

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200510012267.3

[51] Int. Cl.

G06F 19/00 (2006.01)

C12Q 1/68 (2006.01)

C07H 21/02 (2006.01)

A61K 31/7088 (2006.01)

A61K 48/00 (2006.01)

A61P 35/00 (2006.01)

[43] 公开日 2007 年 1 月 31 日

[11] 公开号 CN 1904900A

[51] Int. Cl. (续)

A61P 3/00 (2006.01)

[22] 申请日 2005.7.28

[21] 申请号 200510012267.3

[71] 申请人 中国科学院生物物理研究所

地址 100101 北京市朝阳区北沙滩大屯路 15 号

[72] 发明人 殷勤伟 戴正华 谷同军 冯圣中
丘振戈

[74] 专利代理机构 中科专利商标代理有限责任公司
代理人 程金山

权利要求书 2 页 说明书 88 页 附图 5 页

[54] 发明名称

人类的内源性 siRNA 序列及其应用和筛选方法

[57] 摘要

本发明集 RNA 干扰技术、人体基因库搜索技术、生物计算机信息技术、计算机算法语言和基因工程技术为一体，研发了一种计算机算法语言和相关的应用软件，可以高速地筛选和预测内源性 siRNA 基因和 siRNA 分子。本发明特别涉及一种完整的内源性 siRNA 基因和内源性 siRNA 的发现方法，获得了全新的 255 个来源于编码基因内含子的 siRNAs 和它们的靶 mRNAs。这些 siRNA 分子将可用于研究细胞，组织和器官的发育，分化和生长，探讨基因的功能和表达调控网络，揭示细胞周期控制和细胞凋亡的分子机制。这些 siRNA 分子也可用于研发以相应的 siRNA 为基础的基因药物来预防和治疗多种疾病如病毒感染，肿瘤发生和生长，遗传性和代谢性疾患以及免疫和神经系统疾病。

1. 一种筛选 siRNA 的方法，它包括算法语言设计，计算机软件编写，人体基因文库筛选，内源性 siRNA 预测，和生物学实验鉴定：

1) 用一种新的算法语言来编写筛选和预测内源性 siRNA 基因的软件；
2) 用生物计算机信息技术和人体基因库搜索技术来筛选和预测源于编码基因内含子的 siRNA 的 DNA 序列；

3) 用各种生物学实验对选出的 siRNA 加以鉴定。

2. 一种筛选人体编码基因的内含子中的 siRNA 序列的方法，它包括：

1) 采用动态规划算法从人的内含子序列库中寻找具有反向精确互补的 DNA 结构；

2) 使用多序列排列程式和特定的序列型式测定技术以及人体基因数据库搜索技术来寻找那些位于不同种属的相同基因内的高保守序列；

3) 精选出一段具有 siRNA 典型干环状结构的 DNA 序列；

4) 估价这精选出的 siRNA 序列，标准是采用 ICT-Blast 程序查看这一序列必须有相应的、至少一个以上的靶 mRNA。

3. 权利要求 1 或 2 的方法，其中所述 siRNA 基因片断是一段含有反向精确互补的 DNA 结构，并能转录成 siRNA 前体，经过切割加工互补区能产生一到数个成熟的 siRNA 分子。

4. 权利要求 1 或 2 的方法，其中所述 siRNA 典型干环状结构应含有至少 19-25 个核苷酸长的干茎序列，和 20 个碱基到 120 个碱基不等的环状结构。

5. 权利要求 1 或 2 的方法，其中所述 siRNA 的 DNA 序列是一段编码基因内含子中的 DNA 片段，它能被用来作为合成一个内源性 siRNA 的模板。

6. 权利要求 1 或 2 的方法，其中所述 siRNA 是 19-26 个核苷酸长的单链或双链的 RNA 分子，它们的 3' 段可含有 1-3 个突变的核苷酸碱基。

7. 权利要求 1 或 2 的方法，其中所述 siRNA 分子中的任一或多个核苷酸碱基可被不同的化学基团加以修饰，从而获得较高的稳定性。

8. 权利要求 1 或 2 的方法，其中所述 siRNA 分子能够与一个编码蛋

白质或具有其它生物功能的 RNA 或 DNA 分子特异地杂交，从而干扰其活性或灭活其功能。

9. 权利要求 1 或 2 的方法，其中所述 siRNA 分子的反义链是一系列的能有效地灭活相应的 mRNA 分子的 19-25 个核苷酸长的寡聚核苷酸，包括下列序号的寡聚核苷酸分子：1, 2, 3, 4, 5, 6, 7, 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 到 255。

10. 权利要求 1、2 或 9 的方法，其中所述 siRNA 的 DNA 序列可被整合到质粒或病毒载体中，被增殖和转录加工成成熟的 siRNA 分子。

11. 核苷酸序列，其具有 SEQ ID No. 1-255 中任何一个的核苷酸序列。

12. 权利要求 11 的核苷酸序列在制备和开发治疗肿瘤、遗传性和代谢性疾病等的药物中的应用。

人类的内源性 siRNA 序列及其应用和筛选方法

技术领域

本发明涉及生物信息学,更具体地,本发明涉及一种筛选 siRNA (small interfering RNA, 小分子干扰 RNA) 的方法。本发明还涉及通过该方法筛选的人类的内源性 siRNA 序列及其应用。

背景技术

Micro RNAs (miRNAs) 是一大类~22 核苷酸长的非编码的单链 RNA 分子。在进化上,它们广泛存在于从植物、线虫到人类的细胞中^[1]。最早发现的是 *lin-4* 和它的靶 mRNA, 即 *lin-14*^[2]。Lee 等^[2]1993 年用经典的定位克隆的方法在 *C. elegans*(线虫)中克隆了 *lin-4* 基因, 并通过定点突变发现 *lin-4* 并不编码蛋白, 而是产生一种小 RNA 分子。这种小 RNA 分子能以不完全互补的方式与其靶 mRNA 的 3'非翻译端的特定区域相互作用来抑制 *lin-14* 的表达, 最终导致 *lin-14* 蛋白质合成的减少, 这种现象叫做转译抑制。结果发现 *lin-4* 控制着 *C. elegans* 幼虫由 L1 期向 L2 期的转化。为什么一个只有 22nt 的 RNA 分子, 起着如此重要的调节作用? 当时人们无法解释, 只能认为是一种稀少的个别现象。但是, 2000 年第二个 miRNA *let-7*^[3] 及其人类和果蝇中同源物的发现, 改变了人们的看法, miRNA 可能是一类进化上保守的在生命中起着重要调控作用的分子。它们能有效地抑制相关蛋白质的合成, 能导致靶 mRNA 的降解, 或者其它形式的调节机制来抑制靶基因的表达, 产生基因沉默^[4,5]。近年来发现 miRNA 可能在基因表达调控领域中起着超乎想象的重要作用, miRNA 序列、结构、丰度和表达方式的多样性, 使其可能作为蛋白质编码 mRNA 的强有力的调节子。miRNA 的发现丰富了人们对蛋白质合成控制的认识, 补充了在 RNA 水平对靶 mRNA 分子进行更迅速和有效的调节, 展现了细胞内基因表达调控全方位多层次的网络系统^[6]。miRNA 的发现也是对中心法则中 RNA 次要的中介角色的重要补充, 它将促使生物学家重新思

考细胞遗传调控及其发育等方面的重要问题。

miRNA 的主要功能是细胞在 RNA 水平对基因表达进行调控,其调节与 siRNA 相似,能抑制相关蛋白质的合成,亦能导致靶 mRNA 的裂解,产生基因沉默。近来发现 miRNA 可能在基因表达调控领域中起着超乎想象的重要作用。miRNA 序列、结构、丰度和表达方式的多样性,使其可能作为蛋白质编码 mRNA 的强有力的调节子。miRNA 的发现丰富了人们对蛋白质合成控制的认识,补充了在 RNA 水平对靶 mRNA 分子进行更迅速和有效的调节,展现了细胞内基因表达调控全方位多层次的网络系统。miRNA 的发现也是对中心法则中 RNA 次要的中介角色的重要补充,它将促使生物学家重新思考细胞遗传调控及其发育等方面的重要问题。随着研究的深入,miRNA 将在生命起源和物种进化、基因表达调控的复杂性、疾病发生和发展的机制等方面起到更为深远的作用。同时,miRNA 的研究亦将为 RNAi 技术的应用提供新的依据、思路 and 空间。现在,欧美等发达国家已将有关 miRNA 的研究将广泛而深入地运用到组织器官的定向发育,细胞生长分化的时空调节,信号通路的开启和关闭,细胞周期的监测与调控、学习与记忆,肿瘤的逆分化、肥胖、衰老和死亡,疾病的防治以及有目的的基因表达调控上。这应引起我国科学界的密切关注。

从 miRNA 很容易联想到另外一种小 RNA 分子——siRNA,它们具有很强的关闭基因的功能,能够介导 RNA 干扰(RNA interference ,RNAi)现象,而指导特定的靶 mRNA 降解,siRNA 介导的基因调节系统是细胞内的一个天然的对抗外来遗传物质如病毒,和监控内在基因的时空表达的基因免疫系统。这种系统极有希望成为鉴定基因功能,调控基因表达,和改造基因外显特性的一种战略途径。人们可以仿照大自然的制药模式来设计合成外源性的 19-21 个 siRNA,进行相关基因的药物性治疗。这种 siRNAs 生物战略武器的研发正引起世界研究机构和生物医药公司的高度关注,英国和荷兰联手计划用 RNAi 干扰技术在 2004—2005 年完成全部 35000 个人类基因的沉默。美国计划用 RNAi 干扰技术来 1) 建立 siRNA 分子库:拟找出那些基因的丢失将引起肿瘤,那些基因的丢失将引起肿瘤细胞的死亡、那些基因与人体的脂肪调节有关、那些基因与人体的衰

老有关和那些基因与人体基因组的稳定性有关；2) 筛选治疗肿瘤的新的基因靶点；3) 研发以 RNAi 为基础的基因药物；3) 估价新药的特异性和作用机制。这些研究的直接成果将促使人类对基因表达产物的广泛应用，同时带动农、林、牧、医、药、园艺等相关产业的发展，使得生物产业成为 21 世纪的主导产业。

