

华夏英才基金学术文库

# RNA 干扰：原理与应用

汤 华 主编

科学出版社

北 京

## 内 容 简 介

双链 RNA 可抑制含有特定序列的基因表达,即 RNA 干扰的发现开创了新的生命科学研究领域,目前这个研究领域正在快速发展之中。书中对 RNA 干扰的理论系统、技术体系及应用研究进行整体的构思,以该领域最新的研究成果为基础,除了双链 RNA 干扰之外,重点突出了对新发现的微小 RNA (miRNA) 进展的描述,从而全面系统地描述了 RNA 干扰;结合作者实验中的研究经验,系统地介绍了当今 RNA 干扰在多种生物体系中研究的实用技术方法。

本书适合生物学、医学及相关领域的实验人员、研究生等使用。

### 图书在版编目(CIP)数据

---

RNA 干扰:原理与应用/汤华主编. —北京:科学出版社,2006

(华夏英才基金学术文库)

ISBN 7-03-016331-1

I .R… II .汤… III .核糖核酸-生物技术 IV .Q522

中国版本图书馆 CIP 数据核字(2005)第 115790 号

---

责任编辑:庞在堂 彭克里 席慧/责任校对:包志虹

责任印制:钱玉芬/封面设计:王浩

**科 学 出 版 社 出 版**

北京东黄城根北街 16 号

邮政编码:100717

http://www.sciencep.com

印刷

科学出版社发行 各地新华书店经销

\*

2006 年 9 月第 一 版 开本:B5(720×1000)

2006 年 9 月第一次印刷 印张:31 插页:1

印数:1—3 000 字数:688 000

**定价:75.00 元**

(如有印装质量问题,我社负责调换(环伟))

## 编写人员名单

主 编 汤 华

参编人员 (按姓氏笔画为序)

马卓娅 王 芳 刘 民 汤 华

李 欣 李怡璇 沈志伟 唐晓燕

强 冉 薛晓荣

# 序 言

生命科学的进步是以对自然界新的发现、揭示生命规律为基础的；技术的创新是以科学进步、新的理论为基础的。RNA 干扰理论及技术是生命科学领域近十几年重大的进展之一，它加深了人们对生命规律的认识。由于 RNA 干扰理论及技术的逐步成熟，日益显现出其对生命科学、医药科学等领域的深刻影响，它的广泛应用大大加速了基因或基因组功能的确定，利用 RNA 干扰技术可能开发出治疗疾病的新方法或手段。因而近年来，RNA 干扰技术在其本身理论成熟的同时，迅速扩展到生命科学研究的各个领域及药物研发等领域，尤其在抗病毒、抗肿瘤等疾病研究和治疗方面日益体现出重要价值。

汤华教授根据该领域最新的研究成果并结合自己的经验，精心组织、设计并撰写了《RNA 干扰技术：原理与应用》一书。该书系统地描述了 RNA 干扰，包括微小 RNA (microRNA) 的作用原理及分子机制、可行的技术程序以及应用领域的研究成果，这些内容反映了国际上该领域最前沿的研究水平。我相信此书的出版将为 RNA 干扰技术在我国生命科学及医药领域的应用和基础研究起到积极的推动作用。

郝希山 

中国工程院院士  
天津医科大学校长  
2006 年 6 月 28 日

# 前 言

RNA 不仅作为遗传信息表达的中介体，而且直接参与基因表达的调控过程，从而影响生命体的发育、细胞的分化与功能，这一重要发现是生命科学领域近 20 年来最大的进展之一。近几年来，RNA 干扰，尤其是 siRNA 及新近发现的微小 RNA (miRNA) 研究的发展迅速，目前 RNA 干扰的理论及相关技术已渗透到生命科学的各个领域，引起了越来越多的学者的关注与兴趣。因此，我们根据 RNA 干扰的研究进展并结合本实验室开展这方面工作的经验编写此书。

RNA 干扰现象及分子机制的阐明已逐步形成了新的理论系统，其相关的或由这种理论发展起来的技术方法具有特殊性，RNA 干扰技术作为表观遗传学 (epigenetics) 研究的高效工具在多种生物体的发育、细胞功能、疾病机制等方面得到广泛的应用，并有可能开发成治疗疾病的新方式或新的药物。因而本书按照 RNA 干扰理论 (第 1~7 章)、RNA 干扰技术原理与方法 (第 8~13 章) 和 RNA 干扰的应用研究 (第 14~18 章) 三大部分的顺序编写。

本书是由直接从事 RNA 干扰工作的科研人员承担编写的，汤华参与各章节的编写及总审，其他章节的编写为：李欣 (第 1、2 章)，王芳 (第 2、3、4、5、6、7、13、16 章)，唐晓燕 (第 1、8、9、10、11、12 和 18 章)，李怡璇 (第 12、14 和 16 章)，强冉 (第 12、15 和 16 章)，沈志伟 (第 12、14、17 章)，马卓娅 (第 12、13、15 和 17 章)，薛晓荣 (第 12、14 和 17 章)，刘民 (第 14 章)。此外，唐晓燕、李怡璇、王芳和刘民为本书的校对做了大量的工作，在此表示感谢。

特别感谢中共中央统战部华夏英才基金管理委员会对本书出版的鼎力资助及关怀。中共天津市委、天津市教育卫生工作委员会及天津医科大学各级统战部领导对本书编写给予了热情支持，在此表示诚挚的谢意。

希望本书的出版能对同道们有所裨益，由于 RNA 干扰的理论技术是近几年发展起来的新领域，而且进展非常迅速，因此本书中不当之处在所难免，请读者批评指正。

汤 华

天津医科大学

天津市生命科学中心实验室

2005 年 12 月

# 目 录

序言  
前言

## 第 1 部分 RNA 干扰理论

<b>1 RNA 干扰的发展历史与意义</b> .....	3
1.1 RNAi 的发现 .....	3
1.2 RNAi 的意义 .....	8
<b>2 微小 RNA 概述</b> .....	12
2.1 miRNA 的发现过程 .....	12
2.2 miRNA 的生物学特征 .....	14
2.3 miRNA 的基因组学 .....	17
2.4 miRNA 的成熟及效应过程 .....	22
2.5 miRNA 的功能 .....	25
2.6 展望 .....	29
<b>3 依赖于 RNA 的 RNA 聚合酶</b> .....	35
3.1 RdRP 在基因沉默中作用的发现及相关证据 .....	35
3.2 RdRP 的生化物理性质及其晶体结构 .....	36
3.3 RdRP 与 RNA 病毒 .....	38
3.4 RdRP 和 RNA 沉默 .....	39
3.5 降级性 PCR、过渡性 RNAi 及系统性 RNAi .....	42
3.6 RdRP 作用过程中的问题 .....	45
<b>4 微处理器</b> .....	49
4.1 微处理器的相关蛋白成分及其对 miRNA 的加工 .....	49
4.2 pre-miRNA 从细胞核输出的过程 .....	52
<b>5 干扰性小 RNA 产生的关键性蛋白——Dicer</b> .....	55
5.1 DCR 的发现背景 .....	55
5.2 DCR 的一般生化属性及相关生物学功能 .....	56
5.3 DCR 的作用方式及在不同生物体中的作用 .....	60
5.4 DCR 作用的底物——dsRNA 及 pre-miRNA 的特征及来源 .....	68
5.5 DCR 的加工产物——siRNA 和 miRNA .....	71
5.6 不同生物体中 DCR 的同源物及其基因 .....	72

<b>6 RNA 诱导的沉默复合体</b> .....	78
6.1 RISC 的发现背景 .....	79
6.2 参与 RISC 组装的小 RNA 分子及其可能的组装模式 .....	81
6.3 RISC 的主要成分——AGO 蛋白 .....	85
6.4 RISC 在果蝇体内的组装模型 .....	93
6.5 RNAi 的剪切模型 .....	95
6.6 RISC 指导的靶向性剪切的动力学研究 .....	97
6.7 参与 RISC 组装及效应过程的相关蛋白质及影响因子 .....	99
6.8 RISC 的功能 .....	113
6.9 小结与展望 .....	115
<b>7 细胞核内的 RNA 干扰途径</b> .....	124
7.1 RNA 指导的 DNA 甲基化 .....	124
7.2 RNAi 和异染色质的形成 .....	128
7.3 RNAi 和 DNA 切除 .....	135
7.4 减数分裂性沉默、配对和 RNAi .....	137
7.5 细胞核内 RNAi 途径的作用方式 .....	138
7.6 其他生物体中的相关机制 .....	139
7.7 小结与展望 .....	140
<b>第 2 部分 RNA 干扰技术原理与方法</b>	
<b>8 siRNA 或 dsRNA 的设计与筛选策略</b> .....	149
8.1 重组质粒方法筛选 siRNA .....	151
8.2 高通量筛选 siRNA 的微阵列方法 .....	154
<b>9 化学方法合成 siRNA</b> .....	157
9.1 2'-tBDMS 方法 .....	158
9.2 2'-TOM 方法 .....	159
9.3 2'-ACE 方法 .....	159
<b>10 siRNA 的体外转录及生物合成方法</b> .....	161
10.1 PCR siRNA 合成法 .....	162
10.2 T7 RNA 聚合酶转录法 .....	165
10.3 应用重组人 Dicer 体外合成 siRNA .....	168
<b>11 基因工程方法制备 siRNA</b> .....	170
11.1 表达元件的构建 .....	170
11.2 载体构建 .....	180
<b>12 siRNA 的实验技术方法</b> .....	191
12.1 siRNA 的体外合成及转染方法 .....	191

