



(12) 发明专利申请

(10) 申请公布号 CN 112888785 A

(43) 申请公布日 2021. 06. 01

(21) 申请号 201980065254.2

(22) 申请日 2019.08.02

(30) 优先权数据

2018902840 2018.08.03 AU

2018902896 2018.08.08 AU

PCT/AU2018/051015 2018.09.17 AU

2019900941 2019.03.20 AU

(85) PCT国际申请进入国家阶段日

2021.04.01

(86) PCT国际申请的申请数据

PCT/AU2019/050814 2019.08.02

(87) PCT国际申请的公布数据

W02020/024019 EN 2020.02.06

(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 D·张

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 R·德菲特

(74) 专利代理机构 永新专利商标代理有限公司

72002

代理人 区斌

(51) Int.Cl.

C12N 15/113 (2006.01)

(54) 发明名称

包含非规范碱基对的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。

权利要求书8页 说明书109页

序列表93页 附图38页

1. 一种嵌合的核糖核酸 (RNA) 分子, 其包含双链RNA (dsRNA) 区, 所述双链RNA (dsRNA) 区包含长度为至少20个连续核苷酸的第一有义核糖核苷酸序列和长度为至少20个连续核苷酸的第一反义核糖核苷酸序列, 由此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能够彼此杂交以形成所述dsRNA区, 其中

i) 所述第一有义核糖核苷酸序列由按5'至3'顺序共价连接的第一5'核糖核苷酸、第一RNA序列和第一3'核糖核苷酸组成,

ii) 所述第一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,

iii) 所述第一5'核糖核苷酸与所述第二3'核糖核苷酸碱基配对以形成所述dsRNA区的末端碱基对,

iv) 所述第二5'核糖核苷酸与所述第一3'核糖核苷酸碱基配对以形成所述dsRNA区的末端碱基对,

v) 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约5%至约40%的核糖核苷酸以非规范碱基对碱基配对或不碱基配对,

vi) 所述dsRNA区不包含20个连续的规范碱基对,

vii) 所述RNA分子能够在真核细胞中或体外被加工, 由此切割所述第一反义核糖核苷酸序列, 以产生长度为20-24个核糖核苷酸的短反义RNA (asRNA) 分子,

viii) 所述RNA分子或至少一些所述asRNA分子或两者能够降低真核细胞中靶RNA分子的表达或活性, 以及

ix) 所述RNA分子能够通过在外或在细胞中或在两者中转录而酶促制备。

2. 权利要求1的嵌合RNA分子, 其中所述第一有义核糖核苷酸序列通过第一连接核糖核苷酸序列与所述第一反义核糖核苷酸序列共价连接, 所述第一连接核糖核苷酸序列包含长度为至少4个核苷酸、或4至1000个核糖核苷酸、或4至200个核糖核苷酸、或4至50个核糖核苷酸、或至少10个核苷酸、或10至1000个核糖核苷酸、或10至200个核糖核苷酸、或10至50个核糖核苷酸的环序列, 由此所述第一连接核糖核苷酸序列与所述第二3'核糖核苷酸和所述第一5'核糖核苷酸共价连接, 或优选地与所述第一3'核糖核苷酸和所述第二5'核糖核苷酸共价连接, 使得所述序列包含在RNA的单一连续链中。

3. 权利要求2的嵌合RNA分子, 其中所述RNA分子中的所述环序列包含一个或多个与对所述真核细胞而言是内源的RNA分子互补的结合序列, 和/或所述RNA分子中的所述环序列包含编码多肽或功能性多核苷酸的开放阅读框。

4. 根据权利要求1-3任一项的嵌合RNA分子, 其中所述dsRNA的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约5%至约40%的核糖核苷酸以非规范碱基对碱基配对, 优选地以G:U碱基对碱基配对。

5. 根据权利要求1-4任一项的嵌合RNA分子, 所述第一反义核糖核苷酸序列与所述靶RNA的区域完全互补, 且所述第一有义核糖核苷酸序列与所述靶RNA的区域在序列上的不同之处在于所述靶RNA的区域中的C核苷酸被U核苷酸取代。

6. 根据权利要求1-5任一项的嵌合RNA分子, 其包含第二有义核糖核苷酸序列, 并且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通过第一连接核糖核苷酸序列连接, 所述第一连接核糖核苷酸序列包括长度为至少4个核苷酸的环序列, 由此所述第一

连接核糖核苷酸序列与所述第一3'核糖核苷酸和所述第二5'核糖核苷酸共价连接,并且所述RNA分子进一步包含第二连接核糖核苷酸序列,所述第二连接核糖核苷酸序列包括长度为至少4个核苷酸的环序列,并且其与所述第二3'核糖核苷酸和所述第二有义核糖核苷酸序列共价连接。

7. 根据权利要求1-5任一项的嵌合RNA分子,其包含第二反义核糖核苷酸序列,并且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通过第一连接核糖核苷酸序列连接,所述第一连接核糖核苷酸序列包括长度为至少4个核苷酸的环序列,由此所述第一连接核糖核苷酸序列与所述第二3'核糖核苷酸和所述第一5'核糖核苷酸共价连接,并且所述RNA分子进一步包含第二连接核糖核苷酸序列,所述第二连接核糖核苷酸序列包括长度为至少4个核苷酸的环序列,并且其与所述第二3'核糖核苷酸和所述第二反义核糖核苷酸序列共价连接。

8. 根据权利要求1-5任一项的嵌合RNA分子,其包含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能够互相杂交以形成第二dsRNA区,并且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通过第一连接核糖核苷酸序列连接,所述第一连接核糖核苷酸序列包括长度为至少4个核苷酸的环序列,由此所述第一连接核糖核苷酸序列与所述第一3'核糖核苷酸和所述第二5'核糖核苷酸共价连接,并且所述RNA分子任选地包含第二连接核糖核苷酸序列,所述第二连接核糖核苷酸序列包括长度为至少4个核苷酸的环序列并且其与所述第二3'核糖核苷酸和所述第二有义核糖核苷酸序列共价连接,或其共价连接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。

9. 根据权利要求1-5任一项的嵌合RNA分子,其包含第二有义核糖核苷酸序列和第二反义核糖核苷酸序列,并且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通过第一连接核糖核苷酸序列连接,所述第一连接核糖核苷酸序列包括长度为至少4个核苷酸的环序列,由此所述第一连接核糖核苷酸序列与所述第二3'核糖核苷酸和所述第一5'核糖核苷酸共价连接,并且所述RNA分子进一步包含第二连接核糖核苷酸序列,所述第二连接核糖核苷酸序列包括长度为至少4个核苷酸的环序列并且其与所述第一3'核糖核苷酸和所述第二反义核糖核苷酸序列共价连接,或其共价连接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。

10. 根据权利要求6-9任一项的嵌合RNA分子,其中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各自包含长度为至少20个的连续核苷酸。

11. 根据权利要求6-10任一项的嵌合RNA分子,其中所述第一和第二有义核糖核苷酸序列通过间插核糖核苷酸序列共价连接,所述间插核糖核苷酸序列与所述靶RNA分子在序列上无关或与所述靶RNA分子在序列上相关,或所述第一和第二有义核糖核苷酸序列共价连接而没有间插核糖核苷酸序列。

12. 根据权利要求6-11任一项的嵌合RNA分子,其中所述第一和第二反义核糖核苷酸序列通过间插核糖核苷酸序列共价连接,所述间插核糖核苷酸序列与所述靶RNA分子的互补体在序列上无关或与所述靶RNA分子的互补体在序列上相关,或所述第一和第二反义核糖核苷酸序列共价连接而没有间插核糖核苷酸序列。

13. 根据权利要求6-12任一项的嵌合RNA分子,其中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总体上5%至40%的核糖核苷酸以非规范碱基对碱基配对或不碱基配对,优选地以G:U碱基对碱基配对,其中所述第二dsRNA区不包含20个连续的规范碱基对,并且其中所述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。

14. 根据权利要求6-13任一项的嵌合RNA分子,其中每个连接核糖核苷酸序列的长度独立地在4至约2000个核苷酸之间,优选地每个连接核糖核苷酸序列的长度独立地在4至约1200个核苷酸之间,更优选地每个连接核糖核苷酸序列的长度独立地在4至约200个核苷酸之间,并且最优选地每个连接核糖核苷酸序列的长度独立地在4至约50个核苷酸之间。

15. 根据权利要求6-14任一项的嵌合RNA分子,其还包含5'前导序列或3'尾随序列,或两者。

16. 一种嵌合RNA分子,其包含第一RNA组分和与所述第一RNA组分共价连接的第二RNA组分,

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所述第一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,

其中所述第二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,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,

其中在所述第一RNA组分中,

i) 所述第一有义核糖核苷酸序列由至少20个连续核糖核苷酸组成,其为按5'至3'顺序共价连接的第一5'核糖核苷酸、第一RNA序列和第一3'核糖核苷酸,

ii) 所述第一反义核糖核苷酸序列由至少20个连续核糖核苷酸组成,其为按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,

iii) 所述第一5'核糖核苷酸与所述第二3'核糖核苷酸碱基配对,

iv) 所述第二5'核糖核苷酸与所述第一3'核糖核苷酸碱基配对,

v) 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碱基配对或者不碱基配对,以及

vi) 所述第一dsRNA区不包含20个连续的规范碱基对,

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

(a) 所述嵌合RNA分子或至少一些所述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

(b) 所述第一反义核糖核苷酸序列包含至少20个连续核糖核苷酸的序列,所述至少20个连续核糖核苷酸的序列在序列上与所述靶RNA分子的互补体的区域具有至少50%相同性,优选在序列上与所述靶RNA分子的互补体的区域具有至少90%或100%相同性,或

(c) (a) 和 (b) 二者。

17. 根据权利要求1-16任一项的嵌合RNA分子,其中所述第一反义核糖核苷酸序列的至

少20个连续核糖核苷酸都能够与所述靶RNA分子的第一区域的核苷酸碱基配对。

18. 根据权利要求1-17任一项的嵌合RNA分子,其中所述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分子的区域互补,优选地为完全互补。

19. 权利要求18的嵌合RNA分子,其中所述两个或更多个反义核糖核苷酸序列与相同靶RNA分子的不同区域互补。

20. 权利要求18的嵌合RNA分子,其中所述两个或更多个反义核糖核苷酸序列与不同靶RNA分子的区域互补。

21. 根据权利要求1-20任一项的嵌合RNA分子,其包含发夹RNA (hpRNA) 结构,所述发夹RNA (hpRNA) 结构具有5'端、长度为至少21个核苷酸的有义核糖核苷酸序列、在至少21个连续核苷酸上与所述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、间插环序列和3'端。

22. 根据权利要求1-20任一项的嵌合RNA分子,其包含核糖核苷酸的单链,所述核糖核苷酸的单链具有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个连续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。

23. 根据权利要求1-22任一项的嵌合RNA分子,其中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15%至约30%,或约16%至约25%的核糖核苷酸以非规范碱基对碱基配对或者不碱基配对,优选地以非规范碱基对碱基配对,更优选地以G:U碱基对碱基配对。

24. 根据权利要求1-23任一项的嵌合RNA分子,其中至少50%、至少60%、至少70%、至少80%、至少90%、至少95%、至少97%或100%的所述非规范碱基对为G:U碱基对。

25. 根据权利要求1-24任一项的嵌合RNA分子,其中在所述dsRNA区中少于25%、少于20%、少于15%、少于10%、少于5%、少于1%或没有核苷酸是不碱基配对的。

26. 根据权利要求1-25任一项的嵌合RNA分子,其中在所述dsRNA区中的4个核糖核苷酸中的每一个至6个核糖核苷酸中的每一个形成非规范碱基对或不碱基配对,优选地形成G:U碱基对。

27. 根据权利要求1-26任一项的嵌合RNA分子,其中所述dsRNA区不包含8个连续的规范碱基对。

28. 根据权利要求1-27任一项的嵌合RNA分子,其中所述dsRNA区包含至少8个连续的规范碱基对,优选地为至少8个但不超过12个连续的规范碱基对。

29. 根据权利要求1-28任一项的嵌合RNA分子,其中所述dsRNA区或每个dsRNA区中的所有核糖核苷酸以规范碱基对或非规范碱基对碱基配对。

30. 根据权利要求1-28任一项的嵌合RNA分子,其中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是不碱基配对的。

31. 根据权利要求1-30任一项的嵌合RNA分子,其中所述反义RNA序列在序列上与所述靶RNA分子的区域互补体具有小于100%的相同性,或具有约80%至99.9%的相同性,或具有约90%至98%的相同性,或具有约95%至98%的相同性。

32. 根据权利要求1-30任一项的嵌合RNA分子,其中所述反义RNA序列在序列上与所述靶RNA分子的区域具有100%相同性。

33. 根据权利要求1-32任一项的嵌合RNA分子,其中所述有义和/或反义核糖核苷酸序列,优选地为两者,长度为至少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1000,或约100至约1000,或20至约1000个核苷酸,或20至约500个核苷酸。

34. 根据权利要求1-33任一项的嵌合RNA分子,其中所述有义核糖核苷酸序列中的核糖核苷酸的数目为所述反义核糖核苷酸序列中的核糖核苷酸的数目的约90%至约110%。

35. 根据权利要求1-34任一项的嵌合RNA分子,其中所述有义核糖核苷酸序列中的核糖核苷酸的数目与所述反义核糖核苷酸序列中的核糖核苷酸的数目相同。

36. 根据权利要求1-35任一项的嵌合RNA分子,其中所述嵌合RNA分子进一步包含与所述第一5'核糖核苷酸共价连接的5'延伸序列或与所述第二3'核糖核苷酸共价连接的3'延伸序列,或两者。

37. 根据权利要求1-36任一项的嵌合RNA分子,其中所述嵌合RNA分子进一步包含与所述第二5'核糖核苷酸共价连接的5'延伸序列或与所述第一3'核糖核苷酸共价连接的3'延伸序列,或两者。

38. 根据权利要求1-37任一项的嵌合RNA分子,其包含两个或更多个相同或不同的dsRNA区。

39. 根据权利要求1-38任一项的嵌合RNA分子,其中当在真核细胞中表达时,当与具有以规范碱基对完全碱基配对的相应dsRNA区的类似RNA分子的加工相比时,形成更多长度为22个和/或20个核糖核苷酸的asRNA分子。

40. 一种分离的和/或外源的多核苷酸,其编码根据权利要求1至39任一项的嵌合RNA分子。

41. 权利要求40所述的多核苷酸,其是DNA构建体。

42. 权利要求40或41所述的多核苷酸,其与能够在宿主细胞中或在体外引导所述RNA分子表达的启动子可操作地连接。

43. 权利要求42所述的多核苷酸,其中所述启动子是RNA聚合酶启动子,诸如RNA聚合酶III启动子、RNA聚合酶II启动子,或在体外起作用的启动子。

44. 权利要求40-43任一项的多核苷酸,其编码RNA前体分子,所述RNA前体分子包含内含子,优选地在5'延伸序列或在至少一个环序列中包含内含子,其中所述内含子能够在所述多核苷酸在宿主细胞中或在体外转录期间被剪接出来。

45. 一种包含根据权利要求40-44任一项的多核苷酸的载体。

46. 权利要求45所述的载体,其是病毒载体。

47. 一种宿主细胞,其包含以下的一种或多种或全部:权利要求1至39任一项的嵌合RNA分子、通过加工所述嵌合RNA分子产生的小RNA分子、权利要求40-44任一项的多核苷酸,或权利要求45或权利要求46的载体。

48. 权利要求47的宿主细胞,其为细菌细胞、真菌细胞例如酵母细胞、植物细胞或动物细胞,优选地为植物细胞。

49. 权利要求47或权利要求48的宿主细胞,其是死亡的和/或不能繁殖的。

50. 权利要求42或权利要求43的多核苷酸或根据权利要求47-49任一项的宿主细胞,其编码和/或包含根据权利要求1-39任一项所述的嵌合RNA分子,其中当与编码具有以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的启动子相比时,所述多核苷酸的启动子区具有较低水平的甲基化,例如小于约50%、小于约40%、小于约30%或小于约20%。

51. 根据权利要求50的宿主细胞,优选地为植物细胞或真菌细胞,其包含所述嵌合RNA分子或通过加工所述嵌合RNA分子产生的小RNA分子或二者,其中所述嵌合RNA分子按5'至3'顺序包含第一有义核糖核苷酸序列,包括环序列的第一连接核糖核苷酸序列和第一反义核糖核苷酸序列。

52. 根据权利要求47-51任一项的宿主细胞,其是真核细胞且包含编码根据权利要求1至39中任一项所述的嵌合RNA分子的多核苷酸的至少两个拷贝,并且其中

i) 所述真核细胞中所述靶RNA分子的表达或活性的降低水平与如果细胞具有所述多核苷酸的单个拷贝,所述靶RNA分子的表达或活性的降低水平大致相同或更高,和/或

ii) 当与包含具有以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,所述真核细胞中所述靶RNA分子的表达或活性降低的水平更低。

53. 一种非人生物体,其包含以下的一种或多种或全部:根据权利要求1-39任一项的嵌合RNA分子、通过加工所述嵌合RNA分子产生的小RNA分子、根据权利要求40-44或50任一项的多核苷酸,或权利要求45或权利要求46的载体,或根据权利要求47-52任一项的宿主细胞。

54. 权利要求53所述的非人生物体,其是包含根据权利要求40-44或50任一项的多核苷酸的转基因非人生物体,优选地为植物或真菌。

55. 权利要求54所述的非人生物体,其中所述多核苷酸稳定地整合到所述生物体的基因组中。

56. 一种产生权利要求1至39任一项的嵌合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根据权利要求40-44或50任一项的多核苷酸。

57. 权利要求56所述的方法,其还包括至少部分纯化所述RNA分子。

58. 一种产生权利要求54或权利要求55所述的非人生物体的方法,所述方法包括将根据权利要求40-44或50任一项的多核苷酸导入细胞以使其稳定整合到所述细胞的基因组中,并从所述细胞生成所述非人生物体。

59. 根据权利要求47至52任一项的宿主细胞的提取物,其中所述提取物包含以下的一种或多种:根据权利要求1-39任一项的嵌合RNA分子、通过加工所述嵌合RNA分子产生的小RNA分子、或根据权利要求40-44或50任一项的多核苷酸。

60. 一种组合物,所述组合物包含以下的一种或多种以及一种或多种合适的载体:根据权利要求1-39任一项的嵌合RNA分子、通过加工所述嵌合RNA分子产生的小RNA分子、根据权利要求40-44或50任一项的多核苷酸、权利要求45或权利要求46的载体、根据权利要求47-52任一项的宿主细胞,或权利要求59的提取物。

61. 权利要求60所述的组合物,其为药物组合物。

62. 权利要求60所述的组合物,其适于应用于在田地生长的植物。

63. 根据权利要求60-62任一项的组合物,其进一步包含至少一种化合物,所述至少一

种化合物增强以下的一种或多种的稳定性:所述嵌合RNA分子、通过加工所述嵌合RNA分子产生的RNA分子和多核苷酸,和/或所述至少一种化合物帮助所述RNA分子、嵌合RNA分子或多核苷酸被生物体的细胞摄取。

64. 权利要求63所述的组合物,其中所述化合物是转染促进剂。

65. 一种减少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以下中的一种或多种:根据权利要求1-39任一项的嵌合RNA分子、通过加工所述嵌合RNA分子产生的小RNA分子、根据权利要求40-44或50任一项的多核苷酸、权利要求45或权利要求46的载体、权利要求47-52任一项的宿主细胞、权利要求59的提取物或根据权利要求60-64任一项的组合物。

66. 权利要求65所述的方法,其中将所述嵌合RNA分子或通过加工所述嵌合RNA分子产生的小RNA分子或两者通过局部应用至所述细胞或生物体而与所述细胞或生物体,优选植物细胞、植物、真菌或昆虫接触,或在饲料中提供给所述生物体。

67. 一种减少由害虫或病原体引起的对非人生物体的损害的方法,所述方法包括将以下的一或多种递送至所述害虫或病原体或将以下的一种或多种与所述害虫或病原体接触:根据权利要求1-39任一项的嵌合RNA分子、通过加工所述嵌合RNA分子产生的小RNA分子、根据权利要求40-44或50任一项的多核苷酸、权利要求45或权利要求46的载体、权利要求47-52任一项的宿主细胞、权利要求59的提取物或根据权利要求60-64任一项的组合物。

68. 一种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以下的一种或多种:根据权利要求1-39任一项的嵌合RNA分子、通过加工所述嵌合RNA分子产生的小RNA分子、根据权利要求40-44或50任一项的多核苷酸、权利要求45或权利要求46的载体、根据权利要求47-52任一项的宿主细胞、权利要求59的提取物或根据权利要求60-64任一项的组合物,其中所述RNA分子小RNA分子对所述非人生物体具有有害作用。

69. 权利要求68的方法,其中所述非人生物体是节肢动物或植物。

70. 权利要求69的方法,其中根据权利要求53-55任一项的非人生物体是植物,且所述节肢动物食用所述植物或其部分。

71. 一种在对象中预防或治疗疾病的方法,所述方法包括向所述对象施用以下的一种或多种:根据权利要求1-39任一项的嵌合RNA分子、通过加工所述嵌合RNA分子产生的小RNA分子、根据权利要求40-44或50任一项的多核苷酸、权利要求45或权利要求46的载体、根据权利要求47-52任一项的宿主细胞、权利要求59的提取物或根据权利要求60-64任一项的组合物,其中所述嵌合RNA分子或通过加工所述嵌合RNA分子产生的所述小RNA分子或两者对所述疾病的至少一种症状具有有益效果。

72. 权利要求71的方法,其中所述嵌合RNA分子、多核苷酸、载体或组合物是局部、口服或注射施用的。

73. 权利要求71或72的方法,其中所述对象是脊椎动物。

74. 权利要求73的方法,其中所述脊椎动物是哺乳动物,如人。

75. 根据权利要求1-39任一项的嵌合RNA分子、通过加工所述嵌合RNA分子产生的小RNA分子、根据权利要求40-44或50任一项的多核苷酸、权利要求45或权利要求46的载体、根据权利要求47-52任一项的宿主细胞、权利要求59的提取物或根据权利要求60-64任一项的组合物,其用于在对象中预防或治疗疾病,其中所述嵌合RNA分子或通过加工所述嵌合RNA分

子产生的所述小RNA分子或两者对所述疾病的至少一种症状具有有益效果。

76. 根据权利要求1-39任一项的嵌合RNA分子、通过加工所述嵌合RNA分子产生的小RNA分子、根据权利要求40-44或50任一项的多核苷酸、权利要求45或权利要求46的载体、根据权利要求47-52任一项的宿主细胞、权利要求59的提取物或根据权利要求60-64任一项的组合物在制备用于在对象中预防或治疗疾病的药物中的用途,其中所述嵌合RNA分子或通过加工所述嵌合RNA分子产生的所述小RNA分子或两者对所述疾病的至少一种症状具有有益效果。

77. 一种试剂盒,其包含以下一种或多种:根据权利要求1-39任一项的嵌合RNA分子、通过加工所述嵌合RNA分子产生的小RNA分子、根据权利要求40-44或50任一项的多核苷酸、权利要求45或权利要求46的载体、根据权利要求47-52任一项的宿主细胞、权利要求59的提取物或根据权利要求60-64任一项的组合物。

包含非规范碱基对的RNA分子

技术领域

[0001] 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

[0002] RNA沉默是真核生物中进化保守的基因沉默机制,其由双链RNA (dsRNA) 诱导,所述双链RNA可以是称为发夹结构RNA (hpRNA) 的形式。在基本的RNA沉默途径中,dsRNA被Dicer蛋白加工成短的20-25个核苷酸 (nt) 的小RNA双链体,其中一条链与Argonaute (AGO) 蛋白结合形成RNA诱导的沉默复合体 (RISC)。该沉默复合体使用小RNA作为向导来寻找并结合互补的单链RNA,其中AGO蛋白切割RNA导致其降解。

[0003] 在植物中,存在多种RNA沉默途径,包括microRNA (miRNA)、反式作用小干扰RNA (tasiRNA)、重复相关siRNA (rasiRNA) 和外源(病毒和转基因) siRNA (exosiRNA) 途径。miRNA是从短茎环前体RNA由Dicer-like 1 (DCL1) 在细胞核中加工的20-24nt的小RNA,所述短茎环前体RNA从MIR基因由RNA聚合酶II转录。tasiRNA是主要大小为21nt的阶段性的siRNA,其衍生自长dsRNA的DCL4加工,所述长dsRNA是由RNA依赖性RNA聚合酶6 (RDR6) 从miRNA切割的TAS RNA片段合成的。24-nt rasiRNA由DCL3加工,且前体dsRNA由植物特异性DNA-依赖性RNA聚合酶IV (PolIV) 和RDR2的组合功能从基因组中的重复DNA生成。exosiRNA途径与tasiRNA和rasiRNA途径重叠,并且DCL4和DCL3都参与exosiRNA加工。除了DCL1、DCL3和DCL4外,模式植物拟南芥 (*Arabidopsis thaliana*) 和其他高等植物编码DCL2或等价物,其生成22-nt siRNA,包括22-nt exosiRNA,并在植物中系统性和瞬时性基因沉默中发挥关键作用。所有这些植物小RNA均通过HUA增强子1 (HEN1) 在3'末端核苷酸的2'-羟基处甲基化,并且认为该3'末端2'-O-甲基化稳定植物细胞中的小RNA。miRNA、tasiRNA和exosiRNA在功能上类似于动物细胞中的小RNA,其参与动物中转录后基因沉默或RNA的序列特异性降解。但是,rasiRNA对植物而言是独特的,并具有在同类DNA上引导从头 (de novo) 胞嘧啶甲基化的功能,这种转录基因沉默机制被称为RNA引导的DNA甲基化 (RdDM)。

[0004] dsRNA诱导的RNA沉默已被广泛利用以减少各种真核系统中的基因活性,并且已经开发了许多基因沉默技术。不同的生物体通常适合于不同的基因沉默方法。例如,由于dsRNA诱导的干扰素应答,长dsRNA (长度至少100个碱基对) 不太适合在哺乳动物细胞中诱导RNA沉默,且因此哺乳动物细胞中通常使用较短的dsRNA (少于30个碱基对),而在植物中,具有长dsRNA茎的发夹RNA (hpRNA) 非常有效。在植物中,不同的RNA沉默途径导致了不同的基因沉默技术,例如人工miRNA、人工tasiRNA和病毒诱导的基因沉默技术。然而,迄今为止,RNA沉默技术在植物中的成功应用主要是通过使用长的hpRNA转基因实现的。hpRNA转基因构建体通常包括反向重复,所述反向重复包括由间隔区序列 (形成hpRNA的环) 隔开的靶基因序列的完全互补的有义和反义序列 (当从hpRNA的dsRNA茎转录时),其插入用于在植物细胞中表达的启动子和转录终止子之间。间隔区序列的功能是在构建体制备过程中稳定细菌中的反向重复DNA。产生的hpRNA转录物的dsRNA茎被DCL蛋白加工成引导靶基因沉默的siRNA。hpRNA转基因已被广泛用于敲低基因表达、修饰代谢途径并增强植物的疾病和害虫

抗性以改良作物,且现已报道该技术在作物改良中的许多成功应用(Guo等,2016;Kim等,2019)。

[0005] 但是,最近的研究表明,hpRNA转基因会遭受自我诱导的转录抑制,从而损害了靶基因沉默的稳定性和功效。尽管所有转基因都可能遭受位置或拷贝数依赖性转录沉默的影响,但是hpRNA转基因是独特的,因为它们生成siRNA,siRNA可通过RdDM途径将DNA甲基化引导至其自身序列,并且有可能导致转录自我沉默。

[0006] 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

[0007] 本发明人构思了用于产生RNA分子的遗传构建体的新设计,其包括一个或多个双链RNA区,所述双链RNA区包含多个非规范碱基配对的核苷酸或非碱基配对的核苷酸,或两者,包括具有两个或更多个环序列的形式,在本文中称为环末端dsRNA(ledRNA)。这些RNA分子具有以下一种或多种特征:它们易于合成,在编码它们的遗传构建体转录后在细胞中积累到更高的水平,它们更容易形成dsRNA结构并诱导真核细胞中靶RNA分子有效沉默,并且在植物细胞中加工后可能形成环状RNA分子。当将RNA分子局部应用于植物或喂于动物(例如昆虫)时,RNA分子也是有效的。

[0008] 在第一方面,本发明提供了一种嵌合的核糖核酸(RNA)分子,其包含双链RNA(dsRNA)区,所述双链RNA区包含长度为至少20个连续核苷酸的第一有义核糖核苷酸序列和长度为至少20个连续核苷酸的第一反义核糖核苷酸序列,由此第一有义核糖核苷酸序列和第一反义核糖核苷酸序列能够彼此杂交以形成dsRNA区,其中

[0009] i) 第一有义核糖核苷酸序列由按5'至3'顺序共价连接的第一5'核糖核苷酸、第一RNA序列和第一3'核糖核苷酸组成,

[0010] ii) 第一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,

[0011] iii) 第一5'核糖核苷酸与第二3'核糖核苷酸碱基配对形成dsRNA区的末端碱基对,

[0012] iv) 第二5'核糖核苷酸与第一3'核糖核苷酸碱基配对形成dsRNA区的末端碱基对,

[0013] 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约5%至约40%的核糖核苷酸以非规范碱基对碱基配对或不碱基配对,

[0014] vi) dsRNA区不包含20个连续的规范碱基对,

[0015] vii) RNA分子能够在真核细胞中或体外被加工,从而切割第一反义核糖核苷酸序列,以产生长度为20-24个核糖核苷酸的短反义RNA(asRNA)分子,

[0016] viii) RNA分子或至少一些asRNA分子或两者能够降低真核细胞中靶RNA分子的表达或活性,以及

[0017] ix) RNA分子能够在体外或在细胞中或在两者中通过转录而被酶促地制备。

[0018] 在第一方面的优选实施方案中,第一有义核糖核苷酸序列通过第一连接核糖核苷酸序列与第一反义核糖核苷酸序列共价连接,所述第一连接核糖核苷酸序列包含至少长度为4个核苷酸、或4至1000个核糖核苷酸、或4至200个核糖核苷酸、或4至50个核糖核苷酸、或

至少10个核苷酸、或10至1000个核糖核苷酸、或10至200个核糖核苷酸、或10至50个核糖核苷酸的环序列,由此第一连接核糖核苷酸序列与第一3'核糖核苷酸和第二5'核糖核苷酸共价连接,或与第二3'核糖核苷酸和第一5'核糖核苷酸共价连接,使得所述序列包含在RNA的单一连续链中。在另一个实施方案中,第一连接核糖核苷酸序列与第二3'核糖核苷酸和第一5'核糖核苷酸共价连接,或优选与第一3'核糖核苷酸和第二5'核糖核苷酸共价连接,使得所述序列包含在RNA的单一连续链中。

[0019] 以其最简单的形式,将此类RNA分子称为发夹RNA (hpRNA)。在一个更优选的实施方案中,dsRNA的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约5%至约40%的核糖核苷酸以非规范碱基对碱基配对(优选G:U碱基对)。即,第一有义核糖核苷酸序列的所有核糖核苷酸均以规范碱基对或非规范碱基对与第一反义核糖核苷酸序列的核糖核苷酸碱基配对,由此dsRNA区包含20个连续的碱基对,包括一些非规范碱基对。dsRNA区因此不包含20个连续的规范碱基对。在本发明的hpRNA的更优选的实施方案中,第一反义核糖核苷酸序列与靶RNA的区域完全互补。在该实施方案中,第一有义核糖核苷酸序列与靶RNA的区域在序列上的不同之处在于靶RNA的区域中的C核苷酸被取代为hpRNA中的U核苷酸。此类分子示例于实施例6-11中包含G:U碱基对的发夹RNA。在优选的实施方案中,第一反义核糖核苷酸序列的长度是20至约1000个核苷酸、或20至约500个核苷酸,或如本文所述的其他长度。更优选地, hpRNA在植物细胞或真菌细胞中产生或引入植物细胞或真菌细胞。在这些实施方案中,靶RNA可以是细胞中内源基因的转录物,或者是植物病原体的内源基因的转录物,或害虫(诸如昆虫害虫)的内源基因的转录物。

[0020] 在一个更优选的实施方案中, RNA分子包含第二有义核糖核苷酸序列,并且第一有义核糖核苷酸序列和第一反义核糖核苷酸序列通过第一连接核糖核苷酸序列连接,所述第一连接核糖核苷酸序列包括至少长度为4个核苷酸的环序列,由此第一连接核糖核苷酸序列与第一3'核糖核苷酸和第二5'核糖核苷酸共价连接,并且RNA分子进一步包含第二连接核糖核苷酸序列,所述第二连接核糖核苷酸序列包括长度为至少4个核苷酸的环序列,并且其与第二3'核糖核苷酸和第二有义核糖核苷酸序列共价连接,从而形成1edRNA结构。在一个替代的优选实施方案中, RNA分子包含第二反义核糖核苷酸序列,并且第一有义核糖核苷酸序列和第一反义核糖核苷酸序列通过第一连接核糖核苷酸序列连接,所述第一连接核糖核苷酸序列包括长度为至少4个核苷酸的环序列,由此第一连接核糖核苷酸序列与第二3'核糖核苷酸和第一5'核糖核苷酸共价连接,并且RNA分子进一步包含第二连接核糖核苷酸序列,所述第二连接核糖核苷酸序列包括长度为至少4个核苷酸的环序列,并且其与第二3'核糖核苷酸和第二反义核糖核苷酸序列共价连接,从而形成1edRNA结构。

[0021] 在另一个优选的实施方案中, RNA分子包含第二有义核糖核苷酸序列和第二反义核糖核苷酸序列,其中第二有义核糖核苷酸序列和第二反义核糖核苷酸序列能够互相杂交以形成第二dsRNA区,并且第一有义核糖核苷酸序列和第一反义核糖核苷酸序列通过第一连接核糖核苷酸序列连接,所述第一连接核糖核苷酸序列包括长度为至少4个核苷酸的环序列,由此第一连接核糖核苷酸序列与第一3'核糖核苷酸和第二5'核糖核苷酸共价连接,并且RNA分子进一步或任选地包含第二连接核糖核苷酸序列,所述第二连接核糖核苷酸序列包括长度为至少4个核苷酸的环序列并且其与第二3'核糖核苷酸和第二有义核糖核苷酸序列共价连接,或其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,从而

形成1edRNA结构。在另一个优选的实施方案中, RNA分子包含第二有义核糖核苷酸序列和第二反义核糖核苷酸序列, 并且第一有义核糖核苷酸序列和第一反义核糖核苷酸序列通过第一连接核糖核苷酸序列连接, 所述第一连接核糖核苷酸序列包括长度为至少4个核苷酸的环序列, 由此第一连接核糖核苷酸序列与第二3'核糖核苷酸和第一5'核糖核苷酸共价连接, 并且RNA分子进一步或任选地包含第二连接核糖核苷酸序列, 所述第二连接核糖核苷酸序列包括长度为至少4个核苷酸的环序列并且其与第一3'核糖核苷酸和第二反义核糖核苷酸序列共价连接, 或其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, 从而形成1edRNA结构。在更优选的实施方案中, 第二有义核糖核苷酸序列和第二反义核糖核苷酸序列, 如果存在于RNA分子中, 则各自包含长度为至少20个的连续核苷酸。在这些实施方案中, 第一和第二有义核糖核苷酸序列可以通过间插核糖核苷酸序列共价连接, 所述间插核糖核苷酸序列与靶RNA分子在序列上无关或与靶RNA分子在序列上相关, 或第一和第二有义核糖核苷酸序列共价连接而没有间插核糖核苷酸序列。第一和第二有义核糖核苷酸序列可以形成与靶RNA分子在序列上具有至少50%相同性的一个连续的有义核糖核苷酸区。在进一步的实施方案中, 第一和第二反义核糖核苷酸序列可以通过间插核糖核苷酸序列共价连接, 所述间插核糖核苷酸序列与靶RNA分子的互补体在序列上无关或与靶RNA分子的互补体在序列上相关, 或第一和第二反义核糖核苷酸序列共价连接而没有间插核糖核苷酸序列。第一和第二反义有义核糖核苷酸序列可以形成与靶RNA分子的互补体在序列上具有至少50%相同性的一个连续的反义核糖核苷酸区。在这些实施方案中, RNA分子包含第二有义核糖核苷酸序列和第二反义核糖核苷酸序列, 它们通过碱基配对杂交, 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的总体上5%至40%的核糖核苷酸以非规范碱基对碱基配对或不碱基配对(优选以G:U碱基对碱基配对), 其中第二dsRNA区不包含20个连续的规范碱基对, 并且其中RNA分子能够在真核细胞中或在体外加工, 由此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。

[0022] 在最优选的实施方案中, 考虑整个RNA分子和RNA分子内的每个dsRNA区, 每个有义核糖核苷酸序列及其通过碱基配对杂交的对应的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碱基配对或不碱基配对, 整个RNA分子在任何其dsRNA区内均不包含20个连续的规范碱基对, 并且其中RNA分子能够在真核细胞中或在体外加工, 由此每个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。

[0023] 在优选的实施方案中, 每个连接核糖核苷酸序列的长度独立地在4至约2000个核苷酸之间, 优选地在4至约1200个核苷酸之间, 更优选地在4至约200个核苷酸之间, 且最优选地在4至约50个核苷酸之间。在一个实施方案中, RNA分子还包含5'前导序列或3'尾随序列, 或两者。

[0024] 在第二方面, 本发明提供了一种嵌合的核糖核酸(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

[0025] i) 有义核糖核苷酸序列由按5'至3'顺序共价连接的第一5'核糖核苷酸、第一RNA序列和第一3'核糖核苷酸组成,

[0026] ii) 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，

[0027] iii) 第一5'核糖核苷酸与第二3'核糖核苷酸碱基配对形成dsRNA区的末端碱基对，

[0028] iv) 第二5'核糖核苷酸与第一3'核糖核苷酸碱基配对形成dsRNA区的末端碱基对，

[0029] v) 有义核糖核苷酸序列和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碱基配对或不碱基配对，

[0030] vi) dsRNA区不包含20个连续的规范碱基对，

[0031] vii) RNA分子能够在真核细胞中或体外进行加工，由此切割反义核糖核苷酸序列，以产生长度为20-24个核糖核苷酸的短反义RNA (asRNA) 分子，

[0032] viii) RNA分子或至少一些asRNA分子或两者能够降低真核细胞中靶RNA分子的表达或活性，以及

[0033] ix) RNA分子能够在体外或在细胞中或在两者中通过转录而被酶促制备。

[0034] 如本领域技术人员将意识到的，除了有义核糖核苷酸序列和反义核糖核苷酸序列的长度小于20个连续核苷酸的情况以外，与第一方面有关的每个实施方案都适用于第二方面。

[0035] 在第三方面，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，

[0036] 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个连续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个连续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分子的第一区域杂交，

[0037] 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，

[0038] 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；

[0039] 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

[0040] 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。

[0041] 在第四方面,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,

[0042] 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个连续的核糖核苷酸组成,其中所述第一有义核糖核苷酸序列的至少20个连续的核糖核苷酸与所述第一反义核糖核苷酸序列的至少20个连续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个连续的核糖核苷酸或所述第一反义核糖核苷酸序列的至少20个连续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域相同,

[0043] 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,

[0044] 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,

[0045] 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

[0046] 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。

[0047] 在上述方面的优选的实施方案中,本发明的RNA分子是嵌合RNA分子。

[0048] 在第五方面,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,

[0049] 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,

[0050] 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,

[0051] 其中在所述第一RNA组分中,

[0052] i) 第一有义核糖核苷酸序列由至少20个连续核糖核苷酸组成,其为按5'至3'顺序

共价连接的第一5'核糖核苷酸、第一RNA序列和第一3'核糖核苷酸，

[0053] ii) 第一反义核糖核苷酸序列由至少20个连续核糖核苷酸组成，其为按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，

[0054] iii) 所述第一5'核糖核苷酸与所述第二3'核糖核苷酸碱基配对，

[0055] iv) 所述第二5'核糖核苷酸与所述第一3'核糖核苷酸碱基配对，

[0056] 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碱基配对或者不碱基配对，以及

[0057] vi) 所述第一dsRNA区不包含20个连续的规范碱基对，

[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

[0059] (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

[0060] (b) 所述第一反义核糖核苷酸序列包含至少20个连续核糖核苷酸的序列，所述至少20个连续核糖核苷酸的序列在序列上与所述靶RNA分子的互补体的区域具有至少50%相同性，优选在序列上与所述靶RNA分子的互补体的区域具有至少90%或100%相同性，或

[0061] (c) (a) 和 (b) 二者。

[0062] 在RNA分子具有第一RNA组分的实施方案中，第一RNA组分的第一5'核糖核苷酸和第一3'核糖核苷酸互相碱基配对。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。

[0063] 在优选的实施方案中，第一反义核糖核苷酸序列的至少20个连续的核糖核苷酸都能够与靶RNA分子的第一区域的核苷酸碱基配对。在这种情况下，碱基配对可以是规范的或非规范的，例如具有至少一些G:U碱基对。独立地对于每个G:U碱基对，G可以在靶RNA分子的第一区域中，或者优选地在第一反义核糖核苷酸序列中。可选地，并非第一反义核糖核苷酸序列的至少20个连续核糖核苷酸均与靶RNA分子的第一区域的核苷酸碱基配对。例如，第一反义核糖核苷酸序列的至少20个连续核糖核苷酸中的1、2、3、4或5个核糖核苷酸未与靶RNA分子的第一区域碱基配对。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。

[0064] 在上述方面的一个实施方案中，RNA分子包含一或多个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RNA分子相关，至少部分与靶RNA分子的区域相同或者至少部分与靶RNA分子的区域互补体相同。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连续有义序列的一部分，或与第一和第二RNA组分中的反义序列一起形成一个连续反义序列的一部分。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。

[0065] 在上述方面的实施方案中, 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。

[0066] 在第一至第五方面的实施方案中, 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分子的区域相同, 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域互补, 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相同靶RNA分子的不同区域相同, 其在靶RNA分子中可以是或可以不是连续的。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分子的区域相同。在一个实施方案中, 两个或更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分子是连续的。

[0067] 在第一至第五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分子的区域互补, 优选地为完全互补。靶RNA分子与其互补的区域在靶RNA分子中可以是或不是连续的。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相同靶RNA分子的不同区域互补。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分子的区域互补。在一个优选的实施方案中, 两个或更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分子的互补体是连续的。在优选的实施方案中, 两个或更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碱基对碱基配对。

[0068] 在第一至第五方面的优选实施方案中, RNA分子是单链核糖核苷酸。在最简单形式中, RNA分子包含发夹RNA (hpRNA) 结构, 其具有5'端、长度为至少21个核苷酸的有义核糖核苷酸序列、在至少21个连续的核苷酸上与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、间插环序列和3'端。RNA分子可包含5'-前导序列和/或3'-尾随序列。在另一种形式中, RNA分子包含单链核糖核苷酸, 所述单链核糖核苷酸具有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个连续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, 且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。

[0069] 在第一至第五方面的另一个实施方案中, RNA分子包含与靶RNA的第一区域杂交的第一反义核糖核苷酸序列; 与靶RNA的第二区域杂交的第二反义核糖核苷酸序列, 所述靶

RNA的第二区域与所述靶RNA的第一区域不同;并且所述RNA分子仅包含一个与所述靶RNA具有至少50%序列相同性的有义核糖核苷酸序列,其中所述两个反义序列在所述RNA分子中不连续。在一个实施方案中,靶RNA的第一和第二区域在靶RNA中是连续的。或者,它们不是连续的。

[0070] 在第一至第五方面的另一个实施方案中, RNA分子包含与靶RNA的第一区域具有至少60%相同性的第一有义核糖核苷酸序列;和与靶RNA的第二区域具有至少60%相同性的第二有义核糖核苷酸序列,所述靶RNA的第二区域与所述靶RNA的第一区域不同;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分子中不连续。在一个实施方案中,靶RNA的第一和第二区域在靶RNA分子中是连续的。或者,它们不是连续的。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的相应区域具有至少70%、至少80%、至少90%、至少95%或至少99%相同性,即第一有义序列可以与其靶区具有至少70%相同性,且第二序列与其靶序列具有至少80%相同性,等等。

[0071] 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个连续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个连续碱基对的双链区中,该双链区包括一些非规范碱基对,最优选一些G:U碱基对,所述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。

[0072] 在第三和第四方面的优选实施方案中,第二RNA组分的特征在于:

[0073] i) 第二有义核糖核苷酸序列由至少20个连续核糖核苷酸组成,其为按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,

[0074] ii) 第二反义核糖核苷酸序列由至少20个连续核糖核苷酸组成,其为按5'至3'顺序共价连接的第三5'核糖核苷酸、第四RNA序列和第二3'核糖核苷酸,

[0075] iii) 所述第二5'核糖核苷酸与所述第二3'核糖核苷酸碱基配对,

[0076] iv) 所述第三3'核糖核苷酸与所述第三5'核糖核苷酸碱基配对,

[0077] 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。最优选地,在没有与从第一RNA组分的第一反义序列产生的asRNA组合的情况下或与从第一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。

[0078] 在第五方面的优选实施方案中,第二RNA组分的特征在于:

[0079] i) 第二有义核糖核苷酸序列由至少20个连续核糖核苷酸组成,其为按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,

[0080] ii) 第二反义核糖核苷酸序列由至少20个连续核糖核苷酸组成,其为按5'至3'顺序共价连接的第四5'核糖核苷酸、第四RNA序列和第四3'核糖核苷酸,

[0081] iii) 所述第三5'核糖核苷酸与所述第四3'核糖核苷酸碱基配对,

[0082] iv) 所述第三3'核糖核苷酸与所述第三5'核糖核苷酸碱基配对,

[0083] 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。

[0084] 在上述优选的实施方案的每一个中,更优选第一有义核糖核苷酸序列和第一反义核糖核苷酸序列、和/或第二有义核糖核苷酸序列和第二反义核糖核苷酸序列、和/或每个有义核糖核苷酸序列及其杂交的对应的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碱基配对或者不碱基配对,和/或在互补的有义和反义序列之间形成的dsRNA区不包含20个连续的规范碱基对。更优选地,有义核糖核苷酸序列和其对应的反义核糖核苷酸序列(优选地对于RNA分子的每个dsRNA区)的核糖核苷酸中总体上约12%、约15%、约18%、约21%、约24%、约27%、约30%,或10%至30%、或15%至30%,或甚至更优选16%至25%以非规范碱基对碱基配对,或者不碱基配对。甚至更优选地,RNA分子中的一个或多个dsRNA区的核糖核苷酸中总体上约12%、约15%、约18%、约21%、约24%、约27%、约30%、10%至30%、或15%至30%,或甚至更优选16%至25%以非规范碱基对碱基配对,或者RNA分子中的一个或多个dsRNA区的所有其他核糖核苷酸以规范碱基对碱基配对。在优选的实施方案中,所述第一或第二dsRNA区、或所有dsRNA区中的非规范碱基对的总体上至少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。

[0085] 最优选地,在这些实施方案中,

[0086] (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

[0087] (b) RNA分子中所述第一或第二反义核糖核苷酸序列(优选每个反义核糖核苷酸序列)包含至少20个连续核糖核苷酸的序列,所述至少20个连续核糖核苷酸的序列在序列上与靶RNA分子的互补体的区域具有至少50%相同性、优选具有至少60%相同性、更优选具有至少70%相同性、甚至更优选具有至少80%相同性,最优选与靶RNA分子的互补体的区域具有至少90%相同性或100%相同性,或(a)和(b)二者。

[0088] 在第一至第五方面的实施方案中, 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。

[0089] 在第一至第五方面的实施方案中, 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。

[0090] 在第一至第五方面的实施方案中, 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1200个核糖核苷酸或为4至1000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1000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。在实施方案中, 环的最小长度为20个核苷酸、30个核苷酸、40个核苷酸或50个核苷酸。在一个实施方案中, 真核细胞是脊椎动物细胞或植物细胞, 并且RNA分子的每个环的长度独立地为20至50个核糖核苷酸、或20至40个核糖核苷酸、或20至30个核糖核苷酸。

[0091] 在一个优选的实施方案中, RNA分子中的至少一个环序列包含一个或多个对真核

细胞而言是内源的RNA分子(例如,真核细胞中的miRNA或其他调节RNA)互补的结合序列。如将容易理解的,该特征可以与对于RNA分子的上述任何环长度特征、非规范碱基配对和任何其他特征组合。在一个实施方案中,至少一个环序列包含miRNA的多个结合序列,或多个miRNA的结合序列,或两者。在一个实施方案中,RNA分子中的至少一个环序列包含编码多肽或功能性多核苷酸的开放阅读框。开放阅读框优选可操作地连接至翻译起始序列,由此开放阅读框能够在感兴趣的真核细胞中翻译。例如,翻译起始序列包含内部核糖体进入位点(IRES)或包含在内部核糖体进入位点(IRES)中。所述IRES优选是真核IRES。所翻译的多肽优选长度为50-400个氨基酸残基,或长度为50-300或50-250或50-150个氨基酸残基。当此类RNA分子在植物细胞中产生时,能够被加工形成包含大部分或全部环序列的环状RNA分子,并且能够被翻译以提供高水平的多肽。

[0092] 在第一至第五方面的实施方案中,RNA分子在双链区中没有,或有一个或两个或更多个隆起(bulge)。在本文中,隆起是有义或反义核糖核苷酸序列中的核苷酸,或两个或更多个连续核苷酸,其在dsRNA区中不碱基配对并且在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。

[0093] 在第一至第五方面的实施方案中,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,或两个或更多个连续核苷酸,其在dsRNA区中不碱基配对并且在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。

[0094] 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、或每个dsRNA区包含在双链区中不碱基配对的一个或两个或更多个核苷酸,或独立地包含在双链区中不碱基配对的核苷酸的至多2%、3%、4%、5%、6%、7%、8%、9%或10%。

[0095] 在第一至第五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,例如一种或多种植物物种的真菌病原体的细胞,例如镰刀菌属(*Fusarium*)物种、轮枝孢属(*Verticillium*)物种、或引起白粉病(powdery mildew)的真菌。在一个实施方案中,真核细胞是节肢动物细胞,例如昆虫细胞。优选的昆虫是吸汁液的昆虫,例如蚜虫。例如,昆虫可以是鳞翅目(Lepidopteran)昆虫、鞘翅目(Coleopteran)昆虫或Dipteran昆虫。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。在一个优选的实施方案中,微生物细胞是其中通过从编码RNA分子的遗传构建体转录而产生RNA分子的细胞,其中所述RNA分子基本上或优选主要不在微生物细胞中通过在一个或多个环序列、一个或多个dsRNA区或两者内的切割而被加工。例如,微生物细胞是不具有Dicer酶的酵母细胞或其他真菌细胞。用于产生RNA分子的非常优选的细胞是酿酒酵母(*Saccharomyces cerevisiae*)细胞。微生物细胞可

以是活的,或者可以通过诸如热处理的某种处理杀死,或者可以是干燥粉末的形式。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,例如如果靶RNA是病毒RNA或其他引入的RNA。这类细胞可以针对病毒或其他引入的RNA被预防性地保护。

[0096] 在第一至第五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。

[0097] 因此, 在另一方面, 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五方面和上下文所述的任何实施方案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。

[0098] 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。

[0099] 在第一至第五方面的实施方案中, 形成dsRNA区的有义核糖核苷酸序列和其相应的反义核糖核苷酸序列的核糖核苷酸的总体上约12%、约15%、约18%、约21%、约24%, 或约15%至约30%, 或优选约16%至约25%以非规范碱基对碱基配对或不碱基配对。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。该特征可独立地应用于RNA分子中的一个或多个或所有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10个或9个或优选8个连续的规范碱基对。在另一个实施方案中, dsRNA区包含至少8个连续的规范碱基对, 例如8-12或8-14或8-10个连续的规范碱基对。在优选的实施方案中, dsRNA区中或RNA分子中的所有dsRNA区中的所有核糖核苷酸以规范碱基对或非规范碱基对碱基配对。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碱基配对。

[0100] 在一个实施方案中, RNA分子的一个或多个或所有反义核糖核苷酸序列在序列上与靶RNA分子的区域互补或与两个这样的区域(其在靶RNA分子中可以是或可以不是连续的)具有小于100%相同性, 或约80%至99.9%相同性, 或约90%至98%相同性, 或约95%

至98%相同性、优选98%至99.9%相同性。在优选的实施方案中,一个或多个反义RNA序列在序列上与靶RNA分子的互补体的区域(例如与包含21、23、25、27、30或32个连续核苷酸的区域)具有100%相同性。在一个实施方案中,所述有义或反义核糖核苷酸序列、优选两者长度为至少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个连续核苷酸。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。然而,在dsRNA区中具有超过50个碱基对的RNA分子,例如多达100个或甚至200个碱基对,可用于脊椎动物细胞,条件是dsRNA区在dsRNA区中具有10-30%的核苷酸以G:U碱基对碱基配对。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与之杂交的相应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相应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。

[0101] 在第一至第五方面的实施方案中,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。

[0102] 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优选酿酒酵母细胞中或在其中靶RNA分子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。

[0103] 在第一至第五方面的实施方案中,所述嵌合RNA分子包含序列上相同或优选不同的两个或更多个dsRNA区。

[0104] 在第一至第五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,与具有以规范碱基对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区以规范碱基对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。换句话说,与处理具有相应的dsRNA区的类似RNA分子相比,较少比例的dsRNA分子是通过加工长度为23和/或21个核糖核苷酸的RNA分子形成

的,所述类似RNA分子具有以规范碱基对完全碱基配对的相应dsRNA区。即,根据从RNA分子产生的20-24个核苷酸asRNA的总数的比例,与dsRNA区以规范碱基对完全碱基配对的类似RNA分子相比,这些实施方案的RNA分子不容易加工以提供23-和/或21-核糖核苷酸的短反义RNA。优选地,至少50%的由遗传构建体转录而在细胞中产生的RNA转录物不是通过Dicer加工的。在一个实施方案中,当与加工具有以规范碱基对完全碱基配对的相应dsRNA区的类似RNA分子相比时,当RNA分子在真核细胞中表达即通过在细胞中转录产生时,通过加工RNA分子形成的较大比例的短反义RNA分子在5'末端具有共价连接的超过一个的磷酸盐。即,更大比例的短反义RNA分子具有改变的电荷,这可以在凝胶电泳实验中以分子的移动性改变观察到。

[0105] 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。

[0106] 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,多核苷酸与能够引导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(例如酿酒酵母);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。

[0107] 在一个实施方案中,多核苷酸是嵌合DNA,其依次包含能够启动宿主细胞中RNA分子转录的启动子,启动子可操作地连接至编码RNA分子(优选为hpRNA)的DNA序列,以及转录终止和/或聚腺苷酸化区。在一个优选的实施方案中,RNA分子包含发夹RNA结构,其包含有义核糖核苷酸序列、环序列和反义核糖核苷酸序列,更优选地,其中有义和反义核糖核苷酸序列碱基配对形成dsRNA区,其中dsRNA区中约5%至约40%的核糖核苷酸以非规范碱基对(优选G:U碱基对)碱基配对。在优选的实施方案中,宿主细胞是植物细胞或真菌细胞。

[0108] 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,优选在植物中或真菌细胞中,多核苷酸或载体的启动子区与编码本发明RNA分子的区域可操作地连接,所述启动子区与编码具有以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,较低水平的甲基化为小于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:

[0109] 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

[0110] ii) 当与包含具有以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/或活性降低的水平更低。

[0111] 在另一方面,本发明提供了宿主细胞,其包含本文所述RNA分子、通过加工嵌合RNA分子产生的小RNA分子(长度为20-24nt)、本文所述多核苷酸或包含其的载体。在一个实施方案中,宿主细胞是非人细胞,如细菌细胞、真菌细胞例如酵母细胞如酿酒酵母细胞、植物细胞或非人动物细胞,优选植物细胞。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,宿主细胞是死亡的和/或不能繁殖的。宿主细胞可以是这样的细胞,其中RNA分子通过转录和/或加工产生,或者所述细胞可以是并非其中RNA分子通过转录和/或加工产生的细胞的细胞,例如是包含靶RNA分子的细胞。

[0112] 在一个实施方案中,宿主细胞优选为植物细胞或真菌细胞,其包含嵌合RNA分子或通过加工嵌合RNA分子产生的小RNA分子或二者,其中所述嵌合RNA分子在5'至3'方向包含第一有义核糖核苷酸序列,包括环序列的第一连接核糖核苷酸序列和第一反义核糖核苷酸序列。

[0113] 在另一个实施方案中,宿主细胞是真核细胞,其包含编码根据权利要求1至39中任一项所述的嵌合RNA分子的多核苷酸的至少两个拷贝,并且其中

[0114] i) 所述真核细胞中所述靶RNA分子的表达或活性的降低水平与如果细胞具有所述多核苷酸的单个拷贝,所述靶RNA分子的表达或活性的降低水平大致相同或更高,和/或

[0115] ii) 当与包含具有以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,所述真核细胞中所述靶RNA分子的表达和/或活性降低的水平更低。

[0116] 在另一方面,本发明提供了非人生物体,优选动物或植物,其包含本发明的RNA分子、优选本文所述的嵌合RNA分子、或通过加工所述嵌合RNA分子产生的小RNA分子(长度为20-24nt),或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,优选植物或真菌,是转基因的,因为它包含本发明的多核苷酸。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。本发明还包括动物和植物部分,以及从其获得的产物,其包含RNA分子或通过加工所述嵌合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,例如由此产生的种子、农作物、收获产品和收获后的产品。

[0117] 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。优选地,多核苷酸是编码RNA分子的嵌合DNA分子。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。

[0118] 在另一方面,本发明提供了生产细胞或非人生物体、优选植物细胞、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或RNA分子导入细胞,优选动物细胞、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,细胞是动物细胞,例如人细胞,其可以是培养物中的动物细胞。在一个实施方案中,非人生物体从细胞或子代细胞生成(例如通过再生转基因植物和任选地由其产生子代植物)。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产生RNA分子。在一个实施方案中,所述非人生物体(例如动物或植物)是对害虫或病原体(例如动物害虫或病原体,植物害虫或病原体,优选昆虫害虫或真菌病原体)具有抗性的。在一个实施方案中,方法包括测试包含本发明的多核苷酸或载体或RNA分子的一种或多种非人生物体,优选植物的对害虫或病原体的抗性的步骤。被测试的非人生物体(例如植物)可以是首先引入本发明的多核苷酸或载体或RNA分子的非人生物体(优选植物)的后代,并且因此方法可以包括获得这种后代的步骤。方法可以进一步包括识别和/或选择对害虫或病原体具有抗性的非人生物体(例如动物或植物)的步骤。例如,可以对各自包含本发明的多核苷酸或载体或RNA分子的多种非人类生物体(例如动物或植物)进行测试,以鉴定对所述害虫或病原体具有抗性的非人类生物体,以及鉴定从鉴定出的非人类生物体、动物或植物获得的后代。。

[0119] 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、通过加工所述嵌合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、通过加工嵌合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。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,例如以局部喷洒,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。

[0120] 在一个实施方案中,包含本发明的RNA分子或通过加工嵌合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。

[0121] 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(例如其部分中)靶RNA分子的水平 and/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嵌合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。在本文中,递送可以通过将RNA分子或小RNA分子或其混合物或本发明的多核苷酸或载体喂食、接触、暴露、转化或以其他方式引入细胞或生物体。可以通过使用增加本发明的RNA分子、多核苷酸或载体的摄取的试剂来增强引入,例如借助于转染促进剂、DNA结合多肽或RNA结合多肽、或者可以不添加这样的试剂而进行,例如,通过种植本发明的多核苷酸或载体的转基因的种子并允许种子生长成表达本发明的RNA分子的转基因植物。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 and/或活性,所述靶RNA分子是不同的,例如降低两种或更多种在序列上相关(例如来自基因家族)的靶RNA的水平 and/或活性。因此,在一个实施方案中,通过局部应用至细胞或生物体,或以饲料提供给生物体而将嵌合RNA分子

或通过加工所述嵌合RNA分子产生的小RNA分子或两者与细胞或生物体,优选植物细胞、植物、真菌或昆虫接触。

[0122] 在另一方面,本发明提供了防治非人生物体(例如动物害虫或病原体或植物害虫或病原体)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嵌合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分子和/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,并且昆虫摄食植物或其部分,由此防治昆虫。防治可包括害虫或病原体的存活降低,或害虫或病原体的适应性或繁殖性降低,或两者。防治可以涵盖降低了首先引入RNA分子的害虫或病原体的后代的存活和/或繁殖。

[0123] 本发明的另一方面涉及减少由害虫或病原体引起的对非人生物体例如动物或植物的损害的方法,其包括将本发明的一个或多个RNA分子或通过加工所述嵌合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、由本发明的方法产生的提取物、或本发明的组合物递送至害虫或病原体或与害虫或病原体接触。在一个实施方案中,方法包括播种对于本发明的多核苷酸是转基因的种子,由此所得植物表达该转基因以产生本发明的RNA分子,从而减少由害虫或病原体引起的损害。因此,本发明为农民提供了防治动物或植物的害虫或病原体的手段。本发明扩展到细胞和生物体,例如,动物或植物或其部分,其包含提供给细胞或生物体的本发明的RNA分子、多核苷酸或载体,以及扩展到害虫或病原体,所述害虫或病原体包含RNA分子或通过加工所述嵌合RNA分子产生的小RNA分子(长度为20-24nt)或两者、或本发明的多核苷酸或载体。所述害虫或病原体是活的或者是死的。本发明还涉及包含RNA分子或小RNA分子或两者的后代细胞或生物体。

[0124] 在一个实施方案中,本发明提供了预防或治疗对象的疾病的方法,所述方法包括向对象施用一种或多种本发明的RNA分子或通过加工所述嵌合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分子或两者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,对象是脊椎动物。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。

[0125] 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预防或治疗对象的疾病的用途,其中所述RNA分子或小RNA分子或两者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对象的疾病的药物用途,其中所述RNA分子或由其产生的小RNA分子或两者对所述疾病的至少一种症状具有有益效果。

[0126] 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。试剂盒可进一步包括使用试剂盒的说明书。

[0127] 尽管更广泛地用于转基因表达系统中,如本文所讨论的,也有dsRNA技术的应用依赖于dsRNA分子的大规模生产的需求,例如喷洒农作物以防治疾病和/或害虫。本发明人已经将酿酒酵母鉴定为适合用于大规模生产过程的生物体,因为其中表达的dsRNA分子没有被切割。因此,在另一方面,本发明提供了生产dsRNA分子的过程,所述过程包括

[0128] a) 培养表达一种或多种编码一个或多个dsRNA分子的多核苷酸的酿酒酵母,以及

[0129] b) 收集产生dsRNA分子地酿酒酵母或来自酿酒酵母的dsRNA分子,

[0130] 其中酿酒酵母的培养体积为至少1升。

[0131] dsRNA可以具有任何结构,例如shRNA、miRNA或本发明的dsRNA。

[0132] 在一个实施方案中,酿酒酵母的培养体积为至少10升、至少100升、至少1000升、至少10000升或至少100000升。

[0133] 在一个实施方案中,所述过程产生至少0.1g/升、至少0.5g/升或至少1g/升的本发明的RNA分子。

[0134] 使用所述过程产生的酿酒酵母或从中分离出的dsRNA分子(以纯化或部分纯化(例如提取物)状态)可用于本文所述的方法,例如但不限于减少或下调细胞或生物体中靶RNA分子的水平和/或活性的方法,减少害虫或病原体对非人生物体造成的损害的方法,防治非人生物体的方法或预防或治疗对象中的疾病的方法。

[0135] 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。

[0136] 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。

[0137] 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。

[0138] 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

[0139] 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具有相同性,所述反义序列与所述有义序列互补并且被分成两个区域,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的互补体具有相同性,所述有义序列与所述反义序列互补并且被分成两个区域,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。

[0140] 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。

[0141] 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。

[0142] 图4.ledRNA处理在上述处理区域(1)和未处理区域(3)均诱导GUS下调。

[0143] 图5.ledRNA诱导本氏烟草(*N.benthamiana*)叶中的FAD2.1基因沉默。

[0144] 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。

[0145] 图7.将GUS靶基因区域的核苷酸序列(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。

[0146] 图8.将GUS靶基因区域的核苷酸序列(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。

[0147] 图9.将GUS靶基因区域的核苷酸序列(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。

[0148] 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。

[0149] 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。

[0150] 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。

[0151] 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)。

[0152] 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。

[0153] 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PPGH11和PPGH24的DNA;泳道4-19:16种不同转基因植物的DNA。

[0154] 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PPGH11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。

[0155] 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的移动性。用U6RNA的反义RNA作为泳道上样对照对印迹进行再探测。

[0156] 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。

[0157] 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。

[0158] 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。

[0159] 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。

[0160] 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。

[0161] 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。

[0162] 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靶基因和的互补体的区域的核苷酸序列(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。

[0163] 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靶基因的互补体的区域的核苷酸序列(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。

[0164] 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。

[0165] 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]。

[0166] 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*)。

[0167] 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。

[0168] 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。

[0169] 图31.发夹RNA构建体的启动子和5'区中的DNA甲基化水平。

[0170] 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。

[0171] 图33.G:U hpEIN2品系中的35S启动子仅显示弱甲基化($P<10\%$)。

[0172] 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。

[0173] 图35.48小时时在HeLa细胞中测试哑铃型质粒。

[0174] 图36.dsRNA分子的可能修饰的实施例。

[0175] 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100 μ l的50ng/ μ l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100 μ l含有MpC002、MpRack-1或对照ledGFP的200ng/ μ l ledRNA后,在五天的时间过程中存活的蚜虫的百分比。

[0176] 图38.使用全长有义GUS转录物作为探针的Northern印迹杂交以检测ledGUS和hpGUS RNA。底部的“+”表示高GUS表达;“-”表示低/无GUS表达,即强GUS沉默。

[0177] 图39.Northern印迹杂交检测长hpEIN2和ledEIN2 RNA(上图)和衍生自两个构建体的siRNA(下图)。

[0178] 图40.从GUS hpRNA构建体表达的转录物的茎环结构的示意图。转录物具有互补的有义和反义序列,其碱基配对形成GUS序列特异性dsRNA茎,对于茎具有以碱基对(bp)指示的长度,以及环中核苷酸(nt)的数目。GFP hpRNA构建体编码的转录物形成具有完全规范碱基配对的GFP特异性dsRNA茎(GFPhp[WT])或具有约25%碱基对为G:U碱基对的dsRNA茎(GFPhp[G:U]),具有衍生自GUS编码序列的区域的环。GFPhp转录物的环序列各自包含两个与miR165/miR166互补的序列,并因此提供这些miRNA的结合位点。

[0179] 图41.Northern印迹杂交分析表明,编码hpRNA的转基因在植物细胞中表达时生成不同的环序列片段。(A) GUS靶基因(GUS)和具有1100nt间隔区/环序列的长hpRNA转基因GUShp1100的表达。将编码黄瓜花叶病毒2b RNA沉默抑制物(CMV2b)的构建体包括其中以增强转基因表达。(B) Northern印迹分析显示来自稳定转化的拟南芥植物中两个短hpRNA转基因GUShp93-1和GUShp93-2表达的RNA。将RNA样品用RNase I处理(+)或未处理(-)。两种RNA印迹均与环特异性反义RNA探针杂交。

[0180] 图42.GUShp1100的环在本氏烟草细胞中积累至高水平,并且对RNase R消化具有抗性。

[0181] 图43.表达GUShp1100构建体的转基因酿酒酵母显示对应于全长发夹RNA转录物的单个RNA分子种类。下图显示来自转基因酿酒酵母的RNA样品的Northern印迹杂交。

[0182] 图44.在酿酒酵母中表达的GUShp1100转录物保持全长,且不形成环状环RNA。前四个泳道使用了GUShp1100全长或dsRNA茎的体外转录物,补充有从野生型本氏烟草叶中分离的总RNA。

[0183] 图45.hpRNA环可用作miRNA的有效序列特异性阻遏物。(A) GFPhp[G:U]构建体在转基因拟南芥植物中诱导强的miR165/166抑制表型。(B) Northern印迹杂交,以确定来自转基因拟南芥植物的RNA中GFPhp转录物的丰度。(C) GFPhp环的环状RNA的RT-qPCR分析。

[0184] 序列列表的关键字

[0185] SEQ ID NO:1-GFP 1edRNA的核糖核苷酸序列。

[0186] SEQ ID NO:2-GUS 1edRNA的核糖核苷酸序列。

[0187] SEQ ID NO:3-本氏烟草(N.benthamiana)FAD2.1 1edRNA的核糖核苷酸序列。

[0188] SEQ ID NO:4-编码GFP 1edRNA的核苷酸序列。

[0189] SEQ ID NO:5-编码GUS 1edRNA的核苷酸序列。

[0190] SEQ ID NO:6-编码本氏烟草(N.benthamiana)FAD2.1 1edRNA的核苷酸序列。

[0191] SEQ ID NO:7-编码GFP的核苷酸序列。

[0192] SEQ ID NO:8-编码GUS的核苷酸序列。

[0193] SEQ ID NO:9-编码本氏烟草(N.benthamiana)FAD2.1的核苷酸序列。

[0194] 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。

[0195] 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。

[0196] 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。

[0197] 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。

[0198] 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。

[0199] SEQ ID NO:15-hpGUS[wt]RNA的发夹结构(包括其环)的核糖核苷酸序列。

[0200] SEQ ID NO:16-hpGUS[G:U]RNA的发夹结构(包括其环)的核糖核苷酸。

[0201] SEQ ID NO:17-hpGUS[1:4]RNA的发夹结构(包括其环)的核糖核苷酸。

[0202] SEQ ID NO:18-hpGUS[2:10]RNA的发夹结构(包括其环)的核糖核苷酸。

[0203] 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。

[0204] 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。

[0205] 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。

[0206] 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]。

[0207] 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。

[0208] 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]。

[0209] 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]。

[0210] 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]。

[0211] 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。

[0212] 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。

[0213] 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的互补体的核苷酸序列。

[0214] SEQ ID NO:30-对应于拟南芥 (*A.thaliana*) CHS基因的cDNA的核苷酸804-1103的互补体的核苷酸序列。

[0215] 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)

[0216] 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)

[0217] 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。

[0218] 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。

[0219] 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。

[0220] 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。

[0221] 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。

[0222] 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。

[0223] 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。

[0224] SEQ ID NO:40-EGFPcDNA。

[0225] 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。

[0226] 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。

[0227] 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。

[0228] 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。

[0229] 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。

[0230] 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。

[0231] 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。

[0232] 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]。

[0233] 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]。

[0234] 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]。

[0235] 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。

[0236] SEQ ID No:52-用于扩增200bp GUS有义序列的寡核苷酸引物(GUS-WT-F)

[0237] SEQ ID No:53-用于扩增200bp GUS有义序列的寡核苷酸引物(GUS-WT-R)

[0238] 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)

[0239] 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)

[0240] 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)

[0241] 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)

[0242] 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)

[0243] 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)

[0244] SEQ ID No:60-编码正向引物(35S-F3)的核苷酸序列

[0245] SEQ ID NO:61-编码反向引物(GUSwt-R2)的核苷酸序列

[0246] SEQ ID NO:62-编码正向引物(GUSgu-R2)的核苷酸序列

[0247] SEQ ID NO:63-编码反向引物(GUS4m-R2)的核苷酸序列

[0248] SEQ ID No:64-编码正向引物(35S-F2)的核苷酸序列

[0249] SEQ ID NO:65-编码反向引物(35S-R1)的核苷酸序列

[0250] 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)

