

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12Q 1/68

C12N 15/10



[12] 发明专利申请公开说明书

[21] 申请号 01820900.9

[43] 公开日 2005 年 1 月 19 日

[11] 公开号 CN 1568373A

[22] 申请日 2001.11.29 [21] 申请号 01820900.9

[30] 优先权

[32] 2000.12.1 [33] EP [31] 00126325.0

[32] 2001.3.30 [33] US [31] 60/279,661

[32] 2001.3.30 [33] US [31] PCT/US01/10188

[86] 国际申请 PCT/EP2001/013968 2001.11.29

[87] 国际公布 WO2002/044321 英 2002.6.6

[85] 进入国家阶段日期 2003.6.19

[71] 申请人 马普科技促进协会

地址 德国慕尼黑

共同申请人 欧洲生物分子学实验室

[72] 发明人 T·图舍尔 S·埃尔巴沙

W·兰德科尔 M·威尔姆

R·吕尔曼

[74] 专利代理机构 中国国际贸易促进委员会专利
商标事务所

代理人 程 泳

权利要求书 4 页 说明书 41 页 附图 23 页

[54] 发明名称 介导 RNA 干涉的小 RNA 分子

[57] 摘要

双链 RNA(dsRNA)在多种生物中通过一种被称作 RNA 干涉(RNAi)的过程诱导序列特异的转录后基因沉默。利用果蝇体外系统，我们证明了 19 – 23 nt 短 RNA 片段是 RNAi 的序列特异性介体。短干涉 RNA(siRNA)通过长 dsRNA 的 RNase III 样加工反应产生。化学合成的具有突出 3' 端的 siRNA 双链体在裂解液中介导高效的靶 RNA 切割，切割位点位于指导 siRNA 所覆盖的区域的中心附近。此外，我们还证明了 dsRNA 的加工方向决定是有义还是反义靶 RNA 能被产生的 siRNP 复合物切割。

1. 分离的双链 RNA 分子，其中每条 RNA 链的长度为 19-25 个核苷酸，该 RNA 分子能够靶特异性核酸修饰。
2. 权利要求 1 的 RNA 分子，其中至少一条链具有 1-5 个核苷酸的 3'-突出端。
3. 权利要求 1 或 2 的 RNA 分子，它能靶特异性 RNA 干涉和/或 DNA 甲基化。
4. 权利要求 1-3 中任一项的 RNA 分子，其中每条链的长度为 19-23 个，特别是 20-22 个核苷酸。
5. 权利要求 2-4 中任一项的 RNA 分子，其中 3'-突出端为 1-3 个核苷酸。
6. 权利要求 2-5 中任一项的 RNA 分子，其中为避免降解稳定 3'-突出端。
7. 权利要求 1-6 中任一项的 RNA 分子，它含有至少一个修饰的核苷酸类似物。
8. 权利要求 7 的 RNA 分子，其中修饰的核苷酸类似物选自糖修饰或骨架修饰的核糖核苷酸。
9. 根据权利要求 7 或 8 的 RNA 分子，其中该核苷酸类似物是糖修饰的核糖核苷酸，2'-OH 基被置换为选自 H、OR、R、卤素、SH、SR1、NH2、NHR、NR2 或 CN 的基团，其中 R 是 C1-C6 烷基、烯基或炔基，卤素是 F、Cl、Br 或 I。
10. 权利要求 7 或 8 的 RNA 分子，其中该核苷酸类似物是含有硫代磷酸酯基的骨架修饰的核糖核苷酸。
11. 权利要求 1-10 中任一项的 RNA 分子，其序列与预定的 mRNA 靶分子有至少 50% 的同一性。
12. 权利要求 11 的 RNA 分子，其中同一性至少为 70%。
13. 一种制备权利要求 1-12 中任一项的双链 RNA 分子的方法，包括下列步骤：

(a) 合成两条 RNA 链，每条长度为 19-25 个核苷酸，其中该 RNA 链能够形成双链 RNA 分子，

(b) 在一定条件下结合合成的 RNA 链，其中形成一种双链 RNA 分子，它能靶特异性核酸修饰。

14. 权利要求 13 的方法，其中化学合成 RNA 链。

15. 权利要求 13 的方法，其中酶促合成 RNA 链。

16. 一种介导细胞或生物中靶特异性核酸修饰的方法，包括下列步骤：

(a) 在可发生靶特异性核酸修饰的条件下使细胞或生物接触权利要求 1-12 中任一项的双链 RNA 分子，和

(b) 将双链 RNA 完成的靶特异性核酸修饰导向含有基本对应于该双链 RNA 的序列部分的靶核酸。

17. 权利要求 16 的方法，其中核酸修饰是 RNA 干涉和/或 DNA 甲基化。

18. 权利要求 16 和 17 的方法，其中所述接触包括向可发生靶特异性核酸修饰的靶细胞中导入所述双链 RNA 分子。

19. 权利要求 18 的方法，其中这种导入包括载体介导的输送或注射。

20. 权利要求 16-19 中任一项的方法的用途，用于确定细胞或生物中一种基因的功能。

21. 权利要求 16-19 中任一项的方法的用途，用于调节细胞或生物中一种基因的功能。

22. 权利要求 20 或 21 的用途，其中该基因与一种病理状态有关。

23. 权利要求 22 的用途，其中该基因是一种病原体相关基因。

24. 权利要求 23 的用途，其中该基因是一种病毒基因。

25. 权利要求 22 的用途，其中该基因是一种肿瘤相关基因。

26. 权利要求 22 的用途，其中该基因是一种自身免疫病相关基因。

27. 含有至少一种权利要求 1-12 中任一项的双链 RNA 分子作为活性剂和一种药用载体的药用组合物。

28. 用于诊断的权利要求 27 的组合物。
29. 用于治疗的权利要求 27 的组合物。
30. 一种显示靶基因特异性敲除表型的真核细胞或真核非人类生物，其中用至少一种能够抑制内源靶基因表达的双链 RNA 分子或用编码能够抑制至少一种内源靶基因表达的至少一种双链 RNA 分子的 DNA 转染该细胞或生物。
31. 权利要求 30 的细胞或生物，它是一种哺乳动物细胞。
32. 权利要求 31 的细胞或生物，它是一种人类细胞。
33. 权利要求 30-32 中任一项的细胞或生物，其用至少一种编码靶蛋白或靶蛋白变体或突变形式的外源靶核酸进一步转染，其中该外源靶核酸在核酸水平上不同于内源靶基因，因此与内源靶基因的表达相比，外源靶核酸的表达受双链 RNA 分子的抑制大大减少。
34. 权利要求 33 的细胞或生物，其中外源靶核酸与编码可检测肽或多肽的另一种核酸序列融合。
35. 权利要求 30-34 中任一项的细胞或生物在分析方法中的用途。
36. 权利要求 35 在基因表达谱分析中的用途。
37. 权利要求 35 在蛋白质组分析中的用途。
38. 权利要求 35-37 中任一项的用途，其中对外源靶核酸编码的靶蛋白的变体或突变形式进行分析。
39. 权利要求 38 在鉴定靶蛋白功能域中的用途。
40. 权利要求 35-39 中任一项的用途，其中对至少两种细胞或生物进行比较，它们选自：
 - (i) 一种对照细胞或对照生物，没有靶基因抑制，
 - (ii) 一种细胞或生物，具有靶基因抑制，和
 - (iii) 一种细胞或生物，具有靶基因抑制加外源靶核酸对靶基因的互补。
41. 权利要求 35-40 中任一项的用途，其中的分析包括功能和/或表型分析。
42. 权利要求 30-34 中任一项的细胞在制备方法中的用途。

43. 权利要求 41 的用途，用于从真核细胞中分离蛋白质或蛋白质复合物。

44. 权利要求 43 的用途，用于分离任选地含有核酸的高分子量蛋白质复合物。

45. 权利要求 35-44 中任一项在鉴定和/或表征药理试剂中的用途。

46. 一种系统，用于鉴定和/或表征作用于至少一种靶蛋白的药理试剂，其包括：

(a) 一种真核细胞或真核非人类生物，它能表达至少一种编码所述至少一种靶蛋白的内源靶基因，

(b) 至少一种双链 RNA 分子，它能抑制所述至少一种内源靶基因的表达，和

(c) 一种待测物质或一组待测物质，其中将要鉴定和/或表征这种待测物质或组合的药理学性质。

47. 权利要求 46 的系统，其进一步包括：

(d) 编码靶蛋白或靶蛋白变体或突变形式的至少一种外源靶核酸，其中外源靶核酸在核酸水平上不同于内源靶基因，因此与内源靶基因的表达相比，外源靶核酸的表达受双链 RNA 分子的抑制大大减少。

介导 RNA 干涉的小 RNA 分子

发明领域

本发明涉及介导靶特异的核酸修饰(如 RNA 干涉和/或 DNA 甲基化)所需的双链(ds)RNA 分子的序列和结构特征。

发明背景

向线虫 *C. elegans* 中注射 dsRNA 导致在序列上与注射的 dsRNA 高度同源的基因特异性沉默 (Fire 等人, 1998), 在这一发现后, 提出术语 “RNA 干涉”(RNAi)。之后也在昆虫、青蛙(Oelgeschlager 等人, 2000)和包括小鼠在内的其他动物(Svoboda 等人, 2000; Wianny 和 Zernicka-Goetz, 2000)中观察到 RNAi, 在人类中也可能存在。RNAi 与植物中共抑制和真菌中抑制(quelling)的转录后基因沉默(PTGS)机制密切相关(Catalanotto 等人, 2000; Cogoni 和 Macino, 1999; Dalmay 等人, 2000; Ketting 和 Plasterk, 2000; Mourrain 等人, 2000; Smardon 等人, 2000), RNAi 机制的一些组分也是共抑制引起的转录后沉默所必需的(Catalanotto 等人, 2000; Dernburg 等人, 2000; Ketting 和 Plasterk, 2000)。最近也对这一主题进行了综述(Bass, 2000; Bosher 和 Labouesse, 2000; Fire, 1999; Plasterk 和 Ketting, 2000; Sharp, 1999; Sijen 和 kooter, 2000), 参见 2000 年 3 月 2 日发行的《植物分子生物学》第 43 卷的完整内容。

在植物中, 除了 PTGS 之外, 导入的转基因也能通过 RNA 指导的胞嘧啶 DNA 甲基化导致转录基因沉默(参见 Wassenegger, 2000)。植物中短至 30 bp 的基因组靶标以 RNA 指导的方式甲基化(Pelissier, 2000)。DNA 甲基化在哺乳动物中也存在。

RNAi 和共抑制的自然功能似乎是保护基因组免受移动遗传因子(例如在有活性时在宿主细胞中产生异常 RNA 或 dsRNA 的反转录转座子和病毒)的入侵(Jensen 等人, 1999; Ketting 等人, 1999; Ratcliff

等人, 1999; Tabara 等人, 1999). 特异性 mRNA 降解阻止转座子和病毒的复制, 尽管某些病毒通过表达抑制 PTGS 的蛋白质能克服或阻止这一过程(Lucy 等人, 2000; Voinnet 等人, 2000).

dsRNA 只触发与 dsRNA 相同的区域内的同源 RNA 特异性降解 (Zamore 等人, 2000)。dsRNA 被加工为 21-23 nt RNA 片段, 靶 RNA 切割位点规则地间隔 21-23 nt. 因此提示 21-23 nt 片段是靶识别的指导 RNA(Zamore 等人, 2000)。在细胞裂解前用 dsRNA 转染黑尾果蝇(*D. melanogaster Schneider*)2 细胞, 在由该细胞制备的提取物中也检测到这种短 RNA (Hammond 等人, 2000), 然而, 显示序列特异性核酸活性的级分也含有相当一部分残余 dsRNA。观察到从加工的 dsRNA 中分离的 21-23 nt 片段在一定程度上能介导特异性 mRNA 降解, 这进一步支持了 21-23 nt 片段在指导 mRNA 切割中的作用(Zamore 等人, 2000)。大小相似的 RNA 分子也在显示 PTGS 的植物组织中积累(Hamilton 和 Baulcombe, 1999)。

在此, 我们利用建立的果蝇(*Drosophila*)体外系统(Tuschl 等人, 1999; Zamore 等人, 2000)进一步研究 RNAi 机制。我们证实, 当短 21 nt 和 22 nt RNA 与 3'突出端碱基配对时, 它们作为序列特异性 mRNA 降解的指导 RNA. 在该系统中短 30 bp dsRNA 不能介导 RNAi, 因为它们不再加工为 21 nt 和 22 nt RNA. 此外, 我们还确定了相对于 21 和 22 nt 短干涉 RNA(siRNA)的靶 RNA 切割位点, 并且证实, dsRNA 加工的方向决定了是有义还是反义靶 RNA 能被产生的 siRNP 内切核酸酶复合物切割。此外, siRNA 也可能是转录调节的重要工具, 例如通过引导 DNA 甲基化使哺乳动物基因沉默。

在人类体内细胞培养系统(HeLa 细胞)中进行的其它实验表明, 长度优选地为 19-25 个核苷酸的双链 RNA 分子具有 RNAi 活性。因此, 与果蝇的结果不同, 24 和 25 nt 长双链 DNA 分子对于 RNAi 也是有效的。

发明内容

本发明的目的在于提供能够介导靶特异性 RNA 干涉或其它靶特

异性核酸修饰(如 DNA 甲基化)的新型试剂，这种试剂与现有试剂相比具有更高的效能和安全性。

一种分离的双链 RNA 分子提供了这一问题的解决方案，其中每条 RNA 链的长度为 19-25，特别是 19-23 个核苷酸，其中该 RNA 分子能够介导靶特异性核酸修饰，特别是 RNA 干涉和/或 DNA 甲基化。优选地，至少一条链具有 1-5 个核苷酸、更优选地 1-3 个核苷酸、最优选地 2 个核苷酸的 3'-突出端。另一条链可以是平端，或者具有可达 6 个核苷酸的 3' 突出端。如果 dsRNA 的两条链都正好是 21 或 22 nt，当两端都是平端时(0 nt 突出)，也能观察到一定程度的 RNA 干涉。该 RNA 分子优选地是合成 RNA 分子，基本上不含细胞提取物，例如果蝇胚胎中存在的污染物。此外，该 RNA 分子优选地基本上不含任何非靶特异的污染物，特别是非靶特异的 RNA 分子，例如细胞提取物中存在的污染物。

此外，本发明还涉及分离的双链 RNA 分子的用途，其中每条 RNA 链的长度为 19-25 个核苷酸，用于介导哺乳动物细胞中，特别是人类细胞中靶特异的核酸修饰，特别是 RNAi.

惊奇地发现，合成的短双链 RNA 分子，特别是具有突出 3' 端的，是序列特异的 RNAi 的介体，介导高效的靶-RNA 切割，其中切割位点靠近指导短 RNA 覆盖的区域的中心。

优选地，RNA 分子的每条链的长度为 20-22 个核苷酸(或者在哺乳动物细胞中为 20-25 个核苷酸)，其中每条链的长度可以相同或者不同。3'-突出端的长度优选地达到 1-3 个核苷酸，其中每条链的突出端长度可以相同或者不同。RNA-链优选地含有 3'-羟基。5'-端优选地含有磷酸、二磷酸、三磷酸或羟基。最有效的 dsRNA 由配对的两条 21 nt 链组成，使得 dsRNA 的两端存在 1-3 nt，特别是 2 nt 的 3' 突出端。

siRNA 指导的靶 RNA 切割反应是高度序列特异的。然而，不是 siRNA 的所有位点对于靶识别都有相同的作用。在 siRNA 双链体中心的错配最为关键，基本上消除靶 RNA 切割。相反，与单链靶 RNA 互补的 siRNA 链的 3' 核苷酸(例如位点 21)对于靶识别的特异性就没

有作用。此外，与靶 RNA 具有相同极性的 siRNA 链的未配对 2 nt 3' 突出端的序列对于靶 RNA 切割并不重要，因为只有反义 siRNA 链才能指导靶标识别。因此，从单链突出核苷酸起，只有反义 siRNA 的倒数第二个位点(例如位点 20)需要匹配靶向的有义 mRNA.

令人吃惊的是，本发明的双链 RNA 分子在血清或用于细胞培养的生长培养基中显示很高的体内稳定性。为了进一步提高稳定性，可以稳定 3' 突出端使其免受降解，例如，可以加以选择，使其由嘌呤核苷酸，特别是腺苷或鸟苷核苷酸组成。此外，嘧啶核苷酸置换为修饰类似物，例如尿苷 2 nt 3' 突出端置换为 2'-脱氧胸苷可以耐受，不影响 RNA 干涉的效率。不含 2' 羟基可以显著提高突出端在组织培养基中的核酸酶抗性。

在本发明的一个特别优选的实施方案中，RNA 分子可含有至少一种修饰的核苷酸类似物。该核苷酸类似物可位于基本上不影响靶特异性活性(如 RNAi 介导活性)的位置上，例如位于双链 RNA 分子的 5' 端和/或 3' 端区。特别是，掺入修饰的核苷酸类似物可以稳定突出端。

优选的核苷酸类似物选自糖修饰或骨架修饰的核糖核苷酸。然而应当指出，核碱基修饰的核糖核苷酸也是合适的，即核糖核苷酸，其含有非天然存在的核碱基代替天然存在的核碱基，如 5-位修饰的尿苷或胞苷，例如 5-(2-氨基)丙基尿苷、5-溴尿苷；8-位修饰的腺苷和鸟苷，例如 8-溴鸟苷；去氮杂核苷酸，例如 7-去氮杂-腺苷；O-和 N-烷基化核苷酸，例如 N6-甲基腺苷。在优选的糖修饰的核糖核苷酸中，2' OH 基被替换为选自 H、OR、R、卤素、SH、SR、NH₂、NHR、NR₂ 或 CN 的基团，其中 R 是 C₁-C₆ 烷基、烯基或炔基，卤素是 F、Cl、Br 或 I。在优选的骨架修饰的核糖核苷酸中，与相邻核糖核苷酸连接的磷酯基团被替换为一个修饰的基团，例如硫代磷酸酯 (phosphothioate) 基。应当指出，上述修饰可以组合。

为了介导靶特异的 RNAi 和/或 DNA 甲基化，本发明的双链 RNA 分子的序列必须与核酸靶分子具有足够的同一性。优选地，该序列与 RNA 分子双链部分中的希望的靶分子至少 50%，特别是至少 70% 相

同。在 RNA 分子的双链部分中，同一性更优选地至少 85%，最优先地 100%。一种双链 RNA 分子与预定的核酸靶分子(如 mRNA 靶分子)的同一性可以如下确定：

$$I = \frac{n}{L} \times 100$$

其中 I 是百分同一性，n 是 dsRNA 的双链部分中与靶标相同的核苷酸数，L 是 dsRNA 的双链部分与靶标的序列重叠的长度。

此外，也可包括 3' 突出端，特别是长度为 1-3 个核苷酸的突出端，确定双链 RNA 分子与靶序列的同一性。在此情况下，与靶序列的序列同一性优先地至少 50%，更优先地至少 70%，最优先地至少 85%。例如，来自 3' 突出端的核苷酸和来自双链 5' 和/或 3' 端的可达 2 个核苷酸可以修饰，而活性没有显著丧失。

本发明的双链 RNA 分子可以用包括下列步骤的方法制备：

(a) 合成两条 RNA 链，每条长度为 19-25 个，例如 19-23 个核苷酸，其中该 RNA 链能形成双链 RNA 分子，优先地至少一条链具有 1-5 个核苷酸的 3'-突出端，

(b) 在一定条件下结合合成的 RNA 链，其中形成一种双链 RNA 分子，它能介导靶特异的核酸修饰，特别是 RNA 干涉和/或 DNA 甲基化。

合成 RNA 分子的方法在本领域中众所周知。在本文中，具体是指如 Verma 和 Eckstein (1998) 所述的化学合成法。

由合成的 DNA 模板或者由分离自重组细菌的 DNA 质粒进行酶转录也能制备单链 RNA。一般使用噬菌体 RNA 聚合酶，如 T7、T3 或 SP6 RNA 聚合酶(Milligan 和 Uhlenbeck (1989))。

本发明的另一个方面涉及一种介导细胞或生物中的靶特异性核酸修饰(特别是 RNA 干涉和/或 DNA 甲基化)的方法，包括下列步骤：

(a) 在可以发生靶特异性核酸修饰的条件下使细胞或生物接触本发明的双链 RNA 分子，和

(b) 将双链 RNA 完成的靶特异性核酸修饰导向含有基本对应于双链 RNA 的序列部分的靶核酸。

优选地，接触步骤(a)包括将双链 RNA 分子导入靶细胞中，例如分离的靶细胞，例如细胞培养物，单细胞微生物，或多细胞生物内的一个靶细胞或多个靶细胞。更优选地，该导入步骤包括载体介导的输送，例如利用脂质体载体或通过注射。

本发明的方法可用来确定细胞或生物中一种基因的功能，乃至调节细胞或生物中一种基因的功能，它能够介导 RNA 干涉。细胞优选地是真核细胞或细胞系，例如植物细胞或动物细胞，如哺乳动物细胞，例如胚细胞、多能干细胞、肿瘤细胞，例如畸胎癌细胞或病毒感染的细胞。生物优选地是真核生物，例如植物或动物，如哺乳动物，特别是人类。

本发明的 RNA 分子针对的靶基因可能与病理状态有关。例如，该基因可能是病原体相关基因，例如病毒基因，肿瘤相关基因，或自身免疫病相关基因。靶基因也可能是在重组细胞或遗传改变的生物中表达的异源基因。通过确定或调节，特别是抑制这种基因的功能，可能获得在农业领域或在医学或兽医学领域有价值的信息和治疗效益。

dsRNA 通常作为药用组合物施用。可用已知方法进行施用，其中向体外或体内的希望的靶细胞中导入核酸。常用的基因转移技术包括磷酸钙、DEAE-葡聚糖、电穿孔、显微注射和病毒法(Graham, F.L. 和 van der Eb, A.J. (1973) Virol. 52, 456; McCutchan, J.H. 和 Pagano, J.S. (1968), J. Natl. Cancer Inst. 41, 351; Chu, G. 等人 (1987), Nucl. Acids Res. 15, 1311; Fraley, R. 等人 (1980), J. Biol. Chem. 255, 10431; Capecchi, M.R. (1980), Cell 22, 479). 向细胞中导入 DNA 的一种最新技术是使用阳离子脂质体(Felgner, P.L. 等人(1987), Proc. Natl. Acad. Sci USA 84, 7413). 市售的阳离子脂制剂有，例如 Tfx 50 (Promega) 或 Lipofectamin2000 (Life Technologies)。

