

# 体内与体外 RNA-RNA 相互作用的比较初探

陆启轩<sup>1,2</sup>, 查磊<sup>2,Δ</sup>, 李宗城<sup>2</sup>, 陈临溪<sup>1</sup>, 李伍举<sup>2</sup>, 应晓敏<sup>2</sup>

**【摘要】** 目的 通过比较 *in vitro* 与 *in vivo* 的 RNA-RNA 相互作用 (RNA-RNA interaction, RRI), 探究通过 *in vitro* RRI 推测 *in vivo* RRI 的可靠性。方法 采用 perl 语言编写脚本分析酵母转录组水平 *in vitro* RNA 的二级结构信息, 得到可能的 *in vitro* RRI, 再与酵母的 *in vivo* 小核仁 RNA (snoRNA) -rRNA 相互作用进行比较。结果 发现 *in vitro* snoRNA-rRNA 相互作用与 *in vivo* snoRNA-rRNA 相互作用的重叠率仅为 23.42% (26/111); 而 *in vitro* 测定的 snoRNA 双链片段与 *in vivo* 测定的参与 RRI 的 snoRNA 片段重叠率为 38.78% (19/49); *in vitro* 测定的 rRNA 双链片段与 *in vivo* 测定的参与 RRI 的 rRNA 片段重叠率为 80.70% (46/57)。结论 *in vitro* 和 *in vivo* 条件下 snoRNA-rRNA 的相互作用差异很大, 提示 *in vitro* 条件下测定的 snoRNA-rRNA 的相互作用不能真实反映它们在 *in vivo* 的相互作用。

**【关键词】** RNA-RNA 相互作用; 生物信息学; RNA, 小核仁; 核糖体 RNA

**【中图分类号】** Q522

**【文献标志码】** A

**【文章编号】** 1674-9960(2013)07-0517-04

**DOI:** 10.7644/j.issn.1674-9960.2013.07.010

## A systematic comparison of *in vivo* and *in vitro* RNA-RNA interactions

LU Qi-xuan<sup>1,2</sup>, CHA Lei<sup>2,Δ</sup>, LI Zhong-cheng<sup>2</sup>, CHEN Lin-xi<sup>1\*</sup>, LI Wu-ju<sup>2\*</sup>, YING Xiao-min<sup>2\*</sup>

(1. Institute of Pharmacy and Pharmacology, University of South China, Hengyang, Hunan 421001, China; 2. Institute of Basic Medical Sciences, Academy of Military Medical Sciences, Beijing 100850, China)

\* Co-corresponding authors, YING Xiao-min, Tel: 010-66932301, E-mail: yingxm@bmi.ac.cn; LI Wu-ju, E-mail: liwj@bmi.ac.cn; CHEN Lin-xi, E-mail: chenlinxi@tom.com

**【Abstract】 Objective** To compare *in vitro* and *in vivo* RNA-RNA interactions (RRI) systematically in order to analyze the reliability of obtaining *in vivo* RRI by determining *in vitro* RRI. **Methods** Programs were written using perl language to obtain all the candidate *in vitro* RRIs based on transcriptome-wide *in vitro* RNA secondary structure profile in yeast. The *in vitro* RRIs were then compared with *in vivo* small nucleolar RNA (snoRNA) -rRNA interactions in yeast. **Result** Compared with *in vivo* snoRNA-rRNA interactions, only 23.42% (26/111) of *in vitro* snoRNA-rRNA interactions were reproduced in *in vivo* snoRNA-rRNA interactions, whereas 38.78% (19/49) of *in vitro* snoRNA double-stranded fragments overlapped with *in vivo* snoRNA fragments that participated in snoRNA-rRNA interactions, and 80.70% (46/57) of *in vitro* RDN37-I double-stranded fragments overlapped *in vivo* RDN37-I fragments that participated in snoRNA-rRNA interactions. **Conclusion** There is huge difference between *in vitro* and *in vivo* snoRNA-rRNA interactions. snoRNA-rRNA interactions determined *in vitro* cannot really reflect *in vivo* interactions.

