

# 可变多聚腺苷酸化在肿瘤中的研究进展

周梦浩, 梁华庚\*

华中科技大学同济医学院附属协和医院泌尿外科, 湖北 武汉

收稿日期: 2024年2月27日; 录用日期: 2024年3月21日; 发布日期: 2024年3月28日

## 摘要

可变多聚腺苷酸化(Alternative Polyadenylation, APA)作为一种转录后调控机制, 其在基因表达调控中的作用越来越受到重视。近期的许多研究发现APA在肿瘤的发生发展和耐药机制中发挥着重要作用。本文主要总结了APA在基因表达调控和癌症等疾病中的功能作用和新的APA相关的研究方法和数据库。

## 关键词

可变多聚腺苷酸化, 肿瘤, 基因表达调控, RNA测序

# Advances in Alternative Polyadenylation in Tumors

Menghao Zhou, Huageng Liang\*

Department of Urology, Union Hospital Affiliated to Tongji Medical College, Huazhong University of Science and Technology, Wuhan Hubei

Received: Feb. 27<sup>th</sup>, 2024; accepted: Mar. 21<sup>st</sup>, 2024; published: Mar. 28<sup>th</sup>, 2024

## Abstract

The role of alternative polyadenylation (APA) as a post-transcriptional regulatory mechanism in the regulation of gene expression has received increasing attention. Many recent studies have found that APA plays an important role in tumor development and drug resistance mechanisms. This paper mainly summarizes the functional role of APA in gene expression regulation and diseases such as cancer and new APA-related research methods and databases.

\*通讯作者。

## Keywords

Alternative Polyadenylation (APA), Tumor, Gene Expression Regulation, RNA Sequencing

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 简介

在真核生物的细胞核中, 前体 mRNA (Precursor-mRNA, pre-mRNA)合成后, 需要进行 5'端和 3'端的修饰以及对 pre-mRNA 进行剪接, 才能成为成熟的 mRNA, 被转运到核糖体, 指导蛋白质翻译。需要注意的是, 编码区两侧的 5'和 3'序列不会被翻译, 因此被称为非翻译区(Untranslated Region, UTR)。在这其中, 5'UTR 是核糖体组装进行 mRNA 翻译的主要位点。与之相对应的, 3'UTR 在转录后控制基因表达中发挥着各种作用, 包括但不限于 mRNA 的稳定性、翻译和亚细胞定位[1]。

当 mRNA 识别位于 3'UTR 中的 poly(A)位点(Polyadenylation Sites, PASs)后, 会发生可变多聚腺苷酸化(Alternative Polyadenylation, APA)。APA 是研究最广泛的 3'UTR 加工事件之一, 其结果通常导致 3'UTR 的缩短。早期的研究表明, 3'UTR 的缩短在癌症中普遍存在[2] [3]。

APA 的应用非常普遍, 大约有 70%的人类基因的 3'UTR 中存在 APA, 而约 50%的基因则含有 3 个或更多的 PASs。PAS 的选择是一个动态过程, 主要由顺式元件决定[4] [5]。这种机制在真核生物中似乎是高度保守的, 仅出现在哺乳动物和植物中, 在约 70%的酵母基因中几乎没有发生任何可变剪接(Alternative Splicing, AS)。这证明了其在进化中的重要性[6]。值得注意的是, APA 在非编码 RNA (Non-coding RNA, ncRNA)中的应用也相当普遍[7]。一项对小鼠全基因组的研究发现, 约 79%的 mRNA 基因和 66%的长链非编码 RNA (Long Non-coding RNA, lncRNA)基因中至少存在一种重要的 APA 亚型[8]。

## 2. APA 分类

现有文献已经涉及了使用各种命名法和模式对不同类型的 PAS 进行分类, 但在这里, 我们将它们大致分为 3'UTR-APA 和上游区域-APA (Upstream Regions APA, UR-APA)。当 PAS 大部分位于 3'UTR 时, 被称为 3'UTR-APA, 它改变了 3'UTR 的长度, 但保持了基因产物的一致性。3'UTR-APA 的频繁发生显著影响 mRNA 的多个方面, 包括稳定性、翻译效率等。这凸显了它在基因表达中的重要调控作用。另一类 APA 发生在最后一个外显子的上游, 因此被称为 UR-APA, 其进一步改变了编码蛋白质的可能性[7] [9] [10]。UR-APA 在不同程度上截断蛋白质产物。它包括三个亚类: 末端外显子 APA、内含子 APA、内部外显子 APA。这些 APA 亚型在细胞过程中扮演着重要的角色, 影响了蛋白质的多样性和基因表达的调控[11] [12]。

