

文章编号: 1004-0374(2010)09-0919-06

一种含有端粒重复序列非编码RNA的发现及最新研究进展

刘娟, 邢金良*

(第四军医大学基础部细胞工程中心, 西安710032)

摘要: 端粒是染色体末端的特化结构, 由简单呈串联线性排列的核酸重复序列及相关蛋白质组成。其核酸序列具有高度的保守性, 均富含GC。在人类为TTAGGG的高度重复序列具有维持基因组完整性的作用。端粒功能异常会导致染色体失去稳定性, 促进肿瘤的发生和发展。以往认为端粒附近区域不具有转录活性, 但最近在Science杂志上Azzalin等首次报道了该区域可以转录一种非编码RNA, 即端粒RNA(telomeric RNA)。该分子具有特殊的UUAGGG重复序列, 在调控端粒长度和端粒酶活性上具有重要作用, 在发育、衰老和肿瘤发生发展等研究中已成为热点。该文将对近期有关端粒RNA的研究进展予以综述。

关键词: 端粒RNA; 端粒重复序列; 非编码RNA

中图分类号: Q522; R730.231.3

文献标识码: A

The discovery and research advances of a telomeric repeat-containing non-coding RNA

LIU Juan, XING Jin-liang*

(Cell Engineering Research Center, Department of Cell Biology, Fourth Military Medical University, Xi'an 710032, China)

Abstract: Telomeres, the specialized structure of chromosomes ends, consist of tandem arrays of DNA repeats and related proteins. Telomeres are conserved during biological evolution and rich in GC nucleotide. Telomeres contain TTAGGG repeats in human, which are responsible for maintaining genomic integrity. Telomeres' dysfunction leads to chromosome instability, and contributes to tumor development. It has long been considered that telomeres are transcription-silenced. However, a recent study reported for the first time that a long non-coding RNA was transcribed from this region, named telomeric RNA. The molecule has a special sequence of UUAGGG repeats, plays an important role in the regulation of telomere length and telomerase activity, and has already become a hot spot in the research field of development, aging and tumor progression. This article will review the recent progress of telomeric RNA research.

Key words: telomeric RNA; telomeric repeats; non-coding RNA

端粒(telomere)作为一个染色体末端结构可以防止染色体的降解和染色体之间的融合及重组, 从而维持基因组完整性和稳定性。研究表明: 由于长度缩短造成的端粒功能异常会导致细胞周期阻滞和细胞凋亡的发生; 在DNA损伤检测点功能受损的情况下, 关键性的端粒缩短会引起染色体断裂-融合-再连接而导致染色体失去稳定性, 从而促进肿瘤

的发生^[1-3]。另外, 端粒功能缺陷的细胞常常拥有降低的DNA修复能力和复杂的细胞遗传学的异常^[4-6]。

收稿日期: 2010-03-05; 修回日期: 2010-04-19

基金项目: 国家自然科学基金项目(C171002)

*通讯作者: E-mail: xingjinliang@163.com; Tel: 029-84774573

到目前为止, 已有大量的研究利用临床样本证实端粒的长度缩短与多种恶性肿瘤的发生和发展密切相关, 如乳腺癌^[7]、前列腺癌^[8]、肺癌和膀胱癌^[9]等。然而, 有关端粒缩短的发生及调控机制尚未得到详细阐明。

以往认为端粒附近区域不具有转录活性, Azzalin等^[10]在Science杂志上首次报道了该区域可以转录一种非编码端粒RNA。随后, 几个其他的研究小组相继证实了该RNA分子的存在, 并对其生物学特点和功能进行了观察。研究发现: 哺乳动物的大部分组织中都有端粒RNA的表达, 该转录产物仅定位于细胞核内。端粒RNA参与了端粒结构的形成, 可通过抑制端粒酶活性参与端粒长度的调控。另外与正常组织相比, 该RNA分子在几种肿瘤(包括喉癌、结肠癌和B细胞淋巴瘤)组织中表达显著降低, 提示该分子极有可能与肿瘤的发生和进展有关。鉴于该分子与端粒调控及端粒酶活性潜在的密切关系, 如今端粒RNA已成为重点研究的分子。

