色体上的同一基因座上。通过将半合子转基因植物自身受精,使所产生的一些种子发芽并分析所得植物的目的基因,可以获得纯合转基因植物。

[0461] 还应当理解,含有两个独立分离的外源基因或基因座的两种不同转基因植物也可以杂交(交配),以产生含有两组基因或基因座的后代。适当的F1后代的自交可以产生对于外源基因或基因座而言是纯合的植物。还考虑了与亲本植物的回交和与非转基因植物的异交,以及无性繁殖。类似地,转基因植物可以与包含诸如突变基因的遗传修饰的第二植物和含有所鉴定的转基因和遗传修饰两者的后代杂交。通常用于不同性状和作物的其他育种方法的描述可见于Fehr,In:栽培种发育的育种方法(Breeding Methods for Cultivar Development),WilcoxJ。美国农艺学会(Breeding Methods for Cultivar Development),美国威斯康星州麦迪逊分校(1987)。

[0462] 制剂

[0463] 根据本发明的RNA分子可以作为各种制剂提供。例如,RNA分子可以是固体、软膏剂、凝胶、乳膏剂、粉末、糊剂、混悬剂、胶体、泡沫或气雾剂的形式。固体形式可以包括散剂、粉剂、颗粒剂、微球剂、丸剂、锭剂、片剂、填充膜(包括种子包衣)等,它们可以是水可分散的。在一个实施例中,该组合物是浓缩物的形式。

[0464] 在一个实施例中,RNA分子可以作为局部制剂提供。在一个实施例中,所述制剂稳定制剂中和/或体内的RNA分子。例如,RNA分子可以在脂质制剂中提供。例如,可以在脂质体中提供RNA分子。在一个实施例中,所述制剂包含转染促进剂。

[0465] 在一个实施例中,可将RNA分子掺入适于现场施用的制剂中。在一个实施例中,该田地包括植物。合适的植物包括农作物(例如,谷类和豆类、玉米、小麦、马铃薯、木薯、水稻、高粱、大豆、木薯、大麦或豌豆),或豆科植物)。所述植物可以生长用于生产可食用的根、块茎、叶、茎、花或果实。在一个实施例中,作物植物是谷类植物。谷类植物的实施例包括但不限于小麦、大麦、高粱燕麦和黑麦。在这些实施例中,RNA分子可以配制用于以任何合适的方式施用于植物或植物的任何部分。例如,组合物可以配制用于施用至植物的叶、茎、根、果实、蔬菜、谷粒和/或芽。在一个实施例中,RNA分子被配制用于施用至植物的叶子,并且可喷雾至植物的叶子上。

[0466] 根据所需的制剂,本文所述的RNA分子可以与多种其他试剂一起配制。示例性试剂包括悬浮剂、聚结剂、碱、缓冲剂、苦味剂、芳香剂、防腐剂、推进剂、触变剂、防冻剂和着色剂中的一种或多种。

[0467] 在其他实施例中,RNA分子制剂可包含杀虫剂、农药、杀真菌剂、抗生素、驱虫剂、抗寄生虫剂、抗病毒剂或杀线虫剂。

[0468] 在另一个实施例中,可以将RNA分子掺入药物组合物中。这类组合物通常包括本文所述的RNA分子和药学上可接受的载体。如本文所用,术语“药学上可接受的载体”包括与药物施用相容的溶剂,分散介质,包衣,抗细菌剂和抗真菌剂,等渗剂和吸收延迟剂,等。补充

的活性化合物也可以掺入组合物中。

[0469] 将药物组合物配制成与其预期的给药途径相容。给药途径的实施例包括肠胃外，例如静脉内、真皮内、皮下、吸入、经皮(局部)、经粘膜、口服和直肠给药。

[0470] 在一个实施方案中，活性化合物与保护化合物免于从体内快速消除的载体一起制备，诸如控释制剂，包括植入物和微囊化递送系统。可使用生物可降解的生物相容的聚合物，例如乙烯乙酸乙烯酯、聚酐、聚乙醇酸、胶原、聚原酸酯和聚乳酸。例如，脂质体悬浮液也可用作药学上可接受的载体。这些可以根据本领域技术人员已知的方法制备，例如，如US 4,522,811中所述。

[0471] 根据本发明的RNA分子可以在试剂盒或包装中提供。例如，本文公开的RNA分子可以包装在具有书面说明书的合适容器中，用于产生以上提及的细胞或生物体或治疗病症。

[0472] 防治非人生物体的方法

[0473] 在一个实施例中，根据本发明的RNA分子可以用于防治非人生物体，如昆虫。这些用途涉及使用各种方法施用根据本发明的RNA分子。在一个实施例中，可以提供根据本发明的RNA分子作为供昆虫摄取的昆虫诱饵。在另一个实施例中，RNA分子可以根据需要喷洒在昆虫上。在另一个实施例中，可以将RNA分子喷洒在植物或作物上以保护所述植物或作物免受昆虫的侵害。示例性作物包括棉花、玉米、番茄、鹰嘴豆、木豆、苜蓿、水稻、高粱和豇豆。

[0474] 在一个实施例中，可以提供RNA分子来修饰昆虫行为。在另一个实施例中，可以提供RNA分子以杀死昆虫。在另一个实施例中，可以提供RNA分子以降低昆虫生育力。示例性的昆虫靶标包括家庭昆虫。其他示例性的昆虫靶标包括吸汁液的昆虫，诸如蚜虫(例如桃蚜(Myzus persicae)、麦无网蚜(Metopolophium dirhodum)、禾谷蚜(Rhopalosiphum padi)、大豆蚜(Aphis glycines)、蚕豆蚜虫(Aphis glycines)。另外的示例性昆虫靶标包括蛛形纲动物、蚊子、体外寄生虫、苍蝇、蛛螨、蓟马、扁虱、红禽螨、蚂蚁、蟑螂、白蚁、包括家养蟋蟀的蟋蟀、银鱼、书虱、甲虫、蠓、蚊子和蚤。其他示例性昆虫靶标包括农业害虫。实施例包括树液饲养剂，如椿象和蚜虫，诸如毛虫的咀嚼昆虫，甲虫，蠕虫，诸如蓟马和鼻涕虫的锉磨昆虫，蛾，诸如果蝇的蝇，诸如谷物钻孔虫、象鼻虫和谷物蛾的谷物害虫。

[0475] 在一个实施方案中，所述昆虫是吸汁昆虫。在该实施例中，RNA分子可对MpCOO2和/或MpRack-1具有反义活性。在一个实施方案中，吸汁昆虫是蚜虫。在另一个实施方案中，蚜虫是桃蚜(Myzus persicae)。

[0476] 在一个实施方式中，昆虫靶标是蚂蚁(例如，阿根廷蚁(Linepithema humile))、棉铃虫或玉米穗虫(棉铃虫Helicoverpa armigera)或苍蝇(例如，铜绿蝇(Lucilia cuprina))。在一个实施方案中，靶标昆虫是棉铃虫(Helicoverpa armigera)，并且RNA分子对ABC转运蛋白白基因(ABCwhite)具有反义活性。在另一个实施方案中，靶昆虫是阿根廷蚁(Linepithema humile)，并且RNA分子对于信息素生物合成激活神经肽(PBAN)具有反义活性。在另一个实施方案中，靶昆虫是铜绿蝇(Lucilia cuprina)，并且RNA分子对编码蛋白的一个或多个基因具有反义活性，所述蛋白选自由V型质子ATP酶催化亚基A、RNA酶1/2、几丁质合酶、蜕皮激素受体和γ-微管蛋白1/1-样组成的组。

[0477] 在上述实施方案中，本文公开的组合物和RNA分子可以在分配器中提供。在一个实施例中，该分配器是诱捕器或诱饵。在一个实施例中，该诱捕器和/或诱饵包括含有在此公开的一种或多种RNA分子的诱饵。

[0478] 在一个实施方案中,本发明包括控制昆虫行为的方法,该方法包括将本文公开的RNA分子喷雾、撒粉或以其他方式施用至昆虫。在该实施方案中,RNA分子可以喷雾、撒粉或以其他方式直接施用于昆虫。在另一个实施方案中,可以在昆虫侵染之前将RNA分子喷雾、撒粉或以其他方式施用于植物或作物。

[0479] 在本发明的一个实施方案中,昆虫或蛛形纲动物可属于下列目:蜱螨亚纲(Acari),蛛形纲(Arachnida),虱目(Anoplura),蜚蠊目(Blattodea),鞘翅目(Coleoptera),弹尾目(Collembola),革翅目(Dermaptera),网翅目(Dictyoptera),Dipura,Emboptera,蜉蝣目(Ephemeroptera),蛩蠊目(Grylloba todea),半翅目(Hemiptera),异翅目(Heteroptera,),同翅目(Homoptera),膜翅目(膜翅目),等翅目(Isoptera),鳞翅目(Lepidoptera),食毛目(Mallophaga),长翅目(Mecoptera),神经翅目(Neuroptera),蜻蜓目(Odonata),直翅目(Orthoptera),竹节虫目(Phasmida),Phithiraptera,襀翅目(Plecoptera),原尾目(Protura)、啮虫目(Psocoptera),蚤目(Siphunculata),缨尾目(Thysanura),腹吻亚目(Sternorrhynchoa),翅目(Strepsiptera),缨翅目(Thysanoptera),毛翅目(Trichoptera),缺翅目(Zoraptera)和衣鱼目(Zygentoma)。

[0480] 在本发明的优选但非限制性的实施方案中,昆虫或蛛形纲动物选自由以下组成的组:(1)蜱螨亚纲(Acari):螨类,包括真蜱目(Ixodida)(2)蛛形纲(蛛形纲):蜘蛛属(Araneae)和粉虱属(Opiliones),实施例包括:黑寡妇(Latrodectus mactans)和褐色斜发蛛属(Loxosceles recluse)(3)蚜虫属(Anoplura):虱,例如人体虱(Pediculus humanus)(4)蜚蠊目(Blattodea):蟑螂,包括德国小蠊(Blatella germanica),大蠊属(Periplaneta),包括美洲大蠊(Periplaneta americana)和澳洲蜚蠊(Periplaneta australiasiae),蜚蠊属(Blatta),包括东方蜚蠊(Blatta orientalis)和棕带蜚蠊(Supella longipalpa)。最优选的靶标是德国小蠊(Blatella germanica)。(5)鞘翅目:甲虫,实施例包括:Powderpost甲虫的家族(长蠹总科(Bostricholidae)的家族);大小蠹属(Dendroctonus)物种(黑松节油甲虫,南方松脂甲虫,IPS雕甲虫);地毯甲虫(圆皮蠹属(Anthrenus),毛皮蠹属(Attagenus));家天牛(天牛科(Cerambicyidae)的家族:家天牛(Hylotrupes bajulus);家具窃蠹虫(Anobium punctatum);拟谷盗属(Tribolium)谷斑皮蠹(Trogoderma granarium);齿谷甲虫(Oryzaephilus sarinamensis)等(书虫)(6)革翅目(Dermaptera):蠼螋科的家族(7)双翅目(Diptera):蚊子(Culicidae)和苍蝇(Brachycera),实施例是:按蚊(Anophelinae),例如按蚊属(Anopheles)和蚊亚科(Culicinae),如黄斑伊蚊(Aedes fulvus);虻科(Tabanidae),例如马蝇(Tabanus punctifer),刺舌蝇(Glossina morsitans morsitans),毛蠓科(Psychodidae)和有瓣蝇科(Calyptratae),诸如家蝇(Musca domestica),肉蝇(Sarcophagidae)等(8)异翅亚目(Heteroptera):臭虫,诸如温带臭虫(Cimex lectularius)(9)膜翅目:细腰亚目(Apocrita),包括蚁总科(Formicoidea),蜜蜂总料(Apoidea):红火蚁(Solenopsis invicta),法老蚁(Monomorium pharaonis),弓背蚁属(Camponotus)物种,小黑蚁(Iasius niger),铺道蚁(tetramorium caespitum),红蚁(Myrmica rubra),木蚁属(Formica)物种,举腹蚁属(Crematogaster lineolata)物种,阿根廷蚁(Iridomyrmex humilis),大头蚁属(Pheidole)物种,天鹅绒蚁(Dasymutilla occidentalis)等(10)等翅目(Isoptera):白蚁,

实施例包括:佛罗里达黑翅地下白蚁(Amitermes floridensis),东方地下白蚁(Reticulitermes flavipes),西方地下白蚁(R.hesperus),台湾乳白蚁(Coptotermes formosanus),西方干木白蚁(Incisitermes minor),林木白蚁(Neotermes connexus)和白蚁科(Termitidae)(11)鳞翅目(Lepidoptera):蛾,实施例包括:谷蛾科(Tineidae)和织叶蛾科(Oecophoridae),例如普通衣蛾(Tineola bisselliella),和螟蛾科(Pyrolidae),例如谷蛾(Pyrolisfarinalis)等(12)啮虫目(啮虫目):书虱(Psocids)(13)蚤类(Siphonaptera):蚤类,例如人蚤(Pulex irritans)(14)腹吻亚目(Sternorrhyncha):蚜科(Aphididae)(15)衣鱼目(Zygentoma):银鱼,实施例是:家衣鱼(Thermobia domestica)和蠹鱼(Lepisma saccharina);

[0481] 其他目标昆虫或蛛形纲动物包括家庭昆虫,体外寄生虫和与公共卫生和保健相关的昆虫和/或蛛形纲动物,例如但不限于苍蝇、蛛螨、蓟马、蜱、红家禽螨、蚂蚁(例如通过靶向PBAN)、蟑螂、白蚁、包括家蟋蟀的蟋蟀、银鱼、书虱、甲虫、蠼螋、蚊子和蚤。更优选的靶标是蟑螂(Blattodea),例如但不限于小蠊属(Blatella)物种(例如:德国小蠊(Blatella germanica)),大蠊属(Periplaneta)物种(例如:美洲大蠊(Periplaneta americana),澳洲大蠊(Periplaneta australiasiae);蜚蠊属(Blatta)物种(例如:东方蟑螂(Blatta orientalis)和棕带蜚蠊属(Supella)物种(例如棕带蜚蠊(Supella longipalpa);蚁总科(Formicoidea),例如但不限于火蚁属(Solenopsis)物种(例如红火蚁(Solenopsis invicta)),法老蚁属(Monomorium)物种(例如法老蚁(Monomorium pharaonis)),弓背蚁属(Camponotus)物种(例如弓背蚁(Camponotus)),毛蚁属(lasius)(例如小黑蚁(Iasius niger)),铺道蚁(Tetramorium)属,(例如铺道蚁(tetramorium caespitum)),红蚁属(Myrmica)(例如红蚁(Myrmica rubra)),木蚁属(Formica)物种,举腹蚁属(Crematogaster)(例如举腹蚁(Crematogaster lineolata)),阿根廷蚁属(Crematogaster)(例如阿根廷蚁(Iridomyrmex humilis)),大头蚁属(Pheidole)物种,和天鹅绒属蚁(Dasymutilla)(例如天鹅绒蚁(Dasymutilla occidentalis))白蚁(等翅目(Isoptera)和/或白蚁科(Termitidae)),例如但不限于白蚁属(Termitidae)(例如:佛罗里达黑翅地下白蚁(Amitermes floridensis),散白蚁属(例如:东方地下白蚁(Reticulitermes flavipes),西方地下白蚁(Reticulitermes hesperus),家白蚁属(Coptotermes)(例如台湾乳白蚁(Coptotermes formosanus)),楹白蚁属(Incisitermes)(例如西方干木白蚁(Incisitermes minor)),新白蚁属(Neotermes)(Reticulitermeshesperus)(例如林木白蚁(Neotermes connexus))。

[0482] 在一个实施方案中,靶RNA编码昆虫乳酸合酶。

[0483] 当在植物中递送和/或表达时,本发明的RNA分子可以具有广泛的所需特性,这些特性影响例如农学性状,昆虫抗性(例如通过靶向基因,如MpC002,MpRack-1和ABC转运蛋白基因),疾病抗性(例如通过靶向基因,如LanR),除草剂抗性,不育性,谷粒特性等。靶RNA分子可能参与油、淀粉、碳水化合物、营养素等的代谢,或可能参与蛋白质、肽、脂肪酸、脂质、重组频率(通过靶向基因,如DDM1和FANCM)、蜡、油(通过靶向基因,如TOR)、淀粉、糖、碳水化合物、香料、气味、毒素、类胡萝卜素、激素(通过靶向基因,如EIN2、NCED1和NCED2)、聚合物、类黄酮(通过靶向基因,如查耳酮合酶)、贮藏蛋白、酚酸、生物碱、木质素、单宁、纤维素、糖蛋白、糖脂等的合成。

[0484] 在特定实施例中,所述植物产生:在植物中增加水平的用于产油的酶,诸如芸苔属(Brassicas)植物,例如油菜或向日葵、红花、亚麻、棉花、大豆或玉米;酶,参与植物中的淀粉合成,植物诸如马铃薯,玉米和谷类,如小麦、大麦或水稻;合成的酶,或其本身的蛋白质,天然药物,例如药物或兽用产品。

[0485] 在另一个实施方案中,本发明的RNA分子旨在预防性或治疗性治疗选自以下组成的组的真菌病原体的感染:交链孢霉属(Altemaria)物种;蜜环菌(Armillaria mellae);寡孢节丛孢(Armillaria mellae);禾本科布氏白粉菌(Blumeria graminis)(通过使用如实施例17所述的RNA分子靶向Mlo基因),牛肝菌(Boletus granulatus);灰葡萄孢(Botritis cinerea);蚕豆葡萄孢(Botrytisfabae);白色念珠菌(Candida albicans);麦角菌(Claviceps purpurea);松疱锈病菌(Cronartium ribicola);紫球菌(Epicoccum purpurescens);絮状表皮癣菌(Epidermophyton floccosum);层孔菌(Fomes annosus);尖孢镰刀菌(Fusarium oxysporum);小麦全蚀病菌变种小麦(Gaeumannomyces graminis var.tritici);围小丛壳菌(Glomerella cingulata);胶锈菌(Gymnosporangium juniperi-virginianae);犬小孢子菌(Microsporum canis);桃褐腐病菌(Monilinia fructicola);苜蓿锈菌(Physoderma alfalfae);马铃薯晚疫病菌(Phytopthera infestans);秕糠马拉癣菌(Pityrosporum orbiculare);硫色多孔菌(Polyporus sulphureus);柄锈菌属(Puccinia spp.);酿酒酵母(Saccharomyces cerevisiae);芹菜生壳针孢(Septoria apiicola);红色毛癣菌(Trichophyton rubrum);须癣毛癣菌(T.mentagrophytes);黑粉菌属(Ustilago spp.);苹果黑星菌(Venturia inaequalis);和黄萎病菌(Venturia inaequalis)。

[0486] 要治疗的示例性病症

[0487] 根据本发明的RNA分子可用于各种条件的方法中。在一些实施例中,本发明涉及一种治疗癌症的方法,该方法包括给予在此公开的RNA分子。术语“癌症”是指或描述通常以失调的细胞生长/增殖为特征的哺乳动物生理状况。癌症的实施例包括但不限于癌、淋巴瘤、母细胞瘤、肉瘤和白血病或淋巴恶性肿瘤。此类癌症的更具体实施例包括但不限于鳞状细胞癌(例如上皮鳞状细胞癌),肺癌(包括小细胞肺癌),非小细胞肺癌,肺腺癌和肺鳞癌,腹膜癌,肝细胞癌,胃癌或胃脏癌(包括胃肠癌和胃肠道间质癌),胰腺癌,成胶质细胞瘤,子宫颈癌,卵巢癌癌症,肝癌,膀胱癌,泌脏尿道癌症,肝细胞瘤,乳腺癌,结肠癌,直肠癌,结肠直肠癌,子宫内膜癌或子宫癌,唾液腺癌,肾癌或肾脏癌,前列腺癌,外阴癌,甲状腺癌,肝癌,肛门癌,阴茎癌,黑素瘤,浅表扩散性黑素瘤,恶性雀斑样痣黑素瘤,肢端雀斑样痣黑素瘤,结节性黑素瘤,多发性骨髓瘤和B细胞淋巴瘤(包括低度/滤泡性非霍奇金淋巴瘤(NHL));套细胞淋巴瘤;AIDS相关淋巴瘤;和瓦尔登斯特伦氏巨球蛋白血症);慢性淋巴细胞白血病(CLL);急性淋巴细胞白血病(ALL);毛细胞白血病;慢性粒细胞白血病;和移植后淋巴增殖性疾病(PTLD)以及与瘢痣病、水肿(例如与脑肿瘤相关的水肿)、梅格综合征、脑癌和头颈癌相关的异常血管增殖,以及相关转移。因此,在一个实施例中,本发明涉及治疗基于乳腺、卵巢、结肠、前列腺、肺、脑、皮肤、肝、胃、胰腺或血液的癌症的方法。

[0488] 在其他实施例中,本文所述的方法用于治疗与BRCA1、BRCA2、PALB2、OR RAD51B、RAD51C、RAD51D或相关基因中的突变相关的癌症。在其他实施例中,本文所述的方法用于治疗诸如MSH2、MLH1、PMS2和相关基因的与DNA错配修复相关基因中的突变相关的癌症。在其

他实施例中,本文所述的方法用于治疗诸如BRCA1、MLH1、OR RAD51B、RAD51C、OR RAD51D的具有沉默的DNA修复基因的癌症。

[0489] 在本发明的其他实施例中,本文所述的方法用于杀死具有受损的DNA修复过程的细胞。例如,DNA修复受损的细胞可能异常表达参与DNA修复、DNA合成或同源重组的基因。示例性基因包括XRCC1、ADPRT(PARP-1)、ADPRTL2、(PARP-2)、POLYMERASE BETA、CTPS、MLH1、MSH2、FANCD2、PMS2、p53、p21、PTEN、RPA、RPA1、RPA2、RPA3、XPD、ERCC1、XPF、MMS19、RAD51、RAD51B、RAD51C、RAD51D、DMC1、XRCCR、XRCC3、BRCA1、BRCA2、PALB2、RAD52、RAD54、RAD50、MREU、NB51、WRN、BLM、KU70、KU80、ATM、ATR CPIK1、HK2、FANCA、FANCB、FANCC、FANCD1、FANCD2、FANCE、FANCF、FANCG、FANCC、FANCD1、FANCD2、FANCE、FANCF、FANCG、RAD1,和RAD9。在一个实施例中,本文所述的方法用于杀死具有突变肿瘤抑制基因的细胞。例如,细胞可以在BRCA1或BRCA2中具有一个或多个突变。

[0490] 在本发明的其他实施例中,本文所述的方法用于处理病毒转化的细胞。在本发明的其他实施例中,本文所述的方法用于杀死用潜伏病毒转化的细胞。示例性的潜伏病毒包括CMV、EBV、单纯疱疹病毒(I型和II型)和水痘带状疱疹病毒。在本发明的其他实施例中,本文所述的方法用于治疗由引起癌症、免疫缺陷、肝炎、脑炎、肺炎或呼吸系统疾病的病毒引起的活动性病毒感染。示例性病毒包括细小病毒、痘病毒、疱疹病毒。

[0491] 在本发明的其他实施例中,本文所述的方法用于治疗寨卡(Zika)病毒,科罗拉多蜱热(由RNA病毒Coltivirus引起),西尼罗热(脑炎,由主要出现在中东和非洲的黄病毒引起),黄热病,狂犬病(由弹状病毒科的不同毒株引起),病毒性肝炎,由诺瓦克和诺瓦克样病毒、轮状病毒、杯状病毒和星状病毒引起的胃肠炎(病毒性)-急性病毒性胃肠炎,脊髓灰质炎,由能够经受频繁抗原变异的正粘病毒引起的流感(Flu),麻疹(风疹),副粘病毒科,腮腺炎,呼吸综合征,包括肺炎和急性呼吸综合征,包括由统称为急性呼吸病毒的各种病毒引起的喉炎,和由呼吸道合胞病毒引起的呼吸道疾病。

[0492] 实施例

[0493] 实施例1:材料和方法

[0494] 基因构建体的合成

[0495] 为了设计典型的ledRNA构建体,鉴定了长度为约100-1000个核苷酸、通常为400-600个核苷酸的靶RNA区。在一个实施例中,序列的5'半和侧翼区的约130nt以及类似地侧翼区的3'半和130nt相对于启动子以反义方向取向。这些序列被400-600个核苷酸的有义靶序列中断(图1A)。在所得构建体的5'端之前是启动子,例如T7或SP6 RNA聚合酶启动子,并且将3'端工程化以包括限制酶切割位点以允许体外转录终止。

[0496] 对于在如细菌细胞的细胞中的转录,例如使用诱导型启动子,来引入启动子和终止子序列以促进作为转基因的表达。双链区和环序列长度可以变化。构建体使用标准克隆方法制备或从商业服务提供商订购。

[0497] RNA的合成

[0498] 在用限制酶消化以使DNA在3'端线性化之后,使用RNA聚合酶的转录产生与中央靶序列退火的ledRNAi转录物的5'和3'臂,该分子包含具有单个裂口和末端环的中央茎或双链区。中央序列可以相对于启动子以有义或反义方向取向(分别为图1A,图1B)。

[0499] 对于体外合成,使用适当的限制性酶在3'限制性位点消化构建体的DNA,沉淀,纯

化和定量。根据制造商的说明书使用RNA聚合酶实现RNA合成。使用DEPC处理的水将ledRNA重悬于退火缓冲液(25mM Tris-HCL,pH 8.0,10mM $MgCl_2$)中,以灭活任何痕量的RNA酶。通过纳米液滴分析和凝胶电泳测定由该方法产生的RNA的产量和完整性(分别参见图2)。

[0500] 通过将构建体导入大肠杆菌菌株HT115,实现ledRNA在细菌细胞中的合成。用IPTG(0.4mM)诱导转化的细胞培养物,以表达T7 RNA聚合酶,提供ledRNA构建体的转录。从细菌细胞中提取RNA并纯化基本上如Timmons等(2001)所述。

[0501] 对于Cy3标记的RNA转录,核糖核苷酸(rNTP)混合物含有各10mM的ATP、GTP、CTP,1.62mM UTP和8.74mM Cy3-UTP。转录反应在37℃温育2.5小时。将转录反应(160μl)转移到埃普多夫(Eppendorf)管中,加入17.7μlturbo DNase缓冲液和1μlturbo DNA,并在37℃温育10分钟以消化DNA。然后加入17.7μlTurbo DNAse灭活溶液,混合,室温孵育5min。将混合物离心2分钟,将上清液转移到新的无RNA酶的Eppendorf管中。将1.5μl每个转录反应的样品在凝胶上电泳以测试RNA产物的质量。通常,根据构建体,观察到一个大小为500bp至1000bp的RNA条带。通过向每个试管中加入88.5μl 7.5M乙酸铵和665μl冷100%乙醇将RNA沉淀。将管冷却至-20℃数小时或过夜,然后在4℃下离心30分钟。小心地除去上清液,在-20℃下用1ml 70%乙醇(用无核酸酶的水制备)洗涤RNA微丸并离心。干燥沉淀,并将纯化的RNA重悬于50μl1×RNAi退火缓冲液中。使用纳滴法测量RNA浓度并储存在-80℃直至使用。

[0502] 实施例2:ledRNA的设计

[0503] 如图1A示意性所示,典型的ledRNA分子包含可以被认为是共价连接并与靶RNA具有同一性的两个相邻有义序列的有义序列,与所述有义序列互补并被分成两个区的反义序列,和将所述有义序列与所述反义序列分开的两个环。因此,编码这种形式的ledRNA的DNA构建体按5'至3'顺序包含用于ledRNA编码区转录的启动子,与朝向靶RNA5'端的区具有互补性的第一反义区,第一环序列,有义序列,第二环序列,然后是与靶RNA的3'端区具有互补性的第二反义区,最后是终止转录的手段。在这种排列中,两个反义序列位于有义序列和环序列的侧翼。当转录时,反义序列的两个区与有义序列退火,形成具有两个旁侧环的dsRNA茎。

[0504] 在另一种但相关形式的ledRNA中,有义序列被分为两个区,而两个反义区保持为单一序列(图1B)。因此,编码此第二种形式的ledRNA的DNA构建体按5'至3'顺序包含用于ledRNA编码区转录的启动子,与朝向靶RNA的3'端的区具有同一性的第一有义区,第一环序列,反义序列,第二环序列,然后是与靶RNA的5'端具有同一性的第二有义区,最后是终止转录的手段。在这种排列中,两个有义序列位于反义序列和环序列的侧翼。

[0505] 不希望受理论限制,因为在每个末端的闭合环,所以这些ledRNA结构将比在单链有义和反义RNA之间形成且不具有环的开放末端dsRNA更耐受外切核酸酶,并且也比仅具有单个环的发夹RNA更耐受外切核酸酶。此外,本发明人设想dsRNA茎两端的环将允许Dicer有效接近两端,从而增强dsRNA加工成sRNA和沉默效率。

[0506] 作为第一个实施例,使用T7或SP6 RNA聚合酶制备用于体外转录的遗传构建体,以形成靶向编码GFP或GUS的基因的LedRNA。ledGFP构建体依次包含以下区:对应于GFP编码序列(CDS)(SEQ ID NO:7)的核苷酸358-131的反义序列的前半部分,对应于GFPCDS的核苷酸130-1的第一反义环,对应于GFP CDS的核苷酸131-591的有义序列,对应于GFPCDS的核苷酸731-592的第二反义环,对应于GFPCDS的核苷酸591-359的反义序列的后半部分。

[0507] ledGUS构建体依次包含以下区:对应于GUS CDS(SEQ ID NO:8)的核苷酸609-357的反义序列的前半部分;和对应于GUS CDS的核苷酸356至197的第一反义环,对应于GUS CDS的核苷酸357-860的有义序列,对应于GUS CDS的核苷酸1029-861的第二反义环;对应于GUS CDS的核苷酸861-610的反义序列的后半部分。

[0508] 为了制备单独链的有义/反义GUS dsRNA(常规dsRNA),在pGEM-TEasy载体中的T7和SP6启动子之间连接与GUS CDS的核苷酸357-860对应的相同靶序列。有义链和反义链分别用T7或SP6聚合酶转录,混合转录物并加热混合物使RNA链变性后在退火缓冲液中退火。

[0509] 实施例3:ledRNA的稳定性

[0510] 将ledRNA形成dsRNA结构的能力与末端开放的dsRNA(即没有环,由单独的单链有义和反义RNA退火形成)和长hpRNA进行比较。通过煮沸使ledRNA、长hpRNA以及有义和反义RNA的混合物变性,并使其在退火缓冲液(250mM Tris-HCL,pH 8.0和100mM $MgCl_2$/缓冲液)中退火,然后在1.0%琼脂糖凝胶中在非变性条件下进行电泳。

[0511] 如图2所示,GUS ledRNA和GFP ledRNA都给出双链分子预期的迁移率的显性RNA带,表明预测的ledRNA结构的形成。这与仅显示dsRNA弱带的有义和反义RNA的混合物相反,表明大部分有义和反义RNA不容易彼此退火形成dsRNA。发夹RNA样品给出两个突出条带,表明只有部分转录物形成预测的发夹RNA结构。因此,ledRNA在形成预测的dsRNA结构中是最有效的。

[0512] ledRNA在叶片表面停留和扩散的能力也与dsRNA进行了比较。当施涂到烟叶表面的下部时,在24小时后,可以在未处理的烟叶上部中容易地检测到GUS ledRNA(ledGUS)(图3)。然而,在未处理的叶片上部中不能检测到单独的链GUS dsRNA(dsGUS)(图3)。该结果表明,ledRNA比dsRNA更耐降解,因此能够在植物叶组织内传播。

[0513] 实施例4:通过局部递送测试ledRNA

[0514] 在表达GFP或GUS报告基因的本氏烟(Nicotiana benthamiana)和普通烟草(Nicotiana tabacum)植物中分别测试了ledRNA在局部递送后诱导RNAi的能力。所述GFP和GUS靶序列和所述ledRNA编码构建体的序列分别显示在SEQ ID NO:7、8、4和5中。编码的RNA分子的核糖核苷酸序列如SEQ ID NO:1(GFPLedRNAGFP ledRNA)和2(GUS ledRNA)所示。

[0515] 为了便于将ledRNA可重复且均匀地施涂到叶片表面上,使用软漆刷将在25mM Tris-HCL,pH8.0,10mM $MgCl_2$和Silwet 77(0.05%)中浓度为75-100μg/ml的ledRNA施加到叶片的亚轴表面上。在施涂ledRNA后6小时和3天,取叶片样品用于分析靶向基因沉默。

[0516] 针对本氏烟(N.benthamiana)叶片的GFP和针对普通烟草(Nicotiana tabacum)GUS,ledRNA的施涂导致在处理后6小时,在mRNA(GFP)或蛋白活性(GUS)水平上相应靶基因活性的20-40%和40-50%的明显减少。然而,在该实验中,该降低在处理后3天没有持续。本发明人认为在3天的观察可能是由于转基因对dsRNA处理的一些非特异性应答或ledRNA的耗散量。然而,在单独的实验中,在ledRNA处理后24小时,在处理的和远端未处理的叶片区中都检测到GUS沉默(图4)。

[0517] 实施例5.ledRNA诱导的内源靶基因沉默

[0518] 在另一个实施例中,设计了ledRNA以靶向由内源基因即本氏烟(N.benthamiana)的FAD2.1基因编码的mRNA。靶FAD2.1 mRNA和ledFAD2.1编码构建体的序列分别显示在SEQ ID NO:9和6中。所编码的RNA分子的核糖核苷酸序列提供为SEQ ID NO:3(本氏烟

(N.benthamiana)FAD2.1ledRNA)。

[0519] FAD2.1 ledRNA构建体由以下组成:对应于FAD2.1 CDS(Niben101Scf09417g01008.1)的核苷酸678-379的反义序列的前半部分;对应于FAD2.1 CDS的nt378-242的第一反义环;对应于nt379-979的有义序列;对应于nt1115-980的第二反义环;对应于FAD2.1 CDS的nt979-nt679的反义序列的后半部分。

[0520] 平行使用来自先前实施例的ledGUS RNA作为阴性对照。在第一个实验中,测定靶基因沉默的FAD2.1mRNA水平和累积的C18:1脂肪酸(图5)。还测定了相关基因FAD2.2的活性水平。对于每个样品,将约3μg的总RNA用DNase处理并且使用寡dT引物在50℃下逆转录50分钟。反应在85℃下终止5分钟并且用水稀释至120μl。使用转子基因PCR仪,使用关于管家基因肌动蛋白的基因特异性引物,一式三份分析5μl的每个样品的FAD2.1和FAD2.2 mRNA的相对表达。在随后的实验中,Northern印迹杂交也用于证实通过局部施用的ledFAD2.1 RNA对FAD2.1基因的沉默(图6)。