使人惊讶的是大的医药公司争相加入，如 Novartis, Aventis, GSK, Merck, Ffizer, Abbot Pharmacia, AstraZeneca, P & G, Wyeth, Roche, Bristol-Myers Squibb, Amgen, Schering-Plough, Procter&Gamble, Glaxo-SmithKline, Agouron, Genentech Rigel Pharmaceuticals, 强强联手不断出现，如 Merk-Alnylam, Affymatrix-Qiagen, Lily-ISIS 等。在美国已很难找到一家不使用 RNAi 技术的医药和生物技术公司。时至今日，风险投资对 RNAi 技术已经表现得“垂涎三尺”了。一家位于麻省剑桥的 RNAi 药物开发公司 Alnylam 制药公司成立之初就获得了 1750 万美元的风险投资——要知道，在过去两年里风险投资对刚成立的生物技术公司简直是滴水不施的，而 Alnylam 却能获得如此大笔投资。“毫无疑问，RNA 干扰技术是过去几年中所浮现出来的一项革命性突破，” Alnylam 公司的首席执行官 John Maraganore 说到。John Maraganore 原先是千年制药公司 (Millennium Pharmaceuticals) 一名主要科学家。“RNAi 将成为继重组蛋白和单克隆抗体之后的最主要药物类别。”他说。Ribozyme 公司也获得了 4800 万美元进行 RNAi 的商业化开发。Ribozyme 希望到 2005 年能进行到一期临床试验。Ribozyme 制药公司的首席执行官 Howard Robin 说，RNAi 经过近数年的化学和药理学的研究，“给了公司极大的信心去从事这个最能令人兴奋的医药领域商业化发展”。2004 年 8 月，美国 Acuity 公司向 FDA 提交首个 RNAi 药物的 I 期临床申请。总之，RNAi 基因药物的研发将给生物制药领域带来一场革命性的变革。

有关 RNA 干扰系统的基础研究和医药产业开发已成为世界各国在生物科技领域中抢占的一个最重要的制高点。诺贝尔奖获得者 Sharp 和 Boshor 认为 RNA 干扰研究将是未来十年生物学研究中最激动人心的也是最有可能产生丰富成果领域之一。但至今世界上尚未发现人类内源性 siRNA，所以本发明对人类内源性 siRNA 和其靶基因的揭示是非常重要的

的，它将大大促进小 RNA 的研究，对 RNA 的作用机制提供有价值研究资料，并给人类疾病研究提供新的途径。

发明内容

本发明集 RNA 干扰技术、人体基因库搜索技术、生物计算机信息技术、算法语言和基因工程技术为一体，来研究和开发一类新的内源性的 siRNA 分子。本发明的主要目的在于：

1. 创建一种预测和筛选内源性 siRNA 基因的存在和它们的靶基因的算法语言和方法。
2. 证明编码基因的内含子区存在着内源性 siRNA 分子，它们不同于已发现的 miRNA 分子。
3. 详细地描述 255 个新的内源性 siRNA 分子的序列结构。
4. 生物学验证内源性 siRNA 分子的生物发生。
5. 预测这些内源性 siRNA 分子的靶基因以及它们的可能应用。

发明详述

1. 名词术语

在本发明的上下文中，所用的术语“双链寡聚核苷酸”是一种含有核苷酸的多聚体或寡聚体的双体。如一个双链的 RNA 分子 (dsRNA)，一个双链的 DNA 分子 (dsDNA)，一个双链的 sRNA-cDNA 杂合分子。此术语进一步包括由天然的核苷酸、糖和共价的核苷酸之间的连接，以及经过修饰的或非天然的核苷酸所组成的寡聚核苷酸。这些寡聚体中的每一种类型以及他们的无数的衍生物已被广泛报道。那些被修饰或替代的核苷酸常常优越于自然的核苷酸，如被用来合成相应的寡聚核苷酸，其产物可能有更强的抗酶解性能，更好的被细胞摄入性能和更高的与其靶核酸的亲合性能。

如这里所用的术语“双链的 RNA 分子 (dsRNA)，双链的 DNA 分子 (dsDNA)，双链的 sRNA-cDNA 杂合分子”是指一种核酸双体。它们的每一条链由 19-25 个核苷酸组成。本发明的这种内源性 siRNA 分子能够有效地灭活一个细胞中同源的 RNA 或 DNA 分子。本发明的 siRNA 分子

包括但不限于那些硫代磷酸酯寡聚核苷酸和其它的经过修饰的寡聚核苷酸。

如这里所用的术语“内源性 siRNA 分子”是指一个具有 19-25 个核苷酸长的特定核酸双体。它的正意链与其靶基因的某一 DNA 片段 100% 的相同，它的反义链可与相应的 mRNA 分子杂交，引导 RISC/RITC 特异地抑制这一靶基因分子，而不会沉默其它 RNA 分子的功能。

如这里所用的术语“同源核酸或同源序列”，包括那些能编码蛋白和其它功能的 RNA 的 DNA 分子，从这些 DNA 产生的 RNA 分子包括未成熟的 mRNA，成熟的 mRNA 和其它 RNA 分子以及这些 DNA 分子中的相同的片段。一个内源性 siRNA 分子与靶核酸的相互作用将能影响这个核酸的相应功能。这种由内源性 siRNA 分子介导的靶核酸功能的抑制通常被定义为“RNA 或 DNA 干扰”。RNA 被干扰的功能包括那些 mRNA 的转录，RNA 的转位（从核内到其产生蛋白的地方）RNA 的拼接（产生一种或多种 mRNA 分子），RNA 的转译和其它由 RNA 介导的特殊功能。DNA 被干扰的功能包括，DNA 的复制、转录、修复和重组。这些靶核酸被干扰的最终结果是，合成蛋白或多肽的 mRNA 的降解，其它 RNA 分子的特殊功能的灭活，以及同源 DNA 序列的甲基化。虽然内源性 siRNA 分子能特异地与多个同源的核酸相互作用，具有较高的稳定性和有效性，但本发明的宗旨主要在于揭示人类内源性 siRNA 和其靶基因，促进小 RNA 的研究，对 RNA 的作用机制提供有价值素材，并给人类疾病研究提供新的途径。

2. 算法语言的设计

我们开发出了一种计算机预测的方法，它整合多种算法语言如 Smith-Whatman, HMM, neuronetwork 等于一体，这一算法加入了 RNA 二级结构预测的程序，能够从动力学角度预测 siRNA 基因和 siRNA 分子的稳定性，还加入了系统发育比较，以确定序列的保守性。我们称其为动态系统规划算法。这个方法考虑了以下几个特点：1. 局部位置的反向完全互补配对必须大于 21nt；2. RNA-RNA 双链自由能的预测；3. 靶位点的最低自由能；4. 寻找速度快而准确。

3. 编写预测和筛选软件

内源性 siRNA 是由一个~150nt 的茎环结构前体而来，并且在进化上是保守的，因此可用计算机的方法来识别。我们根据 siRNA 的三个特点开发出来一种新的筛选软件，1.必须来自于编码基因的内含子，产生 siRNA 的保守性，需要形成一个~100nt 的前体；2.在相似的物种中，siRNA 是很保守的；3.在相距较远的物种间，siRNA 是有一定的分歧的。它包括以下三个步骤：1.通过本发明的软件寻找人类编码基因内含子中的反向完全互补配对的序列；2.通过软件辨认此序列是否能形成保守的茎环结构，并给这个结构评价；3.评价 siRNA 在不同物种中的分歧模式。最后，通过生物化学的方法，加以验证。

4. 构建数据库

构建数据库是对内源性 siRNA 进行预测和筛选的必备条件，所以，我们建立了以下相关的数据库：

人体基因库大小 100,000 MB；

内含子数据库规模 100MB；

mRNA 序列库 80MB；

以及其它有关的核酸数据库如：

3'UTR；

5'UTR；

启动子；

MiRNA；

小 RNA 和；

其它物种的 DNA 数据库等。

5. 预测和筛选内源性 siRNA 分子

本发明引入一个能够有效地搜索内源性 siRNA 基因计算方法。鉴定反向精确互补是搜索一个特异的基因组 DNA 序列中是否含有内源性 siRNA 基因的关键步骤。因为内源性 siRNA 基因应含有一个大于 21 核苷酸长的特异序列，它能有效地抑制相应的靶基因的功能。这一序列经过漫长的自然选择与进化的长河，始终高度保留着原有的核苷酸的组成和顺序。因此，内源性 siRNA 基因是作为编码基因调控的一种机制。同时，它也能节省搜索特异的基因组 DNA 序列作为一种基因药物的时间。

虽然人类基因组序列的完成提供了一个研究大多数编码蛋白的基因，tRNA，和 rRNA 的生物信息库，但仍有困难去迅速鉴定那些非编码蛋白的基因。尤其是这些新型的内源性 siRNA 序列长期地被忽视。本发明提出 RNA 的高保守序列，特别是反向精确互补干环状结构可用作为一个重要的识别标志，它可被用来筛选 siRNA 序列。本发明使用我们创建的计算机软件方法如动态规划语言和其它能够预测 RNA 的干环状结构的特殊计算机软件在人体基因组信息库中搜索 siRNA 分子。此外，其他多种方法也是有用的，如计算机寻找具有小分子 RNA 所共有的特征性的结构，微阵列测定在基因间和它们的内含子区内的互补序列，和从与 RNA 结合的蛋白质如脱胺酶（一个 siRNA 结合蛋白）中分离 RNA 分子。总之，在人体基因组信息库中搜索是否存在相应的内源性 siRNA 分子是研究小 RNA 的功能必要步骤。

siRNA 被定义为一类 RNA 分子，它们并不编码一个完整的开放阅读框架，这些来自不同种属的 RNA 分子具有高保守的序列。根据序列保守程度随机地从非编码基因区域或基因的内含子区域内筛选 siRNA 序列是一种可行的方法。因此，本发明提出筛选 siRNA 序列标准是：

RNA 的干环状结构，
高保守的干部区域，
干部区域的长度为 23—45 个核苷酸，
其前体可产生一至几个成熟的 siRNA，
此序列位于编码基因的内含子区。

所有可能的 siRNA 分子都来源于基因间区或基因的内含子区，至今，有关所有种属的基因内含子区的 DNA 序列信息和资料尚不完整或完全缺少。这构成了搜索相应 siRNA 序列的主要障碍。然而，现在我们已研发了专门的软件。这些专门的软件采用的原理是大家熟知的：那就是，一个核苷酸的第一区域，精确互补于同一核苷酸的第二区域，这两个区域以反向平行的方式排列。第一区域中的一个核苷酸能够与第二区域内的核苷酸配对。当第一和第二区域以反向平行的方式排列时，第一区域内至少有 95% 的核苷酸与第二区域内的核苷酸配对。这两个区域的长度大约跨越 23—45 个核苷酸长度。最理想的是第一区域中的所有核苷酸与第

