

DNA 复制 (DNA replication)

特征

- 半保留复制 (Semi-conservative)
- 半不连续复制 (Semi-discontinuous replication)
- 模板和引物依赖 (template- and primer-dependent)
- 5' - 3'方向合成
- 4种dNTP

酶

- DNA解旋酶 (DNA helicase) 利用ATP的化学能在复制叉处分离两条亲本DNA链的酶。**DnaB**是原核生物的解旋酶。
- 单链DNA结合蛋白 (SSBP) 保持单链DNA分离状态，促进DNA复制相关的酶进行复制。**GP32**是T4噬菌体中的SSBP。
- 拓扑异构酶 (Topoisomerases) 缓解为复制而展开螺旋产生的拓扑应变和超螺旋。
 - 拓扑异构酶 I (缺口闭合酶)：引入临时的**单链**断裂 (nicks)，参与DNA超螺旋结构的解除，**不需要ATP**，所需能量来自**拓扑应力**的释放。
 - 拓扑异构酶 II (DNA旋转酶)：引入临时的**双链**断裂，能解除超螺旋结构，也能产生超螺旋结构，反应**需要ATP**提供能量。(引入负超螺旋缓解复制过程中DNA展开产生的正超螺旋)
- 引物酶 (Primease) RNA引物合成。RNA合成无需引物，引物RNA的合成仅需以50个DNA碱基左右为模板。引物酶存在于**引物合成体 (引发体)**，只有当引物合成体把 6 种蛋白 (n, n', n'', DnaB, DnaC与DnaI) 和引物酶合在一起组装后，它才具功能。
- DNA聚合酶 (DNA polymeases)
 - DNA 聚合酶 I：polA基因编码，**3'→5' 外切酶活性 (校对)**，**5'→3' 外切酶活性 (切除RNA引物)**；DNA 修复、RNA 引物切除 (切除RNA引物后填补留下的空隙)
 - DNA 聚合酶 II：polB基因编码，**3'→5' 外切酶活性 (校对)**，DNA 修复
 - DNA 聚合酶 III：polC基因编码，**3'→5' 外切酶活性 (校对)**，复制链延伸 (**复制主力**) 形成先导链和滞后链
- DNA连接酶 (DNA ligase) 将DNA片段连接在一起，发生在冈崎片段间切除RNA引物并填补缺口后，相邻冈崎片段的连接。
- 端粒酶 (Telomerase) 与逆转录酶相关，酶结构中含有一个小的RNA模板。端粒酶的RNA模板被用来在**3'端添加端粒重复序列**。随后细胞以这段被延长的 DNA 为模板合成 RNA 引物，完成最后一段滞后链的合成；最终，这段由端粒酶临时延长的序列会和 RNA 引物一起被切除。
 - 端粒**具有称为**端粒重复序列**的重复序列。在大多数体细胞组织中，端粒酶表达极低甚至完全不表达，随着细胞分裂，端粒会缩短。短端粒可能是细胞衰老的信号。

端粒是**真核生物线性染色体末端**的一段特殊DNA-蛋白复合结构，由高度重复的短DNA序列 (比如人类是TTAGGG 重复序列) 和结合其上的端粒结合蛋白共同组成。

- 端粒 (Telomere) 的三大核心作用**
- 保护染色体末端，维持其稳定性**
端粒能避免染色体末端被细胞识别为DNA双链断裂的信号，防止染色体发生末端融合、降解或异常重组，从而保证基因组完整性。细胞需要特设区分端粒和其他DNA断裂末端，避免错误的修复反应。
 - 实现端粒的延长，抵消复制损耗**
由于DNA复制无法从染色体最末端起始，每次细胞分裂近端粒都会自然缩短。端粒需要通过**端粒酶**等机制进行延长，以维持其长度稳定。端粒长度短基因缺陷会导致细胞衰老，从而诱发细胞的持续分裂能力。
 - 促进减数分裂时染色体的重组，保障同源染色体重组的准确性**
在减数分裂过程中，端粒参与引导染色体的运动与定位，帮助同源染色体完成准确配对和遗传重组，这是物种维持多样性和正常生殖过程的重要基础。

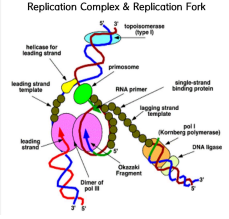
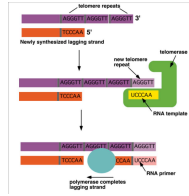
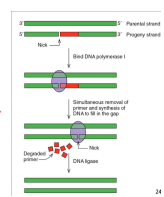
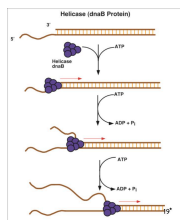
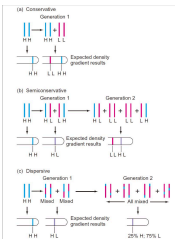
典型原核生物复制叉的总结 (Summary of a typical prokaryotic replication fork)

