

# 选择性多聚腺苷酸化的生物学效应及其调控机制研究进展

徐海冬<sup>1,2,3</sup>, 宁博林<sup>1,2,3</sup>, 牟芳<sup>1,2,3</sup>, 李辉<sup>1,2,3</sup>, 王宁<sup>1,2,3</sup>

1. 农业农村部鸡遗传育种重点实验室, 哈尔滨 150030
2. 黑龙江省普通高等学校动物遗传育种与繁殖重点实验室, 哈尔滨 150030
3. 东北农业大学动物科学技术学院, 哈尔滨 150030

**摘要:** 真核生物基因的前体 mRNA (pre-mRNA) 及一些 lncRNA 在成熟过程中其 3'端会发生剪切和多聚腺苷酸化反应(cleavage and polyadenylation, C/P), C/P 的发生需要多聚腺苷酸化信号(polyadenylation signal, PAS)的存在。选择性多聚腺苷酸化(alternative cleavage and polyadenylation, APA)是指具有多个 PAS 的基因, 在其 mRNA 3'端成熟过程中, 由于选择不同的 PAS, 导致产生出多个 3'UTR 长度和序列组成不同的转录异构体。3'UTR 长度和序列的不同会影响 mRNA 的稳定性、翻译效率、运输和细胞定位等, 因此 APA 是真核生物的一个重要转录后调控方式。近年来, 对大量动物、植物及酵母的基因组测序分析发现, APA 在真核生物广泛存在, 针对 APA 的生物学效应和调控机制开展了一系列研究。目前已鉴定出许多 APA 调控的顺式调控元件和反式作用因子。本文重点介绍了 APA 生物学效应和调控机制的最新研究进展, 并探讨了未来 APA 调控的研究方向。

**关键词:** 基因表达调控; 选择性多聚腺苷酸化; 多聚腺苷酸化信号; 顺式调控元件; 反式作用因子

## Advances of functional consequences and regulation mechanisms of alternative cleavage and polyadenylation

Haidong Xu<sup>1,2,3</sup>, Bolin Ning<sup>1,2,3</sup>, Fang Mu<sup>1,2,3</sup>, Hui Li<sup>1,2,3</sup>, Ning Wang<sup>1,2,3</sup>

1. Key Laboratory of Chicken Genetics and Breeding, Ministry of Agriculture and Rural Affairs, Harbin 150030, China
2. Key Laboratory of Animal Genetics, Breeding and Reproduction, Education Department of Heilongjiang Province, Harbin 150030, China
3. College of Animal Science and Technology, Northeast Agricultural University, Harbin 150030, China

**Abstract:** During the maturation of pre-mRNAs and some lncRNAs, their 3'ends are cleaved and polyadenylated. The cleavage and polyadenylation (C/P) require the presence of a polyadenylation signal (PAS) at the 3'end. Most eukaryotic genes have multiple PASs, resulting in alternative cleavage and polyadenylation (APA). APA leads to transcript

收稿日期: 2020-07-07; 修回日期: 2020-10-19

基金项目: 国家自然科学基金项目(编号: 31572392, 3187131154)和农业部产业体系项目(编号: CARS-41)资助[Supported by the National Natural Science Foundation of China (Nos. 31572392, 3187131154), and the China Agriculture Research System (No. CARS-41)]

作者简介: 徐海冬, 在读博士研究生, 专业方向: 动物遗传育种与繁殖。E-mail: 1689512382@qq.com

通讯作者: 王宁, 博士, 教授, 博士生导师, 研究方向: 动物遗传育种与繁殖。E-mail: wangning@neau.edu.cn

DOI: 10.16288/j.ycz.20-200

网络出版时间: 2020/12/14 14:04:37

URI: <https://kns.cnki.net/kcms/detail/11.1913.R.20201211.1654.003.html>

isoforms with different coding potentials and/or variable 3'UTRs. The 3'UTR affects mRNA stability, translation, transportation, and cellular localization. Therefore, APA is an important mechanism of posttranscriptional gene regulation in eukaryotes. In recent years, whole genome sequencing of animals, plants and yeast has revealed that APA is pervasive in eukaryotes, and the functional consequences and regulation of APA have been studied. To date, many *cis*-acting regulatory elements and *trans*-acting factors for APA regulation have been identified. In this review, we summarize the recent advances in the functional consequences and regulation of APA and discuss the future directions, aiming to provide clues and references for future APA study.

**Keywords:** gene expression regulation; alternative cleavage and polyadenylation; polyadenylation signal; *cis*-acting regulatory element; *trans*-acting factor

真核生物的 mRNA 前体(pre-mRNA)和一些长链非编码 RNA 前体(pre-lncRNA)在成熟过程中其 3'端会被核酸内切酶剪切(cleavage), 然后由模板非依赖的 RNA 聚合酶—poly(A)聚合酶(PAP)在 3'端合成一个多聚腺苷酸尾巴(polyadenylation)。真核生物 mRNA 前体 3'端的剪切和多聚腺苷酸化反应(cleavage and polyadenylation, C/P)需要多聚腺苷酸化信号(polyadenylation signal, PAS), PAS 通常位于剪切位点上游 10~30 个核苷酸处。选择性多聚腺苷酸化(alternative cleavage and polyadenylation, APA)是指含有多个 PAS 的基因, 在 mRNA 3'端成熟过程中, 不同 PAS 的选择利用将导致一个基因产生多个转录异构体, 这些转录异构体具有长短不一的 3'末端<sup>[1]</sup>。近年来, 由于高通量测序技术的发展和研究的深入, 人们发现 APA 在真核生物中普遍存在, 例如人类超过 70%的基因都存在 APA 现象<sup>[2]</sup>。APA 在不同类型的基因中分布不同, 其中, 在蛋白编码基因中分布最为广泛, 其次为 lncRNA 和假基因, 而在小 RNA (small RNA)和微小 RNA (microRNA)中的分布最少<sup>[3]</sup>。大多数 APA 仅会导致 mRNA 含有不同长度的 3'UTR, 但有少数 APA 还会改变 mRNA 编码区序列, 导致编码不同的蛋白。3'UTR 存在大量基因表达的顺式调控元件, 这些顺式调控元件广泛参与 mRNA 的稳定性、翻译、运输和细胞定位<sup>[1]</sup>。APA 使得同一基因的不同转录本具有不同的顺式调控元件, 因而导致不同转录本的表达不同。目前已知 APA 是真核生物基因表达调控的重要调控机制之一, 它在许多生物学过程中发挥重要的调控作用; 其调控异常会导致癌症等多种疾病发生<sup>[4~7]</sup>。近年来, 由于 APA 功能的多样性及其调控的重要性, 针对 APA 的

研究引起了人们广泛关注。本文综述了 APA 调控研究的最新进展, 旨在为深入解析 APA 的调控机制和生物学效应提供参考。

## 1 C/P 反应的顺式调控元件和反式作用因子

C/P 反应是 mRNA 前体成熟的一个重要步骤, 关系到 mRNA 的核外运输和胞质翻译过程。C/P 反应需要位于 mRNA 前体的顺式元件及许多反式作用因子的参与。目前已鉴定出许多 C/P 反应的顺式调控元件、反式作用因子以及多种核酸酶(图 1)<sup>[1,8~10]</sup>。

### 1.1 C/P 反应的顺式调控元件

PAS 是 mRNA 前体进行 C/P 反应的信号, 它位于 mRNA 前体的 3'端。根据 PAS 的序列特征和功能, 可将其组合并归类为核心元件和辅助元件两部分。其中, 核心元件包括距剪切位点上游约 20~22 个核苷酸处的 A-rich 六聚体、剪切位点下游约 10~30 个核苷酸处的 U/GU-rich 元件和临近剪切位点的 CA (或 UA)序列<sup>[9]</sup>。AAUAAA 是典型的 A-rich 六聚体序列, 其单碱基突变所形成的非经典序列仍具有 AAUAAA 的部分活性<sup>[5]</sup>。下游 U/GU-rich 元件的单个核苷酸突变体同样也具有一定的活性。CA (或 UA)是临近剪切位点前常见的两个核苷酸, 它们控制剪切效率。辅助元件包括剪切位点下游的 G-rich 元件和剪切位点上游的 UGUA 元件<sup>[9]</sup>。其中 G-rich 元件影响剪切效率; UGUA 元件促进 C/P 反应<sup>[10]</sup>。此外, 剪切位点上游的 U-rich 元件也参与剪切调控。U-rich

