

分子生物学综合复习题

第1章 绪论

第2章 基因的概念

1. Basic Concepts

- **重复基因(duplicate factor)**: 即在基因组中有多个拷贝的基因——在真核生物基因组中发现这种现象, 真核生物中的重复基因可以达到 30%。重复基因主要是为了满足生物体快速发育的需要。
- **断裂基因/不连续基因(Split genes / interrupted or discontinuous genes)**: 即在基因编码蛋白质的序列中插入与蛋白质编码无关的 DNA 间隔区, 使一个基因分隔成不连续的若干区段。
- **跳跃基因/转座子(Jumping gene / Transposable element)**: 即可移动的或可转移的遗传因子。——这些因子不仅存在于细菌中, 同时也存在于较高等的动物中。
- **假基因(Pseudogenes)**: 是基因组中与编码基因序列非常相似的非功能性基因组 DNA 拷贝, 一般情况都不被转录, 且没有明确生理意义。——根据其来源可分为: 保留了间隔序列的复制假基因(如珠蛋白假基因家族)和缺少间隔序列的已加工假基因。
- **重叠基因(overlapping gene)**: 指在同一段 DNA 顺序上, 由于阅读框架不同或终止早晚不同, 同时编码两个以上基因的现象。

2. Know-How

● DNA 的一级、二级和高级结构

(一) 核酸的一级结构

核酸的碱基顺序是核酸的一级结构。多聚核苷酸是由四种不同的核苷酸单元按特定的顺序组合而成的线性结构聚合物, 因此, 它具有一定的核苷酸顺序, 即碱基顺序。

DNA 的碱基顺序本身就是遗传信息存储的分子形式。生物界物种的多样性即寓于 DNA 分子中四种核苷酸千变万化的不同排列组合之中。

(二) DNA 的二级结构

(1) DNA 分子由两条多聚脱氧核糖核苷酸链(简称 DNA 单链)组成。两条链沿着同一根轴平行盘绕, 形成右手双螺旋结构。螺旋中的两条链方向相反, 即其中一条链的方向为 5'→3', 而另一条链的方向为 3'→5'。

(2) 嘌呤碱和嘧啶碱基位于螺旋的内侧, 磷酸和脱氧核糖基位于螺旋外侧。碱基环平面与螺旋轴垂直, 糖基环平面与碱基环平面成 90°角。

(3) 螺旋横截面的直径约为 2 nm, 每条链相邻两个碱基平面之间的距离为 3.4 nm, 每 10 个核苷酸形成一个螺旋, 其螺距(即螺旋旋转一圈)高度为 3.4 nm。

(4) 两条 DNA 链相互结合以及形成双螺旋的力是链间的碱基对所形成的氢键。碱基的相互结合具有严格的配对规律, 即腺嘌呤(A)与胸腺嘧啶(T)结合, 鸟嘌呤(G)与胞嘧啶(C)结合, 这种配对关系, 称为碱基互补。A 和 T 之间形成两个氢键, G 与 C 之间形成三个氢键。在 DNA 分子中, 嘌呤碱基的总数与嘧啶碱基的总数相等。

DNA 二级结构的多型性

A-DNA(碱基平面不再与螺旋轴垂直, 而是位于大沟中, RNA-RNA/DNA 杂交分子具有这种结构)

B-DNA(典型的 Watson-Crick 双螺旋, 碱基平面基本上与螺旋轴垂直)

C-DNA:

D-DNA

E-DNA

这些不同构象的 DNA 都有共同之处: 都是右手双螺旋; 两条反向平行的核苷酸链通过 Watson-Crick 碱基配对结合在一起; 链的重复单位是单核苷酸; 这些螺旋中都有两个螺旋沟, 分为大沟与小沟, 只是它们的宽窄和深浅程度有所不同。

三股螺旋 DNA 通常是一条同型寡核苷酸与寡嘧啶核苷酸-寡嘌呤核苷酸双螺旋的大沟结合。

郑用链分子生物学综合复习题 08 基础兽医学邹晓冬(基准实验室, HZAU)

(三) DNA 的三级结构

指 DNA 链进一步扭曲盘旋形成超螺旋结构。超螺旋是 DNA 三级结构的主要形式。

所有的 DNA 超螺旋都是由 DNA 拓扑异构酶产生。DNA 分子具有相同的结构, 但链环数不同, 所以称它们为拓扑异构体。拓扑异构酶能够催化它们之间的转换。

DNA 拓扑异构酶(topoisomerase I, II)有两类

(1) 拓扑异构酶 I: 作用是暂时切断一条 DNA 链, 形成酶-DNA 共价中间物而使超螺旋 DNA 松弛化, 然后再将切断的单链 DNA 连接起来, 而不需要任何辅助因子。如: 大肠杆菌的 ϵ 蛋白(MW=110000)就是一种典型的拓扑异构酶 I。

(2) 拓扑异构酶 II: 能将负超螺旋引入 DNA 分子, 该酶能暂时性地切断和重新连接双链 DNA, 同时需要 ATP 水解为 ADP 以供能。如: 大肠杆菌中的 DNA 旋转酶(DNAgyrase)。

环状 DNA 的拓扑学特征

(1) 链环数(linking number, DNA 双链分子的交叉数): 指在双螺旋 DNA 中, 一条链以右手螺旋绕另一条链缠绕的次数, 以 L 表示。

①L 是整数; ②在 cccDNA 中任何拓扑学状态中其值保持不变; ③右手螺旋对 L 取正值。

(2) 初级螺旋圈数(缠绕数 twisting number): 指 DNA 分子中 Watson-Crick 的螺旋数, 以 T 表示。

①可以是非整数; ②是变量; ③右手螺旋时 T 为正值。

(3) 超螺旋数(writhing number)以 W 表示。

①可以是非整数; ②是变量; ③右手螺旋时, W 取负值。

举例说明这三者之间的关系: $L=T+W$

如有松弛的 B-DNA, 共有 420bp, 一条链绕另一链 42 次, 正好形成 42 圈螺旋, 则 $L=T=42$, 所以 $W=0$, 无超螺旋。若固定一端, 另一端按顺时针方向旋转 6 圈, 使双螺旋解开 6 圈, 再将双链连接成闭合环, 而 DNA 仍要保持 B-DNA 的结构, 每个螺旋由 10 个 bp 组成, T 又变成最初的 42, 即有 42 个螺旋, 则 $L=36$, $T=42$, 所以 $W=-6$, 即形成了 6 个负超螺旋(右旋), 在此情况下, T 值保持不变。

(四) 生物体内的超螺旋

(1) 环状 DNA 的超螺旋: 大肠杆菌 DNA 是环状分子, 整个分子长达 1mm 多, 在体内压缩成 1 μ m 直径的小体, 包含了几个独立的超螺旋。

(2) 真核生物的染色体(chromosome)中的 DNA

染色单体由一条连续的 DNA 长链, 经过四级盘旋、折叠而成。——一条 DNA 的长链经过一级结构即形成核小体后, 其长度被压缩 7 倍; 二级结构, 即形成螺旋管后, DNA 长度又被压缩了 6 倍; 三级结构, 即由螺线管形成超螺线管后, DNA 的长度在二级结构的基础上被压缩了 40 倍; 在由三级到四级结构, 即形成染色单体后, DNA 的长度在三级结构的基础上被压缩了 5 倍。因此由一条 DNA 长链, 经过多级螺旋化, 可以使几厘米长的 DNA 与组蛋白等物质共同形成几微米长的染色体, 其长度总共被压缩 8000~10 000 倍。

核小体(nucleosomes)由 H2A、H2B、H3、H4 各两个分子生成的八聚体和由大约 200bpDNA 组成。八聚体在中间, DNA 分子盘绕在外, 而 H1 则在核小体的外面。核小体的组装能导致 DNA 的超螺旋化。

端粒(telomere)是真核生物线形染色体的末端具有的一种特殊结构。真核生物染色体的天然末端, 是细胞必需的遗传组分, 因为它能够补偿染色体末端遗传信息的丢失。端粒是染色体末端的 DNA 重复序列, 作用是保持染色体的完整性。细胞分裂一次, 由于 DNA 复制的原因, 端粒就缩短一点。一旦端粒消耗殆尽, 染色体则易于突变而导致动脉硬化和某些癌症。因此, 端粒和细胞老化有明显的关系。一直以来都知道精、卵细胞的端粒比成年体细胞的都长许多。

● DNA 在体内稳定存在的因素: 氢键、磷酸酯键、碱基堆积力(非特异性结合力)、疏水作用力。

★氢键为弱键, 可加热解链; 磷酸酯键为强键, 需酶促解链。

★0.2 mol/L Na⁺生理盐条件, 可消除 DNA 单链上磷酸基团间的静电斥力。

(1) DNA 分子中非极性的嘌呤、嘧啶环(-NH₂等)使有序的以氢键连接的水分子层绕, 熵值降低。

(2) DNA 分子中非极性碱基的聚集, 使有序的水分子层绕减低到最低的限度, 减少熵值的降低。

(3) 磷酸基团间的静电斥力。

(4) 碱基间的挤压、抵御使其内能增加, 碱基间有序排列的状态破坏(氢键作用力被减弱)。

郑用链分子生物学综合复习题 08 基础兽医学邹晓冬(基准实验室, HZAU)

● DNA 的变性和复性, 核酸分子杂交

(一) 核酸的变性 (Denaturation)

指核酸双螺旋区的多聚核苷酸链间的氢键断裂, 变成单链结构的过程。变性核酸将失去其部分或全部生物活性。核酸的变性并不涉及磷酸二酯键的断裂, 所以它的一级结构(碱基顺序)保持不变。

能够引起核酸变性的因素很多。温度升高、酸碱度改变、甲醛和尿素等的存在均可引起核酸变性。

监测 DNA 是否变性的一个最常用的指标是 DNA 在紫外区 260nm 波长处的吸收光值变化。因为 DNA 变性时, DNA 双链发生解离, 共轭双键更充分暴露, 故 DNA 变性, DNA 在 260nm 处的吸收光度值增加, 并与解链程度有一定的比例关系, 这种关系叫做 DNA 的增色效应。

DNA 的变性从开始到解链完全, 是在一个相当窄的温度内完成的, 在这一范围内, 紫外光吸收值增加达到最大增加值的 50% 时的温度叫做 DNA 的解链温度 (T_m)。一种 DNA 分子的 T_m 值的大小与其所含碱基中的 G+C 的比例相关也与 DNA 分子大小及变性条件有关, G+C 的比例越高, DNA 分子越长, 溶液离子强度越高, T_m 值越大。

常用的 DNA 变性方法

(1) 热变性方法: 缺点是高温可能引起磷酸二酯键的断裂, 得到长短不一的 DNA。

(2) 碱变性方法: pH11.3 时, 可淘汰全部氢键, DNA 完全变成单链。

(3) 维持单链状态: pH>11.3 或盐浓度<0.01mol/L。

(二) DNA 的复性(Renaturation)——复性机制: 10-20bp 成拉链。(有减色效应)

指变性 DNA 在一定条件下恢复成天然 DNA 结构的过程。(热变性的 DNA 可通过缓慢冷却而复性——分子生物学中也称为退火。)一般认为, 比 T_m 值低 25℃ 的温度是 DNA 复性的最佳条件。

★热变性 DNA 在缓慢冷却时可以复性, 快速冷却不能复性。

★DNA 片段越小、浓度越大, 复性越快;

复性条件:

(1) 有足够的盐浓度以消除磷酸基的静电斥力: 0.15 ~ 0.5 mol/L NaCl。

(2) 有足够高的温度以破坏无规则的链内氢键 (但不能太高, 否则配对碱基之间的氢键又难以形成) 一般: 低于 T_m 20 ~ 25 C 的温度。

(三) DNA 分子的杂交 (molecular hybridization)

是用一个 DNA 或 RNA 单链与另一被测 DNA 单链形成双链, 以测定某特异序列是否存在的方法。这种方法已成为遗传学和分子生物学等生命学科中最为普遍和重要的方法之一。

分子杂交的方法多种多样, 但其共同特点是:

①都是应用复性动力学原理;

②都必须有探针 (probe) 的存在。——所谓探针就是用同位素或非同位素如荧光染料, 生物素等标记的短片段特异 DNA 或 RNA 序列。

分子杂交的方法

(1) 原位分子杂交 (in situ hybridization)

①玻片原位杂交: 进行染色体的基因定位, 或观察组织中不同的 RNA 的分布。

②膜上原位杂交: 在基因文库中钓基因。

(2) 斑点杂交 (dot blotting): 也叫狭缝杂交。由 Southern blot 衍生而来。它是将某种生物的总 DNA 直接点在尼龙膜上, 或者通过一个小的长形狭缝印在尼龙膜上, 后将膜变性处理, 预杂交后, 经中和、洗脱、干燥。再将其和探针放入复性缓冲溶液中缓缓复性, 洗涤干燥后进行放射自显影, 观察结果。这种方法是用来初步探测某种生物中含有一种特殊的基因或序列。用以分析 DNA 样品之间的同源性。

(3) Northern 杂交——检验 RNA: 提取某种生物或组织的总 RNA 或 mRNA, 用含有变性剂的琼脂糖凝胶电泳分离 RNA, 分离后再将凝胶上的 RNA 带吸印到尼龙膜上, 在液相中和标记的核酸探针进行杂交, 用此方法可以测定某基因表达的时空特异性

(4) 基因芯片(Gene chip)技术: 是指通过微阵列(Microarray)技术将高密度 DNA 片段阵列通过高速机器人或原位合成方式以一定的顺序或排列方式使其附着在如玻璃片等固相表面, 以荧光标记的 DNA 探针, 借助碱基互补杂交原理, 进行大量的基因表达及监测等方面研究的最新革命性技术。