因此，本发明也涉及一种药用组合物，其含有至少一种上述双链 RNA 分子作为活性剂和一种药用载体。该组合物可在人类医学或兽医

学中用于诊断和治疗用途。

为用于诊断或治疗，该组合物可以是溶液形式，例如可注射溶液，乳膏、软膏、片剂、悬液等。该组合物可以用任何适当的方法施用，例如注射、口服、局部、鼻、直肠给药等。载体可以是任何合适的药用载体。优选地使用能提高 RNA 分子进入靶细胞的效能的载体。这类载体的合适的例子有脂质体，特别是阳离子脂质体。更加优选的施用方法是注射。

RNAi 法的另一个优选用途是对真核细胞或真核非人类生物的功能分析，优选的是哺乳动物细胞或生物，最优先的是人类细胞，例如细胞系，如 HeLa 或 293，或啮齿动物，例如大鼠或小鼠。通过用与预定靶基因同源的适当双链 RNA 分子或编码适当双链 RNA 分子的 DNA 分子转染，能够在靶细胞，例如细胞培养物或靶生物中获得特定敲除表型。惊奇地发现，存在短双链 RNA 分子不会引起宿主细胞或宿主生物的干扰素反应。

因此，本发明的另一个主题是显示靶基因特异的敲除表型的真核细胞或真核非人类生物，其中包含至少一种内源靶基因的表达至少部分缺陷，其中用能够抑制至少一种内源靶基因表达的至少一种双链 RNA 分子或用编码能够抑制至少一种内源靶基因表达的至少一种双链 RNA 分子的 DNA 转染该细胞或生物。应当指出，由于 RNAi 的特异性，本发明能够靶特异地敲除几种不同的内源基因。

细胞或非人类生物，特别是人类细胞或非人类哺乳动物的基因特异性敲除表型可以在分析方法中使用，例如复杂生理过程的功能和/或表型分析，如基因表达谱和/或蛋白质组的分析。例如，人们可以在培养的细胞中制备人类基因的敲除表型，假定它们是备选剪接过程的调节剂。这类基因具体包括 SR 剪接因子家族的成员，如 ASF/SF2、SC35、SRp20、SRp40 或 SRp55。另外也能分析 SR 蛋白对预定的其它剪接基因(如 CD44)的 mRNA 谱的影响。优选地利用基于寡核苷酸的芯片通过高通量方法进行这种分析。

利用基于 RNAi 的敲除技术，可以抑制靶细胞或靶生物中内源靶

基因的表达。内源基因可以用编码靶蛋白或靶蛋白变体或突变形式的外源靶核酸(例如基因或 cDNA)补充，外源核酸任选地可与编码可检测肽或多肽(例如亲和标记物，特别是多亲和标记物)的另一种核酸序列融合。靶基因的变体或突变形式不同于内源靶基因，因为由于单个或多个氨基酸的置换、插入和/或缺失，它们编码在氨基酸水平上与内源基因产物不同的基因产物。变体或突变形式可具有与内源靶基因相同的生物活性。另一方面，变异或突变的靶基因也可能具有不同于内源靶基因的生物活性，例如活性部分缺失、活性完全缺失、活性提高等。

这种互补可以如下完成：共表达外源核酸编码的多肽，例如含有靶蛋白和亲和标记物的融合蛋白和用于敲除靶细胞中内源基因的双链 RNA 分子。这种共表达可以利用一种合适的表达载体或者利用表达载体的组合实现，该载体表达外源核酸编码的多肽(例如标记物修饰的靶蛋白)和双链 RNA 分子。在靶细胞中从头合成的蛋白质和蛋白质复合物将含有外源基因产物，例如修饰的融合蛋白。为了避免 RNAi 双链分子抑制外源基因产物表达，编码外源核酸的核苷酸序列在与双链 RNA 分子同源的序列部分中可能在 DNA 水平上改变(在氨基酸水平上引起或不引起突变)。此外，也可以用来自其它种(例如来自小鼠)的相应核苷酸序列补充内源靶基因。

本发明的细胞或生物的优选用途是对基因表达谱和/或蛋白质组的分析。在一个特别优选的实施方案中，对一种或几种靶蛋白的变体或突变形式进行分析，其中利用如上所述的外源靶核酸将这种变体或突变形式重新导入细胞或生物中。内源基因敲除与利用突变的(例如部分缺失的)外源靶标拯救相组合比使用敲除细胞具有优势。而且，这种方法还特别适用于鉴定靶蛋白的功能域。在另一个优选实施方案中，例如对至少两种细胞或生物的基因表达谱和/或蛋白质组和/或表型特征进行比较。这些生物选自：

- (i) 一种对照细胞或对照生物，没有靶基因抑制，
- (ii) 一种细胞或生物，具有靶基因抑制，和
- (iii) 一种细胞或生物，具有靶基因抑制加外源靶核酸的靶基因互

补。

本发明的方法和细胞也适用于鉴定和/或表征药理试剂的方法，例如从一组待测物质中鉴定新的药理试剂，和/或表征已知药理试剂的作用机制和/或副作用。

因此，本发明也涉及一种鉴定和/或表征作用于至少一种靶蛋白的药理试剂的系统，其包括：

- (a) 一种真核细胞或真核非人类生物，它能表达至少一种编码该靶蛋白的内源靶基因，
- (b) 至少一种双链 RNA 分子，它能抑制所述至少一种内源靶基因的表达，和
- (c) 一种待测物质或一组待测物质，其中将要鉴定和/或表征这种待测物质或组合的药理学性质。

如上所述的系统优选地还包括：

(d) 至少一种编码靶蛋白或其变体或其突变形式的外源靶核酸，其中该外源靶核酸在核酸水平上不同于内源靶基因，因此与内源靶基因的表达相比，外源靶核酸的表达受双链 RNA 分子的抑制大大减少。

此外，RNA 敲除互补法还可用于制备目的，例如从真核细胞，特别是哺乳动物细胞，更具体地是从人类细胞中亲和纯化蛋白质或蛋白质复合物。在本发明的实施方案中，外源靶核酸优选地编码与亲和标记物融合的靶蛋白。

该制备方法可以用于高分子量蛋白质复合物的纯化，其分子量优选地 ≥ 150 kD，更优选地 ≥ 500 kD，任选地可含有核酸，如 RNA。具体例子有：由 U4/U6 snRNP 颗粒的 20 kD、60 kD 和 90 kD 蛋白质组成的异源三聚蛋白质复合物，由分子量 14、49、120、145 和 155 kD 的 5 种蛋白质组成的来自 17S U2 snRNP 的剪接因子 SF3b，含有 U4、U5 和 U6 snRNA 分子和大约 30 种蛋白质的 25S U4/U6/U5 三-snRNP 颗粒，其分子量约为 1.7 MD。

该方法适用于哺乳动物细胞，特别是人类细胞的功能蛋白质组分析。

下面，通过下列附图和实施例更详细地说明本发明。

附图说明

图 1：短至 38 bp 的双链 RNA 能介导 RNAi.

(A) 用于靶向 Pp-Luc mRNA 的 dsRNA 的图示。制备覆盖 29-504 bp 的三个系列的平端 dsRNA。标出了 dsRNA 有义链的第一个核苷酸相对于 Pp-Luc mRNA 起始密码子的位置(p1)。(B) RNA 干涉测定 (Tuschl 等人, 1999)。靶 Pp-Luc 与对照 Rr-luc 活性之比用缓冲液对照(黑条)标准化。dsRNA (5 nM)在果蝇裂解液中 25℃ 预温育 15 分钟，之后加入 7-甲基-鸟苷-加帽的 Pp-luc 和 Rr-luc mRNA (~50 pM)。继续温育 1 小时，然后通过双萤光素酶测定(Promega)分析。数据是至少 4 次独立实验的平均值±标准差。

图 2：29 bp dsRNA 不再加工为 21-23 nt 片段。在果蝇裂解液中内部 ^{32}P -标记的 dsRNA (5 nM)加工后形成 21-23 mer 的时程。标出了 dsRNA 的长度和来源。RNA 分子量标记(M)加于左道中，并标出片段大小。时间为 0 时的双条带是由于 dsRNA 的不完全变性。

图 3：短 dsRNA 只切割 mRNA 靶标一次。

(A) 10 nM 帽 ^{32}P -标记的有义或反义 RNA 与 p133 系列的 10 nM dsRNA 在果蝇裂解液中温育 1 小时，产生的稳定 5'切割产物的变性凝胶电泳。帽标记的靶 RNA 部分核酸酶 T1 消化并部分碱性水解(OH)产生长度标记物。dsRNA 针对的区域在两端表示为黑条。显示了 111 bp 长 dsRNA 的主要切割位点之间的 20-23 nt 间隔。水平箭头表明不是由于 RNAi 引起的非特异性切割。(B) 有义和反义靶 RNA 上切割位点的位置。加帽 177 nt 有义和 180 nt 反义靶 RNA 的序列以反平行方向表示，使得互补序列彼此相对。不同 dsRNA 针对的区域用位于有义与反义靶序列之间的不同颜色的线条标出。切割位点用环形标出：大环为强切割，小环为弱切割。 ^{32}P -放射性标记的磷酸基以星号标记。

图 4：21 和 22 nt RNA 片段通过 RNase III 样机制产生。

(A) dsRNA 加工后~21 nt RNA 的序列。定向克隆 dsRNA 加工产生的~21 nt RNA 片段并测序。来源于 dsRNA 有义链的寡核糖核苷酸

用蓝线表示，来源于反义链的用红线表示。如果在多个克隆中含有相同的序列，则使用粗线，右侧的数字是指频率。dsRNA 介导的靶 RNA 切割位点用橙色环形表示，大环表示强切割，小环表示弱切割(参见图 3B)。有义链之上的环形表示有义靶标内的切割位点，dsRNA 下面的环形表示反义靶标中的切割位点。在来源于 dsRNA 3' 端的~21 nt 片段中鉴定出高达 5 个其它核苷酸。这些核苷酸主要是 C、G 或 A 残基的随机组合，最可能是在 dsRNA 组成链的 T7 转录过程中以非模板方式添加的。(B) ~21 nt RNA 的核苷酸组成的二维 TLC 分析。~21 nt RNA 如下产生：内部放射性标记的 504 bp Pp-luc dsRNA 在果蝇裂解液中温育，凝胶纯化，然后用核酸酶 P1 (最上一行)或核糖核酸酶 T2 (最下一行)消化为单核苷酸。在存在指定 α -³²P 核苷三磷酸之一的情况下，通过转录在内部放射性标记 dsRNA。放射性通过磷成像检测。核苷 5'-一磷酸、核苷 3'-一磷酸、核苷 5',3'-二磷酸和无机磷酸盐分别表示为 pN、Np、pNp 和 P_i。黑色环形表示非放射性载体核苷酸的 UV 吸收点。用 T4 多核苷酸激酶和 γ -³²P-ATP 使核苷 3'-一磷酸 5' 磷酸化，制备放射性标记标准，通过与该标准共迁移鉴定 3',5'-二磷酸(红色环形)。

图 5：合成的 21 和 22 nt RNA 介导靶 RNA 切割。

(A) 对照 52 bp dsRNA 和合成 21 和 22 nt dsRNA 的图示。21 和 22 nt 短干涉 RNA (siRNA) 的有义链用蓝色表示，反义链用红色表示。除了双链体 5 的 22 nt 反义链之外，siRNA 的序列来源于 52 和 111 bp dsRNA 的克隆片段(图 4A)。双链体 6 和 7 中的 siRNA 是 111 bp dsRNA 加工反应所特有的。在双链体 1 和 3 的合成反义链序列中存在两个 3' 突出核苷酸，以绿色表示。对照 52 bp dsRNA 的两条链都通过体外转录制备，一部分转录物可能包括非模板 3' 核苷酸添加。siRNA 双链体指导的靶 RNA 切割位点表示为橙色环形(见图 4A 的图例)，如图 5B 所示确定。(B) 有义和反义靶 RNA 上切割位点的位置。靶 RNA 序列如图 3B 所示。对照 52 bp dsRNA (10 nM) 或 21 和 22 nt RNA 双链体 1-7 (100 nM) 与靶 RNA 一起在果蝇裂解液中 25°C 温育 2.5 小时。稳定的 5' 切割产物在凝胶上分析。切割位点在图 5A 中标出。52 bp dsRNA

针对的区域或有义或反义链在凝胶旁边以黑线标出。切割位点全都位于与 dsRNA 相同的区域内。为了精确确定反义链的切割位点，使用较低百分比的凝胶。

图 6：短 dsRNA 上的长 3' 突出端抑制 RNAi.

(A) 52 bp dsRNA 构建体的图示。有义和反义链的 3' 延伸分别以蓝色和红色表示。靶 RNA 上的切割位点用类似于图 4A 的橙色环形表示，如图 6B 所示确定。(B) 有义和反义靶 RNA 上切割位点的位置。靶 RNA 序列如图 3B 所示。dsRNA (10 nM) 与靶 RNA 一起在果蝇裂解液中 25°C 温育 2.5 小时。稳定的 5' 酶切产物在凝胶上分析。主要切割位点用水平箭头标出，也在图 6A 中表示。52 bp dsRNA 针对的区域在凝胶两侧以黑线表示。

图 7：建议的 RNAi 模型。

预测 RNAi 从 dsRNA (有义链为黑色，反义链为红色) 主要加工为 21 和 22 nt 短干扰 RNA (siRNA) 开始。如果 dsRNA 上存在短的突出 3' 核苷酸，它们可能有利于短 dsRNA 的加工。有待表征的 dsRNA 加工蛋白用绿色和蓝色椭圆形表示，在 dsRNA 上以不对称方式装配。在我们的模型中，假定的蓝色蛋白或蛋白域与 siRNA 链以 3'-5' 方向结合，而假定的绿色蛋白或蛋白域总是结合相对的 siRNA 链。这些蛋白质或亚组与 siRNA 双链体结合，并且保持其方向，这取决于 dsRNA 加工反应的方向。只有与蓝色蛋白结合的 siRNA 序列才能指导靶 RNA 切割。内切核酸酶复合物被称为小干扰核糖核蛋白复合物或 siRNP。在此我们假定，切割 dsRNA 的内切核酸酶也可切割靶 RNA，这可能是通过暂时置换不用于靶识别的被动 siRNA 链。靶 RNA 然后在序列互补指导 siRNA 识别的区域中心切割。

图 8：报道构建体和 siRNA 双链体。

(a) 显示来自质粒 pGL2-对照、pGL-3-对照和 pRL-TK (Promega) 的萤火虫(Pp-luc)和海肾(Rr-luc)萤光素酶报道基因区。显示 SV40 调节元件、HSV 胸苷激酶启动子和两个内含子(线)。GL3 萤光素酶的序列与 GL2 有 95% 相同，但 RL 与这两者完全无关。在转染的哺乳动物

细胞中由 pGL2 表达萤光素酶比由 pGL3 表达大约低 10 倍。SiRNA 双链体针对的区域在萤光素酶基因编码区下用黑线表示。(b) 显示针对 GL2、GL3 和 RL 萤光素酶的 siRNA 双链体的有义(上)和反义(下)序列。GL2 和 GL3 siRNA 双链体只有 3 个单核苷酸置换(以灰色框出)的不同。合成含有反向 GL2 序列的双链体 invGL2 作为非特异对照。2'-脱氧胸昔的 2 nt 3' 突出端表示为 TT; uGL2 类似于 GL2 siRNA, 但是含有核糖尿昔 3' 突出端。

图 9: siRNA 双链体引起的 RNA 干涉。

靶与对照萤光素酶之比用缓冲液对照(bu, 黑线)标准化; 灰线表示 Photinus pyralis (Pp-luc) GL2 或 GL3 萤光素酶与 Renilla reniformis (Rr-luc) RL 萤光素酶之比(左轴), 白线表示 RL 与 GL2 或 GL3 之比(右轴)。图 a、c、e、g、i 描述用 pGL2-对照和 pRL-TK 报道质粒的组合进行的实验, 图 b、d、f、h、j 使用 pGL3-对照和 pRL-TK 报道质粒。用于干涉实验的细胞系在每张图的顶部标出。在标准化之前及在所检测的不同细胞系之间, 缓冲液对照(bu)的 Pp-luc/Rr-luc 比分别为: pGL2/pRL:0.5-10, pGL3/pRL:0.03-1。作图的数据是三次独立实验的平均值±S.D..

图 10: 21 nt siRNA、50 bp 和 500 bp dsRNA 对 HeLa 细胞中萤光素酶表达的影响。

长 dsRNA 的确切长度在线条下面标出。图 a、c、e 描述用 pGL2-对照和 pRL-TK 报道质粒进行的实验, 图 b、d、f 使用 pGL3-对照和 pRL-TK 报道质粒。数据是两次独立实验的平均值±S.D.. (a), (b) 绝对 Pp-luc 表达, 以任意发光单位作图。(c), (d) Rr-luc 表达, 以任意发光单位作图。(e), (f) 标准化靶标与对照萤光素酶之比。siRNA 双链体的萤光素酶活性之比用缓冲液对照(bu, 黑线)标准化; 50 或 500 bp dsRNA 的发光比分别用来自人源化 GFP (hG, 黑线)的 50 或 500 bp dsRNA 的发光比标准化。应当指出, 靶向 GL2 和 GL3 的 49 与 484 bp dsRNA 之间的总体序列差异不足以使 GL2 与 GL3 靶标之间具有特异性(49 bp 片段中有 43 nt 连续同一性, 484 bp 片段中有 239 nt 最长连

续同一性)。

图 11: 21-nt siRNA 双链体 3'突出端的变异。

(A) 实验策略概述。显示了加帽的和聚腺苷酸化的有义靶 mRNA，并且显示了有义和反义 siRNA 的相对位置。根据 8 个不同的反义链，制备了 8 个系列的双链体。siRNA 序列和突出核苷酸的数量以 1 nt 的幅度改变。(B) 在 5 nM 平端 dsRNA 存在下，用黑尾果蝇胚胎裂解液中的对照萤光素酶(*Renilla reniformis*, Rr-luc)标准化的靶萤光素酶(*Photinus pyralis*, Pp-luc)的相对发光。在 dsRNA 存在下测定的发光比用缓冲液对照(bu, 黑线)获得的发光比标准化。标准化的发光比小于 1 表明有特异性干涉。(C-J) 8 个系列 21-nt siRNA 双链体的标准干涉比。siRNA 双链体的序列在条图之上显示。每张图显示由特定反义指导 siRNA 和 5 种不同的有义 siRNA 形成的一组双链体的干涉比。突出核苷酸的数量(3'突出端, 正数; 5'突出端, 负数)在 x 轴上标出。数据点是至少 3 次独立实验的平均值，误差条代表标准差。

图 12: siRNA 双链体有义链长度的变化。

(A) 实验图示。3 条 21-nt 反义链与 8 个有义 siRNA 配对。这些 siRNA 的 3' 端长度改变。反义 siRNA 的 3' 突出端为 1-nt (B)、2-nt (C) 或 3-nt (D)，而每个系列的有义 siRNA 突出端不同。标出了 siRNA 双链体的序列和相应的干涉比。

图 13: 保留 2-nt 3'突出端的 siRNA 双链体的长度变化。

(A) 实验图示。21-nt siRNA 双链体在序列上与图 11H 或图 12C 所示相同。这些 siRNA 双链体延伸到有义 siRNA 的 3' 端(B)或有义 siRNA 的 5' 端(C)。标出了 siRNA 双链体的序列和各自的干涉比。

图 14: siRNA 核糖残基的 2'-羟基的置换。

将 siRNA 双链体链中的 2'-羟基(OH)置换为 2'-脱氧(d)或 2'-O-甲基(Me)。3' 端的 2-nt 和 4-nt 2'-脱氧置换分别表示为 2-nt d 和 4-nt d。尿苷残基置换为 2'-脱氧胸苷。

图 15: 具有 2-nt 3'突出端的 21-nt siRNA 双链体对有义和反义靶 RNA 切割的作图。

(A) ^{32}P (星号)帽标记的有义和反义靶 RNA 和 siRNA 双链体的图示。有义和反义靶 RNA 切割的位置分别在 siRNA 双链体之上和之下用三角形标出。(B) 靶 RNA 切割位点的作图。10 nM 靶标与 100 nM siRNA 双链体在黑尾果蝇胚胎裂解液中温育 2 小时后，在测序凝胶上分析 5'帽标记的底物和 5'切割产物。长度标记物通过靶 RNA 的部分 RNase T1 消化(T1)和部分碱性水解(OH-)产生。图左侧的粗线表示与靶标同向的 siRNA 链 1 和 5 所覆盖的区域。

图 16: 指导 siRNA 的 5'端决定靶 RNA 切割位置。

(A,B) 实验策略图示。所有 siRNA 双链体中的反义 siRNA 相同，但是通过改变 3'端有义链有 18-25 nt 的不同(A)，或者通过改变 5'端有 18-23 nt 的不同(B)。有义和反义靶 RNA 切割位置分别在 siRNA 双链体之上和之下用三角形标出。(C,D) 利用帽标记的有义(上图)或反义(下图)靶 RNA 对靶 RNA 切割的分析。只显示帽标记的 5'切割产物。标出了 siRNA 双链体的序列，有义 siRNA 链的长度在图上部标出。图(C)中用破折号标记的对照道显示在不含 siRNA 情况下温育的靶 RNA。标记物如图 15 所述。(D)下图中的箭头指出相差 1 nt 的靶 RNA 切割位点。

图 17: siRNA 双链体 3'突出端的序列变异。

2-nt 3'突出端(NN, 灰色)如图所示在序列和组成上有变化(T, 2'-脱氧胸昔, dG, 2'-脱氧鸟昔, 星号, 野生型 siRNA 双链体)。标准化的干涉比如图 11 所述确定。野生型序列与图 14 所示相同。

图 18: 靶识别的序列特异性。

显示了错配 siRNA 双链体的序列，修饰的序列片段或单核昔酸以灰色在下面列出。参照双链体(ref)和 siRNA 双链体 1-7 含有 2'-脱氧胸昔 2-nt 突出端。胸昔修饰的参照双链体的沉默率与野生型序列相当(图 17)。标准化的干涉比如图 11 所述确定。

图 19: 保留 2-nt 3'突出端的 siRNA 双链体的长度变化。

siRNA 双链体延伸至有义 siRNA 的 3'侧(A)或有义 siRNA 的 5'侧(B)。标出了 siRNA 双链体的序列和各自的干涉比。对于 HeLa SS6 细