12.2	体外转录制备 dsRNA .....	198
12.3	基因工程方法制备 siRNA 及转染细胞的方法 .....	202
12.4	RNAi 在其他生物体中的应用研究方法 .....	213
<b>13</b>	<b>miRNA 的实验技术方法 .....</b>	<b>234</b>
13.1	miRNA 的提取 .....	234
13.2	miRNA 克隆 .....	238
13.3	miRNA 分子杂交检测技术——Northern 印迹 .....	246
<b>第 3 部分 RNA 干扰的应用研究</b>		
<b>14</b>	<b>RNA 干扰在研究基因组功能方面的应用 .....</b>	<b>259</b>
14.1	动植物中的 RNAi .....	259
14.2	RNAi 在线虫全基因组功能研究中的应用 .....	262
14.3	RNAi 在果蝇研究方面的应用 .....	286
14.4	RNAi 技术在锥虫研究中的应用 .....	294
14.5	PTGS 在植物基因及基因组功能方面的研究 .....	301
<b>15</b>	<b>RNA 干扰在研究胚胎发育方面的应用 .....</b>	<b>319</b>
15.1	RNAi 在小鼠卵母细胞和胚胎中的应用 .....	319
15.2	RNAi 在鸟类胚胎发育中的研究 .....	336
<b>16</b>	<b>RNA 干扰在抗病毒研究中的应用 .....</b>	<b>358</b>
16.1	RNAi 抗 HIV-1 病毒感染 .....	358
16.2	RNAi 抗禽类逆转录病毒感染 .....	373
16.3	RNAi 抗肝炎病毒感染 .....	377
16.4	RNAi 抗流感病毒感染 .....	392
16.5	RNAi 抗 SARS 相关冠状病毒感染 .....	397
16.6	RNAi 抗轮状病毒感染 .....	400
<b>17</b>	<b>RNA 干扰在肿瘤研究中的应用 .....</b>	<b>410</b>
17.1	目前常用的肿瘤相关基因功能的研究方法 .....	410
17.2	RNAi 在肿瘤生物学中的应用 .....	413
17.3	RNAi 在基因治疗中遇到的主要问题 .....	435
<b>18</b>	<b>RNA 干扰在其他领域的治疗研究 .....</b>	<b>444</b>
18.1	RNAi 在心血管疾病中的应用研究 .....	444
18.2	RNAi 在神经系统疾病中的应用研究 .....	446
18.3	RNAi 在内分泌系统疾病中的研究进展 .....	454
18.4	RNAi 在自身免疫性疾病中的应用研究 .....	456
18.5	RNAi 在生物学和医学中的应用前景 .....	458
	<b>索引 .....</b>	<b>465</b>

# 第 1 部分 RNA 干扰理论

传统意义上的 RNA 干扰 (RNAi) 主要是指外源性双链 RNA 在体内引发的基因沉默现象, 而新近发现的微小 RNA 导致的内源性 RNAi 赋予了 RNAi 新的内涵。

目前, 关于 siRNA 和微小 RNA 的产生以及基因沉默过程中涉及的相关成分 (RdRP、Drosha、Dicer、RISC 等) 的研究取得了重大突破。此部分将对 RNA 干扰现象的发现、生物发生及效应过程的分子机制进行系统介绍。

**本部分包括以下 7 章:**

1. RNA 干扰的发展历史与意义
2. 微小 RNA 概述
3. 依赖于 RNA 的 RNA 聚合酶
4. 微处理器
5. 干扰性小 RNA 产生的关键性蛋白——Dicer
6. RNA 诱导的沉默复合体
7. 细胞核内的 RNA 干扰途径

# 1 RNA 干扰的发展历史与意义

1908 年，摩尔根实验室开始利用遗传突变筛选方法在果蝇（*Drosophila*）中分离获得数十种突变体。这种正向（forward）遗传筛选方法为理解细胞和器官的形成及功能作出了巨大的贡献。然而，这种遗传筛选方法仅能在一些理想的有机体中进行。这种有机体具有在实验室条件下能繁殖、生命周期短、繁殖后代甚多等特点，并需数月或数年的工作才能确定特异表型的决定基因。随着分子生物学技术的出现，科学家开始寻找“后门进入”途径，即以序列同源性或生物化学功能为基础来快速确定其基因生物学功能，称之为反向遗传学（reverse genetics，从基因到表型）。

1984 年，Izant 和 Wostrarb 在这个方向上前进了一大步，即用遗传工程构建表达与胸苷激酶（thymidine kinase）基因的反义 RNA 互补的 DNA 重组子转染培养的细胞，发现胸苷激酶在这些细胞中的表达显著减少。此后，研究者陆续发现多种反义 RNA 对各种特异基因活性表达有抑制作用。这种开拓性技术不仅对于研究相关细胞学和那些不能用经典遗传学分析方法研究的整个机体，如爪蟾，具有极其重要的意义，而且对于建立基因功能研究模式做出了突出贡献。但因其作用相对较弱，并且单链 RNA 相对很不稳定，其应用研究未能广泛进行。

而新近发现的 dsRNA 介导对同源靶 RNA 的降解诱导产生了强大且特异性的基因表达抑制或沉默作用，称为 RNA 干扰（RNA interference, RNAi）。这为研究生物体功能基因及表观遗传学（epigenetics）提供了极其有效的工具。

## 1.1 RNAi 的发现

### 1.1.1 RNAi 在植物和真菌中导致转录后的基因沉默

在 20 世纪 90 年代初期，研究者发现在植物和真菌中存在对转基因序列应答的转录后基因沉默方式（post-transcriptional gene silencing, PTGS）。Aapoli 等设计将苯基乙烯酮合成酶（chalcone synthetase）基因转入矮牵牛属植物来提高色素产量时，意外地发现转录量 40% 以上的转基因植物的花呈现白色或杂斑色而非紫色。这种效应不仅是由于所转苯基乙烯酮合成酶基因的作用，而且是其内源性同源基因沉默或受到抑制的结果，因而称之为共抑制（cosuppressing）。1992 年，同样研究色素产生的 Roman 和 Macino 报道在粗糙链孢菌（*Neurospora crassa*）中也发现类似的现象，称之为清除（quelling）。在这两类研究中观察到一些子代

完全恢复了父代表型或隐性状态。在这些研究中都试图高表达基因反而导致其沉默,不过,作用是短暂的。根据这些依据,研究者们形成了这样的理论——DNA甲基化或 RNA 中间体介导了这种应答。关于真菌中基因表达清除来自于转录后过程的研究进一步确定其作用分子很可能是存在于细胞浆中的 RNA 分子。与此相反,一些植物中转录后的共抑制作用可能与基因本身的甲基化有关,导致了基因表达的下调。植物中转录抑制作用可通过基因表达的转录水平和转录后水平来进行调控,后来均称为 RNAi,并得出以下结论:RNA 分子是其介导效应分子。但当时其介导的 RNA 分子特点或结构尚未得到阐明。

### 1.1.2 在线虫 (*Caenorhabditis elegans*) 中发现 RNAi 现象

在体外合成反义 RNA 的研究中,研究者应用同义 RNA 作为特异性的阴性对照,但实验结果让人感到意外——分别注射同义和反义 RNA 产生了类似的表型。在果蝇中的研究证实了这一点,同时注射同义和反义 RNA 都可产生特异性和可重复性的表型。这种相反但互补的分子设计所产生的效应被命名为“RNA 干扰”,以此来区别经典的反义抑制作用。此互补双链 RNA 除能产生特异性抑制基因表达作用外,还被发现能从注射原点扩散到周围的细胞。这提示干扰性 RNA 在线虫中能从起始注射点转移到大多数细胞的组织中,引发系统性应答。利用这种反向遗传学方法,研究人员将线虫浸泡在 RNA 分子中以获得有意义的表型,从而在线虫中对 RNAi 应答的分子机制进行了研究<sup>[1]</sup>。

1998 年,卡内基研究所的 Fine 等首次发现双链 RNA (dsRNA),而非单链反义 RNA 能诱发果蝇体内序列特异性的内源性基因 mRNA 的降解。对于起初观察到的表型作用,研究者认为可能是在体外用质粒做模板进行转录制备同义和反义 RNA 时,产物包含有小量互补的极性 RNA 分子,形成的 dsRNA 和单链 RNA 共同作用导致了上述的效应。结果证实经聚丙烯酰胺凝胶 (PAGE) 纯化出来的单链 RNA 注射入体内所致的表型变化明显减少,而正义链与反义链在体外退火形成的双链 RNA 则引起更为快速、特异的与其双链 RNA 正义链同源的 mRNA 的降解。

RNAi 具有很多的特征,其中之一是已证明了小干扰 RNA (short/small interfering RNA, siRNA) 能引发机体组织细胞功能的变化,且其发挥功能是以酶底物米-曼动力学方式进行的,即一个分子能诱发数十甚至数百个靶 mRNA 个体分子的降解。这一点与反义单链 RNA 的反义抑制作用不同,后者需要过量的反义单链 RNA (antisense RNA, asRNA) 来作用于靶 mRNA。然而,dsRNA 如何能靶定同源的核酸序列? 什么样的模式能解释 dsRNA 的作用机制或潜力?

有研究者利用原位杂交观察到 dsRNA 靶定的 mRNA 不能聚集,当时认为 dsRNA 通过某种方式直接靶定基因,可能干扰转录的起始或延伸,这可解释为什么在单个细胞内一个或两个 dsRNA 分子可导致靶 mRNA 的明显减少。然而,

至今一直找不到明显且完整的证据支持这种理论构想。在 dsRNA 存在的情况下，研究者发现 dsRNA 靶定的基因具有连续性及转录后性的特点。但其编码的 mRNA 在翻译之前即快速降解。支持转录后机制的证据还包括 dsRNA 不能靶定内含子或启动子区域，这提示作用部位在细胞浆内这一机制。此外，尚未发现靶定的细胞内新合成的这种 mRNA 的序列本身有任何改变。这并不奇怪，因为 RNAi 作用是完全可逆的。在果蝇 RNAi 机制的研究中，由于果蝇缺乏明显的甲基化系统，尚不能确定靶定的基因是否被甲基化导致沉默效应。

目前，关于 dsRNA 的作用机制主要有以下三种模式：①在与靶基因序列完全互补的情况下，导致靶序列的降解；②在与靶基因序列不完全互补的情况下，使靶基因翻译受到抑制；③通过甲基化修饰影响染色体的结构（如异染色质形成），从而导致转录水平上的抑制。

