

[文章编号] 1007-7405(2016)02-0107-08

# 美洲鲎 miRNA 的生物信息学挖掘与分析

朱珊珊<sup>1,2</sup>, 翁朝红<sup>1,2</sup>, 肖世俊<sup>1,2</sup>, 韩兆方<sup>1,2</sup>, 谢仰杰<sup>1,2</sup>

(1. 集美大学水产学院, 福建 厦门 361021; 2. 农业部东海海水健康养殖重点实验室, 福建 厦门 361021)

**[摘要]** 利用 miRBase 数据库中已有动物的小 RNA (miRNA), 通过生物信息学方法对美洲鲎 (*Limulus polyphemus*) 表达序列标签 (Expressed Sequences Tag, EST) 和 cDNA 进行序列比对, 挖掘美洲鲎 miRNA, 得到了 72 个 miRNA 前体 (pre-miRNA)、49 个可编码成熟 miRNA, 隶属于 40 个 miRNA 家族。miRNA 家族归类结果表明, 无论在低等脊椎动物还是高等动物中, miRNA 家族的存在都是相当保守的。以 miR-10 基因前体为例, 分析其序列与其他物种同源序列的差异, 显示 miRNA 序列的高度保守性。miR-10 前体序列分析系统进化发现, 物种聚类与传统分类亲缘关系一致。

**[关键词]** miRNA; 美洲鲎; 生物信息学; 挖掘

**[中图分类号]** Q 74

## Bioinformatic Identification of Conserved miRNA in *Limulus polyphemus*

ZHU Shan-shan<sup>1,2</sup>, WENG Zhao-hong<sup>1,2</sup>, XIAO Shi-jun<sup>1,2</sup>, HAN Zhao-fang<sup>1,2</sup>, XIE Yang-jie<sup>1,2</sup>

(1. Fisheries College, Jimei University, Xiamen 361021, China; 2. Key Laboratory of Healthy Mariculture for the East China Sea, Ministry of Agriculture, Xiamen 361021, China)

**Abstract:** In this study, bioinformatics method were used to identify microRNAs of *Limulus Polyphemus*. The ESTs and cDNA sequences of *L. polyphemus* were aligned to the known animal miRNAs from miRBase. 72 miRNA precursors (pre-miRNAs) and 49 mature miRNAs belonging to 40 miRNA families were identified. The classified analysis of miRNA family showed that mature miRNAs and precursor sequences are highly conserved, not only existing in the low class of invertebrates but also in high vertebrates. Analysis of sequences homology with other species based on miR-10 precursors indicated that the sequences of miRNA were fairly conserved. The phylogenetic analysis of miR-10 precursor sequences revealed that clustering mode was consistent with the traditional classification of related species.

**Keywords:** miRNA; *Limulus polyphemus*; bioinformatics; mining

## 0 引言

小 RNA (miRNA 或 microRNA) 广泛存在于病毒、多细胞动物和植物中, 是一类内源性、长度在 20~25 bp 之间的非编码单链 RNA, 它通过与靶基因的 mRNA 互补配对在转录水平上对基因的表达进行负调控, 导致靶基因 mRNA 的降解或者抑制其翻译过程, 从而抑制相关蛋白质的合成。动物、植物、病毒和藻类的基因组中都普遍存在着大量的 miRNA<sup>[1-5]</sup>, 并且在动物基因组中, miRNA 是最常

**[收稿日期]** 2015-08-31

**[修回日期]** 2015-10-12

**[基金项目]** 国家自然科学基金资助项目 (41106144); 福建省自然科学基金资助项目 (2014J01132); 集美大学博士科研启动基金项目 (C613001)

**[作者简介]** 朱珊珊 (1990—), 女, 硕士生, 从事水产养殖方向研究。通信作者: 翁朝红 (1971—), 女, 教授, 从事海洋生物分子技术方向研究, E-mail: wengzhaozhong@jmu.edu.cn。

见的一类基因调控原件。在生物体中,首先形成一个较长的前体 miRNA (pri-miRNA);然后通过细胞核中的核糖核酸酶Ⅲ (RNaseⅢ) 形成特殊的发卡状 RNA 前体 (pre-miRNAs), Pre-miRNAs 再进入细胞质进一步被 RNaseⅢ Dicer 加工成成熟的 miRNA<sup>[6-7]</sup>;最后,成熟的 miRNA 装配形成 miRNA 诱导的沉默复合体 (miRISCs) 进而发挥功能,并且诱导其靶 mRNA 翻译抑制或使与形成 GO 蛋白相关的 mRNA 降解<sup>[8]</sup>。

据估计,动物中约 30% ~ 60% 编码蛋白质的基因受到 miRNA 的调控,涉及已知的所有细胞生物学过程<sup>[9-11]</sup>。以 miR-10 为例,miR-10 是动物中分布最广泛的 miRNA,已有的资料显示,miR-10 家族成员都高度保守。miR-10 通过积极给出促进血管新生信号来调节血管中的内皮细胞生成<sup>[12]</sup>。miR-10 能直接调节 FLT1 和 sFLT1 表达水平,FLT1 是一种细胞表面的蛋白质,能隔绝血管内皮生长因子,sFLT1 是 FLT1 的可溶性剪接变体<sup>[13]</sup>。敲除 miR-10 会导致斑马鱼 (*Barchydanio rerio* var) 幼体胚节间的血管过早终止生长,而 miR-10 过量表达会加速斑马鱼中的血管生成和人类脐静脉内皮细胞生长<sup>[14]</sup>。