[0521] 在2、4和10小时时间点,FAD2.1 mRNA显著降低至在用ledRNA处理的叶组织中几乎检测不到的水平(图5)。在该实验中,不清楚为什么在6小时的时间点,FAD2.1mRNA的水平没有降低那么多。在图6所示的重复实验中,强的FAD2.1下调发生在6和24小时,特别是在24小时时间点。与FAD2.1具有序列同源性的相关FAD2.2基因在2和4小时时间点也通过ledRNA显示下调(图5)。

[0522] 由于FAD2.1和FAD2.2编码脂肪酸Δ12去饱和酶,该去饱和酶使油酸去饱和成亚油酸,因此在用ledRNA处理的叶组织中测定这些脂肪酸的水平。在2、4和6小时的时间点,在ledRNA处理的叶组织中油酸(18:1)积累有明显的增加,这表明降低了FAD2酶的量(图5)。因此,qRT-PCR和脂肪酸组成分析显示ledRNA诱导的FAD2.1基因沉默。

[0523] 实施例6:包含G:U碱基对或错配核苷酸的发夹RNA的设计和测试

[0524] 靶向GUS RNA的修饰的发夹RNA

[0525] 如编码β-葡糖醛酸糖苷酶(GUS)的基因的报道基因提供了简单和方便的测定系统,其可用于测量包括植物细胞的真核细胞中的基因沉默效率(Jefferson等,1987)。因此,本发明人设计、生产和测试了一些修饰的发夹RNA以降低作为靶基因的GUS基因表达的能力,使用基因递送方法将发夹RNA提供给细胞,并将修饰的发夹RNA与常规的发夹RNA进行比较。在实验中用作对照的常规发夹RNA具有200个邻接碱基对长度的双链区,其中所有碱基对是规范碱基对,即G:C和A:U碱基对,在双链区中没有任何G:U碱基对,并且没有任何非碱基配对的核苷酸(错配),靶向GUS mRNA分子与修饰的发夹RNA相同的200nt区。形成双链区的有义和反义序列通过包括PDK内含子的间隔区序列共价连接(Helliwell等,2005;Smith等,2000),在剪接来自初级转录物的内含子后,提供长度为39或45个核苷酸的RNA环(取决于所用的克隆策略)。用于反义序列的DNA片段在5'端侧接XhoI-BamHI限制性酶切位点,在3'端侧接HindIII-KpnI限制性酶切位点,以便于克隆到表达盒中,每个有义序列分别侧接XhoI和KpnI限制位点。对于对照发夹和修饰的发夹,每个发夹RNA的200bp dsRNA区包括200个核苷酸的反义序列,其与来自蛋白质编码区内的野生型GUS序列完全互补。该反义序列,对应于SEQ ID NO:10的核苷酸13-212,是GUS开放阅读框(ORF)的核苷酸804-1003的互补序列(cDNA序列提供为SEQ ID NO:8)。因此GUS靶mRNA的长度大于1900nt。选择用于有义和反义序列的200个核苷酸的长度,其足够小以合理地便于使用合成的寡核苷酸合成DNA片段,

但也足够长以在Dicer加工后提供多个sRNA分子。作为ORF的一部分,该序列不太可能含有隐蔽剪接位点或转录终止位点。

[0526] 遗传构建体的制备

[0527] 分别使用含有XhoI和BamHI位点或HindIII和KpnI的寡核苷酸引物对GUS-WT-F(SEQ ID NO:52)和GUS-WT-R(SEQ ID NO:53)对200bp GUS ORF序列进行PCR扩增,以引入GUS序列的这些限制酶位点5'和3'。将扩增片段插入载体pGEM-T Easy,通过测序来验证正确核苷酸序列。相对于可操作连接的CaMV e35S启动子(Grave,1992),和ocs基因聚腺苷酸化/转录终止子(Ocs-T),通过用BamHI和HindIII消化切下GUS片段,并将其插入pKannibal的BamHI/HindIII位点(Helliwell和Waterhouse,2005),GUS序列以反义方向插入。所得载体被命名为pMBW606并且按5'至3'顺序包含35S::PDK内含子::反义GUS::Ocs-T表达盒。该载体是用作组装四种hpRNA构建体的基础载体的中间载体。

[0528] 仅具有规范碱基对的构建体hpGUS[wt]

[0529] 为了制备仅具有规范碱基对、编码用作对照的发夹RNA分子、命名为hpGUS[wt]的载体,将具有XhoI和KpnI的200bp GUS PCR片段从pGEM-T Easy质粒中切出,并插入pMBW606中35S启动子和PDK内含子之间的XhoI/KpnI位点。这产生了命名为pMBW607的载体,其含有35S::感测GUS[wt]::PDK内含子::反义GUS::OCS-T表达盒。通过用NotI消化切下该盒,并将其插入pART27的NotI位点(Gleave,1992),得到命名为hpGUS[wt]的载体,其编码靶向GUS mRNA的规范碱基配对的发夹RNA。

[0530] 当通过200nt有义和反义序列的杂交自退火时,该发夹具有对应于GUS序列的200个邻接碱基对的双链区。相对于GUS有义序列,表达盒中的有义和反义序列分别侧接位于5'和3'端的BamHI和HindIII限制性酶切位点。当转录时,对应于这些位点的核苷酸也能够杂交,使双链区在每端延伸6bp。在表达盒的转录和PDK内含子从初级转录物的剪接之后,预测在通过Dicer或其他RNA酶进行任何加工之前的发夹RNA结构具有39个核苷酸的环结构。包括其环的发夹RNA结构的核苷酸序列以SEQ ID NO:15提供,并且其折叠自由能被预测为-471.73kcal/mol。因此这是能量稳定的发夹结构。使用"RNAfold"(http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi)基于PDK内含子序列剪接后的核苷酸序列计算自由能。

[0531] 当从具有35S启动子和OCS-T终止子的表达盒转录时,将得到的发夹RNA包埋在较大的RNA分子中,在5'端添加8个核苷酸,在3'端添加大约178个核苷酸,而不考虑在3'端添加任何聚腺苷酸尾。由于修饰的发夹RNA使用相同的启动子-终止子设计,所以这些分子在5'和3'端也具有这些延伸。因此,剪接PDH内含子后的发夹RNA分子的长度为约630个核苷酸。

[0532] 包含G:U碱基对的构建体hpGUS[G:U]

[0533] 通过退火重叠的寡核苷酸GUS-GU-F(SEQ ID NO:54)和GUS-GU-R(SEQ ID NO:55)以及使用高保真LongAmp Taq聚合酶(新英格兰生物实验室(New England Biolabs),目录号M0323)进行的3'末端的PCR延伸,来组装一个DNA片段,该DNA片段包含相同的200个核苷酸有义序列,但相应的野生型GUS区的所有52个胞嘧啶核苷酸(C)被胸腺嘧啶核苷酸(T)取代。将扩增的DNA片段插入pGEM-T Easy载体,通过测序来验证正确核苷酸序列。(SEQ ID NO:11)。然后通过用XhoI和KpnI消化来切出包含修饰序列的DNA片段,并将其插入基本载体

pMBW606的XhoI/KpnI位点。这产生了命名为pMBW608的构建体,其包含表达盒35S::有义GUS[G:U]::PDK内含子::反义GUS::OCS-T。用NotI消化来切下该表达盒,并插入pART27的NotI位点,得到命名为hpGUS[G:U]的载体,其编码G:U碱基配对的发夹RNA分子。

[0534] 该盒编码靶向GUS mRNA的发夹RNA,当通过200nt有义和反义序列的杂交自退火时,其具有52个G:U碱基对(而不是hpGUS[wt]中的G:C碱基对)和148个规范碱基对,即双链区核苷酸的26％涉及G:U碱基对。hpGUS[G:U]中的148个规范碱基对与对照发夹RNA中的相同,在相应位置,包括49个U:A碱基对,45个A:U碱基对和54个G:C碱基对。双链区邻接规范碱基配对的最长延伸为9个碱基对。因此,hpGUS[G:U]的反义核苷酸序列在长度(200nt)和序列上与对照发夹RNAhpGUS[wt]的反义序列相同。在表达盒的转录和PDK内含子从初级转录物的剪接之后,预测在通过Dicer或其他RNA酶进行任何加工之前的发夹RNA结构具有45个核苷酸的环结构。包括其环的发夹结构的核苷酸序列以SEQ ID NO:16提供,并且其折叠自由能被预测为-331.73kcal/mol。对于hpGUS[wt],尽管52个G:U碱基对分别比hpGUS[wt]中的G:C碱基对弱得多,但这因此是能量上稳定的发夹结构。

[0535] 图7显示了修饰的GUS有义序列(SEQ ID NO:11的核苷酸9-208)与GUS靶基因(SEQ ID NO:14)的相应区的比对。

[0536] 包含每第四个核苷酸错配的核苷酸的构建体hpGUS[1:4]

[0537] 设计并组装包含相同的200bp有义序列但其中相应野生型GUS序列的每第4个核苷酸被替换的DNA片段。通过改变C'到G',G'到C',A'到T'和T'到A',使4个核苷酸编组的每编组的每第4个核苷酸(在位置4、8、12、16、20等处的核苷酸)被替换,留下其他核苷酸不变。这些取代都是易位替换,预期其对所得到的发夹RNA结构的去稳定作用大于转变替换。通过将重叠的寡核苷酸GUS-4M-F(SEQ ID NO:56)和GUS-4M-R(SEQ ID NO:57)退火并使用LongAmp Taq聚合酶进行3'端的PCR延伸,来组装DNA片段。将扩增的DNA片段插入pGEM-T Easy载体,通过测序来验证正确核苷酸序列。(SEQ ID NO:12)。然后通过用XhoI和KpnI消化来切出包含修饰序列的DNA片段,并将其插入基本载体pMBW606的XhoI/KpnI位点。这产生了命名为pMBW609的构建体,其包含表达盒35S::有义GUS[1:4]::PDK内含子::反义GUS::OCS-T。用NotI消化来切下该表达盒,并插入pART27的NotI位点,得到命名为hpGUS[1:4]的载体,其编码1:4错配的发夹RNA分子。

[0538] 该盒编码靶向GUS mRNA的发夹RNA,并且当通过有义和反义序列的杂交自退火时,其对于200nt反义序列的50个核苷酸具有错配,包括对于位置200处的核苷酸的错配。除位置200以外,发夹RNA的双链区在长度为199nt的有义和反义序列上具有150个规范碱基对和49个错配核苷酸对,即预测双链区核苷酸的24.6％被错配(不涉及碱基对)。在表达盒的转录和PDK内含子从初级转录物的剪接之后,预测在通过Dicer或其他RNA酶进行任何加工之前的发夹RNA结构具有45个核苷酸的环结构。包括其环的发夹结构的核苷酸序列以SEQ ID NO:17提供,并且其折叠自由能被预测为-214.05kcal/mol。对于hpGUS[wt],这因此是能量上稳定的发夹结构,除了错配的核苷酸。

[0539] 图8显示了修饰的GUS有义序列(SEQ ID NO:12的核苷酸9-208)与GUS靶基因(SEQ ID NO:14)的相应区的比对。

[0540] 构建体hpGUS[2:10],其中10个核苷酸中的9个和10个核苷酸被错配

[0541] 设计并组装包含相同的200bp有义序列但其中相应野生型GUS序列的每第九和第

十个核苷酸被替换的DNA片段。通过改变C'到G',G'到C',A'到T'和T'到A',使10个核苷酸编组的每编组的每第9和第10个核苷酸(在位置9、10、19、20、29、30等处的核苷酸)被替换,留下其他核苷酸不变。通过将重叠的寡核苷酸GUS-10M-F(SEQ ID NO:58)和GUS-10M-R(SEQ ID NO:59)退火并使用LongAmp Taq聚合酶进行3'端的PCR延伸,来组装DNA片段。将扩增的DNA片段插入pGEM-T Easy,通过测序来验证正确核苷酸序列。(SEQ ID NO:13)。然后通过用XhoI和KpnI消化来切出包含修饰序列的DNA片段,并将其插入基本载体pMBW606的XhoI/KpnI位点。这产生了命名为pMBW610的构建体,其包含表达盒35S::有义GUS[2:10]::PDK内含子::反义GUS::OCS-T。用NotI消化来切下该表达盒,并插入pART27的NotI位点,得到命名为hpGUS[2:10]的载体,其编码2:10错配的发夹RNA分子。

[0542] 该盒编码靶向GUS mRNA的发夹RNA,当通过有义和反义序列的杂交自退火时,其对于200nt反义序列的50个核苷酸具有错配,包括对于位置199和位置200处的核苷酸的错配。除位置199和位置200以外,发夹RNA的双链区在长度为198nt的有义和反义序列上具有160个规范碱基对和19个二核苷酸错配,即预测双链区核苷酸的19.2%被错配(不涉及碱基对)。hpGUS[2:10]中的160个碱基对与对照发夹RNA中的相同,在相应位置,包括41个U:A碱基对,34个A:U碱基对和42个G:C和43个G:C碱基对。在表达盒的转录和PDK内含子从初级转录物的剪接之后,预测在通过Dicer或其他RNA酶进行任何加工之前的发夹RNA结构具有45个核苷酸的环结构。包括其环的发夹结构的核苷酸序列以SEQ ID NO:18提供,并且其折叠自由能被预测为-302.78kcal/mol。对于hpGUS[wt],这因此是能量上稳定的发夹结构,除了预期从发夹结构的茎中隆起的错配的核苷酸。

[0543] 图9显示了修饰的GUS有义序列(SEQ ID NO:13的核苷酸9-208)与GUS靶基因(SEQ ID NO:14)的相应区的比对。

[0544] 图10示意性显示了用于表达对照和修饰的发夹RNA的四种遗传构建体。

[0545] 实施例7:检测转基因植物中修饰的发夹RNA

[0546] 用GUS靶基因转化的普通烟草(Nicotiana tabacum)物种的植物用于测试上述四种发夹RNA构建体的效力。具体地,靶植物来自两个纯合的独立的转基因品系PPGH11和PPGH24,各自含有来自载体pWBPPGH的GUS转基因的单拷贝插入,如图11示意性所示。pWBPPGH的T-DNA中的GUS基因具有可操作地连接到来自西葫芦(Cucurbita pepo L.cv.Autumn Gold)的韧皮部蛋白2(PP2)基因的1.3kb长启动子的GUS编码区(SEQ ID NO:8的核苷酸7-1812)(Wang等,1994;Wang,1994)。通过从λ基因组克隆CPP1.3(Wang,1994)中切除PP2启动子加上5'UTR和PP2蛋白编码区的54个核苷酸(编码PP2的前18个氨基酸),来构建构建体pWBPPGH,并融合该具有GUS编码序列的片段从编码GUS第3个氨基酸的核苷酸开始,产生具有GUS活性的N末端融合多肽。pPP2::将GUS:Nos-T盒插入pWBVec2a(Wang等,1998)以产生pWBPPGH,其用于转化普通烟草(Nicotiana tabacum)cv的植物。威斯康辛38使用根癌农杆菌(Agrobacterium tumefaciens)介导的叶盘转化(Ellis等,1987),选择对潮霉素的抗性。两个转基因品系PPGH11和PPGH24的纯合子代植物中的GUS活性相似。在两个转基因植物中的GUS表达不限于韧皮部,而是存在于植物的大部分组织中。因此,在这些植物中,来自PP2启动子的GUS表达似乎是组成型的。选择PP2-GUS植物作为测试植物有两个原因:i)它们给出与35S-GUS植物大约相同的组成性高水平的GUS表达;ii)所述PP2启动子是衍生自西葫芦(Cucurbita pepo)的内源PP2基因启动子,其具有与用于驱动hpRNA转基因表达的35S启

动子不同的序列,因此其不会受到输入的35S启动子的转录共抑制。

[0547] 所有4种发夹RNA构建体(实施例6)用于使用农杆菌(Agrobacterium)介导的叶盘方法(Ellis等,1987)转化PPGH11和PPGH24植物,使用50mg/L卡那霉素作为选择剂。用卡那霉素作为与以前使用的用于导入pWBPPGH的T-DNA的潮霉素不同的试剂的这种选择系统被观察到仅产生转化的植物,没有未转化的植物被再生。将含有来自hpGUS构建体的T-DNA的再生的转基因植物转移到土壤中用于在温室中生长,并在测定GUS活性之前保持约4周。当测定时,转基因植物是健康的和活跃生长的,并且在外观上与未转化的对照植物和亲本PPGHII和PPGH24植物相同。总共,获得了用编码hpGUS[wt]的T-DNA转化的59株转基因植物,获得了用编码hpGUS[G:U]的T-DNA转化的74株植物,获得了用编码hpGUS[1:4]的T-DNA转化的33株植物,获得了用编码hpGUS[2:10]的T-DNA转化的41株植物。

[0548] 使用荧光4-甲基伞形酮基β-D-葡糖苷酸(MUG)测定法(Jefferson等,1987),按照Chen等(2005)所述的改进的动力学方法测定GUS表达水平。通过从每株植物的3个不同叶中采集直径约为1cm的叶片样品,选择膨胀良好、健康且绿色的叶片来测定植物。注意测试植物与对照植物处于相同的生长和发育阶段。每个测定使用每个叶样品提取的5μg蛋白,并如Chen等(2005)所述测量MUG的裂解速率。

[0549] 代表性数据显示在图12中,示出每个独立转基因植物的GUS活性(测定中的MUG单位)。因为hpGUS[wt]构建体的数据显示一些植物显示出强沉默,其中活性降低至少90%,而其他较弱的沉默,在本文中,选择相对于对照植物的10%GUS活性作为用于将植物分类为两类并比较不同构建体的活性水平。

[0550] 使用10%活性水平作为强沉默的基准,编码规范碱基配对的hpGUS[wt]的遗传构建体在所测试的59株转基因植物中的32个中诱导了强GUS沉默(54.2%)。其他27株植物均显示降低的GUS活性,但相对于对照植物保留超过10%的酶活性,因此在本文中被认为显示弱沉默。具有该构建体的转基因植物显示GUS基因沉默程度的宽范围(图12),从小于1%到约80%活性保留,这对于常规发夹结构是典型的(Smith等,2000)。

[0551] 明显相反,hpGUS[G:U]构建体诱导跨越独立转基因品系的一致且均匀的沉默,其中测试的74株植物中的71株(95.9%)显示强烈的GUS沉默。再次不同的是,所有测试的33株hpGUS[1:4]植物显示降低的GUS活性水平,相对于对照植物仅8株(24%)产生<10%的GUS活性,并且其他25株被分类为具有较弱的沉默。这些结果表明该构建体诱导跨越转基因品系的较弱但更均匀的GUS下调水平。hpGUS[2:10]构建体表现得更像hpGUS[wt]构建体,在一些品系中诱导良好的沉默水平(41株中的28株,或68.3%)并且在剩余的13株植物中几乎没有或没有GUS沉默。

[0552] 当仅使用沉默品系(残余活性<10%)进行比较并计算平均GUS活性时,hpGUS[wt]植物显示出最高的平均沉默程度,其次是hpGUS[G:U]植物和hpGUS[2:10]植物(图13)。HPGUS[1:4]植物的GUS活性平均降低最小。GUS沉默的程度显示与衍生自四种不同hpRNA构建体的预测hpRNA结构的热力学稳定性的良好相关性(实施例6)。

[0553] 为了测试这些差异在子代植物中是否持续存在,将含有靶GUS基因(是纯合的)和hpGUS转基因(半合子的)的代表性转基因植物自花授粉。选择来自hpGUS品系的卡那霉素抗性子代植物,因此弃去缺乏hpGUS转基因的任何无效分离子。这确保了hpGUS转基因以纯合或杂合状态存在于所有子代中。测定子代植物的GUS活性,代表性数据示于图14。含有hpGUS

[wt]转基因的后代显然分为两类,即具有强GUS沉默的后代和显示弱沉默或没有沉默的子代。这些类型与上一代的表型很好地相关,表明靶基因沉默的程度是可遗传的。所测试的hpGUS[G:U]品系中的所有植物一致地显示出强沉默,而hpGUS[1:4]系中的植物一致地显示出较弱的沉默。本发明人得出结论,在亲代中观察到的表型通常保持在子代植物中。

[0554] 转化植物的Southern印迹杂交实验

[0555] 在用hpGUS[G:U]构建体产生的大量独立转基因植物中观察到的强基因沉默的均匀性是惊人的和出人意料的。本发明人试图确定除了由hpGUS[G:U]RNA引起的作用之外的任何解释是否引起沉默的均匀性。为了测试多个转基因植物是否产生自预期的独立转化事件,对分离自18株含有hpGUS[G:U]构建体的代表性转基因植物的DNA进行Southern印迹杂交实验。使用Wang等(2008)描述的热酚方法从叶片组织中分离DNA。对于Southern印迹杂交,将来自每个植物样品的约10μg的DNA用HindIII酶消化,通过在TBE缓冲液中的1%琼脂糖凝胶中的凝胶电泳分离,并且使用毛细管方法印迹到Hybond-N+膜上(Sambrook等,1989)。膜在42℃下与来自OCS-T终止子区域的^{32}P标记的DNA片段杂交过夜。选择该探针,因为它与hpGUS[G:U]转基因杂交,但不与不具有OCS-T终止子序列的GUS靶基因杂交。膜在高严格度下洗涤,保留的探针用磷酸成像仪可视化。

[0556] 杂交印迹的放射自显影图显示在图15中。每个泳道显示1-5或1-6个杂交带。没有两个泳道显示相同的模式,即放射自显影显示16个有代表性的hpGUS[G:U]植物各自具有不同的杂交后HindIII片段的模式,因此来自不同的转基因插入。本发明人得出结论,对于hpGUS[G:U]品系观察到的均匀的GUS沉默不是由于植物中相似的转基因插入模式,并且沉默的均匀性是由hpGUS[G:U]RNA的结构引起的。本发明人还得出以下结论:为了获得强的基因沉默,不需要hpGUS[G:U]转基因的多个拷贝;单拷贝的转基因就足够了。

[0557] 转化植物的Northern印迹杂交实验

[0558] 为了确定hpGUS[G:U]RNA是否以与转基因植物中对照发夹RNA相同的方式加工,对从转基因植物的叶片分离的RNA进行Northern印迹杂交实验。进行Northern印迹实验,检测发卡RNA经Dicer-加工后产生的较短RNA(sRNA,长度约21-24个核苷酸)。对分离自转基因HPGUS[wt]和HPGUS[G:U]植物的小RNA进行实验,所述小RNA还含有以(有义)mRNA表达的GUS靶基因。每个构建体选择9个植物用于sRNA分析。对于hpGUS[wt]转基因群体,包括显示弱GUS沉默的植物以及显示强GUS沉默的一些植物。使用热酚法分离小RNA样品(Wang等,2008),并且根据Wang等(2008)的描述进行Northern印迹杂交,在变性条件下进行RNA样品的凝胶电泳。使用的探针是与对应于SEQ ID NO:8的核苷酸804-1003的有义序列或反义序列相对应的32P标记的RNA。

[0559] 图16显示了与反义探针杂交(上版面)以检测衍生自发夹RNA的有义sRNA分子,或与有义探针杂交以检测反义sRNA(下版面)的Northern印迹的放射自显影。在底部,该图显示相对于缺乏hpGUS构建体的对照植物的GUS表达水平的定性评分。在其他实验中,基于与已知长度的RNA相比,sRNA的移动性,观察到与约20-25个核苷酸的小RNA杂交。hpGUS[wt]品系显示sRNA积累量的变化范围。对于有义和反义sRNA都观察到了这一点,尽管反义sRNA条带不如有义条带清晰。由于hpGUS[wt]植物同时含有表达与200nt靶区相对应的有义和反义序列的hpGUS转基因和表达全长有义基因的GUS靶基因,因此可以从发夹RNA或靶mRNA产生有义sRNA。hpGUS[wt]中sRNA水平与GUS沉默程度呈负相关。例如,泳道4和5中代表的两种植

物积累了相对更多的sRNA,但仅显示出中度程度的GUS下调。相反,泳道7和8中呈现的两种植物具有强GUS沉默,但累积相对低水平的sRNA。

[0560] 与hpGUS[wt]植物相反并且与hpGUS[G:U]构建体的相对一致的沉默程度一致,hpGUS[G:U]植物跨越品系累积一致量的反义sRNA。此外,GUS沉默的程度似乎与反义sRNA的量表现出良好的相关性。在这些植物中几乎没有检测到有义sRNA。这是预期的,因为在Northern印迹杂交中使用的RNA探针从野生型GUS序列转录,并且因此具有较低水平的互补性以从hpGUS[G:U]感测sRNA,其中所有C核苷酸被U核苷酸替代,仅允许较低严格性杂交。然而,该实验不排除加工hpGUS[G:U]RNA以产生较少有义sRNA或使它们更快地降解的可能性。

[0561] 重复Northern印迹杂交实验,这次仅使用有义探针检测反义sRNA;放射自显影显示在图17中。再次,来自hpGUS[wt]构建体的反义sRNA的产生与GUS活性负相关(图17的上版面)。强沉默的植物产生高水平的反义sRNA(泳道1、3、5、8和10),而仅显示弱沉默或无沉默的植物在本实验中未产生杂交信号(泳道2、4、6、7和9)。相反,表达hpGUS[G:U]的植物产生低得多但一致量的反义sRNA。观察到表达hpGUS[G:U]的强沉默的植物比表达hpGUS[wt]的强沉默的植物积累低得多的sRNA水平是令人感兴趣的并且提示发明人hpGUS[wt]在植物中通过不同的机制加工,但仍然与hpGUS[wt]构建体一样有效。在该实验中的进一步观察提供了线索,即对于hpGUS[G:U]植物的两个相对微弱的反义带似乎具有与对于来自hpGUS[wt]的反义sRNA带观察到的第二和第四带相同的移动性。这在下面描述的进一步实验中得到证实。本发明人假定来自hpGUS[wt]的sRNA的四个带代表24-、22-、21-和20-mer,并且主要加工hpGUS[G:U]RNA以产生22-和20-mer反义sRNA。

[0562] 来自上述数据的一个重要的明确的结论是hpGUS[G:U]RNA分子被一种或多种Dicer酶加工以产生sRNA,特别是反义sRNA的产生,所述反义sRNA被认为是在如Argonaute的各种蛋白质存在下RNA干扰的介质。所观察到的反义sRNA的产生暗示也产生了有义sRNA,但是该实验没有区分有义sRNA的降解/不稳定或由于与所使用的探针杂交不足而没有检测到有义sRNA。从这些实验,本发明人还得出以下结论:hpGUS[wt]和hpGUS[G:U]RNA分子在它们的加工过程中存在明显的差异。这表明分子被一个或多个Dicers不同地识别。

[0563] 实施例8:表达修饰发夹RNA的转基因植物sRNA分析

[0564] 进行另一个Northern印迹杂交实验以检测来自hpGUS[G:U]植物的反义sRNA,并将其大小与由hpGUS[wt]产生的那些进行比较。放射自显影显示在图18中。此时,与来自hpGUS[wt]的两个主要条带相比,来自hpGUS[G:U]的两个反义sRNA条带的大小差异更加明显。通过比较图18的相邻泳道9和10中条带的迁移率可以最好地看出这一点。该结果证实两种发夹RNA在植物中被一种或多种Dicers不同地加工。

[0565] 为了进一步研究这一点,通过对从植物分离的总的,接头可扩增的sRNA进行深度测序来分析来自hpGUS[wt]和hpGUS[G:U]的小RNA群体。测定了映射于发夹结构RNA双链区的sRNA的频率。并测定其长度分布。结果表明,与hpGUS[wt]构建体相比,hpGUS[G:U]构建体的22-mer反义RNA的频率增加。长度为22nt的sRNA比例的增加表明相对于hpGUS[wt],Dicer-2对hpGUS[G:U]发夹的加工有改变。

[0566] 实施例9:植物转基因的DNA甲基化分析

[0567] 对hpGUS[wt]赋予的GUS沉默程度的可变性的观察以及在hpGUS[wt]植物中检测到反义24-mer sRNA,但在hpGUS[G:U]植物中明显未检测到反义24-mer sRNA,导致本发明人

质疑两种植物群体在靶GUS基因的DNA甲基化水平上是否不同。序列特异性24-mer sRNA被认为参与促进植物中反向重复结构的DNA甲基化(Dong等,2011)。因此,本发明人测试了hpGUS植物中GUS转基因的DNA甲基化水平,特别是发夹编码基因(沉默基因)的35S启动子区的DNA甲基化水平。

[0568] 为此,使用DNA甲基化依赖性核酸内切酶McrBC。McrBC是可商购的核酸内切酶,其在双链DNA的一条或两条链上切割含有甲基胞嘧啶(mC)碱基的DNA(Stewart等,2000)。McrBC识别DNA上的两个位点,这些位点由5'(G或A)mC 3'的两个半位点组成,优选GmC。这些半位点可以被几百个碱基对分开,但最佳的间隔为55至约100bp。在两条链上具有这样的连接的GmC二核苷酸的双链DNA是最好的底物。McrBC活性依赖于一个或两个被甲基化的GC二核苷酸。由于植物DNA可以在CG、CHG或CHH序列中的C处甲基化,其中H代表A、C或T(Zhang等,2018),因此使用McrBC消化DNA并随后PCR扩增基因特异性序列可以用于检测植物基因组中特异性DNA序列中mC的存在或不存在。在该测定中,甲基化的McrBC消化的基因组DNA的PCR扩增与未甲基化的DNA相比产生减少量的扩增产物,但是如果DNA未甲基化,将产生与未处理的DNA相同量的PCR产物。

[0569] 通过标准方法从包含除了靶GUS基因之外的hpGUS[wt]、hpGUS[G:U]或hpGUS[1:4]构建体的植物中分离基因组DNA(Draper和Scott,1988)。根据制造商的说明书,包括存在内切核酸酶活性所需的Mg^{2+}离子和GTP,纯化的DNA样品用McrBC(目录号M0272;马萨诸塞州新英格兰生物实验室)处理。总之,用McrBC在30μl反应体积中消化大约1μg基因组DNA过夜。将消化的DNA样品稀释至100μl,并如下进行PCR扩增所关注的区。

[0570] 将处理的DNA样品用于使用以下引物的PCR反应。对于hpGUS[wt]的35S-GUS连接序列:正向引物(35S-F3),5'-TGGCTCCTACAAATGCCATC-3'(SEQ ID NO:60);反向引物(GUSwt-R2,5'-CARRAACTRTTCRCCCTTCAC-3'(SEQ ID NO:61)。对于hpGUS[G:U]的35S-GUS连接序列:正向引物(GUSgu-R2),5'-CAAAAACTATTCACCCTTCAC-3'(SEQ ID NO:62);反向引物(GUS4m-R2),CACRAARTRTACRCRCTTRAC(SEQ ID NO:63)。对于两种构建体的35S启动子序列:正向引物(35S-F2,反向,5'-GAGGATCTAACAGAACTCGC-3'(SEQ ID NO:64);反向引物(35S-R1),5'-CTCTCCAAATGAAATGAACTTCC-3'(SEQ ID NO:65)。在每种情况下,R=A或G,Y=C或T。用以下循环条件进行PCR反应:94℃1分钟,以下进行35个循环:94℃30秒,55℃退火45秒,68℃延伸1分钟,和最终在68℃延伸5分钟。对PCR扩增产物进行电泳并定量条带的强度。

[0571] 代表性结果示于图19和20中。对于包括包含转录起始位点的35S启动子序列的200bp的35S-GUS连接区,大部分hpGUS[wt]植物显示显著水平的DNA甲基化。在hpGUS[wt]植物群体内,保持较高水平的GUS活性即较少沉默的个体植物似乎具有启动子-GUS有义连接区的更多甲基化。35S启动子区的结果是相似的。相反,大部分hpGUS[G:U]和hpGUS[1:4]植物在35S-GUS连接处显示较弱的DNA甲基化。本发明人认为该近侧启动子序列对于转基因的表达是重要的,并且在该区的甲基化将可能通过转基因的转录基因沉默(TGS)降低沉默构建体的表达。这称为"自身沉默"。

[0572] 关于实施例6-9的一般讨论

[0573] 转基因中反向重复DNA结构的破坏增强其稳定性

[0574] hpGUS[wt]和hpGUS[2:10]转基因植物群体均表现出广泛的靶基因沉默程度。相反,含有hpGUS[G:U]和hpGUS[1:4]植物的两个群体在许多独立的品系中都显示相对均匀的

GUS沉默,其中前者的构建体观察到强的沉默,后者的构建体观察到虽然相对较弱但基因活性仍然显著降低。在来自[G:U]和[1:4]构建体的发夹RNA中,有义和反义序列中约25%的核苷酸参与G:U碱基对或参与在200个核苷酸有义/反义序列上均匀分布的序列错配。由于有义和反义序列之间的序列差异,认为有义和反义"臂"或反向请求结构之间的DNA构建体错配显著破坏反向重复DNA结构。重复的DNA结构可以吸引各种生物体中的DNA甲基化和沉默(Hsieh和Fire,2000)。hpGUS[2:10]构建体也包含有义和反义区的错配,但是有义和反义序列之间的2bp错配的每一个侧接8-bp邻接错配,因此错配可能不会像在[G:U]和[1:4]转基因中那样破坏反向重复DNA结构。因此,由hpGUS[G:U]和hpRNA[1:4]诱导的GUS沉默的均一性可能至少部分是由于破坏了反向重复DNA结构,这导致较低的甲基化并且因此降低了这两种转基因的自身沉默。有义和反义DNA区之间错配的另一个好处是对大肠杆菌中反向重复的克隆有帮助,因为细菌倾向于删除或重排完美的反向重复。