复制方式

- 双向复制 (Bidirectional replication)**: 两个复制叉向远离原点的相反方向移动。
- 单向复制 (Unidirectional replication)**: 只有一个复制叉，沿单一方向移动完成复制。
- 滚环复制 (Rolling circle replication)**: 双链环状DNA中的一条链保持完整，另一条链在特定位置被切开形成缺口，完整链作为模板，从缺口处开始延伸合成新链。

复制起始 (Origin of replication, ORI) 与复制子

- ORI (复制起始子)**：一段独特的DNA序列，DNA复制在该序列中启动。
- Replicon (复制子)**：从同一个复制起点 (ORI) 出发，被复制的全部DNA 区域，就叫一个复制子。



原核生物的基因结构与转录 (transcription)

操纵子 (Operon)：原核生物中的一个基因表达单位，包含**结构基因**与**调控元件**，核心部分——启动子、顺反子、终止子

顺反子 (Cistron)：一段 DNA 序列，包含**编码单一多肽链**的遗传密码，并作为遗传的基本功能单位

多顺反子 mRNA (Polycistronic message)：携带了多个基因编码信息的 mRNA 分子

启动子 (Promoter)：转录起始前，RNA 聚合酶结合的一段 DNA 序列，通常位于基因转录起始位点的上游

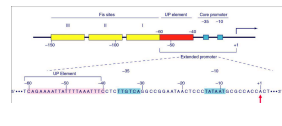
转录起始位点：90% 的基因中，转录起始位点是一个**嘌呤碱基**，且 A+G，起始位点的两侧常是**嘧啶碱基**。典型有 CGT 或 CAT

UP 元件 (UP element)：部分启动子在 -40 ~ -60 bp 处还有一个**额外的增强元件**

序列与共有序列 (保守序列) 越接近，启动子的转录活性 (强度) 越高

例：大肠杆菌 rRNA 基因 (rrnB) 的 P1 启动子是典型的**强启动子**，除了原核生物常见的核心启动子元件，上游 -150 ~ -60 bp 还有三个串联的 Fis 蛋白结合位点，结合**激活蛋白 Fis**，大幅提升转录效率。

两个核心启动子元件：**-10 区 (Pribnow box)**——保守序列为 TATAAT，中心位于转录起始位点上游约 -10 bp 处；**-35 区**——保守序列为 TTTGACA，中心位于转录起始位点上游约 -35 bp 处。



RNA 聚合酶 (RNA polymerase)

组成

- 2 个大亚基：β、β'
- σ 亚基和 α 亚基

金酶 (holoenzyme)：β'、β、α 及 α 组成；**核心酶**：不含 σ 亚基

由于缺少 σ 亚基，**核心酶失去对转录起始位点的特异性识别能力**，只能转录有切口的 DNA 模板，无法转录完整的双链 DNA。

流程

1. RNA 聚合酶全酶在 DNA 上寻找结合位点
2. 形成闭合启动子复合物
3. 紧密结合，解链 DNA 并形成开放启动子复合物
4. 聚合酶停留在启动子上，形成初始转录复合物
5. 启动子清除 (σ 因子释放，核心酶脱离启动子进行转录延伸)

转录起始 (Transcription Initiation)

σ 因子

- α 因子只促进转录起始、不影响延伸，且可被**重复利用**
- 转录起始后，σ 因子与核心酶的**结合变得松散**。荧光共振能量转移技术 (FRET) 表明，σ 因子在**转录起始阶段早期可能并未从核心酶上解离**
- α 因子的两个保守区域参与启动子 -10 区和 -35 区的识别
- β' 亚基促进 σ 因子结合到启动子 -10 区的**非模板链上**
- α 因子单独无法结合 DNA，但与核心酶结合后，会暴露出它的 DNA 结合区域

α 亚基

在 UP 元件的识别中起重要作用，α 亚基突变后，UP 元件对转录不再具有促进作用

转录延伸 (Elongation of transcription)

需要核心酶，DNA 解螺旋依赖拓异构酶

β 亚基

- β 亚基决定了 RNA 聚合酶对**利福平 (Rif 阻断转录起始)**和**利链霉素 (阻断 RNA 链的延伸)**的敏感性
- β 亚基负责确保**磷酸二酯键的形成**

终止子 (Terminator)：一段特定的 DNA 序列，作为信号指示转录过程终止。

转录终止 (Termination of transcription)

- p (rho) 非依赖型终止**：一段反向重复序列，使转录产物的末端能形成**发夹结构**；非模板链上的一串 T 碱基，导致转录产物与模板 DNA 之间形成一连串弱的 U-A 碱基对，削弱转录复合物与模板链的结合力。不依赖 p 因子。
- p 依赖型终止**：需要 p 因子参与；需要反向重复序列，使转录产物末端能形成**发夹结构**；不需要非模板链上的一串 T 碱基序列。