1 端粒RNA是一种最新发现的非编码RNA

端粒是一个存在于真核细胞染色体末端由一定长度串联重复DNA序列(人的重复序列为TTAGGG)构成的特化结构^[11]。该结构与特定蛋白质(如端粒调控因子TRF1和TRF2等)相结合在染色体两端形成“帽子”一样的“T-环”结构, 防止染色体降解以及染色体之间的融合与重组, 进而维持染色体的完整性和稳定性^[12, 13]。端粒具有异染色质特性, 该区域未发现基因的表达。另外, 端粒序列还具有沉默插入到亚端粒区报告基因的特性。因此, 端粒区域长期以来一直被认为是转录沉默的^[14-17]。截至2007年10月, Azzalin等^[10]在Science杂志上首次报道了端粒及亚端粒区域可以转录一种含有端粒重复序列(UUAGGG)的非编码RNA——端粒RNA(telomeric repeat-containing RNA, TERRA), 打破了该区域不具有转录特性的观点。随后, 在Nature Cell Biology、Molecular Cell、Cell和Stem Cell等高水平杂志上相继报道了端粒RNA的研究成果^[18-20]。众所周知, 有关端粒及端粒酶的研究在2009年获得诺贝尔生理学或医学奖。由于端粒RNA与端粒调控及端粒酶活性可能的密切作用关系, 如今该RNA分子已经成为生命科学领域研究的新热点。

2 端粒RNA的生物学特性

2.1 端粒RNA的产生

端粒RNA的转录起始位点位于亚端粒区。该

分子从亚端粒区域向染色体末端方向转录, 转录产物仅定位于细胞核内。两个最主要的端粒RNA研究小组均证实在哺乳动物的大部分组织中都有端粒RNA的表达^[10, 18]。端粒RNA的长度具有较大变化, 介于100 bp~9 kb之间^[10]。RNA聚合酶II是参与端粒RNA转录的主要聚合酶。然而, 应用放线菌素D特异性抑制RNA聚合酶II的活性后, 仍然可以在细胞内检测到端粒RNA的表达, 该结果说明其他种类的RNA聚合酶也可能参与了端粒RNA的转录^[18, 19]。Dejardin等^[21]也证实在端粒附近存在RNA聚合酶I和RNA聚合酶III, 这一发现为上述推测提供了进一步的证据。同样, Luke等^[19]在酵母中也发现端粒可以被RNA聚合酶II转录。然而, 关于端粒RNA产生的很多问题尚未得到解答, 如该RNA分子准确的转录起始位点在何处; 是否每条染色体的末端均可转录; 是否存在特异性染色体末端转录现象; 在特定的器官、组织和细胞类型, 或者特定的生理及病理状态(如各种慢性疾病等)下是否存在特征性的染色体端粒RNA转录谱。

2.2 端粒RNA的加工处理与修饰

同绝大多数RNA聚合酶II的转录产物相类似, 端粒RNA的3'端也可发生多聚腺苷化(加polyA作用)。但是, 目前仍不清楚该分子的多聚腺苷化是转录后直接进行的, 还是先进行剪接然后再加上polyA尾巴。Luke等^[19]发现酵母端粒RNA的多聚腺苷化受到Pap1p和Rat1p(5'-3'外切酶)调控。将酵母中的多聚腺苷化相关基因突变后, 端粒RNA的多聚腺苷化消失, 导致无法检测到端粒RNA, 这个结果也说明polyA尾巴对于维持端粒RNA稳定性具有重要作用; 在酵母的rat1-1突变细胞中端粒RNA增多、端粒长度变短; 过度表达核糖核酸酶H(降解RNA的酶)可防止突变细胞中端粒变短, 表明在这些突变细胞中端粒RNA和端粒DNA以RNA/DNA杂交的形式阻碍了端粒酶在染色体末端的功能。因此, 端粒的转录及其转录产物的降解对调控酵母中端粒酶活性具有重要作用。目前, 关于端粒RNA 5'端的修饰尚未见报道。以往的两项重要研究均发现端粒RNA转录后的降解主要由细胞中的NMD(nonsense-mediated RNA decay)机制介导^[10, 22]。研究结果显示: NMD通路中的UPF1、SMG1和EST1A/SMG6作为主要分子参与了端粒RNA降解的调控; 应用RNAi技术在mRNA水平上分别干涉上述三个分子, 均会导致细胞内端粒RNA水平显著增

