

酵母转座子—Ty 因子及其对基因表达的影响

徐 岗

(中国科学院上海细胞生物学研究所)

基因转座现象首先是在玉米中发现的^[1],以后在许多生物中都发现有这一现象存在。1979年 Cameron 在酵母菌中发现一族中度重复顺序具有转座能力,它们被称为 Ty 因子(Ty element, Ty; Transposon Yeast)^[2]。这一族序列具有许多共同的特点:每段序列长度均在 6 kb 左右,含有一个中心区域称为“ε 序列”,两端分别接有长约 0.33 kb 的同向重复顺序“δ 序列”。Ty 因子对酵母菌基因表达有很大的影响。由于具有转座能力, Ty 可以以一定频率在酵母基因组中移动,当 Ty 插到某个基因上游一定区域内时,即可使该基因的表达水平发生很大变化,多数情况下,可使表达水平提高数十倍。在某些特定条件下, Ty 甚至可促使大段的普通 DNA 序列发生转座。对 Ty 因子转座机制的研究,揭示出这一转座子具有反转录功能,它所含有的遗传信息可从 DNA→RNA→DNA 以反转录方式流动。因此,有人使用“反转录转座子”(Retrotransposon)这一名词来描述这一类具有反转录功能的可移动遗传因子。

一、Ty 因子的 DNA 序列特征

酵母菌中的 Ty 因子可按其同源性大致分为两类:即 Ty 1 和 Ty 2(Ty 2 又称作 Ty 917),每一类中都有多种不同的 Ty 因子。第一类中的各种 Ty 尽管其 DNA 序列的某些位点存在差异,但都与 Cameron 1979 年发现的 Ty 1 序列有很高的同源性,它们在单倍体酵母细胞中约有 30 个拷贝,占细胞总 DNA 量的 0.04%,尽管 Ty 可在基因组中发生位置、数量的变化,其总量似乎是相对恒定的^[3]。第二类中各个序列与 Ty 1 的差异相对较大,但都与 Roeder 在 1980 年发现的 Ty 917 相似,这类 Ty 在细胞中

的拷贝数较少,每个细胞中约有 6 个拷贝^[4,5]。目前所得到的有关 Ty 的资料,大都集中在 Ty1 上。

Ty 的长度一般都在 6 kb 左右,中间是 ε 序列,长约 5.3 kb;在 ε 序列两端分别接有同向重复顺序 δ 序列,其长度均为 0.33 kb(图 1)。由于少数碱基的不同,在 δ 序列和 ε 序列中,各种 Ty 因子的限制性内切酶位点不尽相同,酶切图谱呈现一定的差异^[6]。Ty 1 和 Ty 2 形成异源杂交双链后,可有如图 2 所示的结构^[4]。这表明, Ty 1 和 Ty 2 之间除两个区域(图 2 中的非配对区)不呈现同源性外,相互间仍具有相当高的同源性。

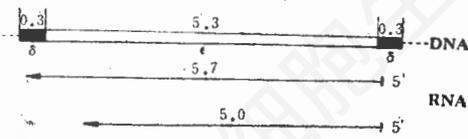


图 1 Ty 因子的一般结构及其转录产物^[6]。



图 2 Ty 1 和 Ty1-17(属于 Ty 2)形成的异源杂交双链示意图^[7]。

图中的数字表示碱基对数。平行线部分为配对区,双弧线部分为非配对区。

对 Ty 因子的序列分析表明, Ty 在 DNA 序列上有许多特点和反转录病毒的前病毒形式十分相似。

本文所用缩写符号:

bp(base pair): 碱基对。kb(kilobase): 千碱基。PRD(Positive Regulatory Determinant): 正调控决定子。

本文承蒙匡达人教授、朱心良副教授指导, 特此致谢。

	序列起点	互补区域
Ty 因子		
Ty 1-17	333	5' TGGTAGCGCCT A T G C T T CGGTTAC 3'
Ty 109	332	-----G-----
Ty 912	335	-----G-----
Ty 1-15	340	-----G-----
tRNA _{met}	3	3' ACCAUCGCGG C G A ... GCCAAAG 5'

图 3 Ty 的 5' δ/ϵ 交界处与 tRNA_{met} 3' 端的互补情况^[8]。

方框内为互补区域。相同的碱基未写出，以“-”表示。几个 Ty 只有一个碱基不同，并且处于非互补区。

在 Ty 5' 侧*(注)的 δ/ϵ 交界处，有一段序列可与酵母菌的 tRNA_{met} 互补配对。这一序列在所有已发现的 Ty 序列中具有相当高的保守性。Ty 1 中的三个 Ty 因子在这一段完全一样，而属于 Ty 2 的 Ty 1-17 这一段中只有一个碱基与 Ty 1 中的不同，且这个碱基正处在可能和酵母 tRNA_{met} 没有作用的区域中(图 3)^[8]。在 Ty 因子的 3' 侧紧靠 3' δ 序列处，已做过序列分析的转座子如 Ty 1-17、Ty 912 和 Ty 109 等，都有一段富含嘌呤的区域。这一段 DNA 顺序通常是：5'TGGGTGGT 3'^[8]。这些序列在反转录病毒中也同样存在，与 tRNA 互补的区域与反转录的起始有关，而 3' 侧的富含嘌呤区则与 cDNA 第二条链的合成有关。同向重复顺序 δ 序列，可从结构特点和功能上将其分为三个部分：U₅、R 和 U₃，它们可分别与反转录病毒长末端重复顺序(LTR)中的同名部分在结构和功能上相对应^[9]，如图 4 所示。