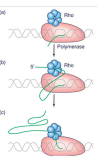
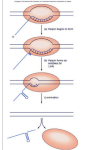
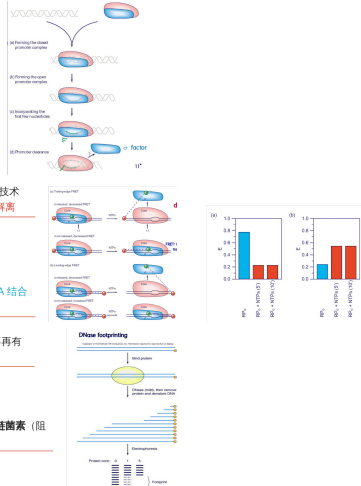
实验发现：p 因子会降低 RNA 合成的净速率；p 因子会缩短 RNA 产物的长度；p 因子将转录产物从 DNA 模板上释放下来

1. 聚合酶在一串弱的 U-A 碱基对处暂停，并且在这些碱基对的上游，一段**发夹结构开始形成**。
2. 发夹结构的形成进一步破坏 RNA-DNA 杂交链的稳定性。不稳定化的机制可能有多种形式——发夹的形成可能会物理性地**将 RNA 拉出聚合酶**；或者它可能直接导致**转录泡坍塌**，将 RNA 从杂交链中释放出来。
3. RNA 产物和聚合酶完全从 DNA 模板上解离，转录终止。

1. 聚合酶在一串弱的 U-A 碱基对处暂停，并且在这些碱基对的上游，一段发夹结构开始形成。

2. 发夹结构的形成进一步破坏 RNA-DNA 杂交链的稳定性。不稳定化的机制可能有多种形式——发夹的形成可能会物理性地将 RNA 拉出聚合酶；或者它可能直接导致转录泡坍塌，将 RNA 从杂交链中释放出来。

3. RNA 产物和聚合酶完全从 DNA 模板上解离，转录终止。



操纵子 (Operon)：原核生物中的一个基因表达单元，包含结构基因和调控元件。每个操纵子都受到**严格调控**，这种调控让细菌的代谢更高效。

操纵区 (操纵序列) (Operator)：原核生物中的一段 DNA 元件，能与特定的阻遏蛋白紧密结合，从而调控相邻基因的表达。

乳糖操纵子 (Lac operon)

三个结构基因

- β-半乳糖苷酶 (lacZ)：将乳糖分解为半乳糖和葡萄糖
- 半乳糖通透酶 (lacY)：将乳糖转运进入细胞内
- 半乳糖苷转乙酰基酶 (lacA)：功能尚未完全明确，参与乳糖的代谢过程

Lac I：编码**乳糖阻遏蛋白 (LacR)**的基因

P (启动子)：乳糖启动子 (RNA 聚合酶的结合位点)

O (操纵区)：乳糖操纵区 (Lac 阻遏蛋白的结合位点)

核心元件——阻遏物和操纵区

负调控机制

- 阻遏物与操纵区结合时，操纵子就处于**关闭状态**；
- 葡萄糖供应充足且乳糖可被利用时，细胞中少量本底表达的乳糖操纵子编码酶，会将乳糖转化为少量**异乳糖**，异乳糖作为诱导剂，通过与阻遏物结合，使**阻遏物从操纵区上解离**，此时操纵子开启表达。

正调控机制

- 分解代谢物激活蛋白 (CAP) 与环腺苷酸 (cAMP) 协同作用，能够促进转录过程。(cAMP 浓度受葡萄糖水平调控，葡萄糖越多，细胞内的 cAMP 浓度越低)

大肠杆菌中负责 L-阿拉伯糖分解代谢的调控系统，核心调控蛋白为 AraC，其依赖阿拉伯糖的存在状态发挥不同作用。

阿拉伯糖操纵子 (Ara operon)

相关元件

- araC 基因
- 操纵序列 araO₂、araO₁
- araC 基因的启动子
- 阿拉伯糖诱导区域 araI (包含 I_x、I₂ 两个结合位点)
- 结构基因 araBAD 的启动子

机制

- 无阿拉伯糖时**，**AraC 蛋白作为阻遏蛋白**，结合到 O₂ 和 I₁ 上，使 210bp 的 DNA 形成环结构，抑制转录。
- 有阿拉伯糖时**，阿拉伯糖作为**诱导物**，与 AraC 蛋白结合，使其构象发生改变，结合到 araI 区的 I₁ 和 I₂ 两个位点；**正向调控**；同时 CAP 蛋白 (结合 cAMP) 也可以结合到 CAP 位点，与 AraC 协同作用，招募 RNA 聚合酶到 araP_{BAD} 启动子上。