螯属节肢动物门 (Arthropoda), 肢口纲 (Merostomata), 剑尾目 (Xiphosurida), 螯科 (Limuroidae), 是一种古老的海洋底栖节肢动物。目前在 miRNA 注册数据库 (miRBase) 中未发现美洲螯或中国螯 miRNA 的序列,其他有螯亚门的动物只有 3 种蜱螨类 (黑脚硬蜱 (*Ixodes scapularis*)、微小扇头蜱 (*Rhipicephalus microplus*) 和棉叶蜱 (*Tetranychus urticae*)) 已有公布序列,已报道还有 2633 个昆虫纲 miRNA 以及其他重要模式生物<sup>[15]</sup>。自从第一个 miRNA (lin-4) 在绣线隐杆线虫 (*Caenorhabditis elegans*) 中被发现,近年来,通过直接克隆、测序和原位杂交等技术手段发现了大量的 miRNA 分子,然而用这些方法来鉴定 miRNA 是低产量和随机的。而动植物中的大部分 miRNA 在进化过程中都是高度保守的,基于已知的 miRNA 序列通过比较基因组学来预测新的 miRNA 同系物已经成为一种有力的方法。目前,通过生物信息学方法在多种动植物体内成功发现了大量的 miRNAs<sup>[16-17]</sup>。但是,对于美洲螯这种有“活化石”之称的水生节肢动物,目前无任何 miRNA 相关资料,也没有通过生物信息学方法预测美洲螯 miRNAs 的报道。本文运用生物信息学的方法对美洲螯的 miRNAs 进行预测和分析,以期对美洲螯 miRNA 的深入研究提供一定的理论依据。

## 1 材料与方法

### 1.1 数据收集与分析

美洲螯表达序列标签 (Expressed Sequence Tag, EST) 和 cDNA (共 8568 条) 从 NCBI 数据库 (<http://www.ncbi.nlm.nih.gov/>) 下载。已知动物的 miRNA 及其前体序列从 miRBase 数据库 (<http://microrna.sanger.ac.uk/sequences/>, 2014 年 6 月) 下载,去除重复序列后,以此作为参考数据库用于预测美洲螯中潜在的 miRNAs。从 NCBI 数据库中下载非冗余蛋白序列 (<ftp://ftp.ncbi.nlm.nih.gov/blast/db/>) 用于删除潜在的蛋白质编码序列。

### 1.2 美洲螯 miRNA 分子预测

将已知动物 miRNA 的非重复成熟序列通过 BLAST N 方法<sup>[18]</sup>与美洲螯的 EST 和 cDNA 序列进行比对, BLAST N 参数设置如下: 1) *E* 值设定为 10; 2) word size 设定为 7; 3) 比对位点与已知 miRNAs 之间不匹配碱基数  $\leq 3$  (包括缺失/插入和转换/颠换)。提取比对位点上下游各 100 个核苷酸大小的片段,采用 BLAST X 方法与 NCBI 中下载的非冗余蛋白序列数据库进行比对分析,删除蛋白质编码序列。利用 RNAfold 软件 (<http://www.tbi.univie.ac.at/~ivo/RNA/index.html>)<sup>[19-20]</sup> 预测茎环结构。用 MiPred 软件 (<http://www.bioinf.seu.edu.cn/miRNA/>)<sup>[21]</sup> 对候选的 pre-miRNA 序列鉴定是否具有 pre-miRNA 二级结构特征,去除伪 pre-miRNA 序列。所获得美洲螯潜在 miRNA 分子必须遵循下列标准: 1) 来自 ESTs 或 cDNA 的同源序列与已知成熟的 miRNA 序列之间不超过 3 个核苷酸差异; 2) miRNA 前体分子能够折叠成合适的发卡结构; 3) 成熟的 miRNA 位于发卡结构的一条臂上,且发卡

结构的环碱基数  $\leq 20$ ; 4) 发卡结构中, miRNA 分子与其互补序列的碱基差异不能多于 6 个; 5) 在 miRNA 成熟区无环状结构; 6) Pre-miRNA 二级结构最低自由能  $\leq 83.736$  kJ/mol。筛选步骤见图 1。

1.3 miRNA 家族归类及保守性分析

从 miRBase 数据库 (<http://www.mirbase.org>) 中下载 miRNA 家族列表, 将预测到的 miRNA 与已知家族成员进行比较, 确定该 miRNA 的家族。并将确定的美洲鲎 miRNA 家族与 miBase 数据库上已发布的斑马鱼、小鼠 (*Mus musculus*)、人类 (*Homo sapiens*)、黑腹果蝇 (*Drosophila melanogaster*) 的 miRNA 家族进行统计分析, 研究美洲鲎与斑马鱼、小鼠、人类、黑腹果蝇的 miRNA 家族的差异性。

