

**Cell Reports Methods, Volume 4**

**Supplemental information**

**PolyAMiner-Bulk is a deep learning-based algorithm  
that decodes alternative polyadenylation dynamics  
from bulk RNA-seq data**

**Venkata Soumith Jonnakuti, Eric J. Wagner, Mirjana Maletić-Savatić, Zhandong Liu, and Hari Krishna Yalamanchili**

## SUPPLEMENTARY TABLE AND FIGURES LEGENDS

**Supplementary Figure S1: Current-generation computational approaches for quantifying 3'UTR length changes from bulk RNA-seq data, related to Introduction, related to Figure 6.**

**Supplementary Figure S2: Representative differential APA gene, DEF8, with a negative PolyAIndex, suggesting 3'UTR shortening, related to Figure 2.** Visualizing the DEF8 read density – as bulk RNA-seq and pseudo-3'UTR-seq read coverage – and plotting its corresponding density proportions as a heatmap reveal that, in the control condition, DEF8 exhibits higher read proportion density in its more distal 3'UTR C/PAS (C/PAS\_2), whereas DEF8 shifts a proportion of its read density towards the more proximal 3'UTR C/PAS (C/PAS\_1) in the knockdown condition.

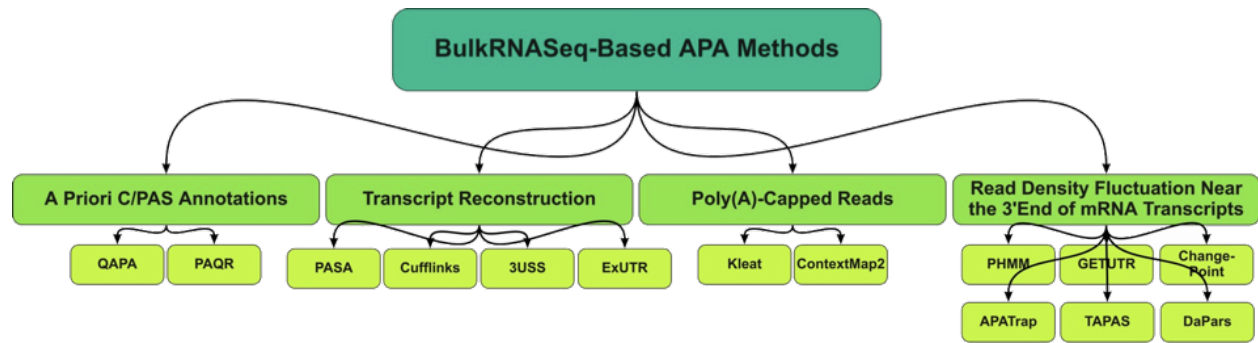
**Supplementary Figure S3: Heatmaps showcasing changes in read density proportions of DAGs containing either 2 or 3 C/PASs, related to Figure 5.** (A) Changes in read density proportions of DAGs containing 2 C/PASs between the control and RBM17 knockdown conditions at a genome-wide scale. (B) Changes in read density proportions of DAGs containing 3 C/PASs between the control and RBM17 knockdown conditions at a genome-wide scale.

**Supplementary Figure S4: Representative differential APA gene comparisons between PolyAMiner-Bulk and other current generation bulk RNA-seq-based APA methods, related to Figure 6.** (A) Representative differential APA gene, CACNG7, identified by PolyAMiner-Bulk but not APAlzyer or DaPars. Compared to the control condition, CACNG7 undergoes 3'UTR shortening in the RBM17 knockdown condition. In the control condition, CACNG7 exhibits higher read proportion density in its distal 3'UTR C/PAS, whereas CACNG7 shifts a proportion of its read density towards the proximal 3'UTR C/PAS in the RBM17 knockdown condition. (B) Representative differential APA gene, ANK1, identified by APAlzyer but not PolyAMiner-Bulk. (C) Representative differential APA gene, ACP1, identified by DaPars but not PolyAMiner-Bulk. PolyAMiner-Bulk does not classify ANK1 nor ACP1 as differential APA genes since the three

samples for each condition do not uniformly undergo changes in read density among the C/PASs. The corresponding read density and heatmap visualizations further corroborate this result.