元件有两类：一类是位于 A-rich 六聚体和剪切位点之间的 U-rich 元件；另一类是位于 A-rich 六聚体上游的 U-rich 元件，这两种 U-rich 元件都可以与 CPSF 复合体互作，促进 C/P 反应<sup>[10]</sup>。

## 1.2 C/P 反应的反式作用因子

目前已知有 4 种蛋白复合体和多个蛋白质单体参与 C/P 反应。参与 C/P 反应的 4 种蛋白质复合体分别是剪切多聚腺苷酸化特异因子(cleavage and polyadenylation specificity factor, CPSF)复合体、剪切激活因子(cleavage stimulation factor, CSTF)复合体、剪切因子复合体 I(cleavage factor I m, CFIm)和剪切因子复合体 II(cleavage factor II m, CFII m)<sup>[1,9,11]</sup>。

CPSF 复合体主要由 CPSF160 (CPSF1)、CPSF100 (CPSF2)、CPSF73 (CPSF3)、CPSF30 (CPSF4)、PAP 交互因子(factor interacting with PAP, FIP1)和 WD 重复域 33 (WD repeat domain 33, WDR33) 6 种蛋白构成<sup>[9,11]</sup>。CPSF160 通过  $\beta$  螺旋结构域特异性结合 AAUAAA<sup>[12]</sup>；CPSF30 的 N 端含有 5 个保守的锌指结构域，其中第 2 个和第 3 个锌指结构域能与 PAS 上游的 U-rich 特异性结合；CPSF73 是一种锌依赖性核酸内切酶，其金属- $\beta$ -内酰胺酶结构域和  $\beta$ -CASP 结构域可精准识别 CA(或 UA)位点并对其临近的位

点进行精准剪切<sup>[9]</sup>。FIP1 的 C 端精氨酸富集区可与 PAS 下游的 U-rich 序列结合，并同时招募 PAP<sup>[9,12]</sup>。WDR33 具有富含  $\beta$  螺旋的 WD40 保守结构域，能与 AAUAAA 直接结合，在蛋白-蛋白和蛋白-RNA 互作中发挥重要功能<sup>[12,13]</sup>。

CSTF 复合体是由 CSTF77 (CSTF1)、CSTF64 (CSTF2)和 CSTF50 (CSTF3) 3 种蛋白质亚基组成的六聚体复合体<sup>[1,11]</sup>。CSTF64 和 CSTF64 $\tau$  (CSTF64 的旁系同源体)的 N 端具有 U-/GU-rich 的识别基序(RNA recognition motif, RRM)<sup>[9]</sup>。CSTF77 蛋白 C 端的脯氨酸富集结构域(monkey-tail, MT)与 RRM 附近的铰链区(hinge 结构域)相互作用，促进 CSTF64 与 RNA 的结合<sup>[9]</sup>；CSTF77 的 N 端具有一个螺旋 HAT 结构域，可分别结合 CSTF64 和 CSTF50，可作为支架蛋白为其他反式作用因子提供结合平台。CSTF50 的保守性较低，仅存在于少数高等动物中，可通过 WD40 结构域与 CSTF77 的 MT 结构域相互作用，也与 RNAP II 的羧基末端域(carboxy-terminal domain, CTD)互作，增强其他反式作用因子的活性<sup>[1]</sup>。

CFIm 复合体主要包括 CFIm25、CFIm59(或 CFIm68)<sup>[9,11]</sup>。CFIm25 具有 RRM 结构域，可特异性识别 PAS 上游的 UGUA 元件，进而招募 CFIm59/CFIm68，形成 CFIm 复合体<sup>[9]</sup>。CFIm59 的 RRM 结

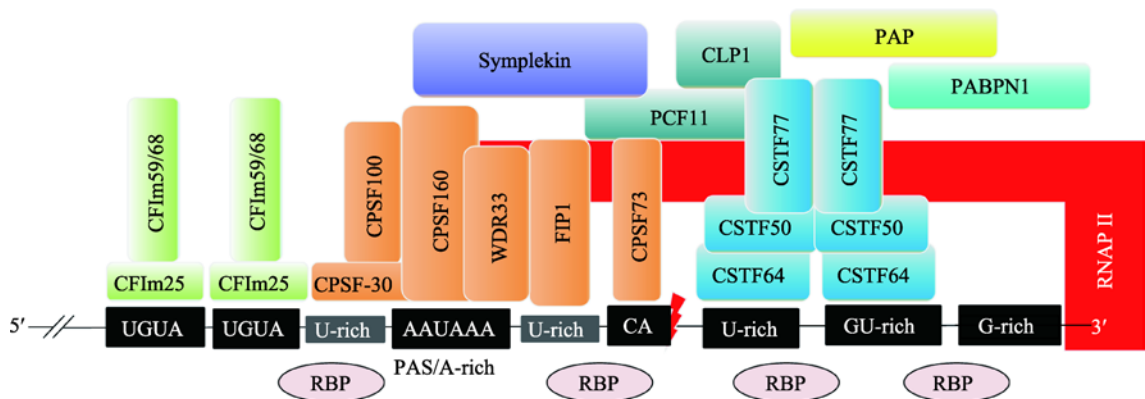


图 1 C/P 反应的顺式调控元件和反式作用因子

Fig. 1 The *cis*-acting regulatory elements and *trans*-acting factors involved in cleavage and polyadenylation (C/P)

CPSF: 剪切多聚腺苷酸化特异因子(cleavage and polyadenylation specificity factor); CSTF: 剪切激活因子(cleavage stimulation factor); CFIm: 剪切因子复合体 I (cleavage factor I m); RBP: RNA 结合蛋白(RNA binding protein); RNAP II: RNA 聚合酶 II(RNA polymerase II); PAP: poly(A)聚合酶(poly(A) polymerase); PABPN1: 核 poly(A)结合蛋白(nuclear poly(A)-binding protein 1); WDR33: WD 重复域 33 (WD repeat domain 33); FIP1: PAP 交互因子(factor interacting with PAP); CLP1: C/P 因子亚基 1 (cleavage factor polyribonucleotide kinase subunit 1); PCF11: C/P 因子亚基(cleavage and polyadenylation factor subunit); PAS: 多聚腺苷酸化信号(polyadenylation signal)。根据参考文献[1, 8~10]总结绘制。

构域与 CFIm25 二聚体结合, 增强 CFIm25 的 RNA 结合活性, 同时也可促进 RNA 环化<sup>[10]</sup>。CFIm 复合物主要包括 PCF11 (PCF11 cleavage and polyadenylation factor subunit) 和 CLP1 (CLP1 cleavage factor polyribonucleotide kinase subunit 1) 两种蛋白质, PCF11 结合在 PAS 的下游, 在 RNAP II 转录终止中发挥作用, CLP1 可作为连接 CPSF 复合物和 CFIm 复合物的衔接蛋白<sup>[11]</sup>。

参与 C/P 反应的蛋白质单体包括偶对蛋白 (symplekin)、RNA 聚合酶 II (RNA polymerase II, RNAP II)、poly(A) 聚合酶 (poly(A) polymerase, PAP) 和核 poly(A) 结合蛋白 (nuclear poly(A)-binding protein 1, PABPN1) 等。RNAP II 的 C 端结构域 (CTD) 能招募多个蛋白复合物, 为 C/P 反应提供酶促反应的平台。

另外, 有些反式作用因子蛋白本身不能直接结合 RNA, 但它们可以作为支架蛋白结合上述各复合物, 从而发挥调控 C/P 的作用, 例如, Symplekin 作为连接 CPSF 和 CSTF 复合物的连接蛋白。