在线虫中，dsRNA 引发的 PTGS 导致了这样一种理论，即在转基因植物中 dsRNA 引发基因沉默的意外发现可能至少部分解释了植物中的 PTGS。此后不久，其倡导者研究阐明在烟草和水稻植物中遗传工程表达 dsRNA 确实呈现了基因沉默。植物和线虫两者进化差异很大，但对 dsRNA 的应答反应出现类似的方式。另外一个共同的特性是干扰性 RNA 能在组织之间扩散，进而激活出现系统沉默效应。不过，这种结果并不意味着所有的转基因诱导 PTGS 都是因体内正义链与反义转录物退火而成的 dsRNA 所致。一种解释模式是高水平同义转录物可能引发通过依赖于 RNA 的 RNA 聚合酶形成拷贝 RNA (cRNA)。因而有研究者认为 dsRNA 或者 cRNA 本身可能是干扰 RNA 的来源。进而，在植物中这种 RNA 可能反馈至基因，并介导了 DNA 甲基化。此后的一系列研究也证实了这种理论。

在线虫中发现 RNAi 后，科学家开始研究 dsRNA 是否在其他生物体中也存在诱发基因沉默的现象。在锥虫 (*Trypanosome*, 一种原生生物)、果蝇和许多其他的动物和植物种系中发现 RNAi 能关闭或减少基因表达。起初，研究者认为 RNAi 在其他脊椎动物中的作用应该是有限的。这是由于对病毒或外源 RNA 应答的依赖于 dsRNA 激活的蛋白激酶 R (double-stranded RNA activated protein kinase R, PKR) 通过触发整个细胞剧烈的反应而最终导致细胞死亡。但是，通过靶定的卵母细胞和 PKR 表达产生之前的小鼠胚胎，证明 dsRNA 同样能引发序列特异性沉默。在斑马鱼研究中也获得类似的结果，但似乎由于非特异性作用，其结果不甚理想。因而，尽管 dsRNA 在极大多数不同种类生物体中都可导致序列特异性沉默，但并不在所有有机体中引发相同的结果。进一步的研究表明，RNAi、PTGS 和清除过程是相似的，机体运用进化保守这种机制来阻止非自我或错误遗传信息的表达，如病毒转位因子编码的蛋白质；异常 RNA 结构，如弥散性 dsRNA 或一些 RNA 病毒表达的复制中间体，提示宿主细胞存在外源入侵物并引发沉默。支持这种 RNAi 和相关共抑制/PTGS 应答功能解释的依据是对

RNA 沉默分子途径成分进行的一些突变导致机体对病毒的敏感性增加。

### 1.1.3 RNAi 作用机制的模式

RNAi 在分析基因功能方面具有巨大的价值，其作用过程中的奇妙机制同样引人注目。dsRNA 可以直接靶定多个拷贝的 mRNA 分子，提示 dsRNA 在体内存在扩增机制。在对果蝇、蠕虫、真菌、植物和哺乳动物细胞的生化与遗传学研究中，研究者获得大量相关证据，在此基础上提出了作用机制的模式。RNA 沉默的应答的核心反应是引发 dsRNA 通过三磷酸腺苷 (ATP) 依赖性 RNA 酶 III (Dicer, DCR) 加工处理成 21~25bp (碱基对)，并具有两个核苷酸突出在其 3' 末端的片段。DCR 含有多个蛋白结构域，包括有依赖于 ATP 的 RNA 解旋酶、PIWI 结构域、两个顺式排列的 RNA 酶 III (RNase III) 结构域及一个 dsRNA 结合结构域。DCR 作用产生的 21~25bp 的产物被称为短或小干扰 RNA 分子 (siRNA)。siRNA 是引发核酸酶降解 mRNA 的引导物，每一个 siRNA 都与 RNA 诱导沉默复合物 (RNA-induced silencing complex, RISC) 相关联，并且研究者推测在 RISC 中 siRNA 进行解链释放出其反义 RNA 通过碱基配对原则与靶 mRNA 结合。这种碱基配对结合导致靶 mRNA 的内源性降解，被切断或产生失去 polyA 尾及 5'-甲基鸟嘌呤帽的片段，释放出的每一片段更易被 RNA 监视机制进一步降解。在靶 mRNA 降解之后，包含有 siRNA 的 RISC 复合体继续靶定另外的 mRNA 分子。

首先，引发 dsRNA 能产生多个 siRNA，一般一个 500bp dsRNA 能产生约 20 个左右的 siRNA，然后 siRNA 整合入 RISC，使复合体发挥催化活性靶定并切割多个 mRNA 分子。这种模式解释了引发 dsRNA 的酶底物米-曼氏动力学过程，包括扩增步骤和催化反应过程。DCR 和 RISC 在可能的 RNAi 作用机制中是进化上最保守的两个成分，并且是核心 RNAi 应答的主要催化反应过程发生的部位。

最近，至少在一些研究中发现扩增反应的存在。次级 siRNA 可能通过称为过渡性 RNAi (transitive RNAi) 的机制产生。其过程可能是 siRNA 或短的反义 RNA 可发挥启动特异性依赖于 RNA 的 RNA 聚合酶 (RdRP) 合成与被靶定序列互补的 RNA (cRNA) 共同重新合成产生 dsRNA 的作用，然后通过 DCR 裂解成新的 siRNA。也有新的证据表明有些 RdRP 也能非常高水平地进行特异性转录或在检测到如病毒入侵的异常 RNA 存在的情况下不需要启动子就能产生 cRNA。最早提出的这种设想是用来解释植物转基因产生的 PTGS 现象。推测对细胞带来的优势是更多的 mRNA 引发并导致更多的靶基因失活，使得细胞更有效地清除感染的病毒。

在线虫、放线菌等生物体中进行 RdRP 突变的研究表明这些蛋白质分子在 RNAi 中起到保守和基础性作用。有意义的是一些突变体可增加机体对病毒的敏感性，以及导致机体发育的缺陷。让人感到奇怪的是至今尚未在果蝇及人类基因

组中发现 RdRP 类似物。那么，这种引发与扩增方式是否是 RNAi 的通用特性？尽管在果蝇细胞提取物中观察到 siRNA 指导的 RNA 合成。然而，3'-OH 被修饰的 siRNA 据推测是不能作为 RNA 合成的启动引物的，但在人 HeLa 细胞和果蝇中却能引发 RNAi 产生。此外，在小鼠卵母细胞中，RNAi 的过程并不需要 RNA 合成。这些证据表明，即使在果蝇及人体细胞中 RNA 扩增对于其 RNAi 的应答也是非必需的。缺乏过渡性 RNAi 的优点是增加其特异性，以致对这种 mRNA 的替换性剪切体 (alternative spliced mRNA) 可能需要分别进行靶定。

#### 1.1.4 RNAi 发挥的生物功能

防止通过 RNA 中介体的病毒或“跳跃”因子的基因表达可能是 RNAi 的主要功能之一，显然这种通过 RNA 中介体裂解的机制对细胞是非常有利的，进而通过引导 RNA 扩增而达到靶定许多不需要的 RNA 分子的目的，从而使细胞更能有效地达到防止引发和靶定这种 mRNA 翻译成功能性蛋白质的作用。对哺乳动物细胞来说，更大的意义在于 siRNA 并不释放 PKR 活性，而是介导序列特异性沉默。目前，以 siRNA 为基础的 RNAi 策略应用于包括哺乳动物在内的多种生物系统中。

RNA 沉默突变体的筛选最早是在放线菌中进行的，最近在线虫和果蝇中也有所发展，并导致确定了几种 RNAi 应答相应蛋白质必需的基因，包括 DCR、RdRP、RISC 及各种解旋酶等。沉默包含的连续遗传学途径分析阐明了这些不同功能成分发挥作用的前后顺序及一些新的功能。有些 RNAi 相关基因缺失突变表现出一些表型，包括增强了对病毒感染的敏感性，增加了“跳跃”因子的移动，发育时序的缺失，有丝分裂和减数分裂的缺失，不育和其他一些发育过程的缺陷。这些表型对进一步增强对 RNAi 功能的理解具有重大意义。

最激动人心的发现之一是 RNAi 调控有机体内源性基因的表达。这种内源性 RNAi 作用是由机体基因组编码的一类小的调控性 RNA (microRNA, miRNA) 所介导的。这种 miRNA 多数情况下与靶 mRNA 的 3' 非翻译区 (3' untranslated region, 3' UTR) 完全互补结合，有些 miRNA 功能上类似 siRNA，整合入 RISC 并导致靶 mRNA 裂解。不过，许多 miRNA 可能通过其他替代途径发挥作用，它与靶序列并不完全互补结合，因而不影响靶 mRNA 稳定性，而是通过干扰翻译过程而阻止编码蛋白质的合成。最近发现，这种干扰性 RNA 与靶 mRNA 的错配及甲基化的部位至关重要，miRNA 在与靶位杂交过程中，其与靶序列的错配发生在结合区中间部位从而形成“U”形区域，影响其靶基因的翻译过程。

假如 siRNA 分子能有效地结合其靶序列 (序列配对及其空间结构决定)，那么 siRNA 分子还可以导致靶 RNA 分子的切割。已有证据表明单一结合位点也可介导翻译过程的抑制或下调。然而，let-7 miRNA 的内源性靶分子的 3' UTR 含有多个与其结合的位点，这提示有些情况下，单一结合的情况可能不足以抑制靶基

因的表达。

### 1.1.5 RNAi 可能引发 DNA 的共价修饰

目前认为 RNAi 在细胞内发挥作用主要包括扩增和催化反应两个主要步骤。而在植物中，发现 RNAi 也能引发 DNA 甲基化，从而导致转录沉默<sup>[3]</sup>。现在证明 siRNA 可能通过直接 DNA 甲基化和在其他一些基因组中的共价修饰来完成转录过程。在植物中，这一活性可能由另外一类 siRNA 介导完成。

最近最有意义的发现之一是在粟酒裂殖酵母 (*S. pombe*) 中 RNAi 机制在合成和维持高度有序的染色质结构与功能中发挥至关重要的作用。缺失 RNAi 过程中的关键基因导致中心粒 DNA 和其他类型的异染色质的表观遗传沉默现象缺失。这些改变伴随着有这些区域的甲基化状态改变进而中心黏合作用丧失，导致核分裂过程中染色体的错误分离。有研究者据此提出理论解释了在野生型细胞中，中心粒重复序列的表达导致 dsRNA 的形成，DCR 将其裂解为 siRNA，siRNA 指导异源染色质位点的 DNA 甲基化。也有报道 siRNA 介导染色体结构更大的改变，在四膜虫中程序性基因组重排导致特异性 DNA 序列的切除。

