

4. 利用启动子分析软件对所有的 miRNA 进行了分析,最后鉴定出了 142 个 miRNA 形成了 53 个簇,由上游的启动子控制表达,其中 28 个簇由多个同源的 miRNA 基因组成,18 个 miRNA 簇形成 8 个不同的同源簇。进一步的分析表明有些簇是通过重复的形式产生多拷贝。
5. 研究结果表明 8 号染色体上存在一个与人具有高度同源的印迹区域,在这个区域发现一个最大的由 41 个 miRNA 组成的 miRNA 基因簇,另外两个簇则位于该簇的上游,分别是由 5 个 miRNA 组成的基因簇以及 C/D snoRNA 基因簇。进一步分析表明这最大的 microRNA 基因簇由三个基因家族组成,这三个家族分别是由不同的 microRNA 通过串联重复而产生的。
6. 研究发现,15 个 miRNA 位于基因组的重复区域。其中 7 个位于 LINE 元件中,其它大多数位于 SINE 中。它们有些完全位于重复元件内,有些则与元件部分重叠。而且还有 2 个 miRNA 位于 rRNA 基因的内含子中。
7. 通过分析发现大多数 miRNA 在动物的进化过程中都是保守的。研究发现,大多数 miRNA 是通过局部串联重复以及非局部重复形式而进化的。串联重复在邻近区产生多个拷贝,而非局部重复发生在不同染色体之间,从而产生基因或者基因簇的迁移。
8. 对所鉴定的 miRNA 靶序列的预测分析表明,大多数 miRNA 所调节的靶序列参与了几乎生物体所有的细胞过程,比如生长发育、代谢、调控因子的调节以及信号转导作用等等。其中参与的最多的功能可能是调节转录因子的表达。同时有些 miRNA 可能参与了一些癌症相关的过程。更为重要的是,我们鉴定出了狗特异性的靶序列,它们可能参与了狗生长过程中的一些特殊过程。最后,我们发现 miR-186 能够调节其宿主蛋白的表达,从而影响其表达量。这是一种自我调节的反馈机制,相互制约,相互影响。

关键词: 微小 RNA(microRNA); 基因组织形式; 同源簇; 反馈抑制

ABSTRACT

MicroRNAs, as a family of non-coding small RNA 18-25nt in length, have been profoundly studied in recent years in plants, animals and virus. As an important regulator in eukaryote organism, many study work had done to prove that microRNA have abroad roles in the development of organism involving every aspect of cell physical activity. The characterization of miRNA is still a hot academic domain for the cause of the identified microRNA far from saturation. The scientists had completed sequencing of dog genome, and through human/dog genomic pair alignment, it is found that dogs and humans have similar levels of overall nucleotide diversity, meanwhile, the species display extreme phenotypic diversity, therefore, all provide a good flat for research disease related with human. More importantly, six miRNA were characterized in miRBase thus far. Above all, these provide a great academic and realistic meaning for identifying miRNAs in dog genome.

As for time- and tissue- speciality of miRNA expression and mis-leading of degradation of mRNA and other non-coding RNAs, there is limitation by using experimental RNomics. Therefore, we apply computational RNomics to predict miRNA in dog based on sequence conservation of mature miRNAs and their precursor displaying hairpin structure, leading to identification of 357 novel miRNAs. The following is the summarization of our result:

1. The loci of all 357 microRNA sequences in genome lie in all 33 chromosomes of genome. Among these miRNAs, 281 sequences belong to single-copy in dog genome and the other 35 sequences belong to multi-copy. The number of sequences in intergenic region is 222, in intron is 125, anti-sense to gene is 50, sense to gene is 75, and in exon is 6, in UTR is 3. All precursors ranged from 65 to 120 nt in length, with average length of 80 nt.
2. We have analyzed whether all miRNA could form stem-loop when joined with flanking sequence in genome through m-fold programs and the results showed that a majority of miRNA have this property. Homologue of human miRNAs had been validated by experiment, indicating that they can express stably. Meanwhile, 8 novel miRNAs were found by searching EST database. Furthermore, we carry out RT-PCR assay of 5 novel miRNA, two of which can obtain the expected result, suggesting that they can express.

3. Trough analysis, we found that the 300 homologue of human miRNA belong to 149 families, and miR-154, miR-17 and let-7 family have 15, 13 and 10, respectively. Most of families only have one members, the case is different from plant. Members of animal family display three mode, identical, closely related or homology.
4. For conforming miRNAs are transcribed under common promoter, we analysis these miRNAs using promoter analysis program, leading to identification of 142 miRNAs grouping into 53 clusters controlled by promoter in upstream, among which 28 clusters is made up of multi homologus miRNAs and 18 miRNA cluster form 8 distinct paralogous clusters. Further analysis indicate that some cluster have multi-copies through duplication.
5. Interestingly, we found a large miRNA cluster made up of 41 miRNAs, which reside in imprinted region of chromosome 8 high conserved between human and dog, and another two clusters lie in upstream of this cluster, a miRNA cluster composed of 5 miRNA and C/D snoRNA cluster.
6. We found that 15 miRNA occur in repetitive region in genome, among which 7 lie in LINE element and the others reside in SINE elements. Some of them completely reside in repetitive region, and some overlap with element. Two miRNAs reside in intron of rRNA gene.
7. A majority of miRNAs are conserved among animal through analysis. Most of miRNAs evolved by local tandem duplication and non-local duplication. Local tandem duplications result in paralogous sequences that are located on the same transcript. Non-local duplications result in paralogous genes or gene clusters on different chromosomes.
8. We predict target of identified miRNAs using 3-UTR. Further analysis indicate that a majority of target involved in almost all cellular processes, for example, development, metabolism, regulatory factor and signal transduction. The most riched GO term is transcriptional factor. Meanwhile, some miRNAs might are involved in processes related with cancer. Most importantly, we identified target gene unique to dog, which regulate special processes during dog development. Finally, we found that miR-186 can regulate expression of its host protein, result in turnout of its protein. This is a self-regulatory feedback mechanism.

Keywords: microRNA; genomic organization; paralogous cluster; feedback inhibition

学位论文独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得 南昌大学 或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名(手写): 周冬根 签字日期: 2007 年 6 月 25 日

学位论文版权使用授权书

本学位论文作者完全了解 南昌大学 有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人授权 南昌大学 可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

(保密的学位论文在解密后适用本授权书)

学位论文作者签名(手写): 周冬根

导师签名(手写): 罗士萍

签字日期: 2007 年 6 月 25 日

签字日期: 2007 年 06 月 26 日

学位论文作者毕业后去向:

工作单位:

电话:

通讯地址:

邮编:

第 1 章 综述

1.1 引言

RNA是生物体内最重要的物质基础之一,它与DNA和蛋白质一起构成生命的框架。但长久以来, RNA一度被认为仅仅是DNA和蛋白质之间的“过渡”,它从DNA 那儿获得自己的顺序,然后将遗传信息转化成蛋白质。然而,一系列的研究表明,这些小分子RNA事实上操纵着许多细胞功能,它可通过互补序列的结合反作用于DNA,从而关闭或调节基因的表达。microRNA (miRNA) 是新近发现的一种非编码的小分子RNA,长度大约在17-25nt,它们对真核生物的生命活动的多方面产生重要的影响,是调节生命活动的一个重要调节分子。目前已经知道它在染色体消除、染色体重组、DNA复制、以及基因调控等方面产生重要的影响。从目前所积累的实验证据可以肯定它涉及到了整个细胞水平的几乎所有事件。

1.2 microRNA 基因

1.2.1 microRNA的发现过程

microRNA 的发现最早要追溯到上世纪九十年代。当时人们是用遗传筛选的方法寻找线虫 (*C.elegans*) 胚胎后发育的缺损体,发现了一个控制线虫幼虫发育时序性的非编码基因 *lin-4*, 该基因产生一对小分子 RNA。其中一个长约 22nt, 另一个约为 61nt, 较长的 RNA 可折叠为茎环样结构,被认为是较短 RNA 的前体^[1-2]。Lin-4 可抑制 *lin-14* 基因表达蛋白质的水平,而 *lin-14* 在发育路径中起着重要的作用。最终的研究结果表明:正是 *lin-4* 和 *lin-14* 的 3' 非翻译区这种直接的但不完全的配对作用是 *lin-4* 调节着 *lin-14* 的蛋白质合成的原因。即 *lin-4* 所编码的小分子 RNA 与 *lin-14* 的 3' 非翻译区的 7 个位点不完全互补结合后,抑制了该基因的蛋白质的翻译;而 *lin-4* 突变体由于 *lin-4* 不能正常的表达,因而 *lin-14* 就会一直处于无抑制状态,它就会一直抑制线虫从幼虫向成虫的过渡,从而出现上述表型^[3-7]。

在相当长时间里, *lin-4* 被认为是内源性的小 RNA 分子的唯一例子。第二个小分子 RNA-*let-7* 由 Ruvkun 实验室发现^[8]。该 miRNA 也为 22nt, 同样在异时性基因路径中发挥作用。与 *lin-4* 相似, *let-7* RNA 识别靶 *lin-41*mRNA 的 3-UTR 序

列, 抑制蛋白 LIN-41 的水平。Lin-4 和 let-7RNA 被命名为小分子时序性 RNA (small temporal RNA) [9-10]。Lin-4 和 let-7 在所有的线虫亚种中呈现出进化上的高度保守性, 其同源物在软体动物、海鞘、小鼠和人类中也能够检测到。这种广泛存在的高度保守性有力的说明了 miRNA 在发育过程中的重要的调节作用。这一观点正已经被线虫的各个器官所鉴定的 miRNA 的功能所证明。Lin-4 现在被认为是 microRNA 分子的奠基性成员 [11-14]。

对 microRNA 指导的基因调节的研究成果集中于寻找更多的 miRNA 和它们所调节的靶分子及其功能等方面的研究。近来的研究表明, miRNA 的功能包括对细胞分化的控制, 细胞死亡, 代谢以及造血干细胞分化等等 [15-18]。

1.2.2 miRNA基因的结构特征

目前在不同物种中发现的绝大多数miRNA 长度一般为22nt, 其中21~23nt 长度的miRNA 占大多数, 约为84%。miRNA 的一个特点是它的前体常形成分子内茎环结构, 而且含有大量的G/C碱基对, 这个前体物经过核酸酶的加工形成成熟的 miRNA [1, 3, 19-21]。

如同线虫的 lin-4 和 let-7 一样, 大部分 miRNA 基因与已注释基因相距较远, 提示他们来自独立的转录单位。然而, 相当多的分子来自 pre-mRNA 中的内含子。这些 miRNA 常与预期的 mRNA 以同一方向排列, 提示这些 miRNA 并没有自己的启动子。这种组织形式为 miRNA 与蛋白的共同表达提供了便利的机制, 同时也解释了 miRNA 与宿主 mRNA 之间的保守关系 [2-4, 22-24]。

然而近来研究表明, 大多数哺乳动物的 miRNA 位于转录单元 (TU) 内。通过即时更新的基因组装配以及 EST 数据库, 人们发现许多以前位于基因间隔区的 miRNA 位于内含子中。有时候由于选择性剪接方式, miRNA 位于同一个 pre-mRNA 的外显子或者是内含子中 [4]。根据 miRNA 基因组位置, miRNA 可以分为下列几种: 1、位于编码蛋白质的转录单元内的内含子中; 2、位于非编码转录单元的内含子中; 3、位于非编码转录单元的外显子中 (见图 1.1)。混合 miRNA 依据剪接模式的不同而成为以上一种基因组组织形式。一些内含子 miRNA 的位置在许多物种中是非常保守的 [3-5, 7-8]。

另外一些 miRNA 基因在基因组中和表达模式中成簇排列, 意味着它们的转录以多顺反子形式的原初转录本进行。在人类基因组中大约 50% 的已知 miRNA 成簇表达, 而且一个簇的 miRNA 经常是彼此相关的。而一个簇中也包含了不相关的 miRNA [1, 13-15]。

目前有研究发现, miRNA 基因可能也来自于进化过程中出现的假基因。在人的 miRNA 基因中, miR-220 和 miR-492 各自位于一个加工的假基因中。这说

明 miRNA 在进化过程中可能出现了重复以及发生了变异,同时也表明了有一部分 miRNA 是来自于假基因中^[25-28]。

几乎所有已经克隆的 miRNA 在相近的动物中保守,如人与小鼠,或秀丽线虫和布氏线虫。许多 miRNA 在动物中具有更为广泛的保守关系。例如,大于三分之一的线虫 miRNA 可以很容易在人的 miRNA 中找到同源物^[3,5-8]。如果比较更远一点的物种,相当多的基因家族的扩大或缩小非常明显,最令人吃惊的例子是 let-7 家族,它在线虫中有四个相同的成员,在人中至少有 15 个,但在果蝇中只有一个^[10,22]。

与植物 miRNA 前体相比,动物 miRNA 前体分子的长度变化较小,大多数在 70-80nt 之间。一部分 miRNA 常以单拷贝形式存在于基因组中。一些 miRNA 的前体分子常以两个或多个拷贝存在于不同染色体或者是同一染色体不同的位置上。某些 miRNA 来自两个或多个仅有微小差异的前体^[2-4,6-7]。

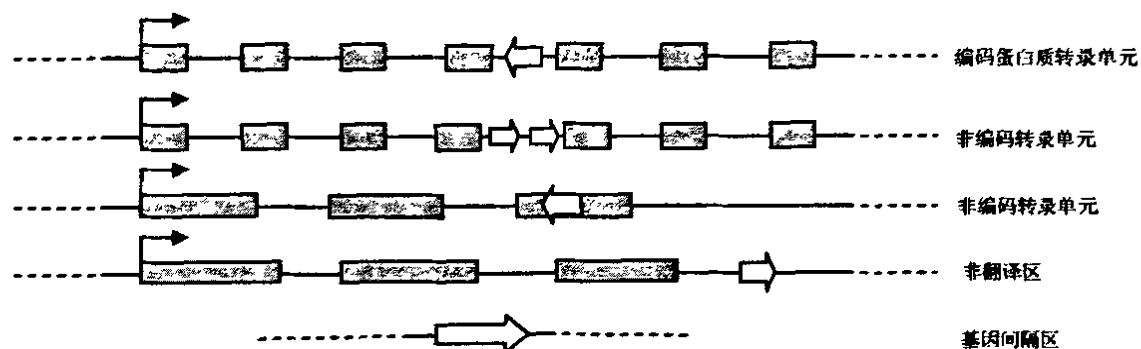


图 1.1 microRNA 的基因组组织形式

Fig 1.1 Genomic organization of micro RNA

1.3 miRNA 的生物合成

1.3.1 miRNA 生物发生的基本概况

成熟的 miRNA 是一种长约 19-25 nt 大小的单链 RNA,而编码 miRNA 的基因则位于基因间区或基因的内含子中。miRNA 的形成需要经历如下阶段:基因的转录、原始转录本(primary miRNA, pri-miRNA)的加工、miRNA 前体分子(miRNA precursor, premiRNA)转运、pre-miRNA 的剪切及 miRNA 链的选择共 5 步,最后形成具有生物学功能的成熟单链 miRNA^[2-4,6-7,12,15,17]。它可以组合进入 RNA 诱导的沉默复合体(RNA-induced silencing complex, RISC)中,进而引起靶 mRNA 的降解或翻译的抑制。而 miRNA 究竟发挥哪种作用,是由 miRNA 与靶 mRNA 之间的互补程度决定的^[29-31]。当二者完全匹配时,引起靶 mRNA 的降解,二者不完全匹配时,

则引起靶mRNA 翻译的抑制。

在miRNA 的生物发生过程中，miRNA 基因的转录是由RNA多聚酶II (pol II) 催化完成的，形成pri-miRNA；然后，pri-miRNA 在细胞核中的一种“微处理器 (microprocessor)” 蛋白质复合体的作用下，产生具有特征性的大约70 nt 大小的发夹状pre-miRNA。生成的pre-miRNA 在核转运受体exportin-5 (Exp-5) 的作用下，转运到细胞浆中。细胞浆中的pre-miRNA 在Dicer 的作用下，剪切产生大约22 nt 大小的双链miRNA (见图1.2) [18-20, 32]。其中只有一条单链可以选择性结合到RISC 中去而成为成熟miRNA。一般而言，成熟的miRNA 多为具有5' 端小突起的那条链。

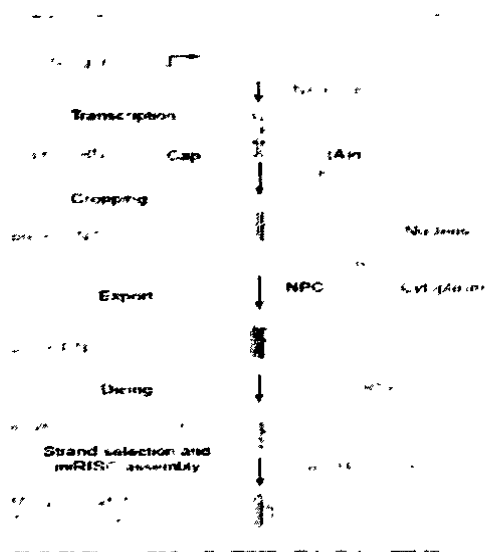


图 1.2 microRNA 的成熟加工过程

Fig 1.2 Maturation and processing of microRNA

1.3.2 miRNA加工的核内过程

miRNA 在细胞核中经历了从最初的miRNA 基因到大约70nt 大小的miRNA 前体阶段。在这个过程中，有多种不同的细胞核内蛋白质的参与，这些蛋白质相互协调作用，保证miRNA 在细胞核内的正常加工。

研究已经发现miRNA 的基因绝大部分定位于基因间区，也有一部分位于基因中的非编码区域，即内含子中。这些miRNA的基因一般由RNA polII进行转录，生成的pri-miRNA 长达几千个碱基，有帽子结构(cap structure)和polyA尾，同时具有一个或几个局部的发夹状结构，以多顺反子形式存在。最近的研究还发现大约50% 的miRNA 基因彼此相邻，簇集存在，带有自己独立的转录启动子。可见，miRNA 的基因也类似于其他的基因，在不同的发育阶段及不同的组织类型中，处在pol II 的精确调控之下^[33-36]。目前，推测由RNA polII 指导的转录可能具有如下的优点：① 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; ② miRNA 和编码蛋白质的基因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^[28, 37]。

从pri-miRNA 到pre-miRNA的加工, 需要细胞核中一种叫做微处理器的复合体的作用。在其作用下, pri-miRNA 被加工成为pre-miRNA, 这一过程在细胞核中完成。细胞核中的微处理器在不同的生物体中, 所包含的成分不尽相同。在线虫、果蝇及哺乳动物体内, 该复合体至少包含有两种蛋白质成分, 分别为Drosha和Pasha (Partner of Drosha)。在人类, Pasha又称为DiGeorge 综合征关键区域基因8 (DiGeorge syndrome critical region gene 8, DGCR8)^[34-36, 38]。

1.3.3 microRNA加工蛋白

Dicer酶是内切核酸酶RNaseIII家族成员, 对它的最早认识是来自对RNA干涉过程中siRNAs的产生的研究, 后来表明它在miRNA 的成熟过程中起着同样重要作用^[31-32]。Dicer 酶包括PAZ 结构域及含有DEXH/DEAH盒的RNA解旋酶结构域。近来通过X射线晶体衍射分析及核磁共振(NMR)获得了PAZ结构域的三维立体结构, 表明它具有异常的OB折叠结构, 其与核酸的亲合性较弱但持久^[35, 37]。研究发现, PAZ结构域具有一条由两个亚结构域所形成的沟, 可能是RNA结合区。其中一个亚结构域具有一个由五条多肽链组成的开放 β 桶芯结构, 两条螺旋位于桶的两个末端, 以及另一条链位于桶的外表面。另一个亚结构域是由连有 α 螺旋的 β 发夹结构组成。因而Dicer酶是通过PAZ结构域与pre-miRNA之间的相互作用而行使切割功能的^[34, 38-40]。

Drosha大约160kDa大小, 全长大约1300个氨基酸, 是RNA酶III (RNase III) 家族成员之一, 包含有一个脯氨酸富集区, 精氨酸-丝氨酸富集区、两个RNA酶III 结构域和C端的一个双链RNA结合结构域(double-stranded RNA-binding domain, dsRBD)^[33-35]。Drosha在线虫、果蝇以及人类中具有保守性。Drosha在果蝇体内形成一种大约500 kDa大小的复合体, 而在人类体内形成的复合体大约650kDa大小。这一复合体即微处理器^[38, 40]。Drosha的主要作用是剪切pri-miRNA形成具有3'端2nt悬垂的pre-miRNA, 此为RNaseIII 作用的主要特征。生成的pre-miRNA的具有3'端2 nt悬垂的末端即为成熟miRNA的一端, 而miRNA另一端是由Dicer在距离已知末端大约22 nt的位点上剪切而成的。可见, Drosha的加工决定着miRNA的真正序列, 其加工的特异性是miRNA的整个生物发生过程中特异性最高的。研究发现, pri-miRNA的三维结构是底物特异性的主要决定因素。剪切位点附近双链的茎部结构和末端的环状结构(>10 nt)也是非常关键的^[38-39]。而且, Drosha复合体可以测量茎部的长度, 一般在距离末端的环状结构约2个螺旋转角

(helical turn) (大约22 nt) 的地方进行剪切。此外, 研究还发现Drosha间接地参与了前核糖体RNA的加工^[40-41]。

微处理器的另一个关键成分是Pasha。Pasha为dsRNA结合蛋白, 参与Drosha对底物的识别。抑制该基因的表达可以使pri-miRNA在细胞核中的堆积增加。DGCR8是第一个被鉴定出来的参与体内miRNA加工过程的Pasha, 进化上具有保守性, 约120 kDa 大小, 包含有两个dsRNA结合结构域和一个已知的可以与脯氨酸富集的多肽相互作用的WW结构域, DGCR8的WW结构域很可能是与Drosha的N端脯氨酸富集区发生相互作用的部位^[37-38]。DGCR8是染色体的22q11.2的预测的30个基因中的一个, 在人类最常见的遗传综合征——DiGeorge综合征中非等位性缺失。所以, 有人推测DiGeorge综合征的发病可能与人类miRNA的加工和表达过程的紊乱有关^[39-41]。

1.3.4 微处理器对pri-miRNA的剪切机制

Drosha对pri-miRNA转录本进行特异性识别的机制尚未完全明了。研究发现Drosha对发夹状茎环结构有一定的亲和力, 而且Drosha也通过一些可以特异识别pri-miRNA的中介因子而招募pri-miRNA到其剪切位点上进行剪切。Pasha即作为一种已经发现的有助于Drosha进行底物特异性识别的蛋白质。Han等提出了关于pri-miRNA剪切机制的“一个加工中心”的模型^[40]。该模型认为Drosha与Pasha形成异二聚体, 对pri-miRNA的茎环状结构部分进行剪切。其中, Drosha结合在远离末端环的双链RNA上, 两个RNase III结构域(RIIIDa和RIIIDb)形成分子内二聚体, 构成一个加工中心, 加工中心中的两个活性位点紧密相邻, 在距离末端环约两个螺旋转角的地方, 在相反方向上剪切RNA链上邻近的磷酸二酯键^[41]。RIIIDa一侧的催化位点催化3'链的剪切, 而RIIIDb一侧的催化位点则催化5'链的剪切。活性中心位点由对应于超耐热菌*Aquifex aeolicus*的RNase III中谷氨酸E40、天冬氨酸D44、D107和E110。而Pasha则结合在靠近末端环状结构的部分, 有助于pri-miRNA的识别及微处理器在底物分子上的定向^[42]。

1.3.5 pre-miRNA的细胞核输出

Drosha对pri-miRNA加工完毕之后, 产生的pre-miRNA需要输出到细胞浆中, 才可以在Dicer作用下生成大约22 nt大小的miRNA双链。这一输出过程需要pre-miRNA在核孔转运受体的作用下穿过位于核膜上的核孔复合体, 而进入细胞浆中^[35,37-38]。现在已经确定进行pre-miRNA输出的主要核孔转运受体为Exp-5。除此之外, Exp-5可以少量地输出tRNA和腺病毒RNAVA。Exp-5对底物进行识别的

特异性是由底物所具有的特征性“微小螺旋”所决定的,与序列无关。该螺旋中包含有大于14bp(碱基对)的茎部和3-8 nt的3'端悬垂,而pre-miRNA所具有的类似的茎部结构和大约2 nt的3'端悬垂使得Exp-5对其的亲合力远远高于其他两种底物,而且pre-miRNA在细胞中的含量相当高,所以很可能pre-miRNA就是Exp-5进行输出的主要底物。Exp5与Ran-GTP以及pre-miRNA形成异三聚体,通过核孔到达胞浆后,Ran-GTP转变为Ran-GDP,释放出pre-miRNA。此外,还发现pre-miRNA与Exp-5结合之后还可以稳定pre-miRNA^[39-40]。

1.4 microRNA的作用机制

miRNA能够通过两种沉默机制中的一种指导RISC下调目的基因的表达:mRNA的切割或翻译抑制。如果mRNA与miRNA具有完美的互补性,miRNA就指导mRNA特异性切割,或者如果两者没有足够的互补性,则miRNA就指导翻译抑制(见图1.3)。

1.4.1 miRNA指导靶mRNA的切割

大多数植物miRNA与靶序列的开放阅读框(ORF)完全匹配,因而在植物中,miRNA主要进入RNA干涉途径以降解靶分子。植物中的RISC是由ARGONAUTE(AGO)蛋白,miRNA或siRNA及相关蛋白组成^[31-34]。AGO蛋白具有与Dicer酶一样的PAZ结构域以及PIWI结构域。PIWI结构域具有与RNaseH类似的结构。据报道,miRNA的3'末端位于PAZ结构域的亲水性结构域内,而靶分子位于PIWI结构域中,两者识别互补之后,RNA酶水解作用于靶分子中与miRNA的第11或第12碱基所对应的残基的磷酸键,随后降解靶分子的5'端序列。因而当miRNA指导切割时,切点的位置是由相关的miRNA的残基所决定,即从配对区的中部向mRNA的5'方向切割^[37]。在植物中,两者完全匹配在大多数情况下是引起切割,但是有时并非如此,miR-172与miR-172样基因与靶分子完全匹配却产生翻译抑制。研究人员发现,经过切割之后,靶分子上与miRNA匹配位点相邻的序列会产生内源性的siRNA,以增强沉默效应。MiRNA切割之后,miRNA仍然是完整的,而且能够指导另外的mRNA的识别和切割^[38-40]。

1.4.2 翻译抑制

通常动物miRNA与靶mRNA的3'-UTR部分序列反向互补结合。与RNAi不同的是,靶mRNA的稳定性不受影响,而其翻译起始后的表达受到影响。这种新的调控机制首次在线虫的lin-14mRNA得到证实,lin-14mRNA受lin-4 miRNA的调控^[38]。

MiRNA可能抑制翻译的延伸或终止,或降解核糖体新合成的新生肽链。然而简单地影响翻译延伸是不可能的,因为未受到抑制的多核糖体的形态与由miRNA抑制的靶mRNA的多核糖体的形态是相同的。当Lin-14蛋白水平下降时,难以区分处于线虫幼虫第一阶段和幼虫后期阶段的lin-14 mRNA的多核糖体外型。具体的调控机制目前还不是很清楚,有待于科学家进一步研究^[39]。多数miRNA是以miRNP形式与靶mRNA结合并使之抑制。最近研究人员发现,处于不稳定状态的mRNA的3'-UTR中的富含AU元件(AREs)能够指导mRNA的降解。参与miRNA加工和功能行使的蛋白,即Dicer-1, Ago1和Ago2是 α 肿瘤坏死因子mRNA快速降解所必需的,该mRNA含有AREs元件。人的miR-16具有与AREs互补的UAAAUAUU序列,它使得具有AREs的RNA易于降解^[37,40]。

1.4.3 靶分子的识别

线虫lin-14的3'-UTR具有与lin-4miRNA的5'区域互补的核心区域,这说明多细胞动物miRNA5'端互补的重要性^[23]。这可从以下事实得到支持:(1)几种无脊椎动物的miRNA的2~8残基与靶mRNA 3'UTR是完全互补的。(2)在第一个确认的无脊椎动物的miRNA与靶分子的互补位点之间,mRNA与miRNA 2~8残基配对的部分在其它种类的同源信息中是完全保守的。一个邻近的至少六个碱基对的螺旋几乎总是出现在这个区域。(3)miRNA的2~8残基在同源的后生动物miRNA中是保守的。(4)预测哺乳动物miRNA的靶分子时,要求跨越miRNA 2~8残基的七聚体具有完全的碱基配对,比要求miRNA的其它任何一个七聚体的配对更有效率^[26-27]。如果不考虑互补位点其它区域的互补程度,与核心区域的错配就可抑制起始时对靶分子的识别,因而阻止了切割或翻译抑制。如果同时允许miRNA的其余部分有足够的配对,将产生切割作用。然而仅由一些侧翼配对来补充核心配对,再与RISC共同作用,足以介导翻译抑制^[31-32,38]。

1.4.4 快速脱腺苷酸化

Wu等人以miR-125b和let-7两种代表性的miRNA为研究对象,以启动子-荧光素酶或 β 球蛋白为报告系统发现它们能够促进mRNA聚腺苷酸尾巴(polyA tail)的去除^[38-39]。研究人员用3'组蛋白茎-环结构取代聚腺苷酸尾巴,结果发现不但可以消除miR-125b对mRNA含量的影响,还可以降低对蛋白质合成的作用。由此可知,miRNA通过降低翻译效率和聚腺苷酸化mRNA的浓度来抑制基因表达远比阻遏翻译强而全面。而且,不像翻译阻遏那样导致mRNA的解体是不可逆转^[40-42]。

1.4.5 miRNA与siRNA的异同

由于miRNA与内源性的siRNA的加工成熟机制大致相同,都需要Dicer的参与,而且功能相似,都具有沉默效应,所以难以通过化学组成或作用机制而将其区分。但是从它们的来源、保守性以及作用的靶分子来看,它们具有如下不同之处:(1) miRNA来自于基因组DNA的基因间隔区或者是断裂基因的内含子中,而siRNA一般来自于mRNA、转座元、病毒或者异染色质DNA。(2) miRNA从形成局部茎环结构的转录物加工而来,而siRNA是通过切割长的双分子RNA二聚体或延长的发夹而产生的。(3)每个miRNA茎环前体分子产生一个miRNA:siRNA二聚体,而每一个siRNA前体分子产生大量的siRNA二聚体。(4) miRNA在亲缘关系比较近的物种间是比较保守的,而内源性的siRNA序列变化较大^[39-41]。

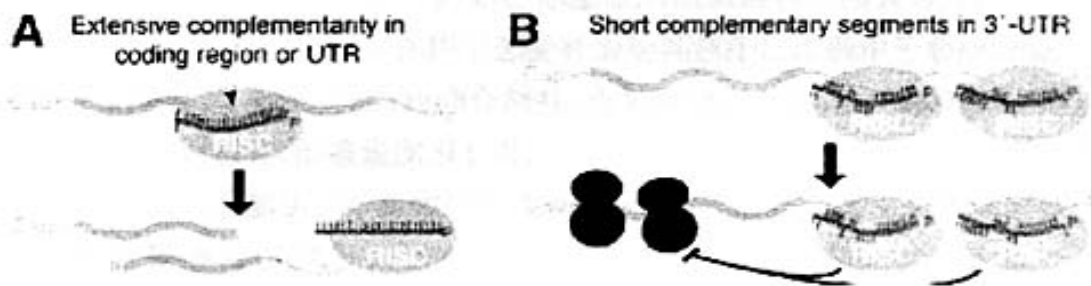


图 1.3 microRNA 的作用机制
Fig 1.3 mechanism of microRNA function

1.5 microRNA 的生物学功能

miRNA 是通过与靶mRNA 的3' -UTR 碱基配对的方式来执行对靶mRNA 的转录翻译抑制的功能。Lin-24 和let-27 在线虫发育中起着时序调控的作用。这两个基因或它们的靶基因发生突变会造成线虫发育的形态突变。