胞，靶向 GL2 萤光素酶的 siRNA 双链体($0.84 \mu\text{g}$)用 pGL2-对照和 pRL-TK 质粒一起转染。为了进行对比，标出了在黑尾果蝇裂解液中检测到的 siRNA 双链体的体外 RNAi 活性。

实施例 1

合成小 RNA 介导的 RNA 干涉

1.1. 实验方法

1.1.1 体外 RNAi

体外 RNAi 和裂解液的制备如前所述进行(Tuschl 等人, 1999; Zamore 等人, 2000)。使用新溶解的肌酸激酶(Roche)对于最佳 ATP 再生至关重要。在 25°C 下, 经过 15 分钟的延长预温育时间, 用 5nM 的 dsRNA 浓度进行 RNAi 翻译测定(图 1), 之后加入体外转录、加帽并且聚腺苷酸化的 Pp-luc 和 Rr-luc 报道 mRNA。继续温育 1 小时, 利用双萤光素酶测定(Promega)和单光 3010C 发光计(PharMingen)分析 Pp-luc 和 Rr-luc 蛋白的相对含量。

1.1.2 RNA 合成

由携带 T7 或 SP6 启动子序列的 PCR 模板体外转录 RNA 使用标准方法, 例如参见(Tuschl 等人, 1998)。合成的 RNA 用 Expedite RNA 亚磷酰胺(Proligo)制备。3'接头寡核苷酸用二甲氧基三苯甲基-1,4-苯二甲醇-琥珀酰-氨基丙基-CPG 合成。寡核糖核苷酸在 $3 \text{ ml } 32\%$ 氨/乙醇(3/1)中 55°C 去保护 4 小时(Expedite RNA)或 55°C 16 小时(3'和 5'接头 DNA/RNA 嵌合寡核苷酸), 然后脱硅烷化, 如前所述凝胶纯化(Tuschl 等人, 1993)。包含长 3'突出端的 dsRNA 制品的 RNA 转录物由有义方向含有 T7 启动子、反义方向含有 SP6 启动子的 PCR 模板产生。用 GCGTAATAACGACTCACTATAGAACAAATTGCTTTACAG (下划线标出 T7 启动子) 作为 5' 引物, 用 ATTTAGGTGACACTATAGGCATAAAGAATTGAAGA (下划线标出 SP6 启动子) 作为 3'引物, 用线性化 Pp-luc 质粒(pGEM-luc 序列)(Tuschl 等人, 1999)作为模板, PCR 扩增有义和反义靶 RNA 的转录模板; T7

转录的有义 RNA 为 177 nt 长，在相对于起始密码子的位点 113-273 之间含有 Pp-luc 序列，之后在 3' 端是 SP6 启动子序列的 17 nt 互补序列。通过只含单启动子序列的两种不同 PCR 产物的转录，制备用于形成平端 dsRNA 的转录物。

利用苯酚/氯仿抽提进行 dsRNA 退火。等摩尔浓度的有义和反义 RNA (50 nM-10 μM, 取决于长度和含量) 在 0.3 M NaOAc (pH 6) 中 90°C 温育 30 秒，然后在室温下用等体积的苯酚/氯仿抽提，之后用氯仿抽提除去残余的苯酚。加入 2.5-3 倍体积的乙醇沉淀产生的 dsRNA。沉淀溶解于裂解缓冲液(100 mM KCl, 30 mM HEPES-KOH, pH 7.4, 2 mM Mg(OAc)₂)中，在 1×TAE 缓冲液中通过标准琼脂糖凝胶电泳证实 dsRNA 的质量。具有 17 nt 和 20 nt 3' 突出端的 52 bp dsRNA (图 6) 在 95°C 下温育 1 分钟退火，迅速冷却到 70°C，然后在 3 个小时内缓慢冷却到室温(50 μl 退火反应，1 μM 标准浓度，300 mM NaCl, 10 mM Tris-HCl, pH 7.5)。然后用苯酚/氯仿抽提 dsRNA，乙醇沉淀，溶解于裂解缓冲液中。

用 1 mM ATP、CTP、GTP、0.1 或 0.2 mM UTP 和 0.2-0.3 μM ³²P-UTP (3000 Ci/mmol)，或者用是放射性标记核昔三磷酸各自的比例而不是 UTP 的比例转录用于制备 dsRNA 的内部 ³²P-放射性标记的 RNA (图 2 和图 4)。靶 RNA 的帽子的标记如前所述进行。靶 RNA 在帽子标记后凝胶纯化。

1.1.3 切割位点作图

标准 RNAi 反应如下进行：预温育 10 nM dsRNA 15 分钟，随后加入 10 nM 帽子标记的靶 RNA。再温育 2 小时(图 2A)或 2.5 小时(图 5B 和 6B)后，通过蛋白酶 K 处理终止反应(Tuschl 等人，1999)。然后在 8% 或 10% 测序凝胶上分析样品。21 和 22 nt 合成 RNA 双链体以 100 nM 终浓度使用(图 5B)。

1.1.4 ~21 nt RNA 的克隆

在不含靶 RNA 的情况下，在果蝇裂解液中温育放射性标记的 dsRNA，产生 21 nt RNA (200 μl 反应，1 小时温育，50 nM dsP111，

或 100 nM dsP52 或 dsP39)。然后用蛋白酶 K 处理反应混合物(Tuschl 等人, 1999), 在变性 15%聚丙烯酰胺凝胶上分离 dsRNA 加工产物。切下大小范围至少为 18-24 nt 的条带, 洗脱到 0.3 M NaCl 中, 在 4℃ 下在硅化试管中过夜。通过乙醇沉淀回收 RNA, 去磷酸化(30 μl 反应, 30 分钟, 50℃, 10 U 碱性磷酸酶, Roche)。通过苯酚/氯仿抽提终止反应, 乙醇沉淀 RNA. 3'接头寡核苷酸(pUUUaaccgcattctcx: 大写, RNA; 小写, DNA; p, 磷酸; x, 4-羟甲基苄基)与去磷酸化的~21 nt RNA 连接(20 μl 反应, 30 分钟, 37℃, 5 μM 3'接头, 50 mM Tris-HCl, pH 7.6, 10 mM MgCl₂, 0.2 mM ATP, 0.1 mg/ml 乙酰 BSA, 15% DMSO, 25 U T4 RNA 连接酶, Amersham-Pharmacia)(Pan 和 Uhlenbeck, 1992)。加入等体积的 8 M 尿素/50 mM EDTA 终止混合物终止连接反应, 直接加样到 15%凝胶上。连接产率超过 50%。从凝胶中回收连接产物, 并 5'-磷酸化(20 μl 反应, 30 分钟, 37℃, 2 mM ATP, 5 U T4 多核苷酸激酶, NEB)。通过苯酚/氯仿抽提终止磷酸化反应, 乙醇沉淀回收 RNA。然后, 5'接头(tactaatacgactcaatAAA: 大写, RNA; 小写, DNA)如上所述与磷酸化连接产物连接。凝胶纯化新的连接产物, 在用作载体的反转录引物(GACTAGCTGGAATTCAAGGATGCGGTAAA, 下划线, EcoRI 位点)存在下, 从凝胶切片中洗脱。反转录(15 μl 反应, 30 分钟, 42℃, 150 U Superscript II 反转录酶, Life Technologies)之后用 CAGCCAACGGAATTCATACGACTCACTAAA(下划线, EcoRI 位点)作为 5'引物和 3' RT 引物进行 PCR. 经苯酚/氯仿抽提和乙醇沉淀纯化 PCR 产物。然后用 EcoRI (NEB)消化 PCR 产物, 用 T4 DNA 连接酶(高浓度, NEB)连环化。在低熔点琼脂糖凝胶上分离大小为 200-800 bp 的多联体, 利用标准的熔化和苯酚抽提法从凝胶中回收, 乙醇沉淀。在标准条件下与 Taq 聚合酶 72℃温育 15 分钟, 补平未配对的末端, 用 TOTO TA 克隆试剂盒(Invitrogen)将 DNA 产物与 pCR2.1-TOPO 载体直接连接。利用 PCR 和 M13-20 和 M13 反向测序引物筛选集落。直接对 PCR 产物进行委托测序(Sequence Laboratories Göttingen

GmbH, 德国). 每个克隆平均获得 4-5 条 21mer 序列。

1.1.5 2D-TLC 分析

放射性标记、凝胶纯化的 siRNA 的核酸酶 P1 消化和 2D-TLC 如 (Zamore 等人, 2000) 所述进行。采用 2 $\mu\text{g}/\mu\text{l}$ 载体 tRNA 和 30 U 核糖核酸酶 T2 (Life Technologies), 在 10 mM 乙酸铵(pH 4.5) 中以 10 μl 反应体积进行核酸酶 T2 消化, 50°C 3 小时。非放射性标准的迁移通过 UV 影像法测定。T2 消化产物与使用 $\gamma^{32}\text{P}$ -ATP 和 T4 多核苷酸激酶 经商品核昔 3'-一磷酸的 5'- ^{32}P -磷酸化制备的标准共迁移, 由此证实核昔-3',5'-二磷酸的身份(数据未显示)。

1.2 结果与讨论

1.2.1 dsRNA 加工为 21 和 22 nt RNA 片段的长度要求

由黑尾果蝇合胞体胚胎制备的裂解液概述了体外 RNAi, 提供了一个用于生化分析 RNAi 机制的新型工具(Tuschl 等人, 1999; Zamore 等人, 2000)。RNAi 对 dsRNA 的长度要求的体外和体内分析揭示, 在降解靶 mRNA 上, 短 dsRNA(<150 bp) 比长 dsRNA 效率更低(Caplen 等人, 2000; Hammond 等人, 2000; Ngo 等人, 1998; Tuschl 等人, 1999)。mRNA 降解效率降低的原因尚不清楚。因此我们在最佳条件下, 在果蝇裂解液中, 检查了靶 RNA 降解对 dsRNA 的精确长度要求 (Zamore 等人, 2000)。合成了几个系列的 dsRNA, 它们针对萤火虫萤光素酶(Pp-luc)报道 RNA。用双萤光素酶测定(Tuschl 等人, 1999)监测靶 RNA 表达的特异性抑制(图 1A 和 1B)。对于短至 38 bp 的 dsRNA, 我们检测到靶 RNA 表达的特异性抑制, 但是 29-36 bp 的 dsRNA 在该方法中无效。此作用与靶位点无关, 与 dsRNA 长度(即长 dsRNA)相关的 Pp-luc mRNA 表达的抑制程度比短 dsRNA 更有效。

曾经提出, dsRNA 加工产生的 21-23 nt RNA 片段是 RNA 干涉和共抑制的介体(Hamilton 和 Baulcombe, 1999; Hammond 等人, 2000; Zamore 等人, 2000)。因此我们分析了大小为 501-29 bp 的 dsRNA 亚组的 21-23 nt 片段形成率。对于 39-501 bp 长 dsRNA, 容易地检测到果蝇裂解液中 21-23 nt 片段的形成(图 2), 但是 29 bp dsRNA 明显延迟。

这一发现与 21-23 nt 片段在指导 mRNA 切割上的作用一致，为 RNAi 缺乏 30 bp dsRNA 提供了一种解释。21-23 mer 形成的长度信赖性可能反映了一种生物学相关的控制机制，其通过规则细胞 RNA 的分子内碱基配对的短结构阻止不希望的 RNAi 激活。

1.2.2 39 bp dsRNA 介导靶 RNA 在单个位点处的切割

向果蝇裂解液中加入 dsRNA 和 5'-加帽的靶 RNA 导致靶 RNA 的序列特异性降解(Tuschl 等人, 1999)。靶 mRNA 只在与 dsRNA 相同的区域内切割，许多靶切割位点相隔 21-23 nt (Zamore 等人, 2000)。因此，特定 dsRNA 的切割位点数预计大约对应于 dsRNA 的长度除以 21。我们对在帽子处 5' 放射性标记的有义和反义靶 RNA 上的靶切割位点作图(Zamore 等人, 2000)(图 3A 和 3B)。在测序凝胶上分离稳定的 5' 切割产物，通过与来自靶 RNA 的部分 RNase T1 和碱性水解梯度比较，确定切割位点。

与以前的发现(Zamore 等人, 2000)一致，所有靶 RNA 切割位点都位于与 dsRNA 相同的区域内。有义或反义靶标只被 39 bp dsRNA 切割一次。每个切割位点都与 dsRNA 覆盖的区域的 5' 端相距 10 nt (图 3B)。与 39 bp dsRNA 有相同 5' 端的 52 bp dsRNA 除了位于第一个位点下游 23 和 24 nt 的两个较弱切割位点之外，在有义靶标上还产生相同的切割位点，距与 dsRNA 相同的区域的 5' 端 10 nt. 反义靶标只切割一次，与各自 dsRNA 覆盖的区域的 5' 端相距 10 nt. 图 1 所示的 38-49 bp dsRNA 的切割位点作图显示，第一个和主要的切割位点总是位于 dsRNA 覆盖的区域的 7-10 nt 下游(数据未显示)。这提示，靶 RNA 切割点取决于 dsRNA 末端，可能意味着加工为 21-23 mer 从双链体末端开始。

较长的 111 bp dsRNA 在有义和反义靶标上的切割位点比预期的更多，大多数似乎以 20-23 nt 的间隔聚簇(图 3A 和 3B)。对于较短的 dsRNA，在有义靶标上的第一个切割位点与 dsRNA 覆盖的区域的 5' 端相距 10 nt，在反义靶标上的第一个切割位点与 dsRNA 覆盖的区域的 5' 端相距 9 nt. 还不清楚是什么导致这种混乱的切割，但是一种可

能性是，较长的 dsRNA 不仅可以从末端开始加工，而且也可在内部加工，或者存在我们还不清楚的一些 dsRNA 加工的特异性决定簇。以前也注意到 21-23 nt 间隔的不规则性(Zamore 等人, 2000)。为了更好地了解 dsRNA 加工和靶 RNA 识别的分子基础，我们决定分析果蝇裂解液中 39、52 和 111 bp dsRNA 加工产生的 21-23 nt 片段的序列。

1.2.3 dsRNA 通过 RNase III 样机制加工为 21 和 22 nt RNA

为了表征 21-23 nt RNA 片段，我们检查了该 RNA 片段的 5' 和 3' 端。对凝胶纯化的 21-23 nt RNA 进行高碘酸盐氧化，随后进行 β -排除，表明存在末端 2' 和 3' 羟基。21-23 mer 对碱性磷酸酶处理也有反应性，表明存在 5' 端磷酸基团。5' 磷酸和 3' 羟基端的存在提示，dsRNA 能被类似于大肠杆菌 RNase III 的酶活性加工(综述见 Dunn, 1982; Nicholson, 1999; Robertson, 1990; Robertson, 1982)。

利用 T4 RNA 连接酶将 3' 和 5' 接头寡核苷酸与纯化的 21-23 mer 连接，进行 21-23 nt RNA 片段的定向克隆。反转录连接产物，PCR 扩增，连环化，克隆并测序。从 39、52 和 111 bp dsRNA 的 dsRNA 加工反应中测序超过 220 个短 RNA(图 4A)。我们发现下列长度分布：1% 18 nt, 5% 19 nt, 12% 20 nt, 45% 21 nt, 28% 22 nt, 6% 23 nt, 2% 24 nt。对加工片段 5' 端核苷酸的序列分析表明，含有 5' 鸟苷的寡核苷酸被低估。这种情况最可能是由于加入了 T4 RNA 连接酶，该酶将 5' 磷酸化鸟苷识别为供体寡核苷酸；在 3' 端未见明显的序列偏差。来源于双链体有义或反义链 3' 端的许多~21 nt 片段含有在使用 T7 RNA 聚合酶合成 RNA 过程中非模板添加核苷酸产生的 3' 核苷酸。有意思的是，也克隆了相当数量的内源果蝇~21 nt RNA，其中一些来自 LTR 和非 LTR 反转录转座子(数据未显示)。这与 RNAi 在转座子沉默中的作用一致。

~21 nt RNA 存在于覆盖完整 dsRNA 序列的聚簇基团中(图 4A)。显然，加工反应通过产生交错的 3' 末端切割 dsRNA，这是 RNase III 切割的另一个特征。对于 39 bp dsRNA，在每条包含突出 3' 端的 dsRNA 组成链中发现两簇~21 nt RNA，而在有义和反义靶标上只检测到一个

切割位点(图 3A 和 3B)。如果~21 nt 片段在介导 mRNA 降解的复合物中作为单链指导 RNA 存在, 可以假定至少存在两个靶切割位点, 但这不是事实。这提示, ~21 nt RNA 可能以双链形式存在于内切核酸酶复合物中, 但是只有一条链能用于靶 RNA 识别和切割。根据~21 nt 双链体与核酸酶复合物结合的方向可以简单地确定~21 nt 链之一在靶切割中的用途。该方向取决于原始 dsRNA 的加工方向。

与 39 bp dsRNA 相比, 52 bp 和 111 bp dsRNA 的~21 mer 簇尚未确定。这些簇覆盖 25-30 nt 的区域, 最可能代表~21 nt 双链体的几个不同的亚群, 因此指导在几个邻近位点的靶切割。这些切割区仍然主要有 20-23 nt 的间隔。决定怎样规则的 dsRNA 才能被加工为~21 nt 片段的原则尚不清楚, 但是以前发现, 使用尿苷能改变切割位点的大约 21-23 nt 的间隔(Zamore 等人, 2000)。大肠杆菌 RNase III 切割 dsRNA 的特异性似乎主要由抗决定簇控制, 即, 在相对于切割位点的特定位置处排除某些特异碱基对(Zhang 和 Nicholson, 1997)。

为了检测加工的~21 nt RNA 片段中存在糖-、碱基-还是帽-修饰, 我们在裂解液中温育放射性标记的 505 bp Pp-luc dsRNA 1 小时, 分离~21 nt 产物, 用 P1 或 T2 核酸酶消化为单核苷酸。然后用 2D 薄层层析法分析核苷酸混合物(图 4B)。P1 或 T2 消化显示, 4 种天然核糖核苷酸都没有修饰。我们以前曾经分析了~21 nt 片段中的腺苷-肌苷转化(温育 2 小时后), 检测到程度较小(<0.7%)的脱氨基作用(Zamore 等人, 2000); 在裂解液中温育较短时间(1 小时)使肌苷部分减至几乎无法检测的水平。可切割磷酸二酯键 3' 的 RNase T2 产生核苷 3'-磷酸和核苷 3',5'-二磷酸, 因而表明存在 5' 端一磷酸。检测到全部 4 种核苷 3',5'-二磷酸, 提示核苷酸间键的切割没有或只有较低的序列特异性。总之, ~21 nt 片段没有修饰, 是由 dsRNA 产生的, 因此在 5' 端存在 5'-一磷酸和 3'-羟基。

1.2.4 合成 21 和 22 nt RNA 介导靶 RNA 切割

对 dsRNA 加工产物的分析表明, ~21 nt 片段是通过具有 RNase III 切割反应所有特征的一种反应产生的(Dunn, 1982; Nicholson, 1999;

Robertson, 1990; Robertson, 1982)。RNase III 在 dsRNA 的两条链上产生两个交错的切口，形成大约 2 nt 的 3' 突出端。我们化学合成了 21 和 22 nt RNA，它们在序列上与某些克隆的~21 nt 片段相同，并且检测了它们介导靶 RNA 降解的能力(图 5A 和 5B)。21 和 22 nt RNA 双链体在裂解液中以 100 nM 浓度温育，浓度比 52 bp 对照 dsRNA 高 10 倍。在这种条件下，易于检测到靶 RNA 切割。21 和 22 nt 双链体的浓度从 100 nM 降到 10 nM 仍能引起靶 RNA 切割。然而，双链体浓度从 100 nM 提高到 1000 nM 不能进一步提高靶切割，这可能是由于裂解液内存在一种限制蛋白因子。

与不介导 RNAi 的 29 或 30 bp dsRNA 不同，具有 2-4 nt 突出 3' 端的 21 和 22 nt dsRNA 介导靶 RNA 的高效降解(双链体 1、3、4、6，图 5A 和 5B)。平端 21 或 22 nt dsRNA(双链体 2、5、7，图 5A 和 5B)降解靶标的能力降低，表明突出的 3' 端对于 RNA-蛋白质核酸酶复合物的重建至关重要。~21 nt 双链体与蛋白质成分的高亲和力结合可能需要单链突出端。5' 端磷酸尽管在 dsRNA 加工后存在，但不是介导靶 RNA 切割所需的，在合成的短 RNA 中缺乏。

合成 21 和 22 nt 双链体指导有义及反义靶标在短双链体覆盖的区域内切割。这是一个重要结果，因为形成两对~21 nt 片段簇的 39 bp dsRNA (图 2)只切割有义或反义靶标一次而不是两次。我们提出~21 nt 双链体的两条链只有一条能指导靶 RNA 切割，核酸酶复合物中~21 nt 双链体的方向取决于 dsRNA 加工的最初方向，从而解释了这一结果。然而，向体外系统输送已经加工好的~21 nt 双链体可以以对称 RNA 双链体的两个可能方向形成活性序列-特异性核酸酶复合物。这导致有义及反义靶标在与 21 nt RNA 双链体相同的区域内切割。

靶切割位点位于与 21 或 22 nt 指导序列互补的第一个核苷酸的 11 或 12 nt 下游，即切割位点靠近 21 或 22 nt RNA 所覆盖的区域的中心(图 4A 和 4B)。替换 22 nt 双链体有义链的两个核苷酸(比较图 5A 中的双链体 1 和 3)只将反义靶标的切割位点替换了两个核苷酸。替换有义和反义链的两个核苷酸使切割位点位移两个核苷酸(比较双链体 1 和

4). 我们预测能设计一对 21 或 22 nt RNA，它们几乎在任何特定位点均能切割靶 RNA。

21 和 22 nt RNA 指导的靶 RNA 切割的特异性似乎是强烈的，因为未检测到异常的切割位点(图 5B)。然而应当指出，21 和 22 nt RNA 双链体 3' 突出端存在的核苷酸在底物识别方面的作用可能比靠近切割位点的核苷酸要小。这是基于以下发现：活性双链体 1 或 3 的 3' 突出端中的 3' 核苷酸(图 5A)不与靶标互补。现在利用 21 和 22 nt RNA 能容易地进行 RNAi 特异性的详细分析。

根据具有突出 3' 端的 21 和 22 nt RNA 介导 RNA 干涉的证据，我们提出将~21 nt RNA 命名为“短干涉 RNA”或 siRNA，将各自的 RNA-蛋白质复合物命名为“小干涉核糖核蛋白颗粒”或 siRNP。