### 1.3 转录与 C/P 反应

真核生物基因的转录延伸、RNA 前体的选择性拼接以及 C/P 反应相偶联。RNAP II 在转录延伸、选择性拼接以及 C/P 反应中都发挥重要作用。RNAP II 的 N 端控制 RNA 的合成, RNAP II 的 CTD 结构域可募集 C/P 反应所需的多种反式作用因子, 为 mRNA 前体的 C/P 反应提供平台。RNAP II 的 CTD 结构域与 C/P 反应的反式作用因子的结合是一个动态过程, 其中 CPSF 复合物中的 WDR33 和 CPSF30 在其他蛋白质的参与下特异性识别 AAUAAA 保守基序<sup>[9]</sup>。同时 CFIm25 和 CSTF64 分别与 PAS 上游 UGUA 元件和下游 U/GU-rich 元件等保守基序特异性结合, 并在 CSTF77、Symplekin、CFIm68/CFIm59 和 FIP1 等蛋白的作用下, CSTF、CPSF 和 CFIm 等形成一个复合体, 随后 CPSF73 对 PAS 下游剪切位点进行精准剪切<sup>[9]</sup>。CPSF 和 FIP1 互作可招募 PAP, 招募的 PAP 会在剪切后的前体 mRNA 的 3' 端缓慢合成 10~15 个腺苷酸残基。此后, PABPN1 会结合到短的 poly(A) 序列上, 同时锚定 PAP, 快速合成 poly(A), 形成长约 200 个腺苷酸的 poly(A) 尾巴<sup>[1]</sup>。转录延伸与 C/P 反应偶联保证了 RNA 的精细化加工,

这是一种快速精确的调控方式, 以最小的耗能方式为 RNA 转录加工提供保障。

### 1.4 胞质多聚腺苷酸化

除了上述细胞核内多聚腺苷酸化, 还存在胞质多聚腺苷酸化 (cytoplasmic polyadenylation)。与核内多聚腺苷酸化不同, 胞质多聚腺苷酸化发生于胞质内成熟的 mRNA, 它不与转录相偶联, 也不存在剪切反应。这类 mRNA 没有翻译活性, 它们的 poly(A) 尾巴通常都很短, 大约有 20 多个核苷酸。胞质多聚腺苷酸化能延长这类 mRNA 的 poly(A) 尾巴, 从而激活其翻译活性。胞质多聚腺苷酸化是调控 mRNA 翻译的一个重要机制, 多见于卵子发生和早期发育胚胎。例如, 在卵母细胞受精后, 卵母细胞质存留的无翻译活性的 mRNA 因发生胞质多聚腺苷酸化而激活其活性, 表达出合子基因组表达所需的酶和转录因子<sup>[14,15]</sup>。在胚胎期晶状体纤维分化过程中, 由于晶状体蛋白相关基因的 mRNA 发生胞质多聚腺苷酸化, 从而提高了这些 mRNA 的翻译效率<sup>[16]</sup>。

与核内多聚腺苷酸化相同, 胞质多聚腺苷酸化也同样受到许多顺式调控元件和反式作用因子的调控, 其中, 顺式调控元件包括胞质多聚腺苷酸化元件 (cytoplasmic polyadenylation element, CPE)、CPE 上游的 Pumilio 结合元件 (pumilio binding element, PBE) 以及 CPE 下游的 PAS<sup>[17,18]</sup>; 已知的反式作用因子有 CPE 结合蛋白 (CPE binding proteins, CPEBs)、Pumilio 蛋白以及 CPSF160/100/30<sup>[19]</sup>。反式作用因子 CPEBs 能结合 CPE 元件, Pumilio 能结合 PBE 元件, 而 CPSF160/100/30 能结合 PAS 元件。非磷酸化的 CPEB1 结合 Pumilio, 招募脱腺苷化酶和翻译起始抑制因子, 进而抑制翻译; 但磷酸化的 CPEB1 会与脱腺苷化酶和翻译起始抑制因子解离, 转而与 CPSF 互作, 从而招募胞质 poly(A) 聚合酶 PAPD4 (GLD-2), 快速合成 80~250 个核苷酸的 poly(A) 尾巴<sup>[17]</sup>。胞质 poly(A) 结合蛋白 (cytoplasmic poly(A)-binding protein 1, PABPC1) 结合 poly(A) 尾巴, 它同时与 mRNA 5 端的帽结构复合物亚基 eIF4G 互作, 形成闭环复合体 (closed loop complex), 从而促进翻译<sup>[19]</sup>。胞质多聚腺苷酸化与核内多聚腺苷酸化的调控有相似之处, 也有不同之处。两者的相似之处, 例如, 两者都需要顺式调控元件 PAS 和反式作用因子 CPSF 和



symplekin 的参与。两者的不同之处,例如,与核内多聚腺苷酸化不同,具有多个 PAS 的 mRNA 在胞质多聚腺苷酸化中,其近端 PAS 可远程调控 mRNA 3' 末端的多聚腺苷酸化<sup>[18]</sup>。

## 2 APA 的生物学效应

根据所在位置不同,PAS 可以分为非编码区 PAS (untranslated region PAS, UTR-PAS)和编码区 PAS (coding region PAS, CR-PAS)两大类<sup>[1,20]</sup>,其中以 UTR-PAS 最为多见。一个基因如果有多个 PAS,不同 PAS 的选择利用会使同一个基因产生 3'末端序列组成和长度均不同的转录本。基因如果选择不同的 UTR-PAS,形成的转录异构体之间 3'UTR 的序列组成和长度均存在差异,但所编码的蛋白质序列不会

改变<sup>[20]</sup>。如果选择位于外显子或内含子的 CR-PAS,所形成的转录本会编码出截短型蛋白质(图 2)<sup>[1,20]</sup>。

真核生物 mRNA 的 3'UTR 存在许多的基因表达顺式调控元件,例如 miRNA、lncRNA 和 RNA 结合蛋白(RNA binding protein, RBP)的结合位点(图 2)。这些调控元件在 mRNA 稳定性、运输、翻译以及细胞定位中发挥重要作用。APA 导致同一个基因产生多个具有不同 3'UTR 的转录本,由于 3'UTR 的差异,使得这些转录本具有不同的 mRNA 稳定性、翻译效率、运输以及细胞定位,并最终导致它们具有不同的生物学功能。