1.5.1 miRNA 在动物中的作用

在动物中,miRNA与靶mRNA进行不完全配对以执行转录后翻译抑制的作用。但是如果动物miRNA与靶序列完全互补,则进入RNA干扰路径指导靶序列降解。近来,在筛选果蝇生长缺陷突变株的过程中发现了Bantam基因座。该基因座编码了果蝇幼虫抑制程序性死亡和促进细胞增殖的miRNA。Bantam miRNA大量表达并抑制Hid mRNA的翻译。Hid是程序性死亡的主要激活因子。Bantam类似于线虫的mir-80/82,这表明mir-80家族能控制线虫的细胞死亡和增殖^[43-45]。动物中的miRNA 还能影响细胞分化。MiR-23与Hes1 基因具有较完美的互补性(大约77%)。Hes1是一种基

本的螺旋-环-螺旋结构,它只在未分化细胞中表达。将人工合成的miR-23加到未分化的NT2神经细胞中,可以发现细胞内的Hes1水平大幅下降。但是将人工合成的miR-23突变体加入到未分化NT2细胞中, Hes1水平保持不变,且其含量与将野生型miR-23加入到已分化的NT2细胞中是一样的。这表明miR-23在翻译水平上抑制Hes1基因的表达,影响细胞分化^[46]。最近Kuang AL 等人发现,在小鼠胚胎干细胞(ES)中编码Dicer酶的基因dicer1缺失后,ES细胞失去了分化功能,同时染色体着丝粒区域的重复序列的沉默效应以及小片段dsRNA的表达量显著降低。这说明miRNA参与了ES细胞的分化,着丝粒区域异染色质的形成^[47]。研究表明,小鼠miR-375的表达会抑制葡萄糖诱导的胰岛素分泌,但是并不改变葡萄糖代谢或细胞内Ca²⁺信号强度。通过RNAi将miR-375的靶分子Mtpn沉默,发现产生的效应与miR-375的作用效果一致^[48]。

有研究以人的一种结缔组织细胞脂肪细胞(adipocyte)为实验对象,从86个候选miRNA分子中筛选出了对其分化有促进作用的mir-143分子,同时运用生物信息学方法预测了mir-143的作用靶点,并结合Western印迹技术,检测了它的预测靶点ERK5 /BMK1蛋白(一种参与细胞增殖和分化的蛋白)的表达情况,发现细胞分化后该蛋白的表达呈上调趋势,但是鉴于miRNA的调节有可能是多位点进行的,因此还不能确定mir-143对该结缔组织细胞分化的影响是否主要是通过调控ERK5/BMK1蛋白水平来实现的^[48-50]。

1.5.2 miRNA在植物中的作用

在植物中,由于miRNA与mRNA互补程度高,故可用来预测植物miRNA的靶位点。有实验证据表明,miRNA能直接指导靶mRNA的切割。植物miRNA大多数是调控发育与细胞分化的基因。MiR-165/166的靶基因PHB与PHV中的序列发生突变后则造成功能获得性突变,phb mRNA表达的组织范围相对于野生型phb mRNA扩大了。最近研究人员发现,植物中的miRNA不仅可以调控细胞周期及细胞分化,还对环境应答和激素应答有一定的作用^[51]。MiR-167和miR-160调节编码植物激素应答因子基因的表达。在拟南芥中,TIR1蛋白因子通过与E3泛蛋白连接酶SCF复合物的吲哚乙酸结合而参与胚轴分化、横向根系形成等生理过程,而MiR-393调控该蛋白的表达。大部分植物miRNA类似于siRNA,介导靶mRNA降解,同时有些miRNA也可能具有转录后翻译抑制作用。利用拟南芥突变体的研究发现拟南芥的一些miRNA序列和转座子有关。同时RNAi对拟南芥DNA的甲基化也有一定的影响,说明在拟南芥中也存在类似于线虫的miRNA影响转座子功能的作用机制^[52-53]。

1.5.3 与人类疾病相关的miRNA

目前还发现一些miRNA分子的表达直接和一些人类疾病相关,尤其是一些神经系统疾病和癌症,如在对缺失SMN(survival of motoneuron)基因而引起杜阿二氏肌萎缩(spinal-muscular atrophy, SMA)的病理研究中发现,组成SMN复合体的两个蛋白质Germin3和Germin4也出现在包含多种miRNA分子的蛋白复合物中,由此推测可能有miRNA参与了这一疾病的发生。关于miRNA与癌症之间的相关性也有报道,如慢性淋巴细胞白血病(chronic lymphocytic leukemia, CLL),在染色体13q14部位的缺失发生率很高。研究发现这一区域编码mir-15以及mir-16分子,而它们在正常人B淋巴细胞中是高表达的,运用生物信息学手段分别对mir-15及mir-16进行了靶点预测,然后通过Western印迹检测发现其预测靶点RARS蛋白的表达水平是与miRNA分子的表达水平相关的,mir-15和mir-16能抑制该蛋白的表达^[54-55]。还发现mir-143和mir-145在结肠癌患者中的表达水平是下调的;mir-155在儿童淋巴瘤中的表达是上调的。Bino等运用生物信息学手段对部分与肿瘤相关miRNA分子进行了靶点预测发现:①miR-15a的靶点可能是CNOT7,它主要表达在结肠癌和主要的肿瘤组织中;②miR-15b的靶点可能是LASS2,这是一种肿瘤转移抑制因子;③miR-143的靶点可能是ING4,它是肿瘤抑制因子p33ING1b的同源体;④miR-155的靶点可能是Gab1,认为它与细胞的增殖有关;⑤miR-145的靶点可能是COL3A1,是在肿瘤细胞中表达上调的一种基因^[56]。最近又发现正常人肺组织中let-7分子的表达水平明显高于肺肿瘤患者中let-7的表达水平,而且相关的癌基因RAS蛋白的表达与之相反,根据生物信息学认定在癌基因RAS mRNA的3'端非编码区含有多个let-7分子的结合位点,从而认为let-7分子可能参与了人肿瘤的调控机制^[57-58]。

1.6 miRNA 的命名和界定

1.6.1 miRNA的界定

最初鉴定新的microRNA基因时主要是采取构建cDNA文库的方法,但是该方法受到了其它一些小分子RNA的影响,以至于产生了假阳性。因而为了将miRNA与siRNA及其它非编码小分子RNA或mRNA片段区分开,Ambros等从RNA的表达和生成的角度提出了miRNA判断方法^[1-2,5]。MiRNA的表达判断标准为:1.采用Northern杂交方法能够检测到~22nt的RNA转录本。2.从小分子RNA构建的cDNA文库中能获得~22nt的序列,这些序列需与基因组序列精确匹配。MiRNA的生成的判断标准

为:1. 具有发夹结构前体,且其中一条臂含有miRNA序列,发夹具有最低自由能,结构中不含大的内环或泡,尤其是大的不对称的泡。2. miRNA序列及其二级结构前体具有进化上的保守性。3. 随着Dicer 酶功能降低,前体物不断积累。一般要综合上述两方面的证据才能判定一个新的miRNA。

1.6.2 miRNA的命名

在对miRNA命名时,一般是在miR的后面加一个阿拉伯数字(如miR-1, miR-2等)。编码miRNA的基因也是用这三个字母表示,大写,中间采用连字符连接,根据生物学惯例使用斜体字(如线虫和果蝇中的*mir-1*,水稻和拟南芥的*MIR156*)。不同物种中,相同的miRNA采用相同的数字。如果同一个基因具有不同的转录本,那么还需在其后面加一小写的字母或数字,如果蝇的miRNA基因簇中的miR-6转录出3个miRNA,则分别记为miR-6-1, miR-6-2, miR-6-3^[2-3,5]。

1.7 microRNA 的研究方法

1.7.1 microRNA的实验RNA组学方法

目前,对microRNA表达的鉴定依然主要依靠以杂交和PCR为主的相关方法,对其功能线索的获得,则主要依赖于在特定细胞或动物模型内外源性抑制或表达相应的microRNA基因,再通过报告基因系统验证后检测特定细胞或动物模型发育过程中的生化或分子变化来实现,并同时利用报告基因系统对microRNA的靶基因进行验证。

1.7.1.1 Northern杂交

Northern杂交为microRNA检测提供了简便而可靠的方法。根据条带分子量的大小以及密度确定microRNA分子是否表达,表达量如何。该技术可以检测至少5ug总RNA中miRNA的水平。Sempere等通过Northern杂交的方法对已知的119个miRNA基因的表达做了检测,发现有一些是组织特异性表达,有一些只是在不同组织中表达量不一样,而脑组织miRNA的表达方式说明这些与脑神经元分化相关^[59]。

1.7.1.2 原位杂交

原位杂交技术用来检测miRNA,更直观的展示出miRNA表达方式,是了解miRNA时空表达谱更方便的方法。可以根据需要选择探针的标记方

式。特异性的原位杂交可以检测细胞、胚胎、组织切片，福尔马林固定石蜡包埋切片中的miRNA水平^[60]。

1.7.1.3 miRNA表达谱芯片

这种技术也是用探针标记检测固相支持物上的目标分子。基因芯片高通量的特点，使得miRNA芯片非常适合检测miRNA表达谱。通过精心设计芯片上的miRNA基因对照序列，可以精确的分析出样品中所有感兴趣的miRNA表达水平。但是miRNA表达谱芯片只能分析已知的miRNA表达水平，不能分析未知的miRNA^[61]。

1.7.1.4 基于核酶保护分析技术

microRNA检测还可以采用一种基于核酶保护分析技术。这种基于液相杂交的新方法操作简便快速，而且灵敏度极高，可以半定量检测少至50ng总RNA模板中的小RNA，灵敏度是Northern杂交的100倍，而且还可以在同一个样品汇总同时检测多个小RNA和长的RNA模板^[62]。

1.7.2 microRNA的计算机RNA组学方法

生物信息学方法是依据在不同的物种中，其成熟的miRNA 具有较大的序列同源性以及前体的茎环结构具有相当大的保守性这一特征在基因组数据库中搜索新的miRNA 基因。该方法根据比较基因组学原理并结合生物信息软件在已测序基因组中进行搜索比对，根据同源性的 高低再进行RNA 二级结构预测，将符合条件的候选miRNA 与已经通过实验鉴定的miRNA 分子进行比较分析，最终确定该物种miRNA 的分布及数量^[63-64]。目前国际上较为普遍使用的两个计算机分析工具是miRseeker和MiRscan，前者已用于果蝇及昆虫基因组候选基因的系统分析，后者则用于线虫和脊椎动物候选基因的分析。这两个工具已经成功鉴定出了大量的miRNA基因并通过了实验证实。由于miRseeker和MiRscan的高灵敏性，它们已用于人类miRNA基因的寻找。由于该方法只能用于已完成基因组测序的物种，而那些未完成测序的物种就无能为力，而且由于miRNA前体长度的可变性，故计算机方法寻找新基因具有一定的遗漏性。所以目前大多数实验室将计算机分析与实验方法结合使用，使得miRNA的数量呈几何级数增长^[65-66]。

1.8 本论文的研究目的与意义

非编码 RNA 在任何生物体内都广泛存在，它们在基因表达调控的各个环节

都发挥着重要的生物功能，因而，分离和鉴定新的非编码 RNA 并阐明其功能是当前 RNA 组学研究的国际前沿。研究表明几乎所有的基因都含有能够与 microRNA 的一些部分相配对的短序列。这些潜在的 microRNA 靶位点在进化过程中被保存下来，这意味着它们在各种物种间的相同基因上存在相同的位点。microRNA 广泛存在于真核生物细胞中，富集于核仁，具有代谢稳定性。它们与特定的蛋白质结合形成核糖核蛋白复合物，是一类在核糖体的生物合成中起重要作用的核酸调控分子。一系列新的 microRNA 的发现及其功能的确定，是九十年代分子生物学取得的重大进展之一。microRNA 的发现为阐明新的表达调控机制提供了新的研究平台。该领域的研究引起人们的广泛关注，如 Smith&Steitz 在 Nature 杂志中所述：“microRNA 组学”引起了分子生物学领域的极大震动。

犬类基因组核苷酸差异水平与人基因组差异水平非常类似，而且犬类具有表型多样性以及进化时间上的短暂性，这为研究人类相关疾病提供了很好的平台，因而鉴定其特有的 microRNA 有着非常重要的现实和理论意义^[67-68, 70]。

因此，本论文采用生物信息学的方法在狗基因组中寻找鉴定 microRNA 分子以及所作用的靶序列，同时系统分析 microRNA 的基因组织形式。

第 2 章 材料与amp;方法

2.1 材料

2.1.1 使用的数据库

- 1) GenBank 的非冗余数据库 nr 和表达序列标签数据库 dbEST
(<http://www.ncbi.nlm.nih.gov>):
- 2) 微小 RNA 数据库 microRNA Database
(<http://microrna.sanger.ac.uk/cgi-bin/sequences/browse.pl>):
- 3) 犬基因组数据库 UCSC database (<http://genome.ucsc.edu/>):
- 4) Ensembl 数据库 Ensembl database (<http://www.ensembl.org/index.html>)
- 5) 哺乳动物非翻译区数据库 UTRResource
(<http://www.ba.itb.cnr.it/UTR/UTRHome.html>):
- 6) Geneontology 数据库 GO database (<http://www.geneontology.org/>):
- 7) NCBI 蛋白质数据库
(<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Protein>)。

2.1.2 供试犬类品种

采用国产肉用狗—松狮狗 (chow chow)

2.1.3 主要仪器和设备

9600 型 PCR 仪: PE 公司
PTC-225 型 PCR 仪: MJ 公司
ZEL BOC XR 凝胶成像系统: BIORAD 公司
Mill-Q Biocel 超纯水器: 美国 Millipore 公司
TGL-16G 型冷冻高速离心机: 上海安亭科学仪器厂
DY-A 型稳压稳流电泳仪: 北京市六一仪器厂
DYYIII 31C 型琼脂糖水平电泳槽: 北京市六一仪器厂
MS1 型涡漩混合器: 德国 IKA 公司
SW-CJ-2FD 超净工作台: 苏宁集团安泰公司

2.2 方法

2.2.1 microRNA的计算机分析

2.2.1.1 microRNA同源基因的搜索

- 1) 以已鉴定的人 microRNA 基因成熟序列为种子序列, 应用 Blat 程序 (<http://genome.ucsc.edu/cgi-bin/hgBlat>)^[71]分析狗基因组数据库, 获得可能的 microRNA 基因;
- 2) 在 Blat 比对未能找到相应的 microRNA 基因的情况下, 分析该基因的前体序列在该物种中是否具有同源, 以进一步确定 microRNA 存在与否;
- 3) 应用在线的 mfold 程序 (<http://www.bioinfo.rpi.edu/applications/mfold/old/rna/>)^[72]分析比对得到的序列, 判断其是否具有 microRNA 的基本结构特征;
- 4) 将可信度高的 microRNA 候选基因作为种子序列, 再次进行 Blastn (<http://www.ncbi.nlm.nih.gov/blast>) 比对, 确定其在基因组的位置。

2.2.1.2 新的microRNA基因的预测

- 1) 从 UCSC 数据库中下载人和狗的基因组比对数据, 将同源性高于 80%的序列放入 RNAz 程序中进行 RNA 分子预测。
- 2) 将预测为 RNA 分子的序列再次放入 mfold 程序中进一步确认。
- 3) 将可信度高的 microRNA 候选基因作为种子序列, 再次进行 Blastn 比对, 确定其在基因组的位置。

2.2.1.3 microRNA的基因组织分析

- 1) McPromoter MM:II 程序^[73] (<http://genes.mit.edu/McPromoter.html>)、Promoter Prediction 2.0^[74] (<http://www.cbs.dtu.dk/services/Promoter/>)、NNPP^[75] (neural network promoter prediction) (http://www.fruitfly.org/seq_tools/promoter.html)用于启动子序列分析预测;

2.2.1.4 microRNA的进化分析

- 1) 以已鉴定的人 microRNA 基因成熟序列为种子序列, 在各个动物物种中寻找其同源物;

- 2) 将相应的 microRNA 同源序列用 Clustal W 进行比对;
- 3) 采用 SplitTree^[75]进行系统进化树分析。

2.2.1.5 microRNA基因的靶序列预测

- 1) 从哺乳动物非翻译区数据库下载狗的 3' 非翻译区 (3' -UTR) 数据;
- 2) 将 3' -UTR 和狗的 microRNA 数据放入 miRanda 程序^[76]进行靶序列的预测;
- 3) 利用 GO 数据库将预测的 microRNA 靶序列进行功能分类。

2.2.2 狗肝中总RNA的提取 (热酚法)

2.2.2.1 试剂和材料

- 试剂 :
- 1) TES 溶液配制:
10 m mol/L Tris-HCl pH 7.5
10 m mol/L EDTA
0.5% SDS
 - 2) 2M NaAC, pH 4.0
 - 3) 水饱和重蒸酚
 - 4) 氯仿: 异戊醇 (49: 1)
 - 5) 3M NaAC , pH 5.2
 - 6) 异丙醇
 - 7) 无水乙醇
 - 8) 70%乙醇

材料 : 新鲜狗肝 (将新鲜狗肝剪成小块放入液氮中保存备用)

2.2.2.2 实验步骤

- 1) 将新鲜的狗肝放入液氮中冷冻, 取出一小块肝脏组织置于灭菌了的冰凉研钵中, 迅速研磨, 并不时地加入液氮以保证肝脏组织处于冷冻状态。取帕在两支 50ml 灭菌了的离心管称取重量, 当彻底研磨成粉末状后, 转入离心管中, 再称取重量, 于是研磨的肝脏组织重量分别为 4.4 g 和 4.3g。
- 2) 向两支离心管中分别加入 9ml TES 溶液和 9ml 酸性酚, 用涡旋震荡器以高速震荡 10 秒后, 于 65° C 水浴中温浴一个小时, 并不时用涡旋震荡器短震荡。
- 3) 冰浴放置 5 分钟。4° C 12000r/min 离心 10 分钟。

- 4) 用移液枪吸取上清至两支 50ml 灭菌了的冰凉离心管中, 每管 12.5ml, 向管中加入等体积的酸性酚, 混匀, 4° C 12000r/min 离心 15 分钟。
- 5) 吸取上清至两支 50ml 灭菌了的冰凉离心管中, 每管 11.5ml, 向管中加入等体积的酸性酚, 混匀, 4° C 12000r/min 离心 15 分钟。
- 6) 吸取上清至一支 50ml 灭菌了的冰凉离心管中, 体积为 20ml, 向管中加入等体积的酸性酚, 混匀, 4° C 12000r/min 离心 10 分钟。
- 7) 吸取上清至灭菌了的 1.5ml 小离心管, 每管 700 μ l, 共 24 管。向各管中加入等体积的酸性酚, 混匀, 4° C 12000r/min 离心 10 分钟。
- 8) 吸取上清至灭菌了的 1.5ml 小离心管, 每管 700 μ l, 共 19 管。向各管中加入等体积的氯仿, 混匀, 4° C 12000r/min 离心 10 分钟。
- 9) 吸取上清至灭菌了的 1.5ml 小离心管, 每管 400 μ l, 共 26 管。向各管中加入 40 μ l 醋酸钠 (PH 5.3) 以及 1000 μ l 冰冷的无水乙醇。冰上放置 90 分钟。
- 10) 4° C 12000r/min 离心 10 分钟。
- 11) 弃去上清, 沉淀在冰冷的 70%乙醇中快速震荡洗涤。4° C 12000r/min 离心 5 分钟。
- 12) 沉淀用无水乙醇洗涤, 晾干。
- 13) 取一支加入 20 μ l DEPC 处理的 3 天水溶解, 用于电泳; 其余的 25 支加入 400 μ l 无水乙醇于 -20° C 保存备用。

2.2.2.3 甲醛变性琼脂糖凝胶电泳检测RNA

试剂: (1) 10 \times MOPS:

MOPS	4.18g
3M NaAC (pH 5.2)	1.66ml
灭菌 0.5M EDTA (pH 8.0)	2ml
加灭菌 3d H ₂ O 至	100ml

将 MOPS 加灭菌 3d H₂O 至 80ml, 调 pH 7.0, 再与 NaAC, EDTA 混合过滤除菌, 加灭菌 3d H₂O 定容至 100ml, 避光保存。

(2) Loading Buffer :

甘油	50%
EDTA (pH 8.0)	1mM
二甲苯青	5%
溴酚蓝	0.25%

凝胶制备:

称取琼脂糖 0.6g 于一洁净的小烧杯, 加入 43.5ml 灭菌 3d H₂O, 用玻璃棒搅拌后放入微波炉内煮沸溶解, 直至溶液透明, 室温冷却至 60° C。向溶液中加入 10×MOPS 5ml, 甲醛 1.5ml, 混匀后倒胶, 注意不能有气泡。

电泳缓冲液制备: 取 30 ml 10×MOPS 加入到 270 ml 灭菌 3d H₂O 中, 混匀。

制备电泳用 RNA 样品:

RNA	5 μl
10×MOPS	1.25 μl
甲醛 (37%)	2.25 μl
Formamide (N, N-二甲基甲酰胺)	6.25 μl
EB(0.5 μg/μl)	0.5 μl

混匀后, 55° C 变性 15 分钟。加入 5 μl 6×溴酚蓝上样缓冲液, 混匀上样, 电泳检测。电泳条件: 80V, 约 30 分钟。

注: MOPS 为 3-(N-玛琳代)丙磺酸

2.2.3 基因组DNA的提取 (玻璃珠法)

2.2.3.1 试剂

- 1) 裂解缓冲液 0.1mol/L Tris-HCl (pH 8.0)、5mmol/L EDTA、1%SDS
- 2) 2.5mol/L NaCl
- 3) 氯仿: 异戊醇=24: 1
- 4) 95%乙醇
- 5) 70%乙醇
- 6) 5mol/L 醋酸铵
- 7) TE 10mmol/L EDTA (pH 8.0)

2.2.3.2 实验步骤

- 1) 将新鲜的狗肝放入液氮中冷冻, 取出一小块肝脏组织置于灭菌了的冰凉研钵中, 迅速研磨, 并不时地加入液氮以保证肝脏组织处于冷冻状态。取一支 50ml 灭菌了的离心管称取重量, 当彻底研磨成粉末状后, 转入离心管中, 再称取重量, 于是研磨的肝脏组织重量为 2.2g。
- 2) 向离心管中加入 5ml 裂解缓冲液, 灭菌石英砂和 500 μl 的 2.5mol/L NaCl。

- 3) 用涡旋震荡器以高速震荡 2 分钟。
- 4) 以 12000r/min 离心 5 分钟。
- 5) 用移液枪吸取上清 4ml 至另一支 50ml 灭菌了的离心管中，加入 4ml Tris 饱和苯酚，混匀，12000r/min 离心 5 分钟。
- 6) 取六支灭菌了的 1.5ml 小离心管，吸取上清至管中，其中一管 700 μ l，五管 600 μ l。向各管中加入与上清等体积的 Tris 饱和苯酚，混匀，12000r/min 离心 5 分钟。
- 7) 再取六支灭菌了的 1.5ml 小离心管，吸取上清至管中，每管 500 μ l。向各管中加入 500 μ l Tris 饱和苯酚，混匀，12000r/min 离心 5 分钟。
- 8) 取四支灭菌了的 1.5ml 小离心管，吸取上清至管中，每管 550 μ l。向各管中加入 550 μ l 氯仿：异戊醇（24：1），混匀，12000r/min 离心 5 分钟。
- 9) 取四支灭菌了的 1.5ml 小离心管，吸取上清至管中，每管 400 μ l。向各管中加入 800 μ l 无水乙醇，冰上放置 40 分钟。
- 10) 室温下 12000r/min 离心 8 分钟沉淀 DNA。
- 11) 弃去上清，用 70%乙醇洗涤沉淀两次。
- 12) 向其中的两支 DNA 加入 30 μ l TE 缓冲液溶解沉淀，并于 -20 $^{\circ}$ C 保存。向另外两支 DNA 中加入 300 μ l TE 缓冲液溶解沉淀，并加入 25 μ l RNase，置于 37 $^{\circ}$ C 水浴中 5 分钟。
- 13) 向两管中加入 300 μ l 5 mol/L 醋酸铵和 600 μ l Tris 饱和苯酚，混匀，12000r/min 离心 5 分钟。
- 14) 取两支灭菌了的 1.5ml 小离心管，吸取上清至管中，每管 300 μ l。向各管中加入等体积的氯仿：异戊醇（24：1），混匀，12000r/min 离心 5 分钟。
- 15) 再取两支灭菌了的 1.5ml 小离心管，吸取上清至管中，每管 250 μ l。向管中加入 500 μ l 无水乙醇，冰上放置 40 分钟。
- 16) 室温下 12000r/min 离心 8 分钟沉淀 DNA。
- 17) 弃去上清，用 70%乙醇洗涤沉淀两次。
- 18) 晾干后，每管加入 20 μ l TE 缓冲液溶解沉淀。
- 19) 制备 1%的胶来检测 DNA 样品，称取 0.2g 琼脂糖于一洁净的小烧杯，加入 20ml 灭菌 3d H₂O，用玻璃棒搅拌后放入微波炉内煮沸溶解，直至溶液透明，室温冷却至 60 $^{\circ}$ C。向溶液中加入 400 μ l 50 \times TAE 溶液，混匀后倒胶。注意不能有气泡。
- 20) 制备电泳用 DNA 样品：3 μ l DNA 样品+2 μ l 6 \times 溴酚兰上样缓冲液，混匀上样，电泳检测。电泳条件：80V，约 30 分钟。

2.2.4 RNA中的DNA的酶解

- 1) 取出一支RNA, 弃去乙醇, 晾干后加入 20 μ l DEPC 处理的 3d 水, 使RNA溶解。瞬时离心。取上次电泳剩余RNA样品 10 μ l 加入其中, 于是总体积为 30 μ l。
- 2) 向管中加入以下试剂:

表 2.1 DNA 酶解

试 剂	用 量
总 RNA	30 μ l (约 50 μ g)
10 \times DNase I Buffer	10 μ l
DNase I (5U/ μ l)	10 μ l
DEPC 处理水	50 μ l

- 3) 共 100 μ l
- 4) 将上述试剂混匀, 37 $^{\circ}$ C 保温 40min。
- 5) 瞬时离心, 加入 400 μ l DEPC 处理水, 混匀。
- 6) 加入 500 μ l 水饱和酚, 混匀, 4 $^{\circ}$ C 12000 r/min 离心 10min。
- 7) 吸取上清至一支 1.5 ml 灭菌了的冰凉离心管中, 体积为 450 μ l, 加入等体积水饱和酚, 混匀, 4 $^{\circ}$ C 12000 r/min 离心 10min。
- 8) 吸取上清至一支 1.5 ml 灭菌了的冰凉离心管中, 体积为 400 μ l, 加入等体积水饱和酚, 混匀, 4 $^{\circ}$ C 12000 r/min 离心 10min。
- 9) 吸取上清至一支 1.5 ml 灭菌了的冰凉离心管中, 体积为 350 μ l, 加入等体积氯仿, 混匀, 4 $^{\circ}$ C 12000 r/min 离心 10min。
- 10) 吸取上清至灭菌了的 1.5ml 小离心管, 每管 25 μ l, 共 10 管。加入 2.5 μ l 醋酸钠, 再加入 75 μ l 无水乙醇, 冰上放置 90 分钟。
- 11) 4 $^{\circ}$ C 13000 r/min 离心 30min。
- 12) 用冰冷的 75%乙醇洗涤沉淀, 4 $^{\circ}$ C 13000 r/min 离心 30min。
- 13) 重复 75%乙醇洗涤 1 次。
- 14) 无水乙醇洗涤 1 次。室温自然干燥。供 RT 反应用。

2.2.5 RNA的逆转录反应

- 1) 在 DNase I 处理后的试管 (RT 管和负对照管) 中加入以下试剂:

表 2.2 RNA 逆转录反应

试 剂	RT(T)管	RT(-)管
DEPC 处理水	12 μ l	13 μ l

5×逆转录酶 Buffer	4 μl	4 μl
10 mM dNTPs(各 10 mM)	1 μl	1 μl
引物 (0.2 μg/μl)	2 μl	2 μl

注: ① RT 管共 19 μl, RT(-)管共 20 μl。

② 先加入 12 μl 或 13 μl DEPC 处理水, 使总 RNA 充分溶解, 然后加入其他试剂。

③ 所用引物的序列: Pn1F: 5' -TgAggTAgTAggTTgTgTg-3'
Pn1R: 5' -gAAggCagTAggTTgTATAg-3'

- 2) 将上述试剂混匀, 65℃变性 5min, 瞬时离心。冰上急冷。
- 3) 在 RT(T)管中加入 1 μl M-MLV 酶, RT(-)管中不加酶。
- 4) 42℃延伸 1 小时。
- 5) 70℃保温 10 分钟后, 冰上急冷。

6 PCR反应及其产物鉴定

(1) 取 3 个 PCR 管, 分别标明 PCR(T)、PCR(-)和 PCR(+). 在这 3 个 PCR 管中分别加入以下试剂:

表 2.3 PCR 反应

试剂	PCR(T)*	PCR(-)*	PCR(+)*
10×Ex Taq Buffer (Mg ²⁺ free)	5 μl	5 μl	5 μl
25mM Mg ²⁺	4 μl	4 μl	4 μl
dNTPs (各 10mM)	1 μl	1 μl	1 μl
Pn1F 引物(10 μM)	1 μl	1 μl	1 μl
Pn1R 引物(10 μM)	1 μl	1 μl	1 μl
TaKaRa Ex Taq	0.25 μl	0.25 μl	0.25 μl
Template	RT(T) 管产物 1 μl	RT(-)管产物 1 μl	基因组 DNA 0.3 μl (本次实验为 70ng)

3dH ₂ O	37 μl	37 μl	37.7 μl
--------------------	-------	-------	---------

(2) 热循环参数

94°C	4min		
94°C	30sec	} 30cycles	
50°C	30sec		
74°C	1min		
74°C	5min		

(3) PCR 产物的鉴定

①3%凝胶制备:

称取 0.6g 琼脂糖于一洁净小烧杯中, 加入 20ml 入 20ml 灭菌 3d H₂O, 用玻璃棒搅拌后放入微波炉内煮沸溶解, 直至溶液透明, 室温冷却至 60° C。向溶液中加入 400 μl 50×TAE 溶液, 混匀后倒胶。注意不能有气泡。

②电泳样品制备:

5 μl PCR 产物+2 μl 16×溴酚兰上样缓冲液, 混匀上样, 电泳检测。电泳条件: 80V, 约 30 分钟。

在对此 PCR 进行电泳检测后, 发现正对照即基因组 DNA 出现条带, 但 T 管没有条带。于是改变 PCR 反应温度, 再进行一次 PCR。

(4) 取 2 个 PCR 管, 分别标明 PCR(T)、PCR(-)。在这 2 个 PCR 管中分别加入以下试剂:

表 2.3 PCR 反应

试剂	PCR(T)*	PCR(-)*
10×Ex Taq Buffer (Mg ²⁺ free)	5 μl	5 μl
25mM Mg ²⁺	4 μl	4 μl
dNTPs (各 10mM)	1 μl	1 μl
Pn1F 引物(10 μM)	1 μl	1 μl
Pn1R 引物(10 μM)	1 μl	1 μl
TaKaRa Ex Taq	0.25 μl	0.25 μl

Template	RT(T) 管产物	RT(-) 管产物
	1 μ l	1 μ l
3dH ₂ O	37 μ l	37 μ l

(5) 热循环参数

94°C	4min	
94°C	30sec	} 30cycles
48°C	30sec	
74°C	1min	
74°C	5min	

(6) PCR 产物鉴定

制备 3%的凝胶对 PCR 产物进行检测。5 μ l PCR 产物+2 μ l 6 \times 溴酚兰上样缓冲液，混匀上样。

在在此次 PCR 进行电泳检测后，发现 T 管出现条带。

第3章 结果

3.1 300个同源 microRNA 基因的发现

目前,在 miRbase 数据库中只有 6 个狗 microRNA 分子被鉴定出来,而根据生物信息学方法预测狗基因组中应该含有 900 到 1900 个 microRNA 分子。为增加狗基因组 microRNA 分子的信息量以及更好的研究 microRNA 的起源,我们用计算机方法搜寻了 UCSC 基因组数据库中狗的基因组数据库和表达序列标签数据库,发现了 300 个与人 microRNA 同源的 microRNA 基因(见表 3.1)。表中同源物的名称是根据目前通用的 microRNA 命名原则并参考当前不同物种中的习惯命名法命名的。而且狗基因组中 microRNA 的数字顺序按照人基因组中的 microRNA 的数字顺序来命名的。为区别物种间的同源物,取该物种种名的前三个字母置于 microRNA 名称前。同一物种内同一 microRNA 的多个拷贝在其后加上 1, 2, 3 等表示;同一物种内属于同一家族的 microRNA 在其后加 a, b, c 等区分。

表 3.1 与人 microRNA 同源分子

Table 3.1 Dog homologues of human microRNA

名称	基因组位置	成熟序列	邻近基因
let-7a-1	chr10:23089154-23089175	tgaggtagtaggttgatagtt	70kb downstream of peroxisome proliferator activated receptor
let-7a-2	chr5:15057802-15057823	tgaggtagtaggttgatagtt	445kb upstream of sortilin-related receptor
let-7b	chr10:23088234-23088255	tgaggtagtaggttggtggtt	70kb downstream of peroxisome proliferator activated receptor
let-7c	chr31:16269069-16269090	tgaggtagtaggttgatggtt	570kb downstream of Ubiquitin carboxyl-terminal hydrolase
let-7d	chr1:100971005-100971025	agaggtagtaggttgcatagt	42kb upstream of protein tyrosine phosphatase domain
let-7e	chr1:108280903-108280923	tgaggtaggaggttgatagtt	8.5kb downstream of hypothetical protein XP_863434
let-7f-1	chr1:100972892-100972913	tgaggtagtagattgtatagtt	40kb upstream of protein tyrosine phosphatase domain
let-7f-2	chrX:45224906-45224927	tgaggtagtagattgtatagtt	Intron 24 of HECT, UBA and WWE domain(anti-sense)
let-7g	chr20:40496653-40496673	tgaggtagtagttgtacagt	Intron 2 of twinfilin-like protein(sense)

miR-1-1	chr24:49527584-49527604	tggaatgtaaagaagtatgta	8.9kb downstream of LOC612405
miR-1-2	chr7:69255574-69255594	tggaatgtaaagaagtatgta	Intron 9 of LOC490521(anti-sense)
miR-7-1	chr1:78565664-78565685	tggaagactagtgattttgtg	Intron 16 of HNRPK gene(sense)
miR-7-2	chr20:57938098-57938119	tggaagactagtgattttgtg	18kb downstream of LOC485032
miR-7-3	chr3:54663965-54663986	tggaagactagtgattttgtg	1.9kb downstream of LOC479036
miR-9-1	chr3:22881834-22881856	tctttggttatctagctgtatga	53kb downstream of LOC479155
miR-9-2	chr3:55272222-55272244	tctttggttatctagctgtatga	37kb downstream of LOC479040
miR-9-3	chr7:44477956-44477978	tctttggttatctagctgtatga	29kb upstream of RHBG
miR-10a	chr9:28168946-28168968	taccctgtagatccgaatttgg	Intron 4 of homeo box B4(sense)
miR-10b	chr36:22985951-22985972	taccctgtagaaccgaatttgg	Intron 1 of Homeobox protein Hox-D4
miR-15a	chr22:4786470-4786491	tagcagcacataatggtttgtg	31kb upstream of potassium channel regulator
miR-15b	chr34:29469909-29469930	tagcagcacatcatggtttaca	Intron 3 of hCAP-C(sense)
miR-16-1	chr22:4786610-4786631	tagcagcacgtaaatttggc g	23kb upstream of potassium channel regulator
miR-16-2	chr34:29470053-29470074	tagcagcacgtaaatttggc g	Intron 3 of hCAP-C(sense)
miR-17-3p	chr22:45426549-45426568	actgcagtggaaggcacttgt	Intron 2 of Glypican-5 precursor(sense)
miR-17-5p-1	chr22:45426512-45426535	caaagtgcttacagtgcaggt agt	Intron 2 of Glypican-5 precursor(sense)
miR-17-5p-2	chrX:107958258-107958279	aaaagtgcttacagtgcaggt agc	54kb upstream of hypothetical protein XP_863430
miR-18a	chr22:45426650-45426671	taagtgcatctagtgcagata	Intron 2 of Glypican-5 precursor(sense)
miR-18b	chrX:107958092-107958113	taagtgcatctagtgcagata	54kb upstream of hypothetical protein XP_863430
miR-19a	chr22:45426832-45426854	tgtgcaaatctatgcaaaactg a	Intron 2 of Glypican-5 precursor(sense)
miR-19b-1	chr22:45427134-45427156	tgtgcaaatccatgcaaaactg a	Intron 2 of Glypican-5 precursor(sense)
miR-19b-2	chrX:107957666-107957688	tgtgcaaatccatgcaaaactg a	55kb upstream of hypothetical protein XP_863430
miR-20a	chr22:45426961-45426983	taaagtgcttatagtgcaggt g	Intron 2 of Glypican-5 precursor(sense)
miR-20b	chrX:107957837-107957859	caaagtgctcacagtgcaggt ag	58kb downstream of nicalin
miR-20c-1	chr22:45426513-45426534	aaagtgcttacagtgcaggt g	Intron 2 of Glypican-5 precursor(sense)
miR-20c-2	chrX:107958258-107958279	aaagtgcttacagtgcaggt	54kb upstream of hypothetical

第 3 章 结果

miR-21	79 chr9:37662765-37662786	g tagcttatcagactgatgttga	protein XP_863430 1.3kb downstream of LOC491120
miR-22	chr9:49177554-49177575	aagctgccagttgaagaactg t	1.4kb upstream of CG6734-PA
miR-23a	chr20:51621784-51621804	atcacattgccaggattcc	4kb downstream of Zinc finger SWIM domain
miR-23b	chr1:74734972-74734992	atcacattgccaggattacc	1.5kb upstream of Aminopeptidase O (AP-O) isoform 2
miR-24-1	chr1:74734230-74734251	tggetcagttcagcaggaaca g	2.2kb upstream of Aminopeptidase O (AP-O) isoform 2
miR-24-2	chr20:51622139-51622160	tggetcagttcagcaggaaca g	3.7kb upstream of Zinc finger SWIM domain
miR-25	chr6:12505782-12505803	cattgcaactgtctcggtctga	Intron 13 of DNA replication licensing factor MCM7(sense)
miR-26a	chr23:10757556-10757576	ttaagtaatccaggataggc	Intron 6 of CTDSP-like(sense)
miR-26b	chr37:28054020-28054041	ttaagtaatccaggataggt	Intron 3 of carboxy-terminal domain(sense)
miR-27a	chr20:51622012-51622032	ttcacagtggttaagtccgc	3.8kb downstream of Zinc finger SWIM domain
miR-27b	chr1:74734738-74734758	ttcacagtggttaagtctgc	1.7kb upstream of Aminopeptidase O (AP-O) isoform 2
miR-28	chr34:23981810-23981831	aaggagctcacagtctattga g	Intron 4 of LIM domain (sense)
miR-29a	chr14:9137232-9137252	tagcaccattgaaatcggt	104kb upstream of Kruppel-like factor 14
miR-29b-1	chr14:9136849-9136871	tagcaccattgaaatcagtt	104kb upstream of Kruppel-like factor 14
miR-29b-2	chr7:9606851-9606873	tagcaccattgaaatcagtt	8.2kb downstream of membrane cofactor protein
miR-29c-1	chr7:9542985-9543004	tagcaccattgaaatcggt	25kb downstream of membrane cofactor protein
miR-29c-2	chr7:9606314-9606333	tagcaccattgaaatcggt	1.3kb downstream of membrane cofactor protein
miR-30a-3p	chr15:5287188-5287209	cttcagtcggatgttgcagc	Intron 5 of nuclear transcription factor Y(sense)
miR-30a-5p -1	chr12:36783527-36783548	tgtaaacatcctcgactggaag	93kb downstream of opioid growth factor receptor-like 1
miR-30a-5p -2	chr15:5287230-5287251	tgtaaacatcctcgactggaag	Intron 5 of nuclear transcription factor Y(sense)
miR-30b	chr13:33940974-33940995	tgtaaacatcctacactcagct	255kb downstream of Zinc finger

第 3 章 结果

miR-30c	chr15:5283970-5283992	tgtaaacaatcctacactctcag c	protein 406 isoform 6 Intron 5 of nuclear transcription factor Y(sense)
miR-30d	chr13:33945242-33945263	tgtaaacaatccccgactggaa g	259kb downstream of Zinc finger protein 406 isoform 6
miR-30e-3p	chr12:36783486-36783507	ctttcagtcggatgtttgcagc	93kb downstream of opioid growth factor receptor-like 1
miR-31	chr11:43934452-43934472	ggcaagatgctggcatagctg	40kb downstream of interferon, alpha 7
miR-32	chr11:67224721-67224741	tattgcacattactaagtgc	Intron 4 of CG2698-PA(anti-sense)
miR-33	chr10:26520455-26520520	gtgcattgtagttgcattg	Intron 3 of sterol regulatory element-binding transcription factor 2
miR-34a	chr5:65484449-65484471	tggcagtgcttagctggttgtt	Intron 4 of G protein-coupled receptor 157(anti-sense)
miR-34b	chr5:24567618-24567640	taggcagtgtaattagctgatt g	14kb upstream of BTG4 protein
miR-34c	chr5:24567022-24567044	aggcagtgtagttagctgattg c	15kb upstream of BTG4 protein
miR-92	chr22:45427250-45427270	tattgcactgtcccggcctg	Intron 2 of Glypican-5 precursor
miR-93	chr6:12505533-12505554	aaagtgctgttcgtgcaggtag	Intron 13 of DNA replication licensing factor MCM7
miR-95	chr3:62686242-62686263	ttcaacgggtattattagca	Intron 15 of Actin-binding LIM protein 2
miR-96	chr14:10022628-10022649	tttggcactagcacattttgc	18kb upstream of Nuclear respiratory factor-1
miR-98	chrX:45224071-45224092	tgaggtagtaagttgattgtt	Intron 24 of HECT, UBA and WWE domain
miR-99a-1	chr31:16268329-16268350	aaccgtagatccgatcttgtg	569kb downstream of Ubiquitin carboxyl-terminal hydrolase
miR-99a-2	chr5:15052369-15052390	aaccgtagatccgatcttgtg	450kb upstream of sortilin-related recepto
miR-99b	chr1:108281080-108281101	caccgtagaacccgaccttgc g	8.1kb downstream of hypothetical protein XP_863434
miR-100-1	chr31:16268329-16268350	aaccgtagatccgatcttgtg	569kb downstream of Ubiquitin carboxyl-terminal hydrolase
miR-100-2	chr5:15052369-15052390	aaccgtagatccgatcttgtg	450kb upstream of sortilin-related recepto
miR-101	chr5:48186773-48186794	tacagtactgtgataactgaag	14.7kb upstream of ribosomal protein L31
miR-103-1	chr24:20465778-20465800	agcagcattgtacaggctat ga	Intron 2 of Pantothenate kinase 2(anti-sense)
miR-103-2	chr4:46381059-46381081	agcagcattgtacaggctat	Intron 5 of pantothenate kinase

miR-105	chrX:123446917-123446936	ga tcaaatgctcagactcctgt	3(anti-sense) 23.5kb downstream of gamma-aminobutyric acid A receptor
miR-106a-1	chr22:45426962-45426983	taaagtgcttatagtcaggtag	Intron 2 of Glypican-5 precursor(sense)
miR-106a-2	chrX:107957837-107957858	aaagtgcttatagtcaggtag	54kb upstream of hypothetical protein XP_863430
miR-106a-3	chrX:107958257-107958280	aaagtgcttatagtcaggtag	54kb upstream of hypothetical protein XP_863430
miR-106b	chr6:12505310-12505330	taaagtgctgacagtgcagat	Intron 13 of DNA replication licensing factor MCM7(sense)
miR-107	chr28:7482406-7482428	agcagcattgtacagggtat	Intron 2 of pantothenate kinase 1(anti-sense)
miR-122a	chr1:20601591-20601613	tggagtgtgacaatgggtttgt	30kb downstream of heart alpha-kinase
miR-124a-1	chr24:49963196-49963217	ttaaggcacgcggtgaatgcc	5.6kb downstream of hypothetical protein XP_863655
miR-124a-2	chr25:31177068-31177089	ttaaggcacgcggtgaatgcc	Intron 3 of hypothetical protein XP_855676
miR-124a-3	chr29:17267181-17267202	ttaaggcacgcggtgaatgcc	164kb upstream of basic helix-loop-helix domain
miR-125a	chr1:108280427-108280449	tcctgagacccttaacctgt	8.1kb downstream of hypothetical protein XP_863434
miR-125b-1	chr31:16314527-16314548	tcctgagaccctaactgtga	586kb upstream of eukaryotic translation elongation factor
miR-125b-2	chr5:15102672-15102693	tcctgagaccctaactgtga	400kb upstream of sortilin-related receptor
miR-126	chr9:52212253-52212273	tcgtaccgtgagtaataatgc	Intron 3 of EGF-like-domain(anti-sense)
miR-127	chr8:72147709-72147730	tcggatccgtctgagcttgct	retrotransposon-like 1(anti-sense)
miR-128a	chr23:8837778-8837798	tcacagtgaaccggtctcttt	Intron 18 of cyclic AMP-regulated phosphoprotein(sense)
miR-128b	chr23:8837778-8837799	tcacagtgaaccggtctctttc	Intron 18 of cyclic AMP-regulated phosphoprotein(sense)
miR-129	chr18:29172954-29172974	cttttgcggtctgggcttgc	50kb upstream of tetratricopeptide repeat domain 17
miR-130a	chr18:41608990-41609011	cagtgcaatgttaaaggcca	3.3kb downstream of yippee-like 4
miR-130b	chr26:33989004-33989025	cagtgcaatgatgaaaggcca	3.6kb downstream of Stromal cell-derived factor 2-like protein

miR-132	chr9:49478540-49478561	taacagtctacagccatggtcg	1 precursor 5.9kb upstream of Hypermethylated in cancer 1 protein
miR-133a-1	chr24:49537020-49537041	ttgtcccccctcaaccagctgt	2kb donstream of hypothetical protein LOC612405
miR-133a-2	chr7:69258694-69258715	ttgtcccccctcaaccagctgt	Intron 9 of mindbomb homolog 1(anti-sense)
miR-133b	chrUn:73267767-7326778 6	ttgtcccccctcaaccagct	unknown
miR-134	chr8:72326236-72326256	tgtgactggtgaccagaggg	2kb downstream of retrotransposon-like 1
miR-135a-1	chr15:40191812-40191834	tatggctttttattcctatgtga	18kb downstream of eukaryotic translation elong factor 1
miR-135a-2	chr20:40457412-40457434	tatggctttttattcctatgtga	20kb downstream of dynein
miR-135b	chr38:4926466-4926487	tatggcttttcattcctatgtg	87kb downstream of kelch-like 18
miR-136	chr8:72149348-72149370	actccattgtttgatgatgga	retrotransposon-like 1(anti-sense)
miR-137	chr6:54695709-54695730	tattgcttaagaatacgcgtag	142kb upstream of DHPDHase
miR-138	chr23:5229083-5229167	agctggtgttgaatc	119kb upstream of hypothetical protein LOC608369
miR-139	chr21:28601506-28601573	tctacagtgcacgtgtct	Intron 2 of cGMP-dependent 3,5-cyclic Phosphodiesterase(sense)
mir-140	chr5:82820552-82820572	agtggttttaccctatggtag	Intron 8 of Nedd4-like E3 ubiquitin-protein ligase WWP2(anti-sense)
miR-141-1	chr27:41110053-41110074	taacactgtctggtaaagatgg	3kb downstream of B-cell receptor-associated protein 37
miR-141-2	chr5:59370309-59370329	taacactgtctggtaacgatg	1.44kb upstream of tubulin tyrosine ligase like family 9
miR-142-3p	chr9:36300840-36300862	tgtagtgtttcctactttatgga	Intron 8 of Peripheral-type benzodiazepine receptor
miR-142-5p	chr9:36300879-36300898	cataaagtagaaagcactac	35bp downstream of Peripheral-type benzodiazepine receptor
miR-143	chr4:62738116-62738136	tgagatgaagcactgtagctc	44kb upstream of Interleukin-17B precursor
miR-144	chr9:46356702-46356723	tacagtatagatgatgtactag	0.7kb downstream of Era G-protein-like 1
miR-145	chr4:62736758-62736781	gtccagttttccaggaaatccc tt	15kb downstream of casein kinase 1
miR-146a	chr4:53451772-53451793	tgagaactgaattccatgggtt	53kb upstream of Securin
miR-146b	chr28:17941066-17941087	tgagaactgaattccataggct	2.8kb downstream of CUE

miR-148a	chr14:42259219-42259240	tcagtgcactacagaactttgt	domain containing 2 154kb upstream of nuclear factor
miR-148b	chr27:3975864-3975885	tcagtgcacacagaactttgt	Intron 8 of Coatomer zeta-1 subunit(anti-sense)
miR-149	chr25:53510451-53510472	tctggctccgtgtcttactcc	Intron 1 of glypican 1 precursor(sense)
miR-150	chr1:109886388-109886409	tctcccaacctgtaccagtg	1.5kb upstream of ribosomal protein S11
miR-151	chr13:38395124-38395145	actagactgaggctccttgag g	Intron 9 of PTK2 protein tyrosine kinase(antisense)
miR-152	chr9:27662436-27662457	tcagtgcacacagaacttg g	24bp downstream of Coatomer zeta-2 subunit
miR-153	chr16:22886312-22886331	ttgcatagtcacaaaagtga	Intron 3 of protein tyrosine phosphatase(anti-sense)
miR-154	chr8:72325841-72325924	gaagttgtctggtggattcg	176kb downstream of retrotransposon-like 1
miR-155	chr31:24096793-24096814	ttaatgctaactgtgatagggg	11kb upstream of mitochondrial ribosomal protein L39
miR-181a-1	chr7:7011718-7011740	aacattcaacgctgtcggtga gt	101kb upstream of protein tyrosine phosphatase
miR-181a-2	chr9:61643462-61643484	aacattcaacgctgtcggtga gt	78kb downstream of olfactomedin-like 2A
miR-181b-1	chr7:7011904-7011925	aacattcattgctgtcggtggg	100kb upstream of protein tyrosine phosphatase
miR-181b-2	chr9:61642230-61642251	aacattcattgctgtcggtggg	77kb downstream of olfactomedin-like 2A
miR-181c	chr20:51594807-51594828	aacattcaacctgtcggtgagt	1.6kb downstream of nanos homolog 3
miR-181d	chr20:51594623-51594645	acattcattgtgtcggtgggt	1.4kb downstream of nanos homolog 3
miR-182	chr14:10026583-10026604	ttggcaatggtagaactcaca	56kb downstream of ubiquitin-conjugating enzyme E2H
miR-183	chr14:10022403-10022425	tatggcactggtagaattcact g	18kb upstream of Nuclear respiratory factor-1
miR-184	chr3:60777008-60777029	tggacggagaactgataagg gt	63kb downstream of cytosolic phosphoprotein DP58
miR-185	chr26:32362589-32362669	tggagagaaaggcagttc	4kb downstream of Ser/Thr-rich protein T10 in DGCR region
miR-186	chr6:77945929-77945951	caaagaattctcctttgggctt	Intron 8 of Zinc finger protein 265(sense)
miR-187	chr7:57243144-57243164	tcgtgtctgtgttcagccg	51kb downstream of hypothetical protein XP_861555
miR-188	chrX:42728545-42728566	catcccttgcattgtggagggt	Intron 1 of Chloride channel

miR-189	chr1:74734267-74734289	gtgcctactgagctgatatcag t	protein 5(sense) 2.2kb upstream of Aminopeptidase O (AP-O) isoform 2
miR-18a	chr22:45426650-45426671	taaggatcatctagtcagata	Intron 2 of Glypican-5 precursor(sense)
miR-18b	chrX:107958092-1079581 13	taaggatcatctagtcagata	54kb upstream of hypothetical protein XP_863430
miR-190	chr30:30424622-30424643	tgatatgttgatatattaggt	Intron 51 of talin 2 isoform 9(sense)
miR-191	chr20:43143422-43143443	caacggaatcccaaaagcag ct	0.2kb upstream of DALR anticodon binding domain
miR-192	chr18:55305795-55305815	ctgacctatgaattgacagcc	3.6kb downstream of CG1241-PA
miR-193a	chr9:44641941-44641961	aactggcctacaaagtcccag	25kb upstream of RAB11 family interacting protein 4
miR-193b	chr6:31906688-31906711	aactggcccacaaagtcccgc ttt	40kb upstream of megakaryoblastic leukemia 2 protein
miR-194	chr38:17895103-17895124	tgtaacagcaactccatgtgga	Intron 13 of mitochondrial isoleucine tRNA synthetase(anti-sense)
miR-195	chr5:35041910-35041930	tagcagcacagaaatattggc	0.4kb downstream of CG32955-PF
miR-196a-1	chr27:4272400-4272420	taggtagtttcatgttgg	0.7kb upstream of Homeobox protein Hox-C10
miR-196a-2	chr9:28222457-28222477	taggtagtttcatgttgg	1.2kb upstream of hypothetical protein XP_858268
miR-196a-3	chr9:28222457-28222477	taggtagtttcatgttgg	1.2kb upstream of hypothetical protein XP_858268
miR-196b	chr14:43299547-43299567	taggtagtttctgttgg	2.3kb upstream of homeobox A10
miR-197	chr6:45331319-45331340	ttcaccaccttctccaccagc	3kb downstream of Guanine nucleotide-binding protein G
miR-199a-1	chr20:53414687-53414709	cccagtggtcagactacctgtt c	Intron 5 of dynamin 2 isoform 2(anti-sense)
miR-199a-2	chr7:29778195-29778217	cccagtggtcagactacctgtt c	Intron 6 of dynamin 3 isoform 3(anti-sense)
miR-199b	chr9:58606432-58606454	cccagtggttagactatctgttc	Intron 7 of dynamin 1(anti-sense)
miR-200a-1	chr27:41110054-41110074	taacctgtctgtaacgatgt	3kb downstream of B-cell receptor-associated protein
miR-200a-2	chr5:59370309-59370330	taacctgtctgtaacgatgt	1.4kb upstream of tubulin tyrosine ligase
miR-200b	chr5:59369637-59369659	taatactgcctgtaaatgatga	2.1kb upstream of tubulin

miR-200c	chr27:41110475-41110496	c taatactgccgggtaatatgatgg	tyrosine ligase like family 3.2kb upstream of Tyrosine-protein phosphatase
miR-202	chr26:22668964-22668985	agaggtatagggcctgggaa aa	Intron 25 of Myosin-18B(sense)
miR-203	chr8:74819632-74819653	gtgaaatgtttaggaccactag	9.7kb downstream of lysophospholipase
miR-204-1	chr1:89887249-89887270	tccctttgcatcctatgcct	Intron 7 of transient receptor potential cation channel(sense)
miR-204-2	chr1:89887249-89887270	tccctttgcatcctttgcct	intron 7 of transient receptor potential cation channel
miR-204-3	chr3:40737600-40737621	tccctttgcatcctttgcct	Intron 9 of transient receptor potential cation channel(sense)
miR-205	chr7:11048078-11048099	tcctcattccaccggagtctg	160kb upstream of calcium/calmodulin-dependent protein
miR-206	chr12:22807601-22807622	tggaatgtaaggaagtgtgtg g	29kb downstream of basic transcription factor 3
miR-208	chr8:6640027-6640048	ataagacgagcaaaaagcttg t	Intron 6 of myosin heavy chain 6 isoform 3(anti-sense)
miR-210	chr18:28664843-28664940	ctgtgcgtgtgacagcggctg a	3.2kb downstream of Protein C11orf13 homolog
miR-211	chr3:40737600-40737621	tccctttgcatcctttgcct	Intron 9 of transient receptor potential cation channel(sense)
miR-212	chr9:49478922-49478942	taacagtctccagtcacggcc	5.5kb upstream of Hypermethylated in cancer 1 protein
miR-213	chr7:7011758-7011779	accatcgaccgttgattgtacc	101kb upstream of protein tyrosine phosphatase
miR-214	chr7:29784059-29784079	acagcaggcacagacaggca g	Intron 7 of dynamin 3(anti-sense)
miR-215	chr38:17894774-17894858	atgacctacgaattgatagac	Intron 13 of mitochondrial isoleucine tRNA synthetase(anti-sense)
miR-216	chr10:59909915-59909935	taatctcagctggcaactgtg	68kb downstream of EGF-containing fibulin-like extracellular matrix protein
miR-217	chr10:59903030-59903053	tactgcatcaggaactgattgg at	61kb downstream of EGF-containing fibulin-like extracellular matrix protein
miR-218-1	chr3:91870836-91870856	ttgtgcttgatctaaccatgt	Intron 12 of Slit homolog 2 protein precursor(anti-sense)
miR-218-2	chr4:46208919-46208939	ttgtgcttgatctaaccatgt	15kb upstream of pantothenate kinase 3

miR-219-1	chr12:5674314-5674334	tgattgtccaaacgcaattct	0.9kb upstream of ring finger protein 1
miR-219-2	chr9:58483186-58483206	tgattgtccaaacgcaattct	2.9kb upstream of CG33276-PA
miR-22	chr9:49177554-49177575	aagctgccagttgaagaactgt	1.4kb upstream of CG6734-PA
miR-220	chr7:80788945-80789055	ccaccaccgtgtccgacacct	Exon of tubulin(anti-sense)
miR-221	chrX:39471452-39471474	agctacattgtctgctgggttc	108kb downstream of Nucleoplasmin 3
miR-222	chrX:39472100-39472123	agctacatctggctactgggtctc	108kb downstream of Nucleoplasmin 3
miR-223	chrX:53762382-53762402	gtgcagttgtcaaatacccc	4.7kb upstream of V-set and immunoglobulin domain
miR-224	chrX:122969698-122969720	caagtcactagtggttccgttta	Intron 3 of GABA(A) receptor(anti-sense)
miR-299-3p	chr8:72298316-72298336	ggtttaccgtcccacatacat	148kb downstream of retrotransposon-like 1
miR-299-5p	chr8:72298287-72298308	tggtttaccgtcccacatacat	148kb downstream of retrotransposon-like 1
miR-301	chr9:37047053-37047075	cagtgcfaatagattgtcaaagc	14kb upstream of hypothetical protein XP_847434
miR-302a	chr32:35353700-35353722	taagtgttccatgttttagtga	Intron 8 of CG12058-PA(anti-sense)
miR-302b-1	chr32:35353526-35353546	taagtgttccatgttttagt	Intron 8 of CG12058-PA(anti-sense)
miR-302b-2	chr32:35353992-35354014	taagtgttccatgttttagtag	Intron 8 of CG12058-PA(anti-sense)
miR-302b-3	chr32:35353994-35354014	taagtgttccatgttttagt	Intron 8 of CG12058-PA(anti-sense)
mir-302c	chr32:35353857-35353879	taagtgttccatgtttcagtgg	Intron 8 of CG12058-PA(anti-sense)
miR-302d	chr32:35353524-35353546	taagtgttccatgttttagtgt	Intron 8 of CG12058-PA(anti-sense)
miR-320	chr25:38005897-38005919	aaaagctgggttgagagggcgaa	Intron 9 of RPC4(anti-sense)
miR-323	chr8:72300261-72300282	gcacattacacggtcgacctct	150kb downstream of retrotransposon-like 1
mir-324	chr5:35188874-35188895	ccactgcccaggtgctgctgg	Intron 11 of VLCAD(anti-sense)
miR-325	chrX:62503532-62503554	cctagtaggtgttcagtaaggt	69kb upstream of ribosomal protein L31 isoform 1
miR-326	chr21:26245546-26245565	cctctgggcccttctccag	21kb upstream of arrestin beta 1 isoform A isoform 2
miR-328-1	chr24:38208579-38208599	ctggcctccctgcccttccg	0.1kb downstream of Rho-related

miR-328-2	chr5:85174096-85174117	ctggccctctctgcccttcggt	GTP-binding protein RhoC Intron 7 of engulfment and cell motility 3, ced-12 homolog(anti-sense)
miR-329	chr8:72301310-72301331	aacacacctggttaacctcttt	151kb downstream of retrotransposon-like 1
miR-330	chr1:112813513-112813535	gcaaagcacacggcctgcagaga	Intron 1 of EMAP-2(sense)
miR-331	chr15:38245640-38245660	gcccctgggcctatcctagaa	8.5kb downstream of Vezatin isoform 3
miR-335	chr14:9397246-9397268	tcaagagcaataacgaaaaatgt	Intron 10 of mesoderm specific transcript isoform b isoform 3
miR-338	chr9:3907824-3907890	tccagcatcagtgattttgtga	Intron 7 of apoptosis-associated tyrosine kinase
miR-340	chr11:4705414-4705436	tccgtctcagttactttatagcc	1.8kb upstream of ring finger protein 130 isoform 1
miR-342	chr8:71462947-71462970	tctcacacagaaatgcacccgtc	Intron 3 of Enah/Vasp-like(sense)
miR-345	chr8:71630674-71630770	tgctgactcttagtccagtgcc	Intron 7 of Mitochondrial carnitine/acylcarnitine carrier protein CACL(anti-sense)
miR-346	chr4:37263622-37263644	tgctgcccgcgatgcctgcctct	Intron 27 of Glutamate receptor delta-1 subunit precursor(anti-sense)
miR-361	chrX:69751313-69751334	ttatcagaatctccaggggtac	Intron 6 of REP-1(anti-sense)
miR-362	chrX:42733842-42733865	aatccttggaacctaggtgtgagt	Intron 1 of Chloride channel protein 5(sense)
miR-363	chrX:107957363-107957384	aattgcacggatccatctgta	55kb upstream of hypothetical protein XP_863430
miR-365-1*	chr6:31901789-31901810	taatgccctaaaaatccttat	45kb upstream of megakaryoblastic leukemia 2 protein
miR-365-2*	chr9:44628573-44628594	taatgccctaaaaatccttat	39kb upstream of RAB11 family interacting protein 4
miR-367	chr32:35353393-35353460	aattgcacttagcaatggtga	Intron 11 of CG12058-PA(anti-sense)
miR-368	chr8:72312886-72312907	acatagaggaaatccacgttt	163kb downstream of retrotransposon-like 1
miR-369-3p	chr8:72336841-72336861	aataatacatggttgctctt	187kb downstream of retrotransposon-like 1
miR-369-5p	chr8:72336806-72336827	agatcgaccgtgtatattcgc	187kb downstream of retrotransposon-like 1
miR-370	chr8:72175566-72175586	gcctgctggggtggaacctgg	26kb downstream of retrotransposon-like 1