由此可见，Ty 因子在其 DNA 结构上，既具有一般转座子的特点，如两端具有长的重复顺序；同时它们还与反转录病毒经反转录后整合到染色体上的前病毒 DNA 非常相似。这些特点，为 Ty 的反转录功能奠定了结构基础。

在 Ty 序列中，还存在着启动子结构、终止子结构和一些与增强子核心区相似的序列，这些将在后面分别说明。

Ty 因子两端的 δ 序列通常是相同的，但

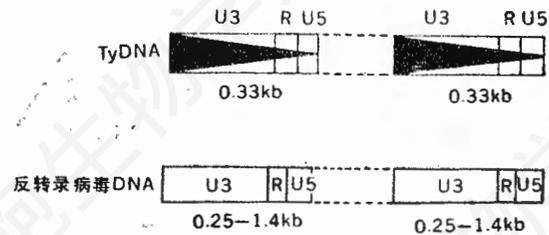


图 4 Ty 的 δ 序列和反转录病毒 LTR 的比较。

上：Ty 的 δ 序列，框中箭头表示 δ 的方向^[2]。下：反转录病毒的 LTR^[10]。

有时也会有个别碱基的差异^[6,11]。 δ 序列既可以在完整的 Ty 序列中出现，也可以单独存在于基因组中，这时它们成为所谓“单 δ 序列”(solo δ -sequence)。单倍体细胞基因组中，约有 100 个单 δ 序列。当以单 δ 序列形式存在时， δ 序列将有较多的变异，其 DNA 顺序的保守性大大低于和 ϵ 序列结合的 δ 序列。

二、Ty 因子的转录及其调控

在酵母菌细胞总 RNA 中，有两种带 poly(A) 的 RNA 是由 Ty DNA 转录而来的。这两种 RNA 的长度分别为 5.7 kb 和 5.0 kb，都称为 Ty RNA，占胞内带 poly(A) RNA 总量的 5—10%。

对 DNA 序列分析后发现，Ty 因子的末端重复顺序 δ 中，含有启动子样和终止子样结构，

*注：Ty 的方向由其转录产物(见后)确定，Ty 的 5' 侧、3' 侧分别对应于 TyRNA 的 5' 和 3' 侧。

且终止子样结构在启动子下游约数十个碱基处。Ty的转录可由5'侧 δ 序列内的启动子控制^[6,9]。由于 δ 在Ty两端同向重复,Ty因子的两端因而分别具有方向相同的启动子和终止子结构。当Ty因子进行转录时,Ty RNA的合成必定首先要通过同侧的终止子结构。但Ty RNA为什么还能继续合成呢?与Ty中相似,反转录病毒的长末端重复顺序LTR(Long Terminal Repeat)中也有这种终止子结构紧跟在启动子下游的现象。Benz等人研究了原病毒(以DNA形式存在的反转录病毒)RNA合成的过程。他们发现,由于DNA的转录产物RNA二级结构的原因,5'侧 δ 内的终止子结构不能终止RNA的合成,只有在3'侧 δ 中,终止子才能有效地发挥终止的功能^[12]。Ty RNA合成可能也是以类似的机制通过第一个终止子的^[6]。

长度为5.7 kb的Ty RNA,其转录从5'侧 δ 内的R区开始,转录起点距 ϵ 序列93—97 bp(不同的Ty略有不同),穿过紧跟其后的终止子样结构,通过整个 ϵ 序列以及3'侧 δ 中的启动子样结构,在3'侧 δ 的R区末端终止。由此可见,5.7 kb Ty RNA几乎包含了Ty因子的全部结构信息(Ty总长约5.9—6 kb),且两端仍是同向重复顺序(即原来 δ 序列中的R区),这与反转录病毒基因组RNA非常相似,参见图5。5.0 kb的Ty RNA其转录起点与5.7 kb Ty RNA相同,但它在中心区 ϵ 序列内终止。至于它如何在 ϵ 序列中实现终止,现在还不清楚。5.7 kb和5.0 kb的Ty RNA从Ty DNA的同一条链上转录得到^[13]。

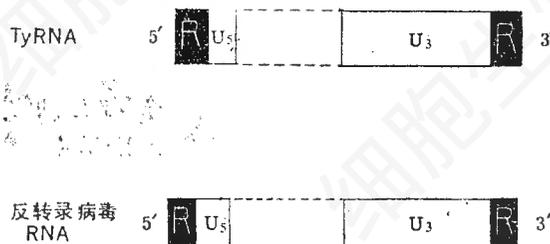


图5 Ty 5.7 kb RNA^[9]和反转录病毒基因组RNA^[10]的结构。

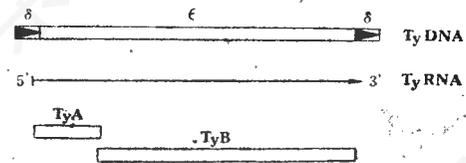


图6 Ty因子的两个阅读框:TyA和TyB。

在Ty RNA上,有两个开放阅读框,分别称为TyA和TyB(图6)。Ty A长约1.3 kb, Ty B长约4.0 kb,两者间有38 bp的重叠(对Ty 2,这一重叠区域为44 bp)^[13,14,15,16]。这两个阅读框在翻译前并不需要剪切,只需由核糖体在两者重叠处进行移框变化,便可将两个阅读框分别翻译出来^[16]。这也是Ty在翻译过程中所具有的一大特点。