高;染色体免疫共沉淀实验进一步表明这些NMD通路蛋白质分子可能直接作用于端粒RNA。

2.3 端粒RNA的转录调控

目前关于端粒RNA转录调控的研究仍处在初步阶段,至今参与端粒RNA转录调控的特异性转录因子尚未得到确认。很多研究证实,哺乳动物端粒含有大量异染色质标志物,如H3K9me3、H4K20me3和HP1,并且亚端粒DNA高度甲基化。如果改变端粒和亚端粒异染色质状态则会造成端粒重组频发^[23-29]。Schoeftner和Blasco^[18]将细胞的*Suv39h*和*Suv4-20h*两种甲基转移酶分别突变后,发现端粒RNA水平显著上调,说明端粒结构更加“开放”进而转录增强。Azzalin等^[30]抑制去乙酰化酶亦可显著上调该RNA分子在人类细胞中的表达水平。Schoeftner和Blasco^[18]还通过免疫共沉淀实验发现端粒结合蛋白TRF1和RNA聚合酶II可发生相互作用。利用RNA干涉技术下调TRF1表达后,端粒RNA的表达水平相应显著下调。然而,由于TRF1仅结合于端粒区而不出现在包含端粒转录起始位点的亚端粒区,并且TRF1的缺失并不影响RNA聚合酶II和端粒区域DNA序列的结合,另外TRF1的过量表达反而降低端粒RNA的表达水平,人们推断TRF1可能并不是传统意义上调控端粒RNA的转录因子,而可能通过间接影响RNA聚合酶II进而发挥调控端粒RNA表达的作用。

3 端粒RNA的生物学功能

3.1 端粒RNA参与端粒结构的形成

首先,Luke等^[19]研究发现端粒RNA可以通过DNA/RNA杂交的方式与端粒DNA发生相互作用。进而,Xu等^[31]证明端粒RNA和端粒DNA在体外可以形成稳定的G四聚体结构(端粒的3'悬突链即鸟嘌呤富集链可以在单链状态和单价阳离子介导的四聚体状态之间保持平衡)。两个研究小组分别在不同的细胞分裂时相(间期和中期)检测到端粒RNA存在于染色体末端结构端粒中^[10,18]。这些研究结果表明,端粒RNA可能是端粒结构的组成部分,在维持端粒结构完整性上有潜在作用,但是由于端粒RNA被发现仅存在于一部分染色体末端,考虑到大部分甚至全部染色体均可能有该RNA分子转录,预示一种可能,即端粒RNA不是端粒永久性的组成部分,可能是端粒组装时的暂时性成分,在特定的染色体末端发挥着调节作用。然而,也许该现象就

是由于仅有部分染色体末端转录或者部分染色体末端转录水平较低导致无法检测引起。另外,尽管发现端粒RNA存在于端粒上,但是端粒RNA是否仅仅与其转录的模板端粒相关联,还是可以从一个端粒移动到另一个端粒尚不得而知。目前也无确切的证据证明存在相关的蛋白质介导端粒RNA和端粒的相互作用。上述推测和问题尚需大量的实验去探讨和验证。

