

# 端粒和端粒酶的发现历史

廖新化

## 1. 染色体末端的两个难题以及端粒的概念

20 世纪 70 年代初，对 DNA 聚合酶特性的深入了解引申出了一个染色体的复制问题。DNA 聚合酶在复制 DNA 的时候必须要有引物来起始，而且它的酶活性具有方向性，只能沿着 DNA 5' 到 3' 的方向合成。染色体复制之初可以由小 RNA 作为引物起始合成，之后细胞的修复机器启动，DNA 聚合酶能够以反链 DNA 为模板，以之前合成的 DNA 为引物，合成新的 DNA 取代染色体中间的 RNA 引物。但是线性染色体最末端的 RNA 引物因为没有另外的引物起始，没有办法被 DNA 取代。所以线性染色体 DNA 每复制一轮，RNA 引物降解后末端都将缩短一个 RNA 引物的长度（图 1，假想的末端隐缩示意图，实际上更为复杂一些）。尽管这个引物不长，但是细胞千千万万代地不断复制，如果不进行补偿，染色体不断缩短，最终就会消失。James Watson（因为发现 DNA 双螺旋结构获得诺奖）最早就明确指出了这个“末端隐缩问题”，并猜想染色体也许可以通过在复制前联体（染色体末端跟末端连起来）的方式来解决末端复制的问题[1]。

早在 1939 年，潜心玉米遗传性状研究的 Barbara McClintock 女士（因为发现玉米的转座子获得诺奖）注意到，在减数分裂后期偶然产生的染色体断裂很容易重新融合起来形成“桥”。在紧接着的有丝分裂中，这种染色体“断裂—融合—桥—断裂”的循环不断继续[2]。既然染色体的断裂末端这么容易相互融合，那么染色体的自然末端，为什么不容易相互融合呢？合理的推测是，染色体的自然末端不同于非正常的 DNA 断裂末端，它应该有一个特殊的结构来避免染色体之间的相互融合。

更早的 1938 年，Hermann Muller（因为发明用 X 射线突变基因而获得诺奖）利用 X 射线照射果蝇产生突变体，注意到染色体的末端跟其它区域的染色体不同，它非常稳定，从未观测到断裂缺失或者倒位（inversion）。他因此先见性地认为染色体的末端比较特殊，它需要被封闭（sealed）起来，并给它一个专有的名称-端粒（telomere，来自希腊词根 telos，末端，和 meros，部分）[3]。

## 2. 端粒 DNA 序列的发现以及人工染色体的发明

那么端粒为什么与众不同呢？简单地，首先是，它的 DNA 序列有没有特殊性？

提到端粒不能不提到一种特殊的模式生物四膜虫 (*Tetrahymena thermophila*)。它对于发现端粒和端粒酶的贡献就像线虫之于发现细胞凋亡一样 (2002 年的诺贝尔医学奖授予了细胞凋亡的研究)。四膜虫有两个细胞核。小核很稳定, 含 5 对染色体, 用于生殖传代。而大核在接合细胞的发育过程中, 染色体断裂成 200-300 个小染色体, rDNA 从染色体上断裂后通过复制更是形成高达~10000 个小染色 (<http://www.ciliate.org>)。

四膜虫的小染色体众多, 也就说端粒可能非常丰富。这就为端粒研究提供了得天独厚的材料。1978 年, Liz 女士利用这种特殊的模式生物纯化了 rDNA, 以 rDNA 为模板通过体外合成掺入 dNTP 的实验, 推断四膜虫的端粒是由许多重复的 5'-CCCCAA-3' 六个碱基序列组成的[4]。第一个谜底揭开了, 哦, 重复序列, 端粒 DNA 果然特殊。序列本身隐隐暗示着解决染色体末端的隐缩问题和保护问题的机制。