二区域中的所有核苷酸互补。如第一区域中的一个腺嘌呤与第二区域中的胸腺嘧啶或尿嘧啶以反向平行的方式形成氢键连接。相似地，第一区域中的一个胞嘧啶能与第二区域中的鸟嘌呤也以反向平行的方式相配对。根据序列保守程度和保守序列的长度，对来自于人体、线虫和果蝇的这一编码基因的内含子区的序列进行测定，结果发现有两类核苷酸的高度保守序列，一类含有 23 个核苷酸长的精确互补序。另一类含有 45 个核苷酸长的精确互补序列。但这两类均含有一个大小不等的环，由 20 到 100 个核苷酸组成。由此可见，编码基因内含子区的一个干环状结构和大于 21 核苷酸长度的高保守 DNA 序列可以作为鉴定一个 siRNA 序列的标准。

6. 靶基因

虽然现在已经找到了过千个 miRNA，并提出它们在细胞增殖、分化、代谢与死亡中发挥着重要的调节作用，但是至今为止，真正确认功能的 miRNA 还是微乎其微，从 1993 年发现第一个 miRNA 到现在，在动物中发现的确认功能的 miRNA 不超过十个。然而，我们搜索了人体 mRNA 数据库发现每个内源性 siRNA 分子都具有一个以上的靶基因，大多数的靶基因可达一百多个，采用 ICT-Blast 程序对它们靶基因进行了进一步的验证。现列举一个实例如下：

表 1. SEQ ID No. 1 的靶基因。

Access No	Gene Name	Region
NM_012468.3	T-cell leukemia/lymphoma 6	5'UTR(861-878)
NM_018672.2	ATP-binding cassette, sub-family A	5'UTR(85-101)
AF083390	WHSC1 protein (WHSC1) mRNA, alternative splice	5'UTR(39-57)
NM_014919.1	Wolf-Hirsch horm syndrome candidate 1	5'UTR(38-55)
NM_198506.1	FLJ44691 protein (FLJ44691),	5'UTR(38-55)
NM_014925.2	KIAA1002 protein	5'UTR(205-222)
XM_498589.1	Hypothetical gene supported by AK128797	5'UTR(2029-2046)
XM_371108.3	Similar to KIAA1314 protein (LOC388462),	5'UTR(1498-1615)
NM_032378	Eukaryotic translation elongation factor 1 delta(EEF1D)	5'UTR(136-153)
AF452411.1	AKT2	5'UTR(72-89)
AF148213	Aggrecanase-1	3'UTR(4085-4102)
NM_130786.2	Alpha-1-B glycoprotein (A1BG)	3' UTR (3136-3153)
NM_017873.2	Ankyrin repeat and SOCS box-containing 6	3'UTR(3063-3080)
NM_181869	APAF1	3'UTR(2405-2512)
NM_003661	APOL1	3'UTR(1666-1683)
NM_030882	APOL2	3'UTR(1674-1690)
NM_053286.1	Aquaporin 6, kidney specific (AQP6),	3'UTR (4428-4445)
AF084555	ARPP-19	3'UTR(4147-4164)
AF082283	BCL10	3'UTR(2390-2407)

NM_001745	CAMLG	3'UTR (1981-1998)
AF196175	Capsaicin receptor	3'UTR(3292-3308)
NM_014959	CARD8	3'UTR(4260-4276)
NM_033357.1	Caspase 8, apoptosis-related cysteine protease	3'UTR(2143-2160)
NM_033355.1	Caspase 8, apoptosis-related cysteine protease	3'UTR(2018-2035)
NM_033356.1	Caspase 8, apoptosis-related cysteine protease	3'UTR(1794-1810)
NM_033332.1	CDC14 cell division cycle 14 homolog B	3'UTR (2682-2699)
NM_138413.2	Chromosome 10 open reading frame 65	3' UTR (1922-1939)
NM_014356.2	Chromosome 6 open reading frame 123 (C6orf123)	3' UTR (785-802)
NM_017437.1	Cleavage and polyadenylation specific factor 2	3'UTR(3988-4015)
AF439325	CRF2/12	3'UTR(4026-4043)
NM_000761.2	Cytochrome D450, family 1	3' UTR(2588-2605)
AF313469	Dendritic cell-associated C-type lectin-1 beta	3'UTR(1148-1165)
NM_024423.1	Desmocollin 3	3' UTR(5038-5055)
NM_001941.2	Desmocollin 3 (DSC3), transcript variant Dsc3a	3' UTR (4994-5010)
AB037161	DSCR6d	3'UTR(1159-1176)
NM_022549.2	Fasciculation and elongation protein zeta 1	3'UTR (1484-1501)
NM_012174.1	F-box and WD-40 domain protein 8 (FBXW8),	3' UTR (3372-3389)
NM_133279.1	FC fragment of IgA	3'UTR (1310-1327)

7. 生物实验证明

见发明的实施例及最佳实施方案。

同现有技术相比所有的优点

同现有技术相比本发明所具有的优点包括：

全新的设计和研制理论：编码基因的内含子中存在着一一定数目的内源性 siRNA 分子，这种分子具有完全反向互补的大于 21 个碱基对的干茎结构，20—100 个碱基构成的环样结构，并有完全互补的大于 21 个碱基长的靶 mRNA。

全新的序列结构：与现有的 ncRNA 和其它小 RNA 包括 miRNA 数据库的 RNA 的序列进行比较，所有的小 RNA 均是新发现的 siRNA 分子。

新的与靶基因的作用模式：与内源性 miRNA 分子不一样，每个内源性 siRNA 分子都具有一个以上的靶 mRNA，并有不同部位靶点如操纵子区，3'UTR，5'UTR，内含子区，和不同的作用模式如操纵子区 DNA 甲基化，降解切割，转译抑制等。

附图说明

图 1. 小 RNA (siRNA/miRNA) 介导的基因对话网络。

图 2. 存放新发现的 siRNA 分子的原始资料表。

图 3. 内源性 siRNA 的前体及成熟的 siRNA

图 4. 内源性 siRNA 对基因操纵子区 DNA 甲基化的作用。

图 5. 内源性 siRNA 作用于一个 mRNA 的 5'UTR 靶区。

图 6. 内源性 siRNA 作用于一个 mRNA 的 3'UTR 靶区。

图 7. 内源性 siRNA 作用于一个 pre-mRNA 的内含子靶区

具体实施方案

下面结合附图对理解本发明作进一步的详细描述，但并非对本发明作限定。

1. 材料与方法

1.1 生物信息学

使用动态规划算法及其我们自主设计和编写的软件在中国科学院计算机研究所的曙光 4000H 单处理器上,从人体内含子数据库中筛选和预测内源性 siRNA 的序列结构。此外,从人体 mRNA 序列库中搜索具有给定特征序列的把序列,再用其它生物信息软件加以验证。关于序列的研究和分析,我们使用的有全文数据库(<http://www.ensembl.org>), NIBC 在线资源 (<http://www.nibc.nlm.nih.gov>), DOE 连接的基因组学会网址 (<http://www.igi.doe.gov>), UCSC 基因组生物信息学网址 (<http://genome.ucsc.edu>), 欧洲生物信息学会的网址 (<http://www.ebi.ac.uk/clustalw/>) 和 折 叠 网 服 务 器 (<http://www.bioinfo.rpi.edu/applications/mfold/old/rna>)。

1.2 样品基因组和 siRNA 质粒的提取与纯化

基因组 DNA 在本实验室利用常规的酚 / 氯仿抽提制备, siRNA 质粒用碱裂解方法提取。RNAi-Ready pSIREN-DNR-DsRed-Express (BD Biosciences Clontech) 载体是用重组 pRNA 载体编码的小 RNA, 这个小 RNA 来自编码基因内含子。用化学法合成寡核苷酸并插入进载体, 此外, 我们也构建了带有荧光素酶或含有荧光蛋白报告基因的复合载体。

1.3 细胞的培养和转染

人体乳腺癌细胞系 MCF-7 和人体宫颈肿瘤细胞系 Hela 是从 ATCC (Rockville, MD) 获得的。细胞在 DMEM 中生长, 培养基中补充有 100U/ml

的青霉素和 100 $\mu\text{g}/\text{ml}$ 的链霉素, 10%热不稳定的 FBS 在 37 $^{\circ}\text{C}$ 下的加湿空气中含有 5%的 CO_2 。瘤细胞培养在六孔盘的每个孔中, 加工后的载体缓冲器独立控制各组。再 Prna 载体 (1 μg) 或空载体中转染, 用 lipofectamine2000 依照产品说明去进行 (Life Technologies, New England)。

1.4 引物设计和探针标记

根据人有关基因的序列的 5'端设计一对引物, PCR 产物经克隆测序鉴定后用于探针标记。引物由上海生工生物工程有限公司合成。引物序列根据所研究的基因来确定 pF 和 pR。用于成熟 siRNA 的 Southern 杂交的一般程序是提取总小 RNA, 脉冲电泳转移到尼龙膜上, 与 PCR-DIG₃ 或同位素标记的探针杂交液杂交。

总 RNA 用 TRI REAGENT (Molecular Research Center) 分离。cDNA 由 2 μg 的总 RNA 通过反转录在 50 $^{\circ}\text{C}$ 下用 SUPERSCRIPT II 反转录酶作用 30min 运送出来的。其结果使 cDNA 主要通过 PCR 扩增和特定引物 (1 μM) 在 20 μl 混合物中。通常, 用 Promega 公司的 TaqDNA 聚合酶进行 35 次扩增循环, PCR 产物加入 1-2%的琼脂进行凝胶电泳。循环规律在每个实验的 PCR 反应线性阶段确定实现最优化。

1.5 荧光素酶分析

双荧光素酶分析是根据制造商的使用指导, 用一个 OptoompI 照度计 (MGM Instruments) 在转染后 24 小时进行的。插入内源性 siRNA 同源片段的 Rr-luc (Promega) 目标载体与只表达内源性 siRNA 的载体共转染细胞, Pp-luc 目标载体共转染 24 小时后得到转染产物, 并可以连续分析荧光素酶的活性, 鉴定内源性 siRNA 对靶基因的抑制作用。