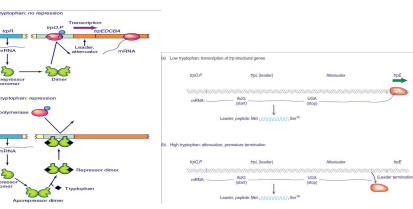
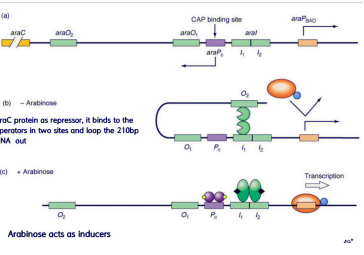
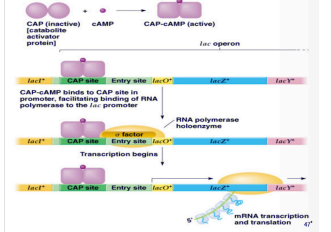
色氨酸操纵子 (Trp operon)

相关元件

- 色氨酸操纵子的调节基因 trpR
- 色氨酸操纵子的操纵序列 (O) 和启动子 (P)
- 色氨酸合成相关的结构基因 trpEDCBA

机制

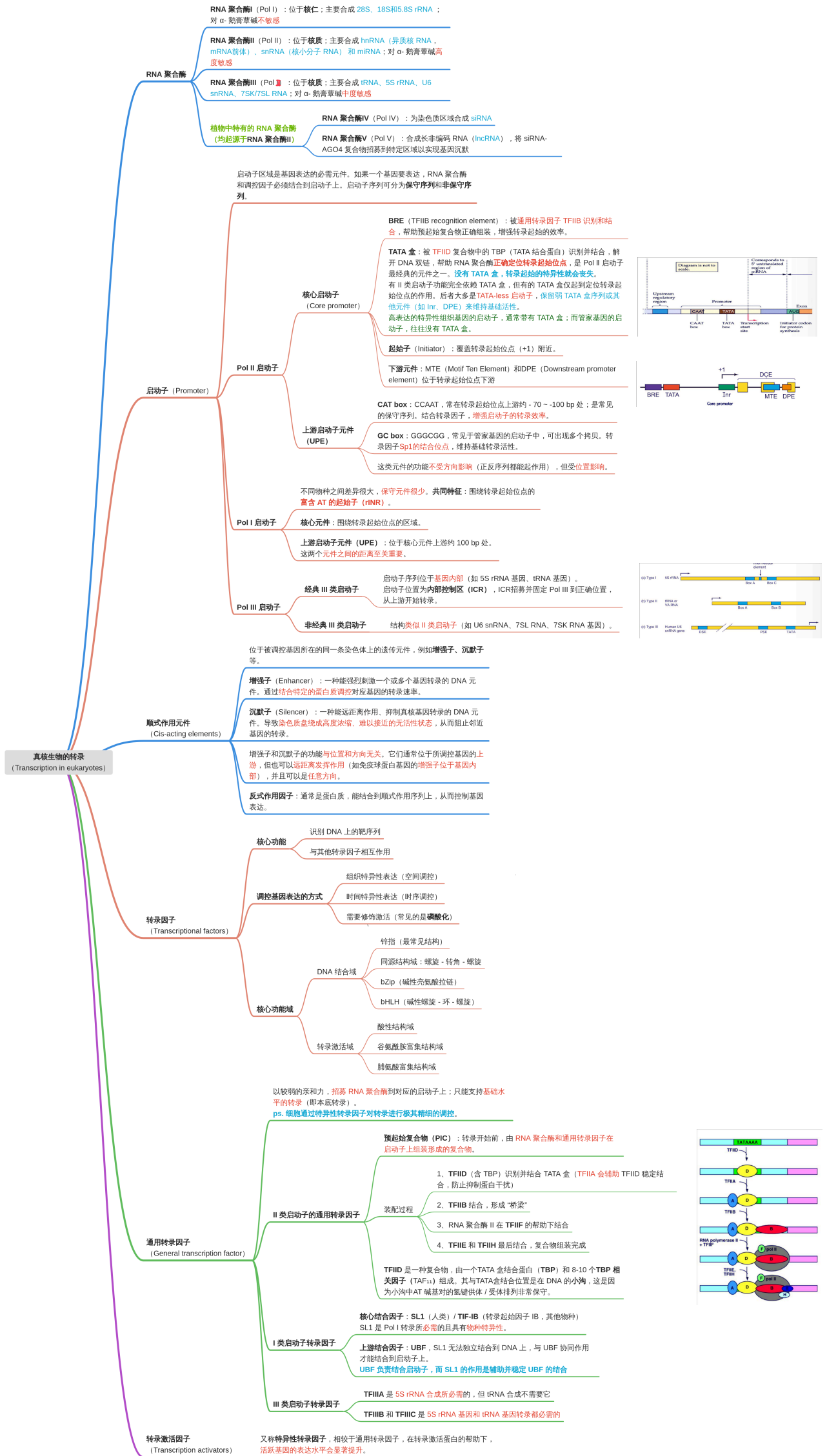
- 低色氨酸时**，trpR 基因转录翻译出**辅阻遏蛋白**单体，并组装成二聚体，但色氨酸不足，**辅阻遏蛋白二聚体无活性**，无法结合操纵序列 trpO。RNA 聚合酶可以顺利结合启动子 trpP，**启动转录**。
- 高色氨酸时**，色氨酸作为**辅阻遏物**，与辅阻遏蛋白二聚体结合，使其构象发生改变，成为有**活性的阻遏蛋白二聚体**，特异性结合到操纵序列 trpO 上，**阻断 RNA 聚合酶与启动子的结合，转录被抑制**。
- 色氨酸操纵子的**衰减调控**：衰减子 (Attenuator) 是位于色氨酸操纵子前导序列中的一段 DNA 序列，它编码了一段**含有两个连续色氨酸密码子**的前导肽。这段肽能通过不同的 RNA 二级结构，决定转录是继续前进，还是提前终止。