miR-371	chr1:106488132-106488196	gtgccgccatttttgagtgt	85kb downstream of myeloid-associated differentiation marker
miR-374-1	chrX:60604633-60604653	tataataacaacctgctaagtg	39kb downstream of Heat shock cognate 71 kDa protein
miR-374-2	chrX:60644810-60644831	ttataataacaacctgataagtg	39kb upstream of hypothetical protein XP_845014
miR-375	chr37:28567574-28567595	ttgttcgttcggctcgcgtga	0.9kb upstream of hypothetical protein XP_545652
mir-376a-1	chr8:72312523-72312543	atcatagaggaaaatccacgt	163kb downstream of retrotransposon-like 1
mir-376a-2	chr8:72313269-72313289	atcatagaggaaaatccacgt	163kb downstream of retrotransposon-like 1
mir-376a-3	chr8:72314029-72314049	atcatagaggaaaatccacgt	164kb downstream of retrotransposon-like 1
mir-376b	chr8:72313651-72313671	atcatagaggaaaatccatgt	164231 downstream of retrotransposon-like 1
miR-377	chr8:72333714-72333735	atcacacaaaggcaactttgt	184kb downstream of retrotransposon-like 1
miR-378	chr4:62479921-62479942	ctcctgactccaggtcctgtgt	Intron 12 of PGC-1-related estrogen receptor alpha coactivator(anti-sense)
miR-379	chr8:72296536-72296602	tgtagactatggaacgta	147kb downstream of retrotransposon-like 1
miR-380-3p	chr8:72299533-72299554	tatgtaatatggccacgtctt	150kb downstream of retrotransposon-like 1
miR-380-5p	chr8:72299498-72299519	tggtgaccatagaacatgcgc	150kb downstream of retrotransposon-like 1
miR-381	chr8:72317764-72317785	tatacaagggcaagctctctgt	168kb downstream of retrotransposon-like 1
miR-382	chr8:72325855-72325876	gaagtgttcgtggtgattcg	176kb downstream of retrotransposon-like 1
miR-383	chr16:41130450-41130471	agatcagaaggtgattgtgct	605kb downstream of sarcoglycan zeta
miR-384	chrX:62449950-62450037	attcctaggcaatgtgtataatg tttgta	308kb downstream of tropomyosin 1 alpha chain
miR-409-3p	chr8:72336546-72336568	cgaatgttgctcgtgaacccct	187kb downstream of retrotransposon-like 1
miR-409-5p	chr8:72336515-72336536	aggttaccgagcaactttgca	187kb downstream of retrotransposon-like 1
miR-410	chr8:72337145-72337165	aatataacacagatggcctgt	187kb downstream of retrotransposon-like 1
miR-421	chrX:60604425-60604509	atcaacagacattaattgggcgc	40kb downstream of Heat shock cognate 71 kDa protein

miR-422a	chr36:18209341-18209362	ctggacttggggtcagaaggc c	66kb upstream of Sp5 transcription factor
miR-422b-1	chr30:24956080-24956100	ctggacttggagtcagaaggc c	48kb downstream of suppressor of hairy wing homolog 4
miR-422b-2	chr4:62479882-62479903	ctggacttggagtcagaaggc c	Intron 12 of PGC-1-related estrogen receptor alpha coactivator
miR-423	chr9:47480984-47481005	agctcggctgaggcccctca g	Intron 1 of CG15747-PA(sense)
miR-424	chrX:108245221-1082452 42	cagcagcaattcatgtttgaa	30kb downstream of hypoxanthine phosphoribosyltransferase
miR-425	chr20:43143941-43143961	atcggaatgtcgtgccc	Intron 1 of DALR anticodon binding domain(sense)
miR-431	chr8:72145691-72145711	tgtcttcagggcctcatgca	retrotransposon-like 1(anti-sense)
miR-432	chr8:72149161-72149183	tcttgagtaggtcattgggtg g	retrotransposon-like 1(anti-sense)
miR-433	chr8:72146623-72146644	atcatgatgggctcctcgggtg	retrotransposon-like 1(anti-sense)
miR-448*	chrX:90623066-90623087	ttgcatatgtaggatgtcccat	Intron 2 of 5-hydroxytryptamine(sense)
miR-449*	chr2:45390408-45390429	tggcagtgattgttagctggt	Intron 10 of cell division cycle 20 homolog(anti-sense)
miR-450*	chrX:108239240-1082392 61	ttttgcgatgtgtcctaata	24kb downstream of hypoxanthine phosphoribosyltransferase
miR-451	chr9:46356528-46356599	aaaccgttaccattactgagttt	69kb downstream of Era G-protein-like 1
miR-452	chrX:122970733-1229707 54	tgtttgcagaggaaactgaga c	Intron 3 of GABA(A) receptor(anti-sense)
miR-455	chr11:71492140-71492161	tatgtgcccttgactacatcg	5.7kb upstream of collagen, type XXVII, alpha 1
miR-483	chr18:49321421-49321442	tcactcctcccctcccgtcttc	Intron 16 of Insulin-like growth factor II precursor(anti-sense)
miR-485-3p	chr8:72327009-72327030	gtcatacacggctctctctctct	177kb downstream of retrotransposon-like 1
miR-485-5p	chr8:72326972-72326993	agaggctggccgtgatgaatt c	177kb downstream of retrotransposon-like 1
miR-486-1	chr16:26912944-26912965	tctgtactgagctgccccga g	0.8kb downstream of ankyrin 1 isoform 3
miR-486-2	chr16:26912984-26913005	tctgtactgagctgccccga g	0.8kb downstream of ankyrin 1 isoform 3
miR-487a	chr8:72323859-72323880	aatcatacaggacatccagtt	174kb downstream of retrotransposon-like 1
miR-487b	chr8:72318265-72318286	aatcgtacagggtcatccact	168kb downstream of

miR-488	chr7:25566778-25566798	cccagataatggcactctcaa	retrotransposon-like 1 Intron 2 of astrotactin isoform 1(sense)
miR-489	chr14:21927495-21927517	agtgacatcacatatacggcagc	Intron 20 of calcitonin receptor(anti-sense)
miR-490	chr16:14579697-14579718	caacctggaggactccatgctg	109kb downstream of Muscarinic acetylcholine receptor
miR-491	chr11:43329133-43329155	agtggggaacccttccatgagga	Intron 2 of CG3520-PA
miR-493-3p	chr8:72137125-72137145	tgaaggtctactgtgtgccag	8.2kb upstream of retrotransposon-like 1
miR-493-5p	chr8:72137084-72137105	ttgtacatgtaggctttcatt	8.2kb upstream of retrotransposon-like 1
miR-494	chr8:72304165-72304188	tgaacatacacgggaaacctctt	154kb downstream of retrotransposon-like 1
mir-495	chr8:72307938-72307960	aaacaacatggtgcacttctt	158kb downstream of retrotransposon-like 1
miR-496	chr8:72332139-72332240	attacatggccaatctc	182kb downstream of retrotransposon-like 1
miR-498	chr4:79294525-79294544	caagccagggggcgttttc	91kb upstream of ribosomal protein S24
mir-499	chr24:27001150-27001172	ttaagactgcagtgatgtttaa	Intron 13 of myosin, heavy polypeptide 7B(sense)
miR-500-3p	chrX:42739112-42739133	atgcacctgggcaaggattcag	Intron 1 of Chloride channel protein 5(sense)
miR-500-5p	chrX:42733383-42733404	atgcacctgggcaaggattctg	Intron 1 of Chloride channel protein 5(sense)
miR-502	chrX:42739075-42739095	atccttgcctatctgggtgcta	Intron 1 of Chloride channel protein 5(sense)
miR-503	chrX:108244927-108244949	tagcagcgggaacagtactgcag	30kb downstream of hypoxanthine phosphoribosyltransferase
miR-504	chrX:111761253-111761273	agaccctggtctgcactctat	Intron 2 of fibroblast growth factor 13(anti-sense)
miR-505	chrX:112934486-112934507	gtcaacactgctggttctctc	11kb upstream of hypothetical protein XP_849787
miR-539	chr8:72319051-72319072	ggagaaattatccttgggtgtg	169kb downstream of retrotransposon-like 1
miR-542-5p	chrX:108240144-108240165	tcggggatcatcatgtcacgag	25kb downstream of hypoxanthine phosphoribosyltransferase
mir-544	chr8:72320443-72320462	attctgcatttttagcaagt	171kb downstream of retrotransposon-like 1
miR-545	chrX:60644587-60644692	atcaacaacatttatcgtg	79kb downstream of Heat shock

miR-551a	chr5:61113643-61113738	gcgaccactcttggttcca	cognate 71 kDa protein 13kb downstream of EGF-like-domain
miR-551b	chr34:36298215-36298309	gcgaccatacttggttcag	75kb upstream of LOC488147
miR-569	chr34:38471240-38471335	agttaatgaatcctggcgagt	Intron 12 of TRAF2 and NCK interacting kinase(sense)
miR-574	chr3:76319711-76319797	cacgctcatgcacacaccac	15kb downstream of CG9590-PA
miR-582	chr2:49362250-49362346	ttacagttgtcaaccagtact	123kb downstream of cAMP-specific 3,5-cyclic posphodiesterase 4D
miR-590	chr6:9187771-9187867	gagcttattcataaaagtacag	Intron 2 of Eukaryotic translation initiation factor 4H(sense)
miR-592	chr14:12132571-12132666	ttgtgcaatatgcgatgatgt	Intron 2 of Metabotropic glutamate receptor 8 Precursor(sense)
miR-599	chr13:4538300-4538394	gttgtgcagttatcaaac	Intron 1 of 60S ribosomal protein L11(anti-sense)
miR-628	chr30:23760021-23760115	tctagtaagagtggcagtcg	Intron 4 of cell cycle progression 1(anti-sense)
miR-652	chrX:86340960-86341057	aatggcgccactagggttg ca	Intron 2 of CG14591-PA(sense)
miR-660	chrX:42737423-42737513	taccattgcatatcggagttg	Intron 1 of Chloride channel protein 5(sense)

* 表示该基因已经在 miRbase 数据库公布。* represents microRNA is identified by miRbase database.

这些 microRNA 前体长度大多数在 65-120nt 之间, 平均长度在 80nt, 但是它们能够折叠成典型的发夹结构。这些 microRNA 分子在狗的基因组中分布是不均一的 (见图 3.1), 其中在 8 号和 X 号染色体上分布最多, 分别为 46 和 42 个分子, 在 19 和 29 号染色体上分别发现一个 miRNA 分子, 而在 17、33、35 号染

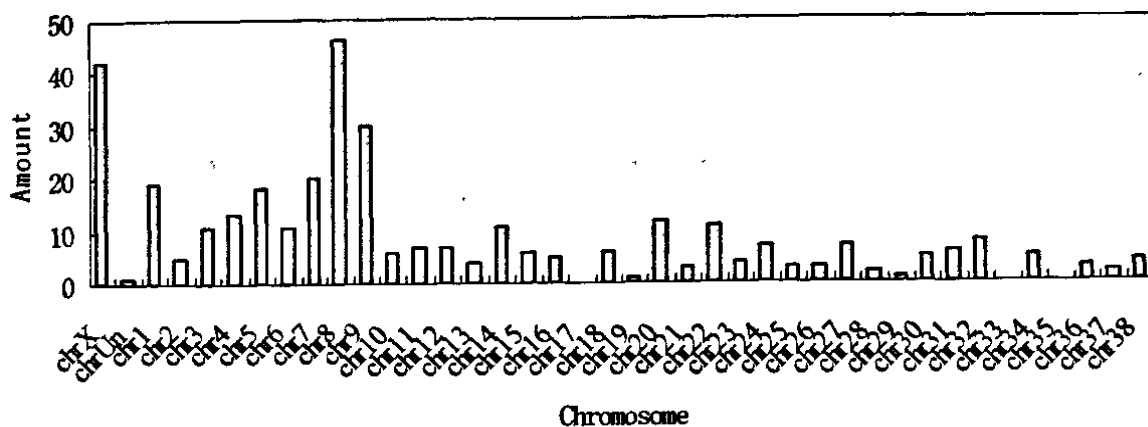


图 3.1 狗基因组 microRNA 的染色体分布

Fig 3.1 distribution of dog microRNA in chromosome

这 300 个与人同源的 miRNA 分子根据 miRBase 的注释进行分类,发现它们属于 149 个基因家族。MiR-154, miR-17 和 let-7 家族分别有 15, 13 和 10 成员,而大多数 miRNA 家族,比如说 miR-153, miR-146, miR-137 家族只有一个成员。

大多数 microRNA 的成熟序列与人的序列是非常保守的,只有一些 microRNA 是相差一两个碱基。miR-143 与人的 miR-143 相差一个碱基,但是它们前体的折叠自由能却是非常接近的,分别为 $56 \text{ kcal} \cdot \text{mol}^{-1}$, $52 \text{ kcal} \cdot \text{mol}^{-1}$ 。miR-489、miR-500 与 miR-107 也有上面类似的情况。

通过同源序列比对表明一些 miRNA 分子的 miRNA*比成熟序列更保守,例如 miR-322、miR-329 和 miR-346 (见图 3.2)。我们发现 miR-322 的 miRNA*上一个碱基的突变伴随着成熟序列上的相应的变化,在狗当中是 A/U 配,而在小鼠却是 G/C 配。

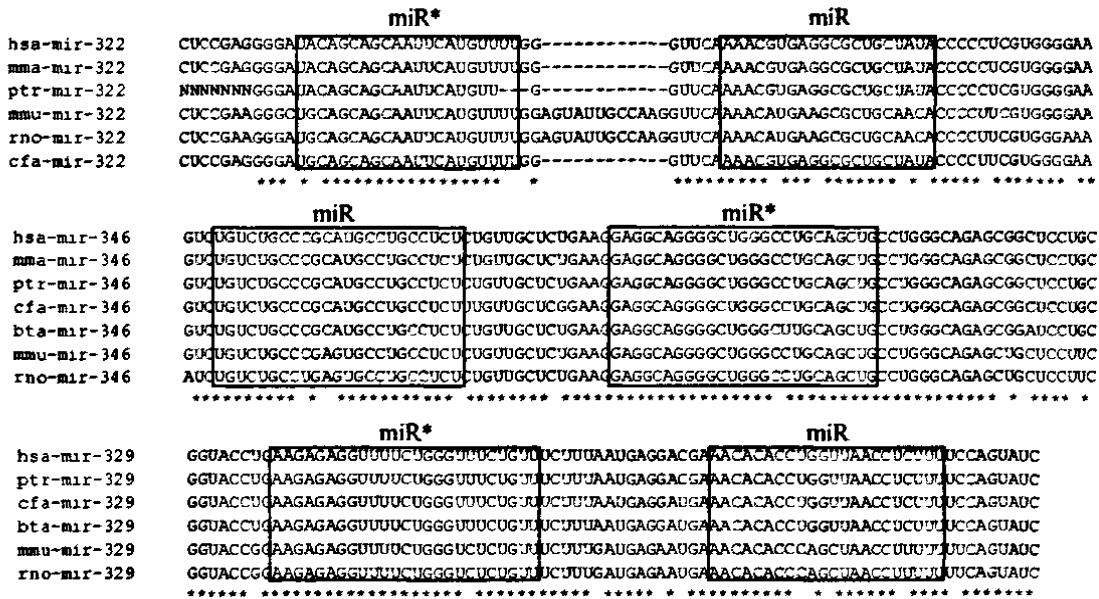


图 3.2 一些动物 microRNA 的多序列比对。miR 代表成熟序列, miR*代表与成熟序列互补的序列。

Fig 3.2 multi sequence alignment of many animal microRNA. miR represent mature sequence, and miR* represent complementary sequence with mature

到目前为止,只存在两个 microRNA 分子在动植物中都是保守存在的。另一方面,我们对植物和动物的 microRNA 序列进行了多序列比对,发现动物 microRNA 前体间的保守性比植物 microRNA 前体之间的保守性要更大(见图 3.3)。

```

cfa-mir-490  AAGTTCATTGTTTCGACACCATGGATCTCCAGGTGGGTCAAGTTTAGAGATACACCAACCTGGAGGACTCCATGCTGTTGAGCTGTTTAC
bta-mir-490  AAGTTCATTGTTTCGACACCATGGATCTCCAGGTGGGTCAAGTTTAGAGATACACCAACCTGGAGGACTCCATGCTGTTGAGCTGTTTAC
mmu-mir-490  AAGTTCATTGTTTCGACACCATGGATCTCCAGGTGGGTCAAGTTTAGAGATACACCAACCTGGAGGACTCCATGCTGTTGAGCTGTTTAC
rno-mir-490  AAGTTCATTGTTTCGACACCATGGATCTCCAGGTGGGTCAAGTTTAGAGATACACCAACCTGGAGGACTCCATGCTGTTGAGCTGTTTAC
ptr-mir-490  AAGTTCATTGTTTCGACACCATGGATCTCCAGGTGGGTCAAGTTTAGAGATACACCAACCTGGAGGACTCCATGCTGTTGAGCTGTTTAC
mma-mir-490  AAGTTCATTGTTTCGACACCATGGATCTCCAGGTGGGTCAAGTTTAGAGATACACCAACCTGGAGGACTCCATGCTGTTGAGCTGTTTAC
hsa-mir-490  AAGTTCATTGTTTCGACACCATGGATCTCCAGGTGGGTCAAGTTTAGAGATACACCAACCTGGAGGACTCCATGCTGTTGAGCTGTTTAC
*****

ath-MIR156a  AAACUGACAGAAGAGAGUGAGC-ACACAAAGGCAAUUU-GCAUADCAUUG--CACUUGCUUCUCU--UGCUGUCUCACUGCUCUUCUGUCAG
ptc-MIR156a  AGACUGACAGAAGAGAGUGAGC-ACACACGAAAGUAUA-UGGUAUGAAGG--CAUUCCAUUGCAGG-GUGUGUCUCACAUUCUCU-CUGUCAG
osa-MIR156a  AGGGUGACAGAAGAGAGUGAGC-ACACGUGGUGUUCUUGCAUAAUADGCUAUGCUUGGAGCUACGCGUGUCACUCUCUCUCUCUGUCAC
zma-MIR156a  GGGCAGACUGCAUCGAUCGACUCGACUCCGAGACGGCGCACGUAUGAAUGAU-GAUGCAGCUGCTGC--UGCUGUCUCACUUCUCUCUCUGUCAG
*****

```

图 3.3 动物 microRNA 和植物 microRNA 前体保守性的比较

Fig 3.3 comparison of conservation between microRNA in animal and microRNA in plant

3. 2 新的 microRNA 分子的鉴定

我们利用 RNAz 软件在狗的基因组中鉴定出 57 个新的 microRNA 分子 (Table3.2), 这些分子在 miRBase 中未曾报道过。这些分子也可以折叠成典型的发夹二级结构(见图 3.4)。这些 miRNA 分子分布在基因组的不同位置。

表 3.2 狗新的 microRNA

Table 3.2 novel microRNA of dog

名称	基因组位置	成熟序列	邻近基因
miR-801*	chrX:108245155-108245263	atgtaagaagtgcattg	30kb downstream of hypoxanthine phosphoribosyltransferase
miR-802*	chrX:107958220-107958284	acctgcactgtaagcacttt	54kb upstream of hypothetical protein XP_863430
miR-803*	chr9:55133289-55133377	tgacagctgcaaggatcattc	Intron 6 of hypothetical protein XP_845481(anti-sense)
miR-804*	chr1:74734963-74735038	tggtgctgtgggttctggca	1.4kb upstream of Aminopeptidase O (AP-O) isoform 2
miR-805*	chrX:56268429-56268511	ttggcttttttcacagagcc	38kb downstream of Ephrin-B1 precursor
miR-806*	chr14:12132568-12132669	tattatgccatgacattgtgt	Intron of Metabotropic glutamate receptor 8 precursor(sense)
miR-807*	chr1:70136070-70136179	tgcagtgaagtcaaatgtgag	Intron of Receptor-type tyrosine-protein phosphatase(anti-sense)
miR-808*	chr1:30897666-30897740	ggtagaaccagggtgaaactc	42kb upstream of myeloblastosis proto-oncogene product
miR-809*	chr12:47506905-47507014	ccactgtgtttttgatggagt	22kb downstream of EEA1

miR-810*	chr12:64887257-64887355	tgaagctgctatgctttgtgc	153kb downstream of EH-domain
miR-811*	chr3:28999740-28999814	tggctcaggaaataccaaggt	Intron of Ras protein-specific guanine factor(sense)
miR-812*	chr4:86009182-86009265	atggacacatattattgtgt	76kb upstream of cadherin 12, type 2 preproprotein
miR-813*	chr3:83307585-83307657	cagacaaaaaacagctgca g	95kb downstream of protocadherin 7 isoform
miR-814*	chr20:40457402-40457479	tggccttttattcctatgtga	1.1kb upstream of CG9886-like isoform 1
miR-815*	chr24:23154817-23154883	ggaacaggccgtcctgactgt	45kb downstream of Protein C20orf54 homolog
miR-816*	chr36:8534376-8534470	ttggcattgtatagacatagc	Intron of axotrophin isoform 5(sense)
miR-817*	chr19:35425673-35425772	tgctcgtccacggctgctg	369kb downstream of ATP-dependent RNA helicase DDX18
miR-818*	chr1:20601544-20601626	tagcagagctgtggagtgtga	30kb downstream of heart alpha-kinase
miR-819*	chr7:65151430-65151513	ttggttgcctttgttcccga	Intron of potassium channel tetramerisation domain(sense)
miR-820*	chr9:35856043-35856126	accgcggtggtgtattggtgt	15kb upstream of mitochondrial ribosomal protein S23
miR-821*	chr9:40290613-40290699	gacccccaccttttgagtct	Intron of acetyl-coenzyme A carboxylase alpha(sense)
miR-822*	chr9:46356696-46356769	tacagtatatgatgatcct	0.7kb downstream of Era G-protein-like 1
miR-823*	chr9:49627794-49627903	gtgtaatgagttcctcagct	Intron of Est1p-like protein A
miR-824*	chr30:8264861-8264953	taagagaatgctgtagaagtg	54kb upstream of hypothetical protein XP_535427
miR-825*	chr30:7445335-7445403	tctgtatcccttaaggcgt	Intron of Homeobox protein Meis2(sense)
miR-826*	chr8:72336492-72336596	cggtgaacccctttcggtat	187kb downstream of retrotransposon-like 1
miR-827*	chr8:8626519-8626595	accgtatctgaaaattcaat	480kb downstream of 40S-ribosomal protein S2
miR-828*	chr27:4211239-4211325	tgaagagccttcagagctgt	0.6kb upstream of Homeobox protein Hox-C4
miR-829*	chr27:4440399-4440495	tgtaacagcactttcatatat	30kb upstream of solute carrier family 7
miR-830*	chr21:28601502-28601583	tcagggttatctacagtgca	Intron of cGMP-dependent 3,5-cyclic phosphodiesterase
miR-831*	chr18:56128863-56128946	gggtttccgggagggtggct	0.4kb upstream of Serine/threonine-protein kinase MARK2
miR-832*	chr7:37286007-37286090	gcgtgcactttactctgtag	Intron of KARP-1-binding protein
miR-833*	chr7:44720392-44720452	tgatgatccccctccccctct	1.8kb downstream of ring finger and KH domain
miR-834*	chr6:67323318-67323409	ttggcactgggacaggacca	342kb downstream of CG5987-PA
miR-835*	chr5:57605645-57605700	ccccagaggccccggtgggt t	73kb downstream of thioesterase, adipose associated isoform BFIT2
miR-836*	chr2:75425386-75425456	tgtgttgtagatgctgggagt	Intron of protein phosphatase 1 regulatory inhibitor
miR-837*	chr31:31631579-31631649	agtggcttgggggtaaggagt	Intron of Mitochondrial 28S ribosomal protein S6
miR-838*	chrX:46459020-46459112	tcctctcctcctaaacctct	6kb downstream of hypothetical protein XP_850380

第 3 章 结果

miR-839*	chrX:33927989-33928101	gaaggtttctggatgctgctcc t	428kb downstream of MID1 interacting G12-like protein
miR-840*	chr9:55133459-55133539	aactgtaataagtttgacagc	Intron of hypothetical protein XP_845481
miR-841*	chr11:49849922-49850044	gaagtgctgtggatttcttgt	426kb downstream of Importin alpha-2 subunit
miR-842*	chr12:64887358-64887428	aagctgctatcgttgtgctt	153kb downstream of EH-domain
miR-843*	chr4:62736711-62736794	tccagtttcccaggaatccct	45kb upstream of Interleukin-17B precursor
miR-844*	chr4:75185650-75185729	aaatatgagttggttgctgg	Intron of Na ⁺ -dependent glutamate transporter
miR-845*	chr4:90471362-90471445	tggcagttggaagatgtact t	19kb downstream of membrane-associated RING-CH protein
miR-846*	chr15:47524170-47524267	taatgtattactctaatacatct	44kb downstream of zinc finger protein 64
miR-847*	chr32:35353423-35353503	aattgcacttagcatcggtg	Intron of CG12058-PA
miR-848*	chr9:3907816-3907899	tccagcatcagtgatttgg	Intron of apoptosis-associated tyrosine kinase
miR-849*	chr9:40290783-40290853	agaataggaaggaggggat gtg	Intron of acetyl-coenzyme A carboxylase alpha
miR-850*	chr8:27380807-27380882	tctgtaactgtcaagtgatt	32kb downstream of MAM domain
miR-851*	chr7:38023936-38024019	acttttattcagcttgagct	71kb upstream of zinc finger protein 238
miR-852*	chr38:4926414-4926514	tatggctttcattcctatgt	67kb upstream of PCTAIRE protein kinase 3
miR-853*	chr7:26064174-26064262	aagagctgtgctcatgtaaca gg	10kb downstream of T-complex protein
miR-854*	chr6:63134759-63134843	ttttattaatggatgctgact	342kb downstream of tubulin
miR-855*	chr2:75425446-75425530	tgtgttgtagtagctgggtt	Intron of protein phosphatase 1 regulatory inhibitor
miR-856*	chr10:23088177-23088260	tgaggtagtaggtgtgtgt	69kb downstream of peroxisome proliferator activated receptor
miR-857*	chrX:123446867-123446949	tgagcatgtgctacggtgtct	23kb downstream of gamma-aminobutyric acid A receptor

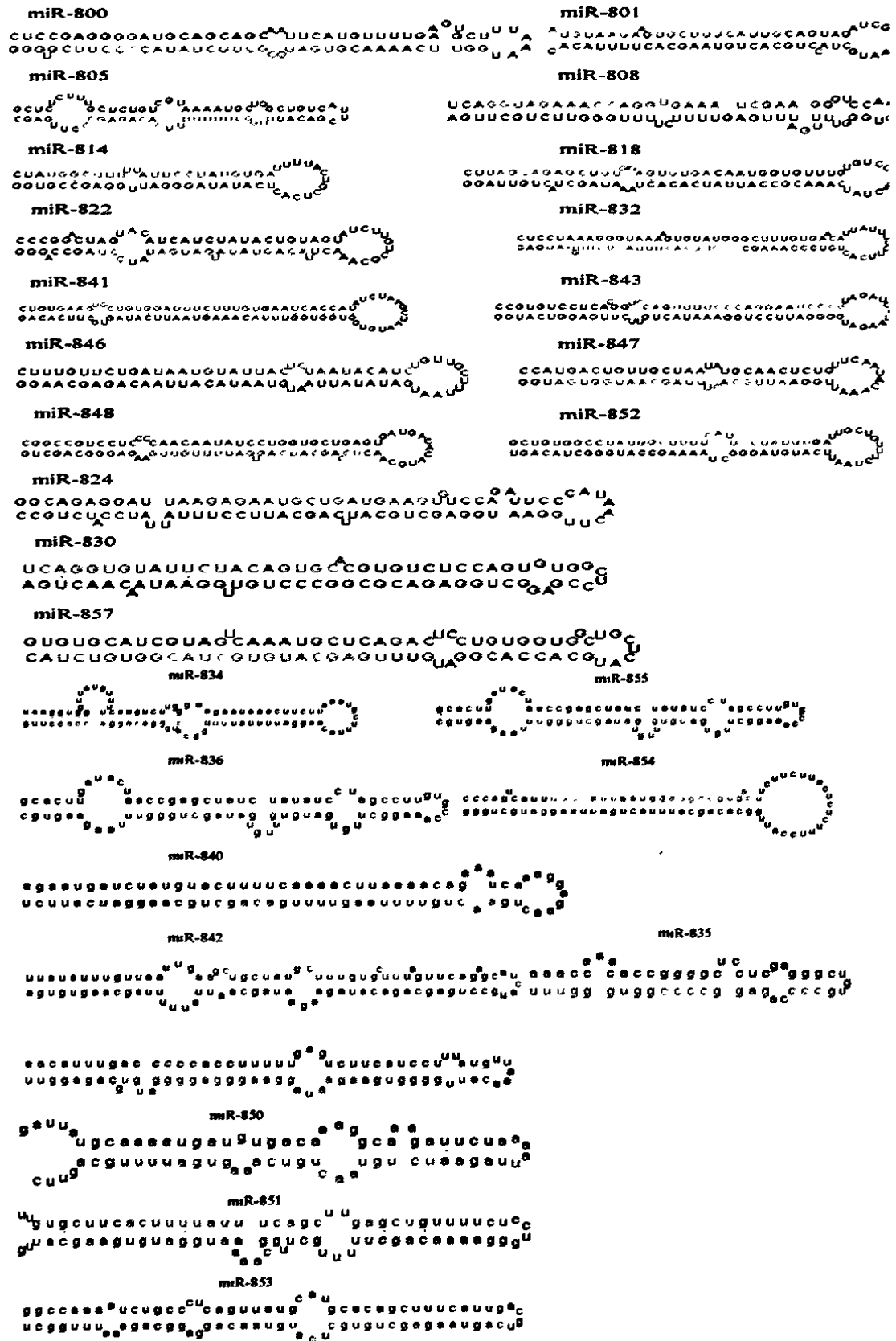


图 3.4 狗的新的 microRNA 的二级结构折叠图

Fig 3.4 secondary structure of novel microRNA in dog

3.3 microRNA 分子的基因组织形式

我们所鉴定出的 357 个 microRNA 分子中, 其中 125 个位于蛋白质编码基因的内含子中, 其它的位于基因间隔区或者是 UTR 区。(见图 3.5)