1.6 Western blotting

蛋白质用以前的方法来提取。总蛋白 (40 μg) 在 10%SDS-聚丙烯酰胺凝胶电泳下分解并有蛋白转运移动的条带跑到聚氯乙烯膜 (Amersham)。膜的封闭用 5%的脱脂乳 TBS 在摇床上过夜, 孵育引物抗体 2 小时。表达的 Actin 基因是用做加载调控的, 是用的抗体包括 CDC6, Actin 和 EEL 等。膜在 1 小时后产生 HRP 偶联的羊抗鼠或兔抗羊第二代抗体。免疫复合物通过放射自显影技术显现。

1.7 甲基化分析

靶基因启动子上的 CpG 岛的甲基化状况分析是通过是甲基化的特殊 PCR(MSP)分析方法如前所述。所有样品中的 DNA 基因组都分离出了 DNAzoL 试剂 (Life Technologies)。MSP 引物根据在靶基因转录起始位点周围的基因组序由 Methprimer(<http://www.urogene.org/methprimer/index1.html>) 查得。亚硫酸盐修饰 DNA,用 PCR 扩增引物来区别甲基化和非甲基化 DNA。Promage 公司的 DNA 聚合酶用于全放大。45 个 PCR 循环可进行更远的分析。

2 结果和讨论

我们假设存在着一个小 RNA (siRNA/miRNA) 介导的基因对话网络。小 RNA 可由内源性基因编码,位于结构基因间区或结构基因的内含子或外显子区,通过 RNA 聚合酶而被转录,其前体含有帽状结构和多聚腺苷酸。簇化的 siRNA/miRNA 前体可由 spliosome 和 Drosha 酶从 pri-miRNA 剪出,在 Dicer 酶的作用下形成成熟的 miRNA 分子。单一的 miRNA 能结合和调节多个不同的 mRNA 分子。同样,多个不同的 siRNA/miRNA 亦可与同一个 mRNA 分子相结合并调控其生物活性。随着编码基因的表达,从内含子和外显子剪切出的 siRNA/miRNA 可以与其上下游的靶子 mRNA 的不同结构区域结合,从而指导转译抑制或切割降解 (图 1)。

为了证明人体编码基因的内含子内存在着 siRNA 调节分子,我们使用动态规划算法及其我们自主设计和编写的软件在中国科学院计算机研究所的曙光 4000H 单处理器上对人体内含子数据库中的 DNA 序列进行了系统的大规模搜索。第一次发现了 255 个新的内源性小 RNA 分子。结果如图 2 所示。

为了证明我们发现的人体编码基因内含子内的 255 个新的内源性小 RNA 分子确实存在,我们用生物学方法研究了它们的发生。图 3 列举了一个实例。Northern Blotting 能够检测到一个来源于 BRCA1 内含子的成熟的 siRNA 的存在 (SEQ ID No. 1)。

为了证明这个来源于 BRCA1 内含子的成熟的 siRNA 具有相应的靶 mRNA 分子,我们搜索了人体 mRNA 数据库,结果发现了大约 130 个靶子 mRNA (表 1)。分析靶序列所在 mRNA 的位子,发现这个 siRNA 能

作用于不同基因的不同位点，包括操纵子区，5'UTR，3'UTR，和内含子区。

为了证明这个 siRNA 作用于不同基因的不同位点能产生不同的基因沉默效应，我们用 MSP 等方法检测了 telomerase 基因表达的状况。于对照组相比，转染 siRNA 的细胞，telomerase 基因表达受到明显抑制（图 4）。同时，用 RT-PCR 检测到在 5'UTR 区段具有靶位点的 ELL 和基因表达也明显下降（图 5）。

此外，用 RT-PCR 对在内含子和 3'UTR 区段具有靶位点的基因表达进行了检测（图 6 和图 7），与对照组相比，没有可检出的差异。继而，我们用 Luciferase 活性高低变化实验显示这个 siRNA 能够对含有其同源序列的 3'UTR 的靶子 mRNA 分子进行抑制，差异高达 50%左右。

参考文献

- 1 David P B. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell*, 2004,116:281~297
- 2 Lee, R C, Feinbaum R L, Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*,1993, 75: 843~854
- 3 Reinhart B J, Slack F J, Basson M, et al. The 21 nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature*, 2000,403: 901~906
- 4 Yin J Q, Wang Y. siRNA-mediated gene regulation system: now and the future. *Int J Mol Med*, 2002, 10: 355~65.
- 5 Tan F L, Yin J Q. Application of RNAi to cancer research and therapy. *Front Biosci*, 2005, 01, 10:1946~1960.
- 6 Yin J Q, Jiang L. MiRNA-directed gene intertalk. *Chin Acad Forum*, 2004, 11 (20) :1-4
- 7 Lee Y, Kim M, Han J, et al. MicroRNA genes are transcribed by RNA polymerase II. *EMBO J*, 2004, 23(20):4051~60.
- 8 Rodriguez A, Griffiths-Jones S, Ashurst JL, et al. Identification of mammalian microRNA host genes and transcription units. *Genome Res*, 2004, 14(10A): 1902~10.
- 9 Lee Y, Ahn C, Han J, et al. The nuclear RNase III Droscha initiates microRNA processing. *Nature*, 2003, 425: 415-419
- 10 Ahmet M.Denli, Bastiaan B.J.Tops, Ronald H.A.Plasterk, et.al. Processing of primary microRNAs by the Microprocessor complex. *Nature*, 2004, 432:231~5
- 11 Gregory R.I., Yan K, Amuthan G, et.al. The microprocessor complex mediates the genesis of microRNAs. *Nature*, 2004, 432:235~40
- 12 HEN Fang & YIN Q. James (2005) Gene expression regulator, MicroRNAs. *Chinese Science Bulletin* 50 (15) 1—12.
- 13 Tan FL& Yin JQ. RNAi, a new therapeutic strategy against viral infection. *Cell Res*. 2004 Dec, 14(6): 460-6.
- 14 殷勤伟 (2004) siRNA 介导的基因沉默。 *生物物理学报* 20, 85—90
- 15 Liu TG, Yin JQ, Shang BY, Min Z, He HW, Jiang JM, Chen F, Zhen YS, Shao RG. Silencing of *hdm2* oncogene by siRNA inhibits p53-dependent human breast cancer. *Cancer Gene The.* 2004 Nov 11(11): 748-56.
- 16 Ma JP, Zhan WH, Wang JP, Peng JS, Gao JS, Yin Q. Specific inhibition of hTERT gene expression by short interfering RNAs in gastric cancer SGC7901 cell. *Zhonghua Wai Ke Za Zhi*. 2004 Nov 22;42(22):1372-6.
- 17 刘铁刚, 尚伯杨, 殷勤伟, 张敏, 何红伟, 甄永苏, 邵荣光. (2004) 多种 siRNA 的抗肿瘤作用研究. *中国新药杂志* 13;6-10

-
- 18 谭余良, 殷勤伟. (2005) SiRNA: 一种新型的治疗肿瘤的药物. *药学学报*. 40 (3): 193-198
 - 19 谭余良, 殷勤伟. (2005) RNAi 干扰技术在基因缺陷模型中的应用. *遗传学报*. 32(4):5-9

序列表

<110> 中科院生物物理研究所
殷勤伟

<120> 人类的内源性 siRNA 序列及其应用和筛选方法

<130> siRNA

<160> 255

<170> PatentIn version 3.3

<210> 1

<211> 22

<212> DNA

<213> 人

<220>

<221> 核苷酸

<222> (1)..(22)

<400> 1

ggcctcccaa agtgctggga tt

22

<210> 2

<211> 22

<212> DNA

<213> 人

<400> 2

aatcccagca cttgggagg cc

22

<210> 3

<211> 52

<212> DNA

<213> 人

<400> 3

ccgcctcgg cctccaaag tgctgggatt acaggcgtga gccaccgcgc cc

52

<210> 4

<211> 52

<212> DNA

<213> 人

<400> 4

gggcgcggtg gctcacgcct gtaatccag cactttggga ggccgaggcg gg 52

<210> 5

<211> 26

<212> DNA

<213> 人

<400> 5

gcctccaaa gtgctgggat tacagg 26

<210> 6

<211> 26

<212> DNA

<213> 人

<400> 6

cctgtaatcc cagcacttg ggaggc 26

<210> 7

<211> 40

<212> DNA

<213> 人

<400> 7

gtacacttg attgtacac acttttagtg tacaatata 40

<210> 8

<211> 40

<212> DNA

<213> 人

<400> 8

tatattgta cactaaaagt gatgtacaat ccaagtgtac 40

<210> 9	
<211> 22	
<212> DNA	
<213> 人	
<400> 9	
tacaattgta tcataaatat at	22
<210> 10	
<211> 22	
<212> DNA	
<213> 人	
<400> 10	
atatattat gatacaattg ta	22
<210> 11	
<211> 26	
<212> DNA	
<213> 人	
<400> 11	
gttttatgg ttttaggtct aacgtt	26
<210> 12	
<211> 26	
<212> DNA	
<213> 人	
<400> 12	
aacgtagac ctaaaccat aaaaac	26
<210> 13	
<211> 24	
<212> DNA	
<213> 人	
<400> 13	
ctcccaaagt gctgggatta cagg	24

<210> 14	
<211> 24	
<212> DNA	
<213> 人	
<400> 14	
cctgtaatcc cagcacttg ggag	24
<210> 15	
<211> 24	
<212> DNA	
<213> 人	
<400> 15	
ctcccaaagt gctgggatta cagg	24
<210> 16	
<211> 24	
<212> DNA	
<213> 人	
<400> 16	
cctgtaatcc cagcacttg ggag	24
<210> 17	
<211> 23	
<212> DNA	
<213> 人	
<400> 17	
ccaggctggt ctgaactcc tgg	23
<210> 18	
<211> 23	
<212> DNA	
<213> 人	
<400> 18	
ccaggagttc aagaccagcc tgg	23

<210> 19	
<211> 32	
<212> DNA	
<213> 人	
<400> 19	
cccgcctcgg cctcccaaag tgctgggatt ac	32
<210> 20	
<211> 32	
<212> DNA	
<213> 人	
<400> 20	
gtaatcccag cacittggga ggccgaggcg gg	32
<210> 21	
<211> 31	
<212> DNA	
<213> 人	
<400> 21	
ccttggcctc ccaaagtgct gggattacag g	31
<210> 22	
<211> 31	
<212> DNA	
<213> 人	
<400> 22	
cctgtaatcc cagcacttgg ggaggccaag g	31
<210> 23	
<211> 22	
<212> DNA	
<213> 人	
<400> 23	
gcctccaaa gtgctgggat ta	22