从TyA和TyB的碱基顺序推导出氨基酸顺序后,可以发现,从一级结构和二级结构的特征来看,TyA与DNA结合蛋白(DNA-binding protein)具有很高的相似性,Ty A的编码产物可能与反转录病毒的gag蛋白相似^[13](gag蛋白与核酸一起构成反转录病毒的核蛋白核心,参见文献[17]和[18])。而TyB则显示出与反转录病毒RSV(Rous Sarcoma Virus)的反转录酶具有很高的同源性^[18,15],并可能还编码整合酶(int)。现在已经知道,Ty因子转座过程中所需的反转录酶就是Ty自己编码的。

Ty RNA的转录,在酵母菌中还要受到许多因素的控制。酵母细胞的接合型状态(Mating type)对Ty RNA的转录有很大影响。在具有接合能力的细胞(如接合型为 α 型或a型的单倍体细胞)中,或a/a型、 α/α 型细胞中,不论在细胞生长的对数期还是稳定期,Ty RNA的水平都较高,占细胞中带Poly(A)RNA总量的5—10%;在非接合型细胞(如a/a型细胞)中,对数期内,Ty RNA的量只相当于接合型细胞中的1/5,在稳定期,则只有1/20^[6]。

一些酵母基因也可对Ty转录发生作用。如SPT 3基因产物是Ty转录的正调控因子。在SPT(Suppressor of Ty)位点突变的酵母细胞,由于没有这个正调控因子,其Ty的表达

将受到抑制。在突变子 *spt 3* 中, 检测不到正常的 5.7 kbTy RNA, 只有少量 5' 端比正常 Ty RNA 少 800 bp 的 Ty 转录产物^[19], 此时 Ty 的转录在原起点下游 800 bp 的地方起始。为什么会产生这个新的起点以及它如何起始, 现在还不清楚。*spt 3* 突变的结果显示出, SPT 3 基因产物对 δ 内的 Ty 启动子具有促进其转录的作用。另一些实验表明, SPT 3 基因产物虽对 Ty 转录有促进作用, 但它们的过量存在却并不能提高 Ty 转录水平^[20]。

还有一些突变子也具有降低 Ty 表达的能力, 如 *roc 1*、*roc 2* 突变^[21] 和不育突变 *ste 7* 等, 都可使 Ty RNA 水平大大降低^[9]。但现在还不知道, 这些突变基因在突变前, 是否也象 SPT 3 那样对 Ty 转录有正调控作用。

三、Ty 因子在基因组中的同源重组

Ty 因子在酵母基因组中, 可以以同源机制(即同源重组)或非同源机制(即转座)发生位置、数量的变动。发生同源重组的频率(10^{-4} — 10^{-5})大大高于转座的频率(10^{-7} — 10^{-8})^[9]。同源重组主要发生在位于基因组中不同位点的两个 Ty 之间, Ty 与散布在基因组中的单 δ 序列也可发生同源重组。这些重组常可导致染色体产生较大的变化, 如基因转位(translocation)、缺失(deletion)或基因倍增(duplication)等。在这一方面, Ty 因子与其它 DNA 片段具有共同的性质, 此时的 Ty 因子只是作为一段同源区域促进染色体 DNA 的重排。

由于同源重组, 酵母基因组中的 Ty 因子可以发生很多的变化。Roeder 和 Fink 把这些变化作了总结^[7], 它们可概括成如下七种情况:

(一) Ty 自身的切失 Ty 因子两端的 δ 序列通常是相同的, 它们可发生 δ - δ 同源重组。这种重组的结果是, Ty 因子的 ϵ 序列和一个 δ 序列被切除丢失, 在原位只剩下一个单 δ 序列(图 7.1)。这很可能是酵母基因组中单 δ 序列的主要来源。

(二) 丢失和转置 同一染色体上存在两

个同源 Ty 时, 若两个 Ty 的方向相同, 这两个 Ty 就成了整段 DNA 两端的同向重复顺序, 此时, 这整个 DNA 区域的行为就如同一个 Ty 因子(图 7.2 a), 可由同源重组丢失其中的一个 Ty 和两 Ty 间的 DNA 片段; 若两 Ty 方向相反, 则可导致两 Ty 间 DNA 片段的倒转(图 7.2 b)。

(三) 基因转位 不同染色体上的两个 Ty 之间也可发生同源重组。这种重组可有两种情况: (1) 产生一条双着丝点染色体和一条无着丝点染色体, (2) 由对等交换而形成基因转位。第一种情况通常是致死的, 因此可检出的只有基因转位(图 7.3)。

(四) 不等交叉 位于二倍体细胞两条同源染色体上的两个 Ty 之间, 可产生不等交叉, 从而造成一条染色体上缺失、另一条染色体上重复的情况(图 7.4)。

(五) 普通 DNA 片段的转座 Ty 因子有时还可促使普通 DNA 片段发生转座。当一段普通 DNA 的两端分别接有呈同向排列的 Ty 序列时, 这一整段 DNA 很容易从该位点切去, 其机制类似于 Ty 自身切失。有时被切去的这一段 DNA 又会出现原来没有这段 DNA 的染色体上, 这意味着发生了转座。这种转座可能与下面将提到的再整合有某种关联(图 7.5)。

(六) 切失和再整合 由同源重组切下的含 Ty 的 DNA 片段, 通常都不含有复制和分配的机构, 因而绝大多数被丢失。但也有少数可以再次整合而回到基因组中。这种整合是通过与染色体上其它同源片段发生交叉而实现的。染色体上一切与 Ty 有同源性的区域, 包括单 δ 都可促进这种再序列, 整合(图 7.6 a, 7.6 b)。