1980 年, 当 Liz 女士在会议上报告她的这一发现的时候, 引起了 Jack Szostak 的极大兴趣。他那时候试图在酿酒酵母 (*Saccharomyces cerevisiae*) 中建构人工线性染色体, 让它能够在细胞中像自然染色体一样复制。但是当环状质粒线性化转入酵母细胞后, 它很快地被降解掉。它的降解是不是因为它的末端没有端粒保护呢? 端粒序列的发现让 Jack Szostak 有机会把线性质粒末端连接上四膜虫的端粒 DNA, 然后再导入酵母细胞。奇迹发生了, 线性质粒不再降解, 它可以在细胞内复制, 人工染色体的想法实现了! ([5] and <http://nobelprize.org>)

值得一提的是, 人工染色体的实现当初也许仅仅是满足人们的异想天开, 但它实际上使 DNA 的大片段克隆成为可能, 后来为人类基因组测序的工作立下了汗马功劳。这也是 Jack Szostak 共同获得诺贝尔奖的重要原因。

1984 年, Liz 实验室通过将酵母端粒克隆到线性人工染色体的方法, 发现酵母的端粒序列是由不太规则的 TG<sub>1-3</sub>/C<sub>1-3</sub>A 重复序列组成的[5, 6]。

### 3. 端粒复制的两个假说以及端粒酶活性的发现

在 1984 年报道酵母端粒序列的同一篇文章中, Liz 实验室发现了一个有趣的现象: 带着四膜虫端粒 DNA 的人工染色体导入到酵母后, 被加上了酵母的端粒而不是四膜虫的端粒序列[6]。由于端粒是由重复序列组成的, 当时人们普遍猜想同源重组是延伸端粒补偿染色体末端隐缩的机制。但是同源重组只能复制出更多本身的序列, 为什么在四膜虫端粒上加的是酵母的端粒序列而不是四膜虫端粒本身的序列呢? 这个现象同源重组是无解的。也许, 可能, 酵母中存在专门的“酶”来复制端粒 DNA。

究竟是重组还是全新的酶? 为了厘清这两个假说, Liz 意识到最重要的是找到这个“酶”。

如前所述，在四膜虫接合细胞的大核发育过程中，大核产生了非常丰富的小染色体，每一个小染色体都被从头加上了端粒。可以推测，如果“酶”的假说成立，此时细胞内的“酶”活性应该是非常高的。

1984年，Carol女士作为博士生加盟了Liz实验室。她们俩精心讨论设计实验，用四膜虫的核抽提液与体外的端粒DNA进行温育，试图在体外检测到这个“酶”活性，看到端粒的延伸。经过不断优化条件，尤其是把底物换成体外合成的高浓度的端粒DNA后，同年的圣诞节，勤奋的Carol同学打开暗盒曝光x光片，终于清楚地看到了“酶”活性。在测序胶的同位素曝光片上，端粒底物明显被从新加上了DNA碱基，而且每六个碱基形成一条很深的带，与四膜虫端粒重复基本单位为六个碱基正好吻合[7]。这种酶活性不依赖于DNA模板，只对四膜虫和酵母的端粒DNA进行延伸，而对随机序列的DNA底物不延伸；并且该活性不依赖于DNA聚合酶[7]。由于同源重组对序列没有特异性的要求并且依赖于DNA聚合酶的活性，至此，她们澄清了这两种假说，证明了有一种“酶”来延伸端粒DNA。这种酶后来被命名为“端粒酶”（telomerase）。

#### 4. 端粒酶 RNA 亚基的发现

紧接着她们开始对端粒酶活性进一步定性。此时Tom Cech（因为发现RNA可以有催化酶活性获得诺奖）正好访问Liz的实验室，她/他们一起做了个简单的实验，就是用RNA酶处理样品，降解样品的RNA，看看端粒酶活性是否受到影响。结果是酶活性竟然消失了，端粒酶活性依赖于RNA[8, 9]。端粒酶会不会是另外一种特殊的RNA催化酶？想必从这个时候，Tom Cech开始被端粒酶深深吸引，并介入了这个领域。当然那时候也知道端粒酶是依赖于蛋白的：用蛋白酶消化后的样品也不具备端粒酶活性[7]。1989年，Carol通过跟踪端粒酶活性，用柱子纯化并克隆了四膜虫的端粒酶RNA亚基。另一个谜底揭开了：RNA亚基有一段RNA序列正好和四膜虫的端粒DNA序列互补，端粒酶正是利用RNA亚基的这段序列作为模板重复复制出端粒DNA[10]。