1.2.5 短 dsRNA 上的 20 nt 3' 突出端抑制 RNAi

我们表示，平端短 dsRNA 似乎从 dsRNA 的末端开始加工。在我们对 RNAi 中 dsRNA 长度依赖性的研究过程中，我们也分析了具有 17-20 nt 突出 3' 端的 dsRNA，我们吃惊地发现它们还不如平端 dsRNA 有效。对于可达 100 bp 的 dsRNA，长 3' 端的抑制作用特别明确，但是更长的 dsRNA 就没有这么明显。根据天然凝胶分析(数据未显示)，这种作用是由于 dsRNA 形成不好。我们检验了长突出 3' 端的抑制作用是否能作为将 dsRNA 加工只定向于短 RNA 双链体两端之一的工具。

我们合成了 52 bp 模型 dsRNA 的 4 种组合：平端、只在有义链上的 3' 延伸、只在反义链上的 3' 延伸和在两条链上的双 3' 延伸，并在裂解液中温育后对靶 RNA 切割位点作图(图 6A 和 6B)。当双链体反义链的 3' 端延伸时，有义靶标的第一个和主要的切割位点丢失，反之亦然，当双链体有义链的 3' 端延伸时，反义靶标的强切割位点丢失。两条链的 3' 延伸使 52 bp dsRNA 实际上失活。~20 nt 3' 延伸使 dsRNA 失活的一种解释是单链 RNA-结合蛋白的结合，这能干扰 dsRNA 加工因子之一在该末端的结合。该结果也与我们的模型一致，在我们的模型中，在装配的 siRNP 中只有 siRNA 双链体的一条链能指导靶 RNA 切割。指导 RNA 切割的链的方向取决于 dsRNA 加工反应的方向。具有 3' 交

错末端可能有利于加工复合物的装配。在有义链 3'端的阻断只允许从反义链的相对 3'端开始 dsRNA 加工。随后产生 siRNP 复合物，其中只有 siRNA 双链体的反义链能指导有义靶 RNA 切割。情况相反时同样如此。

至于较长的 dsRNA (≥ 500 bp, 数据未显示)，长 3'延伸的抑制作用不明显，这提示我们，长 dsRNA 也可能含有内部 dsRNA 加工信号，或者可能由于多种切割因子的结合而协同加工。

1.2.6 dsRNA 指导的 mRNA 切割的模型

新的生物化学数据更新了 dsRNA 如何靶向 mRNA 进行破坏的模型(图 7)。双链 RNA 首先加工为主要长 21 和 22 nt 的短 RNA 双链体，其具有类似于 RNase III 样反应的交错 3'端(Dunn, 1982; Nicholson, 1999; Robertson, 1982)。根据加工 RNA 片段的 21-23 nt 长度，曾经推测 RNase III 样活性可能参与 RNAi (Bass, 2000)。在 RNase III 反应产物中发现，在 siRNA 的末端存在 5'磷酸和 3'羟基，这进一步支持了这一假说(Dunn, 1982; Nicholson, 1999)。细菌 RNase III 和真核同源物——酿酒酵母中的 Rnt1p 和粟酒酵母中的 Pac1p 在核糖体 RNA 和 snRNA 和 snoRNA 的加工中起作用(参见，例如 Chanfreau 等人, 2000)。

关于来自植物、动物或人类的 RNase III 同源物的生物化学，人们知之甚少。主要通过数据库指导的序列分析或 cDNA 的克隆鉴定了两个 RNase III 酶家族。第一个 RNase III 家族的代表是 1327 个氨基酸长的黑尾果蝇蛋白 drosha (Acc. AF116572)。C 端由两个 RNase III 和一个 dsRNA 结合域组成，N 端功能未知。在 *C. elegans* (Acc. AF160248) 和人类(Acc. AF189011)中也发现了密切同源物(Filippov 等人, 2000; Wu 等人, 2000)。drosha 样人 RNase III (Wu 等人, 2000) 最近已经克隆并表征。该基因在人类组织和细胞系中普遍表达，该蛋白质位于细胞的核和核仁中。根据反义抑制研究推断的结果，提出了该蛋白质在 rRNA 加工中的作用。第二类的代表是编码 1822 个氨基酸长的蛋白质的 *C. elegans* 基因 K12H4.8 (Acc. S44849)。该蛋白质含有

一个 N 端 RNA 解旋酶基序，之后是 2 个 RNase III 催化域和一个 dsRNA 结合基序，类似于 *drosha* RNase III 家族。在粟酒酵母(Acc. Q09884)、*A. thaliana* (Acc. FA187317)、黑尾果蝇(Acc. AE003740)和人类(Acc. AB028449)中具有密切同源物(Filippov 等人, 2000; Jacobsen 等人, 1999; Matsuda 等人, 2000)。K12H4.8 RNase III/解旋酶可能是参与 RNAi 的候选物。

在 *C. elegans* 中进行遗传筛选鉴定了 rde-1 和 rde-4，它们是 RNAi 激活所必需的，对转座子移动或共抑制没有影响(Dernburg 等人, 2000; Grishok 等人, 2000; Ketting 和 Plasterk, 2000; Tabara 等人, 1999)。这引出一种假说：这些基因对于 dsRNA 加工非常重要，但是不参与 mRNA 靶降解。这两种基因的功能仍然未知，rde-1 基因产物是类似于兔蛋白 eIF2C 的蛋白质家族的一员(Tabara 等人, 1999)，rde-4 的序列还未有描述。将来对这些蛋白质的生物化学表征将揭示其分子功能。

加工为 siRNA 双链体似乎从平端 dsRNA 或具有短(1-5 nt) 3' 突出端的 dsRNA 的末端开始，以大约 21-23 nt 的幅度加工。短 dsRNA 上的长(~20 nt) 3' 交错末端抑制 RNAi，这可能是通过与单链 RNA 结合蛋白相互作用。在短 dsRNA 侧翼的单链区抑制 RNAi，以及短 30 bp dsRNA 缺乏 siRNA 形成，可以解释 mRNA 中常见的结构区为什么不导致 RNAi 激活。

不希望被理论所束缚，我们假设 dsRNA 加工蛋白或其亚组在加工反应后仍与 siRNA 双链体结合。siRNA 双链体相对于这些蛋白质的方向决定两条互补链的哪一条在指导靶 RNA 降解中起作用。化学合成的 siRNA 双链体指导有义及反义靶 RNA 的切割，因为它们能以两种可能的方向之一与蛋白质成分结合。

合成 21 和 22 nt siRNA 双链体能用于有效的 mRNA 降解，这一重要发现为功能基因组学以及生物医学研究中基因表达的序列特异性调节提供了新的工具。在由于 PKR 反应激活而不能使用长 dsRNA 的哺乳动物系统中，siRNA 可能有效(Clemens, 1997)。因此，siRNA 双

链体是反义或核酶治疗剂的一种新的备选剂。

实施例 2

人组织培养物中的 RNA 干涉

2.1 方法

2.1.1 RNA 制备

用 Expedite RNA 亚磷酰胺和胸昔亚磷酰胺(Proligo, 德国)化学合成 21 nt RNA。合成的寡核苷酸去保护并凝胶纯化(实施例 1), 之后经 Sep-Pak C18 柱(Waters, Milford, MA, USA)纯化(Tuschl, 1993)。针对 GL2 (Acc. X65324)和 GL3 萤光素酶(Acc. U47296)的 siRNA 序列对应于相对于起始密码子第一个核苷酸的编码区 152-173, 针对 RL (Acc. AF025846)的 siRNA 对应于起始密码子之后的区 119-129. 较长的 RNA 用 T7 RNA 聚合酶由 PCR 产物转录, 然后用凝胶和 Sep-Pak 纯化。49 和 484 bp GL2 或 GL3 dsRNA 分别对应于相对于翻译起点的位点 113-161 和 113-596. 50 和 501 bp RL dsRNA 分别对应于位点 118-167 和 118-618. 针对人源化 GFP (hG)的 dsRNA 合成的 PCR 模板从 pAD3 (Kehlenbach, 1998)中扩增, 因此 50 和 501 bp hG dsRNA 分别对应于相对起始密码子的位点 118-167 和 118-618.

为使 siRNA 退火, 20 μ M 单链在退火缓冲液(100 mM 乙酸钾, 30 mM HEPES-KOH pH 7.4, 2 mM 乙酸镁)中 90℃ 温育 1 分钟, 然后 37℃ 1 小时。对于 50 和 500 bp dsRNA, 37℃ 温育步骤延长过夜, 退火反应分别以 8.4 μ M 和 0.84 μ M 的链浓度进行。

2.1.2 细胞培养

S2 细胞在 25℃ 下在补充了 10% FBS、100 单位/ml 青霉素和 100 μ g/ml 链霉素的 Schneider's 果蝇培养基(Life Technologies)中繁殖。293、NIH/3T3、HeLa S3、COS-7 细胞在 37℃ 下在补充了 10% FBS、100 单位/ml 青霉素和 100 μ g/ml 链霉素的 Dulbecco's 改良 Eagle's 培养基中培养。细胞定期传代, 以保持指数生长。转染前 24 小时, 在大约 80% 平铺时, 哺乳动物细胞用胰蛋白酶消化, 用不含抗生素的新鲜

培养基 1:5 稀释($1\text{-}3 \times 10^5$ 细胞/ml), 转移到 24 孔板中(500 μl /孔)。S2 细胞在分裂前不用胰蛋白酶消化。如厂商对于粘附细胞系所述, 用 Lipofectamine 2000 试剂(Life Technologies)进行转染。每孔加入配制于脂质体中的 1.0 μg pGL2-对照(Promega)或 pGL3-对照(Promega)、0.1 μg pRL-TK (Promega)和 0.28 μg siRNA 双链体或 dsRNA; 终体积为每孔 600 μl 。细胞在转染 20 小时后温育, 此后似乎健康。然后用双萤光素酶测定(Promega)监测萤光素酶表达。在 1.1 μg hGFP 编码 pAD3 和 0.28 μg invGL2 inGL2 siRNA 共转染后, 通过对哺乳动物细胞系的荧光显微镜检, 测定转染率, 为 70-90%。报道质粒在 XL-1 Blue (Stratagene) 中扩增, 用 Qiagen EndoFree Maxi 质粒试剂盒纯化。

2.2 结果与讨论

为了检验 siRNA 是否也能在组织培养中介导 RNAi, 我们合成了具有对称 2 nt 3' 突出端的 21 nt siRNA 双链体, 它们针对编码海肾 (*Renilla reniformis*) 和萤火虫 (*Photinus pyralis*, GL2 和 GL3) 萤光素酶两种序列变体的报道基因(图 8a,b)。siRNA 双链体与报道质粒组合 pGL2/pRL 或 pGL3/pRL 利用阳离子脂质体共转染黑尾果蝇 Schneider S2 细胞或哺乳动物细胞。在转染 20 小时后测定萤光素酶活性。在检测的所有细胞系中, 我们在同源 siRNA 双链体存在下观察到报道基因表达特异性减少(图 9a-j)。显然, 绝对萤光素酶表达水平不受非同源 siRNA 影响, 表明 21 nt RNA 双链体没有有害的副作用(对于 HeLa 细胞, 例如图 10a-d)。在黑尾果蝇 Schneider S2 细胞中(图 9a,b), 萤光素酶完全特异性抑制。在报道基因表达强 50-100 倍的哺乳动物细胞中, 特异性抑制不完全(图 9c-j)。同源 siRNA 致使 GL2 表达减少 3-12 倍, GL3 表达减少 9-25 倍, RL 表达减少 1-3 倍。对于 293 细胞, RL siRNA 对 RL 萤光素酶的靶向无效, 但是 GL2 和 GL3 靶特异性反应(图 9i,j)。293 细胞中没有 RL 表达减少可能是由于其表达比检测的其它任何哺乳动物细胞系高 5-20 倍, 和/或由于 RNA 二级结构或结合蛋白质, 靶序列的可接近性受限。然而, 同源 siRNA 双链体对 GL2 和 GL3 萤光素酶的特异靶向表明 RNAi 在 293 细胞中也起作用。

除了 uGL2 之外，所有 siRNA 双链体的 2 nt 3' 突出端都由(2'-脱氧)胸苷组成。3' 突出端中尿苷置换为胸苷在黑尾果蝇体外系统中良好耐受，突出端序列对靶识别而言并不重要。选择胸苷突出端，因为推断它能提高组织培养基中和转染细胞内 siRNA 的核酸酶抗性。实际上，在检测的所有细胞系中，胸苷修饰的 GL2 siRNA 略强于未修饰的 uGL2 siRNA(图 9a, c, e, g, i)。可以想象，进一步修饰 3' 突出端核苷酸可能对 siRNA 双链体的输送和稳定性有其它好处。

在共转染实验中，对于组织培养基的终体积，使用 25 nM siRNA 双链体(图 9, 10)。siRNA 浓度提高为 100 nM 不能提高特异性沉默作用，但是由于质粒 DNA 与 siRNA 之间脂质体包封的竞争，开始影响转染率(数据未显示)。siRNA 浓度降至 1.5 nM 不会降低特异性沉默作用(数据未显示)，即使 siRNA 只比 DNA 质粒浓缩 2-20 倍。这表明，siRNA 是介导基因沉默的非常有效的试剂，在浓度比常规反义或核酶基因靶向实验使用的浓度低几个数量级时，siRNA 是有效的。

为了监测较长 dsRNA 对哺乳动物细胞的作用，制备了与报道基因同源的 50 和 500 bp dsRNA。用来自人源化 GFP (hG)的 dsRNA 作为非特异性对照(Kehlenbach, 1998)。当 dsRNA 与 siRNA 双链体含量相同地(不浓缩)共转染时，报道基因的表达强烈且非特异地降低。以 HeLa 细胞为代表性实例说明了这一结果(图 10a-d)。50 bp dsRNA 共转染使绝对萤光素酶活性非特异地降低 10-20 倍，500 bp dsRNA 共转染使其降低 20-200 倍。对于 COS-7 和 NIH/3T3 细胞也观察到类似的非特异性作用。对于 293 细胞，只有用 500 bp dsRNA 才观察到 10-20 倍非特异性降低。>30 bp dsRNA 使报道基因表达非特异性降低被认为是干扰素反应的一部分。

奇怪的是，尽管报道基因表达强烈地非特异性降低，但我们可以重复地检测到其它序列特异的、dsRNA 介导的沉默。然而，只有在相对报道基因活性用 hG dsRNA 对照标准化时，特异性沉默作用才明显(图 10e, f)。在检测的其它 3 种哺乳动物细胞系中也观察到同源 dsRNA 引起的 2-10 倍的特异性降低(数据未显示)。dsRNA (356-1662 bp)的特异

性沉默作用以前在 CHO-K1 细胞中有报道，但是检测 2-4 倍特异性降低所需的 dsRNA 的量比我们的实验大约高 20 倍(Ui-Tei, 2000)。CHO-K1 细胞的干扰素反应似乎也有缺陷。在另一篇报告中，利用萤光素酶/lacZ 报道分子组合和 829 bp 特异性 lacZ 或 717 bp 非特异性 GFP dsRNA 检测 293、NIH/3T3 和 BHK-21 细胞的 RNAi (Caplen, 2000)。在这种情况下无法检测 RNAi 可能是由于萤光素酶/lacZ 报道分子测定法的低敏感性和靶标与对照 dsRNA 的长度差异。我们的结果综合起来表明，RNAi 在哺乳动物细胞中有活性，但是，如果干扰素系统被>30 bp 的 dsRNA 激活，则沉默作用难以检测。

总之，我们第一次在哺乳动物细胞中证明了 siRNA 介导的基因沉默。短 siRNA 的应用在人类组织培养的基因功能灭活和基因特异治疗剂的研制中具有广阔前景。

实施例 3

RNA 干涉对基因表达的特异性抑制

3.1 材料与方法

3.1.1 RNA 制备和 RNAi 测定

化学 RNA 合成、退火和基于萤光素酶的 RNAi 测定如实施例 1 或 2 或以前的公开文献所述进行(Tuschl 等人, 1999; Zamore 等人, 2000)。所有 siRNA 双链体都针对萤火虫萤光素酶，萤光素酶 mRNA 序列来源于如(Tuschl 等人, 1999)所述的 pGEM-luc (GenBank acc. X65316)。在加入 mRNA 之前，siRNA 双链体在黑尾果蝇 RNAi/翻译反应液中温育 15 分钟。基于翻译的 RNAi 测定至少一式三份进行。

为了对有义靶 RNA 切割作图，产生 177-nt 转录物，其对应于在相对于起始密码子的位点 113-273 之间的萤火虫萤光素酶序列，随后为 SP6 启动子序列的 17-nt 互补序列。为了对反义靶 RNA 切割作图，由模板产生 166-nt 转录物，它是利用 5' 引物 TAATACGACTCACTATAGAGGCCATATCGTTTCATA (下划线标出 T7 启动子) 和 3' 引物 AGAGGGATGGAACCGCTGG 由质粒序列 PCR

扩增的。靶序列对应于在相对于起始密码子的位点 50-215 之间的萤火虫萤光素酶序列的互补序列。鸟苷酰基转移酶标记如前所述进行 (Zamore 等人, 2000)。为了对靶 RNA 切割作图, 在标准条件下 (Zamore 等人, 2000), 100 nM siRNA 双链体与 5-10 nM 靶 RNA 在黑尾果蝇胚胎裂解液中 25°C 温育 2 小时。加入 8 倍体积的蛋白酶 K 缓冲液 (200 mM Tris-HCl pH 7.5, 25 mM EDTA, 300 mM NaCl, 2% w/v 十二烷基硫酸钠) 终止反应。加入蛋白酶 K (E.M. Merck, 溶于水) 至终浓度为 0.6 mg/ml。反应液然后在 65°C 下温育 15 分钟, 用苯酚/氯仿/异戊醇 (25:24:1) 抽提, 用 3 倍体积的乙醇沉淀。样品置于 6% 测序凝胶上。长度标准通过帽标记的有义或反义靶 RNA 的部分 RNase T1 消化和部分碱水解产生。

3.2 结果

3.2.1 21 nt siRNA 双链体中 3' 突出端的变异

如上所述, 位于 siRNA 双链体 3' 末端的 2 个或 3 个未配对核苷酸在靶 RNA 降解上比相应的平端双链体更有效。为了更全面地分析末端核苷酸的功能, 我们合成了 5 种 21 nt 有义 siRNA, 每一个均用相对于靶 RNA 的一个核苷酸展示, 而且合成了 8 种 21 nt 反义 siRNA, 每一个均用相对于靶标的一个核苷酸展示(图 11A)。通过组合有义和反义 siRNA, 产生 8 个系列的 siRNA 双链体, 它们具有合成的突出末端, 覆盖 7 nt 3' 突出端到 4 nt 5' 突出端的范围。用双萤光素酶测定系统测定 siRNA 双链体的干涉(Tuschl 等人, 1999; Zamore 等人, 2000)。siRNA 双链体针对萤火虫萤光素酶 mRNA, 海肾萤光素酶 mRNA 作为内部控制。在 siRNA 双链体存在下测定靶标与对照萤光素酶活性的发光比, 用不含 dsRNA 时的发光比标准化。为了比较, 在图 11B 中显示长 dsRNA (39-504 bp) 的干涉比。长 dsRNA 在 5 nM 浓度下(图 11A), siRNA 双链体在 100 nM 浓度下(图 11C-J) 测定干涉比。选择 100 nM 的 siRNA 浓度, 因为 5 nM 504 bp dsRNA 的完全加工将产生 120 nM 总 siRNA 双链体。

21 nt siRNA 双链体介导 RNAi 的能力依赖于突出核苷酸或形成的

碱基对的数量。含有 4-6 个 3' 突出核苷酸的双链体不能介导 RNAi (图 11C-F)，含有 2 个或 2 个以上 5' 突出核苷酸的双链体同样不能(图 11G-J)。具有 2 nt 3' 突出端的双链体在介导 RNA 干涉上最有效，但是沉默率也是序列依赖性的，具有 2 nt 3' 突出端的不同 siRNA 双链体观察到高达 12 倍的差异(比较图 11D-H)。具有平端、1 nt 5' 突出端或 1-3 nt 3' 突出端的双链体有时有功能。具有 7 nt 3' 突出端的 siRNA 双链体所观察到的较小的沉默作用(图 11C)可能是由于长 3' 突出端的反义作用，而不是由于 RNAi. 长 dsRNA (图 11B)与最有效的 21 nt siRNA 双链体(图 11E, G, H)之间 RNAi 效率的比较表明，100 nM 浓度的单 siRNA 双链体能与 5 nM 504 bp dsRNA 一样有效。

3.2.2 与恒定 21-nt 反义 siRNA 配对的有义 siRNA 的长度变化

为了研究 siRNA 长度对 RNAi 的影响，我们制备了 3 个系列的 siRNA 双链体，使 3 种 21 nt 反义链与 8 种 18-25 nt 有义链结合。在每个 siRNA 双链体系列中，反义 siRNA 的 3' 突出端固定为 1、2 或 3 nt，而有义 siRNA 的 3' 端不同(图 12A). 我们发现，与有义 siRNA 的长度无关，具有反义 siRNA 的 2-nt 3' 突出端的双链体(图 12C)比具有 1-nt 或 3-nt 3' 突出端的双链体更有活性(图 12B, D). 在具有反义 siRNA 的 1-nt 3' 突出端的第一个系列中，含有 21-nt 和 22-nt 有义 siRNA 的双链体(它们分别携带有义 siRNA 的 1-nt 和 2-nt 3' 突出端)最具活性。含有 19-25 nt 有义 siRNA 的双链体也能介导 RNA，但程度较低。类似地，在具有反义 siRNA 的 2-nt 突出端的第二个系列中，具有 2-nt 3' 突出端的 21-nt siRNA 双链体最具活性，其他任何与 18-25 nt 有义 siRNA 的结合都有显著的活性。在具有 3-nt 反义 siRNA 3' 突出端的最后一个系列中，只有含 20-nt 有义 siRNA 和 2-nt 有义 3' 突出端的双链体能减少靶 RNA 表达。这些结果综合起来表明，siRNA 的长度以及 3' 突出端的长度非常重要，具有 2-nt 3' 突出端的 21-nt siRNA 双链体最适合于 RNAi.