[0575] hpRNA的热力学稳定性是影响靶基因沉默程度的重要因素

[0576] 当只比较强沉默的转基因品系时,hpGUS[wt]植物的靶基因下调程度最大,其次是hpGUS[G:U],hpGUS[2:10]和hpGUS[1:4]。RNAFold分析预测hpGUS[wt]发夹RNA结构具有最低的自由能,即最大的稳定性,接着是hpGUS[G:U],hpGUS[2:10]和hpGUS[1:4]发夹。本发明人认为发夹RNA结构越稳定,它可以诱导的靶基因沉默的程度就越大。这也有利于较长的双链RNA结构而不是较短的双链RNA结构。认为有效的Dicer加工需要稳定的双链RNA形成。在此描述的实验结果表明G:U碱基配对构建体相对于诸如hpGUS[1:4]的包含大多数简单错配核苷酸的构建体的另一个重要优点,虽然两种类型的构建体都破坏了反向重复DNA结构,这降低了自身沉默,但是在RNA水平,由于G和U形成碱基对的能力,hpGUS[G:U]RNA更稳定。两种类型的修饰的组合也被认为是有益的,包括双链RNA结构中的G:U碱基对和一些错配的核苷酸,但是在G:U碱基对中涉及的核苷酸比在错配中涉及的核苷酸相对多至少2、3、4或甚至5倍。

[0577] Dicer能高效地加工hpGUS[G:U]RNA

[0578] 这些实验中所回答的一个重要问题是错配的还是G:U碱基配对的hpRNA可以被Dicer加工成小RNA(sRNA)。hpGUS[G:U]植物以及1:4和2:10错配的hpRNA植物中的强沉默,暗示这些发夹RNA结构由Dicer加工。通过易于检测反义sRNA的sRNA Northern印迹杂交对[G:U]分子证实了这一点。此外,HPGUS[G:U]植物中的GUS沉默程度与积累的反义sRNA的量表现出良好的相关性。来自每个品系的两个选择品系的小RNA深度测序分析(仅一个用于hpGUS[wt]),证实hpGUS[G:U]植物,如hpGUS[wt]植物,产生大量的sRNA,而hpGUS[1:4]植物也产生sRNA,但是具有低得多的丰度(图21)。hpGUS[1:4]植物中较低水平的sRNA与GUS沉默效率相对较低相符,这表明dsRNA茎在hpGUS[1:4]RNA中的低热力学稳定性降低了Dicer加工效率。注意到GUS沉默的程度显示与hpGUS[wt]构建体的sRNA水平的相关性相对差,一些强沉默的品系含有相对少量的sRNA。这表明在一些hpGUS[wt]品系中的GUS沉默至少部分归因于转录沉默而不是sRNA引导的PTGS。本发明人认识到通过使用修饰的发夹RNA构建体,特别是G:U构建体,减少了诸如启动子区的参与基因序列甲基化的发夹编码基因的自身沉默。

[0579] G:U和1:4hpRNA转基因显示近端35S启动子区的DNA甲基化降低

[0580] McrBC酶切-PCR分析表明,相对于hpGUS[wt]群体,hpGUS[G:U]和hpGUS[1:4]中的转录起始位点(TSS)附近的240bp 35S序列中的DNA甲基化水平降低。该结果向发明人表明,

由于有义序列中的C至T修饰(在hpGUS[G:U]中)或25％核苷酸错配(在hpGUS[1:4]中),完美的反向重复结构的破坏使hpRNA转基因的转录自身沉默最小化。这与相对于hpGUS[wt]群体在hpGUS[G:U]和hpGUS[1:4]群体中观察到的GUS基因沉默的均匀性一致。本发明人认识到hpGUS[G:U]构建体在降低启动子甲基化方面比hpGUS[1:4]构建体更理想,因此至少因为其在有义序列中具有减少的数目或甚至缺失的胞嘧啶核苷酸,并且因此不吸引可以扩散至启动子的DNA甲基化。

[0581] 实施例10:靶向内源基因的包含G:U碱基对的发夹RNA的设计和测试

[0582] 修饰的发夹RNA靶向EIN2和CHS RNA

[0583] 由于与如上所述的常规发夹RNA相比,G:U修饰的发夹RNA似乎诱导靶基因更一致和均匀的沉默,本发明人希望测试改进的设计是否也会降低内源基因的表达。因此,本发明人设计,生产和测试了靶向EIN2或CHS基因或两者的几种[G:U]修饰的发夹RNA构建体,它们是拟南芥(Arabidopsis thaliana)中的内源基因,被选为用于尝试沉默的示例性靶基因。EIN2基因(SEQ ID NO:19)编码乙烯不敏感蛋白2(EIN2),其是由植物信号分子乙烯调节的信号传导途径中的中心因子,即调节蛋白,并且CHS基因(SEQ ID NO:20)编码查耳酮合酶(CHS),其参与拟南芥(A.thaliana)中花色素苷的产生。产生同时靶向EIN2和CHS基因的另一G:U修饰的构建体,其中转录融合EIN2和CHS序列以产生单一发夹RNA。此外,制备靶向EIN2、CHS或EIN2和CHS两者的三个另外的构建体,其中有义和反义序列中的胞嘧啶碱基均被胸腺嘧啶碱基(本文中称为G:U/U:G构建体)替代,而不仅仅是如针对靶向GUS的修饰的发夹所进行的有义序列。使用基因递送方法向细胞提供发夹RNA,测试修饰的发夹RNA构建体降低内源EIN2基因或EIN2和CHS基因表达的能力。在本实验中用作对照的常规发夹RNA具有单独靶向EIN2或CHS mRNA的长度为200个碱基对的双链RNA区,或包含来自融合在一起作为单一发夹分子的EIN2和CHS基因的每一个的200个碱基对的嵌合双链RNA区。在融合的RNA中,EIN2双链部分邻近发夹环,CHS区远离发夹环。对照发夹RNA的双链区中的所有碱基对都是规范碱基对。

[0584] 构建体制备

[0585] 分别使用寡核苷酸引物对EIN2wt-F(SEQ ID NO:66)和EIN2wt-R(SEQ ID NO:67)或CHSwt-F(SEQ ID NO:68)以及CHSwt-R(SEQ ID NO:69),从拟南芥(Arabidopsis thaliana)Col-0cDNA进行PCR扩增野生型EIN2(SEQ ID NO:19)和CHS cDNA(SEQ ID NO:20)200bp区的DNA片段。关于GUS发夹构建体,将片段插入pGEMT-Easy中(实施例6)。包含200bp经修饰的有义EIN2[G:U](SEQ ID NO:22)和CHS[G:U](SEQ ID NO:24)片段或200bp经修饰的反义EIN2[G:U](SEQ ID NO:25)和经修饰的反义CHS[G:U](SEQ ID NO:26)片段的DNA片段,其各自侧接限制酶位点,通过将相应的寡核苷酸对EIN2gu-F+EIN2gu-R,CHSgu-F+CHSgu-R,asEIN2gu-F+asEIN2gu-R和asCHSgu-F+asCHSgu-R(SEQ ID NO:70-77)退火,随后使用LongAmp Taq聚合酶进行3'端的PCR延伸来组装。所有G:U修饰的PCR片段均克隆入pGEM-T Easy载体,通过测序验证目的核苷酸序列。CHS[wt]::EIN2[wt],CHS[G:U]:EIN2[G:U],asCHS[G:U]::asEIN2[G:U]融合片段通过在pGEM-T Easy质粒的共同XbaI位点连接合适的CHS和EIN2 DNA片段制备。

[0586] 35S::有义片段::PDK内含子::反义片段::OCS-T盒以类似于hpGUS构建体的方式制备。基本上,反义片段通过用HindIII和BamHI消化从相应的pGEM-T Easy质粒中切除,并

插入BamHI和HindIII位点之间的pKannibal中,因此它们相对于35S启动子处于反义方向。然后使用XhoI和KpnI从相应的pGEM-T Easy质粒中切除有义片段,并将其插入适当的含反义克隆的相同位点。然后用noti切除pGEM-T Easy质粒中的所有盒并插入pART27以形成用于植物转化的最终二元载体。

[0587] 图22-25显示了修饰的有义[G:U]和反义[G:U]核苷酸序列与相应的野生型序列的比对,显示了取代的核苷酸的位置。用于发夹RNA的表达盒的设计示意性地示出在图26中。

[0588] 用FOLD程序估算了发卡RNA形成的自由能。这些计算为(kcal/mol):hpEIN2[wt],-453.5;hpEIN2[G:U],-328.1;hpCHS[wt],-507.7;hpCHS[G:U]-328.5;hpEIN2[G:U/U:G],-173.5;hpCHS[G:Y/U:G],-186.0;hpCHS::EIN2[wt],-916.4;hpCHS::EIN2[G:U],-630.9;hpCHS::EIN2[G:U/U:G),-333.8。

[0589] 植物转化

[0590] 使用花浸法将所有EIN2、CHS和嵌合EIN2/CHS构建体用于转化拟南芥(Arabidopsis thaliana)种Col-0植物(Clough和Bent,1998)。为了选择转基因植物,将从农杆菌(Agrobacterium)浸渍的花中收集的种子用氯气灭菌并涂布在含有50mg/L卡那霉素的MS培养基上。所有9个构建体获得了多个转基因品系(表1)。将这些初级转化体(T1代)转移到土壤中,自花授粉并生长至成熟。从这些植物收集的种子(T2种子)用于建立T2植物并筛选对于转基因是纯合的品系。这些用于分析EIN2和CHS沉默。

[0591] 表1.Col-0背景中获得的转基因植物的汇总

[0592]

构建体	获得的转基因品系的数目
hpEIN2[wt]	46
hpCHS[wt]	34
hpEIN2[G:U]	23
hpCHS[G:U]	32
hpEIN2[G:U/U:G]	52
hpCHS[G:U/U:G]	13
hpCHS::EIN2[wt]	28

[0593]

hpCHS::EIN2[G:U]	26
hpCHS::EIN2[G:U/U:G]	20

[0594] EIN2沉默程度分析

[0595] EIN2是拟南芥(A.thaliana)中的基因,其编码参与乙烯感知的受体蛋白。该基因在种子萌发后的幼苗中表达,在植物生长发育过程中也有表达。当在植物中乙烯合成的中间体1-氨基环丙烷-1-羧酸(ACC)存在下在黑暗中发芽时,EIN2突变体幼苗相对于等基因野生型幼苗显示下胚轴伸长。因此,与野生型幼苗相比,通过在包含总黑暗中50μg/L的ACC的MS培养基上发芽种子并测量它们的下胚轴长度来测定转基因植物中的EIN2基因表达和沉

默程度。下胚轴长度是一种易于测量的表型,是EIN2基因表达减少程度的良好指标,表明不同水平的EIN2沉默。根据EIN2沉默的水平,预期具有沉默的EIN2基因表达的植物具有不同程度的下胚轴伸长,在野生型幼苗(短下胚轴)和无效突变幼苗(长下胚轴)之间的范围内。测定来自每个构建体的20株随机选择的独立转化的植物的种子。来自20株含有hpCHS::EIN2[G:U]构建体的一株植物的种子没有发芽。下胚轴长度的数据显示在图27中。

[0596] hpEIN2[wt]品系在EIN2沉默的程度上显示了相当大的范围,其中7个品系(图27中的植物品系2、5、9、10、12、14、16)清楚地显示相对于野生型低水平的沉默或相同的下胚轴长度,并且其他13个品系具有中度至强烈的EIN2沉默。每个独立品系内的单株植物倾向于表现出EIN2沉默程度的范围,如下胚轴长度的差异所示。相反,仅包含hpEIN2[G:U]构建体的两个品系(图27中的植物品系5,18)显示弱的EIN2沉默,其余18显示均匀的强的EIN2沉默。此外,与用hpEIN2[wt]构建体转化的植物相比,在18个品系的每一个中的单个植物似乎具有相对均匀的EIN2沉默。本发明人得出结论,G:U修饰的发夹RNA构建体能够赋予内源基因更一致的较少变化的基因沉默,其比靶向内源RNA的相同区的常规发夹RNA更均匀和更可预测。

[0597] 转基因hpEIN2[wt]和hpEIN2[G:U]群体的EIN2沉默程度与转基因拷贝数的关系也不同。通过子代植物中卡那霉素抗性标记基因的分离比率指示转基因拷贝数-抗性:易感幼苗比率为3:1指示单基因座插入,而更高的比率指示多基因座转基因插入。用hpEIN2[wt]构建体转化的几个多拷贝数品系显示低水平的EIN2沉默,但这不是hpEIN2[G:U]品系的情况,其中单拷贝和多拷贝基因座均显示强EIN2沉默。

[0598] EIN2基因也在CHS::EIN2融合发夹RNA转化的幼苗中沉默。类似于含有单个hpEIN2[G:U]构建体的植物,hpCHS::EIN2[G:U]幼苗清楚地显示与hpCHS::EIN2[wt]幼苗相比,在独立品系之间更均匀的EIN2沉默。对于hpCHS::EIN2[G:U],独立品系内的个体植物之间的沉默也似乎比hpCHS::EIN2[wt]品系更均匀。同时,高度沉默的hpCHS::EIN2[wt]植物的EIN2沉默程度比hpCHS::EIN2[G:U]植物略强,类似于用hpGUS[wt]和hpGUS[G:U]转化的植物之间的比较。沉默程度的比较表明,融合构建体没有诱导比单个hpEIN2[G:U]构建体更强的EIN2沉默,实际上,融合G:U发夹构建体似乎诱导比单个基因靶向的hpEIN2[G:U]构建体稍弱的EIN2沉默。

[0599] 当检查用G:U/U:G构建体转化的植物时,其中将有义和反义序列的胞嘧啶(C)核苷酸修饰为胸腺嘧啶(T)核苷酸,与野生型植物相比,对于所分析的所有20个独立品系观察到下胚轴长度几乎没有增加。这对于hpEIN2[G:U/U:G]和hpCHS::EIN2[G:U/U:G]构建体均观察到。这些结果表明本发明人,具有约46%取代的G:U/U:G碱基配对的发夹RNA构建体不能有效诱导靶基因沉默,可能是因为发夹RNA的碱基配对过于不稳定。发明人认为两个可能的原因可能导致无效。首先,发夹RNA的EIN2双链区在有义和反义序列之间具有200个潜在碱基对的92个G:U碱基对。其次,修饰的反义序列与野生型有义序列的互补序列的比对显示,反义序列中的49C至T置换可能降低反义序列靶向EIN2 mRNA的有效性。本发明人从该实验得出的结论是,至少对于EIN2靶基因,在发夹RNA中可以耐受的并且仍然保持足够的沉默效果的核苷酸取代的数量存在上限。例如,92/200=46%的取代可能是百分比过高的。

[0600] CHS沉默程度分析

[0601] 通过定量逆转录PCR(qRT-PCR)对在体外生长于组织培养基上的从整个植物提取

的RNA测定转基因植物的CHS基因表达水平。用于CHS mRNA的引物为:正向引物(CHS-200-F2),5'-GACATGCCTGGTGCTGACTA-3'(SEQ ID NO:78;反向引物(CHS-200-R2)5'-CCTTAGCGATACGGAGGACA-3'(SEQ ID NO:79)。用作标准参考基因肌动蛋白(Actin)2的引物为:正向引物(Actin2-For)5'-TCCCTCAGCACATTCCAGCA-3'(SEQ ID NO:80),和反向引物(Actin2-Rev)5'-GATCCCATTCATAAAACCCCAG-3'(SEQ ID NO:81)。

[0602] 数据显示,相对于Actin2基因的参考mRNA,在植物中累积的CHS mRNA的水平在50-96%的范围内降低(图28)。

[0603] 与野生型种子的褐色相比,完全缺乏CHS活性的拟南芥(A.thaliana)种子具有淡的种皮颜色。因此,目测转基因植物的种子的种皮颜色。在来自几种植物的种子中观察到种皮颜色的明显减少,而在其他植物中没有观察到种皮颜色的明显减少,尽管在那些植物的叶子中CHS mRNA减少。然而,认为只有当在植物生长期间发育中的种皮中CHS活性几乎完全消除时才显示种皮颜色表型。此外,35S启动子在发育中的种皮中可能没有足够的活性来提供CHS活性降低的水平以提供在无效突变体中见到的淡种子表型。通过使用在种子的种皮中更有活性的启动子可以获得视觉种皮颜色表型的改善。

[0604] 实施例11:分析来自发夹RNA构建体的sRNA

[0605] 对RNA样品进行Northern印迹杂交以检测来自hpEIN2[G:U]植物的反义sRNA,并将其数量及其大小与hpEIN2[wt]产生的sRNA进行比较。该探针是与hpEIN2[wt]构建体中的200个核苷酸有义序列相对应的^{32}P-标记的RNA探针,并且在低严格条件下进行杂交以允许检测更短(20-24个核苷酸)的序列。来自探测的Northern印迹的放射自显影显示在图29中。该实验显示hpEIN2[G:U]发夹RNA被加工成sRNA,并且与hpEIN2[wt]品系的那些相比,在所分析的9个独立转化的hpEIN2[G:U]植物中的积累水平是相对均匀的。类似于对GUS发夹RNA的类似实验,来自hpEIN2[G:U]的两个反义sRNA条带与来自hpEIN2[wt]的主要两个条带的大小差异是十分明显的。通过比较图28的相邻泳道10和11中条带的迁移率可以最好地看出这一点。

[0606] 为了进一步研究这一点,通过对从整个植物分离的总sRNA群体进行深度测序来分析来自hpEIN2[wt]和hpEIN2[G:U]的小RNA群体。测定映射至hpEIN2[wt]和hpEIN2[G:U]的双链区的每个群体的比例。在每个群体中约16000000个读取中,约50,000个sRNA映射到hpEIN2[wt]双链区,而仅约700个映射到hpEIN2[G:U]。这表明从[G:U]发夹产生了更少的sRNA。还观察到EIN2特异性22-mer的比例增加。

[0607] 实施例12 EIN2沉默植物的DNA甲基化分析

[0608] GUS和EIN2沉默结果表明,与具有提供G:U碱基对的修饰的有义序列的构建体相比,具有未修饰的有义序列的hpRNA构建体诱导高度可变水平的靶基因沉默。如上所述,与hpGUS[wt]构建体相比,hpGUS[G:U]构建体的启动子区似乎具有较小的甲基化。为了测试DNA甲基化并比较hpEIN2[wt]和hpEIN2[G:U]转基因植物,使用McrBC方法分析来自每个群体的12株植物在35S启动子和35S-启动子-有义EIN2连接区的DNA甲基化。用于35S启动子区的引物:正向引物(Top-35S-F2),5'-AGAAAATYTTYGTYAAYATGGTGG-3'(SEQ ID NO:82),反向引物(Top-35S-RyGTYAAYATGGTGG-3′)(SEQ ID NO:82),反向引物(Top-35S-R2),5'-TCARTRRARATRTCACATCAATCC-3'(SEQ ID NO:83)。用于35S启动子-有义EIN2连接区的引物:正向引物(Link-35S-F2),5'-YYATYATTGYGATAAAGGAAAGG-3'(SEQ ID NO:84),和反向引物

(Link-EIN2-R2),5'-TAATTRCCACCAARTCATACCC-3'(SEQ ID NO:85)。在这些引物序列的每一个中,Y=C或T并且R=A或G。

[0609] 通过进行实时PCR分析测定DNA甲基化程度的定量。计算每株植物:用McrBC处理基因组DNA后的DNA片段扩增率/不用McrBC处理基因组DNA后的DNA片段扩增率的商。

[0610] 几乎每个hpEIN2[wt]植物在35S启动子处,特别是在35S-EIN2接头处,显示出显著水平的DNA甲基化,但一些高于其他。如图30和31所示,泳道1、4、7、9、11和12中代表的植物品系都显示出如较长下胚轴长度所示的强EIN2沉默。相反,泳道2、3、5、6、8和10中代表的其他6个品系显示相对弱的EIN2沉默,导致较短的下胚轴。这些较弱沉默的品系在启动子和连接序列上显示出更多的DNA甲基化,如在用McrBC预消化基因组DNA时低得多的PCR条带强度所指示的。定量实时PCR(qPCR)测定证实了这些观察结果(图31)。所有12个测试品系在35S启动子区和35S-有义连接区均有一定程度的DNA甲基化。对于hpEIN2[wt]品系2、3、5、6、8和10,甲基化的最大程度,即qPCR测定中的最低商,与通过下胚轴长度测量的沉默程度降低完全相关。这些结果证实,在一些hpEIN2[wt]品系中降低的EIN2沉默与增加的启动子甲基化相关。即使在EIN2沉默的hpEIN2[wt]植物品系中,DNA甲基化水平仍然相当高,特别是35S-有义EIN2连接片段区。当启动子被甲基化时,这被认为引起转录沉默。在沉默构建体的情况下,如此,这是"自我沉默"的形式。

[0611] 与hpEIN2[wt]品系相反,hpEIN2[G:U]品系显示在35S启动子和35S-EIN2连接处较少的DNA甲基化。实际上,这些12个G:U品系中的4个,对应于图30中的泳道1、2、3和7(图31中的泳道13、14、15和20),没有明显的DNA甲基化,如McrBC处理的和未处理的样品之间的相等强度的PCR条带所指示的。当通过qPCR定量这些扩增时,12个品系中的6个品系显示来自McrBC处理的片段几乎没有到没有的减少,因此几乎没有到没有的DNA甲基化-参见图31的下版面,品系13、14、15、18、19和20。这些结果表明,至少在某些品系中,hpEIN2[G:U]构建体相对均匀的EIN2沉默是由于与hpEIN2[wt]相比显着更少的启动子甲基化和转录自沉默。

[0612] 这些结论通过亚硫酸氢盐测序分析转基因植物品系的基因组DNA得到进一步证实。该测定利用以下事实:用亚硫酸氢盐处理DNA将DNA中的未甲基化胞嘧啶碱基转化为尿嘧啶(U)过量,但留下5-甲基胞嘧啶碱基(mC)不受影响。亚硫酸氢盐处理后,在PCR反应中以仅扩增处理的DNA的有义链的方式扩增所限定的目的DNA片段。然后对PCR产物进行批量测序,揭示DNA片段中单个胞嘧啶碱基甲基化的位置和程度。因此,该测定产生了关于DNA片段的甲基化状态的单核苷酸解析信息。

[0613] 通过亚硫酸氢盐测序分析显示hpEIN2[wt]和hpEIN2[G:U]中的每一个的最强水平的EIN2沉默的三种植物品系,对应于图31中的hpEIN2[wt]品系1、7和9以及hpEIN2[G:U]品系13、15和18。这些植物品系显示出最长的下胚轴长度,因此预期每种构建体具有20种品系中最低的DNA甲基化水平。hpEIN2[wt]和hpEIN2[G:U]的结果分别示于图32和33中。当比较时,清楚的是,hpEIN2[wt]植物的35S启动子区和EIN2有义区中的许多胞嘧啶被广泛甲基化。相反,3种hpEIN2[G:U]植物品系在35S启动子区显示低得多的胞嘧啶甲基化水平。

[0614] 实施例13:hpGUS[1:4]构建体的启动子中的DNA甲基化水平

[0615] 当使用如上所述的McrBC和亚硫酸氢盐方法分析从hpGUS[1:4]植物分离的基因组DNA的DNA甲基化时,类似地观察到相对于hpGUS[wt]植物,在35S启动子和35S启动子-GUS有义序列区中胞嘧啶碱基的甲基化较少。

[0616] 关于实施例10-13的一般讨论

[0617] 具有G:U碱基对的双链RNA比常规dsRNA诱导更均匀的基因沉默

[0618] 与GUS构建体类似,hpEIN2[G:U]和hpCHS:EIN2[G:U]均比编码常规发夹RNA的相应hpRNA[wt]构建体诱导更一致和均匀的EIN2沉默。该一致性不仅发生在许多独立的转基因品系之间,而且也发生在各自具有相同转基因插入的转基因品系内的同胞植物之间。除了均匀性之外,hpEIN2[G:U]诱导的EIN2沉默程度接近强沉默的hpEIN2[wt]品系。对CHS基因沉默的分析表明,hpCHS[G:U]构建体有效地降低了50-97%的CHS mRNA水平,但几乎没有植物显示出明显可见的种皮颜色减少的表型。对于种皮颜色中看不到更多可见表型的可能解释是,即使低水平的CHS活性也可能足以产生类黄酮色素。其他可能的解释是35S启动子在发育中的种皮中没有足够的活性以产生表型,或者相比于EIN2序列仅有43个(21.5%)且GUS序列仅有52个(26%),hpCHS[G:U]构建体序列含有65个胞嘧啶取代(32.5%)。此外,CHS序列中的这些胞嘧啶碱基中的许多出现在两个或三个邻接胞嘧啶的组中,因此不是所有的胞嘧啶碱基都需要被取代。当有义链中的所有胞嘧啶被取代时,这导致hpCHS[G:U]RNA中的G:U碱基对比hpEIN2[G:U]和hpGUS[G:U]RNA中的G:U碱基对更多,可能超过最佳。为了验证这一点,使用含有一定范围胞嘧啶取代的序列制备另一组CHS构建体,范围为约5%、10%、15%、20%或25%胞嘧啶碱基取代。测试这些构建体并确定最佳水平。

[0619] hpEIN2[G:U]品系表达更均匀水平的siRNA

[0620] 与较均匀的EIN2基因沉默一致,hpEIN2[G:U]品系以更均匀的水平在独立品系中累积sRNA。这证实用hpGUS构建体的结论是[G:U]修饰的hpRNA被Dicer有效地加工并且能够诱导有效的靶基因沉默。

[0621] 融合构建体也提供基因沉默

[0622] 在实验中包括CHS:EIN2融合构建体的目的是检测两个靶基因是否可以用单个发夹编码构建体沉默。GUS实验表明,发夹结构RNA的自由能及其稳定性与靶基因沉默程度呈正相关。结果表明,CHS:EIN2融合构建体至少在mRNA水平上导致CHS的两种基因的沉默。

[0623] 两种hpRNA构建体hpEIN2[G:U/U:G]和hpCHS:EIN2[G:U/U:G],其中有义和反义序列均从C修饰为T,使得46%的碱基对从规范碱基对转化为G:U碱基对,在大多数转基因植物中仅诱导弱的EIN2沉默或没有EIN2沉默。可能的解释包括i)存在太多的G:U碱基对,导致Dicer加工效率低,和ii)包括太多的G:U碱基对的与靶mRNA结合的sRNA不诱导有效的mRNA切割,或多种因素的组合。

[0624] G:U碱基配对构建体增加的靶基因沉默的均匀性与降低的启动子甲基化有关

[0625] 使用McrBC-消化PCR和亚硫酸氢盐测序的DNA甲基化分析显示所有hpEIN2[wt]植物品系在启动子区显示DNA甲基化,并且甲基化程度与EIN2沉默水平负相关。甚至如通过McrBC-消化PCR所判断的,相对于所有甲基化的胞嘧啶,在35S启动子中显示约40%的DNA甲基化水平。广泛的启动子甲基化被认为是由于在EIN2重复序列处sRNA指导的DNA甲基化,其扩散至邻近的启动子区。与hpRNA[wt]植物品系相比,许多hpEIN2[G:U]品系显示很少至没有启动子甲基化,并且所分析的大部分植物显示较少甲基化的胞嘧啶。如对于hpGUS品系讨论的,多个因素可导致甲基化降低:i)反向重复DNA结构通过在有义序列中将C碱基改变为T碱基而破坏,和ii)有义EIN2序列缺少胞嘧啶,因此不能通过sRNA指导的DNA甲基化而甲基化,和iii)由于具有G:U碱基对的dsRNA区结构的改变导致24-mer RNA产生水平降低,导致

一些Dicer识别的改变,因此Dicer3和/或Dicer4活性和Dicer2活性相对较高。因此,hpEIN2[G:U]转基因可以表现得像正常的,非RNAi转基因(如过表达转基因),并且在一些品系中观察到的启动子甲基化是由于T-DNA插入模式而不是hpRNA转基因的固有反向重复DNA结构。

[0626] 实施例14:用于降低另一内源基因表达的修饰的发夹

[0627] 设计并合成了靶向于其他内源基因的用于产生针对发夹RNA或ledRNA的经修饰的沉默RNA的遗传构建体。这些包括以下内容。

[0628] 拟南芥(A.thaliana)和欧洲油菜(Brassica napus)中的FANCM基因编码范可尼贫血互补组M(FANCM)蛋白,其为DEAD/DEAH盒RNA解旋酶蛋白,登录号NM_001333162和XM_018659358。对应于拟南芥(A.thaliana)的FANCM基因的cDNA的蛋白编码区的核苷酸序列提供于SEQ ID NO:31中,并且对于欧洲油菜(Brassica napus),提供于SEQ ID NO:32中。

[0629] 设计并制备遗传构建体,以表达具有或不具有C到T取代的发夹RNA和靶向拟南芥(A.thaliana)和欧洲油菜(Brassica napus)中的FANCM基因。拟南芥(A.thaliana)中的靶区选择SEQ ID NO:31的核苷酸675-1174(500个核苷酸)。欧洲油菜(B.napus)中的靶区选择SEQ ID NO:32的核苷酸896-1395(500bp)。设计并组装了使用野生型有义序列或修饰的(G:U)有义序列编码发夹RNA的构建体。hpFANCM-At[wt]、hpFANCM-At[G:U]、hpFANCM-Bn[wt]和hpFANCM-Bn[G:U]构建体的核苷酸序列提供于SEQ ID NO:33-36中。为了制备G:U构建体,将有义序列中的所有胞嘧啶碱基替换为胸腺嘧啶碱基-拟南芥(A.thaliana)构建体中的102/500(提供20.4%G:U碱基对),欧洲油菜(B.napus)中109/500(21.8%G:U碱基对)。欧洲油菜(B.napus)G:U修饰的发夹的双链区中邻接规范碱基配对的最长延伸为17个碱基对,第二长为16个邻接碱基对。

[0630] 欧洲油菜(B.napus)DDM1基因编码DNA中甲基化胞嘧啶碱基的甲基转移酶(Zhang等,2018)。对应于欧洲油菜(Brassica napus)的DDM1基因的cDNA的蛋白编码区的核苷酸序列提供于SEQ ID NO:37中。

[0631] 设计并制备遗传构建体,以表达具有或不具有C到T取代的发夹RNA和靶向欧洲油菜(Brassica napus)中的DDM1基因。欧洲油菜(B.napus)基因的两个非邻接靶区选择:SEQ ID NO:37的核苷酸504-815和1885-2074,并直接连接以产生嵌合有义序列。因此有义序列的总长度为502个核苷酸。设计并组装了使用野生型有义序列或修饰的(G:U)有义序列编码发夹RNA的构建体。hpDDM1-Bn[wt]和hpDDM1-Bn[G:U]构建体的核苷酸序列提供于SEQ ID NO:38-39中。为了制备G:U构建体,在欧洲油菜(B.napus)构建体中用胸腺嘧啶替换有义序列中的胞嘧啶-106/502(21.1%G:U碱基对)。G:U修饰的发夹的双链区中邻接规范碱基配对的最长延伸为20个碱基对,第二长为15个邻接碱基对。

[0632] 对于靶向内源基因的另一构建体,设计遗传构建体,以在350个核苷酸的有义序列中的104个C中,在有义序列中表达具有95C至T取代的发夹RNA,在发夹RNA的双链区中提供95/350=27.1%G:U碱基对。即,不是有义序列中的所有C'都被T'取代。特别是,当在有义序列中出现3、4或5个邻接C'时,3个C'中仅有1个或2个,或者4个C'中仅有2个或3个,或者5个邻接C'中仅有2个、3个或4个被T'替代。这提供了双链RNA区中G:U碱基对的更均匀分布。双链区邻接规范碱基配对的最长延伸为15个碱基对,第二长为13个邻接碱基对。

[0633] 设计另一个构建体,其中用C到T或A到G的取代修饰4、5、6或7个核苷酸编组的每个编组中的一个或两个碱基对。其中野生型有义序列具有由T'或G'组成的8个或更多个核苷

酸的一段序列,一个或多个核苷酸在有义链中被取代以在该编组内产生错配核苷酸,或者在反义链中进行C至T或A至G的取代,从而避免从构建体转录的所得发夹RNA的双链区中8个或更多个邻接规范碱基对的双链延伸。

[0634] 实施例15:用于降低动物细胞中基因表达的修饰发夹

[0635] 为了在动物细胞中测试G:U碱基配对形式、ledRNA形式或这两种修饰的组合的修饰的沉默RNA,在以下实验中使用编码增强型绿色荧光蛋白(EGFP)的基因作为模型靶基因。EGFP编码区的核苷酸序列如SEQ ID NO:40所示。选择460个核苷酸的靶区,对应于SEQ ID NO:40的核苷酸131-591。

[0636] 设计并制备表达发夹RNA的称为hpEGFP[wt]的遗传构建体,所述发夹RNA相对于用于表达的启动子按5'至3'顺序包含460个核苷酸的反义EGFP序列,所述反义EGFP序列与EGFP编码区的相应区(核苷酸131-590)完全互补;部分衍生自GUS编码区(相当于GUS ORF的核苷酸802-1042)的312个核苷酸的环序列;和与EGFP编码区的核苷酸131-590在序列上相同的460个核苷酸的有义EGFP序列。编码发夹RNA hpEGFP[wt]的DNA序列(SEQ ID NO:41)在5'端包括NheI限制酶位点,在3'端包括SalI位点,以便克隆到载体pCI中(Promega公司)。该载体适用于哺乳动物细胞转染实验,并提供强CMV启动子/增强子的表达。该构建体还具有插入在NheI位点和反义序列起点之间的T7启动子序列,以提供体外转录,从而使用T7 RNA聚合酶产生发夹RNA。将发夹编码盒插入表达载体pCI中的NheI至SalI位点,由此将RNA编码区可操作地连接至CMV启动子和SV40-晚期聚腺苷酸化/转录终止区。

[0637] 设计并制备相应的发夹构建体,其在有义序列中具有157个C至T取代且在反义序列中没有取代,命名为hpEGFP[G:U](SEQ ID NO:42)。EGFP编码区的靶区为核苷酸131-590。在发夹RNA的茎中,C至T取代以及因此产生的G:U碱基对的百分比为157/460=34.1%。有义和反义序列的长度相同,为460个核苷酸。在基因沉默领域中,由于激活细胞应答(包括干扰素激活)的潜能,通常避免长双链RNA。