- [0251] 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)
- [0252] 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)
- [0253] 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)
- [0254] 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)
- [0255] 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)
- [0256] 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)
- [0257] 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)
- [0258] 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)
- [0259] 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)
- [0260] 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)
- [0261] 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)
- [0262] SEQ ID No:78-编码正向引物(CHS-200-F2)的核苷酸序列
- [0263] SEQ ID NO:79-编码反向引物(CHS-200-R2)的核苷酸序列
- [0264] SEQ ID No:80-编码正向引物(Actin2-For)的核苷酸序列
- [0265] SEQ ID NO:81-编码反向引物(Actin2-Rev)的核苷酸序列
- [0266] SEQ ID NO:82-编码正向引物(Top-35S-F2)的核苷酸序列
- [0267] SEQ ID NO:83-编码反向引物(Top-35S-R2)的核苷酸序列
- [0268] SEQ ID NO:84-编码正向引物(Link-35S-F2)的核苷酸序列
- [0269] SEQ ID NO:85-编码反向引物(Link-EIN2-R2)的核苷酸序列
- [0270] SEQ ID NO:86-有义si22-核糖核苷酸序列
- [0271] SEQ ID NO:87-反义si22-核糖核苷酸序列
- [0272] SEQ ID NO:88-正向引物的核糖核苷酸序列
- [0273] SEQ ID NO:89-反向引物的核糖核苷酸序列
- [0274] SEQ ID NO:90-正向引物的核糖核苷酸序列
- [0275] SEQ ID NO:91-反向引物的核糖核苷酸序列
- [0276] SEQ ID NO:92-dsRNA分子的可能修饰
- [0277] 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。
- [0278] 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。
- [0279] 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。

[0280] 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。

[0281] 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。

[0282] 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。

[0283] 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。

[0284] 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。

[0285] 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。

[0286] 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。

[0287] 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。

[0288] 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。

[0289] SEQ ID NO:105-对应于本氏烟草 (*Nicotiana benthamiana*) TOR基因的cDNA的蛋白质编码区的核苷酸序列。

[0290] SEQ ID NO:106-编码靶向本氏烟草 (*Nicotiana benthamiana*) TOR基因的ledRNA构建体的嵌合DNA的核苷酸序列。

[0291] 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。

[0292] 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。

[0293] 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。

[0294] 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。

[0295] 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。

[0296] 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。

[0297] 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。

[0298] 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。

[0299] 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。

[0300] 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。

[0301] 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。

[0302] 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。

[0303] 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。

[0304] 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。

[0305] 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。

[0306] 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。

[0307] SEQ ID NO:123-对应于桃蚜 (*Myzus persicae*) MpC002 基因的 cDNA 的蛋白编码区的核苷酸序列。

[0308] 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。

[0309] SEQ ID NO:125-编码靶向桃蚜 (*M.persicae*) C002 基因的 ledRNA 的嵌合构建体的核苷酸序列。

[0310] 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。

[0311] SEQ ID NO:127-对应于棉铃虫 (*Helicoverpa armigera*) ABCwhite 基因的 cDNA 的核苷酸序列。

[0312] 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。

[0313] 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。

[0314] 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。

[0315] 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)。

[0316] 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。

[0317] 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。

[0318] 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。

[0319] 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。

[0320] SEQ ID NO:136-TaMlo靶基因 (AF384144)。

[0321] 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。

[0322] SEQ ID NO:138-对应于华东葡萄 (*Vitis pseudoreticulata*) ML0基因 (登录号 KR362912) 的cDNA的蛋白编码区的核苷酸序列。

[0323] SEQ ID NO:139-编码靶向华东葡萄 (*Vitis*) ML0基因的第一ledRNA构建体的嵌合DNA的核苷酸序列。

[0324] SEQ ID NO:140-Cyp51同系物1 (登录号KK764651.1,基因座RSAG8_00934)。

[0325] SEQ ID NO:141-Cyp51同系物2 (登录号KK764892.1,基因座号RSAG8_12664)。

[0326] SEQ ID NO:142-靶向编码Cyp51的基因的编码ledRNA构建体的嵌合DNA的核苷酸序列。

[0327] SEQ ID NO:143-CesA3靶基因 (登录号JN561774.1)。

[0328] SEQ ID NO:144-靶向编码CesA3的基因的编码ledRNA构建体的嵌合DNA的核苷酸序列。

具体实施方式

[0329] 一般技术和定义

[0330] 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)。

[0331] 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 1 and 2, IRL Press (1991), D.M.Glover and B.D.Hames (editors), *DNA Cloning: A Practical Approach*, Volumes 1-4, IRL Press (1995 and 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)。

[0332] 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。

[0333] 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。

[0334] 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。

[0335] 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。

[0336] 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分子的至少一个区域共享互补性。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。

[0337] 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分子的至少一个区域至少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。

[0338] 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的区域中产生错配。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。

[0339] 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。

[0340] 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。

[0341] 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。

[0342] 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 NaH₂PO₄(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 NaH₂PO₄(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。

[0343] 本发明还涵盖跨越连续核糖核苷酸的“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连续的核糖核苷酸碱基对。连续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“连续”用于指一系列核糖核苷酸。包含连续系列的核糖核苷酸将通过一系列连续的磷酸二酯键连接,每个核糖核苷酸直接与下一个键合。

[0344] 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。

[0345] 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RNA单一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。

[0346] 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。

[0347] 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具有有害作用。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。

[0348] 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存在于对象中,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 and/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、炎症性疾病、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。

[0349] 可以根据本发明的RNA分子以及包含该RNA分子的组合物施用至对象。术语例如“对象”、“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,对象是哺乳动物。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,对象是人。例如,对象可以是成人。在另一个实施例中,对象可以是孩子。在另一个实施例中,对象可以是青少年。在另一个实施例中,可以根据本发明的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,可以根据本发明的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。

[0350] 如本文所用,“抗性”或其变形是相对术语,其中RNA分子的存在增加抗性,例如害虫或病原体的繁殖减少,或对生物体的破坏水平降低。

[0351] 如本文所用,术语“在序列上与靶标不相关”是指沿着间插RNA序列的全长具有小于50%相同性的分子。另一方面,术语“在序列上与靶标相关”是沿着间插RNA序列的全长具有50%或更多相同性的分子。

[0352] 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。

[0353] 如本文所用,除非另有相反说明,术语“约”是指指定值的 $\pm 20\%$ 、更优选 $\pm 10\%$ 。

[0354] 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。

[0355] 非规范碱基配对

[0356] 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(连续)反义核糖核苷酸序列的区域杂交,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(连续)有义核糖核苷酸序列的区域杂交,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。

[0357] 在以下实施方案中,如果仅存在一个(连续的)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。

[0358] 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。

[0359] 在一个实施方案中,本发明的RNA分子的dsRNA区包含20个连续碱基对,其中所述20个连续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个连续碱基对,其中20个连续碱基对中的至少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个连续碱基对,其中20个连续碱基对中的至少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个连续碱基对,其中20个连续碱基对中的至少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个连续碱基对,其中20个连续碱基对中的至少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个连续碱基对,其中20个连续碱基对中的至少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个连续碱基对,其中20个连续碱基对中的至少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个连续碱基对,其中20个连续碱基对中的至少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个连续碱基对,其中20个连续碱基对中的至少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个连续碱基对中最多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个连续碱基对中的每一个。

[0360] 在一个实施方案中,本发明的RNA分子的dsRNA区包含21个连续碱基对,其中所述21个连续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个连续碱基对,其中21个连续碱基对中的至少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个连续碱基对,其中21个连续碱基对中的至少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个连续碱基对,其中21个连续碱基对中的至少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个连续碱基对,其中21个连续碱基对中的至少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个连续碱基对,其中21个连续碱基对中的至少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个连续碱基对,其中21个连续碱基对中的至少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个连续碱基对,其中21个连续碱基对中的至少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个连续碱基对,其中21个连续碱基对中的至少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个连续碱基对中最多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个连续碱基对中的每一个。

[0361] 在一个实施方案中,本发明的RNA分子的dsRNA区包含22个连续碱基对,其中所述22个连续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个连续碱基对,其中22个连续碱基对中的至少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个连续碱基对,其中22个连续碱基对中的至少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个连续碱基对,其中22个连续碱基对中的至少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个连续碱基对,其中22个连续碱基对中的至少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个连续碱基对,其中22个连续碱基对中的至少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个连续碱基对,其中22个连续碱基对中的至少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个连续碱基对,其中22个连续碱基对中的至少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个连续碱基对,其中22个连续碱基对中的至少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个连续碱基对中最多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个连续碱基对中的每一个。

[0362] 在一个实施方案中,本发明的RNA分子的dsRNA区包含23个连续碱基对,其中所述23个连续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个连续碱基对,其中23个连续碱基对中的至少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个连续碱基对,其中23个连续碱基对中的至少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个连续碱基对,其中23个连续碱基对中的至少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个连续碱基对,其中23个连续碱基对中的至少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个连续碱基对,其中23个连续碱基对中的至少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个连续碱基对,其中23个连续碱基对中的至少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个连续碱基对,其中23个连续碱基对中的至少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个连续碱基对,其中23个连续碱基对中的至少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个连续碱基对中最多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个连续碱基对中的每一个。

[0363] 在一个实施方案中,本发明的RNA分子的dsRNA区包含24个连续碱基对,其中所述24个连续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个连续碱基对,其中24个连续碱基对中的至少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个连续碱基对,其中24个连续碱基对中的至少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个连续碱基对,其中24个连续碱基对中的至少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个连续碱基对,其中24个连续碱基对中的至少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个连续碱基对,其中24个连续碱基对中的至少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个连续碱基对,其中24个连续碱基对中的至少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个连续碱基对,其中24个连续碱基对中的至少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个连续碱基对,其中24个连续碱基对中的至少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个连续碱基对中最多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个连续碱基对中的每一个。

[0364] 在以下实施方案中,如果仅存在一个(连续的)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个连续的规范碱基对,即20个连续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个连续的规范碱基对。在一个实施方案中,dsRNA区不包含18个连续的规范碱基对。在一个实施方案中,dsRNA区不包含17个连续的规范碱基对。在一个实施方案中,dsRNA区不包含16个连续的规范碱基对。在一个实施方案中,dsRNA区不包含15个连续的规范碱基对。在一个实施方案中,dsRNA区不包含14个连续的规范碱基对。在一个实施方案中,dsRNA区不包含13个连续的规范碱基对。在一个实施方案中,dsRNA区不包含12个连续的规范碱基对。在一个实施方案中,dsRNA区不包含11个连续的规范碱基对。在一个实施方案中,dsRNA区不包含10个连续的规范碱基对。在一个实施方案中,dsRNA区不包含9个连续的规范碱基对。在一个实施方案中,dsRNA区不包含8个连续的规范碱基对。在一个实施方案中,dsRNA区不包含7个连续的规范碱基对。在上述实施方案中,优选RNA分子的dsRNA区或RNA分子中每个dsRNA区的连续规范碱基配对的最长子区是5、6或7个连续规范碱基配对,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个连续碱基对。在一个优选的实施方案中,dsRNA区包含12-19或20个连续碱基对。在一个实施方案中,dsRNA区包含14-19或20个连续碱基对。在这些实施方案中,dsRNA区包含15个连续碱基对。在一个实施方案中,dsRNA区包含16、17、18或19个连续碱基对。在一个实施方案中,dsRNA区包含20个连续碱基对。优选地,在上述实施方案中,连续碱基对包含至少一个G:U碱基对的至少一个非规范碱基对,更优选地,连续碱基对区中的所有非规范碱基对为G:U碱基对。

[0365] 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,在子区中连续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,在子区中连续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,在子区中连续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。

[0368] 在一个实施方案中，dsRNA区包含10个连续碱基对的子区，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个连续碱基对，其中10个连续碱基对中的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个连续碱基对，其中10个连续碱基对中的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个连续碱基对，其中10个连续碱基对中的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个连续碱基对，其中10个连续碱基对中的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个连续碱基对，其中10个连续碱基对中的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个连续碱基对，其中15个连续碱基对中的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个连续碱基对，其中10个连续碱基对中的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个连续碱基对，其中10个连续碱基对中的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个连续碱基对，其中10个连续碱基对中的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个连续碱基对，其中10个连续碱基对中的2-4个是非规范碱基对。在一个实施方案中，RNA分子的一个(连续)或多个或所有dsRNA区中的非规范碱基对不与非碱基对相邻。在另一个实施方案中，非规范碱基对是来自非碱基对的至少2个连续的碱基对。在另一个实施方案中，非规范碱基对是来自非碱基对的至少3、4、5、6、7、8、9、10或更多个连续的碱基对。在一个实施方案中，RNA分子的一个(连续)或多个或所有dsRNA区中的非规范碱基对不与环序列相邻。在另一个实施方案中，非规范碱基对是来自环序列的至少2个连续的碱基对。在另一个实施方案中，非规范碱基对是来自环序列的至少3、4、5、6、7、8、9、10或更多个连续的碱基对。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。

[0369] 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。

[0370] 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。

[0371] 在另一个实施方案中,非规范碱基对的3%至50%是G:U碱基对。在另一个实施方案中,非规范碱基对的5%至30%是G:U碱基对。在另一个实施方案中,非规范碱基对的10%至30%是G:U碱基对。在另一个实施方案中,非规范碱基对的15%至20%是G:U碱基对。

[0372] 在以上实施方案的一个实例中,在RNA分子的一个(连续)或多个或所有dsRNA区中存在至少3个G:U碱基配对。在另一个实例中,存在至少4、5、6、7、8、9或10个G:U碱基配对。在另一个实例中,存在至少3至10个G:U碱基配对。在另一个实例中,存在至少5至10个G:U碱基配对。

[0373] 包含非规范碱基配对的dsRNA区包含20个连续核苷酸的反义序列,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个与相同靶RNA分子互补(即与相同靶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。

[0374] 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。

[0375] 在一个实施方案中,嵌合RNA分子在分子的环的碱基处不包含非规范碱基对。在另一实施方案中,一个、两个、三个、四个、五个或更多个或所有非规范碱基对侧翼为规范碱基对。

[0376] 在一个实施方案中,嵌合RNA分子包含至少一个植物DCL-1切割位点。

[0377] 在一个实施方案中,靶RNA分子不是病毒RNA分子。

[0378] 在一个实施方案中,靶RNA分子不是南非木薯花叶病毒(South African cassava mosaic virus) RNA分子。

[0379] 在一个实施方案中,嵌合RNA分子包含至少一个非碱基对或至少一段非碱基对,其侧翼为规范碱基对、非规范碱基对或规范碱基对和非规范碱基对。例如,这可以是本文所述的隆起。

[0380] 在一个实施方案中,嵌合RNA分子不包含具有大于11个规范碱基对的双链区。

[0381] 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。间隔区序列的一些核

糖核苷酸可以与RNA分子中的其他核糖核苷酸碱基配对,例如与间隔区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区的侧翼为末端环。

[0382] 在一个实施方案中,当本发明RNA分子的dsRNA区在5个连续碱基对的任何子区中具有至少3个非规范碱基对时,非规范碱基对不连续,而是由一个或多个规范碱基对隔开,即dsRNA区不具有3个或更多个连续的非规范碱基对。在一个实施方案中,dsRNA区不具有4个或更多个连续的非规范碱基对。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。

[0383] 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。

[0384] 沉默活性

[0385] 本发明的RNA分子具有反义活性,因为它们包含与靶RNA分子的区域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分子的区域基本上互补。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分子的互补体相比可以具有插入、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分子的第一区域相同的约15、约16、约17、约18、约19或更多个连续核苷酸。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分子的第一区域相同的约20个连续核苷酸。

[0386] “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。

[0387] 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、外部向导序列(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。

[0388] 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域,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。

[0389] 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。

[0390] 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。

[0391] 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(参见例如Shure et al., 1983; Preiss et al., 1987; Gupta et al., 1988; Olive et al., 1989; Bhattacharyya et al., 1990; Dunwell, 2000; Brar et al., 1996; Kishore

and Somerville,1993;US5,530,192和WO 94/18337)。

[0392] 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)。

[0393] 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。

[0394] 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。

[0395] 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。

[0396] 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。

[0397] ledRNA分子

[0398] 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:

[0399] -连接核糖核苷酸序列,其共价连接第一和第二RNA组分;

[0400] -5'前导序列;和,

[0401] -3'尾随序列。

[0402] 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个连续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个连续核糖核苷酸的第一反义核糖核苷酸序列,其中在RNA分子中第一反义核糖核苷酸序列与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分子的第一区域杂交。

[0403] 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个连续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个连续核糖核苷酸的第一反义核糖核苷酸序列,其中所述RNA分子中所述第一反义核糖核苷酸序列与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分子的第一区域的互补体的序列相同。这两个实施方案的该第一RNA组分的实例示意性地显示在图1A的左半部分或图1B的

右半部分中。

[0404] 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,由此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。

[0405] 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。

[0406] 在一个实施方案中,第一有义序列与靶RNA的区域具有基本上的序列相同性,该相同性可以是对于长度小于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个连续核糖核苷酸、优选至少20个连续核糖核苷酸与靶RNA分子的第一区域在序列上具有至少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%相同性。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个连续核糖核苷酸与靶RNA分子的第一区域具有100%相同性。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分子的区域具有100%相同性,其余的核糖核苷酸与靶RNA分子具有至少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%相同性。

[0407] 在一个实施方案中,第一有义核糖核苷酸序列的至少20个连续核糖核苷酸与靶RNA分子的第一区域具有至少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%相同性。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分子的区域具有100%相同性,其余的核糖核苷酸与靶RNA分子具有至少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%相同性。在另一个实施方案中,第一有义核糖核苷酸序列的至少20个连续核糖核苷酸与靶RNA分子的第一区域具有100%相同性。

[0408] 在一个实施方案中,所述第一反义序列与所述靶RNA的区域的互补体具有基本上的序列相同性,该相同性可以是对于所述互补体的长度小于20个核苷酸的序列。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个连续核糖核苷酸、优选至少20个连续核糖核苷酸与靶RNA分子的第一区域的互补体在序列上具有至少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%相同性。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个连续核糖核苷酸与靶RNA分子的第一区域的互补体具有100%相同性。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个核糖核苷酸与靶RNA分子的区域的互补体具有100%相同性,其余的核糖核苷酸与靶RNA分子的互补体具有至少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%相同性。

[0409] 在一个实施方案中,第一反义核糖核苷酸序列的至少20个连续核糖核苷酸与靶RNA分子的第一区域的互补体具有至少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%相同性。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分子的区域的互补体具有100%相同性,其余的核糖核苷酸与靶RNA分子的互补体具有至少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%相同性。在另一个实施方案中,第一反义核糖核苷酸序列的至少20个连续核糖核苷酸与靶RNA分子的第一区域具有100%相同性。

[0410] 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。

[0411] 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。

[0412] 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。

[0413] 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,且因此仅具有两个5'末端核苷酸和两个3'末端核苷酸。在另一个实施方案中,每个核糖核苷酸与两个其他核苷酸共价连接,即RNA分子是环状的并且具有自身互补区,且因此没有5'末端核苷酸和没有3'末端核苷酸。

[0414] 在一个实施方案中,RNA分子的双链区可以包含由有义RNA序列或反义RNA序列或两者中的未配对核苷酸产生的一个或多个隆起。在一个实施方案中,RNA分子包含一系列隆起。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。

[0415] 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。

[0416] 在另一个实施方案中,第一有义核糖核苷酸序列的20个连续核苷酸与第一5'核糖核苷酸共价连接,而没有任何间插核苷酸,并且第一反义核糖核苷酸序列的20个连续核苷酸与第一3'核糖核苷酸共价连接,而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且不留下未碱基配对的5'或3'端,或两者。

[0417] 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区,所述第一有义核糖核苷酸序列在序列上与靶RNA分子的第一区域基本相同。例如,连接核糖核苷酸序列可以作为该区和环之间的间隔区。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分子的第一区域基本相同,所述连接核糖核苷酸序列充当这些序列之间的间隔区。在一个实施方案中,在RNA分子中提供了在序列上与靶RNA分子的第一区域基本相同的至少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。

[0418] 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'前导序列由这样的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,例如通过的脂质基团(诸如胆固醇)、或维生素(诸如生物素),或多肽的连接。这样的修饰可以帮助将RNA分子摄入其中RNA将起作用的真核细胞。

[0419] 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碱基配对,或者除了1、2或3个核糖核苷酸之外全部碱基配对。

[0420] 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苷酸(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。

[0421] 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,第一3'核糖核苷酸和第二5'核糖核苷酸,或第二3'核糖核苷酸和第一5'核糖核苷酸,但并非两者,是共价连接的。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方案。在另一个实施方案中,各自的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。

[0422] 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,其在序列上分别与靶RNA分子或其互补体的区域基本相同。在一个实施方案中,每个发夹具有与相同靶RNA分子的区域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,每个发夹具有与相同靶RNA分子的不同区域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分子的区域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可以不相关的两个靶RNA分子的表达和/或活性。

[0423] 在RNA分子的双发夹结构的每个发夹中,按5'至3'顺序,每个发夹中有义和反义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连续的反义-有义-

有义-反义(图1A),或者是其中两个反义序列是连续的有义-反义-反义-有义(图1B)。

[0424] 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。

[0425] 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(pseudoknot)和四环(tetraloop)。

[0426] 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。

[0427] 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。在另一个较低优选的实施方案中,环包含200、500、700或1000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1000个核糖核苷酸的长度。例如,环可以具有600至1000个核糖核苷酸的长度。例如,环可以具有4至1000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。

[0428] 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。

[0429] 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个连续碱基对的末端,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,较低优选地,环可以在至少20、30、50、100、200、500或更多连续碱基对的末端。

[0430] 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与其完全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分子的区域相同。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,和与其完全碱基

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分子的区域相同。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。

[0431] 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相同靶RNA分子的不同区域相同。例如,所述序列可以与相同靶分子的至少2个、至少3个、至少4个、至少5个、至少6个区域相同。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,两个或更多个有义核糖核苷酸序列在序列上与相同靶RNA分子的相同区域相同。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个区域相同。

[0432] 在另一个实施方案中,两个或更多个有义核糖核苷酸序列没有间插环(间隔区)序列。

[0433] 在一个实施方案中,RNA分子可以包含核糖核苷酸的单链,所述核糖核苷酸的单链具有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个连续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。

[0434] 在另一个实施方案中,RNA组分的连续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“未连接的缺口”。错配和未连接的缺口可位于RNA分子的不同位置。对于实施方案,未连接的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的位置。

[0435] 在一个实施方案中,RNA分子由RNA的单链产生。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。可通过连接包含未连接的缺口的上述RNA分子(例如用RNA连接酶)来产生闭合分子。

[0436] 在另一个实施方案中,RNA分子包含5'-延伸序列或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。

[0437] 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。

[0438] 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:

[0439] -与第一5'核糖核苷酸共价连接的5'延伸序列;

[0440] -与第二3'核糖核苷酸共价连接的3'延伸序列;

[0441] -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;

[0442] -与第二5'核糖核苷酸共价连接的5'延伸序列;

[0443] -与第一3'核糖核苷酸共价连接的3'延伸序列;

[0444] -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。

[0445] 编码RNA分子的核酸

[0446] 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。

[0447] 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。

[0448] 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。

[0449] 在一个实施例中,本文公开的多核苷酸,其编码包含内含子的RNA前体分子,优选在5'延伸序列中或在至少一个环序列中,其中所述内含子能够在多核苷酸在宿主细胞或体外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。

[0450] 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。

[0451] 重组载体

[0452] 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。

[0453] 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。

[0454] 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。

[0455] 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连续,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连续或位于其转录增强的编码序列附近。

[0456] 当存在多个启动子时,每个启动子可以独立地相同或不同。

[0457] 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W0 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W091/02071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。

[0458] 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。

[0459] 表达载体

[0460] 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。

[0461] 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 (植物分子生物学的方法), 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。

[0462] 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。

[0463] 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引导表达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, EF1 α , SV40, PGK1, Ubc, 人 β 肌动蛋白, CAG, TRE, UAS, CaMKII α , 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基因复合体启动子、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。

[0464] 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-H³⁰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。

[0465] 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、Safeners(WO 97/06269),或其也可有利地使用(6)器官特异性启动子。

[0466] 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引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。

[0467] 在一个特别优选的实施方案中,启动子引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。

[0468] 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,启动子优先引导在例如果实的果髓的果实可食用部分中的表达。

[0469] 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)。

[0470] 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。

[0471] 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(*Atmyb5*),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。

[0472] 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)。

[0473] 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Kozziel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。

[0474] 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。

[0475] 转移核酸

[0476] 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。

[0477] 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。

[0478] 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。

[0479] 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。

[0480] 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*)。

[0481] 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)。

[0482] 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。

[0483] 重组细胞

[0484] 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。

[0485] 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。

[0486] 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*)。

[0487] 合适的宿主细胞的另一个实施例是电感受态HT115细胞。

[0488] 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。

[0489] 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。

[0490] 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(*Chlamydomonas 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。

[0491] 转基因植物

[0492] 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。

[0493] 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。

[0494] 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)。植物组织可以是原位转化物(in planta)、器官培养物、组织培养物或细胞培养物。

[0495] 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 $C_{16:0}$ 及其在 C_{18} 脂肪酸被排除C-2位置的去饱和产物。C-1位置被 C_{18} 脂肪酸占据并且在小程度上被 C_{16} 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。

[0496] 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。

[0497] “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。

[0498] 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。

[0499] 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。

[0500] 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。

[0501] 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(*Arachis 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*)、大麻(*Cannabis sativa*)、红花(*Carthamus tinctorius*)，巴西油桃木果(*Caryocar brasiliense*)，椰子(*Cocos nucifera*)，埃塞俄比亚两节芥(*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*) 或本氏烟草 (*Nicotiana benthamiana*), *Oenocarpus bacaba*, 巴套阿酒实棕 (*Oenocarpus bataua*), *Oenocarpus distichus*, 稻属 (*Oryza*), 诸如水稻 (*Oryza sativa*) 和稗稻 (*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 (*Sesamum indicum*), 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 cereale*), 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草 (*Nicotiana benthamiana*)。

[0502] 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*)、帚裂稗草 (*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。

[0503] 在优选的实施方案中, 植物是被子植物。

[0504] 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、罂粟、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。

[0505] 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。

[0506] 转化

[0507] 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。

[0508] 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。

[0509] 在一个实施方案中,可以通过将根据本发明的RNA分子局部应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RNA分子的组合物局部应用于植物或其部分。

[0510] 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,可以使用浸花(原位转化(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。

[0511] 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。

[0512] 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 5,932,479,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。

[0513] 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。

[0514] 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。

[0515] 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。

[0516] 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相同基因座上。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。

[0517] 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), Wilcox J. 美国农艺学会 (Breeding Methods for Cultivar Development), 美国威斯康星州麦迪逊分校 (1987)。

[0518] 制剂

[0519] 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。

[0520] 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。

[0521] 如本文所用，术语“转染促进剂”是指添加至RNA分子以增强进入细胞的摄取的组合物，所述细胞包括但不限于植物细胞、昆虫细胞或真菌细胞。可以使用本领域已知的适用于转染细胞的任何转染促进剂。实例包括阳离子脂质，例如以下的一种或多种：DOTMA (N-[1-(2,3-二油酰氧基)-丙基]-N,N,N-三甲基氯化铵)、DOTAP (1,2-双(油酰氧基)-3-)-3-(三甲基铵)丙烷)、DMRIE (1,2-二肉豆蔻基氧基丙基-3-二甲基-羟乙基溴化铵)、DDAB (二甲基二

十八烷基溴化铵)。脂精胺,特别是DOSPA(2,3-二醇氧基-N-[2(精胺甲酰胺基)乙基]-N,N-二甲基-1-丙胺-三氟乙酸盐)和DOSPER(1,3-二醇氧基-2-(6羧基精胺)-丙基-酰胺和二烷基-和四烷基-四甲基亚精胺,包括但不限于TMTPS(四甲基四棕榈酰亚精胺)、TMTOS(四甲基四油烯基亚精胺)、TMTLS(四甲基四月桂基亚精胺)、TMTMS(四甲基十四烷基亚精胺)和TMDOS(四甲基二油酰精胺)。阳离子脂质任选地与非阳离子脂质,特别是中性脂质结合,例如脂质,例如DOPE(二油酰基磷脂酰乙醇胺)、DPhPE(二植烷酰基磷脂酰乙醇胺)或胆固醇组合。合适的市售转染试剂的非限制性实例包括Lipofectamine(Life Technologies)和Lipofectamine 2000(Life Technologies)。

[0522] 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。

[0523] 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。

[0524] 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。

[0525] 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。

[0526] 将药物组合物配制成为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。

[0527] 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。

[0528] 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。

[0529] 防治非人生物体的方法

[0530] 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。

[0531] 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。

[0532] 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MpC002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*)。

[0533] 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。

[0534] 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。

[0535] 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。

[0536] 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(Gryllobatodea),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,襁翅目(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)。

[0537] 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 reclusa) (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(Blattea), 包括东方蜚蠊(Blattea 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(Cerambycidae) 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): 蠹蝽科的家族 (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(Calypttratae), 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), 蜜蜂总科(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);

[0538] 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、蠹蝽、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(Blattea) 物种(例如: 东方蜚蠊(Blattea

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 (*Iasius*) (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*) (*Reticulitermes hesperus*) (例如林木白蚁 (*Neotermes connexus*))。

[0539] 在一个实施方案中, 靶 RNA 编码昆虫乳酸合酶。

[0540] 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 或 MLO), 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。

[0541] 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。

[0542] 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 mellea*); 寡孢节丛孢 (*Armillaria mellea*); 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 (*Botrytis cinerea*); 蚕豆葡萄孢 (*Botrytis fabae*); 白色念珠菌 (*Candida albicans*); 麦角菌 (*Claviceps purpurea*); 松疱锈病菌 (*Cronartium ribicola*); 紫球菌 (*Epicoccum purpureascens*); 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 (*Phytophthora*

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*)。

[0543] 要治疗的示例性病征

[0544] 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, 泌尿尿道癌症, 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) 以及与癰疽病、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。

[0545] 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。

[0546] 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 CIPK1、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。

[0547] 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。

[0548] 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。

[0549] 实施例

[0550] 实施例1:材料和方法

[0551] 遗传构建体的合成

[0552] 为了设计典型的1edRNA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区域。在一个实施例中,序列的5'半和侧翼区的约130nt以及类似地侧翼区域的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。

[0553] 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。

[0554] RNA的合成

[0555] 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1edRNAi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)。

[0556] 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1edRNA重悬于退火缓冲液(25mM Tris-HCL, pH 8.0, 10mM MgCl₂)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)。

[0557] 通过将构建体导入大肠杆菌菌株HT115,实现1edRNA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1edRNA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。

[0558] 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17.7μl turbo DNase缓冲液和1μl turbo DNA,并在37℃温育10分钟以消化DNA。然后加入17.7μl Turbo 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产品的质量。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50 μ l 1 $\times$ RNAi退火缓冲液中。使用纳滴法测量RNA浓度并储存在-80 $^{\circ}$ C直至使用。

[0559] 实施例2:ledRNA的设计

[0560] 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具有相同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域的反义序列,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。

[0561] 在另一种但相关形式的ledRNA中,有义序列被分为两个区域,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'端的区域具有相同性的第一有义区,第一环序列,反义序列,第二环序列,然后是与靶RNA的5'端具有相同性的第二有义区,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。

[0562] 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。

[0563] 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域: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GFP CDS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GFP CDS的核苷酸731-592的第二反义环,对应于GFP CDS的核苷酸591-359的反义序列的后半部分。

[0564] ledGUS构建体依次包含以下区域: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。

[0565] 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。

[0566] 实施例3:ledRNA的稳定性

[0567] 将ledRNA形成dsRNA结构与末端开放的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 MgCl₂/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。

[0568] 如图2所示,GUS ledRNA和GFP ledRNA都给出双链分子预期的移动性的显性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。

[0569] 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。

[0570] 实施例4:通过局部递送测试ledRNA

[0571] 在表达GFP或GUS报告基因的本氏烟草(*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。

[0572] 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 MgCl₂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。

[0573] 针对本氏烟草(*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)。

[0574] 实施例5. ledRNA诱导的内源靶基因沉默

[0575] 在另一个实施例中,设计了ledRNA以靶向由内源基因即本氏烟草(*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 (本氏烟草(*N. benthamiana*) FAD2.1 ledRNA)。

[0576] 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。

[0577] 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FAD2.1 mRNA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应用的ledFAD2.1 RNA对FAD2.1基因的沉默(图6)。

[0578] 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,FAD2.1 mRNA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)。

[0579] 由于FAD2.1和FAD2.2编码脂肪酸 Δ 12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。

[0580] 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

[0581] 靶向GUS RNA的修饰的发夹RNA

[0582] 如编码 β -葡萄糖醛酸糖苷酶(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个连续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区域。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的互补体(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。

[0583] 遗传构建体的制备

[0584] 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。

[0585] 仅具有规范碱基对的构建体hpGUS[wt]

[0586] 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。

[0587] 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个连续碱基对的双链区。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。

[0588] 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。

[0589] 包含G:U碱基对的构建体hpGUS[G:U]

[0590] 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。

[0591] 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。双链区连续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。

[0592] 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)的相应区域的比对。

[0593] 包含每第四个核苷酸错配的核苷酸的构建体hpGUS[1:4]

[0594] 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。

[0595] 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。

[0596] 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) 的相应区域的比对。

[0597] 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

[0598] 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。

[0599] 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。

[0600] 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)的相应区域的比对。

[0601] 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。

[0602] 实施例7:检测转基因植物中修饰的发夹RNA

[0603] 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。

[0604] 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PPGH11和PPGH24植物相同。总体上,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。

[0605] 使用荧光4-甲基伞形酮基β-D-葡萄糖苷酸(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。

[0606] 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。

[0607] 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)。

[0608] 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。

[0609] 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)。

[0610] 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。

[0611] 转化植物的Southern印迹杂交实验

[0612] 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10 μ g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 $^{\circ}$ C下与来自OCS-T终止子区的 ^{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。

[0613] 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。

[0614] 转化植物的Northern印迹杂交实验

[0615] 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。

[0616] 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条带不如反义条带清晰。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。

[0617] 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。

[0618] 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。

[0619] 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。

[0620] 实施例8:表达修饰发夹RNA的转基因植物sRNA分析

[0621] 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中条带的移动性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。

[0622] 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。

[0623] 实施例9:植物转基因的DNA甲基化分析

[0624] 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。

[0625] 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G^mC。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。

[0626] 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30 μ l反应体积中消化大约1 μ g基因组DNA过夜。将消化的DNA样品稀释至100 μ l, 并如下进行PCR扩增所关注的区域。

[0627] 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'-CARRAACTRTTCRCCTTCAC-3' (SEQ ID NO:61)。对于hpGUS[G:U]的35S-GUS连接序列: 正向引物(GUSgu-R2), 5'-CAAAACTATTACCCCTTCAC-3' (SEQ ID NO:62); 反向引物(GUS4m-R2), CACRAARTTACRCRCTTRAC (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 $^{\circ}$ C 1分钟, 以下进行35个循环: 94 $^{\circ}$ C 30秒, 55 $^{\circ}$ C 退火45秒, 68 $^{\circ}$ C 延伸1分钟, 和最终在68 $^{\circ}$ C 延伸5分钟。对PCR扩增产物进行电泳并定量条带的强度。

[0628] 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”。

[0629] 关于实施例6-9的一般讨论

[0630] 转基因中反向重复DNA结构的破坏增强其稳定性

[0631] 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连续错配, 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。

[0632] hpRNA的热力学稳定性是影响靶基因沉默程度的重要因素

[0633] 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。

[0634] Dicer能高效地加工hpGUS[G:U] RNA

[0635] 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。

[0636] G:U和1:4hpRNA转基因显示近端35S启动子区的DNA甲基化降低

[0637] 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。

[0638] 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

[0639] 修饰的发夹RNA靶向EIN2和CHS RNA

[0640] 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

—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。

[0641] 构建体制备

[0642] 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-0 cDNA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。

[0643] 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。

[0644] 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。

[0645] 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。

[0646] 植物转化

[0647] 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。

[0648] 表1.Co1-0背景中获得的转基因植物的汇总

[0649]	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
	hpCHS::EIN2[G:U]	26
	hpCHS::EIN2[G:U/U:G]	20

[0650] EIN2沉默程度分析

[0651] 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。

[0652] 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。

[0653] 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。

[0654] 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。

[0655] 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体的比对显示, 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。

[0656] CHS沉默程度分析

[0657] 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'-GATCCCATTTCATAAAACCCCAG-3' (SEQ ID NO:81)。

[0658] 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)。

[0659] 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。

[0660] 降低拟南芥(*Arabidopsis thaliana*)中PDS基因的表达

[0661] 选择另一个拟南芥(*Arabidopsis*)基因作为示例性靶基因, 即八氢番茄红素去饱和酶(PDS)基因, 其编码八氢番茄红素去饱和酶, 其在类胡萝卜素生物合成过程中催化八氢番茄红素脱饱和为 ζ -胡萝卜素。预期PDS沉默会导致拟南芥植物的光漂白, 这在视觉上容易观察到。因此, 与靶向450个核苷酸的PDS mRNA序列的传统hpRNA构建体相比, 制备并测试了G:U修饰的hpRNA构建体。450个核苷酸的PDS序列包含82个胞嘧啶(C), 它们被胸苷(T)取代,

导致hpRNA hpPDS[G:U]的dsRNA区中的18.2%的碱基对为G:U碱基对。利用农杆菌介导的转化将编码hpPDS[G:U]的遗传构建体和编码hpPDS[WT]的对照遗传构建体引入拟南芥Col-0生态型。

[0662] 对于hpPDS[WT]和hpPDS[G:U]构建体,分别鉴定了100和172个转基因品系。令人惊讶的是,所有这些品系均显示出在卡那霉素抗性选择性培养基上出现的年轻T1幼苗的子叶中的光漂白,在植物生长的早期阶段,两个转基因群体之间没有明显差异。这些表明两个构建体在诱导子叶中的PDS沉默方面同样有效。但是,一些T1植物发育的真叶不再是光漂白的,而是看起来呈绿色或淡绿色,表明PDS沉默在真叶中释放或减弱了。hpPDS[WT]群体的显示绿色真叶转基因品系所占比例比hpPDS[G:U]群体的高得多。根据强PDS沉默(整个植物中强烈的光漂白)、中度PDS沉默(淡绿色或斑驳的叶子)和弱PDS沉默(完全绿色或弱斑驳的叶子),将转基因植物分为三个不同的类。hpPDS[WT]品系具有弱PDS沉默的植物比例为43%,相比之下hpPDS[G:U]品系为7%。实际上,弱沉默组的所有hpPDS[G:U]品系在真叶上仍表现出轻度斑驳,与之相比,弱沉默的hpPDS[WT]植物大多具有完全绿叶。这些结果表明,G:U修饰的hpRNA构建体在跨独立的转基因群体中提供了比常规(完全规范碱基配对的)hpPDS构建体更加均匀的PDS沉默,这与上述GUS和EIN2沉默测定的结果一致。更重要的是,PDS沉默结果表明以前未曾发现过hpRNA转基因诱导的基因沉默在植物中的发育变异性,并表明hpRNA转基因沉默在子叶中比在真叶中更有效和稳定。根据跨独立品系的均匀基因沉默,PDS沉默结果表明,G:U修饰的hpRNA转基因在发育上比常规hpRNA构建体更稳定,提供了更稳定和持久的沉默。

[0663] 实施例11:分析来自发夹RNA构建体的sRNA

[0664] 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³²P-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]的主要两个条带的移动差异是十分明显的。通过比较图28的相邻泳道10和11中条带的移动性可以最好地看出这一点。

[0665] 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。

[0666] 图29显示,传统的(完全规范碱基配对的)和G:U修饰的hpRNA品系均积累了两个主要的siRNA大小片段。与以前的报道一致,传统hpRNA品系中的主要siRNA与21nt和24nt sRNA大小标记物的迁移相似。但是,来自两个G:U修饰的转基因的两条主要siRNA条带在凝胶上的迁移略快,这表明它们的大小比传统hpRNA转基因的小,或者其末端化学修饰与传统的hpRNA转基因不同。

[0667] 为了研究两种不同类型构建体之间siRNA的大小分布是否可能不同,从一个hpGUS [WT]品系和hpGUS[G:U]的各自两个品系, hpEIN2[WT]和hpEIN2[G:U]中分离出小RNA,并使用Illumina平台进行测序,每个样品大约产生1600万条sRNA读段。还对来自两个高度沉默的hpGUS[1:4]品系的样品进行了测序。确定了映射到发夹RNA的双链区和内含子间隔区的sRNA的数目。还将siRNA映射到靶GUS mRNA和EIN2 mRNA的上游和下游区域,以检测传递性siRNA。测序数据证实, hpGUS[G:U]品系与hpGUS[WT]品系相似,生成大量的siRNA,而hpGUS[1:4]品系也生成siRNA,但丰度要低得多。hpGUS[1:4]品系中较低水平的siRNA与hpGUS[1:4]对GUS沉默的效率相对低相一致,且这表明相对于传统的发夹, dsRNA茎在hpGUS[1:4]RNA中较低的热力学稳定性降低了Dicer的处理效率。尽管在Northern印迹中显示的反义siRNA的移动性发生了明显变化,但在传统的和错配的hpRNA品系之间, siRNA的大小分布没有明显差异,所有样品均显示21-nt sRNA是主要的大小类别。传统和错配的hpGUS品系之间22nt反义siRNA的比例丰度存在细微的差异: hpGUS[G:U]和hpGUS[1:4]品系显示22nt大小类别的比例高于hpGUS[WT]品系。传统和错配hpRNA品系的测序数据的显著特征是,在所有样品中, 24-nt siRNA的丰度均比21-nt siRNA低得多,即有义24-nt的siRNA少约3-21倍,而反义24-nt siRNA少约4-35倍。这与Northern印迹结果显著不同, Northern印迹结果显示两种主要大小类别的量相对相等。有趣的是, hpEIN2[WT]-7和hpEIN2[G:U]-14/15样品在Northern印迹上显示相似的反义siRNA丰度,但是在测序数据中,与hpEIN2[WT]-7品系(134112个读段)相比, hpEIN2[G:U]品系20-24nt反义siRNA总数(17290和29211)要少得多。

[0668] 对于hpGUS[G:U]和hpEIN2[G:U]品系,几乎所有有义siRNA均与hpRNA的G:U修饰的有义序列匹配,而大多数反义siRNA均具有野生型反义序列。这表明这些有义和反义siRNA的绝大部分直接从初始hpRNA[G:U]转录物加工,而不是由于RDR介导的hpRNA或靶RNA转录物的扩增,否则会产生相同模板序列的有义和反义siRNA。与此相一致,从hpRNA转基因的环区(PDK内含子)或GUS或EIN2 mRNA的非靶向的下游区域中仅检测到少量的20-24nt sRNA读段(传递性siRNA)。但是,两个hpGUS[1:4]品系均显示出相对高比例的野生型有义siRNA,这表明这两个品系中的强GUS沉默(对于hpGUS[1:4]群体而言相对罕见)可能涉及RDR扩增。实际上,在hpGUS[1:4]品系中,从hpRNA靶区域下游的靶基因序列中检测到的siRNA量比从dsRNA茎中检测到的siRNA量高,表明这些品系中存在传递性沉默。

[0669] 综上, sRNA测序数据表明,传统的和错配的hpRNA品系的siRNA具有相似的大小分布,除了22-nt大小类别,这表明Northern印迹检测到的差异迁移是由于不同的5'或3'化学修饰。Northern印迹结果与测序数据之间的相对sRNA丰度差异(例如hpEIN2[WT]与hpEIN2[G:U]衍生的siRNA和21-nt与24-nt的差异)表明sRNA文库制备期间,不同的siRNA群体和大小类别可能具有不同的克隆效率。

[0670] 已知植物sRNA在3'末端核苷酸处具有2'-O-甲基,其被认为可以稳定sRNA。先前显示3'甲基化可抑制但不能阻止降低sRNA克隆效率的3'衔接子(adaptor)连接(Ebhardt等, 2005)。因此,在 β -消除试验中, hpRNA[WT]和hpRNA[G:U]衍生的siRNA与高碘酸钠一起使用。该处理不会导致hpRNA[WT]和hpRNA[G:U]衍生的siRNA的凝胶移动性发生变化,表明这两个siRNA群体在3'末端均被甲基化,并且hpRNA[WT]和hpRNA[G:U]衍生的siRNA之间3'化学修饰没有差异。

[0671] 标准的sRNA测序方案基于具有5'单磷酸(允许5'衔接子连接)的sRNA(Lau等,

2001)。假定通过Dicer-处理的sRNA具有5'单磷酸,但在秀丽隐杆线虫(*C.elegans*)中,发现许多siRNA在5'末端具有二磷酸或三磷酸,这改变了sRNA的凝胶移动性并阻止了在标准sRNA克隆程序中sRNA的5'衔接子连接(Pak和Fire 2007)。尚不清楚植物sRNA是否也具有差异的5'磷酸化。hpRNA[WT]和hpRNA[G:U]衍生的siRNA的5'磷酸化状态因此通过用碱性磷酸酶处理总RNA,然后进行Northern印迹杂交来检查。这种处理降低了所有hpRNA衍生的sRNA的凝胶移动性,表明存在5'磷酸化。但是,在磷酸酶处理后,hpRNA[G:U]衍生的siRNA表现出比hpRNA[WT]衍生的siRNA更大的移动性改变,这导致两组去磷酸化的siRNA在凝胶上的相同位置迁移。使用多核苷酸激酶反应用³²P放射性标记21和24-nt sRNA大小标记物的5'端,并因此应具有单磷酸化的5'末端。这表明,与大小标记物在相同位置迁移的hpRNA[WT]衍生的siRNA可能是单磷酸化的siRNA,而hpRNA[G:U]衍生的siRNA迁移较快,其在5'末端具有多于一个磷酸。因此,得出的结论是,植物细胞中由传统的和G:U修饰的hpRNA转基因产生的siRNA具有不同的磷酸化。

[0672] 实施例12 EIN2沉默植物的DNA甲基化分析

[0673] 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'-AGAAATYTTYGTAAAYATGGTGG-3'(SEQ ID NO:82),反向引物(Top-35S-RyGTAAAYATGGTGG-3')(SEQ ID NO:82),反向引物(Top-35S-R2),5'-TCATRRARATRTCACATCAATCC-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。

[0674] 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。

[0675] 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,如在使用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。

[0676] 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。

[0677] 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(5mC)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。

[0678] 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。

[0679] 实施例13:hpGUS[1:4]构建体的启动子中的DNA甲基化水平

[0680] 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。

[0681] 关于实施例10-13的一般讨论

[0682] 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

[0683] 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连续胞嘧啶的组中,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。

[0684] hpEIN2[G:U]品系表达更均匀水平的siRNA

[0685] 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。

[0686] 融合构建体也提供基因沉默

[0687] 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。

[0688] 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。

[0689] 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

[0690] 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引导的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引导的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。

[0691] 实施例14: 用于降低另一内源基因表达的修饰的发夹

[0692] 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。

[0693] 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。

[0694] 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*)中的靶区域选择SEQ ID NO:31的核苷酸675-1174(500个核苷酸)。欧洲油菜(*B. napus*)中的靶区域选择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连续规范碱基配对的最长延伸为17个碱基对, 第二长为16个连续碱基对。

[0695] 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。

[0696] 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*) 基因的两个非连续靶区域选择: 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连续规范碱基配对的最长延伸为20个碱基对, 第二长为15个连续碱基对。

[0697] 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个连续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个连续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。双链区连续规范碱基配对的最长延伸为15个碱基对, 第二长为13个连续碱基对。

[0698] 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个或更多个连续规范碱基对的双链延伸。

[0699] 实施例15: 用于降低动物细胞中基因表达的修饰发夹

[0700] 为了在动物细胞中测试G:U碱基配对形式、1edRNA形式或这两种修饰的组合物修饰的沉默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个核苷酸的靶区域, 对应于SEQ ID NO:40的核苷酸131-591。

[0701] 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T7RNA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。

[0702] 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编码区的靶区域为核苷酸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。

[0703] 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靶区域的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结构被包裹在较大的RNA转录物中,该转录物包括5'上游区和3'下游区。

[0704] 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 $162/460=35.2\%$ 。

[0705] 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MEM α 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°C,5%CO₂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1 μ g/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1 μ l的Lipofectamine 2000混合,两者均在50 μ l OPTI-MEM(美国英杰)中稀释,并在室温下孵育20分钟。然后将复合体加入细胞并孵育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200 μ L 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。

[0706] 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'-gcaagcugaccugaaguuc-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'-TTTAGTATATGTGCTGCCG-3'

(SEQ ID NO:88), 反向引物为5'-ctcgagttccaaaaagctgaccctgaagttcatctctcttgaagatgaacttcagggtcagccaaacaaggcttttctccaa-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'-TTTGTATATATGTGCTGCCG-3' (SEQ ID NO:90), 反向引物是5'-ctcgagttccaaaaaataagtcgcagcagtcacaatctcttgaattgtactgctgcgacttatgaataccgcttcctcctgag-3' (SEQ ID NO:91)。

[0707] 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个连续碱基对和21个连续碱基对的双链区, 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。

[0708] 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)。

[0709] 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诱导的沉默复合体(RISC)效应蛋白中的较长dsRNA的加工。

[0710] 实施例16: 植物中靶向DDM1和FANCM基因的RNA构建体

[0711] 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。

[0712] 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)。

[0713] 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I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)。

[0714] 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)。

[0715] 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,因此总体上具有四个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。

[0716] 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基因的同源区的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。

[0717] 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。

[0718] 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 tumefaciens*) 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。

[0719] 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。

[0720] 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。

[0721] 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。

[0722] 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基因的相同区域的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。

[0723] 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基因的相同区域的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。

[0724] 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 tumefaciens*) 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。

[0725] 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。

[0726] 实施例17:用于真菌疾病抗性的RNA构建体

[0727] 针对大麦和小麦的Mlo基因的LedRNA

[0728] 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。该病主要通过减少头状花序(head)数以及减少籽粒(kernel)大小和重量来影响谷粒产量。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。

[0729] 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。显著地,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)。

[0730] 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具有大约95-97%和98%相同性。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相同性程度并靶向具有高度序列保守性的基因区的ledRNA构建体。

[0731] 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(split sequence)为反义序列且连续序列为有义序列(图1A)。

[0732] 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连续的序列,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。

[0733] 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10 μ g 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,TaMlo mRNA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域是特异性的-在叶片的远端未处理部分中TaMlo mRNA水平没有降低。

[0734] 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。

[0735] 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。

[0736] 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。在目前的制剂中,以低至200ng/ μ l(每片叶子总共2 μ g)的浓度应用RNA显示显著抑制白粉病的损伤,这表明抑制性RNA的量可以大大减少,而同时仍提供对真菌的生长和发育的抑制。此外,在ledRNA处理后1、2、4、7和14天接种叶片以观察保护作用保持多长时间。从处理后第24小时的第一个时间点到处理后第14天的最后一个时间点的整个时间过程中均观察到内源基因有效沉默,此时内源基因仍显示91%的表达降低。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。

[0737] 靶向葡萄(Vitis vinifera)的VvMLO基因的LedRNA

[0738] 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。显著地,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。

[0739] 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VvMLO3、VvMLO4和VvMLO17的ledRNA构建体(根据Feechan et al., Functional Plant Biology, 2008, 35:1255-1266命名)。对于第一个,例如,选择VvMLO3靶基因的860个核苷酸的序列,其对应于SEQ ID NO:138的核苷酸297-1156。制备编码靶向VvMLO3、VvMLO4和VvMLO17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连续序列为有义序列(图1A)。各ledRNA的dsRNA区长为600bp;dsRNA区中的有义序列是不间断的连续的序列,例如对应于SEQ ID NO:138的核苷酸427-1156。编码一种ledRNA的核苷酸序列在本文中提供为SEQ ID NO:139。

[0740] 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 vinifera*) 植物Cabernet Sauvignon品种的叶。随后,将白粉病菌的孢子施用于叶片的处理和未处理区。使用定量RT-PCR观察到靶mRNA的水平降低。随时间跟踪疾病进展。通过以1 μ g/ml应用靶向VvMLO3、VvMLO4或VvMLO11的ledRNA溶液观察到VvMLO4显著下调。

[0741] 靶向真菌基因的LedRNA

[0742] 针对真菌病原体Rhizoctonia solani的Cyp51基因的编码区设计了LedRNA构建体,该基因是麦角固醇的合成以及真菌的存活和生长所必需的。使用上述ledRNA的设计原理制备了遗传构建体,其中分裂序列为反义序列且连续序列为有义序列(图1A)。设计了单一的ledRNA构建体,以靶向来自R. solani的两个基因,其中的ledRNA的dsRNA区域包含来自每个基因的350bp;dsRNA区中的有义序列是不间断的、连续的序列,例如对应于SEQ ID NO:140的核苷酸884-1233和SEQ ID NO:141的核苷酸174-523。本文提供的编码ledRNA之一的核苷酸序列为SEQ ID NO:142。通过体外转录制备了ledRNA,并用R. solani菌丝体接种物将其以每100 μ l培养物5 μ g的浓度应用于培养基。通过读取600nm处的培养物的光密度,在零时和接下来一周的每一天测量真菌的生长。在含有ledRsCyp51的培养物中,R. solani的生长明显低于含有RNA缓冲液或对照ledGFP(其中没有在R. solani中的相应靶)的对照培养物。

[0743] 还设计了ledRNA编码构建体,并制备了针对Phytophthora cinnamomi分离株94.48中的CesA3纤维素合酶基因的编码区。使用上述ledRNA的设计原理制备了遗传构建体,其中分裂序列为反义序列且连续序列为有义序列(图1A)。设计了ledRNA构建体,以靶向Phytophthora cinnamomi的CesA3基因,其中的ledRNA的dsRNA区域包含来自基因编码区的500bp;dsRNA区中的有义序列是不间断的、连续的序列,例如对应于SEQ ID NO:143的核苷酸884-1233。本文提供的编码ledRNA之一的核苷酸序列为SEQ ID NO:144。ledRNA在体外被转录,并以每100 μ l培养物3 μ g的比例应用于培养基。与虚假处理的(仅RNA缓冲液)或ledGFP处理的培养物相比,在用靶向PcCesA3的ledRNA处理的培养物中观察到方向性菌丝体生长的大量损失。方向性生长的损失以及由此产生的无定形球状生长模式使联想到具有破坏的细胞壁生物合成的细胞,并因此与PcCesA3基因的沉默相一致。

[0744] 实施例18:靶向植物中其他基因的RNA构建体

[0745] 靶向拟南芥 (*A.thaliana*) 和本氏烟草 (*N.benthamiana*) Tor基因的LedRNA

[0746] 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。

[0747] 使用分离序列为有义序列且连续序列为反义序列的ledRNA的设计原理,设计并制备靶向本氏烟草 (*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)。靶区域长度为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连续的序列,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样品施涂到本氏烟草 (*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。

[0748] 靶向大麦 (*H.vulgare*) ALS基因的LedRNA

[0749] 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。

[0750] 为了测试ledRNA是否能够通过外源递送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,连续序列是反义序列 (图1B)。靶区域长度为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连续的序列,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。

[0751] 用限制性内切酶MlyI消化编码ledRNA的遗传构建体,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2 μ g。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 (*Hordeum 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%。

[0752] 靶向小麦和大麦的NCED1和NCED2基因的ledRNA

[0753] 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。

[0754] 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。

[0755] 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具有约97%相同性。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连续序列为有义序列(图1A)。靶区域长度为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连续的序列,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。

[0756] 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。靶区域长度为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连续的序列,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。

[0757] 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。

[0758] 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。

[0759] 靶向拟南芥(A.thaliana) EIN2基因的ledRNA

[0760] 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。

[0761] 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。

[0762] 靶向拟南芥 (*A.thaliana*) CHS基因的ledRNA

[0763] 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。

[0764] 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。

[0765] 靶向狭叶羽扇豆 (*Lupinus angustifolius*) 的LanR基因的ledRNA

[0766] 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。

[0767] 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连续序列为有义序列 (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。

[0768] 靶向VRN2基因的LedRNAi赋予小麦春化反应性

[0769] 将鉴定为小麦ZCCT1基因 (Genbank登录号AAS58481.1) 的同系物的在TGACv1_scaffold_374416_5AL、TGACv1_scaffold_320642_4BL和TGACv1_scaffold_342601_4DL中的小麦VRN2A、VRN2B和VRN2D候选基因鉴定为设计ledRNAi构建体的靶标。VRN2B基因的309bp区域用于命名为LedTaVRN2的ledRNAi构建体的dsRNA区域。使用T7 RNA聚合酶通过体外转录产生Led RNAi,并在水中稀释。该溶液用于浸渗小麦谷粒,用于在4℃下萌芽3天。使用春化敏感小麦品种CSIRO W7的种子。将处理过的种子种植在土壤中,并随着时间的推移观察所得植物从营养生长到花发育的转变。记录开花时间,其以从花序鞘中伸出谷穗作为指示,并记录开花时主茎上的叶数。与仅使用缓冲液或非特异性dsRNA对照孵育的种子衍生的植物相比,用LedTaVRN2孵育的种子衍生的植物平均提前至少17天开花。此外,在开花时,用LedTaVRN2孵育的种子衍生的植物平均在主茎上的叶数减少了2.3,这表明专用于产生叶片的节减少,且更多的节专用于花/谷粒。

[0770] 实施例19:靶向昆虫基因的RNA构建体

[0771] 引言

[0772] 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的使用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。

[0773] 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含有dsRNA的人工膳食上饲养蚜虫来递送。

[0774] 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。首先选择2个蚜虫基因作为下调的靶基因,1个编码关键效应蛋白(C002),第二个编码活化蛋白激酶C受体(Rack-1)。C002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)。MpC002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)。

[0775] 材料和方法:蚜虫培养和植物材料

[0776] 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。

[0777] 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100 μ l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。

[0778] 在一个实验中使用的烟叶和萝卜叶收集自22℃在16小时光照/8小时黑暗循环下生长在土壤中的植物。在涉及萝卜叶切开的实验中,切下了附着在茎的片段(长约2cm)上的小萝卜叶(2-4cm²)。为了保持叶片新鲜,将茎插入直径为5cm的皮氏培养皿中的培养基中,该培养基每100毫升水中包含1.5克Bacto琼脂和1.16克Aquasol。用细画笔将蚜虫转移到叶子上。将具有叶和蚜虫的培养皿在20℃以16小时光照/8小时黑暗循环下保持在生长柜中。

[0779] 使用标准方法,通过包含一种或多种T7启动子和T7 RNA聚合酶的DNA模板的体外RNA转录制备双链RNA(dsRNA)。

[0780] MpC002和MpRack-1基因和LedRNA构建体

[0781] 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1edRNA构建体。编码1edRNA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1edRNA。将编码靶向MpC002和MpRack-1基因的1edRNA分子的载体DNA导入大肠杆菌菌株DH5 α ,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。

[0782] 1edRNA分子对减少蚜虫性能的效力

[0783] 为了检查靶向MpC002或MpRack-1基因的1edRNA是否影响蚜虫性能,将每种1edRNA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1edRNA,即1edGFP。

[0784] 在每种1edRNA分子的50ng/ μ l的较低浓度下,从含有MpC002或MpRack-1edRNA的人工饮食饲喂后的蚜虫存活与对照1edGFP没有显著差异。然而,靶向MpC002基因的1edRNA显著($P<0.05$)降低了桃蚜的繁殖率(图37)。与用具有对照1edRNA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/ μ l的较高浓度下,靶向MpC002或MpRack-1的1edRNA增加了成蚜死亡率(图37B)。在24小时后,还观察到包括MpC002或MpRack-1edRNAs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1edRNA的使用能引起蚜虫的死亡,减少蚜虫的繁殖。将各1edRNA的效力与靶向靶基因的相同区的双链RNA分子(dsRNAi)进行比较,所述dsRNAi包含分开但退火的有义和反义RNA链。

[0785] 蚜虫对1edRNA分子的吸收

[0786] 为了跟踪蚜虫内1edRNA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1edRNA。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1edRNA的蚜虫。靶向MpC002或MpRack-1的Cyr3标记的1edRNA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1edRNA分子的有效靶标。

[0787] 1edRNA稳定性

[0788] 为了检测饮食中和从饲喂蚜虫回收的1edRNA的稳定性,在饲喂含有标记的1edRNA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,1edMpC002 RNA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1edRNA降解,但是大部分1edRNA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。1edRNA分子的这种程度的稳定性应当允许1edRNA具有活性并且当外源施用保持活性。

[0789] 植物叶片对标记ledRNA的吸收

[0790] 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 $1\mu\text{g}/\mu\text{l}$) 涂在直径为2cm的圆圈中,并用黑色记号笔标记施涂区。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,经局部应用,在植物组织中发现了ledRNA分子。

[0791] 蚜虫对局部LedRNA的摄取

[0792] 将Cy3标记的ledGFP RNA涂在萝卜叶上,以查看蚜虫是否能够从植物摄取局部应用的ledRNA。将十微升每种Cy3标记的ledGFP (浓度为 $10\mu\text{g}/\mu\text{l}$) 涂在小的切下的萝卜叶上($\sim 2\text{cm}^2$)。对照叶涂有等量的未标记的ledGFP。标记的萝卜叶和对照萝卜叶均被八种不同发育阶段的蚜虫侵染。使用上述针对烟叶的方法捕获叶和蚜虫荧光的图像。在对照叶片和蚜虫中没有可检测到的荧光,而涂有Cy3标记的ledGFP的叶片却是高度荧光的。在具有Cy3标记的ledRNA的叶片上喂食后24小时内,蚜虫在整个身体中均显示出强荧光,但在肠道和腿部比其他身体部位更明显。实验表明,蚜虫能够从局部应用的植物中摄取ledRNA分子。

[0793] 筛选其他蚜虫RNAi靶基因

[0794] 为了鉴定更多的蚜虫靶基因,总共对16种蚜虫基因评估了其作为RNAi靶的适用性。选择的候选基因涉及蚜虫的发育、繁殖、喂食或解毒。通过包含与靶基因mRNA的区相对应的有义和反义序列来靶向每个基因的常规dsRNA(dsRNAi)以每微升饮食2微克RNA的浓度补充到蚜虫人工饮食中。使用对蚜虫存活和繁殖率的影响来确定蚜虫RNAi靶基因的适用性。在研究的16个基因中,有9个基因显示了蚜虫存活和/或繁殖率的降低。除MpC002和MpRack-1外,其他合适的靶基因是编码以下多肽的基因及其在蚜虫中具有的功能类型为以下:微管蛋白(登录号XM_022321900.1,细胞结构),胰岛素相关肽(XM_022313196.1,胚胎发育),V型ATPase E亚基(XM_022312248.1,能量代谢),缺口后凸(XM_022313819.1,生长和发育),蜕皮触发激素(XM_022323100.1,发育-蜕皮),短神经肽F(XM_022314068.1,神经系统)和白细胞激肽(XM_022308286.1,水平衡和食物摄入)。对于大多数基因而言,RNAi对蚜虫繁殖的影响似乎比对存活的影响更稳健和更强烈,即对繁殖的影响更大。

[0795] 外源RNAi对蚜虫的跨世代(Trans-generation)作用

[0796] 为了检查RNAi的作用能持续多长时间,在两龄或三龄发育阶段的蚜虫上喂食补充了靶向MpC002、MpRack-1、MpGhb或补充了对照dsGFP的dsRNAi的人工饲料,持续10天。然后将存活下来的蚜虫转移到切下的萝卜叶上,而不应用RNA。对于所有三个基因,高达6天,每只存活的蚜虫产生的若虫数量显著低于以对照dsGFP RNA分子或水喂食的蚜虫的数量。对于MpC002和MpRack-1dsRNA,萝卜叶上的较低繁殖率至少维持9天。为了研究dsRNAi是否影响后代,将在萝卜叶上三天内出生且未直接以含RNA的饮食喂食的蚜虫移到新鲜切下的萝卜叶上,并监测其存活和产生率持续15天。虽然存活率没有显著差异,但是与出生自喂食具有对照dsGFP或水的饮食的母体蚜虫的蚜虫相比,出生自喂食具有MpC002、MpRack-1或MpGh

dsRNA的饮食的母体蚜虫的蚜虫均产生的蚜虫数量要显著更少。结论是,将dsRNA分子喂给亲本蚜虫引起的影响在后代蚜虫中持续存在。

[0797] 结论

[0798] 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以及鉴定合适的靶基因,蚜虫是吸吮树液的害虫的主要组别,是世界范围的问题。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/ μ l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/ μ l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。正如传统dsRNA的结果所示,ledRNA对蚜虫繁殖的影响可能至少持续两代。

[0799] 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,并且在蚜虫中检测到。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。此外,基于通过减少编码这些dsRNA分子的转基因的自沉默来增加这些dsRNA分子的积累水平,使用在分子的dsRNA区中包含10-30%G:U碱基对的ledRNA[G:U]或发夹[G:U]RNA预期可以提供更好的防治。

[0800] 实施例20:靶向其他昆虫基因的RNA构建体

[0801] 靶向昆虫基因的LedRNA

[0802] 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。

[0803] 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。

[0804] 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,连续序列是反义序列(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连续的序列,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。

[0805] 用限制性内切酶SnaBI消化编码ledRNA的遗传构建体,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。

[0806] 制备在双链茎中具有G:U碱基对的相应的ledRNA构建体,并以规范碱基配对的ledRNA进行比较。

[0807] 靶向蚂蚁中基因的ledRNA

[0808] 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。

[0809] 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,连续序列是反义序列(图1B)。靶区域长度为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连续的序列,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。

[0810] 用限制性内切酶SnaBI消化编码ledRNA的遗传构建体,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*)。

[0811] 靶向铜绿蝇(*L.cuprina*)基因的LedRNA

[0812] 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。属于蚋蝇科(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,连续序列是反义序列(图1B)。在每种情况下,靶区域长度为约600个核苷酸,dsRNA区中的反义序列是不间断的连续序列。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。

[0813] 实施例21.转基因衍生的ledRNA在稳定转化的植物中以高水平积累

[0814] 合成了编码靶向GUS报告基因或拟南芥EIN2基因的mRNA的1edRNA序列的DNA片段,并将其克隆到pART7中以形成p35S:1edRNA:0cs3'聚腺苷酸化区域/终止子表达盒,用于在植物细胞中表达。然后将片段用NotI切下,并插入pART27的NotI位点,以形成用于植物转化的1edGUS和1edEIN2载体。将1edGUS构建体和设计用于生成具有563bp dsRNA茎和1113nt环的长hpRNA的现有hpGUS构建体,分别通过农杆菌介导的转化方法转化进表达GUS的烟草品系PPGH24。来自显示出强GUS沉默或GUS活性几乎没有或没有明显降低的独立转化体的RNA样品用于Northern印迹杂交测定法来检测转基因编码的hpGUS或1edGUS RNA。如图38所示,从1edGUS转化的植物中检测到的杂交信号比从显示出强GUS沉默的hpGUS转化的植物中检测到的杂交信号要强得多(在图38中用“-”表示)。实际上, hpGUS RNA样品的大多数杂交信号都是非特异性背景信号,从对照、未转化植物(WT)的RNA中也观察到了非特异性背景信号。对于1edGUS品系观察到了几个强烈的杂交带,大概是由于全长1edRNA的一些部分加工所致。

[0815] 编码1edGUS的遗传构建体的核苷酸序列示于SEQ ID NO:5。核苷酸1-17对应于用于体外RNA合成的T7 RNA聚合酶启动子,核苷酸18-270对应于GUS反义序列的5'部分,核苷酸271-430对应于环1序列,核苷酸431-933对应于GUS有义序列,核苷酸934-1093对应于环2序列,且核苷酸1094-1343对应于GUS反义序列的3'部分。

[0816] 以相似的方式,通过农杆菌介导的转化,将1edEIN2和hpEIN2构建体分别引入到拟南芥植物的Col-0生态型中。编码hpEIN2[wt]RNA的hpEIN2构建体如前所述,并以反向重复构型包含200bp有义和反义EIN2序列,由PDK内含子隔开。编码1edEIN2的遗传构建体的核苷酸序列示于SEQ ID NO:116。核苷酸37-225对应于EIN2反义序列的5'部分,核苷酸226-373对应于环1序列,核苷酸374-773对应于EIN2有义序列,核苷酸774-893对应于环2序列,且核苷酸894-1085对应于EIN2反义序列的3'部分。核苷酸37-225(反义)与核苷酸374-573(有义)互补,且核苷酸894-1085(反义)与核苷酸574-773(有义)互补。

[0817] 来自初级独立转化体的RNA样品用于Northern印迹杂交分析。如图39所示,对于较大的RNA分子,1edEIN2植物显示出比hpEIN2植物更强烈的杂交信号(图39,上图),表明1edEIN2衍生的RNA积累的水平高于hpEIN2衍生的RNA。对于20-25个核苷酸大小范围内的加工的RNA(siRNA),在1edEIN2植物中检测到的siRNA丰度要比hpEIN2植物更高(图39,下图),并且siRNA的量与较大的RNA分子的丰度很相关。这些结果表明,转基因衍生的1edRNA在一定程度上被Dicer加工成了siRNA,但没有完全被加工。还表明1edRNA转基因比相应的hpRNA转基因生成更多的siRNA。

[0818] 这些结果表明,1edRNA构建体在植物细胞中表达时,与相应的hpRNA构建体相比,导致更高水平的未加工和加工的转录物积累。认为这表明1edRNA分子的稳定性增加。

[0819] 实施例22. 发夹RNA是植物中环状RNA的有效前体

[0820] 环状RNA(circRNA)是共价连接的闭合环,其没有游离5'和3'末端或聚腺苷酸化序列作为3'区。它们通常是非编码的,因为它们不编码多肽,且因此不被翻译。circRNA对RNA酶(特别是对核酸外切酶(例如RNase R))的消化具有相对的抗性。病毒或类病毒起源的circRNA或与病毒相关的卫星RNA早已在植物和动物中被观察到。例如,植物中的亚病毒RNA病原体马铃薯纺锤块茎类病毒(Potato Spindle Tuber Viroid)具有大约360nt大小的环状RNA基因组。在植物中,此类卫星RNA通常在辅助病毒存在下能够被复制。相反,类病毒完

全依赖宿主功能,所述宿主功能包括内源性植物RNA聚合酶用于其复制。

[0821] 将RNA深度测序技术与专门设计的生物信息学工具结合使用,现已从植物和动物基因组中鉴定出大量的cirRNA。已在包括拟南芥、水稻和大豆的植物中鉴定出数千种推定的circRNA,这些植物倾向于表现出组织特异性或生物和非生物胁迫应答性表达模式,但尚未证明植物中circRNA的生物学功能。许多推定的植物circRNA的组织特异性或胁迫应答性表达模式表明它们可能在植物发育和防御应答中具有潜在作用,但这尚未得到证实。

[0822] 关于circRNA的生物发生的共识是,它们是由内含子反向剪接形成,即剪接机制“反向剪接”pre-mRNA,并将剪接的外显子共价连接在一起。因此,内源性内含子剪接机制对于目前的circRNA生物发生模型至关重要。这种生物发生模型主要基于哺乳动物系统的研究,其中大多数外显子circRNA都显示出包含规范的内含子剪接信号,所述规范的内含子剪接信号包括共有的GT/AG内含子边界二核苷酸。在动物中,外显子circRNA侧翼的内含子区域通常含有可转座元件序列的短反向重复序列,并且这导致了互补内含子序列促进circRNA形成的提示。实际上,已经基于天然存在的外显子-内含子序列与含有互补TE重复的可拼接内含子,开发了用于在动物中表达circRNA的载体系统。但是,互补侧翼序列在circRNA形成中的作用在植物中尚不清楚,因为具有这种侧翼内含子序列的已鉴定外显子circRNA的比例非常低,从拟南芥中的0.3%至水稻中的6.2%不等。