### 2.1 APA 产生差异性蛋白亚型

基因利用 CR-PAS 会编码出 C 端不同的蛋白异构体,有些蛋白异构体可能会缺失跨膜区或 DNA 结

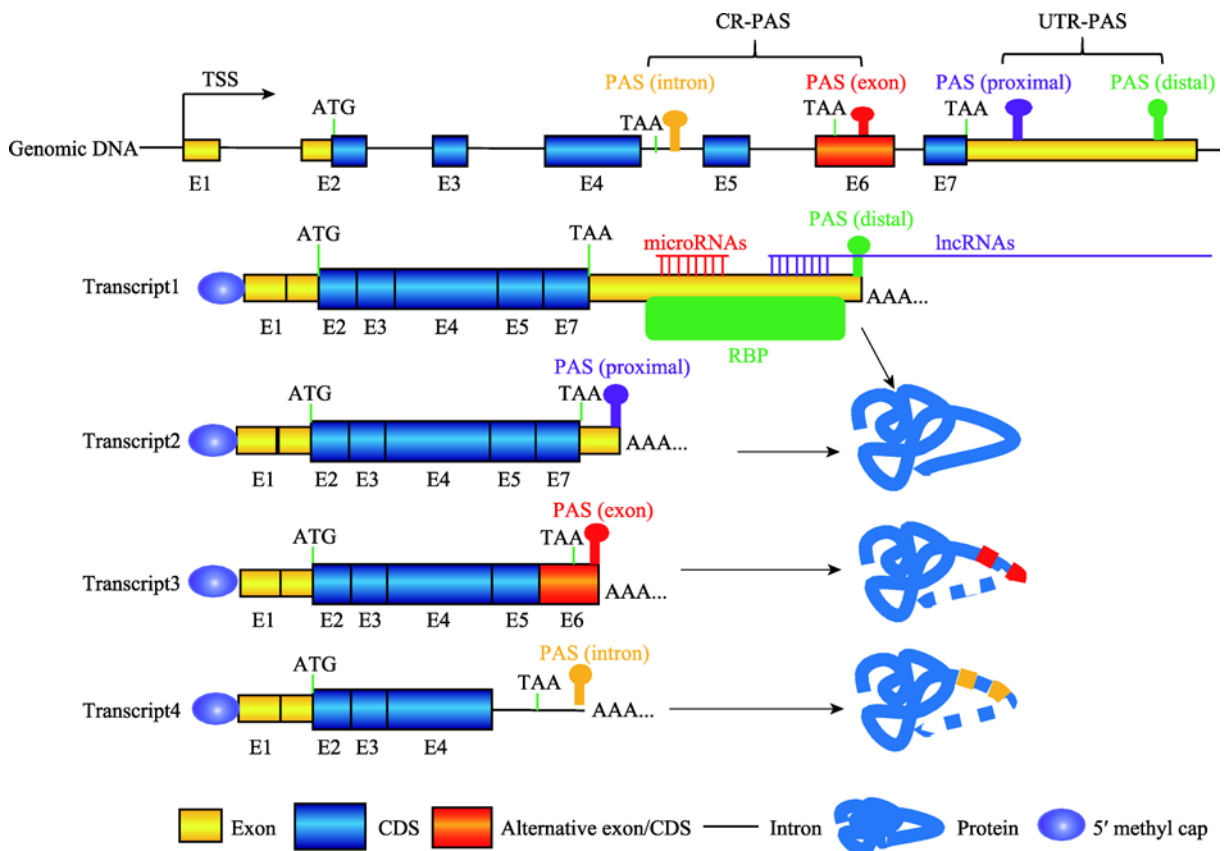


图 2 PAS 类型及其 APA 示意图

Fig. 2 Schematic of different types of PAS and the APA events

PAS: 多聚腺苷酸化信号 (polyadenylation signal); CR-PAS: 编码区 PAS (coding region PAS); UTR-PAS: 非编码区 PAS (untranslated region PAS); TSS: 转录起始位点 (transcriptional start site); CDS: 蛋白质编码区 (coding sequence); RBP: RNA 结合蛋白 (RNA binding protein); lncRNA: 长链非编码 RNA (long non-coding RNA)。

合区等重要功能结构域(图 2), 不同蛋白异构体的功能可能相近也可能相反<sup>[1]</sup>。有研究发现, 鸡生长激素受体(growth hormone receptor, *GHR*)基因存在 CR-PAS 和 UTR-PAS。*GHR* 基因利用 UTR-PAS 会产生完整的 GHR 蛋白, 完整的 GHR 定位于细胞膜, 作为 GH-GHR-IGF1 信号通路的重要分子; 而利用 CR-PAS 则会形成疏水跨膜域缺失的分泌蛋白—生长激素结合蛋白(growth hormone binding protein, GHBP)<sup>[21,22]</sup>, 主要参与维持体内 GH 水平的稳定<sup>[23]</sup>。在 B 细胞和 T 细胞激活过程中, 多数基因选择利用 CR-PAS, 使得编码的蛋白质由膜锚定型转向分泌型<sup>[24]</sup>。视网膜母细胞瘤结合蛋白(retinoblastoma-binding protein 6, RBBP6)利用 CR-PAS 产生截短型蛋白 ISO6, 它与全长的 RBBP6 竞争性参与 PAS 剪切<sup>[25]</sup>。

## 2.2 APA 影响 mRNA 的稳定性

绝大多数 APA 可导致同一个基因转录出多个不同长度的 3'UTR 转录本。因为 3'UTR 存在调控稳定性的顺式调控元件, 如 miRNA 的结合位点(图 2), 因此, 同一个基因的不同转录本具有不同的稳定性<sup>[9]</sup>。一般来说, 短 3'UTR 的转录本具有较高的稳定性和翻译效率, 这类转录本多见于增殖期细胞; 而长的 3'UTR 转录本则有利于转录后的精细化调控, 这类转录本多见于细胞分化和组织发育等复杂生物学过程<sup>[26]</sup>。

干扰素调节因子 5 (interferon regulatory factor 5, *IRF5*)是系统性红斑狼疮的候选基因。双链 RNA 结合蛋白 STAU1 能结合 *IRF5* 长 3'UTR 亚型转录本的 Alu 元件, 导致 *IRF5* mRNA 快速降解; 而 *IRF5* 短 3'UTR 亚型的转录本不存在 Alu 元件, 因此稳定性较高<sup>[27]</sup>。*Pax3* 是调节肌细胞和肝细胞增殖的关键基因, 其长 3'UTR 亚型转录本含有 *miR-206* 结合位点, 而其短 3'UTR 亚型转录本没有, 因此 *Pax3* 的短 3'UTR 亚型转录本可逃避 *miR-206* 介导的基因沉默, 稳定性较高<sup>[28,29]</sup>。

尽管短 3'UTR 亚型转录本可显著提升 mRNA 的翻译效率, 但是, 3'端测序和蛋白定量质谱分析发现小鼠和人的 T 细胞中短 3'UTR 对蛋白质表达丰度的影响有限<sup>[30-32]</sup>, 提示并非所有 APA 引起的短 3'UTR 亚型都可提高 mRNA 的翻译效率。

## 2.3 APA 影响 mRNA 的运输和定位

mRNA 的出核转运和胞质定位受到位于 3'UTR 的定位顺式调控元件(zip codes)和反式作用因子(zip codes 结合蛋白)的调控<sup>[33]</sup>。同一基因 3'UTR 序列长度和组成不同的转录本可能具有不同的定位信号, 这会导致同一个基因的不同转录本在细胞中的分布不同和功能差异。

研究表明, 出核蛋白(Aly/REF export factor, ALYREF)通过与 PABPN1 和 CSTF64 互作, 并结合 mRNA 的 3'UTR 促进 mRNA 出核转运<sup>[34]</sup>。另外, 研究发现, 核 RNA 输出因子 1 (nuclear RNA export factor1, NXF1)可与 CFIm68 互作, 并结合于成熟的长 3'UTR 亚型转录本的 UGUA 元件协助其出核转运<sup>[35,36]</sup>。脑源性神经营养因子(brain-derived neurotrophic factor, *BDNF*)<sup>[37]</sup>和肌醇单磷酸酶(inositol monophosphatase 1)<sup>[38]</sup>的长 3'UTR 亚型转录本主要定位在神经细胞树突中, 而其短 3'UTR 亚型转录本则主要分布在神经元胞体内<sup>[37]</sup>。长 3'UTR 亚型 *CD47* 转录本 3'UTR 的 U-rich 元件可结合 Hu antigen R (HuR), HuR 会招募 SET 蛋白, 促进 SET 与新翻译出的 CD47 蛋白互作, 并同时激活 Rac 家族 GTP 酶 1 (Rac family small GTPase 1, RAC1), 从而使 CD47 向质膜转移; 短 3'UTR 亚型 *CD47* 转录本的 3'UTR 则缺少 U-rich 元件, 不能与 HuR 结合, 无法招募 SET 蛋白和激活 RAC1, 故造成 CD47 蛋白滞留在内质网中<sup>[39]</sup>。*CD44*、整合素  $\alpha 1$  (integrin  $\alpha 1$ , *ITGA1*)和肿瘤坏死因子受体超家族成员 13C (TNF receptor superfamily member 13C, *TNFRSF13C*)的不同蛋白质亚型也存在这种细胞定位机制<sup>[39]</sup>。虽然长 3'UTR 亚型转录本在细胞质中易降解, 但却具有高效的出核方式, 这保证了长 3'UTR 亚型转录本在细胞质中发挥作用, 也保证了其转录后的精细调控。

## 3 APA 调控机制

简单来说, 一个具有多个 PAS 的基因, 其 mRNA 前体在成熟过程中 PAS 的选择利用取决于各 PAS 发生 C/P 反应的时机。当 mRNA 前体具有两个 PAS 时, PAS 的选择涉及 Tp1、Tp2 和 Tt 三个关键时间变量, Tp1 和 Tp2 分别是在近端 PAS 和远端 PAS 进行 C/P

反应的时间,  $T_t$  是转录复合体通过两个 PAS 间的时间<sup>[40]</sup>。如果  $T_{p1} < T_t + T_{p2}$ , 近端 PAS 的利用率增加, 会产生短 3'UTR 亚型转录本; 如果  $T_{p1} > T_t + T_{p2}$ , 则远端 PAS 的利用率增加, 产生长 3'UTR 亚型转录本<sup>[40]</sup>。这 3 个关键时间参数受到顺式调控元件信号强弱和反式作用因子作用强度等多个因素的调控。目前已知 APA 的调控受到 PAS 序列、PAS 序列间的距离、染色质的修饰、转录酶的延伸速度、蛋白质-RNA 互作强度和核心反式作用因子浓度的影响<sup>[1,9]</sup>。