DNA 芯片为生物芯片的一种。

郑用链分子生物学综合复习题 08 基础兽医学邹晓冬（基准实验室，HZAU）

生物芯片包括：DNA 芯片、蛋白质芯片、其它芯片。

按制备方式分：

(1) 原位合成芯片：采用显微光蚀刻等技术在特定部位原位合成寡核苷酸而制备的芯片。探针较短。

(2) DNA 微集阵列：将预先制备的 DNA 片段以显微打印的方式有序地固化于支持物表面而制成的芯片。探针的来源较灵活。

DNA 芯片的应用

(1) DNA 测序、杂交测序 (SBH)；

(2) 基因表达分析；

(3) 基因组研究：作图、测序、基因鉴定、基因功能分析；

(4) 基因诊断：寻找和检测与疾病相关的基因及在 RNA 水平上检测致病基因的表达；

(5) 药物研究与开发。

基因表达谱 (gene expression pattern) 芯片的应用最为广泛，技术上也最成熟。这种芯片可以检测整个基因组范围的众多基因在 mRNA 表达水平的变化。它能对来源于不同个体、不同组织、不同细胞周期、不同发育阶段、不同分化阶段、不同生理病理、不同刺激条件下的组织细胞内基因表达情况进行对比分析。从而对基因群在个体、组织、发育、分化、疾病、刺激等特异性的变化特征和规律进行描述。

● 原核、真核生物基因结构的异同

相同点：

都由能够编码蛋白质的编码区和具有调控作用的非编码区组成。

不同点：

原核细胞编码区是连续的，所有碱基对均能编码蛋白质，而真核细胞编码区是间隔的、不连续的，它有不表达的序列，叫做内含子，编码蛋白质的序列叫做外显子。内含子和非编码区合成非编码序列。

● 真核生物与原核生物基因组特点的异同

相同点：

都由生物基本单位中的所有核酸序列组成，都有重复序列和单一序列，都是生物的遗传物质。

不同点：

1) 真核生物基因分布在多个染色体上，而原核生物只有一个染色体；

2) 真核生物基因组远大于原核生物基因组；

3) 真核生物细胞中 DNA 与组蛋白和大量非组蛋白结合，并有核膜将其与细胞质隔离，结果真核细胞的转录和翻译在时间上和空间上都是分离的，而原核细胞的基因转录和翻译是同步的；

4) 真核生物的基因是不连续的，中间存在不被翻译的内含子序列，而原核生物几乎每一个基因都是完整的连续的 DNA 片段；

5) 基因组中非编码序列远多于编码序列；

6) 存在着重复序列，重复次数从几次到几百万次不等；

7) 真核生物基因组的复制起点多，缺少明显的操纵子结构，而原核生物基因组一般是一个复制子；

8) 真核生物基因组与原核相同，存在转座因子。

● 原核生物与真核生物基因表达调控的异同

相同点：

1) 都包括转录水平的调控和转录后调控，并且以转录水平的调控最为重要。

2) 结构基因中都存在着许多特异的调控成份。

不同点：

1) 原核生物的染色体是裸露的 DNA，染色质结构对基因的表达没有明显的调控；真核生物的染色质 DNA 与组蛋白紧密结合形成的核小体，染色质结构对基因的调控是明显的。

2) 原核生物中有正调控（乳糖操纵子）和负调控；真核生物迄今已知的主要是正调控，而且一个真核基因通常都有多个调控序列，必须有多个激活物同时特异地结合去才能启动基因的转录。

3) 原核生物在多细胞中表达，不同细胞的表达不一样，并且其转录和翻译都在同一地点同时进行；真核生物在单细胞中表达，并且其转录在细胞核中进行，翻译在胞质中进行。

第3章 DNA 复制

1. Basic Concepts

- **复制子(replicon):** 基因的一个单位。为 DNA 分子中能从起始点进行复制的部分。
- **复制体(replisome):** 是参与DNA复制的蛋白质复合物, 包含DNA聚合酶, 引发酶, 解旋酶, 单链结合蛋白和其它辅助因子。
- **半保留复制(semiconservative replication):** DNA 复制的一种方式。每条链都可用作合成互补链的模板, 合成出两分子的双链 DNA, 每个分子都是由一条亲代链和一条新合成的链组成。
- **半不连续复制(semidiscontinuous replication):** DNA 分子复制在分子全长中从若干起点向其两侧进行复制, 然后经 DNA 连接酶将复制成的片段连接为一个完整分子, 可使复制速度大大加快, 加速生长发育的过程。
- **冈崎片断(Okazaki fragment):** 相对比较短的 DNA 链(大约 1000 核苷酸残基), 是在 DNA 的滞后链的不连续合成期间生成的片段。
- **端粒(telomere):** 是真核细胞染色体末端的特殊结构。人端粒由 6 个碱基重复序列(TTAGGG)和结合蛋白组成。端粒有重要的生物学功能, 可稳定染色体的功能, 防止染色体 DNA 降解、末端融合, 保护染色体结构基因, 调节正常细胞生长。
- **端粒酶(telomerase):** 是使端粒延伸的反转录 DNA 合成酶。是个由 RNA 和蛋白质组成的核糖核酸-蛋白复合物。其以 RNA 组分为模板, 蛋白组分具有催化活性, 以端粒 3'末端为引物, 合成端粒重复序列。端粒酶的活性在真核细胞中可检测到, 其功能是合成染色体末端的端粒, 使因每次细胞分裂而逐渐缩短的端粒长度得以补偿, 进而稳定端粒长度。主要特征是用它自身携带的 RNA 作模板, 通过逆转录合成 DNA。端粒酶在细胞中的主要生物学功能是通过其逆转录酶活性复制和延长端粒 DNA 来稳定染色体端粒 DNA 的长度。
- **酵母人工染色体(yeast artificial chromosome, YAC):** 是利用酿酒酵母的染色体的复制元件构建的载体, 其工作环境在酿酒酵母中。它可以克隆和分析大片段的染色体 DNA, 并且可以分离那些在大肠杆菌(Escherichia coli)中不可能得到的序列。将来自其他物种的较大片段 DNA 连接到酵母 DNA 上, 在宿主细胞中外来 DNA 能随着酵母细胞中的其他染色体一起复制。
- **细菌接合(conjugation/mating):** 是细菌通过细胞的暂时沟通和染色体或质粒DNA的转移而导致基因重组的过程。
- **拓扑异构酶(topoisomerase):** 是一种可通过切断 DNA 的一条或两条链中的磷酸二酯键, 然后重新缠绕和封口来改变 DNA 连环数的酶。存在于细胞核内, 他们能够催化 DNA 链的断裂和结合, 从而控制 DNA 的拓扑状态。
- **质粒不相容性(plasmid incompatibility):** 指同种的或亲缘关系相近的两种质粒不能同时稳定地保持在一个细胞内的现象。当两种质粒同时导入同一细胞时, 它们在复制及随后分配到子细胞的过程中彼此竞争, 在一些细胞中, 一种质粒占优势, 而在另一些细胞中另一种质粒却占上风。当细胞生长几代后, 占少数的质粒将会丢失, 因而在细胞后代中只有两种质粒的一种。
- **Klenow fragment (Klenow 片断):** 它具有 3'到 5'的外切酶活性和 5'到 3'的聚合功能, 而去除了完整的 DNA 聚合酶 I 所具有的 5'到 3'外切酶活性, 可用于 DNA 双链末端 3'突出端切平和 5'突出端补平反应。

- **Sanger sequencing (sanger 测序):** 即 Sanger (1977) 发明的双脱氧核糖核酸链末端终止法。它根据核苷酸在某一固定的点开始, 随机在某一特定的碱基处终止, 并且在每个碱基后面进行荧光标记, 产生以 A、T、C、G 结束的四组不同长度的一系列核苷酸, 然后在尿素变性的 PAGE 胶上电泳进行检测, 从而获得可见的 DNA 碱基序列。

Sanger 法测序的原理是: 每个反应含有所有四种脱氧核苷酸三磷酸(dNTP)使之扩增, 并混入限量的一种不同的双脱氧核苷三磷酸(ddNTP)使之终止。由于 ddNTP 缺乏延伸所需要的 3'-OH 基团, 使延长的寡聚核苷酸选择性地在 G、A、T 或 C 处终止, 终止点由反应中相应的双脱氧而定。每一种 dNTPs 和 ddNTPs 的相对浓度可以调整, 使反应得到一组长几个至千以上个, 相差一个碱基一系列片断。它们具有共同的起始点, 但终止在不同的核苷酸上, 可通过高分辨率变性凝胶电泳分离大小不同的片段, 凝胶处理后可用 X-光胶片放射自显影或非同位素标记进行检测。

郑用链分子生物学综合复习题 08 基础兽医学邹晓冬(基准实验室, HZAU)

- **Pyrosequencing (焦磷酸测序):** 是一种对短到中等长度的DNA序列样品进行高通量的、精确和重复性好的分析的技术。是由 4 种酶催化的同一反应体系中的酶级联化学发光反应, 在每一轮测序反应中, 只加入一种三磷酸脱氧核(糖核)苷 (dNTP), 若该dNTP与模板配对, 聚合酶就可以将其掺入到引物链中并释放出等摩尔数的焦磷酸基团 (PPi)。PPi可最终转化为可见光信号, 并由PyrogramTM转化为一个峰值。每个峰值的高度与反应中掺入的核苷酸数目成正比。(然后加入下一种dNTP, 继续DNA链的合成。)

焦磷酸测序的原理是:

第 1 步: 1 个特异性的测序引物和单链DNA模板结合, 然后加入酶混合物(包括DNA聚合酶、ATP硫酸化酶、荧光(素)酶、腺苷三磷酸双磷酸酶)和底物混合物(包括腺苷酰硫酸APS和荧光素。

第 2 步: 向反应体系中加入 1 种dNTP, 如果它刚好能和DNA模板的下一个碱基配对, 则会在DNA 聚合酶的作用下, 添加到测序引物的 3'末端, 同时释放出一个分子的焦磷酸(PPi)。

第 3 步: 在ATP硫酸化酶的作用下, 生成的PPi可以和APS结合形成ATP; 在荧光素酶的催化下, 生成的ATP又可以和荧光素结合形成氧化荧光素, 同时产生可见光。通过CCD光学系统即可获得一个特异的检测峰, 峰值的高低则和相匹配的碱基数成正比。

第 4 步: 反应体系中剩余的dNTP和残留的少量ATP在Apyrase的作用下发生降解。

第 5 步: 加入另一种dNTP, 使第 2-4 步反应重复进行, 根据所获得的峰值图即可读取准确的DNA序列信息。

- **454 sequencing (454 测序):** 由454 生物科学发明, 它使用类似于焦磷酸测序的方法, 是一种依靠生物发光进行DNA序列分析的新技术。在DNA聚合酶, ATP硫酸化酶, 荧光素酶和双磷酸酶的协同作用下, 首先将引物上每一个dNTP的聚合与一次荧光信号释放偶联起来, 通过检测荧光信号释放的有无和强度, 以达到实时测定DNA序列的目的。454 的特点不需要荧光标记的引物或核酸探针, 也不需要进行电泳; 具有分析结果快速、准确、灵敏度高和自动化的特点。

- **Solexa sequencing (Solexa 测序):** 是一种利用单分子阵列的新一代高通量低成本的测序技术, 目前在 DNA 和 RNA 以及甲基化等方面的应用广泛。

Solexa 法测序的原理是: 首先将 DNA 从细胞中提取, 然后将其打断到约 100-200bp 大小, 再将接头连接到片段上, 经 PCR 扩增后制成 Library。随后在含有接头的芯片上将已加入接头的 DNA 片段绑定在 flow cell 上, 经反应, 将不同片段扩增。在下一步反应中, 四种荧光标记的染料应用边合成边测序的原理, 在每个循环过程里, 荧光标记的核苷和聚合酶被加入到单分子阵列中。互补的核苷和核苷酸片断的第一个碱基配对, 通过酶加入到引物上。多余的核苷被移走。这样每个单链 DNA 分子通过互补碱基的配对被延伸, 利用生物发光蛋白, 比如萤火虫的荧光素酶, 可通过碱基加到引物后端时所释放出的焦磷酸盐来提供检测信号。针对每种碱基的特定波长的激光激发结合上的核苷的标记, 这个标记会释放出荧光。荧光信号被 CCD 采集, CCD 快速扫描整个阵列检测特定的结合到每个片断上的碱基。通过上述的结合, 检测可以重复几十个循环, 这样就有可能决定核苷酸片断中的几十个碱基。

- **SOLiD sequencing (SOLiD 测序):** 其全称为 Sequencing by Oligonucleotide Ligation and Detection, 它是通过寡核苷酸连接和检测来进行测序的方法。与聚合酶测序方法不同的是, SOLiD 系统通过连接酶(而不是聚合酶读取序列)并利用专利的逐步连接 (stepwise ligation) 技术来产生高质量的数据, 可应用于全基因组测序、染色质免疫共沉淀 (ChIP)、微生物测序、数字核型分析、临床测序、基因型分析、基因表达分析和小分子 RNA 的发现等等。SOLiD 测序的化学及读出过程与 Sanger 测序及其他边合成边测序方法不同。

SOLiD 法测序的原理是: SOLiD 技术采用四色荧光标记寡核苷酸进行连续连接合成代替聚合酶链接反应为基础, 取代了传统的聚合酶连接反应, 可对单拷贝 DNA 片断进行大规模扩增和高通量并行测序。这种替代能够明显减少因碱基错配而出现的错误, 消除相位不同步的问题, 以获得更高的保真度。另外, SOLiD 系统采用末端配对分析和双碱基编码技术, 在测序过程中对每个碱基判读两遍, 从而减少原始数据错误, 以提供内在的校对功能, 它是目前新一代基因分析技术中准确度最高的。SOLiD 的特点在于其无可比拟的高通量(每轮运行可产生超过 40 亿碱基的可定位数据)、可扩展性(开放式的上样玻片格式和灵活可变的磁珠密度可通过更新操作流程来提高检测通量)、精确性(原始碱基数据的准确度达到 99.94%, 而在 15X 覆盖率时的准确度可高达 99.999%)、方便性(实时跟踪运行状态, 允许在多个步骤中重新进入流程)和运用灵活性(鉴于以上的诸多优势, 所以它应用范围十分广泛)。

2. Know-How

● 如何验证“DNA 半保留复制”、“DNA 半不连续复制”?