[0823] 长发夹RNA (hpRNA) 转基因已被广泛用于诱导植物的基因沉默或RNA干扰 (Wesley等,2001)。hpRNA转基因构建体通常包含反向重复,所述反向重复相对于启动子序列具有互补的有义和反义序列,并且在其之间具有间隔区序列以分离和连接有义和反义序列。间隔区还在载体构建过程中稳定细菌细胞中DNA质粒的反向重复序列结构。因此,预期来自典型hpRNA转基因的RNA转录物将形成茎-环结构,其具有碱基配对的有义和反义序列的双链(ds)茎以及对应于间隔区序列的“环”。由于有义和反义区通过碱基配对退火形成分子的dsRNA区或茎区的能力,此类RNA转录物也称为自互补RNA。

[0824] 来自长hpRNA的环片段在植物细胞中积累并且对RNase R具有抗性

[0825] 制备了转基因,该转基因编码靶向GUS mRNA的长hpRNA,其具有563bp的有义和反义序列以及1113bp的间隔区(图40,GUShp1100)。还制备了第二转基因,其编码靶向相同GUS mRNA的较短的hpRNA,其具有93bp的有义和反义序列以及93bp的间隔区(GUShp93-1)。将两种构建体分别引入到本氏烟草叶细胞中以瞬时表达发夹RNA,并且还用于转化拟南芥植物用于稳定整合和可遗传的转基因表达。如先前报道,两种构建体在引入植物细胞并在植物细胞中表达后均生成预期大小的环序列的不同RNA片段(图41;Wang等人,2008;Shen等人,2015)。在本研究中,发明人希望确定环序列是否被转变为环状RNA。

[0826] 制备具有拟南芥U6启动子而不是35S启动子的第三构建体,用于表达较短的hpRNA(GUShp93-2)。还制备了第四GUS hpRNA构建体,其包括PDK内含子作为间隔区序列(图40中的GUShpPDK)。该构建体编码发夹RNA,其中预计转录后会剪接内含子,留下的环序列要短得多。这些构建体也被引入本氏烟草叶中,以检查是否可以检测到环序列以及它们是否形成环状RNA。这些构建体中的dsRNA茎和环序列均衍生自GUS编码序列,并且未引入已知的内含子序列。在存在或不存在表达靶GUS的构建体的情况下,使用农杆菌介导的浸润向本氏烟草叶中分别引入构建体,以及编码并表达作为病毒抑制蛋白(VSP)的黄瓜花叶病毒2b蛋白的遗传构建体以增强转基因表达。使用Northern印迹杂交测定法分析环片段的积累和大小。

代表性的Northern印迹的放射自显影照片如图42所示。

[0827] 如图42所示,如先前报道的那样,在农杆菌浸润的样品中容易检测到GUS_{hp}1100的长环片段(Shen等人,2015)。为了测试该环片段是否为环状,将RNA样品用RNase R处理并在聚丙烯酰胺凝胶上电泳。RNase R处理使用10 μ g总RNA(或50ng体外转录物)与RNase R缓冲液和水混合,总体积为20 μ l。将混合物在沸水中加热3分钟,在冰上快速冷却,然后加入0.5 μ l RNase R,并将试管在37 $^{\circ}$ C孵育10分钟。使酶失活,并通过用乙醇沉淀来回收残留的RNA。RNase R处理降解了大多数RNA,如凝胶中溴化乙锭染色的材料急剧减少所表明的(图42,下图)。通过所有RNase R处理测定,凝胶中仍可见一些核糖体RNA片段,表明某些RNA种类对RNase R的部分抗性。尽管经过RNase R处理的样品中的总RNA耗竭,但大约1100nt的环片段仍然丰富,与未处理的样品相比仅减少约24%的量。这表明环片段对RNase R消化具有相对抗性,并因此在结构上是环状的。相对于未处理的样品,环RNA量减少了24%,这归因于在商业上获得的RNase R酶中核酸内切酶活性的残留量,或者归因于乙醇沉淀步骤中RNase R消化后RNA回收率的降低。

[0828] 重复RNase R处理测定,其中包含50ng与环序列相对应的体外转录的RNA作为线性RNA对照。另外,用两轮RNase R处理对hpGUS1100浸润的本氏烟草RNA样品进行处理,以更严格地测试RNase R的抗性。观察到,经过一轮RNase R处理后,来自GUS_{hp}1100浸润的本氏烟草叶的环片段的76%保留下来,而仅保留约8.5%的线性体外转录物。两倍的RNase R处理进一步减少了衍生自环的物质,但并未消除它。还应注意的是,来自本氏烟草样品的与环序列相对应的RNA带在凝胶印迹上比在体外转录物显得更大,这与环状RNA一致,据报道,具有相同数量的核苷酸的环状RNA在凝胶电泳中比线性RNA分子迁移得更慢。从这些实验得出结论,约1100个核苷酸的环序列是环状的。

[0829] GUS_{hp}93-1和GUS_{hp}PDK浸润的本氏烟草RNA样品的Northern印迹杂交分析也检测到大小与环序列长度相对应的RNA分子。对于GUS_{hp}93-1和GUS_{hp}93-2构建体,比起35S启动子驱动的GUS_{hp}93-1,由U6启动子引导的GUS_{hp}93-2产生的环片段更多,这表明U6启动子比本氏烟草叶细胞中的35S启动子具有更强的转录活性,或分子以某种方式更稳定。

[0830] GUS_{hp}PDK构建体的间隔区序列包括大小为0.76kb的可剪接的PDK内含子,并因此该构建体的初始转录物包含约0.8kb的环。处理Northern印迹以去除GUS探针,并用全长反义探针针对PDK内含子序列再次探测。PDK探针与未知的RNA种类强烈杂交,其在所有泳道上均观察到强烈的条带。RNase A处理减少但不能完全消除该非特异性条带。尽管如此,尽管片段的丰度看上去相对较弱,但在GUS_{hp}PDK浸润的RNA样品中可以检测到预期大小的PDK内含子特异性条带,这可能是由于内含子序列是从大多数GUS_{hp}PDK初始转录物中剪接出的。为了检查PDK环片段是否为环状,用RNase R处理了GUS_{hp}PDK浸润的本氏烟草叶的RNA。通过RNase R处理几乎完全去除了非特异性杂交带。相反,尽管由于来自非特异性条带的强信号而无法容易地与未处理样品比较丰度,但在RNase R处理后容易检测到PDK内含子条带。综上,这些结果表明hpRNA转录物是环状RNA形成的有效前体,并暗示环状RNA对应于整个环序列。

[0831] RNase R抗性环片段也可在稳定转化的拟南芥植物中积累

[0832] hpGUS347和两个hpGFP构建体(图40)用于转化拟南芥植物的生态型Col-0和表达针对每个构建体选择的转基因的两个植物。hpGUS347构建体作为hpGFP构建体的对照用于

该实验,所述hpGFP构建体被设计为包含用于测试miRNA海绵功能的miR165/166结合位点(在实施例24中讨论)。分析了T2代的转基因植物中从hpGUS347构建体产生的RNA分子的积累,特别是检测了环序列以及它们是否为环状。在来自两个hpGUS347品系的RNase R处理和未处理的RNA样品中均检测到对应于hpGUS347转录物的环的条带。至于农杆菌浸润的本氏烟草组织的RNA样品,与未经处理的样品相比,经RNase R处理的样品的条带强度似乎略有降低,但大多数RNA信号得以保留。使用设计用于检测circRNA的引物进行的RT-qPCR分析证实了RNase R处理的hpGUS347样品中circRNA的存在,与未处理的样品相比,其丰度略有降低。这些结果表明,表达产生发夹RNA的稳定整合的hpRNA转基因也从环序列产生了circRNA。

[0833] 在dsRNA茎环连接处切下hpRNA转录物的环并形成环状RNA

[0834] 为了进一步确认衍生自环序列的RNA分子的环状特性并表征其连接序列,通过RT-PCR使用可以扩增假定的连接序列的寡核苷酸引物从GUShp1100、GUShp93和GUShpPDK浸润的样品中扩增环序列。然后将RT-PCR产物克隆到pGEM-T Easy载体中并测序,确认连接处的核苷酸序列。环状RNA的环切割和连接的核苷酸位置有些变化,其5'位点位于dsRNA茎的3'端内,且3'位点位于环的3'端附近,但5'位点对距离dsRNA茎的3'端10个核苷酸的G核苷酸显示出明显的偏爱。注意到,PDK内含子环状RNA的切下和连接位点遵循与来自GUShp1100和GUShp93 RNA的那些相同的模式,并且位于规范内含子剪接位点之外。结论是,环状RNA的形成是由茎-环结构决定的,而与内含子剪接无关。还得出结论,至少在该实施例中,通过在dsRNA茎的3'端内的5'切割和在环序列的3'端附近的3'切割,加工发夹RNA以释放并环化环序列,其中在切下的序列的5'和3'端之间形成共价键。

[0835] 实施例23. 在酿酒酵母中表达的hpRNA未被加工成环状RNA

[0836] 酵母物种酿酒酵母是一种真核生物体,并拥有与所有真核生物一样的内含子剪接机制。由于目前环状RNA形成的共有模型是基于内含子剪接的,发明人研究了hpRNA是否可以像其在植物细胞中那样在酿酒酵母中形成环状RNA。为了生成表达hpRNA的构建体,将GUShp1100的反向重复区从植物表达载体中切下,并在酵母ADH1启动子的控制下插入酵母表达载体中(图43),并将所得的遗传构建体引入酿酒细胞。如图43所示,对从三个独立的转基因酵母菌株中的每一个提取的RNA的Northern印迹杂交分析检测到一个对应于GUShp1100转录物的高分子量条带。这表明GUShp1100转录物未在酿酒酵母中加工,但仍保持全长。为了证实这一点,比较了酿酒酵母表达的和本氏烟草表达的GUShp1100转录物对RNase R处理的应答。如图44所示,酿酒酵母表达的RNA显示出对RNase R处理高度敏感的高分子量条带,且因此不是环状的。即,酵母RNA样品没有表现出如本氏烟草细胞中产生的衍生自环序列的环状分子。结果表明,在酿酒酵母中表达的GUShp1100转录物未被加工,并保持全长。通过凝胶电泳,酿酒酵母RNA条带的大小似乎比体外GUShp1100转录物大,这可能是由于酿酒酵母表达的RNA中存在而在体外转录物中不存在的5'和3'UTR和poly(A)序列。因此,酿酒酵母中内含子剪接机制的存在不足以允许hpRNA环的加工和环状RNA的形成,如在植物细胞中发生的那样。

[0837] 以类似的方式,将遗传构建体GUShp347引入酿酒酵母中并表达。Northern印迹杂交分析再次表明,hpRNA看起来是全长的,并且似乎没有被加工,至少不具有环序列或dsRNA区的切割。

[0838] 发明人得出的结论是,不具有Dicer酶(Drinneberg et al.,2003)的酿酒酵母和其相关的出芽酵母作为产生全长发夹和ledRNA(包括本文所述的修饰的RNA分子)的生物体是有利的。这样的全长RNA在期望未加工的dsRNA时是有用的,例如用于通过局部应用于昆虫来沉默基因活性。

[0839] 实施例24.hpRNA环可用作抑制miRNA功能的有效“海绵”

[0840] 已经发现动物中的一些环状RNA包含与特定miRNA互补的多个序列,并从而充当那些miRNA的结合位点,称为miRNA“海绵”。发明人测试了由长hpRNA构建体产生的环状RNA是否可以在植物细胞中充当miRNA海绵。设计了两个GFP hpRNA构建体(图40),它们具有相同的GUS序列衍生的间隔区,不同之处在于一个序列经修饰后具有两个拟南芥miR165/166结合位点。该构建体GFPhp[G:U]具有反向重复序列,该序列具有与第二(对照)构建体GFPhp[WT]相同的反义序列,但具有修饰的有义序列,其中所有胞嘧啶核苷酸均被胸腺嘧啶取代。因此,GFPhp[G:U]的转录物将形成对应于GFP序列的dsRNA区,除了大约25%的碱基对是G:U碱基对。另一个构建体GFPhp[WT]编码发夹RNA,其具有与GFPhp[G:U]的发夹相同长度的完全规范碱基配对的dsRNA茎,并用作对照(图40)。包含没有miR165/166结合位点的间隔区的GUS hpRNA构建体GUShp347被包括作为第二对照。

[0841] 使用该构建体分别转化拟南芥,并分别获得这三种构建体的转基因植物。目视检查转化的植物的与miR165/166的减少有关的表型,其中包括将叶子独特地折叠成“喇叭”。如预期的那样,GUShp347转化的植物未显示与miR165/166抑制相关的表型。同样,在GFPhp[WT]转化的植物中没有观察到明确的表型。相反,大多数GFPhp[G:U]植物表现出不同水平的联想到miR165/166抑制的表型,包括喇叭表型。

[0842] 在从GFPhp[G:U]转化植物中提取的RNA上进行Northern印迹杂交,该植物具有一系列轻度、中度和强至重度表型,以检查hpRNA表达的积累。使用的探针是对应于GUS mRNA的全长反义RNA。该探针与GUShp347转录物的有义和邻近环具有822bp的连续序列互补性。该探针与GFPhp转录物的序列互补性较小,所述GFPhp转录物在GUS衍生序列中共有228bp的环区,在位于两个miRNA结合序列侧翼的长度分别为49、109和70bp的三个非连续区域中。如图45B所示,在GFPhp[G:U]植物中检测到非常大量的GFP hpRNA分子,并且在Northern印迹中检测到的RNA分子的量与表型的严重性正相关。GFPhp[WT]植物表现出的hpRNA分子积累水平低,其仅在Northern印迹分析中才可检测到,这与常规的hpRNA转基因相比于G:U修饰的hpRNA转基因而言相对较低的转录水平相符。即,如以上实施例中所示,与相应的hpRNA[WT]转基因相比,hpRNA[G:U]转基因较少受到自我沉默。

[0843] RT-qPCR用于定量衍生自环序列的环状RNA分子的积累。结果表明,GFPhp[G:U]转基因植物中存在大量circRNA,其与全长hpRNA积累水平相关(图45C)。Northern印迹杂交分析检测20-25nt大小范围的小RNA,证实了GFPhp[G:U]植物中miR165/166的下调。减少的程度与hpRNA和circRNA的量以及表型的严重程度相关。使用RT-qPCR对miR165/166靶基因进行表达分析表明,在表现出强烈的miR165/166下调和严重的表型的植物中释放了miR165/166对靶基因的抑制。综上,这些结果表明,hpRNA环可用作特定的miRNA海绵,以抑制植物中的miRNA功能。

[0844] 发明人还设想了将在植物细胞中以高水平产生的环状RNA作为稳定分子的用途,以翻译成产生高水平多肽的工具。为了启动帽非依赖性翻译,理想地使用内部核糖体进入

位点 (IRES)。已经鉴定出许多 IRES 序列。

[0845] 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。

[0846] 本申请要求 2018 年 8 月 3 日提交的 AU2018902840、2018 年 8 月 8 日提交的 AU2018902896、2018 年 9 月 17 日提交的 PCT/AU2018/051015 和 2019 年 3 月 20 日提交的 AU2019900941 的优先权,其公开内容通过引用并入本文。

[0847] 本文讨论和/或引用的所有出版物均全文引入本文。

[0848] 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。

[0849] 参考文献

[0850] Abraham (2002). Cell 111:9-12.

[0851] Acevedo-Garcia et al., (2017). Plant Biotechnology Journal 15:367-378.

[0852] Alvarez et al., (2000). Theor Appl Genet 100:319-327.

[0853] Baumlein et al., (1991). Mol. Gen. Genet. 225:459-467.

[0854] Baumlein et al., (1992). Plant J. 2:233-239.

[0855] Bhattacharyya et al., (1990) Cell 60:155-122.

[0856] Brar et al., (1996) Biotech Genet. Eng Rev 13:167-79.

[0857] Broothaerts et al., (2005). Nature 433:629-633.

[0858] Broun et al., (1998). Plant J. 13:201-210.

[0859] Buchanan-Wollaston, (1994) Plant physiology 105:839-846.

[0860] Busk et al., (1997). Plant J. 11:1285-1295.

[0861] Chen et al., (2005). Functional Plant Biology 32:671-681.

[0862] Chikwamba et al., (2003). Proc. Natl. Acad. Sci. U.S.A. 100:11127-11132.

[0863] Christou and Klee, (2004) Handbook of Plant Biotechnology, John Wiley and Sons.

[0864] Chung et al., (2006). BMC Genomics 7:120.

[0865] Clough and Bent (1998). Plant J. 16:735-743.

[0866] Corrado and Karali (2009). Biotechnol. Adv. 27:733-743.

[0867] Courvalin et al., (1995). Life Sci. 318:1209-1212.

[0868] Dadd and Mittler (1966). Experientia 22:832-833.

[0869] Darji et al., (1997). Cell 91:765-775.

[0870] Dong et al., (2011) Plant J. 68:633-45.

[0871] 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.

[0872] Drinnenberg et al.. (2009) Science 326:544-550.

- [0873] Dunwell (2000) .J Exp Botany 51Spec No:487-496.
- [0874] Ebhardt et al., (2005) Proc.Nat.Acad.Sci.USA 102:13398-13403.
- [0875] Ellerstrom et al., (1996) .Plant Mol.Biol.32:1019-1027.
- [0876] Elliott et al., (2002) .Mol.Plant Microbe Interact.15:1069-1077.
- [0877] Ellis et al., (1987) .EMBO J 6:11-16.
- [0878] Feechan et al., Functional Plant Biology, 2008, 35:1255-1266.
- [0879] Fehr, In: Breeding Methods for Cultivar Development, Wilcox J. ed., American Society of Agronomy, Madison Wis. (1987) .
- [0880] Gan (1995) .Molecular characterization and genetic manipulation of plant senescence. PhD thesis. University of Wisconsin, Madison.
- [0881] Gan and Amasino (1995) .Science 270:1986-1988.
- [0882] Glawe (2008) .Ann.Rev.Phytopathol.46:27-51.
- [0883] Gleave (1992) .Plant Mol Biol 20:1203-1207.
- [0884] Guo et al., (2016) Curr.Genom.17:476-489.
- [0885] Gupta et al., (1988) Plant Mol.Biol.10:215-224.
- [0886] Helliwell and Waterhouse (2005) .Methods in Enzymology 392:24-35.
- [0887] Hershey and Stoner, (1991) Plant Molecular Biol.17:679-690.
- [0888] Hinchey et al., (1988) .Biotechnology 6:915-922.
- [0889] Horvath et al., (2000) .Proc.Natl.Acad.Sci.U.S.A.97:1914-1919.
- [0890] Hsieh and Fire (2000) Annu Rev Genet 14:187-204.
- [0891] Jeddeloh, et al., (1999) .Nat.Genet.22:94-97.
- [0892] Jefferson et al., (1987) .EMBO J 6:3901-3907.
- [0893] Jepson et al., (1994) .Plant Mol.Biol.26:1855-1866.
- [0894] Khan (2017) .Sci Rep 7, 40025
- [0895] Kim et al., (2019) J.Gen.Plant Path. <https://doi.org/10.1007/s10327-019-00865-7>
- [0896] Kishore and Somerville (1993) .Curr Opin Biotechnol.4:152-158.
- [0897] Koziel et al., (1996) .Plant Mol.Biol.32:393-405.
- [0898] Lacroix et al., (2008) .Proc.Natl.Acad.Sci.U.S.A.105:15429-15434.
- [0899] Lau et al., (2001) Science 294:858-862.
- [0900] Li et al., (1996) .FEBS Lett.379:117-121.
- [0901] McCahill et al., (2002) .Molecular Pharmacology 62:1261-1273.
- [0902] McCullough and Schuler (1997) .Nucl Acids Res.25:1071-1077.
- [0903] Mahfouz et al., (2006) .Plant Cell 18:477-490.
- [0904] Matsuoka et al., (1994) .Plant J.6:311-319.
- [0905] Meier et al., (1997) .FEBS Lett.415:91-95.
- [0906] Melamed-Bessudo et al., (2012) Proc.Natl.Acad.Sci.USA 109(16) :e981-988.
- [0907] Millar et al., (2006) .Plant J.45:942-954.
- [0908] Mutti et al., (2006) .J Insect Sci 6:38.

- [0909] Mutti et al., (2008). *Proceedings of the National Academy of Sciences* 105:9965-9969.
- [0910] Olive et al., (1989) *Plant Mol Biol* 12:525-538.
- [0911] Padidam (2003). *Transgenic Res.* 12:101-109.
- [0912] Perrin et al., (2000). *Mol Breed* 6:345-352.
- [0913] Pitino et al., (2011). *PLoS ONE* 6, e25709.
- [0914] Pitino and Hogenhout (2012). *Molecular Plant-Microbe Interactions* 26: 130-139.
- [0915] Potenza et al., (2004). *In Vitro Cell Dev. Biol. Plant* 40:1-22.
- [0916] Powell et al., (1996). *Vaccines* 183, Abstract.
- [0917] Preiss et al., (1987). In: *Tailoring Genes for Crop Improvement* (Bruening et al., eds.), Plenum Press, S.133-152.
- [0918] Sambrook et al., (1989). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbour Laboratory Press.
- [0919] Schaffner (1980). *Proc. Natl. Acad. Sci. U.S.A.* 77:2163-2167.
- [0920] Schwartz et al., (1997). *Science* 276:1872-1874.
- [0921] Scott et al., (2013). *Journal of Insect Physiology* 59:1212-1221.
- [0922] Seddas et al., (2004). *Virology* 325:399-412.
- [0923] Shen et al., (2015) *Frontiers in Plant Science* 6:281.
- [0924] Shiina et al., (1997). *Plant Physiol.* 115:477-483.
- [0925] Shure et al., (1983). *Cell* 35:225-233.
- [0926] Sizemore et al., (1995). *Science* 270:299-302.
- [0927] Slater et al., *Plant Biotechnology-The Genetic Manipulation of Plants*, Oxford University Press (2003).
- [0928] Smith et al., (2000). *Nature* 407:319-320.
- [0929] Stalker et al., (1988). *Science* 242:419-423.
- [0930] Stewart et al., (2000). *J Mol Biol* 298:611-622.
- [0931] Tan et al., (2011). *Plant Physiol.* 156:1577-1588.
- [0932] Thillet et al., (1988). *J. Biol. Chem* 263:12500-12508.
- [0933] Timmons et al., (2001). *Gen* 263:103-112.
- [0934] Ulmasov et al., (1995). *Plant Physiol.* 108:919-927.
- [0935] Wang (1994) Isolation of phloem specific gene promoters for use in genetic engineering of insect resistance in rice. PhD thesis, University of Durham, UK.
- [0936] Wang et al., (1994) *Plant Molecular Biology* 24:159-170.
- [0937] Wang et al., (1998) *Acta Horticulturae* 461:401-407.
- [0938] Wang et al., (2008) *RNA* 14:903-913. Wang et al., (2013). *PLoS Genet* 9, e1003865.
- [0939] Wang et al., (2014). *Nature Biotechnology* 32:9.

- [0940] Weiss (2003). *Int. J. Med. Microbiol.* 293:95:106.
- [0941] Weissbach et al., (1988). In: *Methods for Plant Molecular Biology*, Academic Press, San Diego, Calif.
- [0942] Wijnker et al., (2008). *Trends in Plant Science* 13:640-646.
- [0943] Wesley et al. (2001) *Plant J.* 27:581-590.
- [0944] Yang et al., (1994) *Particle Bombardment Technology for Gene Transfer*, Oxford Press, Oxford, England.
- [0945] Yang et al., (2003). *Planta* 216:597-603.
- [0946] Yelina et al., (2012). *PLoS Genetics* 8 (8):e1002844. doi:10.1371/journal.pgen.1002844
- [0947] Yu et al., (2016). *Pest Management Science* 72:1090-1098.
- [0948] Zhang et al., (2018). *Nat Rev Mol Cell Biol.* 19:489-506.

序列表

<110> 联邦科学技术研究组织

<120> 包含非规范碱基对的RNA分子

<130> 526425PCT

<150> AU 2018902840

<151> 2018-08-03

<150> AU 2018902896

<151> 2018-08-08

<150> PCT/AU2018/051015

<151> 2018-09-17

<150> AU 2019900941

<151> 2019-03-04

<160> 144

<170> PatentIn version 3.5

<210> 1

<211> 1229

<212> RNA

<213> Artificial Sequence

<220>

<223> GFP ledRNA

<400> 1

```

gggugucgcc  cucgaacuuc  accucggcgc  gggucuugua  guugccgucg  uccuugaaga  60
agauggugcg  cuccuggacg  uagccuucgg  gcauggcgga  cuugaagaag  ucgugcugcu  120
ucaugugguc  gggguagcgg  cugaagcacu  gcacgccgua  ggugaaggug  gucacgaggg  180
ugggccaggg  cacgggcagc  uugccggugg  ugcagaugaa  cuucaggguc  agcuugccgu  240
agguggcauc  gccucgccc  ucgccggaca  cgcugaacuu  guggccguuu  acgucgccgu  300
ccagcucgac  caggaugggc  accaccccgg  ugaacagcuc  cucgcccug  cucacuaugg  360
aucaacuagg  gauccccug  aaguucaucu  gcaccaccgg  caagcugccc  gugcccuggc  420
ccaccucgu  gaccaccuuc  accuacggcg  ugcagugcuu  cagccgcuac  cccgaccaca  480
ugaagcagca  cgacuucuu  aaguccgcca  ugcccgaagg  cuacguccag  gagcgacca  540
ucuucuuaa  ggacgacggc  aacuacaaga  cccgcgccga  ggugaaguuc  gagggcgaca  600
cccuggugaa  ccgcaucgag  cugaagggca  ucgacuuaa  ggaggacggc  acauccugg  660
ggcacaagcu  ggaguacaac  uacaacagcc  acaacgucua  uaucauggcc  gacaagcaga  720
agaacggcau  caaggugaac  uucaagauc  gccacaacau  cgaggacggc  agcgugcagc  780
ucgccgacca  cuaccagcag  aacaccccc  ucggcgacgg  ccccugucug  cugccaagcu  840
uuaggugauc  caagcuugau  ccgggcuuu  cuuguacagc  ucguccaugc  cgagagugau  900
cccggcggcg  gucacgaacu  ccagcaggac  caugugaucg  cgcuucucgu  uggggucuuu  960
gcucagggcg  gacugggugc  ucagguagug  guugucgggc  agcagcacgg  ggccgucgcc  1020

```

gaugggggug uucugcuggu aguggucggc gagcugcacg cugccguccu cgauguugug 1080
 gcggaucuuug aaguucaccu ugaugccguu cuucugcuug ucggccauga uauagacguu 1140
 guggcuguug uaguuguacu ccagcuugug cccaggaug uugccguccu ccuugaaguc 1200
 gaugcccuuc agcucgaugc gguucacca 1229

<210> 2

<211> 1326

<212> RNA

<213> Artificial Sequence

<220>

<223> GUS ledRNA

<400> 2

gggaacagac gcgugguuac agucuuugcg gacaugcguc accacgguga uaucguccac 60
 ccagguguuuc ggcguggugu agagcauuac gcugcgauug auuccggcau aguuuaagaa 120
 aucauggaag uaagacugcu uuuuuuugcc guuuuugcug guaauacca uuuccggcg 180
 gauagucugc caguucaguu cguuguucac acaaacggug auacguacac uuuuuccggc 240
 aaauacauac ggcgugacau cggcuucaa uggcguaug ccgcccugau gcuccaucac 300
 uuccugauua uugaccaca cuuugccgua augagugacc gcaucgaaac gcagcacgau 360
 acgcuggccu gccaaccuu ucgguaauaa gacuucgcgc ugauaccaga cgugccguau 420
 guuauugccg ggaaaagugu acguaucacc guuuguguga acaacgaacu gaacuggcag 480
 acuaucccg cgggaauugu gauuaccgac gaaaacggca agaaaaagca gucuuacuuc 540
 caugauuuu uuaacuaugc cggaauccau cgcagcguaa ugcucuacac cacgccgaac 600
 accugggugg acgauaucac cguggugacg caugucgcgc aagacuguaa ccacgcguc 660
 guucccgacu ggagguggu ggccaauugu augucagcg uugaacugcg ugaugcgau 720
 caacaggugg uugcaacugg acaaggcacu agcgggacuu ugcaaguggu gaauccgcac 780
 cucuggcaac cgggugaagg uuaucucuau gaacugugcg ucacagccaa aagccagaca 840
 gagugugaua ucuaccgcu ucgcgucggc auccggucag uggcagugaa gggccaacag 900
 uuccugauua accacaaacc guucuacuuu acuggcuuug gucgucauga agaugcgac 960
 uuacguggca aaggauucga uaacgugcug auggugcacg accacgcau aauggacugg 1020
 auuggggcca acuccuaccg uaccucgcau uaccuuuacg cugaagagau gcucgaugug 1080
 guuaaucagg aacuguugc ccuucacugc cacugaccgg augccgacgc gaagcgggua 1140
 gauaucacac ucugucugc uuugggcugu gacgcacagu ucuaagagau aaccuucacc 1200
 cgguugccag aggugcgga uaccacuug caaaguccg cuagugccu guccaguugc 1260
 aaccaccugu ugauccgcau cacgcaguuc aacgcugaca ucaccuugg ccaccaccug 1320
 ccaguc 1326

<210> 3

<211> 1485

<212> RNA

<213> Artificial Sequence

<220>

<223> FAD2.1 ledRNA

<400> 3

```

gaagaucugu agccucucgc ggucauugua gauugggccg uaagggucau agugacaugc 60
aaagcgauca uaaugucggc cagaaacauu gaaagccaag uacaaaggcc agccaagagu 120
aagggugauc guaagugaaa uaacccggcc ugguggauug uucaaguacu uggaauacca 180
uccgaguugu gauuucggcu uaggcacaaa aaccucaucg cgcucgagug agccaguguu 240
ggaguggugg cgacgaugac uauauuuucca agagaaguag ggcaccauca gagcagagug 300
gaggauaagc ccgacagugu caucaaccca cugguaguca cuaaaaggcau gguggccaca 360
uucgucgca auaacccaaa uaccagugca aacacaaccc ugacaaaucc aguaaaauagg 420
ccaugcaagg uagcaauccu aggcacucug cucugauggu gcccuacuuc ucuuggaaau 480
auagucaucg ucgccaccac uccaacacug gcucacucga gcgcgaugag guuuuugugc 540
cuaagccgaa aucacaacuc ggaugguauu ccaaguacuu gaacaaucca ccaggccggg 600
uuauuucacu uacgaucacc cuuacucuug gcuggccuuu guacuuggcu uucaauguuu 660
cuggccgaca uuauaugcgc uuugcauguc acuaugaccc uuacggccca aucuacaug 720
accgcgagag gcuacagauc uuuccuuucug augcuggagu uauuggagcu gguuaucua 780
uauaucguau ugccuuggua aaagggcuag cuuggcucgu guguauguau ggcguaccac 840
uccuaaucgu gaacggcuuc cuugucuuga ucacuuauuu gcagcacacu caccgucan 900
ugccucacua cgauucaucc gaaugggauu ggcuaagggg agcuuuggca accgucgaca 960
gagacuaugg cauucuaaac aaggucuucc acaacaucac cgauacucac guaguccacc 1020
aucuguucuc gaccaugcca cacucuagag ugaugcuuca ucuuucucca cauagauaca 1080
cucuuuugcu uccuccaca uugccuugaa aaccgggguu ccgucaaaau gguaguaguc 1140
uccgaguaau ggcuugacug cuuuuguugc cuccauugca uuguagugug gcauggucga 1200
gaacagaugg uggacuacgu gaguaucggu gauguugugg aagaccuugu uuagaaugcc 1260
auagucucug ucgacgguug ccaaagcucc ccuagccaa ucccauucgg augaaucgua 1320
gugaggcaau gacgggugag ugugcugcaa auuagugauc aagacaagga agccguucac 1380
gauuaggagu gguacgccau acauacacac gagccaagcu agcccuuuua ccaaggcaau 1440
acgauauagu agauaaccag cuccaauaac uccagcauca gaaag 1485

```

<210> 4

<211> 1258

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA construct encoding GFP ledRNA

<400> 4

```

taatacgact cactataggg tgtcgccctc gaacttcacc tcggcgcggg tctttagttt 60
gccgtgtcc ttgaagaaga tgggtgcgtc ctggacgtag ctttcgggca tggcggactt 120
gaagaagtcg tgctgcttca tgtggtcggg gtagcggctg aagcactgca cgccgtaggt 180
gaaggtggtc acgagggtgg gccagggcac gggcagcttg ccggtggtgc agatgaactt 240
cagggtcagc ttgccgtagg tggcatcgcc ctcgcctcg ccggacacgc tgaacttgtg 300

```

```

gccgtttacg tgcgcgtcca gctcgaccag gatgggcacc accccggtga acagctcctc 360
gcccttgctc actatggatc aactagggat cccctgaag ttcattctgca ccaccggcaa 420
gctgcccgtg ccctggccca ccctcgtgac caccttcacc tacggcgtgc agtgcttcag 480
ccgctacccc gaccacatga agcagcacga cttcttcaag tccgccatgc ccgaaggcta 540
cgtccaggag cgcaccatct tcttcaagga cgacggcaac tacaagacct gcgccgaggt 600
gaagttcgag ggcgacaccc tggatgaacc catcgagctg aaggcatcg acttcaagga 660
ggacggcaac atcctggggc acaagctgga gtacaactac aacagccaca acgtctatat 720
catggccgac aagcagaaga acggcatcaa ggtgaacttc aagatccgcc acaacatcga 780
ggacggcagc gtgcagctcg ccgaccacta ccagcagaac acccccatcg gcgacggccc 840
cgtgctgctg ccaagcttta ggtgatccaa gcttgatccg ggctttactt gtacagctcg 900
tccatgccga gagtgatccc ggcgggcggtc acgaactcca gcaggacct gtgatcgcg 960
ttctcgttgg ggtctttgct cagggcggac tgggtgctca ggtagtggtt gtcgggcagc 1020
agcacggggc cgtcgccgat ggggggtgtt tgctggtagt ggtcggcgag ctgcacgctg 1080
ccgtcctcga tgttgtggcg gatcttgaag ttcacctga tgccgttctt ctgcttgctg 1140
gccatgatat agacgttggt gctgtttagt ttgtactcca gcttggtccc caggatgttg 1200
ccgtcctcct tgaagtcgat gcccttcagc tcgatcggt tcaccattgt cgggatac 1258

```

<210> 5

<211> 1346

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA construct encoding Gus ledRNA

<400> 5

```

taatacgact cactataggg aacagacgcg tggttacagt cttgcgcgac atgcgtcacc 60
acgggtgatat cgtccacca ggtgttcggc gtgggtgtaga gcattacgct gcgatggatt 120
ccggcatagt taaagaaatc atggaagtaa gactgctttt tcttgccgtt ttcgtcggta 180
atcaccattc ccggcgggat agtctgccag ttcagttcgt tgttcacaca aacggtgata 240
cgtacacttt tcccggcaat aacatacggc gtgacatcgg cttcaaattg cgtatagccg 300
ccctgatgct ccatcacttc ctgattattg acccacatt tgccgtaatg agtgaccgca 360
tcgaaacgca gcacgatacg ctggcctgcc caacctttcg gtataaagac ttcgcgctga 420
taccagacgt gccgtatggt attgccggga aaagtgtacg tatcaccgtt tgtgtgaaca 480
acgaactgaa ctggcagact atcccgccgg gaatggtgat taccgacgaa aacggcaaga 540
aaaagcagtc ttacttccat gatttcttta actatgccgg aatccatcgc agcgtaatgc 600
tctacaccac gccgaacacc tgggtggacg atatcaccgt ggtgacgcat gtcgcgcaag 660
actgtaacca cgcgtctgtt cccgactggc aggtgggtggc caatggtgat gtcagcgttg 720
aactgcgtga tcgggatcaa caggtggttg caactggaca aggcactagc gggactttgc 780
aagtggtgaa tccgcacctc tggcaaccgg gtgaaggtta tctctatgaa ctgtgcgtca 840
cagccaaaag ccagacagag tgtgatatct acccgcttcg cgtcggcatc cggtcagtgg 900
cagtgaaggg ccaacagttc ctgattaacc acaaaccgtt ctactttact ggctttggtc 960

```

gtcatgaaga tgcggactta cgtggcaaag gattcgataa cgtgctgatg gtgcacgacc 1020
 acgcattaat ggactggatt ggggccaact cctaccgtac ctcgcattac ctttacgctg 1080
 aagagatgct cgatgtggtt aatcaggaac tgttggccct tctactgccac tgaccgatg 1140
 ccgacgcgaa gcgggtagat atcacactct gtctggcttt tggctgtgac gcacagttca 1200
 tagagataac cttcaccggt ttgccagagg tgcggattca ccacttgcaa agtcccgcta 1260
 gtgccttgtc cagttgcaac cacctgttga tccgcatcac gcagttcaac gctgacatca 1320
 ccattggcca ccacctgcca gtcaac 1346

<210> 6

<211> 1512

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA construct encoding FAD2.1 ledRNA

<400> 6

atttaggtga cactatagaa gatctgtagc ctctcgcgtt cattgtagat tgggccgtaa 60
 gggtcatagt gacatgcaaa gcgatcataa tgtcggccag aaacattgaa agccaagtac 120
 aaaggccagc caagagtaag ggtgatcgta agtgaaataa cccggcctgg tggattgttc 180
 aagtacttgg aataccatcc gagttgtgat ttcggcttag gcacaaaaac ctcatcgcgc 240
 tcgagtgagc cagtgttggg gtggtggcga cgatgactat atttccaaga gaagtagggc 300
 accatcagag cagagtggag gataagcccc acagtgtcat caaccactg gtagtcacta 360
 aaggcatggt ggccacattc gtgcgcaata acccaaatac cagtgcaaac acaaccctga 420
 caaatccagt aaataggcca tgcaaggtag caatcctagg cactctgctc tgatggtgcc 480
 ctacttctct tggaaatata gtcacgtcgc ccaccactcc aacactggct cactcgagcg 540
 cgatgagggt tttgtgccta agccgaaatc acaactcgga tggatttcca agtacttgaa 600
 caatccacca ggccgggtta tttcacttac gatcaccctt actcttggct ggcctttgta 660
 cttggctttc aatgtttctg gccgacatta tgatcgcttt gcatgtcact atgaccctta 720
 cggcccaatc tacaatgacc gcgagaggct acagatcttc ctttctgatg ctggagttat 780
 tggagctggt tatctactat atcgtattgc cttggtaaaa gggctagctt ggctcgtgtg 840
 tatgtatggc gtaccactcc taatcgtgaa cggttctctt gtcttgatca cttatttgca 900
 gcacactcac ccgtcattgc ctactacga ttcacccgaa tgggattggc taaggggagc 960
 tttggcaacc gtcgacagag actatggcat tctaaacaag gtcttccaca acatcaccga 1020
 tactcacgta gtccaccatc tgttctcgac catgccacac tctagagtga tgcttcatct 1080
 ttctccacat agatacactc ttttgettcc ctccacattg ccttgaaaac cggggttccg 1140
 tcaaattggt agtagtctcc gagtaatggc ttgactgctt ttgttgctc cattgcattg 1200
 tagtgtggca tggtcgagaa cagatggtgg actacgtgag tatcgggtgat gttgtggaag 1260
 acctgtttta gaatgccata gtctctgtcg acggttgcca aagctccctc tagccaatcc 1320
 cattcggatg aatcgtagtg aggcaatgac gggtagtgtg gctgcaaata agtgatcaag 1380
 acaaggaagc cgttcacgat taggagtggc acgccataca tacacacgag ccaagctagc 1440
 ccttttacca aggcaatacg atatagtaga taaccagctc caataactcc agcatcagaa 1500

agccccgggac tc 1512

<210> 7

<211> 732

<212> DNA

<213> *Aequorea victoria*

<400> 7

```
atggtgagca agggcgagga gctgttcacc ggggtggtgc ccatcctggt cgagctggac 60
ggcgacgtaa acggccacaa gttcagcgtg tccggcgagg gcgagggcga tgccacctac 120
ggcaagctga ccctgaagtt catctgcacc accggcaagc tgcccgtgcc ctggcccacc 180
ctcgtgacca ctttcaccta cggcgtgcag tgcttcagcc gctaccccga ccacatgaag 240
cagcacgact tcttcaagtc cgccatgccc gaaggtacg tccaggagcg caccatcttc 300
ttcaaggacg acggcaacta caagaccgc gccgaggtga agttcgaggg cgacaccctg 360
gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcaacat cctggggcac 420
aagctggagt acaactacaa cagccacaac gtctatatca tggccgacaa gcagaagaac 480
ggcatcaagg tgaacttcaa gatccgccac aacatcgagg acggcagcgt gcagctcgcc 540
gaccactacc agcagaacac ccccatcggc gacggccccg tgctgctgcc cgacaaccac 600
tacctgagca cccagtccgc cctgagcaaa gacccaacg agaagcgcg tcacatggtc 660
ctgctggagt tcgtgaccgc cgccgggatc actctcggca tggacgagct gtacaagtaa 720
agccccgatac tc 732
```

<210> 8

<211> 1855

<212> DNA

<213> *Escherichia coli*

<400> 8

```
atggtccgtc ctgtagaaac cccaaccgt gaaatcaaaa aactcgacgg cctgtgggca 60
ttcagtctgg atcgcgaaaa ctgtggaatt gatcagcgtt ggtgggaaag cgcgttacia 120
gaaagccggg caattgctgt gccaggcagt tttacgatac agttcgccga tgcagatatt 180
cgtaattatg cgggcaacgt ctggtatcag cgcgaagtct ttataccgaa aggttgggca 240
ggccagcgta tcgtgctgcg tttcgatgcg gtcactcatt acggcaaagt gtgggtcaat 300
aatcaggaag tgatggagca tcagggcggc tatacgccat ttgaagccga tgtcacgccg 360
tatgttattg ccgggaaaaag tgtacgtatc accgtttgtg tgaacaacga actgaactgg 420
cagactatcc cgccgggaat ggtgattacc gacgaaaacg gcaagaaaaa gcagtcttac 480
ttccatgatt tctttaacta tgccggaatc catcgacgcg taatgctcta caccacgccg 540
aacacctggg tggacgatat caccgtggtg acgcatgtcg cgcaagactg taaccacgcg 600
tctgttcccc actggcaggt ggtggccaat ggtgatgtca gcgttgaaact gcgtgatgcg 660
gatcaacagg tgggtgcaac tggacaaggc actagcggga ctttgcaagt ggtgaatccg 720
cacctctggc aaccgggtga aggttatctc tatgaactgt gcgtcacagc caaaagccag 780
acagagtgtg atatctaccc gcttcgcgtc ggcacccgtg cagtggcagt gaagggcaa 840
cagttcctga ttaaccacaa accgttctac ttactggct ttggtcgtca tgaagatgcg 900
```

gacttacgtg gcaaaggatt cgataacgtg ctgatggtgc acgaccacgc attaatggac 960
 tggattgggg ccaactccta ccgtacctcg cattaccctt acgctgaaga gatgctcgac 1020
 tgggcagatg aacatggcat cgtggtgatt gatgaaactg ctgctgtcgg ctttaacctc 1080
 tcttttaggca ttggtttcga agcgggcaac aagccgaaag aactgtacag cgaagaggca 1140
 gtcaacgggg aaactcagca agcgcactta caggcgatta aagagctgat agcgcgtgac 1200
 aaaaaccacc caagcgtggt gatgtggagt attgccaacg aaccggatac ccgtccgcaa 1260
 gtgcacggga atatttcgcc actggcgga gcaacgcgta aactcgaccc gacgcgtccg 1320
 atcacctgcg tcaatgtaat gttctgcgac gtcacaccg ataccatcag cgatctcttt 1380
 gatgtgctgt gcctgaaccg ttattacgga tggtatgtcc aaagcggcga tttggaaacg 1440
 gcagagaagg tactggaaaa agaacttctg gcctggcagg agaaactgca tcagccgatt 1500
 atcatcaccg aatacggcgt ggatacgta gccgggctgc actcaatgta caccgacatg 1560
 tggagtgaag agtatcagtg tgcattggtg gatatgtatc accgcgtctt tgatcgcgtc 1620
 agcgcgtcgc tcggtgaaca ggtatggaat ttcgccgatt ttgcgacctc gcaaggcata 1680
 ttgcgcgttg gcgtaacaa gaaagggatc ttcactcgcg accgcaaacc gaagtcggcg 1740
 gcttttctgc tgcaaaaacg ctggactggc atgaacttcg gtgaaaaacc gcagcaggga 1800
 ggcaacaat gaatcaacaa ctctcctggc gcaccatcgt cggctacagc ctcgg 1855

<210> 9

<211> 1152

<212> DNA

<213> *Nicotiana benthamiana*

<400> 9

atgggagctg gtggtaatat gtctcttgta accagcaaga ctggcgaaaa gaagaatcct 60
 cttgaaaagg taccaacctc aaagcctcct ttcacagttg gtgatataca gaaggccatc 120
 ccacctcact gctttcagcg gtctctcggt cgttcgttct cctatgttgt gtatgacctt 180
 ttactgggtg ccgtcttcta ctacattgcc accattact tccacctcct cccgtcccca 240
 tattgctacc ttgcatggcc tatttactgg atttgtcagg gttgtgtttg cactgggtatt 300
 tgggttattg cgcacgaatg tggccacat gcctttagt actaccagtg ggttgatgac 360
 actgtcgggc ttatcctcca ctctgctctg atgggtgccct acttctcttg gaaatatagt 420
 catcgtcggc accactccaa cactggtcga ctcgagcgcg atgaggtttt tgtgcctaag 480
 ccgaaatcac aactcggatg gtattccaag tacttgaaca atccaccagg ccgggttatt 540
 tcacttacga tcaccttac tcttggctgg ctttgtact tggctttcaa tgtttctggc 600
 cgacattatg atcgctttgc atgtcactat gaccttacg gcccaatcta caatgaccgc 660
 gagaggctac agatcttcct ttctgatgct ggagttattg gagctgggta tctactatat 720
 cgtattgcct tggtaaaaagg gctagcttgg ctctgtgtga tgtatggcgt accactccta 780
 atcgtgaacg gcttccttgt cttgatcact tatttgcagc aactcaccg gtcattgcct 840
 cactacgatt catccgaatg ggattggcta aggggagctt tggcaaccgt cgacagagac 900
 tatggcattc taaacaaggt cttccacaac atcaccgata ctcacgtagt ccaccatctg 960
 ttctcgacca tgccacacta caatgcaatg gaggaacaa aagcagtcaa gccattactc 1020
 ggagactact accaatttga cggaaccccg gttttcaagg caatgtggag ggaagcaaaa 1080

gagtgtatct atgtggagaa agatgaagca tcacaaggca aaggtgtttt ctggtacaaa 1140
aacaaattct ga 1152

<210> 10

<211> 225

<212> DNA

<213> Artificial Sequence

<220>

<223> GUS sense region for constructs encoding hairpin RNA molecules
targeting the GUS mRNA

<400> 10

cctcgaggat cctcgcgtcg gcatccggtc agtggcagtg aagggcgaac agttcctgat 60
taaccacaaa ccgttctact ttactggctt tggtcgtcat gaagatgcgg acttgcgtagg 120
caaaggattc gataacgtgc tgatggtgca cgaccacgca ttaatggact ggattggggc 180
caactcctac cgtacctcgc attaccetta cgaagcttgg tacc 225

<210> 11

<211> 216

<212> DNA

<213> Artificial Sequence

<220>

<223> GUS sense region for the construct encoding the hairpin RNA
molecule hpGUS[G:U]

<400> 11

ccctcgagtt gtgttggtag ttggttagtg gtagtgaagg gtgaatagtt tttgattaat 60
tataaattgt tttattttat tggttttggt tggtatgaag atgtggattt gtgtggtaaa 120
ggatttgata atgtgttgat ggtgtatgat tatgtattaa tggattggat tggggttaat 180
ttttattgta ttttgtatta tttttatggg tacc 216

<210> 12

<211> 216

<212> DNA

<213> Artificial Sequence

<220>

<223> GUS sense region for constructs encoding the hairpin RNA molecule
hpGUS[1:4]

<400> 12

ccctcgagtc ggggccgcaa ccgtcactg ggagtcaagc gcgtacactt cgtgaataag 60
cactaacggt tgtacattag tgggtttcgt cctcaagaac atggggagtt ggggtgccaat 120
ggaatcgta aggtggtgaa ggtccaccac ctcgctttat tggctctcat tcggggcaag 180
tccaacccta cgtcggattt cccataccgg tacc 216

<210> 13

<211> 216

<212> DNA

<213> Artificial Sequence

<220>

<223> GUS sense region for constructs encoding the hairpin RNA molecule hpGUS[2:10]

<400> 13

ccctcgagtc gcgtcgcgat ccggtctctg gcagtgttgg gcgaactctt cctgatatac 60
cacaaagggt tctactaaac tggcttacgt cgtcatctag atgcggtgtt gcgtgggtaa 120
ggattcctta acgtgcacat ggtgcagcac cacgcaaaaa tggactccat tggggcgtag 180
tcctacgcta cctcgtata cccttagcgg tacccc 216

<210> 14

<211> 240

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA sequence of nucleotides 781-1020 of the protein coding region of the GUS gene

<400> 14

gagtgtgata tctacccgct tcgctcggc atccggtcag tggcagtga gggcgaacag 60
ttcctgatta accacaaacc gttctacttt actggctttg gtcgtcatga agatgcggac 120
ttgcgtggca aaggattcga taacgtgctg atggtgcacg accacgcatt aatggactgg 180
attggggcca actcctaccg tacctcgcat tacccttacg ctgaagagat gctcgactgg 240

<210> 15

<211> 463

<212> RNA

<213> Artificial Sequence

<220>

<223> hairpin structure (including its loop) of the hpGUS[wt] RNA

<400> 15

ggauccucgc gucggcaucc ggucaguggc agugaagggc gaacaguucc ugauuaacca 60
caaaccguuc uacuuuacug gcuuuggucg uaugaagau gcggacuugc guggcaaagg 120
auucgauaac gugcugaugg ugcacgacca cgcauuuau gacuggauug gggccaacuc 180
cuaccguacc ucgcauuacc cuuacgaagc uugguacccc agcuuguugg gaagcugggu 240
ucgaaaucga uaagcuucgu aaggguuau cgagguacgg uaggaguugg ccccauucca 300
guccauuaau gcguggucgu gcaccaucag cacguuauugc aauccuuugc cacgcaaguc 360
cgcaucuuca ugacgaccaa agccaguaaa guagaacggu uugugguuaa ucaggaacug 420
uucgccccuuc acugccacug accggaugcc gacgcgagga ucc 463

<210> 16

<211> 457

<212> RNA

<213> Artificial Sequence

<220>

<223> hairpin structure (including its loop) of the hpGUS[G:U] RNA

<400> 16

```
cucgaguugu guugguauuu gguuaguggu agugaagggu gaauaguuuu ugauuaauua 60
uaaaauuguuu uauuuuauug guuuuugguug uuaugaagau guggauuugu gugguaaagg 120
auuugauaaau guguugaugg uguaugauua uguauuaaug gauuggauug ggguaaauuu 180
uuauuguaauu uuguauuuuu uuuaugggua cccagcuug uugggaagcu ggguucgaaa 240
ucgauaagcu ucguaagggu aaugcgaggu acgguaggag uggccccaa uccaguccau 300
uaaugcgugg ucgugcacca ucagcacguu aucgaauccu uggccacgca aguccgcauc 360
uucaugacga ccaaagccag uaaaguagaa cgguuugugg uaaucagga acuguucgcc 420
cuucacugcc acugaccgga ugccgacgcg aggaucc 457
```

<210> 17

<211> 457

<212> RNA

<213> Artificial Sequence

<220>

<223> hairpin structure (including its loop) of the hpGUS[1:4] RNA

<400> 17

```
cucgagucgg guccgcaacc gcucacuggg agucaagcgc guacacuucg ugaauaagca 60
cuaacggguug uacauuagug gguuucgucc ucaagaacau ggggaguugg gugccaaugg 120
aaucguuaag guggugaagg uccaccaccu cgcuuuauug gucugcauuc ggggcaaguc 180
caacccuacg ucggauuucc cauaccgua cccagcuug uugggaagcu ggguucgaaa 240
ucgauaagcu ucguaagggu aaugcgaggu acgguaggag uggccccaa uccaguccau 300
uaaugcgugg ucgugcacca ucagcacguu aucgaauccu uggccacgca aguccgcauc 360
uucaugacga ccaaagccag uaaaguagaa cgguuugugg uaaucagga acuguucgcc 420
cuucacugcc acugaccgga ugccgacgcg aggaucc 457
```

<210> 18

<211> 457

<212> RNA

<213> Artificial Sequence

<220>

<223> hairpin structure (including its loop) of the hpGUS[2:10] RNA

<400> 18

```
cucgagucgc gucgcaucc ggucucuggc aguguugggc gaacucuucc ugauauacca 60
caaagggguuc uacuaaacug gcuuacgucg ucaucuagau gcgguguugc guggguaagg 120
auuccuuaac gugcacaugg ugcagcacca cgcaaaaug gacuccauug gggcguacuc 180
```

cuacgcuacc ucgcuaauacc cuuagcggua cccagcuug uugggaagcu gggaucaaa 240
 ucgauaagcu ucguaaaggu aaugcgaggu acgguaggag uuggcccaaa uccaguccau 300
 uaaugcgugg ucgugcacca ucagcacguu aucgaauccu uugccacgca aguccgcauc 360
 uucaugacga ccaaagccag uaaaguagaa cgguuugugg uuaaucagga acuguucgcc 420
 cuucacugcc acugaccgga ugccgacgcg aggaucc 457

<210> 19

<211> 4851

<212> DNA

<213> *Arabidopsis thaliana*

<400> 19

atctctctct ttcgatggaa ctgagctctt tctctctttt ctctctcttt ctctctctat 60
 ctctatctct cgtagcttga taagagtttc tctcttttga agatccgttt ctctctctct 120
 cactgagact attgttggtta ggtcaacttg cgatcatggc gatttcgaag gtctgaagct 180
 gatttcgaat ggtttgaga tatccgtagt ggttaagcat atggaagtct atgttctgct 240
 cttggttgct ctgttagggc ttctccatt tggaccaact tagctgaatg ttgtatgac 300
 tctctccttg aagcagcaaa taagaagaag gtctggctct taacttaaca tctggttact 360
 agaggaaact tcagctatta ttaggttaaag aaagactgta cagagttgta taacaagtaa 420
 gcgtagagt ggctttgttt gcctcggtag tagaagaacc gactgattcg ttgttggtgtg 480
 ttagcttttg agggaaatcag atttcgcgag ggaaggtgtt ttagatcaaa tctgtgaatt 540
 ttactcaact gaggctttta gtgaaccacg actgtagagt tgacctgaa tcctactctg 600
 agtaattata ttatcagata gatttaggat ggaagctgaa attgtgaatg tgagacctca 660
 gctaggggtt atccagagaa tggttcctgc tctacttctt gtccttttgg tttctgtcgg 720
 atatatatgat cccgggaaat gggttgcaaa tatcgaagga ggtgctcgtt tcgggtatga 780
 cttggtaggca attactctgc ttttcaattt tgccgccatc ttatgccaat atgttgacgc 840
 tcgcataagc gttgtgactg gtaaactctt ggctcagatc tgcaatgaag aatatgacaa 900
 gtggacgtgc atgttcttgg gcattcaggc ggagttctca gcaattctgc tcgaccttac 960
 catggttggt ggagttgcgc atgcacttaa cttttgttt ggggtggagt tatccactgg 1020
 agtggttttg gccgccatgg atgcgttttt atttctgtt ttgcctctt tccttgaaaa 1080
 tggtaggca aatacagtat ccatttactc tgcaggcctg gtattacttc tctatgtatc 1140
 tggcgtcttg ctgagtcagt ctgagatccc actctctatg aatggagtgt taactcggtt 1200
 aaatggagag agcgcatcgc cactgatggg tcttcttggc gcaagcatcg tccctcacia 1260
 tttttatata cattcttatt ttgctgggga aagtacatct tcgtctgatg tcgacaagag 1320
 cagcttggtg caagaccatt tgttcgcat ctttggtgtc ttcagcggac tgtcacttgt 1380
 aaattatgta ttgatgaatg cagcagctaa tgtgtttcac agtactggcc ttgtggtact 1440
 gacttttcac gatgccttgt cactaatgga gcaggtatct atgagtcgc tcattccagt 1500
 ggtctttttg atgctcttgt tcttctctag tcaaattacc gcactagctt gggctttcgg 1560
 tggagaggtc gtcctgcatg acttctgaa gatagaaata cccgcttggc ttcactgtgc 1620
 tacaatcaga attcttgacg ttgctcctgc gctttattgt gtatggacat ctggtgcaga 1680
 cggaatatac cagttactta tattcaccca ggtcttggtg gcaatgatgc ttccttgctc 1740

```

ggtaataaccg cttttccgca ttgcttcgtc gagacaaatc atgggtgtcc ataaaatccc 1800
tcaggttggc gagttcctcg cacttacaac gtttttggga tttctggggt tgaatgttgt 1860
ttttgttggt gagatgggat ttgggagcag tgactgggct ggtggtttga gatggaatac 1920
cgtgatgggc acctcgattc agtacaccac tctgcttgta tcgtcatgtg catccttatg 1980
cctgatactc tggctggcag ccacgccgct gaaatctgcg agtaacagag cggaagctca 2040
aatatggaac atggatgctc aaaatgcttt atcttatcca tctgttcaag aagaggaaat 2100
tgaaaagaaca gaaacaagga ggaacgaaga cgaatcaata gtgcggttgg aaagcagggt 2160
aaaggatcag ttggatacta cgtctgttac tagctcggtc tatgatttgc cagagaacat 2220
tctaatacag gatcaagaaa tccgttcgag ccctccagag gaaagagagt tggatgtaaa 2280
gtactctacc tctcaagtta gtagtcttaa ggaagactct gatgtaaagg aacagtctgt 2340
attgcagtca acagtgggta atgaggtcag tgataaggat ctgattgttg aaacaaagat 2400
ggcgaaaatt gaaccaatga gtctgttgga gaagattgtt agcatggaga ataacagcaa 2460
gtttattgaa aaggatgttg aaggggtttc atgggaaaca gaagaagcta ccaaagctgc 2520
tctacaagc aactttactg tcggatctga tggctctctt tcattccgca gcttaagtgg 2580
ggaaggggga agtgggactg gaagcctttc acggttgcaa ggtttgggac gtgctgcccg 2640
gagacactta tctgcgatcc ttgatgaatt ttggggacat ttatatgatt ttcattgggca 2700
attggttgct gaagccaggg caaagaaact agatcagctg tttggcactg atcaaaagtc 2760
agcctcttct atgaaagcag attcgtttgg aaaagacatt agcagtggat attgcatgtc 2820
accaactgcg aagggaatgg attcacagat gacttcaagt ttatatgatt cactgaagca 2880
gcagaggaca ccgggaagta tcgattcggt gtatggatta caaagagggt cgtcaccgtc 2940
accgttggtc aaccgtatgc agatgttggg tgcatatggt aacaccacta ataataataa 3000
tgcttacgaa ttgagtgaga gaagatactc tagcctgcgt gctccatcat cttcagaggg 3060
ttgggaacac caacaaccag ctacagttca cggataccag atgaagtc atgtagacaa 3120
tttggcaaaa gaaaggcttg aagccttaca atcccgtgga gagatccga catcgagatc 3180
tatggcgctt ggtacattga gctatacaca gcaacttgct ttagccttga aacagaagtc 3240
ccagaatggt ctaaccctg gaccagctcc tgggtttgag aattttgctg ggtctagaag 3300
catatcgca caatctgaaa gatcttatta cgggtgttcca tcttctggca atactgatac 3360
tgtttggcga gcagtagcca atgagaaaaa atatagtagc atgccagata tctcaggatt 3420
gtctatgtcc gcaaggaaca tgcatttacc aaacaacaag agtggatact gggatccgtc 3480
aagtggagga ggagggtatg gtgcgtctta tggtcggtta agcaatgaat catcgttata 3540
ttctaatttg gggtcacggg tgggagtacc ctgacttat gatgacattt ctcaatcaag 3600
aggaggctac agagatgcct acagtttgcc acagagtga acaacaggga ccggatcgct 3660
ttggtccaga cagccctttg agcagtttgg tgtagcggag aggaatggtg ctgttggtga 3720
ggagctcagg aatagatcga atccgatcaa tatagacaac aacgcttctt ctaatgttga 3780
tgcagaggct aagcttcttc agtcgttcag gcaactgtatt cttaaagctta ttaaacttga 3840
aggatccgag tggttgtttg gacaaagcga tggagtgtat gaagaactga ttgaccgggt 3900
agctgcacga gagaagtta tctatgaagc tgaagctcga gaaataaacc aggtgggtca 3960
catgggggag ccactaattt catcggttcc taactgtgga gatggttgcg tttggagagc 4020
tgatttgatt gtgagctttg gagtttggtg cattcaccgt gtccttgact tgtctctcat 4080

```

ggagagtcgg cctgagcttt ggggaaagta cacttacgtt ctcaaccgcc tacagggagt 4140
 gattgatccg gcggttctcaa agctgcggac accaatgaca ccggtgctttt gccttcagat 4200
 tccagcgagc caccagagag cgagtccgac ttcagctaac ggaatgttac ctccggctgc 4260
 aaaaccggct aaaggcaaatt gcacaaccgc agtcacactt cttgatctaa tcaaagacgt 4320
 tgaaatggca atctcttgta gaaaaggccg aaccgggtaca gctgcagggtg atgtggcttt 4380
 cccaaagggg aaagagaatt tggcttcggt tttgaagcgg tataaacgtc ggttatcgaa 4440
 taaaccagta ggtatgaatc aggatggacc cggttcaaga aaaaacgtga ctgcgtacgg 4500
 atcattgggt tgaagaagaa gaacattgtg agaaatctca tgatcaaagt gacgtcgaga 4560
 gggaagccga agaatacaaaa ctctcgcttt tgattgctcc tctgcttcgt taattgtgta 4620
 ttaagaaaag aagaaaaaaa atggattttt gttgcttcag aatttttcgc tcttttttc 4680
 ttaatttgggt tgtaatgta tgtttatata catatatcat catcatagga ccatagctac 4740
 aaaccgaatc cggtttgtgt aattctatgc ggaatcataa agaatcgtc ggtttgaaat 4800
 gttaaattctc ctaaaccgga tctctgcacg tagctgacac atcgacgcta g 4851

<210> 20

<211> 1703

<212> DNA

<213> *Arabidopsis thaliana*

<400> 20

gttgcaaata tataaatcaa tcaaaagatt taaaaccac cattcaatct tggtaaagtaa 60
 cgaaaaaaaa gggaagcaag aagaaccaca gaaaaggggg ctaacaacta gacacgtaga 120
 tcttcatctg cccgtccatc taacctacca cactctcact ttctttttcc cgtgtcagtt 180
 tgttatataa gctctcactc tccggtatat ttccaaatac acctaacttg ttagtagacac 240
 aacagcaaca tcaaactcta ataaacccaa gttggtgtat actataatgg tgatggctgg 300
 tgcttcttct ttggatgaga tcagacaggc tcagagagct gatggacctg caggcatctt 360
 ggctattggc actgctaacc ctgagaacca tgtgcttcag gcggagtatc ctgactacta 420
 cttccgcac accaacagtg aacacatgac cgacctcaag gagaagttca agcgcatgtg 480
 cgacaagtcg acaattcgga aacgtcacat gcactgcac gaggaattcc tcaaggaaaa 540
 cccacacatg tgtgcttaca tggtctcttc tctggacacc agacaggaca tcgtggtggt 600
 cgaagtccct aagctaggca aagaagcggc agtgaaggcc atcaaggagt ggggccagcc 660
 caagtcaaag atcactcatg tcgtcttctg cactacctcc ggctgcgaca tgcttggtgc 720
 tgactaccag ctcaccaagc ttcttgggtc ccgtccttcc gtcaagcgtc tcatgatgta 780
 ccagcaaggt tgcttcgccg gcggtactgt cctccgtatc gctaaggatc tcgccgagaa 840
 caatcgtgga gcacgtgtcc tcgttgtctg ctctgagatc acagccgtta ccttccgtgg 900
 tccctctgac acccaccttg actcctcgt cggtcaggct cttttcagtg atggcgccgc 960
 cgcactcatt gtggggtcgg accctgacac atctgtcggg gagaaacca tctttgagat 1020
 ggtgtctgcc gctcagacca tccttcaga ctctgatggt gccatagacg gacatttgag 1080
 ggaagtgggt ctcacctcc atctcctcaa ggatgttccc ggcctcatct ccaagaacat 1140
 tgtgaagagt ctagacgaag cgtttaaac tttggggata agtgactgga actccctctt 1200
 ctggatagcc caccctggag gtccagcgat cctagaccag gtggagataa agctaggact 1260

aaaggaagag aagatgaggg cgacacgtca cgtgttgagc gagtatggaa acatgtcgag 1320
cgcggtgcgtt ctcttcatac tagacgagat gaggaggaag tcagctaagg atggtgtggc 1380
cacgacagga gaaggggttg agtgggggtgt cttgtttggt ttcggaccag gtctcactgt 1440
ttagacagtc gtcttgacac gcgttcctct ctaaacagaa cgcttgccctt ctatctgcct 1500
acctacctac gcaaaacttt aatcctgtct tatgttttat ataataaat cattatatgt 1560
ttacgcaata attaaggaag aattcatttg atgtgatatg tgatatgtgc tggacaggtc 1620
tattcgactg tttttgtact ctcttttttg tgtcttttta caatattaaa tctatgggtc 1680
ttgaatgaca tcaaactcttt gtt 1703

<210> 21

<211> 229

<212> DNA

<213> *Arabidopsis thaliana*

<400> 21

cctcgaggat cctctagacc tcagctaggg tttatccaga gaatggttcc tgctctactt 60
cctgtccctt tggtttctgt cggatatatt gatcccgga aatgggttgc aaatatcgaa 120
ggaggtgctc gtttcgggta tgacttggtg gcaattactc tgcttttcaa ttttgccgcc 180
atcttatgcc aatatgttgc agctcgcata agcgttaagc ttggtaccc 229

<210> 22

<211> 218

<212> DNA

<213> Artificial Sequence

<220>

<223> fragment comprising the 200nt sense sequence of EIN2

<400> 22

cctcgagtct agattttagt tagggtttat ttagagaatg gtttttgttt tttttttgt 60
ttttttggtt tttgttgat atattgattt tgggaaatgg gttgtaaata ttgaaggagg 120
tgtttgtttt gggtatgatt tggtggtaat tttttgttt ttttaattttg ttgttatttt 180
atgttaatat gttgtagttt gtataagtgt tgggtaccc 218

<210> 23

<211> 230

<212> DNA

<213> *Arabidopsis thaliana*

<400> 23

cctcgaggat ccgttgtctg ctctgagatc acagccgtta cttccgttg tccctctgac 60
accaccttg actccctcgt cggtcaggt cttttcagt atggcgcgc cgcactcatt 120
gtggggtcgg accctgacac atctgtcgga gagaaacca tctttgagat ggtgtctgcc 180
gctcagacca tccttcaga ctctgatggt gctctagaag cttggtaccc 230

<210> 24

<211> 219

<212> DNA

<213> Artificial Sequence

<220>

<223> fragment comprising the 200nt sense sequence of CHS

<400> 24

```
cctcgagggtt gtttgttttg agattatagt tgttatTTTT tgtggTTTT ttgatattta 60
ttttgattttt tttgttggtt aggttttttt tagtgatggt gttgttgat ttattgtggg 120
gttggttttt gatataattt ttggagagaa atttatTTTT gagatggtgt ttgttgttta 180
gattatTTTT ttagattttg atggtgtcta gaggtaccc 219
```

<210> 25

<211> 219

<212> DNA

<213> Artificial Sequence

<220>

<223> fragment comprising the 200nt antisense sequence of EIN2

<400> 25

```
caagcttaat gtttatgtga gttgtaatat attggtataa gatggtggtt aaattgaaaa 60
gtagagtaat tgttattaag ttatatTTTga aatgagtatt ttttttgata tttgtaattt 120
atTTTTTggg attaatatat ttgatagaaa taaaaggat aggaagtaga gtaggaatta 180
ttTTTTTggat aaatttttagt tgaggttcta gaggatccc 219
```

<210> 26

<211> 219

<212> DNA

<213> Artificial Sequence

<220>

<223> fragment comprising the 200nt antisense sequence of CHS

<400> 26

```
caagcttcta gagtattatt agagtttTga aggatggttt gagtggtaga tattatttta 60
aagatgggtt tttttttgat agatgtgtta gggtttgatt ttataatgag tgtggtggtg 120
ttattattga aaagagtttg attgatgagg gagttaaggT ggggtgttaga gggattatgg 180
aaggtaatgg ttgtgatttt agagtagata atggatccc 219
```