<210> 24

<211> 22

<212> DNA

<213> 人

<400> 24

taatcccagc actttgggag gc

22

<210> 25

<211> 24

<212> DNA

<213> 人

<400> 25

tcccaaagtg ctgggattac aggc

24

<210> 26

<211> 24

<212> DNA

<213> 人

<400> 26

gcctgtaatc ccagcacttt ggga

24

<210> 27

<211> 22

<212> DNA

<213> 人

<400> 27

gtgctgggat tacaggcatg ag

22

<210> 28

<211> 22

<212> DNA

<213> 人

<400> 28

ctcatgcctg taatcccagc ac

22

<210> 29	
<211> 21	
<212> DNA	
<213> 人	
<400> 29	
tatatatgta tacatatata t	21
<210> 30	
<211> 21	
<212> DNA	
<213> 人	
<400> 30	
atatatatgt atacatatat a	21
<210> 31	
<211> 21	
<212> DNA	
<213> 人	
<400> 31	
atatatatgt atacatatat a	21
<210> 32	
<211> 21	
<212> DNA	
<213> 人	
<400> 32	
tatatatgta tacatatata t	21
<210> 33	
<211> 21	
<212> DNA	
<213> 人	
<400> 33	
tatatatgta tacatatata t	21

<210> 34	
<211> 21	
<212> DNA	
<213> 人	
<400> 34	
atatatatgt atacatatat a	21
<210> 35	
<211> 21	
<212> DNA	
<213> 人	
<400> 35	
tatatatgta tacatatata t	21
<210> 36	
<211> 21	
<212> DNA	
<213> 人	
<400> 36	
atatatatgt atacatatat a	21
<210> 37	
<211> 21	
<212> DNA	
<213> 人	
<400> 37	
tatatatgta tacatatata t	21
<210> 38	
<211> 21	
<212> DNA	
<213> 人	
<400> 38	
atatatatgt atacatatat a	21

<210> 39

<211> 22

<212> DNA

<213> 人

<400> 39

ctttcactta gcataatgtt tt

22

<210> 40

<211> 22

<212> DNA

<213> 人

<400> 40

aaaacattat gctaagtga ag

22

<210> 41

<211> 48

<212> DNA

<213> 人

<400> 41

gcggtgagcc gagatcgcgc cactgcactc cagcctgggc gacagagc

48

<210> 42

<211> 48

<212> DNA

<213> 人

<400> 42

gctctgtcgc ccaggctgga gtgcagtggc gcgatctcgg ctcaccgc

48

<210> 43

<211> 21

<212> DNA

<213> 人

<400> 43

cctcccaaag tgctgggatt a

21

<210> 44	
<211> 21	
<212> DNA	
<213> 人	
<400> 44	
taatcccagc actttgggag g	21
<210> 45	
<211> 28	
<212> DNA	
<213> 人	
<400> 45	
ggcctcccaa agtgctggga ttacaggc	28
<210> 46	
<211> 28	
<212> DNA	
<213> 人	
<400> 46	
gcctgtaatc ccagacttt gggaggcc	28
<210> 47	
<211> 39	
<212> DNA	
<213> 人	
<400> 47	
gcctcccaaa gtgctgggat tacaggcgtg agccaccgc	39
<210> 48	
<211> 39	
<212> DNA	
<213> 人	
<400> 48	
gcggtggctc acgcctgtaa tcccagact ttgggaggc	39