**【Key words】** RNA-RNA interactions; bioinformatics; snoRNA; rRNA

RNA-RNA 相互作用 (RNA-RNA interaction, RRI) 是与蛋白-蛋白相互作用、蛋白-DNA 相互作用并列的一类重要的分子间相互作用。它在生物体内广泛存在, 最为经典的 RRI 是在翻译过程中 tR-

NA 识别 mRNA 的三联密码子。此外, RRI 更多地以非编码 RNA (non-coding RNA, ncRNA) 与靶 RNA 之间相互作用的形式出现, 如真核生物的微小 RNA (microRNA, miRNA) 与靶 mRNA、小核仁 RNA (small nucleolar RNA, snoRNA) 与靶 rRNA、长链非编码 RNA (long non-coding RNA, lncRNA) 与靶 mRNA, 以及原核生物的小 RNA (small RNA, sRNA) 与靶 mRNA。这些 RRI 发挥着基本而重要的生物学功能, 包括调节细菌外膜蛋白的生成<sup>[1]</sup>、群体感应<sup>[2]</sup>和毒力基因的表达<sup>[3]</sup>, 调控线虫幼虫发育时序<sup>[4]</sup>、癌症的发生和转移<sup>[5,6]</sup>等。

**【基金项目】** 国家 973 计划资助项目 (2010CB912801); 国家自然科学基金资助项目 (31241404, 30901577, 81270420, 31071157)

**【作者简介】** 陆启轩, 硕士研究生, 研究方向: 生物信息学, E-mail: luqixuan6@163.com; Δ 共同第一作者, 查磊, 助理研究员, 研究方向: 生物信息学, E-mail: leicha@bmi.ac.cn

**【作者单位】** 1. 南华大学药物药理研究所, 湖南衡阳 421001; 2. 军事医学科学院基础医学研究所, 北京 100850

**【通讯作者】** 应晓敏, Tel: 010-66932301, E-mail: yingxm@bmi.ac.cn; 李伍举, E-mail: liwj@bmi.ac.cn; 陈临溪, E-mail: chenlinxi@tom.com

研究 RRI 的首要任务是测定 RRI。目前,测定 RRI 主要依靠生物信息学预测和小规模实验验证。由于 RRI 以 RNA 分子间的碱基互补配对的方式结合,因而遵循一些热动力学规律,如存在连续配对碱基对(种子区)、2 个 RNA 分子结合能低于打开相互作用位点自身结构的能量等。依据上述原则,已经提出了一些 *in silico* RRI 预测算法,如真核生物 miRNA-mRNA 相互作用预测算法<sup>[7,8]</sup>、真核生物 snoRNA-rRNA 相互作用预测算法<sup>[9]</sup>。本课题组在细菌 RRI 领域开展了一些研究工作,提出了细菌 sRNA-mRNA 相互作用预测算法 sTarPicker 和 sRNA-Target<sup>[10-12]</sup>。生物信息学方法可以实现转录组水平的 RRI 预测,但其最大的问题是准确性有限,还需要通过实验方法验证预测的 RRI 在体内是否真实发生。

目前测定 RRI 的实验方法均只能针对特定的一对 RNA-RNA 相互作用或与特定蛋白结合的 RNA 之间的相互作用。X 射线衍射晶体分析法(X-ray crystallography)、核磁共振(NMR)<sup>[13]</sup>、补骨脂素交联法( psoralen cross-linking)<sup>[14]</sup>、酵母 RNA 杂交系统(yeast RNA-hybrid system)<sup>[15]</sup>、电泳迁移率变动分析(electrophoretic mobility shift assay, EMSA)<sup>[16]</sup>、荧光报告基因以及遗传学方法可用于分析 RRI,但这些传统方法只能针对特定的一对或几对 RRI 及其作用位点进行研究,且耗时耗力效率较低。2011 年, Kudla 等<sup>[17]</sup>采用交联、拼接和杂合体测序法(cross-linking, ligation and sequencing of hybrids, CLASH)测定了酵母中与 Nop1、Nop56 和 Nop58 三个蛋白形成复合物的 *in vivo* snoRNA-rRNA 相互作用。CLASH 方法采用近几年刚兴起的高通量测序技术测定与特定蛋白结合的 RRI,较传统实验方法在通量上有了很大的提高;然而,对于那些与未知蛋白结合、甚至不依赖蛋白的 RRI,该方法依然有很大的局限性,因此不能实现转录组水平的 RRI 测定。