<210> 27

<211> 300

<212> DNA

<213> *Arabidopsis thaliana*

<400> 27

```
agtaattata ttatcagata gatttaggat ggaagctgaa attgtgaatg tgagacctca 60
gctaggggtt atccagagaa tggttcctgc tctacttctt gtccttttgg tttctgtcgg 120
atatattgat cccgggaaat gggttgcaaa tatcgaagga ggtgctcggt tcgggtatga 180
```

cttgggtggca attactctgc ttttcaattt tgccgccatc ttatgccaat atgttgcagc 240
tcgcataagc gttgtgactg gtaaacactt ggctcagatc tgcaatgaag aatatgacaa 300
<210> 28
<211> 300
<212> DNA
<213> Arabidopsis thaliana
<400> 28
tccgtatcgc taaggatctc gccgagaaca atcgtggagc acgtgtcctc gttgtctgct 60
ctgagatcac agccgttacc ttccgtggtc cctctgacac ccaccttgac tccctcgctc 120
gtcaggctct tttcagtgat ggccgccg cactcattgt ggggtcggac cctgacacat 180
ctgtcggaga gaaacccatc tttgagatgg tgtctgccgc tcagaccatc cttccagact 240
ctgatgggtc catagacgga catttgaggg aagttggtct caccttccat ctctcaagg 300
<210> 29
<211> 240
<212> DNA
<213> Arabidopsis thaliana
<400> 29
tcttcattgc agatctgagc caagtgttta ccagtcacaa cgcttatgcg agctgcaaca 60
tattggcata agatggcggc aaaattgaaa agcagagtaa ttgccacaa gtcatacccg 120
aaacgagcac ctcttcgat atttgcaacc catttcccg gatcaatata tccgacagaa 180
acaaaaagga caggaagtag agcaggaacc attctctgga taaaccctag ctgaggtctc 240
<210> 30
<211> 300
<212> DNA
<213> Arabidopsis thaliana
<400> 30
gatggaaggt gagaccaact tccctcaa atgtccgtctat ggcaccatca gagtctggaa 60
ggatgggtctg agcggcagac accatctcaa agatggggtt ctctccgaca gatgtgtcag 120
gggtccgaccc cacaatgagt gcggcgggc cactactgaa aagagcctga ccgacgaggg 180
agtcaagggt ggtgtcagag ggaccacgga aggtaacggc tgtgatctca gagcagacaa 240
cgaggacacg tgctccacga ttgttctcgg cgagatcctt agcgatacgg aggacagtac 300
<210> 31
<211> 4035
<212> DNA
<213> Arabidopsis thaliana
<400> 31
atgggatcta gggttccaat agaaaccatc gaagaagacg gcgaattcga ttgggaagca 60
gcagtcaaag aaatcgactt ggcttgtctt aaaaccacaa acgcttcttc ttcttcgtca 120
tcccatttca ctcttttggc taatccacca attacggcaa atctcactaa gccacctgcg 180

aagagacaat ctactctcga taaattcatc ggcagaaccg aacataaacc ggagaatcat 240
caagttgttt ccgagtgtgg tgtaacgat aacgataata gtcctttagt tgggattgat 300
cctgaggcag ctaaaacttg gatttatcca gtgaatggga gtgttccttt aagagattat 360
cagtttgcta taacgaagac tgctttgttt tcgaatacat tgggtggcttt gcctacggga 420
cttggtaaaa cgcttatagc tgcggttggt atgtataatt acttcagatg gtttccacaa 480
ggtaaaatag tatttgcggc gccttctagg cctcttgta tgcagcagat tgaggcgtgt 540
cataatattg ttggaatacc acaagaatgg acgattgact tgacgggtca gacatgtcct 600
tcgaaaagag cttttttgtg gaaaagcaaa cgggttttct ttgtcactcc acaagtgtta 660
gagaaggata tacagtcagg aacatgtctt actaactact tggtttgctt ggtgatcgac 720
gaggcacatc gagctttagg gaattattct tattgtgttg tagttcgtga gttgatggcg 780
gtaccgatac agctgagaat actggctctt actgcaactc ctggatcaaa gacacaggcc 840
atccagggtg tcattgataa tttgcagata tccacacttg aatatcgaaa tgagagtgc 900
catgatgttt gcccttatgt ccacgacaga aaattagaag tcatcgaggt tcccttgggt 960
caagatgcag atgatgtatc gaaacgcctg tttcatgtta tacgtccata tgcagtcagg 1020
cttaaaaaact ttgggggtta tctaaataga gatatacaaa ctttaagtcc acacgaagta 1080
cttatggcaa gggataagtt tcgtcaagca cctctaccag gccttcccc tgtaaatcac 1140
ggagatgtag aatcttgctt tgcagctctt atcactcttt atcatattcg taagctcctt 1200
tctagtcatg gaataagacc agcgtatgag atgctagaag agaaattgaa agaagggcc 1260
tttgctaggt tgatgagtaa gaatgaagat attaggatga cgaagctttt gatgcagcaa 1320
aggttgtcac atggagcacc aagcccaaaa ttgtcgaaga tgttagaaat actggttgat 1380
catttcaaag tgaaagatcc gaagacatca cgggtcatta ttttctcaaa tttcagagga 1440
agcgtaaagag acataatgaa cgcattaagt aatattggag atatggtcaa agcaactgag 1500
tttattgggtc aaagttcagg taagacattg aaaggccagt cgcaaaaaat tcagcaggct 1560
gttttgagaa aatttagagc tgggggggtt aatgttattg tcgcaacatc tattggtgaa 1620
gaaggcttgg atatcatgga agttgacctt gttatatgtt ttgatgctaa tgtatctcct 1680
ctgaggatga ttcaacggat gggaagaact ggaaggaaaa ataatggctg agttgtagtt 1740
cttgcttggtg aaggatcaga aaagaacagc tatatgcgaa agcaagcaag tggacgggct 1800
attaaaaaac acatgcggaa tggaggaaca aatagtttta attttcatcc tagtccaagg 1860
atgattcccc atgtttataa gccagaagtt cagcatgttg agttttcaat caagcaattc 1920
gttccacgtg gaaagaaaact acaagaggag tatgccactg agactccagc tttccagaaa 1980
aagcttacac ctgcagagac gcatatgctc gctaagtatt acaacaacc ccatgaggaa 2040
aagttgagag tgccttaat tgcgttcctt cacttcaga cattgccatc caaggtgcac 2100
aaagtaatgc attcacgtca aacaggcatg ttaattgacg ctatgcagca cttgcaagag 2160
ccaacttttt cagaacagag taaaagctt ttcactgagt ttcgagctcc tttgggtgaa 2220
agagaagagc ttgatacagg tctgagggtt actaatgatc caaaagatct acactctgtc 2280
cgtgatttgg aagtcaacac atcacagaga aaggcaaaac aagttgaatc tcccacaagc 2340
accttagaga caacagagaa ggattacgaa gaatcttcac ccacacaccg ttatcttttc 2400
agttcagaat gtgcatccgt tgatactctg gggaacgtct tcgtaatgcc agttcctctt 2460
ttattctttc ctaatgttct ggagtcagac aatacgctc tgcctaaaac agaaaaacaa 2520

cattcttgcc ggaatacatc tcacattgac ttagttccag tagatacttc ggaaaaacat 2580
 cggcaagata atatctcatg caagttaaag gaaagattct cgccagacgg tgccagcgag 2640
 aactagaga ctcatagcct tgtgaaaagg aactccacca gagtaggtga agatgatgta 2700
 gcgaattctg ttggagaaat tgtgttatca tcggatgaag atgactgtga gggattggag 2760
 cttagtccac ggctcactaa cttcatcaag agcggcattg ttccagagtc acctgtctat 2820
 gaccaagggg aagcgaacag agaagaagat cttgaatttc ctcagctttc ttcacccatg 2880
 aggttcagta acgaattggc aggagagtct tctttccctg agagaaaggt tcagcataag 2940
 tgcaacgatt ataacattgt gtctacaacc actgaattga gaactcctca gaaggaggt 3000
 ggtttggcca acggaacaga atgcttggct gtttctccta ttcctgagga ttggagaact 3060
 cccttggcga atctgacaaa cacaaacagc agcgtcgcga aagattggcg ggtgagttct 3120
 ggagaaaagt tagaaaactct tcgacagcct cgcaagtga agagactacg tagacttgga 3180
 gattgctcga gtgctgtaaa ggagaattat cctggatta cagaggcaga ccatatcaga 3240
 tctcgttctc gcggtaaaaa gcacattaga ggtaagaaga agatgatcat ggatgatgat 3300
 gtccaagtct tcattgacga ggaagctgag gtctcttcgg gagcagagat gtcggctgat 3360
 gagaacgaag atgtgactgg cgattcattt gaagatagtt tcatagatga cggaacaatg 3420
 cctacagcaa atactcaagc cgagtctggg aaagttgaca tgatggctgt ttacaggcgt 3480
 tctcttctca gccagtcacc attaccggcg agatttcgtg atttagccgc atcaagtctg 3540
 agtccttatt ctgctggacc cttgacgaga ataaatgaga gcagaagcga ctcagataaa 3600
 tcattgtctt ctcttcgaac accaaaaaca acaaactctg agtcaaacca agatgcaatg 3660
 atgataggaa accttctggg agtacaatc tcgtcagata gccgaaaag gaaatttagc 3720
 ttatgcaact cggcgaatgc ccccgatgatt aacttagaaa gcaagtttgc agctcatgca 3780
 caagccacgg agaaggaaag ccatgaaggc gtgagaagca atgcaggtgc gttagagtac 3840
 aatgatgatg atgatgatgc attctttgcg aactagact ttgatgcaat ggaagcaca 3900
 gccacattgt tattgtcga acagagatcc gaagcaaaag agaaagaaga cgcaacggtt 3960
 atacctaata caggcatgca gagaagtgat ggtatggaga aagatgcacc atcttttgat 4020
 cttggctctgt ggtga 4035

<210> 32

<211> 2310

<212> DNA

<213> Brassica napus

<400> 32

atgtcaaatg aaaataaaaa tataaaaact aaatttcac ctagttcaag gatgattccc 60
 catgtttata agccagaagt tcagcatgtt aagttttcga tcgagcaatt cattccacgt 120
 ggaaagaagc tacaagatga gcttgccact gagactccag ctttcaagaa aaagcttaca 180
 ccggaagaga tggatatgct cgccaagtat ttcaaaccac acgaggaaaa gtggagagtt 240
 tccttgattg ctttccctca cttccaaaca ttgccatcca aagtgcacaa agtaatgcat 300
 tcacgcaaaa caagcatatt aattgatgct atgcagcatc tgcaagagac aactttgaca 360
 gagcaagta aaagtcttct cattaagtat ggagctcctt tggctgaaag agatgagctt 420
 gacgcaggtc tgagggttgg tgatgatccg aaagatttac cctcttccga tgatttgat 480

gtcaacacat cacagagaaa ggcaaaacaa attttagaat ctcccacaag cacatttagag 540
 actacagaga aggatattcga agcatcttca cccacacact gttatctttt cagttcagaa 600
 tgtgcgtccg ttgatactct ggggaaggtc tttgtattgc cggttcctct ctcatctct 660
 tctaattgtac cagggtcaga ctgcgtggga agagaaaaag aactttcttc cccgaataag 720
 tcccacactg acgttggtcc gatagatagt tcctcaaaac atcggcaaga taatatttca 780
 tgcaagttaa agcaaggatt cttgccagat tgtgccaacg agactttgga gtcccaaagc 840
 cttttgaaaa ggcaactccac cgatgtaggt aaaggagata tagagaattg tgctggagaa 900
 attatgatat catcggatga agaagacgac tgtgaggatt tggagcttag tccaaggctc 960
 actaacttca tcaagagtgg cgttggtcca gattcacctg tctatgacca agttgcatac 1020
 gaagcaaaca gagaagaaga cttgatctt ccacacacga gtttaactaa tgaattggca 1080
 gaagagccat cgacacctga gaaaaagggt cacattgctt ctacggccaa tgaattcaga 1140
 actcctcaga aggaagaaga tttagccaac gaaacagaaa gcttcgctgt ttctccaatg 1200
 cctgaggagt ggagaactcc cttggcgaat atcaccaacg caagcagcag cgctagcaaa 1260
 gattggcgcg tgagttcggg agaaaagtca gaaactcttc gacagcctcg caagttgaag 1320
 agacttcgta gacttgaga ttgctcgagt gctgtgaagg agaataatcc tggatttgca 1380
 aagacagacc atatcagatc tcgttctcgc agtgtaaaga acataagagg taaaatgatt 1440
 ctgtatttcc ttttgctctg tgttcaaggc aagaagaaga tacgcgcgga taataatgct 1500
 agaacttca ttgaagcgga agctgagtg tcttcggaat cagaaatgtc ggttgatgag 1560
 aacgtagatt tgaccagcga ttcatttgaa gatagcttca tagatgacgg tacaatgcct 1620
 acagcaaata ctcaagccga gtgtgctaaa gttgacatga tggccgttta caggtatata 1680
 tcgaatcaaa acaagtcttt cttctactat gatttactaa gaatcataag ctatggtttc 1740
 cacagacgtt ctctactcag ccaatcacca ttaccggcaa gatttcgtga ttagctgca 1800
 tcaagtccga gtccttattc ttctggtctc ttgaagacaa taaatgagag cagaagcgac 1860
 tcagataaat cattgtcttc tcttagaacc ccacaaacaa cgaacaacga gtcaaacaag 1920
 gatgcaatgg ccacaggaga ctttcggta gcacaaatct caacagacag ccgaaaaagg 1980
 aaattcagct tatgcaactc agcgaatgtc ccagtgatta acttgaaaa caagtttgaa 2040
 gctcatgcac aagccacgga gaaggaaagc catgaaggtc cgagaagcaa tgcaggtgca 2100
 tcacagtaca aggatgagga tgaagatgat gatgcattct acgcgacact ggactttgat 2160
 gccatggaag cgcatgcgac attgctattg tcgaaacaaa ggtcagaaac gaaaacaaaa 2220
 gaagatgcat cggtgaaacc tcatttgggc aatcagagga atgatggttt gccgaaggat 2280
 gggccatctt ttgatcttgg tttgtggtga 2310
 <210> 33
 <211> 1822
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> hpFANCM-At[wt]
 <400> 33
 ggctcgagaa ccgaattcta atacgactca ctatagggtc aggaacatgt cttactaact 60

```

acttggtttg cttggtgatc gacgaggcac atcgagcttt agggaattat tcttattgtg 120
ttgtagttcg tgagttgatg gcggtaccga tacagctgag aatactggct cttactgcaa 180
ctcctggatc aaagacacag gccatccagg gtatcattga taatttgcag atatccacac 240
ttgaatatcg aaatgagagt gaccatgatg tttgccctta tgtccccgac agaaaattag 300
aagtcacgca ggttcccttg ggtcaagatg cagatgatgt atcgaaacgc ctgtttcatg 360
ttatacgccc atatgcagtc aggcttaaaa actttggggg taatctaaat agagatatatac 420
aaactttaag tccacacgaa gtacttatgg caagggataa gtttcgtcaa gcacctctac 480
caggccttcc ccatgtaaat cacggagatg tagaatcttg ctttgcagct cttatcaggt 540
aaggaaaataa ttatttttctt ttttctttt agtataaaat agttaagtga tgttaattag 600
tatgattata ataatatagt tgttataatt gtgaaaaaat aatttataaa tatattgttt 660
acataaacia catagtaatg taaaaaaata tgacaagtga tgtgtaagac gaagaagata 720
aaagttgaga gtaagtatat tatttttaat gaatttgatc gaacatgtaa gatgatatac 780
tagcattaat atttgtttta atcataatag taattctagc tggtttgatg aattaaatat 840
caatgataaa atactatagt aaaaataaga ataaataaat taaaataata tttttttatg 900
attaatagtt tattatataa ttaaatatct ataccattac taaatatttt agtttaaaag 960
ttaataaata ttttgttaga aattccaatc tgcttgtaat ttatcaataa acaaaatatt 1020
aaataacaag ctaaagtaac aaataatatc aaactaatag aaacagtaat ctaatgtaac 1080
aaaacataat ctaatgctaa tataacaaag cgcaagatct atcattttat atagtattat 1140
tttcaatcaa cattcttatt aatttctaaa taatacttgt agttttatta acttctaaat 1200
ggattgacta ttaattaaat gaattagtcg aacatgaata aacaaggtaa catgatagat 1260
catgtcattg tgttatcatt gatcttacct ttggattgat tacagttgat aagagctgca 1320
aagcaagatt ctacatctcc gtgatttaca tggggaaggc ctggtagagg tgcttgacga 1380
aacttatccc ttgccataag tacttcgtgt ggacttaaag tttgtatatc tctattttaga 1440
ttaaccccaa agtttttaag cctgactgca tatggacgta taacatgaaa caggcgtttc 1500
gatacatcat ctgcatcttg acccaaggga acctcgatga cttctaattt tctgtcgggg 1560
acataagggc aaacatcatg gtcactctca tttcgatatt caagtgtgga tatctgcaaa 1620
ttatcaatga taccctggat ggctgtgtc tttgatccag gagttgcagt aagagccagt 1680
attctcagct gtatcggtac cgccatcaac tcacgaacta caacacaata agaataattc 1740
cctaaagctc gatgtgcctc gtcgatcacc aagcaaacca agtagttagt aagacatgtt 1800
cctgaccccg ggateccaagc tt 1822
<210> 34
<211> 1822
<212> DNA
<213> Artificial Sequence
<220>
<223> hpFANCM-At[G:U]
<400> 34
ggctcgagaa ccgaattcta atacgactca ctatagggtt aggaatatgt tttattaatt 60
atttggtttg tttggtgatt gatgaggtat attgagtttt agggaattat ttttattgtg 120

```

ttgtagtttg tgagttgatg gtggtattga tatagttgag aatattgggt tttattgtaa 180
 tttttggatt aaagatatag gttatttagg gtattattga taattttag atatttatat 240
 ttgaatattg aaatgagagt gattatgatg tttgttttta tgtttttgat agaaaattag 300
 aagttattga ggtttttttg ggtaagatg tagatgatgt attgaaatgt ttgttttatg 360
 ttatatgttt atatgtagtt aggtttaaaa attttgggt taatttaa atagatatat 420
 aaattttaag tttatatgaa gtatttatgg taagggataa gttttgttaa gtatttttat 480
 taggtttttt ttatgtaa attgagatg tagaattttg tttttagtt tttattaggt 540
 aaggaaataa ttattttctt ttttctttt agtataaa atgttaagtga tgtaattag 600
 tatgattata ataatatagt tgttataatt gtgaaaaa attttataa tatattgttt 660
 acataaaca catagtaatg taaaaaata tgacaagtga tgtgtaagac gaagaagata 720
 aaagttgaga gtaagtatat tttttta atgtttgatg gaacatgta gatgatatac 780
 tagcattaat atttgtttta atcataatag taattctagc tggtttgatg aattaaatat 840
 caatgataa atactatagt aaaaataaga ataaataa taaaataa tttttttatg 900
 attaatagtt tattatataa ttaaataat ataccattac taaatatttt agtttaaaag 960
 ttaataaata ttttgtaga aattccaatc tgcttgta ttatcaata acaaaatatt 1020
 aaataacaag ctaaagtaac aaataatc aaactaatg aaacagta ctatgtaac 1080
 aaaacataa ctatgctaa tataacaa cgcaagatc atcattttat atagtattat 1140
 tttcaatcaa cattcttatt aatttctaaa taatacttgt agttttatta acttctaat 1200
 ggattgacta ttaattaaat gaattagtcg aacatgaata aacaagtaa catgatagat 1260
 catgtcattg tgttatcatt gatcttcat ttggattgat tacagttgat aagagctgca 1320
 aagcaagatt ctacatctcc gtgatttaca tggggaaggc ctggtagagg tgcttgacga 1380
 aacttatccc ttgccataag tacttcgtgt ggacttaaag tttgtatatt tctatttaga 1440
 ttaaccccaa agtttttaag cctgactgca tatggacgta taacatgaaa caggcgtttc 1500
 gatacatcat ctgcatctg acccaaggga acctgatga cttctaattt tctgtcgggg 1560
 acataagggc aaacatcatg gtcactctca tttcgatatt caagtgtgga tatctgcaa 1620
 ttatcaatga taccctggat ggctgtgtc tttgatccag gagttgcagt aagagccagt 1680
 attctcagct gtatcggtac cgccatcaac tcacgaacta caacacaata agaataattc 1740
 cctaaagctc gatgtgcctc gtcgatcacc aagcaacca agtagttagt aagacatgtt 1800
 cctgaccccg gtagccaagc tt 1822

<210> 35

<211> 1818

<212> DNA

<213> Artificial Sequence

<220>

<223> hpFANCM-Bn[wt]

<400> 35

ggatccttgg tacctaatac gactcactat aggagaaat tatgatatca tcggatgaag 60
 aagacgactg tgaggatttg gagcttagtc caaggctcac taacttcac aagagtggcg 120
 ttgttccaga ttcacctgtc tatgaccaag ttgcatacga agcaaacaga gaagaagacc 180

ttgatcttcc acacacgagt ttaactaatg aattggcaga agagccatcg acacctgaga 240
 aaaaggttca cattgcttct acggccaatg aattcagaac cccaacgaag gaagaagatt 300
 tagccaacga aacagaaagc ttcgctgttt ctccaatgcc tgaggagtgg agaactccct 360
 tggcgaatat caccaacgca agcagcagcg ctagcaaaga ttggcgcgtg agttcgggag 420
 aaaagtcaga aactcttcga cagcctcgca agttgaagag acttcgtaga cttggagatt 480
 gctcgagtgc tgtgaaggag aataatcctg gtattgcaaa gacagacat atcgtaagga 540
 aataattatt ttcttttttc ctttttagtat aaaatagtta agtgatgtta attagtatga 600
 ttataataat atagttgtta taattgtgaa aaaataatth ataaatatat tgtttacata 660
 aacaacatag taatgtaaaa aaatatgaca agtgatgtgt aagacgaaga agataaaagt 720
 tgagagtaag tatattatth ttaatgaatt tgatcgaaca tgtaagatga tatactagca 780
 ttaatatthg ttttaatcat aatagtaatt ctagctgggt tgatgaatta aatatcaatg 840
 ataaaatact atagtaaaaa taagaataaa taaattaaaa taatatttht ttatgattaa 900
 tagtttatta tataattaaa tatctatacc attactaaat attthtagtht aaaagttaat 960
 aaatattht ttagaaattc caatctgctt gtaatttatc aataaacaaa atattaaata 1020
 acaagctaaa gtaacaaata atatcaaact aatagaaaca gtaatctaata gtaacaaaac 1080
 ataactaata gctaataata caaagcgcaa gatctatcat tttatatagt attattthtca 1140
 atcaacattc ttattaatth cttaaataata cttgtagtht tattaacttc taaatggatt 1200
 gactattaat taaatgaatt agtcgaacat gaataaacia ggtaacatga tagatcatgt 1260
 cattgtgtta tcattgatct tacatttgga ttgattacag gatatggtct gtctthtcaa 1320
 taccaggatt attctccttc acagcactcg agcaatctcc aagtctacga agtctcttca 1380
 acttgcgagg ctgtcgaaga gthtctgact thtctccga actcacgcgc caatctthtgc 1440
 tagcgtctgt gcttgcgttg gtgatattcg ccaaggaggt tctccactcc tcaggcattg 1500
 gagaaacagc gaagcttht gthtctgttg ctaaatcttc thtcttcgtt ggggtthtga 1560
 attcattggc cgtagaagca atgtgaacct thtctcagg tgatgatggc thtcttgcca 1620
 attcattagt taaactctg tgtggaagat caaggtcttc thtctgttht gcttctgatg 1680
 caacttggtc atagacaggt gaatctggaa caacgccact cttgatgaag ttagtgagcc 1740
 ttggactaag ctccaaatcc tcacagtcgt thtcttcac cgatgatatc ataattthtct 1800
 tctagaaagg atccccggg 1818

<210> 36

<211> 1818

<212> DNA

<213> Artificial Sequence

<220>

<223> hpFANCM-Bn[G:U]

<400> 36

ggatccttgg tacctaatac gactcactat agggagaaat tatgatatta ttggatgaag 60
 aagatgattg tgtggatttg gagthtagtht taaggttht taattthatt aagagtgggtg 120
 ttgtthttaga thtatttgtht tatgattaaag ttgtatatga agtaaatagt gaagaagatt 180
 ttgattthttht atatatgagt ttaattaatg aattggtaga agagthattg atattthgaga 240

aaaaggttta tattgttttt atggttaatg aatttagaat tttaatgaag gaagaagatt 300
 tagttaatga aatagaaagt tttgttggtt ttttaatggt tgaggagtgg agaatttttt 360
 tgggtgaatat tattaatgta agtagtagtg ttagtaaaga ttggtgtgtg agtttgggag 420
 aaaagttaga aattttttga tagttttgta agttgaagag atttttaga tttggagatt 480
 gtttgagtgt tgtgaaggag aataattttg gtattgtaaa gatagattat attgtaagga 540
 aataattatt ttcttttttc ctttttagtat aaaatagtta agtgatgta attagtatga 600
 ttataataat atagttgtta taattgtgaa aaaataattt ataaatatat tgtttacata 660
 aacaacatag taatgtaaaa aaatatgaca agtgatgtgt aagacgaaga agataaaaagt 720
 tgagagtaag tatattatgt ttaatgaatt tgatcgaaca tgtaagatga tatactagca 780
 ttaatatgtg ttttaatcat aatagtaatt ctagctgggt tgatgaatta aatatcaatg 840
 ataaaatact atagtaaaaa taagaataaa taaattaaaa taatattttt ttatgattaa 900
 tagtttatta tataattaaa tatctatacc attactaaat attttagttt aaaagttaat 960
 aaatatgttg ttagaaattc caatctgctt gtaattttatc aataaacaaa atattaaata 1020
 acaagctaaa gtaacaaata atatcaaact aatagaaaca gtaatctaata gtaacaaaac 1080
 ataactaat gctaataata caaagcgcaa gatctatcat tttatatagt attattttca 1140
 atcaacattc ttattaatgt ctaaataata cttgtagttt tattaacttc taaatggatt 1200
 gactattaat taaatgaatt agtcgaacat gaataaacia ggtaacatga tagatcatgt 1260
 cattgtgtta tcattgatct tacatttgga ttgattacag gatatggctt gtctttgcaa 1320
 taccaggatt attctccttc acagcactcg agcaatctcc aagtctacga agtctcttca 1380
 acttgcgagg ctgtcgaaga gtttctgact tttctccga actcacgcgc caatctttgc 1440
 tagcgctgct gcttgcttg gtgatattcg ccaagggagt tctccactcc tcaggcattg 1500
 gagaaacagc gaagctttct gtttcgttg ctaaactctt ttccttcgtt ggggttctga 1560
 attcattggc cgtagaagca atgtgaacct tttctcagg tgcgatggc tcttctgcca 1620
 attcattagt taaactcgtg tgtggaagat caaggtcttc ttctctgtt gcttcgtatg 1680
 caacttggc atagacagg gaatctggaa caacgccact cttgatgaag ttagtgagcc 1740
 ttggactaag ctccaaatcc tcacagtcgt cttcttcac cgatgatatc ataatttctc 1800
 tctagaaagg atccccggg 1818

<210> 37

<211> 2274

<212> DNA

<213> Brassica napus

<400> 37

atggttagtc tgcgctccac agaaaacact ccggttcgg aatggccag cgacggcaaa 60
 acggagaaaag atggctccgg cgactcacc acttctgttc tcagcgatga ggaaaactgt 120
 gaagagaaaa ctgctactgt tgctgtagag gaagagatac ttctagccaa gaatggagat 180
 tcgtctctta tctctgaggc catggctcag gaagaagagc agcttctcaa aatccgggaa 240
 gatgaagaga ttgctaaacg tgctgctggc tctggtgaag ctctgatct gaatgatact 300
 cagtttacta aacttgatga gctcttgacc caaaccagc tctactctga gtttctcctt 360
 gagaaaatgg aggatatcac caaaaatggg atagaagtg agacccaaaa ggccgagcct 420

gagcctgagc ctgagcccca gaagaaaggc cgtggacgta aaagaaaggc tgctcctcag 480
 ggcgacagta tgaaggctaa gaaagctgtt gctgctatga tttcaagatc caaagaaggc 540
 cgtgaatctg ccgactcaga tctgacagag gaagaaagag tcatgaaaga gcagggtgaa 600
 cttgttcctc ttctgactgg cggaaagtta aagtcttata agtcaaagg tgtcaaattg 660
 ctgatatcat tgtggcaaaa tggtttgaat ggaattttag ctgatcaaatt gggctcttga 720
 aagacaattc aaaccattgg tttcctatca cacctcaaag gaaatgggtt ggatgggtcca 780
 tatctagtca ttgccccact ctctactctt tcaaactgga tgaatgagat cgctagggtc 840
 acgccttcca ttaatgcaat catttaccat ggagataaga aagaaaggga tgagctcagg 900
 aagaggcaca tgcccagaac tgttggtccg aagtcccta tagtcataac ttcttatgag 960
 gttgctatga atgatgctaa aaagaatctg cggcactatc catggaaata tgttgtgatt 1020
 gatgagggtc acagggtgaa aaaccacaag tgtaaactgc tgaggggagct aagatacttg 1080
 aatatggaga acaaaacttct gctgacagga acacctctgc aaaataattt gtctgagctt 1140
 tggtcactgt tgaattttat tctgcctgac atctttgcat cacatgacga atttgaatca 1200
 tggtttgatt tttctggaaa gaataataat gaagcaacta aggaagaagg agaagagaaa 1260
 agaagagctc aagtgggtgc gaaacttcat aatatactac gacctttcat cctccggaga 1320
 atgaaatgtg atgttgagct ctcaattccc cggaaaaaag agattatcat ctatgctaca 1380
 atgacggacc atcagaagaa gttccaggaa catcttgtga accacacctt ggaagcacac 1440
 attagagatg atactgtccg aggtcatggc ttgaaggga agcttaacaa tcttgctatt 1500
 caacttcgaa agaactgcaa ccatcctgac cttcttgtgg ggcaactaga tggctcatat 1560
 ctctaccac ctttggaaaga cattgtggga cagtgcggtt aattccgctt attggagaga 1620
 ttgcttggtc gggtatttgc caaaaatcac agagtcctta tcttctccca gtggacaaaa 1680
 atactggaca ttatggatta ctacttcagt gagaaggggt ttgaggtttg ccgaatcgac 1740
 ggtagtgtga aactagaaga aaggagaaga cagatccaag aattcaatga tgagaagagc 1800
 aactgcagga tatttcttct cagtaccaga gccggaggac tcggaattaa tcttactgct 1860
 gcagatacat gcatcctcta cgatagcgat tggaaccctc aatggactt gcaagccatg 1920
 gacagatgcc acagaattgg tcagacaaaa cctgttcatg ttacaggct tgcgacggct 1980
 cagtcaatag agggccgagt tctgaaacga gcatacagta agcttaagct ggaacatgtg 2040
 gttattggca aggggcagtt tcatcaagaa cgtgccaagt cttcaacacc gttagaggaa 2100
 gatgacatac tggcgttgct taaggacgac gaaaatgctg aagataaact gatacaaacc 2160
 gacataagcg aggaggatct tgacagggtg cttgaccgta gtgatctgat gattacctta 2220
 ccgggcgaga ctcaagcaca tgaagctttt ccagtgaagg gtccgggttg ggaa 2274
 <210> 38
 <211> 1824
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> hpDDM1-Bn[wt]
 <400> 38
 ggatccttgg tacctaatac gactcactat aggagctgt tgctgctatg atttcaagat 60

ccaaagaagg ccgtgaatct gccgactcag atctgacaga ggaagaaaga gtcataaatg 120
 agcagggtga acttggttct cttctgactg gcggaaagtt aaagtcttat cagctcaaag 180
 gtgtcaaattg gctgatatca ttgtggcaaa atggtttgaa tggaatttta gctgatcaaa 240
 tgggtcttgg aaagacaatt caaacattg gtttcctatc accccaacaa ggaaatgggt 300
 tggatgggtcc atatctagtc attgccccac tctctactct ttcaaaagcg attggaacct 360
 tcaaatggac ttgcaagcca tggacagatg ccacagaatt ggtcagacaa aacctgttca 420
 tgtttacagg cttgcgacgg ctctagtcatt agagggccga gttctgaaac gagcatacag 480
 taagcttaag ctggaacatg tggttattgg caaggggcag tttcatcaag aacgtggtaa 540
 ggaaataatt attttctttt ttccttttag tataaaatag ttaagtatg ttaattagta 600
 tgattataat aatatagttg ttataattgt gaaaaataa ttataaata tattgtttac 660
 ataacaaca tagtaattga aaaaaatatg acaagtatg tgtaagacga agaagataaa 720
 agttgagagt aagtatatta tttttaatga atttgatcga acatgtaaga tgatatacta 780
 gcattaatat ttgttttaat cataatagta attctagctg gtttgatgaa ttaaatatca 840
 atgataaaat actatagtaa aaataagaat aaataaatta aaataatatt tttttatgat 900
 taatagttaa ttatataatt aaatatctat accattacta aatatttttag tttaaaagtt 960
 aataaatatt ttgttagaaa ttccaatctg cttgtaattt atcaataaac aaaatattaa 1020
 ataacaagct aaagtaacaa ataatatcaa actaatagaa acagtaatct aatgtaacaa 1080
 aacataatct aatgctaata taacaaagcg caagatctat cttttatat agtattattt 1140
 tcaatcaaca ttcttattaa tttctaaata atactttag ttttattaac ttctaaatgg 1200
 attgactatt aattaaatga attagtcgaa catgaataaa caagtaaca tgatagatca 1260
 tgtcattgtg ttatcattga tcttacattt ggattgatta cagtcacgtt cttgatgaaa 1320
 ctgccccctg ccaataacca catgttccag cttagctta ctgtatgctc gtttcagaac 1380
 tcggccctct attgactgag ccgtcgcaag cctgtaaaca tgaacagggt ttgtctgacc 1440
 aattctgtgg catctgtcca tggcttgcaa gtccatttga gggttccaat cgcttttgaa 1500
 agagtagaga gtggggcaat gactagatat ggaccatcca acccatttcc ttgttgggt 1560
 gataggaaac caatggtttg aattgtcttt ccaagacca tttgatcagc taaaattcca 1620
 ttcaaaccat ttgcccacaa tgatcagc catttgacac ctttgagctg ataagacttt 1680
 aactttccgc cagtcagaag aggaacaagt tcacctgct ctttcatgac tctttcttcc 1740
 tctgtcagat ctgagtcggc agattcacgg cttcttttg atcttgaaat catagcagca 1800
 acagcttcta gaaaggatcc cggt 1824

<210> 39

<211> 1823

<212> DNA

<213> Artificial Sequence

<220>

<223> hpDDM1-Bn[G:U]

<400> 39

ggatccttgg tacctaatac gactcactat agggttggtt ttgttatgat ttttaagattt 60
 aaagaagggt gtgaatttgt tgatttagat ttgatagagg aagaaagagt tatgaatgag 120

taggggtgaat ttgttttttt tttgattggt ggaaagttaa agttttatta gtttaaaggt 180
 gttaaagtgt tgatattatt gtggtaaaat ggtttgaatg gaattttagt tgattaaatg 240
 ggttttggaa agataattta aattattggt tttttattat attttaaagg aaatgggttg 300
 gatggtttat atttagttat tgttttattt tttatttttt taaaagtgat tggaattttt 360
 aaatggattt gtaagttagt gatagatggt atagaattgg ttagataaaa tttgtttatg 420
 tttatagggt tgtgatgggt tagttaatag aggggtgagt tttgaaatga gtatatagta 480
 agtttaagtt ggaatatgtg gttattggta aggggtagtt ttattaagaa tgtgtggtaa 540
 ggaaataatt attttctttt ttccttttag tataaaatag ttaagtgatg ttaattagta 600
 tgattataat aatatagttg ttataattgt gaaaaataa tttataaata tattgtttac 660
 ataaacaaca tagtaatgta aaaaaatatg acaagtgatg tgtaagacga agaagataaa 720
 agttgagagt aagtatatta tttttaatga atttgatcga acatgtaaga tgatatacta 780
 gcattaatat ttgttttaat cataatagta attctagctg gtttgatgaa ttaaatatca 840
 atgataaaat actatagtaa aaataagaat aaataaatta aaataatatt tttttatgat 900
 taatagttta ttatataatt aaatatctat accattacta aatatttttag tttaaaagtt 960
 aataaatatt ttgttagaaa ttccaatctg cttgtaattt atcaataaac aaaatattaa 1020
 ataacaagct aaagtaacaa ataatatcaa actaatagaa acagtaatct aatgtaacaa 1080
 aacataatct aatgctaata taacaaagcg caagatctat cattttatat agtattattt 1140
 tcaatcaaca ttcttattaa tttctaaata atactttagt ttttattaac ttctaaatgg 1200
 attgactatt aattaaatga attagtcgaa catgaataaa caaggtaaca tgatagatca 1260
 tgtcattgtg ttatcattga tcttacattt ggattgatta cagtcacgtt cttgatgaaa 1320
 ctgccccctg ccaataacca catgttccag ctttaagctta ctgtatgctc gtttcagaac 1380
 tcggccccct attgactgag ccgtcgcaag cctgtaaaca tgaacagggt ttgtctgacc 1440
 aattctgtgg catctgtcca tggcttgcaa gtccatttga gggttccaat cgcttttgaa 1500
 agagtagaga gtggggcaat gactagatat ggaccatcca acccatttcc ttgttgggtg 1560
 gataggaaac caatggtttg aattgtcttt ccaagaccca tttgatcagc taaaattcca 1620
 ttcaaaccat tttgccacaa tgatatcagc catttgacac ctttgagctg ataagacttt 1680
 aactttccgc cagtcagaag aggaacaagt tcacctgct ctttcatgac tctttcttcc 1740
 tctgtcagat ctgagtcggc agattcacgg ctttctttgg atcttgaaat catagcagca 1800
 acagttctag aaaggatccc ggg 1823

<210> 40

<211> 720

<212> DNA

<213> Artificial Sequence

<220>

<223> EGFP

<400> 40

atggtgagca agggcgagga gctgttcacc ggggtggtgc ccatcctggt cgagctggac 60
 ggcgacgtaa acggccacaa gttcagcgtg tccggcgagg gcgaggcgca tgccacctac 120
 ggcaagctga ccctgaagtt catctgcacc accggcaagc tgcccgtgcc ctggcccacc 180

ctcgtgacca ccttcaccta cggcgtgcag tgcttcagcc gctaccccga ccacatgaag 240
 cagcacgact tcttcaagtc cgccatgccc gaaggctacg tccaggagcg caccatcttc 300
 ttcaaggacg acggcaacta caagacccgc gccgaggtga agttcgaggg cgacaccctg 360
 gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcaacat cctggggcac 420
 aagctggagt acaactacaa cagccacaac gtctatatca tggccgacaa gcagaagaac 480
 ggcatacaagg tgaacttcaa gatccgccac aacatcgagg acggcagcgt gcagctcgcc 540
 gaccactacc agcagaacac ccccatcggc gacggccccg tgctgctgcc cgacaaccac 600
 tacctgagca cccagtcgcg cctgagcaaa gaccccaacg agaagcgcgga tcacatggtc 660
 ctgctggagt tcgtgaccgc cgccgggatac actctcggca tggacgagct gtacaagtaa 720

<210> 41

<211> 1262

<212> DNA

<213> Artificial Sequence

<220>

<223> hpEGFP[wt]

<400> 41

gctagctaat acgactcact atagggcagc agcacggggc cgtcgccgat gggggtgttc 60
 tgctggtagt ggtcggcgag ctgcacgctg ccgctctcga tgttggtggcg gatcttgaag 120
 ttcaccttga tgccgttctt ctgcttgctg gccatgatgt atacgttgtg gctgttgaag 180
 ttgtactcca gcttgtgccc caggatgttg ccgctctcct tgaagtcgat gcccttcagc 240
 tcgatgcggg tcaccagggt gtcgccctcg aacttcacct cggcgcgggt cttgtagttg 300
 ccgctcgtcct tgaagaagat ggtgcgctcc tggacgtagc cttcgggcat ggcggacttg 360
 aagaagtcgt gctgcttcat gtggtcgggg tagcggctga agcactgcac gccgtaagcg 420
 aaggtgggtca ctagtgtggg ccagggcacg ggcagcttgc cgggtggtgca gatgaacttc 480
 aggggtctaga ccgcgtcggc atccgggtcag tggcagtga gggcgaacag ttcctgatta 540
 gggggatgaa gctacctggt ccgaaccaca aaccgttcta ctttactggc tttggtcgtc 600
 atgaagatgc ggacttgctg ggcaaaggat tcgataacgt gctgatggtg caccgaccac 660
 cattaatgga ctttaccttt taatggggaa tgaagctacc tgggtccgaac tcctaccgta 720
 cctcgcatta cccttacgct gaagagatgc tcgactgggc agatgaacat ggcacgtat 780
 ttaggtgaca ctatagccct gaagttcatc tgcaccaccg gcaagctgcc cgtgccctgg 840
 cccacactag tgaccacctt cgcttacggc gtgcagtgtc tcagccgcta ccccgaccac 900
 atgaagcagc acgacttctt caagtccgcc atgcccgaag gctacgtcca ggagcgcacc 960
 atcttcttca aggacgacgg caactacaag acccgcgccg aggtgaagtt cgagggcgac 1020
 accctggtga accgcatcga gctgaagggc atcgacttca aggaggacgg caacatcctg 1080
 gggcacaagc tggagtacaa cttcaacagc cacaacgtat acatcatggc cgacaagcag 1140
 aagaacggca tcaagtgtaa cttcaagatc cgccacaaca tcgaggacgg cagcgtgcag 1200
 ctcgccgacc actaccagca gaacaccccc atcggcgacg gccccgtgct gctgccgtcg 1260

ac 1262

<210> 42

<211> 1262

<212> DNA

<213> Artificial Sequence

<220>

<223> hpEGFP[G:U]

<400> 42

```

gctagctaatacgcactcactatagggcagcagcacggggccgtcgccgatgggggtgttc60
tgctggtagtggtcggcgagctgcacgctgccgtcctcga tgttggtggcg gatccttgaag120
ttcaccttga tgccgttcttctgcttgctgcgccatgatgtatagcttgtg gctgttgaag180
ttgtactcca gcttgtgccc caggatgttgccgtcctcct tgaagtcgat gcccttcagc240
tcgatgcggttcaccagggtgtcgccctcg aacttcacctcggcgcggggtctttagattg300
ccgtcgtccttgaagaagatggtgcgctctggacgtagccttcgggcattgcgggacttg360
aagaagtcgtgctgcttcatgtggtcgggtagcggtgaagcactgcac gccgtaagcg420
aagggtgtca ctagtggtggccagggcacggcagcttgc cgggtgtgcatgatgaacttc480
aggggtctagacgcgtcggcatccgggtcagtggcagtgaaggcgcaacagttcctgatta540
gggggatgaagctacctggtccgaaccaca aaccgttctactttactggctttggtcgtc600
atgaagatgcggacttgctggcaaaggat tgcataacgtgctgatggtgcacgaccacg660
cattaatgga ctttaccttttaatggggaa tgaagctacc tgggtccgaactcctaccgta720
cctcgcatcacccttacgctgaagagatgctcgactgggcagatgaacatggcatcgat780
ttaggtgacactatagccttgaagtttatttgtattattgtaagttgtttgtgttttgg840
tttatatttagtattatttttgtttatggtgtgtagtggtttagttgttattttgattat900
atgaagtagtatgatttttttaagtttggtatgtttgaaggttatgttttagagtgtatt960
atgtttttttaaggatgatgtaattataagatttggtgttgaggatgaagtttgagggtgat1020
atgtttgtgaattgtattgattgaagggtattgattttaggaggatgtaatatattttg1080
gggtataagttggagtataattttaatagtataatgtatatattatggttgataagtag1140
aagaatgggttaaggtgaattttaagattgtttataattgaggatgtagtgtagtag1200
ttgtgtgattattattagtagaatattttattgggtgatgtttgtgtgtgtgtgtgtc1260
ac1262

```

<210> 43

<211> 1259

<212> DNA

<213> Artificial Sequence

<220>

<223> ledEGFP[wt]

<400> 43

```

gctagctaatacgcactcactatagggtgtgccctcgaacttcacctcggcgcggttctt60
gtagttgccgtcgtccttgaagaagatggtgcgctcctggacgtagccttcgggcatggc120
ggacttgaagaagtcgtgctgcttcatgtgtcgggtagcggtgaagcactgcacgcc180
gtaggtgaaggtggtcacgagggtggccaggcacgggcagcttgccggtggtgcagat240

```

gaacttcagg gtctagaccg cgtcggcatc cggtcagtgg cagtgaaggg cgaacagttc 300
 ctgattaggg ggatgaagct acctggtccg aaccacaaac cgttctactt tactggcttt 360
 ggtcgtcatg aagatgcgga cttgcgtggc aaaggattcg accctgaagt tcatctgcac 420
 caccggcaag ctgcccgtgc cctggccac cctcgtgacc accttcacct acggcgtgca 480
 gtgcttcagc cgctaccccc accacatgaa gcagcacgac ttcttcaagt ccgccatgcc 540
 cgaaggctac gtccaggagc gcaccatctt cttcaaggac gacggcaact acaagacccg 600
 cgccgaggtg aagttcgagg gcgacaccct ggtgaaccgc atcgagctga agggcatcga 660
 cttcaaggag gacggcaaca tcctggggca caagctggag tacaactaca acagccacaa 720
 cgtctatata atggccgaca agcagaagaa cggcatcaag gtgaacttca agatccgcca 780
 caacatcgag gacggcagcg tgcagctcgc cgaccactac cagcagaaca ccccatcg 840
 cgacggcccc gtgtgtgtgc ctaacgtgct gatggtgcac gaccacgcat taatggactt 900
 taccttttaa tggggaatga agctacctgg tccgaactcc taccgtacct cgcattacc 960
 ttacgtgaa gagatgctcg actgggcaga tgaacatggc atcgtggcag cagcacgggg 1020
 ccgtcgccga tgggggtgtt ctgctggtag tggtcggcga gctgcacgct gccgtcctcg 1080
 atgttggtgc ggatcttgaa gttcaccttg atgccgttct tctgcttgct ggccatgata 1140
 tagacgttgt ggctgttgta gttgtactcc agcttggtcc ccaggatgtt gccgtcctcc 1200
 ttgaagtcca tgcccttcag ctcgatgcgg ttcaccattg tcgggataca ctcgctgcac 1259

<210> 44

<211> 1259

<212> DNA

<213> Artificial Sequence

<220>

<223> ledEGFP[G:U]

<400> 44

gctagctaata acgactcact atagggtgtc gccctcgaac ttcacctcgg cgcggttctt 60
 gtagttgccg tcgtccttga agaagatggg gcgctcctgg acgtagcctt cgggcatggc 120
 ggacttgaag aagtcgtgct gttcatgtg gtctgggtag cggctgaagc actgcacgcc 180
 gtaggtgaag gtggtcacga ggggtggcca gggcacgggc agcttgccgg tgggtcagat 240
 gaacttcagg gtctagaccg cgtcggcatc cggtcagtgg cagtgaaggg cgaacagttc 300
 ctgattaggg ggatgaagct acctggtccg aaccacaaac cgttctactt tactggcttt 360
 ggtcgtcatg aagatgcgga cttgcgtggc aaaggattcg attttgaagt ttatttgtat 420
 tattggtgaag ttgtttgtgt tttggtttat tttgtgatt atttttattt atgggtgtgta 480
 gtgttttagt tgttattttg attatatgaa gtagtatgat ttttttaagt ttgttatgtt 540
 tgaaggttat gtttaggagt gtattatttt ttttaaggat gatggtatt ataagatttg 600
 tgttgaggtg aagtttgagg gtgatatttt ggtgaattgt attgagttga agggatttga 660
 ttttaaggag gatggttaata ttttggggta taagttggag tataattata atagttataa 720
 tgtttatatt atggttgata agtagaagaa tggatttaag gtgaatttta agatttgtta 780
 taatattgag gatggtagt tgtagtttgt tgattattat tagtagaata tttttatttg 840
 tgatggtttt gtgttgttgt ttaacgtgct gatggtgcac gaccacgcat taatggactt 900

taccttttaa tggggaatga agctacctgg tccgaactcc tacctacct cgcattaccc 960
ttacgctgaa gagatgctcg actgggcaga tgaacatggc atcgtggcag cagcacgggg 1020
ccgtcgccga tgggggtgtt ctgctggtag tggtcggcga gctgcacgct gccgtcctcg 1080
atgttggtggc ggatcttgaa gttcaccttg atgccgttct tctgcttgtc ggccatgata 1140
tagacgttgt ggctgttgta gttgtactcc agcttggtcc ccaggatgtt gccgtcctcc 1200
ttgaagtcga tgcccttcag ctcgatgcgg ttcaccattg tcgggataca ctcgtcgac 1259

<210> 45

<211> 200

<212> DNA

<213> Artificial Sequence

<220>

<223> hpGUS[G:U]

<400> 45

ttgtgttggg atttggttag tggtagtgaa gggtagtag tttttgatta attataaatt 60
gttttatattt attggttttg gttgttatga agatgtggat ttgtgttgta aaggatttga 120
taatgtgttg atggtgtatg attatgtatt aatggattgg attggggta atttttattg 180
tattttgtat tatttttatg 200

<210> 46

<211> 200

<212> DNA

<213> Artificial Sequence

<220>

<223> hpGUS[1:4]

<400> 46

tcgggtccgc aaccgctcac tgggagtcaa gcgcgtacac ttcgtgaata agcactaacg 60
gttgtacatt agtgggtttc gtcctcaaga acatggggag ttgggtgcca atggaatcgt 120
taaggtgggtg aaggtccacc acctcgcttt attggtctgc attcggggca agtccaaccc 180
tacgtcggat ttcccatacc 200

<210> 47

<211> 200

<212> DNA

<213> Artificial Sequence

<220>

<223> hpGUS[2:10]

<400> 47

tcgcgtcgcg atccggtctc tggcagtgtt gggcgaactc ttcctgatat accacaaagg 60
gttctactaa actggcttac gtcgtcatct agatgcggtg ttgcgtgggt aaggattcct 120
taacgtgcac atggtgcagc accacgcaaa aatggactcc attggggcgt actcctacgc 180
tacctcgcta tacccttagc 200

<210> 48

<211> 200

<212> DNA

<213> Artificial Sequence

<220>

<223> hpEIN2[G:U]

<400> 48

gatttttagtt agggtttatt tagagaatgg tttttgtttt attttttggtt tttttggttt 60
ttgttgata tattgatttt gggaaatggg ttgtaaata tgaaggaggt gtttgttttg 120
ggatatgattt ggtgtaatt attttgtttt ttaattttgt tgttatttta tgtaatatg 180
ttgtagtttg tataagtgtt 200

<210> 49

<211> 200

<212> DNA

<213> Artificial Sequence

<220>

<223> hpCHS[G:U]

<400> 49

gttgtttggt ttgagattat agttgttatt ttttggtgtt ttttgatat ttattttgat 60
ttttttgttg gttaggtttt ttttagtgat ggtgttggtg tatttattgt ggggttggtat 120
tttgatata ttgttgaga gaaatttatt tttgagatgg tgtttggtgt ttagattatt 180
tttttagatt ttgatggtgt 200

<210> 50

<211> 200

<212> DNA

<213> Artificial Sequence

<220>

<223> hpEIN2[G:U/U:G]

<400> 50

aatgtttatg tgagttgtaa tatattggta taagatgggt gtaaaattga aaagtagagt 60
aattgttatt aagttatatt tgaaatgagt attttttttg atatttgtaa ttattttttt 120
gggattaata tatttgatag aaattaaaag gataggaagt agagtaggaa ttattttttg 180
gataaatttt agttgaggtt 200

<210> 51

<211> 199

<212> DNA

<213> Artificial Sequence

<220>

<223> hpCHS[G:U/U:G]

<400> 51

gtattattag agtttggaag gatggtttga gtggtagata ttattttaaa gatggggtttt 60
tttttgatag atgtgttagg gtttgatttt ataatgagtg tgggtggtgtt attattgaaa 120
agagtttgat tgatgaggga gttaagggtg gtgttagagg gattatggaa ggtaatgggtt 180
gtgatttttag agtagataa 199

<210> 52

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> GUS-WT-F

<400> 52

cctcgaggat cctcgcgtcg gcatccggtc 30

<210> 53

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> GUS-WT-R

<400> 53

gggtaccaag cttcgtaagg gtaatgcgag gta 33

<210> 54

<211> 118

<212> DNA

<213> Artificial Sequence

<220>

<223> GUS-GU-F

<400> 54

ccctcgagtt gtgttggtat ttggttagtg gtagtgaagg gtgaatagtt ttgattaat 60
tataaattgt tttattttat tggttttggt tgttatgaag atgtggattt gtgtggta 118

<210> 55

<211> 118

<212> DNA

<213> Artificial Sequence

<220>

<223> GUS-GU-R

<400> 55

ggggtagcca taaaaataat acaaaataca ataaaaatta accccaatcc aatccattaa 60
tacataatca tacaccatca acacattatc aaatccttta ccacacaaat ccacatct 118

<210> 56
<211> 118
<212> DNA
<213> Artificial Sequence
<220>
<223> GUS-4M-F
<400> 56
ccctcgagtc ggggccgcaa ccgctcactg ggagtcaagc gcgtacactt cgtgaataag 60
cactaacggt tgtacattag tgggtttcgt cctcaagaac atggggagtt gggtgcca 118
<210> 57
<211> 118
<212> DNA
<213> Artificial Sequence
<220>
<223> GUS-4M-R
<400> 57
ggggtaccgg tatgggaaat ccgacgtagg gttggacttg ccccgaatgc agaccaataa 60
agcgagggtg tggaccttca ccaccttaac gattccattg gcaccaact ccccatgt 118
<210> 58
<211> 118
<212> DNA
<213> Artificial Sequence
<220>
<223> GUS-10M-F
<400> 58
ccctcgagtc gcgtcgcat ccggtctctg gcagtgttg gcgaactctt cctgatatac 60
cacaaagggt tctactaaac tggcttacgt cgtcatctag atgcggtgtt gcgtgggt 118
<210> 59
<211> 118
<212> DNA
<213> Artificial Sequence
<220>
<223> GUS-10M-R
<400> 59
ggggtaccgc taagggtata gcgaggtagc gtaggagtag gcccgaatgg agtccatttt 60
tgcgtggtgc tgcaccatgt gcacgttaag gaatcctac ccacgaaca ccgcatct 118
<210> 60
<211> 20
<212> DNA

<213> Artificial Sequence
<220>
<223> Forward primer (35S-F3)
<400> 60
tggctcctac aaatgccatc 20
<210> 61
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> Reverse primer (GUSwt-R2)
<220>
<221> R
<222> (3) .. (4)
<223> A or G
<220>
<221> R
<222> (9) .. (9)
<223> A or G
<220>
<221> R
<222> (13) .. (13)
<223> A or G
<400> 61
carraactrt tcrcccttca c 21
<210> 62
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> Forward primer (GUSgu-R2)
<400> 62
caaaaactat tcacccttca c 21
<210> 63
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> reverse primer (GUS4m-R2)

<220>
<221> R
<222> (4) .. (4)
<223> A or G
<220>
<221> R
<222> (7) .. (7)
<223> A or G
<220>
<221> R
<222> (9) .. (9)
<223> A or G
<220>
<221> R
<222> (13) .. (13)
<223> A or G
<220>
<221> R
<222> (15) .. (15)
<223> A or G
<220>
<221> R
<222> (19) .. (19)
<223> A or G
<400> 63
cacraarttrt acrcrccttra c 21
<210> 64
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Forward primer (35S-F2)
<400> 64
gaggatctaa cagaactcgc 20
<210> 65
<211> 23
<212> DNA
<213> Artificial Sequence
<220>

<223> reverse primer (35S-R1)
<400> 65
ctctccaaat gaaatgaact tcc 23
<210> 66
<211> 36
<212> DNA
<213> Artificial Sequence
<220>
<223> EIN2wt-F
<400> 66
cctcgaggat cctctagacc tcagctaggg tttatc 36
<210> 67
<211> 33
<212> DNA
<213> Artificial Sequence
<220>
<223> EIN2wt-R
<400> 67
gggtaccaag cttaacgctt atgcgagctg caa 33
<210> 68
<211> 32
<212> DNA
<213> Artificial Sequence
<220>
<223> CHSwt-F
<400> 68
cctcgaggat ccgttgtctg ctctgagatc ac 32
<210> 69
<211> 38
<212> DNA
<213> Artificial Sequence
<220>
<223> CHSwt-R
<400> 69
gggtaccaag cttctagagc accatcagag tctggaag 38
<210> 70
<211> 119
<212> DNA
<213> Artificial Sequence

<220>

<223> EIN2_{gu}-F

<400> 70

cctcgagtct agatttttagt tagggtttat ttagagaatg gtttttggtt tattttttgt 60
ttttttgggt tttgttggat atattgattt tgggaaatgg gttgtaaata ttgaaggag 119

<210> 71

<211> 119

<212> DNA

<213> Artificial Sequence

<220>

<223> EIN2_{gu}-R

<400> 71

gggtaccaac acttatacaa actacaacat attaacataa aataacaaca aaattaaaaa 60
acaaaataat taccacaaaa tcatacccaa aacaaacacc tccttcaata tttacaacc 119

<210> 72

<211> 119

<212> DNA

<213> Artificial Sequence

<220>

<223> CHS_{gu}-F

<400> 72

cctcgagggt gtttggtttg agattatagt tgttattttt tgtgggtttt ttgatattta 60
ttttgatattt tttgttgggt aggttttttt tagtgatggg gttgttgtat ttattgtgg 119

<210> 73

<211> 119

<212> DNA

<213> Artificial Sequence

<220>

<223> CHS_{gu}-R

<400> 73

gggtacctct agacaccatc aaaatctaaa aaaataatct aaacaacaaa caccatctca 60
aaaataaatt tctctccaac aaatatatca aaatccaacc ccacaataaa tacaacaac 119

<210> 74

<211> 120

<212> DNA

<213> Artificial Sequence

<220>

<223> asEIN2_{gu}-F

<400> 74

caagcttaat gtttatgtga gttgtaatat attggtataa gatggtggta aaattgaaaa 60
gtagagtaat tggtattaag ttatatattga aatgagtatt ttttttgata tttgtaattt 120

<210> 75

<211> 120

<212> DNA

<213> Artificial Sequence

<220>

<223> asEIN2gu-R

<400> 75

gggatcctct agaacctcaa ctaaaattta tccaaaaaat aattcctact ctacttccta 60
tcctttttaat ttctatcaaa tatattaatc ccaaaaaata aattacaaat atcaaaaaaaaa 120

<210> 76

<211> 120

<212> DNA

<213> Artificial Sequence

<220>

<223> asCHSgu-F

<400> 76

caagcttcta gagtattatt agagtttgga aggatgggtt gagtggtaga tattatttta 60
aagatgggtt tttttttgat agatgtgta gggtttgatt ttataatgag tgtggtggtg 120

<210> 77

<211> 120

<212> DNA

<213> Artificial Sequence

<220>

<223> asCHSgu-R

<400> 77

gggatccatt atctactcta aaatcacaac cattaccttc cataatccct ctaacaccca 60
ccttaactcc ctcatcaatc aaactctttt caataataac accaccacac tcattataaa 120

<210> 78

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> CHS-200-F2

<400> 78

gacatgcctg gtgctgacta 20

<210> 79

<211> 20

<212> DNA
<213> Artificial Sequence
<220>
<223> CHS-200-R2
<400> 79
ccttagcgat acggaggaca 20
<210> 80
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Actin2-For
<400> 80
tccctcagca cattccagca 20
<210> 81
<211> 22
<212> DNA
<213> Artificial Sequence
<220>
<223> Actin2-Rev
<400> 81
gatcccatc ataaaacccc ag 22
<210> 82
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
<223> Top-35S-F2
<220>
<221> Y
<222> (8) .. (8)
<223> C or T
<220>
<221> Y
<222> (11) .. (11)
<223> C or T
<220>
<221> Y
<222> (14) .. (14)

<223> C or T
<220>
<221> Y
<222> (17) .. (17)
<223> C or T
<400> 82
agaaaatytt ygtyaayatg gtgg 24
<210> 83
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
<223> Top-35S-R2
<220>
<221> R
<222> (4) .. (4)
<223> A or G
<220>
<221> R
<222> (6) .. (7)
<223> A or G
<220>
<221> R
<222> (9) .. (9)
<223> A or G
<220>
<221> R
<222> (12) .. (12)
<223> A or G
<400> 83
tcarrrara trtcacatca atcc 24
<210> 84
<211> 23
<212> DNA
<213> Artificial Sequence
<220>
<223> Link-35S-F2
<220>
<221> Y

<222> (1) .. (2)
<223> C or T
<220>
<221> Y
<222> (5) .. (5)
<223> C or T
<220>
<221> Y
<222> (10) .. (10)
<223> C or T
<400> 84
yyatyattgy gataaaggaa agg 23
<210> 85
<211> 22
<212> DNA
<213> Artificial Sequence
<220>
<223> Link-EIN2-R2
<220>
<221> R
<222> (6) .. (6)
<223> A or G
<220>
<221> R
<222> (14) .. (14)
<223> A or G
<400> 85
taattrccac caartcatac cc 22
<210> 86
<211> 22
<212> RNA
<213> Artificial Sequence
<220>
<223> sense si22
<400> 86
gcaagcugac ccugaaguuc au 22
<210> 87
<211> 22
<212> RNA

<213> Artificial Sequence

<220>

<223> antisense si22

<400> 87

gaacuucagg gucagcuugc cg 22

<210> 88

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Forward primer

<400> 88

ttttagtata tgtgctgccg 20

<210> 89

<211> 83

<212> DNA

<213> Artificial Sequence

<220>

<223> reverse primer

<400> 89

ctcgagttcc aaaaaagctg accctgaagt tcattctctt tgaagatgaa cttcagggtc 60

agccaaacaa ggcttttctc caa 83

<210> 90

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> forward primer

<400> 90

ttttagtata tgtgctgccg 20

<210> 91

<211> 83

<212> DNA

<213> Artificial Sequence

<220>

<223> reverse primer

<400> 91

ctcgagttcc aaaaaaataa gtcgcagcag tacaattctt tgaattgtac tgctgcgact 60

tatgaatacc gcttcctcct gag 83

<210> 92

<211> 88

<212> RNA

<213> Artificial Sequence

<220>

<223> dsRNA molecules

<400> 92

aagugauuuug ugaauuggauu cgguuaaguu agugauaggu aucacgcugg ccauuacuga 60
 caugaccgga uucguucgcg agucacuu 88

<210> 93

<211> 2653

<212> DNA

<213> Brassica napus

<400> 93

agaaaaatcg aaaaatgtga cagtgcgtct tttcacttaa taccctcggt ttgaatttgc 60
 tctcggaag cgtctgagag agtggtcggg gatttctccc gccgcttggg gttttttccg 120
 ttaccggaat atccttctcc tccgatgggt agtctgcgt ccacagaaaa cactccggct 180
 tcggaaatgg ccagcgacgg caaaacggag aaagatggct ccggcgactc acccacttct 240
 gttctcagcg atgaggaaaa ctgtgaagag aaaactgcta ctgttgctgt agaggaagag 300
 atacttctag ccaagaatgg agattcgtct cttatctctg aggccatggc tcaggaagaa 360
 gagcagcttc tcaaaatccg ggaagatgaa gagattgcta aacgtgctgc tggctctggt 420
 gaagctcctg atctgaatga tactcagttt actaaacttg atgagctctt gacccaaacc 480
 cagctctact ctgagtttct ccttgagaaa atggaggata tcaccaaaaa tgggatagaa 540
 ggtgagaccc aaaaggccga gcctgagcct gagcctgagc ccgagaagaa aggccgtgga 600
 cgtaaaagaa aggctgctcc tcagggcgac agtatgaagg ctaagaaagc tgttgctgct 660
 atgatttcaa gatccaaaga aggccgtgaa tctgccgact cagatctgac agaggaagaa 720
 agagtcataa aagagcaggg tgaacttggt cctcttctga ctggcggaag gttaaagtct 780
 tatcagctca aaggtgtcaa atggctgata tcattgtggc aaaatgggtt gaatggaatt 840
 ttagctgac aaatgggtct tggaaagaca attcaaacca ttggtttctc atcacacctc 900
 aaaggaaatg ggttgatgg tccatatcta gtcattgccc cactctctac tctttcaaac 960
 tggatgaatg agatcgctag gttcacgcct tccattaatg caatcattta ccatggagat 1020
 aagaaagaaa gggatgagct caggaagagg cacatgcccc gaactgttgg tccgaagttc 1080
 cctatagtc taacttctta tgaggttgct atgaatgatg ctaaaaagaa tctgcggcac 1140
 tatccatgga aatatgttgt gattgatgag ggtcacaggt tgaaaaacca caagtgtaaa 1200
 ctgctgaggg agctaagata cttgaatatg gagaacaaac ttctgctgcc aggaacacct 1260
 ctgcaaaata atttgtctga gtttcgggtc ctgttggaatt ttattctgcc tgacatcttt 1320
 gcatcacatg acgaatttga atcatgggtt gatttttctt gaaagaataa taatgaagca 1380
 actaaggaag aaggagaaga gaaaagaaga gctcaagtgg ttgcgaaact tcataatata 1440
 ctacgacctt tcacctccg gagaatgaaa tgtgatgttg agctctcact tccccgaaa 1500

```

aaagagatta tcatctatgc tacaatgacg gaccatcaga agaagttcca ggaacatctt 1560
gtgaaccaca ccttggaagc acacattaga gatgatactg tccgaggtca tggcttgaag 1620
ggaaagctta acaatcttgc tattcaactt cgaaagaact gcaaccatcc tgaccttctt 1680
gtggggcaac tagatggctc atatctctac ccacctttgg aagacattgt gggacagtgc 1740
ggtaaattcc gcttattgga gagattgctt gttcggttat ttgccaaaaa tcacagagtc 1800
cttatcttct cccagtggac aaaaatactg gacattatgg attactactt cagtgagaag 1860
gggtttgagg tttgccgaat cgacggtagt gtgaaactag aagaaaggag aagacagatc 1920
caagaattca atgatgagaa gagcaactgc aggatatttc ttctcagtac cagagccgga 1980
ggactcggaa ttaatcttac tgctgcagat acatgcatcc tctacgatag cgattggaac 2040
cctcaaatgg acttgcaagc catggacaga tgccacagaa ttggtcagac aaaacctgtt 2100
catgtttaca ggcttgcgac ggtcagtc atagagggcc gagttctgaa acgagcatac 2160
agtaagctta agctggaaca tgtggttatt ggcaaggggc agtttcatca agaacgtgcc 2220
aagtcttcaa caccgttaga ggaagatgac atactggcgt tgcttaagga cgacgaaaat 2280
gctgaagata aactgataca aaccgacata agcgaggagg atcttgacag ggtgcttgac 2340
cgtagtgatc tgatgattac cttaccgggc gagactcaag cacatgaagc ttttcagtg 2400
aaggggccgg gttgggaagt ggtctcgtct agctcagctg gagggatgct gtcttcctc 2460
aacagttaga accactcttt gcaaaaccac ttcggtgtgt ttttttttcc ggaacataac 2520
cggttacttt tgctgctac tcggaagttt taacttgaaa ccttggaac atctgatgaa 2580
aacaattgcg gatattatgt tattagacta tttatttatg ctttttgaaa tttggcagta 2640
attttttagt taa 2653

```

<210> 94

<211> 1773

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a hairpin RNAi (hpRNA) construct targeting a DDM1 gene of *B. napus*

<400> 94

```

gggagctgtt gctgctatga tttcaagatc caaagaaggc cgtgaatctg ccgactcaga 60
tctgacagag gaagaaagag tcatgaatga gcagggtgaa cttgttcctc ttctgactgg 120
cgaaaagtta aagtcttata agctcaaagg tgtcaaatgg ctgatatcat tgtggcaaaa 180
tggtttgaat ggaatttttag ctgatcaaat ggtctttgga aagacaattc aaaccattgg 240
tttcctatca cccaacaag gaaatgggtt ggatgggtcca tatctagtca ttgccccact 300
ctctactctt tcaaaaagcga ttggaaccct caaatggact tgcaagccat ggacagatgc 360
cacagaattg gtcagacaaa acctgttcat gtttacaggc ttgcgacggc tcagtcaata 420
gagggccgag ttctgaaacg agcatacagt aagcttaagc tggaacatgt ggttattggc 480
aaggggcagt ttcatacaaga acgtggtaag gaaataatta ttttcttttt tccttttagt 540
ataaaatagt taagtatgt taattagtat gattataata atatagttgt tataattgtg 600
aaaaaataat ttataaatat attgtttaca taaacaacat agtaatgtaa aaaaatatga 660

```

```

caagtgatgt gtaagacgaa gaagataaaa gttgagagta agtatattat ttttaatgaa 720
tttgatcgaa catgtaagat gatatactag cattaatatt tgttttaatc ataatagtaa 780
ttctagctgg tttgatgaat taaatatcaa tgataaaata ctatagtaaa aataagaata 840
aataaattaa aataatattt ttttatgatt aatagtttat tatataatta aatatctata 900
ccattactaa atatttttagt ttaaaagtta ataaatattt tgttagaaat tccaatctgc 960
ttgtaattta tcaataaaca aaatattaaa taacaagcta aagtaacaaa taatatcaaa 1020
ctaatagaaa cagtaatcta atgtaacaaa acataatcta atgctaatat aacaaagcgc 1080
aagatctatc attttatata gtattatttt caatcaacat tcttattaat ttctaaataa 1140
tacttgtagt tttattaaact tctaaatgga ttgactatta attaaatgaa ttagtcgaac 1200
atgaataaac aaggtaacat gatagatcat gtcatttgtt tatcattgat cttacatttg 1260
gattgattac agtcacgttc ttgatgaaac tgccccctgc caataaccac atgttccagc 1320
ttaagcttac tgtatgctcg tttcagaact cggccctcta ttgactgagc cgtcgcaagc 1380
ctgtaaacat gaacaggttt tgtctgacca attctgtggc atctgtccat ggcttgcaag 1440
tccatttgag ggttccaatc gcttttgaaa gagtagagag tggggcaatg actagatatg 1500
gaccatccaa cccatttcct tgttgggtg ataggaaacc aatggtttga attgtctttc 1560
caagacccat ttgatcagct aaaattccat tcaaaccatt ttgccacaat gatatcagcc 1620
atttgacacc tttgagctga taagacttta actttccgcc agtcagaaga ggaacaagtt 1680
caccctgctc tttcatgact ctttcttcct ctgtcagatc tgagtcggca gattcacggc 1740
cttctttgga tcttgaatc atagcagcaa cag 1773

```

<210> 95

<211> 1773

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a hairpin RNAi (hpRNA) construct with G:U basepairs, targeting a DDM1 gene of *B. napus*

<400> 95

```

gggttggtgt tggtatgatt ttaagattta aagaagggtg tgaatttggt gatttagatt 60
tgatagagga agaaagagtt atgaatgagt aggggtgaatt tgtttttttt ttgattgggtg 120
gaaagttaaa gttttattag tttaaagggtg ttaaatgggt gatattattg tggtaaaatg 180
gtttgaatgg aatttttagt gattaaatgg gttttggaaa gataatttaa attattgggtt 240
ttttattata ttttaaagga aatgggttgg atgggtttata tttagttatt gttttatttt 300
ttattttttt aaaagtgatt ggaattttta aatggatttg taagttagtg atagatgtta 360
tagaattggg tagataaaat ttgtttatgt ttatagggtt gtgatgggtt agttaataga 420
gggttgagtt ttgaaatgag tatatagtaa gtttaagttg gaatatgtgg ttattggtta 480
ggggtagttt tattaagaat gtgtggtaag gaaataatta ttttcttttt tccttttagt 540
ataaaatagt taagtgatgt taattagtat gattataata atatagttgt tataattgtg 600
aaaaaataat ttataaatat attgtttaca taaacaacat agtaatgtaa aaaaatatga 660
caagtgatgt gtaagacgaa gaagataaaa gttgagagta agtatattat ttttaatgaa 720

```

ttgatcgaa catgtaagat gatatactag cattaatatt tgttttaatc ataatagtaa 780
 ttctagctgg ttgatgaat taaatatcaa tgataaaata ctatagtaaa aataagaata 840
 aataaattaa aataatattt ttttatgatt aatagtttat tatataatta aatatctata 900
 ccattactaa atatttttagt ttaaaagtta ataaatattt tgtagaaat tccaatctgc 960
 ttgtaattta tcaataaaca aaatattaaa taacaagcta aagtaacaaa taatatcaaa 1020
 ctaatagaaa cagtaatcta atgtaacaaa acataatcta atgctaatat aacaaagcgc 1080
 aagatctatc attttatata gtattatttt caatcaacat tcttattaat ttctaaataa 1140
 tacttgtagt tttattaact tctaaatgga ttgactatta attaaatgaa ttagtcgaac 1200
 atgaataaac aaggtaacat gatagatcat gtcatttgtt tatcattgat cttacatttg 1260
 gattgattac agtcacgttc ttgatgaaac tgccccctgc caataaccac atgttccagc 1320
 ttaagcttac tgtatgctcg tttcagaact cggccctcta ttgactgagc cgtcgcaagc 1380
 ctgtaaacat gaacagggtt tgtctgacca attctgtggc atctgtccat ggcttgcaag 1440
 tccatttgag ggttccaatc gcttttgaaa gagtagagag tggggcaatg actagatatg 1500
 gaccatccaa cccatttcct tgttgggtg ataggaaacc aatggtttga attgtctttc 1560
 caagaccat ttgatcagct aaaattccat tcaaaccatt ttgccacaat gatatcagcc 1620
 atttgacacc ttgagctga taagacttta actttccgcc agtcagaaga ggaacaagtt 1680
 caccctgctc tttcatgact ctttcttcct ctgtcagatc tgagtcggca gattcacggc 1740
 cttctttgga tcttgaaatc atagcagcaa cag 1773

<210> 96

<211> 1260

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct, targeting a DDM1 gene of B. napus