1.4 miR-10 前体同源性分析

采用 Clustal W 方法对 miR-10 前体序列与其同源序列进行多序列比对。

1.5 miRNA 前体序列系统进化分析

分别从 miRbase 数据库中下载紫球海胆 (*Strongylocentrotus purpuratus*)、绿海胆 (*Lytechinus variegatus*)、海星 (*Patiria miniata*)、文昌鱼 (*Branchiostoma lanceolatum*)、绿安蜥蜴 (*Anolis carolinensis*)、眼镜蛇王 (*Ophiophagus hannah*)、斑点叉尾鲷 (*Ictalurus punctatus*)、斑马鱼、斑胸草雀 (*Poephila guttata*)、大棕蝠 (*Eptesicus fuscus*)、黑腹果蝇、拟果蝇 (*Drosophila simulans*)、日本血吸虫 (*Schistosoma japonicum*)、曼氏血吸虫 (*Schistosoma mansoni*)、细粒棘球绦虫 (*Echinococcus granulosus*) 和多房棘球绦虫 (*Echinococcus multilocularis*) 等物种的 miR-10 前体序列, 使用软件 MEGA 4.0 对美洲鲎与这些物种的 miR-10 前体序列进行比对分析, 并用邻接法 (Neighbor-Joining, NJ) 建立系统发育树, 初步分析美洲鲎 miRNA 在不同物种间的保守性及进化关系。

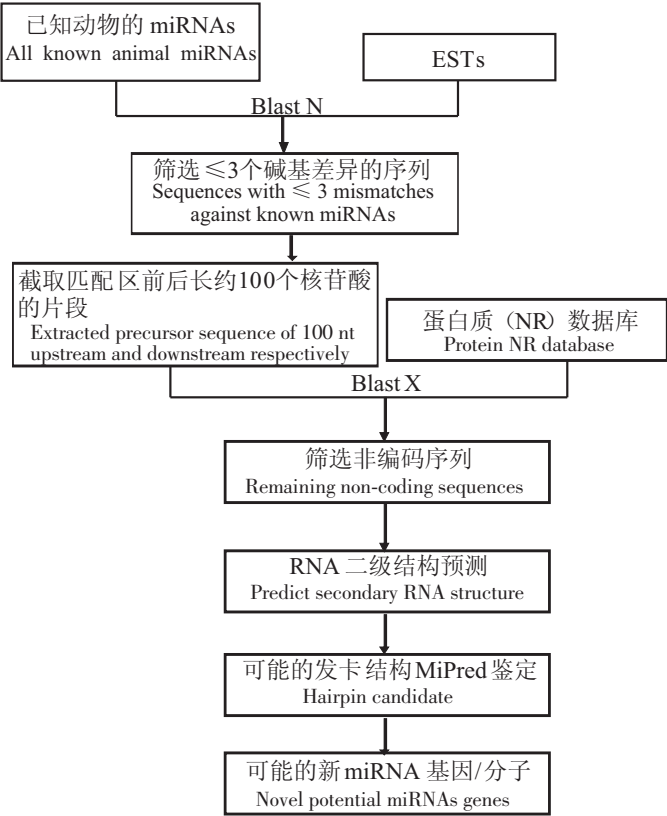


图 1 miRNA 筛选流程图

Fig.1 Flowchart of miRNA prediction in *L. polyphemus*

2 结果

2.1 美洲鲎中潜在 miRNA 预测结果

本研究通过比较基因组学和生物信息学方法对美洲鲎 miRNA 进行了挖掘, 筛选出 miRNA 前体 72 个、miRNA 49 个, 隶属于 40 个 miRNA 家族。成熟 miRNA 长度为 19 ~ 23 个核苷酸。49 条成熟 miRNA 中, 17 条位于前体序列茎环结构的 5' 端, 其他 32 条位于 3' 端。

有关美洲鲎 miRNA 的家族、名字、长度及其位置等具体信息见表 1。

2.2 新预测美洲鲎 miRNA 前体特点分析

72 条 pre-miRNA 的长度为 46 ~ 102 个核苷酸, 平均为 74 个核苷酸; 二级结构折叠最小自由能为  $-2.61 \times 10^6 \sim -9.09 \times 10^5$  J/mol, 平均为  $-1.76 \times 10^7$  J/mol。碱基组成 U + A 的含量 (平均为 68.4%) 远高于 G + C 含量 (平均为 31.6%)。72 条 pre-miRNA 都具有典型的茎环发卡结构。部分 pre-miRNA 的二级结构见图 2。