目前 RRI 研究已经受限于实验技术的发展。如能对 RRI 进行转录组水平的测定,则能极大地促进我们对包含染色质、蛋白、RNA 在内的三元分子调控网络的认识。

目前,转录组水平 *in vivo* RRI 的测定还存在很大的技术难点。例如,如何解析所有 RNA 结合蛋白、不依赖蛋白的 RRI 如何固化等。然而,转录组水平 *in vitro* RNA 二级结构的测定近几年有了很大的突破,Kertesz 等<sup>[18]</sup>和 Underwood 等<sup>[19]</sup>两个实验室分别提出了采用高通量测序技术测定酵母和小鼠细胞转录组水平 *in vitro* RNA 的二级结构。通过对这些数据的生物信息学分析,可以得到所有可能的

RNA 分子内和分子间相互作用,使得测定 *in vitro* RRI 已经可以实现。那么,能否通过测定 *in vitro* RRI 推测 *in vivo* RRI,而且,在实际应用中经常采用 *in vitro* 实验验证 RRI,但有关 *in vitro* 和 *in vivo* RRI 的系统比较,目前国内外尚无相关报道。为此,本文首次从计算生物学角度对近 2 年发表的 2 组分别来自酵母 *in vitro* 和 *in vivo* RRI 进行了比较分析,初步探究 *in vitro* 和 *in vivo* RRI 的差异。

## 1 材料与方法

### 1.1 数据

*in vivo* RRI 数据来自 Kudla 等测定的酵母 *in vivo* 条件下与 Nop1、Nop56 和 Nop58 结合的 snoRNA-rRNA 相互作用。Kudla 等对 HTP 标记的小核仁核糖核蛋白(snoRNP)的酵母进行紫外线照射交联后,通过 RNase-IT 酶(RNase A 与 T1 混合酶)消化降解游离 RNA,再利用 T4 连接酶连接 snoRNP 保护的 RNA 片段形成嵌合体,采用高通量测序测定嵌合体片段并定位分析,得到 *in vivo* 相互作用的 snoRNA-rRNA 片段。测得的 *in vivo* RRI 共计 253 对,涉及 46 个 snoRNA 和 1 个 rRNA(RDN37-1)。

*in vitro* RRI 数据来自 Kertesz 等测定的酵母转录组水平 *in vitro* RNA 二级结构。Kertesz 等提取酵母总 RNA,变性后复性,而后分别加入单链剪切酶 RNase S1 和双链剪切酶 RNase V1 对 RNA 进行剪切,再采用高通量测序技术对剪切产物进行测序,得到 RNA 每个碱基二级结构状态的 PARS(parallel analysis of RNA structure)分值。PARS 分值为正数,表示该碱基双链的情况多于单链的情况,分值越大,双链情况越多;反之,则表示该碱基单链的情况多于双链的情况,分值越小,单链情况越多。Kertesz 等共测定了酵母中 3083 个 RNA 的 PARS 分值,包括 58 个 snoRNA、5 个 rRNA、14 个 tRNA、3 个 snRNA、3 个 ncRNA 和 3000 个 mRNA。PARS 分值从 <http://genie.weizmann.ac.il/pubs/PARS10> 下载。通过生物信息学分析(详见“1.2 方法”),得到所有可能的 *in vitro* RRI。

本研究中,取 *in vivo* 和 *in vitro* 两组数据共同存在的 snoRNA 和 rRNA 进行比较分析。在 253 对 *in vivo* snoRNA-rRNA 相互作用中,241 对涉及的 snoRNA(40 个)和 rRNA(1 个,RDN37-1)是两组数据均测得的。在 *in vitro* 实验中,Kertesz 等在对测序读段基因定位时,统一定位到成熟 RNA 序列上,其中,RDN37-1 rRNA 是长度为 5354 nt 的所谓“成熟体”序列。然而,RDN37-1 实际长度为 6858 nt,还包含 2 个外转录间隔区(external transcribed spacers, 5'

ETS: 700 nt, 3' ETS: 211 nt) 和 2 个内转录间隔区 (internal transcribed spacers, ITS1: 361 nt, ITS2: 232 nt), 也即 *in vivo* 实验中片段定位的 RDN37-1 序列。鉴于此, 将 *in vivo* RRI 数据中 RDN37-1 的作用位点位于 ETS 和 ITS 的 19 对相互作用去除, 得到 222 对 snoRNA-rRNA 相互作用。由于 *in vitro* 技术未能测定位于 snoRNA 和 RDN37-1 序列末端碱基的二级结构信息, 因此, 相关的 111 对 snoRNA-rRNA 也被去除, 剩余 111 对 RRI 用于后续的 *in vivo* 与 *in vitro* RRI 比较分析。