## 1.2 RNAi 的意义

### 1.2.1 RNAi 在功能性基因研究中的应用

随着基因组结构及测序工作的完成，接下来的基因功能确定是生命科学的最具挑战的领域。传统的基因敲除、转基因或突变的方法存在周期长、不能大规模同步进行等缺点。RNAi 的发现及作为反向遗传学技术的应用似乎是最完美的技术，其最合理的用途可能是用来分析由不同的基因组项目推测的数以千计的基因功能。RNAi 被广泛应用在线虫及果蝇的基因功能的研究中，主要是通过大量研究对某一特异性遗传途径进行解析或通过研究系统性功能基因组中每个基因的方式来进行。RNAi 在基因组水平上用来筛选确定导致某一特异表型或改变转基因表达的基因。

或许至今最宏伟的计划是由英国和荷兰科学家联合启动的设计针对人类基因组中每个基因的 siRNA 项目，然后使用它们靶定肿瘤细胞系中的每个基因来确定细胞恶性转化所需的基因。更多的靶定途径是将微阵列分析及 RNAi 技术相结合，来选择性靶定在一些细胞和组织类型中上调和过表达的基因，使用这一途径已确定了控制结肠癌增殖相关的基因。

或许 RNAi 最有意义的应用途径之一是在那些最好的遗传学工具不能使用的有机体中进行基因功能性研究。RNAi 提供的反向遗传学方法给一些常规遗传学方式不很适合研究的有机体带来新的研究途径。RNAi 也能使种系之间如果蝇和

其他昆虫，线虫和其他种系之间更容易进行比较性分析。

#### 1.2.1.1 系统性与细胞为基础的 RNAi

早期研究观察到 dsRNA 并不能在所有生物种系中引起 RNAi 应答。斑马鱼对注射的 dsRNA 趋向于引起非特异性致死性应答反应，但通过注射 siRNA 而非 dsRNA 可减轻其非特异应答。而其他一些种系，如海胆 (sea urchin) 和许多线虫则似乎对 RNAi 是耐受的。当发挥 RNAi 的能力在多种种系的进化进程中丧失时，有可能有些 RNAi 耐受的种系丧失了系统性水平上进行应答的能力但保持了细胞水平上的自动应答能力。然而，导入 (delivery) 可能是 RNAi 技术应用的最大障碍。在具有管道形态植物系统中，干扰性 RNA 分子可能通过韧皮部和水分子运输系统进行扩散。线虫则有能力摄入和转运外源性提供的 dsRNA，尽管并不是所有细胞都具有这种能力，尤其那些缺乏转运子表达的细胞。

有效的 dsRNA 或 siRNA 转运至哺乳动物细胞的策略包括转染、再转染及利用病毒载体的转基因途径；由 RNA 聚合酶 III 启动表达的发夹结构及相关方法已成为 RNAi 发展起来的治疗方法。已有好几种方法能有效地将 siRNA 转运至哺乳动物体内，如病毒和非病毒转运系统都已发展起来。从小鼠尾静脉注射 siRNA 引起肝脏细胞中靶 mRNA 和蛋白质水平的减少。这些结果表明至少在一些组织和细胞中 RNAi 的转运并不是从前认为的难以克服的困难。快速、安全、有效及组织特异的 siRNA 转运可能是 siRNA 作为治疗药物开发最大的困难。

#### 1.2.1.2 快速、经济、高效的 RNAi 技术

RNAi 现象存在的广泛性远远超过人们的预期，对此问题的深入研究结果将为进化的观点提供有力佐证。虽然为破坏特异性基因活性，可利用广泛应用的反向遗传学方法，包括同源重组、靶定缺失、“跳跃”因子介导的基因破坏、转染的反义和修饰的反义寡聚核苷酸等技术，但与这几种使基因功能丧失或降低的技术相比，RNAi 技术具有更明显的优点：它比反义 RNA 技术更有效，更容易产生功能丧失或降低。而且通过与细胞特异性启动子及可诱导系统结合使用，可以在发育的不同时期或不同器官中有选择地进行，与基因敲除技术造成的功能永久性缺失相比，这是更受科学家偏爱的。

与传统的遗传学途径相比，RNAi 可能具有多方面的优势。在多数情况下，应用 RNAi 可使研究人员在较短时间里确定特定基因的功能，而使用传统的方法可能至少需要数月或数年。另一方面，从遗传筛选的角度来讲 RNAi 可以更为准确地获得表型，尤其是那些具有多种作用或在发育早期表现作用的基因。而传统的方式，仅是推测有表型的突变体才能分离出来。此外，使用 RNAi 很容易获得双基因或多基因“突变体”，尤其在关于定位相关的多基因表型方面。

RNAi 技术也是一种经济的基因功能分析方法。传统的分析方法需要繁琐的

准备材料并进行方法的筛选、制改图谱、走样、测序比对应，而 RNAi 技术要简便很多。并且，目前在基因数据库中多数基因或序列的功能尚不清楚，因而可利用快速又经济的 RNAi 技术对这些基因作一功能分类或初步分类。

在发育过程中的不同阶段，特异性消除基因的功能是 RNAi 最大优点之一。假如使用合适的启动子，RNAi 在一些特定细胞类型中可能消除或减少特异基因的活性。这可能是一种更具正确的选择。使用组织特异性启动子驱动特异基因相关 siRNA 的表达，导致相应组织靶基因功能清除从而产生靶定破坏基因的遗传杂合子。这种策略使研究人员更精确指导各种组织特异基因表达的缺乏发生。

转基因表达 dsRNA 可能使 RNAi 技术更具优势，因为相对 dsRNA 的应答是瞬间的，通常持续几小时至几天。假如需要更稳定的突变体，可应用转基因技术，利用通用的可诱导组织特异性的启动子产生 dsRNA。根据使用目的的不同来选择不同的表达方法，使 RNAi 瞬时特性更具优势，如 RNAi 可产生一系列表型特征，常常在单一实验中，可发生功能轻微变化至完全清除的不同情况。

### 1.2.2 RNAi 在生物技术和生物医学中的应用前景

RNAi 是一种核酸序列特异性阻止基因功能的有效工具，可产生基因功能的消除或通过减少功能表型从而有助于加深对基因功能的理解。RNA 诱导的沉默同样在生物技术和生物医药领域具有明显的应用价值<sup>[4]</sup>。应用 siRNA 靶定雄性生殖相关基因产生了不孕的雄性体，因而成为正在研究的生物避孕的方法。转基因诱导 RNA 沉默正用于抑制基因的表达，对抑制内源性基因表达或抵抗病毒的感染具有潜在商用价值。对于含 dsRNA 的食物是否对群体或环境有什么副作用，其安全性尚需进一步研究。

siRNA 具有作为治疗性药物开发的潜力。药学应用研究主要侧重在抗病毒和抗肿瘤领域的应用，多数在细胞水平和动物水平进行。SIR27，一种针对人血管内皮生长因子（VEGF）受体的 siRNA 已在美国进入一期临床试验。

RNAi 为基础的抗病毒感染研究包括人乳头瘤病毒、乙型肝炎病毒、丙型肝炎病毒、鼠类白血病病毒、流感病毒以及人获得性免疫缺陷病毒（HIV）<sup>[5]</sup>等。当然，病毒也可能存在抵抗宿主 RNAi 干扰的机制。

使用 RNAi 来发现新的药物靶位及确定靶位点的研究已成为药物开发的领域之一。所报道的最振奋人心的研究之一是靶定在结肠癌细胞中过度表达的基因——*survivin*（存活素基因），导致肿瘤生长的抑制。同时，有研究发现用 RNAi 技术阻止肿瘤相关基因的表达可以有效地抑制肿瘤细胞的生长。RNAi 为基础的医学应用正在快速进行中，以质粒或病毒为载体表达发夹状结构 siRNA 的基因沉默方式已被广泛应用<sup>[6]</sup>。当然，有效的转运策略仍然是大家感兴趣的主要领域之一。

RNAi 为基础的药物开发或医学治疗方式的发展仍需注意其发挥作用的特异

性问题。siRNA 与靶序列之间单一碱基的错误能明显减弱 RNAi 的作用。错配的数目及发生的位置对 siRNA 作用效果的影响仍有待于阐明。如果错配仍有一定作用，使人担心 siRNA 在有效抑制靶基因表达同时，也可能影响相似但不完全相同的其他基因的功能。同样在同一基因序列中，靶定不同部位的 siRNA，其作用可能不完全一致。这种作用差异的机制尚不完全清楚。可能与 siRNA 能否顺利与 mRNA 序列区域结合有关。结合 siRNA 实验和基因谱分析可以确定 siRNA 特异性和有效性水平。此外，最新研究显示，向成年小鼠静脉注射可以表达特异性发夹状 RNA (short hairpin RNA, shRNA) 的腺病毒载体后，可以在肝细胞中长时间表达 shRNA，但同时也发现这种高浓度、长时间的 shRNA 长表达可以导致肝脏的损伤，呈剂量依赖性，部分可致小鼠死亡。目前认为这种损伤与肝脏来源的 miRNA 的下调有关，可能是内源性小 RNA 途径过饱和所致。因此，在将 RNAi 应用于临床进行治疗时，必须严格控制 shRNA 的剂量<sup>[7]</sup>。

目前，RNAi 已变为热门领域，RNA 的潜在应用价值十分巨大。可以预见，以 RNAi 为基础的技术策略应用于临床治疗（包括药物与生物治疗）可能为时不远。

(李欣 唐晓燕 汤华)

### 参 考 文 献

- 1 Gunter M, Thomas T. Mechanisms of gene silencing by double-stranded RNA. *Nature*, 2004, 431: 343~349
- 2 Fire A, Xu S, Montgomery MK, et al. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, 1998, 391: 806~811
- 3 Robin A. RNAi and Heterochromatin—a Hushed-up affair. *Science*, 2002, 297: 1818~1819
- 4 Edward T. RNA as drug and antidote. *Nature*, 2002, 419: 23~24
- 5 Gregory JH, John JR. Unlocking the potential of the human genome with RNA interference. *Nature*, 2004, 431: 371~378
- 6 Derek MD, Carl DN, Phillip A. Killing the messenger: short RNAs that silence gene expression. *Nature*, 2003, 4: 457~467
- 7 Grimm D, Streetz KL, Jopling CL, et al. Fatality in mice due to oversaturation of cellular/short hairpin RNA pathways. *Nature*, 2006, 441: 537~541