低 Trp 时，核糖体遇到前导肽中的两个连续色氨酸密码子时，无法继续翻译，停在了前导序列的 I₁ 区上，2 区及刚转录出来的 3 区配对，形成**2-3 茎环 (没有多聚 U 尾，不是终止信号)**，因此 RNA 聚合酶通过衰减子区域，**继续转录**。

高 Trp 时，核糖体可以通过 I₁、I₂ 和 2 区部分碱基配对后，3 区只能和刚转录出来的 4 区配对，形成**3-4 茎环**，其后直接连有多聚 U 尾，是标准的**Rho 非依赖型转录终止信号**。

前导序列 DNA 有 4 段互补茎环序列：1. 2. 3. 4.



内含子 (intron)：基因转录部分中的一个区域，该区域在转录后会被切除。
外显子 (Exon)：基因中的一个区域，最终会出现在该基因的成熟转录本中。

RNA 剪接：从未成熟 RNA 中切除内含子，并将外显子连接在一起以形成最终产物的过程。
对 mRNA：先是合成初级转录产物（仍含有从基因中复制来的内含子的 mRNA 前体），称为 **hmRNA**（异质性核 RNA），再是剪接后 mRNA 的成熟。

剪接信号：内含子的前两个碱基为 **GU**，内含子的最后两个碱基为 **AG**。

- 剪接机制**
- 分支点 A 的 2'-OH 攻击 5' 外显子内含子边界的磷酸二酯键，将 5' 外显子作为离去基团推出，并形成中间的**套索结构**。
 - 5' 外显子的 3'-OH 攻击 3' 内含子外显子边界的磷酸二酯键，将内含子（套索形式）置换出来，并将两个外显子连接在一起。

剪接体是由 **RNA 和多种组成颗粒**，在真核细胞的细胞核内负责执行 mRNA 的剪接。

结构：snRNA 和蛋白质
SnRNPs：核小分子核糖核蛋白
Smurps：snRNA（识别关键的剪接信号）+ 蛋白质的复合物
剪接涉及多种 **snRNP**：U1、U2、U4、U5 和 U6。

剪接体 (Spliceosome)

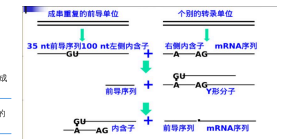
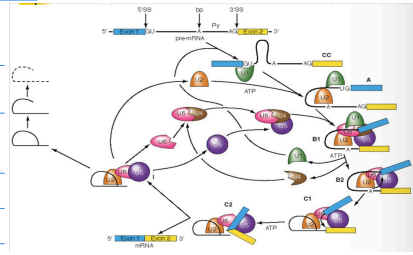
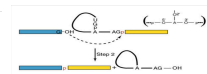
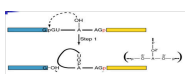
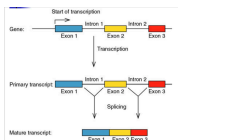
U1 snRNP：U1 snRNA 的序列与 5' 剪接位点和 3' 剪接位点的共有序列互补，U1 snRNA 与这些剪接位点进行碱基配对。其中，U1 snRNA 与剪接体 mRNA 的 5' 剪接位点之间的碱基配对是剪接所必需的，但不是充分条件。

U6 snRNP：通过 U6 snRNA 的碱基配对与内含子的 5' 端结合，这种碱基配对发生在套索中间体形成之前，该结合对剪接过程是必不可少的。在剪接过程中，U6 还与 U2 结合。