## 3. APA 因子

pre-mRNA 3'端加工复合物由四个亚单位组成。裂解和聚腺苷酸化特异因子(Cleavage and Polyadenylation Specificity Factor, CPSF)组成员包括 CPSF1 (又称 CPSF160)、CPSF2 (又称 CPSF100)、CPSF3 (又称 CPSF73)、CPSF4 (又称 CPSF30)、FIP1 (又称 FIP1L1)和 WDR33。研究表明, CPSF1 在 pre-mRNA 3'端的形成中发挥着至关重要的作用。在拟南芥中, CPSF2 被发现具有锚定 PASs 并介导转录终止的能力[13]。作为一种 pre-mRNA 3'端加工内切酶, CPSF3 参与了转录本周期的终止, 包括 RNA 的裂解[14] [15]。

FIP1 通过与多聚腺苷酸聚合酶的相互作用来调节 PAS。WDR33 是 3'端加工过程中的关键组分之一, 它在与 AAUAAA 的结合中扮演主要角色[16] [17]。

切割激活因子(Cleavage Stimulation Factor, CSTF)由 CSTF1、CSTF2 和 CSTF3 组成。CSTF 可以加强 CPSF 对 PAS 的识别能力。Yang 等人的研究发现 CSTF1 在 DNA 损伤反应过程中的染色质重塑环节发挥了重要作用[18] [19]。CSTF2 可以直接与 RNA 相互作用。其有一个 CSTF2t 的对等物, 两者的功能在一定程度上相同, 敲除两者都会引起 APA 的明显变化[20] [21]。CSTF3 也是 APA 和核定位的一个关键成分[22]。mRNA 3'端的加工处理一般主要由 CSTF1、CSTF2 和 CSTF3 负责[22] [23]。

切割因子 I (Cleavage Factor I, CFI)和切割因子 II (Cleavage Factor II, CFII)是 APA 的调控因子, 也是哺乳动物裂解机制的两个核心成分[10] [24]。在 3'UTR 长度的调节中, CFI 具有十分重要的作用。具体来说, CFI 可以选择性地与末端外显子中的远端 PAS 结合, 从而提升远端 PAS 的使用率。Rüegsegger 等人的研究表明 CFI 可以加强 CPSF 与 PAS 作用的稳定性[25]。此外, 在 HEK293 细胞中, CFI (尤其是 CFIm25 和 CFIm68)功能缺失会导致整个转录组的近端 PAS 使用率增加[26] [27]。CFII 是 3'端处理机制中特征最少的成分。CFII 由 CLP1 和 PCF11 组成。CFII 与 RNA 的亲合力主要被 PCF11 影响, 而 CFII 的裂解能力则取决于 CLP1 [28] [29]。

## 4. APA 相关技术

### 4.1. PolyAID

基因组中 PASs 的分布应与局部基因结构进行共同演化。否则, 虚假的多腺苷酸化可能导致过早的转录终止并生成异常蛋白质。为了深入了解跨越人类基因组的 PASs 优化的机制, 研究者开发了深度机器学习模型, 以前所未有的核苷酸水平分辨率识别全基因组的潜在 PAS, 并计算它们在基因组背景中的强度和使用情况。该模型定量地测量了位置特异性基序的重要性以及它们在 PAS 形成和剪切异质性中的相互作用[30]。

### 4.2. stAPAMiner

空间转录组学(Spatial Transcriptomics, ST)技术为破译转录组景观的空间背景提供了机会。stAPAMiner 的工具包可以从 ST 数据中挖掘 APA 的空间模式。从 ST 数据中识别并量化了 APA 位点。其中, 设计了一个基于 k-近邻算法的估算模型来恢复 APA 信号, 然后确定了具有 APA 使用变异空间模式的 APA 基因。stAPAMiner 利用 ST 的强大功能, 以空间分辨率探索空间 APA 模式的转录图谱。该工具包可在 <https://github.com/BMILAB/stAPAMiner> 和 <https://ngdc.cnbc.ac.cn/biocode/tools/BT007320> 上获取[31]。

### 4.3. REPAC

REPAC 是一个从 RNA 测序(RNA Sequencing, RNA-seq)数据分析 APA 的框架。REPAC 利用注释 PASs 上游 50 bp 窗口的表达量估计值, 拟合一个广义线性回归模型, 以评估不同条件下发生的 PASs 使用率差异[32]。

### 4.4. ImmAPA

免疫相关 APA 事件(Immune-related APA Event, ImmAPA)评分管道, 这是一种综合算法, 用于描述 APA 事件在肿瘤免疫相关通路中的调控作用。在 ImmAPAs 中, 排名第一的 COL1A1 3'UTR 使用与预后恶化和肿瘤免疫逃脱密切相关。此外, 通过机器学习方法构建的免疫检查点抑制剂(Immune Checkpoint Blockade, ICB)相关 ImmAPA 评分模型可有效预测免疫疗法的疗效[33]。

#### 4.5. APARENT2

残差神经网络模型 APARENT2 能更准确地从 DNA 序列推断 3'端剪切和多聚腺苷酸化。该模型推广到 APA 的情况, 适用于 PASs 数量可变的 APA [34]。