### 3.1 APA 顺式调控

PAS 是 C/P 反应的顺式调控元件, PAS 序列信号的强弱影响其在 APA 中是否被选择和利用。远端 PAS 常具有经典的 AAUAAA、U/GU-rich 等元件<sup>[26,40]</sup>, 而近端 PAS 常含有这些元件的非经典序列, 其剪切活性较远端 PAS 低, 故在 APA 中通常选择和利用远端 PAS, 这种现象普遍存在于在果蝇、小鼠和人的大多数基因<sup>[2]</sup>。

两个 PAS 间的序列也会影响 PAS 的选择。研究发现, 在两个 PAS 间插入外源片段可提高近端 PAS 的利用<sup>[41]</sup>。两个 PAS 之间的 DNA 甲基化也会影响 PAS 的选择利用。研究发现, 作为甲基化特异绝缘蛋白的 CCCTC 结合因子(CCCTC-binding factor, CTCF)会与 DNA 甲基化位点结合, 并招募黏连蛋白复合物(cohesin complex)形成染色质环状结构, 阻止 RNAP II 的转录延伸, 从而促进近端 PAS 的选择利用<sup>[42]</sup>。

染色质空间结构也可影响 PAS 的选择, 处在染色质较为松散区域的 PAS, 易被选择和利用。研究发现, 酿酒酵母的组蛋白 H3K4 和 H3K36 甲基化酶 SET1 和 SET2 可调控 APA, 这是由于这两个甲基化酶修饰组蛋白 H3 后, H3K4me1 和 H3K36me3 可增加核小体乙酰化, 从而减少了核小体的位阻效应, 使染色质处于松散的开放状态, 有利于选择远端 PAS<sup>[43]</sup>。还有研究发现 APA 的调控也与 mRNA 3'UTR 的腺嘌呤的甲基化( $N^6$ -methyladenosine,  $m^6A$ )和 RNA 二级结构有关<sup>[44,45]</sup>。

### 3.2 APA 反式调控

APA 受到前述参与 C/P 反应的 4 个蛋白质复合

物的核心蛋白因子、RNA 拼接因子以及其他一些 RNA 结合蛋白等的调控。

其中, RNAP II 是 APA 的重要反式调控因子。黑腹果蝇中有一种突变型 RNAP II, 它的转录延伸效率较低, 导致 *polo* 基因短 3'UTR 亚型转录本的表达量增加<sup>[46]</sup>。在紫外线诱导的酵母 DNA 损伤修复模型中, 紫外线引起的 RNAP II 的 C 端高度磷酸化, 在降低转录延伸效率的同时还会影响剪接体蛋白复合物的结合<sup>[47]</sup>, 结果导致 *RPB2* 基因利用近端 PAS, 产生短 3'UTR 亚型转录本<sup>[48]</sup>。

反式作用因子浓度的改变是调控 APA 的一个重要因素。有研究发现 CSTF64/CSTF64 $\tau$  或 FIP1 的缺失可增加多数基因远端高活性 PAS 的选择利用<sup>[49-51]</sup>, 敲低 CFI $m25$  和 CFI $m68$  会增加多数基因近端低活性 PAS 的选择利用<sup>[52]</sup>。在紫外线诱导 DNA 损伤修复的模型中, BRCA1/BARD1 结合并隔绝 CSTF50, 使得 CSTF50 的浓度显著降低, 造成较多基因选择利用远端 PAS<sup>[53]</sup>。胚胎干细胞(ESCs)分化过程中, FIP1 和 CSTF 各亚基的蛋白水平呈逐渐下降的趋势, 同时较多的基因选择利用远端 PAS。与此相一致, FIP1 缺失会导致 ESC 细胞较多基因选择利用远端 PAS<sup>[50]</sup>。降低 RNA 结合蛋白 Nudt21 会促进细胞重编程相关基因选择利用近端 PAS, 形成短 3'UTR 亚型转录本, 提高其稳定性和编码能力, 促进已分化细胞发生去分化, 形成具有多能性的祖细胞<sup>[54]</sup>。

反式作用因子和顺式调控元件的互作强度减弱或丧失也是影响 APA 的重要因素。Robert 等<sup>[27]</sup>发现 *IRF5* 基因近端 PAS 的经典序列 AAUAAA 突变为非经典 AAGAAA 后, CPSF160 与其互作的能力下降, 导致由选择近端 PAS 转而选择远端 PAS, 从而产生长 3'UTR 亚型的 mRNA。异质性核糖核蛋白 F(heterogeneous nuclear ribonucleoproteins F, hnRNP) 和 H/H<sup>[49]</sup>、多聚嘧啶序列结合蛋白(PTB)<sup>[55]</sup>和性别致死因子(sex lethal)<sup>[56]</sup>可结合 CR-PAS 下游的 GU-rich 元件, 竞争性抑制 CSTF64 与 GU-rich 元件结合, 导致 UTR-PAS 型选择利用增加。有报道发现 PABPN1 与近端 PAS 结合并抑制 CPSF160 与其互作, 导致远端 PAS 的选择利用增加<sup>[57,58]</sup>。异质性核糖核蛋白 K(heterogeneous nuclear ribonucleoprotein K, hnRNP K)可阻止 CFI $m68$  与 CFI $m25$  结合, 抑制 CFI $m$  复合物与 *NEAT1* 的近端 PAS 上游的 UGUA 结合, 导致远



端 PAS 的选择利用增加, 从而促进长 3'UTR 亚型 lncRNA *NEAT1\_2* 的表达<sup>[59,60]</sup>。 $\alpha$ CP 胞嘧啶结合蛋白 ( $\alpha$ CP poly(C) binding protein,  $\alpha$ CP PCBP) 能结合 C-rich 元件, 促进 C-rich 元件临近的 PAS 的选择利用<sup>[61]</sup>。

许多 APA 的调控因子自身也受到 APA 的调控。例如人和果蝇 *CSTF77* 基因利用 CR-PAS 产生没有编码能力的 mRNA<sup>[62]</sup>, 导致 *CSTF77* 的表达降低, *CSTF77* 表达下降使其对 CR-PAS 的选择利用减少, UTR-PAS 选择利用增加, 产生具有编码能力的长亚型转录本, 提高细胞内 *CSTF77* 蛋白水平<sup>[63]</sup>, 而高水平的 *CSTF77* 会促进 CR-PAS 的选择利用, 产生没有编码能力的 mRNA。有报导表明 *CSTF64* 和 *CSTF64 $\tau$*  也存在这种表达互抑现象<sup>[64,65]</sup>。这种通过 APA 调节 mRNA 长度的机制可以维持 C/P 反应所需蛋白的生理浓度, 并为其他基因的 APA 调控提供了保障。

## 4 APA 异常与疾病

目前已知绝大多数的真核生物基因都存在 APA。作为基因表达的一个重要调控机制, APA 的异常与人类疾病密切相关。目前发现, 血液、免疫、神经、癌症等多种疾病与 APA 的异常相关<sup>[4-7]</sup>。研究发现, 在地中海贫血患者中,  $\alpha 2$ -globin 和  $\beta$ -globin 基因近端 PAS 的突变和缺失造成了远端 PAS 选择利用的增加<sup>[4]</sup>; 同样发现, 在静脉血栓患者中, 近端 PAS 突变引起的凝血酶原(prothrombin)基因远端 PAS 的利用增加<sup>[66]</sup>。在急性粒细胞白血病患者中, *CPSF160* 的表达显著高于健康人群, 造成白血病相关的融合基因产生高稳定性的短 3'UTR 亚型转录本, 促进白血病细胞生长<sup>[67]</sup>。在动脉粥样硬化、II 型糖尿病、结肠炎和关节炎等慢性炎症疾病患者中, 炎症反应因子 *NLRP3* 的短 3'UTR 亚型转录本的表达水平较高, *NLRP3* 的短 3'UTR 亚型转录本可逃避 *Tristetraprolin*(TTP)和 *miR-223* 的负调控, 导致 *NLRP3* 蛋白表达升高, 诱导机体发生炎症反应<sup>[68]</sup>。胶质母细胞瘤的 *CFIm25* 的表达增加, 导致致癌基因产生高稳定性的短 3'UTR 亚型转录本, 增强了恶性胶质瘤细胞的增殖和侵袭能力<sup>[69]</sup>。在肌萎缩性脊髓侧索硬