[0638] 设计并制备称为ledEGFP[wt]以表达ledRNA的ledRNA构建体,所述ledRNA相对于用于表达的启动子按5'至3'顺序包含与EGFP编码序列的核苷酸131-358完全互补的228个核苷酸的反义EGFP序列;150个核苷酸的环序列;460个核苷酸的有义EGFP序列,其序列与EGFP编码区的核苷酸131-590相同(SEQ ID NO:40);144个核苷酸的环序列;和232个核苷酸的反义序列,其与EGFP编码序列的核苷酸359-590完全互补,侧翼接NheI和SalI限制性位点(SEQ ID NO:43)。因此编码的ledRNA是图1A所示的类型。当通过一个有义和两个反义序列之间的碱基配对自退火时,ledRNA结构具有对应于EGFP靶区的460个碱基对的双链区,其中两个反义序列不直接彼此共价连接,而是在对应于核苷酸358和359的末端之间具有"间隙"或"裂口"。在CMV启动子和SV40晚期聚腺苷酸化/转录终止区的序列中,LedRNA结构被包埋在较大的RNA转录物中,该转录物包括5'上游区和3'下游区。

[0639] 设计并制备相应的ledRNA构建体,其在有义序列中具有162个C至T取代且在反义序列中没有取代,命名为ledEGFP[G:U](SEQ ID NO:44)。在每种情况下,EGFP编码区中的靶区相对于以ATG起始密码子起始的蛋白质编码区为核苷酸131-590(SEQ ID NO:40)。在ledRNA的茎中,C至T取代以及因此产生的G:U碱基对的百分比为162/460=35.2%。

[0640] 通过将载体转染到细胞中,测试编码hpEGFP[wt]、hpEGFP[G:U]、ledEGFP[wt]和ledEGFP[G:U]沉默RNA的质粒在CHO、HeLa和VERO细胞中的基因沉默活性。通过将测试质粒

与GFP表达质粒共转染进行测定。所有测定一式三份进行。将CHO细胞(中国仓鼠卵巢细胞)和VERO细胞(非洲绿猴肾细胞)以每孔1×10^5个细胞的密度接种到24孔板中。CHO细胞在MEMα修饰(美国西格玛(Sigma))中生长,HeLa和VERO细胞在DMEM(美国英杰(Invitrogen))中生长。两种基础培养基均补充有10%胎牛血清,2mM谷氨酰胺,10mMHepes,1.5g/L碳酸氢钠,0.01%青霉素和0.01%链霉素。细胞在37℃,5%CO_2下生长。然后使用Lipofectamine 2000(脂质体)用质粒DNA或作为EGFP沉默对照的siRNA以1μg/孔转染细胞。简单地说,将测试siRNA或质粒与GFP报告质粒(pGFP N1)组合,然后与1μl的Lipofectamine 2000混合,两者均在50μl OPTI-MEM(美国英杰)中稀释,并在室温下孵育20分钟。然后将复合物加入细胞并孵育4小时。替换细胞培养基并将细胞培养72小时。然后对细胞进行流式细胞计数以测量GFP沉默。简单地说,将待分析的细胞用胰蛋白酶消化,在PBSA中洗涤,再悬浮于PBSA中的200μL 0.01%叠氮化钠和2%FCS中,并使用FACScalibur(美国BD公司(Becton Dickinson))流式细胞仪进行分析。使用CELLQuest软件(美国BD公司)进行数据分析并以平均荧光强度(MFI)报告为具有报告子和非相关(阴性对照)shRNA的对照细胞的百分比。

[0641] 称为si22的anti-GFP siRNA获自凯杰公司(Qiagen)(美国)。si22的anti-GFP siRNA序列为有义5'-gcaagcugacccugaaguucau-3'(SEQ ID NO:86)和反义5'-gaacuucagggucagcuugccg-3'(SEQ ID NO:87)。使用小鼠U6序列作为模板,通过一步PCR反应产生命名为pshGFP的阳性对照遗传构建体。正向引物为5'-TTTTAGTATATGTGCTGCCG-3'(SEQ ID NO:88),反向引物为5'-ctcgagttccaaaaaagctgaccctgaagttcatctctcttgaagatgaacttcagggtcagccaaacaaggcttttctccaa-3'(SEQ ID NO:89)。将包含全长表达盒的扩增产物连接到pGEM-T Easy中。还通过相同的PCR方法构建了非相关shRNA对照质粒。对于该构建,正向引物是5'-TTTTAGTATATGTGCTGCCG-3'(SEQ ID NO:90),反向引物是5'-ctcgagttccaaaaaaataagtcgcagcagtacaatctcttgaattgtactgctgcgacttatgaataccgcttcctcctgag-3'(SEQ ID NO:91)。

[0642] 来自一个实验的结果数据示于图34中。与不相关的shRNA对照相比,对于si22和pshGFP阳性对照,在VERO和CHO细胞中观察到EGFP活性的明显降低(RNA沉默)。这些阳性对照是充分验证的小dsRNA分子(si22)或编码shRNA(pshGFP),已知其在哺乳动物细胞中具有很强的沉默活性。对照RNA分子分别具有20个邻接碱基对和21个邻接碱基对的双链区,仅使用规范碱基对并且在双链区中没有任何错配核苷酸,并且长度在20-30个碱基对的范围内,通常用于哺乳动物细胞。相反,hpRNA和ledRNA构建体表达具有长dsRNA区的分子。在两种细胞类型中观察到所有四种构建体在显著程度上特异性沉默EGFP表达(图34)。包含G:U取代显著提高了CHO细胞中两种构建体的沉默。在VERO细胞中,仅观察到ledEGFP[G:U]构建体相对于ledEGFP[wt]的沉默显著改善。

[0643] 在使用HeLa(人)细胞并在转染后48小时测定EGFP活性的第二个实验中,获得了类似的结果(图35)。

[0644] 值得注意的是,使用hpRNA和ledRNA效应分子在哺乳动物细胞中观察到基因沉默,因为它们具有比常规20至30bp大小范围更长的双链区。还清楚的是,替代核苷酸以产生G:U碱基对的修饰显著增强了这些较长dsRNA分子的基因沉默作用。这种作用可能是由于这些结构更接近地类似于在真核细胞中观察到的内源性priRNA(miRNA的前体),并且因此改善了用于装载到RNA诱导的沉默复合物(RISC)效应蛋白中的较长dsRNA的加工。

[0645] 实施例16:植物中靶向DDM1和FANCM基因的RNA构建体

[0646] 本发明人考虑了提高可产生新的遗传图谱和多样性(遗传增益)的速率的方法,并探索植物中期望的性能性状。一种被认为是找到增加植物有性生殖期间发生的重组率的方法。植物育种者依靠重组事件产生不同的遗传(等位基因)组合,他们可以搜寻与性能获得相关的所需遗传图谱。然而,在每个育种步骤中的重组事件的数目相对于可以探索的可能的遗传图谱的数目是极低的。此外,控制这些事件在基因组中发生的位置的元素尚未被很好地理解。因此,本发明人考虑是否通过转基因方法外源或内源递送的ledRNA可用于改变植物中的重组率,以允许快速增加遗传多样性并使育种群体内更快的遗传增益成为可能。

[0647] 植物的表观基因组受到组织、包装和稳定基因组的DNA和相关蛋白质的一系列不同化学修饰的影响。这些修饰还调节重组发生的位置,紧密的基因组包装是重组的强抑制剂(Yelina等,2012;Melamed-Bessudo等2012)。脱氧DNA甲基化1(DDM1)是一种调节DNA甲基化和基因组包装的酶。该基因的突变可以改变重组事件的位置(Yelina等,2012;Melamed-Bessudo等2012)。

[0648] 减数分裂期间的重组事件受到严格调控,每个染色体上只发生1-2个事件,以确保中期1的正确染色体分离。通过酶SPOII通过DNA的双链断裂(DSB)引发重组事件(Wijnker等,2008)。这导致沿着染色体的数百个DSB。尽管这些DSB中的一些导致交换,但大部分在重组事件发生之前被DNA修复酶修复。此外,存在许多抑制DSB发展成交叉的负调节剂。在本发明人考虑的初始方法中,将编码ledRNA分子或常规发夹RNA分子的遗传构建体作为比较引入拟南芥(A.thaliana)植物,其靶向编码可能潜在地影响重组率的蛋白质因子的基因,例如FANCONI贫血补充M组(FANCONI ANEMIA COMPLEMENTATION GROUP M)(FANCM)。

[0649] 拟南芥(A.thaliana)的DDM1基因的核苷酸序列由登录号AF143940提供(Jeddeloh等,1999)。已经显示DDM1基因表达的减少降低了拟南芥(A.thaliana)中的DNA甲基化并且增加了交叉事件的数量和位置。(Melamed-Bessudo和Levy,2012)。

[0650] 欧洲油菜(Brassica napus)是异源四倍体物种,并且在A和C亚基因组的每一个上,在染色体A7、A9、C7和C9上具有两个DDM1基因,因此总共具有四个DDM1基因。这些基因命名为BnaA07g37430D-1、BnaC07g16550D-1、BnaA09g52610D-1和BnaC09g07810D-1。欧洲油菜(Brassica napus)的DDM1基因BnaA07g37430D-1的核苷酸序列由登录号XR_001278527提供(SEQ ID NO:93)。设计并制备靶向4个基因的500个核苷酸区的发夹RNA构建体,所述基因对应于SEQ ID NO:93的核苷酸650-959和2029-2218。基于基因之间的序列保守性,用于设计hpRNA和ledRNA构建体的核苷酸区靶向存在于欧洲油菜(B.napus)中的所有四种DDM1基因BnaA07g37430D-1、BnaC07g16550D-1、BnaA09g52610D-1和BnaC09g07810D-1。hpRNA构建体的元件顺序为启动子-有义序列-环序列,包含Hellsgate载体的内含子-反义序列-转录终止子/聚腺苷酸化区。编码hpRNA的嵌合DNA的核苷酸序列提供为SEQ ID NO:94。

[0651] 制备编码靶向相同的500个核苷酸区并具有相同结构的发夹RNA的第二发夹RNA构建体,除了有义序列的97个胞嘧啶核苷酸(C)被胸腺嘧啶核苷酸(T5-T7)替代。当所述嵌合DNA被转录并且所述G:U取代的hpRNA被自身退火时,这提供了97/500=19.4%的dsRNA区中的核苷酸被碱基配对为G:U碱基对。编码G:U修饰的hpRNA的嵌合DNA的核苷酸序列提供为SEQ ID NO:95。此外,制备了编码靶向欧洲油菜(B.napus)的DDM1基因的相同区的ledRNA的嵌合DNA。该编码ledRNA的嵌合DNA的核苷酸序列提供为SEQ ID NO:96。

[0652] 为了通过体外转录产生RNA,用在编码区之后立即切割的限制酶HincII切割DNA制备物,用RNA聚合酶T7体外转录,纯化RNA,然后在水性缓冲液中浓缩。使用ledRNA靶向欧洲油菜(B.napus)子叶中的内源性DDM1转录物。小心地切下来自在组织培养基上无菌生长的5日龄幼苗的子叶,并将其置于含有2ml Ms液体培养基(包含2%(w/v)蔗糖)以及113μg ledRNA或100ul水性缓冲液作为对照的皮氏培养皿中。用于处理的MS液体介质含有Silwett-77,一种表面活性剂(0.5μl在60ml中)。在摇床上轻柔振荡温育皮氏培养皿,使得子叶浸泡在含有ledRNA的溶液中。应用ledRNA后5小时和7小时收获样品。在平行实验中,用10μg的ledRNA或缓冲液涂覆子叶的上表面并在湿组织纸上温育。在应用ledRNA后7小时收集样品。

[0653] 此外,为了靶向欧洲油菜(B.napus)的生殖组织中的DDM1内源性转录物,在存在或不存在等分的根癌农杆菌(Agrobacterium tumafecians)菌株AGL 1细胞悬浮液即活AGL 1细胞的情况下,将甘蓝,油菜花芽暴露于ledRNA。含有或不含AGL1细胞的水性缓冲液用作各自的对照。AGL1在含有25mg/ml利福平(rifampicin)的10ml LB液体培养基中于28℃生长2天。以3000rpm离心5分钟收获细胞。洗涤细胞微丸,并将细胞重悬于2ml液体MS培养基中。将花芽在含有2ml的MS液体培养基(包括在50ml的MS液体培养基中的0.5μl的Silwett-77)以及62μg的ledRNA或62μg+50μl的AGL1培养物的皮氏培养皿中温育。作为对照,使用50μl缓冲液或50μl缓冲液+50μl AGL1培养物。将样品在摇床上轻柔振荡温育7小时。每个处理使用三个生物复制品。

[0654] 将处理的和对照的子叶和花芽在无菌蒸馏水中洗涤两次,使用组织纸除去表面水并用液氮速冻。从处理和对照组织中分离RNA,用DNase处理除去基因组DNA并定量。使用等量的来自经ledRNA处理的样品及其相应对照的总RNA合成第一链cDNA。使用实时荧光定量PCR(qRT-PCR法)检测DDM1的表达。

[0655] 在用ledRNA浸泡的处理后的子叶中,与对照相比,DDM1转录物丰度在5小时减少了约83-86%,在7小时进一步减少了91%。类似地,在用ledRNA包被的子叶中观察到与对照相比DDM1 mRNA水平降低约78-85%。在没有农杆菌(Agrobacterium)细胞的情况下,与对照相比,用ledRNA处理的花芽中没有检测到DDM1 mRNA丰度的差异。然而,与其相应的对照相比,在农杆菌(Agrobacterium)存在下用ledRNA处理的花芽中观察到DDM1转录物水平降低约60-75%。当没有农杆菌(Agrobacterium)的对照与具有农杆菌(Agrobacterium)的对照比较时,未检测到DDM1转录物水平的显著差异,表明农杆菌(Agrobacterium)细胞本身不引起DDM1转录物的减少。总之,这些结果表明,ledRNA能够降低子叶和花芽中的内源DDM1转录物水平,而活的农杆菌(Agrobacterium)细胞似乎促进ledRNA进入花芽。也可以通过物理方法如刺穿花芽外层、离心或真空渗透或这些方法的组合实现ledRNA的这种可接近性。

[0656] 某些拟南芥(Arabidopsis thaliana)突变体,例如zip4突变体,缺乏减数分裂杂交,导致染色体同系物的错误分离并且因此降低生育力并且导致较短的长角果(果实),这些短角果可以从视觉上与野生型的短角果区分开。降低FANCM基因表达可以逆转zip4突变体的表型。

[0657] 拟南芥(A.thaliana)的FANCM基因的核苷酸序列由登录号NM_001333162提供(SEQ ID NO:97)。设计并制备靶向基因的500个核苷酸区的发夹RNA构建体,所述基因对应于SEQ ID NO:97的核苷酸和853-1352。构建体的元件顺序为启动子-有义序列-环序列,包含

Hellsgate载体的内含子-反义序列-转录终止子/聚腺苷酸化区。编码hpRNA的嵌合DNA的核苷酸序列提供为SEQ ID NO:98。制备编码靶向相同的500个核苷酸区的类似发夹RNA的第二发夹RNA构建体,除了有义序列的102个胞嘧啶核苷酸(C)被胸腺嘧啶核苷酸(T5-T7)替代。当所述嵌合DNA被转录并且因此所述G:U取代的hpRNA被自身退火时,这提供了102/500＝20.4％的dsRNA区中的核苷酸被碱基配对为G:U碱基对。编码G:U修饰的hpRNA的嵌合DNA的核苷酸序列提供为SEQ ID NO:99。此外,制备了编码靶向拟南芥(A.thaliana)的FANCM基因的相同区的ledRNA的嵌合DNA。该编码ledRNA的嵌合DNA的核苷酸序列提供为SEQ ID NO:100。

[0658] 欧洲油菜(B.napus)在其A和C亚基因组中的每一个上具有FANCM基因,命名为BnaA05g18180D-1和BnaC05g27760D-1。欧洲油菜(B.napus)的FANCM基因之一的核苷酸序列由登录号XM_022719486.1提供SEQ ID NO:101)。设计编码发夹RNA的嵌合DNA并使其靶向基因的503个核苷酸区,对应于SEQ ID NO:101的核苷酸2847-3349。构建体的元件顺序为启动子-有义序列-环序列,包含Hellsgate载体的内含子-反义序列-转录终止子/聚腺苷酸化区。编码hpRNA的嵌合DNA的核苷酸序列提供为SEQ ID NO:102。制备编码靶向相同的503个核苷酸区的类似发夹RNA的第二发夹RNA构建体,除了有义序列的107个胞嘧啶核苷酸(C)被胸腺嘧啶核苷酸(T5-T7)替代。当所述嵌合DNA被转录并且所述G:U取代的hpRNA被自身退火时,这提供了107/500＝21.4％的dsRNA区中的核苷酸被碱基配对为G:U碱基对。编码G:U修饰的hpRNA的嵌合DNA的核苷酸序列提供为SEQ ID NO:103。此外,制备了编码靶向欧洲油菜(B.napus)的FANCM基因的相同区的ledRNA的嵌合DNA。该编码ledRNA的嵌合DNA的核苷酸序列提供为SEQ ID NO:104。

[0659] 为了通过体外转录产生RNA,用在编码区之后立即切割的限制酶HincII切割DNA制备物,用RNA聚合酶T7体外转录,纯化RNA,然后在水性缓冲液中浓缩。将ledRNA与根癌农杆菌(Agrobacterium tumefacians)AGL1一起用于靶向拟南芥(A.thaliana)的zip4突变体的减数分裂前芽中的FANCM转录物。由于减弱的杂交形成,与野生型长角果相比,zip4突变体的长角果更短,易于肉眼观察,因此导致生育力降低。抑制zip4突变体中的FANCM已显示恢复生育力和恢复长角果长度。

[0660] 在每种情况下在表面活性剂的存在下,在该情况下是Silwett-77,使含有减数分裂前芽的拟南芥(A.thaliana)zip4花序与靶向FANCM的ledRNA以及AGL1或以AGL1作为对照的缓冲液接触。一旦种子定植完成,将从减数分裂前芽发育的长角果切除以确定种子数。在来自ledRNA处理的样品的15个长角果中,两个长角果显示10个种子,一个长角果具有9个种子,而对照长角果中的种子数目为3-6个。这些结果表明观察到的种子数的增加是由于ledRNA抑制FANCM转录物水平,从而导致减数分裂交换数增加和生育力增加。

[0661] 实施例17:用于真菌疾病抗性的RNA构建体

[0662] 针对大麦和小麦的Mlo基因的LedRNA

[0663] 禾谷类植物的真菌病、白粉病、是由大麦中的子囊大麦白粉病菌(Blumeria graminis f.sp.hordei)和小麦中相关的小麦白粉病菌(Blumeria graminis f.sp.tritici)引起的。禾本科布氏白粉菌(B.graminis)是白粉菌目(Erysiphales)专性生物营养真菌病原体(Glawe,2008),其需要植物宿主用于繁殖,涉及真菌和宿主细胞之间的紧密相互作用以使真菌从植物获得营养。在真菌子囊孢子或分生孢子接触表面后,真菌最

初侵染叶、叶鞘或耳的表皮层。叶片在侵染后保持绿色和活性一段时间,然后变成粉末状,菌丝团生长,叶片逐渐变色并死亡。随着疾病的发展,真菌菌丝体可能会出现微小的黑点,这些黑点是真菌有性子实体。白粉病在世界范围内分布,并且在寒冷、潮湿的气候中最有害。该病主要通过减少穗数、减少籽粒大小和重量来影响粮食产量。目前,病害防治是通过向作物喷洒杀真菌剂或通过生长抗性栽培种进行的,所述杀真菌剂需要在条件寒冷潮湿时频繁施用,成本高。此外,澳大利亚小麦白粉病已经出现了抗药性。

[0664] 大麦和小麦的Mlo基因编码Mlo多肽,该多肽通过不明机制赋予对禾本科布氏白粉菌(B.graminis)的易感性。存在多种由植物独有的Mlo基因家族编码的密切相关的MLO蛋白。每个基因编码定位于质膜中的具有未知生化活性的7个跨膜结构域蛋白。显着地,只有该家族内的特异性Mlo基因能够作为白粉病易感基因,并且这些基因编码在Mlo蛋白的细胞质C-末端结构域内具有保守基序的多肽。Mlo多肽作为白粉病易感因子的作用机制尚不清楚。还没有报道天然小麦mlo突变体的出现,这可能是由于小麦的多倍体性质。然而,人工产生的mlo突变体显示出一些对该疾病的抗性,但通常显示出显著降低的谷物产量或过早的叶衰老(Wang等;Acevedo-Garcia等,2017)。

[0665] 六倍体小麦具有3个Mlo基因同源序列,分别命名为位于染色体5AL、4BL,和4DL上的TaMlo-A1、TaMlo-B1和TaMlo-D1(Elliott等,2002)。对应于这些基因的cDNA的核苷酸序列可以以下登录号获得:TaMlo-A1、AF361933和AX063298;TaMlo-B1、AF361932、AX063294和AF384145;以及TaMlo-D1、AX063296。A、B、D基因组上的基因的核苷酸序列与编码的多肽的氨基酸序列分别有大约95-97%和98%相同。三个基因均在植物叶片中表达,表达量随植物生长成熟而增加。因此,本发明人设计并制备了能够降低所有三种基因的表达,利用基因之间的序列同一性程度并靶向具有高度序列保守性的基因区的ledRNA构建体。

[0666] 制备编码靶向所有三个TaMlo-A1、TaMlo-B1和TaMlo-D1基因的ledRNA构建体的嵌合DNA。使用上述用于ledRNA的设计原理制备遗传构建体,其中分裂序列为反义序列且邻接序列为有义序列(图1A)。

[0667] 选择TaMlo靶基因的500bp核苷酸序列,对应于与SEQ ID NO:136的1403-1569融合的核苷酸916-1248。各ledRNA的dsRNA区长度为500bp;dsRNA区中的有义序列是不间断的邻接的序列,例如对应于与SEQ ID NO:136的1403-1569融合的核苷酸916-1248。编码ledRNA的核苷酸序列在本文中提供为SEQ ID NO:137。

[0668] 用T7RNA聚合酶体外转录制备LedRNA,纯化后悬浮于缓冲液中。使用漆刷将10μg ledRNA/叶施用于处于Zadoks 23生长阶段的小麦植物的叶区。作为对照,仅使用缓冲液模拟处理一些叶。收获处理过的和对照的叶片样品并提取RNA。对提取的RNA的QPCR分析显示,作为三种TaMlo mRNA的组合,TaMlomRNA水平降低了95.7%。还处理并测定处于Z73生长阶段的植物。通过QPCR,它们显示相对于对照叶片样品,TaMlo基因表达降低了91%。在处理过的叶片区中观察到的TaMlo基因表达的降低对处理过的区是特异性的-在叶片的远端未处理部分中TaMlo mRNA水平没有降低。

[0669] 在大麦mlo突变体中,观察到多种疾病防御相关基因的表达增加。因此,通过QPCR测定ledRNA处理的小麦叶片中编码PR4、PR10、β-1,3-葡聚糖酶、几丁质酶、生殖细胞和ADP-核糖基化因子的防御相关基因的水平。这些基因中没有一个在处理的叶区中相对于对照叶区在表达水平上显著改变。

[0670] 为了测试ledRNA通过降低Mlo基因表达增加抗病性的能力,将白粉病菌的孢子施用于叶片的处理和未处理区。将叶片从小麦植物分离,如之前用ledRNA处理并保持在培养基上(每升水50mg苯并咪唑和1g琼脂)以防止叶片在光下衰老。24h后接种白粉菌孢子,病程5-24d。相对于未接受ledRNA的对照叶片,处理过的叶片显示几乎没有到没有真菌菌丝体生长,并且没有叶片变色,所述ledRNA显示由变色区包围的广泛菌丝体生长。

[0671] 在进一步的实验中,施用较低水平的ledRNA来鉴定有效的最小水平的ledRNA。此外,在ledRNA处理后2、4、7或更多天接种叶片以观察保护作用保持多长时间。还将用ledRNA制剂喷洒整株植物,并在接种真菌病害剂后测试其抗病性。

[0672] 靶向葡萄(Vitisvinifera)的VvMLO基因的LedRNA

[0673] 葡萄(Vitis vinifera)和华东葡萄(Vitis pseudoreticulata)的MLO基因编码MLO多肽,这赋予对由子囊菌白粉菌(Erysiphe necator)引起的真菌病白粉病的易感性。白粉菌(E.necator)是专性生物营养真菌病原体,其需要植物宿主用于繁殖,涉及真菌和宿主细胞之间的紧密相互作用以使真菌从植物获得营养。存在多个由基因家族编码的紧密相关的MLO蛋白,所有这些蛋白都是植物独有的,并且编码定位于质膜中的生物化学活性未知的7个跨膜结构域蛋白。显着地,只有该家族内的特异性MLO基因能够作为白粉病易感基因,并且这些基因编码在Mlo蛋白的细胞质C-末端结构域内具有保守基序的多肽。MLO多肽作为白粉病易感因子的作用机制尚不清楚。

[0674] 如下设计并制备靶向葡萄属(Vitis)物种的三种不同但相关的MLO基因,即VvMLO6、VvMLO11和VvMLO15的ledRNA构建体。对于第一个,例如,选择VvMLO6靶基因的600个核苷酸的序列,其对应于SEQ ID NO:138的核苷酸556-1155。制备编码靶向VvMLO6、VvMLO7和VvMLO11基因的三个ledRNA构建体的嵌合DNA。使用上述用于ledRNA的设计原理制备遗传构建体,其中分裂序列为反义序列且邻接序列为有义序列(图1A)。各ledRNA的dsRNA区长为600bp;dsRNA区中的有义序列是不间断的邻接的序列,例如对应于SEQ ID NO:138的核苷酸556-1155。编码一种ledRNA的核苷酸序列在本文中提供为SEQ ID NO:139。

[0675] ledRNA是通过体外转录制备的,并且单独或作为所有三种的混合物施用于葡萄(Vitis)叶。随后,将白粉病菌的孢子施用于叶片的处理和未处理区。使用QPCR观察到靶mRNA的下调,并随时间跟踪疾病进展。

[0676] 实施例18:靶向植物中其他基因的RNA构建体

[0677] 靶向拟南芥(A.thaliana)和本氏烟(N.benthamiana)Tor基因的LedRNA

[0678] 雷帕霉素靶标(TOR)基因编码丝氨酸-苏氨酸蛋白激酶多肽,其控制真核细胞中的许多细胞功能,例如响应于各种激素、应激和营养可用性。已知它是调节翻译机制以优化用于生长的细胞资源的主要调节剂(Abraham,2002)。至少在动物和酵母中,TOR多肽被抗真菌药雷帕霉素灭活,导致其被指定为雷帕霉素靶标。在植物中,TOR对于发育中的种子的胚胎发育至关重要,如TOR中的纯合突变体的致死性所示(Mahfouz等人,2006),并且参与生长线索与细胞代谢的耦合。TOR基因表达的下调被认为导致脂肪酸合成增加,导致植物组织中脂质含量增加。

[0679] 使用分离序列为有义序列且邻接序列为反义序列的ledRNA的设计原理,设计并制备靶向本氏烟(Nicotiana benthamiana)TOR基因的ledRNA构建体,cDNA蛋白编码区的核苷酸序列提供为SEQ ID NO:105(图1B)。靶区长度为603个核苷酸,对应于SEQ ID NO:105的核

苷酸2595-3197。ledRNA的dsRNA区长度为603bp;dsRNA区中的反义序列是不间断的邻接的序列,对应于SEQ ID NO:105的核苷酸2595-3197的互补体。编码ledRNA的核苷酸序列在本文中提供为SEQ ID NO:106。用限制酶MlyI切割编码ledRNA构建体的遗传构建体的DNA制备物,所述限制酶MlyI在编码区之后立即切割DNA,用RNA聚合酶SP6体外转录并纯化RNA,然后在水性缓冲液中浓缩。将ledRNA样品施涂到本氏烟(N.benthamiana)的上表面。在2天和4天后,收获处理的叶片样品,干燥,并通过定量气相色谱法(GC)测定总脂肪酸含量。用TOR ledRNA处理的叶片样品显示总脂肪酸(TFA)含量从对照(未处理)样品中观察到的2.5-3.0%(TFA重量/干重)增加到ledRNA处理的样品的3.5-4.0%。这表示TFA含量相对于对照增加17%至60%,表明在ledRNA处理的组织中TOR基因表达降低。

[0680] 靶向大麦(H.vulgare)ALS基因的LedRNA

[0681] 乙酰乳酸合酶(ALS)基因编码在植物和微生物中发现的酶(EC 2.2.1.6),其催化支链氨基酸亮氨酸、缬氨酸和异亮氨酸合成中的第一步。ALS酶催化丙酮酸转化为乙酰乳酸,然后乙酰乳酸被其他酶进一步转化为支链氨基酸。ALS抑制剂用作除草剂,例如磺酰脲、咪唑啉酮、三唑并嘧啶、嘧啶基氧基苯甲酸酯和磺酰氨基羰基三唑啉酮类除草剂。

[0682] 为了测试ledRNA是否能通过外源递送RNA至植物而降低ALS基因表达,设计并制备了编码ledRNA的遗传构建体,其靶向大麦(Hordeum vulgare)中的ALS基因。大麦(H.vulgare)ALS基因序列在本文中提供为SEQ ID NO:107(登录号LT601589)。使用用于ledRNA的设计原理制备遗传构建体,其中分裂序列是有义序列,邻接序列是反义序列(图1B)。靶区长度为606个核苷酸,对应于SEQ ID NO:107的核苷酸1333-1938。ledRNA的dsRNA区长度为606bp;dsRNA区中的反义序列是不间断的邻接的序列,对应于SEQ ID NO:107的核苷酸1333-1938的互补体。编码ledRNA的核苷酸序列在本文中提供为SEQ ID NO:108。编码区在SP6 RNA聚合酶启动子的调控下进行体外转录。

[0683] 用限制性内切酶MlyI消化编码ledRNA的基因构建体,在ledRNA编码区的下游切割,并根据转录试剂盒的说明用RNA聚合酶SP6进行体外转录。将RNA应用于大麦植物叶片的上表面。从处理过的叶片样品中提取RNA(24小时后)。对RNA样品进行定量逆转录-PCR(QPCR)测定。结果显示,在ledRNA处理的组织中ALS mRNA水平降低。(提取处理和未处理植物的总RNA,DNase处理,定量并使用引物CTTGCCAATCTCAGCTGGATC逆转录2μg。使用正向引物TAAGGCTGACCTGTTGCTTGC和反向引物CTTGCCAATCTCAGCTGGATC作为定量PCR的模板。将ALS mRNA表达相对于大麦(Horendeum chilense)分离物H1番茄红素-环化酶基因进行标准化。在LED处理的植物中ALS表达降低了82%。

[0684] 靶向小麦和大麦的NCED1和NCED2基因的ledRNA

[0685] 在植物中,由类胡萝卜素前体合成植物激素脱落酸(ABA),其合成途径中的第一个关键步骤是由9-顺式环氧类胡萝卜素双加氧酶(NCED)催化,其酶将9-顺式叶黄素裂解为黄体素(Schwartz等,1997)。已知激素ABA促进种子中的休眠(Millar等,2006),并且参与其他过程,如胁迫响应。认为增加NCED基因的表达增加ABA浓度,从而促进休眠。小麦、大麦等作物中存在两种由不同的同源基因编码的NCED同功酶,分别命名为NCED1和NCED2。

[0686] 对于ABA的分解,酶ABA-8-羟化酶(ABA8OH-2,也称为CYP707A2)将ABA羟化,这是其分解代谢的一个步骤,从而导致了休眠和种子萌发的中断。

[0687] 靶向编码在大麦(Hordeum vulgare)HvNCED1(登记号AK361999,SEQ ID NO:109)

或HvNCED2(登记号AB239298;SEQ ID NO:110)和小麦中相应的同源基因的ledRNA构建体被设计用于在大麦和小麦植物中进行转基因表达。这些构建体使用小麦和大麦NCED1和NCED2基因的高度保守区,小麦和大麦核苷酸序列在保守区中约97%相同。使用上述用于ledRNA的设计原理制备遗传构建体,其中分裂序列为反义序列且邻接序列为有义序列(图1A)。靶区长度为602个核苷酸,对应于SEQ ID NO:109的核苷酸435-1035。LedRNA的dsRNA区长度为602bp;dsRNA区中的有义序列是不间断的邻接的序列,对应于SEQ ID NO:110的核苷酸435-1035。编码NCED1和NCED2 ledRNA的核苷酸序列在本文中提供为SEQ ID NO:111和112。

[0688] 以相似的方式,制备靶向小麦(T.aestivum)和大麦(H.vulgare)的ABA-OH-2基因(登录号DQ145933,SEQ ID NO:113)的ledRNA构建体。靶区长度为600个核苷酸,对应于SEQ ID NO:113的核苷酸639-1238。ledRNA的dsRNA区长度为600bp;dsRNA区中的有义序列是不间断的邻接的序列,对应于SEQ ID NO:113的核苷酸639-1238。编码ledRNA的嵌合DNA的核苷酸序列提供为SEQ ID NO:114。

[0689] 将编码ledRNA的嵌合DNA插入处于Ubi基因启动子控制下的表达载体中,所述启动子在包括发育中的种子在内的大多数组织中组成型表达。切下表达盒并插入到二元载体中。这些用于产生转化的小麦植物。

[0690] 转基因小麦植株生长至成熟,由其获得种子,并分析NCED或ABA-OH-2基因的表达降低以及与基因表达降低相对应的对种子休眠的影响。预期休眠程度改变的表型范围。为了调节改变的表型的程度,产生修饰的遗传构建体,用于在双链RNA区中表达具有G:U碱基对的ledRNA,特别是用于其中ledRNA的双链区中15-25%的核苷酸参与G:U碱基对的ledRNA,作为双链区中核苷酸总数的百分比。

[0691] 靶向拟南芥(A.thaliana)EIN2基因的ledRNA

[0692] 如实施例10所述,拟南芥(A.thaliana)的EIN2基因编码参与乙烯感知的受体蛋白。当在ACC上发芽时,EIN2突变体幼苗相对于野生型幼苗显示下胚轴伸长。由于该基因在种子发芽后不久在幼苗中表达,因此相对于预先形成的RNA的外源递送,通过转基因手段递送ledRNA被认为是测试EIN2下调程度的最适合的途径。

[0693] 设计靶向拟南芥(Arabidopsis thaliana)EIN2基因(SEQ ID NO:115)的ledRNA构建体,其靶向靶基因mRNA的400个核苷酸区。通过将编码ledRNA的序列(SEQ ID NO:116)插入包含35S启动子的载体以在拟南芥(A.thaliana)植物中表达ledRNA来制备构建体。产生转基因拟南芥(A.thaliana)植物并通过QPCR测试EIN2基因表达的降低和在ACC存在下测试下胚轴长度测定。在一些转基因品系的植物中观察到EIN2表达水平降低和下胚轴长度增加。