端粒和端粒酶领域的领军人物多数是女士。公平起见，不妨多介绍几位。Virginia Zakian女士实验室的Daniel Gottschling发现端粒区域具有TPE效应（Telomere Position Effect，端粒位置效应），也就是置入端粒区域的基因会被silenced（沉默），不表达[11]。我们来看看Daniel Gottschling自己成立实验室后是如何利用TPE现象来设计实验，筛选出酵母的RNA亚基基因的，同时也见识一下酵母这个非常强大的遗传学模式生物。

首先把URA和ADE两个基因通过遗传重组的方法置入端粒区域。URA基因使酵母能够自己合成uracil（尿嘧啶）；缺少ADE基因则会让酵母累积红色素，使酵母克隆变成红色。

由于端粒的 TPE 效应, *URA* 和 *ADE* 不表达, 酵母只能在含有 uracil 的培养基中生长, 且克隆是红色的。在这个遗传改造过的酵母中转入酵母的 cDNA 表达库, 可以预计, 某些调控端粒长度的基因通过过表达可以改变端粒的长度。如果端粒长度变得足够短的话, TPE 效应就会消失, *URA* 和 *ADE* 两个基因就会启动表达。那么这种短端粒的酵母就能够在不含有 uracil 的培养基中生长, 并且酵母克隆显示出白色 (图 2)。

用这两个指标进行筛选, 可以筛选到一系列让端粒变短的基因。其中一个基因比较特殊, 任何阅读框都只能阅读一小段蛋白序列, 看起来更像个 RNA 基因。深入研究发现这个基因含有酵母端粒序列的互补序列, 而且对这一序列进行突变, 酵母的端粒序列也发生相应的改变-这个基因正是编码酵母端粒酶 RNA 亚基的基因[12]。这个发现有运气的成分, 因为一般情况下, 端粒正调控基因过表达, 端粒应该变长才对。但是恰恰相反, RNA 亚基的过表达, 端粒反而变得很短。

此后的 1995 年, 同样是 Liz 实验室报道了酵母端粒酶的活性[13]。

## 5. 端粒酶催化亚基的发现

到 RNA 亚基被揭示为止, 谜团就只剩下端粒酶的蛋白质亚基了。端粒酶既然能够利用 RNA 模板亚基来复制 DNA, 那么很容易推测这个蛋白亚基可能具有 RNA dependent DNA polymerase 活性 (依赖于 RNA 的 DNA 聚合酶活性), 也就是逆转录酶的活性。更进一步地说, 它的蛋白序列里应该包含逆转录酶特有的结构域。尽管没有实验证据, 这个谜底通过逻辑推理实际上已经猜到了一半, 很多人都想彻底地揭开它, 不同实验室的竞争也变得激烈起来。

1989 年, 端粒和端粒酶领域的另外一位女杰, Jack Szostak 实验室的 Vicki Lundblad 利用设计精巧的遗传学筛选方法, 从酵母中筛选到了 *EST1* 基因 (这个遗传学方法描述起来比 Daniel Gottschling 的更为复杂, 这里就不介绍了, 有兴趣的去读原始文献)。敲除 *EST1* 基因, 端粒会随着酵母的传代逐渐缩短, 最后缩短到一定程度时, 酵母就衰老死亡[14]。

Est1 蛋白从酵母的表型看来很像是端粒酶的蛋白催化亚基。Vicki Lundblad 和 Liz 在 90 年的 Cell 杂志上大胆猜测 Est1 蛋白含有逆转录酶的结构域[15]。Virginia Zakian 实验室在 95 年的 cell 上也报道 Est1 蛋白是酵母端粒酶的体外活性所必需的[16]。不过发表在顶尖杂志的工作并不一定是好工作, 科学也有谬误的时候, 尤其是在竞争激烈, 人们已经感觉得接近于揭开谜底的时候容易急躁, 匆忙发表了一些不够 solid (可靠) 的数据。这一次运气不站在她们一边, 谜底猜错了。