<400> 96

ggggtgatag gaaaccaatg gtttgaattg tctttccaag acccatttga tcagctaaaa 60
 ttccattcaa accattttgc cacaatgata tcagccattt gacacctttg agctgataag 120
 actttaactt tccgccagtc agaagaggaa caagttcacc ctgctctttc atgactcttt 180
 cttcctctgt cagatctgag tcggcagatt cacggccttc tttggatctt gaaatcatag 240
 cagcaacagc tttcttagcc ttcatactgt cgccctgagg agcagccttt cttttacgtc 300
 cacggccttt cttctcgggc tcaggetcag gctcaggctc ggctttttgg gtctcacctt 360
 ctatccccatt tttggtgata gctgttgctg ctatgatttc aagatccaaa gaaggccgtg 420
 aatctgccga ctcagatctg acagaggaag aaagagtcat gaatgagcag ggtgaacttg 480
 ttctctttct gactggcgga aagttaaagt cttatcagct caaaggtgtc aaatggctga 540
 tatcattgtg gcaaaaatgg ttgaatggaa ttttagctga tcaaatgggt cttggaaaga 600
 caattcaaac cattggtttc ctatcacccc aacaaggaaa tgggttgat ggtccatatt 660
 tagtcattgc cccactctct actctttcaa aagcgattgg aaccctcaaa tggacttgca 720
 agccatggac agatgccaca gaattggtca gacaaaacct gttcatgttt acaggcttgc 780

gacggctcag tcaatagagg gccgagttct gaaacgagca tacagtaagc ttaagctgga 840
 acatgtgggtt attggcaagg ggcagtttca tcaagaacgt cactacggtc aagcacctg 900
 tcaagatcct cctcgcttat gtcggtttgt atcagtttat cttcagcatt ttcgtcgtcc 960
 ttaagcaacg ccagtatgtc atcttcctct aacgggtgtg aagacttggc acgttcttga 1020
 tgaaactgcc ccttgccaat aaccacatgt tccagcttaa gcttactgta tgctcgtttc 1080
 agaactcggc cctctattga ctgagccgtc gcaagcctgt aaacatgaac aggttttgtc 1140
 tgaccaattc tgtggcatct gtccatggct tgcaagtcca tttgagggtt ccaatcgctt 1200
 ttgaaagagt agagagtggg gcaatgacta gatatggacc atccaacca tttccttgtt 1260

<210> 97

<211> 4421

<212> DNA

<213> *Arabidopsis thaliana*

<400> 97

tcttccaaaa tttgcccgcc attctctgtg tctcttcgtc taagggtttc ttccaaagaa 60
 cgacgacaaa accactgaac ctaaaatccg aaatccaaaa gattcatgcg aaaaaatcgt 120
 taaagagtac caatttcaga aagttacatt cttacagaga acaagtaatt cccagaaat 180
 gggatctagg gttccaatag aaaccatcga agaagacggc gaattcgatt gggaagcagc 240
 agtcaaagaa atcgacttgg cttgtcttaa aaccacaaac gcttcttctt cttcgtcatc 300
 ccatttcact cctttggcta atccaccaat tacggcaaat ctactaagc cacctgcgaa 360
 gagacaatct actctcgata aattcatcgg cagaaccgaa cataaaccgg agaatcatca 420
 agttgtttcc gagtgtgggtg ttaacgataa cgataatagt ccttttagttg ggattgatcc 480
 tgaggcagct aaaacttggg tttatccagt gaatgggagt gttcctttaa gagattatca 540
 gtttgctata acgaagactg ctttgttttc gaatacattg gtggctttgc ctacgggact 600
 tggtaaaacg cttatagctg cggttgttat gtataattac ttcagatggg ttccacaagg 660
 taaaatagta tttgcggcgc cttctaggcc tcttgtgatg cagcagattg aggcgtgtca 720
 taatatgtgt ggaataccac aagaatggac gattgacttg acgggtcaga catgtccttc 780
 gaaaagagct tttttgtgga aaagcaaacg ggttttcttt gtcactccac aagtgttaga 840
 gaaggatata cagtcaggaa catgtcttac taactacttg gtttgcttgg tgatcgacga 900
 ggcacatcga gctttaggga attattctta ttgtgttgta gttcgtgagt tgatggcggt 960
 accgatacag ctgagaatac tggctcttac tgcaactcct ggatcaaaga cacaggccat 1020
 ccagggtatc attgataatt tgcagatata cacacttgaa tatcgaaatg agagtgacca 1080
 tgatgtttgc ccttatgtcc acgacagaaa attagaagtc atcgaggttc ctttgggtca 1140
 agatgcagat gatgtatcga aacgcctgtt tcatgttata cgtccatatg cagtcaggct 1200
 taaaaacttt ggggttaatc taaatagaga tatacaaact ttaagtccac acgaagtact 1260
 tatggcaagg gataagtttc gtcaagcacc tctaccaggc cttcccatg taaatcacgg 1320
 agatgtagaa tcttgctttg cagctcttat cactctttat catattcgta agctcctttc 1380
 tagtcatgga ataagaccag cgtatgagat gctagaagag aaattgaaag aagggccatt 1440
 tgctagggtg atgagtaaga atgaagatat taggatgacg aagcttttga tgcagcaaag 1500
 gttgtcacat ggagcaccaa gcccaaaatt gtcgaagatg ttagaaatac tggttgatca 1560

```

tttcaaagtg aaagatccga agacatcacg ggtcattatt ttctcaaatt tcagaggaag 1620
cgtaagagac ataatgaacg cattaagtaa tattggagat atgggtcaaag caactgagtt 1680
tattgggtcaa agttcaggta agacattgaa aggccagtcg caaaaaattc agcaggctgt 1740
tttggagaaa tttagagctg gggggttcaa tgttattgtc gcaacatcta ttggtgaaga 1800
aggcttggat atcatggaag ttgacctagt tatatgtttt gatgctaata tatctcctct 1860
gaggatgatt caacggatgg gaagaactgg aaggaaaaat aatggtcgag ttgtagtctt 1920
tgcttgtgaa ggatcagaaa agaacagcta tatgcgaaag caagcaagtg gacgggctat 1980
taaaaaacac atgcggaatg gaggaacaaa tagttttaat tttcatccta gtccaaggat 2040
gattcccat gtttataagc cagaagttca gcatgttgag ttttcaatca agcaattcgt 2100
tccacgtgga aagaaactac aagaggagta tgccactgag actccagctt tccagaaaaa 2160
gcttacacct gcagagacgc atatgctcgc taagtattac aacaaccccg atgaggaaaa 2220
gttgagagtg tccttaattg cgttccctca cttccagaca ttgccatcca aggtgcacaa 2280
agtaatgcat tcacgtcaaa caggcatgtt aattgacgt atgcagcact tgcaagagcc 2340
aactttttca gaacagagta aaagcttctt cactgagttt cgagctcctt tgggtgaaag 2400
agaagagctt gatacaggtc tgagggttac taatgatcca aaagatctac actctgtccg 2460
tgatttgga gtcaacacat cacagagaaa ggcaaaacaa gttgaatctc ccacaagcac 2520
cttagagaca acagagaagg attacgaaga atcttcaccc acacaccgtt atcttttcag 2580
ttcagaatgt gcatccgttg atactctggg gaacgtcttc gtaatgccag ttctcttttt 2640
attctttcct aatgttctgg agtcagacaa tacgcctctg cctaaaacag aaaaacaaca 2700
ttcttgccgg aatacatctc acattgactt agttccagta gatacttcgg aaaaacatcg 2760
gcaagataat atctcatgca agttaaagga aagattctcg ccagacggtg ccagcgagac 2820
actagagact catagccttg tgaaaaggaa ctccaccaga gtaggtgaag atgatgtagc 2880
gaattctgtt ggagaaattg tgttatcatc ggatgaagat gactgtgagg gattggagct 2940
tagtccacgg ctactaact tcatcaagag cggcattgtt ccagagtcac ctgtctatga 3000
ccaaggggaa gcgaacagag aagaagatct tgaatttctt cagctttctt cacccatgag 3060
gttcagtaac gaattggcag gagagtcttc tttccctgag agaaaggttc agcataagt 3120
caacgattat aacattgtgt ctacaaccac tgaattgaga actcctcaga aggaggtagg 3180
tttggccaac ggaacagaat gcttggtgtt ttctctatt cctgaggatt ggagaactcc 3240
cttggcgaat ctgacaaaca caaacagcag cgctcgcaaa gattggcggg tgagttctgg 3300
agaaaagtta gaaactcttc gacagcctcg caagttgaag agactacgta gacttgagga 3360
ttgctcgagt gctgtaaaagg agaattatcc tggtattaca gaggcagacc atatcagatc 3420
tcgttctcgc ggtaaaaaagc acattagagg taagaagaag atgatcatgg atgatgatgt 3480
ccaagtcttc attgacgagg aagctgaggt ctcttcggga gcagagatgt cggctgatga 3540
gaacgaagat gtgactggcg attcatttga agatagtttc atagatgacg gaacaatgcc 3600
tacagcaaat actcaagccg agtctggtaa agttgacatg atggctgttt acaggcgttc 3660
tcttctcagc cagtcaccat taccggcgag atttctgtat ttagccgcat caagtctgag 3720
tccttattct gctggaccct tgacgagaat aatgagagc agaagcgact cagataaatc 3780
attgtcttct cttcgaacac caaaaacaac aaactctgag tcaaaccaag atgcaatgat 3840
gataggaaac ctttcggtag taaaaatctc gtcagatagc cggaaaagga aatttagctt 3900

```

atgcaactcg gcgaatgccc ccgtgattaa cttagaaagc aagtttgcag ctcatgcaca 3960
 agccacggag aaggaaagcc atgaaggcgt gagaagcaat gcaggtgcgt tagagtacaa 4020
 tgatgatgat gatgatgcat tctttgcgac actagacttt gatgcaatgg aagcacaagc 4080
 cacattgtta ttgtcgaaac agagatccga agcaaaagag aaagaagacg caacggttat 4140
 acctaatcca ggcatgcaga gaagtgatgg tatggagaaa gatgcacat cttttgatct 4200
 tgggtctgtg tgattcttct ttcatacgaa gatactaagt tatgtatata gattgacaaa 4260
 ggagacagta gagcataggc atttggatgt atgttttgtg tattaagttt aggtatatcc 4320
 tattgaagta cagtgcctaa ggcagtgcac atgggttaaat caagggttaat gcctcaattc 4380
 gttgaaccct ttaagtaatg acacaaatat gactacatcg g 4421

<210> 98

<211> 1771

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a hairpin RNAi (hpRNA) construct targeting a FANCM gene of *A. thaliana*

<400> 98

gggtcaggaa catgtcttac taactacttg gtttgcttgg tgatcgacga ggcacatcga 60
 gcttttagga attattctta ttgtgttgta gttcgtgagt tgatggcggg accgatacag 120
 ctgagaatac tggctcttac tgcaactcct ggatcaaaga cacaggccat ccagggtatc 180
 attgataatt tgcagatatc cacacttgaa tatcgaaatg agagtgacca tgatgtttgc 240
 ccttatgtcc ccgacagaaa attagaagtc atcgaggttc ccttgggtca agatgcagat 300
 gatgtatcga aacgcctggt tcatgttata cgtccatatg cagtcaggct taaaaacttt 360
 ggggttaatc taaatagaga tatacaaact ttaagtccac acgaagtact tatggcaagg 420
 gataagtttc gtcaagcacc tctaccaggc cttcccatg taaatcacgg agatgtagaa 480
 tcttgctttg cagctcttat caggtaagga aataattatt ttcttttttc ctttttagtat 540
 aaaatagtta agtgatgtta attagtatga ttataataat atagttgtta taattgtgaa 600
 aaaataattt ataaatatat tgtttacata aacaacatag taatgtaaaa aaatatgaca 660
 agtgatgtgt aagacgaaga agataaaagt tgagagtaag tatattattt ttaatgaatt 720
 tgatcgaaca tgtaagatga tatactagca ttaatatttg ttttaatcat aatagtaatt 780
 ctagctgggt tgatgaatta aatatcaatg ataaaatact atagtaaaaa taagaataaa 840
 taaattaaaa taatattttt ttatgattaa tagtttatta tataattaaa tatctatacc 900
 attactaaat atttttagttt aaaagttaat aaatatattg ttagaaattc caatctgctt 960
 gtaatttatc aataaataaa atattaaata acaagctaaa gtaacaaata atatcaaact 1020
 aatagaaaca gtaatctaata gtaacaaaac ataactaat gctaataata caaagcgcaa 1080
 gatctatcat tttatatagt attattttca atcaacattc ttattaattt ctaaataata 1140
 cttgtagttt tattaacttc taaatggatt gactattaat taaatgaatt agtcgaacat 1200
 gaataaacia ggtaacatga tagatcatgt cattgtgtta tcattgatct tacatttgga 1260
 ttgattacag ttgataagag ctgcaaagca agattctaca tctccgtgat ttacatgggg 1320

aaggcctggt agagggtgctt gacgaaactt atcccttgcc ataagtactt cgtgtggact 1380
 taaagtttgt atatctctat ttagattaac cccaaagttt ttaagcctga ctgcatatgg 1440
 acgtataaca tgaaacaggc gtttcgatac atcatctgca tcttgacca agggaacctc 1500
 gatgacttct aattttctgt cggggacata agggcaaaca tcatggtcac tctcatttcg 1560
 atattcaagt gtggatatct gcaaattatc aatgataccc tggatggcct gtgtctttga 1620
 tccaggagtt gcagtaagag ccagtattct cagctgtatc ggtaccgcca tcaactcacg 1680
 aactacaaca caataagaat aattccctaa agctcgatgt gcctcgtcga tcaccaagca 1740
 aaccaagtag ttagtaagac atgttctga c 1771

<210> 99

<211> 1771

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a hairpin RNAi (hpRNA) construct with G:U basepairs, targeting a FANCM gene of *A. thaliana*

<400> 99

gggttaggaa tatgttttat taattatttg gtttgtttgg tgattgatga ggtatattga 60
 gttttaggga attattttta ttgtgttgta gtttgtgagt tgatggtggt attgatatag 120
 ttgagaatat tggtttttat tgtaattttt ggattaaaga tataggttat ttagggattt 180
 attgataatt ttagatatt tatatttgaa tattgaaatg agagtgatta tgatgtttgt 240
 ttttatgttt ttgatagaaa attagaagtt attgaggttt ttttgggtta agatgtagat 300
 gatgtattga aatgtttgtt ttatgttata tgtttatatg tagttaggtt taaaaatttt 360
 ggggttaatt taaatagaga tatataaatt ttaagtttat atgaagtatt tatggtaagg 420
 gataagtttt gtttaagtatt tttattaggt tttttttatg taaattatgg agatgtagaa 480
 ttttgttttg tagtttttat taggtaagga aataattatt ttcttttttc ctttttagtat 540
 aaaatagtta agtgatgtta attagtatga ttataataat atagttgtta taattgtgaa 600
 aaaataattt ataaatatat tgtttacata aacaacatag taatgtaaaa aaatatgaca 660
 agtgatgtgt aagacgaaga agataaaagt tgagagtaag tatattattt ttaatgaatt 720
 tgatcgaaca tgtaagatga tatactagca ttaatatttg ttttaatcat aatagtaatt 780
 ctactgggtt tgatgaatta aatatcaatg ataaaatact atagtaaaaa taagaataaa 840
 taaattaaaa taatattttt ttatgattaa tagtttatta tataattaaa tatctatacc 900
 attactaaat atttttagttt aaaagttaat aaatatattg ttagaaattc caatctgctt 960
 gtaatttatc aataaacaaa atattaaata acaagctaaa gtaacaaata atatcaaact 1020
 aatagaaaca gtaatctaata gtaacaaaac ataactaat gctaataata caaagcgcaa 1080
 gatctatcat tttatatagt attattttca atcaacattc ttattaattt ctaaataata 1140
 cttgtagttt tattaacttc taaatggatt gactattaat taaatgaatt agtcgaacat 1200
 gaataaacaa ggtaacatga tagatcatgt cattgtgtta tcattgatct tacatttgga 1260
 ttgattacag ttgataagag ctgcaaagca agattctaca tctccgtgat ttacatgggg 1320
 aaggcctggt agagggtgctt gacgaaactt atcccttgcc ataagtactt cgtgtggact 1380

taaagtttgt atatctctat ttagattaac cccaaagttt ttaagcctga ctgcatatgg 1440
acgtataaca tgaaacaggc gtttcgatac atcatctgca tcttgacca agggaacctc 1500
gatgacttct aattttctgt cggggacata agggcaaaca tcatggtcac tctcatttcg 1560
atattcaagt gtggatatct gcaaattatc aatgataccc tggatggcct gtgtctttga 1620
tccaggagtt gcagtaagag ccagtattct cagctgtatc ggtaccgcca tcaactcacg 1680
aactacaaca caataagaat aattccctaa agctcgatgt gcctcgtcga tcaccaagca 1740
aaccaagtag ttagtaagac atgttcctga c 1771

<210> 100

<211> 1259

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct, targeting a FANCM gene of *A. thaliana*

<400> 100

gggacataag ggcaaacatc atggtcactc tcatttcgat attcaagtgt ggatatctgc 60
aaattatcaa tgataccctg gatggcctgt gtctttgatc caggagttgc agtaagagcc 120
agtattctca gctgtatcgg taccgccatc aactcacgaa ctacaacaca ataagaataa 180
ttccctaaag ctcgatgtgc ctcgatgatc accaagcaaa ccaagtagtt agtaagacat 240
gttcctgact gtatatcctt ctctaacact tgtggagtga caaagaaaac ccgtttgctt 300
ttccacaaaa aagctctttt cgaaggacat gtctgaccg tcaagtcaat cgtccattct 360
tgtgggtattc caacaatatg tcaggaacat gtcttactaa ctacttggtt tgcttggtga 420
tcgacgaggc acatcgagct ttagggaatt attcttattg tgttgtagtt cgtgagttga 480
tggcgggtacc gatacagctg agaatactgg ctcttactgc aactcctgga tcaaagacac 540
aggccatcca gggatatcatt gataatttgc agatatccac acttgaatat cgaaatgaga 600
gtgaccatga tgtttgccct tatgtccccg acagaaaatt agaagtcacg gaggttcctt 660
tggttcaaga tgcagatgat gtatcgaaac gcctgtttca tgttatacgt ccatatgcag 720
tcaggcttaa aaactttggg gttaatctaa atagagatat acaaacttta agtccacacg 780
aagtacttat ggcaagggat aagtttcgtc aagcacctct accaggcctt ccccatgtaa 840
atcacggaga tgtagaatct tgctttgcag ctcttactat tcgtcatcct aatatcttca 900
ttcttactca tcaacctagc aaatggcctt tctttcaatt tctcttctag catctcatac 960
gctggtctta ttccatgact agaaaggagc ttacgaatat gataaagagt gataagagct 1020
gcaaagcaag attctacatc tccgtgattt acatggggaa ggcttggtag aggtgcttga 1080
cgaaacttat cccttgccat aagtaactcg tgtggactta aagtttgtat atctctatct 1140
agattaaccc caaagttttt aagcctgact gcatatggac gtataacatg aaacaggcgt 1200
ttcgatacat catctgcac ttgacccaag ggaacctcga tgacttctaa ttttctgtc 1259

<210> 101

<211> 4228

<212> DNA

<213> Brassica napus

<400> 101

```
tccaaaattg gttttgcccc ccaatgtggc ttcggcgagg gtttcttcca caaaacccca 60
ctcaacctaa aatctgattc ggcgagaaac gctgtctact tatctcacgc gaaaagaaag 120
gcgtagatcc accctaaact aaaacagagc atcaagtga atgggacccg agtttccgat 180
cgaactcggt gaagaagaag atggattcga ttgggaagca gcagtcagag aaatcgactt 240
ggcttgcctc aaatccttaa acccttcttc ttcttcttcg acccatttca ccaacggcaa 300
tggcactaaa cctgctaaaa gacaatctac tcttgatcga ttcacgcaa gagccgacca 360
caagcctcct cctccgtatc ctctgttgt ttccgacccg agtttccgagt gtggtactaa 420
cgacaacact cccagcgtcg ggattgatcc tgagacagct aaaacttgga tttatccaat 480
gaacgttcct ctaagagatt atcagtttgc tataacgaag actgctttgt tttcaaacac 540
attagttgct ttaccaacag gccttggtaa aacgctcata gctgcagttg taatgtataa 600
ttacttcaga tggtttccac aaggtaaaat tgtctttgcc gcaccttcta ggctcttgt 660
gatgcagcag attgaggcct gccataatat cgtggggata ccacaagaat ggacgattga 720
cttgacgggt cagacttgcc cttccaaaag agcttccttg tggaaaacca aaagggtttt 780
cttcgtcact ccacaagttc ttgagaagga tatacagtca ggaacgtgtg ttaccaactg 840
cttggtttgc ttggtgatcg acgaggcaca tcgagcttta gggaattatt cttatttgt 900
tgtagttcgt gagttgatgg cagtaccagt gcagttgaga atattggctc ttactgcaac 960
tcctggatca aagacacagg ccatacaggg tatccttgat aatttgcaga tatcaacact 1020
tgaatatcga aacgagagtg accatgatgt ctgcccttat gtccacgaca gaaaagtaga 1080
actaatcgag gttcccttgg gtaaagatgc agatgaggt tctaaacgcc tattagatgt 1140
tatacgtcca tatgctgtca ggcttaaaaa tttcggggtc attctaagca gggattatca 1200
aactttgagt ccacacgaat tacttatggc aagggataag tttcgtgaag cacctgtacc 1260
aggcattccc catataagtc acggagatgt agaactttgc tttgcagctc ttatcacgct 1320
ttatcacatt cgcaagcttc tttctagtca tggaataagg ccagcgtatg agatgcttga 1380
agaaaaactt caggaagggc catttgctag gttgatgagt aagaatgaag atattagat 1440
gacgaagctt ttgatgcagc aaaggttgtc gaacggagca ccaagcccga aattgtccaa 1500
gatgttggag attctagttg atcactacaa aataaaagat ccgaggacat cacgggtcat 1560
tattttctcg aatttcagag gaagcgtaag agacataatg gacgcattaa gtaatattga 1620
agatgttgtc aaagcaactg agttttattg tcaaagttca ggtaagacac tgaagggaca 1680
gtcgcaaaaa gttcagcaag ctgttctgga gaaattttaga tctggtgggt ttaatgttat 1740
tgttgcaaca tctatcggcg aagaaggctt ggatatcatg gaagtcgact tagttatatg 1800
ttttgatgct aatgtatccc ctctgaggat gatccaacgc atgggaagaa ctggaaggaa 1860
aaataatggc cgagttgtag ttcttgcttg tgaaggatct gaaaagaata gctatatgcg 1920
aaagaaagca aatggccaag ccattaaaaa acacatgcgg aatggaggaa tgaatagttt 1980
taattttcat ctagtccaa ggatgattcc catgtttat aagccagaag ttcagcatgt 2040
taagttttcg atcgagcaat tcattccacg tggaaagaag ctacaagatg agcctgccac 2100
tgagactcca gctttcaaga aaaagcttac accggaagag atggatatgc tcgccaagta 2160
tttcaaacc aacgaggaaa agtggagagt ttccttgatt gctttccctc acttccaaac 2220
```

attgccatcc aaagtgcaca aagtaatgca ttcacgccaa acaagcatat taattgatgc 2280
tatgcagcat ctgcaagaga caactttgac agagcaaagt aaaagtttct tcattaagta 2340
tggagctcct ttggctgaaa gagatgagct tgacgcaggt ctgagggttg gtgatgatcc 2400
gaaaggtaaa tttagtctca atgatttgga tggcaacaca tcacagagaa aggcaaaaca 2460
aattttagaa tctcccacaa gcacattaga gactacagag aaggatttcg aagcatcttc 2520
acccacacac tgttatcttt tcagttcaga atgtgcgtcc gttgatactc tggggaaggt 2580
ctttgtattg ccggttcctc tctcattctc ttctaattga ccagggtcag actgcgtggg 2640
aagagaaaaa gaactttctt ccccgaataa gtcccacact gacgttggtc cgatagatag 2700
ttcctcaaaa catcggcaag ataatatctt atgcaagtta aagcaaggat tcttgccaga 2760
ttgtgccaac gagactttgg agtcccaaag cttttgaaa aggcaactcca ccgatgtagg 2820
taaaggagat atagagaatt gtgctggaga aattatgata tcatcggatg aagaagacga 2880
ctgtgaggat ttggagctta gtccaaggct cactaacttc atcaagagtg gcgttggttc 2940
agattcacct gtctatgacc aaggagttgc atacgaagca aacagagaag aagacctga 3000
tcttccaccc acgagtttaa ctaatgaatt ggcagaagag ccatcgacac ctgagaaaaa 3060
ggttcacatt gcttctacgg ccaatgaatt cagaactcct cagaaggaag aagatttagc 3120
caacgaaaca gaaagcttcg ctgtttctcc aatgcctgag gagtggagaa ctcccttggc 3180
gaatatcacc aacgcaagca gcagcgctag caaagattgg cgcgtgagtt cgggagaaaa 3240
gtcagaaact cttcgacagc ctcgcaagtt gaagagactt cgtagacttg gagattgctc 3300
gagtgtgtg aaggagaata atcctggtat tgcaaagaca gaccatatca gatctcgttc 3360
tcgcagtgt aagaacataa gaggcaagaa gaagatacgc gcgataata atgctagaat 3420
cttcattgaa gcggaagctg aggtgtcttc ggaatcagaa atgtcggttg atgagaacgt 3480
agatttgacc agcgattcat ttgaagatag cttcatagat gacggtacaa tgcctacagc 3540
aaatactcaa gccgagtgtg cttaaagttga catgatggcc gtttacagac gttctctact 3600
cagccaatca ccattaccgg caagatttcg tgatgtagct gcatcaagtc cgagtcctta 3660
ttcttctggt ctcttgaaga caataaatga gagcagaagc gactcagata aatcattgtc 3720
ttctcttaga accccacaaa caacgaacaa cgagtcaaac aaggatgcag tggccacagg 3780
agactttttg gtagcacaaa tctcaacaga cagccggaaa aggaaattca gcttatgcaa 3840
ctcagcgaat gtcccagtga ttaacttgga aaacaagttt gaagctcatg cacaagccac 3900
ggagaaggaa agccatgaag gtccgagaag caatgcaggt gcatcacagt acaaggatga 3960
ggatgaagat gatgatgcat tctacgcgac actggacttt gatgccatgg aagcgcatgc 4020
gacattgcta ttgtcgaaac aaaggtcaga aacgaaaaca aaagaagatg catcggtgaa 4080
acctcatctt ggcaatcaga ggaatgatgg tttgccgaag gatgggccat cttttgatct 4140
tggtttgtgg tgattattct cctattaagt taaagtgtat aaaggttgac atttggatgt 4200
atgttttgtg tatttagttt gtgtcata 4228

<210> 102

<211> 1769

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a hairpin RNAi (hpRNA) construct targeting a FANCM gene of *B. napus*

<400> 102

```

gggagaaatt atgatatcat cggatgaaga agacgactgt gtggatttgg agcttagtcc 60
aaggctcact aacttcatca agagtggcgt tgttcagat tcacctgtct atgaccaagt 120
tgcatacgaa gcaaacagtg aagaagacct tgatcttcca cacacgagtt taactaatga 180
attggcagaa gagccatcga cacctgagaa aaaggttcac attgcttcta cggccaatga 240
attcagaacc ccaacgaagg aagaagattt agccaacgaa acagaaagct tcgctgtttc 300
tccaatgcct gaggagtgga gaactccctt ggcgaaatc accaacgcaa gcagcagcgc 360
tagcaaaagt tggcgcgtga gttcgggaga aaagtcagaa actcttcgac agcctcgcaa 420
gttgaagaga cttcgtagac ttggagattg ctcgagtgt gtgaaggaga ataactctgg 480
tattgcaaag acagaccata tcgtaaggaa ataattatct tcttttttcc ttttagtata 540
aaatagttaa gtgatgttaa ttagtatgat tataataata tagttgttat aattgtgaaa 600
aaataattta taaatatatt gtttacataa acaacatagt aatgtaaaaa aatatgacaa 660
gtgatgtgta agacgaagaa gataaaaagt gagagtaagt atattatctt taatgaattt 720
gatcgaacat gtaagatgat atactagcat taatatttgt tttaatcata atagtaattc 780
tagctggttt gatgaattaa atatcaatga taaaatacta tagtaaaaat aagaataaat 840
aaattaaaaa aatatctttt tatgattaat agtttattat ataattaaat atctatacca 900
ttactaaata ttttagttta aaagttaata aatatcttgt tagaaattcc aatctgcttg 960
taatttatca ataaacaaaa tattaaataa caagctaaag taacaaataa tatcaaacta 1020
atagaaacag taatctaagt taacaaaaca taatctaagt ctaatataac aaagcgcaag 1080
atctatcatt ttatatagta ttattttcaa tcaacattct tattaatttc taaataatac 1140
ttgtagtttt attaacttct aaatggattg actattaatt aaatgaatta gtcgaacatg 1200
aataaacaag gtaacatgat agatcatgtc attgtgttat cattgatctt acatttggat 1260
tgattacagg atatggtctg tctttgcaat accaggatta ttctccttca cagcactcga 1320
gcaatctcca agtctacgaa gtctcttcaa cttgcgaggc tgtcgaagag tttctgactt 1380
ttctcccga ctcacgcgcc aatctttgct agcgtgtgtg cttgcgttgg tgatattcgc 1440
caaggagatt ctccactcct caggcattgg agaaacagcg aagctttctg tttcgttggc 1500
taaactcttct tccttcgttg gggttctgaa ttcattggcc gtagaagcaa tgtgaacctt 1560
tttctcaggt gtcgatggct cttctgcaa ttcattagtt aaactcgtgt gtggaagatc 1620
aaggtcttct tctctgtttg cttcgtatgc aacttggtea tagacaggtg aatctggaac 1680
aacgccactc ttgatgaagt tagtgagcct tggactaage tccaaatcct cacagtcgtc 1740
ttcttcatcc gatgatatca taatttctc 1769

```

<210> 103

<211> 1769

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a hairpin RNAi (hpRNA) construct with G:U

basepairs, targeting a FANCM gene of *B. napus*

<400> 103

```

gggagaaatt atgatattat tggatgaaga agatgattgt gtggatttgg agtttagttt 60
aaggtttatt aattttatta agagtgggtg tgttttagat ttatttgttt atgattaagt 120
tgtatatgaa gtaaatagtg aagaagattt tgatttttta tatatgagtt taattaatga 180
attggtagaa gagttattga ttttgagaa aaaggtttat attgttttta tggttaatga 240
atttagaatt ttaatgaagg aagaagattt agttaatgaa atagaaagtt ttgttgtttt 300
tttaatgttt gaggagtgga gaattttttt ggtgaatatt attaatgtaa gtagtagtgt 360
tagtaaagat tgggtgtgtg gtttgggaga aaagttagaa attttttgat agttttgtaa 420
gttgaagaga tttttagat ttggagattg tttgagtgtt gtgaaggaga ataattttgg 480
tattgtaaag atagattata ttgtaaggaa ataattattt tcttttttcc ttttagtata 540
aaatagttaa gtgatgttaa ttagtatgat tataataata tagttgttat aattgtgaaa 600
aaataattta taaatatatt gtttacataa acaacatagt aatgtaaaaa aatatgacaa 660
gtgatgtgta agacgaagaa gataaaaagt gagagtaagt atattatttt taatgaattt 720
gatcgaacat gtaagatgat atactagcat taatatttgt tttaatcata atagtaattc 780
tagctggttt gatgaattaa atatcaatga taaaatacta tagtaaaaaat aagaataaat 840
aaattaaaaat aatatttttt tatgattaat agtttattat ataattaaat atctatacca 900
ttactaaata ttttagttta aaagttaata aatattttgt tagaaattcc aatctgcttg 960
taatttatca ataaacaaaa tattaaataa caagctaaag taacaaataa tatcaaacta 1020
atagaaacag taatctaattg taacaaaaca taatctaattg ctaataatac aaagcgcaag 1080
atctatcatt ttatatagta ttattttcaa tcaacattct tattaatttc taaataatac 1140
ttgtagtttt attaacttct aaatggattg actattaatt aaatgaatta gtcgaacatg 1200
aataaacaag gtaacatgat agatcatgtc attgtgttat cattgatctt acatttgatg 1260
tgattacagg atatggtctg tctttgcaat accaggatta ttctccttca cagcactcga 1320
gcaatctcca agtctacgaa gtctcttcaa cttgcgaggc tgtcgaagag tttctgactt 1380
ttctcccgaa ctacgcgcc aatctttgct agcgctgctg cttgcgttgg tgatattcgc 1440
caaggagatt ctccactcct caggcattgg agaaacagcg aagctttctg tttcgttggc 1500
taaactcttct tccttcgttg gggttctgaa ttcattggcc gtagaagcaa tgtgaacctt 1560
tttctcaggt gtcgatggct cttctgcaa ttcattagtt aaactcgtgt gtggaagatc 1620
aaggtcttct tctctgttg cttcgtatgc aacttggtca tagacaggtg aatctggaac 1680
aacgccactc ttgatgaagt tagtgagcct tggactaage tccaaatcct cacagtcgtc 1740
ttcttcatcc gatgatatca taatttctc 1769

```

<210> 104

<211> 1259

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct, targeting a FANCM gene of *B. napus*

<400> 104

gggttctgaa ttcataggcc gtagaagcaa tgtgaacctt tttctcaggt gtcgatggct 60
 cttctgccaa ttcatagatt aaactcgtgt gtggaagatc aaggtcttct tctctgtttg 120
 cttcgtatgc aacttgggtca tagacagggtg aatctggaac aacgccactc ttgatgaagt 180
 tagtgagcct tggactaagc tccaaatcct cacagtcgtc ttcttcatcc gatgatata 240
 taattttctcc agcacaattc tctatatctc ctttacctac atcgggtggag tgccttttca 300
 aaaggctttg ggactccaaa gtctcgttgg cacaatctgg caagaatcct tgctttaact 360
 tgcataaaat attatcttgg agaaattatg atatcatcgg atgaagaaga cgactgtgtg 420
 gatttggagc ttagtccaag gctcactaac ttcataaga gtggcgttgt tccagattca 480
 cctgtctatg accaagttgc atacgaagca aacagtgaag aagacctga tcttccacac 540
 acgagtttaa ctaatgaatt ggacagaagc ccatcgacac ctgagaaaaa ggttcacatt 600
 gcttctacgg ccaatgaatt cagaacccca acgaaggaag aagatttagc caacgaaaca 660
 gaaagcttcg ctgtttctcc aatgcctgag gaggggagaa ctcccttggc gaatatcacc 720
 aacgaagca gcagcgctag caaagattgg cgcgtgagtt cgggagaaaa gtcagaaact 780
 cttcgacagc ctgcgaagtt gaagagactt cgtagacttg gagattgctc gaggctgtgtg 840
 aaggagaata atcctggtat tgcaaaagca gaccatatcc agcttccgct tcaatgaaga 900
 ttctagcatt attatccgcg cgtatcttct tcttgcttg aacacagagc aaaaggaaat 960
 acagaatcat ttacctctt atgttcttta cactgcgaga acgagatctg atatggtctg 1020
 tctttgcaat accaggatta ttctccttca cagcactcga gcaatctcca agtctacgaa 1080
 gtctcttcaa cttgcgaggc tgtcgaagag tttctgactt ttctcccgaa ctcacgcgcc 1140
 aatctttgct agcgtctgtg cttgcgttgg tgatatctgc caaggaggtt ctccactcct 1200
 caggcattgg agaaacagcg aagctttctg tttcgttggc taaatcttct tccttcgtt 1259

<210> 105

<211> 7074

<212> DNA

<213> *Nicotiana benthamiana*

<400> 105

atgcgagctg tttttgaaac aaagcgtgat cgtgaaatct tggtccttgc tagtaaagtt 60
 ctgggtcacc tagctagatc tggcgggtgca atgactgcag atgaagtggga acgtcagata 120
 aaagttgcac taggatggct tcgtggtgaa agaattgagt atcgtttctt tgctgctgtc 180
 ttaatatattga aggaaaatggc ggaaaatgct tcaactgttt tcaatgttca tgtgccggac 240
 tttgtggagg ttgtttgggt tgctctgaag gatccaacat tggctgttcg agagaaggct 300
 gttgaggcat tgcgtgcctg ctttcgcgtt attgaaaagc gcgagacacg atggcgtgtt 360
 cagtggatatt ataagatgtt tgaggtacc caagatggat tgaccagaac tgcgcctgtt 420
 catagtatac atggctccct tctcgcagtg ggagagctgc taaggaatac aggagagttc 480
 atgatgtcaa gatacaggga ggttgacagaa attgttataa gatacttggga gcaccagat 540
 cgcctagttc ggctcagcat aacttctcta cttcctcgaa ttgctcattt cctgcgtgat 600
 cgatttgtga ctaactactt aacgatatgc atgaatcata tacttcatgt ccttaaaata 660
 cctgcagaac gtgccagtgg gttcattgct cttggggaga tggctggtgc tctggatggt 720

```

gaactcatta actatttgcc gacaataacc tctcacttgc gtgatgcgat tgctccccgt 780
agaggcaggc cctcatttga ggctctggca tgtgttgga atattgctaa agcaatggga 840
cctgcaatgg agcctcatgt tcgtggtctc ttggatgcta tgttttctgc tgggctttcc 900
ctgacactag tggaagcctt ggagcaaata actgaaagca ttccatcttt gttgccgacc 960
attcaagatc ggcttcttga atgtatttca gcaattctct ccagatctaa tcatgcactc 1020
tcaagacaat caactgctat gagtcgagga catattgcaa cagttacccc ccaagtacca 1080
gaactgagtg gtcctgcact agttcaactt gctttgcaga ctcttgctcg ttttaatttc 1140
aagggccatg atcttcttga gtttgcaagg gagtctgttg ttgtgtattt agaagatgag 1200
gatggagcta cacgaaaaga tgctgcgcta tgttgctgca aactagtagc aaattctttc 1260
ttggcgatat cttctaccca gtttagtcct agtagaatca atcgtgccag tggaaagcga 1320
cgctgacttg ttgaagagat tgtgcaaaaa cttctcatcg ctgctgttgc cgacgctgat 1380
gttactgttc ggcaattcaat tttttcttct ctgtatgctg atggaggatt cgatgagttt 1440
ctagctcagg ctgatagttt gacagctata tttgccactt taaatgatga ggattttgaa 1500
gttcgtgact atgcaatttc actagctggg agactatctg aaaagaatcc agcatatgtt 1560
cttcagcac ttcgtcgcca tcttattcag ctgttaactt acctagagca aagtgcagat 1620
aataaatgta aagaagagag tgcaaagtta ttgggttgct tgattcgcaa ttgtgaacga 1680
ttagtcttc catacattgc tcccatacac aaggctcttg ttgcgaaact ctgtgaaggc 1740
acaggagtca atgcgaatag tggcattatt agtggagttc tagtgactgt tggagatctt 1800
gccagagtgg gtggctttgc catgcggcag tatatttcag aacttatgcc attaatcggt 1860
gaagctctac tggatggggc agctgccacc aaacgtgaag tggccgtttc aacacttggg 1920
caagttgtac agagtacagg atatgtcata actccataca atgagtatcc tcagttgctt 1980
ggtttactct tgaaactgct caatggtgaa ctggcttggg caaccagaag agaggttttg 2040
aaggttctcg gcatcatggg tgcattagat ccccatgtgc acaagcgcaa tcagcaaagc 2100
ttacccggat cccatggtga agttacccgg gtgactgggtg atcctgggtca acatatcaga 2160
tcaatggatg aattgcctat ggatctttgg ccctccttg caacatctga agattattat 2220
tccactgttg ctatcaactc actcatgcgg atactcaggg atccatctct gtcaagttac 2280
caccagaaag tggttggatc tcttatgttt atttcaagt ccatgggcct tggctgtgtc 2340
ccttatttgc ctaaggtttt gcctgatctc tttcacattg tacgaacatg tgaggatggg 2400
ctaaaagaat ttataacatg gaagcttggg accttgggtat ctattgtccg ccagcacatc 2460
cgtaagtatc tgccagagtt actctctctg atatcagaaa tatgggtcatc tttcagcttg 2520
cctgttgcta acagacctgt tcacattgct cctattttgc atctcgtgga gcaactttgc 2580
ttggctctca acgatgaatt tagaaagtac cttgctgata tacttcctg ctgtattcaa 2640
gttcttactg atgcagagag gtttagtgac tacacatacg ttattcctat tctccacaca 2700
cttgaagttt ttggtgggac attagatgag catatgcac tgcttcttcc tgcacttatt 2760
cggttgttta aattggatgc ttcagtagaa gtaagacgcg gtgcaatcaa aactctcaca 2820
agattgatac ctcgtgtgca ggtcactgga cacatatctt ctcttggtgca tcacttgaag 2880
cttgtcttgg acgggaacaa agaagagctc aggaaggatg ctgttgatgc actttgttgt 2940
ctagctcatg ctcttggaga ggacttcacc atttttatct attctattca caagcttatg 3000
gttaaacata ggctgcagca caaggaattt gaagaaatcc gaggacgact ggaaaaacgt 3060

```

gagccactga ttttggggag caccgcagct cagagattaa atcggcggtt cccagttgag 3120
 gtcacagctg atcctttgag tgatggagag aatgagcact acgaggttgg gacggacatg 3180
 cataagcagc ttaaaagcca tcaggttaat gatggtagat tgcgtaccgc tggtaggct 3240
 tctcaacgaa gcactaaaga ggattgggca gattggatga ggcatttcag cattgaactt 3300
 ctgaaagaat cacctagtcc agcattgcga acttgtgcaa gactcgtca actgcagcct 3360
 tttgttgggc gagagttggt tgctgcaggt tttgttagct gctggtcaca acttaatgag 3420
 gctagtcaaa ggcagctagt acgtagtcta gaaatggcat tttcgtctcc aaatatccct 3480
 cctgaaattc ttgtacact tctgaacttg gcggagttta tggaacacga tgagagaccc 3540
 cttcctattg atatccgtct gcttgggtgct cttgcggaga agtgtcgagc atttgcaaag 3600
 gccctacact acaaggaaat ggaatttgaa ggcgcacttt caaataggag ggatgcaaat 3660
 cctgttgctg tagttgaagc tctaattccat ataaataatc aattacatca acatgaggca 3720
 gctgttgga tattaacata tgctcagcag ctttggggg ttcaattgaa ggagtcattg 3780
 tacgagaaat tgcaacgctg ggatgatgct cttaaagcat aactgctaa ggcgtcaca 3840
 gcttcgagtc cacatcttgc tttggatgct actttagggc gtatgcgatg ctttgcgtct 3900
 ctagctcggg gggaggagct taacaatctt tgtaaggaat actggacacc agctgagcca 3960
 gcagctcgac tggaaatggc accaatggct gctagtgcgg cctggaacat gggtagtggt 4020
 gatcagatgg cagagtatgt ttctcggctt gatgatggtg atgaaaccaa actgcgagtc 4080
 ttgggaaata ccgctgccag tggcgatgga agtagtaatg gcaccttttt cagggtctgtt 4140
 cttctagttc ggcgagggaa gtacgatgaa gcacgtgaat atgttgaaag agcaaggaaa 4200
 tgtttggcga ccgagctcgc tgcaactggt cttgagagct atgaacgtgc ttacagcaac 4260
 atgggtccgtg ttcagcagct ttctgaatta gaagaggtga ttgaatactg tactcttctt 4320
 atgggaaacc ctgttgctga aggaagaaga gctcttggtt gcaatatgtg gaatgagcgc 4380
 ataaagggtg caaaaagaaa tggtgaggtt tggcaagtac ttttagctgt gagggcactt 4440
 gtattgcctc ctacagaaga cattgaaaca tggatcaaat ttgcatcact ttgccgaag 4500
 aatggcagaa ttagccaagc tagatctaca ttggttaaac ttttacagtt cgatccagaa 4560
 tcaactcctg caactgtgcg gtatcatggt cccctcagg tgatgctagc atacttaaag 4620
 taccaatggt cacttggcga ggatcataag cgaaaggaag ctttgctag gttgcaggac 4680
 cttgccatgg acctctcaag aacagcagct cttcaaccag tattgcagaa tggattagtt 4740
 gcttcttctg gtgtgccact tgttgctcgt gtatatctca gactcggcac ttggaagtgg 4800
 gcactttctc ctggtttgga tgatgattct attcaagaaa ttcttagtgc atttacaat 4860
 gctactcact gtgcaacgaa gtgggggaaag gcatggcata cctgggcact tttcaatacc 4920
 gcagtgatgt ctcatcac tctgagaggt tttgcgaata ttgcttcaca gtttgttggt 4980
 gctgccgtaa ctggttattt tcaactctata gcatgcggag cacatgctaa ggggtgttgat 5040
 gatagtttac aggatattct tcgtcttctt actttgtggt tcaaccatgg agctacttcg 5100
 gatgtccaaa tggcattgca gaaaggattc actcatgtta acatcaacac atggttggtt 5160
 gttttacctc agattattgc acggatacat tcaaataacc atgctgtcag agaactgata 5220
 caatccttgc tagtgcaat tggacagagt catccacagg ctcttatgta tccgttctt 5280
 gtggcatgta agtcaattag caatttgcgc agagctgcgg ctcaagaagt ggttgataaa 5340
 gttagacagc acagcggcgt actcgttgat caggcccaac ttgtctcaaa ggagcttatc 5400

agggttgcaa tactgtggca tgaaatgtgg catgaggcac tggaagaggc cagccgttta 5460
 tattttggcg aacacaacat tgagggcatg ctgaaggtgt tagagcctct gcatgaaatg 5520
 cttgaggaag gagcgatgag gaacaatacc actataaagg agaaagcatt catccaggca 5580
 taccgtcttg agttgttga ggcgtatgaa tgttgatga agtatcggag aactggtaaa 5640
 gatgctgaat taacgcaggc ttgggatctc tattatcatg tattcaggcg gatagataag 5700
 cagcttcaaa cactcacaac cctggatttg cagtctgttt ccccgagtt actggagtgt 5760
 cgaaatttgg agctagctgt tcctggaact tatatagcag atgcaccagt ggtgacaatt 5820
 gcatcatttg caccccaact tgttgtaatt acatccaaac aacggcctcg aaaattgaca 5880
 atccatggga gtgatggaga agactatgct tttttgctca aagggcacga agatctacgc 5940
 caagatgaac gtgtcatgca gttgtttggg ctggttaata ctttgctcga gaattcaaga 6000
 aagactgcag agaaagattt atcaattcaa cgatatgctg tcattccatt gtcccctaatt 6060
 agtggactga taggatgggt tccaaattgc gacaccttg accagcttat tcgagaatat 6120
 agggatgccc ggaagatcac cctaaatcaa gagcataaat tgatgctgag ttttgaccg 6180
 gattatgata atttgccact tattgctaag gtggaggtgt ttgaatatgc tttgcaaat 6240
 acagaaggga atgacttacc aagggttctt tgggttaaaga gtcgtacttc tgaagtctgg 6300
 ctggacagaa gaacaaatta tacaagaagt ttggctgtca tgagtatggt tggataccta 6360
 cttggtctgg gtgatcgaca tcctagtaac ctcatgcttc accgatacag tgggaagatt 6420
 ctgcatattg actttggaga ttgctttgaa gcttcaatga atcgggagaa gtttccagag 6480
 aaggttccct ttcgactcac tagaatgctt gtaaaagcaa tggaggtttag tggatatagag 6540
 ggaaatttcc ggtcaacatg tgagaatgta atgcaagttc tccgactgca taaagatagt 6600
 gttatggcta tgatggaggc ctttgttcac gatccactta taaattggcg tcttttcaac 6660
 ttcaatgaag ttccgcaaat gtccgcactt gccagtgcac atgtccctcc tgttgatgaac 6720
 agtgaggaat cttcttcaaa tagagagctt cttcagccac aaagggtgac aaggagaga 6780
 gaactgcttc aggcggtcaa tcaattaggt gatgccaatg aggttctaaa tgaacgtgct 6840
 gtggctgtta tggctcgaat gagtaataaa ctacaggac gtgattttgc tgctacttct 6900
 acatctgcga gctctctaca acatgcactg gaccacagta cgtaatttc tggagagacg 6960
 cgtgaagctg atcatggttt atcagtgaac ctacaagtcc aaaaacttat tcaacaagcg 7020
 tcgtctcatg aaaatctttg ccaaaattat gttgggtggt gtccattttg gtag 7074

<210> 106

<211> 1513

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct targeting a TOR gene of *N. benthamiana*

<400> 106

tgaatttagg tgacactata gaacaaagaa gagctcagga aggatgctgt tgatgcactt 60
 tgttgcttag ctcatgctct tggagaggac ttcaccattt ttattcattc tattcacaag 120
 cttatgggta aacataggct gcagcacaag gaatttgaag aaatccgagg acgactggaa 180

aaacgtgagc cactgatttt ggggagcacc gcagctcaga gattaaatcg gcggttccca 240
 gttgagggtca tcagtgatcc tttgagtgat ggagagaatg agcactacga ggttgggacg 300
 gacatgcata agcagcttaa aagccatcag gttaatgatg gtagattgcg taccgctggt 360
 gaggccttctc aacgaagcac taaagaggat tgggcagagt ggatgaggca tttcagcatt 420
 gaacttctga aagaatcacc tagtccagca ttgcgaactt ttaagctgct tatgcatgtc 480
 cgtcccaacc tcgtagtgct cattctctcc atcactcaaa ggatcactga tgacctcaac 540
 tgggaaccgc cgattttaatc tctgagctgc ggtgctcccc aaaatcagtg gctcacgttt 600
 ttccagtcgt cctcggattt cttcaaattc cttgtgctgc agcctatgtt taaccataag 660
 cttgtgaata gaatgaataa aaatggtgaa gtctctctca agagcatgag ctagacaaca 720
 aagtgcatac acagcatcct tctgagctc ttctttgttc cgtccaaga caagcttcaa 780
 gtgatgcaca agagaagata tgtgtccagt gacctgcaca cgaggtatca atcttgtgag 840
 agttttgatt gcaccgcgtc ttactttctac tgaagcatcc aatttaaaca accgaataag 900
 tgcaggaaga agcagatgca tatgtctatc taatgtccca ccaaaaactt caagtgtgtg 960
 gagaatagga ataacgtatg tgtagtcact aaacctctct gcacagtaa gaacttgaat 1020
 acagcaggga agtatatcag caaggtactt tctaaattca cacatccgta agtatctgcc 1080
 agagttactc tctctgatat cagaaatatg gtcacttttc agcttgccctg ttgctaacag 1140
 acctgttcac attgtctcta ttttgcatct cgtggagcaa ctttgcttgg ctctcaacga 1200
 tgaatttaga aagtaccttg ctgatatact tccctgctgt attcaagttc ttactgatgc 1260
 agagagggtt agtgactaca catacgttat tcctattctc cacacacttg aagtttttgg 1320
 tgggacatta gatgagcata tgcactgtct tcttctgca cttattcggt tgtttaaatt 1380
 ggatgcttca gtagaagtaa gacgcggtgc aatcaaaaact ctcacaagat tgatacctcg 1440
 tgtgcaggtc actggacaca tatcttctct tgtgcatcac ttgaagcttg tcttggacgg 1500
 cccgggactc gaa 1513

<210> 107

<211> 1941

<212> DNA

<213> *Hordeum vulgare*

<400> 107

atggccgcag ccacctccgc cgccgtcgca ttctcgggcg ccgccgccgc cgccgcggcc 60
 ttaccaagc ccgccctcca tctctctccg cgccaccagc ccgcctcgcg ccgcgcgctc 120
 cccgcccgcg tcgtcagggtg ctgcgccgcg tccccgccg ccaccacggc cggcctctcc 180
 cccacctctc tccggccgtg ggggccctcc gagccccgca agggcgccga catctctgtc 240
 gaggcgctcg agcgctgcgg catcgctcac gtcttcgctt accccggcgg cgcgctccatg 300
 gagatccacc aggcgtcac gcgtcgccc gtcatcacca accacctctt ccgccacgag 360
 cagggggagg cgttcgcagc gtccgggtac gcacgcgct cggccgcgt cggcgtctgc 420
 gtcgccacct ccggccccgg ggccaccaac ctgctctccg cgctcgccga cgctctctc 480
 gactccatcc ccatggtcgc catcacgggc caggtccac gccgcatgat cggcacggac 540
 gcgttccagg agacgccc atgtggaggtc acgcgtcca tcaccaagca caactacctg 600
 gtccttgacg tggaggacat cccccgcgtc atccaggaag ctttcttctt cgctctctct 660

ggccgccccg ggccctgtgct ggttgatata cccaaggaca tccagcagca gatggccgtg 720
 cctgttttggg acacgccgat gagtttgcca ggttacatcg cccgcctgcc caagccacca 780
 tctactgaat cgcttgagca ggtcctgcgc ctggttggcg aggcacggcg cccgattctg 840
 tatgtttggtg gcggctgcgc tgcattctggc gaggagtgc gccgctttgt tgagctcact 900
 ggaattccag ttacaactac tctgatgggc cttggcaact tccccagtga cgaccactg 960
 tcaactgcga tgcttgggat gcatggtacc gtgtatgcaa attatgcagt agataaggct 1020
 gacctgttgc ttgcatttgg tgtgcggttt gatgatcgcg tgactgggaa aattgaggct 1080
 tttgcaagca ggtccaagat tgtgcacatt gacattgac cagctgagat tggcaagaac 1140
 aagcagccac atgtctccat ttgtgcagat gttaagcttg ctttacaggg gttgaatggt 1200
 ctattaagtg gcagcaaaagc acaacagggt ctagattttg gtccatggca caaggagtgt 1260
 gatcagcaga agaggagatt tcctctagga tacaagactt ttggtgaggc aatcccaccg 1320
 cagtatgcta tccagggtact ggatgagctg acaaaagggg aggcgattat tgccacaggt 1380
 gttgggcagc atcagatgtg ggcggtcag tattacactt acaagcgcc acgtcagtgg 1440
 ctgtcttcgt ctggtttggg ggcaatggga tttgggttgc cagctgcagc tggcgcttct 1500
 gtggccaacc caggtgtcac agttgttgac attgatgggg atggtagttt cctcatgaac 1560
 attcaggagt tggcgttgat ccgtattgag aacctcccag tgaagtgat gatattgaac 1620
 aaccagcacc tgggaatggt ggtgcagtgg gaggataggt tttacaaggc caaccggcg 1680
 cacacatacc ttggcaaccc agaaaatgag agtgagatat atccagattt tgtgacgatt 1740
 gctaaaggat tcaacgttcc ggcaattcgt gtgacaaaga agagtgaagt cagtgcagct 1800
 atcaagaaga tgcttgagac cccagggccg tacctgctgg atatcattgt cccgcatcag 1860
 gagcacgtgc tgcctatgat cccaagcggg ggtgctttca aggacatgat catggagggt 1920
 gatggcagga cctcgtatta a 1941

<210> 108

<211> 1505

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA targeting the ALS gene of barley (H. vulgare)

<400> 108

agatttaggt gacactatag aatggtggtg cagtgggagg ataggtttta caaggccaac 60
 cgggcgcaca cataccttgg caaccagaa aatgagagt agatataatcc agattttgtg 120
 acgattgcta aaggattcaa cgttccggca gtctgtgtga caaagaagag tgaagtcagt 180
 gcagctatca agaagatgct tgagaccca gggccgtacc tgctggatat cattgtcccc 240
 catcaggagc acgtgctgcc tatgatccca aacggtggtg ctttcaagga catgatcatg 300
 gagggatgat gcaggacct gtattaatct gaatttcgac ctacaagacc tacaagtgtg 360
 acatgcgcaa tcagcatgat gcctgcgtgt tgtatcaact actagggtt cgactgtgaa 420
 ccatgcgttt ttctagttaa cttgtttcat tcatataaga ggtcctgcca tcaccctcca 480
 tgatcatgtc cttgaaagca ccaccgtttg ggatcatagg cagcacgtgc tcctgatgcg 540

ggacaatgat atccagcagg tacggccctg gggctctcaag catcttcttg atagctgcac 600
 tgacttcact cttctttgtc acacgaactg ccggaacgtt gaatccttta gcaatcgta 660
 caaaatctgg atatatctca ctctcatitt ctgggttgcc aaggtatgtg tgcgcccgtg 720
 tggccttgta aaacctatcc tcccactgca ccaccattcc caggtgctgg ttgttcaata 780
 tcatcacctt cactgggagg ttctcaatac ggatcaacgc caactcctga atgttcatga 840
 ggaaactacc atcccatca atgtcaacaa ctgtgacacc tgggttggcc acagaagcgc 900
 cagctgcagc tggcaaccca aatcccattg cccccaacc agacgaagac agccactgac 960
 gtggccgctt gtaagtgtaa tactgagccg cccacatctg atgctgcca acacctgtg 1020
 caataatcgc ctcccccttt gtcagctcat ccagtacctg aatggtctat taagtggcag 1080
 caaagcaca caggtctag attttggctc atggcacaag gagttggatc agcagaagag 1140
 ggagtttctt ctaggataca agacttttgg tgaggcaatc ccaccgcagt atgctatcca 1200
 ggtactggat gagctgacaa aaggggaggc gattattgcc acaggtgttg ggcagcatca 1260
 gatgtggcg gctcagtatt acaattacaa gcggccacgt cagtggctgt ctctgtctgg 1320
 tttgggggca atgggatttg ggttgccagc tgcagctggc gcttctgttg ccaaccagg 1380
 tgtcacagtt gttgacattg atgggatgg tagtttctc atgaacattc aggagttggc 1440
 gttgatccgt attgagaacc tcccagtga ggtgatgata ttgaacaacc agcacctggc 1500
 ccggg 1505

<210> 109

<211> 1824

<212> DNA

<213> *Hordeum vulgare*

<400> 109

atgcagactc tgtcggcgca gcccctcgcc tctctctctt cgatacagcg ccaccatggg 60
 cgccgacgcg gccccggctc cgtccgggtc gctccccgcg cggccgcccgc ggctgccgcc 120
 acgtccacca gcacggcccc ctgcggcgcg tacgtctcgt cgccgtccac gaggaaggtg 180
 cccgggtacg agcagtcgtc gccgcctgcc attgcctcgc cgcagaagca ggggagcagc 240
 ggcggcgagg gcgagcagag cctcaacttc ttccagcgcg cggcggcccgc ggcgctcgac 300
 gcgttcgagg aggggttcat caacaatgtc ctggagcggc cccacgcgct gccgcgcacg 360
 gccgacccgg ccgtgcagat cgccggcaac ttgcccccg tcggcgagca gcccccgctg 420
 cgcgccctca cggctctccg ccgcateccg ccttcatca acggcgtcta cgcccgaac 480
 ggcgccaacc cctgcttcga gccacggcc ggccaccacc tcttcgacgg cgacggcatg 540
 gtccacgcca tccgcatccg aaacggcgcc gccgagtcct acgcctgccg cttaccagag 600
 accgcccgc tctcccagga gcgcgccgcg gggaggcccc tcttcctaa gaccatcggc 660
 gagctccacg gccactctgg catcgcgagg ctggccctct tctacgcgcg cggcgccctgc 720
 ggccctgctc acccgctcca cggcactggt gttgccaacg ccggcctcgt ctacttcaac 780
 ggccgcctcc tcgccatgtc cgaggacgac ctcccgatc aggtccgcgt caccgcccgt 840
 ggcgacctcg agaccgtcgg ccgtacgac ttcgacggcc agctcgactg cgccatgatc 900
 gcgcacccca agctcgaccc tgtctccggc gagctcttcg cgctcagcta cgatgtcatc 960
 aagaagccgt acctcaagta cttctacttc cagcccgacg gcaccaagtc cgccgacgtc 1020

gagatcgagc tcgaccagcc caccatgatc cagcacttcg ccatcaccga gaacttcgtc 1080
 gtcgtgcccc accaccagat ggtgttcaag ctcgccgaga tgttccgcgg cggctcgccg 1140
 gtgatgctcg acaaggagaa gacctccgcg ttccggcgtc tcccaaagta cgccaaggac 1200
 tcgtcggaga tgatgtgggt ggacgtgccg gactgcttct gtttccacct ctggaactcg 1260
 tgggaggagc cggagacgga cgaggtgggt gtgatcggct cctgcacgac ccccgagac 1320
 tccatcttca acgacacgga cgaccacctc gagagcgtgc tcaccgagat ccggctcaac 1380
 acgcgcaccg gcgagtccac gcggcgggcc atcctgccgc tggagagcca ggtgaacctc 1440
 gaggtcggca tggatgaaccg caacatgctg ggccggaaga cgaggtacgc ctacctggcc 1500
 gtggccgagc cgtggcccaa ggtgtccggg ttccgcaagg tggacctggt gaccggcgag 1560
 ctgaccaagt tcgagtacgg cgagggccgg ttccggcgcg agccgtgctt cgtgcccattg 1620
 gacggcgagc acgcgcgccc cggcgccgag gacgacggct acgtgctctc ctctgtgcgc 1680
 gacgaggacg ccggcacatc cgagctcctg gtcgtcaacg ccgccgacat gcggctcgag 1740
 gccaccgtgc agctgccgtc ccgggtcccc tatggcttcc acggcacatt catcggcgac 1800
 gccgacctcg acgcccagca ctaa 1824

<210> 110

<211> 1779

<212> DNA

<213> *Hordeum vulgare*

<400> 110

atgcagacac tcacagcgtc cagctcggtc tcctccatac agcggcaccg gccgcacccc 60
 gcgggcccgc ggtccagctc ggtcaccttc tccgcccgcg ccgtcagctc cgcgccgcgc 120
 gcgcccggcac cgtcccgggt cgtgcgcggc gccgacgcgg cggccgcca gcccctcatt 180
 gccgtcccca agccgcccgc cgtggagagg caggagaaga agtcaactt cttccagcgc 240
 gccgcgggtca cggcgctcga cgcgttcgag gaaggatttg tggccaacgt gctcgagcgc 300
 ccgcacggcc tctccaggac ggtcgacccc gcggtgcaga tcgcccggca cttcgcgctt 360
 gtcggggaga cactcctgt gcaggcgtg cccgtgaccg accgcatccc cccgttcattc 420
 aacggcgtgt acgcccgcaa cggcgccaac ccgcacttcg acccgtcgc cgggcaccac 480
 ctgttcgacg gcgacggcat ggtgcacgct ctgcgcattc gcaacggcgt cgccgagacc 540
 tacgcctccc gcttcaccga gacggagcgc ctgcagcagg agcgcgcgct ggggcgccc 600
 atgttcccca aggccattgg tgagctccat ggccactctg ggatcgcgcg cttgtctctg 660
 ttctacgcgc gcgcggcctg cggcctcatt gacccctcgc gcggcaccgg cgtggccaac 720
 gccgcctgg tctacttcaa cggccacctc ctgcctatgt ccgaggacga catcccgtag 780
 cacgtccgcg tcaccgacga cggcgacctc cagaccgtcg gccgctacga cttcgacggg 840
 cagctcgagt gcccattgat cgcgcacccc aaactcgacc ccgccaccgg ggagctccac 900
 gcgctcagct acgacgtcat caagaagcct tacctgaagt acttctactt cgcgcccgac 960
 ggcaccaagt cggccgacgt cgagatcccc ctggaccagc ccaccatgat ccacgacttc 1020
 gccatcaccg agaattacgt ggtcgtgccc gaccaccagg tgggtgttcaa gctgcaggag 1080
 atgctgcgcg gcggctcgcc cgtggtgctc gacaaggaga agacgtcccg cttcggcgtg 1140
 ctgcccgaagt gcgccgccga cgcgtcggag atggtgtggg tggacgtgcc ggactgcttc 1200

tgcttccacc tctggaacgc gtgggaggag gaggagaccg acgaggtggt ggtgatcggc 1260
 tcctgcatga cccccgccga ctccatcttc aacgagtcgg acgagtgcct cgagagcgtg 1320
 ctcacggaga tccgcctcaa caccgcacc ggcgagtcca cgcggcgccc catcctggcg 1380
 ctgtcagagc aggtgaacct ggaggtcggc atggtgaact ccaacctgct gggccgcaag 1440
 acgcggtacg cctacctggc cgtggccgag ccgtggccca aggtgtccgg cttcgccaag 1500
 gtcgacctgg ccacgggcga gtcacaaaa ttcgagtacg gcgagggccg gttcggcggc 1560
 gagccctgct tcgtgcccac ggacccggcc acgtcccgcg gcgaggacga cgggtacatt 1620
 ctcaccttcg tgcacgacga ggccgcccgc acgtcggagc tgctggtggt caatgccgcc 1680
 gacatgcggc tggaggcgac catccagctg ccgtcccgcg tgccatacgg gttccacggc 1740
 accttcatca cggcaagga gtcgaatcc caggcctga 1779

<210> 111

<211> 1500

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct targeting the NCED1 genes of barley *Hordeum vulgare* and wheat *Triticum aestivum*

<400> 111

taatacgact cactataggg tcgacgaggc cgcaggcgcc gcgcgcgtag aagagggcca 60
 gcctcgcgat gccagagtgg ccgtggagct cgccgatggt cttagggaag acgggcctcc 120
 ccgcggcgcg ctctgggag aggcggggcg tctcggtaga gcggcaggcg taggactcgg 180
 cggcgccggt tcggatgcgg atggcgtgga ccatgccgtc gccgtcgaag aggtggtggc 240
 cggccgtggg ctggaagcag gggttggcgc cgttgcgggc gtagacgccg ttgatgaagg 300
 gcgggatgcg gccggagacc gtgaggcgcc gcacgggggg ctgctcgccg acgggggcga 360
 agttgccggc gatctgcacg gccgggtcgg ccgtgcgcgg cagcgcgtgg ggccgctcca 420
 ggacattggt gatgaacccc tcctcgaacg cgtcgagctc cggccgcata ccgcccttca 480
 tcaacggcgt ctacggccgc aacggcgcca acccctgctt cgagcccacg gccggccacc 540
 acctcttcga cggcgacggc atggtccacg ccatccgcat ccgaaacggc gccgccgagt 600
 cctacgcctg ccgcttcacc gagaccgccc gcctctccca ggagcgcgcc gcggggaggc 660
 ccgtcttccc taagaccatc ggcgagctcc acggccaact tggcatcgcg aggctggccc 720
 tctttacgc gcgcggcgcc tgcggcctcg tcgaccgta ccacggcact ggtgttgcca 780
 acgccggcct cgtctacttc aacggccgcc tctcgccat gtccgaggac gacctccgt 840
 accaggtccg cgtcaccgcc ggtggcgacc tcgagaccgt cggccgctac gacttcgacg 900
 gccagctcga ctgcgccatg atcgcgcacc ccaagctcga cctgtctcc ggcgagctct 960
 tcgcgctcag ctacgatgtc atcaagaagc cgtacctcaa gtacttttac ttcacgccc 1020
 acggcaccaa gtccgccgac gtcgagatcg agctcgacga agcgggaggt cttctccttg 1080
 tcgagcatca ccggcgagcc gccgcggaac atctcggcga gcttgaacac catctggtgg 1140
 tcgggcacga cgacgaagtt ctcggtgatg gcgaagtcgt ggatcatggt gggctggtcg 1200
 agctcgatct cgacgtcggc ggacttggtg ccgtcgggcg tgaagtagaa gtacttgagg 1260

tacggcttct tgatgacatc gtagctgagc gcgaagagct cgccggagac agggctcgagc 1320
 ttgggggtgcg cgatcatggc gcagtcgagc tggccgtcga agtcgtagcg gccgacggtc 1380
 tcgaggctcgc caccggcggc gacgcggacc tggtagggga ggtcgtcctc ggacatggcg 1440
 aggaggcggc cggtgaagta gacgaggccg gcgttggcaa caccagtgcc gtggtacgta 1500

<210> 112

<211> 1500

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct targeting the NCED2
 genes of barley *Hordeum vulgare* and wheat *Triticum aestivum*

<400> 112

taatacgaact cactataggg ttggcgccgt tgcgggcgta cacgccgttg atgaacgggg 60
 ggatgcggtc ggtcacgggc agcgcttgcg caggaggtgt ctccccgaca ggcgcaagt 120
 tgccggcgat ctgcaccgcg gggctcgaccg tcctggagag gccgtgcggg cgctcgagca 180
 cggttgccac aaatccttcc tcgaacgcgt cgagcgccgt gaccgcggcg cgctggaaga 240
 agttgagctt cttctcctgc ctctccacgg cgggcggtt ggggacggca atgaggggct 300
 tggcgggcgc cgctcggcg ccgcgcacga accgggacgg tgccggcgcg cgcggcgcgg 360
 agctgacggc gcgggcggag aaggtgaccg agctggaccg gcggcccgcg gggcgcgcc 420
 ggtgccgctg tatggaggag accgagctgg acgctgtcga cgcggcgccc gccaaagccc 480
 tcattgccgt ccccaagccg cccgccgtgg agaggcagga gaagaagctc aacttcttcc 540
 agcgcgccgc ggtcacggcg ctgcacgcgt tcgaggaagg atttgtggcc aacgtgctcg 600
 agcggccgca cggcctctcc aggacggctg acccgcggt gcagatcgcc ggcaacttcg 660
 cgctgtcgcg ggagacacct cctgtgcagg cgctgcccgt gaccgaccgc atcccccggt 720
 tcataacagg cggtgacgcc cgcaacggcg ccaaccgta cttcgacccc gtcgccgggc 780
 accacctgtt cgacggcgac ggcatggtgc acgctctgcg catccgcaac ggctcgccg 840
 agacctacgc ctcccgctt accgagacgg agcgcttgcg gcaggagcg cgctggggc 900
 gccgatgtt ccccaaggcc attggtgagc tccatggcca ctctgggatc gcgcgccttg 960
 ctctgttcta cgcgcgcgcg gctgcgggc tcctgacccc ctgcgcggc accggcggtg 1020
 ccaacgccgg cctggtctac ttcaacggcc acctctccc cgggtggcggg gtcgagtttg 1080
 gggcgcgga tcatggggca ctcgagctgc ccgtcgaagt cgtagcggcc gacggtcttg 1140
 aggtcgccgt cgctcggtgac gcggacgtgg tacgggatgt cgtcctcgga catggcgagg 1200
 aggtggccgt tgaagtagac caggccggcg ttggccacgc cggtgccgcg cgaggggctc 1260
 atgaggccgc agccgcgcg cgcgtagaac agagcaaggc gcgcgatccc agagtggcca 1320
 tggagctcac caatggcctt ggggaacatc gggcgcccca gcgcgcgctc ctgctgcagg 1380
 cgctccgtct cgggtgaagcg ggagcgtag gtctcgcgca cgccgttgcg gatgcgcaga 1440
 gcgtgcacca tgccgtcgcc gtcgaacagg tgggtcccg cgacggggtc gaagtacgta 1500

<210> 113

<211> 1521

<212> DNA

<213> *Hordeum vulgare*

<400> 113

```

atggccttct tcctcctcct gtgcatactc gtctctgtgg ccatcgtgtc ctacgcccac 60
cacgcaatcc ggcggaggcg ccagggtgtc gtcataggcc gtcatagaca ggccgcccctc 120
aagctgcccc ccggtcccat gggcctgcct tacgtcggcg agaccctgca gctctactcc 180
caggacccca gcgtcttcct ctctccaag cagaagcggc acggcgagat cttcaagacg 240
cacctcctgg ggtgcccgtg cgtgatgctg gcgagcccgg aggcgggcgcg cttcgtgctg 300
gtgtcgcggg cccacctctt caagccgacg taccgcgga gcaaggagcg cctcatcggc 360
ccgtcggcgc tcttcttcca ccaggcgac taccacctcc gcctccgccc gctcgtccag 420
ggcccgtcgc gcccggaggc cctgcgcaag ctctgtccgg acatcgaggc cgccgttcgc 480
tccacgctcg ccgctgggc ggacggcgac gtgcacgca ctttccacgc catgaagagg 540
ctctcgttcg acgtcggcat cgtgacgatc ttcgccgggc ggctggacga gcggcggaag 600
gaggagctca ggcggaacta cgccgtcgtg gagaaaggct acaactcctt cccaacagc 660
ttccccggga cgctatacta caaggcgatc caggcgaggc ggcggctgaa cgccgtgctg 720
agcgacgtcg tgcacgagcg tagggagcgg ggcgagcacg gcgacgacct cctcggctgc 780
ctcatgcggt cgccggcccg cgccgacgac gccgacgacg agggcgcgct gctgacggac 840
gagcaggctc ccgacaacgt catcggcgtg ctgttcgcgg cgcaggacac gacggccagc 900
gtgctcacct ggatcgtcaa gtacctccac gaccgcccga agctgctcga ggccgtcagg 960
gcggagcacg cggcgatcca cgaggccaac gacggcgga ggcggccgct gacatgggcg 1020
cagacgagga gcatgacgct gacgcacagg gtgattttgg agagcctaag gatggccagc 1080
atcatctcct tcacgttcag ggaggccgtg gccgacgtgg agtacaagg gtttcttacc 1140
cccaaggggt ggaaggatgat gccgctcttc aggaacatcc atcacagccc ggactacttc 1200
caggatccac acaagttcga cccttcgcga ttcaagggtg cgccgcggcc gaacaccttc 1260
acscggttcg ggagcggggt gcacgcgtgc ccggggaacg agctggccaa gctcgagatg 1320
ctgggtgctc tccaccacct ggtcaccggc tacaggtggg aggttggttg atcgagcgac 1380
gacgtcgagt acagcccatt cccggttccc cgccatggcc tgctcgccag ggtacggcga 1440
gatgacggcg tctgcgcggg taggaagggg tgcccactg atgaagatga caactacgac 1500
gacgacgaag tgatagtgtg a 1521