表 1 新发现的美洲鲎 miRNA  
Tab.1 New miRNAs found in *limulus polyphemus*

编号 Number	名称 Name	成熟序列 Mature Sequence	长度 Length /nt	家族 Family	基因组编号 ID Genome number	位置 Position
1	lpo-miR-466i	UGUGUGUGUGUGUGUGUGUG	20	467	> gi569598286	5'
2	lpo-miR-1187	UAUGUGUGUGUGUAUGUGUGUAA	23	1187	> gi569933568	3'
3	lpo-miR-1277	UAUAUAUAUAUAUGUACGUAU	21	1277	> gi569936442	3'
4	lpo-miR-2325c	GGUUGUUUUUUUUCUUUUUC	21	2325	> gi569947340	3'
5	lpo-miR-10	ACCCUGUAGAUCCGAAUUUGU	21	10	> gi569969021	5'
6	lpo-miR-375	UUUGUUCGUUUGGCUCGAGUUA	22	375	> gi51640560	3'
7	lpo-miR-219b	CACAAGAAUUGCGUUUGGACAA	22	219	> gi570011497	3'
8	lpo-miR-2361	GUUGUGUUGUUUUUUUUUGU	21	2361	> gi569987049	5'
9	lpo-miR-4705	GCAAUCGCCAAGUGUUUGA	19	4705	> gi283514006	5'
10	lpo-miR-101-1	AGUUAUCACAUUGUGUUGCU	20	101	> gi283500714	3'
11	lpo-miR-338	CUGCAGUCACAGGAUAUUGUU	21	338	> gi283519678	3'
13	lpo-miR-4000a	UUUUUAUUUAAUUUCCAU	19	4000	> gi283543172	5'
14	lpo-miR-3755	UGUAUAUAUAGAUGUAUAU	19	3755	> gi283539177	5'
15	lpo-miR-1006	UGUAUAUAUAGAUGUAUAU	19	1006	> gi283524161	5'
16	lpo-miR-3248	UUACACAUUUUAUUUAAAC	19	3248	> gi283539183	3'
17	lpo-miR-5011	GUGUAUAUAUAUAGGCAUGC	20	5011	Contig140	3'
18	lpo-miR-297a	UGUAUGUGUGUGUGUAUGU	20	297	> gi283495901	5'
19	lpo-miR-7562	GUGAUCACGUGUGUGUGUG	20	7562	> gi283495488	5'
20	lpo-miR-553	AAAAAGUCAGGGUUUGUUU	19	553	> gi283512795	3'
21	lpo-miR-754b	AUCUUAUCAUCAGUCAUA	19	754	> gi283543729	3'
22	lpo-miR-4115	AAGAACUGGAGUGUAGCAA	19	4115	> gi   283499990   gb   FN226743.1	3'
23	lpo-miR-548d	AAGAAAUUAUGUAUUUGC	19	548	> gi283523149	3'
24	lpo-miR-548aj	AACAGAAAUUCAGUUUUU	19	548	> gi283543354	3'
25	lpo-miR-6961	UAUGAGCGGAGGAAGAAAGGAGGA	23	6961	Contig240	3'
26	lpo-miR-3350	UGUAUUCUAAUUUAAUCUUU	20	3350	> gi283501288	5'
27	lpo-miR-1	UAAUUUUUUUAAAUUCAUA	19	1	> gi283495867	5'
28	lpo-miR-466	UAUGUGUAUAUUUAUAUAU	19	466	> gi283541692	5'
29	lpo-miR-466f	UGUGUGUGUGUGUGUAUGU	20	467	> gi283514139	3'
30	lpo-miR-466m	GCGUGUGAAUGUUUGUUGU	20	467	> gi283522950	3'
31	lpo-miR-2177	AGAAAUUUAAUGUUUAAU	19	2177	> gi283511467	3'
32	lpo-miR-2218	ACAAAACUUCUCUAUUCUGGC	21	2218	> gi114106997	3'
33	lpo-miR-34	CACGUGAUCUUGGUUGUUGU	21	34	> gi283501082	3'
34	lpo-miR-D14	AAAAUAUAUAACAGACAU	19	D14	> gi283511211	5'
35	lpo-miR-669b	GUGUGUGUGUGUGUGUAUG	20	467	> gi283514139	3'
36	lpo-miR-669m	UGUGCGUGCAUGUGUAUAU	20	467	> gi283512684	3'
37	lpo-miR-669o	AGUGUGCGUGCAUGUGUAU	20	467	> gi283512684	3'
38	lpo-miR-190	AAAAUAUGAUUAAUCCUG	19	190	> gi283525373	3'
39	lpo-miR-2370-3p	AGAAAGUAAAGUUGCAUUA	19	2370	> gi283543991	3'
40	lpo-miR-5317a	CUUGUUUCAGUAUCGUCGU	19	5317	> gi283543932	5'
41	lpo-miR-2807a	GAAAUUGGUAUCCUGCUG	19	2807	Contig625	5'
42	lpo-miR-723	UUUAAAAUAUAUUACUUU	19	723	> gi283522784	3'
43	lpo-miR-964	GAAGGUUCCUGAAGCUUAC	19	964	Contig627	3'
44	lpo-miR-467b	UGUGUGUGUGUGUGUGUAU	22	467	> gi283514139	3'
45	lpo-miR-2193	AUGUGUGUGUGUAUUGUGU	20	2193	> gi283495935	5'
46	lpo-miR-3327	UAUGUAAUUUUUAUGUUGU	20	3327	> gi283500386	5'
47	lpo-miR-2359	AUUUUGUUGUUUAUUUUU	20	2359	> gi283511211	3'
48	lpo-miR-3811d	CUCCACUGACUGUUGGACA	19	3811	> gi283519491	3'
49	lpo-miR-7649	CAGAAGUGUUAUAGACAUG	19	7649	Contig812	3'



2.3 miRNA 家族归类及保守性分析

美洲鲎中保守的 miRNA 在其他物种中同样存在。本文分析了美洲鲎的重要 miRNA 家族与斑马鱼、小鼠、人类、黑腹果蝇中同源基因的保守性。图 3 显示了 18 个 miRNA 家族在几种两侧对称动物的分布情况。鲎与其他几种两侧对称动物之间共有 miRNA 家族的数量特点是: 与人共有的 miRNA 家族数目最多, 有 12 个共有的 miRNA 家族, 其次是小鼠有 9 个, 然后依次是斑马鱼 (8 个) 和黑腹果蝇 (5 个)。结果表明, 美洲鲎作为一种无脊椎动物, 它不但与无脊椎动物 (如果蝇) 有相同的家族, 而且与高等脊椎动物 (如人) 具有共有的家族, 说明 miRNA 在进化上具有非常高的保守性。