3.2.3 具有恒定 2-nt 3' 突出端的 siRNA 双链体的长度变化

然后我们检查了通过保留对称 2-nt 3' 突出端同时改变两条 siRNA

链长度的影响(图 13A)。制备两个系列的 siRNA 双链体，包括图 11H 的 21-nt siRNA 双链体作为参照。通过在有义 siRNA 的 3' 端(图 13B)或在反义 siRNA 的 3' 端(图 13C)延长碱基配对的片段，双链体的长度在 20-25 bp 间不等。20-23 bp 双链体引起靶萤光素酶活性的特异性抑制，但是 21-nt siRNA 双链体比其它任何双链体有效性至少高 8 倍。24-和 25-nt siRNA 双链体不引起任何可检测的干涉。由于双链体两端的变化产生类似的作用，序列特异的影响较小。

3.2.4 2'-脱氧和 2'-O-甲基修饰的 siRNA 双链体

为了评估 siRNA 核糖残基对于 RNAi 的重要性，检测了含有 21-nt siRNA 和 2-nt 3' 突出端以及 2'-脱氧或 2'-O-甲基修饰的链的双链体(图 14)。2-nt 3' 突出端置换为 2'-脱氧核苷酸没有影响，甚至置换配对区内与突出端相邻的另外两个核糖核苷酸产生具有明显活性的 siRNA。因此，siRNA 双链体的 42 nt 中的 8 nt 置换为 DNA 残基，活性没有损失。然而，一条或两条 siRNA 链完全置换为 2'-脱氧残基消除了 RNAi，置换为 2'-O-甲基残基同样如此。

3.2.5 靶 RNA 切割位点的确定

以前确定了 22-nt siRNA 双链体和 21-nt/22-nt 双链体的靶 RNA 切割位点。发现靶 RNA 切割位点位于 siRNA 双链体覆盖的区域的中心，在与 21-或 22-nt siRNA 指导序列互补的第一个核苷酸的 11 或 12 nt 下游。具有 2-nt 3' 突出端的 5 种不同 21-nt siRNA 双链体(图 15A)与 5' 帽标记的有义或反义靶 RNA 在黑尾果蝇裂解液中温育(Tuschl 等人，1999；Zamore 等人，2000)。5' 切割产物在测序凝胶上分析(图 15B)。切割的有义靶 RNA 的量与通过基于翻译的实验测定的 siRNA 双链体的效率有关，siRNA 双链体 1、2、4(图 15B 和 11H, G, E)切割靶 RNA 比双链体 3 和 5 更快(图 15B 和 11F, D)。显然，5' 切割产物和输入靶 RNA 的总放射性随着时间的增长不是恒定的，而且 5' 切割产物不积累。推测起来，由于缺乏 5'-帽子的任一条 poly(A)尾，切割产物一从 siRNA-内切核酸酶复合物上释放，即快速降解。

有义和反义靶 RNA 的切割位点都位于 siRNA 双链体覆盖的区域

的中间。根据双链体沿靶序列的 1-nt 置换来看，5 种不同双链体产生的每种靶标的切割位点有 1-nt 不同。靶标精确地在靶位点 11 nt 下游切割，靶位点与序列互补的指导 siRNA 的 3'-核苷酸互补(图 15A, B)。

为了确定是指导 siRNA 的 5' 端还是 3' 端控制靶 RNA 切割，我们设计了实验策略，在图 16A 和 B 中概述。一种 21-nt 反义 siRNA，在该研究中保持不变，与 5' 或 3' 端修饰的有义 siRNA 配对。有义和反义靶 RNA 的切割位点如上所述确定。有义 siRNA 的 3' 端的改变(监测 1-nt 5' 突出端到 6-nt 3' 突出端)不影响有义或反义靶 RNA 的切割位点(图 16C)。有义 siRNA 的 5' 端的改变不影响有义靶 RNA 的切割(图 16D, 上图)，预计是因为反义 siRNA 未改变。然而，反义靶 RNA 切割受到影响，并且强烈依赖于有义 siRNA 的 5' 端(图 16D, 下图)。当有义 siRNA 大小为 20 或 21 nt 时，只切割反义靶标，切割位点时相差 1-nt，表明靶识别的 siRNA 的 5' 端控制靶 RNA 的切割。当从与指导 siRNA 的 5' 核苷酸配对的靶核苷酸开始以下游方向计数时，该位点位于核苷酸 10 与 11 之间(参见图 15A)。

3.2.6 序列效应和 3' 突出端的 2'-脱氧置换

对于 siRNA 的功能，2-nt 3' 突出端是优选的。我们希望了解突出核苷酸的序列对于靶识别是否有作用，或者这是否只是内切核酸酶复合物(RISC 或 siRNP)重建所需的一个特征。我们合成了有义和反义 siRNA，其具有 AA、CC、GG、UU 和 UG 3' 突出端，包括 2'-脱氧修饰 TdG 和 TT。野生型 siRNA 在有义 3' 突出端中含有 AA，在反义 3' 突出端中含有 UG (AA/UG)。所有 siRNA 双链体在干涉测定中都起作用，使靶表达至少降低 5 倍(图 17)。最有效的 siRNA 双链体是序列型 NN/UG、NN/UU、NN、TdG 和 NN/TT (N, 任何核苷酸)，使靶表达降低 10 倍以上。含有 AA、CC 或 GG 的反义 siRNA 3' 突出端的 siRNA 双链体比野生型序列 UG 或突变 UU 活性低 2-4 倍。这种 RNAi 效率的降低可能是由于倒数第二个 3' 核苷酸对序列特异性靶识别的作用，因为 3' 端核苷酸由 G 变为 U 没有影响。

有义 siRNA 3' 突出端序列的改变没有显示任何依赖序列的作用，

这是预料到的，因为有义 siRNA 对于有义靶 mRNA 的识别不一定有作用。

3.2.7 靶识别的序列特异性

为了检验靶识别的序列特异性，我们向 siRNA 双链体的配对片段内引入序列改变，并测定沉默率。通过颠倒 3-或 4-nt 长的短片段或点突变，引入序列改变(图 18)。为了避免干扰碱基配对的 siRNA 双链体结构，一条 siRNA 链的序列改变在互补 siRNA 链中补偿。为了降低合成成本，所有 2-nt 3' 突出端序列都是 TT (T, 2'-脱氧胸苷)。TT/TT 参照 siRNA 双链体在 RNAi 方面与野生型 siRNA 双链体 AA/UG 相当(图 17)。用基于翻译的发光测定法定量介导报道 mRNA 破坏的能力。显示含有反向序列片段的 siRNA 双链体靶向萤火虫萤光素酶报道分子的能力显著降低(图 18)。位于反义 siRNA 3' 端与中心之间的序列改变完全消除了靶 RNA 识别，但是靠近反义 siRNA 5' 端的突变显示程度较低的沉默。与预测的靶 RNA 切割位点正对的 A/U 碱基对颠换，或者远离预测位点的一个核苷酸的颠换，阻止靶 RNA 的切割，因此表明，siRNA 双链体中心的单突变区别错配的靶标。

3.3 讨论

siRNA 是不仅在昆虫细胞中，而且在哺乳动物细胞中，用于灭活基因表达的有价值的试剂，具有广阔的治疗应用前景。我们系统分析了在黑尾果蝇胚胎裂解液中促进高效靶 RNA 降解所需的 siRNA 双链体的结构决定簇，从而提出了最有效的 siRNA 双链体的设计原则。完美的 siRNA 双链体能够沉默基因表达，假定使用含量相当的总 RNA，其效率相当于 500 bp dsRNA。

3.4 siRNA 使用指南

高效沉默的 siRNA 双链体优选地由 21-nt 反义 siRNA 组成，应当选择它们形成具有 2-nt 3' 突出端的 19 bp 双螺旋。2-nt 3' 突出核糖核苷酸的 2'-脱氧置换不影响 RNAi，但是有助于降低 RNA 合成的成本，并且可提高 siRNA 双链体的 RNase 抗性。然而，更广泛的 2'-脱氧或 2'-O-甲基修饰降低了 siRNA 介导 RNAi 的能力，这可能是通过干扰

SiRNAP 装配的蛋白质结合。

靶标识别是一个高度序列特异性的过程，由与靶标互补的 siRNA 介导。指导 siRNA 的 3'-核苷酸对于靶标识别的特异性没有作用，而 3' 突出端的倒数第二个核苷酸影响靶 RNA 切割，错配使 RNAi 降低 2-4 倍。指导 siRNA 的 5' 端似乎也比 3' 端更能容许错配的靶 RNA 识别。位于靶 RNA 切割位点对面的 siRNA 中心的核苷酸是重要的特异性决定簇，甚至单核苷酸改变也能使 RNAi 降至不可检测的水平。这提示，在基因靶向实验中，siRNA 双链体能区别突变的或多态性等位基因，这可能成为未来治疗发展的重要特征。

当有义和反义 siRNA 与内切核酸酶复合物或其定型复合物的蛋白质成分结合时，认为两者起不同的作用；该复合物中 siRNA 双链体的相对方向决定了哪条链能用于靶标识别。合成 siRNA 双链体在双螺旋结构上二重对称，但序列不对称。siRNA 双链体在黑尾果蝇裂解液中与 RNAi 蛋白的结合将形成两种不对称复合物。在这种假定的复合物中，有义和反义 siRNA 的手性环境不同，因此功能不同。这种预测显然不适用于回文 siRNA 序列，或者能结合为同型二聚体的 RNAi 蛋白。为了使可影响有义与反义靶向 siRNP 之比的序列效应达到最小，我们建议使用含有相同 3' 突出序列的 siRNA 序列。我们推荐将有义 siRNA 突出端的序列调节为反义 3' 突出端的序列，因为有义 siRNA 在一般的敲除实验中没有靶标。有义和反义切割的 siRNP 重建的不对称性可能(部分地)导致 RNAi 效率的改变，用该研究使用的具有 2-nt 3' 突出端的多种 21-nt siRNA 双链体观察到这种改变(图 14)。此外，靶位点处的核苷酸序列和/或靶 RNA 结构的可接近性可能导致这些 siRNA 双链体的效率的改变。

参考文献

Bass, B. L. (2000). 双链 RNA 作为基因沉默的模板。 Cell 101, 235-238.

Bosher, J. M. 和 Labouesse, M. (2000). RNA 干涉：遗传魔杖和遗

传看家狗. *Nat. Cell Biol.* 2, E31-36.

Caplen, N. J., Fleenor, J., Fire, A. 和 Morgan, R. A. (2000). 培养的果蝇细胞中 dsRNA-介导的基因沉默: 用于分析 RNA 干涉的一种组织培养模型. *Gene* 252, 95-105.

Catalanotto, C., Azzalin, G., Macino, G. 和 Cogoni, C. (2000). 蠕虫和真菌中的基因沉默. *Nature* 404, 245.

Chanfreau, G., Buckle, M. 和 Jacquier, A. (2000). 酿酒酵母 RNase III 对一类保守的 RNA 四环的识别. *Proc. Natl. Acad. Sci. USA* 97, 3142-3147.

Clemens, M. J. (1997). PKR—由双链 RNA 调节的一种蛋白激酶. *Int. J. Biochem. Cell Biol.* 29, 945-949.

Cogoni, C. 和 Macino, G. (1999). 植物和真菌中依赖同源性的基因沉默: 同一主题的大量变异. *Curr. Opin. Microbiol.* 2, 657-662.

Dalmay, T., Hamilton, A., Rudd, S., Angell, S. 和 Baulcombe, D. C. (2000). 转基因介导的而不是病毒介导的转录后基因沉默需要 *Arabidopsis* 的一种 RNA-依赖性 RNA 聚合酶基因. *Cell* 101, 543-553.

Dernburg, A. F., Zalevsky, J., Colaiacovo, M. P. 和 Villeneuve, A. M. (2000). *C. elegans* 种系中转基因介导的共抑制. *Genes & Dev.* 14, 1578-1583.

Dunn J. J. (1982). 核糖核酸酶 III. 《酶》, 第 15 卷, 部分 B, P. D. Boyer 编著 (New York: Academic Press), pp. 485-499.

Filippov, V., Solovyev, V., Filippova, M. 和 Gill, S. S. (2000). 真核生物中的一种新型 RNase III 家族蛋白. *Gene* 245, 213-221.

Fire, A. (1999). RNA-触发的基因沉默. *Trends Genet.* 15, 358-363.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. 和 Mello, C. C. (1998). *Caenorhabditis elegans* 中双链 RNA 引起的强烈特异性遗传干涉. *Nature* 391, 806-811.

Grishok, A., Tabara, H. 和 Mello, C. C. (2000). *C. elegans* 中 RNAi

遗传的基因要求。 *Science* 287, 2494-2497.

Hamilton, A. J. 和 Baulcombe, D. C. (1999). 植物转录后基因沉默中的一个小反义 RNA 种。 *Science* 286, 950-952.

Hammond, S. M., Bernstein, E., Beach, D. 和 Hannon, G. J. (2000). 在果蝇细胞中, 一种 RNA 指导的核酸酶介导转录后基因沉默。 *Nature* 404, 293-296.

Jacobsen, S. E., Running, M. P. 和 M., M. E. (1999). *Arabidopsis* 中 RNA 解旋酶/RNase III 基因的破坏导致花分生组织中无控制的细胞分裂。 *Development* 126, 5231-5243.

Jensen, S., Gassama, M. P. 和 Heidmann, T. (1999). 依赖同源性的基因沉默对转座因子的抑制(Taming)。 *Nat. Genet.* 21, 209-212.

Kehlenbach, R. H., Dickmanns, A. & Gerace, L. (1998). 包括 Ran 和 CRM1 在内的核胞质穿梭因子介导 NFAT 的体外核输出。 *J. Cell Biol.* 141, 863-874.

Kennerdell, J. R. 和 Carthew, R. W. (1998). dsRNA-介导的遗传干涉证明无翅途径中 frizzled 和 frizzled 2 作用的用途。 *Cell* 95, 1017-1026.

Ketting, R. F., Haverkamp, T. H., van Luenen, H. G. 和 Plasterk, R. H. (1999). 转座子沉默和 RNA 干涉需要的 *C. elegans* 的 Mut-7 是 Werner 综合征解旋酶和 RNaseD 的同源物。 *Cell* 99, 133-141.

Ketting, R. F. 和 Plasterk, R. H. (2000). *C. elegans* 中共抑制与 RNA 干涉之间的遗传联系。 *Nature* 404, 296-298.

Lucy, A. P., Guo, H. S., Li, W. X. 和 Ding, S. W. (2000). 位于核中的植物病毒蛋白对转录后基因沉默的抑制。 *EMBO J.* 19, 1672-1680.

Matsuda, S., Ichigotani, Y., Okuda, T., Irimura, T., Nakatsugawa, S. 和 Hamaguchi, M. (2000). 编码推断的 RNA 解旋酶的一种新人类基因(HERNA)的分子克隆和表征。 *Biochim. Biophys. Acta* 31, 1-2.

Milligan, J.F. 和 Uhlenbeck, O.C. (1989). 利用 T7 RNA 聚合酶合成小 RNAs。 *Methods Enzymol.* 180, 51-62.

Mourrain, P., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J. B., Jouette, D., Lacombe, A. M., Nikic, S., Picault, N., Remoue, K., Sanial, M., Vo, T. A. 和 Vaucheret, H. (2000). 转录后基因沉默和天然病毒抗性需要 *Arabidopsis SGS2* 和 *SGS3* 基因. *Cell* 101, 533-542.

Ngo, H., Tschudi, C., Gull, K. 和 Ullu, E. (1998). 在布氏锥虫 (*Trypanosoma brucei*) 中双链 RNA 诱导 mRNA 降解. *Proc. Natl. Acad. Sci. USA* 95, 14687-14692.

Nicholson, A. W. (1999). 细胞核糖核酸酶的功能、机制和调节. *FEMS Microbiol. Rev.* 23, 371-390.

Oelgeschlager, M., Larrain, J., Geissert, D. 和 De Robertis, E. M. (2000). 进化上保守的 BMP-结合蛋白缠绕的原肠胚形成促进 BMP 信号发生. *Nature* 405, 757-763.

Pan, T. 和 Uhlenbeck, O. C. (1992). 经历 Pb^{2+} 自切割的 RNAs 的体外选择. *Biochemistry* 31, 3887-3895.

Pelissier, T. 和 Wassenegger, M. (2000). 一种 30 bp 的 DNA 靶标对于 RNA-指导的甲基化是足够的. *RNA* 6, 55-65.

Plasterk, R. H. 和 Ketting, R. F. (2000). 基因沉默. *Curr. Opin. Genet. Dev.* 10, 562-567.

Ratcliff, F. G., MacFarlane, S. A. 和 Baulcombe, D. C. (1999). 不含 DNA 的基因沉默. 病毒间 RNA 介导的交叉保护. *Plant Cell* 11, 1207-1216.

Robertson, H. D. (1990). 大肠杆菌核糖核酸酶 III. *Methods Enzymol.* 181, 189-202.

Robertson, H. D. (1982). 大肠杆菌核糖核酸酶 III 切割位点. *Cell* 30, 669-672.

Romaniuk, E., McLaughlin, L. W., Neilson, T. 和 Romaniuk, P. J. (1982). 受体寡核糖核苷酸序列对 T4 RNA 连接酶反应的影响. *Eur J Biochem* 125, 639-643.

Sharp, P. A. (1999). RNAi 和双链 RNA. *Genes & Dev.* 13, 139-141.

Sijen, T. 和 Kooter, J. M. (2000). 转录后基因沉默: RNAs 是攻击不是防御? *Bioessays* 22, 520-531.

Smardon, A., Spoerke, J., Stacey, S., Klein, M., Mackin, N. 和 Maine, E. (2000). 在 *C. elegans* 中, EGO-1 与 RNA 指导的聚合酶和种系发育中的功能和 RNA 干涉有关. *Curr. Biol.* 10, 169-178.

Svoboda, P., Stein, P., Hayashi, H. 和 Schultz, R. M. (2000). RNA 干涉所致小鼠卵母细胞中静止母本 mRNAs 的选择性减少. *Development* 127, 4147-4156.

Tabara, H., Sarkissian, M., Kelly, W. G., Fleenor, J., Grishok, A., Timmons, L., Fire, A. 和 Mello, C. C. (1999). *C. elegans* 中的 rde-1 基因、RNA 干涉和转座子沉默. *Cell* 99, 123-132.

Tuschl, T., Ng, M. M., Pieken, W., Benseler, F. 和 Eckstein, F. (1993). 核心鸟苷的外环碱基功能基团对于锤头状核酶活性的重要性. *Biochemistry* 32, 11658-11668.

Tuschl, T., Sharp, P. A. 和 Bartel, D. P. (1998). 从部分随机化的 U2 和 U6 snRNA 文库中体外选择新型核酶. *EMBO J.* 17, 2637-2650.

Tuschl, T., Zamore, P. D., Lehmann, R., Bartel, D. P. 和 Sharp, P. A. (1999). 体外双链 RNA 引起的定向 mRNA 降解. *Genes & Dev.* 13, 3191-3197.

Ui-Tei, K., Zenno, S., Miyata, Y. 和 Saigo, K. (2000). 利用萤火虫萤光素酶基因作为靶标对果蝇和中国仓鼠培养细胞中 RNA 干涉的敏感测定. *FEBS Letters* 479, 79-82.

Verma, S. 和 Eckstein, F. (1999). 修饰的寡核苷酸: 使用者的合成和策略. *Annu. Rev. Biochem.* 67, 99-134.

Voinnet, O., Lederer, C. 和 Baulcombe, D. C. (2000). 在烟草 *Nicotiana benthamiana* 中, 一种病毒运动蛋白阻止基因沉默信号的扩散. *Cell* 103, 157-167.

Wassenegger, M. (2000). RNA-指导的 DNA 甲基化. *Plant Mol.*

Biol. 43, 203-220.

Wianny, F. 和 Zernicka-Goetz, M. (2000). 早期小鼠发育过程中双链 RNA 对基因功能的特异性干涉。 Nat. Cell Biol. 2, 70-75.

Wu, H., Xu, H., Miraglia, L. J. 和 Crooke, S. T. (2000). 人 RNase III 是一种参与前核糖体 RNA 加工的 160 kDa 蛋白质。 J. Biol. Chem. 17, 17.

Yang, D., Lu, H. 和 Erickson, J.W. (2000) 在果蝇胚胎中, 在 RNAi 期间, 加工的小 dsRNAs 可以介导序列特异的 mRNA 降解的证据。 Curr. Biol., 10, 1191-1200.

Zamore, P. D., Tuschl, T., Sharp, P. A. 和 Bartel, D. P. (2000). RNAi: 双链 RNA 指导 mRNA 以 21-23 个核苷酸的间隔依赖 ATP 地切割。 Cell 101, 25-33.

Zhang, K. 和 Nicholson, A. W. (1997). 双螺旋序列抗决定簇对核糖核酸酶 III 加工的调节。 Proc. Natl. Acad. Sci. USA 94, 13437-13441.