DNA半保留复制的证明: 1957年由Mathew Meselson和Franklin Stahl设计。

- (1)把大肠杆菌培养在含有单一氮源 ^{15}N 的培养基上培养多代后, DNA 呈重密度;
- (2)转移到仅含有 ^{14}N 的培养基上增殖一代;
- (3)从这些细胞制备的 DNA 在 CSCI 梯度离心中形成一条带, 它的密度表明这些 DNA 分子为一条含 ^{15}N 和另一条含 ^{14}N 的杂合链;
- (4)让细胞再在 ^{14}N 介质中增殖两代, 提取的 DNA 在 CSCI 中形成两条带, 一条是重、轻的杂和链, 另一条为纯轻链。

由此证明 DNA 复制是半保留的不是全保留的。

DNA 半不连续复制的证明: 1968 年由日本学者冈崎等提出。

用含 ^3H 的dT标记用T4噬菌体感染的大肠杆菌, 在短时间内对DNA进行分离, 得到的均为DNA小片段; 而过一段时间后对DNA进行分离, 则检测到 DNA大片段。当使用DNA连接酶的缺失变异株时, 检测到大量DNA片段的积累, 由此证明DNA复制中有小片段合成, 即可验证DNA的半不连续复制。

● 质粒复制原点(ori)的克隆与鉴定

假设该质粒来自细菌B(非E. coli), 将该质粒经过适当酶切, 克隆到一个克隆载体(E. coli的质粒, 携带某抗生素R的抗性基因)中, 转化E. coli建立一个该质粒的小文库, 再电转化到细菌B中, 涂含R的培养基平板, 阳性克隆(可以在该质粒宿主B中复制)进行测序, 可初步获得该质粒的ori序列, 然后进一步精确定位(外切酶酶切逐步缩短, 再转化验证)。——重组质粒转化法。

● 质粒拷贝数: 调控、测定

质粒拷贝数是指在正常生长条件下, 每个细菌细胞或每条染色体所对应的平均质粒数。

在质粒复制子的调控下, 质粒拷贝数可随细菌培养条件的变化在一个较窄的范围内波动; 生长条件恒定时, 质粒增殖的速度与宿主细胞增殖的速度完全一致, 拷贝数保持不变。

低拷贝数的质粒 DNA 在宿主细胞分裂前只能复制 1~2 次, 而高拷贝数质粒可以在细胞分裂前复制成 10~200 拷贝。低拷贝数的质粒一般是严紧型质粒(stringent plasmid), 它们的复制随细菌染色体的复制同步进行; 高拷贝数的质粒称为松弛型质粒(relaxed plasmid), 独立于细菌细胞而自主复制。

质粒的拷贝数依赖于各种胞内外因素: 质粒拷贝数首先决定于自身的遗传性, 控制质粒拷贝数的基因位于一个包括 DNA 复制起点在内的质粒 DNA 的区域内。其次, 质粒的拷贝数也受细胞生长条件的影响。细胞的生长速率增大时, 质粒的拷贝数下降, 主要是因为在高比的生长速率下, 质粒的复制速度跟不上细胞分裂的速度, 从而质粒拷贝数不断下降。此外, 培养时的温度和培养基的组成都对质粒拷贝数有影响, 比如当营养物限制或缺乏时会使质粒的拷贝数下降。

测定胞内质粒拷贝数的方法, 大致可以归纳为两类:

- (1)是直接进行测定的方法, 包括一些物理分离的方法和杂交的方法;
- (2)是间接进行测定的方法。

物理分离的方法就是将质粒 DNA 与染色体 DNA 进行分离并定量。将质粒 DNA 与染色体 DNA 分开主要是基于质粒 DNA 的两大特性: ①质粒 DNA 是以共价闭合环状的形式存在, 与染色体 DNA 的线状形式不同; ②质粒 DNA 的分子量比染色体 DNA 要小得多。这些方法主要包括: 氯化铯-溴化乙锭平衡梯度离心法、凝胶电泳法、高效液相色谱法以及毛细管电泳法等。

核酸杂交法是根据同源核酸序列可以配对的原理, 人工制备出与特定片段有广泛顺序一致性的 DNA 或 RNA 探针, 通过观察是否有 DNA-DNA 或 DNA-RNA 杂交分子的生成来检测特定基因的方法。通常用来测定质粒拷贝数的杂交方法有 Dot-blot 和 Southern-blot 两种。

间接测定质粒拷贝数, 即测定基因的剂量, 不是象其它方法那样直接测定质粒的浓度, 而是通过确定质粒所编码的某个基因产物的功能来反映质粒的含量。这种方法一般都是通过测定某个酶的活性来实现的, 最常用的是测定 β -内酰胺酶的活性。采用这种方法的前提是确证所分析的表现型与基因剂量之间是呈线性关系, 并且这个基因定位在质粒上。

郑用链分子生物学综合复习题 08 基础兽医学邹晓冬(基准实验室, HZAU)

● F 质粒的复制与接合转移

接合转移是两个完整的细菌细胞通过性菌毛直接接触,由供体细菌将质粒 DNA 转移给受体细菌的过程。接合性质粒有 F 质粒、R 质粒等。F 质粒是 1952~1953 年间由莱德伯格等人最先发现的一种质粒。

F 质粒(决定细菌的性别)的复制与接合转移

有性菌毛的细菌有 F 质粒,相当于雄性菌,以 F⁺表示;无性菌毛的细菌无 F 质粒,相当于雌性菌,以 F⁻表示。F⁺菌的性菌毛末端与 F⁻菌表面上的受体结合,结合后性菌毛渐渐缩短,使两菌紧靠在一起, F⁺菌中 F 质粒的一股 DNA 链断开,逐渐由细胞连接处伸入 F⁻菌,继而以滚环模式进行复制。所以在受体菌获得 F 质粒时,供体菌并不失去 F 质粒;受体菌在获得 F 质粒后即变为 F⁺菌,也长出性菌毛。

● 线性 DNA 末端的复制

所有已知的核酸聚合酶,无论是 DNA 聚合酶还是 RNA 聚合酶都只从 5'端向 3'端移动,新链的合成方向与聚合酶移动方向一致,即只能是 5'→3';而对于 DNA 的合成必需一段引物的存在,体内 DNA 复制时,由一段 RNA 引物起始 DNA 合成,起始后它必须切除,切除后,5'端如何起始呢?

4 种解决途径:

- (1) 通过将线性复制子转变为环状或多聚分子。如 λ 噬菌体:滚环产生多聚复制子);
- (2) 某种蛋白质可能会介入,在真正的末端上启动。几种线性病毒核酸具有与 5'端碱基共价结合的蛋白质,其中了解最清楚的例子是腺病毒 DNA)。
- (3) DNA 可形成特殊的结构,如在末端形成发夹。使分子没有游离末端。草履虫 (Paramecium) 的线性线粒体 DNA 的复制中就形成了一种交联;
- (4) 末端是可变的,而不是精确确定的。真核生物染色体可能采用这种方式,在这种情况下, DNA 末端的短重复序列的拷贝数改变(如端粒的复制)。

● 常见的三种 DNA 复制模式的特点

(1) Cairns 模型(Cairns model):即 θ 模型,是环状双链 DNA 的一种复制模型。

细菌 DNA 的复制遵循环状 DNA 分子双向复制的原则,首先在复制点形成一个复制“泡”,随之沿着环的两个方向进行复制,泡逐渐扩大,形成像希腊字母“θ”的形状,故环状 DNA 的双向复制模式称为 θ 模型,最后由一个 DNA 环复制为两个子环。

➤ 最好归纳为半保留半不连续复制的情况。(θ 复制相对较少)

(2) 滚环模型(rolling-circle model):即 σ 模型,是环状单链 DNA 的一种复制模型。

某些病毒或细菌的双链环状 DNA 或以某种方式转变的双链环状 DNA,在复制时由对正链复制原点处进行单链特异性切割,所形成的 5'端即从双股 DNA 置换出来,并为 SSB 所覆盖,这时的 DNA 聚合酶 III 以 3'-OH 为引物,以负链为模板,从 3'-OH 基端逐步增添脱氧核苷酸,随着复制的进行,5'端长度即不断增加,这一过程中,单链尾巴的延伸伴随着双链 DNA 的绕轴转动,故称为滚环式复制。

真核 rDNA 的扩增、F 因子 DNA 的转移、λ 裂解途经的复制以及 φX174、M13、G4 等环状单链 DNA 噬菌体的第二阶段复制都是进行滚环复制的。

滚环复制又叫 σ 复制,有以下特点:

- ① 共价延伸。先在一条亲代链(+)链)上切一缺口,然后以另一环状负链为模板,在正链切口上的 3'-OH 上延伸,新合成的链和正链连在一道;
- ② 模板链和新合成的链分开。3'端不断延环状模板合成,5'端脱离负环,随着复制,游离的 5'端越来越长,和环状模板完全分开;
- ③ 不需 RNA 引物,在正链 3'-OH 上延伸;
- ④ 只有一个复制叉;
- ⑤ 形成多联体(concatemer):5'游离的部分经半保留复制形成双链,并含有多个基因组,串联在一起称为多联体。λ 的增殖、接合,以及真核生物 rDNA 的扩增都是以这种形式。

(3) D 环模型(D-Loop model):是线粒体、叶绿体 DNA 的一种复制模型。

线粒体 DNA 的双股链由于浮力密度的不同而分为轻链(L 链)和重链(H 链),它有两个单向复制叉,两条链的复制原点不在同一点上,而且两个复制原点的激活有先有后,复制从 H 链的原点开始,以 L 链为模板,新合成的 H 链即转换为原来的 H 链,所形成的结构为取代环,或 D-环,故称 D 环复制。

郑用链分子生物学综合复习题 08 基础兽医学邹晓冬(基准实验室, HZAU)

D 环复制又叫 σ 复制, 有以下特点:

- (1) mtDNA 的每条链有 1 个独立的复制起始点序列;
- (2) H 链在 D 环中启动复制;
- (3) 当第一个 fork 移动使 L 链的起始点暴露时, L 链启动复制;
- (4) 叶绿体 DNA 采用相同的复制机制;
- (5) mtDNA 的复制是随机的;
- (6) 线粒体通过增加与其数量呈比例的基因组数目来完成其 DNA 复制;
- (7) 不同 mtDNA 复制次数可能不同, 这可能导致子代线粒体中等位基因的表现度改变;
- (8) 线粒体分离到子细胞中也是随机的。

● **DNA Polymerase I: 结构、功能与应用**

大肠杆菌 DNA 聚合酶 I 是一种依赖于 DNA 的 DNA 聚合酶, 分子量 109kd。它有 5'→3' DNA 聚合酶活性, 5'→3' 及 3'→5' 外切酶活性。

DNAPol I 在空间结构上近似球体, 直径约 65Å。在酶的纵轴上有一个约 20Å 的深沟(cleft), 带有正电荷, 这是该酶的活性中心位置, 在此位置上至少有 6 个结合位点:

- 1) 模板 DNA 结合位点;
- 2) DNA 生长链或引物结合位点;
- 3) 引物末端结合位点, 用以专一引物或 DNA 生长链的 3'-OH;
- 4) 脱氧核苷三磷酸结合位点;
- 5) 5'→3' 外切酶活性位点, 用以结合生长链前方的 5'-端脱氧核苷酸并切除之;
- 6) 3'→5' 外切酶活性位点, 用以结合和切除生长链上未配对的 3'-端核苷酸。

DNA 聚合酶 I 的功能:

① 5'→3' 的聚合作用。但不是复制染色体而是修补 DNA, 填补 DNA 上的空隙或是切除 RNA 引物后留下的空隙。

② 5'→3' 外切酶活性。切除受损伤的 DNA。它在切口平移(nick translation)中应用。

③ 3'→5' 的外切酶活性。消除在聚合作用中掺入的错误核苷酸——切口平移(nick translation); 链的置换及模板转换(template-switching)等。

④ 内切酶活性。

DNA 聚合酶 I 的应用

① 在基因工程技术中, 常用于补平限制酶切割 DNA 后产生的 3'凹端, 以进行末端标记, 从而合成 cDNA 第二链;

② 在体外诱变中从单链模板中合成双链 DNA, 用于双脱氧末端终止法进行 DNA 测序;

③ 用于 PCR。

● **DNA 甲基化: 形式、意义、检测方法**

DNA 甲基化是指在 DNA 甲基化转移酶的作用下, 在基因组 CPG 二核苷酸的胞嘧啶 5' 碳位共价键结合一个甲基基团。

DNA 甲基化主要形式: 5-甲基胞嘧啶(5-mC)、少量的 N6-甲基嘌呤(N6-mA)及 7-甲基鸟嘌呤(7-mG)。

DNA 甲基化的意义: 能引起染色质结构、DNA 构象、DNA 稳定性及 DNA 与蛋白质相互作用方式的改变, 从而控制基因表达。

DNA 甲基化的检测方法:

(1) 重亚硫酸盐(亚硫酸氢盐)修饰: 未甲基化的胞嘧啶转化成尿嘧啶; 甲基化的胞嘧啶不发生变化。

(2) 用甲基化特异性引物对进行 PCR 扩增: 未甲基化引物对("U"primers, Tube 1); 甲基化引物对("M"primer, Tube 2)。

第4章 转录

1. Basic Concepts

- **模板链(template strand):**是 DNA 双链中按碱基配对规律能指引转录生成 RNA 的一股单链, 也称有意义链或 Watson 链。
- **编码链(coding strand):** 是 DNA 双链中与模板链相对的单链, 不进行转录, 也称反义链或 Crick 链。
- **转录单位(transcription unit):** 是转录后形成一个 RNA 分子的一段 DNA 序列。
- **转录因子(transcription factor):** 是指能够结合在某基因上游特异核苷酸序列上的蛋白质, 这些蛋白质能调控其基因的转录。它是转录起始过程中 RNA 聚合酶所需的辅助因子。
- **启动子(promoter):** 是DNA模板上专一地与RNA聚合酶结合并决定转录从何处起始的部位, 也决定基因的转录效率。(启动子可以被RNA聚合酶辨认, 并开始转录。)
- **增强子(enhancer):** 是增加同它连锁的基因转录频率的 DNA 序列。它是通过启动子来增加转录的, 有效的增强子可以位于基因的 5'端, 也可位于基因的 3'端, 有的还可位于基因的内含子中。
- **沉默子(silencer):** 是参与基因表达负调控的一种元件, 是在研究 t 淋巴细胞的 t 抗原受体(tcr)基因表达调控时发现的。
- **终止子(terminator):** 是一段位于基因或操纵组末端的DNA片段, 可中断转录作用。
- **抗终止作用(anti-termination):** 指有的蛋白因子能作用于终止序列, 减弱或取消终止子的作用。
- **操纵子(operon):** 是启动基因、终止基因和一系列紧密连锁的结构基因的总称。——很多功能上相关的结构基因串联排列在染色体上, 会由一个共同的控制区来操纵这些基因的表达。
- **hnRNA(heterogeneous nuclear RNA):** 为存在于真核生物细胞核中的不稳定、大小不均的一组高分子 RNA 的总称。占细胞全部 RNA 之百分之几, 在核内主要存在于核仁的外侧。
- **snRNA(Small nuclear RNA):** 是细胞核内的小 RNA, 它是真核生物转录后加工过程中 RNA 剪接体(spliceosome)的主要成分, 参与 mRNA 前体的加工过程。
- **scRNA(Small cytoplasmic RNA):** 是细胞浆中的小 RNA, 它参与蛋白质的合成和运输。
- **snoRNAs(Small nucleolar RNAs):**核仁小分子 RNA, 是一类广泛分布于真核生物细胞核仁的小分子非编码 RNA, 具有保守的结构元件。
- **snRNPs(Small nuclear ribonucleoproteins):**是 snRNA 和蛋白质结合形成的小核核糖核蛋白颗粒。它和剪接作用有密切关系。有些 snRNPs 也可与供体、受体剪接位点及分支顺序相互补。
- **scRNPs(Small cytoplasmic ribonucleoproteins):**是 scRNA 与蛋白质结合形成的小核糖体蛋白颗粒, 是小分子细胞浆核糖核蛋白。