化症(amyotrophic lateral sclerosis, ALS)和额颞叶痴呆(frontotemporal dementia, FTD)患者中, 抑微管装配蛋白(*stathmin-2*)基因利用 CR-PAS 所产生的转录本不能翻译出具有功能的蛋白, 导致运动神经元轴突的再生被抑制<sup>[70]</sup>。在阿尔兹海默症(Alzheimer's disease, AD)患者中, 微管相关蛋白的 *Tau* 基因选择利用近端 PAS, 避免了 *miR-34a* 的抑制作用, 导致 *TAU* 蛋白的表达增加, 加剧神经元内 *TAU* 蛋白凝聚, 造成神经纤维缠结<sup>[71]</sup>。在帕金森疾病(Parkinson's disease, PD)患者中, 细胞内多巴胺会促进  $\alpha$ -突触蛋白( $\alpha$ -synuclein)基因产生长 3'UTR 亚型转录本, 增强了  $\alpha$ -突触蛋白转录本的稳定性和蛋白表达水平, 造成  $\alpha$ -突触蛋白由突触末端向线粒体和胞体转移, 蓄积的  $\alpha$ -突触蛋白形成路易小体(lewy body)<sup>[72]</sup>。慢性淋巴细胞白血病和多发性骨髓瘤细胞的 3'端转录组测序分析发现, 肿瘤细胞的抑癌基因较多地选择利用 CR-PAS, 产生缺失 DNA 结合结构域或跨膜结构域的蛋白, 这些蛋白往往丧失抑癌作用<sup>[73,74]</sup>。

## 5 结语与展望

近来大量的研究表明, APA 是真核生物中广泛存在的一个重要基因表达调控机制, 但目前人们对于许多基因 APA 的生物学效应和调控机制了解还不够, 特别是对特定组织和细胞在特定发育阶段的 APA 及其 APA 的生物学效应和调控尚不清楚。未来有必要开展以下 4 个方面的工作: (1)建立高通量、高准确率的 APA 测序分析平台。虽然目前已建立了多种 3'端富集测序分析技术, 但是这些方法仍存在不足之处, 分析结果不够准确。未来, 有必要采用第三代测序和单细胞测序技术开展特定组织或细胞在特定发育阶段的 APA 研究。(2)开展 APA 的调控机制研究。尽管目前人们已鉴定出了许多 APA 的顺式调控元件和反式作用因子, 但还需进一步开展不同组织和细胞的特异性 APA 顺式调控元件和反式作用因子(包括非编码 RNA)的鉴定和功能分析。(3)开展表观遗传修饰对 APA 的调控研究。目前已有研究发现, DNA、RNA 和组蛋白的表观修饰影响基因的 APA, 但其作用机制尚不清楚, 未来有必要开展 DNA



甲基化、组蛋白修饰以及 RNA 前体甲基化等表观修饰对 APA 的影响和作用机制的研究。(4)开展 APA 的生物学效应和调控机制的体内研究。目前, APA 研究主要局限于体外试验, APA 在体内真正的生物学效应和调控机制还有待阐明, 未来需要利用基因敲除和基因编辑等技术, 开展体内 APA 的生物学效应和调控机制研究。随着 APA 研究的不断深入, 相信许多生物学过程的分子作用机制和一些人类疾病的发生机制终将会被阐明。

## 参考文献(References):

- [1] Tian B, Manley JL. Alternative polyadenylation of mRNA precursors. *Nat Rev Mol Cell Biol*, 2017, 18(1): 18–30. [\[DOI\]](#)
- [2] Derti A, Garrett-Engle P, Macisaac KD, Stevens RC, Sriram S, Chen RH, Rohl CA, Johnson JM, Babak T. A quantitative atlas of polyadenylation in five mammals. *Genome Res*, 2012, 22(6): 1173–1183. [\[DOI\]](#)
- [3] Brutman JN, Zhou X, Zhang YZ, Michal J, Stark B, Jiang ZH, Davis JF. Mapping diet-induced alternative polyadenylation of hypothalamic transcripts in the obese rat. *Physiol Behav*, 2018, 188: 173–180. [\[DOI\]](#)
- [4] Curinha A, Oliveira Braz S, Pereira-Castro I, Cruz A, Moreira A. Implications of polyadenylation in health and disease. *Nucleus*, 2014, 5(6): 508–519. [\[DOI\]](#)
- [5] Gruber AJ, Zavolan M. Alternative cleavage and polyadenylation in health and disease. *Nat Rev Genet*, 2019, 20(10): 599–614. [\[DOI\]](#)
- [6] Patel R, Brophy C, Hickling M, Neve J, Furger A. Alternative cleavage and polyadenylation of genes associated with protein turnover and mitochondrial function are deregulated in Parkinson's, Alzheimer's and ALS disease. *BMC Med Genomics*, 2019, 12(1): 60. [\[DOI\]](#)
- [7] Nourse J, Spada S, Danckwardt S. Emerging roles of RNA 3'-end cleavage and polyadenylation in pathogenesis, diagnosis and therapy of human disorders. *Biomolecules*, 2020, 10(6): 915. [\[DOI\]](#)
- [8] Jafari Najaf Abadi MH, Shafabakhsh R, Asemi Z, Mirzaei HR, Sahebhasagh R, Mirzaei H, Hamblin MR. CFIm25 and alternative polyadenylation: conflicting roles in cancer. *Cancer Lett*, 2019, 459: 112–121. [\[DOI\]](#)
- [9] Neve J, Patel R, Wang ZQ, Louey A, Furger AM. Cleavage and polyadenylation: ending the message expands gene regulation. *RNA Biololgy*, 2017, 14(7): 865–890. [\[DOI\]](#)
- [10] Zheng DH, Tian B. RNA-binding proteins in regulation of alternative cleavage and polyadenylation. *Adv Exp Med Biol*, 2014, 825: 97–127. [\[DOI\]](#)
- [11] Knipe DM, Howley PM. *Fields virology*. 6th ed. 2013: Lippincott Williams & Wilkins Publishing. [\[DOI\]](#)
- [12] Chan SL, Huppertz I, Yao CG, Weng LJ, Moresco JJ, Yates JR 3rd, Ule J, Manley JL, Shi YS. CPSF30 and Wdr33 directly bind to AAUAAA in mammalian mRNA 3' processing. *Genes Dev*, 2014, 28(21): 2370–2380. [\[DOI\]](#)
- [13] Hamilton K, Sun YD, Tong L. Biophysical characterizations of the recognition of the AAUAAA polyadenylation signal. *RNA*, 2019, 25(12): 1673–1680. [\[DOI\]](#)
- [14] Sha QQ, Zhang J, Fan HY. A story of birth and death: mRNA translation and clearance at the onset of maternal-to-zygotic transition in mammals. *Biol Reprod*, 2019, 101(3): 579–590. [\[DOI\]](#)
- [15] Winata CL, Łapiński M, Pryszcz L, Vaz C, Bin Ismail MH, Nama S, Hajan HS, Lee SGP, Korzh V, Sampath P, Tanavde V, Mathavan S. Cytoplasmic polyadenylation-mediated translational control of maternal mRNAs directs maternal-to-zygotic transition. *Development*, 2018, 145(1): dev159566. [\[DOI\]](#)
- [16] Shao M, Lu T, Zhang C, Zhang YZ, Kong SH, Shi DL. Rbm24 controls poly(A) tail length and translation efficiency of *crystallin* mRNAs in the lens via cytoplasmic polyadenylation. *Proc Natl Acad Sci USA*, 2020, 117(13): 7245–7254. [\[DOI\]](#)
- [17] Villalba A, Coll O, Gebauer F. Cytoplasmic polyadenylation and translational control. *Curr Opin Genet Dev*, 2011, 21(4): 452–457. [\[DOI\]](#)
- [18] Dai XX, Jiang JC, Sha QQ, Jiang Y, Ou XH, Fan HY. A combinatorial code for mRNA 3'-UTR-mediated translational control in the mouse oocyte. *Nucleic Acids Res*, 2019, 47(1): 328–340. [\[DOI\]](#)
- [19] Charlesworth A, Meijer HA, de Moor CH. Specificity factors in cytoplasmic polyadenylation. *Wiley Interdiscip Rev RNA*, 2013, 4(4): 437–461. [\[DOI\]](#)
- [20] Chen W, Jia Q, Song YF, Fu HH, Wei G, Ni T. Alternative polyadenylation: methods, findings, and impacts. *Genomics Proteomics Bioinformatics*, 2017, 15(5): 287–300. [\[DOI\]](#)
- [21] Lau SL. Molecular characterization of the chicken growth hormone receptor gene. Hong Kong: Hong Kong University, 2005. [\[DOI\]](#)
- [22] Lau JS, Yip CW, Law KM, Leung FC. Cloning and characterization of chicken growth hormone binding protein (cGHBP). *Domest Anim Endocrinol*, 2007, 33(1): 107–121. [\[DOI\]](#)