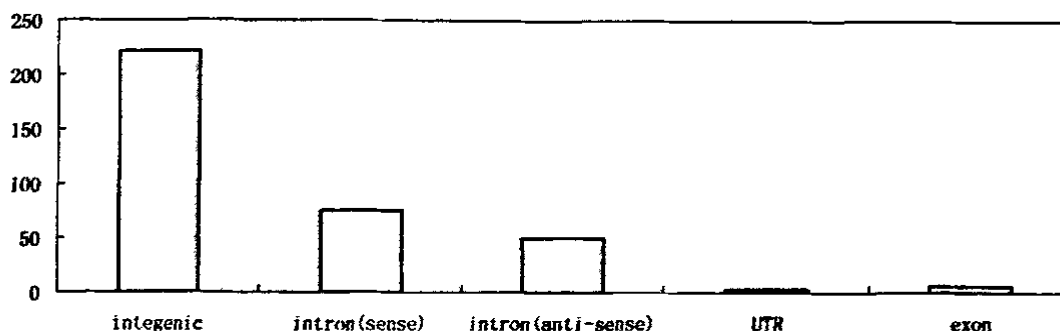


图 3.5 microRNA 的基因组织形式

Fig 3.5 Genomic organization of microRNA

通过基因组位置研究表明大多数同源的 microRNA 分子在人和狗的位置是非常相似的, 同源序列可以位于编码垂直同源基因的不同内含子中, 或者是位于类似邻近基因的基因间隔区。然而有些 miRNA 分子具有一些差异。狗的 miR-34a 位于编码 G 蛋白耦联受体基因的内含子中, 但是人的 miR-34a 却位于基因间隔区, 靠近编码 G 蛋白耦联受体基因。miR-181b 也存在类似这样的情况。

通过 EST 数据库搜索, 我们发现几个 miRNA 分子分别位于一个转录单位内, 而这个转录单位却反方向位于一个宿主基因内。miR-208 反方向位于 MYH6 基因的内含子内, 但是它同时正向位于一个 EST (BU744457) 内的内含子中, 这个 EST 在正向上与 MYH6 基因转录单元重叠。另外, miR-361 反向位于 CHM 基因的内含子中, 但同时反向位于一个多聚腺苷酸化的不包含内含子的 mRNA (AB012223) 中。

通过 microRNA 启动子序列分析, 我们发现 142 个 microRNA 分子组织成 50 个簇 (Table 3.3), 这些簇位于内含子或者是基因间隔区。

表 3.3 狗 microRNA 基因簇

Table 3.3 microRNA cluster in dog genome

No.	cluster
1	Let-7d let-7f-1
2	miR-125a let-7e miR-99b
3	miR-24-1 miR-189 miR-27b miR-23b
4	miR-30e-5p miR-30-3p
5	Let-7a Let-7b
6	miR-183 miR-96
7	miR-29a miR-29b

8	miR-30a-3p miR-30a-5p
9	miR-181c miR-181d
10	miR-23a miR-27a miR-24-2
11	miR-17-5p miR-17-3p miR-20c miR-18a miR-19a4 miR-20a miR-106a miR-19b miR-92
12	miR-15a miR-16
13	miR-141 miR-200a miR-200c
14	miR-99a miR-100 let-7c
15	miR-367 miR-302a miR-302b miR-302c miR-302d
16	miR-15b miR-16
17	miR-215 miR-194
18	miR-422b miR-378
19	miR-34b miR-34c
20	miR-200a miR-200b miR-429
21	miR-106b miR-93 miR-25
22	miR-181a-2 miR-213 miR-181b
23	miR-493-5p miR-493-3p
24	miR-432 miR-136
25	miR-431 miR-433 miR-127 miR-432
26	miR-299-5p miR-299-3p miR-380-5p miR-380-3p
27	miR-323 miR-329
28	miR-376a-1 miR-376b miR-368 miR-376a-2 miR-376a-3
29	miR-381 miR-487b miR-539
30	miR-154 miR-134 miR-382
31	miR-485-5p miR-485-3p
32	miR-496 miR-377
33	miR-409-5p miR-409-3p miR-369-5p miR-369-3p miR-410
34	miR-451 miR-822
35	miR-132 miR-212
36	miR-181a miR-181b
37	miR-363 miR-19b miR-106a miR-20b miR-17-5p miR-18b
38	miR-450 miR-542-5p
39	miR-503 miR-424
40	miR-224 miR-452
41	miR-221 miR-222
42	miR-500-5p miR-362
43	miR-502 miR-500-3p
44	miR-98 let-7f
45	miR-421 miR-374
46	miR-191 miR-425
47	miR-143 miR-145
48	miR-127 miR-136
49	miR-323 miR-329
50	miR-29b miR-29c

这些簇的序列比对表明，一些 miRNA 簇是由同源的 miRNA 分子所组成，另外一些 miRNA 簇具有多个拷贝，其中 28 个簇含有同源的 microRNA 分子，18 个簇形成 8 个不同的同源簇群（见图 3.6）。

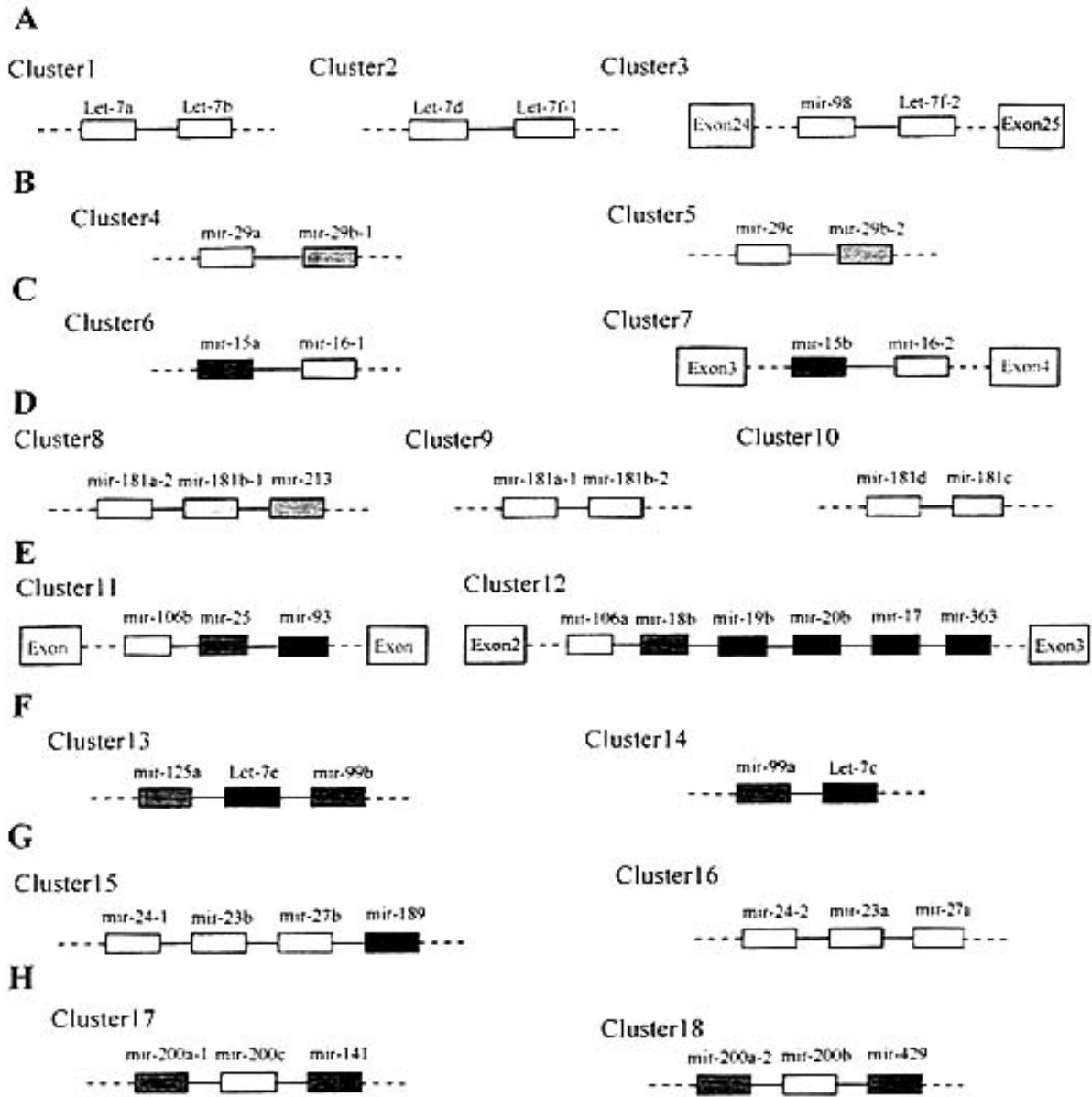


图 3.6 狗基因组中同源簇的基因组织形式。microRNA 基因用长方形表示，同源基因用相同的颜色表示。

Fig 3.6 Characterization of genomic organization of paralogous clusters in dog. miRNA genes are represented schematically by colored boxes. All are drawn to scale except the exons. Parologue of each miRNA gene in different cluster is denoted by rectangle with same color.

值得一提的是，我们在8号染色体上发现一个很大的保守区，这个区域在人和小鼠中被注释为印迹区。在这个区域中，存在两个 microRNA 簇、一个 box C/D snoRNA 基因簇以及几个蛋白质基因（见图 3.7）。其中一个较大的 miRNA 基因簇由 41 个 miRNA 基因组成，在这个簇中，通过序列分析，可以分成三个家族以及几个单拷贝的 miRNA 分子。这三个家族可能是通过串联重复产生的。另外一个 miRNA 基因簇则位于反转座子基因中，只是反向的，这预示着这个簇可能具有调节该反转座子的作用。

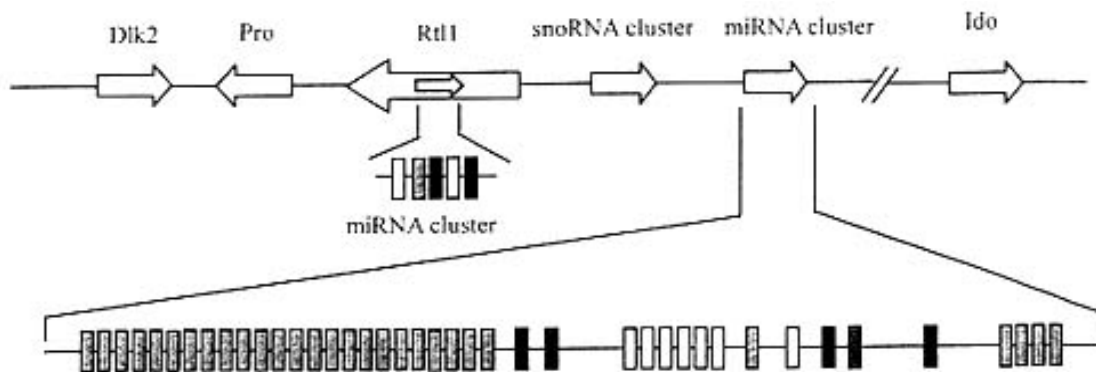


图 3.7 印迹区内最大的 microRNA 基因簇的基因组图谱。蛋白质编码基因用箭头表示，snoRNA 基因簇用箭头表示，而 microRNA 基因簇用有颜色的长方形表示。

Fig 3.7 Genomic map of the imprinted region holding the largest cluster. The position of several protein-coding genes is indicated by arrow, Rtl1, retrotransposon-like 1, vertical bars (small nucleolar RNAs). The largest cluster is denoted by colored rectangle, different color represent A-, B-,C-type family. These single copies are represented by different color.

植物当中，一些 microRNA 是来自重复元件或转座子中，所以我们在狗基因组中寻找是否也有 miRNA 分子存在于重复元件中。通过分析，我们鉴定出 13 个 miRNA 分子位于重复元件中，其中 7 个来自于长散在元件（Long interspersed nuclear element, LINE），其它则位于短散在元件（Short interspersed nuclear element, SINE）中。大多数与重复相关的 miRNA 分子的前体是通过两个相邻的 LINE 元件以反向的方式形成的。而 miR-325 是一个 LINE 元件和一个 SINE 元件反向形成的。有趣的是，miR-378 和 miR-422b 是分别由两个 SINE 元件形成的（见图 3.8）。同时我们也发现到，类似于果蝇 miR-308，狗的 miR-837 位于线粒体 28S 核糖体蛋白的第一个内含子内，同时小鼠的同源序列也位于核糖体蛋白的内含子中。

生。这个簇在除了节足动物门没有找到同源物之外，在其它较高等的动物中都有发现（见图 3.9）。

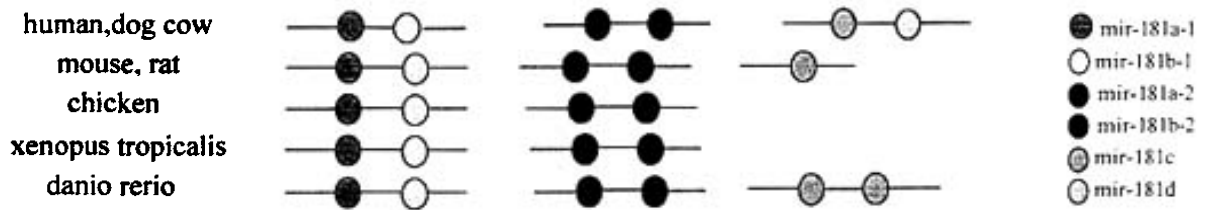


图 3.9 miR-181 基因家族的进化历史。该家族具有 6 个成员，形成 3 个基因簇。在斑马鱼中，具有 miR-181c 的两个拷贝，而在其它物种中只有一个拷贝。

Fig 3.9 Evolution history of the mir-181 family. in human, there are six copies, grouped in three cluster. In zebrafish, there are two copy of mir-18c., which is absent in amphibian and aves.

miR-144 簇是比较古老的，它是由 miR-144 和 miR-451 组成的，但是在线虫中却没有发现。而且我们发现 miR-144 和 miR-451 没有序列相似性，所以它们的进化枝分离的比较远（见图 3.10）。

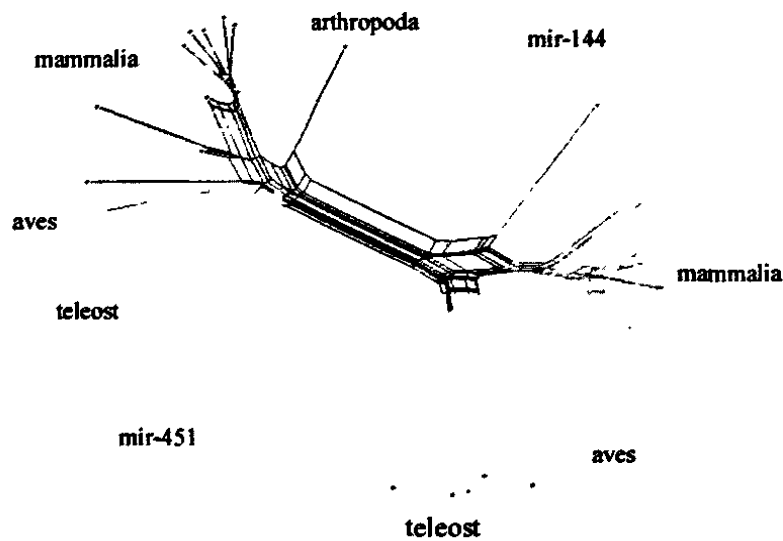


图 3.10 miR-144 基因簇的进化网络。

Fig 3.10 Phylogenetic network of mir-144 cluster.

miR-141 簇是由同源的 miR-141 和 miR-200 组成的，它是通过串联重复产生的。通过分析发现，这个簇是先于脊索动物的产生而产生的，并且有两个拷贝，然而我们并没在线虫或节肢动物中找到该簇。

miR-34 家族在较高等动物中具有三个成员，而在节足动物中只有一个。进化网络表明这个家族在进化中形成四个独立的分支，但是 miR-34a、miR-34b 和 miR-34c 关系比较近，这说明 miR-34 是比较古老的（见图 3.11）。

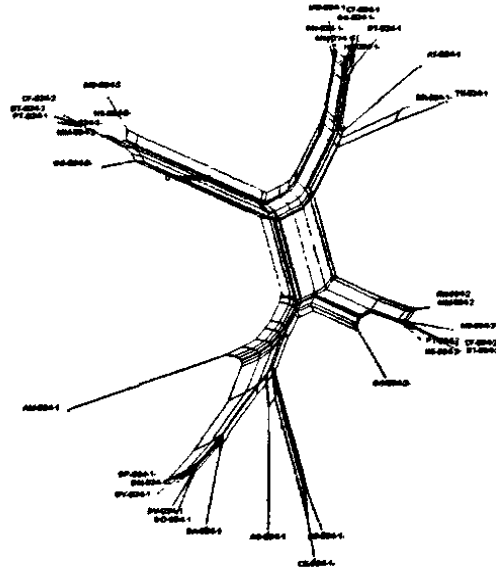


图 3.11 miR-34 家族的进化树分析。miR-34 家族在进化过程中可以明显的分为四个分支。

Fig 3.11 Gene tree of mir-34 cluster. mir-34-1 seems to evolve rapidly. This cluster form four distinct clades during evolution.

在很多情况下，我们并不能只通过序列信息和位置来研究一个簇的进化历史。例如，miR-23 簇是由 miR-23、miR-24 和 miR-27 组成的，它们之间是没有序列相关的。四足动物具有两个簇的拷贝，分别由三个 miRNA 基因组成，而斑马鱼具有五个簇的拷贝，河豚鱼具有四个拷贝，它们一般位于不同的染色体上。通过分析，我们推测出 miR-23 簇两种不同的进化历史。第一种进化历史，首先在硬骨鱼中由 miR-24 和 miR-23a 或者 miR-23b 组成的簇经过一个重复，产生簇的两个拷贝，然后随着进化，在脊椎动物和斑马鱼中获得了 miR-27，而河豚鱼却丢失了由三个基因组成的基因簇的一个拷贝。同时两个成员的簇在四足动物中丢失了。第二种进化历史，在脊椎动物中，存在 miR-23a 或 miR-23b、miR-24 和 miR-27 组成的基因簇，经过重复产生两个拷贝，而后在硬骨鱼中 miR-27 丢失了（见图 3.12）。

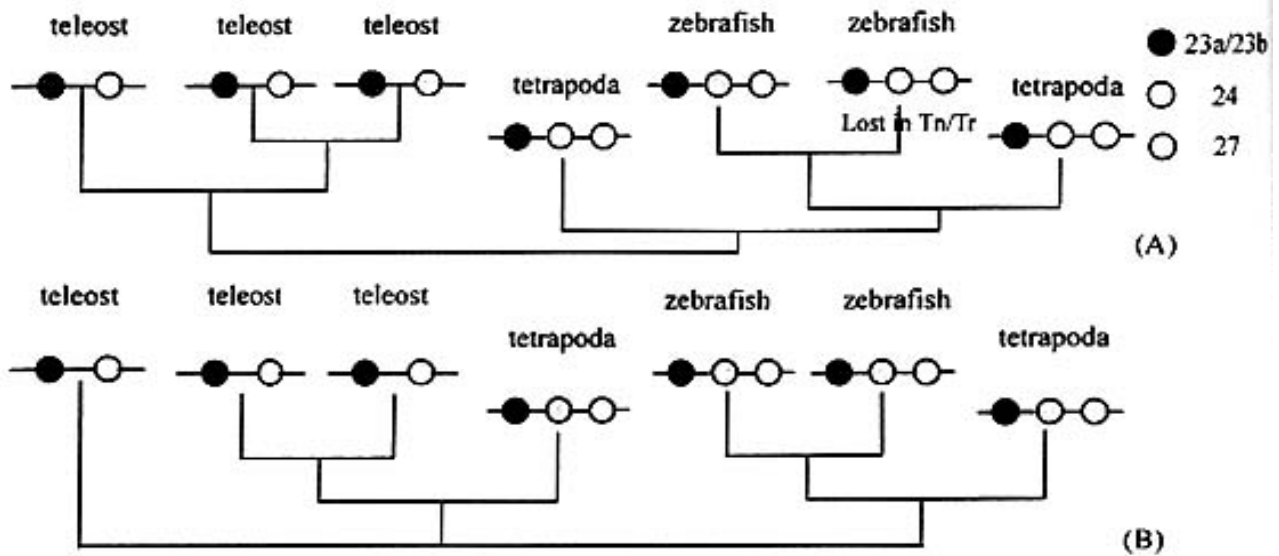


图 3.12 miR-23 基因簇的两种可能的进化路径。

Fig.3.12 The two most possible constructions for the history of the mir-23 cluster.

在分析中我们发现，在进化过程中，有些 miRNA 分子在哺乳动物中缺失。miR-133 簇就是一个例子。这个家族在脊椎动物中很早就通过串联重复产生。在斑马鱼中，它有 7 个拷贝，而在人中只有 3 个拷贝。

3.5 microRNA 靶序列的鉴定

利用鉴定的 miRNA 序列，我们在狗基因组中筛选蛋白质编码基因的 3' 非翻译区进而寻找潜在的靶位点，最终找到了 865 个靶位点。

表 3.4 microRNA 基因的靶序列

Table 3.4 Target gene of identified microRNA

miRNA	Score	MFE	Target gene
let-7a	120	-20.52	thyroid stimulating hormone receptor
let-7a	121	-20.93	leucine rich repeat containing 6 (testis)
let-7a	123	-27.22	retinaldehyde binding protein 1
let-7a	123	-22.3	max interactor 1
let-7a	124	-24.01	collagen alpha 2(iv) chain precursor
let-7a	126	-27.33	transmembrane protein 43
let-7a	126	-26.04	chromodomain helicase dna binding protein 9
let-7c	120	-24.85	deoxyguanosine kinase
let-7c	123	-27.22	zinc finger, cchc domain containing 11
let-7c	127	-25.43	g1/s-specific cyclin d3
let-7d	121	-24.24	similar to cg9667-pa
let-7d	122	-28.74	baculoviral iap repeat-containing protein 6

let-7d	123	-23.92	rgs19-interacting protein 1
let-7d	126	-23.71	hypothetical protein xp_850502
let-7e	120	-27.14	tRNA-splicing endonuclease subunit sen34
let-7e	120	-23.71	importin alpha-2 subunit
let-7e	121	-25.32	30 kda protein
let-7e	122	-26.66	f10b5.9
let-7e	127	-25.89	ring finger protein 137
let-7e	131	-30.63	upf0120 protein dkfzp564c186
let-7f	120	-22.52	netrin-g1 ligand
let-7f	121	-22.52	immunoglobulin superfamily, member 4d
let-7f	122	-27.19	interferon regulatory factor 2 binding protein 1
let-7g	120	-24.86	coatomer protein complex, subunit beta 2
let-7g	121	-26.75	breast cancer metastasis-suppressor 1-like
mir-1	120	-20.69	probable ma-dependent helicase p68
mir-7	120	-22.38	bax inhibitor-1
mir-7	121	-23.97	de-etiolated 1
mir-7	121	-23.51	cg14536-pa
mir-7	122	-23.04	hypothetical protein xp_861067
mir-7	122	-21.32	5-3 exoribonuclease 2
mir-7	122	-21.32	y48b6a.3
mir-7	122	-22.27	threonine synthase-like 1
mir-7	125	-21.28	upstream binding protein 1
mir-7	126	-21.46	hypothetical protein xp_533386
mir-7	126	-22.08	probable histone acetyltransferase myst1
mir-7	129	-25.5	nuclear autoantigen
mir-9	120	-22.16	annexin a5
mir-9	122	-25.88	juxtaposed with another zinc finger gene 1
mir-9	123	-20.09	cop9 (constitutive photomorphogenic)homolog, subunit 2
mir-9	123	-21.27	max interactor 1
mir-9	128	-24.03	pyruvate dehydrogenase [lipoamide]kinase isozyme 1, mitochondrial precursor
mir-9	131	-27.04	zinc finger ccch type containing 11a
mir-9	134	-27.42	b cell lymphoma/leukemia 10
mir-10a	124	-29.58	serine/threonine kinase 35
mir-10b	120	-26.87	nei-like 2
mir-15b	120	-21.98	activin receptor type ii precursor
mir-15b	122	-23.79	dual specificity mitogen-activated protein kinase kinase 1
mir-15b	123	-22.02	importin 8
mir-15b	133	-26.61	chromatin modifying protein 1
mir-16	122	-26.52	activin receptor type ii precursor (actr-ii)
mir-16	124	-25.9	son protein
mir-17-3p	123	-25.96	lanc-like protein 2
mir-17-5p	123	-20.84	angiopoietin-like 7

mir-17-5p	124	-21.49	voltage-dependent anion channel 2
mir-17-5p	126	-23	cofilin 2
mir-17-5p	128	-21.66	frizzled 3 precursor
mir-17-5p	128	-24.72	gamma-aminobutyric-acid de receptor beta-3 subunit precursor
mir-17-5p	129	-22.42	sorting nexin 12
mir-17-5p	130	-23.75	dipeptidyl-peptidase i precursor
mir-17-5p	131	-26.92	proteasome 26s atpase subunit 6
mir-17-5p	131	-23.51	ring finger protein 127
mir-17-5p	138	-29.23	tripartite motif protein 8
mir-18a	123	-20.13	na/cl-dependent taurine transporter
mir-18a	128	-25.04	intersectin 1
mir-18a	129	-23.75	neurobeachin
mir-20a	120	-21.69	similar to jagunal homolog 1
mir-20a	121	-21.13	sorting nexin 12
mir-20a	122	-22.82	proteasome 26s atpase subunit 6
mir-20a	122	-21.95	hla-b associated transcript 4
mir-20a	123	-20.9	heterogeneous nuclear ribonucleoprotein l-like
mir-20a	124	-23.61	tripartite motif protein 8
mir-20a	136	-22.03	hypothetical protein xp_547102
mir-22	120	-24.26	hypothetical protein xp_849994
mir-22	126	-25.18	kh domain containing, ma binding, signal transduction associated 3
mir-22	127	-24.3	ubiquitin conjugation factor e4 a
mir-23a	123	-22.9	neurofilament, light polypeptide
mir-23a	125	-22.17	regulator of g-protein signalling 12
mir-24	120	-28.8	calreticulin
mir-24	121	-26.84	nuclear factor 1 a-type
mir-24	122	-24.02	procollagen, type v, alpha 1
mir-24	126	-28.37	wilms tumour 1-associating protein
mir-24	127	-26.24	lamin b receptor
mir-25	120	-21.13	mitochondrial ribosomal protein l49
mir-25	123	-28.32	pyruvate dehydrogenase [lipoamide]]kinase isozyme 4, mitochondrial precursor
mir-25	128	-26.51	heat shock protein 8
mir-25	128	-23.16	nuclear factor 1 b-type
mir-27a	125	-23.6	mitochondrial import receptor subunit tom22 homolog
mir-27a	129	-25.29	ret proto-oncogene
mir-28	120	-21.92	cg8253-pa
mir-28	122	-23.2	trna-splicing endonuclease subunit sen34
mir-28	122	-27.06	cell division cycle 10
mir-28	122	-24.77	junctionophilin 1
mir-28	123	-21.37	thyroid stimulating hormone receptor
mir-28	123	-26.93	gpi-linked nad(p)(+)-arginine adp-ribosyltransferase 1 precursor