2. Know-How

● *E. coli* RNA 聚合酶的结构与功能

RNA聚合酶(RNA polymerase简称RNAP)或称核糖核酸聚合酶, 是一种负责以DNA或RNA为模板催化RNA合成的酶。主要指DNA指导的RNA聚合酶, 即以DNA为模板, 以4种核苷三磷酸(ATP, GTP, CTP和UTP)为底物, 从5'到3'合成RNA的酶。其催化方式与DNA聚合酶相似, 但不具备有校正作用的外切核酸酶活性, 聚合反应也不需引物。

(1) RNA 聚合酶的结构:

- 全酶($\alpha 2 \beta \beta' \sigma$): 能同启动子部位结合, 并催化转录的起始。
- 核心酶($\alpha 2 \beta \beta'$): 催化 RNA 的(延长)合成。

(2) 亚基的功能:

- ★ α 亚基可能与识别启动子区域的顺序有关。
- ★ β 亚基是 RNA 聚合酶的催化亚基, 可与 RNA 聚合酶抑制剂结合, 从而导致酶活性的丧失。
- ★ β' 亚基的功能与 RNA 聚合酶同模板 DNA 的结合有关。用酶解的方法将 β' 亚基除去, 导致该酶同 DNA 结合能力下降。

★ σ 亚基(因子)与 RNA 聚合酶对启动子部位的识别有关。一旦 RNA 聚合酶结合在正确的转录起始部位上, 且合成第一个磷酸二酯键后, σ 因子便解离出来。不同的 σ 因子识别不同的启动子, 基因的表达受不同的启动子控制。

郑用链分子生物学综合复习题 08 基础兽医学邹晓冬(基准实验室, HZAU)

● RNA 探针的体外标记方法

RNA 探针的体外标记方法主要基于一类新型载体 pSP 和 pGEM, 这类载体在多克隆位点两侧分别带有 SP6 启动子和 T7 启动子, 在 SP6RNA 聚合酶或 T7RNA 聚合酶作用下可以进行 RNA 转录, 如果在多克隆位点接头中插入了外源 DNA 片段, 则可以此 DNA 两条链中的一条为模板转录生成 RNA。这种体外转录反应效率很高, 在 1h 内可合成近 10 μ g 的 RNA 产生, 只要在底物中加入适量的放射性或生物素标记的 NTP, 则所合成的 RNA 可得到高效标记。该方法能有效地控制探针的长度并可提高标记物的利用率。

体外转录法标记 RNA 探针可用于商品化转录质粒来制备。这些质粒中应该包括 SP6、T3、或 T7 RNA 聚合酶的 RNA 启动位点, 而这些启动位点则与多克隆位点 (multiple cloning sites, MCS) 相邻。

● 真核生物 RNA 聚合酶的分类、核定位、转录产物

分类: 真核生物的 RNA 聚合酶主要有三种, 即 RNA 聚合酶 I、RNA 聚合酶 II 和 RNA 聚合酶 III。

核定位:

(1) RNA 聚合酶 I 定位在核仁, 负责 rRNA 基因的转录。

(2) RNA 聚合酶 II 定位在核质, 负责转录产生 mRNA 前体, 即核不均一 RNA-hnRNA, 需要“TATA”框。

(3) RNA 聚合酶 III 定位在胞质, 负责 tRNA 基因如 5SrRNA 基因和其他小 RNA 的生成。

转录产物:

(1) RNA 聚合酶 I 转录产物是 45-S 核蛋白体 RNA(45S-rRNA), 经剪接修饰生成除 5SrRNA 外的各种核蛋白体 RNA (rRNA), 由 rRNA 与蛋白质组成的核蛋白体即核糖体, 是蛋白质合成的场所。

(2) RNA 聚合酶 II 转录生成 hnRNA 和 mRNA, 是真核生物中最活跃的 RNA 聚合酶。RNA 聚合酶 II 转录在核内转录生成不均一核 RNA (hnRNA), 然后加工成 mRNA 并输送给细胞质的蛋白质合成体系。mRNA 是蛋白质合成的模板, 为各种 RNA 中寿命最短、最不稳定的, 需经常重新合成。在此意义上说, RNA 酶 II 是真核生物中最活跃的 RNA 聚合酶, 在转录从起始过渡到延长有重要作用。

(3) RNA 聚合酶 III 转录的产物都是 小分子量的 RNA, 如 tRNA、5SrRNA 和 snRNA。转运 RNA (tRNA) 负责转运氨基酸, 大小都在 100 核苷酸以下。小分子核内 RNA (snRNA) 有多种、由 90~300 核苷酸组成, 参与 RNA 剪接过程。

Sigma 因子与 Rho 因子

Sigma (σ) 因子: 是原核生物 RNA 聚合酶全酶的成份。Sigma 因子本身并无催化功能, 但是它对在正确启动子位点起始转录是必要的。

sigma 因子决定 RNA 聚合酶识别特异性, 在识别转录起始区中起关键作用; 原核生物细胞仅有一种 RNA 聚合酶, 核心酶参与转录延长, 全酶具转录起始作用。在转录起始阶段, sigma 因子识别特异启动子序列; 不同的 sigma 因子决定特异基因的转录激活。

Rho 因子: 是原核生物转录终止因子, 有 ATP 酶和解螺旋酶两种活性。

在原核生物中、转录与翻译是同时进行的, 意味着核糖体追赶着 DNA 指导的 RNA 聚合酶。Rho 因子是一个依赖于 RNA 的 ATP 水解酶, 能够在转录过程中将在转录泡中的 RNA-DNA 杂合体分开。因此它顺着转录的方向(5'→3')追赶 DNA 指导的 RNA 聚合酶。若核糖体正好妨碍了它的前进, 则依赖于肋 σ 因子的终止反应不会发生。所以在细菌转录终止中很少涉及到 Rho 因子。

● 原核生物与真核生物 promoter 的结构特点

原核生物启动子

原核生物启动子是由两段彼此分开且又高度保守的核苷酸序列组成。对 mRNA 的合成极为重要。

启动子区域包括:

(1) Pribnow 盒: 位于转录起始位点上游 5-10bp, 一般由 6~8 个碱基组成, 富含 A 和 T, 故又称为 TATA 盒或 -10 区。启动子来源不同, Pribnow 盒的碱基顺序稍有变化。

(2) -35 区: 位于转录起始位点上游 35bp 处, 一般由 10 个碱基组成。

启动子有强弱之分, 虽然原核细胞仅靠一种 RNA 聚合酶就能负责所有 RNA 的合成, 但它却不能识别真核基因的启动子。为了表达真核基因, 必须将其克隆在原核启动子的下游, 才在原核表达系统中被转录。在原核生物表达系统中, 通常使用的可调控的强启动子有 lac (乳糖启动子)、trp (色氨酸启动子)、PL 和 PR(λ 噬菌体的左向和右向启动子)以及 tac (乳糖和色氨酸的杂合启动子)等。

郑用链分子生物学综合复习题 08 基础兽医学邹晓冬(基准实验室, HZAU)

真核生物启动子:

真核基因启动子是由基因转录起始位点(+1)及其5'上游近端大约100~200bp以内(或下游100bp)的一组具有独立功能的DNA序列组成。每个元件长度约为7~20bp,是决定RNA聚合酶转录起始和转录频率的关键元件。

启动子区域包括:

A、核心启动子:包括转录起始位点(+1)(一般是A或G)及转录起始位点上游-25/-30bp处富含TA的典型元件TATA框。它是使RNA聚合酶II转录正常起始所必需的、最少的DNA序列。

B、上游启动子元件:包括通常-70bp附近的CAAT框(GGCCAATCT)和GC框(GGGCGG)等,能通过TFII-D复合物调节转录起始的频率,提高转录效率。

原核与真核生物启动子结构的差异

真核生物的不同启动子间不像原核那样有明显共同一致的序列,而且单靠RNA聚合酶难以结合DNA而起动转录,而是需要多种蛋白质因子的相互协调作用。并且真核生物的不同启动子序列也很不相同,要比原核更复杂、序列也更长。(真核生物的基因调控区、启动区范围均比较大。)

● **Promoter 的鉴定与活性测定**

转化的外源基因在生物受体细胞中能否正常表达是基因工程所关注的重要问题,因为具有经济性的外源基因在一定水平上有效表达是评价其经济价值的重要指标。当外源基因导入生物体后,其表达受许多因素的影响,但启动子在决定基因表达方面具有决定性的作用,启动子调控基因表达能力的大小决定了其控制的基因的表达强度。因此采用有效的启动子能够显著地改变外源基因在植物中的表达水平。

启动子的鉴定与活性测定:

利用绿色荧光蛋白(GFP)基因作为报告基因,对启动子在受体细胞中调控基因表达的活性进行检测:采用基因枪法对受试细胞进行遗传转化,培养24h后在荧光显微镜下观察GFP基因的表达情况,并通过软件测定不同培养条件下GFP基因表达后的荧光强度,以较为准确的鉴定该启动子的活性大小,并达到量化的目的。

——GFP是一些腔肠动物所特有的生物荧光素蛋白,GFP在其他原核或真核细胞中,在紫外光或蓝光激发下都发出绿光,即发射团的形成无物种特异性,也不需要特异的辅助因子。它简单、省时、可靠。另外,基因枪转化方法也没有宿主限制。

● **转录起始位点?**

转录起始位点是指与新生RNA链第一个核苷酸相对应的DNA链上的碱基,研究表明常为一个嘌呤(A或G)。

寻找并确定起始位点的方法:

(1) 引物延伸分析(primer extension analysis)

引物延伸分析用于定量mRNA的量,测定低丰度的mRNA的种类。另外引物延伸实验可标定转录产物的5'-端,确定转录的精确起始。特异的末端标记的引物退火到RNA链的互补区域,随后用RNA作为模板,用反转录酶延伸引物得到cDNA,用变性聚丙烯酰胺凝胶电泳分析cDNA。cDNA的长度为引物的标记的核苷酸到RNA-5'末端间的碱基数,所得cDNA的量和目的RNA的起始量成正比。(理想的引物是ssDNA,长度为20-40个核苷酸,与要分析的转录产物的5'末端区域互补。延伸产物小于150个核苷酸,可得到最佳分离效果。)

(2) 核酸酶S1作图(Mapping RNA with Nuclease S1)

三种不同的核酸酶——S1核酸酶、RNA酶、外切核酸酶VII被用来进行RNA定量,确定内含子位置,及鉴定在克隆的DNA模板上mRNA的5'-末端和3'-末端的位置。含有目的mRNA的RNA样品与互补的DNA或RNA探针在利于形成杂合体的条件下温育。在反应的末期,用核酸酶来降解未杂合的单链RNA和DNA,余下的DNA-RNA或RNA-RNA杂合体用凝胶电泳分离,接着用自显影或Southern杂交来观察。当探针的摩尔数在杂交反应中过量时,信号强度与样品中目的mRNA的浓度成正比。用过量探针与一系列定量的靶序列杂交作出标准曲线,从而可准确估算样品的浓度。在真核系统中,S1核酸酶消化过的5'-末端通常代表转录起始点,而3'-末端代表聚腺苷酸化的位点。同样的策略可用来对拼接位点的5'-末端和3'-末端作图。

● **RNA 转录后加工的常见形式及其机制**

转录后的加工 (posttranscriptional modification) 是指将各种前体 RNA 分子加工成成熟的各种 RNA。在原核和真核生物中 mRNA, tRNA 和 rRNA 都是来自于转录, 在真核中还有第四种 RNA 分子即 snRNA, 同样也是转录的产物。在细胞内, 由 RNA 聚合酶合成的原初转录物往往需要一系列的变化, 包括链的裂解、5' 和 3' 末端的切除和特殊结构的形成、核苷的修饰、以及拼接和编辑等过程, 才转变为成熟的 RNA 分子。

三种形式:

- ①减少部分片段: 如切除 5' 端前导序列, 3' 端尾巴和中部的内含子;
- ②增加部分片段: 5' 加帽, 3' 加 poly(A), 通过归巢插入内含子;
- ③修饰: 对某些碱基进行甲基化等, 从而指导 RNA (gRNA) 为模板在 mRNA 上插入或删除一些碱基, 以增加信息量、校正遗传信息和调控表达。

RNA 转录后加工的机制:

(1) 原核生物中 RNA 的加工

mRNA 一般不进行转录后加工, 一经转录通常立即进行翻译。

rRNA 的基因与某些 tRNA 基因组成混合操纵子, 转录后需切割, 断链成为 rRNA 和 tRNA, 如大肠杆菌 rRNA 前体的加工: rRNA 前体先经甲基化修饰, 再被切割。

tRNA 前体加工包括: 由内切酶在 tRNA 两端切断; 3' 端切去附加顺序, 进行修剪; 3' 端加上-CCA-OH; 核苷酸修饰和异构化, ψ 和 T 由 U 修饰而来。