(七) Ty 序列的置换 一个单独的 Ty 因子可以和另一条染色体上含有两个 Ty 序列的 DNA 片段发生同源重组(如图 7.7 所示)。第一条染色体上的那个单独 Ty 因子其左端 δ 序列与第二条染色 DNA 片段最左端的 δ 序列配对, 右端也同样配对。这样就产生一个环, 环中含有第二条染色体的 Ty 序列和两 Ty 间的

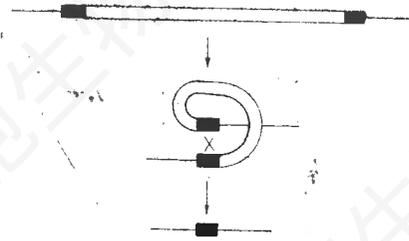


图 7.1 Ty 两端 δ -序列同源重组, 造成 Ty 的切失, 在原位只留一个单 δ -序列。

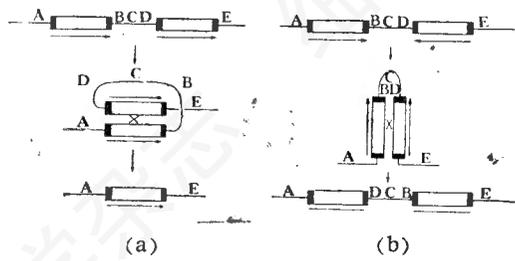


图 7.2 切失和倒置

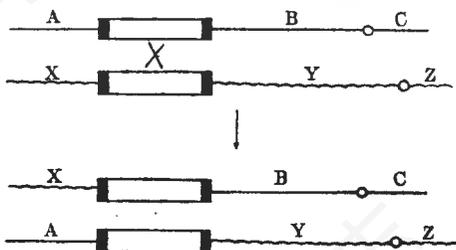


图 7.3 由 Ty-Ty 同源重组对等交叉造成转位

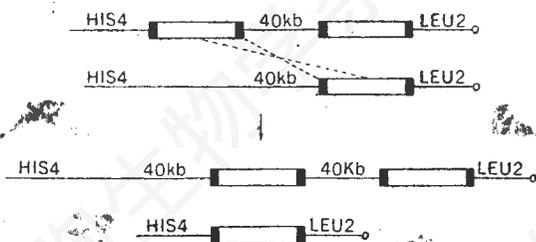


图 7.4 不等交叉造成重复和缺失

普通 DNA 序列。最终, 第二条染色体上的该段 DNA 被第一条染色体的 Ty 序列所置换。

四、Ty 因子的转座

Ty 因子在酵母基因组中可以从一个位点转移到另一个新的位点。接受 Ty 插入的染色体位点称为“靶位点”。与同源重组不同, 转座

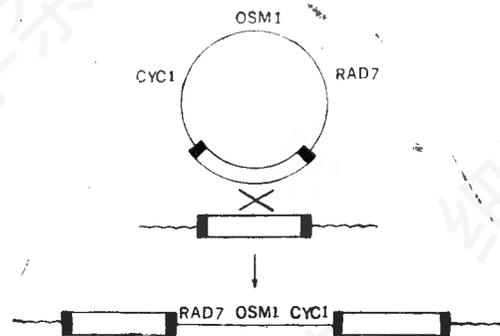


图 7.5 Ty 造成大片段 DNA 转座

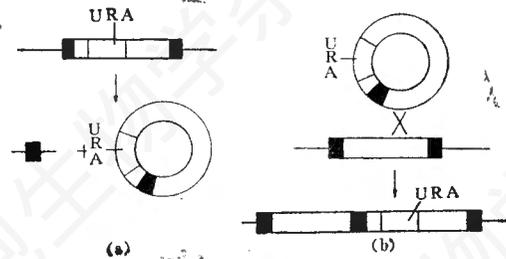


图 7.6 切失(a)和再整合(b)。URA 为标记基因。

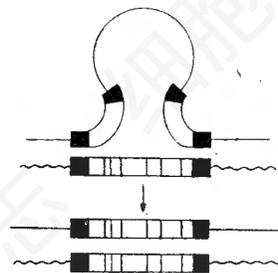


图 7.7 Ty 序列的置换, Ty 序列中的线条指示特征酶切位点。

图 7(1-7) 与 Ty 有关的同源重组^[7]。

□: Ty 的 ϵ 序列。■: Ty 的 δ 序列。

—(线~~~~): 染色体 DNA。

×: 交叉发生的位点。

的靶位点与 Ty 因子一般没有同源性。不同的 Ty 因子其靶位点的 DNA 顺序不尽相同, 如 Ty 1-B 10 是: 5'-GAAAC-3'; 而 Ty 1-D 15 则为: 5'-ATTTT-3'^[22]。许多实验表明, Ty 的转座并不需要同源性的介导。

和其它生物中的转座子相似, 当 Ty 转座

到新的位点时,其靶位点DNA将产生一小段倍增重复。这样,Ty插入后,其两端是呈同向重复的靶序列。对Ty因子,这一段靶序列的长度为5bp。这一长度与其它真核生物中可移动因子的靶序列长度完全一致。而一般在原核生物中,转座子的倍增靶序列长度为3—12bp。