图 1A

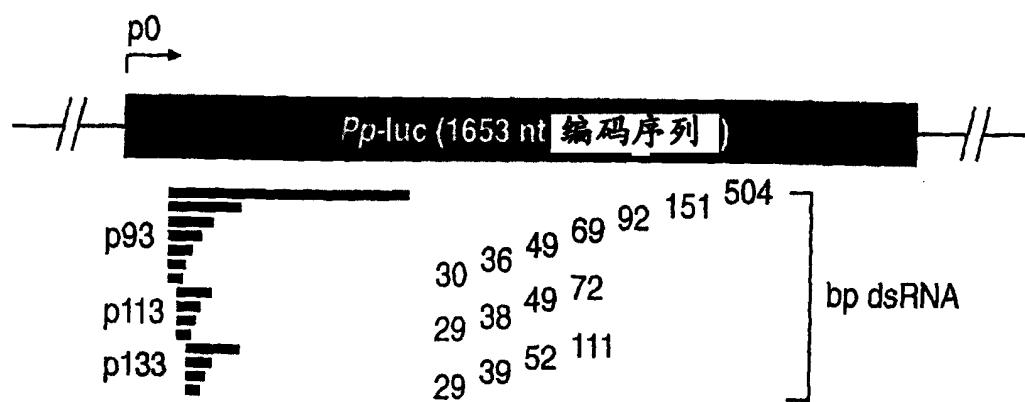


图 1B

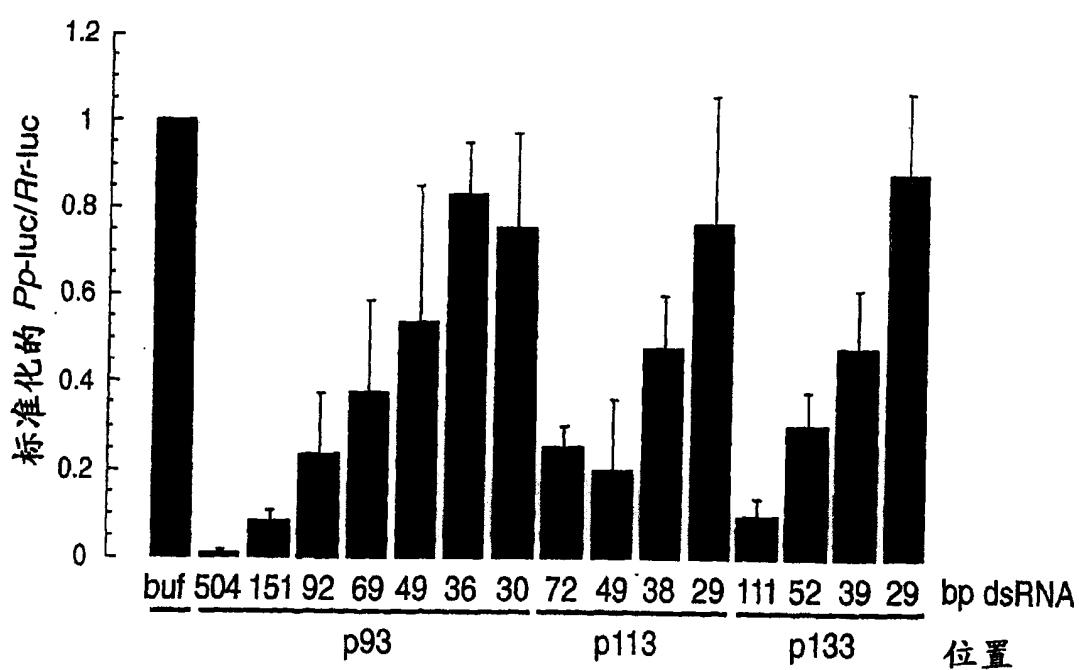


图 2

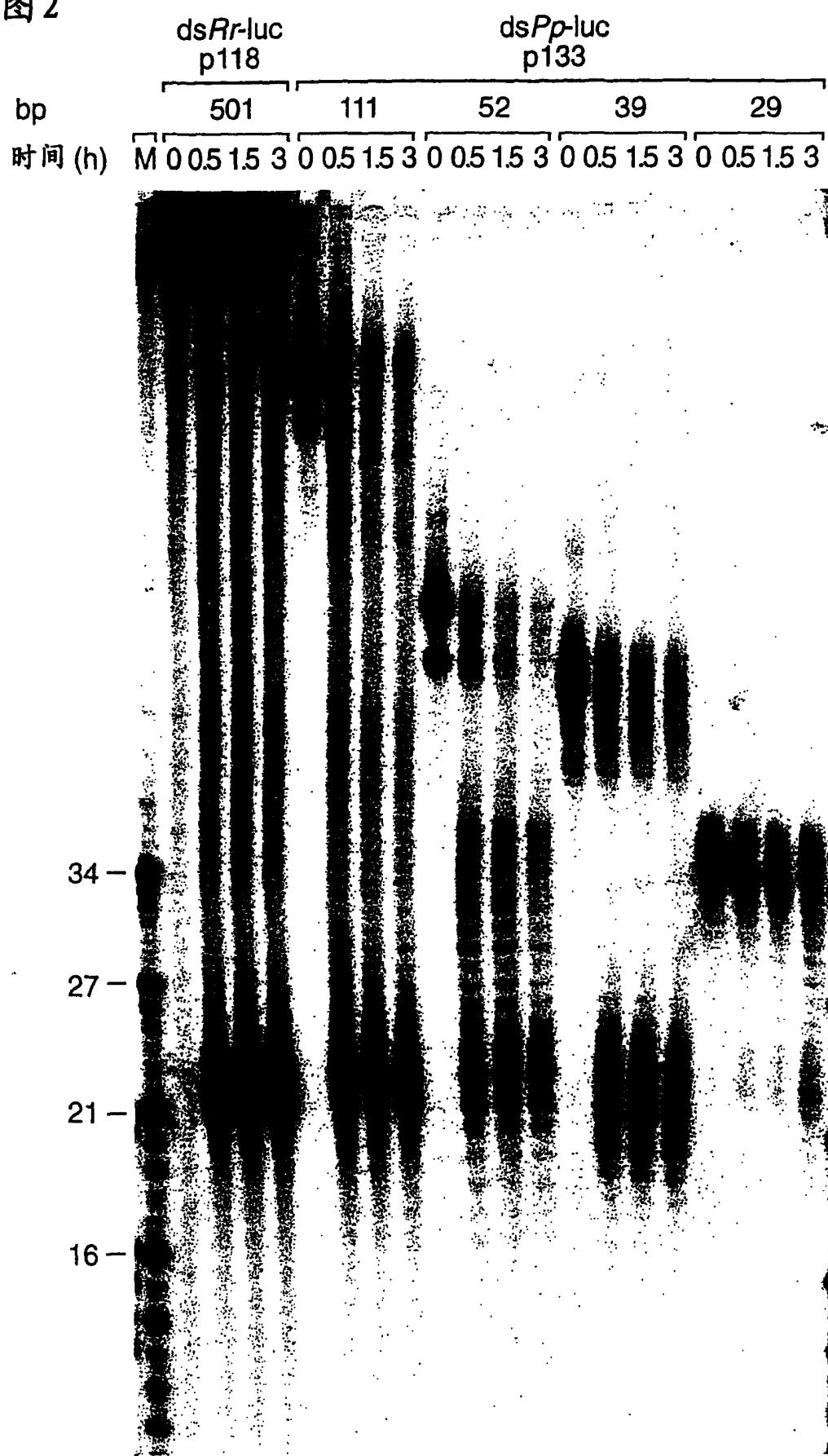


图 3a

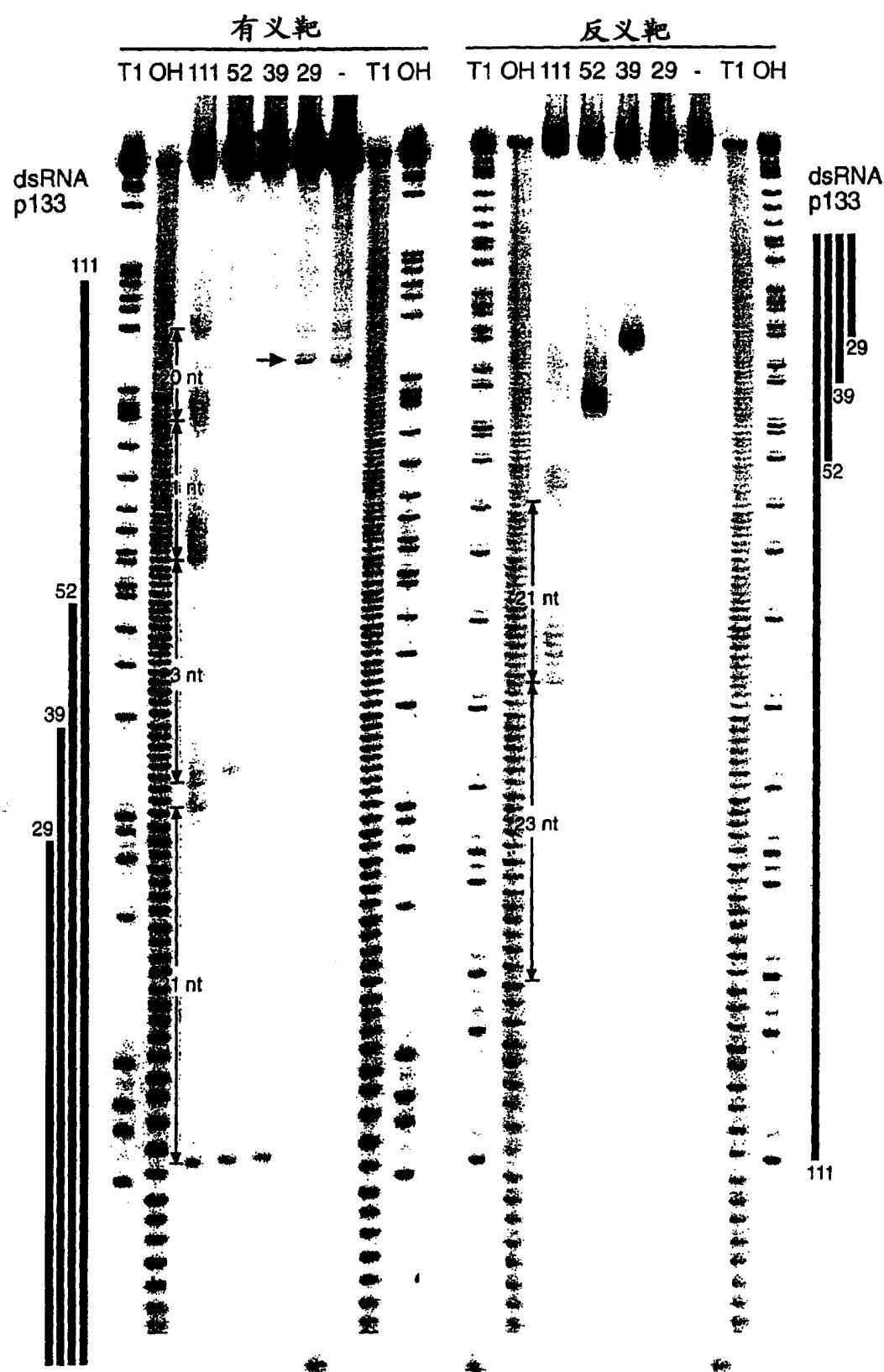


图 3B

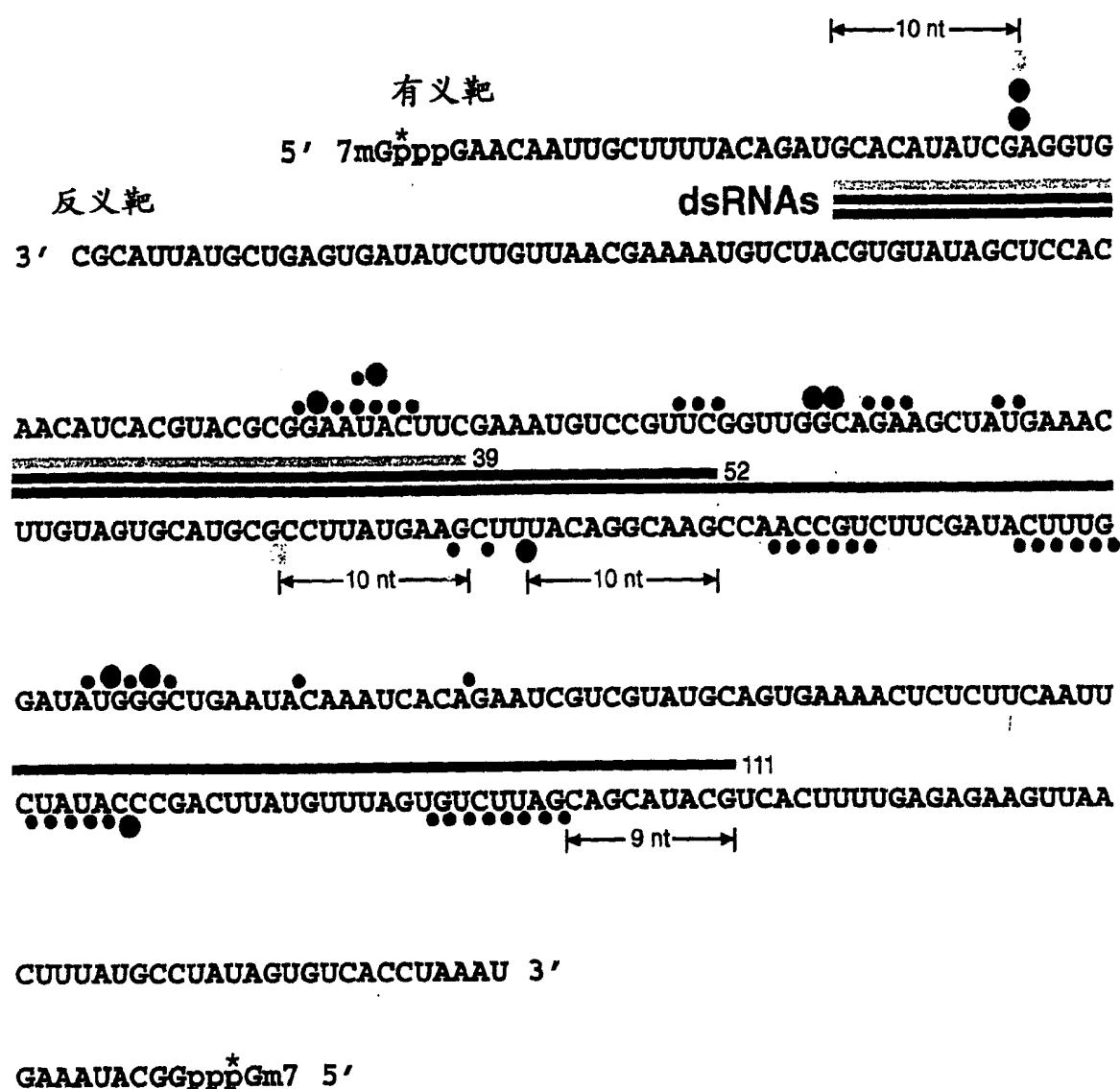


图 4A

39 bp dsPp-luc p133

5' GCACAUUAUCGAGGUGAACAUCAUCGUACGCGGAAUACUUC
3' CGUGUAUAGCUCCACUUGUAGUGCAUGC GCCUUUAUGAAG

52 bp dsPp-luc p133

5' GCACAUUAUCGAGGUGAACAUCAUCGUACGCGGAAUACUUCGAAAUGUCCGUUC
3' CGUGUAUAGCUCCACUUGUAGUGCAUGC GCCUUUAUGAAGCUUUACAGGCAAG

111 bp dsPp-luc p133

5' GCACAUUAUCGAGGUGAACAUCAUCGUACGCGGAAUACUUCGAAAUGUCCGUUCGGU
3' CGUGUAUAGCUCCACUUGUAGUGCAUGC GCCUUUAUGAAGCUUUACAGGCAAGCCA