[0694] 靶向拟南芥(A.thaliana)CHS基因的ledRNA

[0695] 植物中的查耳酮合酶(CHS)基因编码催化4-香豆酰-CoA和丙二酰-CoA转化为柚皮素查耳酮的酶,该酶是类黄酮生物合成中的第一个有保证的酶。类黄酮是一类主要在植物中发现的有机化合物,参与防御机制和胁迫耐受性。

[0696] 设计靶向拟南芥(Arabidopsis thaliana)CHS基因(SEQ ID NO:117)的ledRNA构建体,其靶向靶基因mRNA的338个核苷酸区。通过将编码ledRNA的DNA序列(SEQ ID NO:118)插入包含35S启动子的载体以在拟南芥(A.thaliana)植物中表达ledRNA来制备构建体。转基因拟南芥(A.thaliana)植物通过用二元载体中的遗传构建体转化产生,并通过QPCR测试

CHS基因表达的降低和类黄酮产生的降低。在一些转基因品系的植物中,例如在转基因种子的种皮中观察到CHS表达水平的降低和类黄酮水平的降低。

[0697] 靶向狭叶羽扇豆(Lupinus angustifolius)的LanR基因的ledRNA

[0698] 狭叶羽扇豆(Lupinus angustifolius L.)的LanR基因编码与赋予烟草花叶病毒(TMV)抗性的烟草N基因序列相关的多肽。

[0699] 设计并制备一种用于产生靶向狭叶羽扇豆(L.angustifolius)的LanR基因的ledRNA分子的嵌合DNA(登录号:XM_019604347,SEQ ID NO:119)。使用上述用于ledRNA的设计原理制备遗传构建体,其中分裂序列为反义序列且邻接序列为有义序列(图1A)。编码ledRNA的核苷酸序列在本文中提供为SEQ ID NO:120。通过体外转录产生LedRNA,纯化和浓缩,并将等分的RNA施用到含有LanR基因狭叶羽扇豆(L.angustifolius)植物的叶上。将病毒样品施用于处理和未处理的植物,几天后比较疾病症状。

[0700] 实施例19:靶向昆虫基因的RNA构建体

[0701] 引言

[0702] 蚜虫是吸汁昆虫,其通过饲喂植物汁液直接地引起,并且在一些情况下,通过传播在植物中引起疾病的各种病毒而间接地引起对植物的实质性和有时严重的损害。虽然在一些情况下Bt毒素在保护作物植物免受咀嚼昆虫方面是有效的,但它通常对于吸汁昆虫不是有效的。含有抗性基因的植物栽培种可能是防治蚜虫的有效方法,然而,大多数抗性基因对某些蚜虫物种或生物型是高度特异性的,并且由于新生物型通过遗传或表观遗传变化而迅速进化,因此抗性常常是过度出现的。此外,抗性基因在许多作物中是不可达的,或对于某些常见蚜虫物种,如侵染广泛宿主物种的绿桃蚜虫,可能不存在。目前主要通过频繁施用导致蚜虫抗药性的农药来防治蚜虫。例如,在澳大利亚仅一种农药作用模式组仍然有效对抗桃蚜,因为桃蚜已设法获得对所有其他注册的杀虫剂的抗性。

[0703] 在一些研究中已经显示RNAi介导的基因沉默可用作许多蚜虫物种中的研究工具,综述参见Scott等,2013;Yu等,2016,但未显示有效保护植物免受蚜虫攻击。在这些研究中,靶向参与蚜虫生长和发育、侵染或饲养过程的关键基因的dsRNA通过直接注射至蚜虫或通过在含有dsRNA的人工膳食上饲养蚜虫来递送。

[0704] 为了测试诸如本文所述的ledRNA分子的经修饰的RNAi分子防制吸汁昆虫的潜力,本发明人出于若干原因选择了桃蚜(Myzus persicae)作为吸汁昆虫的模型。首先,桃蚜是一种多食昆虫,在世界范围内广泛地侵染宿主植物种类,包括主要的粮食作物和园艺作物。其次,桃蚜是诸如甜菜西黄病毒的一些破坏性病毒传播的主要原因,而甜菜西黄病毒在一些油菜种植区已经造成了严重的危害。选择2个蚜虫基因作为下调的靶基因,1个编码关键效应蛋白(COO2),第二个编码活化蛋白激酶C受体(Rack-1)。COO2蛋白是蚜虫唾液腺蛋白,其对于蚜虫取食其宿主植物是必需的(Mutti等,2006;Mutti等,2008年)。Rack1是结合活化的蛋白激酶C的细胞内受体,所述活化的蛋白激酶C是主要参与信号转导级联的酶(McCahill等,2002;Seddas等,2004)。MpCOO2主要在蚜虫唾液腺中表达,并且MpRack1主要在肠中表达。在以前的研究中,通过直接注射或人工饮食饲养使用RNAi导致所测试的几种蚜虫物种的死亡(Pitino等,2011;Pitino和Hogenhout,2012;Yu等,2016)。

[0705] 材料和方法:蚜虫培养和植物材料

[0706] 在澳大利亚西部采集桃蚜(Myzus persicae)。在每次实验之前,在饲养室中在环

境光下在萝卜植物(Raphanus sativus L.)上饲养蚜虫。将蚜虫转移到具有精细漆刷的实验人工饮食笼中。

[0707] 用于蚜虫饲养的人工饮食的组分与Dadd和Mittler(1966)中描述的相同。用于蚜虫人工饮食的装置使用具有1cm直径和1cm高度的塑料管。将100μl含有或不含ledRNA的人造蚜虫饮食包封在两层石蜡膜之间以产生饮食小袋。在该小袋的顶部,有一个供蚜虫四处移动的室,并且通过将它们的管心针刺穿该拉伸的石蜡膜的顶层来从饮食中进食。使用细漆刷将8只一龄或二龄若虫轻轻转移到蚜虫室中。该实验在20℃的生长箱中进行。

[0708] 在一个实验中使用的烟叶收集自22℃在16小时光照/8小时黑暗循环下生长在土壤中的植物。

[0709] MpC002和MpRack-1基因和LedRNA构建体

[0710] 所测试的作为靶基因的桃蚜MpC002和MpRack-1基因与Pitino等(2011;2012)所描述的相同。从NCBI网站获得两个基因的DNA序列,MpC002(>MYZPE13164_0_v1.0_000024990.1|894nt)和MpRack-1(>MYZPE13164_0_v1.0_000198310.1|960nt)。这两个基因的cDNA序列在本文中提供为SEQ ID NO:123和124。以与前面实施例所述相同的方式设计ledRNA构建体。编码ledRNA分子的DNA序列在本文中提供为SEQ ID NO:125和126,用作转录模板以合成ledRNA。将编码靶向MpC002和MpRack-1基因的ledRNA分子的载体DNA导入大肠杆菌菌株DH5α,以制备用于体外RNA转录的质粒DNA,并导入大肠杆菌菌株HT115,以用于体内(细菌中)转录。

[0711] ledRNA分子对减少蚜虫性能的效力

[0712] 为了检查靶向MpC002或MpRack-1基因的ledRNA是否影响蚜虫性能,将每种ledRNA通过如实施例1所述的人工饮食途径递送至蚜虫。在每个实验中,设置10个生物复制品;每个生物复制品有8个1-2龄桃蚜若虫。每个实验中的对照使用等浓度的不相关的ledRNA,即ledGFP。

[0713] 在每种ledRNA分子的50ng/μl的较低浓度下,从含有MpC002或MpRack-1ledRNA的人工饮食饲喂后的蚜虫存活与对照ledGFP没有显著差异。然而,靶向MpC002基因的ledRNA显著(P<0.05)降低了桃蚜的繁殖率(图37)。与用具有对照ledRNA的对照饮食维持的成年蚜虫产生的若虫数量相比,每个成蚜产生的若虫平均数量减少约75%。在200ng/μl的较高浓度下,靶向MpC002或MpRack-1的ledRNA增加了成蚜死亡率(图37B)。在24小时后,还观察到包括MpC002或MpRack-1ledRNAs在内的饮食中蚜虫的存活率降低,并在实验的5天中持续进行。结果表明,靶向蚜虫必需基因的ledRNA的使用能引起蚜虫的死亡,减少蚜虫的繁殖。将各ledRNA的效力与靶向靶基因的相同区的发夹RNA分子(hpRNAi)进行比较。

[0714] 蚜虫对ledRNA分子的吸收

[0715] 为了跟踪蚜虫内ledRNA的摄入和分布,在实施例1所述的合成过程中,用Cya3(花青染料标记的核苷酸三磷酸)标记靶向MpC002或MpRack-1基因的ledRNA。已经报道Cyr3标记对常规dsRNA分子的生物学功能没有影响,因此可以用作荧光检测的标记。使用Leica EL6000微系统仪器使用共聚焦显微镜检查已经喂饲标记的ledRNA的蚜虫。靶向MpC002或MpRack-1的Cyr3标记的ledRNA可在人工饮食喂饲数小时内、随后在生殖系统中、甚至在作为已经饲养的成虫后代的新生若虫中检测到。结果表明,蚜虫对消化系统功能或繁殖起关键作用的基因可以通过喂饲成为ledRNA分子的有效靶标。

[0716] ledRNA稳定性

[0717] 为了检测饮食中和从饲喂蚜虫回收的ledRNA的稳定性,在饲喂含有标记的ledRNA分子的饮食后从人工饮食和蚜虫蜜露回收RNA。将RNA样品在凝胶上电泳并通过荧光检测进行检查。在饲喂前,ledMpC002RNA在琼脂糖凝胶上清楚地显示出约700bp的单一产物。从人工饮食中回收的RNA显示大小为100-700bp的RNA涂片,表明在室温下暴露于饮食25天后有一些降解,但仍大部分完整。从蚜虫蜜露中回收的RNA在350-700bp的RNA范围内显示荧光,因此在很大程度上也是完整的。尽管有部分ledRNA降解,但是大部分ledRNA分子能够在人工饮食以及蚜虫蜜露中保持完整相当长的时间。ledRNA分子的这种程度的稳定性应当允许ledRNA具有活性并且当外源施用时保持活性。

[0718] 植物叶片对标记ledRNA的吸收

[0719] 将Cy3标记的ledMpC002 RNA涂在烟草叶片的上表面,以观察其是否能够穿透叶片组织。将10微升Cy3标记的ledMpC002(浓度为1μg/μl)涂在直径为2cm的圆圈中,并用黑色记号笔标记施涂区域。使用Leica EL 6000微系统仪,在5小时的时间段内捕获了在525nm激发下的叶片荧光图像,将着色的组织与未着色的组织进行比较。在施用后1小时内,可在叶肉组织中清楚地检测到Cy3标记物,因此Cy3标记物清楚地穿透叶表面上的蜡质表皮层。荧光水平在2小时时增加,并保持到第5小时的时间点。不清楚ledRNA分子是进入细胞还是进入细胞核。然而,由于吸汁昆虫特别是从植物叶和茎的韧皮部筛分子取食,所以RNA传递到植物细胞中对于蚜虫基因的沉默不是必需的。实验表明,经局部给药,在植物组织中发现了ledRNA分子。

[0720] 结论

[0721] 这项研究的目的是使用ledRNA设计来测试外源RNAi在防治蚜虫方面的应用,该蚜虫是主要是一种吸吮树液的害虫,是世界范围的大问题。已知蚜虫具有RNAi机制来加工外源RNA(Scott等,2013;Yu等,2016)。在此,通过含有靶向MpC002或MpRack-1基因的ledRNA分子的人工饮食口服递送能够引起蚜虫死亡并减少蚜虫的繁殖。这些分子针对两种不同的靶基因进行了测试,一种编码效应蛋白C002,另一种编码活化蛋白激酶的受体(Rack-1),这对于桃蚜(Myzus persicae)的摄食和发育必不可少。当添加到浓度低至50ng/μl的人工饮食中时,靶向这些基因的ledRNA分子显著降低蚜虫繁殖。在200ng/μl的较高浓度下,ledRNA也增加蚜虫死亡率。当使用Cy3标记研究ledRNA摄取时,在喂饲人工饮食数小时内,在蚜虫皮中,随后在生殖系统中,甚至在作为喂饲成虫子代的新生若虫中观察到ledRNA分子。

[0722] 还显示,在至少三周和半周的人工饮食中,ledRNA分子大部分保持完整。在蚜虫蜜露中也发现了大部分完整的ledRNA分子,蚜虫蜜露是蚜虫的排泄产物。在植物叶片上施用标记的ledRNA,它可以进入蚜虫取食的韧皮部。总而言之,这些结果表明ledRNA用于防治蚜虫和其他吸汁昆虫的强烈潜力,包括通过经由饮食的外源递送,提供用于管理蚜虫和其他吸汁昆虫的实用方法。这些RNA分子还可以在转基因植物中表达,使用有利于韧皮部组织中RNA合成的启动子来防治蚜虫和其他吸汁昆虫。

[0723] 实施例20:靶向其他昆虫基因的RNA构建体

[0724] 靶向昆虫基因的LedRNA

[0725] 棉铃虫(Helicoverpa armigera)是鳞翅目的昆虫害虫,也称为棉铃虫或玉米螟。棉铃虫(H.armigera)的幼虫对包括许多重要栽培作物在内的多种植物取食,并造成每年价

值数十亿美元的相当大的作物损害。幼虫是多食的和世界性的害虫,可以取食广泛的植物物种,包括棉花、玉米、番茄、鹰嘴豆、木豆、苜蓿、水稻、高粱和豇豆。

[0726] 选择棉铃虫(H.armigera)ABC转运蛋白白基因(ABCwhite)作为具有容易检测表型的靶基因,以测试昆虫幼虫中的ledRNA和ledRNA(G:U)构建体。ABC转运蛋白属于ATP结合盒式转运蛋白超家族-例如,在棉铃虫(Helicoverpa)基因组中鉴定出54个不同的ABC转运蛋白基因。ABC转运蛋白编码携带跨越膜的多种分子中的任何一种或多种的膜结合蛋白。蛋白质使用由ATP水解释放的能量将分子转运穿过膜。一些ABC转运蛋白与棉铃虫H.armigera中的植物次生代谢产物降解有关(Khan等,2017)。ABCwhite蛋白将眼色素和蝶啶途径前体转运到眼中的色素颗粒中,敲除突变体呈现白眼。

[0727] ABCwhite基因的核苷酸序列提供为SEQ ID NO:127(登录号:KU754476)。为了测试ledRNA是否能够通过外源递送RNA到幼虫食物中来降低ABCwhite基因表达,设计并制备了编码ledRNA的遗传构建体。使用用于ledRNA的设计原理制备遗传构建体,其中分裂序列是有义序列,邻接序列是反义序列(图1B)。靶区长度为603个核苷酸,对应于SEQ ID NO:127的核苷酸496-1097。ledRNA的dsRNA区长度为603bp;dsRNA区中的反义序列是不间断的邻接的序列,对应于SEQ ID NO:127的核苷酸496-1097的互补体。编码ledRNA的核苷酸序列在本文中提供为SEQ ID NO:128。编码区在T7 RNA聚合酶启动子的调控下进行体外转录。

[0728] 用限制性内切酶SnaBI消化编码ledRNA的基因构建体,在ledRNA编码区的下游切割,并根据转录试剂盒的说明用RNA聚合酶T7进行体外转录。将RNA加入人工饮食中,提供给棉铃虫(H.armigera)幼虫。

[0729] 制备在双链茎中具有G:U碱基对的相应的ledRNA构建体,并与规范碱基配对的ledRNA进行比较。

[0730] 靶向蚂蚁中基因的ledRNA

[0731] 阿根廷蚁(Linepithema humile),通常称为阿根廷蚁,是一种在几个大陆广泛流行的昆虫害虫。选择编码信息素生物合成激活神经肽(PBAN)神经肽样(LOC105673224)的阿根廷蚁(L.humile)基因作为靶基因,参与信息素与昆虫之间的通讯。

[0732] PBAN基因的核苷酸序列提供为SEQ ID NO:129(登录号:XM_012368710)。为了测试ledRNA是否能够通过将该RNA以诱饵的形式外源递送到饮食中来降低PBAN基因表达,设计并制备了靶向该基因的编码ledRNA的遗传构建体。使用用于ledRNA的设计原理制备遗传构建体,其中分裂序列是有义序列,邻接序列是反义序列(图1B)。靶区长度为540个核苷酸,对应于SEQ ID NO:129的核苷酸136-675。ledRNA的dsRNA区长度为540bp;dsRNA区中的反义序列是不间断的邻接的序列,对应于SEQ ID NO:129的核苷酸136-675的互补体。编码ledRNA的核苷酸序列在本文中提供为SEQ ID NO:130。编码区在T7 RNA聚合酶启动子的调控下进行体外转录。

[0733] 用限制性内切酶SnaBI消化编码ledRNA的基因构建体,在ledRNA编码区的下游切割,并根据转录试剂盒的说明用RNA聚合酶T7进行体外转录。将RNA包被在玉米粉上,用于口服递送到阿根廷蚁(L.humile)。

[0734] 靶向铜绿蝇(L.cuprina)基因的LedRNA

[0735] 铜绿蝇(Lucilia cuprina)是一种昆虫害虫,更普遍地称为澳洲绵羊绿头蝇。属于蚜蝇科(Calliphoridae),为昆虫双翅目(Diptera)的一员。选择了五个靶基因以进行

ledRNA构建体测试,即编码铜绿蝇(L.cuprina)的V型质子ATPase催化亚基A(登录号XM_023443547)、RNAse 1/2(登录号XM_023448015)、几丁质合酶(登录号XM_023449557)、蜕皮激素受体(EcR;登录号U75355)、和γ-微管蛋白1/1样(登录号XM_023449717)的基因。使用ledRNA的设计原理制备每个遗传构建体,其中分裂序列是有义序列,邻接序列是反义序列(图1B)。在每种情况下,靶区长度为约600个核苷酸,dsRNA区中的反义序列是不间断的邻接序列。编码靶向ATP酶-A基因的ledRNA的核苷酸序列在本文中提供为SEQ ID NO:131。编码靶向RNA酶1/2基因的ledRNA的核苷酸序列在本文中提供为SEQ ID NO:132。编码靶向几丁质合酶基因的ledRNA的核苷酸序列在本文中提供为SEQ ID NO:133。编码靶向EcR基因的ledRNA的核苷酸序列在本文中提供为SEQ ID NO:134。编码靶向γ-微管蛋白1/1-样基因的ledRNA的核苷酸序列在本文中提供为SEQ ID NO:135。在每个构建体中,编码区在T7 RNA聚合酶启动子的调控下进行体外转录。

[0736] 本领域技术人员应当理解,在不脱离广泛描述的本发明的精神或范围的情况下,可以对具体实施例中所示的本发明进行多种变化和/或修改。因此,本实施例在所有方面都被认为是说明性的而非限制性的。

[0737] 本申请要求2017年9月15日提交的AU 2017903773,2018年8月3日提交的AU2018902840和2018年8月8日提交的AU2018902896的优先权,其公开内容通过引用并入本文。

[0738] 本文讨论和/或引用的所有出版物均全文引入本文。

[0739] 包括在本说明书中的文件、动作、材料、装置、物品等的任何讨论仅仅是为了提供本发明的上下文。不应视为承认这些内容中的任何或全部形成现有技术基础的一部分或者是与本发明相关的领域中的公知常识,因为其存在于本申请的每个权利要求的优先权日之前。

[0740] 参考文献

[0741] Abraham(2002).Cell 111:9-12.

[0742] Acevedo-Garcia et al.,(2017).Plant Biotechnology Journal 15:367-378.

[0743] Alvarez et al.,(2000).Theor Appl Genet 100:319-327.

[0744] Baumlein et al.,(1991).Mol.Gen.Genet.225:459-467.

[0745] Baumlein et al.,(1992).Plant J.2:233-239.

[0746] Bhattacharyya et al.,(1990)Cell 60:155-122.

[0747] Brar et al.,(1996)Biotech Genet.Eng Rev 13:167-79.

[0748] Broothaerts et al.,(2005).Nature 433:629-633.

[0749] Broun et al.,(1998).Plant J.13:201-210.

[0750] Busk et al.,(1997).Plant J.11:1285-1295.

[0751] Chen et al.,(2005).Functional Plant Biology 32:671-681.

[0752] Chikwamba et al.,(2003).Proc.Natl.Acad.Sci.U.S.A.100:11127-11132.

[0753] Christou and Klee,Handbook of Plant Biotechnology,John Wiley and Sons (2004).

[0754] Chung et al.,(2006).BMC Genomics 7:120.

[0755] Clough and Bent(1998).Plant J.16:735-743.

[0756] Corrado and Karali(2009).Biotechnol.Adv.27:733-743.

[0757] Courvalin et al.,(1995).Life Sci.318:1209-1212.

[0758] Dadd and Mittler(1966).Experientia 22:832-833.

[0759] Darji et al.,(1997).Cell 91:765-775.

[0760] Dong et al.,(2011)Plant J.68:633-45.

[0761] Draper and Scott(1988).In:J.Draper et al.,(Eds.),Plant Genetic Transformation and Gene Expression:A Laboratory Manual,Alden Press,Oxford, pp.199-236.

[0762] Dunwell(2000).J Exp Botany 51Spec No:487-496.

[0763] Ellerstrom et al.,(1996).Plant Mol.Biol.32:1019-1027.

[0764] Elliott et al.,(2002).Mol.Plant Microbe Interact.15:1069-1077.

[0765] Ellis et al.,(1987).EMBO J 6:11-16.

[0766] Fehr,In:Breeding Methods for Cultivar Development,Wilcox J.ed., American Society of Agronomy,Madison Wis.(1987).

[0767] Gan(1995).Molecular characterization and genetic manipulation of plant senescence.PhD thesis.University of Wisconsin,Madison.

[0768] Gan and Amasino(1995).Science 270:1986-1988.

[0769] Glawe(2008).Ann.Rev.Phytopathol.46:27-51.

[0770] Gleave(1992).Plant Mol Biol 20:1203-1207.

[0771] Gupta et al.,(1988)Plant Mol.Biol.10:215-224.

[0772] Helliwell and Waterhouse(2005).Methods in Enzymology 392:24-35.

[0773] Hinchee et al.,(1988).Biotechnology 6:915-922.

[0774] Horvath et al.,(2000).Proc.Natl.Acad.Sci.U.S.A.97:1914-1919.

[0775] Hsieh and Fire(2000)Annu Rev Genet 14:187-204.

[0776] Jeddeloh,et al.,(1999).Nat.Genet.22:94-97.

[0777] Jefferson et al.,(1987).EMBO J 6:3901-3907.

[0778] Jepson et al.,(1994).Plant Mol.Biol.26:1855-1866.

[0779] Khan(2017).Sci Rep 7,40025

[0780] Kishore and Somerville(1993).Curr Opin Biotechnol.4:152-158.

[0781] Koziel et al.,(1996).Plant Mol.Biol.32:393-405.

[0782] Lacroix et al.,(2008).Proc.Natl.Acad.Sci.U.S.A.105:15429-15434.

[0783] Li et al.,(1996).FEBS Lett.379:117-121.

[0784] McCahill et al.,(2002).Molecular Pharmacology 62:1261-1273.

[0785] McCullough and Schuler(1997).Nucl Acids Res.25:1071-1077.

[0786] Mahfouz et al.,(2006).Plant Cell 18:477-490.

[0787] Matsuoka et al.,(1994).Plant J.6:311-319.

[0788] Meier et al.,(1997).FEBS Lett.415:91-95.

[0789] Melamed-Bessudo et al.,(2012)Proc.Natl.Acad.Sci.USA 109(16):e981-988.

[0790] Millar et al.,(2006).Plant J.45:942-954.

[0791] Mutti et al.,(2006).J Insect Sci 6:38.

[0792] Mutti et al.,(2008).Proceedings of the National Academy of Sciences 105:9965-9969.

[0793] Olive et al.,(1989)Plant Mol Biol 12:525-538.

[0794] Padidam(2003).Transgenic Res.12:101-109.

[0795] Perrin et al.,(2000).Mol Breed 6:345-352.

[0796] Pitino et al.,(2011).PLoS ONE 6,e25709.

[0797] Pitino and Hogenhout(2012).Molecular Plant-Microbe Interactions 26:130-139.

[0798] Potenza et al.,(2004).In Vitro Cell Dev.Biol.Plant 40:1-22.

[0799] Powell et al.,(1996).Vaccines 183,Abstract.

[0800] Preiss et al.,(1987).In:Tailoring Genes for Crop Improvement(Bruening et al.,eds.),Plenum Press,S.133-152.

[0801] Sambrook et al.,(1989).Molecular Cloning:A Laboratory Manual,Cold Spring Harbour Laboratory Press.

[0802] Schaffner(1980).Proc.Natl.Acad.Sci.U.S.A.77:2163-2167.

[0803] Schwartz et al.,(1997).Science 276:1872-1874.

[0804] Scott et al.,(2013).Journal of Insect Physiology 59:1212-1221.

[0805] Seddas et al.,(2004).Virology 325:399-412.

[0806] Shiina et al.,(1997).Plant Physiol.115:477-483.

[0807] Shure et al.,(1983).Cell 35:225-233.

[0808] Sizemore et al.,(1995).Science 270:299-302.

[0809] Slater et al.,Plant Biotechnology-The Genetic Manipulation of Plants,Oxford University Press(2003).

[0810] Smith et al.,(2000).Nature 407:319-320.

[0811] Stalker et al.,(1988).Science 242:419-423.

[0812] Stewart et al.,(2000).J Mol Biol 298:611-622.

[0813] Tan et al.,(2011).Plant Physiol.156:1577-1588.

[0814] Thillet et al.,(1988).J.Biol.Chem 263:12500-12508.

[0815] Timmons et al.,(2001).Gen 263:103-112.

[0816] Ulmasov et al.,(1995).Plant Physiol.108:919-927.

[0817] Wang MB.Isolation of phloem specific gene promoters for use in genetic engineering of insect resistance in rice.PhD thesis,University of Durham,UK.

[0818] Wang et al.,(1994)Plant Molecular Biology 24:159-170.

[0819] Wang et al.,(1998)Acta Horticulturae 461:401-407.

[0820] Wang et al.,(2008)RNA 14:903-913.Wang et al.,(2013).PLoS Genet 9,e1003865.

[0821] Wang et al.,(2014).Nature Biotechnology 32:9.

[0822] Weiss(2003).Int.J.Med.Microbiol.293:95:106.

[0823] Weissbach et al.,(1988).In:Methods for Plant Molecular Biology, Academic Press,San Diego,Calif.

[0824] Wijnker et al.,(2008).Trends in Plant Science 13:640-646.

[0825] Yang et al.,Particle Bombardment Technology for Gene Transfer,Oxford Press,Oxford,England(1994).

[0826] Yang et al.,(2003).Planta 216:597-603.

[0827] Yelina et al.,(2012).PLoS Genetics 8(8):e1002844.doi:10.1371/journal.pgen.1002844

[0828] Yu et al.,(2016).Pest Management Science 72:1090-1098.

[0829] Zhang et al.,(2018).Nat Rev Mol Cell Biol.19:489-506.

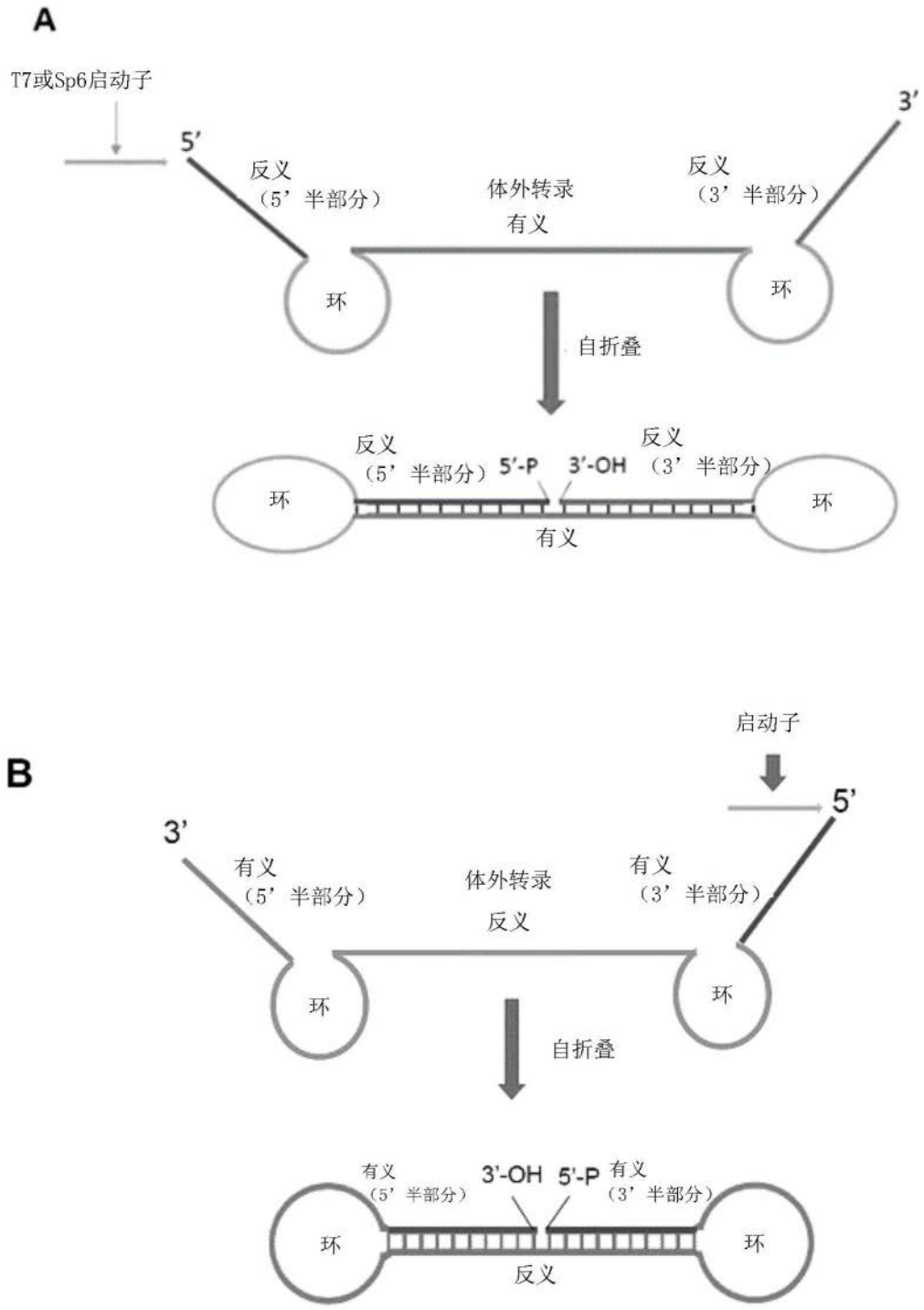

A
T7或Sp6启动子
5'
反义
（5'半部分）
体外转录
有义
反义
（3'半部分）
3'
环
环
自折叠
反义
（5'半部分）
5'-P
3'-OH
反义
（3'半部分）
环
环
有义
B
启动子
5'
3'
有义
（5'半部分）
体外转录
反义
有义
（3'半部分）
环
环
自折叠
有义
（5'半部分）
3'-OH
5'-P
有义
（3'半部分）
环
环
反义

图1

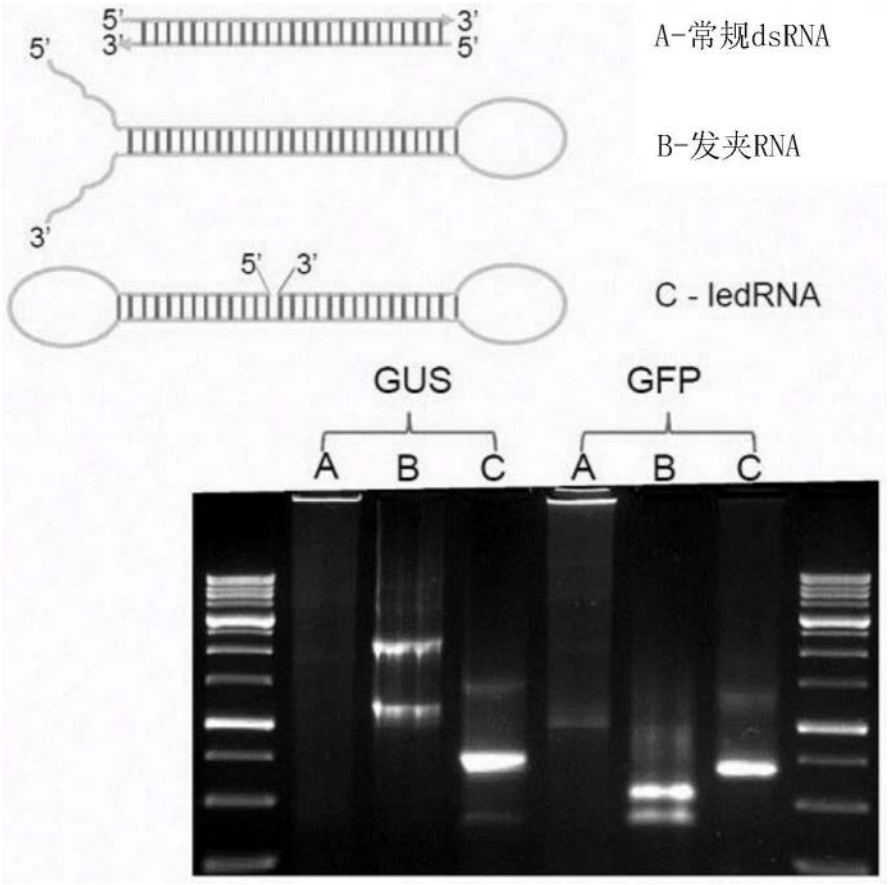


图2

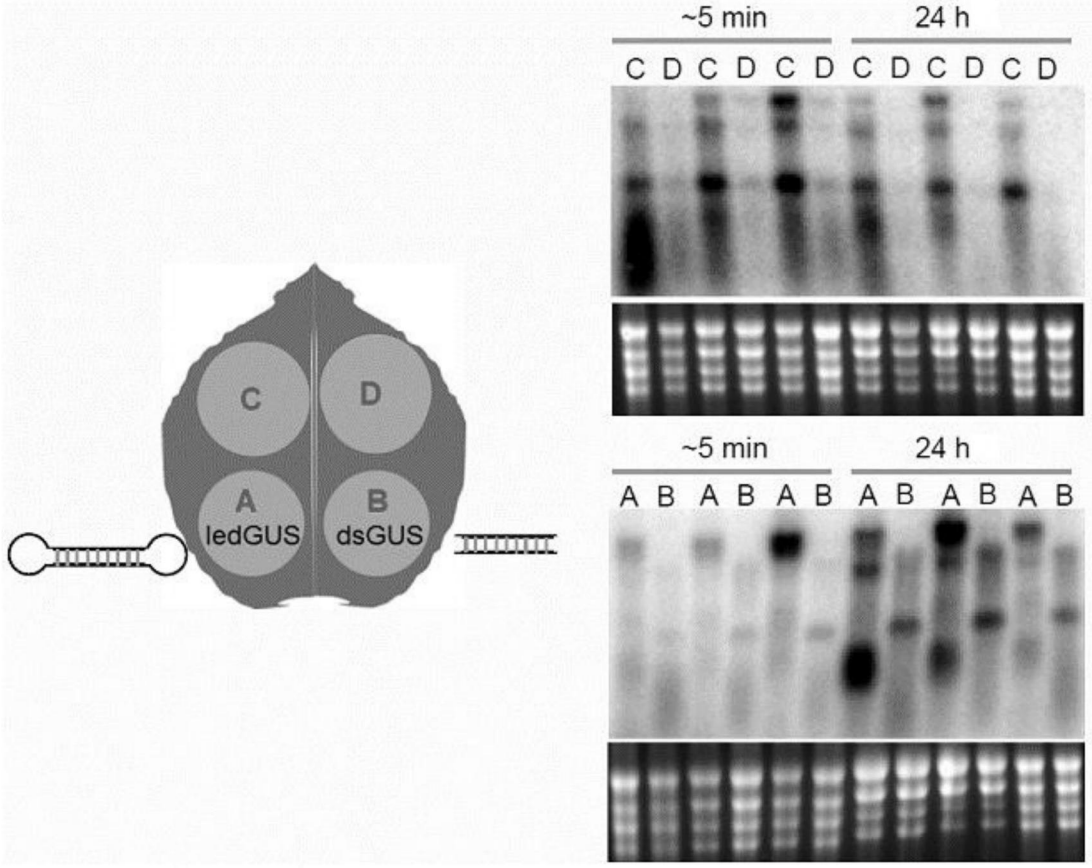

C
D
A
ledGUS
B
dsGUS
~5 min
24 h
C D C D C D C D C D C D
~5 min
24 h
A B A B A B A B A B A B