2.4 miR-10 前体序列同源性分析

美洲鲎 miR-10 前体序列长度为 69 bp, 其中成熟 miRNA 序列长度为 21 bp, 其 miR-10 成熟序列为: ACCCUGUAGAUCCGAAUUUGU, 种子区序列为 ACCCUGU; 下载其他物种的 miR-10 家族的 pre-miRNA 进行序列比对, 分析结果 (见图 4) 显示, 美洲鲎 (lpo)、斑马鱼 (dre)、红鳍东方鲀 (fru)、人 (hsa)、鼠 (mmu)、文昌鱼 (bfl) 和拟暗果蝇 (dps) 的 miR-10 家族的 pre-miRNA 序列具有高度相似性和同源性, 特别是成熟区部分具有高度保守性, 只有若干个碱基的差异, 而种子区碱基则完全相同。

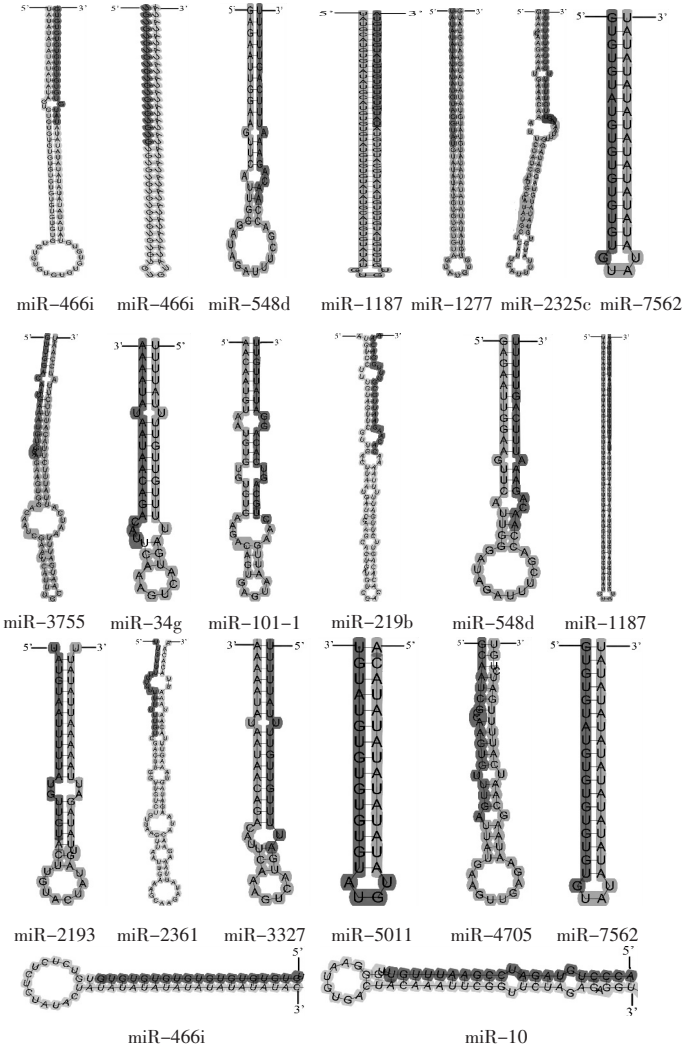
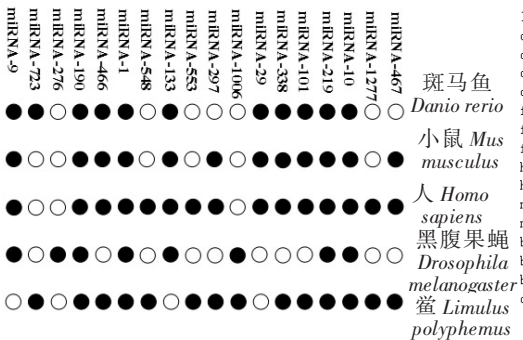
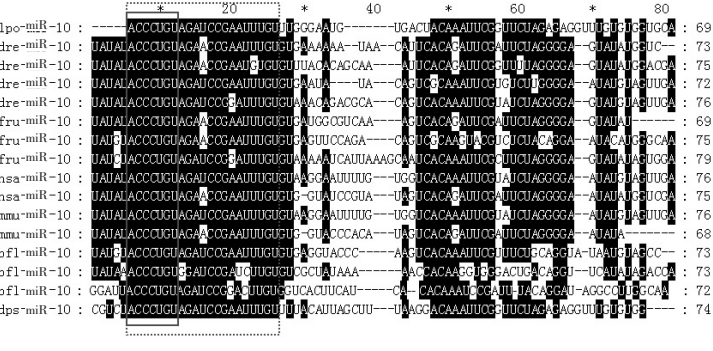


图 2 部分 pre-miRNA 的二级结构 (深色部分为成熟序列区)  
Fig.2 The secondary structure of some pre-miRNAs in *L.polyphemus* (deep color area is part of mature sequences)



说明: 黑色表示 miRNA 家族存在, 白色表示 miRNA 家族缺失。  
Notes: the black indicated existence of the miRNA family; the white indicated deletion of the miRNA family.