## 2 微小 RNA 概述

微小 RNA (microRNA, miRNA) 是由内源性发夹 (hairpin) 结构转录产物衍生而来的一种长为 19 ~ 25nt 的单链 RNA (single-stranded RNA, ssRNA)<sup>[1~3]</sup>。随后 miRNA 可以作为一种引导性分子通过碱基配对与靶 mRNA 结合从而在转录后水平引起靶 mRNA 的剪切或是翻译的抑制。在每种高等生物中, 存在 200 种以上的 miRNA, 可见 miRNA 是最大的基因家族之一, 约占基因组的 1%<sup>[2]</sup>。最近的研究表明, miRNA 在不同的调节途径中发挥关键的作用, 包括发育时序的控制、造血细胞的分化、细胞凋亡、细胞增殖以及器官的形成等方面。最近的研究还发现部分双链 DNA 病毒, 如 EB (Epstein-Barr) 病毒、卡波氏肉瘤相关疱疹病毒 (Kaposi sarcoma-associated herpesvirus, KSHV) 等都可以编码产生 miRNA<sup>[4,5]</sup>, 从而推测 miRNA 在病毒的致病性及病毒与宿主的相互作用方面可能有一定的作用。miRNA 与其靶分子形成一种复杂的调控性网络。例如, 单一一种 miRNA 可以结合并调节多种不同的 mRNA 靶分子, 同样, 几种不同的 miRNA 也可以结合并协调调节同一种靶 mRNA<sup>[6]</sup>。最新研究表明, 人类超过 1/3 的基因都可以被 miRNA 所靶定<sup>[7]</sup>。因此, 通过表达在每种细胞类型中的 miRNA 的作用, 就可以影响成千上万种 mRNA 的功能<sup>[8]</sup>。目前对这种复杂的调节性网络的理解尚处于起步阶段。在这里, 我们将重点对 miRNA 的基本概念, miRNA 的基因家族的基因结构、miRNA 的生物发生过程及 miRNA 的功能等方面进行详细的描述。

### 2.1 miRNA 的发现过程

1993 年, Ambros 等<sup>[9]</sup>发现, 控制线虫幼虫发育时序的一个已知基因 *lin-4* 并不编码产生蛋白质, 而是产生一对小 RNA。一种 RNA 长约 22nt, 而另一种约 61nt。推测较长的一个可以折叠形成一种茎环状结构, 从而成为较短的一种 RNA 的前体。随后, Ambros 等<sup>[9,10]</sup>注意到这种 *lin-4* RNA 与 *lin-14* 基因的 3' UTR 区存在多个互补位点。而这些互补的位点正好位于先前认为的可以由 *lin-4* 产物介导的 *lin-14* 抑制的 3' UTR 区<sup>[11]</sup>, 随后的研究进一步证明了这些互补位点在 *lin-4* 对 *lin-14* 的调节中有着重要的意义, 同时也发现这种调节可以减少 *lin-14* 蛋白的产量, 但是并不影响 *lin-14* mRNA 的水平。总的看来, 这些发现支持一种模式, 那就是 *lin-4* RNA 与 *lin-14* 的 3' UTR 区互补配对, 进而特异性的抑制 *lin-14* mRNA 的翻译, 以此作为一种调节形式, 影响线虫幼虫第一期向第二期的过渡<sup>[9,10]</sup>。

较短的 *lin-4* RNA 现在被认为是一种叫做 miRNA 的基因家族的成员之

—<sup>[12~14]</sup>，随着越来越多的 miRNA 成员，miRNA 调节的靶分子以及 miRNA 功能的发现，miRNA 指导的基因调节的广度及调节的重要性逐渐成为了科学家们研究的热点。最近的研究发现，miRNA 功能涉及细胞增殖的控制、细胞的死亡、脂代谢、线虫中的神经元的形成、哺乳动物中造血细胞系分化的调节以及植物中叶和花发育的控制等各个方面<sup>[2]</sup>。应用计算机的方法对 miRNA 调节的信息进行搜索，发现上述的功能只代表了 miRNA 功能很小的一部分<sup>[6,15~17]</sup>。

在发现 *lin-4* RNA 之后最初的几年中，这类微小调节性 RNA 的基因组学几乎没有任何的进展，*lin-4* RNA 样的 RNA 分子既没发现存在于线虫以外的其他的生物体中，而且在线虫中也没发现有类似于 *lin-4* RNA 的非编码 RNA 分子存在。这种情况直到 *let-7* 发现之后才有了较大的改变。*let-7* 是存在于线虫中的又一种约 22nt 的小调节性 RNA，是在线虫的另一种发育阶段所发现的。研究发现 *let-7* RNA 可以促进晚期幼虫向成虫的过渡<sup>[18,19]</sup>。随后，在人类和果蝇的基因组中，很快的发现了 *let-7* 基因的同源物，而且在人类和果蝇体内也检测到了 *let-7* RNA 的存在<sup>[20]</sup>。

由于 *lin-4* 和 *let-7* 在发育时序的控制中发挥共同的作用，这两种 RNA 最初被统一称为小瞬时 RNA (small temporal RNA, stRNA)<sup>[20]</sup>。但是，在后来不到一年的时间内，三个实验室从线虫、果蝇及人类细胞中克隆出了上百种微小的非编码 RNA，果蝇中约 20 种，线虫中约 60 种，而人类约 30 种<sup>[12~14]</sup>。这些小 RNA 类似于 *lin-4* 和 *let-7*，均为约 22nt 大小，推测其都来源于其茎环状前体的一条臂上，而且都具有一定的进化保守性。但是，与 *let-7* 和 *lin-4* 不同的是，这些新发现的约 22nt 的分子并不是表达在不同的发育阶段，而是表达在特定的细胞类型中。于是，另一个新的名词——microRNA 就产生了，它不仅包括 stRNA，而且包括所有其他的具有相似特征但是功能未知的小 RNA<sup>[12~14]</sup>。

近几年来 miRNA 克隆快速进展使得发现了大量的其他的 miRNA 基因，于是英国的 Sanger 中心专门成立了 miRNA 的数据库 (<http://www.sanger.ac.uk/Software/Rfam/mirna/>)，以对发现的 miRNA 进行统一的归类 and 编号。所有 miRNA 以“miR”作为前缀，在其后加上唯一的标志性的识别号码，如 miR-2、miR-89 等，而编码 miRNA 的基因也用同样的三字母前缀，但应当注意的是，根据生物体惯用的不同，有的生物体中需要大写，用连字符。如在线虫和果蝇体内为 miR-1，而在拟南芥中，应为 MIR156。识别号码按顺序给出，即使在不同的生物体中，同样的 miRNA 有相同的识别号码，有时非常相近的同源物，也可以有相同的识别号码，这是由研究者们根据具体情况决定的。例如，果蝇体内的 miR-1 和线虫及人类的 miR-1 有一个单核苷酸的不同，但是，它们有着相同的序列号。同一物种中相同或非常相近的 miRNA 序列可以有相同的识别号码，用字母或数字后缀表示区别。如果蝇中的 mir-13a 和 mir-13b 指的是序列有轻微差异的约 22nt 的转录本 (又称“转录物”)；而 mir-6-1 和 mir-6-2 则为相同的转录

本<sup>[1]</sup>。

截至 2006 年 6 月 12 日, 英国的 Sanger 中心的 miRNA 注册网 (<http://microRNA.sanger.ac.uk/cgi-bin/sequences/browse.pl>) 上已注册在线的 miRNA 已经达到上千种, 其中包括线虫 (*C. elegans*) 的 114 种 miRNA, 人类 (*Homo sapien*) 的 462 种, 黑腹果蝇 (*Drosophila melanogaster*) 的 78 种, 原鸡 (*Gallus gallus*) 的 144 种, 拟南芥 (*Arabidopsis thaliana*) 的 118 种, 斑马鱼 (*Danio rerio*) 的 372 种。此外, 目前发现有 6 种病毒可编码产生 miRNA, 包括 EB 病毒的 23 种 miRNA, 人巨细胞病毒的 11 种 miRNA, KSHV 的 13 种, 小鼠  $\gamma$  疱疹病毒 (mouse gamma herpesvirus) 的 9 种, 恒河猴淋巴隐病毒 (Rhesus lymphocryptovirus) 的 16 种以及猿猴病毒 40 (Simian virus 40) 的 1 种。随着 cDNA 克隆和计算机预测方法的扩展, 该数字仍在不断地增加。

## 2.2 miRNA 的生物学特征

目前认为, miRNA 是一种长约 22nt 的单链 RNA, 其 3' 端可有 1~2 个碱基长度的变化, 是由一种具有局部发夹状结构的内源性转录本经 RNase III——DCR (Dicer) 加工而来的。它广泛存在于真核生物中, 不编码蛋白质。miRNA 的核苷酸组成具有一定的特征性: 末端具有类似于 siRNA 的特征性结构——5' 磷酸和 3' 羟基。其序列特征为: 5' 端第一个碱基对 U 有强烈的倾向性, 而对 G 有抗性, 但第二到第四个碱基缺乏 U, 一般来讲, 除第四个碱基外, 其他位置碱基通常都缺乏 C<sup>[21]</sup>。这些参数变化并不存在于所有的 RNA 片段中, 是 miRNA 独有的碱基组成。因而 miRNA 和 1.1.3 节提到的 siRNA 在来源、产生、结构及功能等方面既有共性的一面, 也有不同之处。现将二者之间的主要区别列在下表中 (表 2-1)。

表 2-1 siRNA 和 miRNA 之间的区别

项 目	siRNA	miRNA
机制	多为外源性, 如病毒感染、转座子或人工插入 dsRNA 之后诱导产生或人工导入, 也存在部分内源性 siRNA, 如重复相关 siRNA (rasiRNA)	内源性, 机体正常的调控机制
基因座位	无固定的基因座位	有固定的基因座位
直接来源	长链 dsRNA	发夹状前体分子 pre-miRNA
分子特征	双链 RNA, 3' 端有 2 个非配对碱基, 通常为 UU	单链 RNA
对靶 RNA 特异性	较高, 一个突变容易引起 RNAi 沉默效应的改变	相对较低, 一个突变不影响 miRNA 的效应