图3

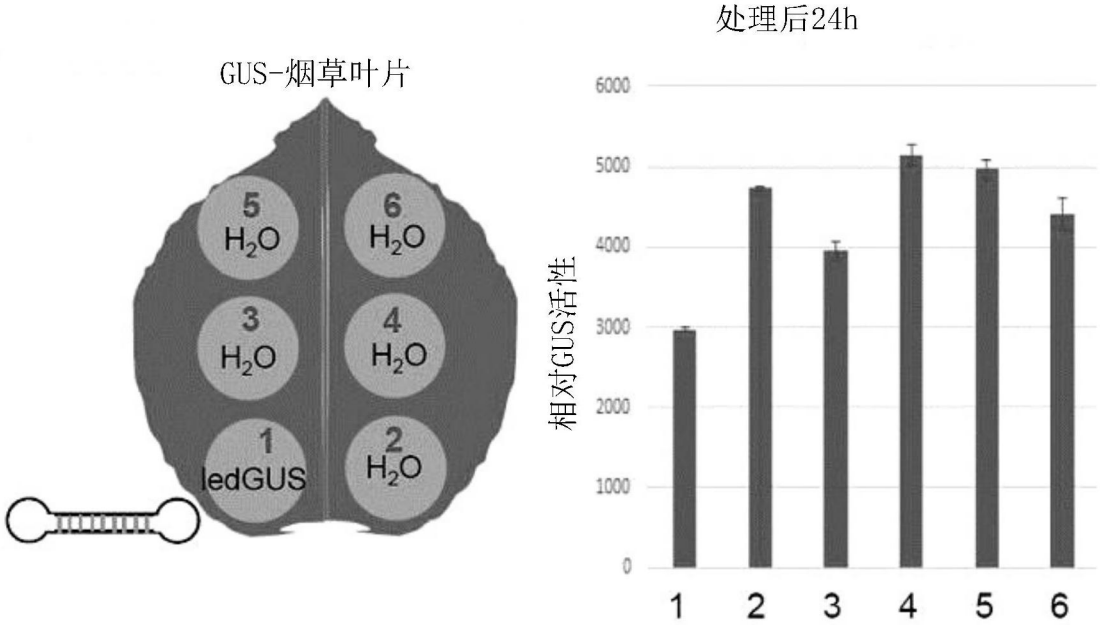

GUS-烟草叶片
5
H2O
6
H2O
3
H2O
4
H2O
1
ledGUS
2
H2O
处理后24h
相对GUS活性
6000
5000
4000
3000
2000
1000
0
1
2
3
4
5
6

图4

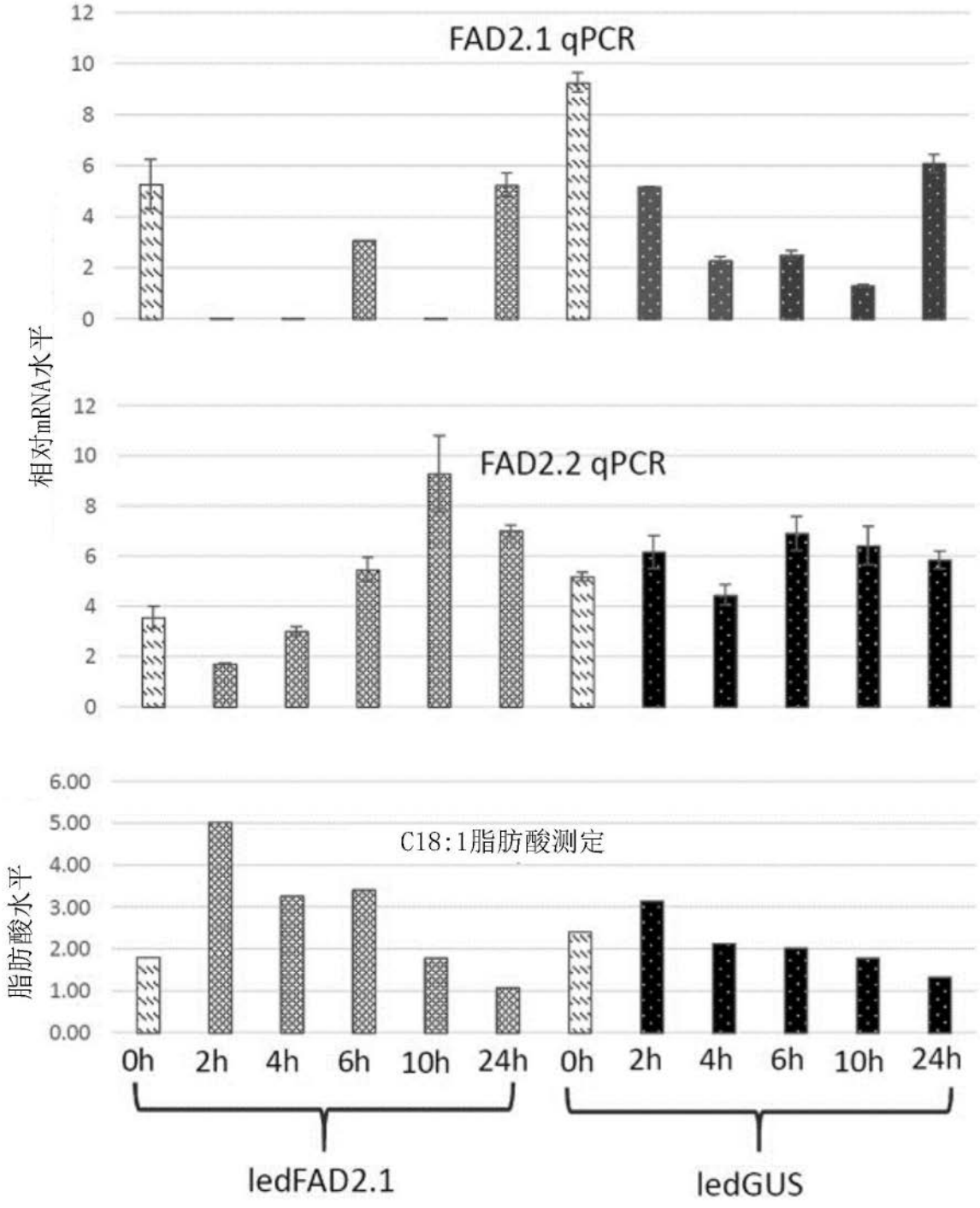

FAD2.1 qPCR
FAD2.2 qPCR
相对mRNA水平
C18:1脂肪酸测定
脂肪酸水平
0h
2h
4h
6h
10h
24h
ledFAD2.1
ledGUS

图5

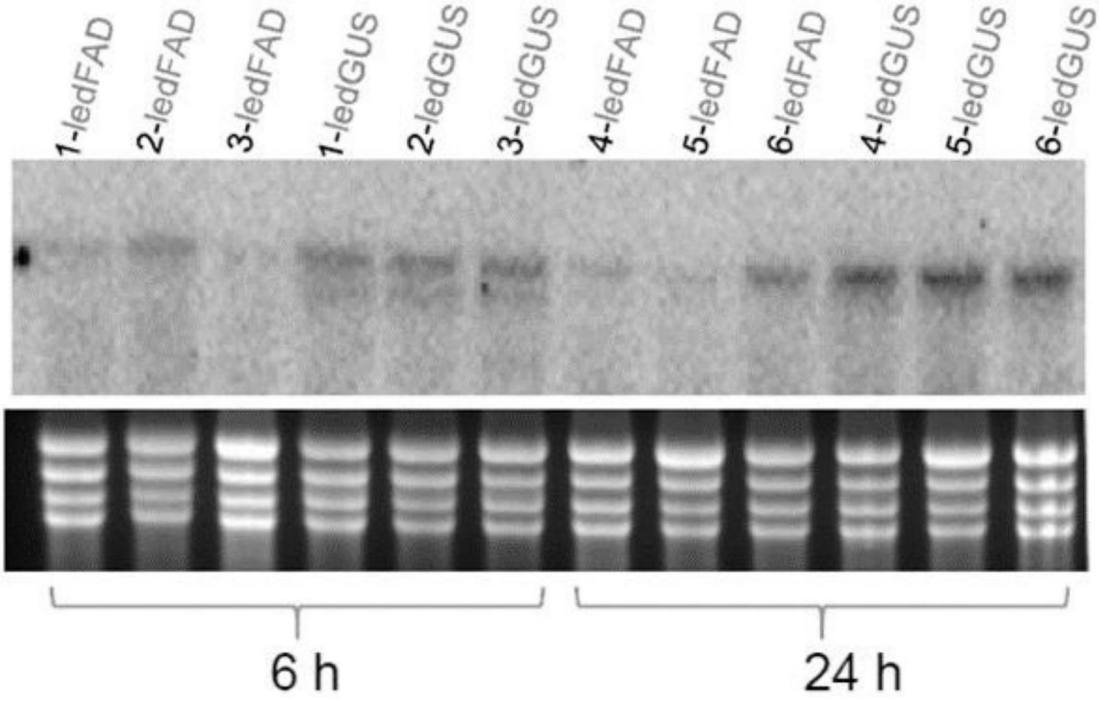


图6

与WT GUS ORF比对

```
GUS             GAGTGTGATATCTACCCGCTTCGCGTCGGCATCCGGTCAGTGGCAGTGAAGGGCGAACAG 840
GUS-GU          ----------------------TTGTGTTGGTATTTGGTTAGTGGTAGTGAAGGGTGAATAG 40
                                      * * ** ** **  *** ***** ********* *** **

GUS             TTCCTGATTAACCACAAACCGTTCTACTTTACTGGCTTTGGTCGTCATGAAGATGCGGAC 900
GUS-GU          TTTTTGATTAATTATAAATTGTTTTATTTTATTGGTTTTGGTTGTTATGAAGATGTGGAT 100
                **  *******  * ***  *** ** **** *** ****** ** ********* ***

GUS             TTGCGTGGCAAAGGATTCGATAACGTGCTGATGGTGCACGACCACGCATTAATGGACTGG 960
GUS-GU          TTGTGTGGTAAAGGATTTGATAATGTGTTGATGGTGTATGATTATGTATTAATGGATTGG 160
                *** **** ******** ***** *** ******** * **  * * ********* ***

GUS             ATTGGGGCCAACTCCTACCGTACCTCGCATTACCCTTACGCTGAAGAGATGCTCGACTGG 1020
GUS-GU          ATTGGGGTTAATTTTTATTGTATTTTGTATTATTTTTATG-------------------- 200
                *******  ** *  **  ***  * * ****   *** *
```

图7

```
GUS             GAGTGTGATATCTACCCGCTTCGCGTCGGCATCCGGTCAGTGGCAGTGAAGGGCGAACAG 840
GUS-4M          --------------------TCGGGTCCGCAACCGCTCACTGGGAGTCAAGCGCGTACAC 40
                                    *** *** ***:*** *** *** *** *** ***:***

GUS             TTCCTGATTAACCACAAACCGTTCTACTTTACTGGCTTTGGTCGTCATGAAGATGCGGAC 900
GUS-4M          TTCGTGAATAAGCACTAACGGTTGTACATTAGTGGGTTTCGTCCTCAAGAACATGGGGAG 100
                *** ***:*** ***:*** *** ***:*** *** *** *** ***:*** *** ***

GUS             TTGCGTGGCAAAGGATTCGATAACGTGCTGATGGTGCACGACCACGCATTAATGGACTGG 960
GUS-4M          TTGGGTGCCAATGGAATCGTTAAGGTGGTGAAGGTCCACCACCTCGCTTTATTGGTCTGC 160
                *** *** ***:***:***:*** *** ***:*** *** ***:***:***:***:***

GUS             ATTGGGGCCAACTCCTACCGTACCTCGCATTACCCTTACGCTGAAGAGATGCTCGACTGG 1020
GUS-4M          ATTCGGGGCAAGTCCAACCCTACGTCGGATTTCCCATACC-------------------- 200
                *** *** *** ***:*** *** *** ***:***:***
```

图8

与WT GUS ORF比对

```
GUS             GAGTGTGATATCTACCCGCTTCGCGTCGGCATCCGGTCAGTGGCAGTGAAGGGCGAACAG 840
GUS-10M         --------------------TCGCGTCGCGATCCGGTCTCTGGCAGTGTTGGGCGAACTC 40
                                    ********  ********: ********::********:

GUS             TTCCTGATTAACCACAAACCGTTCTACTTTACTGGCTTTGGTCGTCATGAAGATGCGGAC 900
GUS-10M         TTCCTGATATACCACAAAGGGTTCTACTAAACTGGCTTACGTCGTCATCTAGATGCGGTG 100
                ********::********  ********::********: ******** :********:

GUS             TTGCGTGGCAAAGGATTCGATAACGTGCTGATGGTGCACGACCACGCATTAATGGACTGG 960
GUS-10M         TTGCGTGGGTAAGGATTCCTTAACGTGCACATGGTGCAGCACCACGCAAAAATGGACTCC 160
                ******** :******** :********: ********  ********::********

GUS             ATTGGGGCCAACTCCTACCGTACCTCGCATTACCCTTACGCTGAAGAGATGCTCGACTGG 1020
GUS-10M         ATTGGGGCGTACTCCTACGCTACCTCGCTATACCCTTAGC-------------------- 200
                ******** :********  ********::********
```

图9

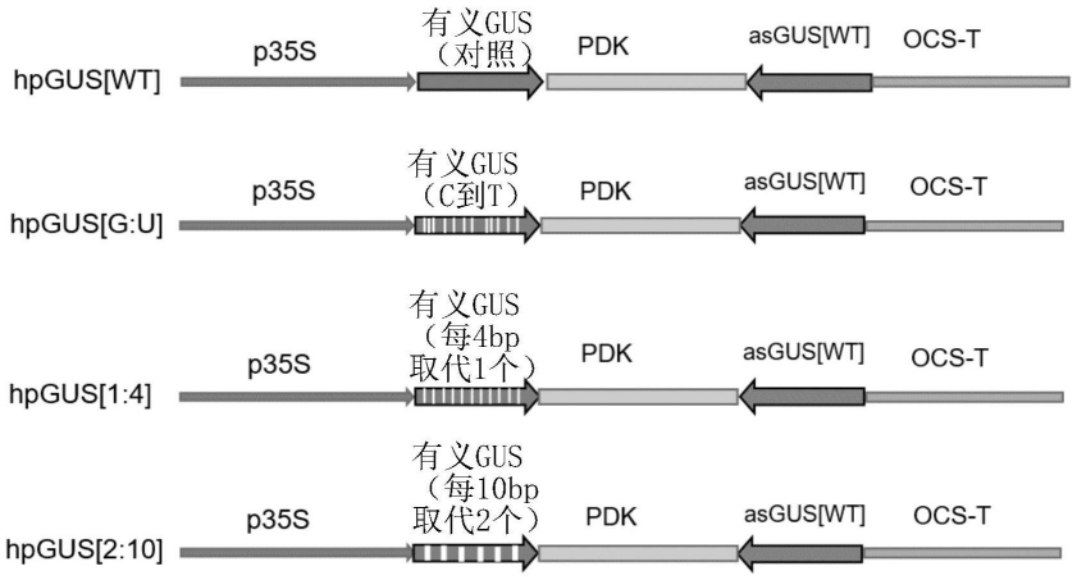


图10

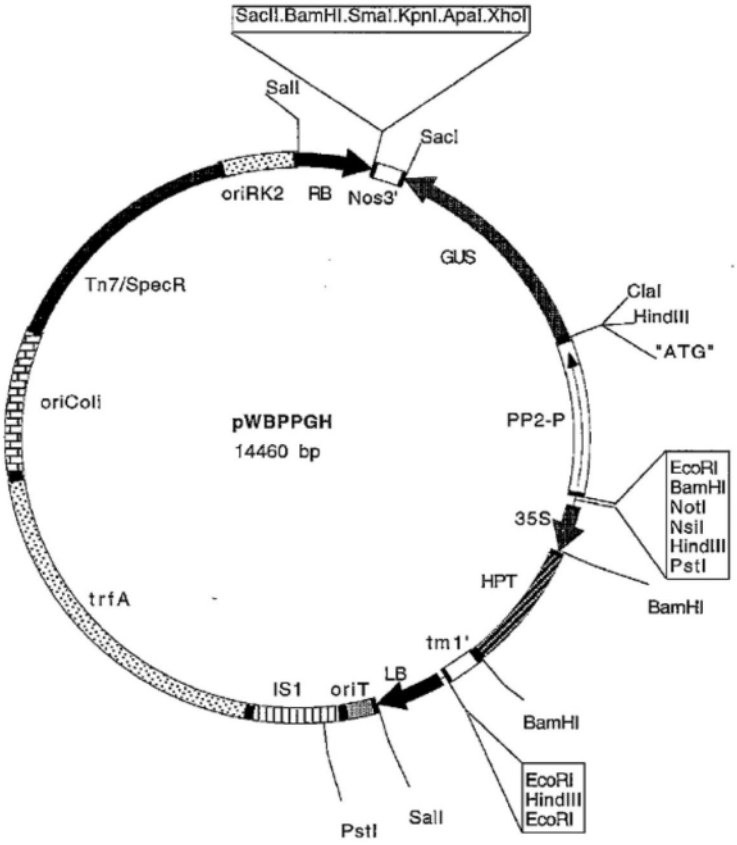

SacII.BamHI.SmaI.KpnI.ApaI.XhoI
SalI
SacI
oriRK2
RB
Nos3'
GUS
Tn7/SpecR
ClaI
HindIII
"ATG"
oriColi
pWBPPGH
14460 bp
PP2-P
EcoRI
BamHI
NotI
NsiI
HindIII
PstI
35S
HPT
BamHI
trfA
tm1'
LB
IS1
oriT
BamHI
EcoRI
HindIII
EcoRI
PstI
SalI

图11

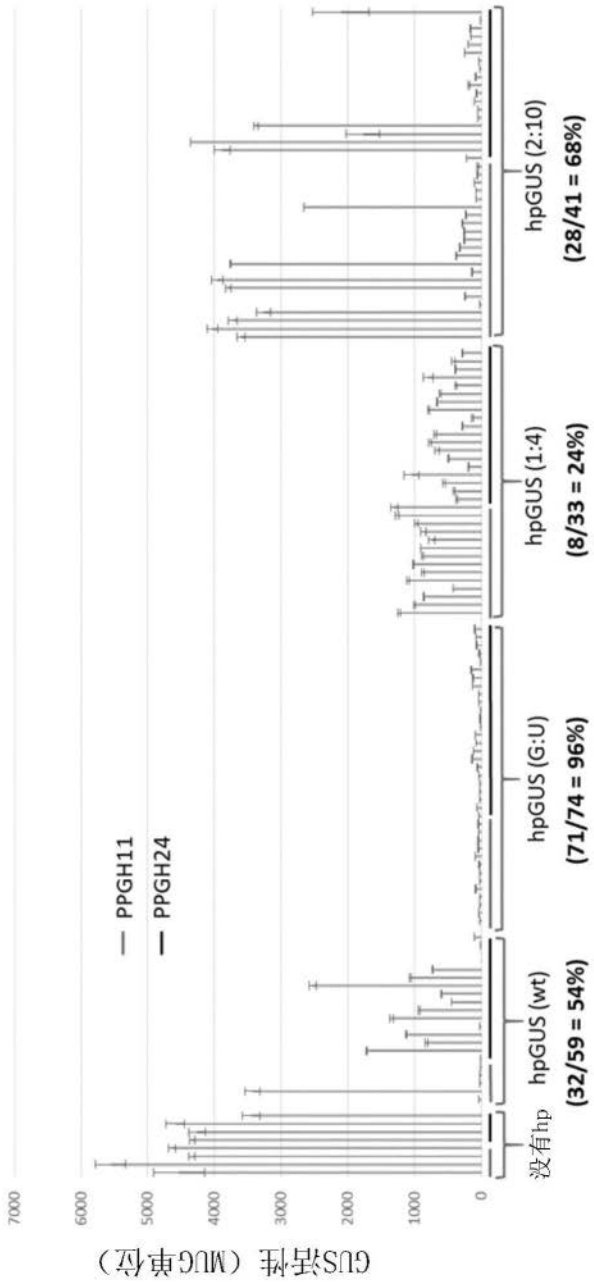

GUS活性（MUG单位）
7000
6000
5000
4000
3000
2000
1000
0
PPGH11
PPGH24
没有hp
hpGUS (wt)
(32/59 = 54%)
hpGUS (G:U)
(71/74 = 96%)
hpGUS (1:4)
(8/33 = 24%)
hpGUS (2:10)
(28/41 = 68%)

图12

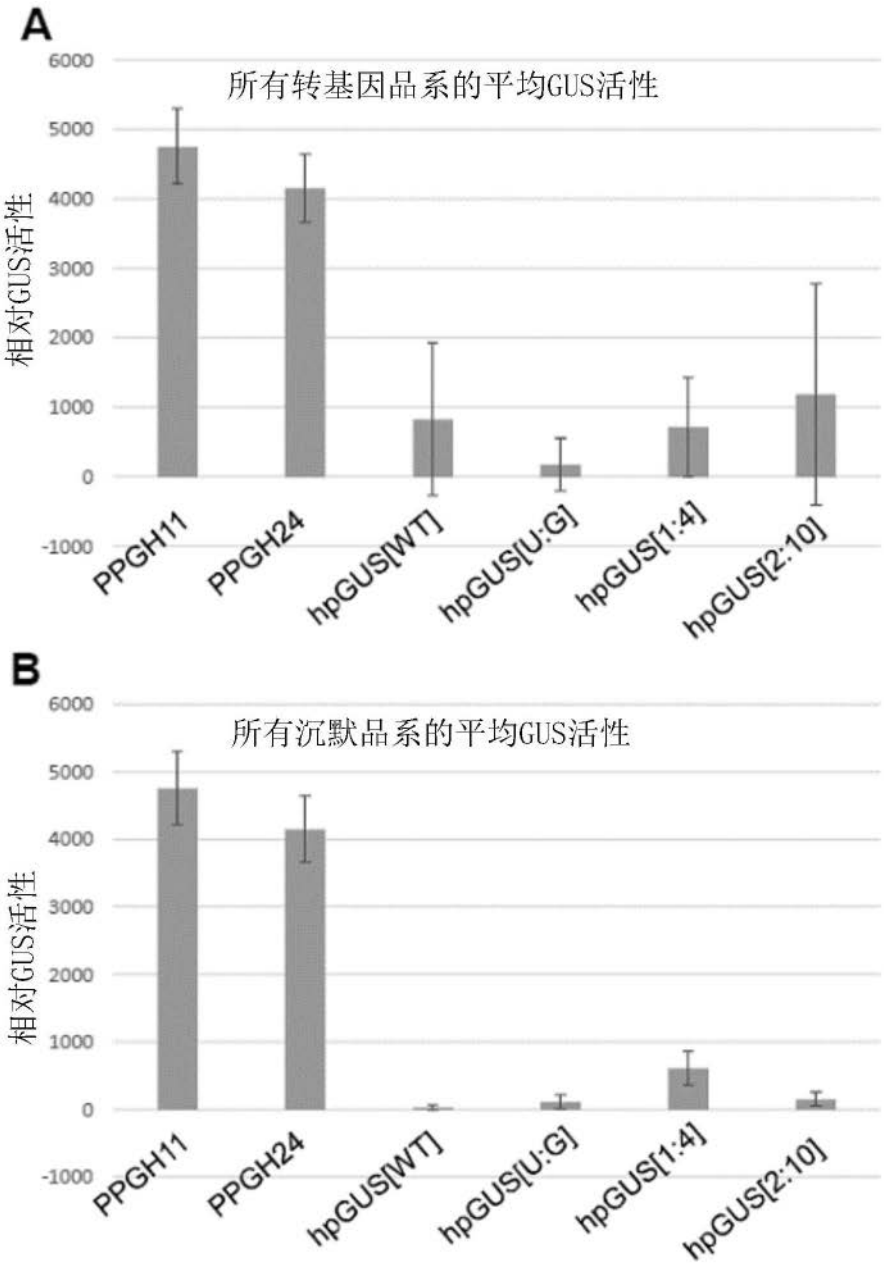

A
所有转基因品系的平均GUS活性
相对GUS活性
6000
5000
4000
3000
2000
1000
0
-1000
PPGH11
PPGH24
hpGUS[WT]
hpGUS[U:G]
hpGUS[1:4]
hpGUS[2:10]
B
所有沉默品系的平均GUS活性
相对GUS活性
6000
5000
4000
3000
2000
1000
0
-1000
PPGH11
PPGH24
hpGUS[WT]
hpGUS[U:G]
hpGUS[1:4]
hpGUS[2:10]

图13

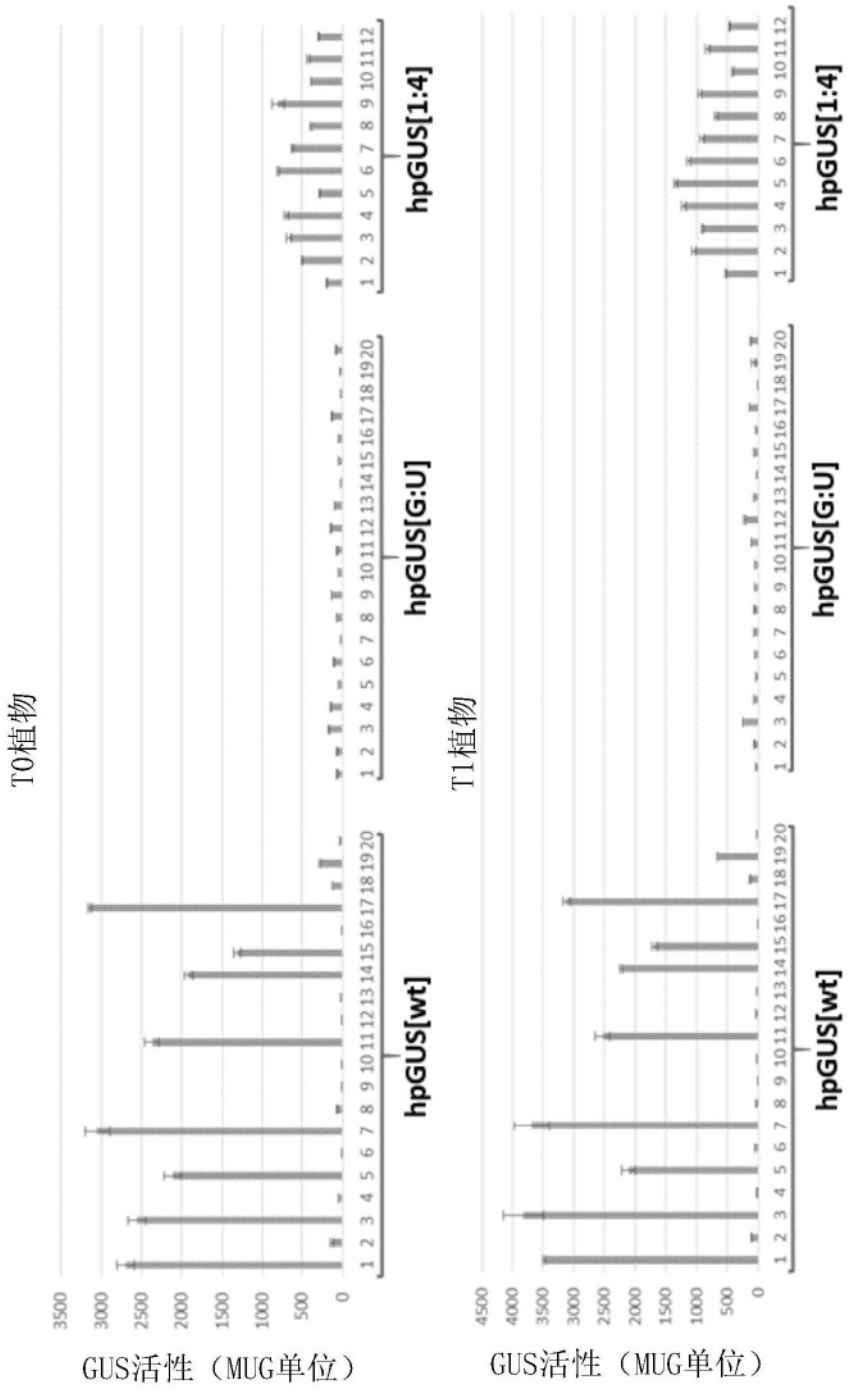


图14

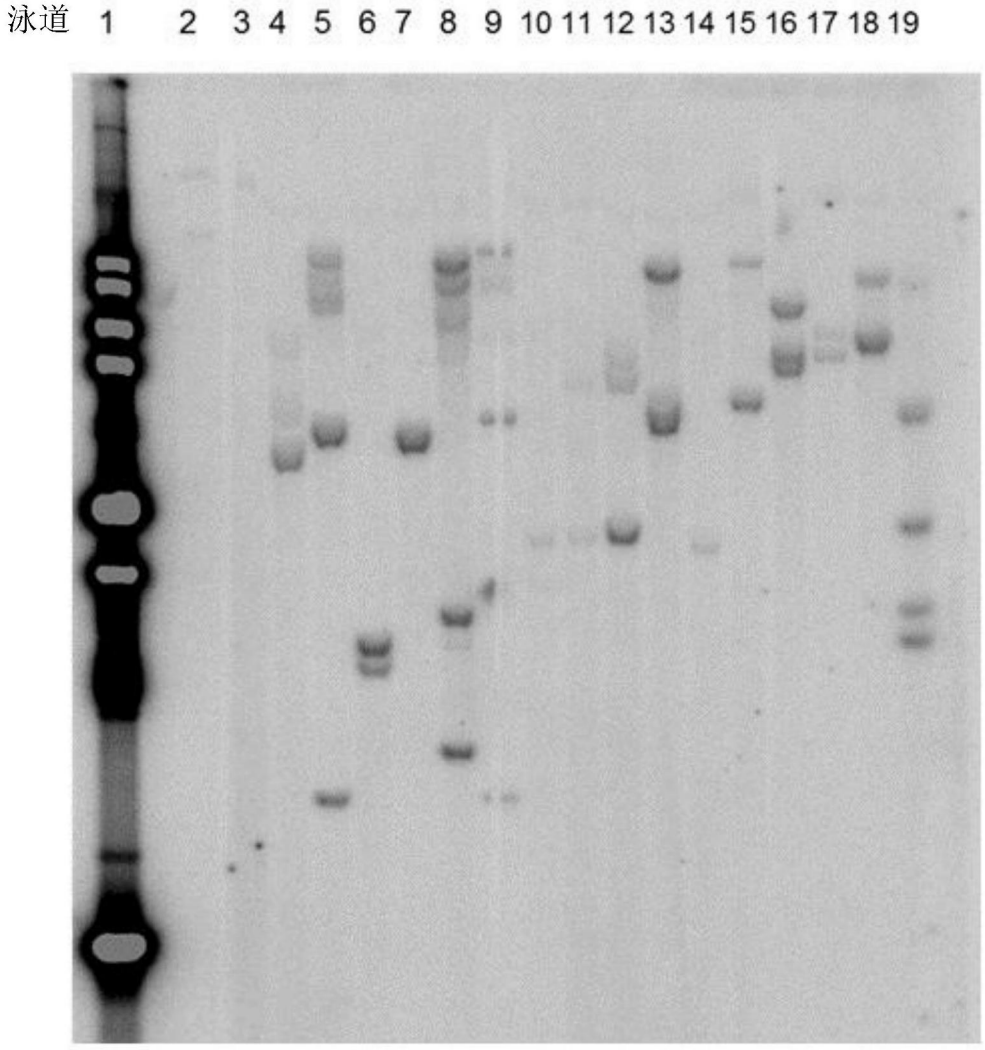


图15

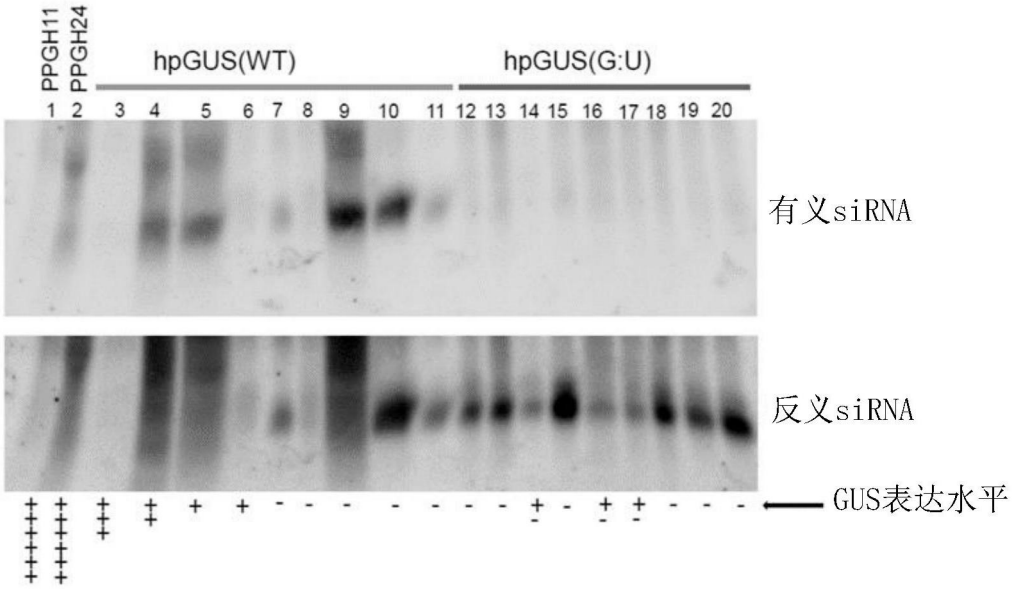

PPGH11
PPGH24
hpGUS(WT)
hpGUS(G:U)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
有义siRNA
反义siRNA
GUS表达水平