(2) 真核生物中 RNA 的加工

大多数真核基因都是断裂基因, 转录产物需通过拼接, 去除插入部分 (即内含子), 使编码区 (外显子) 成为连续序列。RNA 编码序列改变称为编辑, RNA 编码和读码方式的改变称为再编码。由于存在选择性拼接、编辑和再编码, 一个基因可以产生多种蛋白质。

真核生物 mRNA 前体的加工: 必须经复杂过程, 还有 5' 端加帽子, 3' 端加 polyA。

rRNA 和 tRNA 的加工: 也需剪接、编辑、内部甲基化、修饰异构化、戴帽和加尾。

● **逆转录酶的结构、功能与应用**

逆转录酶 (reverse transcriptase) 又称 RNA 指导下的 DNA 聚合酶、RNA 合成酶、RNA 复制酶, 它是以 RNA 为模板指导三磷酸脱氧核苷酸合成互补 DNA (cDNA) 的酶。

逆转录酶的结构

- (1) 哺乳类 C 型病毒的反转录酶和鼠类 B 型病毒的反转录酶都是一条多肽链。
- (2) 鸟类 RNA 病毒的反转录酶由两个结构相关的亚单位组成。
- (3) 真核生物中的反转录酶则结构多样。

逆转录酶的功能

以 dNTP 为底物, 以 RNA 为模板, 以 tRNA (主要是色氨酸 tRNA) 为引物, 在 tRNA 3'-OH 末端上, 按 5'→3' 方向, 合成一条与 RNA 模板互补的 DNA 单链, 这条 DNA 单链叫做互补 DNA, 它与 RNA 模板形成 RNA-DNA 杂交体。随后在逆转录酶的作用下, 水解掉 RNA 链, 再以 cDNA 为模板合成第二条 DNA 链。至此, 完成由 RNA 指导的 DNA 合成过程。

逆转录酶是多功能酶, 有 3 种酶活力:

- (1) 形成杂合分子 RNA-DNA, 为 RNA 指导下 DNA 聚合活力。
- (2) 形成双链 DNA, 为 DNA 指导下的 DNA 聚合活力。
- (3) 水解 RNA-DNA 杂合分子中的 RNA, 为核酸外切酶活力。

逆转录酶无校正功能, 因此错误率较高。(逆转录酶只对病毒本身的 RNA 起作用。)

逆转录酶的应用

- (1) 可作为合成某些特定 RNA 的互补 DNA 的工具酶。
- (2) 可用于 DNA 的序列分析和克隆重组 DNA。

反转录酶已成为一种重要的工具酶。用组织细胞提取 mRNA 并以它为模板, 在反转录酶的作用下, 合成出互补的 DNA (cDNA), 由此可构建出 cDNA 文库 (cDNA library), 从中筛选特异的目的基因, 这是在基因工程技术中最常用的获得目的基因的方法。同时, 它也可用来标记 cDNA 而作为放射性的分子探针。

第5章 翻译

1. Basic Concepts

- **单顺反子 (monocistronic) mRNA**: 指只编码一条多肽链的 mRNA 分子, 即一个编码基因转录生成一个 mRNA 分子、经翻译生成一条多肽链, 每个基因转录有各自的调节元件。见于真核生物。
- **多顺反子 (polycistronic) mRNA**: 指可编码多条肽链的 mRNA 分子, 即包含多个为 RNA 编码或多肽链编码的序列。见于大多数原核生物。
- **密码子的简并性(degeneracy)**: 指同一种氨基酸具有两个或更多个密码子的现象。
- **Two-fold degenerate codon (二重简并位点)**: 指那些处于密码子位置, 可被翻译成同一种氨基酸的两种不同的核苷酸替换成另外两个核苷酸后, 则会导致氨基酸的翻译发生变化的密码子位点。
- **Four-fold degenerate codon (四重简并位点)**: 指那些改变了一个核苷酸为任何其它三个核苷酸中的一个, 而对核糖体将氨基酸插入到蛋白质不产生任何影响的密码子位点。
- **同义密码子(synonym)**: 为同一种氨基酸编码的各个密码子。
- **Codon family(密码子家族)**: 为同一种氨基酸编码的所有密码子, 它们相互间仅在第三密码子位置上有差异。
- **3'-UTR (3' Untranslated Region, 3'非编码区)**: 为从终止密码子至转录终止的一段非翻译区。是调控翻译的一个重要组成部分。
- **5'-UTR (5' Untranslated Region, 5'非编码区)**: 为从转录起始位点至起始密码子的一段非翻译区。是调控翻译的一个重要组成部分。
- **密码子偏性(codon bias)**: 指生物体中编码同一种氨基酸同义密码子的非均衡使用现象, 由于这一现象与遗传信息的载体分子 DNA 和生物功能分子蛋白质相关联, 所以具有重要的生物学意义。
- **同工(isoaccepting) tRNA**: 指携带有相同氨基酸的 tRNA。
- **RBS (ribosomal binding site, 核糖体结合位点)**: 是起始密码子AUG上游的一段非翻译区序列。(RBS 中有SD序列。)——通常将RBS等同为SD序列, 它们之间几乎无差异。
- **SD序列 (SD sequence)**: 是原核mRNA的起始部位, 由一段富含嘌呤的特殊核苷酸顺序组成。(它位于起始密码子上游约 4~7 个核苷酸之前, 是一段富含嘌呤的 5' ...AGGAGG...3' 短小序列, 它可与 16S rRNA 3' ...UCCUCC...5' 完全互补。以促使核糖体结合至mRNA, 利于翻译的起始。)
- **信号肽(signal peptide)**: 是能启动蛋白质运转的任何一段多肽。绝大部分被运入内质网内腔蛋白质都带有一个信号肽。
- **蛋白质的自我剪接(self-splicing)**: 指蛋白质具有自我催化能力, 能将自身的某些部位切除的现象。
- **Intein (蛋白内含子又称为内蛋白子)**: 是近年发现的一种新的翻译后加工的产物。为一种在蛋白水平上通过自我剪接所进行的蛋白质转译后加工的方式。它是 1994 年由 Perler 等首先提出的。它是一个自我拼接的蛋白质元件, 类似于基因组中的内含子 intron 在 RNA 的剪接中所起的重要作用。
- **Extein (蛋白外显子又称为外蛋白子)**: 与蛋白内含子相对应。蛋白内含子和蛋白内外显子是由一个开放读码框(ORF)所编码的, 它们共同产生一个mRNA分子, 在产生前体蛋白后再去掉内蛋白子, 余下的外蛋白子以肽键的方式连接在一起成为成熟的蛋白。
- **Proteomics (蛋白质组学)**: 指在大规模水平上研究蛋白质的特征, 包括蛋白质的表达水平, 翻译后的修饰, 蛋白与蛋白相互作用等, 由此获得蛋白质水平上的关于疾病发生, 细胞代谢等过程的整体而全面的认识。
- **MALDI (Matrix-assisted laser desorption ionization, 基质辅助激光解吸附质谱技术)**: 其基本原理是将分析物分散在基质分子中并形成晶体, 当用激光照射晶体时, 由于基质分子经辐射所吸收的能量, 导致能量蓄积并迅速产热, 从而使基质晶体升华, 致使基质和分析物膨胀并进入气相。(MALDI所产生的质谱图多为单电荷离子, 因而质谱图中的离子与多肽和蛋白质的质量有一一对应关系。MALDI产生的离子常用飞行时间检测器 (Time-of-Flight, TOF) 来检测质量数, 因此合称MALDI-TOF, 适合对蛋白质、多肽、核酸和多糖等生物大分子的研究。)
- **TOF (time of flight, 激光传感器)**: 其原理就是一个旋转得反射镜, 将激光光束或者超声波按一定间隔反射出去, 然后根据旋转所得角度和时间差来得到不同角度得距离值。(距离=速度×时间)

郑用链分子生物学综合复习题 08 基础兽医学邹晓冬(基准实验室, HZAU)

- **SPR (Surface Plasmon Resonance, 表面等离子体共振):** 是一种物理光学现象。表面等离子体共振检测是一种利用表面等离子体波 (SPW , Surface Plasma Wave) 进行检测的技术, 它具有无须标记、高速化、专一性、灵敏度高以及大量平行筛选等优点。表面等离子体 (SP) 是沿着金属和电介质间界面传播的电磁波形成的。当平行表面的偏振光以称之为表面等离子体共振角入射在界面上, 发生衰减全反射时, 入射光被耦合入表面等离子体内, 光能大量被吸收, 在这个角度上由于表面等离子体共振引起界面反射光显著减少。由于 SPR 对金属表面电介质的折射率非常敏感, 不同电介质其表面等离子体共振角不同。同种电介质, 其附在金属表面的量不同, 则 SPR 的响应强度不同。基于这种原理的生物传感器通常将一种具特异识别属性的分子即配体固定于金属膜表面, 监控溶液中的被分析物与该配体的结合过程。在复合物形成或解离过程中, 金属膜表面溶液的折射率发生变化, 随即被 SPR 生物传感器检测出来等等。

2. Know-How

- **翻译过程中所需的 3 种 RNA 的结构特点和功能**

(1) mRNA 的结构特点和功能

结构特点: 5'-端有 7-甲基鸟苷三磷酸 (m⁷GpppN) 帽子。

3'-端有多聚腺苷酸(polyA)尾巴。

功能: 为蛋白质的合成提供模板。

(2) tRNA 的结构特点和功能

结构特点: 各种 tRNA 的一级结构互不相同; 但它们的二级结构都呈三叶草形, 这种三叶草形结构的主要特征是: 含有四个螺旋区、三个环和一个附加叉。四个螺旋区构成四个臂, 其中含有 3'末端的螺旋区称为氨基酸臂, 因为此臂的 3'-末端都是 C-C-A-OH 序列, 可与氨基酸连接。三个环分别用 I、II、III 表示。环 I 含有 5, 6 二氢尿嘧啶, 称为二氢尿嘧啶环 (DHU 环)。环 II 顶端含有由三个碱基组成的反密码子, 称为反密码环; 反密码子可识别 mRNA 分子上的密码子, 在蛋白质生物合成中起重要的翻译作用。环 III 含有胸苷 (T)、假尿苷 (ψ)、胞苷 (C), 称为 TψC 环; 此环可能与结合核糖体有关。tRNA 在二级结构的基础上进一步折叠成为倒“L”字母形的三级结构。

功能: 转运氨基酸时用于合成蛋白质。

(3) rRNA 的结构特点和功能

结构特点: rRNA 是单链, 它包含不等量的 A 与 U、G 与 C, 但是有广泛的双链区域。在双链区, 碱基因氢键相连, 表现为发夹式螺旋。它一般与核糖体蛋白质结合在一起, 形成核糖体, 如果把 rRNA 从核糖体上除掉, 核糖体的结构就会发生塌陷。原核生物的核糖体所含的 rRNA 有 5S、16S 及 23S 三种。S 为沉降系数 (sedimentation coefficient), 当用超速离心测定一个粒子的沉淀速度时, 此速度与粒子的大小直径成比例。5S 含有 120 个核苷酸, 16S 含有 1540 个核苷酸, 而 23S 含有 2900 个核苷酸。而真核生物有 4 种 rRNA, 它们分子大小分别是 5S、5.8S、18S 和 28S, 分别具有大约 120、160、1900 和 4700 个核苷酸。

功能: 与核糖体蛋白共同构成核糖体——蛋白质的合成部位, 参与蛋白质的合成。

- **如何正确理解通用密码子的“三性”: 通用性、不重叠性、简并性**

在 mRNA 分子中每相邻的三个核苷酸编成一组, 在蛋白质合成时, 代表某一种氨基酸, 称为密码子, 亦称三联体密码。

(1) 通用性: 指不同的生物密码子基本相同, 即共用一套密码子。除线粒体的个别密码外, 生物界通用一套遗传密码, 细菌、动物和植物等不同物种之间, 蛋白质合成机制及其 mRNA 都是可以互换的。例如, 真核生物的基因可以在原核生物中表达, 反之亦然。

(2) 不重叠性: 指两个密码子间没有标点符号, 读码必须按照一定的读码框架, 并且从正确的起点开始, 一个不漏地一直读到终止信号。对于特定的三联体密码而言, 其中的每个核苷酸都具有不重叠性。例如, 如果 RNA 分子 UCAGACUGC 的密码解读顺序为: UCA、GAC、UGC, 则它不可以同时解读为 UCA、CAG、AGA、GAC……等。不重叠性使密码解读简单而准确无误。

(3) 简并性: 指绝大多数氨基酸具有 2 个以上不同的密码子。编码相同氨基酸的密码子称同义密码子。由于简并性, 某些 DNA 碱基变化不会引起相应蛋白质的氨基酸序列改变, 这对维持物种的稳定性有重要意义。

郑用链分子生物学综合复习题 08 基础兽医学邹晓冬(基准实验室, HZAU)

● 密码子偏性(codon bias)及其应用

密码子偏性是指生物体中编码同一种氨基酸同义密码子的非均衡使用现象, (某一物种或某一基因通常倾向于使用一种或几种特定的同义密码子) 由于这一现象与遗传信息的载体分子 DNA 和生物功能分子蛋白质相关联, 所以具有重要的生物学意义。

应用: 密码子偏性相关的生物学领域

(1) 通过对已知基因密码子偏性模式和不同物种密码子偏性的分析, 可预测外源基因的最适宿主及在宿主中的表达水平, 或者通过基因工程手段采用最优密码子以提高基因表达水平, 从而提高其在宿主中的表达水平。

(2) 有助于更好的了解转录和翻译进程中的调控机制 (与某些物种的蛋白表达水平相关)。

(3) 揭示有关物种间或某一物种的基因家族间的基因进化规律;

(4) 增加转录及翻译的效率(如提高准确性)及减少翻译过程中消耗。

(5) 密码子偏性的分析常对许多实验操作起指导和辅助作用, 如: 鉴定编码区 制备基因克隆的寡核苷酸探针, 基因芯片设计等。

● 蛋白质合成的起始 (How is the first AA added?)