3.2 端粒RNA通过抑制端粒酶活性参与端粒长度的调控

由于理论上端粒RNA的UUAGGG重复序列可以与端粒酶的RNA模板互补配对,因此人们猜测端粒RNA有可能通过结合端粒酶模板区域抑制其活性。为了证明这个假设,Schoeftner和Blasco^[18]首先从培养细胞提取总蛋白,然后分别与(UUAGGG)₄或(CCCUAA)₄RNA探针孵育,随后通过TRAP试剂盒检测端粒酶活性,发现(UUAGGG)₄RNA探针与小鼠胚胎细胞和人类HeLa细胞的蛋白提取物孵育可导致端粒酶活性完全消失;相反,(CCCUAA)₄RNA探针对端粒酶活性无明显影响。基于此,我们提出端粒RNA可能通过与端粒酶中RNA模板结合,形成RNA二聚体,继而抑制端粒酶的活性。Luke等^[19]研究发现端粒RNA可以通过DNA/RNA杂交的方式与端粒DNA发生相互作用;Yehezkel等^[32]也推测其机制为:端粒RNA与端粒DNA以RNA/DNA的方式杂交结合,阻碍DNA复制叉前进,并阻止端粒酶靠近染色体末端,进而导致端粒丢失和严重的DNA损伤。Xu等^[31]发现端粒的3'悬突链和端粒RNA可形成G四聚体。在此基础上,我们假设端粒RNA阻碍端粒酶延长端粒的方式有以下三种:与端粒酶RNA模板或端粒DNA杂交(图1);与端粒的3'悬突链形成G四聚体(图2)。端粒RNA的特殊结构使得上述三种推测均有可能性。

3.3 端粒RNA参与细胞分化及发育

在未分化的胚胎干细胞中,女性的2条X染色体和男性的X、Y染色体上均可检测到端粒RNA的存在。但是,胚胎干细胞分化后,端粒RNA存在的染色体位置发生显著变化,仅存在于女性的一条Xi染色体和男性的Y染色体上^[18,20,33]。另外,Marion等^[20]在由成纤维细胞诱导而来的多能干细胞中检测到端粒RNA的水平显著上升。这些结果提示,端粒RNA在细胞分化发育过程中可能扮演了重要的角色,但其作用机制需进一步研究。

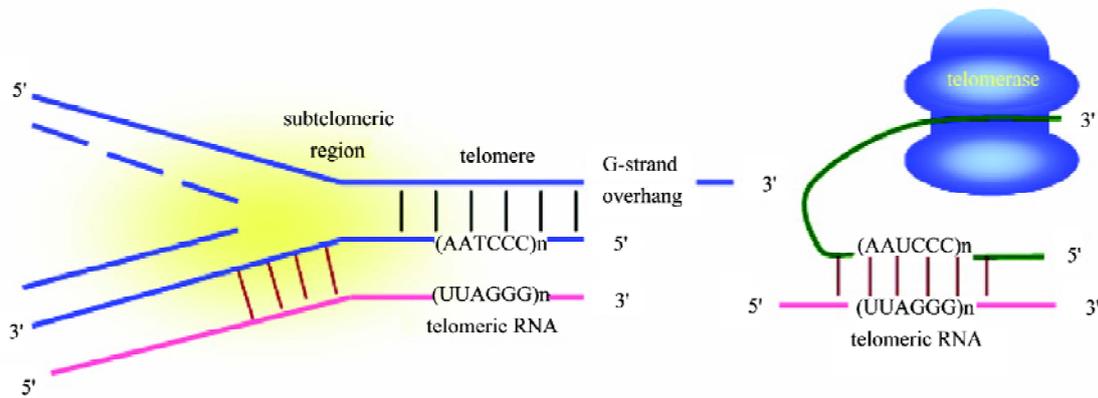


图1 端粒RNA与端粒酶或端粒的结合形式

蓝色线条代表染色体DNA链，粉红色线条代表端粒RNA，黄色区域为亚端粒区，最右边蓝色哑铃状图形为端粒酶，其中的绿色线条为端粒酶的RNA模板。由于端粒RNA包含UUAGGG重复序列，故端粒RNA既可以与端粒DNA中CCCTAA重复序列结合，也可与端粒酶RNA模板的CCC UAA重复序列结合，从而阻止端粒酶延长端粒，甚至阻碍DNA复制