那么端粒酶的催化亚基是什么呢？1996年，Tom Cech 实验室用生化的手段纯化了四膜虫端粒酶复合体，其中有一个蛋白根据分子量命名为 p123 [17]。同一时期，Vicki Lundblad 实验室改进了她的遗传学筛选方法，筛选到了几个与酵母端粒复制密切相关的基因，命名为 *EST2*，*EST3* 和 *EST4*（也叫 *CDC13*）[18]。这个改进版的筛选方法非常厉害，它把酵母端粒酶全酶的亚基一网打尽。值得一提的是尽管这个结果只发表在“影响因子”不太高的专业杂志 *Genetics*（遗传学）上，它的影响却非常深远，此后很长一段时间乃至到现在，酵母端粒和端粒酶领域很大一部分精力都集中在阐述这几个基因的功能上。

用生化方法纯化出来的四膜虫 p123 蛋白，以及用遗传学方法筛选出来的酵母 Est2 蛋白后来都被 Tom Cech 实验室证明是端粒酶的催化亚基：它们含有逆转录酶的结构域，如果对该结构域的关键氨基酸进行突变，则端粒酶活性消失[19]。同年的 1997 年稍晚于 Tom Cech，第一个发现癌基因 *RAS* 的 Robert Weinberg 也参与了这个工作，他们也报道了酵母和人的端粒酶催化亚基[20, 21]。此后，人们用体外转录和翻译系统共表达了端粒酶的催化亚基和 RNA 亚基，在体外重建了端粒酶活性，证明这两个核心亚基的存在是端粒酶活性的必需且完备条件（sufficiency）[22]。

至此，有关染色体末端隐缩问题和保护问题的谜底终于全部揭开了。端粒和端粒酶的一系列发现完美地解释了这两个问题：染色体末端的 DNA 由简单重复的端粒序列构成，端粒（本文为了简洁称端粒 DNA 为端粒，其实端粒的准确定义是端粒 DNA 和结合蛋白形成的复合体）保护着染色体末端，使之区别于一般的断裂染色体末端，而不被各种酶降解，相互之间不会融合。端粒酶负责端粒的复制，端粒酶的催化亚基利用端粒酶自己的 RNA 亚基作为模板通过转位不断重复复制出端粒 DNA，从而补偿在染色体复制过程中的末端隐缩，保证染色体的完全复制（图 3）。

这的确是非常完美的发现旅程。然而生物体在展示内在的简单性的同时，还会显示出它内在的多样性和复杂性。需要指出的是本文描述的只是端粒和端粒酶领域发现的主线。实际上，细胞穿过亿万年的时光，在漫长的进化中尝试了各种可能性。端粒酶复制只是其中一种最为普遍的解决染色体末端隐缩问题的方式。回到最初的猜想，当时人们猜测同源重组延伸端粒的假说也并没有错，细胞实际上也可以通过同源重组的方式延伸端粒[23]。裂殖酵母可以通过染色体头尾相联，环化的方式来避开染色体末端隐缩的问题，在缺乏端粒酶和端粒的情况下生存传代[24]。这与 James Watson 提出的染色体联体的猜想颇有异曲同工之妙。果蝇能通过转座子的不断复制延伸端粒[25]。而病毒更是无所不用其极，它们能够利用蛋白[26]

或者 tRNA[27]作为引物来起始基因组 DNA 的合成,从而使自己的染色体解决复制的隐缩问题。

不得不赞叹,进化,或者生物体,真是天才啊!