作为合成蛋白质的原料, AA 必须首先经过活化而成为氨酰 tRNA, 才能进入核糖体进入肽链的合成。

氨基酸的激活: 在合成酶和 ATP 的作用下, 氨基酸被激活且转移到 tRNA 分子上。

(1) 参与成分为氨基酸, tRNA, 氨酰tRNA合成酶, ATP, Mg^{2+} 。

(2) 氨酰 tRNA 合成酶对氨基酸和 tRNA 的识别是专一性的, 从而引出第二遗传密码系统 (第二遗传密码系统: 指的是氨酰 tRNA 合成酶与 tRNA 之间的相互识别作用, 它不仅可识别特异的氨基酸, 而且也识别特定的 tRNA, 以反映它在蛋白质合成的精确度中的关键作用), 以为起始复合物的形成提供先决条件。

而起始复合物的形成就意味着起始的开始。

起始复合物的形成过程:

(1) 30S 与 IF_1 和 IF_3 结合, IF_3 防止 30S 与 50S 结合, 结合 mRNA (AUG) 进入 P 位;

(2) GTP 与甲酰甲硫氨酰 tRNA 结合 IF_2 进入, fmet tRNA 进入 P 位, 其上反密码子与 AUG 配对;

(3) IF_2 的 GTP 酶活性水解 GTP 为 GDP 和 P_i , 并释放 IF_1 、 IF_2 、 IF_3 ;

(4) 50S 亚基由于 IF_3 释放而得以结合进去;

(5) 形成一个完整的 70S 核糖体起始复合物, A 位等待下一个氨酰 tRNA 进入。

注:

核糖体上有两个 tRNA 结合位点: P 位和 A 位。

P 位: 肽酰结合位点, 起始 tRNA 结合处。

A 位: 氨酰接受位, 延长用的 tRNA 进入此位。

起始复合物的组分: 甲酰甲硫氨酰 tRNA, AUG, 30S, 50S、起始因子 IF_1 、 IF_2 、 IF_3 , GTP 和 Mg^{2+} 。

蛋白质生物合成的分子机制

蛋白质生物合成 (protein biosynthesis): 指生物按照从脱氧核糖核酸 (DNA) 转录得到的信使核糖核酸 (mRNA) 上的遗传信息合成蛋白质的过程。由于 mRNA 上的遗传信息是以密码 (见遗传密码) 形式存在的, 只有合成为蛋白质才能表达出生物性状, 因此将蛋白质生物合成比拟为转译或翻译。蛋白质生物合成包括氨基酸的活化及其与专一转移核糖核酸 (tRNA) 的连接; 肽链的合成 (包括起始、延伸和终止) 和新生肽链加工成为成熟的蛋白质 3 大步骤。其中心环节是肽链的合成。蛋白质生物合成需核糖体、mRNA、tRNA、氨酰转移核糖核酸 (氨酰 tRNA) 合成酶、可溶性蛋白质因子等大约 200 多种生物大分子协同作用来完成。

肽链合成方向: 由 H_2N -到-COOH 即不断从 $-NH_2$ 向-COOH 端逐个加上 AA。

(1) 氨基酸的激活: 氨基酸与相应的 tRNA 相连接的过程;

(2) 起始复合物的形成: 核糖体 30s 小亚基和 50s 大亚基结合成 70s 复合物;

(3) 肽链的延长: 在 RNA 指导下合成蛋白质;

(4) 终止: 结束肽链的合成;

(5) 合成后对所得前体进行的加工并最终成为有生物活性的蛋白质。

● 分别比较原核生物、真核生物转录和翻译的控制元件

真核生物基因的结构和表达控制元件与原核生物有很大的不同。真核生物由于外显子与内含子镶嵌排列, 转录产生的 RNA 须切除内含子拼接外显子才能最后表达, 因此真核生物的基因是断裂的。真核生物的基因不能直接在原核生物表达, 只有将加工成熟的 mRNA 经逆转录合成互补的 DNA (cDNA), 再接上原核生物的表达控制元件, 才能在原核生物中表达。另外 mRNA 很不稳定, 容易被 RNA 酶分解, 因此真核生物须建立 cDNA 文库来进行克隆和表达研究。

(所谓 cDNA 文库是指细胞全部 mRNA 逆转录成 cDNA 并被克隆的总和。)

顺式作用元件 (cis-acting element): 又称分子内作用元件, 是 DNA 与 DNA 作用。是指对基因表达有调节活性的特异 DNA 序列, 其活性只影响与其自身同处在一个 DNA 分子上的基因。顺式作用元件本身不编码任何蛋白质, 仅提供一个作用位点, 要与反式作用因子相互作用而起作用。

——真核以正调控为主; 原核以负调控为主。

(1) 原核生物中, 大多数基因表达通过操纵子模型进行调控, 其顺式作用元件主要由启动基因、操纵基因和调节基因组成。

(2) 真核生物中, 与基因表达调控有关的顺式作用元件主要有启动子、增强子和沉默子。

反式作用因子 (trans-acting factor): 又称分子间作用因子、转录因子、转录调节因子或转录调节蛋白, 是蛋白质对 DNA 的作用。指一些和顺式作用元件结合的可扩散性蛋白。大多数真核转录调节因子由某一基因表达后, 可通过另一基因的特异的顺式作用元件相互作用, 从而激活另一基因的转录。反式作用因子有两个重要的功能结构域: DNA 结合结构域和转录活化结构域, 它们是其发挥转录调控功能的必需结构。反式作用因子可被诱导合成, 其活性也受多种因素的调节。

(1) 原核生物中的反式作用因子主要分为特异因子、激活蛋白和阻遏蛋白。

(2) 真核生物中的反式作用因子主要分为基本转录因子、转录激活因子和转录抑制因子。

● 蛋白质的转运与细胞定位

真核生物不但有细胞核、细胞质和细胞膜, 而且有许多具有膜结构的细胞器, 所以结构比较复杂。那么, 在细胞质内合成的蛋白质是如何到达细胞的不同部位的呢?

首先要取决于蛋白质本身的结构, 也就是它是否具有转运的信号, 同时, 定位于不同部位的蛋白质具有不同信号。对于核内蛋白质, 线粒体内蛋白质和叶绿体内蛋白质的转运机制及其信号还不甚清楚, 而较为清楚的是细胞表面蛋白质、分泌蛋白质和溶酶体蛋白质的转运途径及其信号。

细胞表面蛋白质、分泌蛋白质和溶酶体蛋白质(三者有时统称为分泌蛋白质)都具有与原核生物跨膜蛋白相似的信号肽。这些信号肽位于 N 端, 长 15 个~30 个氨基酸残基, 大部分为疏水氨基酸, 在最靠 N 端的几个氨基酸残基中, 总是有带正电荷的赖氨酸或精氨酸。像原核生物的信号肽一样, 能够形成包括两个 α -螺旋的发夹结构。这一结构不是插进细胞膜中, 而是在信号识别颗粒(signal recognition particle, 简称 SRP) 的帮助下插入到粗面内质网的膜中。

上述三类蛋白质进入内质网后, 在内质网腔内进行肽链的折叠, 并进行了初步的糖基化。然后产生转运小泡依次进入顺面(cis)高尔基体、中间高尔基体和反面(trans)高尔基体。在反面高尔基体中完成糖基化过程并进行分拣(sorting)。分拣系统所识别的也是蛋白质分子上的信号, 但这种信号不是蛋白质的一级结构, 而是其空间结构的某一区域, 这种信号称为信号补丁(signal patch)。不言而喻, 信号补丁上的氨基酸残基在一级结构中可能相距很远。进入分泌贮备小泡(secretory storage vesicles)和溶酶体的蛋白质具有这种信号补丁, 而进入细胞膜和细胞表面的蛋白质可能不需要这样的信号, 可能是通过一种称为"主流机制"(bulkflow)的途径由内质网经过高尔基复合体(由上述三种高尔基体组成)一直到达细胞表面。

蛋白质在到达细胞膜表面(可能还有各种细胞器膜表面)是怎样固定于膜的表面的呢? 通过一种非蛋白质锚定于膜的表面。这种非蛋白锚一般为脂类, 共价连接于蛋白质的羧端(为切除了原初产物羧端 15 个~30 个氨基酸的信号序列以后产生的羧端)。这种非蛋白质锚的主体部分为磷脂酰肌醇(简称 PI), 再通过己聚糖和乙醇胺的磷酸酯与蛋白质的竣端残基以酰胺键相连, 因此这类锚称为糖基-磷脂酰肌醇锚(glycosyl-phosphatidylinositol, 简称 GPI)。膜表面蛋白质接 GPI 锚的机制和场所都还不清楚, 估计可能发生在粗面内质网的腔面内膜上。GPI 的作用不单是将这些蛋白质固定于膜的表面, 而且可能是信号传导系统的成员。

● **蛋白质的翻译后修饰：磷酸化、泛素化、糖基化**

蛋白质翻译后修饰：指在 mRNA 被翻译成蛋白质后，对蛋白质上个别氨基酸残基进行共价修饰的过程。常见的有泛素化、磷酸化、糖基化、脂基化、甲基化和乙酰化等。

(1) 泛素化

泛素化是单个或多个泛素在泛素激活酶、泛素结合酶及泛素蛋白质连接酶的作用下共价修饰底物蛋白质的过程。泛素由 76 个氨基酸组成，高度保守，普遍存在于真核细胞内。共价结合泛素的蛋白质能被蛋白酶识别并降解，这是细胞内短寿命蛋白和一些异常蛋白降解的普遍途径。与消化道内进行的蛋白质水解不同，从泛素与蛋白的结合到将蛋白水解成小的肽段，整个水解过程需要能量参与——泛素-蛋白酶系统是一个对于真核细胞非常重要的调节系统。

泛素化对于细胞分化、细胞器的生物合成、细胞凋亡、DNA 修复、新蛋白生成、调控细胞增殖、蛋白质运输、免疫应答和应激反应等生理过程都起到很重要的作用。

(2) 磷酸化

磷酸化是通过蛋白质磷酸化激酶将 ATP 的磷酸基转移到蛋白的特定位点上的过程。大部分细胞过程实际上是被可逆的蛋白磷酸化所调控的，至少有 30% 的蛋白被磷酸化修饰，磷酸化的作用位点为蛋白上的 Ser, Thr, Tyr 残基。在磷酸化调节过程中，细胞的形态和功能都发生改变。

可逆的磷酸化过程几乎涉及所有的生理及病理过程，如细胞信号转导、肿瘤发生、新陈代谢、神经活动、肌肉收缩以及细胞的增殖、发育和分化等。

(3) 糖基化

蛋白质的糖基化是低聚糖以糖苷的形式与蛋白上特定的氨基酸残基共价结合的过程。

蛋白质糖基化可以按照氨基酸和糖的连接方式分为四类：O 位糖基化、N 位糖基化、C 位甘露糖化以及 GPI 锚定连接，蛋白质的糖基化影响蛋白的功能，在许多生物过程中起着重要的作用，如免疫保护、病毒的复制、细胞生长、细胞与细胞之间的黏附、炎症的产生等。很多蛋白，如转录因子、核小孔蛋白、热休克蛋白、RNA 聚合酶 II、致癌基因翻译产物、酶等，都发现了糖基化这种翻译后修饰方式，糖基化异常经常导致疾病的发生。

● **如何筛选和鉴定蛋白质 P 的受体或配体 X？**

所谓蛋白质“受体”和“配体”即两种可以相互作用并通过相互作用发挥生物学功能的两种蛋白质，通常用已知的一方钓取另一方。方法很多，如双杂交、pull-down、SPR 等技术均可。

几种研究蛋白质的相互作用的方法：

(1) 酵母双杂交系统：是当前广泛用于蛋白质相互作用组学研究的一种重要方法。其原理是当靶蛋白和诱饵蛋白特异结合后，诱饵蛋白结合于报道基因的启动子，启动报道基因在酵母细胞内的表达，如果检测到报道基因的表达产物，则说明两者之间有相互作用，反之则两者之间没有相互作用。将这种技术微量化、阵列化后则可用于大规模蛋白质之间相互作用的研究。在实际工作中，人们根据需要发展了单杂交系统、三杂交系统和反向杂交系统等。Angermayr 等设计了一个 SOS 蛋白介导的双杂交系统。可以研究膜蛋白的功能，丰富了酵母双杂交系统的功能。此外，酵母双杂交系统的作用也已扩展至对蛋白质的鉴定。

(2) pull-down 技术：蛋白质相互作用的类型有牢固型相互作用和暂时型相互作用两种。牢固型相互作用以多亚基蛋白复合体常见，最好通过免疫共沉淀 (Co-IP)、Pull-down 技术或 Far-western 法研究。Pull-down 技术用固相化的、已标记的饵蛋白或标签蛋白 (生物素-、PolyHis-或 GST-)，从细胞裂解液中钓出与之相互作用的蛋白。通过 Pull-down 技术可以确定已知的蛋白与钓出蛋白或已纯化的相关蛋白间的相互作用关系，从体外传路或翻译体系中检测出蛋白相互作用关系。

(3) SPR (Surface Plasmon Resonance, 等离子共振技术)：已成为蛋白质相互作用研究中的新手段。它的原理是利用一种纳米级的薄膜吸附上“诱饵蛋白”，当待测蛋白与诱饵蛋白结合后，薄膜的共振性质会发生改变，通过检测便可知这两种蛋白的结合情况。SPR 技术的优点是不需标记物或染料，反应过程可实时监控。测定快速且安全，还可用于检测蛋白——核酸及其它生物大分子之间的相互作用。

● **蛋白质组学关键技术的原理：**

目前蛋白质组学研究在表达蛋白质组学方面研究的最为广泛，其分析通常有三个步骤：第一步、运用 2-DE 技术分离样品中的蛋白质；第二步、应用质谱技术或 N 末端测序鉴定 2-DE 分离的蛋白质；第三步、应用生物信息学技术存储、处理、比较获得的数据。

(1) 2-DE (Two-dimensional gel electrophoresis, 双相凝胶电泳)