在 111 对 RRI 中, *in vivo* 测得的 snoRNA 和 rRNA 作用片段分别有 49 和 57 个互不相同的片段。

## 1.2 方法

根据 *in vitro* 测得的 RNA 二级结构 PARS 值, 并依据以下 2 个原则判断可能的 RRI。

(1) 双链片段判定原则: 以 PARS 值  $\geq 2$  的碱基为中心, 分别向 5' 和 3' 方向延伸, 要求延伸碱基的 PARS 值  $\geq 0$ , 且连续 PARS 值等于 0 的碱基数  $\leq 4$ , 这样所得到的片段, 认为是形成双链的片段。

(2) *in vitro* RRI 判定原则: 任何两个双链片段, 都是可能的 RRI。这是一个极限宽松的 RRI 判定原则。之所以采用这一原则, 就是为了考察在极限宽松条件下 *in vitro* 和 *in vivo* RRI 的差异是否明显。如果在极限宽松条件下 *in vitro* 与 *in vivo* RRI 的差异都很大, 则任何更为严格的 *in vitro* RRI 判定原则都只会增加 *in vitro* 和 *in vivo* RRI 的差异。

此外 *in vitro* 和 *in vivo* RRI 是否一致的判定原则: 若 *in vitro* RRI 中的 snoRNA 和 rRNA 与 *in vivo* RRI 中的 snoRNA 和 rRNA 分别重叠  $\geq 5$  个碱基, 则认为它们的 RRI 一致。

## 2 结果

对得到的 111 对 *in vitro* 和 *in vivo* snoRNA-RDN37-1 相互作用进行比较分析, 发现 26 对 snoRNA-RDN37-1 在 *in vitro* 和 *in vivo* RRI 中一致, 重复率为 23.42% (图 1)。这一结果提示, snoRNA 与 RDN37-1 的相互作用在体内 (*in vivo*) 与体外 (*in vitro*) 差异很大。

为考察 *in vitro* 与 *in vivo* snoRNA-RDN37-1 相互作用差异巨大的原因, 进一步分析了参与 snoRNA-RDN37-1 相互作用的 2 个片段, 即 snoRNA 和 RDN37-1 各自在 *in vitro* 和 *in vivo* 条件下测定的二级结构的一致性。结果发现, 在 48 个参与 *in vivo* snoRNA-RDN37-1 相互作用的 snoRNA 片段中, 19 个片段与 *in vitro* snoRNA-RDN37-1 相互作用中的 snoRNA 片段一致, 重复率为 38.78% (图 1 图 2A),

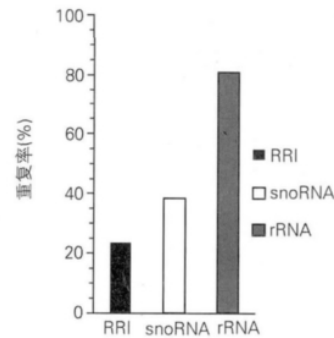


图 1 RRI, snoRNA 和 rRNA 在 *in vitro* 与 *in vivo* RRI 中的重复率

略高于 snoRNA-RDN37-1 相互作用的重复率; 而在 57 个参与 *in vivo* snoRNA-RDN37-1 相互作用的 RDN37-1 片段中, 46 个片段与 *in vitro* snoRNA-RDN37-1 相互作用中的 RDN37-1 片段一致, 重复率高达 80.70% (图 1, 图 2B), 远高于 snoRNA-rRNA 相互作用的重复率。这一结果提示, *in vitro* 和 *in vivo* snoRNA-rRNA 相互作用的巨大差异很大程度上缘于 snoRNA 在 *in vitro* 与 *in vivo* 中二级结构的巨大差异, 同时也说明 RNA 二级结构对 RRI 具有重要影响。