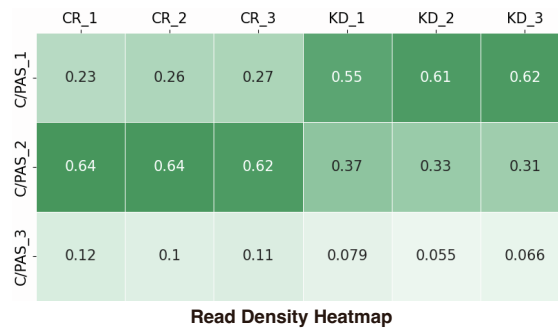
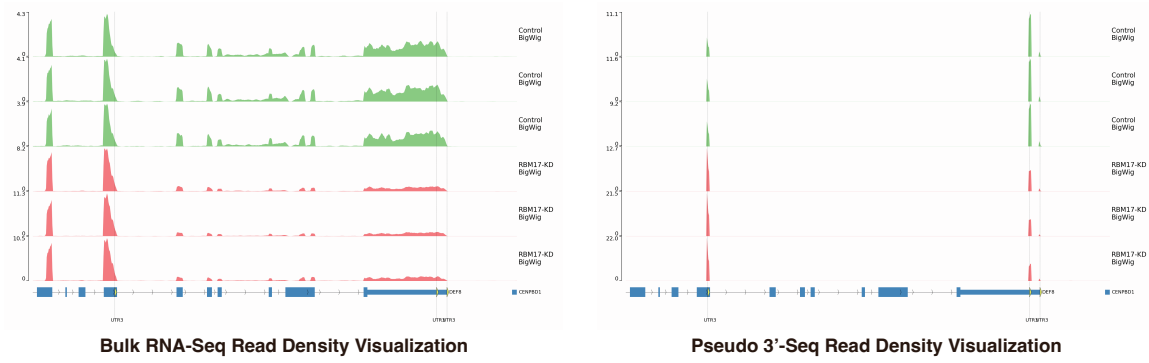
**Supplementary Table S1: Feature comparison chart of PolyAMiner-Bulk and current-generation computational approaches quantifying 3'UTR length changes from bulk RNA-seq data, related to Figure 6.**