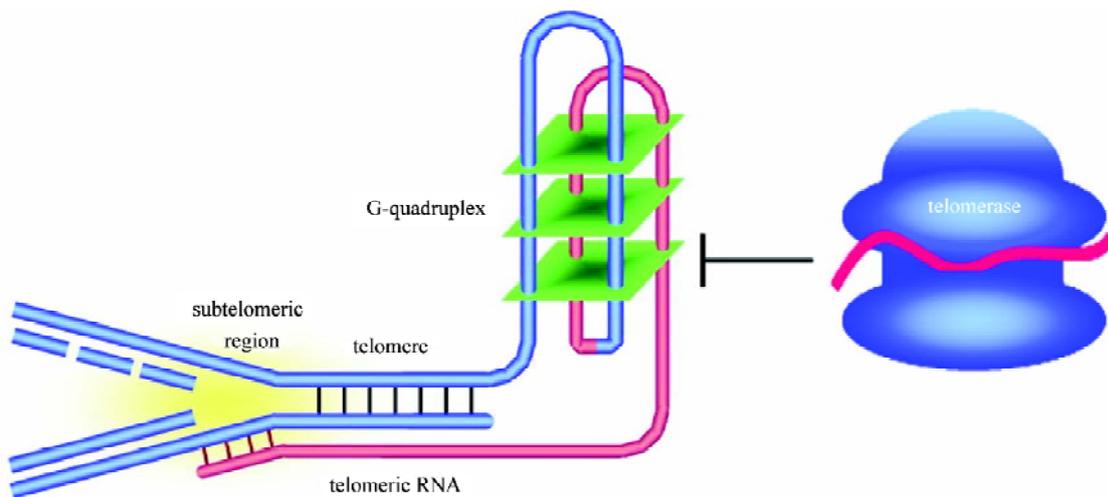


图2 端粒RNA与端粒形成的G-四聚体结构

蓝色线条代表染色体DNA链，粉红色线条代表端粒RNA，黄色区域为亚端粒区，绿色部分为G四聚体，最右边蓝色哑铃状物体为端粒酶，其中的红色线条为端粒酶的RNA模板。端粒RNA与端粒的3'悬突链形成的G四聚体可以阻止端粒酶靠近端粒末端，从而阻止端粒延长

3.4 端粒RNA导致端粒丢失和DNA损伤

3.4.1 端粒RNA降解异常可造成端粒丢失

人类细胞中NMD介导RNA降解机制的关键分子UPF1、hEST1A或SMG1的缺失可以显著上调端粒RNA的表达水平，进而造成端粒快速丢失或严重的DNA损伤^[32, 34]。这些结果提示：NMD因子可能通过调节端粒RNA水平维持正常的染色体末端DNA复制以及端粒长度。

3.4.2 端粒RNA的负反馈回路缺陷可引起DNA损伤自我加强的负反馈回路广泛存在于生物系统。

例如，酵母着丝粒以及哺乳动物失活的X染色体，在异染色质形成过程中存在RNA介导的负反馈回路。Yehezkel等^[32]研究发现端粒RNA通过自我加强的负反馈回路，同样参与了端粒异染色质的形成。缺乏DNMT3b和DNA甲基化是人类ICF综合征(immunodeficiency centromeric instability and facial anomalies)基因损伤的主要原因。在ICF细胞中，端粒RNA的水平异常增多及端粒长度明显缩短造成端粒RNA的负反馈回路中断，从而不能维持亚端粒异染色质的稳定性。这些发现说明RNA介导的负反馈