- [23] Dehkhoda F, Lee CMM, Medina J, Brooks AJ. The growth hormone receptor: mechanism of receptor activation, cell signaling, and physiological aspects. *Front Endocrinol*, 2018, 9: 35. [DOI]
- [24] Cheng LC, Zheng DH, Baljinnyam E, Sun FZ, Ogami K, Yeung PL, Hoque M, Lu CW, Manley JL, Tian B. Widespread transcript shortening through alternative polyadenylation in secretory cell differentiation. *Nat Commun*, 2020, 11(1): 3182. [DOI]
- [25] Di Giammartino DC, Li WC, Ogami K, Yashinskii JJ, Hoque M, Tian B, Manley JL. RBBP6 isoforms regulate the human polyadenylation machinery and modulate expression of mRNAs with AU-rich 3' UTRs. *Genes Dev*, 2014, 28(20): 2248–2260. [DOI]
- [26] Turner RE, Pattison AD, Beilharz TH. Alternative polyadenylation in the regulation and dysregulation of gene expression. *Semin Cell Dev Biol*, 2018, 75: 61–69. [DOI]
- [27] Graham RR, Kyogoku C, Sigurdsson S, Vlasova IA, Davies LRL, Baechler EC, Plenge RM, Koeuth T, Ortmann WA, Hom G, Bauer JW, Gillett C, Burt N, Cunningham Graham DS, Onofrio R, Petri M, Gunnarsson I, Svenungsson E, Rönnblom L, Nordmark G, Gregersen PK, Moser K, Gaffney PM, Criswell LA, Vyse TJ, Syvänen AC, Bohjanen PR, Daly MJ, Behrens TW, Altshuler D. Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus. *Proc Natl Acad Sci USA*, 2007, 104(16): 6758–6763. [DOI]
- [28] Boutet SC, Cheung TH, Quach NL, Liu L, Prescott SL, Edalati A, Iori K, Rando TA. Alternative polyadenylation mediates microRNA regulation of muscle stem cell function. *Cell Stem Cell*, 2012, 10(3): 327–336. [DOI]
- [29] de Morree A, Klein JDD, Gan Q, Farup J, Urtasun A, Kanugovi A, Bilen B, van Velthoven CTJ, Quarta M, Rando TA. Alternative polyadenylation of Pax3 controls muscle stem cell fate and muscle function. *Science*, 2019, 366(6466): 734–738. [DOI]
- [30] Gruber AR, Martin G, Müller P, Schmidt A, Gruber AJ, Gumieny R, Mittal N, Jayachandran R, Pieters J, Keller W, van Nimwegen E, Zavolan M. Global 3' UTR shortening has a limited effect on protein abundance in proliferating T cells. *Nat Commun*, 2014, 5: 5465. [DOI]
- [31] Spies N, Burge CB, Bartel DP. 3' UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. *Genome Res*, 2013, 23(12): 2078–2090. [DOI]
- [32] Spangenberg L, Shigunov P, Abud APR, Cofré AR, Stimamiglio MA, Kuligovski C, Zych J, Schittini AV, Costa ADT, Rebelatto CK, Brofman PRS, Goldenberg S, Correa A, Naya H, Dallagiovanna B. Polysome profiling shows extensive posttranscriptional regulation during human adipocyte stem cell differentiation into adipocytes. *Stem Cell Res*, 2013, 11(2): 902–912. [DOI]
- [33] Jambhekar A, Derisi JL. Cis-acting determinants of asymmetric, cytoplasmic RNA transport. *RNA*, 2007, 13(5): 625–642. [DOI]
- [34] Shi M, Zhang H, Wu XD, He ZS, Wang LT, Yin SY, Tian B, Li GH, Cheng H. ALYREF mainly binds to the 5' and the 3' regions of the mRNA *in vivo*. *Nucleic Acids Res*, 2017, 45(16): 9640–9653. [DOI]
- [35] Ruepp MD, Aringhieri C, Vivarelli S, Cardinale S, Paro S, Schümperli D, Barabino SML. Mammalian pre-mRNA 3' end processing factor CF I m 68 functions in mRNA export. *Mol Biol Cell*, 2009, 20(24): 5211–5223. [DOI]
- [36] Chen SL, Wang RJ, Zheng DH, Zhang H, Chang XY, Wang K, Li WC, Fan J, Tian B, Cheng H. The mRNA export receptor NXF1 coordinates transcriptional dynamics, alternative polyadenylation, and mRNA export. *Mol Cell*, 2019, 74(1): 118–131. e7. [DOI]
- [37] An JJ, Gharami K, Liao GY, Woo NH, Lau AG, Vanevski F, Torre ER, Jones KR, Feng Y, Lu B, Xu BJ. Distinct role of long 3' UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. *Cell*, 2008, 134(1): 175–187. [DOI]
- [38] Andreassi C, Riccio A. To localize or not to localize: mRNA fate is in 3'UTR ends. *Trends Cell Biol*, 2009, 19(9): 465–474. [DOI]
- [39] Berkovits BD, Mayr C. Alternative 3' UTRs act as scaffolds to regulate membrane protein localization. *Nature*, 2015, 522(7556): 363–367. [DOI]
- [40] Davis R, Shi YS. The polyadenylation code: a unified model for the regulation of mRNA alternative polyadenylation. *J Zhejiang Univ Sci B*, 2014, 15(5): 429–437. [DOI]
- [41] Marsollier AC, Joubert R, Mariot V, Dumonceaux J. Targeting the polyadenylation signal of pre-mRNA: a new gene silencing approach for facioscapulohumeral dystrophy. *Int J Mol Sci*, 2018, 19(5): 1347. [DOI]
- [42] Nanavaty V, Abrash EW, Hong CJ, Park S, Fink EE, Li ZY, Sweet TJ, Bhasin JM, Singuri S, Lee BH, Hwang TH, Ting AH. DNA methylation regulates alternative polyadenylation via CTCF and the cohesin complex. *Mol Cell*, 2020, 78(4): 752–764. e6. [DOI]