#### 4.6. DeeReCT-APA

APA 深度调控代码和工具(Deep Regulatory Code and Tools for Alternative Polyadenylation, DeeReCT-APA)用于定量预测给定基因中所有替代 PAS 的使用情况。为了适应不同基因可能具有不同数量的 PAS, DeeReCT-APA 将问题视为目标长度可变的回归任务。基于卷积神经网络-长短期记忆(Convolutional Neural Network-long Short-term Memory, CNN-LSTM)架构, DeeReCT-APA 利用 CNN 层提取序列特征, 使用双向 LSTM 明确模拟竞争 PAS 之间的相互作用, 并输出代表基因所有 PAS 使用水平的百分比分数。代码和数据可在 <https://github.com/lzx325/DeeReCT-APA-repo> 上获取[35]。

#### 4.7. QuantifyPoly (A)

加权密度峰聚类方法(QuantifyPoly (A))可以准确量化全基因组的多腺苷酸化选择。在已发表的动物和植物 3'端测序数据集上应用 QuantifyPoly (A), 它们的多腺苷酸化图谱被重塑为无数新的 PAS 群。这些新型 PAS 群大多在不同的生物样本中显示出显著的动态使用情况, 或与反式作用因子的结合位点相关联。这些 PAS 群的上游序列富含聚腺苷酸化信号 UGUA、UAAA 和/或 AAUAAA, 其方式与物种有关。PAS 群也具有物种特异性, 植物的微异质性通常高于动物。QuantifyPoly (A)广泛适用于任何类型的 3'端测序数据和物种, 可准确量化和构建复杂而动态的多腺苷酸化图谱[36]。

#### 4.8. Aptardi

Aptardi (从 RNA-Seq 数据和 DNA 序列信息进行 APA 转录组分析)在机器学习范式中利用 DNA 序列和 RNA 测序预测表达的 PAS。Aptardi 将 DNA 核苷酸序列、基因组对齐的 RNA-Seq 数据和初始转录组作为输入数据。程序会对这些初始转录本进行评估, 以确定生物样本中表达的 PAS, 并相应地优化转录本的 3'末端部分。Aptardi 模型的平均精确度是标准转录组装配程序的两倍。Aptardi 模型的召回率(算法检测到的真实 PAS 的比例)提高了三倍以上[37]。

#### 4.9. MAAPER

越来越多的 RNA-seq 方法, 尤其是用于单细胞转录组分析的方法, 会产生接近 PAS 的读数, 称为近位点读数。而 MAAPER 利用近位点读数进行 APA 分析。MAAPER 预测 PAS 的准确度和灵敏度都很高, 并能以强大的统计学方法检查不同类型的 APA 事件[38]。

#### 4.10. SCAPTURE

SCAPTURE 可以从基于 3'标签的单细胞 RNA 测序(Single-cell RNA Sequencing, scRNA-seq)中识别、评估和量化裂解位点和 PAS。SCAPTURE 以高灵敏度和准确性在单细胞中从头检测 PASs, 从而能检测到以前未注释的 PASs。量化的替代 PAS 转录本完善了基因表达之外的细胞身份分析, 丰富了从 scRNA-seq 数据中提取的信息[39]。

#### 4.11. scDaPars

scRNA-seq 的 APA 动态分析(Dynamic Analysis of APA from Single-cell RNA-seq, scDaPars)可使用 3'末端或全长 scRNA-seq 数据在单细胞和单基因分辨率下精确量化 APA 事件。在真实数据和模拟数据的验

证结果表明, scDaPars 能恢复因单细胞中测序的 mRNA 数量较少而导致的缺失 APA 事件。在应用于癌症和人类内胚层分化数据时, scDaPars 不仅揭示了细胞类型特异性 APA 调控, 还识别出了在传统基因表达分析中一般无法发现的细胞亚群[40]。

## 5. APA 相关数据库

### 5.1. ipaQTL-atlas

ipaQTL-atlas (<http://bioinfo.szbl.ac.cn/ipaQTL>)是一个内含子多腺苷酸化的综合门户网站。ipaQTL-atlas 基于对 GTEx 数据库中 838 位个体的 15,170 个 RNA-seq 数据的分析, 包含约 98 万个与内含子 APA 事件相关的 SNPs。ipaQTL-atlas 提供了访问内含子多腺苷酸化的信息的一站式门户, 可极大地推动 APA 相关疾病易感基因的发现[41]。

### 5.2. scAPAdb

scAPAdb (<http://www.bmibig.cn/scAPAdb>)提供了一个全面的、人工编辑的单细胞水平 PAS、APA 事件和 poly (A)信号图谱。目前, 其收集了来自超过 360 个 scRNA-seq 实验的 APA 信息, 涵盖 6 个物种, 包括人类、小鼠和其他几个植物物种。此外, 该数据库还提供数据批量下载, 用户可以通过基因名称、基因功能等各种关键字查询数据库[42]。

### 5.3. scAPAAtlas

scAPAAtlas (<http://www.bioailab.com:3838/scAPAAtlas>)用于探索不同细胞类型的 APA, 并解释潜在的生物学功能。基于从 24 个人类组织和 25 个小鼠正常组织中获取的 scRNA-seq 数据, 研究人员系统地识别了不同细胞类型的细胞特异性 APA 事件, 并研究了 APA 与基因表达水平之间的相关性。同时估算了细胞类型特异性 APA 事件与 microRNA 或 RNA 结合蛋白(RNA Binding Protein, RBP)之间的相互作用[43]。