```

<210> 114

<211> 1506

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct targeting the ABA-OH-2 genes of barley *Hordeum vulgare* and wheat *Triticum aestivum*

<400> 114

```

taatacgact cactataggg cggtcgtgga ggtacttgac gatccaggtg agcacgctgg 60
ccgtcgtgtc ctgcgccgag aacagcacgc cgatgacgtt gtcggcgacc tgctcgtccg 120

```

tcagcagcgc gccctcgtcg tcggcgtcgt cgccgccggc ccgcgaccgc atgaggcagc 180
 cgaggaggtc gtcgccgtgc tcgccccgct ccctacgctc gtgcacgacg tcgctcagca 240
 cgccgttcag ccgccgcctc gcctggatcg ccttgtagta tagcgtcccg gggaagctgt 300
 tggggaagga gttgtagcct ttctccacga cggcgtagtt ccgcctgagc tcctccttcc 360
 gccgctcgtc cagccgcccc ccgaagatcg tcacgatgcc gacgtcgaac gagagcctct 420
 tcatggcgtg gaaagtgtg gcgacgtgc cgtccgccta caactccttc cccaacagct 480
 tccccgggac gctatactac aaggcgatcc aggcgaggcg gcggctgaac ggcgtgctga 540
 gcgacgtcgt gcacgagcgt agggagcggg gcgagcacgg cgacgacctc ctgggctgcc 600
 tcatgcggtc gcgggccggc ggcgacgacg ccgacgacga gggcgcgctg ctgacggacg 660
 agcaggtcgc cgacaacgtc atcggcgtgc tgttcgcggc gcaggacacg acggccagcg 720
 tgctcacctg gatcgtcaag tacctccacg accgccgta gctgctcgag gccgtcaggg 780
 cggagcacgc ggcatccac gaggccaacg acggcgggag gcggccgctg acatgggcgc 840
 agacgaggag catgacgtg acgcacaggg tgattttgga gaggcctaagg atggccagca 900
 tcatctcctt cacgttcagg gaggccgtgg ccgacgtgga gtacaaaggg tttcttatcc 960
 ccaaggggtg gaaggtgatg ccgctcttca ggaacatcca tcacagcccg gactacttcc 1020
 aggatccaca caagttcgac ctttcgcgat tcaaggtcgc tcgatccaac aacctccac 1080
 ctgtagccgg tgaccagggt gtggatgagc accagcatct cgagcttggc cagctcgttc 1140
 cccgggcacg cgtgcacccc gctcccgaac ggagtgaagg tgttcggccg cggcgccacc 1200
 ttgaatcgcg aagggtcgaa cttgtgtgga tcctggaagt agtccgggct gtgatggatg 1260
 ttctgaaga gcggcatcac cttccacccc ttggggataa gaaacccttt gtactccacg 1320
 tcggccacgg cctccctgaa cgtgaaggag atgatgctgg ccattccttag gctctccaaa 1380
 atcacctgtg gcgtcagcgt catgtcctc gtctgcgcc atgtcagcgg ccgcctcccg 1440
 ccgtcgttgg cctcgtggat cgccgcgtgc tccgcctga cggcctcgag cagctacgta 1500
 ggtacc 1506

<210> 115

<211> 3885

<212> DNA

<213> *Arabidopsis thaliana*

<400> 115

atggaagctg aaattgtgaa tgtgagacct cagctagggt ttatccagag aatggttcct 60
 gctctacttc ctgtcctttt ggtttctgtc ggatatattg atcccgggaa atgggttgca 120
 aatatcgaag gaggtgctcg tttcgggtat gacttggtgg caattactct gcttttcaat 180
 tttgccgcca tcttatgcca atatgttgca gctcgcataa gcgttgtgac tggtaaacac 240
 ttggctcaga tctgcaatga agaatatgac aagtggacgt gcatgttctt gggcattcag 300
 gcggagttct cagcaattct gctcgacctt accatggttg tgggagttgc gcatgcactt 360
 aaccttttgt ttgggggtgga gttatccact ggagtgtttt tggccgcat ggatgcgttt 420
 ttatttcctg ttttcgcctc tttccttgaa aatggtatgg caaatacagt atccatttac 480
 tctgcaggcc tggatattact tctctatgta tctggcgtct tgctgagtca gtctgagatc 540
 ccactctcta tgaatggagt gttaactcgg ttaaatggag agagcgcatt cgcactgatg 600

```

ggtcttcttg gcgcaagcat cgtccctcac aatTTTTata tccattctta ttttgctggg 660
gaaagtacat cttcgtctga tgtcgacaag agcagcttgt gtcaagacca tttgttcgcc 720
atctttgggtg tcttcagcgg actgtcactt gtaaattatg tattgatgaa tgcagcagct 780
aatgtgtttc acagtactgg ccttgtggta ctgacttttc acgatgcctt gtcactaatg 840
gagcaggtat ttatgagtcc gctcattcca gtggcttttt tgatgctctt gttcttctct 900
agtcaaatTA cgcactagc ttgggctttc ggtggagagg tcgtcctgca tgacttcctg 960
aagatagaaa taccgccttg gcttcactcg gctacaatca gaattcttgc agttgctcct 1020
gcgctttatt gtgtatggac atctggtgca gacggaatat accagttact tatattcacc 1080
caggtcttgg tggcaatgat gcttccttgc tcggtaatac cgcttttccg cattgcttcg 1140
tcgagacaaa tcatgggtgt ccataaaatc cctcaggttg gcgagttcct cgcacttaca 1200
acgttttttg gatttctggg gttgaatgtt gtttttgttg ttgagatggg atttgggagc 1260
agtgactggg ctggtggttt gagatggaat accgtgatgg gcacctcgat tcagtacacc 1320
actctgcttg tatcgtcatg tgcattctta tgctgatac tctggctggc agccacgccg 1380
ctgaaatctg cgagtaacag agcggagct caaatatgga acatggatgc tcaaaatgct 1440
ttatcttata catctgttca agaagaggaa attgaaagaa cagaaacaag gaggaacgaa 1500
gacgaatcaa tagtgcggtt ggaaagcagg gtaaaggatc agttggatac tacgtctgtt 1560
actagctcgg tctatgattt gccagagaac attctaatac cggatcaaga aatccgttcg 1620
agccctccag aggaaagaga gttggatgta aagtactcta cctctcaagt tagtagtctt 1680
aaggaagact ctgatgtaaa ggaacagtct gtattgcagt caacagtggg taatgaggtc 1740
agtgataagg atctgattgt tgaaacaaag atggcgaaaa ttgaaccaat gagtcctgtg 1800
gagaagattg ttagcatgga gaataacagc aagtttattg aaaaggatgt tgaaggggtt 1860
tcatgggaaa cagaagaagc taccaaagct gctcctacaa gcaactttac tgtcggatct 1920
gatggtcctc cttcattccg cagcttaagt ggggaagggg gaagtgggac tggaagcctt 1980
tcacggttgc aaggtttggg acgtgctgcc cggagacact tatctgcgat ccttgatgaa 2040
ttttggggac atttatatga tttcatggg caattggttg ctgaagccag ggcaaagaaa 2100
ctagatcagc tgtttggcac tgatcaaaag tcagcctctt ctatgaaagc agattcgttt 2160
ggaaaagaca ttagcagtgg atattgcatg tcaccaactg cgaagggaat ggattcacag 2220
atgacttcaa gtttatatga ttcactgaag cagcagagga caccgggaag tatcgattcg 2280
ttgtatggat tacaaagagg ttcgtcaccg tcaccgttgg tcaaccgtat gcagatgttg 2340
ggtgcatatg gtaacaccac taataataat aatgcttacg aattgagtga gagaagatac 2400
tctagcctgc gtgctccatc atcttcagag ggttgggaac accaacaacc agctacagtt 2460
cacggatacc agatgaagtc atatgtagac aatttggcaa aagaaaggct tgaagcctta 2520
caatcccgtg gagagatccc gacatcgaga tctatggcgc ttggtacatt gagctataca 2580
cagcaacttg ctttagcctt gaaacagaag tcccagaatg gtctaacccc tggaccagct 2640
cctgggtttg agaattttgc tgggtctaga agcatatcgc gacaatctga aagatcttat 2700
tacggtgttc catcttctgg caatactgat actgttggcg cagcagtagc caatgagaaa 2760
aaatatagta gcatgccaga tatctcagga ttgtctatgt ccgcaaggaa catgcattta 2820
ccaacaaca agagtggata ctgggatccg tcaagtggag gaggagggtg tggtgcgtct 2880
tatggtcggg taagcaatga atcatcgta tattctaatt tggggtcacg ggtgggagta 2940

```

ccctcgactt atgatgacat ttctcaatca agaggaggct acagagatgc ctacagtttg 3000
 ccacagagtg caacaacagg gaccggatcg ctttggtcca gacagccctt tgagcagttt 3060
 ggtgtagcgg agaggaatgg tgctgttggg gaggagctca ggaatagatc gaatccgatc 3120
 aatatagaca acaacgcttc ttctaattgt gatgcagagg ctaagcttct tcagtcgttc 3180
 aggcactgta ttctaaagct tattaaactt gaaggatccg agtggttggt tggacaaagc 3240
 gatggagttg atgaagaact gattgaccgg gtagctgcac gagagaagtt tatctatgaa 3300
 gctgaagctc gagaaaataaa ccagggtgggt cacatggggg agccactaat ttcacgggtt 3360
 cctaactgtg gagatggttg cgtttggaga gctgatttga ttgtgagctt tggagtttgg 3420
 tgcattcacc gtgtccttga cttgtctctc atggagagtc ggcctgagct ttggggaaag 3480
 tacacttacg ttctcaaccg cctacaggga gtgattgatc cggcgttctc aaagctgcgg 3540
 acaccaatga caccgtgctt ttgccttcag attccagcga gccaccagag agcgagtccg 3600
 acttcagcta acggaatgtt acctccggct gcaaaaccgg ctaaaggcaa atgcacaacc 3660
 gcagtcacac ttcttgatct aatcaaagac gttgaaatgg caatctcttg tagaaaaggc 3720
 cgaaccggta cagctgcagg tgatgtggct ttcccaaagg ggaaagagaa tttggcttcg 3780
 gttttgaagc ggtataaacg tcggttatcg aataaaccag taggtatgaa tcaggatgga 3840
 cccggttcaa gaaaaaacgt gactgcgtac ggatcattgg gttga 3885

<210> 116

<211> 1080

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct targeting the EIN2 gene of *A. thaliana*

<400> 116

taatacgact cactataggg agctgcaaca tattggcata agatggcggc aaaattgaaa 60
 agcagagtaa ttgccaccaa gtcatacccg aaacgagcac ctccttcgat atttgcaacc 120
 catttcccg gatcaatata tccgacagaa accaaaagga caggaagtag agcaggaacc 180
 attctctgga taaaccctag ctgaggcttc acattcacia tttcagcttc catcctaaat 240
 ctatctgata atataattac tcagagtagg attcaaggta aactctacag tcgtggttca 300
 ctaaaagcct cagttgagta aaattcacag atttgatcta aaacacctga atgtgagacc 360
 tcagctaggg tttatccaga gaatggttcc tgctctactt cctgtccttt tggtttctgt 420
 cgatatatt gatcccgga aatgggttgc aaatatcgaa ggaggtgctc gtttcgggta 480
 tgacttggtg gcaattactc tgcttttcaa ttttgccgcc atcttatgcc aatatgttgc 540
 agctcccaac agcgttgtga ctggtaaaca cttggctcag atctgcaatg aagaatatga 600
 caagtggacg tgcattgtct tgggcattca ggcggagttc tcagcaattc tgctcgacct 660
 taccatgggt gtgggagttg cgcattgact taaccttttg tttgggggtg agttatccac 720
 tggagtgttt ttggccgcca tggatgcgag tgggatctca gactgactca gcaagacgcc 780
 agatacatag agaagtaata ccaggcctgc agagtaaag gatactgtat ttgccatacc 840
 attttcaagg aaagaggcga aaacaggaaa taaaaacgca tccatggcgg ccaaaaacac 900

tccagtggat aactccaccc caaacaaaag gttaagtgca tgcgcaactc ccacaacat 960
 ggtaaggctc agcagaattg ctgagaactc cgcctgaatg cccaagaaca tgcacgtcca 1020
 cttgtcatat tcttcattgc agatctgagc caagtgttta ccagtcacaa cgctgttgac 1080

<210> 117

<211> 1188

<212> DNA

<213> *Arabidopsis thaliana*

<400> 117

atgggtgatgg ctgggtgctt ttctttggat gagatcagac aggctcagag agctgatgga 60
 cctgcaggca tcttggttat tggcactgct aacctgaga accatgtgct tcaggcggag 120
 tatcctgact actacttccg catcaccaac agtgaacaca tgaccgacct caaggagaag 180
 ttcaagcgca tgtgcgacaa gtcgacaatt cggaaacgtc acatgcatct gacggaggaa 240
 ttctcaagg aaaaccaca catgtgtgct tacatggctc cttctctgga caccagacag 300
 gacatcgtgg tggtcgaagt ccctaagcta ggcaaagaag cggcagtga ggccatcaag 360
 gagtggggcc agcccaagtc aaagatcact catgtcgtct tctgactac ctccggcgctc 420
 gacatgcctg gtgctgacta ccagctcacc aagcttcttg gtctccgtcc ttccgtcaag 480
 cgtctcatga tgtaccagca aggttgcttc gccggcggtc ctgtcctccg tatcgctaag 540
 gatctcgccg agaacaatcg tggagcacgt gtctcgttg tctgctctga gatcacagcc 600
 gttaccttcc gtggtccctc tgacaccac cttgactccc tcgtcgggtc ggctcttttc 660
 agtcatggcg ccgccgact cattgtgggg tcggaccctg acacatctgt cggagagaaa 720
 cccatctttg agatgggtgc tgccgctcag accatccttc cagactctga tggtgccata 780
 gacggacatt tgagggaagt tggctcacc ttccatctcc tcaaggatgt tcccggcctc 840
 atctccaaga acattgtgaa gactctagac gaagcgttta aacctttggg gataagtga 900
 tggaactccc tcttctggat agcccaccct ggaggtccag cgatcctaga ccaggtggag 960
 ataaagctag gactaaagga agagaagatg agggcgacac gtcacgtgtt gagcgagtat 1020
 ggaaacatgt cgagcgcgtg cgttctcttc atactagacg agatgaggag gaagtcagct 1080
 aaggatggtg tggccacgac aggagaaggg ttggagtggg gtgtcttggt tggtttcgga 1140
 ccaggtctca ctgttgagac agtcgtcttg cacagcgttc ctctctaa 1188

<210> 118

<211> 1080

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct targeting the CHS gene of *A. thaliana*

<400> 118

taatacgact cactataggg gccacacat ccttagctga cttctctc atctcgtcta 60
 gtatgaagag aacgcacgcg ctgcacatgt ttccatactc gctcaacacg tgacgtgtcg 120
 ccctcatctt ctcttccttt agtcctagct ttatctccac ctggtctagg atcgttgac 180

ctccaggggtg ggctatccag aagagggagt tccagtcact tatccccaaa ggttttaaacg 240
 cttcgtctag actcttcaca atgttcttgg agatgaggcc gggaacatcc ttgaggagat 300
 ggaagggtgag accaacttcc ctcaaagtgc cgtctatggc accatcagac tggaactccc 360
 tcttctggat agccccacct ggaggtccag cgatcctaga ccaggtggag ataaagctag 420
 gactaaagga agagaagatg agggcgacac gtcacgtgtt gagcgagtat ggaaacatgt 480
 cgagcgcgtg cgttctcttc atactagacg agatgaggag gaagtcagct aaggatgggtg 540
 tggccccgac aggagaaggg ttggagtggg gtgtcttgtt tggtttcgga ccaggtctca 600
 ctgttgagac agtcgtcttg cacagcgttc ctctctaaac agaacgcttg ctttctatct 660
 gcctacctac ctacgcaaaa ctttaatect gtcttatgtt ttatataata taatcattat 720
 atgtttacgc aataattaag gaagaatgac atttccaaac aaagatttga tgtcattcaa 780
 gacccataga tttaatatg taaaaagaca caaaaaagag agtacaaaaa cagtcgaata 840
 gacctgtcca gcacatatca catatcacat caaatgcatt cttccttaat tattgcgtaa 900
 acatataatg attatattat ataaaacata agacaggatt aaagttttgc gtaggtagggt 960
 aggcagatag aaggcaagcg ttctgttttag agaggaacgc tgtgcaagac gactgtctca 1020
 acagtgaagc ctggtccgaa accaaacaag acaccccact ccaacccttc tctgtcaac 1080

<210> 119

<211> 3456

<212> DNA

<213> *Lupinus angustifolius*

<400> 119

atgttgactc ttcaaccac acatgagtca agtagtcaat accctcctca tacacttata 60
 gctgagacct gtcacttga ttatctgtac tatactaate aaagttctct aattatgtca 120
 cttggagaat catccctgca atggaaatac catgttttct tgagtttttag gggaggtgac 180
 acccgcttaa gcttcactaa tcacttatat gctgcgttgg tgcgaaaagg aatcattact 240
 ttccgagatg acaaacaact tcacaaagga gatgccattt ctcaacatct gcatcaatca 300
 atccaacagt ctctagctgc cattgttgtt atctcggaga actatgcttc ttccacttgg 360
 tgttttgatg agctaaaact aattcttgaa tcgagaatag atgtttttcc agtcttttat 420
 ggtgtcactc cttctgatgt tcgataccag aaaaatagtt ttgctgaggc tttcaataaa 480
 catgttgtaa gatttgaaca agatgaagag aaagtgcata aatggagaga ttgcttgaaa 540
 gaagttgctg atttttctgg atgggagtcc aaggacatgg ctgaagcaga actcattgaa 600
 gatgttattg aaaaggatg gataaaacta caacaaaat tgccatccta caatgaagga 660
 gtggttgat ttgattcaag ggtgaagaaa atgatttcac ttttaagcat aggatcaca 720
 gatattcggg ttatcgggat atgggggatg gctggaactg gaaaaacaat tcttgctaga 780
 gtaatctacg aaacaataag tagccaattt gagattaaat gtttccttct taatgttaga 840
 gaggtttctc aaacatctga tggattgggt tccttataaa gaaaacttct ttctaccctt 900
 aagataagca acctagaaat tgatgatttg tatgatggaa agaagaaaat tatgaacctt 960
 ttgtgcaaca aaagtgttct tcttgcctt gatgacatta gtcatttaag tcagctagag 1020
 aatttggtc aaactaaagg ttggtttggt ccatgcagca gaggataat aacaacaaa 1080
 gatatgcact tactagtatc acatggtgcg tgtgagaagt atgagatgag aatcttaaat 1140

gaaagtctct cctttcaact cttcagccag aaagcattca gaagagataa acctccagag 1200
 ggttatattag aaataactaa aagtatggtc aaatatgctg gaggtcttcc tttggcactt 1260
 aaagtgttgg gttcttttgt ttgtggaaga agtctcagtc agtggaagga tgctttggat 1320
 aagataaaac aagttctgcc gaaagacatt ttgaacacac taataatagg ttatgatgga 1380
 ctagaagatg cagaaaagac tttgttttta gatattgctt tcttctttac aggacggtcg 1440
 aaaattgaag tgatacaggt attggcagat tgtggcctta atccaacaat tggaataagt 1500
 cttcttattg aaagatctct agtaagttgt tgtggaggaa ttttggaat gcatgattta 1560
 cttcaagaaa tgggtagaaa tattgtatat caagaatctc cggatgatgc aagcagacgc 1620
 agtaggttat gctctttaga agatattaac cgagtattca gaaaaaaca gggaaccaat 1680
 atcattcaag gaatagttct gaaatcaagt gacctatgtg aagcatattg gcatcctgaa 1740
 gccttctcaa aaatggataa tcttagagta ctcatcattt tgtgtgattt gcaccttccc 1800
 ctcggcctca aatgtctctc tagttcatta aaacttcttg aatggaaggg atatcctttg 1860
 gaatatctac catttggcct gcaactgcta gaacttgctc acttgaaaat gcattgcagc 1920
 aaacttaaac aactttggaa tggaactcaa attttcagag agctaaaatc aattgatctc 1980
 agtgattcca gagatctaatt tcaaactcca gatatttctg aggttccatg tcttgagagt 2040
 ttagttttga aaggttgtaa aaaccttggt gaggttcac aatctgttgc aaagcacaag 2100
 aatgttgcta tactagacct ggaaggttgc atcagtctta agaccctgcc aagaaaattg 2160
 gagatgaatg ctttggaaaa gttcattctc tccggctgct cacaaattaa aaaccttccc 2220
 gaatttgggg agagtatgga atgtctatct atgcttaatt taagagattg cacaagtctt 2280
 gtttctcttc cacagagtgt tcgaaacatg aaatccttta gagatctcaa tatccatggt 2340
 tgctcaaaat tgtttaagct gacaaacaat tcaaatgaaa ataatgtcgt ggaagaaatt 2400
 gatgagactg aaacaggtag gagagaagtg cattcatcat ggagcttttc tctccttact 2460
 gagaaagtgt ttgatttcgt aaagtatcca gtagcatgg actcgaagtt gccttctctc 2520
 tcaagtttcc ctcggttgaa gaaattagat atgggcaact gtaatctcag tgatggacca 2580
 attatagatc atattggaca tttaacatca ctggaagtgt tatatttagc tgggaacaac 2640
 tttgttgacc ttacagcaag cattggtaac ctttctcggc tacaacgcct tggtttatat 2700
 aaatgccgaa gacttaggac attgcctgag cttccacca gtgtatgcca gttacttatg 2760
 aacgactgca ctcaactgga acctatgta tttgacacac aaataatttt gaaaatatatt 2820
 gaggcaaata gatggagcct gacacgcgaa ttgtggttcc tgattccagg gagtgaaatc 2880
 ccagcatggt ttgagcatca agattatttt agcctgaaac caagtttagc gcctttcgat 2940
 tatcacgagg agtatgcttt tattgtttca acaatagtaa acatccctga ctattgcctt 3000
 tcaagtgatt ggataggaat tattgtatgc tttttactgg aaagtggttt aaaggcagac 3060
 ctacacagac atattcgtag aagtcggtc acgatcggat ggtcttttaa agatcccgat 3120
 gcgaaaacgg tttacccctt acgttcaact aaacgtcgtt ggacacattt caaaggcaat 3180
 cacctattga ttactacttt tggaagtgat catagaatat acaagcacta cttacttgt 3240
 ggcaaaagca aagtgcatt gatattttgt ggtgagaata tttgcaagtg cgggaagcta 3300
 aagctgaaaa actgtgggat ccgtgtgatt tgtaaggaag atggtgtatc gcgtagaggc 3360
 gaggaacga gtgaagtga ggtgccttcc acttcagttg aatctgatgt tcacaaaca 3420
 tcacgaataa ctgaaattac agatgaatat gaataa 3456

<210> 120

<211> 1280

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct targeting the L.
angustifolius N-like gene

<400> 120

```

taatacgact cactataggg tatcgaacat cagaaggagt gacaccataa aagactggaa 60
aaacatctat tctcgattca agaattagtt ttagctcatc caaacaccaa gtggaagagt 120
atttccattg cagggatgat tctccaagt acataattag agaactttga ttagtatagt 180
acagataatc aaagtgacag gtctcagcta taagtgtatg aggagggtat tgactacttg 240
actcatgtgt gggttgaaga gtcaacatat agcttcacgg atcccacagt ttttcagctt 300
tagcttcccc cacttgcaaa tattctcacc acaaaatata aattgcactt tgcttttgcc 360
acaagttaag tagtgcttgt atattctatg atcacctgtg tgactcttca acccacacat 420
gagtcaagta gtcaataccc tcctcataca cttatagctg agacctgtca ctttgattat 480
ctgtactata ctaatcaaag ttctctaatt atgtcacttg gagaatcatc cctgcaatgg 540
aaatactctt ccacttggtg tttggatgag ctaaaactaa ttcttgaatc gagaatagat 600
gtttttccag tcttttatgg tgtcactcct tctgatgttc gatacccgta gttttgctga 660
ggctttcaat aaacatgttg taagatttga acaagatgaa gagaaagtgc agtttgagca 720
tcaagattat tttagcctga aaccaagttt agcgccttgc gattatcacg aggagtatgc 780
ttttattggt tcaacaatag taaacatccc tgactattgc ctttcaagtg attggatagg 840
aattattgta tgctttttac tggaaagtgg tttaaaggca gacctacaca gacatatccc 900
aaaagtagta atcaataggt gattgccttt gaaatgtgtc caacgacgtt tagtgaagcg 960
taaggggtaa accgtttctg catcgggatc tttaaaagac catccgatcg tgaccggact 1020
tctacgaata tgtctgtgta ggtctgcctt taaaccactt tccagtaaaa agcatacaat 1080
aattcctatc caatcacttg aaaggcaata gtcagggatg ttactattg ttgaaacaat 1140
aaaagcatac tcctcgtgat aatcgaaagg cgctaaactt ggtttcaggc taaaataatc 1200
ttgatgctca aactgcactt tctcttcata ttgttcaaat cttacaacat gtttattgaa 1260
agcctcagca aaactacgta 1280

```

<210> 121

<211> 1527

<212> DNA

<213> Vitis pseudoreticulata

<400> 121

```

atggctggcg acgaggagac gacgacgacg gcagcaacac ttgaaacaac gtccacttgg 60
gctgttgccct ctgtttgctt tattttgatt gcactctcca tacttattga gcatgccctc 120
catctcttag ccaagtactt caacaagaag cggaggaggt ctctcattca tgctcttaac 180
aacgtcaaat cggagttgat gctcttgggg ttcgtctctt tgttgctgac tgtgtgccaa 240

```

aagtatatg cgaagatttg tatcccaagg agcgtaggtg aaacttttct tccctgcaag 300
accttgacag aaagtgattc agaagaagaa accaaatgcg aagagcaggg aaagatgtct 360
ttgctgtcta gacaaggcgt ggaggaacta caatacttaa ttttcgtgct ggccttcttc 420
cattccctct actgcgtcct cacattcggt cttgggatgg ccaagatgaa gaaatgggag 480
tcctgggagg cagaaacaag aacactggaa tatcagttta caaatgatcc acggagggtc 540
aggctcatcc atcagacatc atttggaag caacatctga gatattggag tgagcatcag 600
atacttcgtt ggccggcttg ttttattcgg cagttctatc catccgtctc caaagtggat 660
tacttgactc ttagacatgg gttcattatg gccattttg cagaaggaag caactatgac 720
ttccaaaagt atataaaaag agctttggaa aaagactttg gagggtggg gggaggaagt 780
ttctgggttt ggagtttctc catgcttttt gtgttcttca atgctcaagt attttacaac 840
tatttatggc taccctttat tccattgggtg atgctgttgt tggttggaac aaagctacag 900
ggcattataa ctaagatgtg cttagatagc catgataaag ctctcgttgt tagaggaact 960
ttgcttgta ggcccagtga tcacttcttc tggtttgaa aaccggaatt gctctacat 1020
cttatgcact ttatatgtt tcagaactct tttcaactgg cgttctttac atggacttgg 1080
tacaaatttg gattcagatc atgcttccat gatacaactg aggatatcgt cataaggctt 1140
gtcatgggtg tgtagtaca actcctttgt ggctacgtga cactgcctct gtatgccctg 1200
gtcacgcaga tggggacatc aatgaggaca attgtcttta ctgaggaggt cgttgaaggt 1260
ctgaacagat ggagaaggaa agccaagaaa aacatagcac gcaggaacaa cactcagct 1320
cgtccctccc tggatgcttc actcgacaat tcaccttctt ttaacactct ggatacttct 1380
ttctctgtag acctcgatca gccatcatca gatgctgggt atttgactgt tgaaatatca 1440
gatgaagaga cggtcgctac taaacggcca gaaccgcgtc agaagttggg atcttttgag 1500
ggtttcgact cgtgcaaaac atcataa 1527

<210> 122

<211> 1480

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a first ledRNA construct targeting a Vitis MLO gene

<400> 122

taatacgact cactataggg tagccataaa tagttgtaaa atacttgagc attgaagaac 60
acaaaaagca tggagaaaact ccaaaccag aaacttctc ccaccaccac tccaaagtct 120
ttttccaaag ctctttttat atacttttgg aagtcatagt tgcttcttc tgcaaaatgg 180
gccataatga acccatgtct aagagtcaag taatccactt tggagacgga tggatagaac 240
tgctgaataa aacaaaccgg ccaacgaagt atccgatgct cactccaata tctcagatgt 300
tgctttccaa atgatgtctg atggatgagc gtgaacctcc gtggatcatt tgtaaactga 360
tattccagtg ttcttgtttc tgctcccag gactcccatt tcttcattct ggccatccca 420
agaccgaatg tgaggacgca gtagagcaca tcatgttgaa agcaacatct gagatattgg 480
agtgagcatc ggatacttcg ttggccggtt tgttttattc agcagttcta tccatccgtc 540

tccaaagtgg attacttgac tcttagacat gggttcatta tggcccattt tgcagaagga 600
 agcaactatg acttccaaaa gtatataaaa agagcttttg aaaaagactt tggagtgggtg 660
 gtgggaggaa gtttctgggt ttggagtttc tccatgcttt ttgtgttctt caatgctcaa 720
 gtatttttaca actattttatg gctacccgta attccattgg tgatgctggt gttggttgga 780
 acaaagctac agggcattat aactaagatg tgcctagata gccatgataa agctctcgtt 840
 gtttagaggaa ctttgcttgt caggccccagt gatcacttct tctggtttgg aaaaccggaa 900
 ttgctcctac atcttatgca ctttatattg tttcagaact cttttcaact ggcgttcttt 960
 acatggactt ggtacaaatt tggattcaga tcatgcttcc atgatacaac tgaggatatac 1020
 gtcataaggc ttgtcatggg tgtgttatgg ctttccttct ccactgttcc agaccttcaa 1080
 cgactccctc agtaaaagaca attgtcctca ttgatgtccc catctgcgtg accagggcat 1140
 acagaggcag tgtcacgtag ccacaaagga gttgtactaa cacacccatg acaagcctta 1200
 tgacgatatac ctcagttgta tcatggaagc atgatctgaa tccaaatttg taccaagtcc 1260
 atgtaaagaa cgccagttga aaagagttct gaaacaatat aaagtgcata agatgtagga 1320
 gcaattccgg ttttccaaac cagaagaagt gatcactggg cctgacaagc aaagttcctc 1380
 taacaacgag agctttatca tggctatcta ggcacatctt agttataatg ccctgtagct 1440
 ttgttccaac caacaacagc atcaccaatg gaattacgta 1480

<210> 123

<211> 894

<212> DNA

<213> *Myzus persicae*

<400> 123

atgttcaaac acttgtgcaa taccgtttca caaagtataa aacctagtag ttttttatca 60
 aaagtttggt caaacaaata tctcgtcgtg ccgtaccgga tagcgatttt taacaacatg 120
 ggaagttaca aattgtacct ggccgtcatg gcaatagctg tcatagctgc agttcaggaa 180
 attagttgca aggttcagac ttccgaacag gacgatgatc aggaaggata ttacgatgat 240
 gagggaggag tgaacgataa tcaggagaa gagaacgata atcagggaga agagaacgat 300
 aatcaggagag aagagaacga taatcaggga gaagagaagg aagaagtttc cgaaccagag 360
 atggagcacc atcagtgcga agaatacaaa tcgaagatct ggaacgatgc atttagcaac 420
 ccgaaggcta tgaacctgat gaaactgacg tttaatacag ctaaggaatt gggctccaac 480
 gaagtgtgct cggacacgac ccgggcctta tttaacttcg tcgatgtgat ggccaccagc 540
 ccgtacgccc acttctcgtc aggtatgttt aacaagatgg tggcgtttat tttgaggag 600
 gtggacacga catcggacaa atttaaagag acgaagcagg tggtcgaccg tatctcgaaa 660
 actccagaga tccgtgacta tatcaggaa tcggccgcca agaccgtcga cttgctcaag 720
 gaaccaaga ttagagcacg actgttcaga gtgatgaaag ctttcgagag tctgataaaa 780
 ccaaacgaaa acgaagcatt aatcaaacag aagattaagg ggttaaccaa tgctcccgtc 840
 aagttagcca aggttgccat gaaaacggtt ggacgtttct ttagacattt ttaa 894

<210> 124

<211> 960

<212> DNA

<213> *Myzus persicae*

<400> 124

```

atgactgaga caatgcaact ccgtgggtacc cttcgtgggc ataatggttg ggttacgcag 60
atcgccacca atccgatcca cactgacatg attctgtctt gttcacgaga caagaccttg 120
attgtttggg atctgacacg tgatgagctc aactatggta tccccaagaa acgtttgtac 180
ggacattcgc acttcgtcag cgacgtcggt ctttcatcag atggtaacta cgctctttcc 240
ggttcttggg ataagactct tcgtctgtgg gatttggtg ctggacgtac cactcgtcgt 300
tttgaagacc acaccaagga tgtattgagc gttgccttct ctgctgacaa ccgtcaaadc 360
gtttctggaa gtcgggacaa gactatcaag ttgtggaata ctttggctga gtgcaaatac 420
actattcagg atgatggaca tagcgattgg gtatcatgtg tacggttctc tcctaataac 480
cataacccaa tcattgtgag tgctggttgg gacaagggtg tcaaggtagt gaacttaact 540
aactgccgca tcaagaccaa ccattatgga cacactggat accttaacac cgttactgtt 600
tcacctgatg gttctttgtg tgcttcagga ggaaaagatt gcaaagctat gttatgggat 660
cttaatgacg gcaaacactt gcacacactg gaccataacg atatcattga agctttgtgc 720
tttagcccca accgttactg gttgtgcgct gcatttggac catcaatcaa aatttgggat 780
ttggaaagca aagaaatggt tgaggaaact cgcccagaag ttgtatctca atcacagaat 840
agcaataaccg aaccacccag atgtctgtca cttgcatggt caactgatgg acaaacattg 900
tttgcctgat actcagacaa taacattaga gtttggcaag tgtctgtcag tgctcggtta 960

```

<210> 125

<211> 1401

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric construct encoding the ledRNA targeting *M.persicae* C002 gene

<400> 125

```

gaatttctaat acgactcact atagggtgct ccattctctg ttcggaaact tcttccttct 60
cttctccctg attatcggtc tcttctccct gattatcggt ctcttctccc tgattatcgt 120
tctcttctcc ctgattatcg ttcactctc cctcatcgc gtaatactct tctgatcat 180
cgtcctgttc ggaagtctga accttgcaac taatttctg aactgcagct atgacagcta 240
ttgccatgac ggccaggtag aatttgtaac ttcccatggt gttaaaaatc gctatccggt 300
acggcacgac gagatatttg tttgaacaaa cttttgataa aaaactacta ggttttatac 360
tttgtgaaac ggtattgcac aagtgtttga acataagaga gttggaagtt acaaattgta 420
cctggccgtc atggcaatag ctgtcatagc tgcagttcag gaaattagtt gcaaggttca 480
gacttccgaa caggacgatg atcaggaagg atattacgat gatgaggag gagtgaacga 540
taatcaggga gaagagaacg ataatacagg agaagagaac gataatcagg gagaagagaa 600
cgataatcag ggagaagaga aggaagaagt ttccgaacca gagatggagc acccaacagt 660
gcgaagaata caaatcgaag atctggaacg atgcatttag caaccgaag gctatgaacc 720
tgatgaaact gacgtttaat acagctaagg aattgggctc caacgaagtg tgctcggaca 780

```

cgacccgggc cttatttaac ttcgtcgatg tgatggccac cagcccgtac gccacttct 840
 cgctaggtat gtttaacaag atggtggcgt ttattttgag ggaggtggac acgacatcg 900
 acaatctgaa cagtcgtgct ctaatcttgg gttccttgag caagtggacg gtcttggcgg 960
 ccgagttcct gatatagtca cggatctctg gagttttcga gatacgggcg accacctgct 1020
 tcgtctcttt aaatttgtcc gatgtcgtgt ccacctccct caaaataaac gccaccatct 1080
 tgtaaacaat acctagcgag aagtgggcgt acgggctggt ggccatcaca tcgacgaagt 1140
 taaataaggc ccgggtcgtg tccgagcaca cttcgttga gcccaattcc ttagctgtat 1200
 taaacgtcag tttcatcagg ttcatagcct tcgggttgc aaatgcatcg ttccagatct 1260
 tcgatttgta ttcttcgcac tgtaacaag cttagcatat ccatgatatc tgtagtttt 1320
 tttcctgaaa gagcgccgc cctagcataa cccgcgggg cctcttcggg ggtctcgcg 1380
 ggttttttgc tgaaaggatc c 1401

<210> 126

<211> 1401

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric construct encoding the ledRNA targeting M.persicae

Rack-1 gene

<400> 126

gaattctaata acgactcact atagggttat ggatattagg agagaaccgt acacatgata 60
 cccaatcgct atgtccatca tcctgaatag tgtatttgca ctcagccaaa gtattccaca 120
 acttgatagt cttgtccga cttccagaaa cgatttgacg gttgtcagca gagaaggcaa 180
 cgctcaatac atccttgggtg tggcttctaa aacgacgagt ggtacgtcca gcagccaaat 240
 cccacagacg aagagtctta tcccaagaac cggaaagagc gtagttacca tctgatgaaa 300
 gaacgacgtc gctgacgaag tgcgaatgtc cgtacaaacg tttcttgggg ataccatagt 360
 tgagctcatc acgtgtcaga tcccaacaa tcaaggtctt gtcccggttc ttgggataag 420
 actcttcgtc tgtgggattt ggctgctgga cgtaccactc gtcgttttga agaccacacc 480
 aaggatgtat tgagcgttgc cttctctgct gacaaccgtc aaatcgtttc tggaagtcgg 540
 gacaagacta tcaagttgtg gaatactttg gctgagtga aatacactat tcaggatgat 600
 ggacatagcg attgggtatc atgtgtacgg ttctctccta atatccataa cccaaccatt 660
 gtgagtgtc gttgggacaa ggttgtcaag gtatggaact taactaactg ccgcatcaag 720
 accaaccatt atggacacac tggatacctt aacaccgtta ctgtttcacc tgatggttct 780
 ttgtgtgctt caggaggaaa agattgcaaa gctatgttat gggatcttaa tgacggcaaa 840
 cacttgca cactggacca taacgatatc attgaagctt tgtgcttttag cccaaccgt 900
 tacacagaca tctgggtggt tcggtattgc tattctgtga ttgagataca acttctgggc 960
 gaagttcctc aaccatttct ttgtttcca aatcccaaat tttgattgat ggtccaaatg 1020
 cagcgcaaa ccagtaacgg ttggggctaa agcacaagc ttcaatgata tcgttatggt 1080
 ccagtgtgtg caagtgtttg ccgtcattaa gatcccataa catagctttg caatcttttc 1140
 ctctgaagc acacaaagaa ccatcaggtg aaacagtaac ggtgttaagg tatccagtgt 1200

gtccataatg gttgggtcttg atgcggcagt tagttaagtt ccataccttg acaaccttgt 1260
 cccaaccagc actcacaatg gttaacccgg gtagcatatc catgatatct gttagttttt 1320
 ttcctgaaag agcggccgcc ctagcataac cccgcggggc ctcttcgggg gtctcgcggg 1380
 gttttttgct gaaaggatcc c 1401

<210> 127

<211> 2396

<212> DNA

<213> *Helicoverpa armigera*

<400> 127

agacattgat tagtgagctc caaactccgt acgtacgttc ttagtttagt ttgttcgttc 60
 gtattgtcgc agtcacatcg ctccggtgcc cgcttcgaca tttcccgcga aaagtgcagt 120
 aacatatccg tgatctgtgt gaatatgtca gtgacttttt taaattaatt ttttaatagc 180
 aaaattgtga tcgaaggaaat ttttacaaga tgacggctgg gaatgaagag catgagcctc 240
 taattacatc gtctgtcgac aatcagcgtg tggcctacag taattcacca ccgatgacc 300
 gcacaccaga atcttcttcc ccacgcggca gtggcggaga agtaacgcta gccataccat 360
 cacaccgcaa ctatggagcc atcggaggcg tggagaaggt cacatacacc tgggcagaca 420
 tcaatgcctt tgctactgaa tccaggtcta ggtcccgaag gatttggaac ttctggaagc 480
 cctccgccag tggcatgttc cagcaaagga aacagttgtt gaggaatgta aatggagccg 540
 cctaccaggc cgaactgtc gccatcatgg gatcctccgg tgccgggaag accacactcc 600
 tcaacactct gaccttccgc actccaagcg ggggtgctgtc cagtggcact cgagcactga 660
 acggccagcc tgctaccctt gaggcgttat cagcactgtc tgcgtatgtt cagcagcagg 720
 atctgttcat tggcacgctg actgtgaagg agcatttagt attccaggct atgggtgcga 780
 tggaccgaca tataccgtat gcgcagcgca tgaggagagt tcaagagggtt attactgagt 840
 tggcgctaac aaaatgccag aacacagtga taggcatccc tgggcggctg aagggtatct 900
 ccggcgggga gatgaagagg ctgtccttcg ccagcgaggt gtcacggat ccaccgtca 960
 tgttctgcga tgaaccacc tctggactcg attcttttat ggcgcagaat gttatacagg 1020
 tactgaaagg tctcgcacaa aaaggcaaga cagtcgtatg cacgatccac cagccgtctt 1080
 cggagctgta cgcatgttc gataagctgc tcatcatggc agacgggaag gtcgccttcc 1140
 tcggctcccc tgatcaggct aatgatttct ttaaagacct aggagcagcg tgctctccta 1200
 actacaacc agcggaccac ttcattcaac tcttggcggg agtgccgggc agggaggaga 1260
 ccacgcgcac cactatcgat actgtctgca cggcattcgc gcgctctgag gtcggctgca 1320
 agattgctgc agaagctgaa aatgcactct actttgagcg caagatatcg cagggtctggg 1380
 cggacccggc gtggtctgaa gccacggcta tccgcgcgcg ccgctcgccg tacaaggcgt 1440
 cgtggtgcgc gcagttccgc gcggtgctgt ggcgtcgtg gctgtccgtc actaaggagc 1500
 ccatgctcat caaagtgcgc ttctacaga ctattatggt atcgatcctg atcggcgtga 1560
 tctacttcgg gcagcacctg gaccaggacg gcgtgatgaa catcaacggc gccatttca 1620
 tgttcctcac caacatgacc ttccagaaca tcttcgtgt tattaacgta ttctgctcag 1680
 aactgccaat attcatacga gaacaccact ccgggatgta tcgagctgac gtgtacttcc 1740
 tatcgaagac gttagccgaa gcacctgtgt tcgccacat accacttgtg ttcaccacca 1800

tagcatacta catgataggg ctgaaccctg aacctaagcg gttctttata gcgtccggtt 1860
 tggctgccct gattactaac gttgctacgt cgtttggtta cctgatatcg tgtgccagca 1920
 acagcgtgag catggcagcg tcagtgggac ctcccatcat catccccttc atgttggttcg 1980
 gaggtttctt cctcaacact ggctccgtac caccatacct gggctggata tcgtacctgt 2040
 cctggttcca ctacggcaac gaagcgtgc tggtaacca gtggtctgga gtggaaacca 2100
 tcgcctgcac ccgggagaac ttcacctgtc ccgcctctgg gcaggtcgtc ttggatactc 2160
 tttagcttttc tgaggatgac ttcacaatgg acgtgggtgaa catgatccta cttttcatcg 2220
 gcttcagatt tttggcgtat ctgcgtctct tgtaccgcgc tcgccgaggc aagtgagtct 2280
 taggtacaaa atgctgcgag aatgggccat atgaaggaag aatgttgaat aaatagtgtg 2340
 attattttagg atgtaaggag tcaatggaga ttgataaat aaaacaattt ataccg 2396

<210> 128

<211> 1480

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct targeting a ABC
 transporter white gene of Helicoverpa armigera

<400> 128

taatacgact cactataggg tatatgtcgg tccatccgca ccatagcctg gaatactaaa 60
 tgctccttca cagtcagcgt gccaatgaac agatcctgct gctgaacata cgcagacagt 120
 gctgataacg cctcaggggt agcaggctgg ccgttcagtg ctcgagtgcc actggacagc 180
 accccgcttg gagtgcggaa ggtcagagtg ttgaggagtg tggctttccc ggcaccggag 240
 gatcccatga tggcgagcag ttgcctggg taggcggctc catttacatt cctcaacaac 300
 tgtttccttt gctggaacat gccactggcg gagggcttcc agaagttcca aatccttcgg 360
 gacctagacc tggattcagt agcaaaggca ttgatgtctg cccagggtgta tgtgaccttc 420
 tccacgcctc cgatggctcc atagtgttgc cagcaaagga aacagttggt gaggaatgta 480
 aatggagccg cctacccagg cgaactgctc gccatcatgg gatcctccgg tgccgggaag 540
 accacactcc tcaacactct gaccttccgc actccaagcg gggctgtgtc cagtggcact 600
 cgagcactga acggccagcc tgctaccctt gaggcgttat cagcactgtc tgcgtatgtt 660
 cagcagcagg atctgttcat tggcacgctg actgtgaagg agcatttagt attccaggct 720
 atggtgcgga tggaccgaca tatacccgta tgcgcagcgc atgaggagag ttcaagaggt 780
 tattactgag ttggcgctaa caaaatgcc gaacacagtg ataggcatcc ctgggcggct 840
 gaagggtatc tccggcgggg agatgaagag gctgtccttc gccagcgagg tgctcacgga 900
 tccaccgctc atgttctgcg atgaaccac ctctggactc gattctttta tggcgagaa 960
 tggtatacag gtactgaaag gtctcgaca aaaaggcaag acagtcgtat gcacgatcca 1020
 ccagccgtct tcggagctgt acgcgatgat gaagtgttc gctgggttgt agttaggagg 1080
 acacgtgct cctaggtctt taaagaaatc attagcctga tcaggggagc cgaggaaggc 1140
 gaccttcccg tctgccatga tgagcagctt atcgaacatc gcgtacagct ccgaagacgg 1200
 ctggtggatc gtgcatacga ctgtcttgcc tttttgtgcg agacctttca gtacctgtat 1260

aacattctgc gccataaaag aatcgagtcc agaggtgggt tcatcgcaga acatgagcgg 1320
 tggatccgtg agcacctcgc tggcgaagga cagcctcttc atctccccgc cggagatacc 1380
 cttcagccgc ccagggatgc ctatcactgt gttctggcat tttgttagcg ccaactcagt 1440
 aataacctct tgaactctcc tcatgcgctg cgcatacgta 1480

<210> 129

<211> 811

<212> DNA

<213> *Linepithema humile*

<400> 129

agagagaacg atgaggacaa tgagatggaa aaaacaacaa cgtcccaacg tcccttcgac 60
 gacgccattc caccagccct ataaaacccc gaggatcacc ggcgteccaa cattactcgg 120
 tcagagtctc gaggaacgcc gtgtccgaga tgatcatcac caggaaccgc atcaaccgcg 180
 caactctaata ctgcgttctg gcgtcgtggc tttgcttggc gtctcgcgct tccgccgaat 240
 acgaatcgcg ggagatgtcg aacggcggac cgggcgtcga cgcctcgtgc atcgagggca 300
 agtgcataaa gcgcaccgcc acgcaggatg ctaccgccag catgtgggtc ggccccgcgtt 360
 tgggaagacg gcgcagatcg gacgagaagc aggaagtga ttccgagata caggctcttg 420
 cggaagcctt ggatagcggg cgtttggccc tatttgccat tccagctaac gacaagagac 480
 aaccgactca atttacaccg cgactggggc gaggatcaga cgaggaccta tcctcctacg 540
 gagacgcgat tgagaggaac gagatcgacg atcgtatatt acccgcggtta ttcgcgccgc 600
 gtttaggacg acgaattcct tggtcaccgt cgccgagact tggacgcca ttacgcagca 660
 ttttgcgaaa aatgtaggcg ccgtcgaaag attattatca aaagttacaa atgaagagt 720
 atctcgtaga cctgcgcgtg aagatgaaat aacaactaaa attatagcac tattaagaca 780
 taaagaaata aagtactgat gtttatttgt a 811

<210> 130

<211> 1360

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct targeting a PBAN gene in Argentine ants

<400> 130

taatacgact cactataggg aattcacttc ctgcttctcg tccgatctgc gccgtcttcc 60
 caaacgcggg ccgaaccaca tgctggcggg agcatcctgc gtggcgggtc gcttcatgca 120
 cttgccctcg atgcacgagg cgctcgacgc cggctccggc ttcgacatct cccgcgattc 180
 gtattcggcg gaagcgcgag acgccaagca aagccacgac gccagaacgc agattagagt 240
 tgcgcgggtg atgcggttcc tggatgatgat catctcggac acggcggttc tcgagactct 300
 gaccgagtaa tgttgggacg ccgatgatcc tcgggggttt atagggttg tggaatggcg 360
 tcgtcgaagg gacgttggga cgttggtgtt ttttccatct cattgtcctc atcgttcacg 420
 ccgtgtccga gatgatcatc accaggaacc gcatcaaccg cgcaactcta atctgcgttc 480

tggcgtcgtg gctttgcttg gcgtctcgcg cttccgccga atacgaatcg cgggagatgt 540
 cgaacggcgcg accggggcgtc gacgcctcgt gcatcgaggg caagtgcattg aagcgcaccg 600
 ccacgcagga tgctaccgcc agcatgtggt tcggccccgcg tttgggaaga cggcgcagat 660
 cggacgagaa gcaggaagtg aattccccgta atacaggctc tggcggaagc cttggatagc 720
 gggcgtttgg ccctatttgc cattccagct aacgacaaga gacaaccgac tcaattttaca 780
 ccgcgactgg ggcgaggatc agacgaggac ctatcctcct acggagacgc gattgagagg 840
 aacgagatcg acgatcgtat attaccgcg ttattcgcgc cgcgtttagg acgacgaatt 900
 ccttggtcac cgtcgccgag acttggacgc caattacgca gcattttgcg aaaaatgaaa 960
 catcagtact ttattttctt atgtcttaat agtgcataa ttttagttgt tatttcatct 1020
 tcacgcgcag gtctacgaga tcaactttca tttgtaactt ttgataataa tctttcgacg 1080
 gcgcctacat ttttcgcaaa atgtcgcgta attggcgctc aagtctcggc gacggtgacc 1140
 aaggaattcg tcgtcctaaa cgcggcgcg ataacgcggg taatatacga tcgtcgatct 1200
 cgttcctctc aatcgcgtct ccgtaggagg ataggtctc gtctgatcct cgcgccagtc 1260
 gcggtgtaaa ttgagtcggt tgtctcttgt cgtagctgg aatggcaaat agggccaaac 1320
 gcccgctatc caaggcttcc gccagagcct gtattacgta 1360

<210> 131

<211> 1480

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct targeting a gene encoding V-type proton ATPase catalytic subunit A of *L. cuprina*

<400> 131

taatacgact cactataggg aagttcttgt catagaaatc atccaaagca cgcattgtatt 60
 tggagtagga aatcaaccaa ttgatggagg ggaaatgttt acgttgggcc aatttcttgt 120
 ccaaacccca gaacacttgt acgataccca aagtggcaga agtaacggga tcagagaaat 180
 caccaccagg aggagataca gcaccgacaa tggaaacgga accttcacgt tcagggttac 240
 ccaaacactt gacacgacca gcacgttcgt agaaggaggc caaacgggca cccaagtagg 300
 ctgggtaacc ggaatcggca ggcatttcag ccaaacgacc agaaatttca cgaagagctt 360
 cggcccaacg ggaggtagaa tcagccatca tagatacgtt gtaaccata tcacggaagt 420
 attcagacaa ggtaataaccg gtataaacga ttccggttac ccagcctact tgggtgcccc 480
 tttggcctcc ttctacgaac gtgctggtcg tgtcaagtgt ttgggtaacc ctgaacgtga 540
 aggttccgtt tccattgtcg gtgctgtatc tctcctggt ggtgatttct ctgatcccg 600
 tacttctgcc actttgggta tcgtacaagt gttctggggg ttggacaaga aattggccca 660
 acgtaaacad ttcccctcca tcaattgggt gatttctac tccaaataca tgcgtgcttt 720
 ggatgatttc tatgacaaga acttccccgta attcgtacca ttgcgtacca aggtcaagga 780
 aatcttgcaa gaagaagaag atttgtccga aattgtacaa ttggtcggta aggttcatt 840
 ggccgaaact gacaagatca ccttgaagt cgccaaattg ctttaaggacg atttcttgca 900
 acagaactcc tactcatcat acgacagatt ctgccccttc tacaagagtg tgggtatgtt 960

gaagaacatc attgccttct acgacttggc tcgtcactcc gtcgaatcca ccgctcaatc 1020
 tgaaaacaaa atcacctgga atgtcattct gaaagcctgt tgtaaattctt cgtgtaattg 1080
 ttcaaagtca gccttgatct tggcttcacc gtccttaacg ggatccttga atttcatgga 1140
 agacaattgg tacataatgt tacccatagc ttcacggatg acattccagg tgattttgtt 1200
 ttcagattga gcggtggatt cgacggagtg acgagccaag tcgtagaagg caatgatgtt 1260
 cttcaacata cccacactct tgtagaaggg gcagaatctg tcgtatgatg agtaggagtt 1320
 ctgttgcaag aaatcgtcct taagcaattt ggcgacttcc aaggtgatct tgtcagtttc 1380
 ggccaatgaa gccttaccga ccaattgtac aatttcggac aaatcttctt cttcttgcaa 1440
 gatttccttg accttggtac gcaatggtac gaattacgta 1480

<210> 132

<211> 1480

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct targeting a gene
 encoding RNase 1/2 of *L. cuprina*

<400> 132

taatacgact cactataggg aataatttgt ggtagacata gcgggttact tcctcatgtt 60
 cgtaaagca gacctggtat tgtctcatga aacgggaaga tgacaattca aaaccaacat 120
 tgacgagagt agtgccacca ttgcagctgc tgcccactt cttagctaca aaagcaggcc 180
 aactggtgca aacaagactg ggcaggaggt gctgaacacc attgacttta aaagtgggtgc 240
 cactaacaca ggtagcgata tgggtcttgc ctgataaagg atgagcaaag ccaactggtac 300
 aatggatctc aatatcttgc ccagcagcca catcaattct tccagtatca gaaaaagggt 360
 agagttcagt agtgccagggt ttgatataca agggttgttt agccttaaga ccaccgcgaa 420
 tgggtatgga acaaccacca ctgcgtggaa tattgagatc cattgtacca gtggctttgc 480
 tcacccctta tcaggcaaga cccatatcgc tacctgtgtt agtggcacca cttttaaagt 540
 caatgggtgtt cagcactccc tgcccagtct tgtttgcacc agttggcctg cttttgtagc 600
 taagaagtcg ggcagcagct gcaatggtgg cactactctc gtcaatgttg gttttgaatt 660
 gtcactcttc cgtttcatga gacaatacca ggtctgcttt aacgaacatg aggaagtaac 720
 ccgctatgtc taccacaaat tattcccgta cccaacagcg tgccactttc ctattcatta 780
 atgcagctcc ccagtggcaa gttttcaatg ccggttaatt ggctcgtgta gaggatgggtg 840
 tacgcgcctg ggtgtccaaa aataaaatca atgttcgatg ctataccggt gtttatgggtg 900
 tcaccactct acccaacaaa gagggacgtg agactcctct atatttgtct cgtgatgcc 960
 ataataatgg tttgattcct gttcccaaata tatacttccg tgttggttata caacctgcc 1020
 ccaataagggt tattgttttc gttggtgtca caggcataag aataaccagc agtgatatca 1080
 gttttcttcc aactaatata gttaacctta tcactgacat ctttgcaaat aatatagtcc 1140
 tttttgattt gttccaaagt caaatgggga ttgttgacac caacgaaaac aataccctta 1200
 ttgggtggcag gttgtataac cacacggaag tataatttgg gaacaggaat caaaccatta 1260
 ttattggcat cacgagacaa atatagagga gtctcacgtc cctctttgtt gggtagagtg 1320

gtgacaccat aaacaccggt atagcatcga acattgattt tatttttggga caccagcg 1380
 cgtacaccat cctctacacg agcccaatta cggcattga aaacttgcca ctggggagct 1440
 gcattaatga ataggaaagt ggcacgctgt tgggtacgta 1480

<210> 133

<211> 1480

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct targeting a gene
 encoding chitin synthase of *L. cuprina*

<400> 133

taatacgact cactataggg tggttgatt tttatattaa caccatacac ctcaaagca 60
 gctttttcaa tgctgtaaat caatattttt acatattcat tgagaggagg atttttcgga 120
 ttctccacat atttttcatc cagtataaag gcatcatcaa agaaaatatt agctaaaatg 180
 taaaaaaaaaag aaaacttaat tgatagatac tcttcattcc aatccacaat tcgtctattt 240
 aatgttatac attgttcaac caaaaaacca caataccacg gacaaatgaa cagtttttcg 300
 gtgggtaaat ttttatcatt ttttgacgc catatatgat ttgttatcca caattgtgac 360
 agccaccaca acaaccatat ccaaagataa tctttagcga ccacattaaa aaaactccat 420
 gtatcgtgac cgaagcccag tgacgttgat aaaaatttac ccaccgaaaa actgttcatt 480
 tgtccgtggt attgtggttt tttggtgaa caatgtataa cattaaatag acgaattgtg 540
 gattggaatg aagagtatct atcaattaag ttttcttttt ttacattttt agctaataat 600
 ttctttgatg atgcctttat actggatgaa aaatatgtgg agaatccgaa aaatcctcct 660
 ctcaatgaat atgtaaaaaat attgattacc agcattgaaa aagctgcatt tgaggtgtat 720
 ggtgttaata taaaaatcaa gccacccgta aaaattgaaa caccttatgg cggtcgtttg 780
 gtgtggacac tgcctggctc ctcaaagatg attgcccatt taaaaacaa agataaaaata 840
 cgacataaga aacgctggct acaggttatg tacatgtact atttgttggg ttttcgtata 900
 atggaattgg aatcagtatc ggccaagcgt aaggcagtga tagcagaaaa tacatttttg 960
 ctggctcttg atggtgatat tgactttcaa ccgcaggcag tgcaactggt aatagaccgt 1020
 atgaaggcca tagatgaatt aggtgctagc caggactaca taaaacacaa ccaataacat 1080
 gctctgttgc tttttgcaac caatgacctg tagcgtattc gaagatttga taccaaacca 1140
 tagggcctct accaactgga tgaatacgac cacaggcagc acctaattca tctatggcct 1200
 tcatacggct tattaacagt tgcaactgct gcggttgaaa gtcaatatca ccatcaagag 1260
 ccagcaaaaaa tgtattttct gctatcaact ccttacgctt ggccgatact gattccaatt 1320
 ccattatacg aaaacccaac aaatagtaca tgtacataac ctgtgaccag cgtttcttat 1380
 gtcgtatttt atctttgttt tttaaatggg caatcatctt tgagcgacca ggcagtgtcc 1440
 acaccaaacg accgccataa ggtgtttcaa tttttacgta 1480

<210> 134

<211> 1481

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct targeting a gene encoding ecdysone receptor (EcR) of *L. cuprina*

<400> 134

```

taatacgact cactataggg tgaaagatca tcacgacctg atgatatgga attagctgat 60
tcagatctgg gtgtatgatg catcatacta ctgttattca tatgatgatg gtgatgatga 120
tttaatccat tactgatgat actgtgaatg ccaatattgg catgaataac ttggccattt 180
tgtgaggcct gtaacgagtt taactgttgg gcattgctaa cgatattggg accattaata 240
ttgaccgata agccaccacc accgccgcca ataccatgt ggccatttgt gtggtgggaa 300
ctgctattac tgtgattact gttgctgttg tgggtgaaat gattgtggct gtgattgtga 360
ttattcactt gactgccacc accaccacc agaccattga gtgaagtcac accgggtaca 420
ccaccaccac ctccacctcc tccaacaaat cacagtaata gcagttccca ccacacaaat 480
ggccacatgg gtattggcgg cggtggtggg ggcttatcgg tcaatattaa tgggtcccaat 540
atcggttagca atgcccaca gttaaactcg ttacaggcct caaaaatgg ccaagttatt 600
catgccaata ttggcattca cagtatcatc agtaatggat taaatcatca tcaccatcat 660
catatgaata acagtagtat gatgcatcat acaccagat ctgaatcagc taattccata 720
tcatcaggtc gtgatgatct ttcacccgta tccaccaaat caccacctta gtggttcgaa 780
acacttggtg tccatttggt gagaccgcgc cagtggaaaa cattatgggg tctacagttg 840
tgagggttgt aaagggttct tcaaacgtac cgtacgcaag gacttgacat atgcttgtcg 900
tgaggacaga aattgcatta tagataaacg acaaagaaat cgttgccagt attgtcgtaa 960
tcaaaagtgt ttagcttggt gcatgaaacg cgaagcggtc caagaggaa gacaacgtgg 1020
tactcgtgct gctaacgcta gagctgcctt ttgctcggct tcaatgatgc gttctatagt 1080
gagatcacgt aatgaactgc tgggtttaa gtcttctccg ccagcaccaa ccacattgct 1140
taccaccaca ccacctcctc caccaccgcc agcaccagca gctctagcgt tagcagcacg 1200
agtaccacgt tgtcgttcct cttggaccgc ttcgcgtttc atgccacaag ctaaacactt 1260
ttgataacga caatactggc aacgatttct ttgtcgttta tctataatgc aatttctgtc 1320
ctcacgacaa gcataatgtc agtccttgcg tacggtagct ttgaagaacc ctttacaacc 1380
ctcacaactg tagaccccat aatgttttcc actggcgagg tctccacaaa tggaacacaa 1440
gtgttttcgaa ccactaaggg ggtgatttgg tggatacgta g 1481

```

<210> 135

<211> 1481

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct targeting a gene encoding gamma-tubulin 1/1-like of *L. cuprina*

<400> 135

```

taatacgact cactataggg aaaacgctgt aggtttgtat aagtttctta ggaaaacgat 60

```

cagacaaacg ttccataata taagagccca tgccggaacc agtaccaccg gctatagaat 120
 ggcatagaac aaatccctcc aaggaatcac tgccatctgc ctcacgatca ataatgtcaa 180
 aaatttcctc ttgtaatttt tcaccttgac tatagccgga agcccaattg ttgccggcac 240
 caccaccatg tttagacaag taaacatttt cgggattata taacttggca tagggtgaac 300
 tcataatggg gtgtataact cgcggctcca aatccaaaag tacggcacgt ggtatatagt 360
 gatcatcgtc agcctgataa aagaatacat ccttgcgac tactccatct gtagcaaaat 420
 cctctaacac tccactaggt gaaatgctat acacaccatt atgagttcac cctatgccaa 480
 gttatataat cccgaaaatg tttacttgtc taaacatggg ggtgggtgccg gcaacaattg 540
 ggcttccggc tatagtcaag gtgaaaaatt acaagaggaa atttttgaca ttattgatcg 600
 tgaggcagat ggcatgatt ccttggaggg atttgttcta tgccattcta tagccggtgg 660
 tactggttcc ggcatgggct cttatattat ggaacgtttg tctgatcggt ttctaagaa 720
 acttatacaa acctacagcg ttttcccgta ccacaacct acgttatect tcatatatga 780
 ataataattt gataggattg acggcacctt tgatacctac cccccaatta cattttctaa 840
 tgaccgggta tactcctcta actacagata gtgatcccaa tttgaatata cgcaaaacta 900
 cggtagtaga tggtatgaga cgttttattgc aacccaaaaa tatgatgggt tcatcgggtc 960
 cggataaagc aaatattcat tggtatattt ccatattaaa tattatacag ggtgaagtag 1020
 atcccactca agtccacaaa tctctactga ttggccatca ttaggcccgga aactttatga 1080
 ttactttgta tatatggaga acttctggac aaggctactt gtatactggc cggaccccag 1140
 ggtatgaatt gagctaattt gcgttcacgt atacgttgta gagatttggt gacttgagtg 1200
 ggatctactt caccctgtat aatattttaat atggaaatat aacaatgaat atttgcttta 1260
 tccggaccgc atgaaacat catatttttg ggttgcaata aacgtctcat aacatctagt 1320
 accgtagttt tgcgtatatt caaattggga tcaactatct tagtttagagg agtataaccg 1380
 gtcattagaa aatgtaattg ggggtaggt atcaaagggt cgtcaatcc tatcaaatga 1440
 ttattcatat atgaaggata acgtaggggt gtggtacgta g 1481

<210> 136

<211> 1605

<212> DNA

<213> *Triticum aestivum*

<400> 136

atggc aaagg acgacgggta cccccggcg cggacgtgc cggagacgcc gtctggggcg 60
 gtggcgctgg tcttcgccgt catgatcatc gtctccgtcc tcctggagca cgcgctccac 120
 aagctcggcc attggttcca caagcggcac aagaacgcgc tggcggaggc gctggagaag 180
 atgaaggcgg agctgatgct ggtgggattc atctcgtgc tgctcgccgt cacgcaggac 240
 ccaatctccg ggatatgcat ctcccagaag gccgccagca tcatgcgcc ctgcaagggtg 300
 gaaccgggtt cgtcaagag caagtacaag gactactact gcgccaaga gggcaagggtg 360
 gcgctcatgt ccacgggcag cctgcaccag ctccacatat tcattcttct gctagccgtc 420
 ttccatgtca cctacagcgt catcatcatg gctctaagcc gtctcaagat gagaacatgg 480
 aagaaatggg agacagagac cgcctccttg gaataaccgt tcgcaaatga tcctgcgcgg 540
 ttccgcttca cgcaccagac gtcgttcgtg aagcggcacc tgggcctgtc cagcaccccc 600

ggcgtcagat ggggtgggtggc cttcttcagg cagttcttca ggtcgggtcac caaggtggac 660
 tacctcatct tgagggcagg cttcatcaac gcgcacttgt cgcagaacag caagttcgac 720
 ttccacaagt acatcaagag gtccatggag gacgacttca aagtcgtcgt tggcatcagc 780
 ctccccgtgt gggctgtggc gatcctcacc ctcttccttg atatcgacgg gatcggcaca 840
 ctcacctggg tttctttcat ccctctcatc atcctcttgt gtgttggaac caagctagag 900
 atgatcatca tggagatggc cctggagatc caggaccggt cgagcgtcat caagggggca 960
 cccgtgggtcg agcccagcaa caagtcttc tggttccacc gccccgactg ggtcctcttc 1020
 ttcatacacc tgacgtgtt ccagaacgcg tttcagatgg cacatttcgt gtggacagtg 1080
 gccacgcccg gcttgaagga ctgcttccat atgaacatcg ggctgagcat catgaaggctc 1140
 gtgctggggc tggctctcca gttcctgtgc agctacatca cttccccct ctacgcgcta 1200
 gtcacacaga tgggatcaaa catgaagagg tccatctttg acgagcagac agccaaggcg 1260
 ctgaccaact ggcggaacac ggccaaggag aagaagaagg tccgagacac ggacatgctg 1320
 atggcgcgaga tgatcggcga cgcaacaccc agccgaggca cgtccccgat gcctagccgg 1380
 ggctcatcgc cgggtcacct gttcagaag ggcatgggac ggtctgacga tccccagagc 1440
 gcaccgacct cgccaaggac catggaggag gctagggaca tgtaccggt tgtggtggcg 1500
 catcctgtac acagactaaa tcctgtgac aggcggagggt cggctctcttc atcagccctc 1560
 gatgccgaca tccccagcgc agatttttcc ttcagccagg gatga 1605

<210> 137

<211> 1277

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct, targeting Mlo from *Triticum aestivum*

<400> 137

taatacgact cactataggg tgcccccttg atgacgctcg accggtcctg gatctccagg 60
 gccatctcca tgatgatcat ctctagcttg gttccaacac acaagaggat gatgagaggg 120
 atgaaagaaa cccaggtagg tgtgccgatc ccgtcgatat caaggaagag ggtgaggatc 180
 gccacagccc acagcgggag gctgatgcca acgacgactt tgaagtcgtc ctccatggac 240
 ctcttgatgt acttggtgaa gtcgaacttg ctgttatgcg acaaatgcgc gttgatgaag 300
 cctgccctca aggtgaggta gtccaccttg gtgaccgacc tgaagaactg cctgaagaag 360
 gccaccaccc atctgacgcc ggggggtgctg gagaggggtt cgacttccac aagtacatca 420
 agagggtccat ggaggacgac ttcaaagtcg tcgttggcat cagcctcccg ctgtgggctg 480
 tggcgatcct caccctcttc cttgatatcg acgggatcgg cacactcacc tgggtttctt 540
 tcatccctct catcatcctc ttgtgtgttg gaaccaagct agagatgac atcatggaga 600
 tggccctgga gatccaggac cggtcgagcg tcatcaaggg ggcacccgac gtcgagccca 660
 gcaacaagtt cttctggttc caccgccccg actgggtcct cttcttcata cacctgacgc 720
 tgtttccaga gtcacacaga tgggatcaaa catgaagagg tccatcttcg acgagcagac 780
 agccaaggcg ctgaccaact ggcggaacac ggccaaggag aagaagaagg tccgagacac 840

ggacatgctg atggcgcaga tgatcggcga cgcgacgccc agccgaggca cgtccccac 900
 cacaaccggg tacatgtccc tagcctcctc catggtcctt ggcgaggtcg gtgcgctctg 960
 gggatcgtca gaccgtccca tgcccttctg aagcaggtgc accggcgatg agccccggct 1020
 aggcacgcgg gacgtgcctc ggctgggctg cgcgtcgccg atcatctgcg ccatcagcat 1080
 gtccgtgtct cggaccttct tcttctcctt ggccgtgttc cgccagttgg tcagcgcctt 1140
 ggctgtctgc tcgtcgaaga tggacctctt catgtttgat cccatctgtg tgacttctgg 1200
 aacagcgtca ggtgtatgaa gaagaggacc cagtcggggc ggtggaacca gaagaacttg 1260
 ttgctgggct cgacgtc 1277