项 目	siRNA	miRNA
作用方式	RNAi 途径	miRNA 途径
生物发生的亚细胞定位	细胞质	从细胞核到细胞质
加工过程	对称的来源于双链 RNA 的前体	不对称加工, miRNA 仅是 pre-miRNA 的一条臂
Dicer 蛋白	在存在 Dicer 同源物的情况下, Dicer 同源物有不同的分工	同左
AGO 蛋白	真正发挥作用的 AGO 蛋白存在一定的分工	同左
RISC 中的效应分子	单链分子, 是 siRNA 的引导链	为 miRNA 分子本身
互补性	多为完全互补	多为不完全互补, 存在错配现象
作用方式	多为剪切靶 mRNA	多为抑制靶 mRNA 的翻译
对 RNA 的影响	降解靶 mRNA, 影响 mRNA 的稳定性	在 RNA 代谢的各个层面进行调控; 与 mRNA 的稳定性无关
作用位置	可作用于 mRNA 的任何部位	主要作用于靶 mRNA 的 3'-UTR 区
重要特性	特异性、高效性	高度的保守性、组织特异性和发育阶段特异性
生物学意义	生物抵抗外来核酸片断入侵的一种方式, 原始活性为抑制转座子活性和病毒感染	参与生物调节的各个层面, 包括发育、分化、凋亡等

但是, 随着研究的深入, siRNA 和 miRNA 的界限已越来越不明确了, 两者加工和表达机制上的相似性, 包括二者同为 DCR 加工的产物, 同为 RISC 的组分, 二者的存在都需要 AGO 蛋白的存在以及二者作用的发挥均依赖于小 RNA 与靶 mRNA 之间的互补程度, 进而剪切或抑制翻译等, 说明其沉默机制中存在部分的重叠, 使人们推测它们都是早期 RNA 在进化上的遗留物, 并逐步被蛋白质调控所代替。现在争论较多的是这两种机制的从属关系和进化顺序, 以前曾有人认为 siRNA 是 miRNA 的补充, 因为 miRNA 参与正常情况下生长发育的调控, 而 siRNA 不参与动物体的正常生长, 只有在外源性 dsRNA 诱导下才产生。到底哪个出现得更早些, 是 RNA 的彻底降解还是表达的抑制? 大部分科学家更倾向于在早期的生命进化中, 类似于 siRNA 的调控机制出现得更早, 并且它可能是早期基因调控的主要方式。因为外源基因的入侵在早期的生物进化中就出现了, 采取这种方式就能比较有效地破坏外源基因并且防止转座子转移, 但是这种基因的抑制方式过于彻底, 导致调节的不可逆, 而且费力费时, 所以在进化过程中可能逐步产生了 miRNA 来代替它, 从而使 miRNA 成为发挥调控作用的主导分子<sup>[21]</sup>。

为了避免将 siRNA 或其他 RNA 也归为 miRNA, Ambros 等<sup>[1]</sup>根据 miRNA 的生物学特征, 从表达和生物发生等方面规定了一系列 miRNA 必须符合的标准。现将其逐一描述如下:

### 1) 表达标准。

A. 小 RNA 的表达必须通过与 RNA 样本的杂交来确证, 最常用的是 Northern 印迹。当然, 其他的检测方法也是可行的, 包括有 RT-PCR、引物延伸分析、RNA 酶保护分析及微阵列分析等。然而, Northern 印迹通常是首选的确证 miRNA 的方法, 因为这种方法通常可以同时检测到成熟的 miRNA (约 22nt) 和 miRNA 的发夹状前体 (约 70nt) 的存在。

B. 克隆得到的 cDNA 文库中的约 22nt 的序列必须与其来源的生物体的基因组序列精确匹配。

### 2) 生物发生标准。

C. 小 RNA 序列通常存在于发夹状前体的一条臂上, 这种发夹状前体缺乏内部大的茎环状结构或转角, 尤其不能有大的非对称性的转角。这种前体分子在动物中长 60~80nt, 而在植物中变异较大, 有的可以长达几百个核苷酸。在这条标准中, 发夹状的前体必须以最低自由能的形式进行折叠, 这一点可以用 mfold 程序或其他传统的 RNA 折叠程序进行预测, 而且在发夹结构的茎部最少有 16bp 配对的碱基, 其中必须包含有约 22nt 的 miRNA 序列和其他的发夹部分的序列。

D. 小 RNA 序列在生物进化上具有高度的保守性。这种序列的保守性通常在前体发夹状结构中也可以见到, 但是, 对较成熟的 miRNA 片段而言, 其保守性相对较低。这种保守性的发夹结构应当符合标准 C 中的最小配对要求, 但是并不要求其折叠的能量最低。

E. 如果在 DCR 作用弱化的情况下, miRNA 的前体分子在细胞中堆积, 则进一步加强了 miRNA 存在的证据。但是, 鉴于在某些细胞中弱化 DCR 作用的技术存在一定难度, 这一标准并不常用。

因为单独的表达证据 (A 或 B) 都不能排除 siRNA, 而且发夹状的结构 (C 或 D) 并不是 miRNA 生物发生过程中所独有的特征, 也不是 DCR 加工的唯一特征 (E)。因此, miRNA 表达和生物发生的证据都是必需的。最理想的状态是, miRNA 是体内的约 22nt 的非对称性的产物, 而且由 DCR 加工具有进化保守性的发夹状前体分子而来, 即 A+D+E。但是, 如果没有 DCR 加工的证据, A+D 也足够了。应当注意到在表达标准中, A 相对强于 B。相似的是, 对发夹状前体相关的标准而言, D 相对强于 C。D 的进化保守性是进行前体结构预测的前提。因此, 标准 C 应当有强烈的表达证据的支持, 即 A+C; 而标准 B 应当有强烈的生物发生证据的支持, 即 B+D。当约 22nt 的 miRNA 并没有被克隆或检测到, 但是如果可以检测到其具有生物进化保守性的发夹状前体, 而且在 DCR 功能弱化之后该前体分子 (D+E), 则该候选基因仍然可以被归为 miRNA。但是,

后一种情况应当被视为特殊情况，因为其他类型的包含有发夹状结构的 RNA 也可以被 DCR 进行代谢。

已经确证的 miRNA 在其他物种中的同源物不一定要符合上述的严格的标准，即使没有实验的确证，只要是已知的 miRNA 的非常近的同源物就符合标准 D，那么就可以被认为是 miRNA。如果一种序列反复多次在 cDNA 克隆中被克隆到（符合标准 B），那么，即使该序列和基因组不完全匹配，但是和一种已知的 miRNA 非常相似，则可以被认为是这种已知的 miRNA 的变异体。不符合上述要求的小 RNA 分子则被视为 siRNA 或其他种类的小 RNA 分子。

## 2.3 miRNA 的基因组学

### 2.3.1 miRNA 的基因及其转录

早期的研究指出尽管有少部分的 miRNA 存在于已知基因的内含子中，包括其正义链和反义链方向，但是大部分的 miRNA 在基因组水平定位于基因间区（距离已知或预测的基因大于 1kb 的位置）<sup>[12,13]</sup>。这就使得人们开始假设 miRNA 基因是一种自主转录单元。另外一个很有趣的现象是已知的 miRNA 中有约 50% 与其他的 miRNA 相邻<sup>[12,13]</sup>，这就增加了一种可能性，即这些聚集的 miRNA 可能由单一个多顺反子转录单元转录而来。最新的关于 miRNA 基因表达的研究表明，miRNA 基因可以有其自己本身的转录启动子<sup>[22,23]</sup>，而且这种聚集的 miRNA 来源于多顺反子形式的原转录本（primary transcript of miRNA, pri-miRNA）<sup>[13]</sup>。尽管目前尚不能排除小部分的 miRNA 基因可能由其他的 RNA 聚合酶转录而来的可能性，但是目前一般认为 miRNA 基因由 RNA 聚合酶 II（Pol II）转录而来<sup>[11,12]</sup>。起初认为 miRNA 的转录是由 Pol III 介导的，因为大多数的小 RNA，如 tRNA 等都是由 Pol III 进行转录的，但是，发现有的时候 pri-miRNA 可以长达几千个碱基，而且在其末端有超过四个的连续的尿嘧啶，这将会终止 Pol III 的转录<sup>[24]</sup>。几个包含有 miRNA 序列的具有多聚腺苷酸（polyA）尾的转录本以及邻近的一个 mRNA 已经用表达序列标签（expressed sequence tag, EST）分析和基因结构分析鉴定了出来<sup>[25~27]</sup>。这也从另一个侧面揭示了 miRNA 与其他基因类似，在发育过程及不同的组织中，处在 Pol II 的精确调控之下<sup>[28,29]</sup>。此外，一个具有功能的成熟的 miRNA 也可以由包含有 pri-miRNA 片段的质粒在异源 Pol II 启动子的作用下产生<sup>[23,30]</sup>。最近的研究报道了三方面直接的证据。首先，pri-miRNA 具有帽子结构（cap structure）和 polyA 尾<sup>[23]</sup>。第二，miRNA 转录本活性对一定浓度的专门抑制 Pol II 的  $\alpha$ -鹅膏蕈碱（ $\alpha$ -amanitin）敏感，而对其他浓度的专门抑制 Pol I 或 Pol III 的  $\alpha$ -鹅膏蕈碱不敏感<sup>[22]</sup>。第三，染色质免疫沉淀实验（chromatin immunoprecipitation analyse）已经证明了 miR-23a~27a~24-2 的启动子与 Pol