### 5.4. 3'aQTL-atlas

3'aQTL-atlas (3'UTR APA Quantitative Trait Loci, 3'aQTLs)提供了一份全面的列表 (<https://wlcob.oit.uci.edu/3aQTLatlas>), 其中包含约 149 万个与目标基因 APA 相关的单核苷酸多态性 (Single-Nucleotide Polymorphisms, SNPs), 这些 SNPs 是基于 GTEx 数据库中的 15,201 个 RNA-seq 样本。它还包括按基因/SNP 跨组织的 3'aQTL 搜索、3'aQTL 基因组浏览器、3'aQTL 方框图和全基因组关联研究 (Genome-wide association studies, GWAS)-3'aQTL 共定位事件可视化。3'aQTL-atlas 旨在将 APA 确立为一种新兴的分子表型, 以解释大部分 GWAS 风险 SNPs, 从而对 APA 的遗传基础以及人类性状和疾病中与 APA 相关的易感基因有重要的新见解[44]。

## 6. APA 与肿瘤

人类蛋白质编码基因中约 70%的 miRNA 靶标和约 11%的腺苷酸/尿苷酸富集区域位于 3'UTR 中[45]。在对整体 APA 景观的系统调查中, 发现肿瘤样本中的总体 APA 比匹配的正常样本短, 并且癌细胞系中的 APA 比肿瘤样本中的缩短更广泛[46]。3'UTR 的缩短导致 miRNA 结合位点的减少和一些 mRNA 不稳定元件的消失。在正常细胞中, 原癌基因在蛋白质编码区使用远端 PAS, 转录本通常由 miRNA 和/或 RBPs 调控[47]。随着 APA 事件的发生, 一旦细胞选择近端 PAS 产生短的 3'UTR, 就有可能消除 miRNA 和/或 RBP 的结合位点, 导致 mRNA 失去正常控制并诱导癌变[48]。

目前已经有许多关于 APA 调节因子在不同肿瘤中发挥作用的研究。比如说, Chen 等人的研究发现 PABPN1 调控 mRNA 的 APA 抑制膀胱癌进展[49]; Xiong 等人的研究发现 PABPN1 通过抑制 SGPL1 和



CREG1 的 APA 促进透明细胞肾细胞癌的进展[50]; Tan 等人的研究发现 NUDT21 缺失诱导的 MORC2APA 促进了 KIRC 癌症的发生[51]。除此之外, 还有研究发现 APA 调节因子在肿瘤的耐药机制方面也起着关键的作用。在 B 细胞祖细胞急性淋巴细胞白血病(B cell progenitor acute lymphoblastic leukemia, B-ALL)中, NUDT21 通过 APA 限制 CD19 水平, 降低 B-ALL 对嵌合抗原受体 T 细胞疗法(Chimeric Antigen Receptor T-Cell Immunotherapy, CAR-T)和双特异性 T 细胞啮合疗法 blinatumomab (靶向表面糖蛋白 CD19)的敏感性。在肺癌中, PD-L1 的剪接变体 PD-L1-vInt4 可能是癌症对抗 PD-L1 疗法产生耐药性的原因之一[52]。乳腺癌对新辅助化疗(Neoadjuvant Chemotherapy, NAC)的耐药性是由一种新型 p62 mRNA 异构体驱动的, 它能摆脱 miRNA 介导的抑制, 并导致 p62 蛋白表达增加[53]。

## 7. 总结

随着高通量测序技术的发展, 科研人员开始对 APA 的调控方式、APA 的功能以及 APA 在肿瘤中的作用有了更清晰的认识。越来越多的证据表明, APA 是基因表达的一个新的调控机制, 并且现在研究中许多计算工具和数据库被开发, 目的是用于更准确高效地检测 APA 事件。其中大多数是从标准 RNA-seq 数据中推断 PAS 的使用信息。另外一些则利用深度学习模型预测出不同生物条件下的新型 APA 事件。总之, 这些工具能够有效帮助科研人员研究全基因组 APA 图谱, 并且在 APA 调控基因表达和功能多样性的理解方面也起到了重要作用。许多疾病(包括癌症)的病理生理学过程, 都会发生广泛的 APA, APA 事件正在成为极具潜力的临床生物标志物。尽管目前的研究丰富了我们对于 APA 的认识, 但其某些功能我们依旧不够了解, 如 PAS 与不同 RBPs 的亲合力和 APA 调控的更多其他细节等。因此, 对于 APA 的调控、APA 对生物过程的影响以及 APA 在肿瘤耐药性中的作用机制的进一步研究仍是十分必要的。