2-DE 在 1975 年出现，是一项广泛应用于分离细胞、组织、或其他生物样品中蛋白质混合物的技术。它根据蛋白质不同的特点分两相分离蛋白质。

第一相是等电聚焦 (IEF) 电泳，根据蛋白质等电点的不同进行分离：蛋白质是两性分子，根据其周围环境 pH 可以带正电荷、负电荷或静电荷为零；等电点 (pI) 是蛋白质所带静电荷为零时的 pH，周围 pH 小于其 pI 时，蛋白质带正电荷，大于其 pI 时蛋白质带负电荷；IEF 时，蛋白质处于一个 pH 梯度中，在电场的作用下，蛋白质将移向其静电荷为零的点，静电荷为正的蛋白将移向负极，静电荷为负的将移向正极，直到到达其等电点，如果蛋白质在其等电点附近扩散，那么它将带上电荷重新移回等电点。这就是 IEF 的聚焦效应，它可以在等电点附近浓集蛋白，从而分离电荷差别极微的蛋白。

第二相是 SDS 聚丙烯酰胺凝胶电泳 (SDS-PAGE)，根据蛋白质的分子量不同进行分离。此相是在包含 SDS 的聚丙烯酰胺凝胶中进行。SDS 是一种阴离子去污剂，它能缠绕在多肽骨架上使蛋白质带负电，所带电荷与蛋白质的分子量成正比，在 SDS 聚丙烯酰胺凝胶中蛋白质分子量的对数与它在胶中移动的距离基本成线性关系。

经过 2-DE 以后，二维平面上每一个点一般代表了一种蛋白，这样成千种不同的蛋白即可被分离，有关蛋白质的等电点、分子量及每种蛋白的数量信息也可以得到。

(2) MS (Mass Spectrometry, 质谱)

对 2-DE 所产生的上千个蛋白用传统的方法如 Edman 降解法等进行分析将是一个很艰巨的任务。质谱技术的发展解决了这一难题。它需要三个步骤，首先通过离子化装置将分子转化为气态离子，接着通过质谱分析器按照质荷比 (m/z) 的不同进行分离，最后转化到离子检测装置。目前，用来分析蛋白质和肽的样品离子化技术主要包括基质辅助激光解吸收离子化质谱 (MALDI) 和电子喷射离子化质谱 (ESI)。MALDI 通常与飞行时间质谱 (TOF) 相结合。TOF 主要用来测量分析物飞过固定的路径所需的时间。另一种鉴别蛋白质的方法是串联质谱 (MS/MS)。在这种情况下，经质谱分析的肽段进一步断裂并再次进行质谱分析，这样可得到肽序列的部分信息。

(3) SPR (Surface Plasmon Resonance, 表面等离子体共振)：（见名词解释）

第6章 基因表达调控

原核基因表达的调控主要表现为对生长环境变化的反应和适应;真核基因表达的调控主要是指 mRNA 的形成与使用的调节与控制,主要调控环节是转录水平。原核生物以负调控为主,真核生物以正调控为主。

● **原核生物基因表达调控:**原核生物在对外环境突然变化的反应中,是通过诱导或阻遏合成一些相应的蛋白质来调整与外环境之间的关系。

操纵子学说:1961年,法国科学家莫诺(J·L·Monod, 1910-1976)与雅可布(F·Jacob)发表“蛋白质合成中的遗传调节机制”一文,提出操纵子学说。莫诺与雅可布最初发现的是大肠杆菌的乳糖操纵子。

乳糖可作为培养大肠杆菌的能源。大肠杆菌能产生一种酶(叫做“半乳糖苷酶”),能够催化乳糖分解为半乳糖和葡萄糖,以便作进一步的代谢利用。编码半乳糖苷酶的基因(简称 z)是一个结构基因。这个结构基因与操纵基因共同组成操纵子。操纵基因受一种叫作阻遏蛋白的调控。当阻遏蛋白结合到操纵基因之上时,乳糖会起诱导作用,它与阻遏蛋白结合,使之从操纵基因上脱落下来。这时,操纵基因开启,相邻的结构基因也表现活性,细菌就能分解并利用乳糖了,这样,乳糖便成了诱导半乳糖苷酶产生的诱导物。

上述内容表明,大肠杆菌的乳糖操纵子是一个十分巧妙的自动控制系统:当培养基中含有充分的乳糖,同时不含葡萄糖时,细菌便会自动产生半乳糖苷酶来分解乳糖,以资利用。当培养基中不含乳糖时,细菌便自动关闭乳糖操纵子,以免浪费物质和能量。

六十年代中期,在操纵子中发现了另一个开关基因,称为启动基因。启动基因位于操纵基因之前,二者紧密相邻。启动基因由环腺苷酸(cAMP)启动,而环腺苷酸能被葡萄糖所抑制。这样,葡萄糖便通过抑制环腺苷酸而间接抑制启动基因,使结构基因失活,停止合成半乳糖苷酶。

由此可知,结构基因同时受两个开关基因——操纵基因与启动基因的调控。只有当这两个开关都处于开启状态时,结构基因才能活化。当培养基中同时存在葡萄糖和乳糖时,葡萄糖通过抑制环腺苷酸而间接抑制启动基因,并进而抑制结构基因,使细菌不产生半乳糖苷酶。这种情况下,细菌便会自动优先利用葡萄糖,因为葡萄糖是比乳糖更好的能源。

乳糖操纵子:是一种可诱导操纵子。

(1) 乳糖操纵子的结构:

大肠杆菌的乳糖操纵子含 Z、Y 及 A 三个结构基因,分别编码 β -半乳糖苷酶、透酶、乙酰基转移酶,此外还有一个操纵序列 O、一个启动序列 P 及一个调节基因 I。I 基因编码一种阻遏蛋白,后者与 O 序列结合,使操纵子受阻遏而处于转录失活状态。在启动序列 P 上游还有一个分解(代谢)物基因激活蛋白 CAP 结合位点,由 P 序列、O 序列和 CAP 结合位点共同构成 LAC 操纵子的调控区。

(2) 阻遏蛋白的负性调节:

当没有乳糖存在时,乳糖操纵子处于阻遏状态。此时, I 基因列在 P 启动序列操纵下表达的乳糖阻遏蛋白与 O 序列结合,故阻断转录启动。阻遏蛋白的阻遏作用并非绝对,偶有阻遏蛋白与 O 序列解聚。因此,每个细胞中可能会有寥寥数分子 β 半乳糖苷酶、透酶生成。

当有乳糖存在时,乳糖操纵子即被诱导。真正的诱导剂并非乳糖本身。乳糖经透酶催化、转运进入细胞,再经原先存在于细胞中的少数 β -半乳糖苷酶催化,转变为别乳糖。后者作为诱导剂分子结合阻遏蛋白,使蛋白构型变化,导致阻遏蛋白与 O 序列解离、发生转录,使 β -半乳糖苷酶分子增加 1000 倍。

(3) CAP 的正性调节:分解代谢物基因激活蛋白 CAP 是同二聚体,在其分子内有 DNA 结合区及 cAMP 结合位点。

当没有葡萄糖及 cAMP 浓度较高时, cAMP 与 CAP 结合,这时 CAP 结合在乳糖启动序列附近的 CAP 位点,可刺激 RNA 转录活性,使之提高 50 倍;

当有葡萄糖存在时, cAMP 浓度降低, cAMP 与 CAP 结合受阻,因此乳糖操纵子表达下降。

由此可见,对乳糖操纵子来说 CAP 是正性调节因素,乳糖阻遏蛋白是负性调节因素。

(4) 对调节机制的解释:大肠杆菌根据碳源性质选择代谢方式。

当有葡萄糖存在时,细菌优先选择葡萄糖供应能量。葡萄糖通过降低 cAMP 浓度,阻碍 cAMP 与 CAP 结合而抑制乳糖操纵子转录,使细菌只能利用葡萄糖。

当没有葡萄糖而只有乳糖存在时,阻遏蛋白与 O 序列解聚, CAP 结合 cAMP 后与乳糖操纵子的 CAP 位点,激活转录,使得细菌利用乳糖作为能量来源。

郑用链分子生物学综合复习题 08 基础兽医学邹晓冬(基准实验室, HZAU)

色氨酸操纵子: 是一种可阻遏操纵子。

(1) 色氨酸操纵子模型结构:

5 种结构基因: trpE、D、C、B、A;

调控结构: 启动子、操纵基因、前导序列、弱化子;

阻遏物 trpR 基因: 与 trp 操纵子相距较远。

(2) 阻遏物对色氨酸操纵子的负调控: trpR 基因编码无辅基阻遏物与色氨酸结合形成有活性的色氨酸阻遏物, 然后与操纵子结合而阻止转录。

色氨酸不足: 阻遏物三维空间结构发生变化, 不能与操纵子结合, 操纵元开始转录;

色氨酸浓度升高: 色氨酸与阻遏物结合, 空间结构发生变化, 可与操纵子结合, 阻止转录。

(3) 衰减作用对色氨酸操纵子的调控: 色氨酸操纵子转录的衰减作用通过位于 L 基因的衰减子使转录终止, 衰减子中两个相邻的色氨酸密码子及原核生物中转录与翻译的偶联是产生衰减作用的基础, 在高浓度色氨酸环境中, 衰减子的部分序列的转录产物能形成 ρ 因子不依赖的转录终止结构, 使转录停止。

● **真核生物基因表达调控**: 真核生物转录的激活与被转录区域的染色质结构变化有关; 真核生物基因表达以正调控为主, 真核生物的转录和翻译不偶联。核心途径: 环境信号转导-染色质活化-转录的激活。

启动子: 与原核启动子的含义相同, 是指 RNA 聚合酶结合并启动转录的 DNA 序列。但真核同启动子间不像原核那样有明显共同一致的序列, 而且单靠 RNA 聚合酶难以结合 DNA 而启动转录, 它需要多种蛋白质因子的相互协调作用, 不同蛋白质因子又能与不同 DNA 序列相互作用, 不同基因转录起始及其调控所需的蛋白因子也不完全相同, 因而不同启动子序列也很不相同。真核生物有 3 类 RNA 聚合酶, 负责转录 3 类不同的启动子, 分别为 I 类、II 类和 III 类。

(1) I 类启动子: 由 RNA 聚合酶 I 负责转录的 rRNA 基因, 启动子 (I 类) 比较单一, 由转录起始位点附近的两部分序列构成。第一部分是核心启动子 (core promoter), 由 -45—+20 位核苷酸组成, 单独存在时就足以起始转录。另一部分由 -170—107 位序列组成, 称为上游调控元件, 能有效地增强转录效率。

(2) II 类启动子: 由 RNA 聚合酶 II 负责转录的 II 类基因包括所有蛋白质编码基因和部分 snRNA 基因。后者的启动子结构与 III 类基因启动子中的第三种类型相似, 编码蛋白质的 II 类基因启动子在结构上有共同的保守序列。

(3) III 类启动子: 由 RNA 聚合酶 III 负责转录的是 5SrRNA、tRNA 和某些核内小分子 RNA (snRNA), 其启动子组成较复杂, 又可被分为三个亚类。两亚类 5S rRNA 和 tRNA 基因的启动子是内部启动子, 位于转录起始位点的下游, 都由两部分组成。第三亚类启动子由三个部分组成, 位于转录起始位点上游。

➤ **核心启动子 (core promoter)**: 指 RNA 聚合酶起始转录所必需的最小的 DNA 序列, 包括转录起始点及其上游 -25/-30bp 处的 TATA 盒。它单独起作用时只能确定转录起始位点和产生基础水平的转录。

➤ **上游启动子元件 (upstream promoter elements, UPE)**: 包括通常位于 -70bp 附近的 CAAT 盒和 GC 盒、以及距转录起始点更远的上游元件。这些元件与相应的蛋白因子结合能提高或改变转录效率。不同基因具有不同的上游启动子元件, 其位置也不相同, 这使得不同的基因表达分别有不同的调控。

增强子: 是一种能够提高转录效率的顺式调控元件, 最早是在 SV40 病毒中发现的长约 200bp 的一段 DNA, 可使旁侧的基因转录提高 100 倍, 其后在多种真核生物, 甚至在原核生物中都发现了增强子。增强子通常占 100-200bp 长度, 也和启动子一样由若干组件构成, 基本核心组件常为 8-12bp, 可以单拷贝或多拷贝串连形式存在。增强子的作用有以下特点:

(1) 增强子提高同一条 DNA 链上基因转录效率, 可以远距离作用, 通常可距离 1-4kb、个别情况下离开所调控的基因 30kb 仍能发挥作用, 而且在基因的上游或下游都能起作用。

(2) 增强子作用与其序列的正反方向无关, 将增强子方向倒置依然能起作用。而将启动子倒就不能起作用, 可见增强子与启动子是很不相同的。

(3) 增强子要有启动子才能发挥作用, 没有启动子存在, 增强子不能表现活性。但增强子对启动子没有严格的专一性, 同一增强子可以影响不同类型启动子的转录。例如当含有增强子的病毒基因组整合入宿主细胞基因组时, 能够增强整合区附近宿主某些基因的转录; 当增强子随某些染色体段落移位时, 也能提高移到的新位置周围基因的转录。使某些癌基因转录表达增强, 可能是肿瘤发生的因素之一。

(4) 增强子的作用机理虽然还不明确, 但与其他顺式调控元件一样, 必须与特定的蛋白质因结合后

郑用链分子生物学综合复习题 08 基础兽医学邹晓冬（基准实验室，HZAU）

才能发挥增强转录的作用。增强子一般具有组织或细胞特异性，许多增强子只在某些细胞或组织中表现活性，是由这些细胞或组织中具有的特异性蛋白质因子所决定的。

反义 RNA (anti-sense RNA): 是指与目的 DNA 或 RNA 序列互补的 DNA 或 RNA 片段。

反义 RNA 与特定的 mRNA 结合的位点通常是 SD 序列、起始密码子 AUG 和部分 N 端的密码子，从而抑制 mRNA 的翻译，所以又称这类 RNA 为干扰 mRNA 的互补 RNA。

在原核生物中反义 RNA 具有多种功能，例如调控质粒的复制及其接合转移，抑制某些转位因子的转位，对某些噬菌体溶菌-溶源状态的控制等。

由于核糖体不能翻译双链的 RNA，所以反义 RNA 与 mRNA 特异性的互补结合，即抑制了该 mRNA 的翻译。通过反义 RNA 控制 mRNA 的翻译是原核生物基因表达调控的一种方式，最早是在 E.coli 的产肠杆菌素的 Col E1 质粒中发现的，许多实验证明在真核生物中也存在反义 RNA。