UGGCAGAACUAUGAACGAUAUGGGCUGAAUACAAACACAGAACUGUCGUUAUGC
ACCGUCUUUCGAUACUUUGCUAUACCCGACUUAGUUUAGUGUCUUAGCAGCAUACG

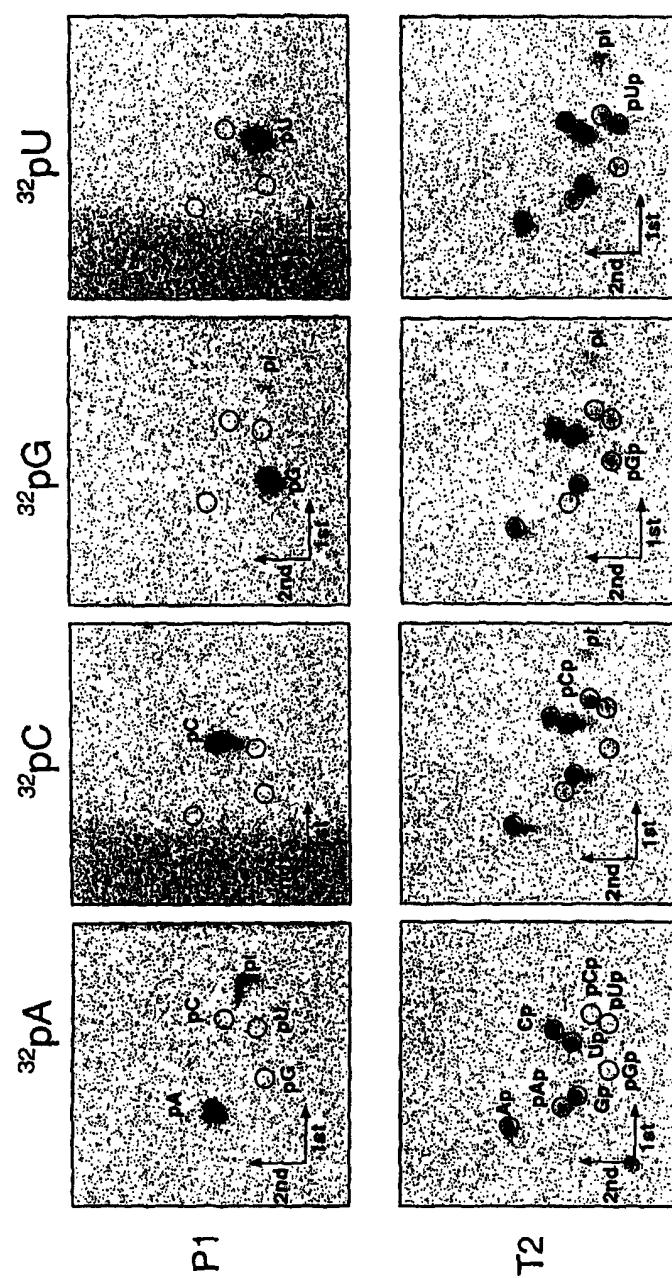


图 4B

图 5A

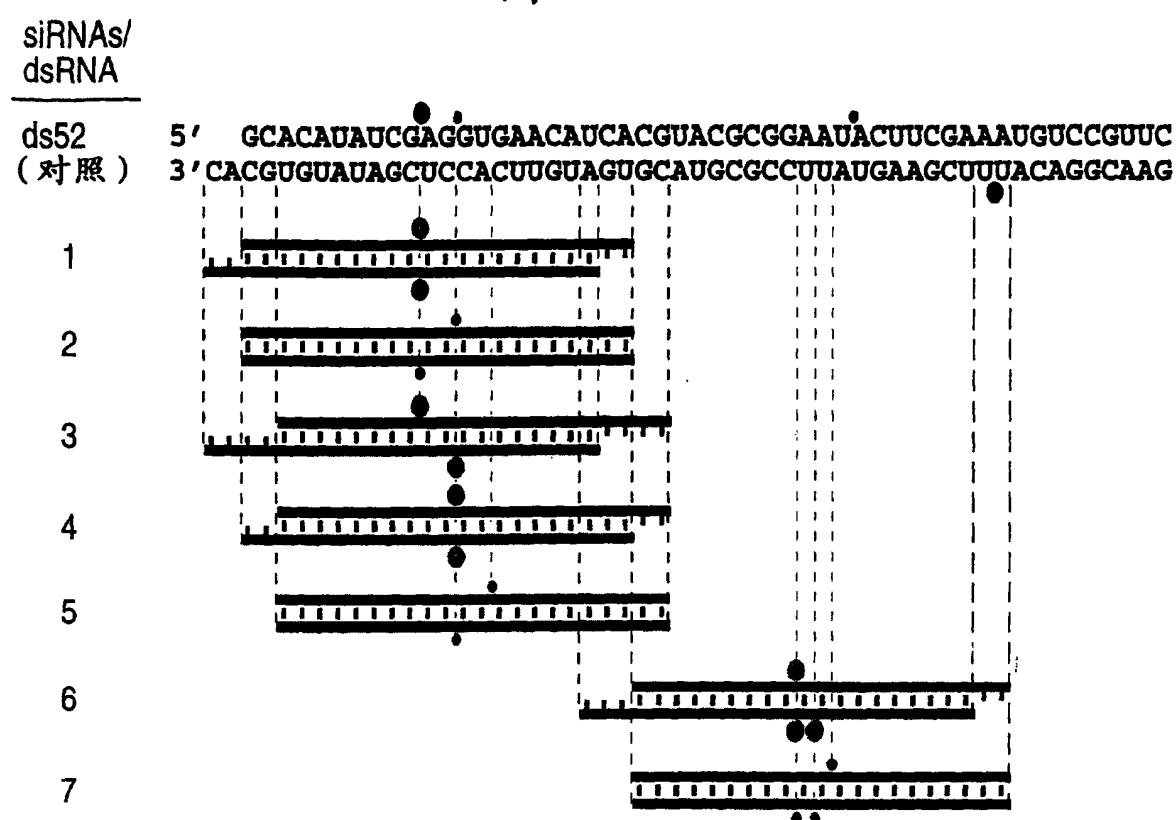


图 5B

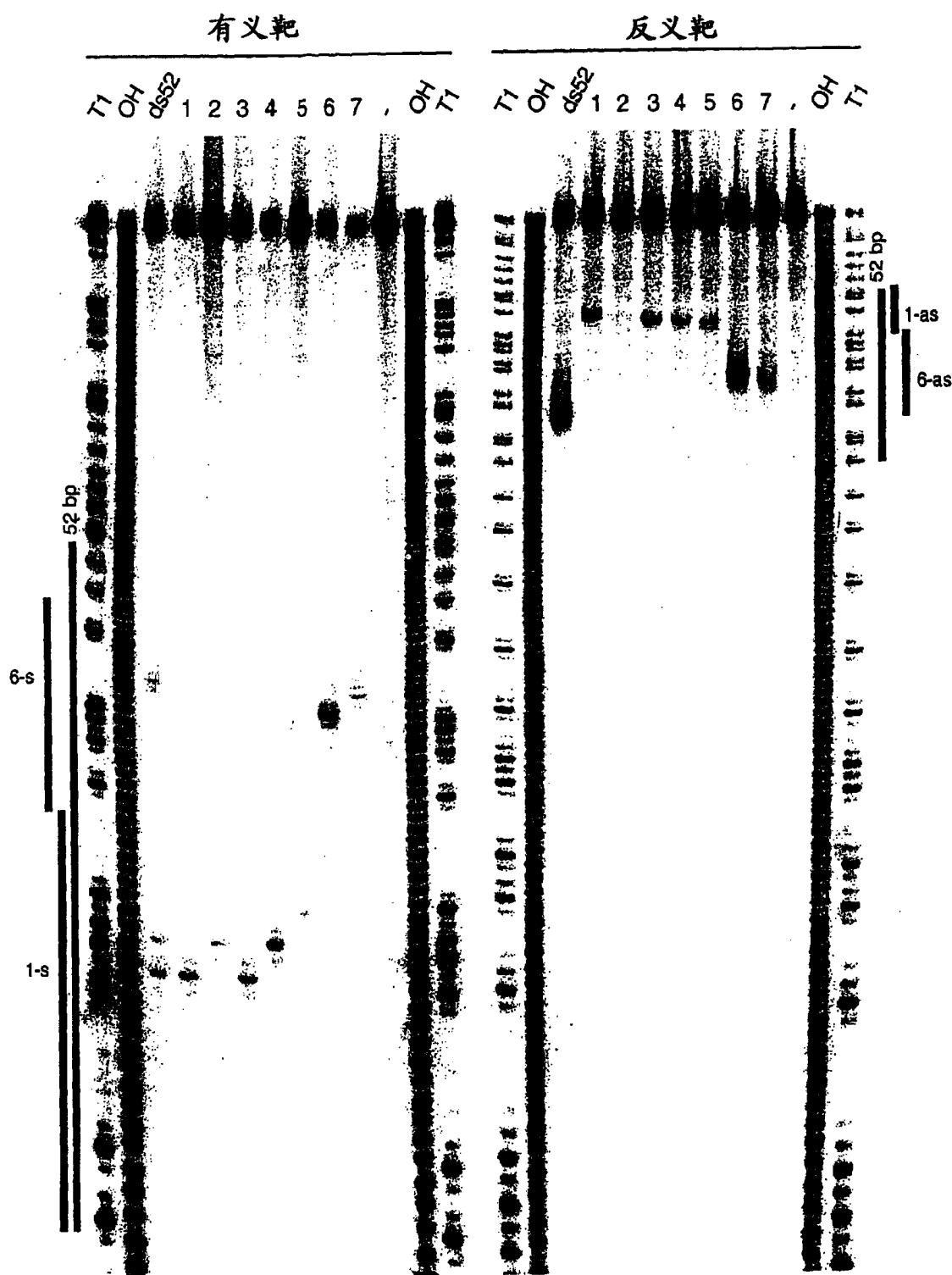


图 6A

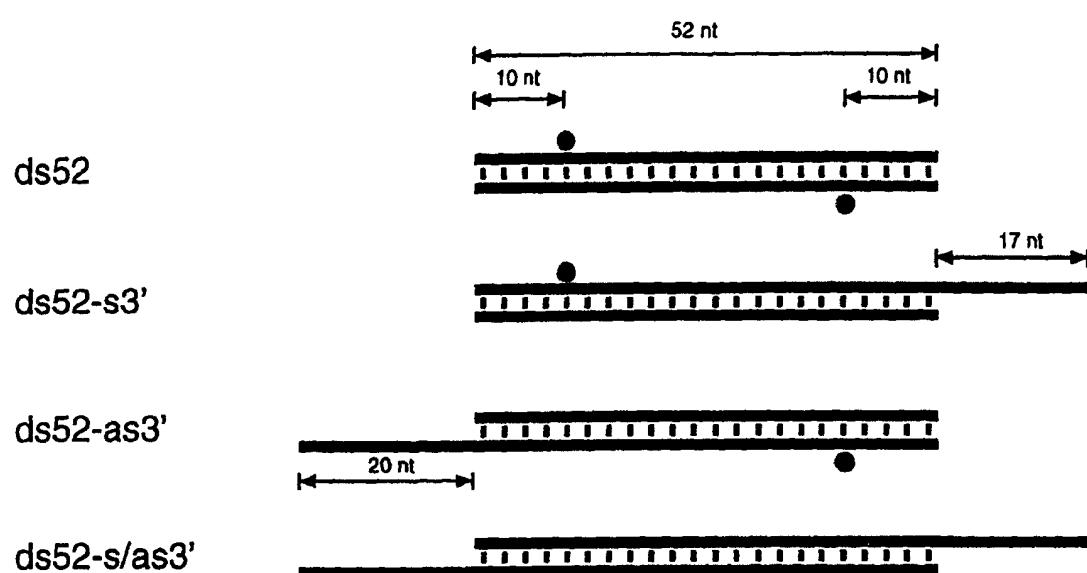


图 6B

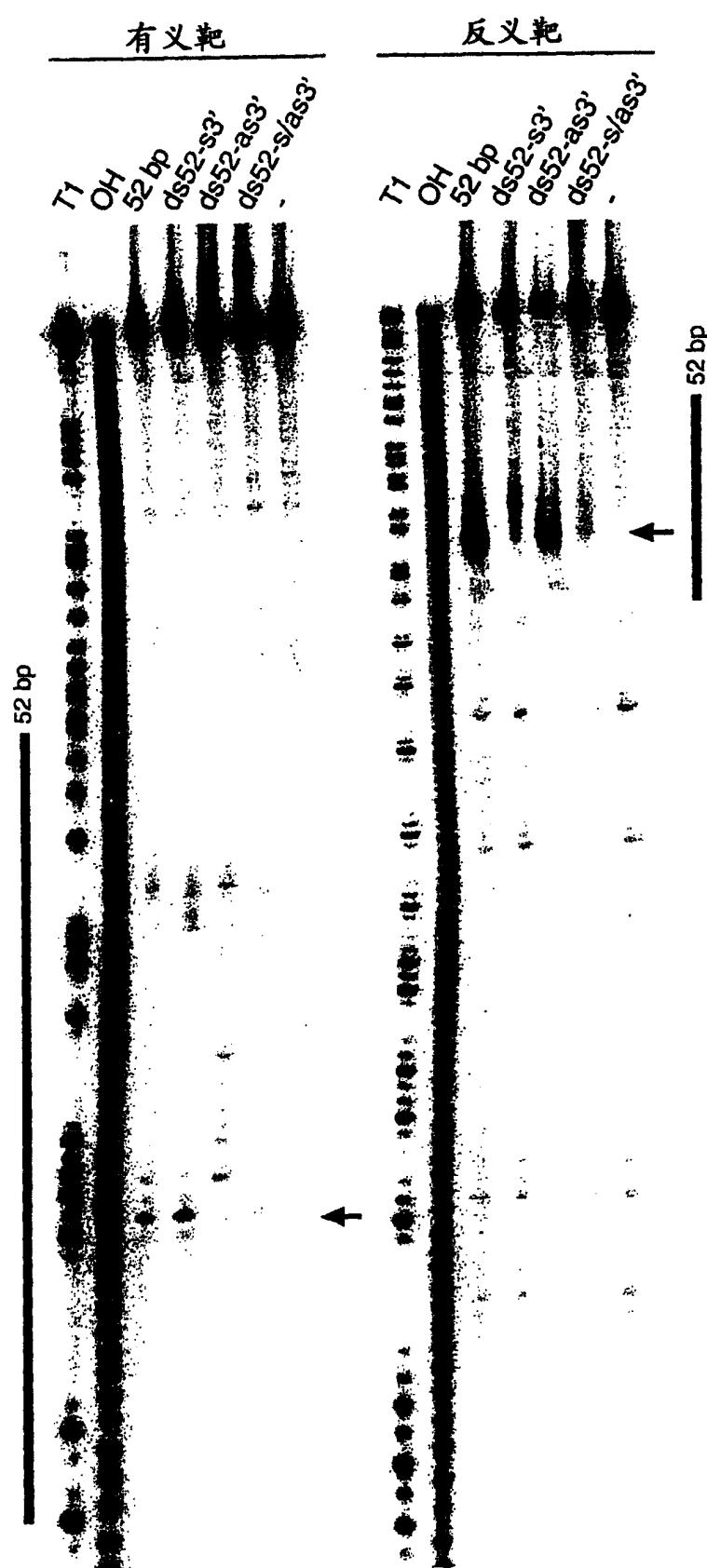


图 7

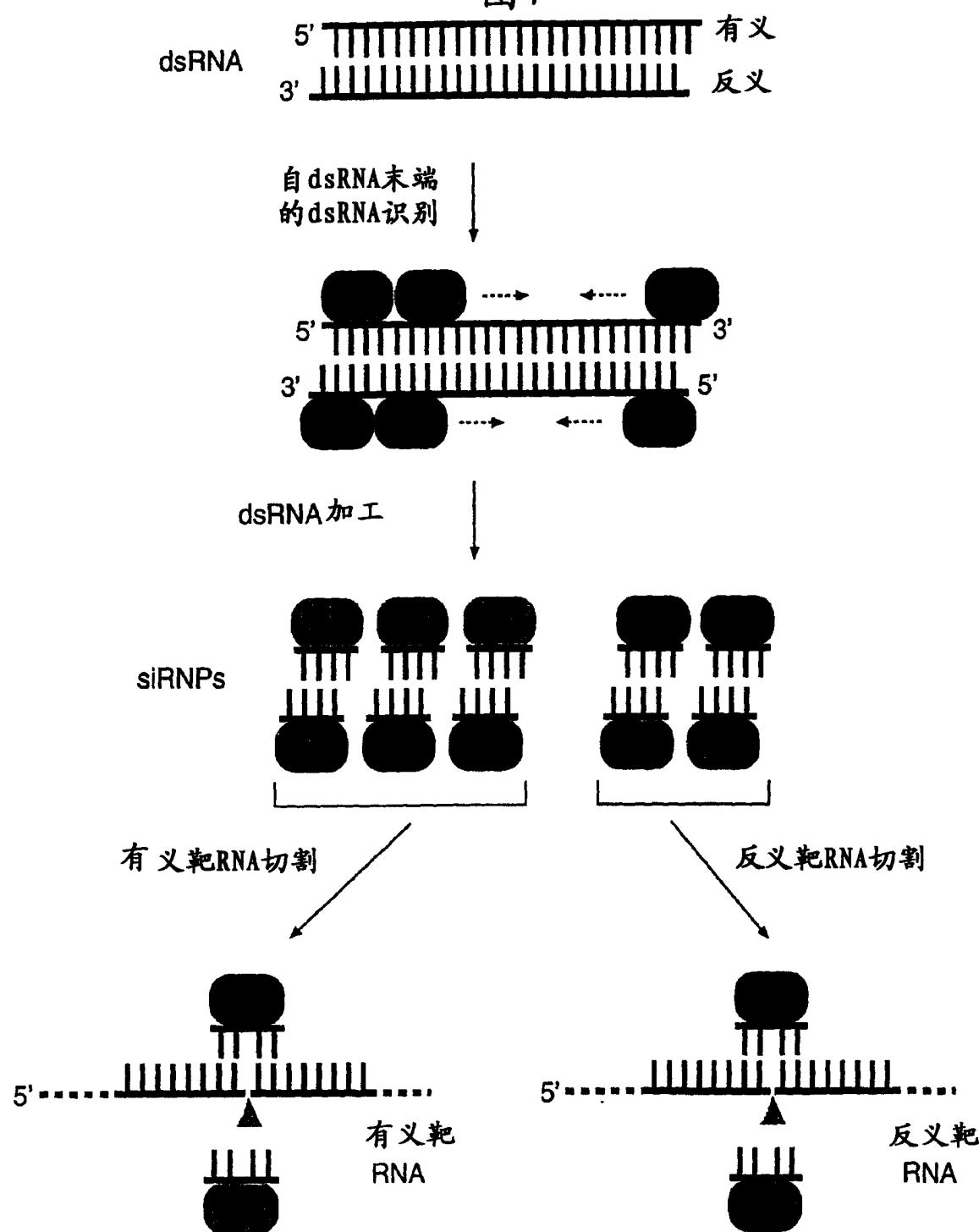
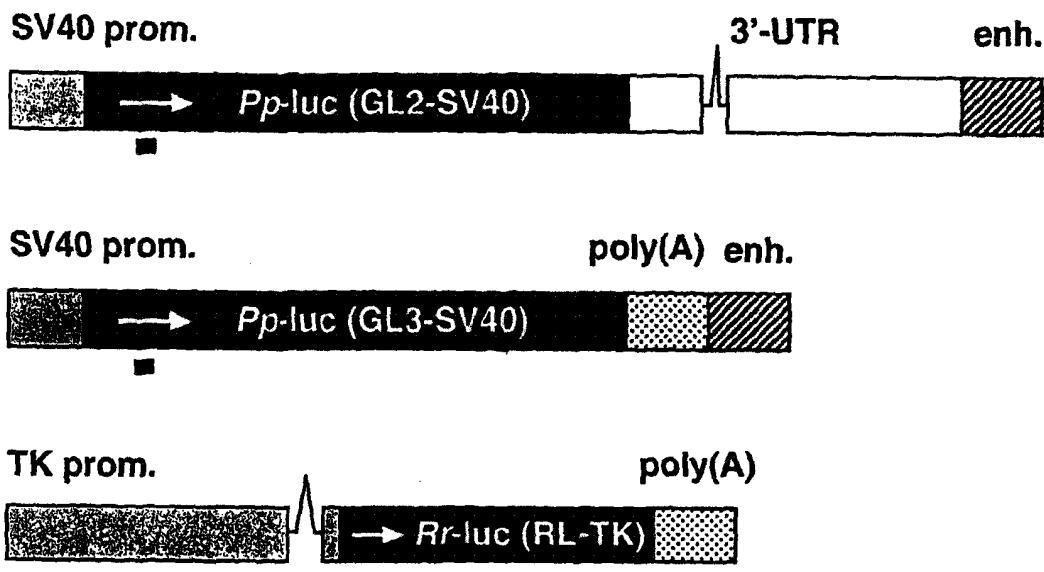


图8

a



b

siRNA

双链体

uGL2	5' CGUACGCGGA <u>A</u> UACUUUCGA <u>D</u> UUG <u>C</u> AUGCG <u>C</u> U <u>A</u> UGAAGCU 5'
GL2	5' CGUACGCGGA <u>A</u> UACUUUCGATT TT <u>G</u> GAUGCG <u>C</u> U <u>A</u> UGAAGCU 5'
GL3	5' C <u>GUACGCG</u> U <u>GAGU</u> ACUUUCGATT TT <u>G</u> GAUGCG <u>A</u> CU <u>A</u> UGAAGCU 5'
invGL2	5' AGCUUCAUAAGGC <u>G</u> CAUGCTT TTUC <u>G</u> GAAGU <u>A</u> UCC <u>G</u> CGUACG 5'
RL	5' AA <u>ACA</u> AUG <u>CAG</u> AAA <u>AUGCUG</u> TT TT <u>UUG</u> UAC <u>GUAC</u> UUUU <u>ACG</u> AC 5'

图 9

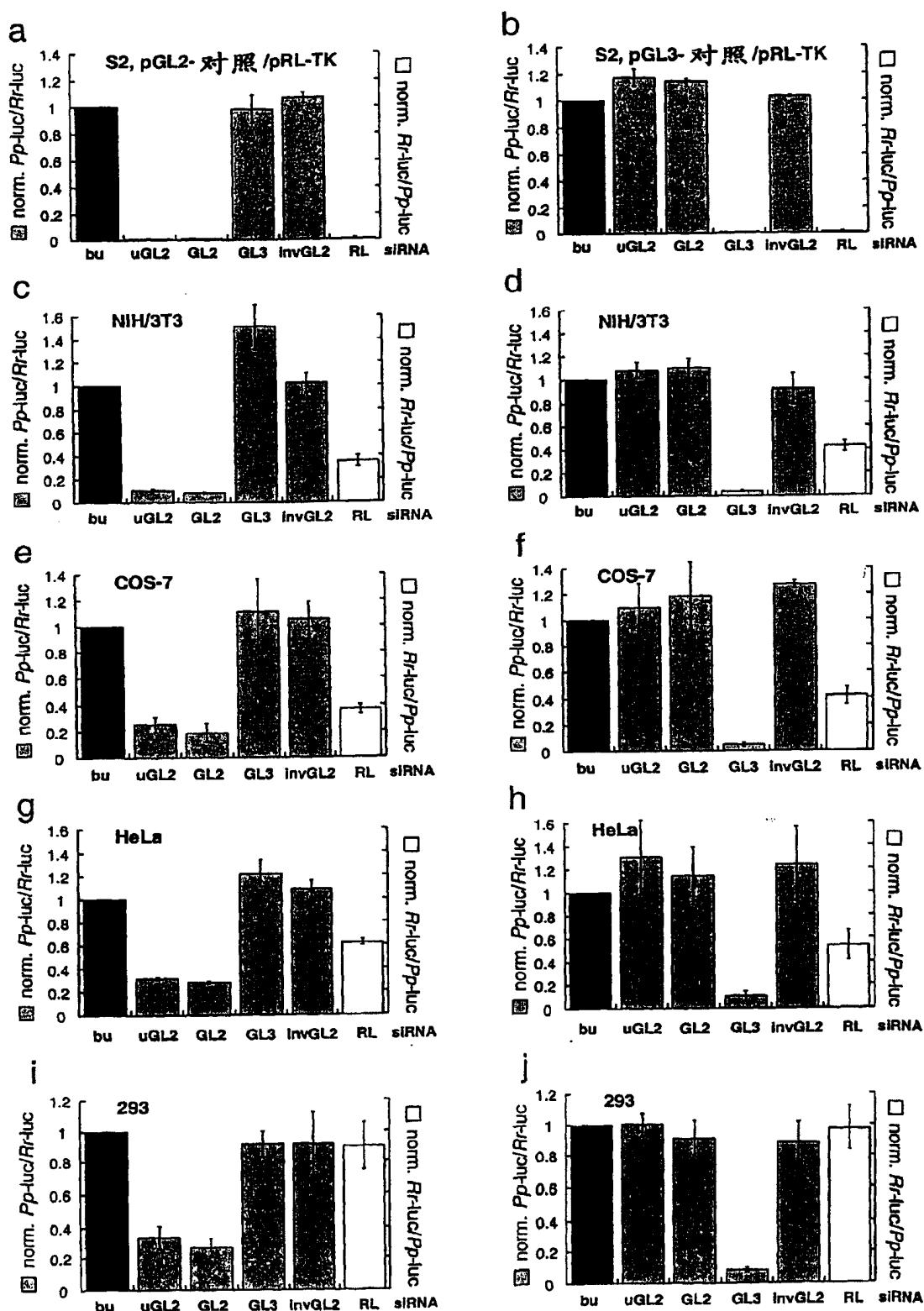


图 10

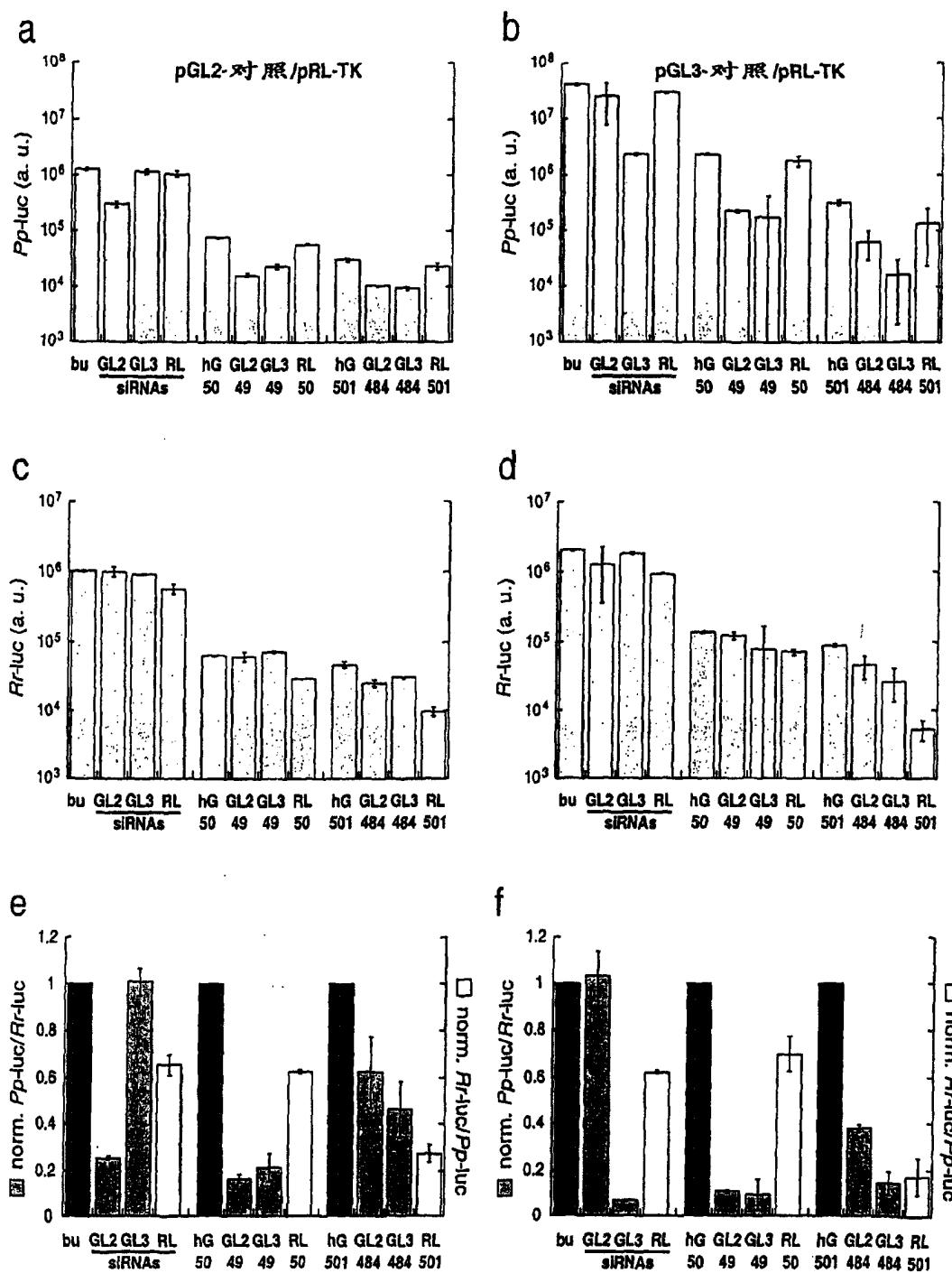


图 11

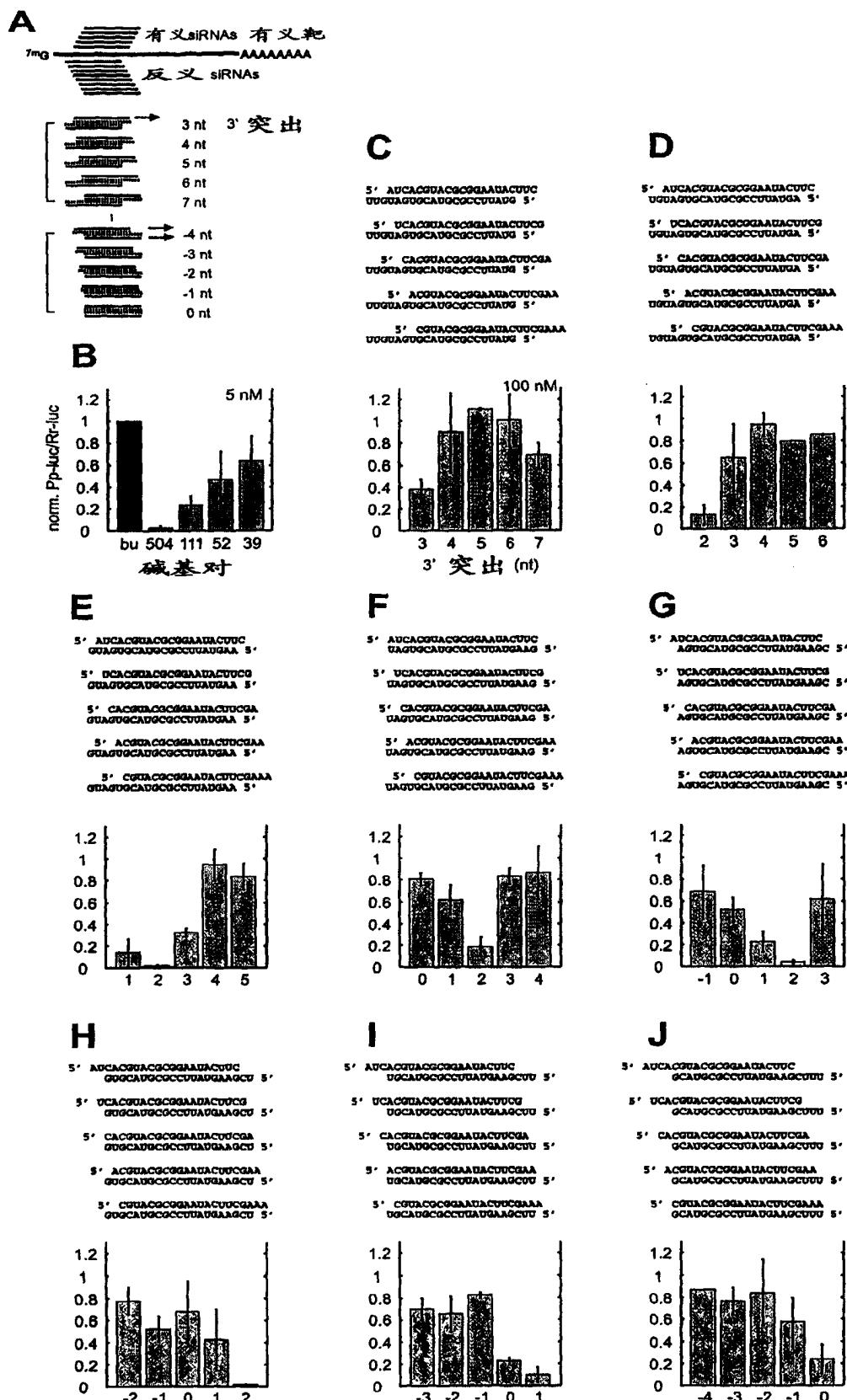
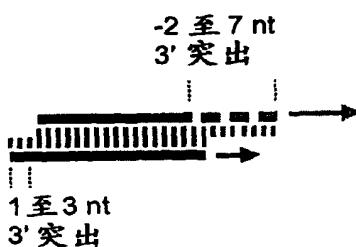