图 3 几种两侧对称动物 miRNA 家族的分布情况  
Fig.3 Distribution of miRNA across Bilateria



说明: 虚线框内为成熟区; 实线框内为种子区。  
Notes: mature sequences located inside of the dashed box; seed sequences inside of the solid line box.

图 4 不同物种 miR-10 家族 pre-miRNA 的序列联配  
Fig.4 Pre-miRNA alignments of miR-10 family from different species

2.5 美洲鲎 miR-10 前体序列系统发育分析

从 miRbase 数据库下载的其他物种（紫球海胆、绿海胆、海星、文昌鱼、绿安蜥蜴、眼镜蛇王、斑点叉尾鲷、斑马鱼、斑胸草雀、大棕蝠、黑腹果蝇、拟果蝇、日本血吸虫、曼氏血吸虫、细粒棘球绦虫、多房棘球绦虫）的 miR-10 前体序列比对分析，并构建进化树（见图 5）。结果表明，美洲鲎与同为节肢动物的 3 种果蝇（果蝇、黑腹果蝇、拟果蝇）聚成一支，4 种扁形动物（日本血吸虫、曼氏血吸虫、细粒棘球绦虫、多房棘球绦虫）聚成一支，棘皮动物、鱼类、爬行类同样按其自身各自门类聚成一支。由 miR-10 前体序列构建的系统进化树与传统系统分类法物种亲缘关系完全符合，同一个科属的物种聚在一起。美洲鲎与同为节肢动物门的果蝇聚成一支。

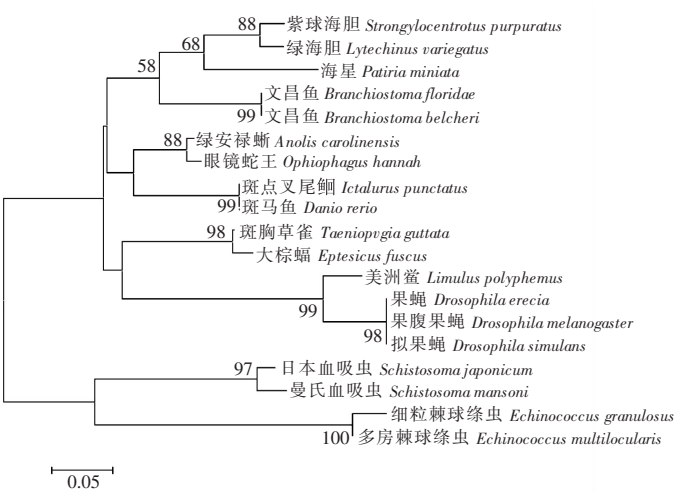


图 5 miR-10 前体序列进化树

Fig.5 The phylogenetic tree of miR-10 precursor sequences

3 讨论

3.1 miRNA 计算机预测

关于 miRNAs 的研究在沉寂近十年后，又重新成为人们的研究热点。通过实验方法和生物信息学方法已经克隆到了上万种 miRNAs。miRNA 能够通过克隆<sup>[22-23]</sup>、高通量测序<sup>[24-25]</sup>以及生物信息学方法<sup>[26-27]</sup>预测。如今，通过高通量测序，获得了很多动植物的 miRNA，但美洲鲎 miRNA 未见报道。然而，miRNA 只在特定的发展阶段和特定的组织中表达。因此，直接克隆和测序获得的 miRNAs 有一定的局限性<sup>[28-30]</sup>。

miRNA 具有保守性，但不是每一个 miRNA 都是严格保守的，不同物种间 miRNA 碱基序列、二级折叠结构可能存在一定的差异，所以本文研究内容限于具有物种间保守性的 miRNA。今后对于生物信息学预测的 miRNA 需要通过基因克隆、Northern blot、RT-PCR 等生物学实验方法进行验证。而研究 miRNA 特定生理阶段的表达以及发现某物种特异的 miRNA 可以通过 microRNA 高通量测序进行。

3.2 miRNA 的保守性分析

本研究以 miR-10 家族为例，分析 miRNA 的保守性，结果表明，miRNA 前体在分类地位相差很远的物种中具有高度的保守性和同源性，特别是成熟区序列，说明 miRNA 的调控机制是一种古老的细胞生命活动调控途径，在非常低等的动物中已经出现，在进化高等的物种仍然保留。而且 miRNA 的调控在细胞生命活动中广泛存在，据推测，人类基因组上约有三分之一的基因都受到 miRNA 的调控<sup>[31]</sup>。对 miRNA 的功能研究发现，miRNA 参与了包括细胞增殖、分化、凋亡、癌症发生等各种生理病理过程<sup>[32]</sup>。

本研究通过生物信息学方法获得了 49 条 miRNA，对这些 miRNA 的生物信息学分析，加深了人们对海生节肢动物系统发生、基因表达调控等各方面的理解，为后续研究美洲鲎生理、胚胎发育、生殖等生命活动调控奠定了基础，对这些 miRNA 的生物信息学预测结果的进一步利用可以从实验验证和扫描其组织表达入手。作为无脊椎动物中进化较为原始的一类动物，美洲鲎在免疫、胚胎发育、细胞学等各个领域都是研究海生节肢动物系统发生的重要参照，同时也是研究动物进化史的珍贵材料。