由于转座,可使得Ty在许多新的位点出现,进而导致细胞表型的突变。如突变株his 4-912和his 4-917等就是Ty转座到几个不同位点而导致的His⁺突变^[5,23]。类似的还有ADH 2、CyC 7-H 2等突变。Ty转座时一般都插到染色体基因的上游或下游。至今尚未发现插到基因编码区的情况,只有少数插到启动子内^[24]。因此,有人假设,基因外的DNA可能有某些特性可促进Ty的插入,而编码区DNA则会抑制这种插入;同样也存在另一种可能,即Ty插入编码区可使基因损伤而有害于细胞生存,因而暂时无法测到^[6]。

在某些染色体上,转座活动进行得特别活跃。第3号染色体的左臂,就有一个Ty转座的“热点”。在这一不太大的区域内,有三个完整的Ty序列,两个完整的单 δ 序列,以及多个退化的单 δ 序列。序列分析表明,那众多的 δ 序列(包括已退化的 δ 序列),是由多次与转座有关的插入、重组、缺失造成的^[25]。这一区域内同样也可以发生很多与Ty或 δ 有关的同源重组。这一区域含有两个tRNA基因:tRNA^{Gln}和tRNA^{Leu}的基因。Ty因子和 δ 序列在染色体上一般都与tRNA很接近^[2,24],因此,第3号染色体的这一区域之所以成为转座“热点”,很可能是这些tRNA基因为转座提供了合适的环境^[25]。

关于Ty因子转座的机制,近年来已有了突破性的进展,Boeke等人的工作^[9],揭示了Ty是经由反转录而完成转座过程的。

Boeke及其同事将带有内含子(intron)的核糖体蛋白51基因(rp51)插到TyH3中作为标记基因,然后把含rp51的TyH3置于半乳

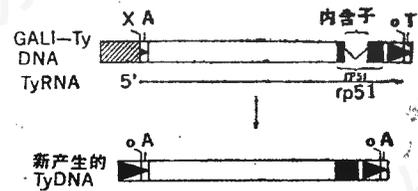


图8 Ty转座情况示意图。^[9]

在GAL1-TyDNA中,Ty5' δ 的2/3被GAL1启动子取代(图中▨区),转座后,缺失的5' δ 被修复,rp51基因中的内含子被切除。

×,XhoI内切酶切点。A、T指示 δ 中的碱基。

O,缺失的XhoI位点。

糖启动子(GAL1 promoter)的控制之下。用半乳糖诱导后,这一标记的TyH3可以发生转座。转入新位点的TyH3经结构分析,发现rp51中的内含子已被切除(图8)。同时发现,如使TymRNA过量生成,可大大提高转座频率;如果TymRNA含量降低,将同时伴有转座频率的下降。这表明,TymRNA对转座是至关重要的。

由于TymRNA的起点和终点都在Ty两端的 δ 区域内,因此TymRNA的两端与TyDNA相比将各有一段缺失:5'端缺U3区,3'端缺U5区。若Ty mRNA是Ty转座的中介物,则必定要有恢复两端缺失的过程。对Ty序列末端分析后,Boeke发现,Ty5'端的U3区是以3' δ 的U3区为模板复制的,而3'端U5区是以5' δ 的U5为复制模板的(参见图4、图8)^[9]。这一点与反转录病毒十分相似。

还有人观察到,当Ty因子强烈表达时,酵母细胞质中可看到大量病毒样颗粒,虽然这些颗粒未显示有感染其它细胞的活性,但它们显现出含有大量从TyDNA来的RNA和一些反转录酶活性^[26]。

所有这些结果,都明显地支持反转录的假说,即Ty首先需转录产生TymRNA,由这些mRNA反转录产生新的TyDNA,再经整合等途径进入染色体上新的位点,从而完成转座过程^[9]。因此,Boeke等人构建了新的名词“反转录转座子”(Retrotransposon)来描述Ty这一类具有反转录能力的转座子。

五、Ty 对酵母基因表达的影响及其作用机制

Ty 因子这一家族中的众多成员,对酵母基因表达有着多种多样的影响。它们可直接影响某个基因的表达,使该基因产物大量增加或显著减少;也可将原来不相关的某些基因的表达置于细胞接合型控制(MAT control)之下,使该基因产物的量随细胞接合型而变化;由 Ty 同源重组而导致的缺失、重复等染色体变化,又可产生新的相应的突变。因此, Ty 对研究酵母基因表达及其调控有着十分重要的意义。

(一) Ty 因子和 ROAM 效应

酵母细胞有两种接合型: α 型(MAT α)和 a 型(MAT a)。这两种细胞可以相互接合,形成 MAT 位点为杂合型的二倍体细胞——a/ α 型细胞。a/ α 型可以形成芽孢,但不再具有接合能力。前面已经说过, Ty 本身的表达是受酵母细胞接合型控制的。Ty 还具有这样的能力,即把原来与接合型无关的基因也置于接合型的控制之下。

由 Ty 插入而形成的突变子中,受 Ty 影响的那个基因,其表达水平也将会受到细胞接合型的控制,而不管该基因原先是否受接合型的影响。大多数情况下,在接合型细胞(α 型或 a 型单倍体细胞)中,该基因表达的水平大大高于非接合型细胞(如 a/ α 型或带不育基因 ste 7 的细胞)中,这种效应称为 ROAM*(注)^[27] 效应。如由 Ty 插入引起细胞色素表达变化的 CYC 7-H 2 菌株,其细胞色素 C 的表达在 α 型或 a 型细胞中,比野生型高 20 倍;而在 a/a 型细胞和某些不育细胞中,只能提高 1—4 倍^[11,27]。其它许多突变株,象精氨酸代谢突变株、尿素利用突变株等,它们都与 Ty 插入有关,因而其相应的基因产物都可随细胞接合型变化,呈现与上面类似的现象^[6]。对某些突变,如 Ty 在 ADH 2 位点(ADH: Alcohol dehydrogenase gene)引起的突变,只有在以非发酵性碳水化合物如甘油为碳源时,其 ROAM