## 3 讨论

在实验研究中 *in vitro* 的实验结果常用于说明生物学问题, 而 *in vivo* 的实验结果往往更能反映生物体的真实作用情况。众所周知, *in vitro* 与 *in vivo* 的环境不同, 势必有着不同的生物反应。本文从计算生物学的角度, 对 *in vitro* 与 *in vivo* snoRNA-rRNA 相互作用进行了分析, 发现 *in vitro* 与 *in vivo* snoRNA-rRNA 的相互作用的确存在很大差异, *in vitro* 分析得到的 RRI 仅重现了 *in vivo* 中 snoRNA-rRNA 相互作用的 23.42%。这一结果揭示, *in vitro* 与 *in vivo* 的 RRI 差异很大。因此, 不能仅根据 *in vitro* 的研究结果来说明生物学问题。

本文对 *in vitro* 和 *in vivo* 条件下二级结构的进一步分析结果表明, *in vitro* 和 *in vivo* RRI 的巨大差异很大程度上可能是由于 snoRNA 在 *in vitro* 和 *in vivo* 条件下结构的巨大差异导致的。snoRNA 是长度在 60 ~ 300 nt 的小 RNA 分子, 结构不稳定, *in vivo* 条件下是以与蛋白质结合形成复合物 snoRNP 的形式存在, 而在 *in vitro* 蛋白条件改变的情况下, 结构与 *in vivo* 条件下差别很大。事实上, 既往的研究也表明, snoRNA 是先与蛋白形成 snoRNP, 而后作用于 rRNA, 并指导其甲基化等修饰<sup>[20]</sup>; 而 rRNA 是长度大于 5000 nt 的大 RNA 分子, 结构比较稳定, 在

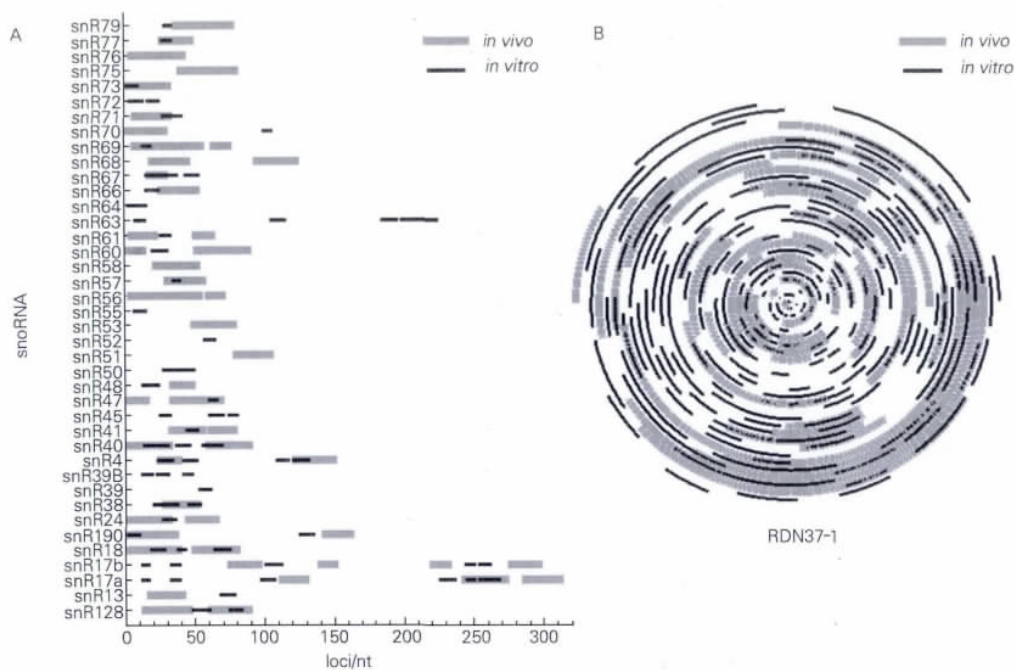


图2 snoRNA与RDN37-1在*in vitro*与*in vivo*实验中双链片段的重叠情况

A. snoRNA在*in vitro*与*in vivo*条件下双链片段的重叠情况,横坐标为snoRNA的位点,纵坐标为40个snoRNA(根据snoRNA在*in vitro*与*in vivo*条件下双链片段的位点信息,采用Matlab绘图);B. RDN37-1在*in vitro*与*in vivo*条件下双链片段的重叠情况,图以RDN37-1起始位点为圆点的螺旋(根据RDN37-1 rRNA在*in vitro*与*in vivo*条件下双链片段的位点信息,采用Matlab绘图);loci/nt片段定位/核苷酸