## 参考文献

- [1] Mayr, C. (2019) What Are 3' UTRs Doing? *Cold Spring Harbor Perspectives in Biology*, **11**, a034728. <https://doi.org/10.1101/cshperspect.a034728>
- [2] Sandberg, R., Neilson, J.R., Sarma, A., Sharp, P.A. and Burge, C.B. (2008) Proliferating Cells Express MRNAs with Shortened 3' Untranslated Regions and Fewer MicroRNA Target Sites. *Science*, **320**, 1643-1647. <https://doi.org/10.1126/science.1155390>
- [3] Mayr, C. and Bartel, D.P. (2009) Widespread Shortening of 3'UTRs by Alternative Cleavage and Polyadenylation Activates Oncogenes in Cancer Cells. *Cell*, **138**, 673-684. <https://doi.org/10.1016/j.cell.2009.06.016>
- [4] Wang, R., Nambiar, R., Zheng, D. and Tian, B. (2018) PolyA\_DB 3 Catalogs Cleavage and Polyadenylation Sites Identified by Deep Sequencing in Multiple Genomes. *Nucleic Acids Research*, **46**, D315-D319. <https://doi.org/10.1093/nar/gkx1000>
- [5] Herrmann, C.J., Schmidt, R., Kanitz, A., Artimo, P., Gruber, A.J. and Zavolan, M. (2020) PolyASite 2.0: A Consolidated Atlas of Polyadenylation Sites from 3' End Sequencing. *Nucleic Acids Research*, **48**, D174-D179. <https://doi.org/10.1093/nar/gkz918>
- [6] Marini, F., Scherzinger, D. and Danckwardt, S. (2021) TREND-DB-A Transcriptome-Wide Atlas of the Dynamic Landscape of Alternative Polyadenylation. *Nucleic Acids Research*, **49**, D243-D253. <https://doi.org/10.1093/nar/gkaa722>
- [7] Turner, R.E., Pattison, A.D. and Beilharz, T.H. (2018) Alternative Polyadenylation in the Regulation and Dysregulation of Gene Expression. *Seminars in Cell & Developmental Biology*, **75**, 61-69. <https://doi.org/10.1016/j.semcdb.2017.08.056>
- [8] Hong, W., Ruan, H., Zhang, Z., Ye, Y., Liu, Y., Li, S., Jing, Y., Zhang, H., Diao, L., Liang, H. and Han, L. (2020) APAAtlas: Decoding Alternative Polyadenylation across Human Tissues. *Nucleic Acids Research*, **48**, D34-D39. <https://doi.org/10.1093/nar/gkz876>
- [9] Tian, B., Pan, Z. and Lee, J.Y. (2007) Widespread mRNA Polyadenylation Events in Introns Indicate Dynamic Interplay Between Polyadenylation and Splicing. *Genome Research*, **17**, 156-165. <https://doi.org/10.1101/gr.5532707>
- [10] Zhang, Y., Liu, L., Qiu, Q., Zhou, Q., Ding, J., Lu, Y. and Liu, P. (2021) Alternative Polyadenylation: Methods, Mechanism, Function, and Role in Cancer. *Journal of Experimental & Clinical Cancer Research*, **40**, Article No. 51.