II 相关<sup>[22]</sup>。Pol II 指导的转录可能具有如下的好处，例如，miRNA 基因的转录可以被不同的 Pol II 调节因子所控制，从而使得在不同的发育阶段、不同的细胞类型中表达特异的 miRNA。而且，miRNA 和编码蛋白质的基因的表达可能存在彼此协调的作用，尤其是二者同时位于一个转录本时更为明显。最近关于 miRNA 基因定位与已知的转录单元之间的关系的研究为认识 miRNA 的生物发生过程掀开了新的篇章<sup>[26]</sup>。Rodriguez 等<sup>[26]</sup>通过分析 miRNA 基因组定位与已知的转录单元之间的关系，发现约 70% (232 个中有 161 个) 的哺乳动物的 miRNA 基因定位于确定的转录单元上。其中 117 种 miRNA 基因存在于内含子的有义链方向上，30 种 miRNA 与非编码 RNA 的外显子重叠，而 14 种 miRNA 依据剪接形式的不同而可以与外显子或内含子重叠 (称之为混合的 miRNA)。在这 117 种内含子的 miRNA 中，90 种 miRNA 位于蛋白质编码基因的内含子中，而 27 种 miRNA 位于非编码 RNA 的内含子中。因此，根据基因组定位的不同，miRNA 基因可以分为如下几类<sup>[31]</sup>：第一，非编码转录单元的外显子 miRNA，如 miR-155；第二，非编码转录单元的内含子 miRNA，如 miR-15a 位于非编码 RNA 基因 *DLEU2* 的内含子区；第三，蛋白质编码转录单元的内含子 miRNA，例如，miR-25 位于编码 DNA 复制允许因子 MCM7 转录本的第 13 个内含子中。混合的 miRNA 基因根据其剪接形式的不同而可以归为上述三类中的任一类。一个明显的问题是内含子或外显子的定位是否影响 miRNA 的生物发生过程，或者说剪接和 pri-miRNA 的加工之间是否存在联系？目前一般认为剪接发生在 pri-miRNA 的加工之前，而且生成的内含子套索被加工然后释放出 miRNA 前体 (miRNA precursor, pre-miRNA)。这一点类似于小核仁 RNA (small nucleolar RNA, snoRNA) 的产生。在线虫中，let-7 上游的反式剪接是 pri-miRNA 的加工所必需<sup>[17]</sup>，反式剪接反应是用剪接前导序列 1 (spliced leader 1, SL1) 替代了 let-7 茎环上游的序列，改变了茎环结构附近的二级结构从而使得剪接的转录本可以更有效的被加工。但是，这一例子并不可能代表 miRNA 成熟的一般模型，这是由于 miRNA 序列通常位于远离剪接界 (splice junction) 的位置。另一个有趣的现象是同时包含有蛋白质编码序列和 miRNA 序列的单一一个转录本。最近研究发现，同时包含有萤光素酶 cDNA 和 miR-21 前体序列的转录单元可以产生萤光素酶蛋白和 miR-21<sup>[23]</sup>。那么，究竟是单一一种原转录本可以产生 miRNA 和蛋白质，还是每个原转录本将不得不在 miRNA 途径和 mRNA 途径中进行选择，至今尚不清楚。

mRNA 和 miRNA 通常具有相似的表达过程，这说明这些 miRNA 可以作为长的宿主转录单元的一部分进行转录<sup>[26]</sup>。目前，只有很少一部分 miRNA 的启动子在实验中被鉴定出来<sup>[22,23,27,32]</sup>。那么关于不同种类的 miRNA 启动子的一般特征有待进一步研究。找到一些 miRNA 特异性的元件对 miRNA 基因的预测和 miRNA 生物发生过程的预测都将有着不可估量的作用。而且，目前已经发现，在几乎所有的线虫 miRNA 的上游存在着一种高度保守型的序列，为约 200bp 大

小的“CTCCGCCC”模序，但是这一序列在转录控制中的重要性尚待进一步研究<sup>[33]</sup>。在其他生物中尚未发现类似模序的存在<sup>[33]</sup>。此外，还发现 miRNA 可以受到转录水平及转录后水平的调节。前者表现为 miRNA 受到转录因子的调节，C-myc 为一种调节细胞增殖的转录因子，它可以激活人类 13 号染色体上的包含有 6 个 miRNA 的基因簇的表达<sup>[34]</sup>，而且也发现在人类胚胎干细胞中存在一套特异的 miRNA，其中有很多受到同样的调节并且成簇存在<sup>[35]</sup>。而 miR-182 仅限于在特定的组织类型中表达，而其前体分子则广泛存在于各种组织中的现象则体现了 miRNA 受到了转录后水平的调控<sup>[36]</sup>。

### 2.3.2 miRNA 的表达形式

许多种 miRNA 都有其独特的表达形式，例如，线虫中的 lin-4 和 let-7 RNA 的同源物在发育阶段有着阶段特异性的表达<sup>[12~14]</sup>，而有的则具有组织特异性的表达，如 miR-1 专一的表达于哺乳动物的心脏中<sup>[14,37]</sup>，而 miR-223 则主要表达于小鼠骨髓的粒细胞和巨噬细胞中<sup>[38]</sup>；mir-35-mir-42 则主要表达于线虫的胚胎中<sup>[13]</sup>。而且 mir-290-mir-295 只在小鼠的胚胎干细胞中表达，而在分化的细胞中不表达<sup>[15]</sup>。表达芯片技术已经被用来检测 miRNA 并且已经揭示出了在不同的发育阶段或哺乳动物的脑的不同区域中 miRNA 都有不同的表达形式。由于上述的 miRNA 表达的阶段特异性和组织特异性，有理由认为线虫不同发育阶段的每种细胞类型都可能有不同的 miRNA 表达谱，这就使得生物体的转录调控的网络更为复杂。

miRNA 表达的另外一个显著的特点是细胞中某一种 miRNA 表达的绝对量。例如，miR-2、miR-52 和 miR-58 在每一个成体蠕虫细胞中平均大于 50 000 个分子<sup>[39]</sup>，这一表达水平相当高，而这种高表达水平是由于转录加强还是由于降解减慢至今尚不清楚。还有一些 miRNA 表达水平相当低，例如，miR-124 在每个成体的蠕虫细胞中只有 800 个分子<sup>[39]</sup>。这种表达平均较低的水平可能是由于该 miRNA 在多种细胞中的表达水平较低但是在某些特定的细胞中表达很高所致。小鼠中的 miR-124 专一的表达于脑组织中，更支持后一种解释<sup>[37]</sup>。

### 2.3.3 应用生物信息学方法对 miRNA 进行预测

人们都很好奇为什么 miRNA 这么大的一个基因家族没有早被发现，很清楚并不是因为其种类稀少的缘故。miRNA 及其相关的蛋白质复合体似乎是占到细胞中的核蛋白复合体的大部分。但是，miRNA 的表达受限于阶段特异性和组织特异性，而且 miRNA 表达量相对较低，因而在克隆过程中一直被忽略。这样，计算机的方法就在一定程度上弥补了 miRNA 基因鉴定的方法。早期应用计算机的方法进行同源性搜索，而找到了 miRNA 基因的直系同源物 (ortholog) 和共生同源物 (paralog)<sup>[12~14]</sup>。另一个相似的方法是寻找已知的 miRNA 基因附近的其他的茎环状结构，从中找出那些可能代表基因簇的其他的基因<sup>[13,33]</sup>。因为新发现的好多的

miRNA基因都是串联排列与操纵子样的基因簇中的，而且序列的差异使得应用一般的方法具有一定的难度，因此这种方法对寻找 miRNA 基因是非常重要的。

不依赖于已知基因的同源性或相似性的基因寻找方法已经被应用于整个基因组中<sup>[19]</sup>。首先由位于推测的蛋白质编码区外开始，寻找可能形成茎环结构的序列，然后参考其进化保守性及与已知的 miRNA 基因的类似性，对这些候选的 miRNA 茎环结构进行评分。目前应用较为成功的敏感的评分工具为 MIRscan 和 miRseeker。前者已经成功地应用于线虫和脊椎动物<sup>[39]</sup>，而后者则成功地应用于昆虫<sup>[41]</sup>。二者都鉴定出了数十种后来被确证正确的 miRNA。鉴于其相对高的敏感性，这些方法被成功地用来预测人类、线虫及果蝇等生物的基因组中的 miRNA 基因，每种生物中这些计算机预测出来的 miRNA 的数量占到了全基因组的 1%，可见 miRNA 家族是相当大的一个基因家族。

但是，上述的这些方法还不能发现那些同源性不高的 miRNA。最近，Nam 等<sup>[42]</sup>在上述模型的基础上，提出了一种新的寻找 miRNA 基因的概率共分析模型——ProMiR。这种模型同时考虑了 miRNA 前体的结构和序列特点，采用配对的隐马尔科夫模型（hidden Markov model, HMM）来鉴定同源关系近或远的 miRNA。这一模型将 miRNA 基因的序列和结构在概率框架下联系起来，通过检测 Drosha 剪切位点的信号，验证 miRNA 基因和成熟 miRNA 区域是否存在。通过对人类第 16、17、18、19 号染色体进行基因组筛查，ProMiR 至少成功地发现了 23 种新的 miRNA 候选基因。而且，这些 miRNA 的候选基因没有显示出与已知 miRNA 明显的序列相似性。用 RNAi 的方法抑制 Drosha 的表达后进行定量 PCR，23 个候选基因中的 9 个是可以被 Drosha 加工的转录本。这揭示出 ProMiR 预测的 miRNA 基因至少有 40% 的准确性。

这种推测暗示着在真核细胞中存在着大量的 miRNA 基因，到目前为止，注册在线的 miRNA 的基因以及含有 miRNA 的生物的种类正在不断地增加，而且其扩展的范围已经涉及了病毒的基因组。目前已经发现了 6 种病毒可以编码病毒的 miRNA，从而对理解病毒的致病性及其与宿主的相互作用提供了一个新的理解角度。

#### 2.3.4 应用 miRNA 表达谱微阵列技术检测基因组范围 miRNA 的表达

目前检测 miRNA 的技术较多，包括 Northern 印迹、斑点印迹法、RNA 酶保护分析、引物延伸分析和定量 PCR 等。但是可以同时检测多种 miRNA 的技术相对有限，大范围的 cDNA 克隆可以提供不同样本中 miRNA 的相对表达水平，但是这种技术相对耗时耗力，而且很难确定所有已知 miRNA 的表达水平。最近研究出的 miRNA 表达谱微阵列技术因其高通量、高灵敏度而被广泛应用于同时检测某一样本中多个 miRNA 表达水平的变化。miRNA 是生物体内一类大的基因家族，每种生物体至少编码上百种 miRNA 分子，而每种 miRNA 分子可以同