第3章 结果

mir-28	125	-21.79	h3 histone, family 3b
mir-28	125	-28.35	lipidosin
mir-28	129	-24.87	cg9987-pa
mir-28	130	-27.68	h-rev107-like protein 5
mir-29a	121	-23.08	transcriptional regulator atrx
mir-29a	125	-22.49	cerebral endothelial cell adhesion molecule 1
mir-30a-3p	120	-32.68	protein c20orf36
mir-30a-3p	120	-26.89	recombining binding protein suppressor of hairless
mir-30a-3p	129	-30.26	islet cell autoantigen 1
mir-30a-5p	122	-28.46	collagen alpha 1(iv) chain precursor
mir-30a-5p	123	-22.21	pappalysin-2 precursor
mir-30a-5p	125	-20.59	hypothetical protein xp_536264
mir-30b	120	-24.76	ras-related protein rab-18
mir-30b	123	-27.84	proprotein convertase subtilisin/kexin type 5 precursor
mir-30b	127	-20.35	stromal antigen 2
mir-30b	141	-30.27	embryonic ectoderm development
mir-30c	130	-26.06	heat-shock protein 105 kda
mir-30d	120	-20.58	zinc finger, matrin type 5
mir-30d	122	-28.46	collagen alpha 1(iv) chain precursor
mir-30d	128	-23.66	hypothetical protein xp_536264
mir-31	121	-23.19	p53-inducible cell-survival factor
mir-31	123	-23	myristoylated alanine-rich c-kinase substrate
mir-31	124	-25.29	ankyrin 2
mir-31	126	-26.88	protocadherin 9
mir-31	130	-29.85	peroxisomal membrane protein 11b
mir-31	135	-30.31	apolipoprotein c-ii precursor
mir-32	121	-23.87	cytoplasmic polyadenylation element binding protein 2
mir-32	121	-22.37	hypothetical protein xp_858448
mir-34a	120	-24.75	senataxin
mir-34a	121	-29.09	pral family protein 2
mir-34a	122	-26.2	na ⁺ -dependent glutamate transporter
mir-34a	122	-25.45	target of egr1, member 1
mir-34a	123	-25.5	cop9 constitutive photomorphogenic homolog subunit 7b
mir-34a	123	-24.99	heterogeneous nuclear ribonucleoprotein r
mir-34a	124	-27.96	non-pou domain containing, octamer-binding
mir-34a	124	-27.69	ras p21 protein activator 3
mir-34a	125	-29.62	cg11430-pb
mir-34a	126	-22.02	ankyrin 2
mir-34a	131	-26.67	phosphoglycerate kinase 1
mir-34a	137	-30.56	heat shock 70kda protein 14
mir-92	120	-28.2	tumor protein d54
mir-92	124	-21.6	hypothetical protein xp_854863
mir-93	120	-21.08	cg4203-pa
mir-93	122	-21.13	enhancer of polycomb homolog 2

mir-93	122	-26.42	fln29 gene product
mir-93	123	-25.14	coated vesicle membrane protein
mir-93	124	-20.42	necdin-like 2
mir-93	126	-27.57	emsy protein
mir-93	127	-26.97	rim binding protein 2
mir-95	120	-23.48	surviving
mir-95	135	-22.16	antiquitin
mir-96	121	-26.11	dna repair protein rad51 homolog 4
mir-96	122	-21.12	rhophilin, rho gtpase binding protein 2
mir-96	127	-20.83	atp-binding cassette, sub-family a member 1
mir-96	127	-24.49	muskelin 1, intracellular mediator containing kelch motifs
mir-98	120	-24.01	pre-b-cell leukemia transcription factor 3
mir-98	122	-24.73	wd repeat and socs box containing protein 2
mir-98	122	-22.85	collagen alpha 2(iv) chain precursor
mir-98	123	-20.09	lim domain kinase 2
mir-98	123	-21.42	cg11137-pa
mir-98	123	-22.99	mixed lineage kinase-related kinase mrk-beta
mir-98	124	-22.04	transmembrane protein 41b
mir-98	124	-21.14	endothelin receptor type a
mir-98	126	-20.62	max interactor 1
mir-98	126	-24.22	rab14, member ras oncogene family
mir-98	127	-21.42	pre-b-cell leukemia transcription factor 3
mir-98	135	-21.33	mitogen-activated protein kinase kinase kinase kinase 3
mir-101	120	-25.49	snare protein ykt6
mir-101	122	-22.18	paralemmin 2
mir-101	127	-22.96	archain
mir-101	127	-26.61	mad homolog 2
mir-101	128	-21.5	hiv-1 rev binding protein
mir-103	120	-21.5	thyroid hormone receptor-associated protein complex 240 kda component
mir-103	120	-22.2	b7-2 protein
mir-103	121	-24.58	tissue inhibitor of metalloproteinase 3
mir-103	121	-22.09	set and mynd domain containing protein 2
mir-103	122	-24.37	peroxisomal biogenesis factor 19
mir-103	122	-25.48	cg6903-pa
mir-103	123	-23.85	muscleblind-like 1
mir-103	123	-22.88	cg2924-pc
mir-103	123	-22.35	mos polyadenylation response element
mir-103	126	-23.25	cullin-5
mir-103	128	-21.03	k+ channel tetramerization protein
mir-103	130	-26.5	interferon-induced 35 kda protein
mir-106b	120	-21.69	jagunal homolog 1
mir-106b	121	-25.14	ns1-associated protein 1
mir-106b	121	-21.13	sorting nexin 12

mir-106b	122	-22.82	proteasome 26s atpase subunit 6
mir-106b	123	-20.9	heterogeneous nuclear ribonucleoprotein l-like
mir-106b	136	-22.03	hypothetical protein xp_547102
mir-107	120	-22.8	endosulfine alpha
mir-107	120	-21.04	rab geranylgeranyltransferase, beta subunit
mir-107	121	-21.63	myotubularin related protein 4
mir-107	123	-26.16	protein kinase, cgmp-dependent, type i
mir-107	124	-25.91	peroxisomal biogenesis factor 19
mir-107	125	-26.15	glycerol kinase
mir-107	125	-24.92	interferon-induced 35 kda protein
mir-10a	124	-29.58	serine/threonine kinase 35
mir-10b	120	-26.87	nei-like 2
mir-122a	120	-26.75	hypothetical protein xp_847170
mir-122a	120	-25.39	protein tyrosine phosphatase,receptor type, g precursor
mir-122a	121	-21.98	dickkopf related protein-3 precursor
mir-122a	123	-27.14	start domain containing 7
mir-122a	123	-24.89	glucose-6-phosphate 1-dehydrogenase
mir-122a	134	-26.98	sprouty-related protein 1 with evh-1 domain
mir-125a	128	-33.98	ataxin 2-binding protein 1
mir-128a	121	-26.62	serine/threonine protein phosphatase 2a, 55 kda regulatory subunit b
mir-128a	121	-23.5	emsy protein
mir-128a	122	-23.22	leucine-rich repeat transmembrane neuronal protein 3 precursor
mir-128a	126	-25.95	sar1a gene homolog 1
mir-129	120	-20.85	26s proteasome non-atpase regulatory subunit 7
mir-129	131	-31.28	werner helicase interacting protein
mir-130a	121	-21.14	yeast ino80-like protein
mir-130a	121	-20.77	putative homeodomain transcription factor 2
mir-130a	122	-22.08	ubiquitin-conjugating enzyme e2 d2
mir-130a	133	-24.71	trinucleotide repeat containing 6a
mir-130a	125	-22.38	sec14 and spectrin domains 1
mir-132	120	-25.33	solute carrier family 25, member 28
mir-132	120	-24.09	calcium channel, voltage-dependent,beta 4 subunit
mir-132	121	-25.79	transmembrane emp24 protein transport domain containing 9
mir-132	127	-28.16	e74-like factor 2 (ets domain transcription factor)
mir133a	121	-27.89	c34c12.4
mir133a	121	-30.23	r3h domain protein kiaa1002
mir133a	121	-28.95	testis-specific serine kinase 4
mir133a	122	-32.59	opioid receptor, sigma 1
mir133a	124	-32	actin-related protein 2/3 complex subunit 1b
mir133a	128	-31.96	swi/snf-related de matrix-associated actin-dependent regulator of chromatin d1

mir-134	120	-26	sterile alpha motif domain containing 9-like
mir-134	125	-25.9	phosphoinositide-3-kinase, class 2, alpha polypeptide
mir-135a	121	-20.12	hypothetical protein xp_846956
mir-135a	122	-26.57	transcription factor spl
mir-135a	123	-20.06	eh domain binding protein 1
mir-135a	123	-23.25	fas-associated factor 1
mir-135a	123	-23.75	ring finger protein 37
mir-135a	124	-24.67	high-mobility group 20a
mir-135a	125	-23.19	lutropin-choriogonadotropic hormone receptor precursor
mir-135a	126	-23.21	acetyl-coa synthetase 2-like
mir-135a	128	-23.3	ae binding protein 2
mir-135a	129	-21.91	potassium channel tetramerisation domain containing 10
mir-135a	140	-26.25	dna-(apurinic or apyrimidinic site)lyase
mir-136	121	-21.9	prefoldin 5
mir-136	121	-20.31	protein phosphatase 1, regulatory (inhibitor) subunit 14c
mir-136	122	-24.85	prostaglandin-endoperoxide synthase 1
mir-136	126	-20.79	myosin regulatory light polypeptide 9
mir-136	129	-23.16	forkhead box o3a
mir-137	124	-20.21	horma domain containing 1
mir-140	121	-24.2	guanine nucleotide-binding protein g(i)/g(s)/g(o) gamma-3 subunit
mir-140	129	-22.48	cytokine receptor-like factor 3
mir-141	120	-22.68	stromal cell-derived factor 2 precursor (sdf-2)
mir-141	121	-21.11	transcriptional regulator atrx
mir-141	122	-20.31	phytanoyl-coa hydroxylase interacting protein-like
mir-141	122	-23.79	hypothetical protein xp_848099
mir-141	125	-24.33	tbc1 domain family, member 15
mir-141	127	-22.08	amyloid beta a4 precursor protein-binding family b member 2
mir-141	127	-20.31	chromodomain helicase dna binding protein 9
mir-141	128	-25.38	tyrosine-protein kinase jak1
mir-141	129	-25.19	splicing factor u2af 35 kda subunit
mir-141	129	-27.6	polyadenylate-binding protein 4 (poly(a)-binding protein 4
mir-142-3p	132	-27.14	splicing factor, arginine/serine-rich 10
mir-142-3p	128	-21.11	serologically defined colon cancer antigen 1
mir-142-3p	127	-23.54	RNA binding motif protein 25
mir-142-3p	124	-24.57	Autophagy protein 16-like
mir-142-3p	122	-20.62	hypothetical protein XP_849242
mir-142-3p	122	-21.79	scavenger receptor class B, member 1
mir-142-3p	122	-22.87	interleukin-21
mir-142-3p	121	-23.43	Sad1/unc-84 protein-like 1
mir-142-3p	121	-20.96	sorting nexin 3
mir-142-3p	121	-23.93	mitochondrial ribosomal protein L18

mir-143	122	-26.9	nuclear receptor subfamily 1, group h, member 4
mir-143	123	-23.09	retinoblastoma binding protein 7
mir-143	123	-25.44	ubiquitin-conjugating enzyme e2 variant 2
mir-143	127	-25.86	heterogeneous nuclear ribonucleoprotein a2/b1
mir-144	120	-22.21	probable rna-dependent helicase p68
mir-144	120	-20.51	zinc finger x-chromosomal protein
mir-144	122	-21.8	cg12341-pa
mir-144	123	-22.08	polyhomeotic like 3
mir-144	123	-20.32	adaptor-related protein complex ap-1,sigma 3
mir-144	124	-20.52	zinc finger protein 650
mir-144	131	-25.35	pyruvate dehydrogenase [lipoamide]]kinase isozyme 4, mitochondrial precursor
mir-145	122	-33.46	chromosome x open reading frame 40
mir-145	126	-30.75	camp responsive element binding protein 3-like 4
mir-145	131	-32.77	desmoglein-1
mir-145	142	-29.98	tyrosine-protein kinase jak 1
mir-146a	121	-23.38	cg11986-pa
mir-146a	122	-20.75	vacuolar protein sorting 54
mir-146a	123	-21.87	actin-binding lim protein 1
mir-146a	128	-21.05	cell cycle progression 2 protein
mir-148a	120	-25.64	guanylate cyclase activator 1a
mir-148a	121	-20.11	activated rna polymerase ii transcriptional coactivator p15
mir-148a	124	-27.52	far upstream element binding protein 1
mir-148a	133	-30.07	lymphocyte cytosolic protein 2
mir-149	120	-29.59	hypothetical protein xp_536264
mir-149	120	-29.79	nucleolar gtp-binding protein 1
mir-149	120	-30.41	ubiquinol-cytochrome c reductase subunit
mir-149	122	-30.86	heterogeneous nuclear ribonucleoprotein u
mir-149	122	-36.01	paraneoplastic antigen ma1
mir-149	122	-29.88	zinc finger protein 289, id1 regulated
mir-149	131	-37.61	cg7550-pa
mir-150	120	-30.27	adipocyte complement related protein of 30 kda
mir-150	121	-36.29	heterogeneous nuclear ribonucleoprotein a0
mir-150	121	-31.96	hypothetical protein xp_533819
mir-150	122	-26.9	origin recognition complex, subunit 2
mir-150	123	-34.86	hypothetical protein xp_845035
mir-150	126	-31.81	inner centromere protein
mir-150	129	-33.88	ubiquitin specific protease,proto-oncogene
mir-150	130	-30.78	follicle-stimulating hormone-related protein 1 precursor
mir-151	122	-28.06	olfactomedin-like 2a
mir-151	123	-20.39	rna binding motif protein 5
mir-152	121	-26.19	tar dna binding protein
mir-152	121	-23.4	glycoprotein m6a
mir-152	122	-31.2	k562 cell-derived leucine-zipper-like protein 1

mir-152	122	-28.85	protein c5orf5
mir-152	124	-24.7	interleukin-13 receptor alpha-1 chain precursor
mir-152	133	-32.69	josephin 1
mir-181a	121	-20.01	intersectin 2
mir-181a	123	-20.99	dnaj (hsp40) homolog, subfamily b, member 14
mir-181a	126	-26.94	malonyl coa-acyl carrier protein transacylase, mitochondrial precursor
mir-181a	131	-23.65	ccr4-not transcription complex, subunit 4
mir-181b	121	-24.37	huntingtin interacting protein b
mir-181b	122	-25.87	f26e4.12
mir-181b	123	-25.44	clathrin heavy chain
mir-181b	124	-22.72	ariadne-1 protein homolog
mir-181b	126	-27.25	atp-binding cassette, sub-family a member 1
mir-181b	127	-27.43	glutamate receptor, metabotropic 1
mir-181b	129	-29.67	cylindromatosis (turban tumor syndrome)
mir-181c	124	-25	ornithine carbamoyltransferase precursor
mir-181c	126	-25.99	lin-7 homolog c
mir-181c	129	-25.36	y47d9a.2a
mir-181d	136	-29.55	huntingtin interacting protein b
mir-181d	130	-27.55	atp-binding cassette, sub-family a member 1
mir-181d	130	-28.9	cylindromatosis (turban tumor syndrome)
mir-181d	127	-25.82	solute carrier family 2, facilitated glucose transporter, member 8
mir-181d	127	-24.53	ns1-associated protein 1
mir-181d	126	-25.87	glutamate receptor, metabotropic 1
mir-181d	124	-24.03	nuclear receptor co-repressor/hdac3 complex subunit
mir-181d	124	-26.04	cg10286-pa
mir-181d	122	-24.09	testis specific, 10
mir-181d	120	-25.2	inner nuclear membrane protein man1
mir-182	120	-23.21	bag-family molecular chaperone regulator-4
mir-182	121	-25.43	basic helix-loop-helix domain containing, class b3
mir-182	122	-20.23	intersectin 2
mir-182	123	-22.63	juxtaposed with another zinc finger gene 1
mir-182	128	-27.45	brain-derived neurotrophic factor
mir-183	120	-29.26	ariadne ubiquitin-conjugating enzyme e2 binding protein homolog 1
mir-183	121	-22.64	fatty aldehyde dehydrogenase
mir-183	132	-24.48	wee1-like protein kinase
mir-183	133	-28.2	thyroid hormone receptor associated protein 2
mir-184	120	-24.89	bromodomain containing 8
mir-184	121	-23.44	cg6903-pa
mir-184	121	-24.19	mothers against decapentaplegic homolog 5
mir-184	121	-25.26	nipsnap1 protein
mir-184	122	-23.72	sh3 domain containing ring finger 2

mir-184	122	-30.93	ephrin receptor ephb2 precursor
mir-184	125	-27.19	glycerol-3-phosphate dehydrogenase 1-like
mir-184	128	-22.32	pabp1-dependent poly a-specific ribonuclease subunit pan3
mir-186	120	-22.56	cg11796-pa
mir-186	121	-23.55	tpr domain, ankyrin-repeat and coiled-coil-containing
mir-186	122	-24.39	tsc22 domain family 3
mir-186	122	-23.37	protein c20orf103 precursor
mir-186	123	-20.55	actin-binding lim protein 1
mir-186	123	-20.82	casein kinase ii alpha 1 subunit
mir-186	124	-21.16	homeodomain-interacting protein kinase 3
mir-186	125	-22.51	armadillo repeat-containing protein
mir-186	125	-25.05	origin recognition complex subunit 5
mir-186	127	-24.02	zinc finger 265
mir-186	131	-20.75	bcl-2-associated transcription factor 1
mir-186	136	-25.59	ribosome biogenesis protein bms1 homolog
mir-188	120	-25.39	sporulation-induced transcript 4-associated protein
mir-188	120	-29.67	galactocerebrosidase
mir-188	121	-24.32	enhancer of polycomb 1
mir-188	122	-27.98	myeloid/lymphoid or mixed-lineage leukemia protein 2
mir-188	130	-29.76	hypothetical protein xp_854965
mir-188	135	-39.57	adenosine a1 receptor
mir-188	146	-39.67	protein-l-isoaspartate(d-aspartate)o-methyltransferase
mir-189	120	-24.09	implantation-associated protein
mir-189	120	-22.5	rna binding protein 5
mir-189	124	-28.52	positive cofactor 2,glutamine/q-rich-associated protein
mir-189	124	-24.97	protein kinase, amp-activated, alpha 1 catalytic subunit
mir-189	138	-31.38	leupaxin
mir-18a	123	-20.13	na/cl-dependent taurine transporter
mir-18a	128	-25.04	intersectin 1
mir-18a	129	-23.75	neurobeachin
mir-190	122	-22.47	probable atp-dependent rna helicase ddx6
mir-190	130	-20.18	alkyldihydroxyacetonephosphate synthase, peroxisomal precursor
mir-191	124	-24.92	tubulin, alpha 1
mir-191	124	-21.52	paired related homeobox 1 lipoamide acyltransferase component of branched-chain
mir-192	120	-25.98	alpha-keto acid dehydrogenase complex,mitochondrial precursor
mir-192	121	-20.68	f33h2.2
mir-192	124	-22.07	periphilin 1
mir-192	128	-28.27	dna-directed rna polymerase ii 140 kda polypeptide
mir-194	122	-22.67	cg14182-pa
mir-194	136	-26.39	flotillin-1

mir-195	121	-24.59	rab-interacting factor
mir-195	125	-21.01	phosphoserine aminotransferase
mir-195	127	-23.95	hydroxymethylglutaryl-coa synthase,mitochondrial precursor
mir-195	130	-28.44	atpase, h+ transporting, v1 subunit b
mir-195	130	-22.92	semaphorin 6d precursor
mir-196a	120	-21.31	retinoblastoma-binding protein 1
mir-196a	121	-22.15	mitochondrial tumor suppressor 1
mir-196a	122	-21.82	upf2 regulator of nonsense transcripts homolog
mir-196a	123	-24.28	hypothetical protein xp_854965
mir-196a	125	-23.66	solute carrier family 4, sodium bicarbonate transporter-like, member 10
mir-196a	126	-23.64	splicing factor, arginine/serine-rich 1 (asf/sf2)
mir-196a	131	-20.57	sprouty-related protein 1 with evh-1 domain
mir-196a	134	-22.18	cug triplet repeat, ma binding protein 2
mir-197	122	-38.08	chromodomain helicase dna binding protein 3
mir-197	126	-33.88	nucleosome assembly protein 1-like 4
mir-197	127	-32.8	latrophilin 1 precursor
mir-197	129	-33.31	cg3625-pb
mir-197	129	-34.11	pumilio homolog 2
mir-197	133	-39.22	exportin-7 (ran-binding protein 16)
mir-197	144	-35.81	rab5b, member ras oncogene family
mir-199a	123	-28	zinc finger protein 91
mir-199a	126	-36.97	heavy neurofilament protein
mir-199a	127	-26.02	ma binding motif protein 18
mir-199a	128	-28.27	ribophorin ii precursor
mir-199a	130	-32.95	nad(p)h dehydrogenase [quinone] 1(quinone reductase 1)
mir-199a	132	-23.68	udp-glucose dehydrogenase
mir-200a	122	-25.88	polyadenylate-binding protein 4
mir-200a	123	-25.03	stromal cell-derived factor 2 precursor (sdf-2)
mir-200a	123	-22.3	hypothetical protein xp_848099
mir-200a	124	-22.99	phytanoyl-coa hydroxylase interacting protein-like
mir-200a	125	-20.86	amyloid beta a4 precursor protein-binding family b member 2
mir-200a	125	-22.65	calpastatin
mir-200a	127	-25.67	splicing factor u2af 35 kda subunit
mir-202	120	-23.33	non imprinted in prader-willi/angelman syndrome 2
mir-202	120	-24.59	zinc finger and btb domain containing 10
mir-202	121	-25.59	tight junction protein 1
mir-202	122	-25.18	cg9005-pa
mir-202	125	-28.77	cd19 antigen
mir-204	120	-26.73	peroxisome proliferator activated receptor gamma coactivator 1 alpha
mir-204	121	-28.14	hypothetical protein xp_533238

mir-204	121	-23.12	hypothetical protein xp_863615
mir-204	125	-30.57	atp synthase, h ⁺ transporting, mitochondrial f0 complex, subunit c precursor
mir-204	126	-29.08	f-box only protein 32
mir-204	127	-23.03	signal transducer and activator of transcription 5b
mir-204	127	-28.79	heat shock protein hsp 90-beta
mir-205	120	-26.38	xenotropic and polytropic retrovirus receptor
mir-205	120	-25.39	transforming growth factor beta regulated gene 1 (predicted)
mir-205	121	-31.35	probable rna-dependent helicase p68 (dead-box protein p68) (dead-box protein 5)
mir-205	121	-33.16	adp-ribosylation factor 2
mir-205	122	-27.11	cytochrome b561
mir-205	122	-31.15	probable low-affinity copper uptake protein 2
mir-205	123	-30.63	myozenin 1
mir-205	126	-29.92	leucine zipper and cttnbip1 domain containing
mir-205	128	-35.52	gdp-l-fucose synthetase
mir-205	129	-31.28	guanine nucleotide binding protein-like 3 (nucleolar)-like
mir-205	130	-29.54	tatd dnase domain containing 2
mir-205	141	-33.69	chloride channel 3
mir-206	121	-20.76	hypothetical protein xp_532944
mir-206	121	-20.38	carbonic anhydrase-related protein 10
mir-206	122	-23.08	peripheral plasma membrane protein cask
mir-206	123	-21.65	kinesin family member 23
mir-206	125	-24.24	cg15160-pa
mir-206	129	-24.14	vesicle amine transport protein 1
mir-208	123	-21.7	hypothetical protein XP_537341
mir-211	121	-30.42	swi/snf related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 3
mir-211	121	-28.13	hypothetical protein xp_533238
mir-211	121	-24.31	digeorge syndrome criticalde region protein 6
mir-211	122	-23.39	signal transducer and activator of transcription 5b
mir-211	122	-32.14	hypothetical protein xp_854863
mir-211	123	-27.56	cyclin fold protein 1
mir-211	123	-32.1	hmba-inducible
mir-211	123	-32.85	heterogeneous nuclear ribonucleoprotein u
mir-211	125	-30.13	peroxisome proliferator activated receptor gamma coactivator 1 alpha
mir-212	121	-33.33	liprin-beta 2
mir-212	122	-27.11	egfr-coamplified and overexpressed protein

mir-212	125	-22.71	hypothetical protein xp_536835
mir-214	120	-27.22	camp-dependent protein kinase inhibitor alpha
mir-214	120	-25	olfactomedin related er localized protein
mir-214	120	-27.59	gy-box
mir-214	121	-31.69	g protein-coupled receptor 157
mir-214	122	-26.99	calcium/calmodulin-dependent protein kinase ii gamma
mir-214	123	-26.55	malonyl coa-acyl carrier protein transacylase, mitochondrial precursor
mir-214	124	-27.34	skeletal muscle and kidney enriched inositol phosphatase
mir-214	124	-29.65	ubiquinol-cytochrome c reductase subunit
mir-214	128	-27.33	dihydropteridine reductase
mir-214	128	-32.19	karyopherin (importin) alpha 3
mir-216	120	-20.36	ankyrin repeat and soxs box-containing 9
mir-216	123	-25.52	epoxide hydrolase 2, cytoplasmic
mir-216	127	-24.89	5-amp-activated protein kinase,gamma-1 subunit
mir-216	131	-25.13	comm domain containing de protein 8
mir-217	120	-23.13	rab5 interacting protein 2
mir-217	121	-21.67	glutathione s-transferase a4
mir-217	123	-27.72	tousled-like kinase 2
mir-217	124	-27.36	afg3-like protein 2
mir-217	127	-28.51	spermatid perinuclear rna-binding protein
mir-217	128	-25.98	malic enzyme 3, nadp(+)-dependent,mitochondrial
mir-220	120	-29.01	solute carrier family 7
mir-210	125	-31.21	mam domain containing glycosylphosphatidylinositol anchor 1
mir-221	120	-24.91	ubiquitin-conjugating enzyme e2d 3(ubc4/5 homolog, yeast)
mir-221	122	-22	eukaryotic translation elongation factor 1 alpha 1
mir-221	123	-21.71	hypothetical protein xp_534674
mir-221	125	-21.28	cg17593-pa
mir-221	127	-33.1	dead-box protein 3, x-chromosomal
mir-221	128	-30	f54c1.5a
mir-221	128	-22.75	eukaryotic translation elongation factor 1 alpha 1
mir-222	120	-31.75	trna-splicing endonuclease subunit sen34
mir-222	125	-30.56	dna replication licensing factor mcm6
mir-299-3p	122	-24.08	catenin (cadherin-associated protein), beta 1, 88kda
mir-222	134	-34.22	leptin receptor gene-related protein
mir-302a	125	-23.22	interleukin 8
mir-302a	127	-20.76	ccr4-not transcription complex subunit 4
mir-302a	127	-21.82	plasma paf acetylhydrolase
mir-302a	128	-20.91	cg6171-pa
mir-302a	129	-26.52	acyl-coa synthetase long-chain family member 1
mir-302a	130	-29.76	hematopoietic antigen cd38
mir-302a	131	-21.94	cg17652-pa

mir-302a	132	-22.32	calmodulin 1
mir-302a	132	-21.65	membrane associated dna binding protein
mir-302a	134	-25.77	son protein
mir-324	121	-31.29	golgi reassembly stacking protein 1
mir-324	122	-33.23	f-box only protein 31
mir-324	124	-34.14	sterol carrier protein 2
mir-324	126	-34.08	6-phosphofructokinase, muscle type
mir-324	127	-31.09	transmembrane protein 9 precursor
mir-324	127	-38.41	interleukin-13 receptor alpha-1 chain precursor
mir-324	132	-36.3	matrix metalloproteinase 2 preproprotein
mir-328	121	-30.68	semaphorin 4b precursor
mir-328	122	-31.01	cerebellin 1 precursor protein
mir-328	123	-35.39	pdgfra associated protein 1
mir-328	123	-37.29	hypothetical protein xp_863626
mir-328	123	-35.28	vesicle amine transport protein 1
mir-328	124	-31.1	ariadne-2 protein homolog (ari-2)(triad1 protein)
mir-328	124	-33.16	rab11b, member ras oncogene family
mir-328	124	-35.4	mitochondrial import receptor subunit tom22 homolog
mir-328	126	-35	nucleobindin 1 precursor (calnuc)
mir-328	126	-37.67	heart- and neural crest derivatives-expressed protein 1
mir-328	126	-35.28	membralin
mir-328	126	-37.45	glucose-6-phosphate 1-dehydrogenase
mir-328	127	-40.83	paxillin
mir-328	127	-38.59	cg7289-pa
mir-328	130	-35.48	atp synthase mitochondrial f1 complex assembly factor 2
mir-328	132	-41.92	sodium- and chloride-dependent transporter xtrp2
mir-331	122	-32.58	protein c20orf22
mir-331	123	-30.76	e2f transcription factor 7
mir-338	120	-30.26	pou domain, class 2, transcription factor 2
mir-338	122	-27.53	cellular nucleic acid binding protein 1
mir-338	123	-24.98	ns1-associated protein 1
mir-338	125	-27.86	family with sequence similarity 32,member a
mir-338	125	-22.83	proto-oncogene protein c-fos
mir-338	126	-24.18	chromosome 22 open reading frame 13
mir-338	127	-26.42	arl-6 interacting protein-1 (aip-1)
mir-338	128	-23.22	pdz and lim domain 5 isoform enh2
mir-338	130	-25.78	hypothetical protein xp_851492
mir-338	130	-29.33	l-lactate dehydrogenase a chain
mir-338	132	-25.41	septin-2
mir-338	135	-27.39	rab14, member ras oncogene family
mir-339	120	-33.91	px19-like protein
mir-339	120	-31.57	tetracycline transporter-like protein

mir-339	120	-27.42	hepatoma-derived growth factor
mir-339	122	-29.35	hla-b associated transcript 2
mir-339	122	-29.6	riken cdna d130038b21
mir-339	127	-35.68	short transient receptor potential channel 3
mir-339	127	-34.42	mucin-1 precursor
mir-339	127	-37.25	isocitrate dehydrogenase 3 (nad ⁺)alpha
mir-339	130	-34.61	serum amyloid a-3 protein precursor
mir-339	130	-32.08	electron transfer flavoprotein-ubiquinone oxidoreductase, mitochondrial precursor
mir-339	136	-38.92	riken cdna 2010005j08
mir-339	138	-37.86	gpi-linked nad(p)(+)-arginine adp-ribosyltransferase 1 precursor
mir-342	123	-24.34	serine (or cysteine) proteinase inhibitor, clade b (ovalbumin), member 1
mir-342	125	-29.11	cortactin
mir-342	130	-27.05	mof4 family associated protein 1
mir-345	123	-27.2	potassium inwardly-rectifying channel j15
mir-345	124	-24.56	ccr4-not transcription complex,subunit 6-like
mir-346	120	-30.44	24-dehydrocholesterol reductase precursor
mir-346	121	-30.26	rna-binding protein 4
mir-346	122	-31.23	isocitrate dehydrogenase [nad]subunit gamma, mitochondrial precursor
mir-346	123	-33.67	dickkopf related protein-3 precursor
mir-346	125	-38.74	malonyl coa-acyl carrier protein transacylase, mitochondrial precursor
mir-346	127	-33.08	transcription factor coe4
mir-346	139	-42.55	transcription factor mafb
mir-362	121	-21.15	nucleoporin 88
mir-362	123	-25.69	adenylate cyclase, type vii
mir-362	123	-31.47	peroxisomal acyl-coenzyme a thioester hydrolase 1
mir-362	124	-26.8	bromodomain and wd repeat domain containing 1
mir-362	126	-29.2	nad(p) transhydrogenase,mitochondrial precursor
mir-362	127	-25.79	coagulation factor xiii a chain precursor
mir-362	131	-23.41	rna binding motif protein 22
mir-367	120	-22.94	dihydrouridine synthase 2-like
mir-367	124	-22.69	rhomboid, veinlet-like 7
mir-367	125	-25.02	cell growth regulator with ring finger domain 1
mir-367	127	-23.31	calnexin
mir-367	129	-23.46	glycoprotein m6a
mir-368	120	-20.65	grpe protein homolog 1, mitochondrial precursor
mir-368	121	-26.62	peroxisomal membrane protein pmp34
mir-368	124	-22.68	riken cdna 0610040d20
mir-368	125	-24.39	ubiquilin 1
mir-368	127	-24.74	signal transducing adaptor molecule 1

mir-368	131	-22.89	melanoma-associated antigen h1
mir-368	131	-24.88	rna binding motif protein 14
mir-370	122	-33.89	microfibrillar-associated protein 4
mir-370	123	-32.93	kelch-like 21
mir-370	124	-34.69	cg6903-pa
mir-370	124	-33.71	interferon stimulated exonuclease gene 20-like 2
mir-370	128	-34.47	igf-ii mrna-binding protein 1
mir-370	129	-31.96	hypothetical protein xp_858504
mir-370	134	-36.26	protein c2orf17 homolog
mir-371	121	-24.01	complement component c8 gamma chain precursor
mir-371	121	-23.61	cd36 antigen
mir-371	129	-24.97	nad(p) transhydrogenase,mitochondrial precursor
mir-371	130	-28.12	whsc111 protein
mir-375	121	-24.14	cytoplasmic polyadenylation element binding protein 4
mir-375	121	-21.65	neuromodulin
mir-375	122	-25.02	oxysterol binding protein-related protein 6
mir-375	129	-30.23	proteinase activated receptor 3 precursor
mir-376	122	-25.22	heme binding protein 1
mir-377	123	-20.08	dual specificity mitogen-activated protein kinase kinase 4
mir-378	120	-35.34	drebrin 1
mir-378	121	-29.27	egl nine (c.elegans) homolog 2
mir-378	121	-29.22	ribosome receptor
mir-378	121	-31.24	sorting nexin-1
mir-378	122	-31.99	pyruvate dehydrogenase [lipoamide]]kinase
mir-378	123	-30.18	rna-binding protein raly
mir-378	123	-31.86	hypothetical protein xp_849102
mir-378	124	-34.16	snrpn upstream reading frame protein
mir-378	127	-36.29	lysosomal alpha-glucosidase precursor
mir-378	127	-35.04	arp1 actin-related protein 1 homolog a,contractin alpha
mir-378	129	-34.8	mitochondrial translational release factor 1-like
mir-381	120	-27.64	protein kinase c, beta type
mir-381	121	-23.32	cg12713-pa
mir-381	124	-23.98	cytoplasmic polyadenylation element binding protein 4
mir-381	126	-26.63	tyrosine-protein kinase receptor tyro3 precursor
mir-382	120	-24.36	ankyrin repeat domain 28
mir-382	121	-22.66	mannosyl-oligosaccharide 1,2-alpha-mannosidase ic
mir-382	122	-23.99	sparc related modular calcium-binding protein 1 precursor
mir-382	124	-25.15	histidyl-trna synthetase
mir-382	124	-31.34	centrin 2
mir-382	129	-28.88	serta domain-containing protein 1
mir-383	120	-21.58	l-plastin

mir-383	120	-20.35	ribosomal protein l37
mir-383	120	-22.2	copine-6
mir-383	120	-25.98	round spermatid basic protein 1
mir-383	121	-20.51	c-type lectin, superfamily member 14
mir-383	121	-22.4	lipoamide acyltransferase component of branched-chain alpha-keto acid dehydrogenase complex,mitochondrial precursor
mir-383	122	-21.02	g-protein coupled receptor family c group 5 member b precursor
mir-383	122	-21.55	ankyrin 2
mir-383	124	-26.64	v-rel reticuloendotheliosis viral oncogene homolog a
mir-383	125	-23.45	phd finger protein 20-like 1
mir-383	127	-26.55	protocadherin 9 precursor
mir-383	130	-29.03	thioredoxin, mitochondrial precursor
mir-383	131	-27.13	cell division control protein 2 homolog
mir-383	133	-30.03	mitogen-activated protein kinase 1
mir-422a	120	-29.91	hypothetical protein xp_849105
mir-422a	120	-22.09	peptidyl-prolyl cis-trans isomerase like 1
mir-422a	121	-35.34	tax1
mir-422a	123	-30.19	dual specificity mitogen-activated protein kinase kinase 1
mir-422a	126	-36.6	calcium/calmodulin-dependent protein kinase ii gamma
mir-422a	128	-34.91	24-dehydrocholesterol reductase precursor
mir-422a	129	-33.02	solute carrier family 26, member 6
mir-422a	130	-31.06	fus interacting protein
mir-422a	132	-32.81	hypothetical protein xp_858504
mir-422a	133	-35.72	diabetes related ankyrin repeat protein
mir-423	120	-32.4	neuron specific calcium-binding protein hippocalcin
mir-424	120	-24.29	karyopherin (importin) alpha 3
mir-424	120	-21.13	retinoblastoma-associated protein 140
mir-424	122	-26.91	serine/threonine kinase 19
mir-424	122	-21.01	ubiquitin carboxyl-terminal hydrolase 7
mir-424	122	-23.41	tyrosine-protein phosphatase,non-receptor type 3
mir-424	122	-23.14	f16a11.1
mir-424	128	-27.5	tata box binding protein-like protein 1
mir-424	129	-22.5	carbonic anhydrase xiv precursor
mir-424	131	-22.49	ubiquitin protein ligase e3a
mir-424	131	-23.14	u7 snrna-associated sm-like protein lsm11
mir-425	123	-30.96	zinc transporter like 2
mir-431	122	-27.87	dolichyldiphosphatase 1
mir-431	130	-30.46	heat-shock protein beta-7
mir-432	120	-26.62	cd164 antigen, sialomucin
mir-432	120	-32.78	lectin, galactoside-binding, soluble, 9

mir-432	120	-25.49	chloride channel protein 6
mir-432	121	-22.56	transmembrane protein 41b
mir-432	121	-23.49	p53-inducible cell-survival factor
mir-432	122	-26.01	peptidyl prolyl isomerase h
mir-432	122	-28.72	cell death activator cide-b
mir-432	123	-26.57	dead-box protein 3, x-chromosomal
mir-432	124	-28.96	transcription factor ap-2 beta
mir-432	124	-33.4	calpain-3 (calpain 13) (calpain p94)
mir-432	126	-30.04	mesoderm development candidate 2
mir-432	127	-36.21	peripherin
mir-432	129	-34.88	putative transcription factor znf131
mir-432	131	-28.81	ef-hand calcium binding domain 2
mir-448	120	-21.24	m-phase phosphoprotein 6
mir-448	122	-22.02	vacuolar protein sorting 35
mir-448	122	-24.29	rab8b, member ras oncogene family
mir-448	123	-20.87	protein fam34a
mir-448	123	-21.62	sh3-domain kinase binding protein 1
mir-448	124	-20.97	calcitonin receptor-stimulating peptide-1
mir-448	125	-24.36	zinc finger protein 521
mir-448	125	-22.42	a disintegrin and metalloproteinase domain 22 preproprotein
mir-448	126	-22.4	ets variant gene 5
mir-448	127	-27.52	delta-like protein 4 precursor
mir-449	120	-23.2	protein phosphatase 1, regulatory (inhibitor) subunit 2
mir-449	120	-26.61	erythrocyte protein band 4.1-like 1
mir-449	120	-23.49	heat-shock protein 105 kda
mir-449	121	-25.36	mannosyl (alpha-1,3)-glycoprotein beta-1,4-n-acetylglucosaminyltransferase
mir-449	121	-22.81	mitogen-activated protein kinase 10
mir-449	123	-22.99	transmembrane protein 22
mir-449	124	-22.67	cg5020-pa
mir-449	124	-23.37	dead box polypeptide 17
mir-449	125	-29	heat shock 70kda protein 14
mir-449	129	-26.15	electron transfer flavoprotein-ubiquinone oxidoreductase, mitochondrial precursor
mir-450	123	-24.58	g protein-activated inward rectifier potassium channel 1
mir-450	123	-20.16	small edrk-rich factor 2
mir-450	126	-20.44	adp-ribosylation factor-like 2 binding protein
mir-450	141	-21.15	fibulin 2 precursor
mir-451	121	-21.9	kinesin family member 1c
mir-452	121	-21.22	connective tissue growth factor
mir-452	121	-22.86	mammalian ependymin related protein-1 precursor (merp-1)
mir-452	121	-20.64	elav-like 2