*in vitro*和*in vivo*条件下结构差别不大。

总之,*in vitro*实验测定的RRI不能真实反映*in vivo* RRI,对*in vitro* RRI的结果要谨慎对待,需要用*in vivo*的实验方法验证分析。要实现转录组水平*in vivo* RRI的测定,必须发展新的高通量*in vivo* RRI测定方法。

#### 【参考文献】

- Guillier M, Gottesman S. Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs [J]. Mol Microbiol, 2006, 59(1): 231-247.
- Tu KC, Bassler BL. Multiple small RNAs act additively to integrate sensory information and control quorum sensing in *Vibrio harveyi* [J]. Genes Dev, 2007, 21(2): 221-233.
- Toledo-Arana A, Repoila F, Cossart P. Small noncoding RNAs controlling pathogenesis [J]. Curr Opin Microbiol, 2007, 10(2): 182-188.
- Reinhart BJ, Slack FJ, Basson M, et al. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans* [J]. Nature, 2000, 403(6772): 901-906.
- Tavazoie SF, Alarcón C, Oskarsson T, et al. Endogenous human microRNAs that suppress breast cancer metastasis [J]. Nature, 2008, 451(7175): 147-152.
- Huang Q, Gumireddy K, Schrier M, et al. The microRNAs miR-373 and miR-520c promote tumour invasion and metastasis [J]. Nat Cell Biol, 2008, 10(2): 202-210.
- Kertesz M, Iovino N, Unnerstall U, et al. The role of site accessibility in microRNA target recognition [J]. Nat Genet, 2007, 39(10): 1278-1284.
- Yousef M, Jung S, Kossenkov AV, et al. Naïve Bayes for microRNA target predictions-machine learning for microRNA targets [J]. Bioinformatics, 2007, 23(22): 2987-2992.
- Kehr S, Bartschat S, Stadler PF, et al. PLEXY: Efficient target prediction for box C/D snoRNAs [J]. Bioinformatics, 2011, 27(2): 279-280.
- Ying X, Cao Y, Wu J, et al. sTarPicker: a method for efficient prediction of bacterial sRNA targets based on a two-step model for hybridization [J]. PLoS One, 2011, 6(7): e22705.
- Zhao Y, Li H, Hou Y, et al. Construction of two mathematical models for prediction of bacterial sRNA targets [J]. Biochem Biophys Res Commun, 2008, 372(2): 346-350.
- Cao Y, Zhao Y, Cha L, et al. sRNATarget: a web server for prediction of bacterial sRNA targets [J]. Bioinformatics, 2009, 3(8): 364.
- Westhof E. RNA modeling, naturally [J]. Proc Natl Acad Sci USA, 2012, 109(8): 2691-2692.
- Calvet JP, Pederson T. Base-pairing interactions between small nuclear RNAs and nuclear RNA precursors as revealed by psoralen cross-linking *in vivo* [J]. Cell, 1981, 26(3): 363-370.
- Piganeau N, Schauer UE, Schroeder R. A yeast RNA-hybrid system for the detection of RNA-RNA interactions *in vivo* [J]. RNA, 2006, 12(1): 177-184.
- Morita T, Maki K, Aiba H. Detection of sRNA-mRNA Interactions by Electrophoretic Mobility Shift Assay [M]//Keiler KC. Bacterial Regulatory RNA. Totowa, New Jersey: Humana Press, 2012: 235-244.
- Kudla G, Granneman S, Hahn D, et al. Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast [J]. Proc Natl Acad Sci USA, 2011, 108(24): 10010-10015.
- Kertesz M, Wan Y, Mazor E, et al. Genome-wide measurement of RNA secondary structure in yeast [J]. Nature, 2010, 467(7311): 103-107.
- Underwood JG, Uzilov AV, Katzman S, et al. FragSeq: transcriptome-wide RNA structure probing using high-throughput sequencing [J]. Nat Methods, 2010, 7(12): 995-1001.
- Bachellet JP, Cavallé J, Hüttenhofer A. The expanding snoRNA world [J]. Biochimie, 2002, 84(8): 775-790.

(孙承媛 编辑 2013-03-01 收稿)