- [43] Michaels KK, Mohd Mostafa S, Ruiz Capella J, Moore CL. Regulation of alternative polyadenylation in the yeast *Saccharomyces cerevisiae* by histone H3K4 and H3K36 methyltransferases. *Nucleic Acids Res*, 2020, 48(10): 5407–5425. [DOI]
- [44] Yue YN, Liu J, Cui XL, Cao J, Luo GZ, Zhang ZZ, Cheng T, Gao MS, Shu X, Ma HH, Wang FQ, Wang XX, Shen B, Wang YZ, Feng XH, He C, Liu JZ. VIRMA mediates preferential m<sup>6</sup>A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. *Cell Discov*, 2018, 4: 10. [DOI]
- [45] Patraquim P, Warnefors M, Alonso CR. Evolution of Hox post-transcriptional regulation by alternative polyadenylation and microRNA modulation within 12 *Drosophila* genomes. *Mol Biol Evol*, 2011, 28(9): 2453–2460. [DOI]
- [46] Pinto PAB, Henriques T, Freitas MO, Martins T, Domingues RG, Wyrzykowska PS, Coelho PA, Carmo AM, Sunkel CE, Proudfoot NJ, Moreira A. RNA polymerase II kinetics in polo polyadenylation signal selection. *EMBO J*, 2011, 30(12): 2431–2444. [DOI]
- [47] Maita H, Nakagawa S. What is the switch for coupling transcription and splicing? RNA Polymerase II C-terminal domain phosphorylation, phase separation and beyond. *Wiley Interdiscip Rev RNA*, 2020, 11(1): e1574. [DOI]
- [48] Yu LJ, Volkert MR. UV damage regulates alternative polyadenylation of the RPB2 gene in yeast. *Nucleic Acids Res*, 2013, 41(5): 3104–3114. [DOI]
- [49] Nazim M, Masuda A, Rahman MA, Nasrin F, Takeda JI, Ohe K, Ohkawara B, Ito M, Ohno K. Competitive regulation of alternative splicing and alternative polyadenylation by hnRNP H and CstF64 determines acetylcholinesterase isoforms. *Nucleic Acids Res*, 2017, 45(3): 1455–1468. [DOI]
- [50] Lackford B, Yao CG, Charles GM, Weng LJ, Zheng XF, Choi EA, Xie XH, Wan J, Xing Y, Freudenberg JM, Yang PY, Jothi R, Hu G, Shi YS. Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. *EMBO J*, 2014, 33(8): 878–889. [DOI]
- [51] Hwang HW, Park CY, Goodarzi H, Fak JJ, Mele A, Moore MJ, Saito Y, Darnell RB. PAPERCLIP identifies microRNA targets and a role of CstF64/64tau in promoting non-canonical poly(A) site usage. *Cell Rep*, 2016, 15(2): 423–435. [DOI]
- [52] Zhou ZJ, Qu J, He L, Zhu Y, Yang SZ, Zhang F, Guo T, Peng H, Chen P, Zhou Y. Stiff matrix instigates type I collagen biogenesis by mammalian cleavage factor I complex-mediated alternative polyadenylation. *JCI Insight*, 2020, 5(3): e133972. [DOI]
- [53] Fontana GA, Rigamonti A, Lenzken SC, Filosa G, Alvarez R, Calogero R, Bianchi ME, Barabino SML. Oxidative stress controls the choice of alternative last exons via a Brahma-BRCA1-CstF pathway. *Nucleic Acids Res*, 2017, 45(2): 902–914. [DOI]
- [54] Brumbaugh J, Di Stefano B, Wang XY, Borkent M, Forouzmand E, Clowers KJ, Ji F, Schwarz BA, Kalocsay M, Elledge SJ, Chen Y, Sadreyev RI, Gygi SP, Hu G, Shi YS, Hochedlinger K. Nudt21 controls cell fate by connecting alternative polyadenylation to chromatin signaling. *Cell*, 2018, 172(1–2): 106–120. e21. [DOI]
- [55] Roy D, Bhanja Chowdhury J, Ghosh S. Polypyrimidine tract binding protein (PTB) associates with intronic and exonic domains to squelch nuclear export of unspliced RNA. *FEBS Lett*, 2013, 587(23): 3802–3807. [DOI]
- [56] Gawande B, Robida MD, Rahn A, Singh R. *Drosophila* Sex-lethal protein mediates polyadenylation switching in the female germline. *EMBO J*, 2006, 25(6): 1263–1272. [DOI]
- [57] Jenal M, Elkon R, Loayza-Puch F, van Haften G, Kühn U, Menzies FM, Oude Vrielink JAF, Bos AJ, Drost J, Rooijers K, Rubinsztein DC, Agami R. The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell*, 2012, 149(3): 538–553. [DOI]
- [58] Mohibi S, Zhang J, Chen XB. PABPN1, a target of p63, modulates keratinocyte differentiation through regulation of p63α mRNA translation. *J Invest Dermatol*, 2020, 131(140): 2166–2177. [DOI]
- [59] Naganuma T, Nakagawa S, Tanigawa A, Sasaki YF, Goshima N, Hirose T. Alternative 3'-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *EMBO J*, 2012, 31(20): 4020–4034. [DOI]
- [60] Naganuma T, Hirose T. Paraspeckle formation during the biogenesis of long non-coding RNAs. *RNA Biol*, 2013, 10(3): 456–461. [DOI]
- [61] Ji XJ, Wan J, Vishnu M, Xing Y, Liebhaber SA. αCP Poly(C) binding proteins act as global regulators of alternative polyadenylation. *Mol Cell Biol*, 2013, 33(13): 2560–2573. [DOI]
- [62] Juge F, Audibert A, Benoit B, Simonelig M. Tissue-specific autoregulation of *Drosophila* suppressor of forked by alternative poly(A) site utilization leads to accumulation of the suppressor of forked protein in mitotically active cells. *RNA*, 2000, 6(11): 1529–1538. [DOI]
- [63] Luo WT, Ji Z, Pan ZH, You B, Hoque M, Li WC,

- Gunderson SI, Tian B. The conserved intronic cleavage and polyadenylation site of CstF-77 gene imparts control of 3' end processing activity through feedback auto-regulation and by U1 snRNP. *Plos Genet*, 2013, 9(7): e1003613. [DOI]
- [64] Yao CG, Biesinger J, Wan J, Weng LJ, Xing Y, Xie XH, Shi YS. Transcriptome-wide analyses of CstF64-RNA interactions in global regulation of mRNA alternative polyadenylation. *Proc Natl Acad Sci USA*, 2012, 109(46): 18773–18778. [DOI]
- [65] Yao CG, Choi EA, Weng LJ, Xie XH, Wan J, Xing Y, Moresco JJ, Tu PG, Yates JR, Shi YS. Overlapping and distinct functions of CstF64 and CstF64 $\tau$  in mammalian mRNA 3' processing. *RNA*, 2013, 19(12): 1781–1790. [DOI]
- [66] Ceelie H, Spaargaren-van Riel CC, Bertina RM, Vos HL. G20210A is a functional mutation in the prothrombin gene; effect on protein levels and 3'-end formation. *J Thromb Haemost*, 2004, 2(1): 119–127. [DOI]
- [67] Shima T, Davis AG, Miyauchi S, Kochi Y, Johnson DT, Stoner SA, Junichiro Y, Miyamoto T, Zhou JH, Ball ED, Akashi K, Zhang DE. CPSF1 regulates AML1-ETO fusion gene polyadenylation and stability in t(8; 21) acute myelogenous leukemia. *Blood*, 2017, 130(Suppl.1): 2498. [DOI]
- [68] Haneklaus M, O'neil JD, Clark AR, Masters SL, O'Neill LAJ. The RNA-binding protein Tristetraprolin (TTP) is a critical negative regulator of the NLRP3 inflammasome. *J Biol Chem*, 2017, 292(17): 6869–6881. [DOI]
- [69] Masamha CP, Xia Z, Yang JX, Albrecht TR, Li M, Shyu AB, Li W, Wagner EJ. CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. *Nature*, 2014, 510(7505): 412–416. [DOI]
- [70] Melamed Z, López-Erauskin J, Baughn MW, Zhang OY, Drenner K, Sun Y, Freyermuth F, McMahon MA, Beccari MS, Artates JW, Ohkubo T, Rodriguez M, Lin NW, Wu DM, Bennett CF, Rigo F, Da Cruz S, Ravits J, Lagier-Tourenne C, Cleveland DW. Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration. *Nat Neurosci*, 2019, 22(2): 180–190. [DOI]
- [71] Dickson JR, Kruse C, Montagna DR, Finsen B, Wolfe MS. Alternative polyadenylation and miR-34 family members regulate tau expression. *J Neurochem*, 2013, 127(6): 739–749. [DOI]
- [72] Rhinn H, Qiang L, Yamashita T, Rhee D, Zolin A, Vanti W, Abeliovich A. Alternative  $\alpha$ -synuclein transcript usage as a convergent mechanism in Parkinson's disease pathology. *Nat Commun*, 2012, 3: 1084. [DOI]
- [73] Lee SH, Singh I, Tisdale S, Abdel-Wahab O, Leslie CS, Mayr C. Widespread intronic polyadenylation inactivates tumour suppressor genes in leukaemia. *Nature*, 2018, 561(7721): 127–131. [DOI]
- [74] Singh I, Lee SH, Sperling AS, Samur MK, Tai YT, Fulciniti M, Munshi NC, Mayr C, Leslie CS. Widespread intronic polyadenylation diversifies immune cell transcriptomes. *Nat Commun*, 2018, 9(1): 1716. [DOI]

(责任编辑: 杨昭庆)