<210> 49
<211> 24
<212> DNA
<213> 人

<400> 49
agcctcccaa agtgctggga ttac 24

<400> 50
gtaatcccag cacttgga ggct 24

<210> 51
<211> 22
<212> DNA
<213> 人

<400> 51
gggattacag gcgtgagcca cc 22

<210> 52
<211> 22
<212> DNA
<213> 人

<400> 52
ggtggctcac gcctgtaac cc 22

<210> 53

<211> 23

<212> DNA

<213> 人

<400> 53

acagagcgag actccgtctc aaa

23

<210> 54

<211> 23

<212> DNA

<213> 人

<400> 54

tttgagacgg agtctcgctc tgt

23

<210> 55

<211> 46

<212> DNA

<213> 人

<400> 55

gcctggccag gttttatgct ttttaagaca atttgggct gggcac

46

<210> 56

<211> 46

<212> DNA

<213> 人

<400> 56

gtgcccagcc caaaattgct taaaaagca taaacctgg ccaggc

46

<210> 57

<211> 23

<212> DNA

<213> 人

<400> 57

cctcggcctc ccaaagtgct ggg

23

<210> 58

<211> 23

<212> DNA

<213> 人

<400> 58

cccagcactt tgggaggccg agg

23

<210> 59

<211> 23

<212> DNA

<213> 人

<400> 59

cctcggcctc ccaaagtgct ggg

23

<210> 60

<211> 23

<212> DNA

<213> 人

<400> 60

cccagcactt tgggaggccg agg

23

<210> 61

<211> 24

<212> DNA

<213> 人

<400> 61

taatccgcc gcctcggcct ccca

24

<210> 62

<211> 24

<212> DNA

<213> 人

<400> 62

tgggaggccg aggcgggcgg atta

24

<210> 63

<211> 28

<212> DNA

<213> 人

<400> 63

gcctccaaa gtgctgggat tacaggcg

28

<210> 64

<211> 28

<212> DNA

<213> 人

<400> 64

cgctgtaat cccagcactt tgggaggc

28

<210> 65

<211> 25

<212> DNA

<213> 人

<400> 65

ctccactaaa aatacaaaaa ttagc

25

<210> 66

<211> 25

<212> DNA

<213> 人

<400> 66

gctaatttt gtatttttag tggag

25

<210> 67

<211> 30

<212> DNA

<213> 人

<400> 67

ctcggcctcc caaagtgctg ggattacagg

30

<210> 68

<211> 30

<212> DNA

<213> 人

<400> 68

cctgtaatcc cagcactttg ggaggccgag

30

<210> 69

<211> 26

<212> DNA

<213> 人

<400> 69

gcctcccaaa gtgctgggat tacagg

26

<210> 70

<211> 26

<212> DNA

<213> 人

<400> 70

cctgtaatcc cagcacttig ggaggc

26

<210> 71

<211> 21

<212> DNA

<213> 人

<400> 71

gcctccaaa gtgctgggat t

21

<210> 72

<211> 21

<212> DNA

<213> 人

<400> 72

aatcccagca ctttgggagg c

21

<210> 73

<211> 25

<212> DNA

<213> 人

<400> 73

ctccactaaa aatacaaaaa ttagc

25

<210> 74

<211> 25

<212> DNA

<213> 人

<400> 74

gctaattttt gtatttttag tggag

25

<210> 75

<211> 25

<212> DNA

<213> 人

<400> 75

atatatataa atatatttat ataaa

25

<210> 76

<211> 25

<212> DNA

<213> 人

<400> 76

tttatataaa tatattata tatat

25

<210> 77

<211> 24

<212> DNA

<213> 人

<400> 77

ctcccaaagt gctgggatta cagg

24

<210> 78

<211> 24

<212> DNA

<213> 人

<400> 78

ctgtaatcc cagcacttg ggag

24

<210> 79

<211> 23

<212> DNA

<213> 人

<400> 79

cccaaagtc tgggattaca ggc

23

<210> 80

<211> 23

<212> DNA

<213> 人

<400> 80

gcctgtaatc ccagcacttt ggg

23

<210> 81

<211> 24

<212> DNA

<213> 人

<400> 81

gcctcccaaa gtgctgggat taca

24

<210> 82

<211> 24

<212> DNA

<213> 人

<400> 82

tgtaatcca gcacttggg aggc

24

<210> 83

<211> 24

<212> DNA

<213> 人

<400> 83

gcctcccaa gtgctgggat taca

24

<210> 84

<211> 24

<212> DNA

<213> 人

<400> 84

tgtaatcca gcacttggg aggc

24

<210> 85

<211> 23

<212> DNA

<213> 人

<400> 85

tcccaaagt ttgggattac agg

23

<210> 86

<211> 23

<212> DNA

<213> 人

<400> 86

cctgtaatcc caacacttg gga

23

<210> 87

<211> 134

<212> DNA

<213> 人

<400> 87

catttgcgat tatttcatgt aactgaacc tattgcttca aaaatagcac agaacgaatg 60

ttgattctgt ttcttcatt aaaaaatgtg ctactgtgca tttttccca ctctacgac 120

tttaaaggcg cgcg

134

<210> 88

<211> 27

<212> DNA

<213> 人

<400> 88

gctgggatta caggcgtgag ccaccac

27

<210> 89

<211> 27

<212> DNA

<213> 人

<400> 89

gtggtggctc acgcctgtaa tcccagc

27

<210> 90

<211> 21

<212> DNA

<213> 人

<400> 90

tatatatagt agttatatat a

21

<210> 91

<211> 21

<212> DNA

<213> 人

<400> 91

tatatataac tactatatat a

21

<210> 92

<211> 29

<212> DNA

<213> 人

<400> 92

tatattata tataaatata tttatatat

29

<210> 93

<211> 29

<212> DNA

<213> 人

<400> 93

atatataaat atatttatat ataaatata

29

<210> 94

<211> 28

<212> DNA

<213> 人

<400> 94

tcccaaagtg ctgggattac aggcgtga

28

<210> 95

<211> 28

<212> DNA

<213> 人

<400> 95

tcacgcctgt aatcccagca ctttgga

28

<210> 96

<211> 26

<212> DNA

<213> 人

<400> 96

tgctgggatt acaggcgtga gccact

26

<210> 97

<211> 26

<212> DNA

<213> 人

<400> 97

agtggtcac gcctgtaac ccagca

26

<210> 98

<211> 41

<212> DNA

<213> 人

<400> 98

ctcggcctcc caaagtgctg ggattacagg catgagccac t

41

<210> 99

<211> 41

<212> DNA

<213> 人

<400> 99

agtggtcat gcctgtaac ccagcactt gggaggccga g

41

<210> 100

<211> 25

<212> DNA

<213> 人

<400> 100

ctcccaaagt gctgggatta caggc

25

<210> 101

<211> 25

<212> DNA

<213> 人

<400> 101

gcctgtaac ccagcactt gggag

25

<210> 102

<211> 26

<212> DNA

<213> 人

<400> 102

atatattata tgatatatta tattat 26

<210> 103

<211> 26

<212> DNA

<213> 人

<400> 103

ataatataat atatcatata atatat 26

<210> 104

<211> 45

<212> DNA

<213> 人

<400> 104

atatattata ttatatgata tattatatga tatattatat tatat 45

<210> 105

<211> 45

<212> DNA

<213> 人

<400> 105

atataatata atatacata taatatatca tataatataa tatat 45

<210> 106

<211> 40

<212> DNA

<213> 人

<400> 106

atatattata tgatatatta tattatatga tatattatat

40

<210> 107

<211> 40

<212> DNA

<213> 人

<400> 107

atataatata tcatataata taatatatca tataatatat

40

<210> 108

<211> 27

<212> DNA

<213> 人

<400> 108

tatattatat gatattat attatat

27

<210> 109

<211> 27

<212> DNA

<213> 人

<400> 109

atataatata atatacata taatata

27

<210> 110

<211> 25

<212> DNA

<213> 人

<400> 110

atatattata tgatatatta tatta

25

<210> 111

<211> 25

<212> DNA

<213> 人

<400> 111

taatataata tatcatataa tatat

25

<210> 112

<211> 25

<212> DNA

<213> 人

<400> 112

atatattata tgatatatta tatta

25

<210> 113

<211> 25

<212> DNA

<213> 人

<400> 113

taatataata tatcatataa tatat

25

<210> 114

<211> 29

<212> DNA

<213> 人

<400> 114

aagtgctggg attacaggcg tgagccacc

29

<210> 115

<211> 29

<212> DNA

<213> 人

<400> 115

ggtggctcac gcctgtaac ccagcactt

29

<210> 116

<211> 29

<212> DNA

<213> 人

<400> 116

aagtgctggg attacaggcg tgagccacc

29

<210> 117

<211> 29

<212> DNA

<213> 人

<400> 117

ggtggctcac gcctgtaatc ccagcactt

29

<210> 118

<211> 29

<212> DNA

<213> 人

<400> 118

aagtgctggg attacaggcg tgagccacc

29

<210> 119

<211> 29

<212> DNA

<213> 人

<400> 119

ggtggctcac gcctgtaatc ccagcactt

29

<210> 120

<211> 29

<212> DNA

<213> 人

<400> 120

aagtgctggg attacaggcg tgagccacc

29

<210> 121

<211> 29

<212> DNA

<213> 人

<400> 121

ggtggctcac gcctgtaatc ccagcactt

29

<210> 122

<211> 24

<212> DNA

<213> 人

<400> 122

gggttcacc atgttgcca ggct 24

<210> 123

<211> 24

<212> DNA

<213> 人

<400> 123

agcctggcca acatggtgaa accc 24

<210> 124

<211> 23

<212> DNA

<213> 人

<400> 124

agaatatttt ggattcttaa taa 23

<210> 125

<211> 23

<212> DNA

<213> 人

<400> 125

ttattaagaa tccaaaat tct 23

<210> 126

<211> 27

<212> DNA

<213> 人

<400> 126

tggcctccca aagtgctggg attacag

27

<210> 127

<211> 27

<212> DNA

<213> 人

<400> 127

ctgtaatccc agcactttgg gaggcca

27

<210> 128

<211> 21

<212> DNA

<213> 人

<400> 128

ctttaggagg ccgaggcagg a

21

<210> 129

<211> 21

<212> DNA

<213> 人

<400> 129

tcctgcctcg gcctcctaaa g

21

<210> 130

<211> 24

<212> DNA

<213> 人

<400> 130

gcctcggcct cccaaagtgc tggg

24

<210> 131

<211> 24

<212> DNA

<213> 人

<400> 131

cccagcactt tgggaggccg aggc

24

<210> 132