3.3 miRNA 的多样性

本文筛选到的 49 个 miRNA 成员代表着 40 个 miRNA 家族，但这 49 个成员并不是均匀地分布在 40 个家族里。其中：miR467 家族最多，有 7 个家族成员；miR10 家族有 2 个家族成员；miR548 家族

有2个家族成员;其余 miRNA 家族均只有1个家族成员。这表明 miRNA467 家族种类具有多样性,预计可能在美洲鲎各种生理活动调控上有重要作用。

美洲鲎的 miRNA 前体在结构上存在多样性。同一 miRNA 的前体结构有多个,比如 miR-466i。不同 miRNA 的前体结构也不相同。这也反映在不同 miRNA 的形成以及在基因表达调控中的作用不同上。美洲鲎 miRNA 的多样性还表现在成熟序列在前体中位置上的多样性,miR-8466i、miR-10、miR-2361和 miR-4705 等17个 miRNA 位于其前体的5'端,其余32个 miRNA 位于其前体的3'端上。

### 3.4 miRNA 功能举例分析

本研究筛选所得的 miRNA 中,以 lpo-miR-34 为例进行分析。miR-34 家族在进化上相对保守,2007 年被发现为 p53 的调控基因,它可以抑制癌细胞生长和扩散<sup>[33-34]</sup>。在其他物种中, lpo-miR-34 参与细胞内凝固抑制剂 II 型、丝氨酸蛋白水解酶、血细胞转谷氨酰胺酶这3种生化进程所需蛋白质的调控。而这些生化进程主要与血液凝固的复杂机制有关,推测 lpo-miR-34 主要参与美洲鲎血液凝固生化过程相关基因的调控。

以筛选所得的 lpo-miR-375 为例分析。miR-375 最早是在人的基因组中发现的,在胰腺和大脑中特异性表达<sup>[35]</sup>。在胰腺器官生成时,miR-375 的表达量大量增加。斑马鱼中 miR-375 功能缺失实验的结果证明,miR-375 对  $\beta$  细胞的形成至关重要。敲除 miR-375 后的小鼠,其  $\beta$  细胞大量减少,同时  $\alpha$  细胞大量增加。miR-375 的靶基因是 *MTPN*, *MTPN* 基因能通过编码蛋白质来调控激素分泌和胞外分泌。miR-375 也可以负向调控 *PDK-1* 基因<sup>[36]</sup>。由此推测 miR-375 可能和美洲鲎组织生长有关。

## [ 参 考 文 献 ]

- [1] AVESSON L, REIMEGARD J, WAGNER E G, et al. MicroRNAs in Amoebozoa: deep sequencing of the small RNA population in the social amoeba *Dictyostelium discoideum* reveals developmentally regulated microRNAs. *RNA*, 2012, 18(10): 1771-1782. DOI:10.1261/rna.033175.112.
- [2] BEREZIKOV E, GURYEV V, VAN DE BELT J, et al. Phylogenetic shadowing and computational identification of human microRNA genes. *Cell*, 2005, 120(1): 21-24. DOI:10.1016/j.cell.2004.12.031.
- [3] COLAIACOVO M, SUBACCHI A, BAGNARESI P, et al. A computational-based update on microRNAs and their targets in barley (*Hordeum vulgare* L.). *BMC Genomics*, 2010, 11(1): 1-15. DOI:10.1186/1471-2164-11-595.
- [4] LAGOS-QUINTANA M, RAUHUT R, MEYER J, et al. New microRNAs from mouse and human. *RNA*, 2003, 9(2): 175-179. DOI:10.1261/rna.2146903.
- [5] REESE T A, XIA J, JOHNSON L S, et al. Identification of novel microRNA-like molecules generated from herpesvirus and host tRNA transcripts. *Journal of Virology*, 2010, 84(19): 10344-10353. DOI:10.1128/JVI.00707-10.
- [6] LEE Y, JEON K, LEE J T, et al. MicroRNA maturation: stepwise processing and subcellular localization. *Embo Journal*, 2002, 21(17): 4663-4670. DOI:10.1093/emboj/cdf476.
- [7] LEE Y S, NAKAHARA K, PHAM J W, et al. Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell*, 2004, 117(1): 69-81. DOI:10.1016/s0092-8674(04)00261-2.
- [8] KROL J, LOEDIGE I, FILIPOWICZ W. The widespread regulation of microRNA biogenesis, function and decay. *Nature Reviews Genetics*, 2010, 11: 597-610. DOI:10.1038/nrg2843.
- [9] BUSHATI N, COHEN S M. MicroRNA functions. *Annual Review of Cell and Developmental Biology*, 2007, 23: 175-205.
- [10] RANA T M. Illuminating the silence: understanding the structure and function of small RNAs. *Nature Reviews Molecular Cell Biology*, 2007, 8(1): 23-36. DOI:10.1038/nrm2085.
- [11] FRIEDMAN R C, FARH K K H, BURGE C B, et al. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Research*, 2009, 19(1): 92-105. DOI:10.1101/gr.082701.108.
- [12] NAGUIBNEVA I, AMEYAR-ZAZOUA M, POLESSKAYA A, et al. The microRNA miR-181 targets the homeobox protein Hox-A11 during mammalian myoblast differentiation. *Nature Cell Biology*, 2006, 8(3): 278-284.
- [13] WIENHOLDS E, KLOOSTERMAN W P, MISKA E, et al. MicroRNA expression in zebrafish embryonic development.