<210> 138

<211> 1527

<212> DNA

<213> *Vitis pseudoreticulata*

<400> 138

atggctggcg acgaggagac gacgacgacg gcagcaacac ttgaaacaac gtccacttgg 60
 gctgttgctt ctgtttgctt tattttgatt gcactctcca tacttattga gcatgccctc 120
 catctcttag ccaagtactt caacaagaag cggaggaggt ctctcattca tgctcttaac 180
 aacgtcaaat cggagttagt gctcttgggg ttcgtctctt tgttgctgac tgtgtgcca 240
 aagtatatgg cgaagatttg tatcccaagg agcgtaggtg aaacttttct tccctgcaag 300
 accttgacag aaagtgattc agaagaagaa accaaatgcg aagagcaggg aaagatgtct 360
 ttgctgtcta gacaaggcgt ggaggaacta caatacttaa ttttcgtgct ggccttcttc 420
 cattccctct actgcgtcct cacattcggt cttgggatgg ccaagatgaa gaaatgggag 480
 tcctgggagg cagaaacaag aacactggaa tatcagttta caaatgatcc acggaggttc 540
 aggctcatcc atcagacatc atttggaag caacatctga gatattggag tgagcatcag 600
 atacttcgtt ggccggcttg ttttattcgg cagttctatc catccgtctc caaagtggat 660
 tacttgactc ttagacatgg gttcattatg gccattttg cagaaggaag caactatgac 720
 ttccaaaagt atataaaaag agctttggaa aaagactttg gagtgggtgg gggaggaagt 780
 ttctgggttt ggagtttctc catgcttttt gtgttcttca atgctcaagt attttacaac 840
 tattttatggc taccctttat tccattgggt atgctgttgt tggttggaac aaagctacag 900
 ggcattataa ctaagatgtg cttagatagc catgataaag ctctcgttgt tagaggaact 960
 ttgcttgtca ggcccagtga tcacttcttc tggtttgaa aaccggaatt gctcctacat 1020
 cttatgcact ttatattgtt tcagaactct tttcaactgg cgttctttac atggacttgg 1080
 taaaaatttg gattcagatc atgcttccat gatacaactg aggatatcgt cataaggctt 1140
 gtcacgggtg tgtagtaca actcctttgt ggctacgtga cactgcctct gtatgccctg 1200
 gtcacgcaga tggggacatc aatgaggaca attgtcttta ctgaggagt cgttgaagg 1260
 ctgaacagat ggagaaggaa agccaagaaa aacatagcac gcaggaacaa ccatcagct 1320
 cgtccctccc tggatgcttc actcgacaat tcaccttctt ttaacactct ggatacttct 1380
 ttctctgtag acctcgatca gccatcatca gatgctggtt atttgactgt tgaaatatca 1440
 gatgaagaga cggctgctac taaacggcca gaaccgcgtc agaagttggg atcttttgag 1500
 ggtttcgact cgtgcaaaac atcataa 1527

<210> 139

<211> 1480

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric DNA encoding a ledRNA construct targeting a Vitis ML0 gene

<400> 139

```

taatacgaact cactataggg tagccataaa tagttgtaaa atacttgagc attgaagaac 60
acaaaaagca tggagaaaact ccaaaccag aaacttctc ccaccaccac tccaaagtct 120
ttttccaaag ctctttttat ataacttttg aagtcatagt tgcttcttc tgcaaatgg 180
gccataatga acccatgtct aagagtcaag taatccactt tggagacgga tggatagaac 240
tgctgaataa aacaaaccgg ccaacgaagt atccgatgct cactccaata tctcagatgt 300
tgctttccaa atgatgtctg atggatgagc gtgaacctcc gtggatcatt tgtaaactga 360
tattccagtg ttcttgtttc tgcctcccag gactcccatt tcttcattct ggccatccca 420
agaccgaatg tgaggacgca gtagagcaca tcatttgga agcaacatct gagatattgg 480
agtgagcatc ggatacttcg ttggccggtt tgttttattc agcagttcta tccatccgtc 540
tccaaagtgg attacttgac tcttagacat gggttcatta tggcccattt tgcagaagga 600
agcaactatg acttccaaaa gtatataaaa agagctttgg aaaaagactt tggagtgggtg 660
gtgggaggaa gtttctgggt ttggagtctc tccatgcttt ttgtgttctt caatgctcaa 720
gtattttaca actatttatg gctaccgta attccattgg tgatgctgtt gttggttgga 780
acaaagctac agggcattat aactaagatg tgcctagata gccatgataa agctctcgtt 840
gtagaggaa ctttgcttgt caggccagat gatcattct tctggtttgg aaaaccgga 900
ttgctcctac atcttatgca ctttatattg tttcagaact cttttcaact ggcgttcttt 960
acatggactt ggtacaaatt tggattcaga tcatgcttcc atgatacaac tgaggatatt 1020
gtcataaggc ttgtcatggg tgtgttatgg ctttccttct ccatctgttc agaccttcaa 1080
cgactccctc agtaaagaca attgtcctca ttgatgtccc catctgcgtg accagggcat 1140
acagaggcag tgtcacgtag ccacaaagga gttgtactaa cacacccatg acaagcctta 1200
tgacgatatt ctcagttgta tcatggaagc atgatctgaa tccaaatttg taccaagtcc 1260
atgtaaaaga cgccagttga aaagagttct gaaacaatat aaagtgcata agatgtagga 1320
gcaattccgg ttttccaaac cagaagaagt gatcactggg cctgacaagc aaagttcctc 1380
taacaacgag agctttatca tggttatcta ggcacatctt agttataatg ccctgtagct 1440
ttgttccaac caacaacagc atcaccaatg gaattacgta 1480

```

<210> 140

<211> 1531

<212> DNA

<213> Rhizoctonia solani

<400> 140

```

cattccatcg ctgctacagt tgataccccg cgacgcgtcc ttaccgcctc ttgtgttcca 60

```

ctgggttcca attattggct ccgctatcgc gtacgggtgat gaccctcttg catttctctt 120
 ttcgtgccgc gaaaagtacg gcgacctgtt caccttcggtt cttctcggcc gcaaaatgac 180
 cgtcgcgttg ggcccaaagg gtagtaattt tatcctggga ggaaaacttt cccaagtctc 240
 agccgaggaa gcctacaccc accttaccac tccagtcttt ggcaaggatg ttgtctatga 300
 cgtcccgaac catgtactca tggagcaaaa gaagtttgtc aagttcggac ttaccaccga 360
 gaacttccga gcctacgtcg atatgatcgt agacgagacc gtgaacaacc tcattcgtaa 420
 ggagctctcc cctgaaaact gccacgcga ctcccagggc tgggggtgct tccatgcatt 480
 caaaaagctg gccgagctca cgattctcac cgctcgcgc acgctgcaag gcaacgaagt 540
 tcgctccaac cttgacaaaa gcttcgcaga attgtatcag gacctgatg gcggcttcac 600
 ccctatcaac ttcctcttcc ccaaccttc gctccctagt tactggcgtc gagaccgtgc 660
 tcaaaagaag atgagcgact tttacgtgaa cattattgag aaacgtaagg cacagagtca 720
 aggggatgag catgatatga ttgtgtcttt gttgaatcag acctacaaag atggccgcgc 780
 gctcagcgac cgtgaggtag ctcatatcat gattgcactc ctcatggccg gtcagcacac 840
 tagttctgca acttcttctt ggacgttctt tcacctgcgc gatcgccctg atattgccga 900
 gaaattgtac gaggaacaag tcaaagtctt tggtaatgcc gatggttcga tgcgccccctt 960
 gaactacgag gaggttgaagg atttgccagt actgagcgca gttatccgtg agaccttgcg 1020
 catgcatccg cctatccaca gtatcatgcg caaatgcatt aacgatatgg ttgttccggc 1080
 gtccctggct gccccaccg gcaaggctaa cgagggccga acctacgttg ttcccaaggg 1140
 ccactatctt ctggcatctc ccgccgttgc acaagtcgac cctcgtgtat ggcgtgacgc 1200
 cgacaagtgg gatcctctgc gctggttgga cccaacggga gctgctgctc aggccggttc 1260
 tttgtacaac gacgagcacg gtgaaacggg tgactatggt tgggggtgccg ttagtaaggg 1320
 taccgagagc ccttatcagc catttggtgc tggtaggcat aggtgcattg gtgaacaatt 1380
 cgcaaata caactcggcg ccattttgtc tactataatc cggaacatgg agatgcgtat 1440
 cgaaaagcac gttcctgacc acaattacca tactatgata gtccctgccta aagatccctg 1500
 tggtatcagg ttcaagctcc gcaccaaggc g 1531

<210> 141

<211> 1617

<212> DNA

<213> *Rhizoctonia solani*

<400> 141

atgaaccagt tcgcatcttg cccatggctc gaatctgcca ctttcattcc tttactgggc 60
 gcttcatgcg tgatcttgat ggccacctgt gcgtgcatcc ttttgaatgt catcgcccaa 120
 ttggttatac cccctgatcc gtcattgcca cctcaggttt tctacgttct accgtatatt 180
 ggctcggcca ttgaatacgg taaggatcca ataggtttct tatcatctgg caggagaaag 240
 tatggggacg tttttacctt cgttctcttg ggacgccgag tgaccgtcgc gcttgggtccc 300
 aagggaagta atttcgtcct tgggggaaag ttgtcacagg tctcggccga agaagcgtac 360
 acacacttga ctacacctgt cttcggcaag ggcgtgattt tcgatgttcc aaatcatgtc 420
 tttatggaac agaagaagtt catcaaatcc ggcctcacia ctgaaaatct acgcgcctac 480
 gtgaacatga tatccgagga gactaccaca ttctgaata aagacttggc tgatacctgt 540

cgtggaaagg aatgggggag gtttcatgta cttgacactc tggctgggct tacgatcctg 600
 accgcctcga ggacgctcca aggagagaa gtgcggctgt ctctggataa gaccttttct 660
 caggtttata aggatttga tgggggattc acaccctga accttatgtt cgccaatctc 720
 cctctgcccc gttactggag gagggaccgt gctcaacgga aaatgagcga tttttatgtg 780
 gacattatca ggaatcgcca agaggaacat cgggattctg aacatgacat gatctctgca 840
 ttagcatcga gagagtacaa ggacggttct cctctaggcg accgcgagat cgctcacttg 900
 atgatagcct tactcatggc tgggcagcac accagctctt cgaccggttc ttgggcattg 960
 ctacacttag ccgatcgccc agatgttgta aaacaactac ttgcagagca agaagaagtg 1020
 cttggtaacg aagatggaaa cttacgacct ctaaccttcg agggcctcca aaaactcccc 1080
 gttctcaact cggttattcg cgaaacttta cgtattcatc cgcccattca tagtatcatg 1140
 cgcaaatgca tagacgatat tgttgctccg gctactctcg cctccccaag ttcggactcg 1200
 acttacatcg taccaaaaagg acattttctc ctgcctctc cggtcactc gcaagtcgac 1260
 ccagacgttt ggttcagcgc gagcgaatgg gaccactcac gatggctaga tccaaacgga 1320
 gtggccgctc aagccgagtc actctacctg ggtgaccaag gtgaaaaagt cgactatggg 1380
 tggggtgtgg taagcaaagg gaccgagagc ccataccagc cattcggtgc tggaaggcat 1440
 cgatgtatcg gtgagaagtt cgcttatgta caacttggga cgattctgtc gactgttgtg 1500
 agaacaattg agatgggggt ggactcgggt gttccggcgc acaactatca taccatgatt 1560
 gttcagccga aggagccctg catgattcag ttcaggttcc gggataggca aaggag 1617

<210> 142

<211> 1823

<212> DNA

<213> Artificial Sequence

<220>

<223> Nucleotide sequence of a chimeric DNA encoding a ledRNA construct targeting a gene encoding Cyp51

<400> 142

aagaattcta atacgactca ctatagggtc caaccagcgc agaggatccc acttgtcggc 60
 gtcacgccat acacgagggt cgacttgtgc aacggcgga gatgccagaa gatagtggcc 120
 cttgggaaca acgtagggtc ggccctcggt agccttgccg gtgggggcag ccagggacgc 180
 cggaacaacc atatcgtaa tgcatttgcg catgatactg tggataggcg gatgcatgcg 240
 caaggtctca cggataactg cgctcagtac tggcaaatcc ttcaactcct cgtagttcaa 300
 ggggcgcac gaaccatcgg cattacaaa gactttgact tgttctctgt acaatttctc 360
 ggcaatatca gggcgatcgg cgaggtgaag aagcgtccag gaagaagttg cagaactagt 420
 gtgctgaccg gccatgagga gtgcaatcat gatatgagct acctcacggt cgctgagcgc 480
 gcggccatct ttgtaggctt gattcaacaa agcagcgcgc tgatattgcc gagaaattgt 540
 acgaggaaca agtcaaagtc tttggtaatg ccgatggttc gatgcgcccc ttgaactacg 600
 aggagttgaa ggatttgcca gtactgagcg cagttatccg tgagaccttg cgcatgcac 660
 cgcttatcca cagtatcatg cgcaaatgca ttaacgatat ggttggttccg gcgtccctgg 720
 ctgccccac cggaaggct aacgagggcc gaacctacgt tgttcccaag ggccactatc 780

ttctggcatc tcccgccgtt gcacaagtcg accctcgtgt atggcgtgac gccgacaagt 840
 gggatcctct gcgctggttg gacccgtata ttggctcggc cattgaatac ggtaaggatc 900
 caataggttt cttatcatct ggcaggagaa agtatgggga cgtttttacc ttcgttctct 960
 tgggacgccg agtgaccgtc gcgcttggtc ccaagggaag taatttcgtc cttgggggaa 1020
 agttgtcaca ggtctcggcc gaagaagcgt acacacactt gactacacct gtcttcggca 1080
 agggcgtgat tttcgatgtt ccaaatacatg tctttatgga acagaagaag ttcatacaat 1140
 ccggcctcac aactgaaaat ctacgcgcct acgtgaacat gatatccgag gagactacca 1200
 cattcctgaa taaagctgag aaaaggtctt atccagagac gaccgcactt ctctgccttg 1260
 gagcgtcctc gaggcggtca ggatcgtaag cccagccaga gtgtcaagta catgaaacct 1320
 cccccattcc tttccacgac aggtatcagc caagtcttta ttcaggaatg tggtagtctc 1380
 ctcggatata atgttcacgt aggcgcgtag attttcagtt gtgaggccgg atttgatgaa 1440
 cttcttctgt tccataaaga catgatttgg aacatcgaaa atcacgcctt tgccgaagac 1500
 aggtgtagtc aagtgtgtgt acgtttcttc ggccgagacc tgtgacaact ttcccccaag 1560
 gacgaaatta ctcccttggt gaccaagcgc gacggtcact cggcgtccca agagaacgaa 1620
 ggtaaaaacg tccccatact ttctcctgcc agatgataag aaacctattg gatccttacc 1680
 gtattcaatg gccgagccaa tatacgtagg taccgggta gcatatccat gatattctgtt 1740
 agtttttttc ctgaaagagc ggccgccta gcataacccc gcggggcctc ttcgggggtc 1800
 tcgcgggggtt ttttgctgaa aga 1823

<210> 143

<211> 3554

<212> DNA

<213> *Phytophthora cinnamomi*

<400> 143

atggggctca ccggcgcggg catcatcgcc tccgtcgtgg gcatcctggg cggcgtgtcg 60
 ctgtcctgcg gcggctggtc gtcgctgtcc ctcggcgctc gctcgtcttt cgtgacgacg 120
 cagttcctct cgcccttcgc catgtacgtg ccgcactgac catcactcct gctccaattc 180
 tgaaacgagg gcgcactatt ggggtgcgtt tcggttctaa ttttggaacg ttccagacac 240
 taatttgatt ttccgctgtt gtgatattcc acctcagggg attcgtggtc gccttcaccg 300
 ccatctcgtc gctgacaagc accaacgagt ggatcgccgt ggccggccggc ggccggcgcg 360
 gcttcgtgat cgccctcatc gtgggcttca tgacgggtctt cggcccgtac atcctgatcc 420
 tcgtcacagg cggcatcatc gcctgctacc tgctgctcgt ggacgcgtac gacggcgtga 480
 aactcttccc gtcagacaay cagctggcgc gccaggagtt cgtcatcgcc ttcattgatca 540
 tcttcgagct cgtgtgctgc tcgtcgtcca agacgtcgga gctggagaac caccgcttca 600
 agtacatcat cttctcggcc atcacgggcg gctggatggc ctcggacggc gtgtcccgc 660
 tcatcgactc gacggctgtt ctgtcggacg tggcctacac ctccatccag gacggcggt 720
 cggccgcgct caagggcacg gacagcagcg cgcagacgt catgttctc atctggggcg 780
 ccgtcttcgt ggtcggcggc ctcaaccagc tctccatgcg ctggggactc atgtgctaca 840
 accgcgttgg tgcgcacgcc cagctcggcc ccgtcagga gcagatgccg gagctgccc 900
 cgggcgccac gctgcctgcc cagaccatga acgagcgcgt gcgcctcgtg tgcgagaact 960

gtttcgccac ggtgcccagc ggcacggcct tctgtaccga gtgtggtgag gcaatgcctt 1020
 cggacgacgc cgacccgaac gtcagcatct cgcaggcgca gatgccgtcg gtcacgatga 1080
 acaacaagtc tcaagtgtcc gaccgctggc agcagggtgc gcaccgcacg tacctcagca 1140
 cgacgtcggtt tgtcgacccc aagcacgcca aggagggcgg cgtgagcatg aaggacaaca 1200
 gccgcagcat ccgcttcatg gactcgggtg tgcagggccc cgacggcaag atgagccagt 1260
 acaacgactc gatcgccggc gtgcgcaact attacgagcc ttcattccgg tcgttcgcca 1320
 tgtcgacctc ctcgatcgcc aaccgcgccg ctgagcccgt tgacacgccc aacatccgca 1380
 agtacaagat gtctggtagt ggcatgttcc acgtcttcta cttcggtacg gctgccaccg 1440
 gtatcttctg gctgtactac ttgactacga tgtacccgca gcagtacttc tgcgaccacg 1500
 cccgccccac gcttccctgc agtgagctgc ccagcagcga gatttcgggc tgctacagtt 1560
 caacgggtcaa cttcgactcc tcttcggcg agggttactg catcaagaac gtgccgttca 1620
 tgtcgtggct catgtacgcg atgatgatct ttagcgagtt cctcaactac ttcttaggtc 1680
 tgctgttcaa cttcagtatg tggcgctcga ttcgtcgtgg cggccgttac atgaacgact 1740
 tcaaaccgcc tatcccgaaa gagcagtggc cgaccgtcga catcttcttg tgtcactaca 1800
 tggaacctgt gacggactcc atggctacgc tgaagaactg tcttgcatg cagtaccctc 1860
 cggagctgct gcacattttc atccttgatg atggttacgc caagtctgtg tgggacgcca 1920
 acaaccactt caaggttacg gtcaacacca aggtgattga gatttgtggt gacctgcgtg 1980
 gcgacgtcgc tcgcatcatg cacgagcgcg tggtcggccc tgtgcaggac gatcagtcct 2040
 tgaagacgtg gcgtcgccag cacagctctg tgcgtgagct ccgcaaagag ggaagcaagg 2100
 gcgtgcagcg tcgcgactgt gctgttggtt cactgtcggc cgactacgac taccgtgacc 2160
 gcggtatccc gcgtgtgact ttcatcggtc gcatgaagcc cgaaacgcac cactccaagg 2220
 ctggtaacat caacaacgcc ctcttcaacg aaggtgccga tggcaagtac ttgctgattc 2280
 tggataatga tatgaagccg caccggaagt tcttgcttgc cgtgctgccg ttctttcttct 2340
 cggagggcga ggctgtggac ggcggaggcc gccagtacag tgacgacatt tcctggaacc 2400
 aggtgtcgta cgtgcagact cctcagtact tcgaggacac gcccagctg accatcatgg 2460
 gagacccgtg tggacacaag aacaccattt tcttcgacgc tgtacagtgt ggtcgtgatg 2520
 gtttcgactc ggcagctttc gccggtacca acgccgtttt ccgtcgccag gctttcgact 2580
 ccatcggtgg cattcagtac ggtaccaga cggaagatgc ctacacgggt aacgtgctgc 2640
 acatttctgg ctgggactcg gtgtacttcc gcaaggattt cgagggtgat gccaggacc 2700
 gcattcgtct gtgcgaaggt gccgtgcccg aaacggtcgc tgctgccatg ggtcagaaga 2760
 agcgttgggc caaggggtgc gtgcagattc tgctcatgaa gaatgagagc gaggtcgacc 2820
 cggactggcg tccgccgcgt gtgctgccc cggaccgaa gccggcgctt gcgttccgc 2880
 gcaagatgtt cttctacgac tcggtgctct acccgttcgg ttccattccc gctctgtgtt 2940
 acgtggcgat cgctatttac tacctgtgta cgggtgacgc tcccatctac gctcgtggta 3000
 ccaagttcct gtactctttc ttgcccgtga cgttctgccg ttgggtactc aacctgctgg 3060
 ccaaccgcgc cgtcgacaac aacgatgtgt ggcgtgccc gcagacctgg ttctcttct 3120
 ccttcatcac gatgatggct attgtggagg ctattcaggc gcgtgtgacg ggcaaagaca 3180
 agtcgtgggc caacacgggt gccggtcaga agacgtcgtg gacagagatc cccaacgtgc 3240
 tcttcttctt cacgctgctc tttagtcaac tgggtggcgt gattcggttc tttgagtacg 3300

agaacgccac gaacccgtgg aactacgtgt ctgctatgtt cttcggcttc ttcgtgatga 3360
 gtcagttcta ccccatggtc aagatgagta tcacggagta ctgtggttgg gaccacacgg 3420
 ccgcgacctt tacggccaac gtgttcggct cgctgctggg ggtgtacatc gtggtgttcg 3480
 tgcagctgtg gcaggtctac tacgagggca acctgcagac ggcccagggt acatcaggtt 3540
 ccacttcgtc ttag 3554

<210> 144

<211> 1409

<212> DNA

<213> Artificial Sequence

<220>

<223> Nucleotide sequence of a chimeric DNA encoding a ledRNA construct targeting a gene encoding CesA3

<400> 144

aagaattcta atacgactca ctatagggtc ggcgtcgtcc gagggcattg cctcaccaca 60
 ctcggtacag aaggccgtgc cgctgggcac cgtggcgaaa cagttctcgc acacgaggcg 120
 cacgcgtcgc ttcattggtct gggcaggcag cgtggcgccc gtgggcagct ccggcatctg 180
 ctccctgacg gggccgagct gggcgtgcgc accaacgcgg ttgtagcaca tgagtcccca 240
 gcgcatggag agctggttga ggccgcccac cacgaagacg gcgccccaga tgaggaacat 300
 gagcgtctgc gcgtgctgt cgatgccctt gagcgcggcc gagccgccgt cctggatgga 360
 ggtgtaggcc acgtccgaca gaacagccgt cgagtcgatg aggcgtcgtg gtcggcggcc 420
 tcaaccagct ctccatgcgc tggggactca tgtgtacaa cgcggttggt gcgcacgccc 480
 agctcggccc cgtcaggag cagatgccgg agctgcccac gggcgccacg ctgcctgccc 540
 agaccatgaa cgagcgcgtg gcctcgtgt gcgagaactg tttcgccacg gtgcccagcg 600
 gcacggcctt ctgtaccgag tgtggtgagg caatgcctc ggacgacgcc gaccggtacg 660
 tcagcatctc gcaggcgcag atgccgtcgg tcacgatgaa caacaagtct caagtgcccg 720
 accgctggca gcagggtccg caccgcacgt acctcagcac gacgtcgttt gtcgacccca 780
 agcacgcca ggagggcggc gtgagcatga aggacaacag ccgcagcatc cgcttcatgg 840
 actcgggtgt gcaggggccc gacggcaaga tgagccagta caacgactcg atcgccggcg 900
 tgcgcagacg tggaacatgc cactaccaga catcttgtag ttgcggatgt tggcggtgtc 960
 aacgggctca gcggcgcggt tggcgatcga gtaggtcgac atggcgaacg accggaatga 1020
 aggctcgtaa tagttgcgca cgccggcgat cgagtcgttg tactggctca tcttgccgtc 1080
 ggggccctgc acacccgagt ccatgaagcg gatgctgcgg ctgttgctct tcatgctcac 1140
 gccgccctcc ttggcgtgct tggggtcgac aaacgacgtc gtgctgaggt acgtgcggtg 1200
 cggcacctgc tgccagcggc cgggcaattg agacttggtt ttcacgtgta ccgacggcat 1260
 ctgcgcctgc gagatgctga cgtacgtagg taccgggta gcatatccat gatattctgtt 1320
 agtttttttc ctgaaagagc ggccgccta gcataacccc gcggggcctc ttcgggggtc 1380
 tcgcgggggt ttttgctgaa agaagctta 1409

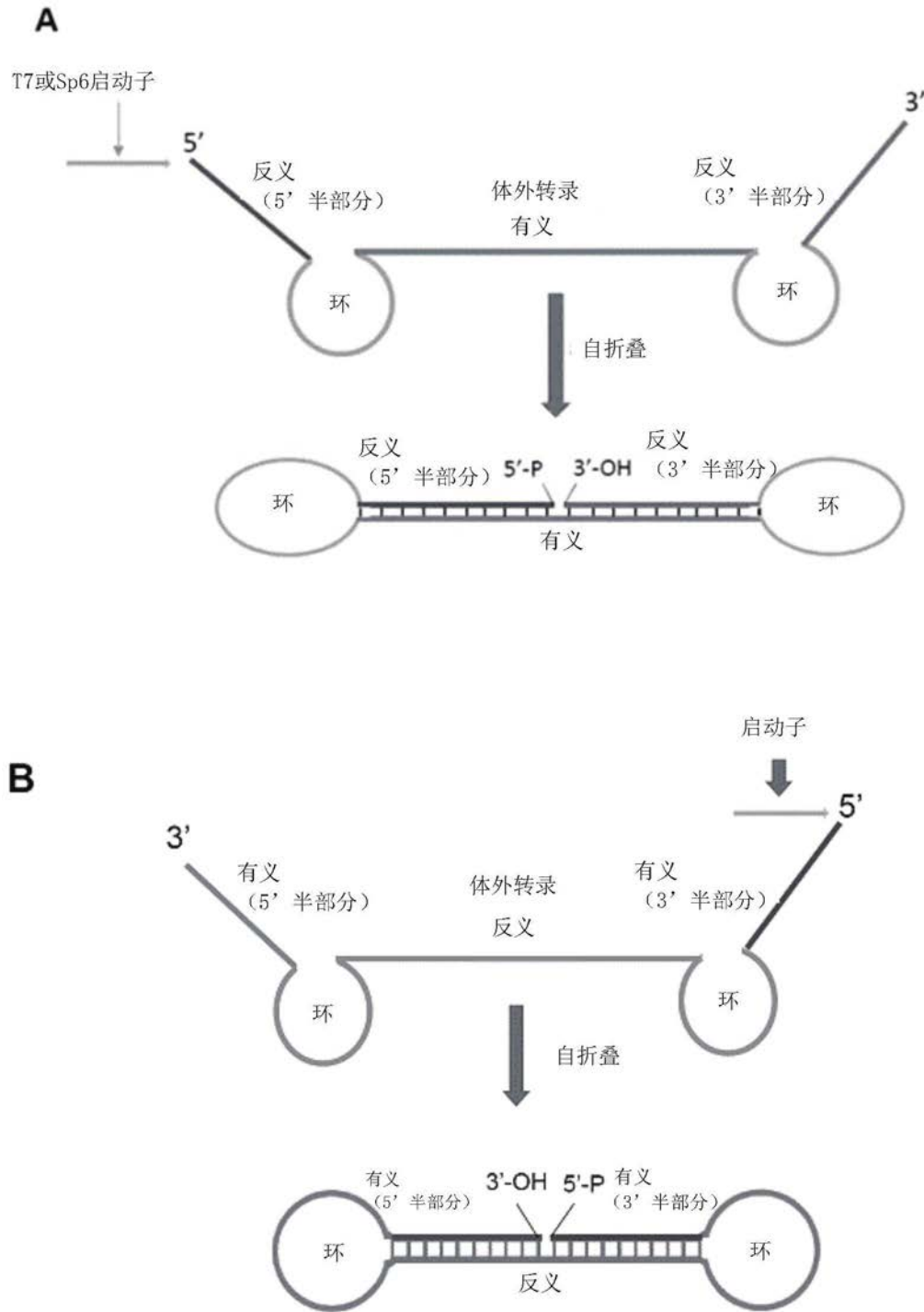


图1

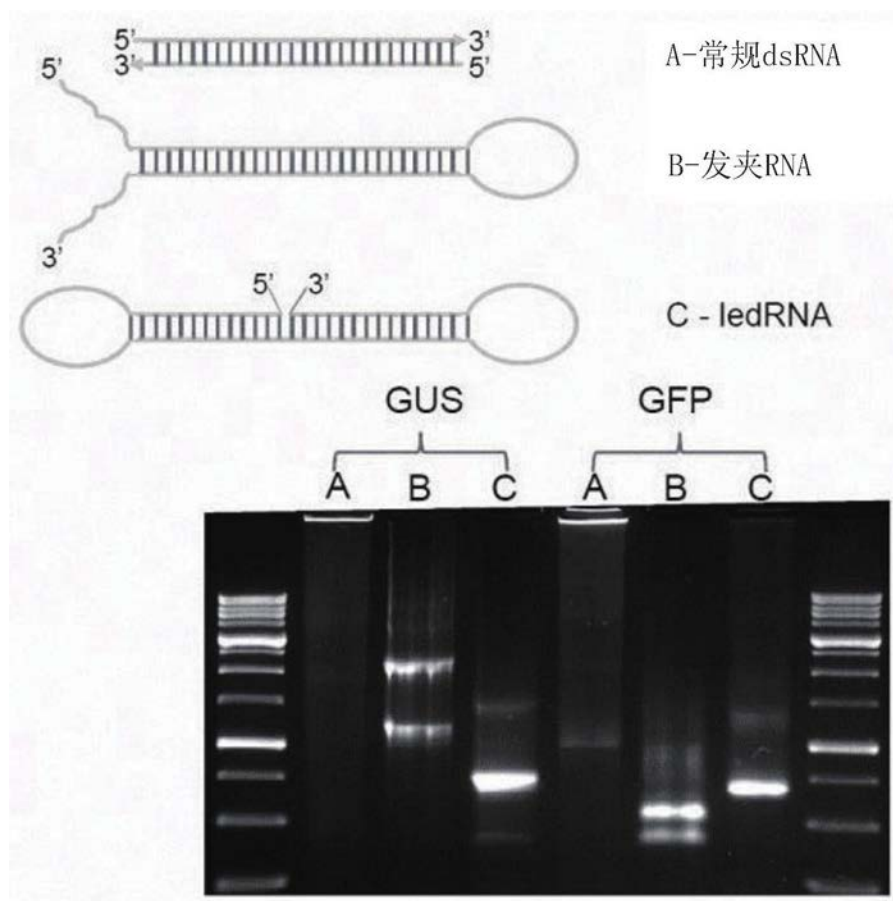


图2

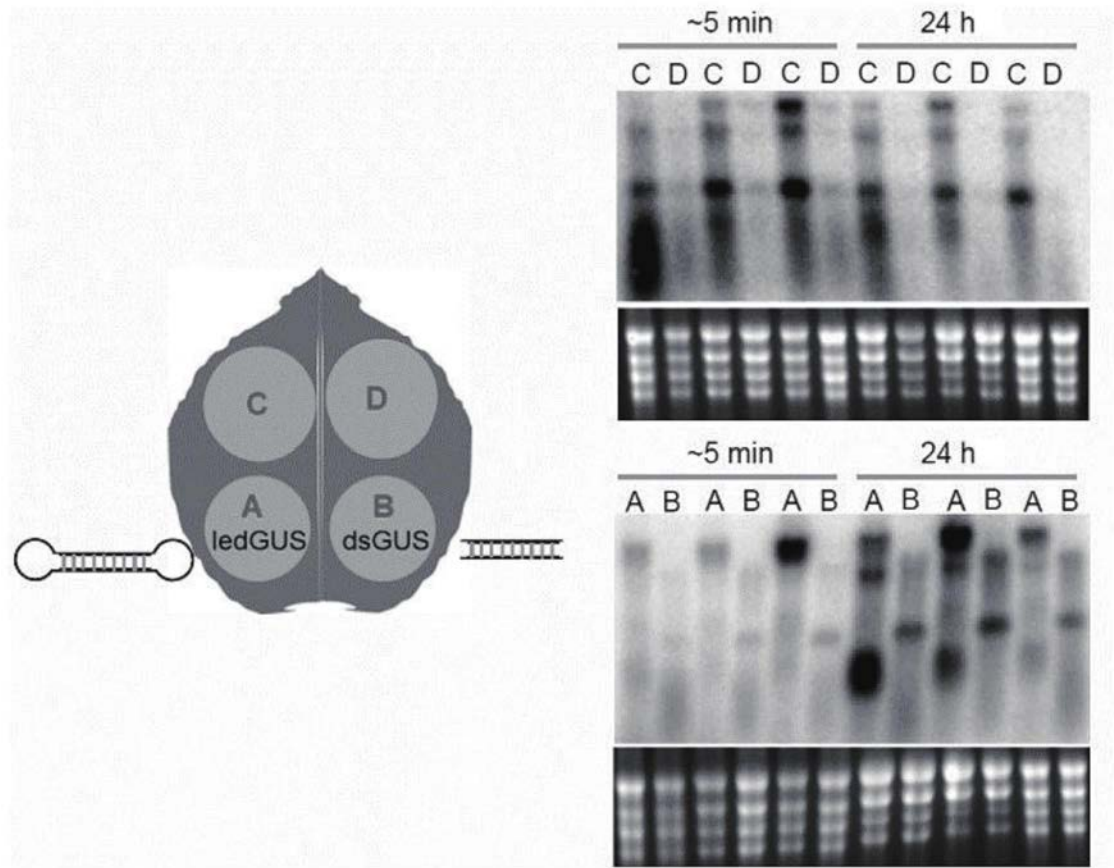


图3

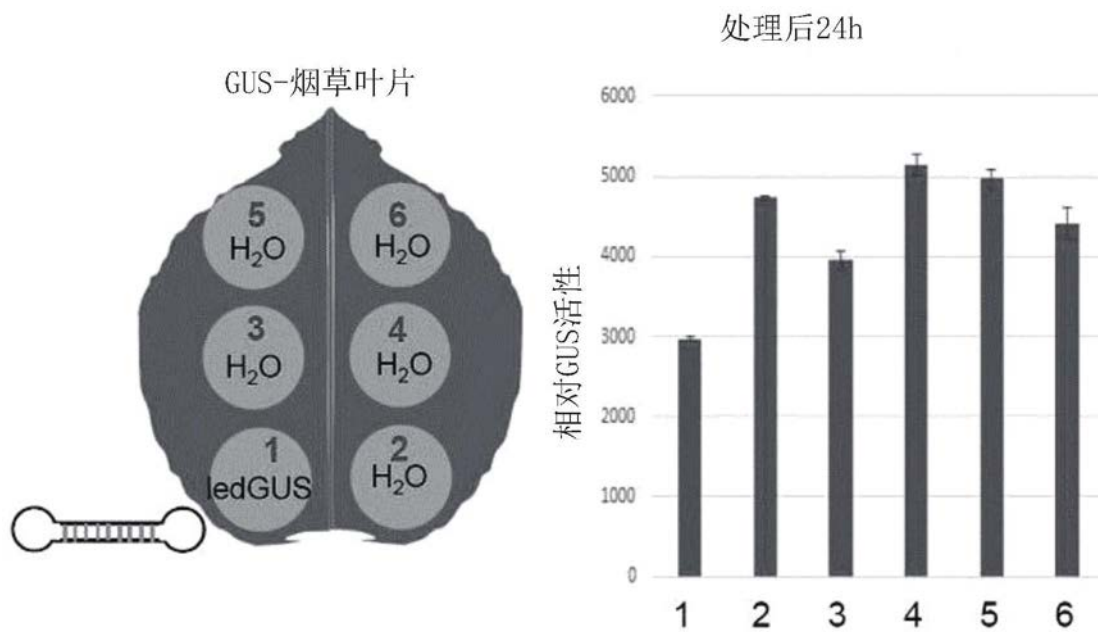


图4

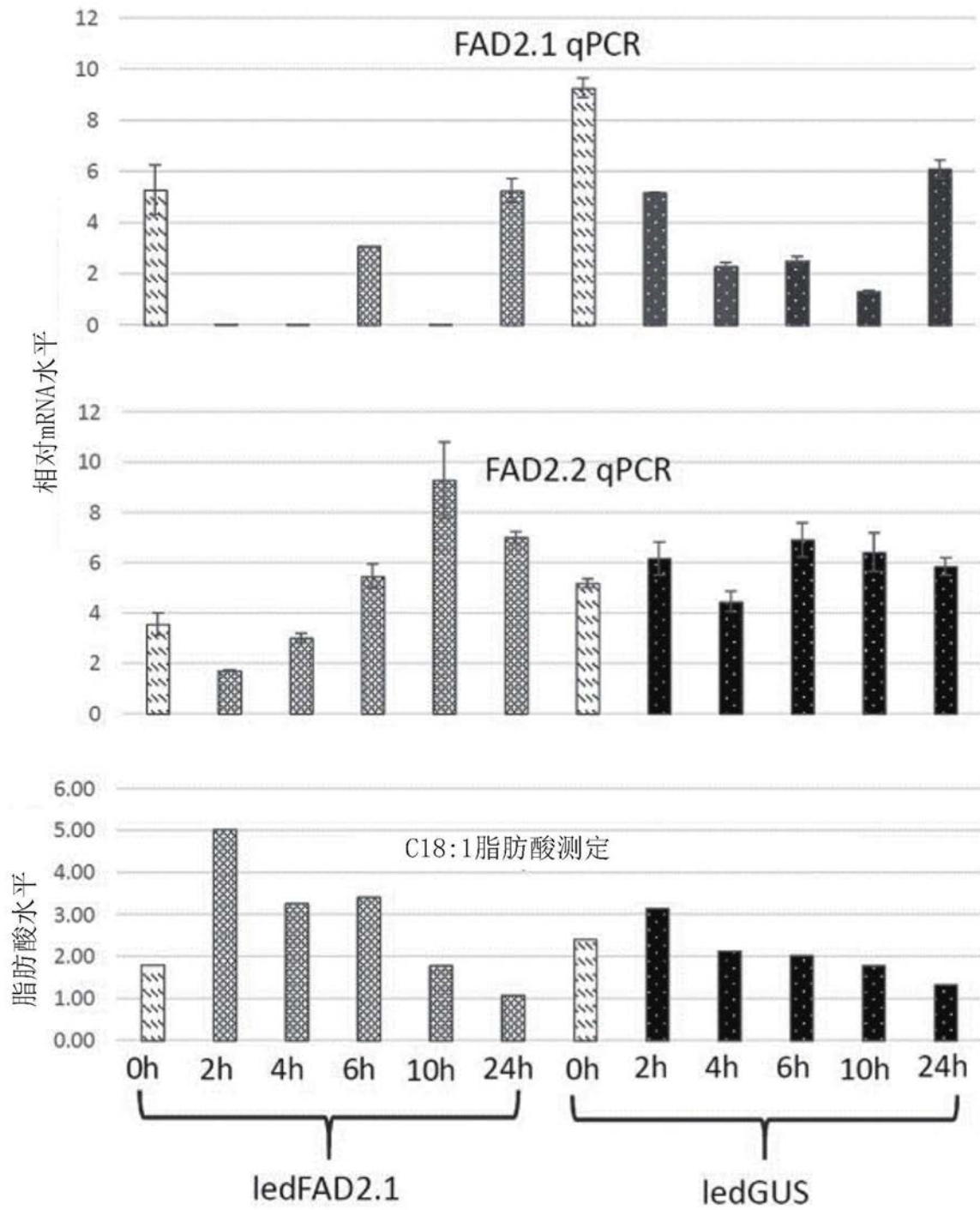


图5

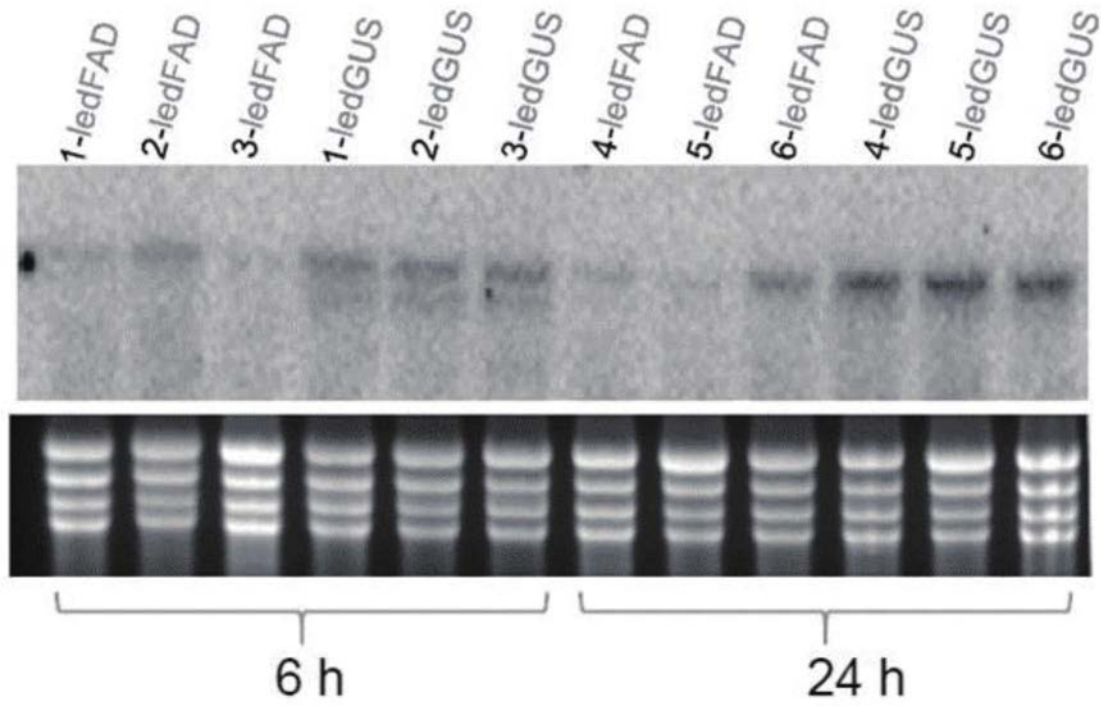


图6

与WT GUS ORF比对

GUS	GAGTGTGATATCTACCCGCTTCGCGTCGGCATCCGGTCAGTGGCAGTGAAGGGCGAACAG	840
GUS-GU	-----TTGTGTGGTATTTGGTTAGTGGTAGTGAAGGGTGAATAG	40
	* * * * *	
GUS	TTCCTGATTAACCACAAACCGTTCTACTTTACTGGCTTTGGTCGTCATGAAGATGCGGAC	900
GUS-GU	TTTTTGATTAATTATAAATTGTTTTATTTTATTGGTTTGGTTGTATGAAGATGTGGAT	100
	** * * * *	
GUS	TTGCGTGGCAAAGGATTCGATAACGTGCTGATGGTGCACGACCACGCATTAATGGACTGG	960
GUS-GU	TTGTGTGGTAAAGGATTTGATAATGTGTGATGGTGTATGATTATGTATTAATGGATTGG	160
	** * * * *	
GUS	ATTGGGGCCAACTCCTACCGTACCTCGCATTACCCTTACGCTGAAGAGATGCTCGACTGG	1020
GUS-GU	ATTGGGGTTAATTTTATTGTATTTGTATTTTATG-----	200
	* * * * *	

图7

```

GUS      GAGTGTGATATCTACCCGCTTCGCGTCGGCATCCGGTCAGTGGCAGTGAAGGGCGAACAG 840
GUS-4M   -----TCGGGTCCGCAACCGCTCACTGGGAGTCAAGCGCGTACAC 40
          *** ***:***:*** ***:*** ***:***:***

GUS      TTCCTGATTAACCACAAACCGTTCTACTTTACTGGCTTTGGTCGTCATGAAGATGCGGAC 900
GUS-4M   TTCGTGAATAAGCACTAACGGTTGTACATTAGTGGGTTTCGTCCTCAAGAACATGGGGAG 100
          *** ***:*** ***:*** ***:*** ***:***:*** ***:*** ***:***

GUS      TTGCGTGGCAAAGGATTTCGATAACGTGCTGATGGTGACGACCACGCATTAATGGACTGG 960
GUS-4M   TTGGGTGCCAATGGAATCGTTAAGGTGGTGAAGGTCCACCACCTCGCTTTATTGGTCTGC 160
          *** ***:***:***:***:*** ***:*** ***:***:***:***:***:***

GUS      ATTGGGGCCAACTCCTACCGTACCTCGCATTACCCTTACGCTGAAGAGATGCTCGACTGG 1020
GUS-4M   ATTCTGGGGCAAGTCCAACCCTACGTCGGATTTCCTATACC----- 200
          *** ***:*** ***:*** ***:*** ***:***:***

```

图8

与WT GUS ORF比对

```

GUS      GAGTGTGATATCTACCCGCTTCGCGTCGGCATCCGGTCAGTGGCAGTGAAGGGCGAACAG 840
GUS-10M  -----TCGCGTCGCGATCCGGTCTCTGGCAGTGTGGGCGAACTC 40
          *****:*****:*****:*****:

GUS      TTCCTGATTAACCACAAACCGTTCTACTTTACTGGCTTTGGTCGTCATGAAGATGCGGAC 900
GUS-10M  TTCCTGATATACCACAAAGGGTTCTACTAACTGGCTTACGTCGTCATCTAGATGCGGTG 100
          *****:***** *****:*****:*****:*****:

GUS      TTGCGTGGCAAAGGATTTCGATAACGTGCTGATGGTGACGACCACGCATTAATGGACTGG 960
GUS-10M  TTGCGTGGGTAAAGATTCTCTTAACGTGCACATGGTGACGACCACGCAAAATGGACTCC 160
          *****:*****:*****:***** *****:*****

GUS      ATTGGGGCCAACTCCTACCGTACCTCGCATTACCCTTACGCTGAAGAGATGCTCGACTGG 1020
GUS-10M  ATTGGGGCGTACTCCTACGCTACCTCGCTATACCCTTAGC----- 200
          *****:***** *****:*****

```

图9

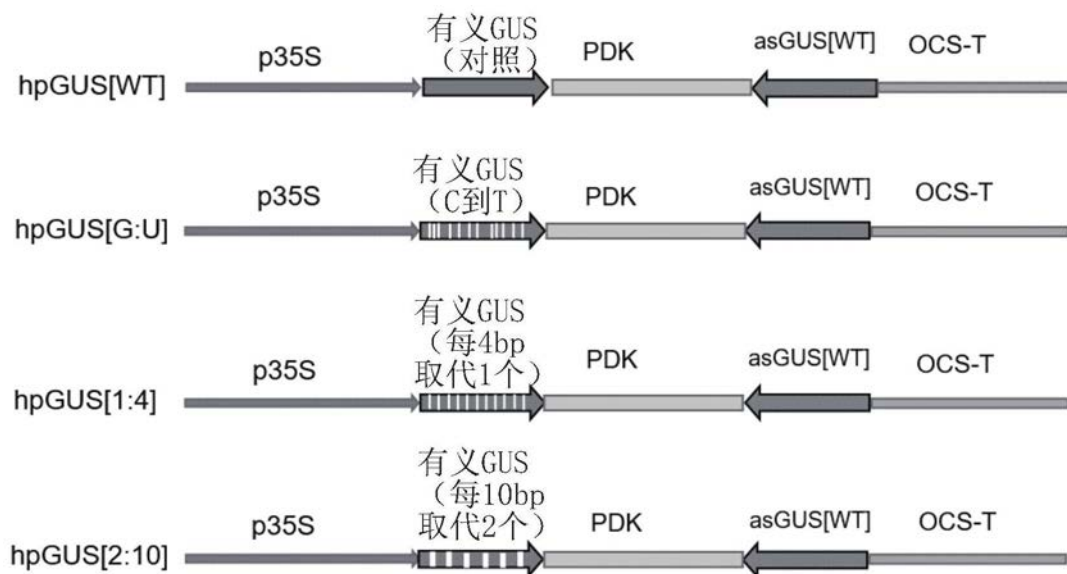


图10

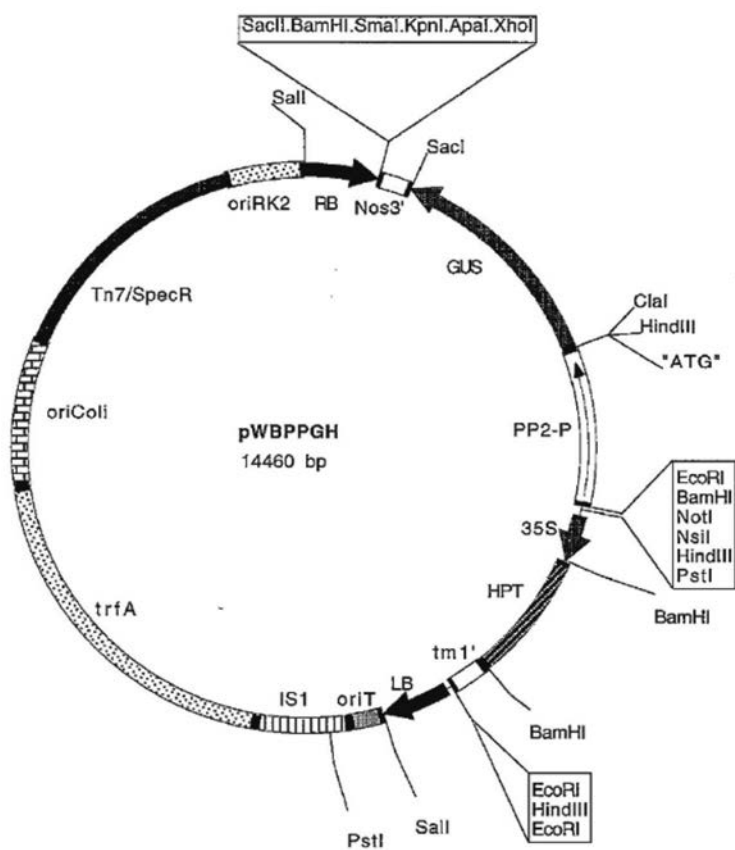


图11

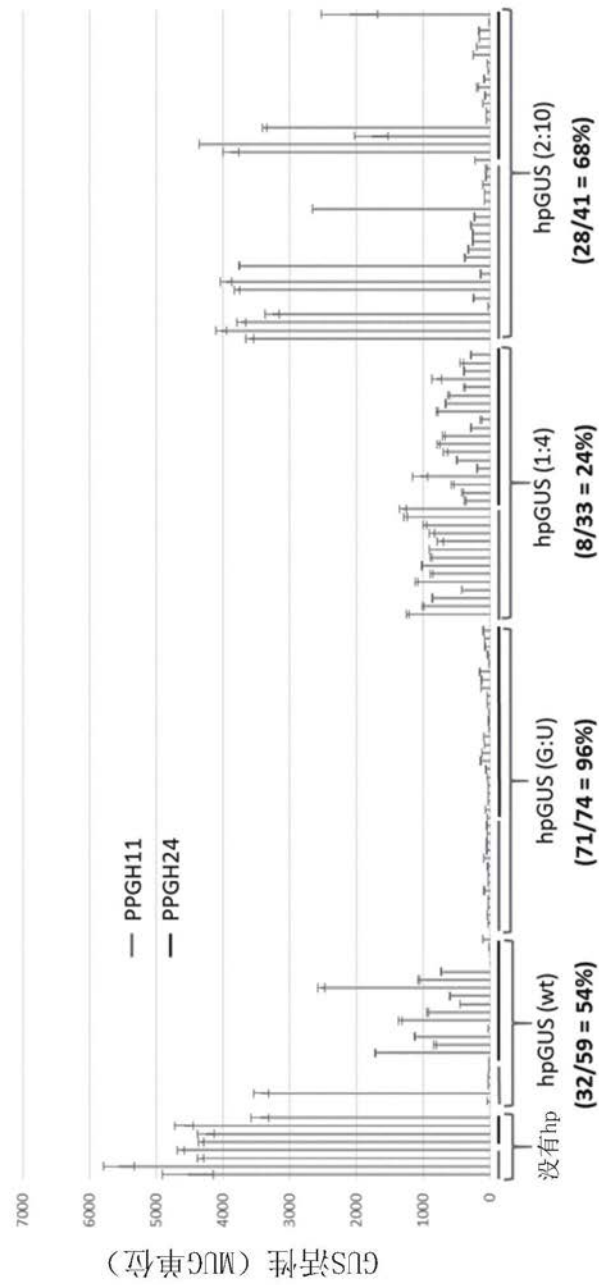


图12

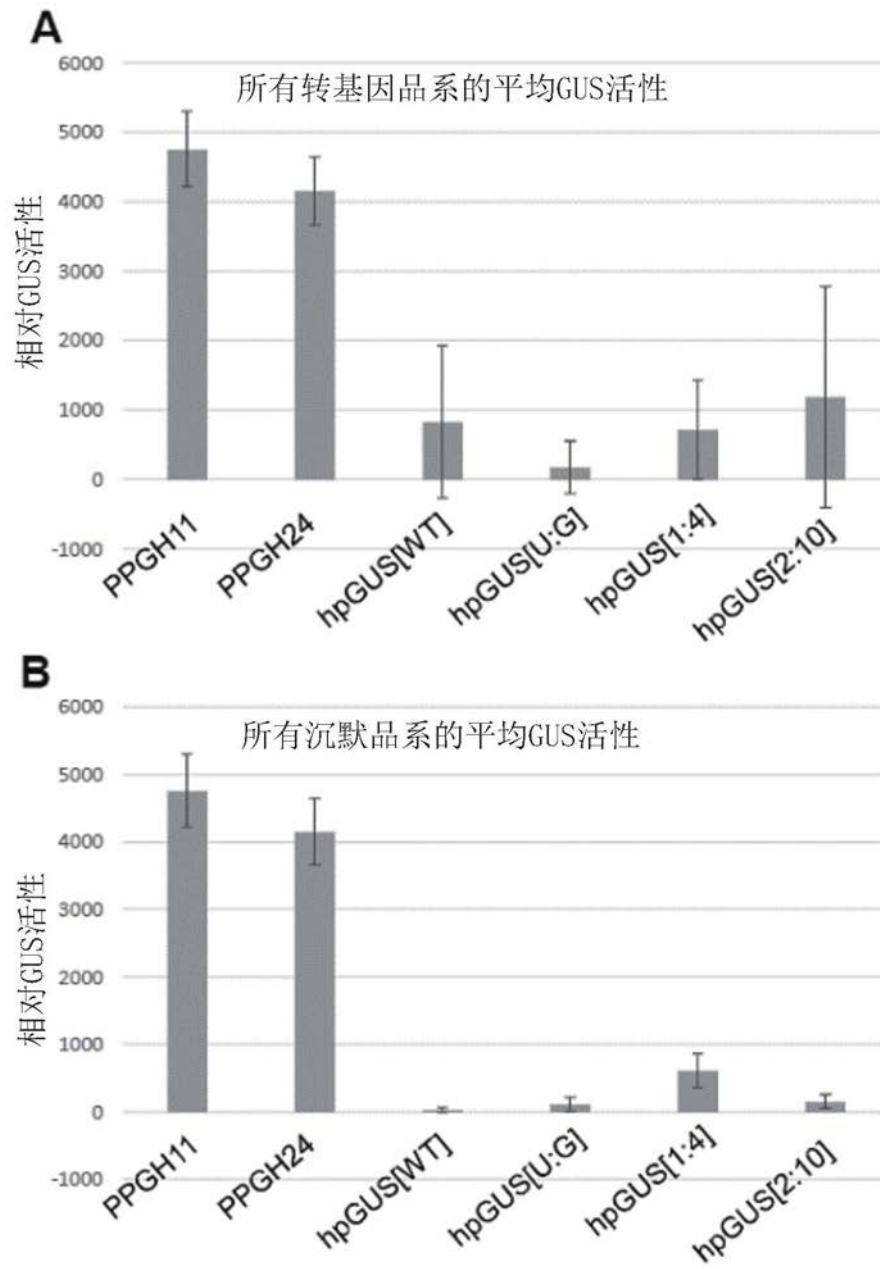


图13

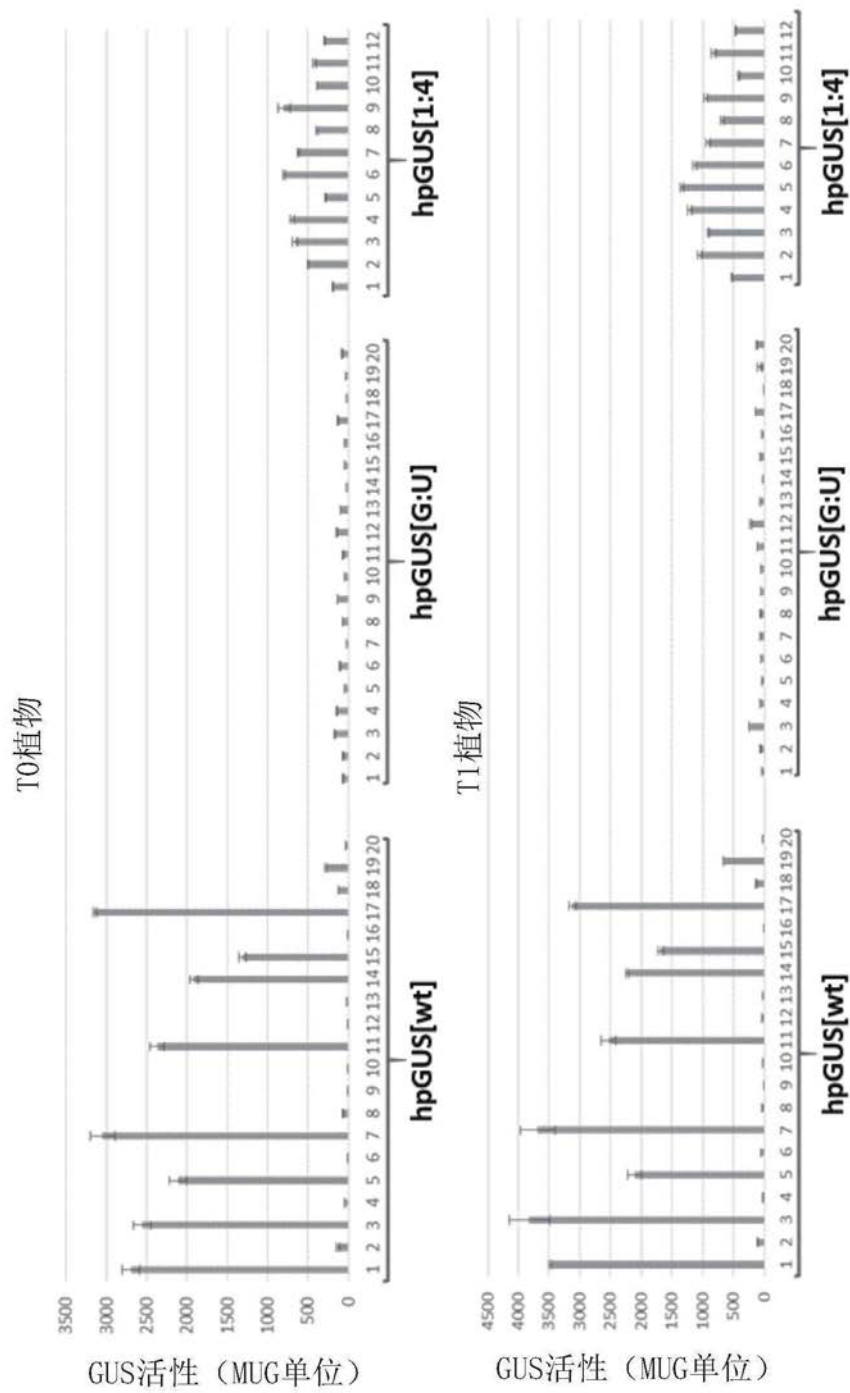


图14

泳道 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

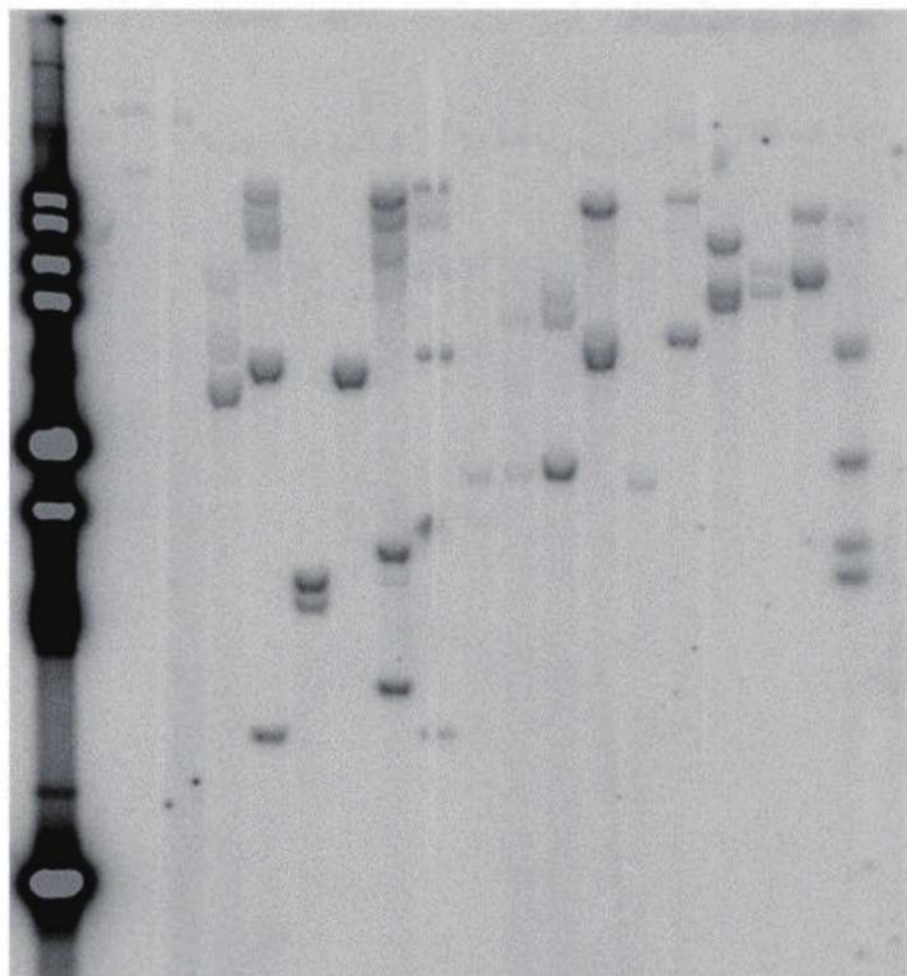


图15

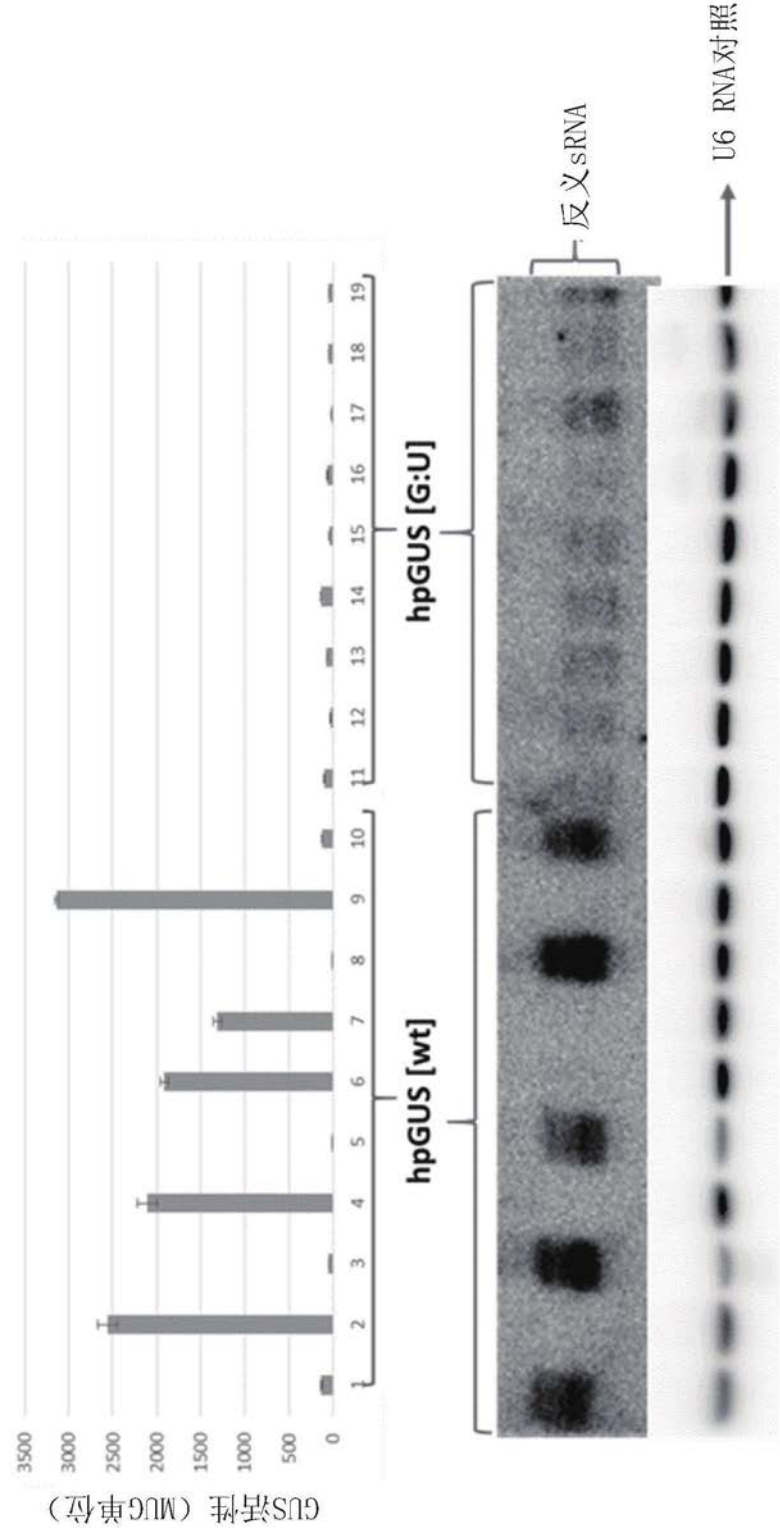


图17

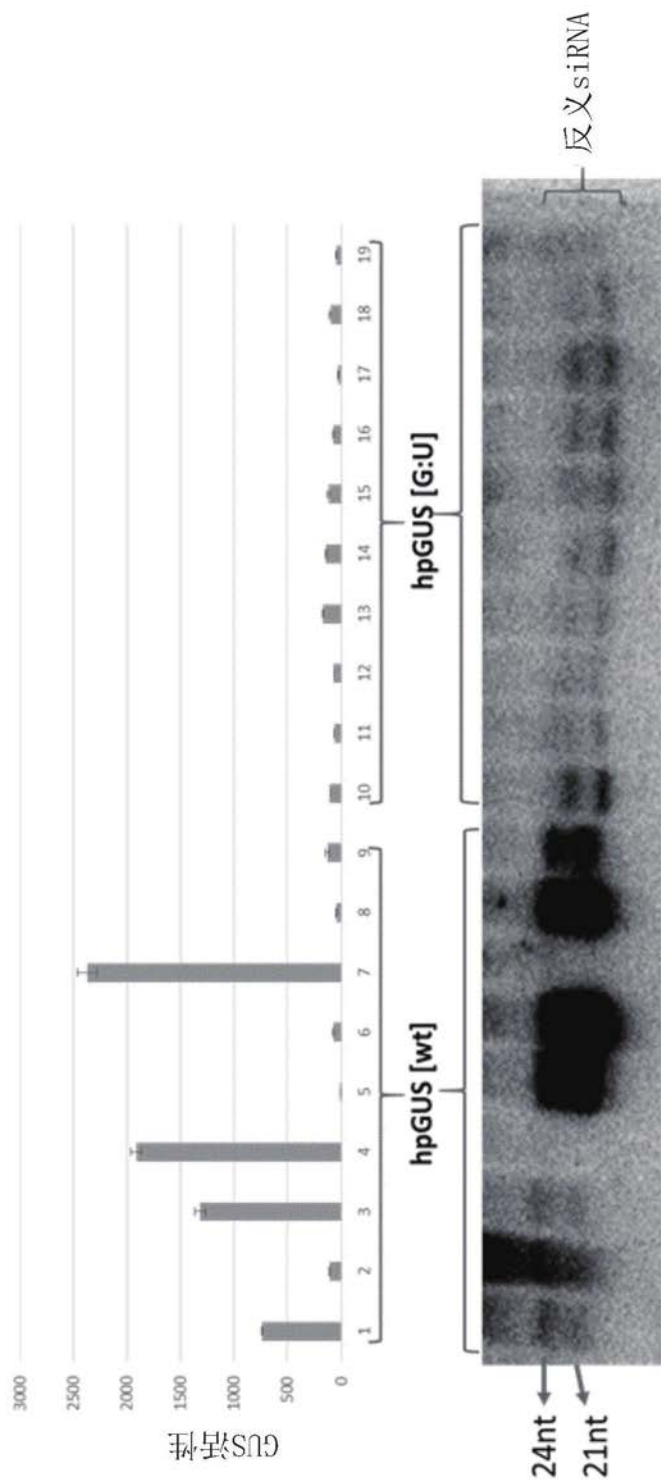


图18

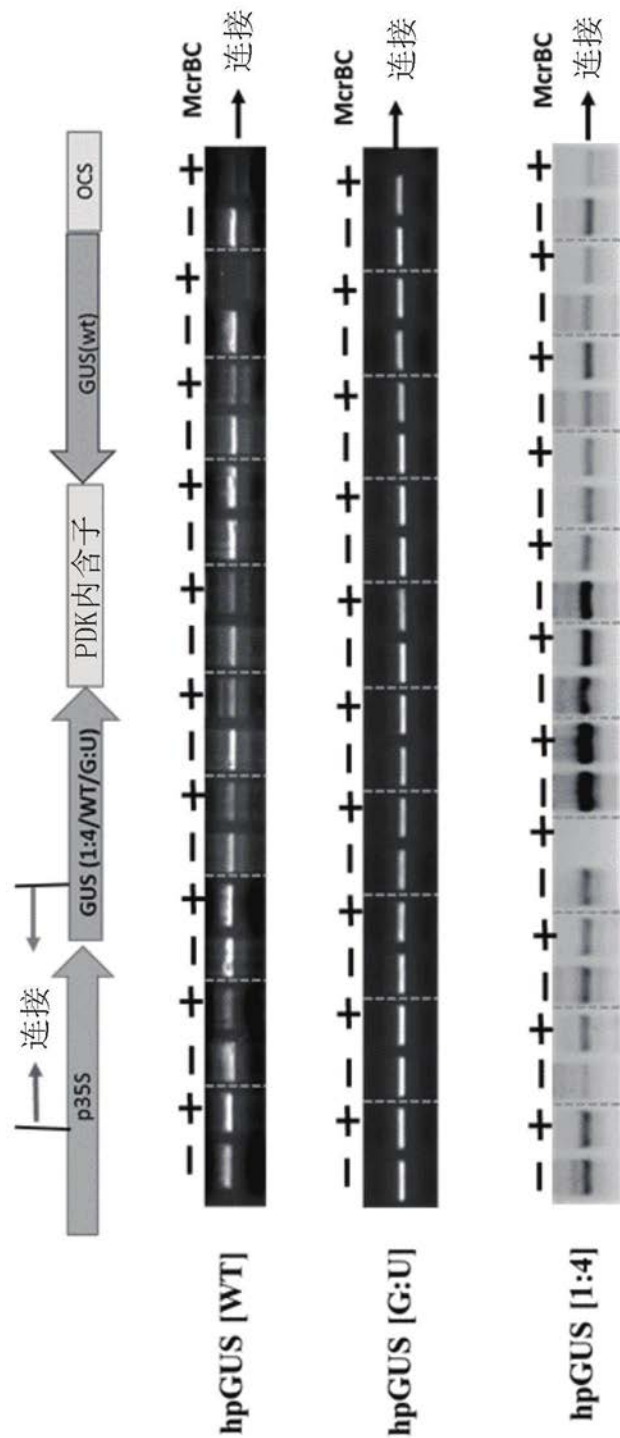


图19

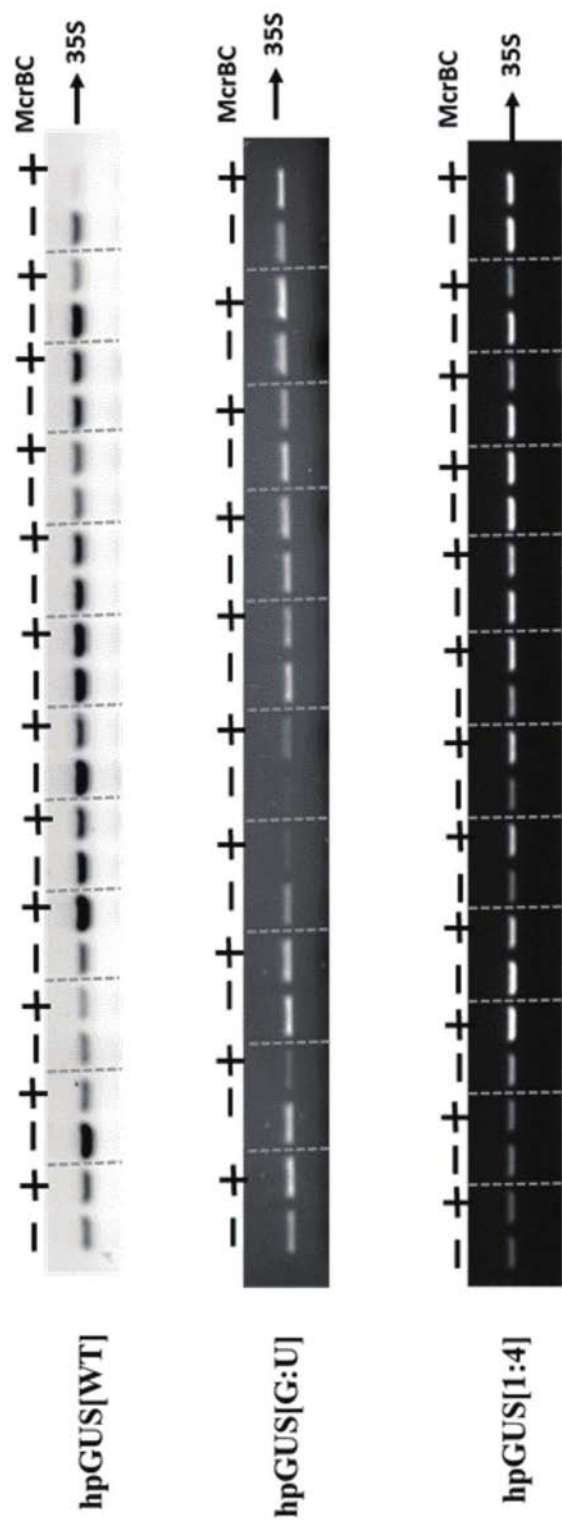


图20

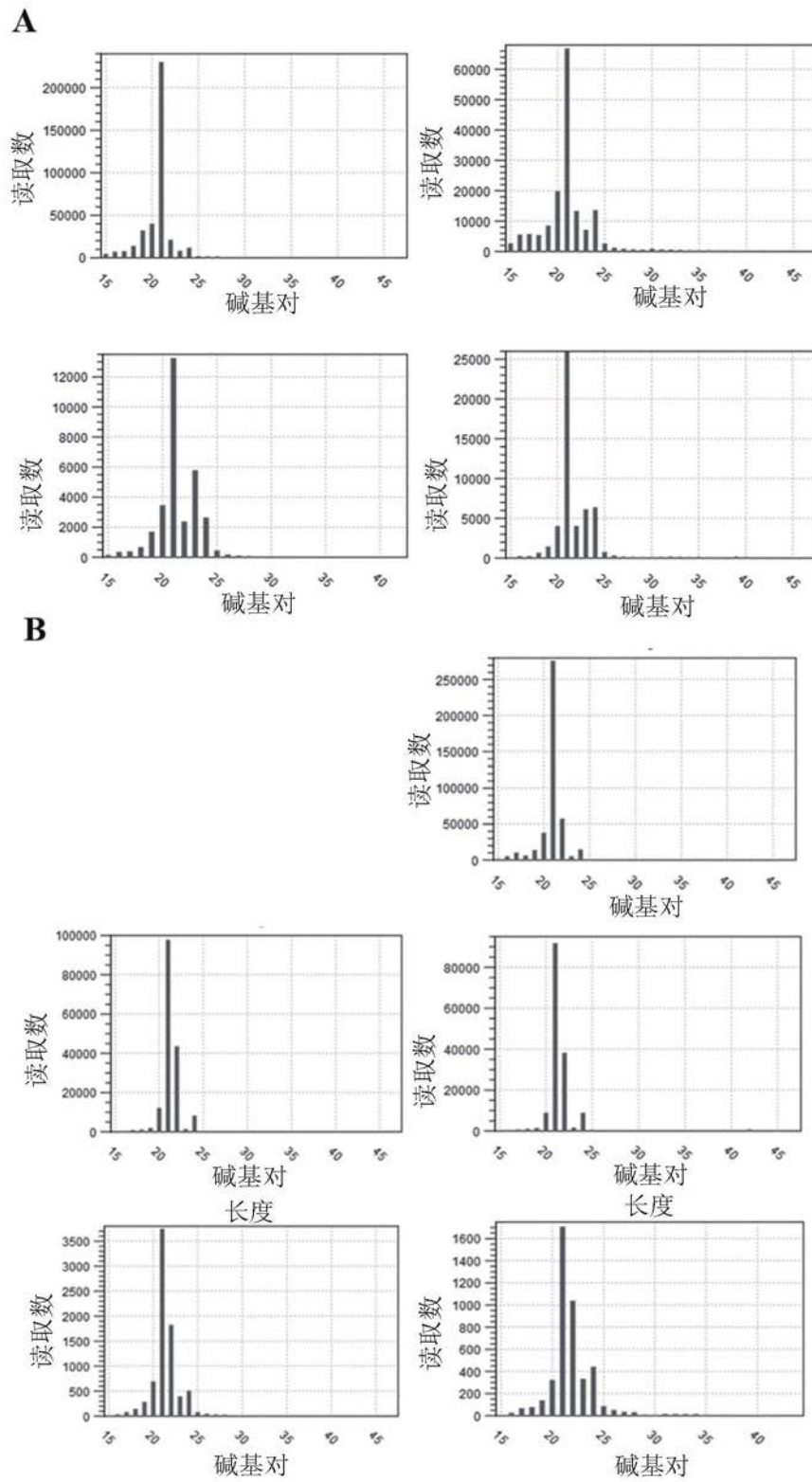


图21

```

EIN2-GU      -----GATTTTA
EIN2-cDNA    AGTAATTATATTATCAGATAGATTTAGGATGGAAGCTGAAATTGTGAATGTGAGACCTCA
                                     ** * *

EIN2-GU      GTTAGGGTTTATTTAGAGAATGGTTTTTGTGTTTTATTTTTTGTGTTTTTGGTTTTTGTGTTG
EIN2-cDNA    GCTAGGGTTTATCCAGAGAATGGTTCCTGCTCTACTTCCTGTCTCTTTGGTTTCTGTCTCGG
* ***** * * * * * * * * * * * * * * * * * * * * * * * * * * * *

EIN2-GU      ATATATTGATTTTGGGAAATGGGTTGTAAATATTGAAGGAGGTGTTGTTTTGGGTATGA
EIN2-cDNA    ATATATTGATCCCGGAAATGGGTTGCAAATATCGAAGGAGGTGCTCGTTTCGGGTATGA
***** * * * * * * * * * * * * * * * * * * * * * * * * * * * *

EIN2-GU      TTTGGTGGTAATTATTTTGTGTTTTTAATTTTGTGTTATTTTATGTTAATATGTTGTAGT
EIN2-cDNA    CTTGGTGGCAATTACTCTGCTTTTCAATTTTGCCGCCATCTTATGCCAATATGTTGCAGC
* * * * * * * * * * * * * * * * * * * * * * * * * * * * *

EIN2-GU      TTGTATAAGTGTT-----
EIN2-cDNA    TCGCATAAGCGTTGTGACTGGTAAACACTTGGCTCAGATCTGCAATGAAGAATATGACAA
* * * * * * * *