U2 snRNP：U2 snRNA 与剪接分支点的保守序列进行碱基配对，U2 还与 U6 形成重要的碱基配对，形成一个称为**催化位点**（帮助这些 snRNP 正确定向以进行剪接）的区域以及**催化位点**（在哺乳动物中，U2 的 5' 端与 U6 的 2' 端配对）。

U5 snRNP：与一个外显子的最后一个核苷酸和下一个外显子的第一个核苷酸结合。这推测是为了将两个外显子对齐，以便进行剪接。

U4 snRNP：与 U6 snRNA 形成碱基配对，一直结合 U6，直到剪接反应需要 U6 时才解离。



可变剪接 (选择性剪接, Alternative splicing)：同一个前体 RNA (pre-mRNA) 可以通过不同的方式进行剪接。如通过可变剪接，抗体分子有不同形式。

反式剪接 (Trans-splicing)

外显子根本不属于一个基因，甚至可能不在同一条染色体上的剪接。

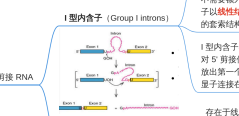
平内含子内的分支点 A 攻击前导外显子与其平内含子之间的连接处。形成一个 **Y 型内含子-外显子中间体**。

前导外显子攻击分支内含子与编码外显子之间的**剪接位点**，产生剪接后的成熟 mRNA + Y 型内含子。

自剪接 RNA (核酶 RNA, Self splicing RNAs)

部分 RNA 无需剪接体或其他蛋白质的辅助，即可完成自剪接。

包含内含子的自剪接 RNA



原生动物的肌肉中**核糖体 RNA** 前体的加工过程。

不同于分支点 A（剪接体 / II 型内含子），I 型内含子以**游离核苷酸**（GMP/GDP/GTP）作为起始攻击基团，作用于 5' 外显子 / 内含子边界。不需要额外能量，仅利用自身的 3'-OH 基团引发亲核攻击。被切除的内含子以**线性结构释放**，而非 II 型自剪接内含子和核糖体 mRNA 剪接中常见的套索结构。

I 型内含子在体外无需蛋白质辅助即可被切除。**第一步**中，由一个鸟苷酸对 5' 剪接位点发起攻击启动，该鸟苷酸被连接到内含子的 5' 端，同时释放出第一个外显子；**第二步**中，第一个外显子攻击 3' 剪接位点，将两个外显子连接在一起，并释放出线性的内含子。

存在于线粒体和叶绿体的基因中。

剪接机制与**反式剪接** mRNA 剪接（**剪接体剪接**）非常相似，但是**无需蛋白质参与，也不需要额外的能量输入**。本质上磷酸二酯键总数不变，位置发生变化转移。

IRNA 剪接

上述剪接反应分两步：1. 内含子中的核糖核苷酸或游核核苷酸对 5' 剪接位点发起攻击；2. 第一个外显子对第二个外显子发起攻击，将两者连接在一起。

IRNA 的剪接也包含两步：rRNA 内切核酸酶切除内含子，RNA 连接酶将两个半 rRNA 片段连接在一起。

转录后加工 (Post-transcriptional events)

加帽 (Capping)

一个**甲基化的鸟苷酸**，通过 **5'-5' 三磷酸连接键**，连接到 mRNA 的 5' 端。该过程发生在**细胞核内**。

类型

- Cap 0：7-甲基鸟苷酸 (m⁷G)，无 2'-O-甲基化修饰
- Cap 1：7-甲基鸟苷酸，且第一个碱基的核糖上存在 2'-O-甲基化修饰
- Cap 2：7-甲基鸟苷酸，且第一个和第二个碱基的核糖上均存在 2'-O-甲基化修饰

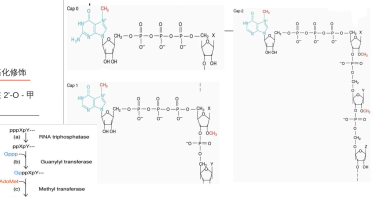
步骤

- 前体 mRNA 5' 端碱基的 γ -磷酸基团被去除
- 鸟苷酸转移酶添加 GTP 的 **GMP 部分**，形成**三磷酸连接键**
- 帽子结构中的鸟嘌呤碱基在 **N7 位**发生**甲基化**
- 其他甲基转移酶为后续核苷酸的核糖 2'-OH 添加甲基基团

甲基供体为：**S-腺苷甲硫氨酸**

功能

- 保护 RNA 不被降解
- 增强 mRNA 的可翻译性，提高翻译效率
- 促进 mRNA 从细胞核运输到细胞质
- 保证前体 mRNA 的正确剪接



加尾 (Polyadenylation)

多聚腺苷酸化：向 RNA 分子的 3' 端添加多聚腺苷酸 (poly(A)) 尾巴的过程。发生在**转录后阶段**，由 **poly(A) 聚合酶** 催化完成。长度约为 250 个核苷酸。**有保守序列 (识别信号)**：**AAUAAA**