第 3 章 结果

mir-452	122	-26.14	sodium channel, voltage-gated, type ix, alpha
mir-452	124	-24.48	erythrocyte protein band 4.1-like 1
mir-452	125	-23.17	lysosomal pro-x carboxypeptidase precursor
mir-452	125	-20.91	mhc class i dla-64
mir-452	126	-26.04	adenosine kinase
mir-452	129	-26.26	x-box binding protein 1
mir-452	130	-28.07	renal sodium-dependent phosphate transport protein 2
mir-452	130	-26.99	rab6 interacting protein 1
mir-452	130	-23.42	round spermatid basic protein 1
mir-452	131	-26.32	breast cancer metastasis-suppressor 1-like
mir-455	121	-23.95	ovarian tumor locus protein
mir-455	124	-25.18	phosphatidylinositol glycan, class s
mir-483	120	-36.8	importin alpha-7 subunit
mir-483	120	-33.19	orm1-like 3
mir-483	121	-34.6	inositol-trisphosphate 3-kinase a
mir-483	122	-31.34	chromodomain helicase dna binding protein 4
mir-483	122	-37.04	cyclin-dependent kinase 4 inhibitor d
mir-483	123	-39.1	ce10 protein
mir-483	123	-33.8	v-ets erythroblastosis virus e26 oncogene like
mir-483	125	-37.2	zinc finger swim domain containing protein 6
mir-483	127	-36.3	hypothetical protein xp_844816
mir-483	127	-33.97	glutaredoxin (thioltransferase)(ttase)
mir-483	128	-37.4	tankyrase, trf1-interacting ankyrin-related adp-ribose polymerase 2
mir-483	134	-38.23	adenylate kinase 2
mir-485-5p	120	-24.69	peroxiredoxin 1
mir-485-5p	121	-21.73	solute carrier family 15 (h+/peptide transporter), member 2
mir-485-5p	124	-25.61	vacuolar protein sorting 33a
mir-485-5p	124	-30.69	active breakpoint cluster region-related protein
mir-483	136	-41.64	zinc finger protein 659
mir-486	121	-37.84	structure specific recognition protein 1
mir-486	125	-26.6	regulatory subunit pr 53 of protein phosphatase 2a
mir-487a	124	-20.05	zinc finger protein 292
mir-487a	127	-27.2	thyroid hormone receptor-associated protein complex 100 kda component
mir-488	120	-24.45	calumenin precursor
mir-488	121	-27.24	hypothetical protein xp_848566
mir-488	121	-23.62	hypothetical protein xp_851516
mir-488	121	-31.18	pdgfa associated protein 1
mir-488	121	-23.24	chromatin modifying protein 5

第 3 章 结果

mir-488	122	-21.81	band 4.1-like protein 3
mir-488	127	-25.75	heterogeneous nuclear ribonucleoprotein k
mir-489	120	-25.19	talin 2
mir-489	122	-22.3	upf0287 protein dc13
mir-489	123	-23.13	t12a2.1
mir-489	123	-21.13	rna binding protein 5
mir-489	124	-25.03	myocyte-specific enhancer factor 2a
mir-489	125	-34.3	beta-adrenergic receptor kinase 1
mir-489	135	-24.5	sorting nexin 3
mir-490	120	-24.39	voltage-dependent anion-selective channel protein 1
mir-490	120	-25.77	toll interacting protein
mir-490	120	-29.42	membralin
mir-490	122	-29.75	tousled-like kinase 2
mir-490	130	-31.3	cg11178-pb
mir-490	133	-35.41	cg7351-pa
mir-491	120	-21.65	afg3-like protein 2
mir-491	120	-21.43	restin-like 2
mir-491	121	-29.88	syntaxin 4a (placental)
mir-491	121	-32.18	myosin-10
mir-491	124	-24.88	probable ubiquitin ligase protein mycbp2
mir-491	124	-31.54	latrophilin 1 precursor
mir-491	130	-32.59	f-box only protein 32
miR-493-3p	125	-25.2	vasohibin-like protein
mir-493-5p	124	-26.67	low density lipoprotein receptor-related protein 6
mir-493-5p	131	-22.85	zinc finger protein 198
mir-493-5p	134	-28.83	seryl-trna synthetase
mir-491	130	-31.68	serine/threonine-protein kinase wnk4
mir-494	120	-21.1	signal transducer and activator of transcription 5b
mir-494	121	-25.16	h2a histone family, member v
mir-494	123	-21.89	crk-like protein
mir-494	124	-25.01	thymus high mobility group box protein tox
mir-494	125	-24.71	phosphoglycerate mutase 1
mir-494	130	-22.47	ns1-associated protein 1
mir-494	130	-24.57	tetratricopeptide repeat domain 14
mir-494	132	-24.73	zinc finger cch-type domain containing 7
mir-495	121	-21.09	tankyrase, trf1-interacting ankyrin-related adp-ribose polymerase 2
mir-495	122	-21.9	ankyrin repeat and socs box-containing protein 2 (predicted)
mir-495	126	-22.08	praja 2, ring-h2 motif containing
mir-495	126	-23.54	fk506-binding protein 2 precursor
mir-495	128	-22.62	riken cdna 1810036i24
mir-495	128	-24.89	cysteine and histidine-rich domain (chord)-containing, zinc-binding protein 1
mir-498	120	-27.89	prkca-binding protein

第 3 章 结果

mir-498	120	-24.25	g patch domain containing protein 3
mir-498	121	-23.51	cg4452-pa
mir-498	121	-27.69	cg14591-pa
mir-498	122	-20.85	n-acetylgalactosaminyltransferase 7
mir-498	122	-26.15	cg9005-pa
mir-498	123	-23.81	cg3040-pa
mir-498	123	-23.62	rna-binding protein with multiple splicing
mir-498	123	-20.76	secretogranin-1 precursor (secretogranin i) (sgi)
mir-498	123	-24.86	carboxypeptidase a2 precursor
mir-498	124	-24.21	cdc14 homolog a
mir-498	124	-27.14	glycerol-3-phosphate acyltransferase,mitochondrial precursor
mir-498	125	-27.66	calsyntenin 1
mir-498	125	-30.36	calcium channel, voltage-dependent, alpha 2/delta 3 subunit
mir-498	126	-23.98	nascent polypeptide-associated complex alpha polypeptide
mir-498	126	-21.92	potassium channel tetramerisation domain containing 9
mir-498	127	-31.61	diacylglycerol kinase, delta 130kda
mir-498	127	-25.18	cg7593-pa
mir-498	128	-27.48	protein phosphatase 1, regulatory (inhibitor) subunit 14c
mir-498	130	-26.64	eukaryotic translation initiation factor 4 gamma 2
mir-498	131	-26.65	hypothetical protein xp_863623
mir-498	132	-28.33	zk546.17
mir-498	132	-28.07	mitogen-activated protein kinase 1
mir-498	137	-28.74	butyrate response factor 1
mir-498	137	-31.22	suppressor of actin mutations 2-like
mir-499	120	-20.42	zinc finger protein 106 homolog
mir-499	122	-22.8	hmg-box transcription factor bbx
mir-499	122	-22.8	hmg-box transcription factor bbx
mir-499	123	-22.39	plasma membrane calcium atpase 1
mir-499	123	-22.39	plasma membrane calcium atpase 1
mir-499	123	-22.39	plasma membrane calcium atpase 1
mir-499	123	-22.39	plasma membrane calcium atpase 1
mir-499	128	-20.84	adp-ribosylation factor 2
mir-499	128	-22.02	wd repeat domain 55
mir-499	128	-20.84	adp-ribosylation factor 2
mir-499	128	-20.84	adp-ribosylation factor 2
mir-499	129	-23.61	ephrin receptor epha7
mir-499	131	-22.29	hypothetical protein xp_536264
mir-499	131	-21.98	thyroid hormone receptor-associated protein complex 240 kda component
mir-500	121	-25.32	nadp-dependent malic enzyme (nadp-me)
mir-500	122	-24.4	protein kinase, d1 type
mir-500	123	-28.27	plastin 3
mir-500	123	-28.27	plastin 3

mir-500	123	-28.27	plastin 3
mir-500	123	-28.27	plastin 3
mir-500	123	-28.27	plastin 3
mir-500	123	-28.27	plastin 3
mir-500	123	-28.27	plastin 3
mir-500	127	-27.29	capping protein (actin filament) muscle z-line, alpha 1
mir-500	127	-22.59	ubiquitin-conjugating enzyme e2r 2
mir-500	127	-27.29	capping protein (actin filament) muscle z-line, alpha 1
mir-500	130	-34	tbc1 domain family, member 10b
mir-502	120	-29.32	leucine zipper and cttnbip1 domain containing
mir-502	120	-28.18	class i cytokine receptor
mir-502	121	-27.27	palmitoyl-protein thioesterase 2 precursor
mir-502	122	-28.02	cg33084-pb
mir-502	122	-28.7	ribosomal protein s6 kinase polypeptide 3
mir-502	123	-27.27	thioredoxin, mitochondrial precursor (mt-trx) (mtrx) (thioredoxin-2)
mir-502	123	-24.53	glutathione-requiring prostaglandin d synthase
mir-502	123	-29.03	hypothetical protein xp_847396
mir-502	124	-25.28	disabled homolog 2 (differentially expressed protein 2) (doc-2)
mir-502	126	-27.57	inositol 1,4,5-triphosphate receptor,type 1
mir-502	126	-30.69	peptidylprolyl isomerase-like 2
mir-502	126	-29.24	membralin
mir-502	131	-31.16	wd and tetratricopeptide repeats protein 1
mir-502	133	-29.94	rna-binding protein 6
mir-503	121	-27.27	palmitoyl-protein thioesterase 2 precursor
mir-503	122	-28.7	ribosomal protein s6 kinase polypeptide 3
mir-503	123	-27.27	thioredoxin, mitochondrial precursor
mir-503	123	-24.53	glutathione-requiring prostaglandin d synthase
mir-503	126	-29.24	membralin
mir-503	131	-31.16	wd and tetratricopeptide repeats protein 1
mir-503	133	-29.94	rna-binding protein 6
mir-504	120	-27	sh3 domain-binding protein 5
mir-504	122	-29.15	hypothetical protein xp_843757
mir-504	122	-29.15	hypothetical protein xp_852264
mir-504	124	-27.71	immunoglobulin superfamily containing leucine-rich repeat
mir-504	130	-29.34	hypothetical protein xp_542329
mir-504	131	-30.85	nucleosome assembly protein 1-like 4
mir-505	120	-22.44	procollagen c-endopeptidase enhancer 2 precursor
mir-505	123	-21.99	protein c2orf17 homolog
mir-505	123	-27.57	karp-1-binding protein
mir-505	124	-22.7	atlastin-like
mir-505	124	-25.5	t-complex protein 1, eta subunit

第 3 章 结果

mir-505	124	-27.12	death effector domain-containing protein
mir-532	121	-28.15	phosphoglycerate kinase 1
mir-532	122	-27.23	rna-binding protein 4
mir-532	126	-32.03	g protein-coupled receptor associated sorting protein 1
mir-532	132	-30.85	secretogranin iii
mir-539	120	-23.74	er lumen protein retaining receptor 2 (kdel receptor 2)
mir-539	122	-22.29	cortactin
mir-539	135	-24.96	tripartite motif protein 8
mir-544	124	-23.75	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit
mir-544	124	-23.75	guanylate cyclase e
mir-545	128	-20.83	hypothetical protein xp_539021
mir-545	133	-20.86	peroxisomal sarcosine oxidase
mir-560	127	-28.58	cg8580-pa
mir-560	127	-31.41	pyruvate dehydrogenase (lipoamide)beta
mir-560	135	-36.49	2-aminoadipic 6-semialdehyde dehydrogenase
mir-560	129	-31.65	cytochrome c oxidase polypeptide viii-liver, mitochondrial precursor
mir-574	120	-24.85	myocyte-specific enhancer factor 2a
mir-574	122	-27.1	serum response factor (srf)
mir-574	124	-30.74	alpha-endosulfine
mir-574	126	-27.05	cysteine rich transmembrane bmp regulator 1
mir-574	126	-28.06	1-acyl-sn-glycerol-3-phosphate acyltransferase delta
mir-574	126	-25.78	nuclear respiratory factor-1
mir-574	132	-30.86	synaptotagmin-1
mir-574	128	-33.64	prosaposin
mir-582	121	-21.36	transmembrane protein 41b
mir-582	130	-20.44	lyric/3d3
mir-582	132	-22.15	hect domain containing 1
mir-582	134	-24.89	thymus high mobility group box protein tox
mir-592	120	-20.81	hematopoietic stem/progenitor cells 176
mir-592	124	-21.9	twinfilin
mir-652	125	-25.44	ankyrin repeat domain 13c
mir-592	133	-22.16	synaptotagmin xiv
mir-660	120	-29.22	elongation factor 1-gamma (ef-1-gamma) (eef-1b gamma)
mir-660	122	-22.94	aftiphilin protein

大多数靶序列只有一个 miRNA 序列与之互补，而少数靶序列有多个 microRNA 和它互补，调控该基因的表达。（见图 3.13）

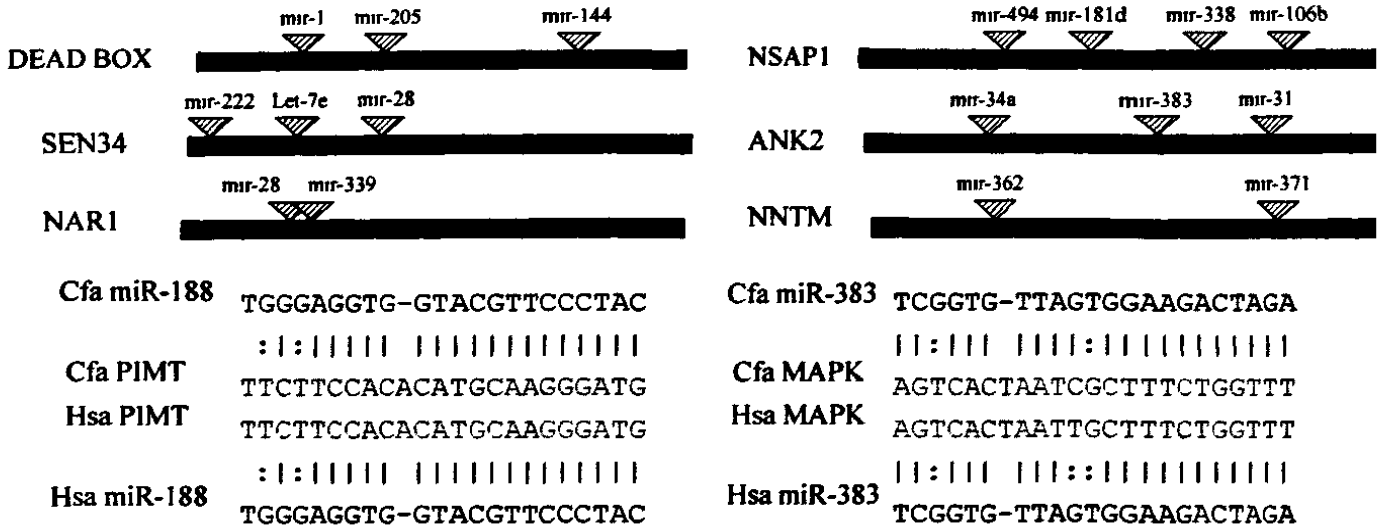


图 3.13 microRNA 作用的一些靶位点。倒三角形表示 microRNA 作用的位置。

Fig 3.13 Potential miRNA target Sites in the UTRs of Selected Genes. The positions of target sites for specific miRNAs (triangles above rectangles) are, in general, distributed nonuniformly. Detailed alignments between the miRNA and the predicted target sites (arbitrary selection) illustrate, in general, stronger match density at the 5' end of miRNAs than at the 3' end, as required by the algorithm and as observed in experimentally validated targets. The non-conserved nucleotides in the target sites are highlighted in gray.

有趣的是，通过计算机寻找，我们找到了一些靶序列与 miRNA 能够完全互补配对，表明这些 miRNA 能够指导 mRNA 的切割（见图 3.14）。



图 3.14 一些 microRNA 能够与靶序列完全互补。通常只有一个错配。

Fig 3.14 Target transcript have nearly perfect complementarity with miRNAs. Only one mismatch occurred

我们利用 GO 数据库来分析 miRNA 的靶序列的功能分类以及分布，结果表明靶基因参与了几乎全部的生物功能（见图 3.15）。参与功能最多的是蛋白质结合活性，这表明 miRNA 作为转录因子调节子而发挥调节作用。

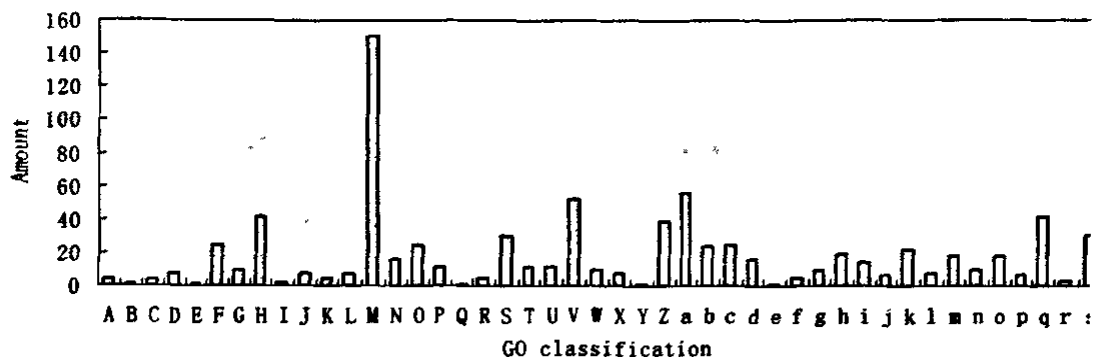


图 3.15 靶基因的功能分类

Fig 3.15 Distribution of target gene according to GO term. The most enriched GO term is ubiquitin-protein ligase activity. Some GO term also have been found in dog genome.

我们研究发现 miR-98、miR-442a/miR-449 和 miR-27a 能够协同调节一个信号转导通路——MAP 信号传导路径。他们分别在效应的不同阶段调节相应的激酶。miR-98 调节 MAPKKK，而 miR-442a/miR-449 调节 MAPKK，最后 miR-27a 调节 MAPK，最终影响效应的传递(见图 3.16)。

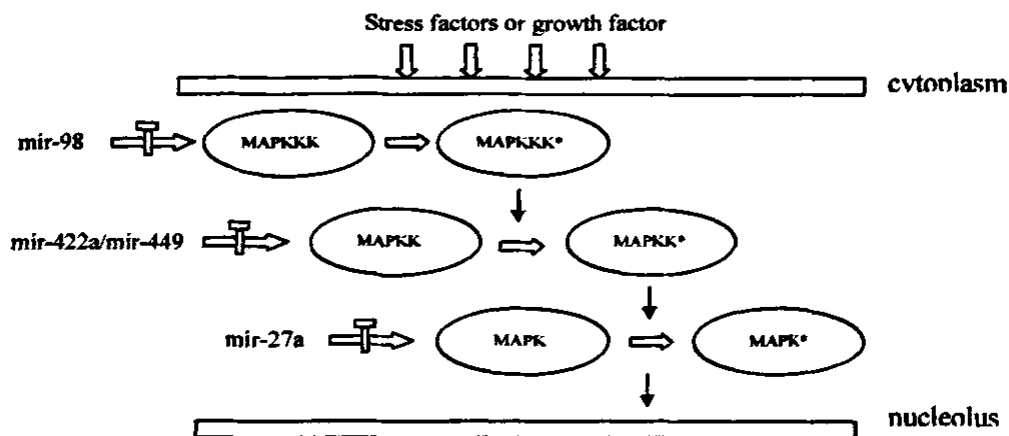


图 3.16 相应的 microRNA 分子能够在 MAP 信号转导途径中调节不同的激酶。

Fig 3.16 The corresponding microRNA regulate different kinase in MAP signal transduction pathway. Members of this family activate each other by adding phosphate groups to serine/threonine amino acids.

一些 miRNA 分子的靶序列可能与癌症发生有关。miR-16 和 miR-15 能够和活化素受体基因以及 SON 基因完全匹配,而这两个基因在结肠癌中被下调表达。有些靶序列参与一些细胞过程,比如糖和脂肪代谢,泛素蛋白降解过程以及转运过程。miR-122a 作用于葡萄糖-6-磷酸脱氢酶以及酪氨酸磷酸化酶,miR-183 作用脂肪酸脱氢酶,miR-199a 调节 UDP 葡萄糖脱氢酶。同时我们发现 miR-107 能够调节溶酶体发生因子,而该基因参与毒素代谢过程。miR-25、miR-204 和 miR-130 能够调节热休克蛋白。非常有趣的是,miR-186 位于锌指蛋白 265 的第 8 个内含子中,以同一方向,而其能够与锌指蛋白 265 的 3' 非翻译区不完全互补。

值得一提的是,我们发现 7 个长度大约在 0.9kb 到 2.2kb 的靶序列,而它们与任何已知蛋白都没有同源性,表明它们是狗特异性的靶基因。

3.6 microRNA 的表达分析

对于本文所鉴定的 357 个 microRNA 分子,300 个是与人有高度同源的,而这些基因已经通过 Northern 杂交或者是 cDNA 克隆实验得到证实,它们可以稳定表达。而对于这 57 个新的 miRNA 序列,其中 14 个位于蛋白质基因的内含子中,可以随宿主基因的表达而表达。对于这些位于基因间隔区的 miRNA,我们通过 EST 数据库搜索,发现 7 个 EST 序列与 miRNA 匹配,表明他们稳定表达,还发现 2 个 miRNA 在多个组织中表达。对于剩余的 36 个 miRNA,我们采取 RT-PCR 的方法来鉴定是否表达,我们挑取二级结构比较好的 5 个 miRNA 进行 PCR 分析。

3.6.1 狗总RNA提取结果

根据热酸性酚法提取狗总 RNA,取待测 RNA 样品各 1 微升,用 1.2%甲醛变性凝胶电泳确定样品 RNA 质量。全部 RNA 样品贮藏于-20℃。如图 3.17 所示, RNA 分布为三条区带,前两条区带(分别对应于 25s, 18s rRNA)较亮,后一条区带(对应于 5.8s rRNA)较暗,且无明显拖尾,可知总 RNA 提取质量较好,大分子 RNA 被提出, RNA 降解少,利于进一步实验。

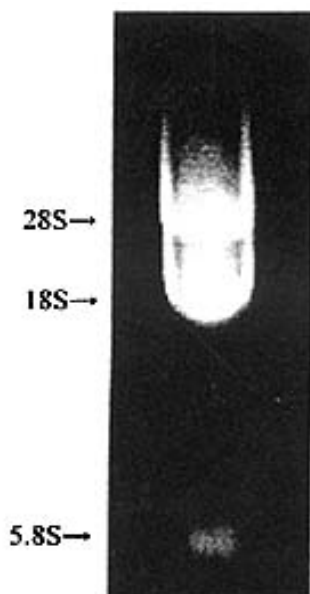


图 3.17 狗总 RNA 电泳图谱

Fig 3.17 extraction of RNA in dog

3.6.2 狗基因组DNA提取结果

根据玻璃珠法提取狗总 DNA，取待测 DNA 样品各 1 微升，用 1%琼脂糖凝胶电泳确定样品 DNA 碱基的大小，如图 3.18 所示，全部 DNA 样品贮藏于-20℃。



图 3.18 狗基因组电泳图谱

Fig 3.18 extraction of genomic DNA in dog

3.6.3 RT-PCR结果

分别用 miR-850 以及 miR-856 的 Pn1F/Pn1R 引物对狗总 RNA 进行 RT-PCR, 通过负对照检测总 RNA 中的 DNA 是否酶解充分, 通过正对照确定引物的特异性, 并结合其扩增条带判断 microRNA 的表达情况。图 3.19 为逆转录结果, RT 带只有一条且与正对照带分子量大小一致, 分别为 72nt 和 65nt。说明这两个 microRNA 能够在肝组织中稳定表达

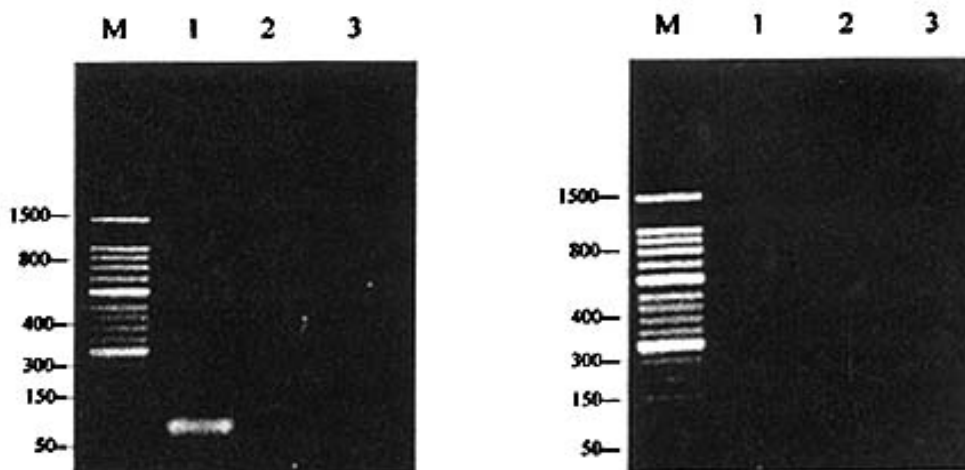


图 3.19 新的基因间隔区的 microRNA 的 RT-PCR 分析。

Fig 3.19 RT-PCR assay of novel intergenic miRNA gene. We predicted the length of miR-856 precursor, about 84nt, and the mature sequence is in 5' arm. After RT-PCR and gel electrophoresis, a fragment with about 80nt in length is obtained(left), the same conclusion holds for miR-850.miR-856 sequence: tgaggtagtaggtgtgtggt, forward primer: Pn1F: 5'-TgAggTAgTAggTTgTgTg-3', reverse primer: Pn1R: 5'-gAAggCAgTAggTTgTATAg-3'.miR-850 sequence: tctgtaactgtcaaagtgatt, forward primer: 5'-AAATgATgTgACAAAggC-3', reverse primer: AATCACTTTgACAgTTACAg M: marker; RT, RT-PCR product; -, RNA-guided PCR product; +, genomic DNA-guided PCR product