- Science, 2005, 309(5732): 310-311. DOI:10.1126/science.1114519.
- [14] LIVAK K J, SCHMITTGEN T D. Analysis of relative gene expression data using real-time quantitative PCR and the method. *Methods*, 2001, 25(4): 402-408. DOI:10.1006/meth.2001.1262.
- [15] 柳承璋. 淡水枝角水蚤 (*Daphnia pulex*) 微小 RNA (miRNA) 的生物信息学发掘与分析. *海洋与湖沼*, 2013, 44(4): 837-845.
- [16] BONNET E, WUYTS J, ROUZE P, et al. Detection of 91 potential in plant conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important target genes. *Proceedings of the National Academy of Sciences of the United States of America*, 2004, 101(31): 11511-11516. DOI:10.1073/pnas.0404025101.
- [17] LIM L P, GLASNER M E, YEKTA S, et al. Vertebrate microRNA genes. *Science*, 2003, 299(5612): 1540.
- [18] ALTSCHUL S F, GISH W, MILLER W, et al. Basic local alignment search tool. *Journal of Molecular Biology*, 1990, 215: 403-410. DOI:10.1016/s0022-2836(05)80360-2.
- [19] HOFACKER I L. RNA secondary structure server. *Nucleic Acids Research*, 2003, 31(13): 3429-3431.
- [20] ZUKER M, STIEGLER P. Optimal computer folding of large RNA sequences using thermo-dynamics and auxiliary information. *Nucleic Acids Research*, 1981, 9(1): 133-148. DOI:10.1093/nar/9.1.133.
- [21] JIANG P, WU H, WANG W, et al. MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features. *Nucleic Acids Research*, 2007, 35: 339-344. DOI:10.1093/nar/gkm368.
- [22] RAMACHANDRA R K, SALEM M, GAHR S, et al. Cloning and characterization of microRNAs from rainbow trout (*Oncorhynchus mykiss*): their expression during early embryonic development. *BMC Developmental Biology*, 2008, 8: 41. DOI:10.1186/1471-213X-8-41.
- [23] SUNKAR R, GIRKE T, JAIN P K, et al. Cloning and characterization of MicroRNAs from rice. *Plant Cell*, 2005, 17(5): 1397-1411. DOI:10.1105/tpc.105.031682.
- [24] CHI W, TONG C, GAN X, et al. Characterization and comparative profiling of MiRNA transcriptomes in bighead carp and silver carp. *PLoS One*, 2011, 6(8): e23549. DOI:10.1371/journal.pone.0023549.
- [25] SOARES A R, PEREIRA P M, SANTOS B, et al. Parallel DNA pyrosequencing unveils new zebrafish microRNAs. *BMC Genomics*, 2009, 10(17): 195. DOI:10.1186/1471-2164-10-195.
- [26] JONES-RHOADES M W, BARTEL D P. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Molecular Cell*, 2004, 14(6): 787-799. DOI:10.1016/j.molcel.2004.05.027.
- [27] LI L, XU J Z, YANG D Y, et al. Computational approaches for microRNA studies: a review. *Mammalian Genome Official Journal of the International Mammalian Genome Society*, 2010, 21(1/2): 1-12. DOI:10.1007/s00335-009-9241-2.
- [28] BRAMEIER M. Genome-wide comparative analysis of microRNAs in three non-human primates. *BMC Research Notes*, 2010, 3(1): 64. DOI:10.1186/1756-0500-3-64.
- [29] WEBER M J. New human and mouse microRNA genes found by homology search. *FEBS Journal*, 2005, 272(1): 59-73.
- [30] ZHANG B H, PAN X P, WANG Q L, et al. Identification and characterization of new plant microRNAs using EST analysis. *Cell Research*, 2005, 15(5): 336-360. DOI:10.1038/sj.cr.7290302.
- [31] 安丽娜, 董斐斐, 王国坤, 等. MicroRNA 与动脉粥样硬化. *国际心血管病杂志*, 2013, 40(4): 210-212.
- [32] 李青, 荆清. microRNA: 心血管发育及疾病中的重要调控因子. *生命科学*, 2010, 22(7): 661-667.
- [33] BONAUEER A, CARMONA G, IWASAKI M, et al. MicroRNA-92a controls angiogenesis and functional recovery of ischemic tissues in mice. *Science*, 2009, 324(5935): 1710-1713. DOI:10.1126/science.1174381.
- [34] XIAO C, SRINIVASAN L, CALADO D P, et al. Lymphoproliferative disease and autoimmunity in mice with increased miR-17-92 expression in lymphocytes. *Nature Immunology*, 2008, 9(4): 405-414. DOI:10.1038/ni1575.
- [35] BEJARANO F, SMIBERT P, LAI E C. MiR-9a prevents apoptosis during wing development by repressing *Drosophila* LIM-only. *Developmental Biology*, 2010, 338(1): 63-73. DOI:10.1016/j.ydbio.2009.11.025.
- [36] BIRYUKOVA I, ASMAR J, ABDESSELEM H, et al. *Drosophila* miR-9a regulates wing development via fine-tuning expression of the LIM only factor, dLMO. *Developmental Biology*, 2009, 327(2): 487-496. DOI:10.1016/j.ydbio.2008.12.036.

(责任编辑 马建华 英文审校 刘静雯)