步骤

- 对转录本进行剪切，切割位点位于 RNA 编码区的末端
- poly(A) 聚合酶向 mRNA 的 3' 端添加多聚腺苷酸尾巴
- 多余 RNA 的降解

功能

- 保护 mRNA
- 增强 mRNA 的可翻译性

rRNA 与 tRNA

RNA 加工 (RNA Processing)：从 RNA 前体上**切除多余片段的**过程。与剪接类似，就是去除不必要的 RNA 片段，但核心差异是剪接会将不同的 RNA 片段拼接在一起。

rRNA

- 真核生物 rRNA 加工**：涉及 **snRNA**（小核仁 RNA）和 **snRNP**（小核仁核糖核蛋白）；E1、E2 (snRNPs)：负责 **18S rRNA** 的成熟；E3 (snRNP)：负责 **5.8S rRNA** 上游的切割
- 原核生物 rRNA 加工**：涉及 **RNase III** 和 **RNase E** 两种核糖酶。原核生物的 rRNA 前体同时包含 **snRNA** 以及全部三种 rRNA，每个转录单位为 16S、23S、5S rRNA 以及 **snRNA** 基因所组成。
- rRNA 前体基因的转录**：rRNA 基因之间被称为**非转录间隔区 (ITS)** 的序列间隔。ITS 与转录间隔区不同，它们会留在前体 rRNA 加工成为成熟 rRNA 的过程中被切除。
- 总结**
 - rRNA 在真核生物的**核仁**中合成，最初以**前体形式**存在，必须经过**加工**才能释放出**成熟的 rRNA**
 - 所有真核生物中，前体 RNA 中 rRNA 的**顺序都是 18S — 5.8S — 28S**
 - snRNA 在这些加工步骤中发挥着至关重要的作用

tRNA

- 切除多聚腺苷酸 (RNase III)**
- 形成成熟的 5' 末端** (RNase P)：RNase P 包含两个亚基，其中一个亚基由 RNA 构成
- 形成成熟的 3' 末端**：RNase D、RNase BN、RNase T 等

RNA 编辑

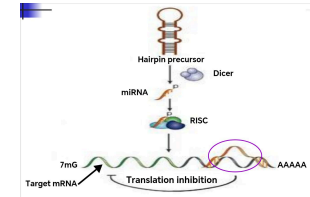
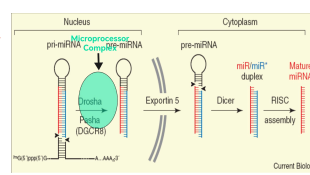
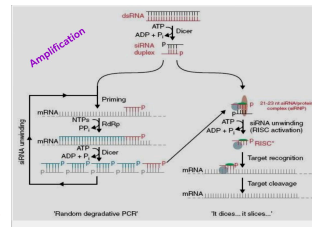
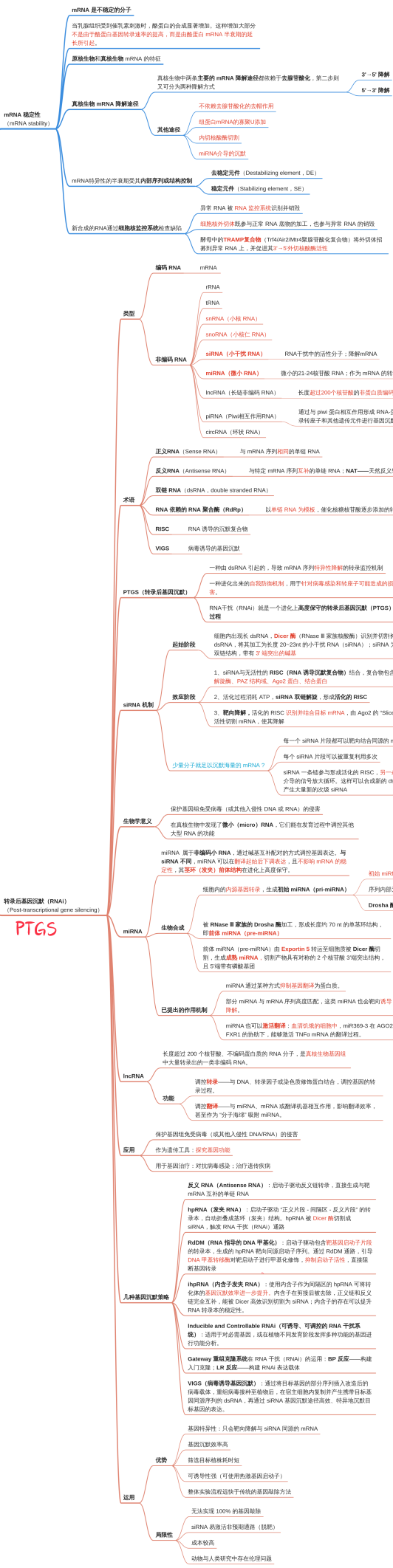
在脊椎动物中，成熟 mRNA 的成熟过程中会添加或删除一些 U，这被称为 RNA 编辑。