<211> 29

<212> DNA

<213> 人

) <400> 132

ctccagcctg ggcaacaaga gcaagactc

29

<210> 133

<211> 29

<212> DNA

<213> 人

<400> 133

gagtcttgct ctgttgccc aggctggag

29

<210> 134

<211> 29

<212> DNA

<213> 人

<400> 134

ctccagcctg ggcaacaaga gcaagactc

29

<210> 135

<211> 29

<212> DNA

<213> 人

<400> 135

gagtcttgct ctgttgccc aggctggag

29

<210> 136

<211> 41

<212> DNA

<213> 人

<400> 136

ctcttataaa tagtgacca tcaccctata ccactaaat t

41

<210> 137

<211> 41

<212> DNA

<213> 人

<400> 137

aatttagtg gtatagggtg atggccact attataaga g

41

<210> 138

<211> 25

<212> DNA

<213> 人

<400> 138

gattacaggt gtgagccacc gcacc

25

<210> 139

<211> 25

<212> DNA

<213> 人

<400> 139

ggtgcggtgg ctcacacctg taatc

25

<210> 140

<211> 22

<212> DNA

<213> 人

<400> 140

ggcctcccaa agtgctggga tt

22

<210> 141

<211> 22

<212> DNA

<213> 人

<400> 141

aatcccagca ctttgggagg cc

22

<210> 142

<211> 25

<212> DNA

<213> 人

<400> 142

gcctccaaa gtgctgggat tacag

25

<210> 143

<211> 25

<212> DNA

<213> 人

<400> 143

ctgtaatccc agcacttgg gaggc

25

<210> 144

<211> 35

<212> DNA

<213> 人

<400> 144

ggaccacctc ctgtgcccc ctctctgcc atggc

35

<210> 145

<211> 35

<212> DNA

<213> 人

<400> 145

gccatggcag aggagggggc acaaggaggt ggtcc

35

<210> 146

<211> 22

<212> DNA

<213> 人

<400> 146

ggatgagcaa agaaagtggt tt

22

<210> 147

<211> 22

<212> DNA

<213> 人

<400> 147

aaaccacttt ctttgctcat cc

22

<210> 148

<211> 29

<212> DNA

<213> 人

<400> 148

cagcctgggc aacagagtga gactctgtc

29

<210> 149

<211> 29

<212> DNA

<213> 人

<400> 149

gacagagtct cactctgttg cccaggctg

29

<210> 150

<211> 54

<212> DNA

<213> 人

<400> 150

atgtattgta tattacattt gaaatatata ttgaaatata tattgttgag tata

54

<210> 151

<211> 54

<212> DNA

<213> 人

<400> 151

tatactcaac aatatatatt tcaatatata ttccaatgt aatatacaat acat

54

<210> 152

<211> 22

<212> DNA

<213> 人

<400> 152

atattttta tatatataaa aa 22

<210> 153

<211> 22

<212> DNA

<213> 人

<400> 153

ttttatata tataaaaaat at 22

<210> 154

<211> 30

<212> DNA

<213> 人

<400> 154

agtgatatt gtgttgta aattatataa 30

<210> 155

<211> 30

<212> DNA

<213> 人

<400> 155

ttatataatt taacaacac aatatcact 30

<210> 156

<211> 26

<212> DNA

<213> 人

<400> 156

actgcactcc agcctgggtg acagag

26

<210> 157

<211> 26

<212> DNA

<213> 人

<400> 157

ctctgtcacc caggctggag tgcagt

26

<210> 158

<211> 22

<212> DNA

<213> 人

<400> 158

atatgtatat gcatatacat at

22

<210> 159

<211> 22

<212> DNA

<213> 人

<400> 159

atatgtatat gcatatacat at

22

<210> 160

<211> 25

<212> DNA

<213> 人

<400> 160

tgggattaca ggcatgagcc actgt

25

<210> 161

<211> 25

<212> DNA

<213> 人

<400> 161

acagtggtc atgcctgtaa tccca

25

<210> 162

<211> 26

<212> DNA

<213> 人

<400> 162

gcctccaaa gtgctgggat tacagg

26

<210> 163

<211> 26

<212> DNA

<213> 人

<400> 163

cctgtaatcc cagcacttg ggaggc

26

<210> 164

<211> 23

<212> DNA

<213> 人

<400> 164

agcctcctga gtagctggga tta

23

<210> 165

<211> 23

<212> DNA

<213> 人

<400> 165

taatcccagc tactcaggag gct

23

<210> 166

<211> 24

<212> DNA

<213> 人

<400> 166

atatgtatac atatgcatat acag

24

<210> 167

<211> 23

<212> DNA

<213> 人

<400> 167

tgtatatgca tatgtataca tat

23

<210> 168

<211> 28

<212> DNA

<213> 人

<400> 168

aagtgctggg attacaggca tgagccac

28

<210> 169

<211> 28

<212> DNA

<213> 人

<400> 169

gtggetcatg cctgtaatcc cagcactt

28

<210> 170

<211> 30

<212> DNA

<213> 人

<400> 170

agtgcagtgg cgcgatctcg gctcactgca

30

<210> 171

<211> 30

<212> DNA

<213> 人

<400> 171

tgcagtgagc cgagatcgcg ccactgcact

30

<210> 172

<211> 24

<212> DNA

<213> 人

<400> 172

gcctcccaa gtgctgggat taca

24

<210> 173

<211> 24

<212> DNA

<213> 人

<400> 173

tgtaatcca gcacttggg aggc

24

<210> 174

<211> 29

<212> DNA

<213> 人

<400> 174

ccaaagtgt gggattacag gcatgagcc

29

<210> 175

<211> 29

<212> DNA

<213> 人

<400> 175

ggctcatgcc tgtaatcca gcacttgg

29

<210> 176

<211> 22

<212> DNA

<213> 人

<400> 176

tgggattaca ggcgtgagcc ac

22

<210> 177

<211> 22

<212> DNA

<213> 人

<400> 177

gtggetcacg cctgtaatcc ca

22

<210> 178

<211> 31

<212> DNA

<213> 人

<400> 178

cctcggcctc ccaaagtgt gggattacag g

31

<210> 179

<211> 31

<212> DNA

<213> 人

<400> 179

cctgtaatcc cagcacttg ggaggccgag g

31

<210> 180

<211> 36

<212> DNA

<213> 人

<400> 180

gcccacctcg gcctccaaa gtgctgggat tacagg

36

<210> 181

<211> 36

<212> DNA

<213> 人

<400> 181

cctgtaatcc cagcacttg ggaggccgag gtgggc

36

<210> 182

<211> 24

<212> DNA

<213> 人

<400> 182

tcccaaagtg ctgggattac aggc

24

<210> 183

<211> 24

<212> DNA

<213> 人

<400> 183

gcctgtaatc ccagcacttt ggga

24

<210> 184

<211> 26

<212> DNA

<213> 人

<400> 184

ctccagcctg ggcaacagag cgagac

26

<210> 185

<211> 26

<212> DNA

<213> 人

<400> 185

gtctcgctct gttgccagc ctggag

26

<210> 186

<211> 63

<212> DNA

<213> 人

<400> 186

gaggcggagc ttgcagtgag ccgagatcgc tgggctgcac tccagcctgg gcgacagagc 60

gag

63

<210> 187

<211> 63

<212> DNA

<213> 人

<400> 187

ctcgtctgt cgcccaggct ggagtgcagc ccagcgatct cggctcactg caagctccgc 60

ctc

63

<210> 188

<211> 30

<212> DNA

<213> 人

<400> 188

aaagtgctgg gattacaggc atgagccacc 30

<210> 189

<211> 30

<212> DNA

<213> 人

<400> 189

ggtggctcat gcctgtaatc ccagcacttt

30

<210> 190

<211> 26

<212> DNA

<213> 人

<400> 190

gcctcccaaa gtgctgggat tacagg

26

<210> 191

<211> 26

<212> DNA

<213> 人

<400> 191

cctgtaatcc cagcactttg ggaggc

26

<210> 192

<211> 34

<212> DNA

<213> 人

<400> 192

tatataatta tatattatat ataatatata atta

34

<210> 193

<211> 34

<212> DNA

<213> 人

<400> 193

taattatata ttatatataa tatataatta tata

34

<210> 194

<211> 28

<212> DNA

<213> 人

<400> 194

atatataata tataattata tattatat

28

<210> 195

<211> 28

<212> DNA

<213> 人

<400> 195

atataatata taattatata ttatatat

28

<210> 196

<211> 31

<212> DNA

<213> 人

<400> 196

atatataata tataattata tattatatat t

31

<210> 197

<211> 31

<212> DNA

<213> 人

<400> 197

aatatataat atataattat atattatata t

31

<210> 198

<211> 28

<212> DNA

<213> 人

<400> 198

atatataata tataattata tattatat

28

<210> 199

<211> 28

<212> DNA

<213> 人

<400> 199

atataatata taattatata ttatatat

28

<210> 200

<211> 24

<212> DNA

<213> 人

<400> 200

gcctcccaaa gtgctgggat taca

24

<210> 201

<211> 24

<212> DNA

<213> 人

<400> 201

tgtaatccca gcacttggg aggc

24

<210> 202

<211> 22

<212> DNA

<213> 人

<400> 202

tgggattaca ggcatagagcc ac

22

<210> 203

<211> 22

<212> DNA

<213> 人

<400> 203

gtggctcatg cctgtaatcc ca

22

<210> 204

<211> 24

<212> DNA

<213> 人

<400> 204

ttttgtatt tttagtagag acgg

24

<210> 205

<211> 24

<212> DNA

<213> 人

<400> 205

ccgtctctac taaaaataca aaaa

24

<210> 206

<211> 37

<212> DNA

<213> 人

<400> 206

cctcagcctc ccaaagtgct gggattacag gcgtgag

37

<210> 207

<211> 37

<212> DNA

<213> 人

<400> 207

ctcagcctg taatcccagc actttgggag gctgagg

37

<210> 208

<211> 21

<212> DNA

<213> 人

<400> 208

gcccacctcg gcctccaaa g

21

<210> 209

<211> 21

<212> DNA

<213> 人

<400> 209

ctttgggagg ccgaggtggg c

21

<210> 210

<211> 37

<212> DNA

<213> 人

<400> 210

cctcagcctc ccaaagtgct gggattacag gcgtgag

37

<210> 211

<211> 37

<212> DNA

<213> 人

<400> 211

ctcagcctg taatcccagc actttgggag gctgagg

37

<210> 212

<211> 21

<212> DNA

<213> 人