效应才可清楚地显现^[28]。

在有些突变子中, Ty 引起 ROAM 效应会受到一些突变基因的抑制。Dubois 等得到两个不连锁的 roc 突变(roc: ROAM mutation Control)^[21]。这些 roc 突变可显著降低由 ROAM 效应引起的基因过量表达。它们是通过什么方式降低表达的,现在还不清楚,但已知这些 roc 突变并不影响其所在细胞的接合能力。还有一些不育突变如 ste 7 突变,也可降低由 Ty 插入而引起的基因过量表达^[6]。

有人曾假设有一种“正调控决定子”PRD (Positive Regulatory Determinant),它们只存在于单倍体细胞中,可促进酵母细胞接合过程中所需基因的表达^[27]。这些正决定子 PRD 可能也可促进受 Ty 影响的基因的表达:在二倍体细胞以及某些不育突变子中,由于没有这些正决定子或有不育突变的存在,那些受 Ty 影响的基因只能以较低的水平表达^[6]。用体外缺失的方法,人们研究了引起细胞色素 C 突变的 TyCYC 7-H 2,发现这个 Ty 因子 ϵ 序列中的一个限制性内切酶位点(Sal I 位点)及其附近区域,对引起细胞色素 C 过量产生是必需的;同时,这一区域和 ROAM 效应也有密切关系^[29]。因此,很可能是 Ty 因子上的某一段序列,如前面所说的 Sal I 切点及附近区域,为某个或某些与细胞接合型有关的反式作用因子(如 PRD)提供了作用部位,从而使受 Ty 影响的基因进一步受到细胞接合型的控制,产生 ROAM 效应。

(二) Ty 序列对邻近基因的影响及其增强子作用

Ty 因子可插到酵母基因的上游或下游。当它插到基因上游一定区域时,可大大促进下游基因的表达。

由 Ty 插入而引起 ADH 2 位点不同突变的菌株,它们 ADH 2 表达的水平 and 调控情况都

*注:ROAM(Regulated Overproducing Alleles responding to Matting signals);受接合信号调控的过量表达基因。

有所不同, Ty 的插入位点也不尽相同, 一般在距野生型 ADH2 转录起点上游 71—156bp 的范围内。但这些突变子中 ADH2 转录起点却都与野生型完全一样^[61]。另有一个过量产生细胞色素 C 的突变子 CYP 3-4, 在该基因野生型转录起点上游 192 bp 处有一个 Ty 插入, 这一突变子的转录起点也和野生型完全一样^[30]。真核生物中, 基因 5' 侧调控区内的“TATA”序列对转录起始有重要作用。当 Ty 插到“TATA”序列与转录起点之间时, 通常将导致该基因产物的下降, 如 ADH 2-8° 突变子就是这种情况^[61]。这表明, Ty 并不能起始下游基因的转录。

上面的实验清楚地表明, Ty 因子既不能为下游基因提供启动子, 也不能改变下游基因的转录起点, 它是通过间接的顺式作用方式在基因上游较远的区域内对之实现控制的。这种作用与某些增强子如 SV 40 中的 72 bp 重复序列相似^[34]。很可能 Ty 就是以增强子(enhancer)形式来起作用的。但与 SV 40 中的 72 bp 序列不同, Ty 的取向对其增强子作用非常重要^[20], 只有当 Ty 与受影响基因呈相反取向时, 即 Ty 的转录方向与该基因相反对, 才有明显的增强子作用。

Ty 的增强子作用与其方向有关但与 Ty 自身的转录无关。反向插在 CYC 7 基因 5' 侧的 Ty 1 序列, 不论其转录程度如何, 均可得到细胞色素 C(CYC 7 基因产物)的过量产生^[32]。Ty 因子本身转录的启动子在重复顺序 δ 中, 但 δ 序列的 2/3 缺失后, 并不影响 Ty 激活邻近基因^[20]。因此, Ty 内部的启动子甚至于 δ 序列, 对其增强子作用都不是必需的, 重要的区域一定在 ϵ 序列内。

对 SV 40 中以及其它一些已知增强子的分析, 发现有一个核心区域: 5'-TGGLLLG-3' (L 为 A 或 T) 对增强子功能是必需的^[33, 34]。在第一类 Ty 因子(Ty 1)中, 可发现至少 4 个这样的序列, 且它们都位于 Ty 1 的 ϵ 区域中^[20]。这种结构很可能就是 Ty 增强子作用的结构基础。