第4章 讨论

4.1 计算机RNA组学是鉴定microRNA快速、高效的方法

目前, 鉴定新的microRNA主要有两种途径: 计算机方法和实验方法。这两种方法分别称为计算机RNA组学和实验RNA组学, 它们各具特色, 相辅相成。实验RNA组学能够找到结构不典型的microRNA、具有时间和组织特异性表达的microRNA以及物种特异的microRNA, 结果可以体现microRNA表达丰度, 是鉴定microRNA的有效方法。但实验RNA组学却存在实验周期长, 一些表达丰度低的microRNA分子难以鉴定的缺陷。计算机RNA组学弥补了实验RNA组学的这些不足, 它是建立在大量基因组序列已经测定的基础上。它能快速、高效地找到表达丰度低以及目前未报道的microRNA, 结果往往可以体现microRNA的基因组织形式和表达调控等信息, 有利于microRNA数据的积累和系统分析。目前国内外已经开发出针对不同物种中预测microRNA的生物信息学软件, 并且都得到了很好的结果。由于基于microRNA保守性的比较基因组学方法不受microRNA序列特征的限制, 在搜索microRNA时表现出无可比拟的优越性^[69-70]。本文考虑到microRNA的序列特征和保守性等多方面的特点, 综合运用多种生物信息学方法, 充分利用公共基因数据库的丰富资源, 鉴定了357个新的microRNA, 极大地丰富了microRNA的数量, 并为进一步分析基因组织形式和进化规律提供了必要的数据基础。

我们这里采取两种生物信息学手段(同源性搜索和RNAz方法)来预测microRNA序列, 这两种方法各有其优缺点^[42-43]。同源性搜索需要已知的microRNA作为种子序列, 可以得到其它物种中同源性较高的同源microRNA序列。这种方法适合于寻找已报道microRNA的同源物。而RNAz方法则是利用全基因组比较基因组学的方法, 根据物种基因组的保守区域分析来预测microRNA序列, 它不需要种子序列, 但是它需要基因组测序完全。RNAz方法能够寻找到目前还没有报道的microRNA。该方法已经在预测非编码RNA方面被认为是比较优秀的^[45, 69-70]。

根据目前国际上生物信息学的推测, 动物基因组中含有的microRNA数量应该占到该基因组编码基因数量的0.5%到1%之间^[76]。由于在狗基因组中能够编码蛋白质的基因大约有19000个, 所以在基因组中应该有900-1900个microRNA基因。因此在狗基因组中还有大量的microRNA还没有被发现, 有待于进一步研究。

4.2 microRNA 具有多样的基因组织形式

目前对于microRNA的基因组织形式的研究已经开展的比较多。microRNA基因可以位于编码基因的内含子中,也可以位于非翻译区,同时也可以位于假设蛋白的外显子中。最近有研究表明, microRNA位于假基因内, 这进一步说明假基因是有功能的, 并不是以前所认为的“垃圾”DNA^[77]。同时研究报道, 有些microRNA是位于转座元件或者是重复元件内, 这一点也在本文得到证实。我们在本文所发现的microRNA大多数是位于内含子中, 这与植物microRNA相反, 植物microRNA大部分是位于基因间隔区。

在本文中, 357个microRNA分子中有2/5的序列组织成基因簇, 并且成簇的microRNA是由上游的启动子共同转录的, 这使的这些microRNA能够高效共同表达。而且有些簇在动物中是高度保守的。在动物中microRNA成簇现象非常普遍, 而在植物中microRNA大多是形成家族, 这种microRNA在动植物中的基因组织形式差异表明了microRNA在动植物中基因表达和进化上的本来差异^[78]。同时我们发现一些microRNA簇具有多个拷贝, 这说明了microRNA簇存在重复现象。这种现象可以用以下模式来加以解释, 在进化过程中, 经过第一次重复, 某一个miRNA基因的拷贝无意中被损坏, 另外一个拷贝可以持续发挥作用, 而接下来的一系列重复可能是由于不同组织或者是不同阶段的特异性表达而需要的^[79]。

狗基因组中一些相关的microRNA位于同源基因的相应的内含子, 从而形成家族。这些家族成员的成熟序列有些是相似, 有些则是密切相关的。因而这些内含子microRNA可能是通过宿主基因重复后而保存下来的, 同时也说明这些microRNA是比较保守的, 而且在这个家族中可能还存在还没被发现的microRNA。

我们在8号染色体发现的最大的microRNA簇是位于印迹区, 而且是母性表达, 这说明这个簇可能与动物发育密切相关。同时这个簇由三个小的基因家族组成, 这三个家族间具有多个拷贝, 这说明这三个家族分别由一个microRNA经过多次局部串联重复而产生。更有趣的是, 我们在其上游发现了一个由5个microRNA基因所组成的基因簇, 这个簇位于反转座子基因上, 方向相反。也就是说这个簇与反转座子基因反向互补, 这说明这些microRNA可以作为siRNA作用而沉默反转座子基因的表达, 起到RNA干扰作用。在动物中, 大多数microRNA与靶序列不完全互补, 只是抑制靶基因的翻译, 而不具有RNA干扰效应。这个例子进一步说明动物中也具有类似于植物和病毒中的RNA作用, 只是依赖于microRNA与靶序列的互补情况^[80]。

4.3 microRNA 进化的机制

本论文通过系统的搜索动物的microRNA横向同源物和垂直同源物来研究动物microRNA的进化过程。目前我们发现在植物和动物中并没有一个microRNA是两大生物界同源的。我们这里研究了一些microRNA的进化历史,结果表明microRNA的进化过程是一个正在进行的过程。它可以由以下过程产生,局部串联重复和非局部重复。

串联重复是microRNA基因簇最经常发生的一种进化机制。在低等脊椎动物中,这种重复机制占据主要部分。miR-134基因簇就是在低等动物中经过局部重复而产生多个拷贝的。非局部重复基本上是在基因组范围内的重复,它们在不同染色体上产生簇的拷贝^[81]。

4.4 miR-186 与靶序列共表达的进化机制

microRNA通过与3'非翻译区互补配对调节靶序列,而且它所调节的靶序列与microRNA并没有序列相似性,并不象siRNA所调节的靶基因与siRNA具有相似性。更进一步的是, microRNA与靶基因共表达的例子目前也没报道过。我们这里所报道的miR-186位于其靶序列的内含子中,而且是同一方向。这说明miR-186是随着宿主基因的表达而表达,通过宿主基因的mRNA前体的剪接而加工成熟的。我们分析认为,随着宿主基因的转录而产生一个pre-mRNA,继而内含子被剪接下来,产生一个miR-186。而miR-186又作用于其宿主基因,使得宿主基因的表达受到影响。它们之间形成一个相互制约相互影响的作用机制,这样以维持这两者的表达量的平衡,这也体现了生物进化过程中的协调机制^[82]。

4.5 microRNA 的功能多样性

通过计算机分析,我们发现了865个靶基因。大多数靶基因与microRNA是不完全互补的,预示着他们被翻译抑制。然而也有些是完全互补的,同时也有些microRNA是与蛋白编码基因的3'UTR区反向互补,这说明这些miRNA可能起到siRNA的作用而将靶序列降解。

由于一些miRNA在许多物种是非常保守的,甚至是亲缘关系较远的物种也具有保守性,说明这些miRNA在功能上的重要性。靶序列预测表明一些miRNA分子能够单独作用于多个靶位点,同时多个miRNA也可能作用于同一个mRNA,这些在本论文也得到了证实。以上这些结果说明miRNA能够与mRNA形成调控网络。

miRNA的长度较短,因此获得性或缺失性的功能突变的积累发生较少,各种突变、启动子的超甲基化或扩增都很容易引起miRNA的激活或失活。miRNA通过调控mRNA的转译或其稳定性来参与正常细胞稳定状态的维持,因此miRNA的异常表达则可能导致其相应靶基因的异常。miRNA活性丧失使靶基因过表达,而miRNA激活则会导致靶基因下调,这些过程可能参与细胞增殖、凋亡、病毒侵袭和血管发生等过程^[36-37, 83]。

在我们所鉴定的靶序列中,具有最多的功能是转录因子调节子,这正是microRNA所具有的基本功能,其主要是在发育过程调节相关过程。同时我们也发现microRNA也参与了代谢,生物防御,以及信号传导方面的生物过程。同时一些microRNA还可以调节TBP蛋白以及剪接因子U2AF的表达,而这两个蛋白是转录水平上的必需因子,这说明microRNA能够在转录水平上间接调节其它基因的表达^[84]。

同时我们发现,有一些靶序列在数据库中并没有找到相似性的序列,也没有注释,这说明这些序列是狗基因组特异性的基因,可能在狗的生长发育过程中发挥着不可或缺的作用,有一定的研究价值。

4.6 microRNA 表达的时空特异性

对于这些新鉴定的microRNA,我们发现7个microRNA分子存在EST数据库,并且它们分别在不同的组织中表达,主要是在肝脏和肾脏中。其中一个分子在5个组织中都存在。这说明microRNA具有组织特异性表达^[1-2, 4]。另一方面,我们对5个二级结构较好的microRNA分子设计引物进行前体RT-PCR分析,但是结果只有2个microRNA得到预期条带,这表明这两个microRNA分子是在肝脏组织中稳定表达,另外三个分子可能是在别的组织中或者是别的时间阶段表达,因而没得到条带。这正是反映了microA的表达特性^[5, 85]。

致谢

本论文是在罗玉萍教授、李思光教授的悉心指导下完成的。从论文的选题、试验研究到论文写作，都凝聚了他们的心血。我非常庆幸能成为他们的学生，三年来我不仅从他们那里学到了大量的科学知识，开拓了研究思路，更重要的是他们豁达的胸怀，对科学严谨认真、孜孜以求的治学态度，兢兢业业、一丝不苟的工作精神，都将对我今后的工作和生活、学习产生积极的影响。在三年里，两位恩师在生活和学习等方面一直给予我无微不至的关怀、关爱和帮助，这些年的时光将是我人生中非常美好的记忆，并使我终生受益！

论文的顺利完成还要感谢龚熹、徐玲师姐的大力帮助以及冯唐楷、温剑、彭洪海、匡燕云、雷剑等其他同学，感谢他们的技术支持和热情帮助。此外，还要感谢南昌大学生命科学院及学校所有对我付出关心的老师和同学！

感谢含辛茹苦将我养育成人的父亲、母亲！

周冬根
二零零七年四月

参考文献

- [1] Ambros V, Bartel B, Bartel DP et al. A uniform system for microRNA annotation. *RNA*, 2003, Vol 9: 277–279.
- [2] Doench JG, Sharp PA. Specificity of microRNA target selection in translational repression. *Genes Dev*, 2004, Vol 18: 504–511.
- [3] Schwarz DS, Du Hutvagner GT, Xu Z et al. Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, 2003, Vol 115: 199–208.
- [4] Lai EC. microRNAs: runts of the genome assert themselves. *Curr Biol*, 2003, Vol 13: R925–R936.
- [5] Kim J, Krichevsky A, Grad Y et al. Identification of many microRNAs that copurify with polyribosomes in mammalian neurons. *Proc Natl Acad Sci USA*, 2004, Vol 101:360–365.
- [6] Lindow M, Krogh A. Computational evidence for hundreds of non-conserved plant microRNAs. *BMC Genomics*, 2005, Vol 13:116–119.
- [7] Sunkar R, Girke T, Jain PK et al. Cloning and characterization of microRNAs from rice. *Plant Cell*, 2005, Vol 17(5):1397–411.
- [8] Shahi P, Loukianiouk S, Bohne-Lang A et al. Argonaute—a database for gene regulation by mammalian microRNAs. *Nucleic Acids Res.*, 2006, Vol 34(Database issue):D115–8.
- [9] Watanabe Y, Yachie N, Numata K et al. Computational analysis of microRNA targets in *Caenorhabditis elegans*. *Gene*, 2006, Vol 3:365:2–10.
- [10] Bartel, D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 2004, Vol 116:281–297.
- [11] Ambros, V. The functions of animal microRNAs. *Nature*, 2004, Vol 431:350–355
- [12] Bejanin S, Cervini R, Mallet J et al. A unique gene organization for two cholinergic markers, choline acetyltransferase and a putative vesicular transporter of acetylcholine. *J Biol Chem*, 1994, Vol 269: 21944–21947.
- [13] Conrad C, Vianna C, Freeman M et al. A polymorphic gene nested within an intron of the tau gene: implications for Alzheimer's disease. *Proc Natl Acad Sci USA*, 2002, Vol 99: 7751–7756.
- [14] Biemar F, Zinzen R, Ronshaugen M et al. Spatial regulation of microRNA gene expression in the *Drosophila* embryo. *Proc Natl Acad Sci U S A*, 2005, Vol 102(44):15907–11.
- [15] Cavaille J, Vitali P, Basyuk E et al. A novel brain-specific box C/D small nucleolar RNA processed from tandemly repeated introns of a noncoding RNA gene in rats. *J Biol Chem.*, 2001, Vol 276:26374–26383.
- [16] Cavaille J, Seitz H, Paulsen M et al. Identification of tandemly-repeated C/D snoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader–Willi / Angelman

- syndrome region. *Hum Mol Genet*, 2002, Vol 11:1527–1538.
- [17] Tycowski KT, Shu MD & Steitz JA. A mammalian gene with introns instead of exons generating stable RNA products. *Nature*, 1996, Vol 379: 464–466.
- [18] Lin, S.P., Youngson, N., Takada, S. et al. Asymmetric regulation of imprinting on the maternal and paternal chromosomes at the Dlk1-Gtl2 imprinted cluster on mouse chromosome 12. *Nat. Genet.*, 2003, Vol 35: 97–102.
- [19] Bakhetia M, Charlton WL, Urwin PE et al. RNA interference and plant parasitic nematodes. *Trends Plant Sci.*, 2005, Vol 10(8):362-367.
- [20] Padmanabhan V, Callas P, Philips G et al. DNA replication regulation protein Mcm7 as a marker of proliferation in prostate cancer. *J Clin Pathol.*, 2004, Vol 57(10):1057-1062.
- [21] Hayashita Y, Osada H, Tatematsu Y et al. A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. *Cancer Res.*, 2005, Vol 65(21):9628-9632.
- [22] Smalheiser NR, Torvik VI. Mammalian microRNAs derived from genomic repeats. *Trends Genet.*, 2005, Vol 21(6):322-326.
- [23] Tanzer A, Amemiya CT, Kim CB et al. Evolution of MicroRNAs Located Within *Hox* Gene Clusters. *J Exp Zool: Mol Dev Evol*, 2005, Vol 304B:75-85.
- [24] Smit AFA, Hubley R, Green P (1996). RepeatMasker Open-3.0.
- [25] Oakley BR. An abundance of tubulins. *Trends Cell Biol.*, 2000, Vol 10:537-542.
- [26] Keeling P, Doolittle W. Alpha-tubulin from early-diverging eukaryotic lineages and the evolution of the tubulin family. *Mol Biol Evol*, 1996, Vol 13:1297-1305.
- [27] Wilde CD, Crowther CE, Cripe TP et al. Evidence that a human beta-tubulin pseudogene is derived from its corresponding mRNA. *Nature*, 1982, Vol 297:83-84.
- [28] Lemischka I, Sharp PA. The sequences of an expressed rat alpha-tubulin gene and a pseudogene with an inserted repetitive element. *Nature*, 1982, Vol 300:330-335.
- [29] Lee MG, Lewis S, Wilde CD et al. Evolutionary history of a multigene family: an expressed human beta-tubulin gene and three processed pseudogenes. *Cell*, 1983, Vol 33:477-487.
- [30] Lewis SA, Cowan NJ. Tubulin pseudogenes as markers for hominoid divergence. *J Mol Biol*, 1986, Vol 187:623-626.
- [31] Lim LP, Glasner ME, Yekta S et al. Vertebrate microRNA genes. *Science*, 2003, Vol 299:1540-1540.
- [32] Lavorgna G, Dahary D, Lehner B et al. In search of antisense. *Trends Biochem Sci*, 2004, Vol 29:134-138
- [33] Lee Y, Jeon K, Lee JT et al. MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J*, 2002, Vol 21:4663-4670.
- [34] Mourelatos Z, Dostie J, Paushkin S et al. miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev*, 2002, Vol 16:720-728.
- [35] Lagos-Quintana M, Rauhut R, Meyer J et al. New microRNAs from mouse and human. *RNA*, 2003, Vol 9:175-179.

- [36] Lai EC, Tomancak P, Williams RW et al. Computational identification of *Drosophila* microRNA genes. *Genome Biol*, 2003, 4:R42.
- [37] Reinhart BJ, Slack FJ, Basson M et al. The 21-nucleotide RNA *let-7* regulates developmental timing in *Caenorhabditis elegans*. *Nature*, 2000, Vol 403:901-906.
- [38] Pasquinelli AE, Reinhart BJ, Slack F et al. Conservation of the sequence and temporal expression of *let-7* heterochronic regulatory RNA. *Nature*, 2000, Vol 408:86-89
- [39] Pasquinelli AE, McCoy A, Jiménez E et al. Expression of the 22 nucleotide *let-7* heterochronic RNA throughout the Metazoa: a role in life history evolution? *Evol Dev*, 2003, Vol 5:372-378.
- [40] Allen E, Xie Z, Gustafson AM et al. Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*. *Nat Genet.*, 2004, Vol 36(12):1282-1290
- [41] Holland PWH, Garcia-Fernández J, Williams NA et al. Gene duplication and the origins of vertebrate development. *Development*, 1994, Vol 12:125-133.
- [42] Amores A, Force A, Yan YL et al. Zebrafish *Hox* clusters and vertebrate genome evolution. *Science*, 1998, Vol 282:1711-1714.
- [43] Spring J. Genome duplication strikes back. *Nat Genet*, 2002, Vol 31:128-129.
- [44] Tanzer A, Stadler PF. Molecular Evolution of a MicroRNA Cluster. *J Mol Biol*, 2004, Vol 339:327-335.
- [45] Bentwich I, Avniel AA, Karov Y et al. Identification of hundreds of conserved and nonconserved human microRNAs. *Nat Genet.*, 2005, Vol 37:766-770.
- [46] Houbaviy HB, Dennis L, Jaenisch R et al. Characterization of a highly variable eutherian microRNA gene. *RNA*, 2005, Vol 11:1245-1257.
- [47] Huson DH. SplitsTree:: analyzing and visualizing evolutionary data. *Bioinformatics*, 1998, Vol 14:68-73.
- [48] Harfe BD. MicroRNAs in vertebrate development. *Curr Opin Genet Dev.*, 2005, Vol 15(4):410-5.
- [49] Hake LE, Richter JD. CPEB is a specificity factor that mediates cytoplasmic polyadenylation during *Xenopus* oocyte maturation. *Cell*, 1994, Vol 79:617-627.
- [50] Mendez R, Richter JD. Translational control by CPEB: A means to the end. *Nat Rev Mol Cell Biol.*, 2001, Vol 2: 521-529.
- [51] Richter JD. Think globally, translate locally: What mitotic spindles and neuronal synapses have in common. *Proc Natl Acad Sci U S A.*, 2001, Vol 98: 7069-7071.
- [52] Huang YS, Carson JH, Barbarese E et al. Facilitation of dendritic mRNA transport by CPEB. *Genes Dev.*, 2003, Vol 17: 638-653.
- [53] Mendez R, Hake LE, Andresson T et al. Phosphorylation of CPE binding factor by Eg2 regulates translation of *c-mos* mRNA. *Nature*, 2000, Vol 404: 302-307.
- [54] Mendez R, Barnard D, Richter JD. Differential mRNA translation and meiotic progression require Cdc2-mediated CPEB destruction. *EMBO J.*, 2002, Vol 21:1833-1844.
- [55] Calin GA, Dumitru CD, Shimizu M et al. Frequent deletions and down-regulation of

- micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A.*, 2002, Vol 99: 15524–15529
- [56] Lagos-Quintana M, Rauhut R, Meyer J et al. New microRNAs from mouse and human. *RNA*, 2003, Vol 9: 175–179.
- [57] Calin GA, Sevignani C, Dumitru CD et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc Natl Acad Sci U S A.*, 2004, Vol 101: 2999–3004.
- [58] Michael MZ, O'Connor SM, van Holst Pellekaan NG et al. Reduced accumulation of specific microRNAs in colorectal neoplasia. *Mol Cancer Res.*, 2003, Vol 1: 882–891.
- [59] Metzler M, Wilda M, Busch K et al. High expression of precursor microRNA-155/BIC RNA in children with Burkitt lymphoma. *Genes Chromosomes Cancer*, 2004, Vol 39: 167–169.
- [60] Wang Y, Hung C, Koh D et al. Differential expression of Hox A5 in human colon cancer cell differentiation: A quantitative study using real-time RT-PCR. *Int J Oncol.*, 2001, Vol 18: 617–622.
- [61] Hommes DW, Peppelenbosch MP, van Deventer SJ. Mitogen activated protein (MAP) kinase signal transduction pathways and novel anti-inflammatory targets. *Gut*. Jan., 2003, Vol 52(1):144-51
- [62] Kyriakis JM, Avruch J. Mammalian mitogen-activated protein kinase signal transduction pathways activated by stress and inflammation. *Physiol Rev.*, 2001, Vol 81:807–69
- [63] Zhang, B.H., Pan, X.P., Cannon, C.H. et al. Conservation and divergence of plant microRNA genes. *Plant Journal*, 2006, Vol 46:243–259.
- [64] Zhang, B.H., Pan, X.P., Cox, S.B. et al. Evidence that miRNAs are different from other RNAs. *Cellular and Molecular Life Sciences*, 2006, Vol 63:246–254.
- [65] Washietl S, Hofacker IL, Stadler PF. Fast and reliable prediction of noncoding RNAs. *Proc Natl Acad Sci U S A*, 2005, Vol 102(7):2454-2459
- [66] Enright AJ, John B, Gaul U et al. miRanda (v1.0b) *Genome Biology*, 2003, Vol 5(1):R1.
- [67] Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, 1993, Vol 75: 843–854.
- [68] Reinhart BJ, Slack FJ, Basson M et al. The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature*, 2000, Vol 403: 901–906.
- [69] Lee RC, Ambros V. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science*, 2001, Vol 294: 862–864.
- [70] Ambros V. MicroRNA pathways in flies and worms: Growth, death, fat, stress, and timing. *Cell*, 2003, Vol 113: 673–676.
- [71] Du T, Zamore PD. microPrimer: the biogenesis and function of microRNA. *Development*, 2005, Vol 132(21):4645-4652
- [72] Kim VN, Nam JW. Genomics of microRNA. *Trends Genet*, 2006, Vol 22(3):165-173
- [73] Sewer A, Paul N, Landgraf P et al. Identification of clustered microRNAs using an ab initio prediction method. *BMC Bioinformatics*, 2005, Vol 7:267.

- [74] Bentwich I. Prediction and validation of microRNAs and their targets. *FEBS Lett.*, 2005, Vol 579(26):5904-10
- [75] Burgler C, Macdonald PM. Prediction and verification of microRNA targets by MovingTargets, a highly adaptable prediction method. *BMC Genomics*, 2005, Vol 6(1):88
- [76] Doench JG, Sharp PA. Specificity of microRNA target selection in translational repression. *Genes Dev.*, 2004, Vol 8: 504-511.
- [77] Tuschl T, Zamore PD, Lehmann R et al. Targeted mRNA degradation by double-stranded RNA in vitro. *Genes Dev.*, 1999, Vol 13: 3191-3197.
- [78] Elbashir SM, Lendeckel W, Tuschl T. RNA interference is mediated by 21 and 22 nt RNAs. *Genes Dev.*, 2001, Vol 15: 188-200.
- [79] Hutvagner G, Zamore PD. A microRNA in a multiple-turnover RNAi enzyme complex. *Science*, 2002a, Vol 297: 2056-2060.
- [80] Llave C, Xie Z, Kasschau KD et al. Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. *Science*, 2002b, Vol 297: 2053-2056.
- [81] Martinez J, Patkaniowska A, Urlaub H et al. Singlestranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell*, 2002, Vol 110:563-574.
- [82] Elaine A, Ostrander, Robert K et al. The canine genome. *Genome Research*, 2005, Vol 15:1706-1716.
- [83] Parker, H.G., Kim, L.V. et al. Genetic structure of the purebred domestic dog. *Science*, 2004, Vol 304: 1160-1164.
- [84] Chomczynski, P., Sacchi, N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.*, 1987, Vol 162:732-735.
- [85] Ken C. Pang, Martin C. Frith, John S. Mattick Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *TRENDS in Genetics*, 2006, Vol.2 .

附录 A: 人基因组中尚未发现同源序列的人小分子 RNA

编号	前体序列	成熟序列	长度
miR-801	atgtaagaagtgcttacattgcagtagatctcaaaaagctacctgcactgt aagcacttttaca	atgtaagaagtgcttaca ttg	21
miR-802	cttttgctccttttctccccctgatggggttcgtagaatttatgacactctat agggctggggctgaggggggtgggaag	cttttgctccttttctccc	21
miR-803	agaatgatctatgtactttcaaaacttaaacagaaatcaaaggaaaact gaactgtttaagtttgacagctgcagggatcattct	tgacagctgcagggatc attc	21
miR-804	tggctgcttgggttcctggcatgctgattgtgacttaagattaaaatcacat tgccagggattaccacgca	tggctgcttgggttcctg gca	21
miR-805	gctctcttgctctgtagtagaattctggctgctcattcggcatttggcttttct cacggagcccttgagc	tggcttttctcacggag cc	21
miR-806	tgatattatgccatgacattgtgtcaatatgcatgatgtgtgtgatggcac agcgtcatcacgtggtagcgaacatcatgacgtaagacgicacaacgc c	tattatgccatgacattgt gt	21
miR-807	tggctctgcagtgaaagtaaatgtaggatcctgacatcaatttgttccatg gaaataattgagatgtcaatgtttccgattatgactgtaccataggagca gatgg	tgcagtgaaagtaaatgt agg	21
miR-808	tcaggtagaagccagggtgaaactcaaagggtccaaactgggttttgcttt gagttccttttgggtctgcttga	ggtagaagccagggtga aactc	21
miR-814	cgetgttctctatggctttttattcctatgtgattctactgctcactcatatagg gattggagccgtggcgacggcg	tggctttttattcctatgtg a	21
miR-818	cttagcagagctgtggagtgtgacaatgggtttgtgtctaaactatcaaac gccattatcacactaaatagctactgctagg	tagcagagctgtggagt gtga	21
miR-819	gagctgtttgggtgctttgtttccggagcgtccaaccagcacaccgc ctccggcaaaactgcgagcaattcaaagcagctc	tgggtgtgctttgtttccc gga	22
miR-820	tcactctgatgatgttatggataaatcattaaggaggcagctcaccgcg gtgatgtattgggttaatagatcatcagagtga	accgcggtgatgtattg gtgt	21
miR-822	cccggactagtagatcatctatactgtagtctcatcgcaaaacttacagta tatgatgatatcccagccaggg	tacagtatatgatgatac cc	21
miR-824	tggcagaggattaagtgaatgctgatgaagtgtccagattcccataacttg gaatggagctgcattcagcattcctttattatccatctgcca	taagtgaatgctgatgaa gtg	21
miR-826	gcctctccatgggtactcggggagagggtaccggagcaactttgcatctgg acgacgaatgttctcgggtgaaccccttttcggtatcaaatccaccaggg aggc	cgggtgaaccccttttcgg tat	21
miR-828	atctggagagcctctcagagctgttcagttgaggagccaaaagaaaatc aaaatgaactttcagttcagagaggcagctctataggt	tggagagcctctcagag ctgt	21
miR-830	tcagggtattctacagtgcacgtgtctccagtgtggctcggaggctgga gacgcggccctgttggagtaacaactgaagcc	tcagggtattctacagt gca	21
miR-833	ggtggggaggggggtgggaagaagggatgtctgtctgtctccccctc	tgctgtccccctccccct	20

附录 A: 人基因组中尚未发现同源序列的人小分子 RNA

	ccccctcact	ct	
miR-841	ttttgttggtcataggtaccaggcittacacaaactgtgaagtgctgtgga	gaagtgctgtggatttct	22
	ttctttgtgaatcaccatactaaagctaagtgggtgggtttacaaagtaat	tgt	
	tcatagtgcttcacag		
miR-842	tattgttaattgaagtgctatgctttgtactttatttcagagcatcatgcctg	tgaagtgctatgctttgt	21
	agcagacatagaaaatagcaattacttttagcaaagtga	ac	
miR-843	tgtcctcaggtccagtttcccagggaatcccttagatgctaagatgggga	tccagtttcccaggaat	22
	ttcctggaaatactgttcttgaggtcatgg	ccct	
miR-846	caccagagtgccttgcttaataatgtattactctaataatctgctgctttaat	ctaataatgtattactcta	21
	gatataatgtatacattatcagagcaaggacaactgatg	at	
miR-847	ccattactgtgctaataatgcaactctgttgaatataaattggaattgcacttt	attgcactttagcaatggt	21
	agcaatggtgatgg	ga	

附录 B: microRNA 与靶序列完全互补

			5' CAGACTCTGGGGGAGTGGAGGG 3'
miR-574	LOC479240	glycosphingolipid metabolism	3' CACCCACACACGTA CTGGCAC 5'
			5' GTGGGTGTGGGTGTGGGTGTG 3'
miR-197	LOC483670	positive regulation of cell proliferation	3' CGACCCACCTCTCCACCACTT 5'
			5' CCTGGGGGGATGGGGTGGTGA 3'

附录 C: 缩略词

缩写	英文全称	中文全称
UTR	Untranslated region	非翻译区
miRNA	microRNA	微小 RNA
Blast	Basic local alignment search tool	基本局部 (同源序列排列) 比对搜索工具
BLAT	Balst-like alignment tool	类似于 balst 比对工具
bp	base pair	碱基对
DEPC	Diethyl pyrocarbonate	焦碳酸二乙酯
DNA	deoxyribonucleic acid	脱氧核糖核酸
cDNA	complementary DNA	互补 DNA
dNTP	deoxyribonucleoside triphosphate	脱氧核糖核苷三磷酸
EB	Ethidium bromide	溴化乙锭
EDTA	ethylene diaminetetraacetic acid	乙二胺四乙酸
Pre-mRNA	precursor mRNA	前体 mRNA
μg	microgram	微克
mol	molar	摩尔
ml	milliliter	毫升
ncRNA	non-coding RNA	非编码 RNA
nr	non-redundant database	非冗余数据库
nt	nucleotide	核苷酸
OD	Optical Density	光密度
ORF	open reading frame	开放读码框
PCR	polymerase chain reaction	聚合酶链式反应
RNA	ribonucleic acid	核糖核酸
ETS	External transcribed spacer	外转录区
ITS	Internal transcribed spacer	转录间隔区
RT	reverse transcription	反转录
SDS	sodium dodecyl sulfate	十二烷基硫酸钠
microRNP	micro RNP	微小分子 RNA 核蛋白体颗粒
Pri-miRNA	Primary microRNA	microRNA 原初转录本
snoRNA	small nucleolar RNA	核仁小分子 RNA
snmRNA	small non-message RNA	小分子非信使 RNA
siRNA	Small interfering RNA	干扰 RNA
RNAi	RNA intererrnce	RNA 干扰
RISC	RNA-induced silencing complex	RNA 诱导沉默复合体
stRNA	Small temporal RNA	小分子时序 RNA
Tris	tris(hydroxymethyl)-Aminomethane	三(羟甲基)氨基甲烷
IGR	Intergenetic region	基因间隔区
μl	microliter	微升

攻读学位期间的研究成果

已发表论文:

- [1] 周冬根, 罗玉萍*, 李思光. 《microRNA 的结构、生物合成及功能》, 生物技术通报, 2005 (10)

待发表论文:

- [1] Zhou Donggen, Li Siguang, Wen Jian, Gong Xi, Xu Ling, You Yuping*. 《Genome-wide computational analyses of microRNA genes from *Canis familiaris*》