#### 附:本文提到的端粒和端粒酶发现大事记

- 1938 年, Hermann Muller 注意到染色体末端的特殊性并提出端粒的概念
- 1939 年, Barbara McClintock 发现玉米细胞的染色体断裂末端容易融合
- 1972 年, James Watson 提出染色体复制的末端隐缩问题
- 1978 年, 报道四膜虫的端粒序列
- 1982 年, 端粒的发现导致人工染色体的发明
- 1984 年, 报道酵母的端粒序列
- 1985 年, 报道四膜虫的端粒酶活性
- 1989 年, 报道四膜虫端粒酶的 RNA 亚基
- 1994 年, 报道酵母端粒酶的 RNA 亚基
- 1995 年, 报道酵母端粒酶活性
- 1996 年, 纯化了四膜虫端粒酶的催化亚基  
遗传筛选到酵母端粒酶的催化亚基
- 1997 年, 证明了四膜虫和酵母端粒酶的催化亚基

#### 参考文献

1. Watson, J.D. (1972). Origin of concatemeric T7 DNA. *Nature: New biology* 239, 197-201.
2. McClintock, B. (1939). The Behavior in Successive Nuclear Divisions of a Chromosome Broken at Meiosis. *Proceedings of the National Academy of Sciences of the United States of America* 25, 405-416.
3. Muller, H.J. (1938). The remaking of chromosomes. *Collecting Net* 13, 182-198.
4. Blackburn, E.H., and Gall, J.G. (1978). A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in *Tetrahymena*. *J Mol Biol* 120, 33-53.
5. Szostak, J.W., and Blackburn, E.H. (1982). Cloning yeast telomeres on linear plasmid vectors. *Cell* 29, 245-255.
6. Champay, J., Szostak, J.W., and Blackburn, E.H. (1984). DNA sequences of telomeres maintained in yeast. *Nature* 310, 154-157.
7. Greider, C.W., and Blackburn, E.H. (1985). Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. *Cell* 43, 405-413.
8. Greider, C.W., and Blackburn, E.H. (1987). The telomere terminal transferase of *Tetrahymena* is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell* 51, 887-898.
9. Greider, C.W., and Blackburn, E.H. (2004). Tracking telomerase. *Cell* 116, S83-86, 81 p following S86.
10. Greider, C.W., and Blackburn, E.H. (1989). A telomeric sequence in the RNA of *Tetrahymena* telomerase required for telomere repeat synthesis. *Nature* 337, 331-337.
11. Gottschling, D.E., Aparicio, O.M., Billington, B.L., and Zakian, V.A. (1990). Position effect at *S. cerevisiae* telomeres: reversible repression of Pol II transcription. *Cell* 63, 751-762.
12. Singer, M.S., and Gottschling, D.E. (1994). TLC1: template RNA component of *Saccharomyces cerevisiae* telomerase. *Science* 266, 404-409.