图16

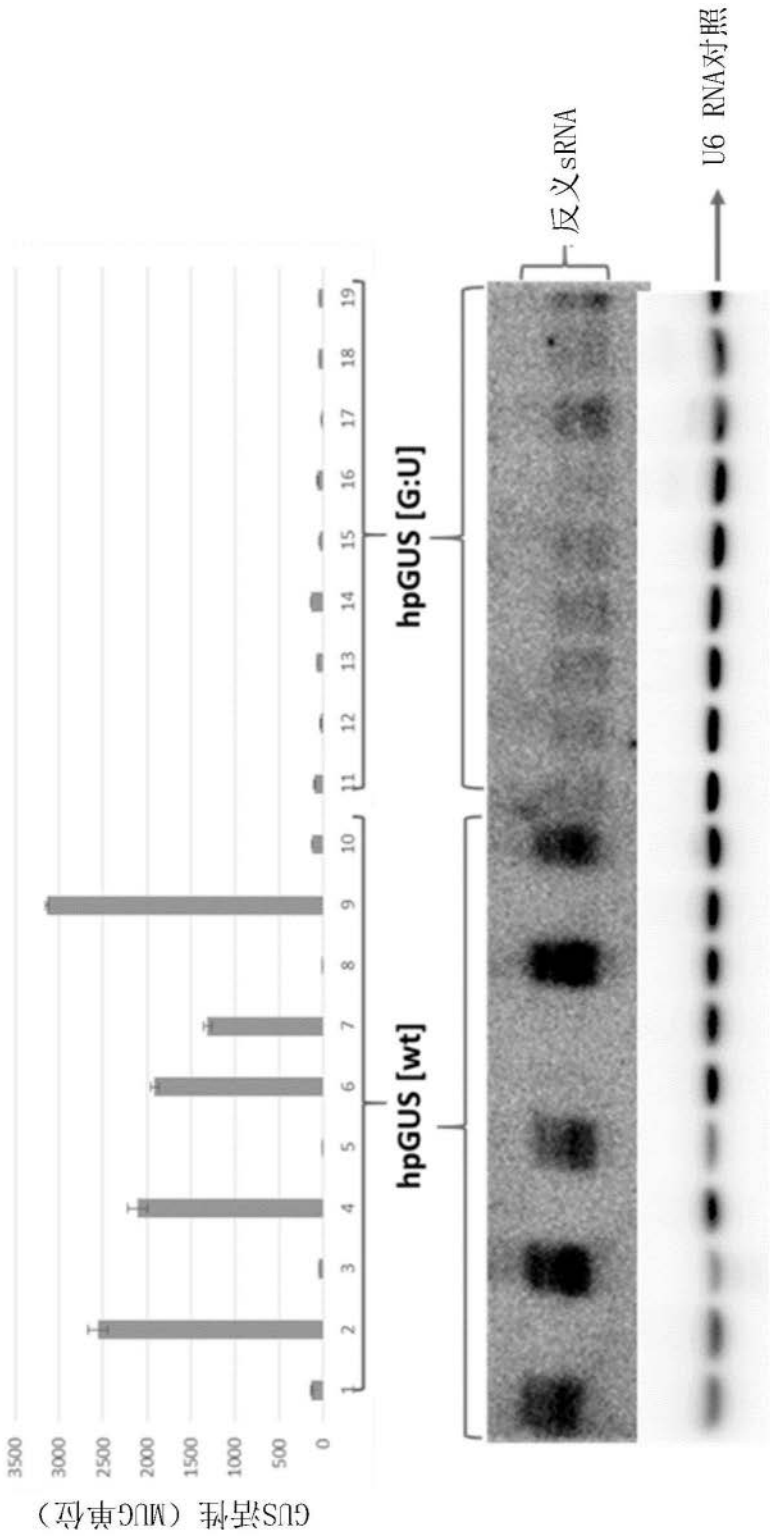

GUS活性（MUG单位）
3500
3000
2500
2000
1500
1000
500
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
hpGUS [wt]
hpGUS [G:U]
反义sRNA
U6 RNA对照

图17

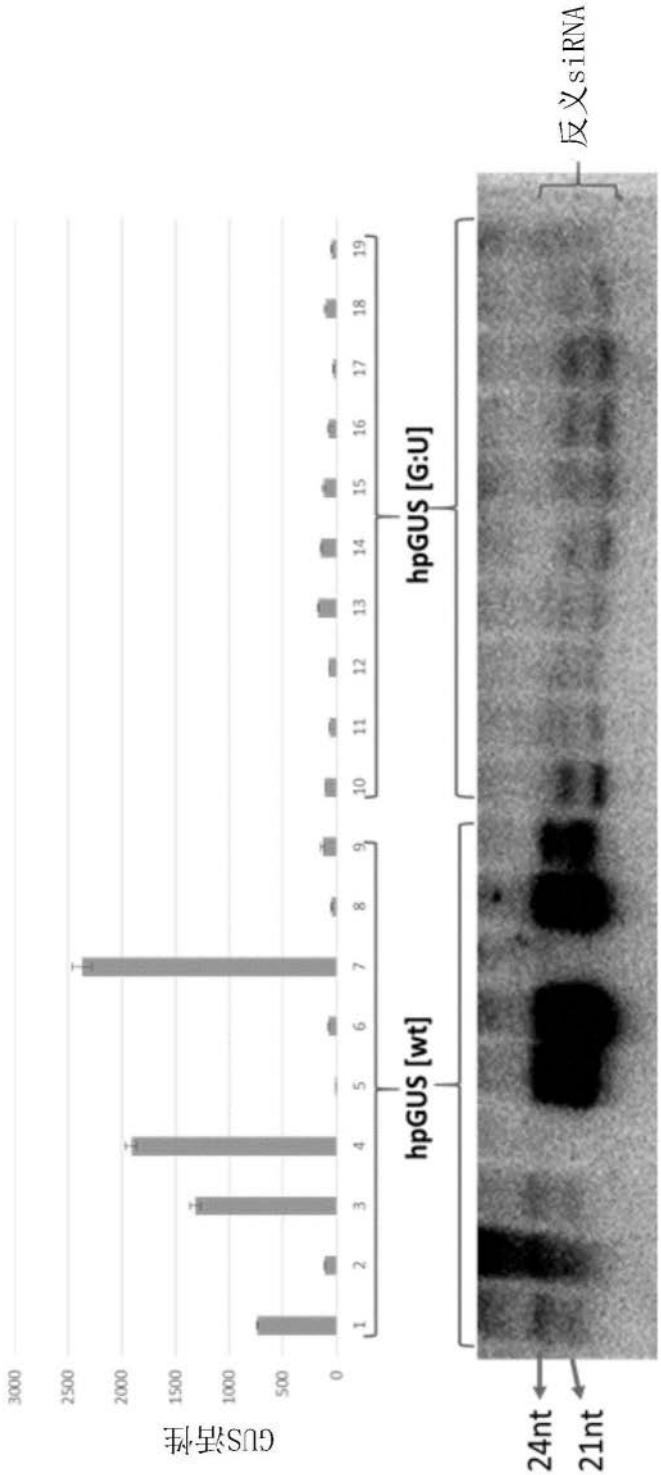

反义siRNA
hpGUS [G:U]
hpGUS [wt]
GUS活性
3000
2500
2000
1500
1000
500
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
24nt
21nt

图18

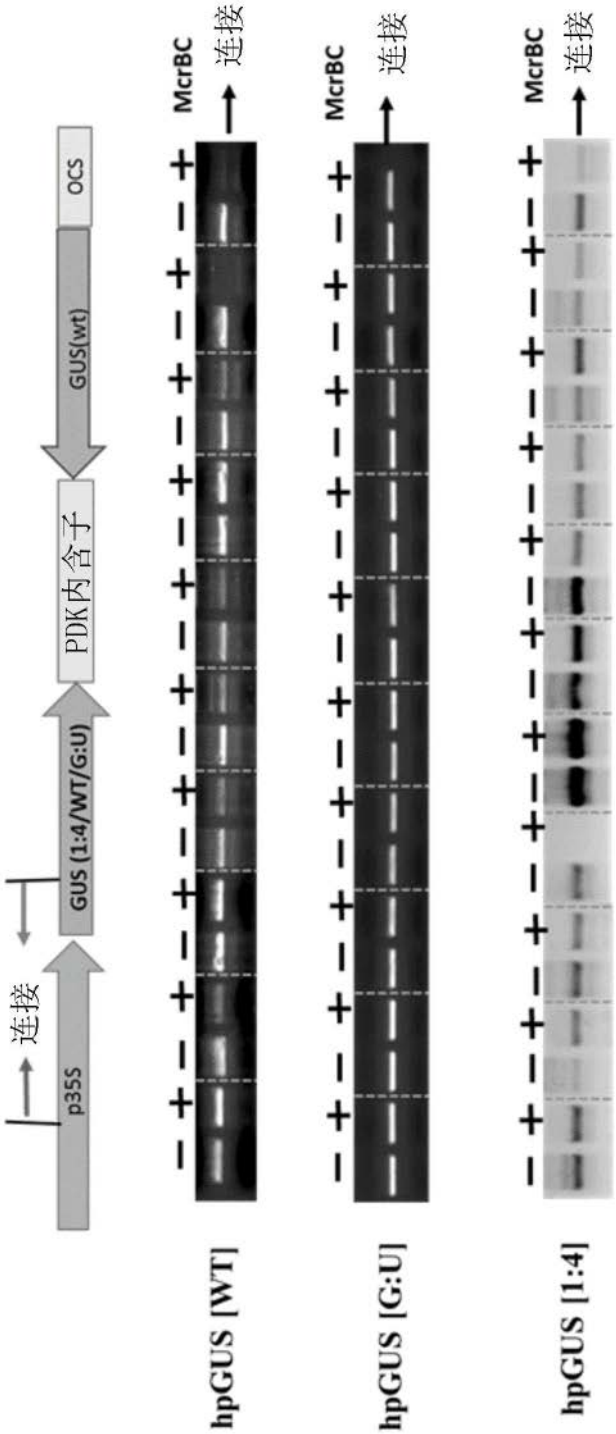

OCS
GUS(wt)
PDK内含子
GUS (1:4/WT/G:U)
连接
p35S
McrBC
连接
hpGUS [WT]
McrBC
连接
hpGUS [G:U]
McrBC
连接
hpGUS [1:4]

图19

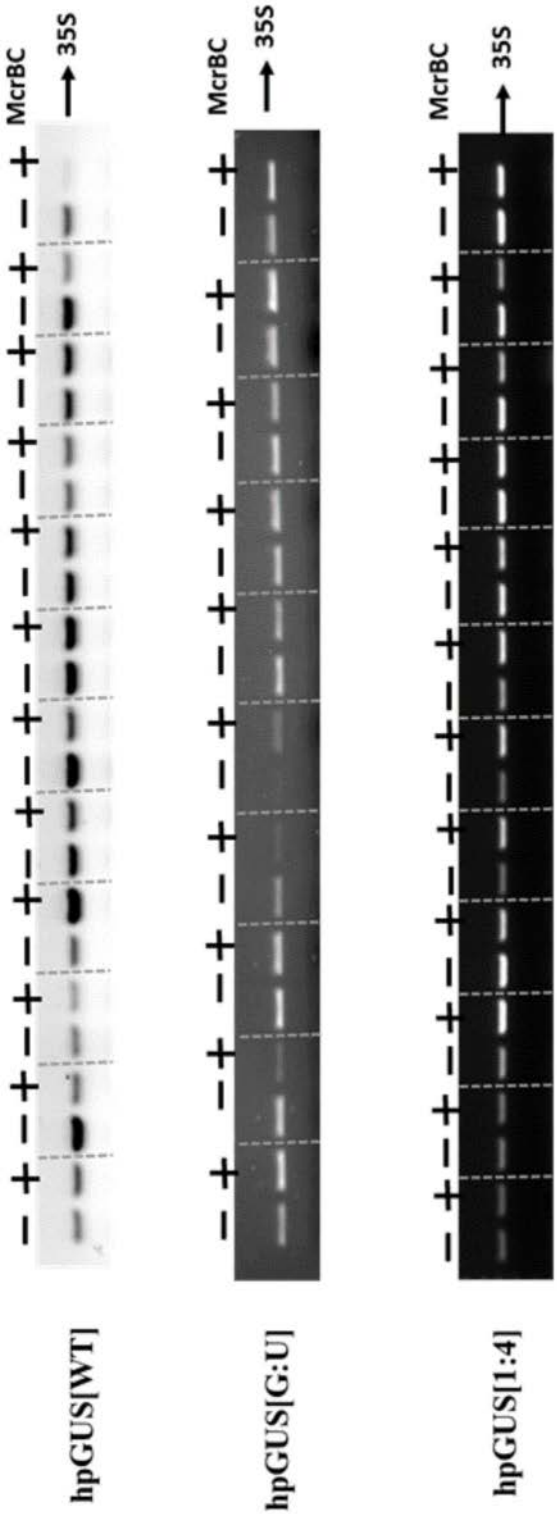


图20

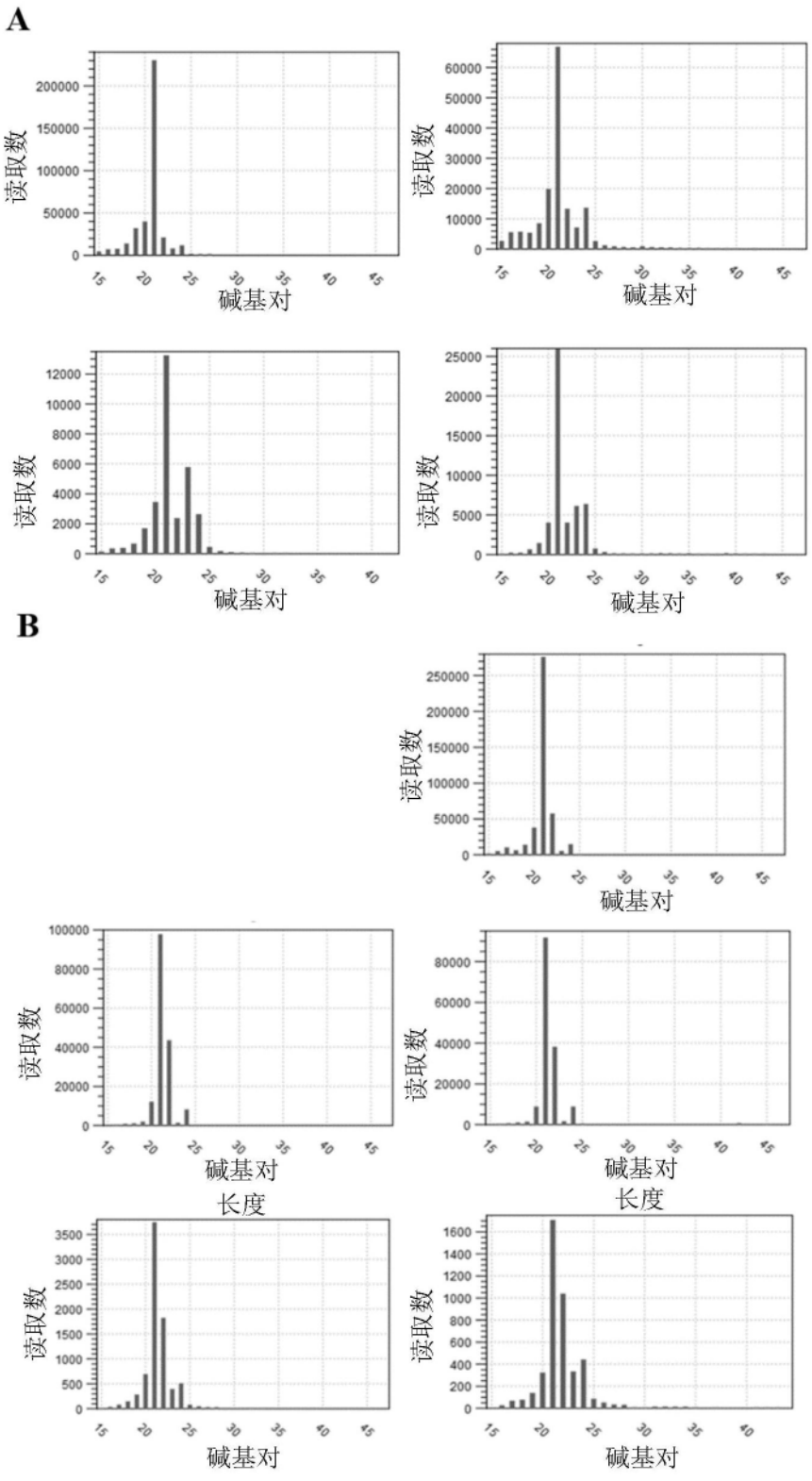


图21

```
EIN2-GU         ----------------------------------------------------GATTTTA
EIN2-cDNA       AGTAATTATATTATCAGATAGATTTAGGATGGAAGCTGAAATTGTGAATGTGAGACCTCA
                                                                   **  * *

EIN2-GU         GTTAGGGTTTATTTAGAGAATGGTTTTTGTTTTATTTTTTGTTTTTTTGGTTTTTGTTGG
EIN2-cDNA       GCTAGGGTTTATCCAGAGAATGGTTCCTGCTCTACTTCCTGTCCTTTTGGTTTCTGTCGG
                * **********  ***********  ** * ** **  ***  ********* *** **

EIN2-GU         ATATATTGATTTTGGGAAATGGGTTGTAAATATTGAAGGAGGTGTTTGTTTTGGGTATGA
EIN2-cDNA       ATATATTGATCCCGGGAAATGGGTTGCAAATATCGAAGGAGGTGCTCGTTTCGGGTATGA
                **********   ************* ****** ********** * **** ********

EIN2-GU         TTTGGTGGTAATTATTTTGTTTTTTAATTTTGTTGTTATTTTATGTTAATATGTTGTAGT
EIN2-cDNA       CTTGGTGGCAATTACTCTGCTTTTCAATTTTGCCGCCATCTTATGCCAATATGTTGCAGC
                 ******* ***** * ** **** *******  *  ** *****  ********* **

EIN2-GU         TTGTATAAGTGTT-----------------------------------------------
EIN2-cDNA       TCGCATAAGCGTTGTGACTGGTAAACACTTGGCTCAGATCTGCAATGAAGAATATGACAA
                * * ***** ***
```

图22

```
CHS-GU          ----------------------------------------------------GTTGTTTGTT
CHS-cDNA        TCCGTATCGCTAAGGATCTCGCCGAGAACAATCGTGGAGCACGTGTCCTCGTTGTCTGCT
                                                                   ***** ** *

CHS-GU          TTGAGATTATAGTTGTTATTTTTTGTGGTTTTTTTTGATATTTATTTTGATTTTTTTGTTG
CHS-cDNA        CTGAGATCACAGCCGTTACCTTCCGTGGTCCCTCTGACACCCACCTTGACTCCCTCGTCG
                 ****** * **  ****  **  *****   * *** *   *  **** *   * ** *

CHS-GU          GTTAGGTTTTTTTTAGTGATGGTGTTGTTGTATTTATTGTGGGGTTGGATTTTGATATAT
CHS-cDNA        GTCAGGCTCTTTTCAGTGATGGCGCCGCCGCACTCATTGTGGGGTCGGACCCTGACACAT
                ** *** * **** ******** *  *  * * * ********** ***   *** * **

CHS-GU          TTGTTGGAGAGAAATTTATTTTTGAGATGGTGTTTGTTGTTTAGATTATTTTTTTAGATT
CHS-cDNA        CTGTCGGAGAGAAACCCATCTTTGAGATGGTGTCTGCCGCTCAGACCATCCTTCCAGACT
                 *** *********   ** ************** **  * * ***  **  **  *** *

CHS-GU          TTGATGGTGT-------------------------------------------------
CHS-cDNA        CTGATGGTGCCATAGACGGACATTTGAGGGAAGTTGGTCTCACCTTCCATCTCCTCAAGG
                 ********
```

图23

反义EIN2-GU与反义EIN2 cDNA比对

```
As-EIN2-GU      ------------------------------------AATGTTTATGTGAGTTGTAATA
asEIN2-cDNA     TCTTCATTGCAGATCTGAGCCAAGTGTTTACCAGTCACAACGCTTATGCGAGCTGCAACA
                                                   ** * ***** *** ** ** *

As-EIN2-GU      TATTGGTATAAGATGGTGGTAAAATTGAAAAGTAGAGTAATTGTTATTAAGTTATATTTG
asEIN2-cDNA     TATTGGCATAAGATGGCGGCAAAATTGAAAAGCAGAGTAATTGCCACCAAGTCATACCCG
                ****** ********* ** ************ **********  *  **** ***   *

As-EIN2-GU      AAATGAGTATTTTTTTTGATATTTGTAATTTATTTTTTGGGATTAATATATTTGATAGAA
asEIN2-cDNA     AAACGAGCACCTCCTTCGATATTTGCAACCCATTTCCCGGGATCAATATATCCGACAGAA
                *** *** *  *  ** ******** **   ****   ***** *******  ** ****

As-EIN2-GU      ATTAAAAGGATAGGAAGTAGAGTAGGAATTATTTTTTGGATAAATTTTAGTTGAGGTT--
asEIN2-cDNA     ACCAAAAGGACAGGAAGTAGAGCAGGAACCATTCTCTGGATAAACCCTAGCTGAGGTCTC
                *  ******* *********** *****  *** * ********   *** ******
```

图24

反义CHS-GU与反义CHS cDNA比对

```
As-CHS-GU       ----------------------------------------GTATTATTAGAGTTTGGAA
asCHS-cDNA      GATGGAAGGTGAGACCAACTTCCCTCAAATGTCCGTCTATGGCACCATCAGAGTCTGGAA
                                                       * *  ** ***** *****

As-CHS-GU       GGATGGTTTGAGTGGTAGATATTATTTTAAAGATGGGTTTTTTTTTGATAGATGTGTTAG
asCHS-cDNA      GGATGGTCTGAGCGGCAGACACCATCTCAAAGATGGGTTTCTCTCCGACAGATGTGTCAG
                ******* **** ** *** *  ** * ************ * *  ** ******** **

As-CHS-GU       GGTTTGATTTTATAATGAGTGTGGTGGTGTTATTATTGAAAAGAGTTTGATTGATGAGGG
asCHS-cDNA      GGTCCGACCCCACAATGAGTGCGGCGGCGCCATCACTGAAAAGAGCCTGACCGACGAGGG
                ***  **    * ******** ** ** *  ** * *********  ***  ** *****

As-CHS-GU       AGTTAAGGTGGGTGTTAGAGGGATTATGGAAGGTAATGGTTGTGATTTTAGAGTAGATAA
asCHS-cDNA      AGTCAAGGTGGGTGTCAGAGGGACCACGGAAGGTAACGGCTGTGATCTCAGAGCAGACAA
                *** ************ *******  * ********** ** ****** * **** *** **
```

图25

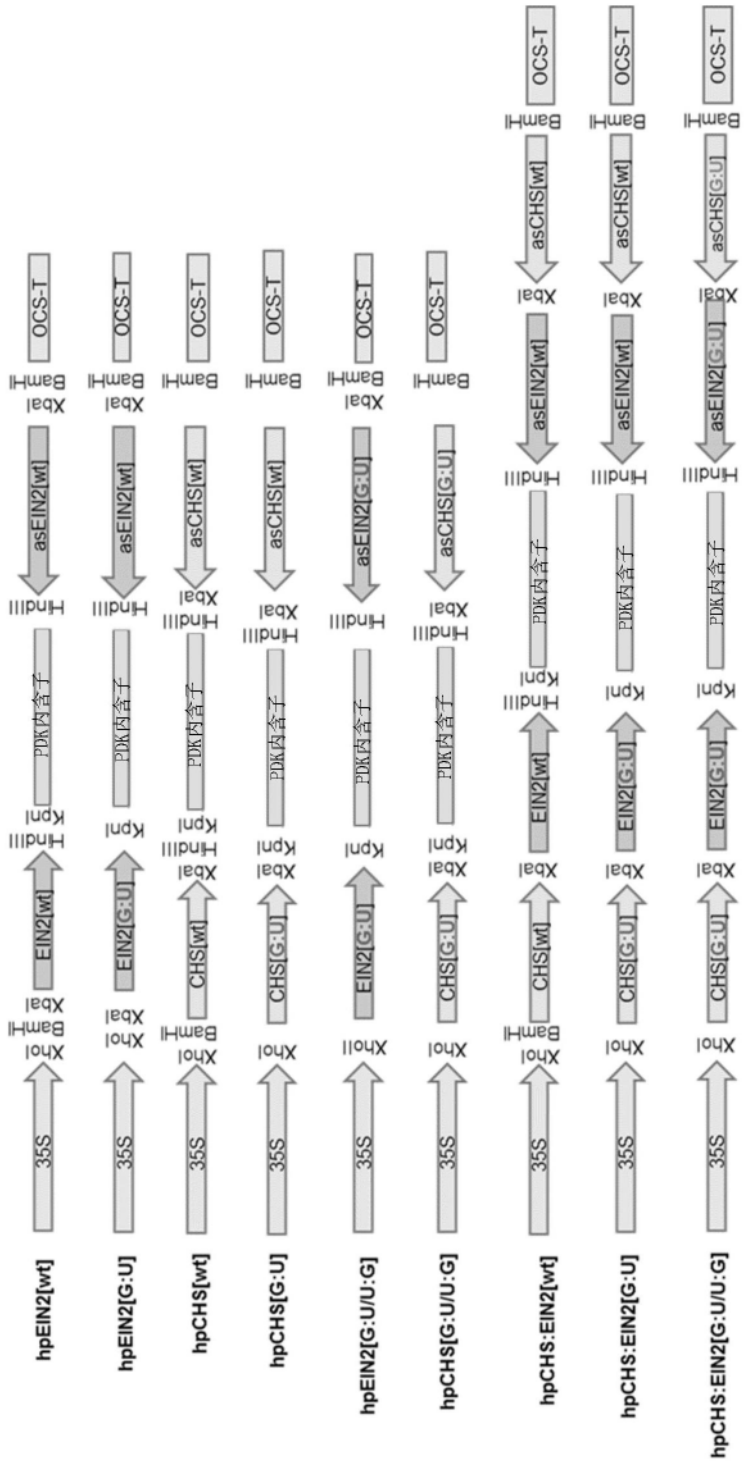

hpEIN2[wt]
35S
XhoI
EIN2[wt]
XbaI
BamHI
PDK内含子
HindIII
KpnI
asEIN2[wt]
XbaI
BamHI
OCS-T
hpEIN2[G:U]
35S
XhoI
EIN2[G:U]
XbaI
BamHI
PDK内含子
KpnI
HindIII
asEIN2[wt]
XbaI
BamHI
OCS-T
hpCHS[wt]
35S
XhoI
CHS[wt]
BamHI
PDK内含子
XbaI
HindIII
KpnI
asCHS[wt]
XbaI
BamHI
OCS-T
hpCHS[G:U]
35S
XhoI
CHS[G:U]
XbaI
PDK内含子
KpnI
HindIII
asCHS[wt]
XbaI
BamHI
OCS-T
hpEIN2[G:U/U:G]
35S
XhoII
EIN2[G:U]
XbaI
PDK内含子
KpnI
HindIII
asEIN2[G:U]
XbaI
BamHI
OCS-T
hpCHS[G:U/U:G]
35S
XhoI
CHS[G:U]
XbaI
PDK内含子
KpnI
HindIII
asCHS[G:U]
XbaI
BamHI
OCS-T
hpCHS:EIN2[wt]
35S
XhoI
CHS[wt]
BamHI
XbaI
EIN2[wt]
XbaI
PDK内含子
KpnI
HindIII
asEIN2[wt]
HindIII
BamHI
OCS-T
hpCHS:EIN2[G:U]
35S
XhoI
CHS[G:U]
XbaI
EIN2[G:U]
XbaI
PDK内含子
KpnI
HindIII
asEIN2[wt]
HindIII
BamHI
OCS-T
hpCHS:EIN2[G:U/U:G]
35S
XhoI
CHS[G:U]
XbaI
EIN2[G:U]
XbaI
PDK内含子
KpnI
HindIII
asEIN2[G:U]
HindIII
BamHI
OCS-T

图26

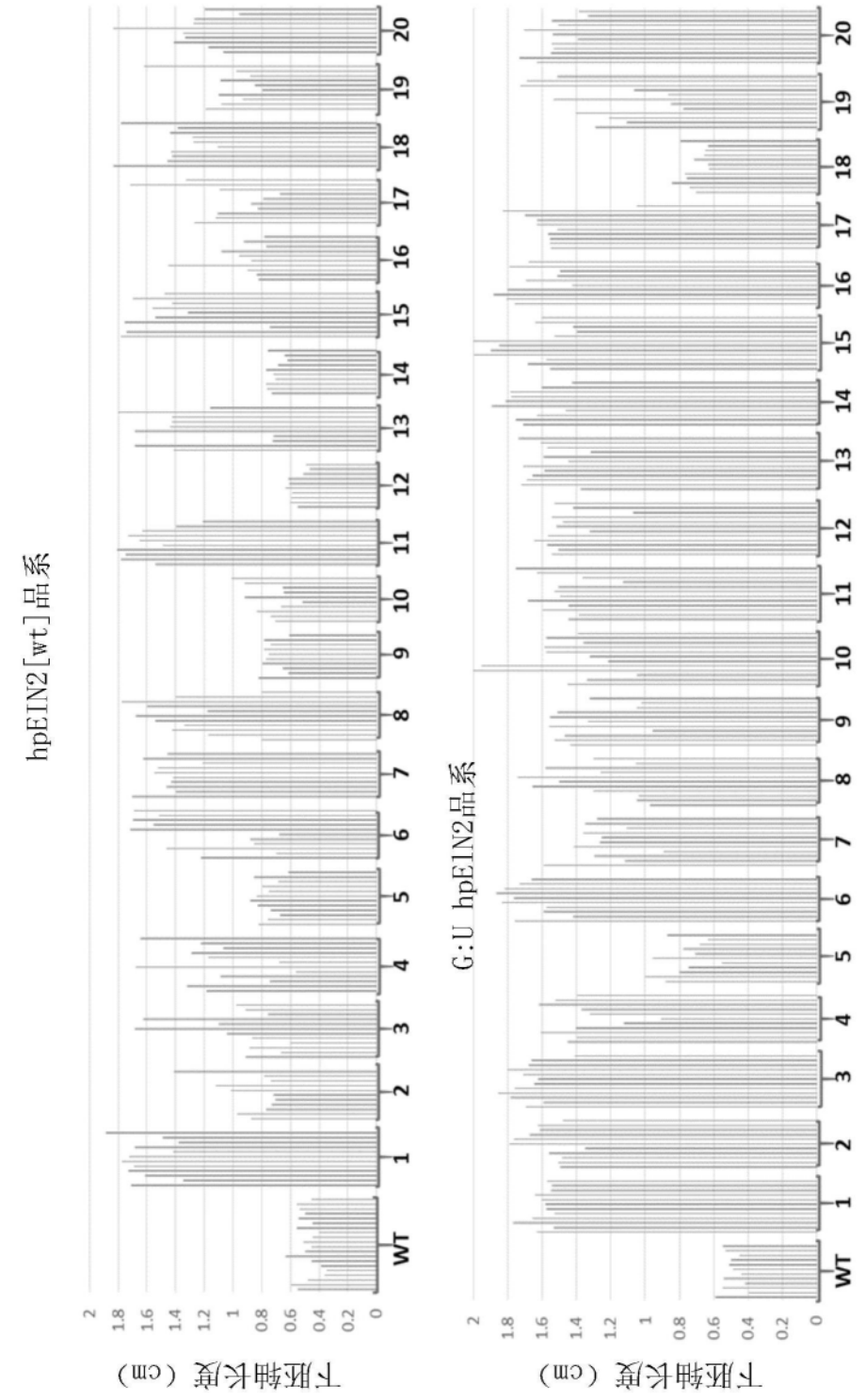

hpEIN2[wt]品系
下胚轴长度（cm）
WT
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
G:U hpEIN2品系
下胚轴长度（cm）