机制缺陷造成严重的端粒功能失活和染色体不稳定性。

4 端粒 RNA 与肿瘤发生发展的关系

由于端粒 RNA 是最近发现的重要分子,有关研究主要集中在其序列特点以及转录调控等生物学特性方面,加之由于缺乏实用的端粒 RNA 检测方法以及未得到肿瘤学家足够关注等原因,该分子与肿瘤发生发展的关系研究在国内外未得到广泛开展。以往有限的研究发现:在人类喉癌、结肠癌和 B 细胞淋巴瘤组织中,其端粒 RNA 水平比相应正常组织的端粒 RNA 水平显著降低,并且随着肿瘤恶性程度的升高,端粒 RNA 水平逐渐降低。这些结果结合端粒 RNA 功能研究的初步发现提示该分子极有可能与肿瘤的发生和进展有关^[18]。

5 展望

Northern blot 是通常采用的端粒 RNA 检测方法,实时定量 PCR 仅仅在个别染色体特异性端粒 RNA 的检测中有过应用。这些方法或者操作复杂、重复性差,或者通用性不高导致应用(尤其是在大样本的定量检测中)受到较大限制。随着人类基因组计划的完成和测序技术的巨大进展,大部分染色体末端靠近端粒的亚端粒区域序列已得到解析,加上最新发展的第二代高通量测序技术的出现,使得利用生物信息学技术分析全部染色体转录的端粒 RNA 序列特征成为可能。基于这些序列特点,设计相应的引物和杂交探针,利用 PCR 和芯片技术可以实现端粒 RNA 表达的高效检测,进而解决端粒 RNA 相关检测方法的瓶颈问题。

以往关于端粒 RNA 的研究主要集中在其生物学特性方面,包括该分子长度和序列组成特点以及转录方式和调控等。而该分子与疾病,尤其是与肿瘤发生发展的重要关系方面的研究在国内外尚未广泛开展。随着高通量检测方法的建立,人们可以全面探讨该 RNA 分子在肿瘤组织中的表达分布特征及临床相关性,分析其潜在的生物学作用,最后在细胞和动物模型中进行验证和开展分子机制的研究。

[参 考 文 献]

- [1] Brunori M, Luciano P, Geli V, et al. The telomerase cycle: normal and pathological aspects. *J Mol Med*, 2005, 83(4): 244-57
- [2] Shammam MA, Koley H, Munshi NC, et al. Growth arrest, apoptosis, and telomere shortening of Barrett's-associated adenocarcinoma cells by a telomerase inhibitor. *Gastroenterology*, 2004, 126(5): 1337-46
- [3] Counter CM. The roles of telomeres and telomerase in cell life span. *Mutat Res*, 1996, 366(1): 45-63
- [4] Williams ES, Klingler R, Bailey SM, et al. Telomere dysfunction and DNA-PKcs deficiency: characterization and consequence. *Cancer Res*, 2009, 69(5): 2100-7
- [5] Jiang H, Schiffer E, Rudolph KL, et al. Proteins induced by telomere dysfunction and DNA damage represent biomarkers of human aging and disease. *Proc Natl Acad Sci USA*, 2008, 105(32): 11299-304
- [6] Herbig U, Ferreira M, Sedivy JM, et al. Cellular senescence in aging primates. *Science*, 2006, 311(5765): 1257
- [7] Shen J, Terry MB, Santella RM, et al. Short telomere length and breast cancer risk: a study in sister sets. *Cancer Res*, 2007, 67(11): 5538-44
- [8] Meeker AK, Hicks JL, Platz EA, et al. Telomere shortening is an early somatic DNA alteration in human prostate tumorigenesis. *Cancer Res*, 2002, 62(22): 6405-9
- [9] Wu X, Amos CI, Zhu Y, et al. Telomere dysfunction: a potential cancer predisposition factor. *J Natl Cancer Inst*, 2003, 95(16): 1211-8
- [10] Azzalin CM, Reichenbach P, Lingner J, et al. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science*, 2007, 318(5851): 798-801
- [11] Blackburn EH. Telomeres. *Trends Biochem Sci*, 1991, 16(10): 378-81
- [12] Karlseder J. Telomere repeat binding factors: keeping the ends in check. *Cancer Lett*, 2003, 194(2): 189-97
- [13] Wang RC, Smogorzewska A, de Lange T. Homologous recombination generates T-loop-sized deletions at human telomeres. *Cell*, 2004, 119(3): 355-68
- [14] Baur JA, Zou Y, Wright WE, et al. Telomere position effect in human cells. *Science*, 2001, 292(5524): 2075-7
- [15] Gao Q, Reynolds GE, Murnane JP, et al. Telomeric transgenes are silenced in adult mouse tissues and embryonic fibroblasts, but are expressed in embryonic stem cells. *Stem Cells*, 2007, 25(12): 3085-92
- [16] Gottschling DE, Aparicio OM, Zakian VA, et al. Position effect at *S. cerevisiae* telomeres: reversible repression of Pol II transcription. *Cell*, 1990, 63(4): 751-62
- [17] Pedram M, Sprung CN, Murnane JP, et al. Telomere position effect and silencing of transgenes near telomeres in the mouse. *Mol Cell Biol*, 2006, 26(5): 1865-78
- [18] Schoeftner S, Blasco MA. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat Cell Biol*, 2008, 10(2): 228-36
- [19] Luke B, Panza A, Lingner J, et al. The Rat1p 50 to 30 exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in *Saccharomyces cerevisiae*. *Mol Cell*, 2008, 32(4): 465-77
- [20] Marion RM, Strati K, Blasco MA, et al. Telomeres acquire embryonic stem cell characteristics in induced pluripotent stem cells. *Cell Stem Cell*, 2009, 4(2): 141-54
- [21] Dejjardin J, Kingston RE. Purification of proteins associated with specific genomic loci. *Cell*, 2009, 136(1): 175-86
- [22] Chawla R, Azzalin CM. The telomeric transcriptome and