<400> 212

gcccacctcg gcctccaaa g

21

<210> 213

<211> 21

<212> DNA

<213> 人

<400> 213

ctttgggagg ccgaggtggg c

21

<210> 214

<211> 21

<212> DNA

<213> 人

<400> 214

gggattacag gtgtgagcca c

21

<210> 215

<211> 21

<212> DNA

<213> 人

<400> 215

gtggctcaca cctgtaatcc c

21

<210> 216

<211> 24

<212> DNA

<213> 人

<400> 216

ttttgtatt tttagtagag acgg

24

<210> 217

<211> 24

<212> DNA

<213> 人

<400> 217

ccgtctctac taaaataca aaaa

24

<210> 218

<211> 36

<212> DNA

<213> 人

<400> 218

cctcagcctc ccaaagtgct gggattacag gcgtga

36

<210> 219

<211> 36

<212> DNA

<213> 人

<400> 219

tcacgcctgt aatcccagca cttgggagg ctgagg

36

<210> 220

<211> 36

<212> DNA

<213> 人

<400> 220

cctcagcctc ccaaagtgt gggattacag gcgtga

36

<210> 221

<211> 36

<212> DNA

<213> 人

<400> 221

tcacgcctgt aatcccagca cttgggagg ctgagg

36

<210> 222

<211> 32

<212> DNA

<213> 人

<400> 222

ggcctcccaa agtgctggga ttacaggcgt ga

32

<210> 223

<211> 32

<212> DNA

<213> 人

<400> 223

tcacgcctgt aatcccagca ctttgggagg cc

32

<210> 224

<211> 21

<212> DNA

<213> 人

<400> 224

ctcccaaagt gctgggatta c

21

<210> 225

<211> 21

<212> DNA

<213> 人

<400> 225

gtaatcccag cactttggga g

21

<210> 226

<211> 52

<212> DNA

<213> 人

<400> 226

atatatatac gtatatatat atacacatat atatatacgt atatatatat at

52

<210> 227

<211> 52

<212> DNA

<213> 人

<400> 227

atatatatat atacgtatat atatatgtgt atatatatat acgtatatat at

52

<210> 228

<211> 40

<212> DNA

<213> 人

<400> 228

tatatatata tacacatata tatatacgtatatata

40

<210> 229

<211> 40

<212> DNA

<213> 人

<400> 229

tatatatata tacgtatata tataatgtgta tatatatata 40

<210> 230

<211> 52

<212> DNA

<213> 人

<400> 230

atatatatac gtatatatat atacacatat atatatacgt atatatatat at 52

<210> 231

<211> 52

<212> DNA

<213> 人

<400> 231

atatatatat atacgtatat atatatgtgt atatatatat acgtatatat at 52

<210> 232

<211> 40

<212> DNA

<213> 人

<400> 232

tatatatata tacacatata tatatacgta tatatatata

40

<210> 233

<211> 40

<212> DNA

<213> 人

<400> 233

tatatatata tacgtatata tatatgtgta tatatatata

40

<210> 234

<211> 21

<212> DNA

<213> 人

<400> 234

cttggectcc caaagtgtg g

21

<210> 235

<211> 21

<212> DNA

<213> 人

<400> 235

ccagcacttt gggaggccaa g

21

<210> 236

<211> 23

<212> DNA

<213> 人

<400> 236

gctctgttgc ccaggctgca gtg

23

<210> 237

<211> 23

<212> DNA

<213> 人

<400> 237

cactgcagcc tgggcaacag agc

23

<210> 238

<211> 23

<212> DNA

<213> 人

<400> 238

gctctgttgc ccaggctgca gtg

23

<210> 239

<211> 23

<212> DNA

<213> 人

<400> 239

cactgcagcc tgggcaacag agc

23

<210> 240

<211> 28

<212> DNA

<213> 人

<400> 240

aagtgctggg attacaggca tgagccac

28

<210> 241

<211> 28

<212> DNA

<213> 人

<400> 241

gtggctcatg cctgtaatcc cagcactt

28

<210> 242

<211> 25

<212> DNA

<213> 人

<400> 242

ccatgtccat gtgcaactca acctc

25

<210> 243

<211> 25

<212> DNA

<213> 人

<400> 243

gaggttgagt tgcacatgga catgg

25

<210> 244

<211> 22

<212> DNA

<213> 人

<400> 244

ccttggcctc ccaaagtgct gg

22

<210> 245

<211> 22

<212> DNA

<213> 人

<400> 245

ccagcacttt gggaggccaa gg

22

<210> 246

<211> 24

<212> DNA

<213> 人

<400> 246

gatcacacca ctgcactcca gcct

24

<210> 247

<211> 24

<212> DNA

<213> 人

<400> 247

aggctggagt gcagtgggtg gatc

24

<210> 248

<211> 23

<212> DNA

<213> 人

<400> 248

tcccaaagtg ctggggttac agg

23

<210> 249

<211> 23

<212> DNA

<213> 人

<400> 249

cctgtaaccc cagcacttg gga

23

<210> 250

<211> 21

<212> DNA

<213> 人

<400> 250

gagaccagcc tgaccaacat g

21

<210> 251

<211> 21

<212> DNA

<213> 人

<400> 251

catgttggtc aggctggtct c

21

<210> 252

<211> 29

<212> DNA

<213> 人

<400> 252

gctctgtcgc ccaggctgga gtgcagtgg

29

<210> 253

<211> 29

<212> DNA

<213> 人

<400> 253

ccactgcact ccagcctggg cgacagagc

29

<210> 254

<211> 21

<212> DNA

<213> 人

<400> 254

aggttggtgc aaaagtgatt g

21

<210> 255

<211> 21

<212> DNA

<213> 人

<400> 255

caatcacttt tgcaccaacc t

21

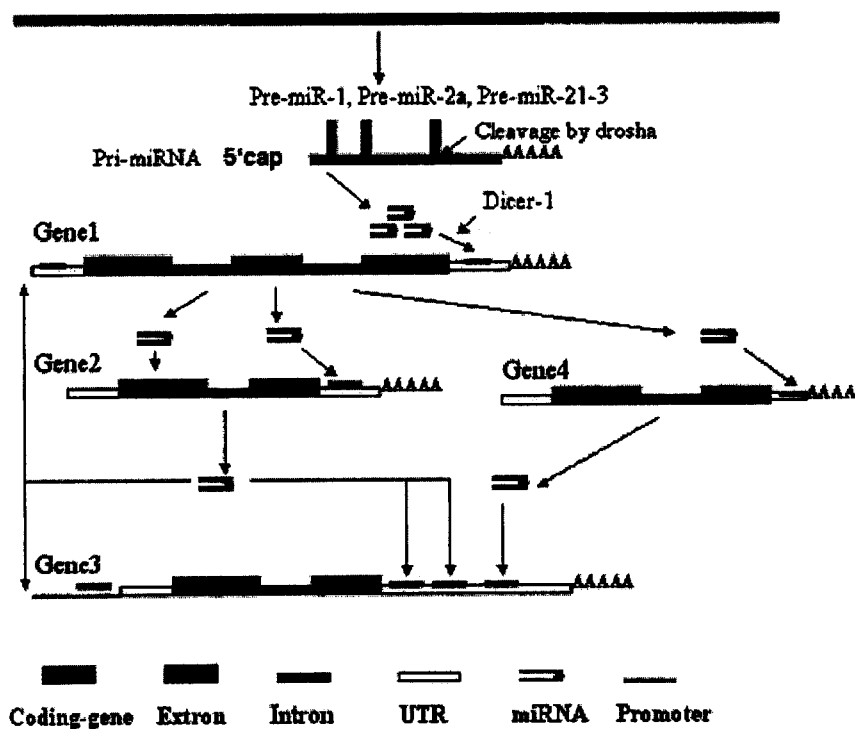


图 1

genome	intlocation	intradi	spaleca	mirname	mirname	mirname
5321T_AC002483	60834..68236	20	r658..680	MM_024081.4	homo sapiens transmembrane gamma-carboxyglutamic acid protein 4	Query: 1 gctcagc
5321T_AC002483	60834..68236	20	r658..680	MM_024081.4	homo sapiens transmembrane gamma-carboxyglutamic acid protein 4	Query: 1 gctcagc
5321T_AC002483	60834..68236	20	r658..680	MM_032933.3	homo sapiens hypothetical protein MGC11386 (MGC11386)	Query: 1 gctcagc
5321T_AC002483	60834..68236	20	r658..680	MM_020880.2	homo sapiens zinc finger protein 530 (ZNF530)	Query: 1 gctcagc
5321T_AC002483	60834..68236	20	r658..680	MM_018066.2	homo sapiens hypothetical protein FLJ10349 (FLJ10349)	Query: 1 gctcagc
5321T_AC002483	60834..68236	20	r658..680	MM_014356.2	homo sapiens chromosome 6 open reading frame 123 (C6orf123)	Query: 1 gctcagc
5321T_AC002483	60834..68236	20	r658..680	MM_014356.2	homo sapiens chromosome 6 open reading frame 123 (C6orf123)	Query: 1 gctcagc
5321T_AC002483	60834..68236	20	r658..680	MM_012287.3	homo sapiens centaurin	Query: 1 gctcagc
5321T_AC002483	60834..68236	20	r658..680	MM_004171.2	homo sapiens solute carrier family 1 (clial high affinity glucose	Query: 1 gctcagc
5321T_AC002483	60834..68236	20	r658..680	MM_130807.2	homo sapiens NDB1	Query: 1 gctcagc
5321T_AC002483	60834..68236	20	r658..680	MM_130807.2	homo sapiens NDB1	Query: 1 gctcagc
5321T_AC002483	60834..68236	20	1520..542	MM_017595.2	homo sapiens solute carrier family 2 (facilitated glucose	Query: 1 tgcagca
5321T_AC002483	60834..68236	20	1520..542	MM_020237.1	homo sapiens chromosome 8 open reading frame 17 (C8orf17)	Query: 1 tgcagca
5321T_AC002483	60834..68236	20	1520..542	MM_018962.1	homo sapiens Down syndrome critical region gene 6 (DSCR6)	Query: 1 tgcagca
5321T_AC002483	60834..68236	20	1520..542	MM_017958.1	homo sapiens plectroin homology domain containing	Query: 1 tgcagca
5321T_AC002483	60834..68236	20	1520..542	MM_017774.1	homo sapiens GMS regulatory subunit associated protein 1-like 1	Query: 1 tgcagca
5321T_AC002483	60834..68236	20	1520..542	MM_017713.1	homo sapiens hypothetical protein FLJ20211 (FLJ20211)	Query: 1 tgcagca
5321T_AC002483	60834..68236	20	1520..542	MM_017643.1	homo sapiens abd domain containing 1 (NBT91)	Query: 1 tgcagca
5321T_AC002483	60834..68236	20	1520..542	MM_018347.1	homo sapiens chromosome 20 open reading frame 29 (C20orf29)	Query: 1 tgcagca
5321T_AC002483	60834..68236	20	1520..542	MM_018304.1	homo sapiens hypothetical protein FLJ11029 (FLJ11029)	Query: 1 tgcagca
5321T_AC002483	60834..68236	20	1520..542	MM_018304.1	homo sapiens hypothetical protein FLJ11029 (FLJ11029)	Query: 3 ctgagat
5321T_AC002483	60834..68236	20	1520..542	MM_018138.1	homo sapiens hypothetical protein FLJ10560 (FLJ10560)	Query: 1 tgcagca
5321T_AC002483	60834..68236	20	1520..542	MM_018097.1	homo sapiens chromosome 15 open reading frame 25 (C15orf25)	Query: 1 tgcagca
5321T_AC002483	60834..68236	20	1520..542	MM_018097.1	homo sapiens chromosome 15 open reading frame 25 (C15orf25)	Query: 4 tgcagca

图 2

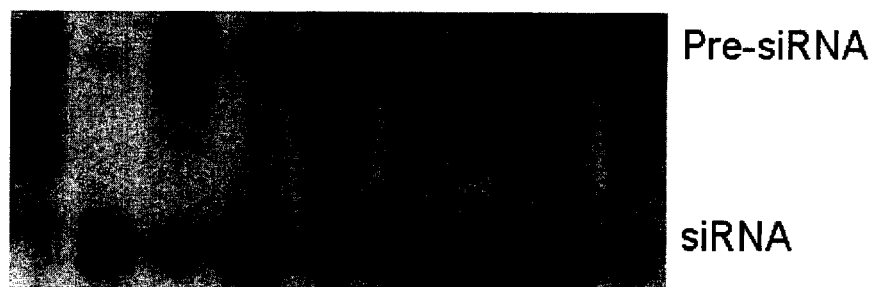


图 3

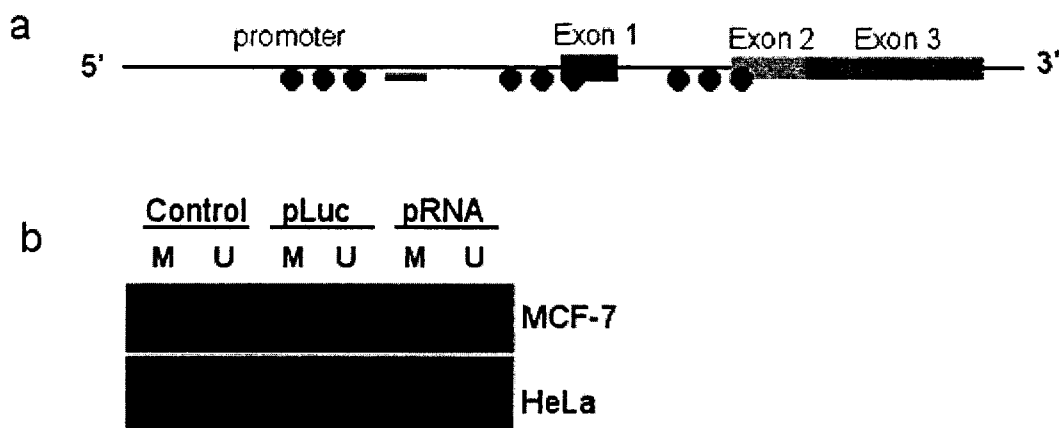


图 4

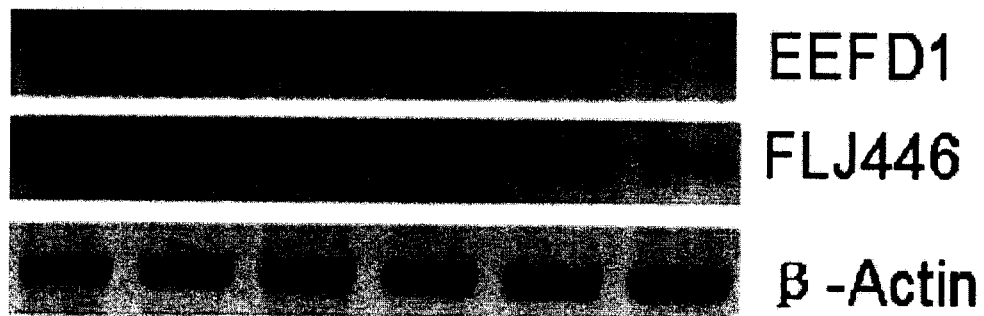


图 5

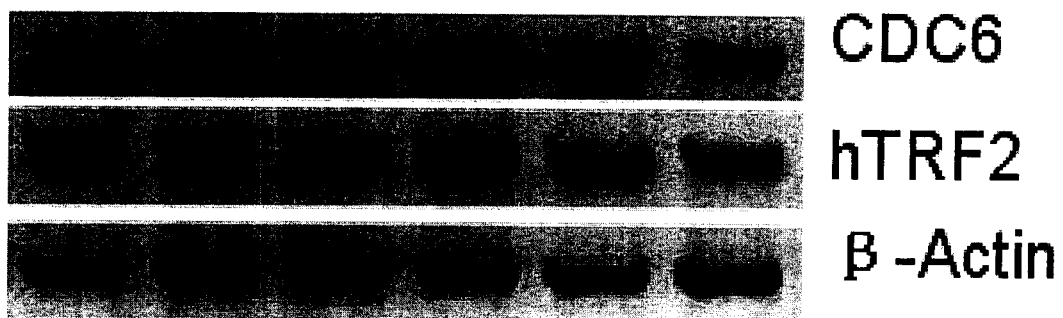


图 6

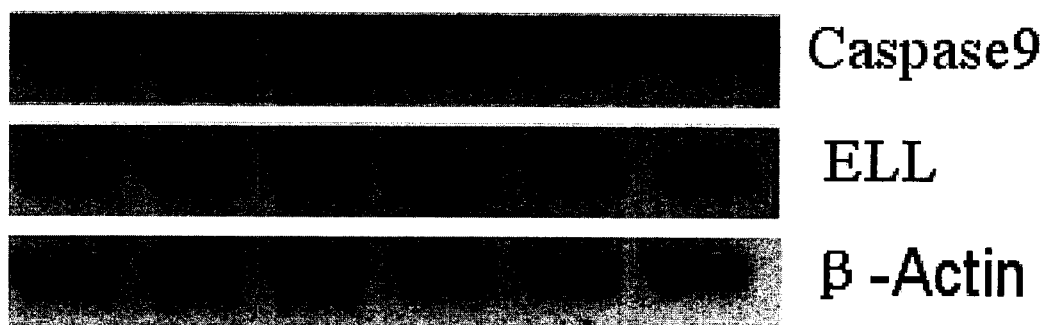


图 7