(三) 基因表达的不稳定性

与 Ty 相关的突变在很多情况下是不稳定的, 会以一定的频率自发地进一步产生基因表达的变化。与这些变化同时发生的, 是 Ty 因子的丢失而在原位留下一个单 δ 序列。其它类型的 DNA 变化如转位、倒置甚至转座, 也可伴随着基因表达的进一步变化而发生。如, Ty 导致的组氨酸缺陷型(His⁻表型)his 4-912 和 his 4-917, 可以较高的频率(10^{-5} 和 10^{-4})回复变成 HIS⁺ 表型^[27]。但这些回复突变子与野生型不同, 它们都是冷敏感型, 且其 HIS⁺ 性状表现为隐性(即在二倍体 his/HIS 中, 此突变子表现为 His⁻)。HIS 4 基因上游原来是 Ty 序列的地方, 在回复突变子中仅有一个单 δ 序列^[5, 35]。在 his 4-912 回复突变的纯合二倍体中, 还可观察到更复杂的 DNA 重排, 如转座、转位、倒置以及缺失等^[5, 35]。其它一些突变, 如 Ty 在 ADH 2 位点引起的突变, 也是不稳定的, 其 ADH 2 基因产物过量产生的表型, 也可以一定的频率(10^{-3} 到 10^{-7})自发突变成 ADH 2⁻ 表型^[36], 这些自发突变也表现出与野生型不同的表型。

以上这些回复突变的产生, 与 Ty 的同源重组有很大关系。Ty 两端的 δ 序列发生同源重组可造成 Ty 的切失, 并在原位留下一个单 δ 序列。Ty 与 Ty 之间、Ty 与其它 δ 之间的同源重组则可导致转位、缺失、倒置等染色体上较大范围的重排(参见图 7)。当然, 上面的那些回复突变也可能和新的 Ty 转座事件有关。

为什么回复突变后其性状会与野生型不同呢? 有一个实验为回答这个问题提供了线索^[37]。把 ADH 2 编码区上游约 170 bp 处起阻遏作用的一段 DNA 去掉, 或把这段 DNA 放到上游更远的地方, 就可使原来为诱导表达的 ADH 2 基因呈现组成型表达。Ty 因子插到 ADH 2 基因前面, 正好可使该阻遏顺序向上游移动约 6 kb。在 Ty 引起的 HIS 4 和 LYS 2 突变中也有类似的情况。由此, 人们指出了如图 9 所

示的模型,以解释这种现象^[38]。前面所提到的正调控决定子(PRDs)也许就在这里起作用。由于当只留有 δ 序列时,其活化作用消失,这更证明 δ 序列对Ty的增强作用贡献不大,很可能还在某种程度上抑制下游基因的转录^[9]。

由Ty引起的突变并不总是不稳定的。当Ty本身发生一些变化,如一端 δ 序列缺失,一部分Ty序列被丢失或取代,使Ty本身不易发生转座或自身同源重组,由Ty引起的基因表达改变即可保留下来^[32]。此时的Ty就成了退化的Ty序列。这种特性很可能在酵母基因组的进化中起过重要作用。利用这一点,我们可以用整合型载体,把经体外改变的Ty和需要表达的基因一起整合到酵母染色体上,从而得到稳定高效表达的菌株,这在酵母基因工程上将会有极其重要的应用价值。

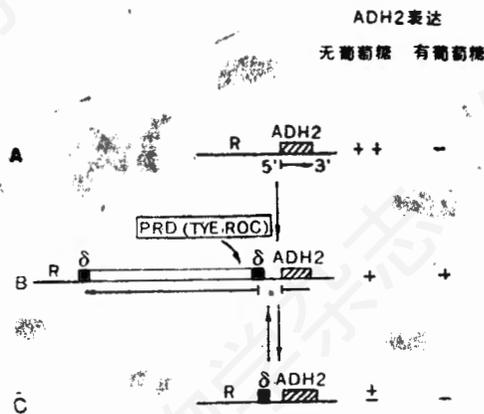


图9 Ty对ADH2基因表达的作用模型^[38]

- A. 野生型ADH2。有葡萄糖存在时,阻遏区R可抑制其表达。
 B. Ty插入后使R区远离,造成ADH2的组成型表达。这种表达至少有一部分受PRD的控制。
 C. Ty切失后留下一个单 δ 序列。不同的 δ 序列或 δ 序列插入位点不同,将使ADH2减少表达或完全不表达。

六、总结

Ty序列作为中度重复顺序,在酵母基因组中有较高的含量(0.04%),且它的转录产物Ty RNA在poly(A)⁺RNA中比例也很高(5—

10%),因此,Ty对酵母细胞的影响是十分显著的。现在已经知道,Ty对基因表达具有明显的促进作用,其作用机制类似于SV40的增强子。Ty因子作为可移动遗传因子,可能对酵母细胞基因的横向传播或推动酵母基因组的进化起着重要作用。这种可移动因子可使细胞表型产生较多的变化,在适当条件下,这些变化可保留下来,这就大大提高了酵母细胞适应环境的能力,促进了基因组的进化。同时,Ty因子也为用基因工程的方法获得新的应用性菌株提供了有利手段。

在酵母基因组中,还有一些成份可能具有转座能力。如Del Rey等人发现的一种称为Sigma的重复顺序,在单倍体中约有20—25个拷贝^[39]。尽管还没有直接的证据说明它们可以转座,但已发现它们在不同菌株中位置有很大不同,同时发现Sigma序列的两端具有反向重复顺序。另外,控制细胞接合型的MAT位点可能也有转座能力^[40];在HIS4C位点,也有报道发现有与Ty无关的转座行为^[41,42,43]。

真核生物基因组中,通常有大量的重复顺序,这些重复顺序往往对基因活性有调控作用,这也是真核基因组的一大特点。Ty因子是酵母中的中度重复顺序,它的转座能力、增强子作用以及影响基因表达的其它功能,对我们理解更高等真核生物基因表达调控也许具有启迪意义。