- <https://doi.org/10.1186/s13046-021-01852-7>
- [11] Tian, B. and Manley, J.L. (2017) Alternative Polyadenylation of mRNA Precursors. *Nature Reviews Molecular Cell Biology*, **18**, 18-30. <https://doi.org/10.1038/nrm.2016.116>
- [12] Yuan, F., Hankey, W., Wagner, E.J., Li, W. and Wang, Q. (2021) Alternative Polyadenylation of mRNA and Its Role in Cancer. *Genes & Diseases*, **8**, 61-72. <https://doi.org/10.1016/j.gendis.2019.10.011>
- [13] Lin, J., Xu, R., Wu, X., Shen, Y. and Li, Q.Q. (2017) Role of Cleavage and Polyadenylation Specificity Factor 100: Anchoring Poly(A) Sites and Modulating Transcription Termination. *The Plant Journal*, **91**, 829-839. <https://doi.org/10.1111/tpj.13611>
- [14] Mandel, C.R., Kaneko, S., Zhang, H., Gebauer, D., Vethantham, V., Manley, J.L. and Tong, L. (2006) Polyadenylation Factor CPSF-73 Is the Pre-mRNA 3'-End-Processing Endonuclease. *Nature*, **444**, 953-956. <https://doi.org/10.1038/nature05363>
- [15] Eaton, J.D., Davidson, L., Bauer, D.L.V., Natsume, T., Kanemaki, M.T. and West, S. (2018) Xrn2 Accelerates Termination by RNA Polymerase II, Which Is Underpinned by CPSF73 Activity. *Genes & Development*, **32**, 127-139. <https://doi.org/10.1101/gad.308528.117>
- [16] Chan, S.L., Huppertz, I., Yao, C., Weng, L., Moresco, J.J., Yates III, J.R., Ule, J., Manley, J.L. and Shi, Y. (2014) CPSF30 and Wdr33 Directly Bind to AAUAAA in Mammalian mRNA 3' Processing. *Genes & Development*, **28**, 2370-2380. <https://doi.org/10.1101/gad.250993.114>
- [17] Schönemann, L., Kühn, U., Martin, G., Schäfer, P., Gruber, A.R., Keller, W., Zavolan, M. and Wahle, E. (2014) Reconstitution of CPSF Active in Polyadenylation: Recognition of the Polyadenylation Signal by WDR33. *Genes & Development*, **28**, 2381-2393. <https://doi.org/10.1101/gad.250985.114>
- [18] Yang, W., Hsu, P.L., Yang, F., Song, J.E. and Varani, G. (2018) Reconstitution of the CstF Complex Unveils a Regulatory Role for CstF-50 in Recognition of 3'-End Processing Signals. *Nucleic Acids Research*, **46**, 493-503. <https://doi.org/10.1093/nar/gkx1177>
- [19] Fonseca, D., Baquero, J., Murphy, M.R., Aruggoda, G., Varriano, S., Sapienza, C., Mashadova, O., Rahman, S. and Kleiman, F.E. (2018) mRNA Processing Factor CstF-50 and Ubiquitin Escort Factor P97 Are BRCA1/BARD1 Co-factors Involved in Chromatin Remodeling During the DNA Damage Response. *Molecular and Cellular Biology*, **38**, e00364-17. <https://doi.org/10.1128/MCB.00364-17>
- [20] Hwang, H.W., Park, C.Y., Goodarzi, H., Fak, J.J., Mele, A., Moore, M.J., Saito, Y. and Darnell, R.B. (2016) PAPERCLIP Identifies MicroRNA Targets and a Role of CstF64/64tau in Promoting Non-Canonical Poly(A) Site Usage. *Cell Reports*, **15**, 423-435. <https://doi.org/10.1016/j.celrep.2016.03.023>
- [21] Takagaki, Y., Seipelt, R.L., Peterson, M.L. and Manley, J.L. (1996) The Polyadenylation Factor CstF-64 Regulates Alternative Processing of IgM Heavy Chain Pre-mRNA During B Cell Differentiation. *Cell*, **87**, 941-952. [https://doi.org/10.1016/S0092-8674\(00\)82000-0](https://doi.org/10.1016/S0092-8674(00)82000-0)
- [22] Hockert, J.A., Yeh, H.J. and MacDonald, C.C. (2010) The Hinge Domain of the Cleavage Stimulation Factor Protein CstF-64 Is Essential for CstF-77 Interaction, Nuclear Localization, and Polyadenylation. *Journal of Biological Chemistry*, **285**, 695-704. <https://doi.org/10.1074/jbc.M109.061705>
- [23] Grozdanov, P.N., Masoumzadeh, E., Latham, M.P. and MacDonald, C.C. (2018) The Structural Basis of CstF-77 Modulation of Cleavage and Polyadenylation through Stimulation of CstF-64 Activity. *Nucleic Acids Research*, **46**, 12022-12039. <https://doi.org/10.1093/nar/gky862>
- [24] Rügsegger, U., Blank, D. and Keller, W. (1998) Human Pre-mRNA Cleavage Factor Im Is Related to Spliceosomal SR Proteins and Can Be Reconstituted *in Vitro* from Recombinant Subunits. *Molecular Cell*, **1**, 243-253. [https://doi.org/10.1016/S1097-2765\(00\)80025-8](https://doi.org/10.1016/S1097-2765(00)80025-8)
- [25] Rügsegger, U., Beyer, K. and Keller, W. (1996) Purification and Characterization of Human Cleavage Factor Im Involved in the 3' End Processing of Messenger RNA Precursors. *Journal of Biological Chemistry*, **271**, 6107-6113. <https://doi.org/10.1074/jbc.271.11.6107>
- [26] Martin, G., Gruber, A.R., Keller, W. and Zavolan, M. (2012) Genome-Wide Analysis of Pre-mRNA 3' End Processing Reveals a Decisive Role of Human Cleavage Factor I in the Regulation of 3' UTR Length. *Cell Reports*, **1**, 753-763. <https://doi.org/10.1016/j.celrep.2012.05.003>
- [27] Gruber, A.R., Martin, G., Keller, W. and Zavolan, M. (2012) Cleavage Factor Im Is a Key Regulator of 3' UTR Length. *RNA Biology*, **9**, 1405-1412. <https://doi.org/10.4161/rna.22570>
- [28] Gruber, A.J. and Zavolan, M. (2019) Alternative Cleavage and Polyadenylation in Health and Disease. *Nature Reviews Genetics*, **20**, 599-614. <https://doi.org/10.1038/s41576-019-0145-z>
- [29] Schäfer, P., Tüting, C., Schönemann, L., Kühn, U., Treiber, T., Treiber, N., Ihling, C., Graber, A., Keller, W., Meister, G., Sinz, A. and Wahle, E. (2018) Reconstitution of Mammalian Cleavage Factor II Involved in 3' Processing of mRNA Precursors. *RNA*, **24**, 1721-1737. <https://doi.org/10.1261/rna.068056.118>