时调控多个 mRNA 的表达，同一 mRNA 也可以同时受多个 miRNA 的调控。也就是说，某一基因表达的状态是多个 miRNA 共同作用的结果，因此，应用 miRNA 微阵列检测技术来同时检测多个 miRNA 的表达水平就显示出了其极大的优势，它不仅可以同时体现某一正常或疾病状态的多个 miRNA 的表达水平，而且可以高通量的分析不同状态下同一 miRNA 的表达水平，这些都有助于真实挖掘 miRNA 的生物学作用。目前，miRNA 表达谱微阵列检测技术已经成功应用于筛选疾病相关或发育分化相关的 miRNA，如造血干细胞中可能与发育相关的 miRNA 的发现<sup>[31]</sup>，某些参与脑分化的 miRNA 的发现<sup>[43]</sup>等，同时应用 miRNA 微阵列技术也可以研究某些肿瘤中 miRNA 的表达水平<sup>[44,45]</sup>，并通过统计学分析从中寻找出有一定统计学意义的表型，以达到从组织和分子信号水平将肿瘤与正常组织区分开来的目的。目前这种方法已经被用来检测肺癌，并发现 miRNA 的表达情况与包括 I 期肿瘤在内的肺腺癌的生存相关。单因素分析表明，hsa-miR-155 高表达和 hsa-let-7a-2 低表达与预后不良有关；多因素分析表明，hsa-miR-155 高表达是预后不良的因素<sup>[46]</sup>。可见，通过 miRNA 表达谱微阵列技术使得发现一种或一簇疾病相关或发育相关的 miRNA，或者使得通过检测 miRNA 的表达水平为疾病的诊断及预后提供一定的线索成为可能。

此外，最新的研究将 miRNA 表达谱微阵列技术和生物信息学技术相结合，可以高效且特异地发现生物体中的新 miRNA<sup>[47]</sup>。在研究人  $\gamma$  疱疹病毒时，首先应用生物信息学的技术预测了病毒基因组中可能形成发夹结构的序列，然后合成其前体分子进行芯片杂交，阳性结果用 Northern 印迹进行验证，结果发现 EB 病毒除了编码原来发现的 5 种 miRNA 之外，还编码 18 种新的 miRNA。可见，结合这些方法可以大大加快 miRNA 发现的进程，从而为更全面更详细的了解 miRNA 的生物学意义奠定一定的基础。

### 2.3.5 植物中 miRNA 基因组学的特殊性

已经成功地从植物中克隆到了 21~24nt 大小的小 RNA<sup>[16]</sup>，这些小 RNA 分子与线虫中的 miRNA 非常的类似，也是内源性的约 22nt 的 RNA，多来源于发夹状前体的一条臂，具有进化上的保守性，而且来源于已知基因以外的区域<sup>[16]</sup>。故而也将其归类为 miRNA。目前为止，已经鉴定出了拟南芥中存在有 117 种 miRNA，这是 miRNA 克隆和计算机方法应用的结果。动物和植物中都发现这一类小 RNA 分子的存在，使得人们开始设想这类调节性小 RNA 分子可能是生物体中最重要的一类调节性分子。

但是，植物和动物中的 miRNA 在一些方面存在着不同，而且这些不同似乎与 miRNA 的生物发生过程相关。最大的不同是 miRNA 前体的茎环结构：植物中 miRNA 前体的茎环结构的大小通常可变，而且一般要大于动物中 miRNA 的前体分子<sup>[16]</sup>。此外还存在一些细微的区别，包括植物中 miRNA 与茎环结构的另

一条臂之间的配对相对多一些，即在植物中 miRNA 与靶 mRNA 之间接近完全匹配；植物中约 21nt 大小的 miRNA 相对多见，而动物中 22~23nt 的 miRNA 则偏多一些；而且植物中的 miRNA 的 5' 端多为 U<sup>[13,16]</sup>。上述的不同以及在植物和动物的 miRNA 中几乎没有保守性，使得人们开始推测 miRNA 基因独立的起源于这两种多细胞种系中。即使在两种起源的情况下，植物和动物中均存在有 miRNA 基因，表明或许这两种来源发生在多细胞生物发生之前并在多细胞生物的形成中可能发挥一定的作用。

## 2.4 miRNA 的成熟及效应过程

### 2.4.1 miRNA 的成熟

从最初在细胞核中生成 pri-miRNA 开始，到生成成熟 miRNA 需要经过一系列生物学过程，包括有 pre-miRNA 的产生，pre-miRNA 的输出以及成熟的 miRNA 的生成。具体过程如图 2-1 所示。

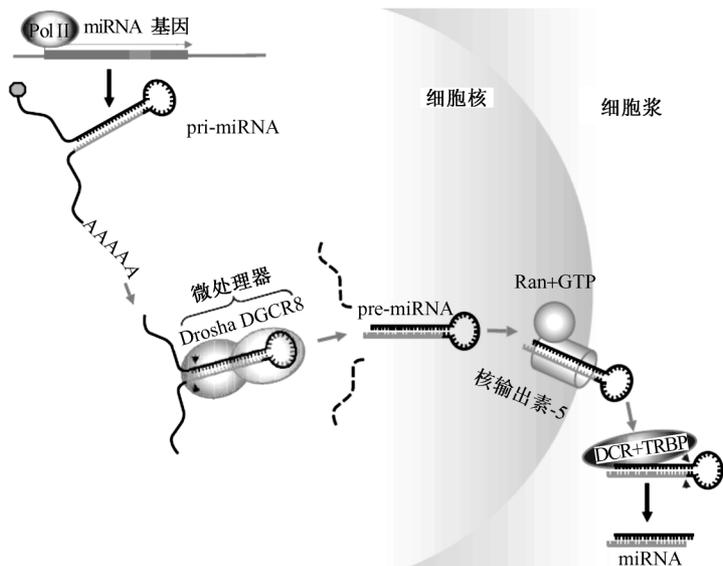


图 2-1 人类细胞中 miRNA 生物发生过程

miRNA 的生物发生过程包含有如下步骤：包含有一个或几个 miRNA 的长的 pri-miRNA 在 RNA Pol II 作用下转录而来，其具有 5' 端的帽子结构和 3' 端的 polyA 尾。然后，pri-miRNA 在微处理器（Drosha 及 DGCR8）作用下，识别不同的 pri-miRNA 的二级结构，并在其根部进行剪切，形成约 70nt 的 pre-miRNA。这些过程均发生在细胞核中。然后，pre-miRNA 在 Exp-5 的作用下，从细胞核中转运到细胞浆中，然后在细胞浆 DCR 复合物作用下产生成熟的 miRNA 分子<sup>[48]</sup>

整个 miRNA 的生物发生过程贯穿了细胞核和细胞浆。pri-miRNA 是在细胞核中，由 RNA Pol II 转录而来的，具有一个或多个发夹状结构，呈多顺反子形式排列。然后，生成的 pri-miRNA 在细胞核中生成约 70nt 的茎环状前体分子 pre-miRNA。这一过程需要 RNase III 的作用，拟南芥中发挥作用的为 DCL1，动物、线虫和果蝇中发挥作用的为 Drosha 蛋白。产生的 pre-miRNA 分子中的一端即为将来成熟的 miRNA 的一个末端。然后，生成的 pre-miRNA 从细胞核转运到细胞浆中，这一过程需要核输出蛋白核输出素-5 (exportin-5, Exp-5) 的作用。最后，pre-miRNA 将在另一种 RNase III——DCR 作用下，被剪切产生成熟的 miRNA 的另一端，这样，就产生了 miRNA 的双链中间体结构——miRNA: miRNA\* (miRNA 的反义链)。然后，其中的一条被降解而另一条即为成熟的 miRNA。最新的研究结果显示 pre-miRNA 的剪切是在 RNA 诱导的沉默复合体 (RNA-induced silencing complex, RISC) 中 DCR 的作用下完成的<sup>[49]</sup>，也就是说 miRNA 的成熟与效应过程是伴随发生的。然后 RISC 中的 Argonaute (AGO) 蛋白剪切 miRNA\*，而使 miRNA 成熟<sup>[50]</sup>。这些过程的细节将在 4.2 节和 5.4 节中给予详细的阐述，在此将不予重复。

## 2.4.2 miRNA 的效应过程

效应过程中形成的 RISC 复合体称为 miRISC，即 miRNA 诱导的沉默复合物。miRISC 中成熟的 miRNA 链即可与其靶 mRNA 结合而调控靶基因的表达。这里，需要注意的有两点：第一，组装进入的 RISC 中的 miRNA 尽管从理论上而言，是随机选择的结果，但是，由于两条链的稳定性不同，所以使得 miRNA 的链具有一定的倾向性，一般是较不稳定的一条链被组装进入 RISC 中（多为 5' 端有突起的那条链）。第二，miRNA 究竟在体内发挥剪切 mRNA 还是抑制翻译的作用，取决于 miRNA 与靶序列结合位点处匹配的程度，而用于识别 miRNA 的进入何种生化途径的 AGO 蛋白家族中的很多成员都是 RISC 的组分<sup>[51]</sup>。目前认为 miRNA 与靶 mRNA 的作用模式包括以下三种：①二者不完全互补，即二者不完全配对结合时，主要影响翻译过程，如线虫的 lin-4；目前有研究发现二者不完全互补时，还可能加速 mRNA 的脱腺苷化 (deadenylation) 作用，从而造成了 mRNA 的不稳定，于是易于被细胞中的 3' 外切核酸酶降解而致 mRNA 水平降低<sup>[40]</sup>；②二者完全互补，即二者完全配对结合后，类似 siRNA 与靶 mRNA 的结合，特异性的切割 mRNA，如 miR39/miR171；③上述两种模式均具备。当其与靶 mRNA 完全互补配对时，直接剪切靶 mRNA，而不完全互补配对时起调节基因翻译的作用。如 let-7 在果蝇体内与靶 mRNA 完全匹配，所以引起靶 mRNA 的剪切，而在线虫体内，则与靶分子不完全匹配，从而只起到抑制翻译的作用。可见，miRNA 就是通过与靶 mRNA 之间的序列的互补程度的不同，而发挥对靶 mRNA 的基因调节作用。不同生物体中 miRNA 发挥作用所涉及的具体过程存在