13. Cohn, M., and Blackburn, E.H. (1995). Telomerase in yeast. *Science* 269, 396-400.
14. Lundblad, V., and Szostak, J.W. (1989). A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell* 57, 633-643.
15. Lundblad, V., and Blackburn, E.H. (1990). RNA-dependent polymerase motifs in EST1: tentative identification of a protein component of an essential yeast telomerase. *Cell* 60, 529-530.
16. Lin, J.J., and Zakian, V.A. (1995). An in vitro assay for *Saccharomyces* telomerase requires EST1. *Cell* 81, 1127-1135.
17. Lingner, J., and Cech, T.R. (1996). Purification of telomerase from *Euplotes aediculatus*: requirement of a primer 3' overhang. *Proc Natl Acad Sci U S A* 93, 10712-10717.
18. Lendvay, T.S., Morris, D.K., Sah, J., Balasubramanian, B., and Lundblad, V. (1996). Senescence mutants of *Saccharomyces cerevisiae* with a defect in telomere replication identify three additional EST genes. *Genetics* 144, 1399-1412.
19. Lingner, J., Hughes, T.R., Shevchenko, A., Mann, M., Lundblad, V., and Cech, T.R. (1997). Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* 276, 561-567.
20. Counter, C.M., Meyerson, M., Eaton, E.N., and Weinberg, R.A. (1997). The catalytic subunit of yeast telomerase. *Proceedings of the National Academy of Sciences of the United States of America* 94, 9202-9207.
21. Meyerson, M., Counter, C.M., Eaton, E.N., Ellisen, L.W., Steiner, P., Caddle, S.D., Ziaugra, L., Beijersbergen, R.L., Davidoff, M.J., Liu, Q., et al. (1997). hEST2, the putative human telomerase catalytic subunit gene, is up-regulated in tumor cells and during immortalization. *Cell* 90, 785-795.
22. Weinrich, S.L., Pruzan, R., Ma, L., Ouellette, M., Tesmer, V.M., Holt, S.E., Bodnar, A.G., Lichtsteiner, S., Kim, N.W., Trager, J.B., et al. (1997). Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTERT. *Nature genetics* 17, 498-502.
23. Kass-Eisler, A., and Greider, C.W. (2000). Recombination in telomere-length maintenance. *Trends in biochemical sciences* 25, 200-204.
24. Sadaie, M., Naito, T., and Ishikawa, F. (2003). Stable inheritance of telomere chromatin structure and function in the absence of telomeric repeats. *Genes & development* 17, 2271-2282.
25. Biessmann, H., and Mason, J.M. (2003). Telomerase-independent mechanisms of telomere elongation. *Cell Mol Life Sci* 60, 2325-2333.
26. King, A.J., Teertstra, W.R., and van der Vliet, P.C. (1997). Dissociation of the protein primer and DNA polymerase after initiation of adenovirus DNA replication. *The Journal of biological chemistry* 272, 24617-24623.
27. Litvak, S., Sarih-Cottin, L., Fournier, M., Andreola, M., and Tarrago-Litvak, L. (1994). Priming of HIV replication by tRNA(Lys3): role of reverse transcriptase. *Trends in biochemical sciences* 19, 114-118.