- [30] Stroup, E.K. and Ji, Z. (2023) Deep Learning of Human Polyadenylation Sites at Nucleotide Resolution Reveals Molecular Determinants of Site Usage and Relevance in Disease. *Nature Communications*, **14**, Article No. 7378. <https://doi.org/10.1038/s41467-023-43266-3>
- [31] Ji, G., Tang, Q., Zhu, S., Zhu, J., Ye, P., Xia, S. and Wu, X. (2023) StAPaminer: Mining Spatial Patterns of Alternative Polyadenylation for Spatially Resolved Transcriptomic Studies. *Genomics, Proteomics & Bioinformatics*, **21**, 601-618. <https://doi.org/10.1016/j.gpb.2023.01.003>
- [32] Imada, E.L., Wilks, C., Langmead, B. and Marchionni, L. (2023) REPAC: Analysis of Alternative Polyadenylation from RNA-Sequencing Data. *Genome Biology*, **24**, Article No. 22. <https://doi.org/10.1186/s13059-023-02865-5>
- [33] Wang, G., Xie, Z., Su, J., Chen, M., Du, Y., Gao, Q., Zhang, G., Zhang, H., Chen, X., Liu, H., Han, L. and Ye, Y. (2022) Characterization of Immune-Related Alternative Polyadenylation Events in Cancer Immunotherapy. *Cancer Research*, **82**, 3474-3485. <https://doi.org/10.1158/0008-5472.CAN-22-1417>
- [34] Linder, J., Koplik, S.E., Kundaje, A. and Seelig, G. (2022) Deciphering the Impact of Genetic Variation on Human Polyadenylation Using APARENT2. *Genome Biology*, **23**, Article No. 232. <https://doi.org/10.1186/s13059-022-02799-4>
- [35] Li, Z., Li, Y., Zhang, B., Li, Y., Long, Y., Zhou, J., Zou, X., Zhang, M., Hu, Y., Chen, W. and Gao, X. (2022) DeeReCT-APA: Prediction of Alternative Polyadenylation Site Usage Through Deep Learning. *Genomics, Proteomics & Bioinformatics*, **20**, 483-495. <https://doi.org/10.1016/j.gpb.2020.05.004>
- [36] Ye, C., Zhao, D., Ye, W., Wu, X., Ji, G., Li, Q.Q. and Lin, J. (2021) QuantifyPoly(A): Reshaping Alternative Polyadenylation Landscapes of Eukaryotes with Weighted Density Peak Clustering. *Briefings in Bioinformatics*, **22**, bbab268. <https://doi.org/10.1093/bib/bbab268>
- [37] Lusk, R., Stene, E., Banaei-Kashani, F., Tabakoff, B., Kechris, K. and Saba, L.M. (2021) Aptardi Predicts Polyadenylation Sites in Sample-Specific Transcriptomes Using High-Throughput RNA Sequencing and DNA Sequence. *Nature Communications*, **12**, Article No. 1652. <https://doi.org/10.1038/s41467-021-21894-x>
- [38] Lin, E., Liu, X., Liu, Y., Zhang, Z., Xie, L., Tian, K., Liu, J. and Yu, Y. (2021) Roles of the Dynamic Tumor Immune Microenvironment in the Individualized Treatment of Advanced Clear Cell Renal Cell Carcinoma. *Frontiers in Immunology*, **12**, Article 653358. <https://doi.org/10.3389/fimmu.2021.653358>
- [39] Li, G.W., Nan, F., Yuan, G.H., Liu, C.X., Liu, X., Chen, L.L., Tian, B. and Yang, L. (2021) SCAPTURE: A Deep Learning-Embedded Pipeline That Captures Polyadenylation Information from 3' Tag-Based RNA-Seq of Single Cells. *Genome Biology*, **22**, Article No. 221. <https://doi.org/10.1186/s13059-021-02437-5>
- [40] Gao, Y., Li, L., Amos, C.I. and Li, W. (2021) Analysis of Alternative Polyadenylation from Single-Cell RNA-Seq Using ScDaPars Reveals Cell Subpopulations Invisible to Gene Expression. *Genome Research*, **31**, 1856-1866. <https://doi.org/10.1101/gr.271346.120>
- [41] Ma, X., Cheng, S., Ding, R., Zhao, Z., Zou, X., Guang, S., Wang, Q., Jing, H., Yu, C., Ni, T. and Li, L. (2023) IpaQTL-Atlas: An Atlas of Intronic Polyadenylation Quantitative Trait Loci across Human Tissues. *Nucleic Acids Research*, **51**, D1046-D1052. <https://doi.org/10.1093/nar/gkac736>
- [42] Zhu, S., Lian, Q., Ye, W., Qin, W., Wu, Z., Ji, G. and Wu, X. (2022) ScAPAdb: A Comprehensive Database of Alternative Polyadenylation at Single-Cell Resolution. *Nucleic Acids Research*, **50**, D365-D370. <https://doi.org/10.1093/nar/gkab795>
- [43] Yang, X., Tong, Y., Liu, G., Yuan, J. and Yang, Y. (2022) ScAPAAtlas: An Atlas of Alternative Polyadenylation across Cell Types in Human and Mouse. *Nucleic Acids Research*, **50**, D356-D364. <https://doi.org/10.1093/nar/gkab917>
- [44] Cui, Y., Peng, F., Wang, D., Li, Y., Li, J.S., Li, L. and Li, W. (2022) 3'aQTL-Atlas: An Atlas of 3'UTR Alternative Polyadenylation Quantitative Trait Loci across Human Normal Tissues. *Nucleic Acids Research*, **50**, D39-D45. <https://doi.org/10.1093/nar/gkab740>
- [45] Lin, Y., Li, Z., Ozsolak, F., Kim, S.W., Arango-Argoty, G., Liu, T.T., Tenenbaum, S.A., Bailey, T., Monaghan, A.P., Milos, P.M. and John, B. (2012) An In-Depth Map of Polyadenylation Sites in Cancer. *Nucleic Acids Research*, **40**, 8460-8471. <https://doi.org/10.1093/nar/gks637>
- [46] Xiang, Y., Ye, Y., Lou, Y., Yang, Y., Cai, C., Zhang, Z., Mills, T., Chen, N.Y., Kim, Y., Muge Ozguc, F., Diao, L., Karmouty-Quintana, H., Xia, Y., Kellems, R.E., Chen, Z., Blackburn, M.R., Yoo, S.H., Shyu, A.B., Mills, G.B. and Han, L. (2018) Comprehensive Characterization of Alternative Polyadenylation in Human Cancer. *Journal of the National Cancer Institute*, **110**, 379-389. <https://doi.org/10.1093/jnci/djx223>
- [47] Di Giammartino, D.C., Li, W., Ogami, K., Yashinski, J.J., Hoque, M., Tian, B. and Manley, J.L. (2014) RBBP6 Isoforms Regulate the Human Polyadenylation Machinery and Modulate Expression of MRNAs with AU-Rich 3' UTRs. *Genes & Development*, **28**, 2248-2260. <https://doi.org/10.1101/gad.245787.114>
- [48] Fu, Y., Chen, L., Chen, C., Ge, Y., Kang, M., Song, Z., Li, J., Feng, Y., Huo, Z., He, G., Hou, M., Chen, S. and Xu, A. (2018) Crosstalk between Alternative Polyadenylation and MiRNAs in the Regulation of Protein Translational Effi-