通过人工合成反义 RNA 基因，并将其导入细胞内转录成反义 RNA，即能抑制某特定基因的表达，阻断该基因的功能，有助于了解该基因对细胞生长和分化的作用。同时暗示了该方法对肿瘤实施基因治疗的可能性。

RNA 干涉 (RNA interference): 是指外源 dsRNA 引发生物体内的基因的同源序列降解，从而表现出的基因转录后的沉默现象，它与植物中的共抑制和真菌中的基因压制可能具有相同的作用机制。在这一过程中，需要 eIF2c 类似蛋白因子、RNA 螺旋酶、RNA 依赖性 RNA 多聚酶、核糖核酸酶、ATP 和转膜蛋白参与。

RNA 干涉可以用于功能基因组学研究，也可用于克服转基因生物体的基因沉默现象，使外源基因在遗传改良生物中能更好地表达，还用于基因治疗，抑制有害基因的表达等。

第7章 DNA 的损伤与修复

1. Basic Concepts

- **突变 (mutation)**: 指因 DNA 损伤而导致的 DNA 分子结构的变化。
- **点突变 (point mutation)**: 指 DNA 单一碱基的变异。
- **突变热点 (hot spots of mutation)**: 指 DNA 中突变率高的区域。
- **回复突变 (reverse mutation)**: 指那些从突变型等位基因变为野生型等位基因的突变。
- **抑制突变 (suppression mutation)**: 指当发生突变时, 生物体采取了一种措施, 原来与此密码子对应的反密码子也突变为与密码子配对, 使阅读后的产物和突变后的产物一样, 而对突变进行了抑制。
- **遗传重组 (genetic recombination)**: 指任何造成基因型变化的基因交流过程。
- **同源重组 (homologous recombination)**: 指大片段同源 DNA 序列之间的交换。
- **位点特异性重组 (site-specific recombination)**: 是发生在两条 DNA 链特异位点上的重组。重组的发生需一段同源序列即特异性位点(又称附着点)和位点特异性的蛋白因子即重组酶参与催化。
- **转座重组 (transposition recombination)**: 指转座的机制依赖 DNA 的交错剪切和复制, 但不依赖于同源序列的重组。
- **重组热点 (hot spots of recombination)**: 指染色体的交换频率高的区域。
- **内含子归巢 (intron homing)**: 指在 I 类和 II 类的某些内含子中含有开放读框, 可产生具有三种功能的蛋白, 从而使内含子(或以其原来的 DNA 形式, 或作为 RNA 的 DNA 拷贝)移动, 插到一个新的靶位点的现象。
- **Holliday intermediate (Holliday 中间体)**: 也即 Holliday 交叉, 指参与同源重组的 DNA 分子之间通过形成“交叉”, 并造成彼此之间发生交换, 最终所形成的一个“四螺旋”中间物结构。(一个 DNA 的一条链断裂、并与另一个 DNA 对应的链连接, 形成 Holliday 中间体)
- **Chi seq (Chi 序列)**: 指一个能提供 E.coli 中 RecA 介导遗传重组热点的八聚体序列。

2. Know-How

● DNA 的突变有哪些类型?

- (1) 点突变: DNA 单一碱基的变异。
- (2) 缺失: 一个或一段核苷酸的消失。
- (3) 插入: 一个或一段核苷酸插入原 DNA 链。
- (4) 倒位或转位: DNA 重组使其中一段核苷酸倒置, 或从一处迁移到另一处。
- (5) 双链断裂。

● 点突变有哪些类型?

- (1) 转换 (transition): 嘌呤与嘌呤、嘧啶与嘧啶之间替换。
- (2) 颠换 (transversion): 嘌呤与嘧啶之间的替代。

● 请列举突变产生的原因(自发, 诱发)?

突变可以是自发的, 即由于未纠正的复制错误或自然的变化, 或者可能是某种化学或物理因素的诱导下产生的。

(1) 在 DNA 复制过程中, 碱基类似物通过与正常的碱基配对相互作用插入 DNA 中, 但它会在接下来的复制过程中形成错误的碱基对, 从而产生突变。

(2) 化学诱变剂改变 DNA 中存在的碱基, 它能够改变碱基的配对性质, 或者导致易错 DNA 修复。嵌入剂插入 DNA 中 2 个相邻的碱基中, 在 DNA 复制时产生移码突变。

(3) 突变还可能由插入基因中的转座因子引起, 或者当增变基因突变时, 导致其它基因产生突变。

● 真核细胞有哪些 DNA 错配修复机制?

- (1) 复制错误的修复;
- (2) 错配修复;
- (3) 切除修复;
- (4) 重组修复;
- (5) SOS 修复。

郑用链分子生物学综合复习题 08 基础兽医学邹晓冬(基准实验室, HZAU)

● 简述三种 DNA 突变检测的方法(测序、RFLP、PCR-RFLP)?

(1) DNA测序(DNA sequencing)是指分析特定DNA片段的碱基序列,也就是腺嘌呤(A)、胸腺嘧啶(T)、胞嘧啶(C)与鸟嘌呤的(G)排列方式。RNA测序则通常将RNA提取后,反转录为DNA后使用DNA测序的方法进行测序。目前应用最广泛的是由Frederick Sanger发明的Sanger双脱氧链终止法(Chain Termination Method)。新的测序方法,例如454 生物科学的方法和Sanger(桑格)双脱氧链终止法是Sanger于1975年发明的。测序过程需要先做一个PCR反应。PCR过程中,双脱氧核糖核酸可能随机的被加入到正在合成中的DNA片段里。由于双脱氧核糖核酸多脱了一个氧原子,一旦它被加入到DNA链上,这个DNA链就不能继续增加长度。最终的结果是获得所有可能获得的、不同长度的DNA片段。目前最普遍最先进的方法,是将双脱氧核糖核酸进行不同荧光标记。将PCR反应获得的总DNA通过毛细管电泳分离,跑到最末端的DNA就可以在激光的作用下发出荧光。由于ddATP, ddGTP, ddCTP, ddTTP(4种双脱氧核糖核酸)荧光标记不同,计算机可以自动根据颜色判断该位置上碱基究竟是A, T, G, C中的哪一个。

(2) RFLP(Restriction Fragment Length Polymorphism, 限制性内切酶片段长度多态性)是指基因型之间限制性片段长度的差异,这种差异是由限制性酶切位点上碱基的插入、缺失、失重排或点突变所引起的。RFLP是发展最早的分子标记技术。RFLP技术的原理是:检测DNA在限制性内切酶酶切后形成的特定DNA片段的大小。因此凡是可以引起酶切位点变异的突变如点突变(新产生和去除酶切位点)和一段DNA的重新组织(如插入和缺失造成酶切位点间的长度发生变化)等均可导致RFLP的产生。此技术及其从中发展出来的一些变型均包括以下基本步骤:DNA的提取、用限制性内切酶酶切DNA、用凝胶电泳分开DNA片段、把DNA片段转移到滤膜上、利用放射性标记的探针显示特定的DNA片段(通过Southern杂交)、分析结果。由于线粒体和叶绿体DNA较小,前三个步骤就完全可能检测出DNA片段的差异,所以往往不必要后面的几个步骤。

(3) PCR-RFLP(polymerase chain reaction - restriction fragment length polymorphism, 聚合酶链式反应-限制性片段长度多态)是基于限制性核酸内切酶识别位点的特异性一项检测方法(因限制性核酸内切酶不能识别所有的碱基变化)。PCR-RFLP分析技术是在PCR技术基础上发展起来的。PCR-RFLP技术的原理是:采用聚合酶链式反应(PCR)扩增目的DNA片段,然后将待检测的DNA片段用限制性内切酶酶切,限制性内切酶识别并切割特异的序列,然后将酶切后的产物进行电泳,再由限制酶图谱分析此段序列的特异切位点,藉由片段的多样性来比对不同来源基因序列的差异性。

● 遗传重组有哪几种类型?

(1) 同源重组(homologous recombination)或普遍性重组(generalized recombination):涉及到大片段同源DNA序列之间的交换。在真核生物减数分裂过程中发生的这种重组是一种交互重组的类型。

(2) 位点特异性重组(site-specific recombination):这种重组的特点是重组发生在特异位点,此位点含有短的同源序列,供位点特异性重组酶识别;重组过程涉及到蛋白的催化的交错切割。最为典型的是 λ 噬菌体在att位点整合到E.coli的基因组中。沙门氏菌鞭毛的相变位中H2基因和rh1基因上游一段995bp序列的倒位。Mu噬菌体中G片断的倒位也涉及短的重复序列,Hin/Gin蛋白的催化以及交错切割也属此类。

(3) 转座重组(transposition recombination):转座的机制依赖DNA的交错剪切和复制,但不依赖于同源序列。转座涉及转座酶,解离酶识别重组分子中的短特异序列,共分为复制型、非复制及保守型三种类型。转座的过程中会形成共合体。两个转座因子之间的重组会引起缺失和倒位。

(4) 模板选择(copy choice)性重组:适于RNA病毒,在这种重组中,聚合酶以一个模板转换到另一个模板来合成RNA,结果新合成的分子将含有两个不同亲本的遗传信息。内含子归巢也属于此种类型。

(5) 特殊重组(special recombination):在哺乳动物的体细胞中,免疫球蛋白成熟的V.J.C(或V.D.JC)区要进行DNA重组,这种重组涉及短的同源序列,以及特殊的剪切、修补、连接和插入碱基。重组由重组激活基编码的RAG 1,2蛋白催化。

(6) 同源特异重组(homologous-special recombination):酵母交配型的转变和以上有的类型都有相似之处,但又有明显的不同。如交配型转变同样涉同源大片断DNA的同源配对,但重组位点仅特异地Z/Y交界区,而且伴随着复制过程;切割时仅受体位点双链被切,这些特点很像复制型转座,但与转座重组的根本区别在于:①位点特异;②在结构上无反向重复,也不形成靶位点的正向重复;③不涉及转座酶和解离酶,因此应将其另列为一个类型,称为同源特异重组。

● 简述同源重组的生物学意义？

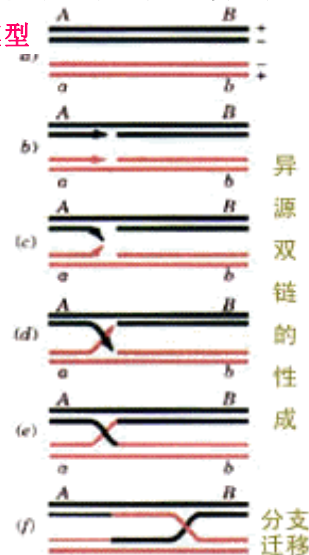
- (1) 损伤修复：能使遭受损害的染色体，得以利用与自身相似，且可使未受伤害的另一条染色体，来进行DNA修复作用。
- (2) 适应环境：使各生物能够更好的适应环境，帮助生存。
- (3) 加速进化：以加速生物由低等到高等、由简单到复杂的进化过程。

● 图示 Holliday 中间体的形成及解决的过程？

1) Holliday 双链侵入模型

其中——

所有模型和切断、链置换、连接等过程。



2) 单链侵入模型（链转移模型）
Meselson-Radding模型

切口

取代

侵入

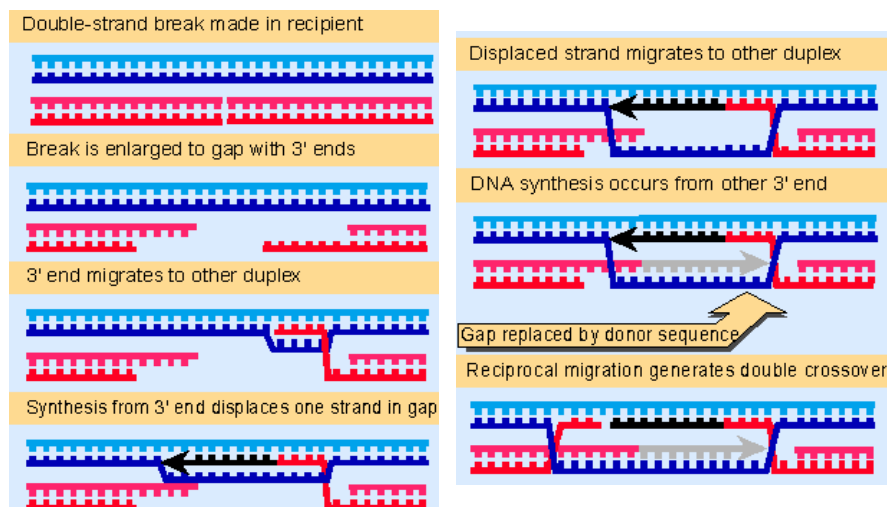
Loop泡的切断

吸收碎链

异构化

分枝迁移

3) 双链断裂修复模型



● 简述重组酶 recA, recBCD, RuvA/B, C 等的功能？

RecA 蛋白是 recA 基因的产物，为含 4 个亚基的四聚体。在重组中，能促使两个同源 DNA 分子的碱基配对，形成杂种分子。反之，两条完整的同源双链螺旋 DNA，即使和 RecA 蛋白一起混合，也不会发生重组反应。RecA 蛋白能特异地识别单链 DNA，并能将之与同源双螺旋中的互补顺序“退火”，同时将另一条链排挤出去，形成所谓 D 环（D-loop）。

RecBCD 由 3 个不同的次单位所组成，分别是 RecB、RecC 与 RecD（其中 RecB 与 RecD 是解旋酶）。在大肠杆菌体内用来起始 DNA 同源重组。RecBCD 同时也作为模式酵素，用来当作实验中所使用的单分子荧光，以了解蛋白质与 DNA 的交互作用。

ruvA 和 ruvB 基因的产物可促进异源双链的形成。RuvA 蛋白可以识别 Holliday 连接体的结构，RuvB 是一个腺苷三磷酸酶并可能提供支链迁移的动力。RuvAB 蛋白复合体可以使支链以 10~20 bp 的速度迁移。ruvC 编码一种能够特异地识别 Holliday 连接体的核酸内切酶，此酶可以在体外切断此种交叉而拆分重组中间体。一个含 4 个碱基的序列提供了 RuvC 拆分 Holliday 连接体的热点。这一序列决定了拆分过程中哪一对 DNA 链被切断，从而决定发生的是补丁型重组还是剪接型重组。