- SMG proteins at the crossroads. *Cytogenet Genome Res*, 2008, 122(3-4): 194-201
- [23] Blasco MA. The epigenetic regulation of mammalian telomeres. *Nat Rev Genet*, 2007, 8(4): 299-309
- [24] Benetti R, Gonzalo S, Jaco I, et al. Suv4-20h deficiency results in telomere elongation and derepression of telomere recombination. *J Cell Biol*, 2007, 178(6): 925-36
- [25] Benetti R, Garcia-Cao M, Blasco MA. Telomere length regulates the epigenetic status of mammalian telomeres and subtelomeres. *Nat Genet*, 2007, 39(2): 243-50
- [26] Gonzalo S, Jaco I, Fraga MF, et al. DNA methyltransferases control telomere length and telomere recombination in mammalian cells. *Nat Cell Biol*, 2006, 8(4): 416-24
- [27] Gonzalo S, Garcia-Cao M, Fraga MF, et al. Role of the RB1 family in stabilizing histone methylation at constitutive heterochromatin. *Nat Cell Biol*, 2005, 7(4): 420-8
- [28] Garcia-Cao M, Jenuwein T, Blasco MA, et al. Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. *Nat Genet*, 2004, 36(1): 94-9
- [29] Garcia-Cao M, Gonzalo S, Blasco MA, et al. A role for the Rb family of proteins in controlling telomere length. *Nat Genet*, 2002, 32(3): 415-9
- [30] Azzalin CM, Lingner J. Telomeres: the silence is broken. *Cell Cycle*, 2008, 7(9): 1161-5
- [31] Xu Y, Kimura T, Komiyama M. Human telomere RNA and DNA form an intermolecular G-quadruplex. *Nucleic Acids Symp Ser*, 2008, (52): 169-70
- [32] Yehezkel S, Segev Y, Selig S, et al. Hypomethylation of subtelomeric regions in ICF syndrome is associated with abnormally short telomeres and enhanced transcription from telomeric regions. *Hum Mol Genet*, 2008, 17(18): 2776-89
- [33] Zhang L, Ogawa Y, Lee JT, et al. Telomeric RNAs mark sex chromosomes in stem cells. *Genetics*, 2009, 182(3): 685-98
- [34] Azzalin CM, Lingner J. The human RNA surveillance factor UPF1 is required for S phase progression and genome stability. *Curr Biol*, 2006, 16(4): 433-9