```

图22

```

CHS-GU      -----GTTGTTTGTT
CHS-cDNA    TCCGTATCGCTAAGGATCTCGCCGAGAACAATCGTGGAGCACGTGTCTCTCGTTGTCTGCT
                                     ***** * * *

CHS-GU      TTGAGATTATAGTTGTTATTTTTTGTGGTTTTTTTGATATTTATTTTGATTTTTTTGTG
CHS-cDNA    CTGAGATCACAGCCGTTACCTTCCGTGGTCCCTCTGACACCCACCTTGACTCCCTCGTCG
***** * * * * * * * * * * * * * * * * * * * * * * * * * * * *

CHS-GU      GTTAGGTTTTTTTTAGTGATGGTGTGTTGTTGTTATTTATTGTGGGGTTGGATTTTGATATAT
CHS-cDNA    GTCAGGCTCTTTTCAGTGATGGCGCCGCCGCACTCATTGTGGGGTCGGACCCTGACACAT
** *** * * * * * * * * * * * * * * * * * * * * * * * * * * * *

CHS-GU      TTGTTGGAGAGAAATTTATTTTTTGAGATGGTGTGTTGTTGTTTAGATTATTTTTTTAGATT
CHS-cDNA    CTGTCGGAGAGAAACCCATCTTTGAGATGGTGTCTGCCGCTCAGACCATCCTTCCAGACT
*** ***** * * * * * * * * * * * * * * * * * * * * * * * * * * * *

CHS-GU      TTGATGGTGT-----
CHS-cDNA    CTGATGGTGCCATAGACGGACATTTGAGGGAAGTTGGTCTCACCTTCCATCTCTCAAGG
*****

```

图23

反义EIN2-GU与反义EIN2 cDNA比对

```

As-EIN2-GU      -----AATGTTTATGTGAGTTGTAATA
asEIN2-cDNA     TCTTCATTGCAGATCTGAGCCAAGTGTTCACCAGTCACAACGCTTATGCGAGCTGCAACA
                      * * * * *
As-EIN2-GU      TATTGGTATAAGATGGTGGTAAAATTGAAAAGTAGAGTAATTGTTATTAAGTTATATTTG
asEIN2-cDNA     TATTGGCATAAGATGGCGGCAAAATTGAAAAGCAGAGTAATTGCCACCAAGTCATACCCG
                      * * * * *
As-EIN2-GU      AAATGAGTATTTTTTTTGATATTTGTAATTTATTTTTTGGGATTAATATATTTGATAGAA
asEIN2-cDNA     AAACGAGCACCTCCTTCGATATTTGCAACCCATTTCCCGGGATCAATATATCCGACAGAA
                      * * * * *
As-EIN2-GU      ATTAAAAGGATAGGAAGTAGAGTAGGAATTATTTTTTGGATAAATTTAGTTGAGGTT--
asEIN2-cDNA     ACCAAAAGGACAGGAAGTAGAGCAGGAACCATTTCTCTGGATAAACCCTAGCTGAGGTCTC
                      * * * * *

```

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      GGTTTGATTTTATAATGAGTGTGGTGGTGTATTATTGAAAAGAGTTTGATTGATGAGGG
asCHS-cDNA     GGTCCGACCCCAATGAGTGCGGCGGCGCCATCACTGAAAAGAGCCTGACCGACGAGGG
                      * * * * *
As-CHS-GU      AGTTAAGGTGGGTGTTAGAGGGATTATGGAAGGTAATGGTTGTGATTTTAGAGTAGATAA
asCHS-cDNA     AGTCAAGGTGGGTGTCAGAGGGACCACGGAAGGTAACGGCTGTGATCTCAGAGCAGACAA
                      * * * * *

```

图25

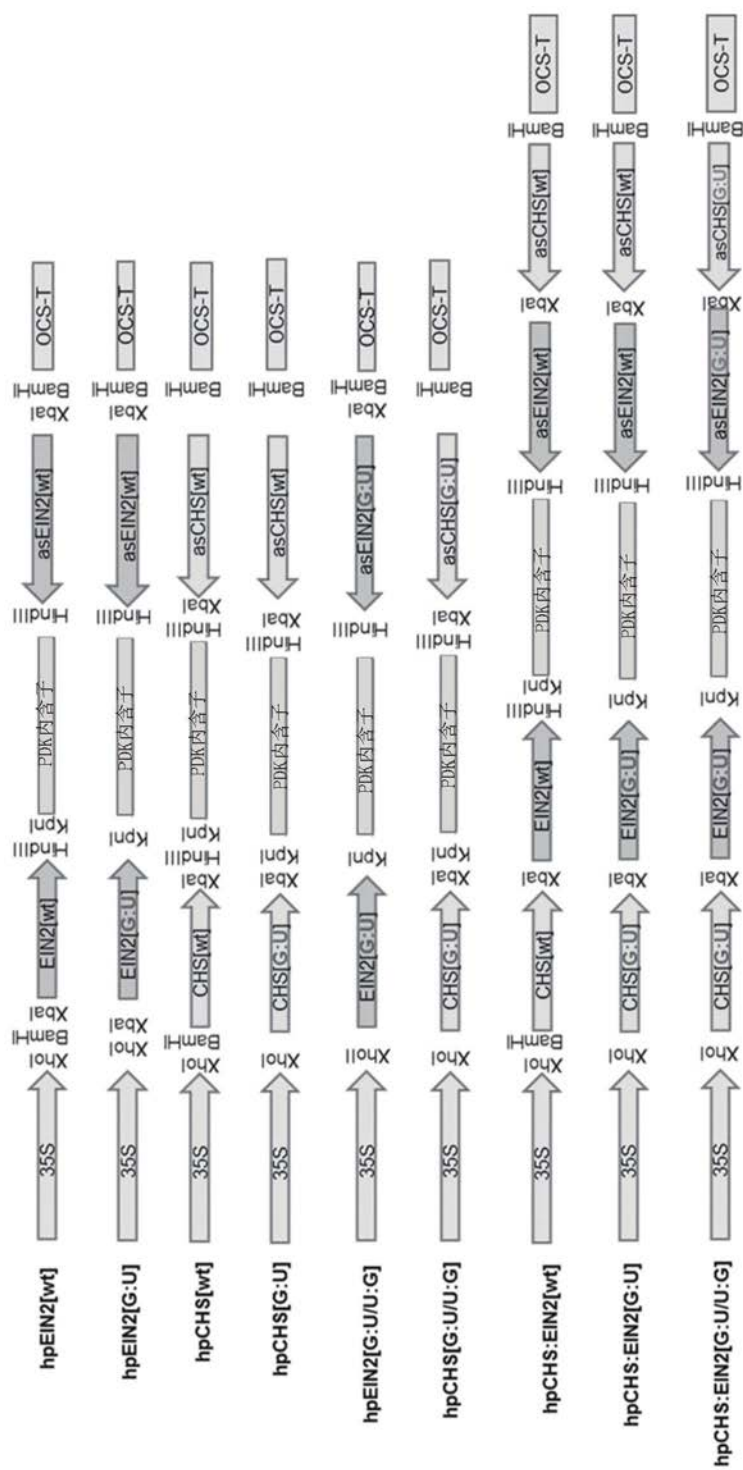


图26

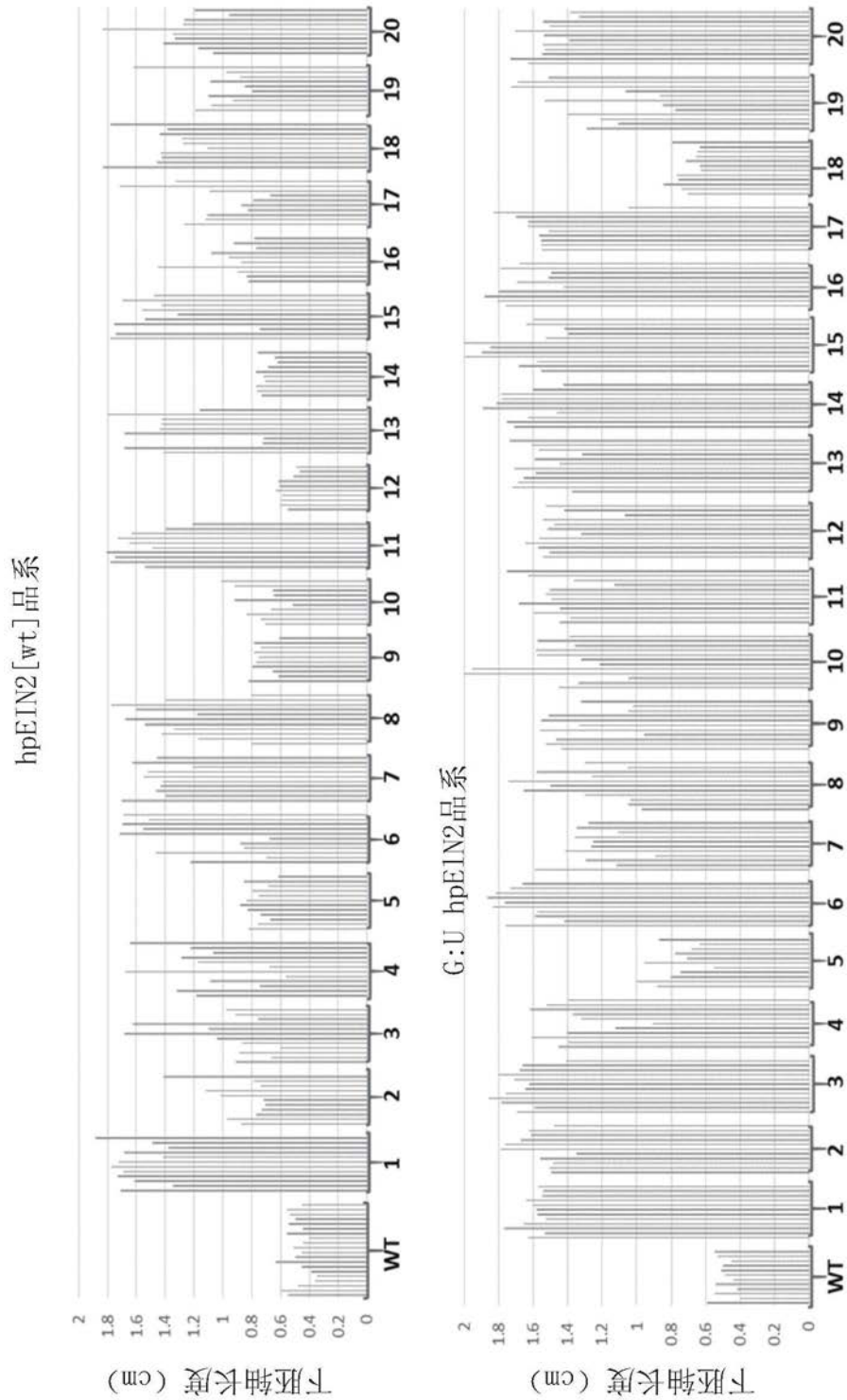


图27

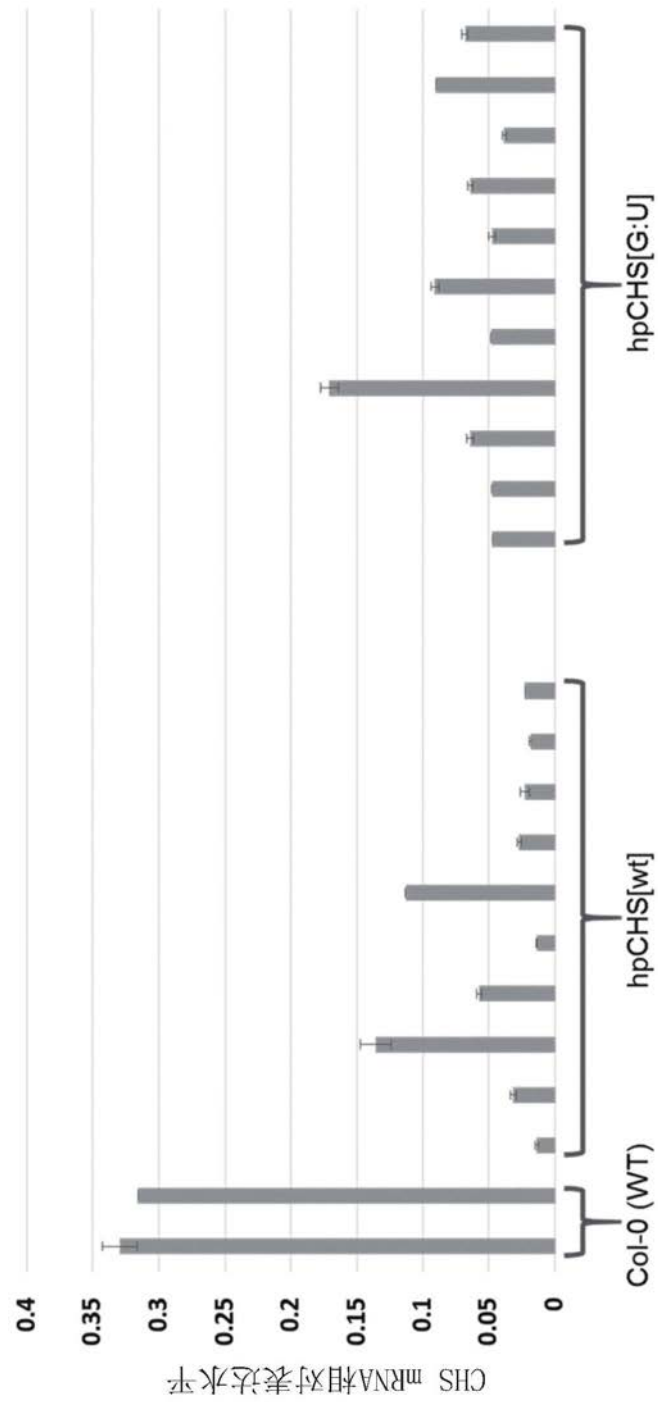


图28

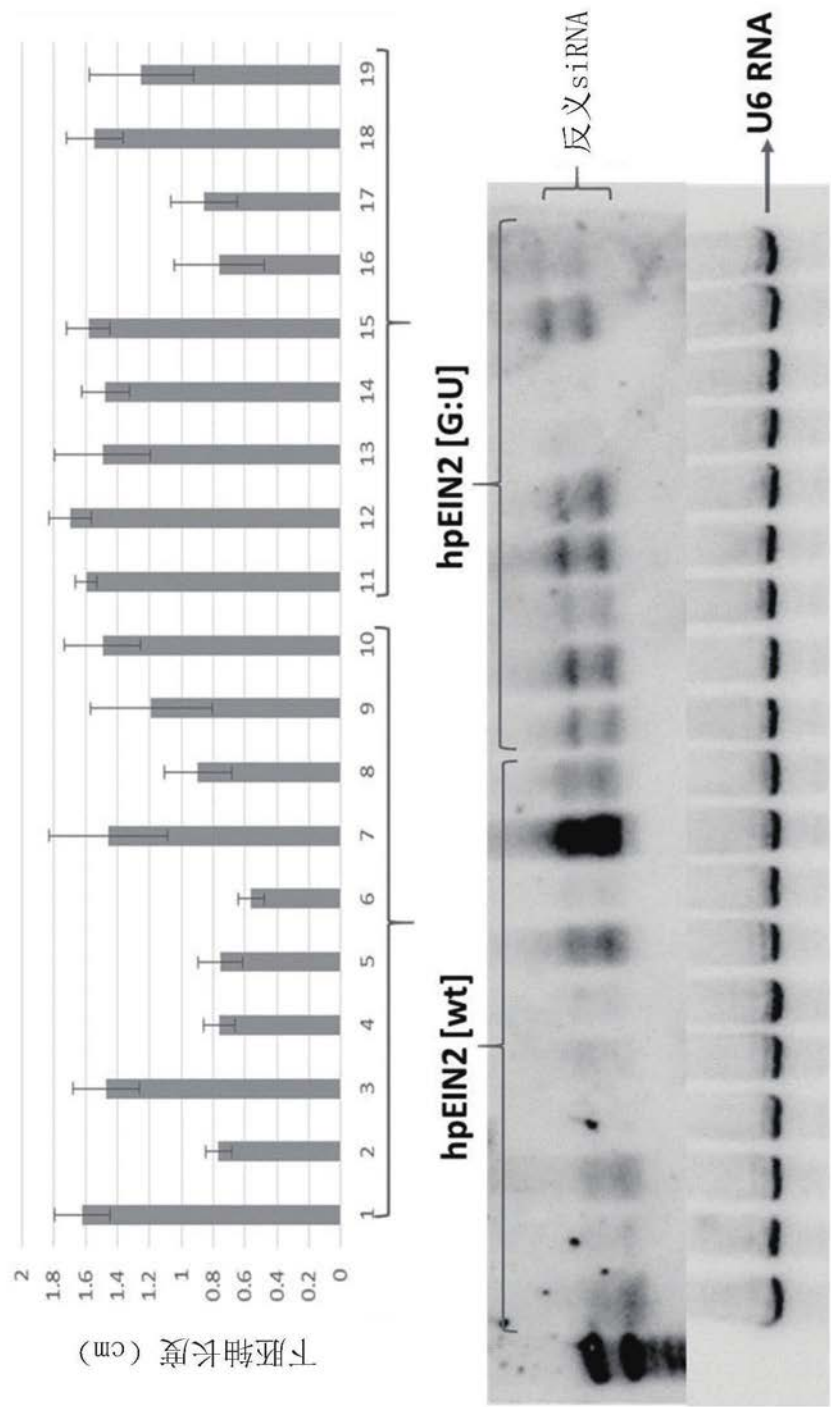


图29

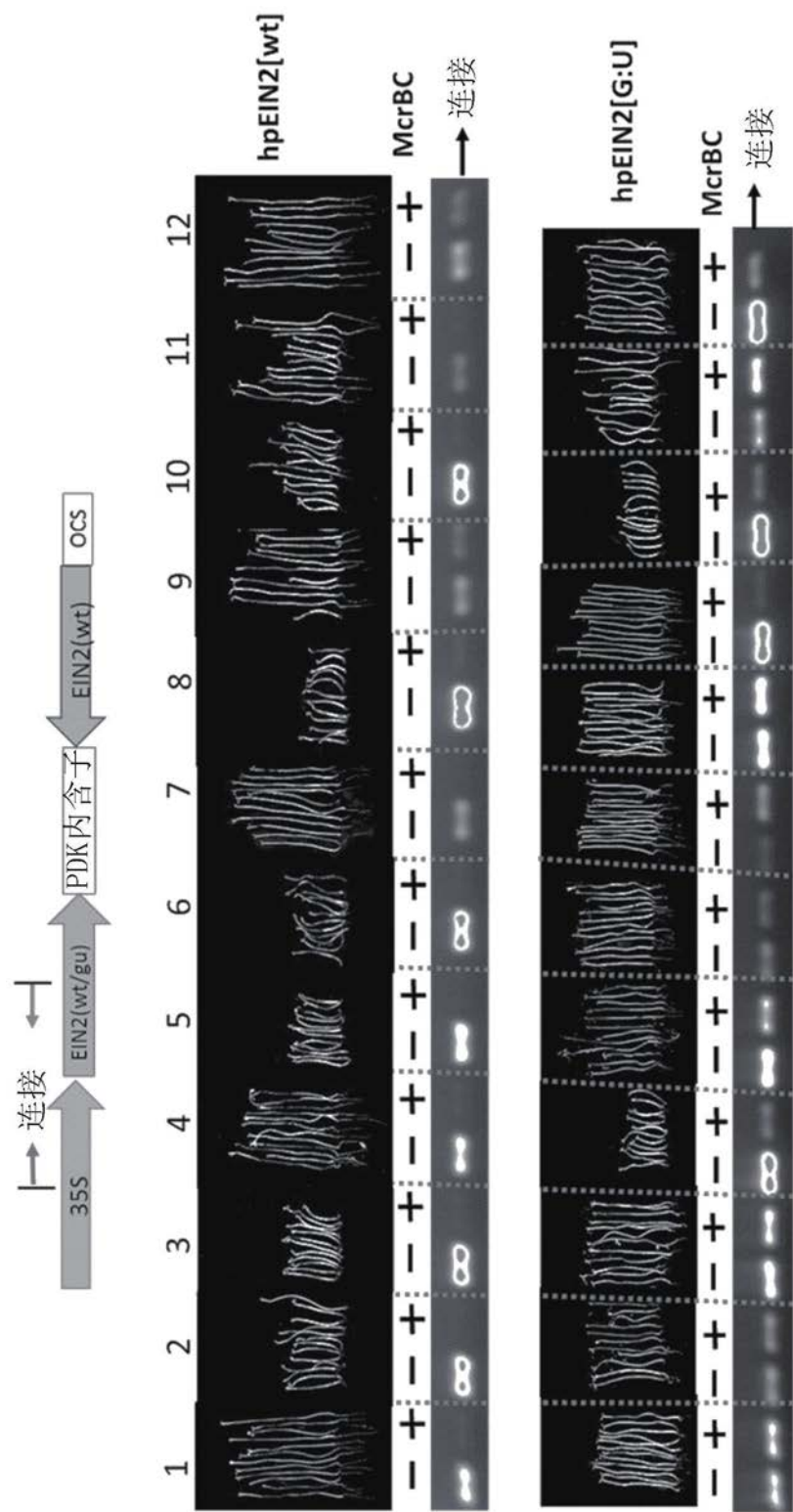


图30

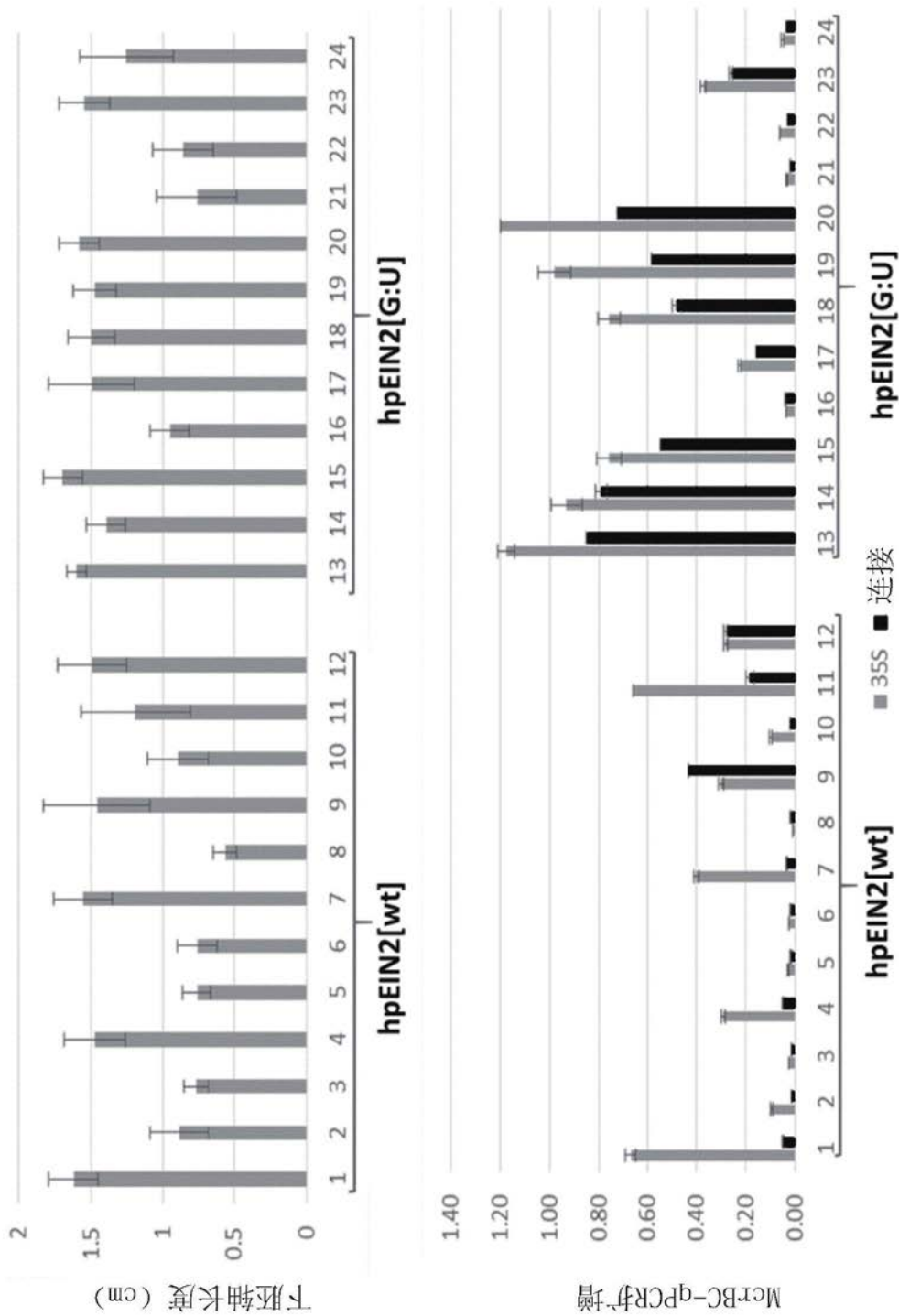


图31

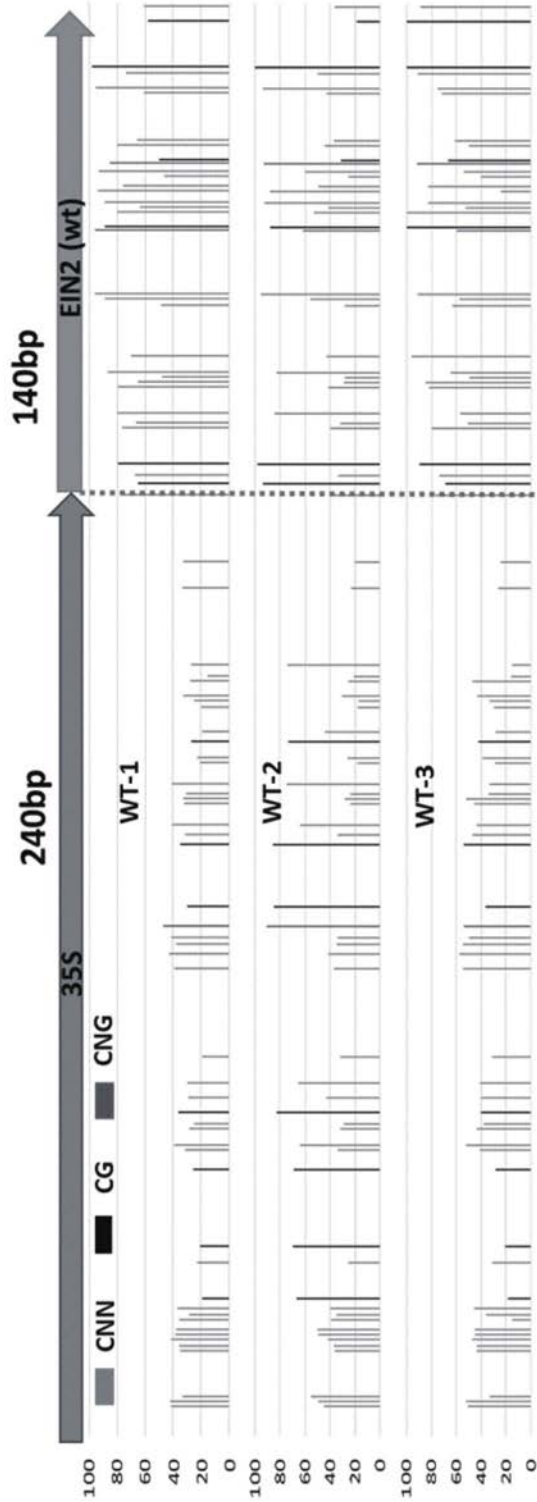


图32

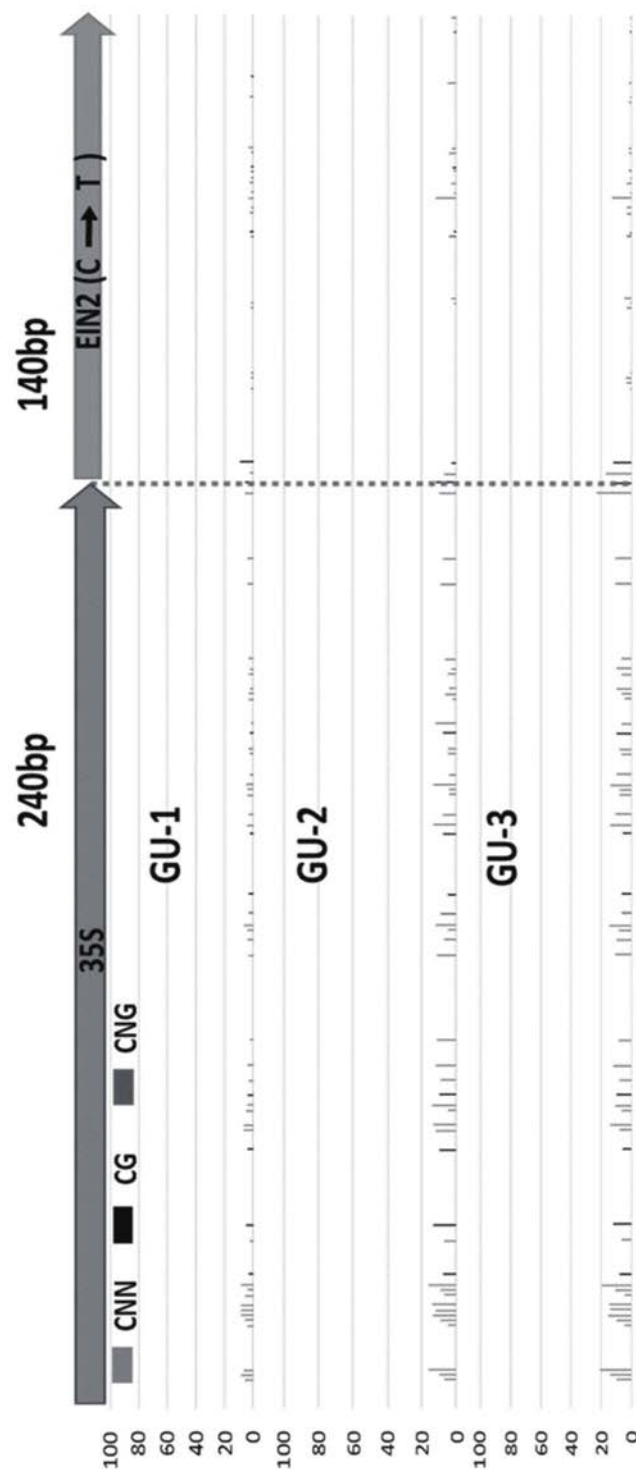


图33

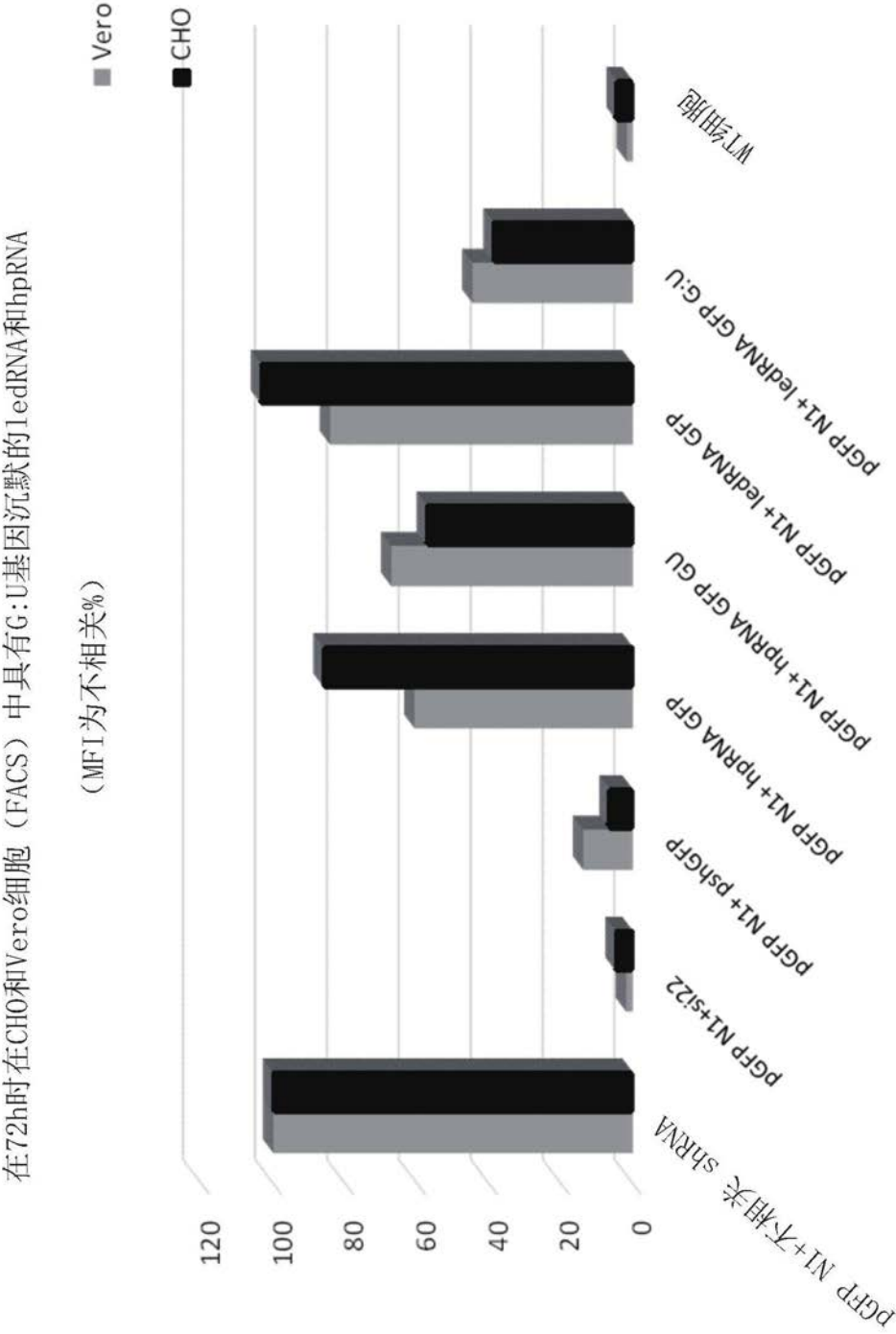


图34

06-2018 哑铃型质粒通过FACS在He1a细胞中48h时的测试
MFI为不相关%

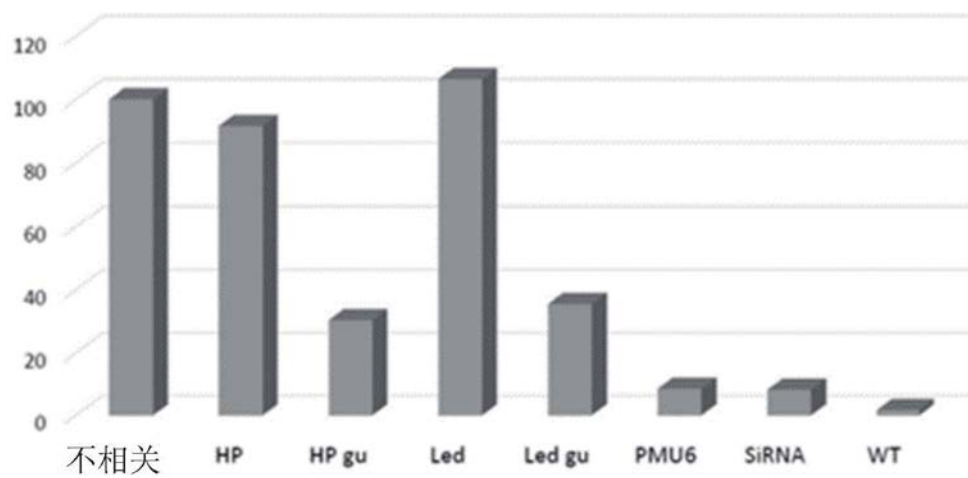


图35

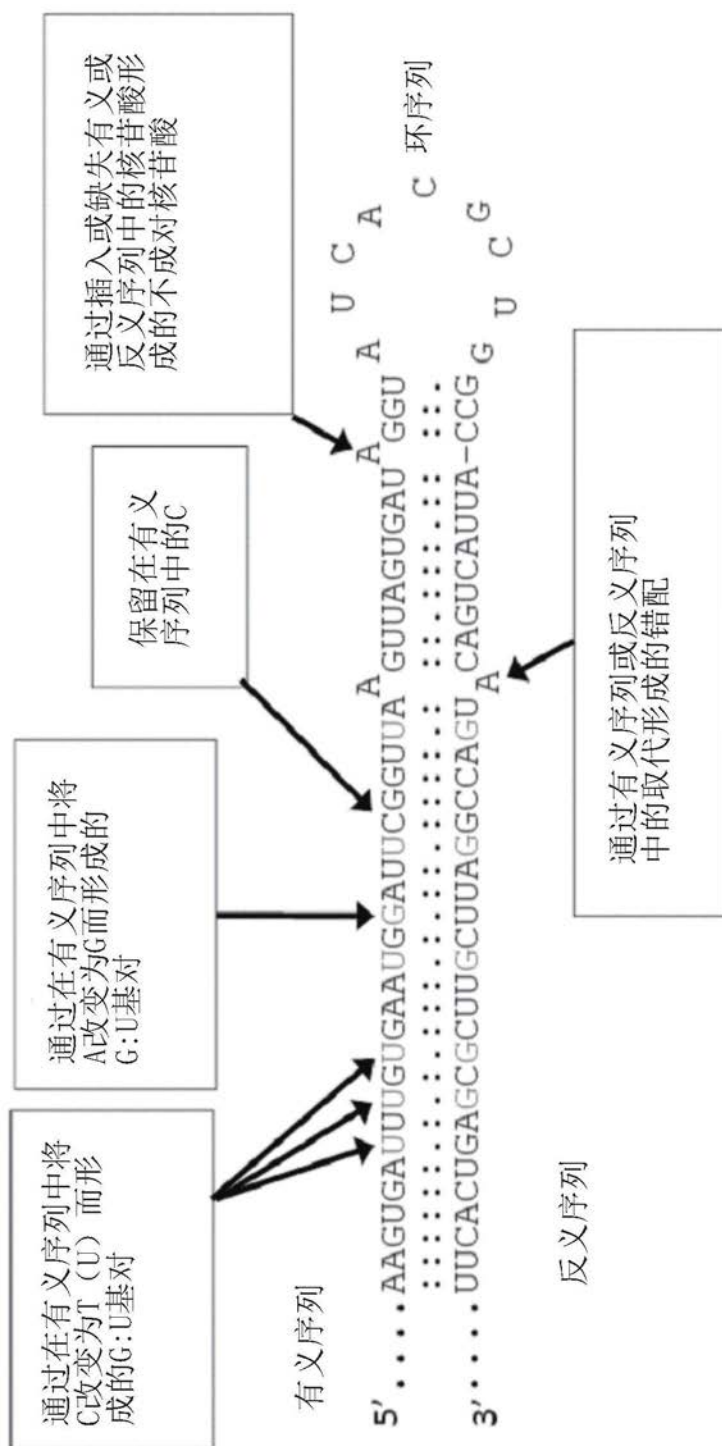


图36

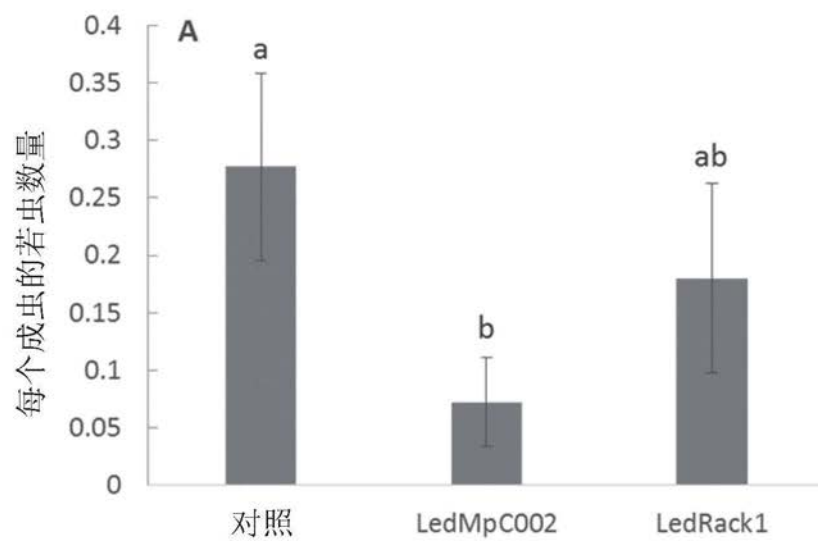
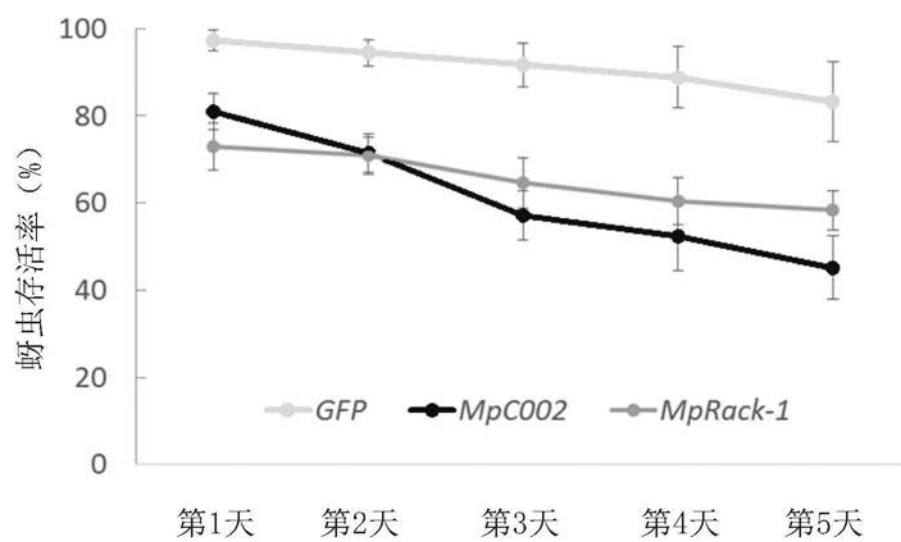
A**B**

图37

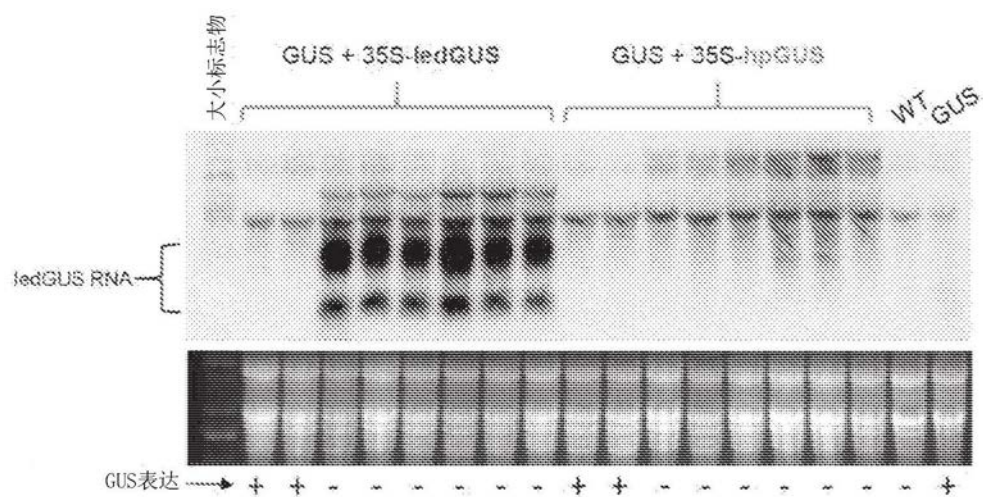


图38

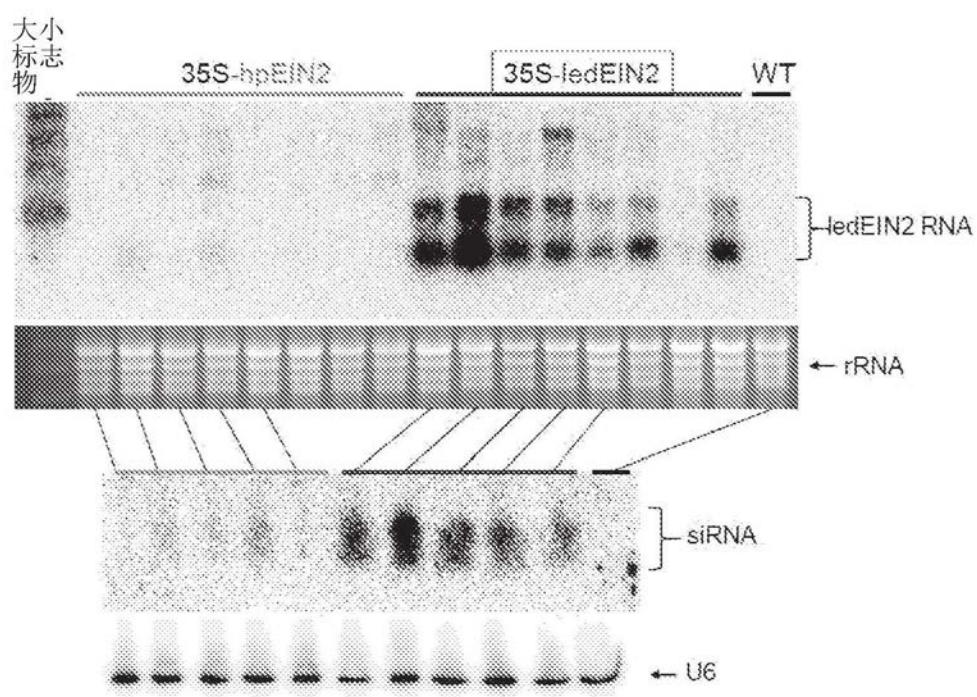


图39

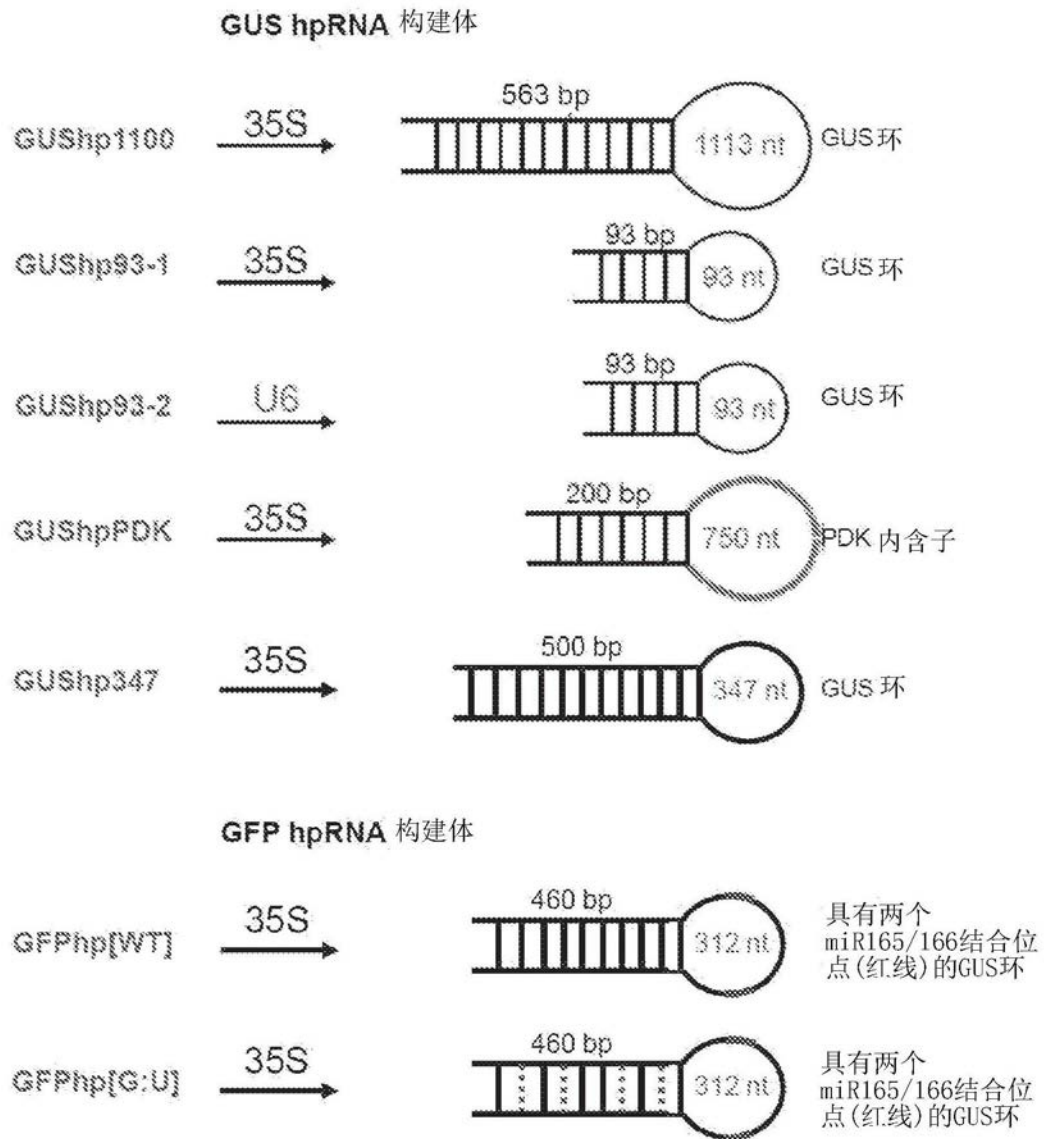


图40

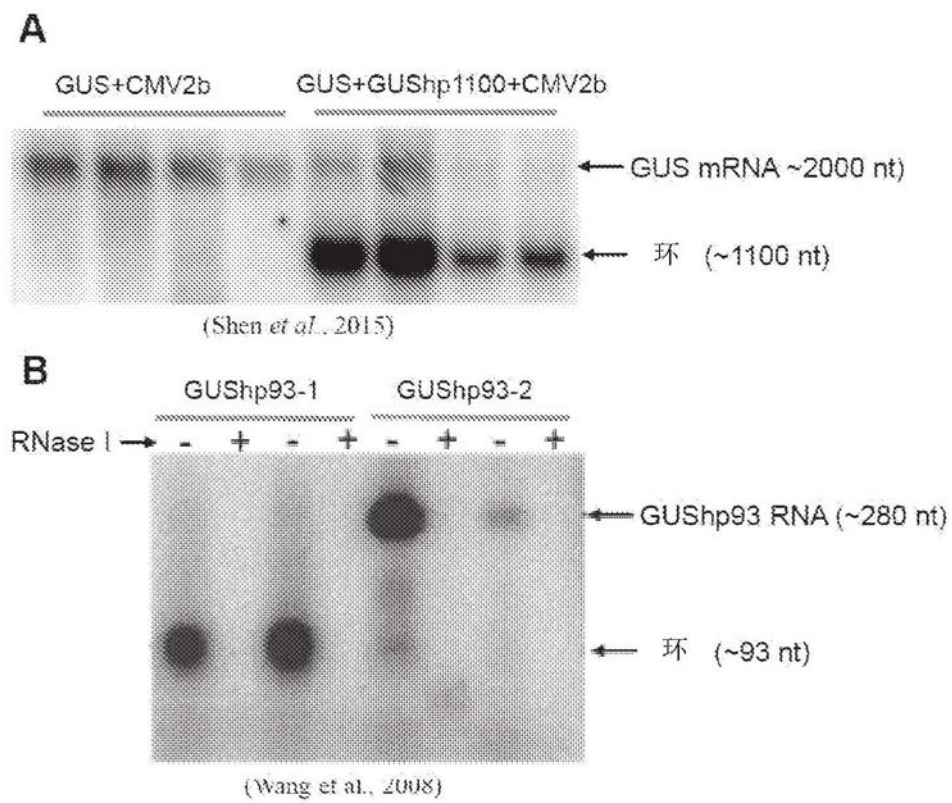


图41

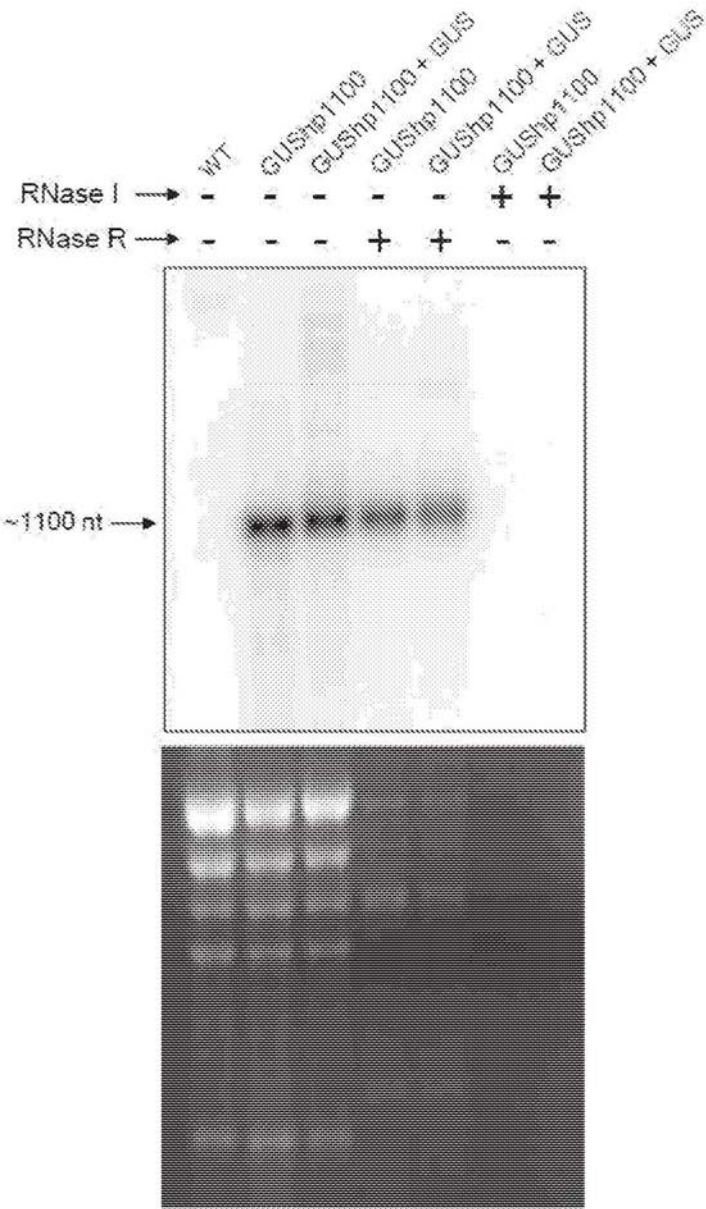


图42

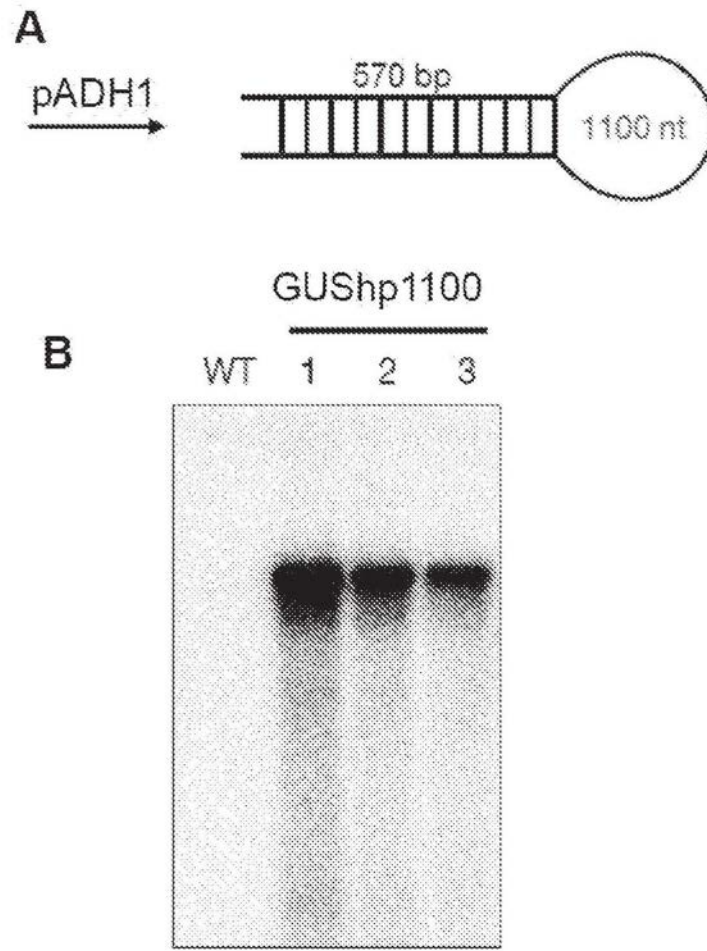


图43

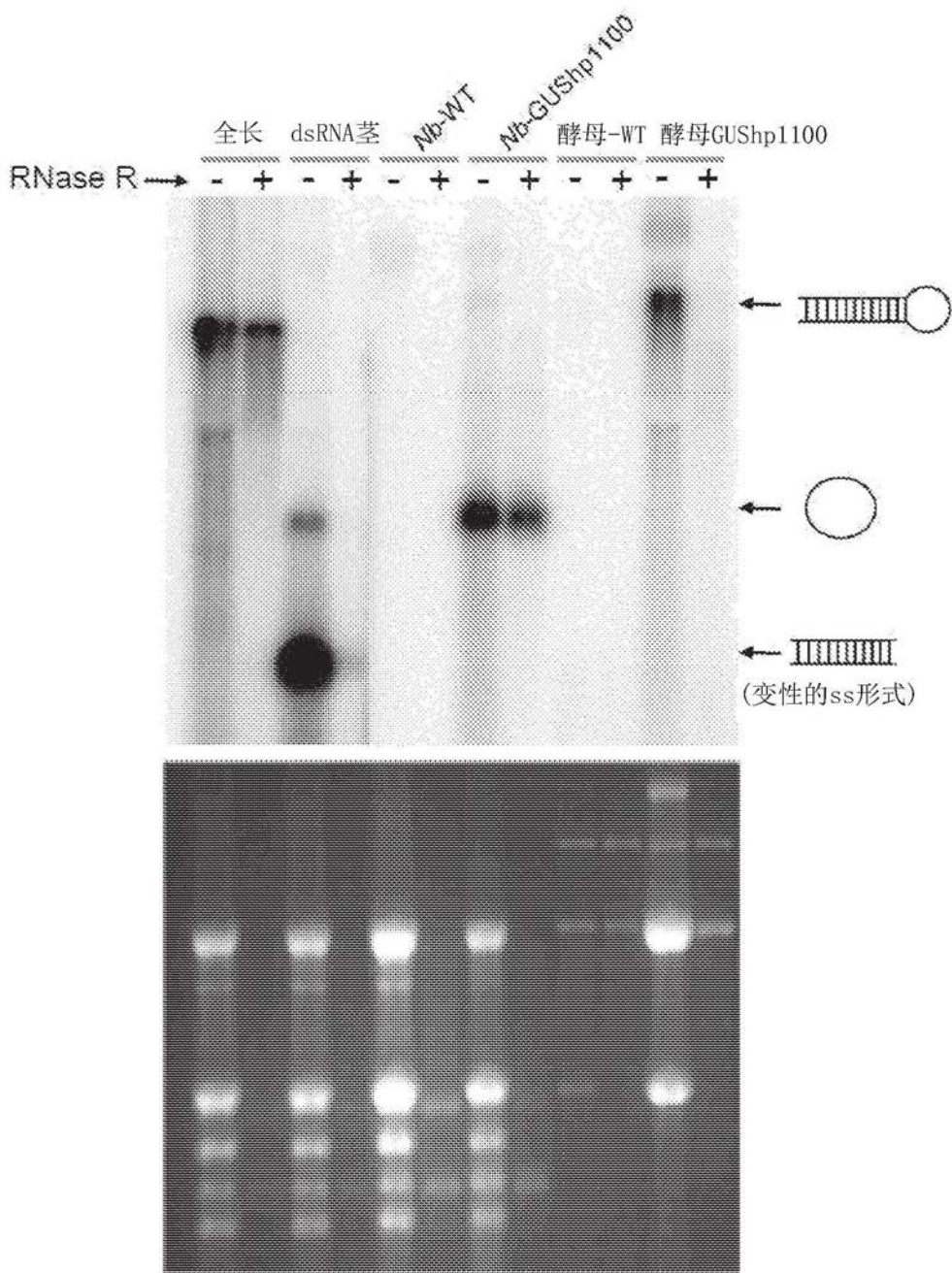


图44

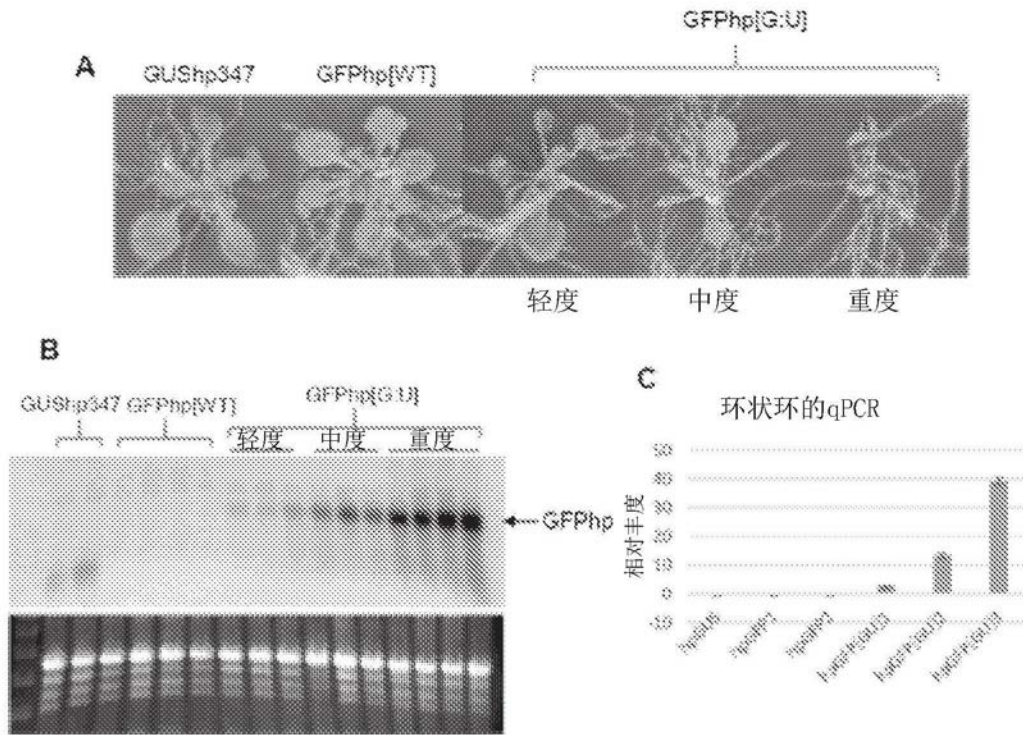


图45