摘要

酵母转座子Ty因子是酵母基因组中的一组中度重复顺序,其平均长度为6 kb,两端为短的同向重复序列(δ -序列),在一级结构上与反转录病毒的原病毒形式非常相似。Ty因子自身还可以编码三个蛋白产物,其结构与功能分别和反转录病毒的gag、rev、int基因产物相关联。因此,很可能它们之间在起源上有着某种联系。Ty因子在基因组中可以用同源重组的方式或非同源重组的方式(即转座)发生位

置和数目的变动。在移动的同时,它们也明显地影响着酵母基因的表达,有时,Ty 因子可以用增强子的方式来大大提高某个酵母基因的表达水平。

参 考 文 献

- [1] McClintock, B., 1952. *Cold Spring Harber Symp. Quant. Biol.*, 16; 13—47.
- [2] Cameron, J. R., 1979. *Cell*, 16; 739—751.
- [3] Klein, H. L. and Peters, T. D., 1984. *Mol. Cell. Biol.*, 4; 329—339.
- [4] Kingsman, A. J., 1981. *J. Mol. Biol.*, 145; 619—632.
- [5] Roeder, G. S. and Fink, G. R., 1980. *Cell*, 21; 239—249.
- [6] Williamson, V. M., 1983. In "INTERNATIONAL REVIEW OF CYTOLOGY", 83; 1—25.
- [7] Roeder, G. S. and Fink, G. R., 1983. In "Mobile Genetic Elements", ed. by Shapiro, J. A., pp. 299—326. New York.
- [8] Warmington, J. R., 1985. *Nucleic Acid Res.*, 13 (18); 6679—6693.
- [9] Boeke, J. D., 1985. *Cell*, 40; 495—500.
- [10] Lewin, B., 1985. In "Genes" 2nd Edition, ed. by Cell, New York.
- [11] Rothstein, R. J. and Sherman, F., 1980. *Genetics*, 94; 891.
- [12] Benz, E. W., 1980. *Nature*, 288; 665—669.
- [13] Clare, J., 1985. *Proc. Natl. Acad. Sci. U. S. A.*, 82; 2829—2833.
- [14] Fulton, A. M., 1985. *Nucleic Acid Res.*, 13; 4097—4112.
- [15] Hauber, J., 1985. *Nucleic Acid Res.*, 13; 2745—2758.
- [16] Wilson, W., 1986. *Nucleic Acid Res.*, 14; 7001—7006.
- [17] Shilinjack, T. K., 1981. *Nature*, 293; 543—548.
- [18] Varmus, H. E., 1983. In "Mobile Genetic Elements", Academic press, New York.
- [19] Winston, F., et al., 1984. *Cell*, 39 (3 Part 2); 675—682.
- [20] Winston, F., et al., 1986. *Nucleic Acid Res.*, 14 (17); 6885—6900.
- [21] Dubois, E., Jacobs, E. and Jauniaux, J. C., 1982. *EMBO J.*, 1; 1133—1140.
- [22] Jurg Gafner, 1980. *Nature*, 268; 414—418.
- [23] Roeder, G. S., et al., 1980. *Science*, 209; 1375—1380.
- [24] Eibel, H. and Philippon, P., 1982. *Int. Conf. Yeast Genet. Mol. Biol.*, 11 th 17. (abstr.).
- [25] Warmington, J. R., 1986. *Nucleic Acid Res.*, 14 (8); 3475—3485.
- [26] David, C. LaPorte, 1986. *TIBS*, 11; 273.
- [27] Errede, B., Cardello, T. S., Sherman, F., Dubois, E., Deschamps, J., and Wiame, J. M., 1980. *Cell*, 22; 427—436.
- [28] Young, T., et al., 1982. In "Genetic Engineering of Microorganism for chemicals", Plenum, New York.
- [29] Errede, B., Company, M., Ferchak, J. D. et al., 1985. *Proc. Natl. Acad. Sci. U. S. A.*, 82; 5423—5427.
- [30] Montgomery, D. L., Boss, J. M., et al., 1982. *J. Biol. Chem.*, 257; 7756—7761.
- [31] Gruss, P., 1984. *DNA*, 3; 1—5.
- [32] Company, M. and Errede, B., 1986. *Mol. Cell. Bio.*, 6(10); 3299—3311.
- [33] Conteras, R. et al., 1982. *Nature*, 300; 500—505.
- [34] Weiher, H., et al., 1983. *Science*, 219; 626—631.
- [35] Chaleff, D. T. and Fink, G. R., 1980. *Cell*, 21; 227—237.
- [36] Ciriacy, M. and Williamson, V. M., 1981. *Mol. Gen. Genet.*, 182; 159—163.
- [37] Beier, D. and Young, E. T., 1982. *Nature*, 300; 724—728.
- [38] Williamson, V. M., et al., 1983. *Mol. Cell. Biol.*, 3; 20—31.
- [39] Del Rey, F. J., Donahue, T. F. and Fink, G. R., 1982. *Proc. Natl. Acad. Sci. U. S. A.*, 79; 4138—4142.
- [40] Herskowitz, I. and Oshina, Y., 1981. In "The Molecular Biology of Yeast *Scaccharomyces*" ed. by Strathern, J. N. et al., Cold Spring Harber, New York.
- [41] DeBruijn, F. and Greer, H., 1981. *Mol. Cell. Biol.*, 1; 381—386.
- [42] Greer, H. and Fink, G. R., 1979. *Proc. Natl. Acad. Sci. U. S. A.*, 76; 4006—4010.
- [43] Greer, H., Igo, M. and DeBruijn, F., 1980. *Cold Spring Harber Symp. Quant. Biol.*, 45; 567—574.