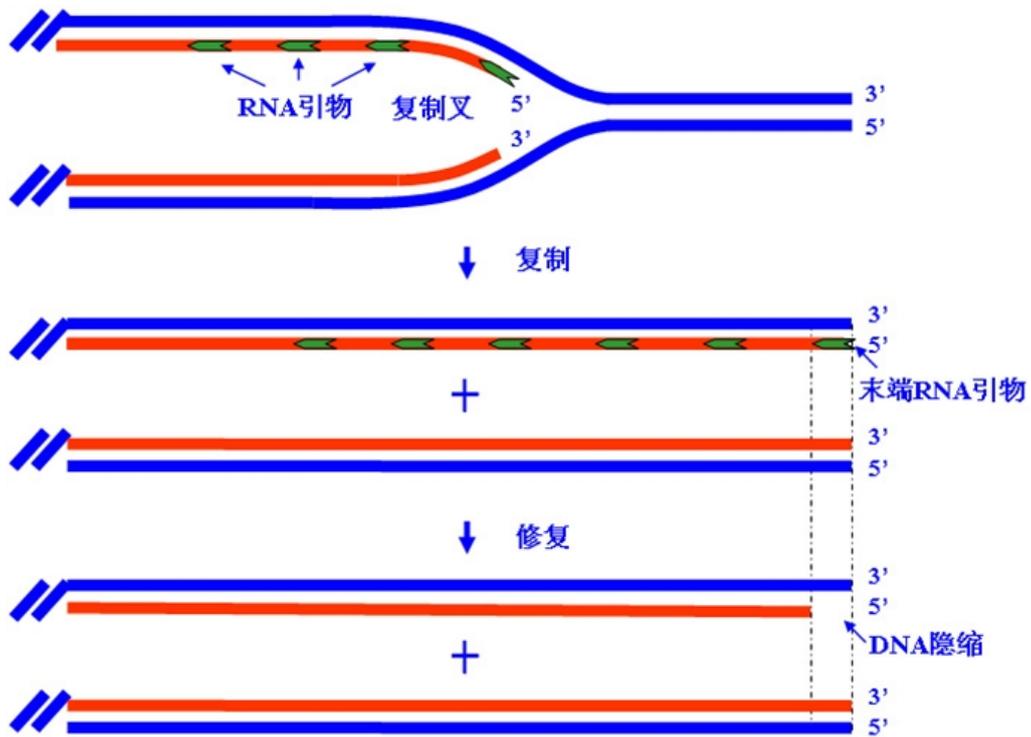


图1 染色体复制的末端隐缩问题

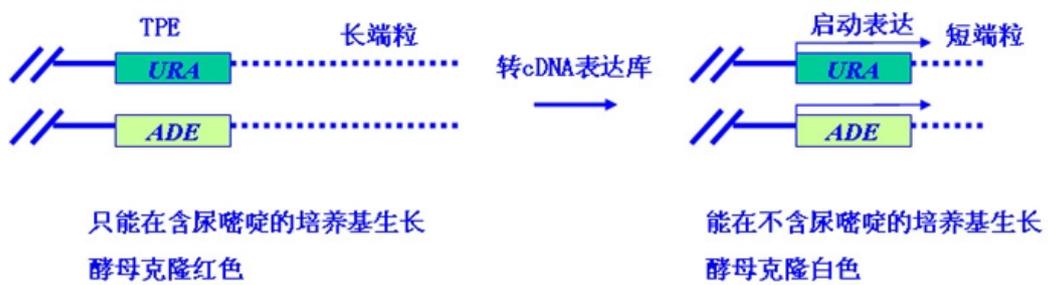


图2 酵母端粒调控基因的遗传筛选

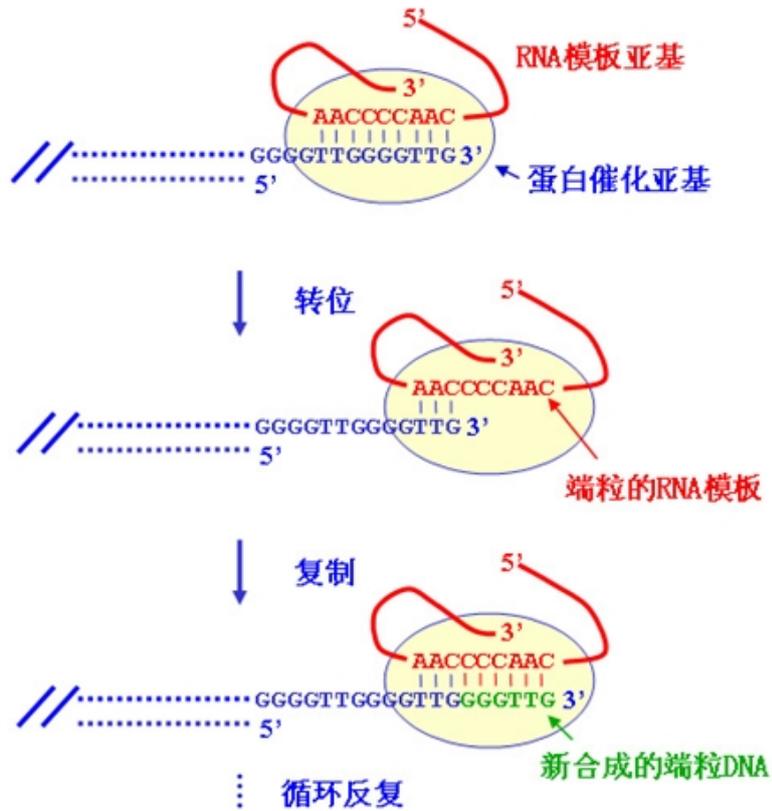


图3 四膜虫端粒酶对端粒DNA的复制

(作者注明：2009年诺贝尔医学或生理学奖授予端粒和端粒酶领域的研究。本文是基于个人理解来整理的端粒和端粒酶的发现历史，因为知识时间有限，其中必有偏差和谬误的地方，关键之处还是以原始文献为主。本人之所以赶这趟诺贝尔奖热，花大量的时间进行文献阅读和整理，是因为它提供了一次极好的向公众传播科学思想的机会。由于端粒和端粒酶领域的一系列发现贯穿着“发现现象/问题”-“提出概念/模型”-“实验验证”的思路，重现这个思路对科学工作者是有启发意义的。本文也提供了一个很好的教学案例。物当尽其用，欢迎转载传播。)