- ciency. *Genome Research*, **28**, 1656-1663. <https://doi.org/10.1101/gr.231506.117>
- [49] Chen, L., Dong, W., Zhou, M., Yang, C., Xiong, M., Kazobinka, G., Chen, Z., Xing, Y. and Hou, T. (2023) PABPN1 Regulates mRNA Alternative Polyadenylation to Inhibit Bladder Cancer Progression. *Cell & Bioscience*, **13**, Article No. 45. <https://doi.org/10.1186/s13578-023-00997-6>
- [50] Xiong, M., Liu, C., Li, W., Jiang, H., Long, W., Zhou, M., Yang, C., Kazobinka, G., Sun, Y., Zhao, J. and Hou, T. (2023) PABPN1 Promotes Clear Cell Renal Cell Carcinoma Progression by Suppressing the Alternative Polyadenylation of SGPL1 and CREG1. *Carcinogenesis*, **44**, 576-586. <https://doi.org/10.1093/carcin/bgad049>
- [51] Tan, Y., Zheng, T., Su, Z., Chen, M., Chen, S., Zhang, R., Wang, R., Li, K. and Na, N. (2023) Alternative Polyadenylation Reprogramming of MORC2 Induced by NUDT21 Loss Promotes KIRC Carcinogenesis. *JCI Insight*, **8**, e162893. <https://doi.org/10.1172/jci.insight.162893>
- [52] Witkowski, M.T., Lee, S., Wang, E., Lee, A.K., Talbot, A., Ma, C., Tsopoulidis, N., Brumbaugh, J., Zhao, Y., Roberts, K.G., Hogg, S.J., Nomikou, S., Ghebrechristos, Y.E., Thandapani, P., Mullighan, C.G., Hochedlinger, K., Chen, W., Abdel-Wahab, O., Eyquem, J. and Aifantis, I. (2022) NUDT21 Limits CD19 Levels through Alternative MRNA Polyadenylation in B Cell Acute Lymphoblastic Leukemia. *Nature Immunology*, **23**, 1424-1432. <https://doi.org/10.1038/s41590-022-01314-y>
- [53] Guo, Q., Wang, H., Duan, J., Luo, W., Zhao, R., Shen, Y., Wang, B., Tao, S., Sun, Y., Ye, Q., Bi, X., Yuan, H., Wu, Q., Lobie, P.E., Zhu, T., Tan, S., Huang, X. and Wu, Z. (2022) An Alternatively Spliced P62 Isoform Confers Resistance to Chemotherapy in Breast Cancer. *Cancer Research*, **82**, 4001-4015. <https://doi.org/10.1158/0008-5472.CAN-22-0909>