图12

A

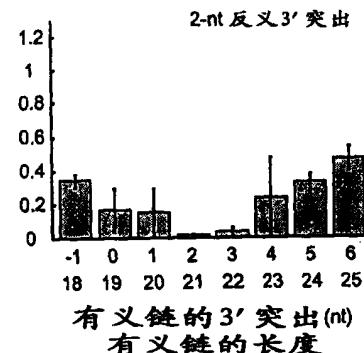
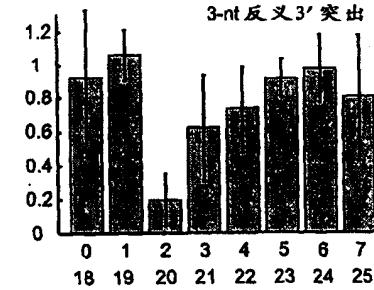
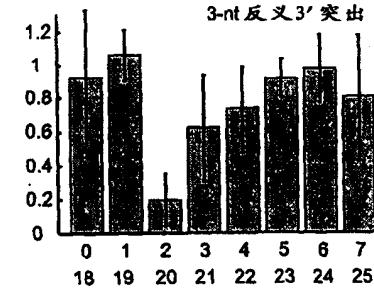
有义 siRNA (18-25 nt)
反义 siRNA (21 nt)

**B**

5' CGUACGCGGAUAUACUUUCG
UGCAUGCGCCUUUAUGAAGCUU 5'
5' CGUACGCGGAUAUACUUCGAA
UGCAUGCGCCUUUAUGAAGCUU 5'
5' CGUACGCGGAUAUACUUCGAA
UGCAUGCGCCUUUAUGAAGCUU 5'
5' CGUACGCGGAUAUACUUCGAAA
UGCAUGCGCCUUUAUGAAGCUU 5'

5' CGUACGCGGAUAUACUUUCG
GUGCAUGCGCCUUUAUGAAGCU 5'
5' CGUACGCGGAUAUACUUCGAA
GUGCAUGCGCCUUUAUGAAGCU 5'
5' CGUACGCGGAUAUACUUCGAA
GUGCAUGCGCCUUUAUGAAGCU 5'
5' CGUACGCGGAUAUACUUCGAAA
GUGCAUGCGCCUUUAUGAAGCU 5'
5' CGUACGCGGAUAUACUUCGAAA
GUGCAUGCGCCUUUAUGAAGCU 5'
5' CGUACGCGGAUAUACUUCGAAA
GUGCAUGCGCCUUUAUGAAGCU 5'
5' CGUACGCGGAUAUACUUCGAAA
GUGCAUGCGCCUUUAUGAAGCU 5'

5' CGUACGCGGAUAUACUUUCG
AGUGCAUGCGCCUUUAUGAAGC 5'
5' CGUACGCGGAUAUACUUCGAA
AGUGCAUGCGCCUUUAUGAAGC 5'
5' CGUACGCGGAUAUACUUCGAA
AGUGCAUGCGCCUUUAUGAAGC 5'
5' CGUACGCGGAUAUACUUCGAAA
AGUGCAUGCGCCUUUAUGAAGC 5'
5' CGUACGCGGAUAUACUUCGAAA
AGUGCAUGCGCCUUUAUGAAGC 5'
5' CGUACGCGGAUAUACUUCGAAA
AGUGCAUGCGCCUUUAUGAAGC 5'
5' CGUACGCGGAUAUACUUCGAAA
AGUGCAUGCGCCUUUAUGAAGC 5'

有义链的 3' 突出(nt)
有义链的长度

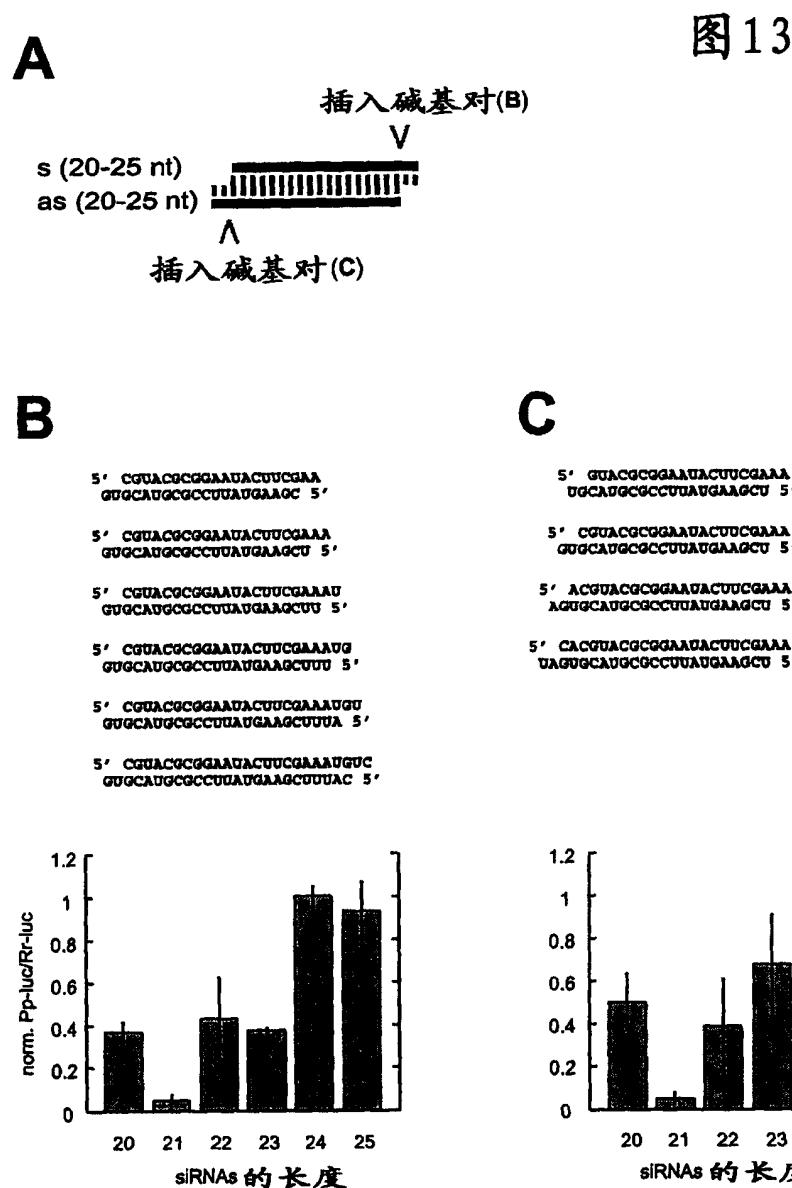


图 14

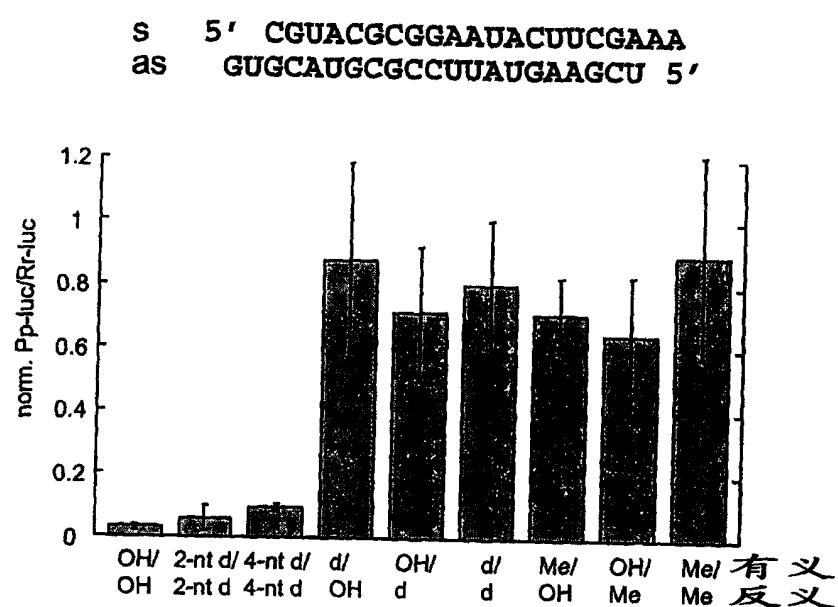


图 15

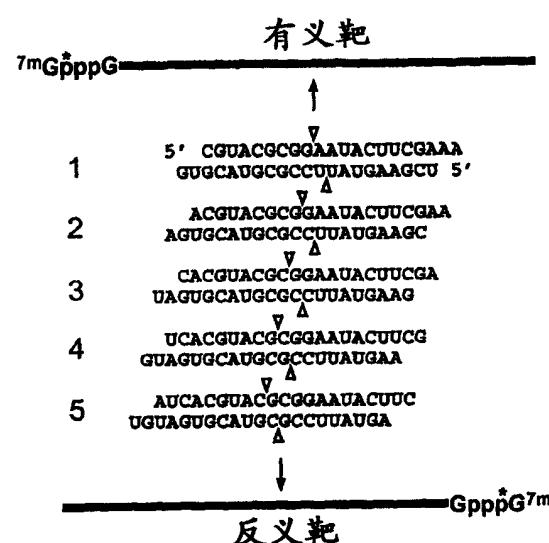
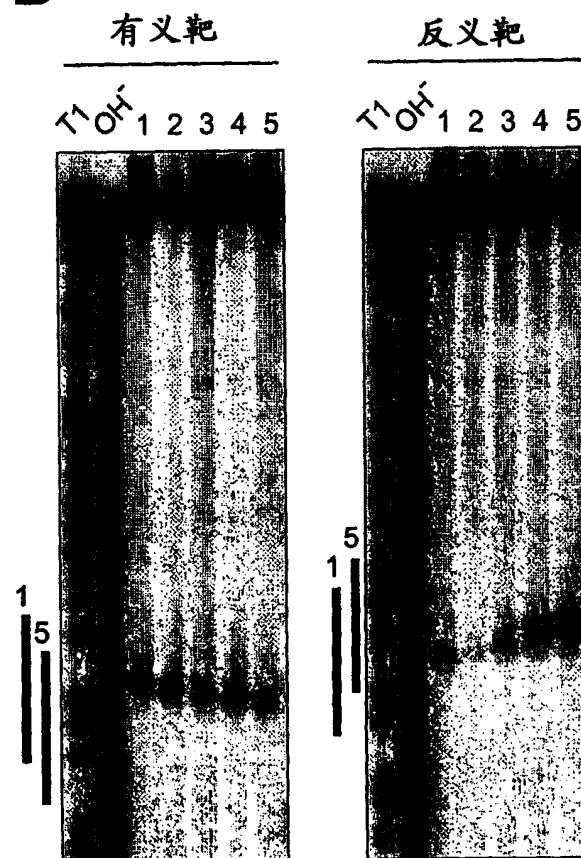
A**B**

图 16

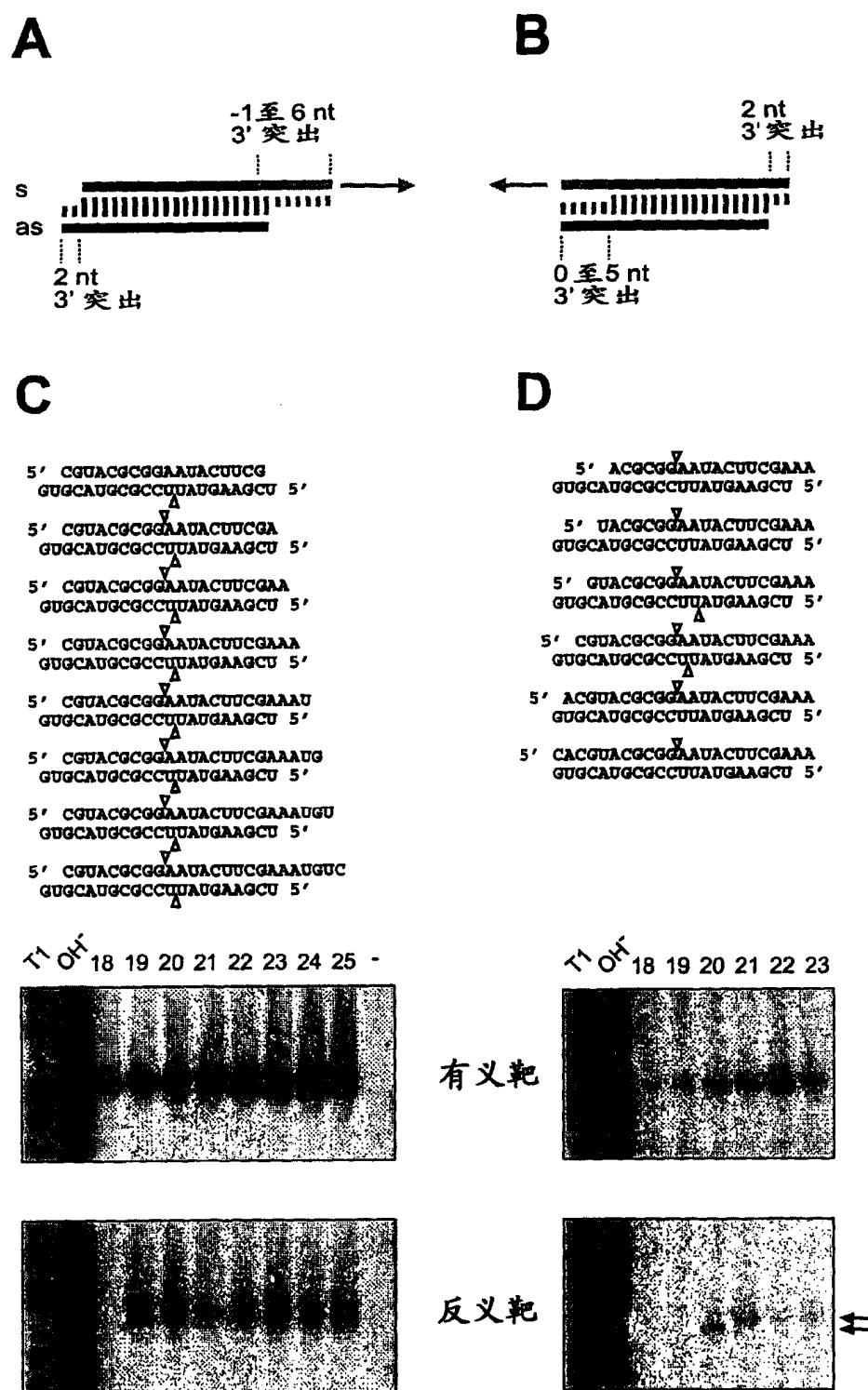
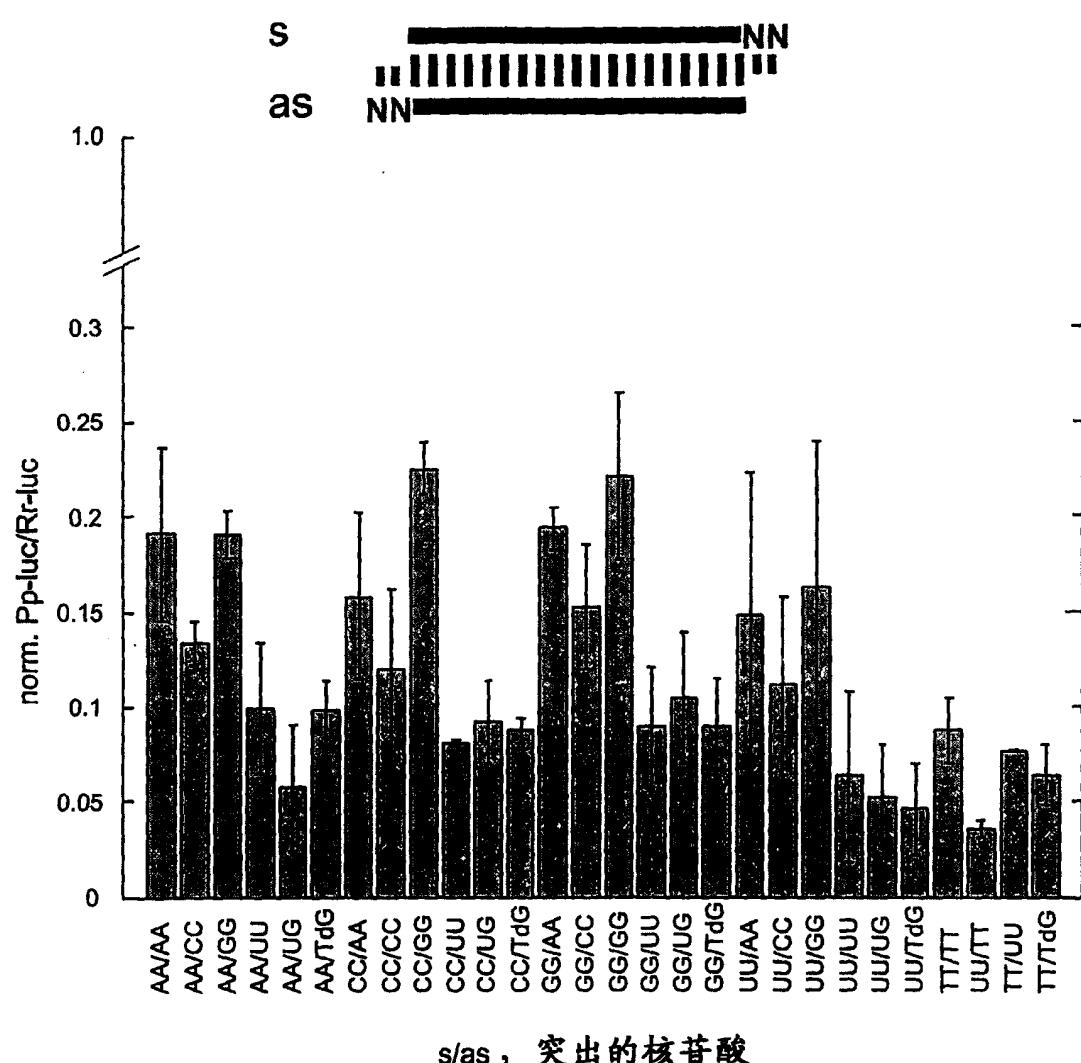


图 17



s/as , 突出的核苷酸

图 18

ref 5' CGUACGCGGAAUACUUUCGATT
TTGCAUGCGCCUUUAUGAAGCU 5'

 1 5' [REDACTED] CGCGGAAUACUUUCGATT
TTGCAUGCGCCUUUAUGAAGCU 5'

 2 5' CGUACGCGGAAUACUUUCGATT
TTGCAU[REDACTED] CUUAUGAAGCU 5'

 3 5' CGUACGCGGAAUACUUUCGATT
TTGCAUGCGC[REDACTED] UGAAGCU 5'

 4 5' CGUACGCGGAAU[REDACTED] CGATT
TTGCAUGCGCCUU[REDACTED] GCU 5'

 5 5' CGUACGCGGAAUACUUUCGATT
TTGCAUGCGCCUU[REDACTED] GCU 5'

 6 5' CGUACGCGGAAUACUUUCGATT
TTGCAUGCGCCUU[REDACTED] GCU 5'

 7 5' CGUACGCGGAAUACUUUCGATT
TTGCAUGCGCCUU[REDACTED] GCU 5'

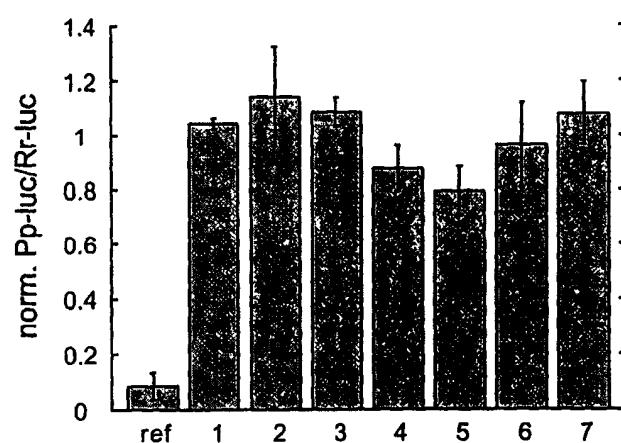


图 19