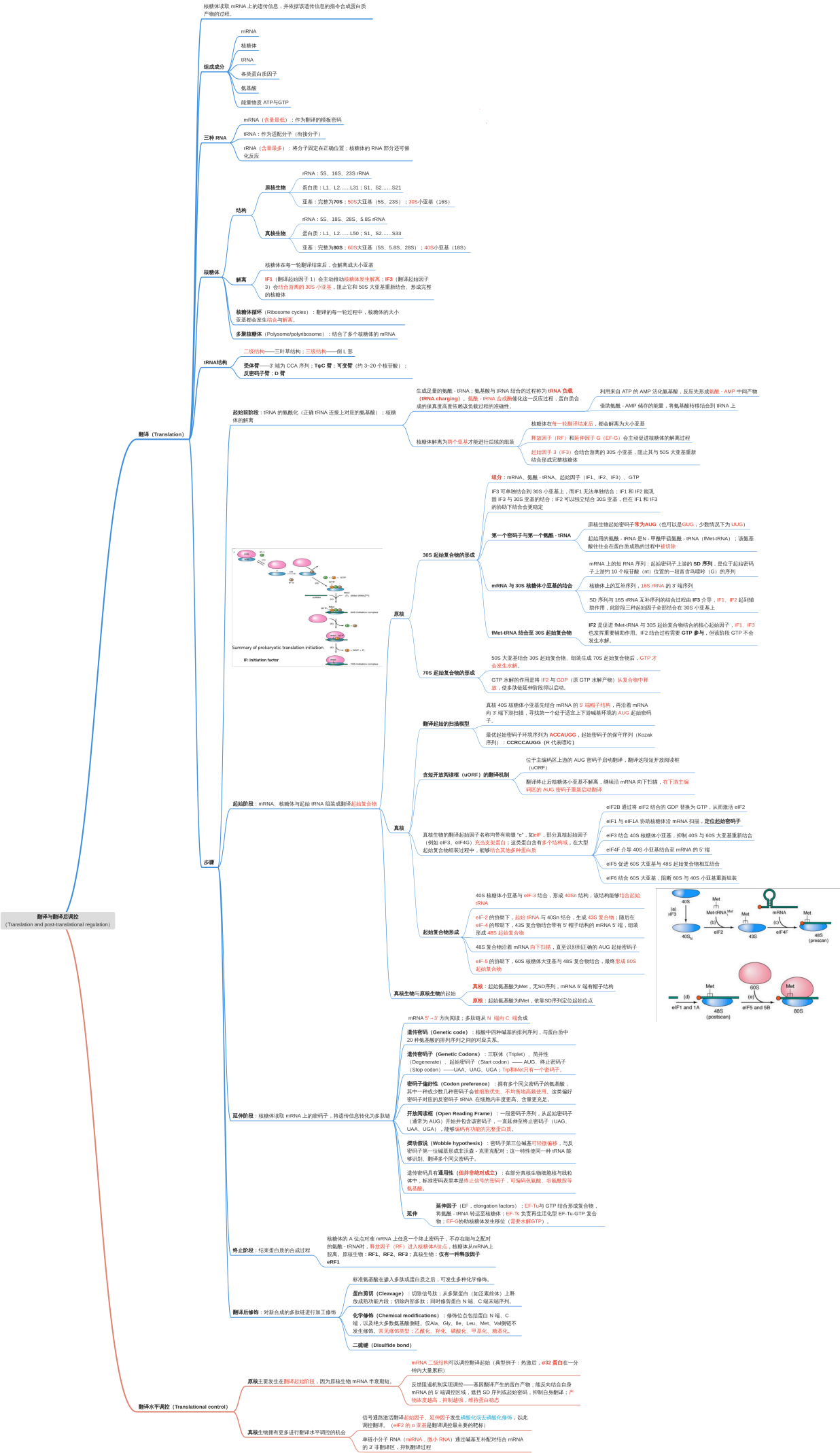
流程

- gRNA 与前体 mRNA 的一个区域杂交
- gRNA-I 的其余部分指导前体 mRNA 一部分的编辑。由于插入 UMP，前体 mRNA 长度增加。
- 新的 gRNA-II 通过与前体 mRNA 新编辑区域的 5' 端杂交，取代 gRNA-I。
- gRNA-II 指导前体 mRNA 下一部分的编辑。
- 用额外的 gRNA 重复上述步骤，直到 RNA 被完全编辑。

基因表达的转录后调控
(Posttranscriptional Control of Gene Expression)

PTGS





功能基因组学 (Functional Genomics)