图27

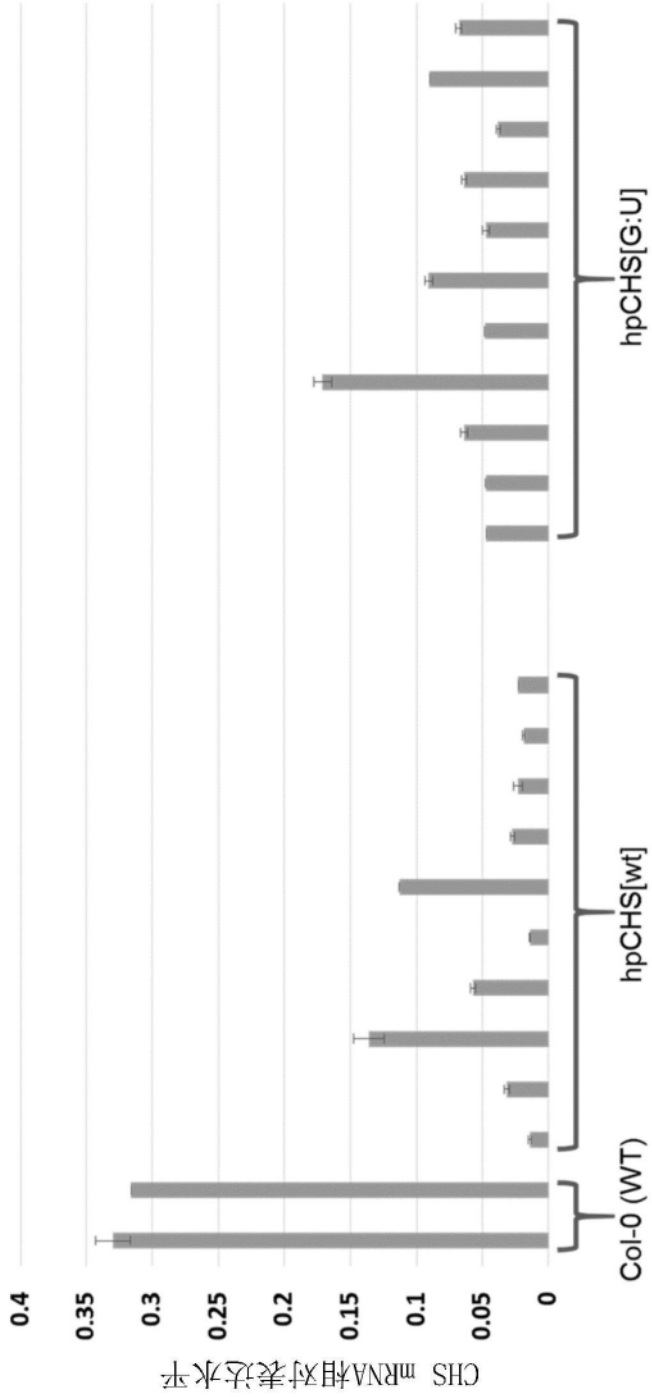

CHS mRNA相对表达水平
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Col-0 (WT)
hpCHS[wt]
hpCHS[G:U]

图28

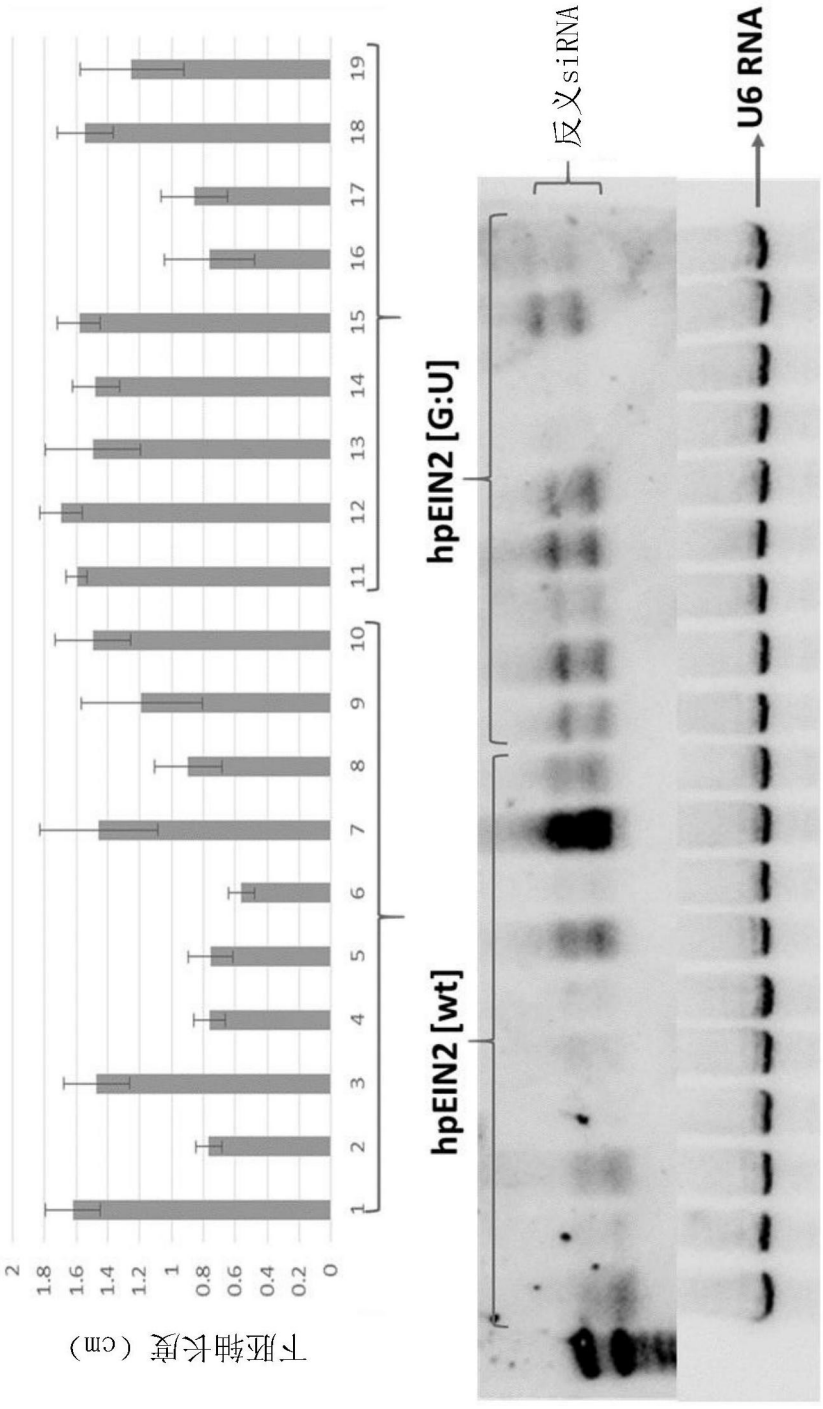

下胚轴长度（cm）
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
hpEIN2 [wt]
hpEIN2 [G:U]
反义siRNA
U6 RNA

图29

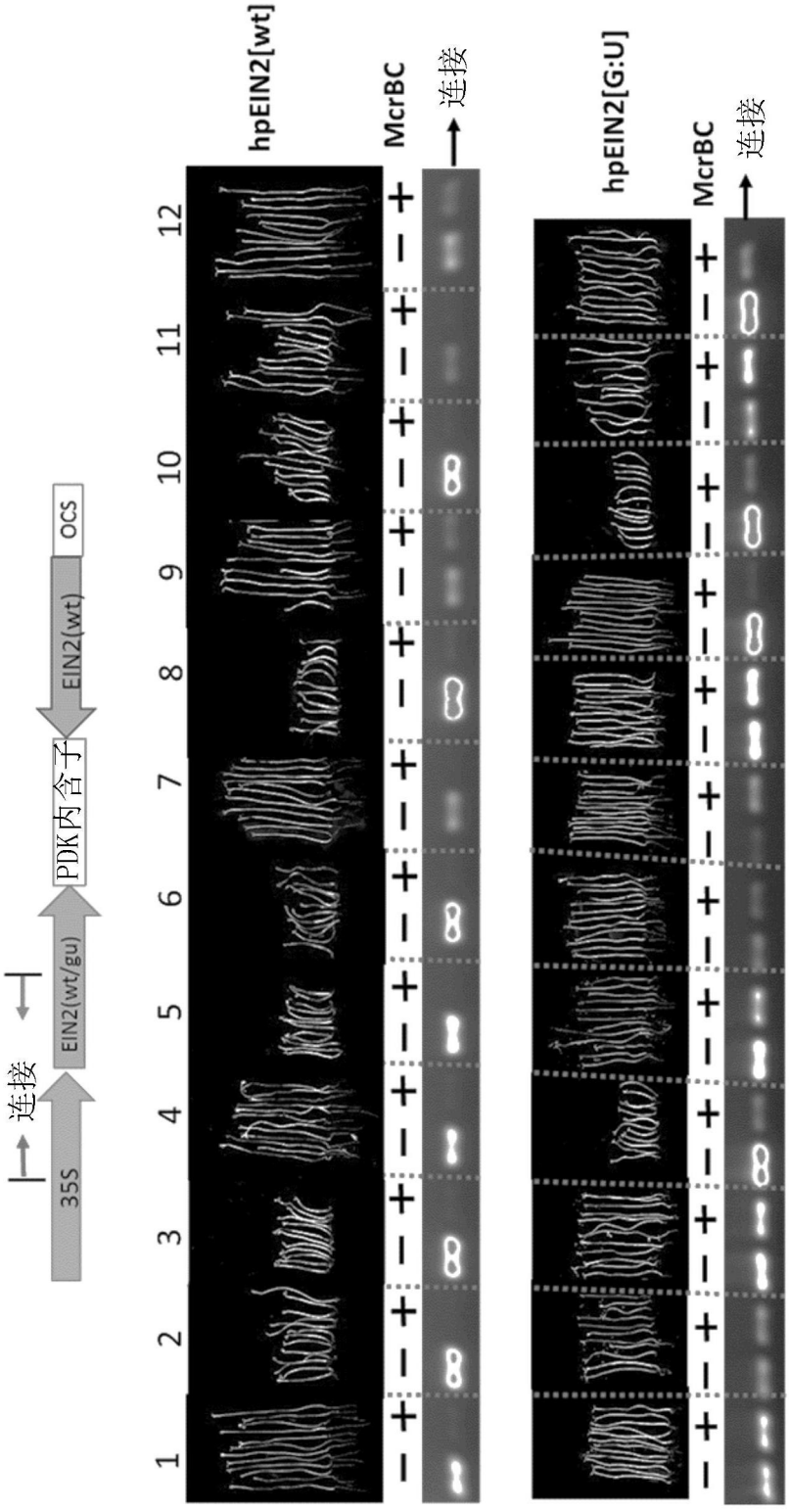

35S
连接
EIN2(wt/gu)
PDK内含子
EIN2(wt)
OCS
hpEIN2[wt]
McrBC
连接
hpEIN2[G:U]
McrBC
连接
1 2 3 4 5 6 7 8 9 10 11 12

图30

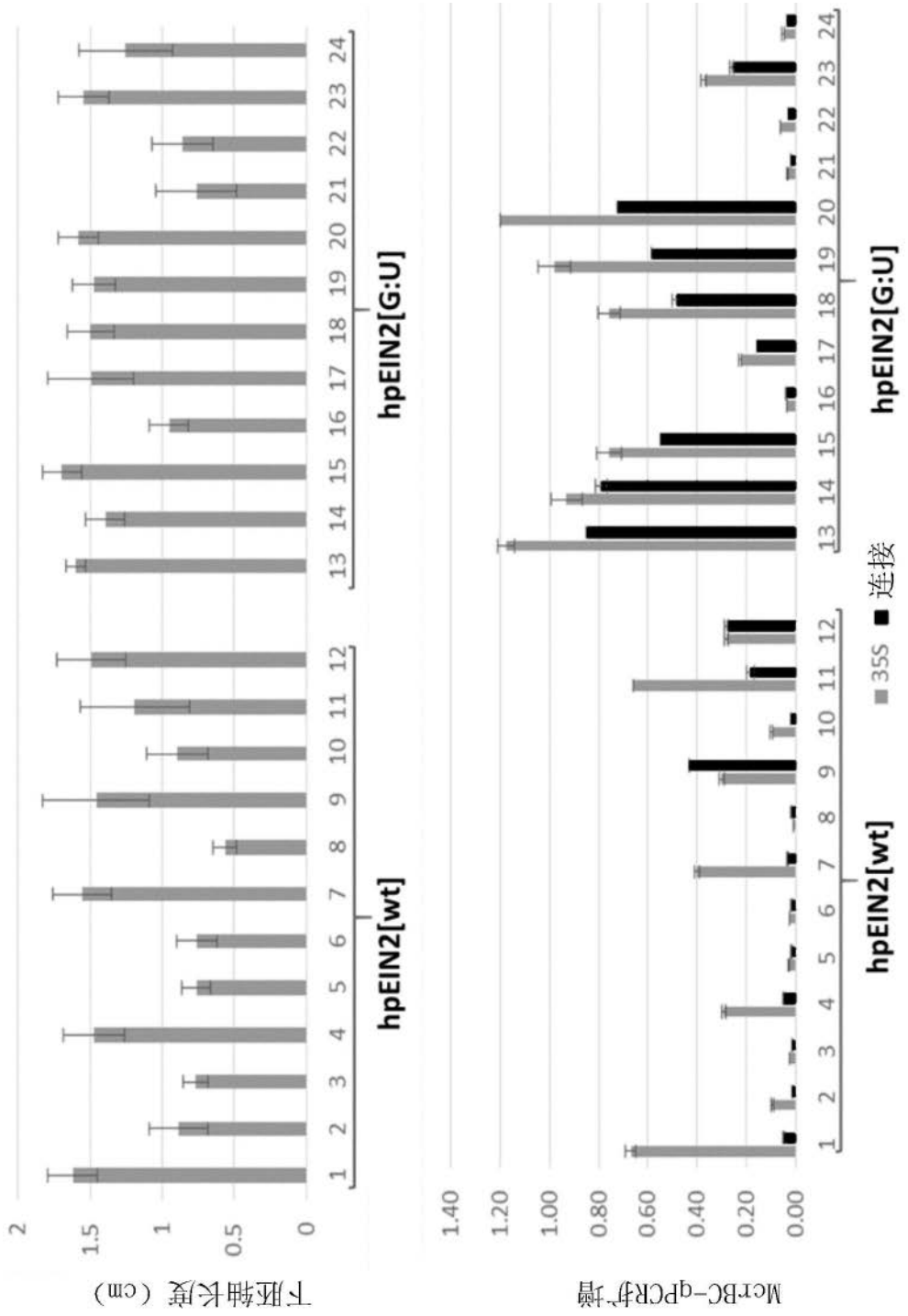

下胚轴长度（cm）
2
1.5
1
0.5
0
1 2 3 4 5 6 7 8 9 10 11 12
hpEIN2[wt]
13 14 15 16 17 18 19 20 21 22 23 24
hpEIN2[G:U]
McrBC-qPCR扩增
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
35S
连接
1 2 3 4 5 6 7 8 9 10 11 12
hpEIN2[wt]
13 14 15 16 17 18 19 20 21 22 23 24
hpEIN2[G:U]

图31

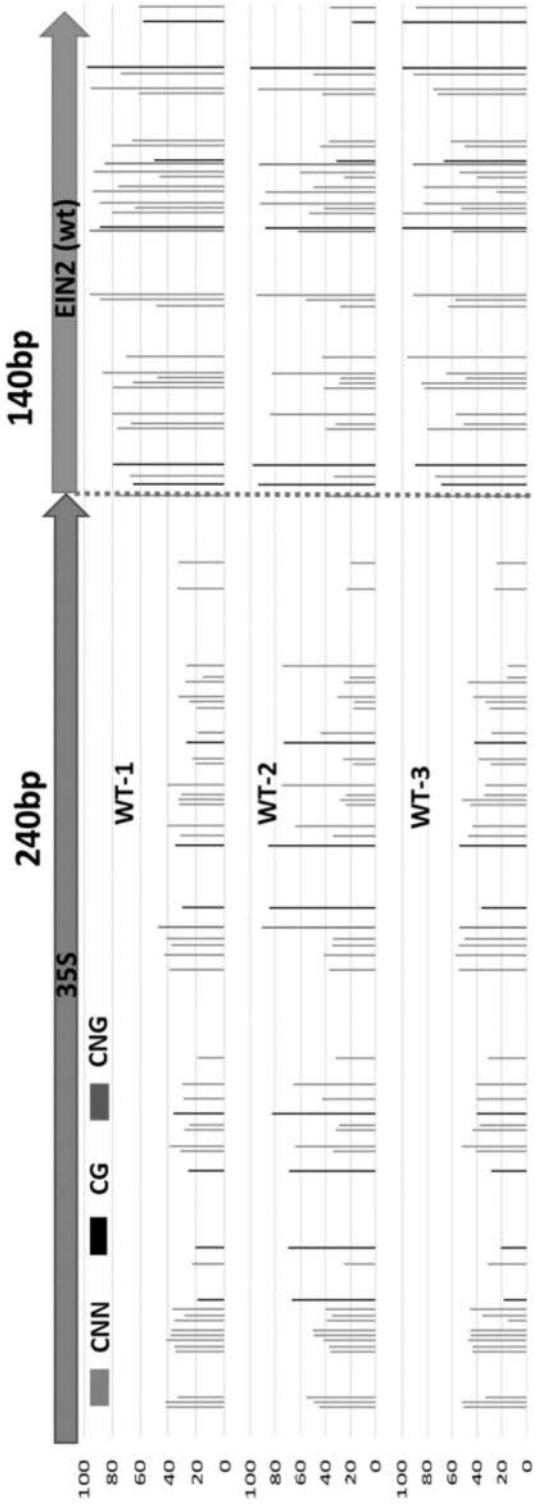


图32

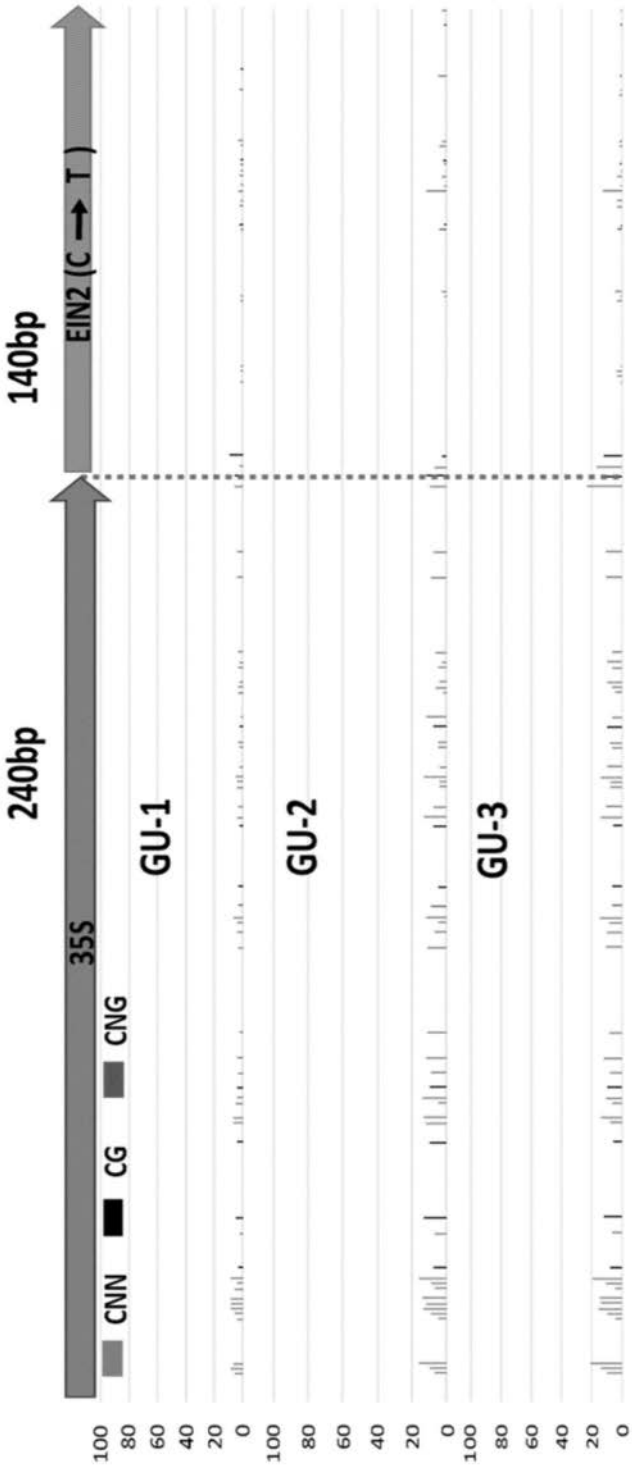


图33

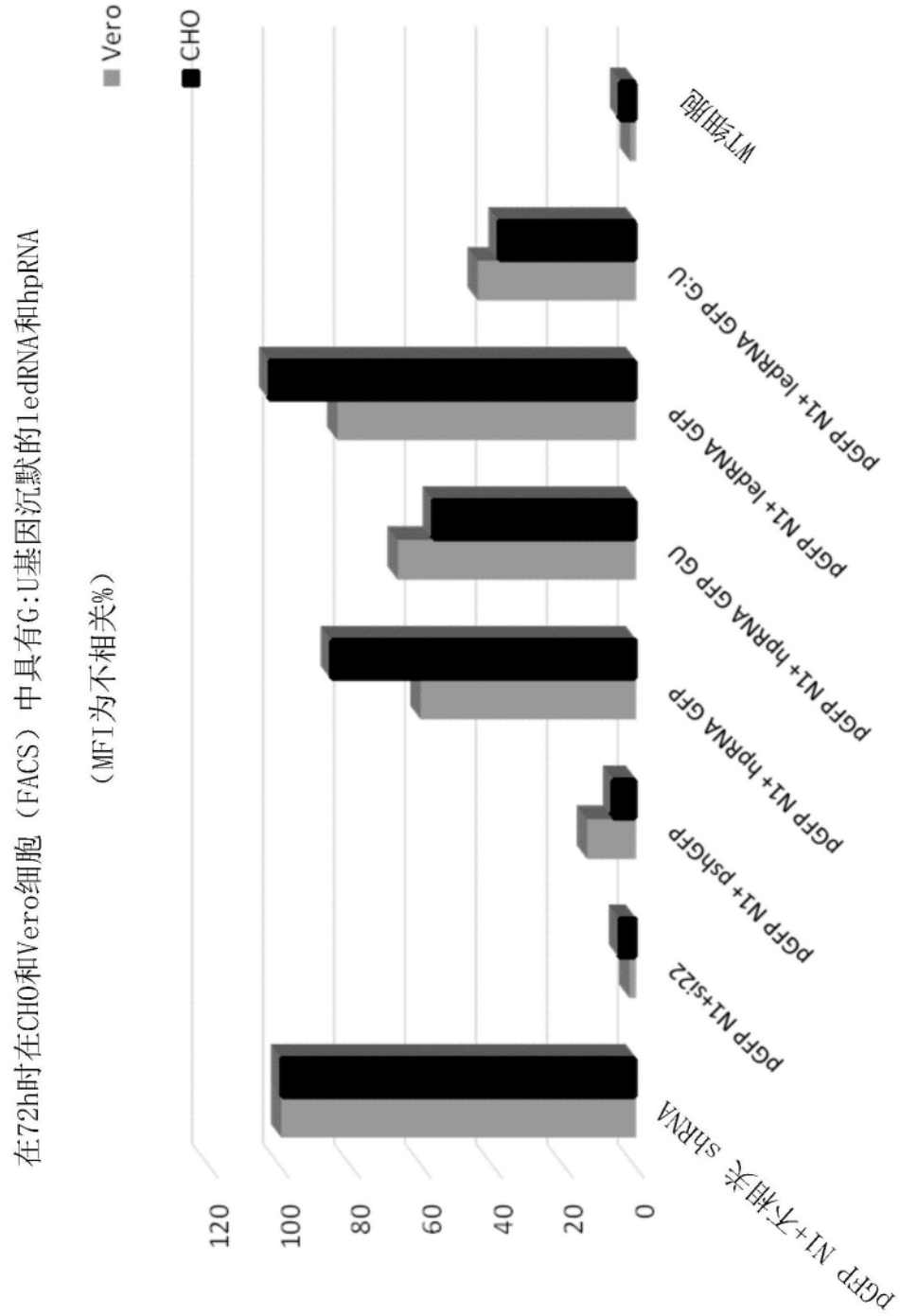

在72h时在CHO和Vero细胞（FACS）中具有G:U基因沉默的ledRNA和hpRNA
（MFI为不相关%）
Vero
CHO
120
100
80
60
40
20
0
pGFP N1+无相关 shRNA
pGFP N1+siL2
pGFP N1+pshGFP
pGFP N1+ hpRNA GFP
pGFP N1+ hpRNA GFP GU
pGFP N1+ ledRNA GFP
pGFP N1+ ledRNA GFP G:U
WT细胞

图34

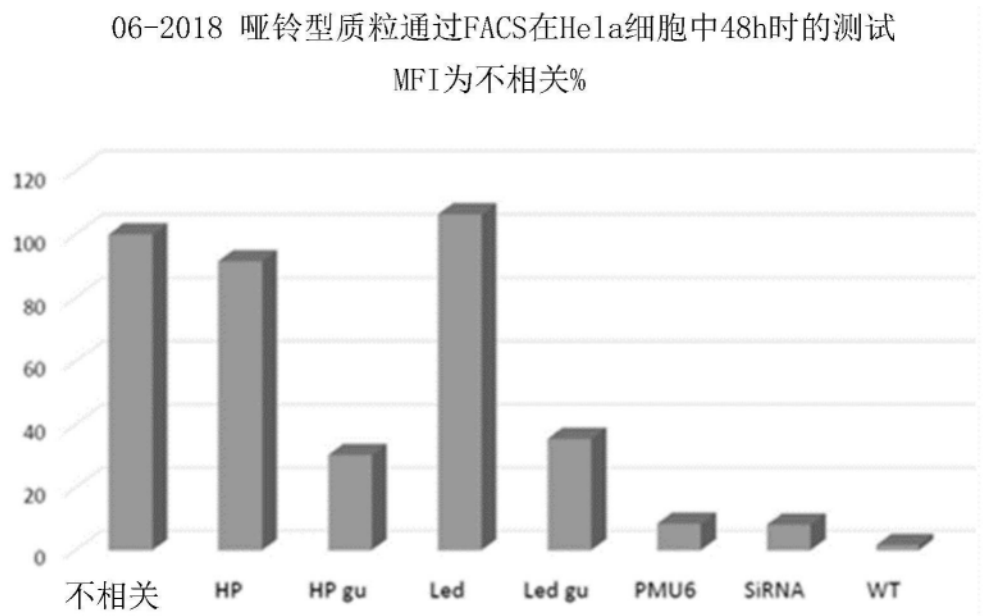

06-2018 哑铃型质粒通过FACS在Hela细胞中48h时的测试
MFI为不相关%
120
100
80
60
40
20
0
不相关
HP
HP gu
Led
Led gu
PMU6
SiRNA
WT

图35

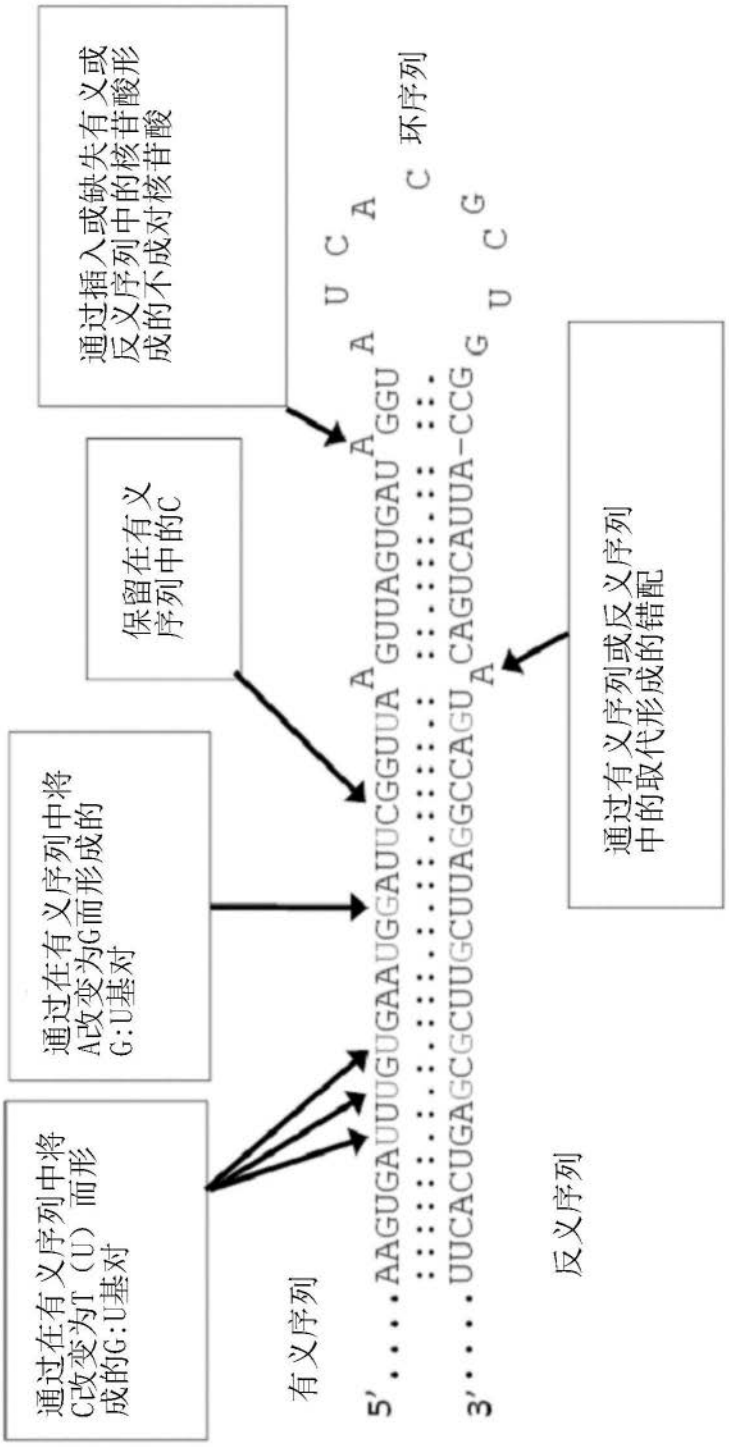

通过插入或缺失有义或反义序列中的核苷酸形成的不成对核苷酸
保留在有义序列中的C
通过在有义序列中将A改变为G而形成的G:U基对
通过在有义序列中将C改变为T（U）而形成的G:U基对
通过有义序列或反义序列中的取代形成的错配
环序列
有义序列
反义序列
5'
3'

图36

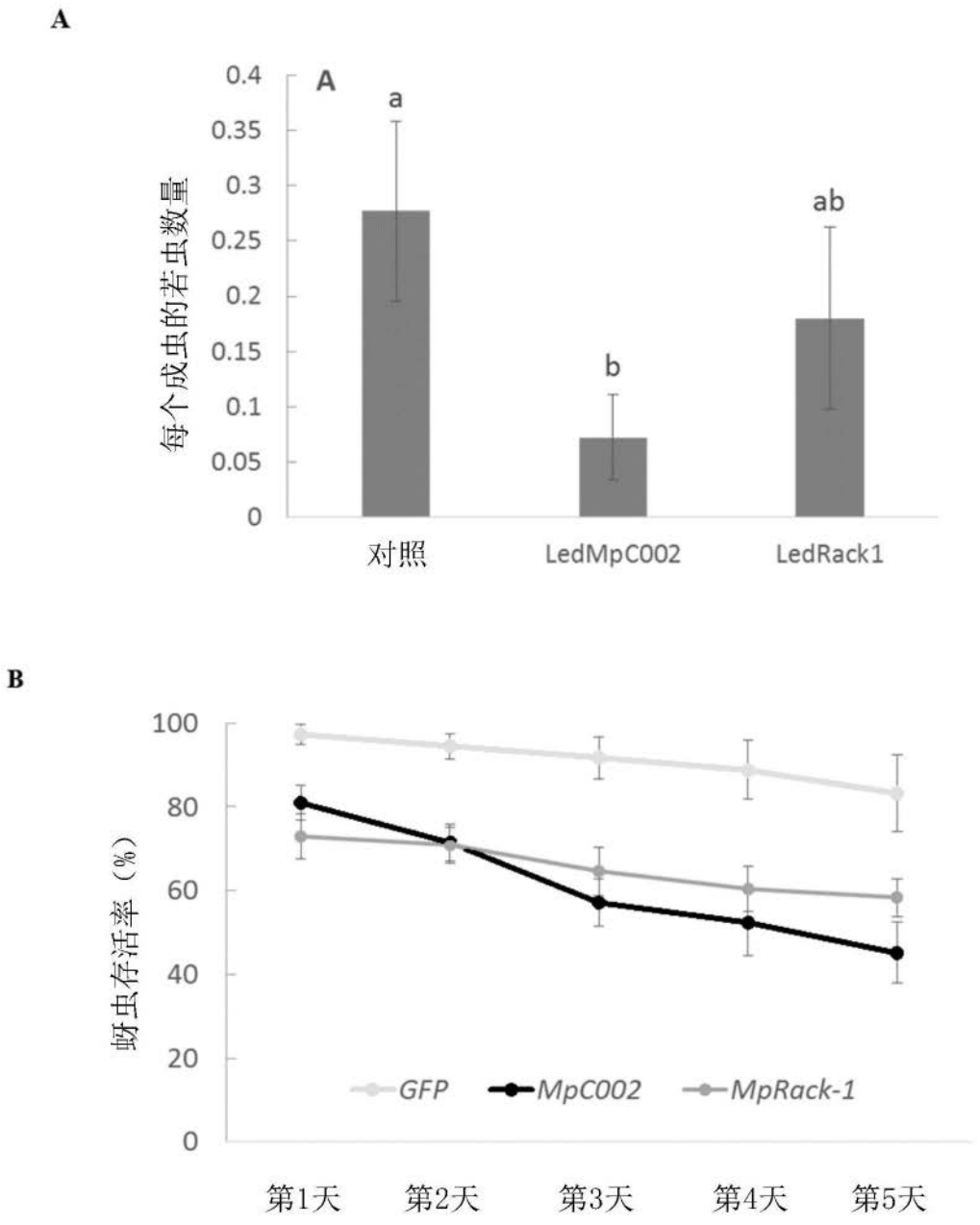

A
A
a
ab
b
每个成虫的若虫数量
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
对照
LedMpC002
LedRack1
B
蚜虫存活率（%）
100
80
60
40
20
0
GFP
MpC002
MpRack-1
第1天
第2天
第3天
第4天
第5天

图37