SUPPLEMENTARY FIGURE S1

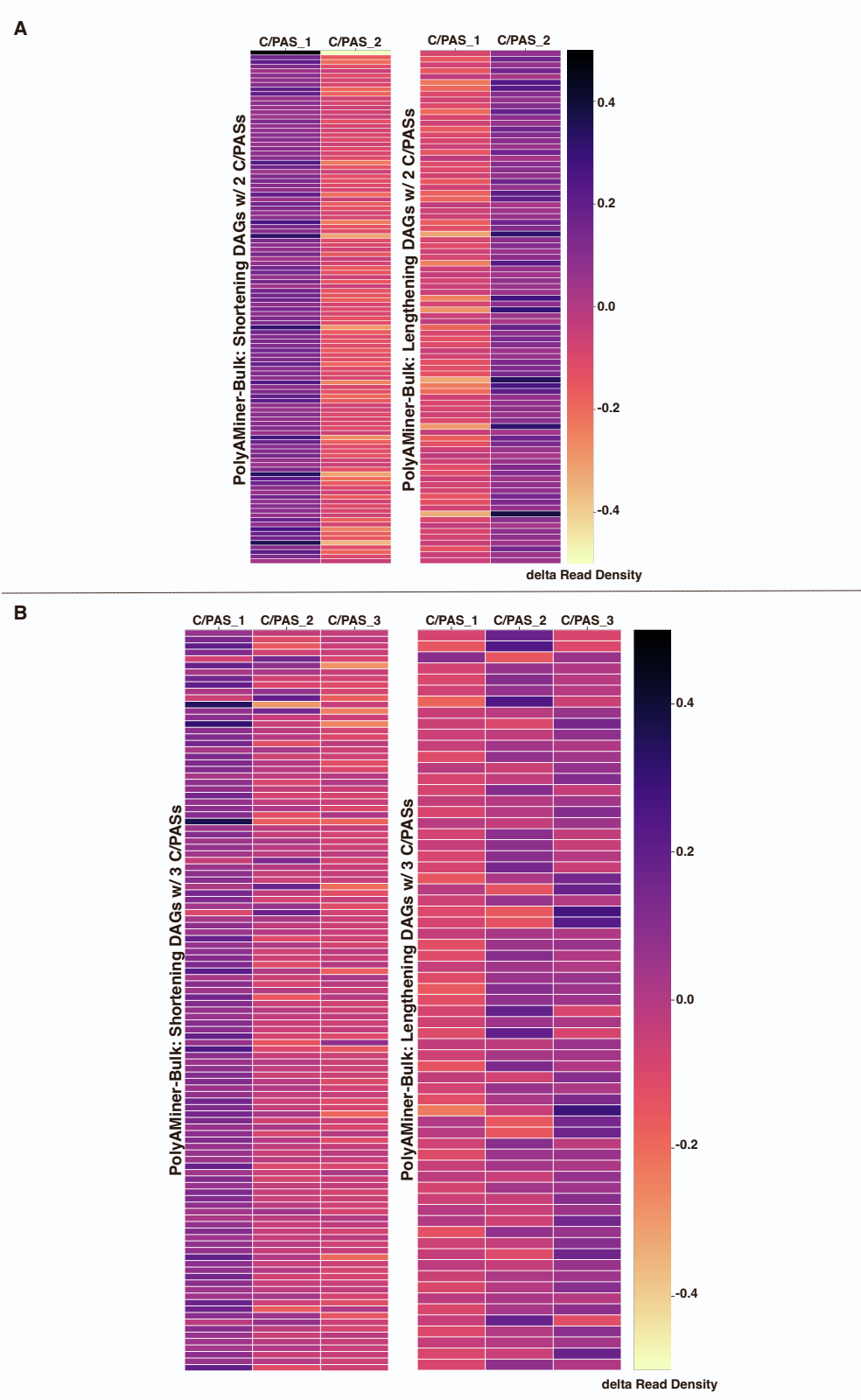


## SUPPLEMENTARY FIGURE S2

### DEF8 = Representative Differential APA Gene (DAG)

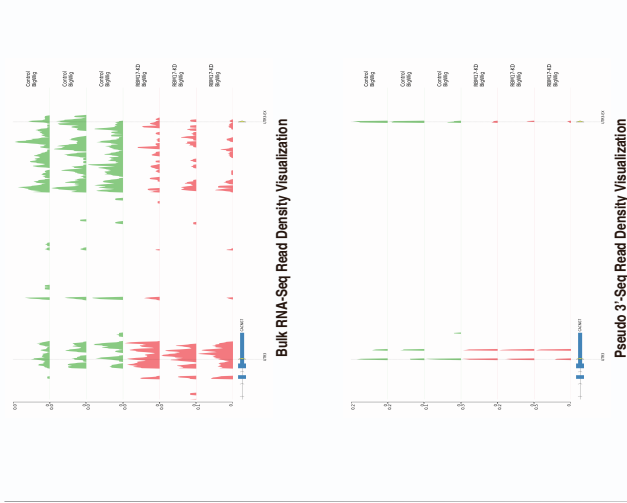


SUPPLEMENTARY FIGURE S3

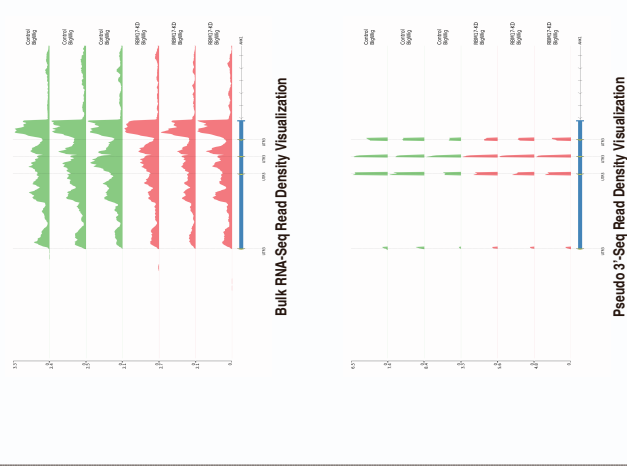


# SUPPLEMENTARY FIGURE S4

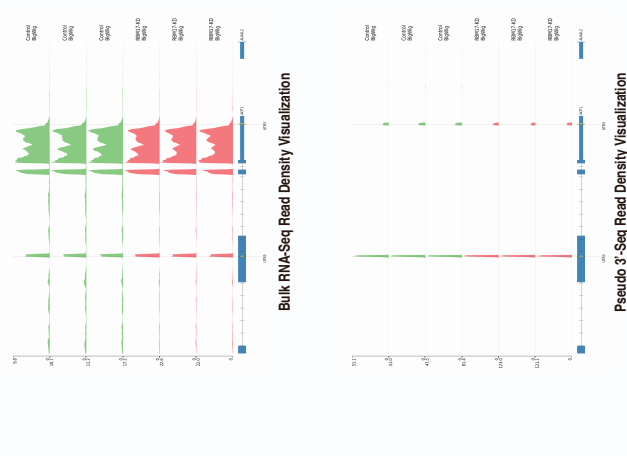
**A** CACNG7 = DAG that is identified by PolyAminer-Bulk but not APAlyzer or DaPars



**B** ANK1 = DAG that is identified by APAlyzer but not PolyAminer-Bulk



**C** ACP1 = DAG that is identified by DaPars but not PolyAminer-Bulk



## Supplementary Table S1

Feature	PolyAMiner-Bulk	APalyzer	APAtrap	DaPars	QAPA	Roar	TAPAS
<i>De novo</i> C/PAS identification	YES	NO	YES	YES	NO	NO	YES
Reference Database	PolyASite & PolyA_DB	PolyA_DB only	N/A	N/A	PolyASite/GENCODE	PolyA_DB & APASdb	N/A
Deep Learning Model	YES	NO	NO	NO	NO	NO	NO
Intra-distal and intra-proximal APA quantification	YES	NO	NO	NO	NO	NO	NO
3'UTR APA	YES	YES	YES	YES	YES	YES	YES
IPA	YES	YES	NO	NO	NO	NO	NO
Visualization (Volcano Plot)	YES	YES	NO	NO	NO	NO	NO
Visualization (IGV Read Density)	YES	NO	NO	NO	NO	NO	NO
Visualization (Heatmap)	YES	NO	NO	NO	NO	NO	NO
Reference	This study	(Wang, et al., 2020)	(Ye, et al., 2018)	(Xia, et al., 2014)	(Ha, et al., 2018)	(Grassi, et al., 2016)	(Arefeen, et al., 2018)

### Program download site:

PolyAMiner-Bulk: <https://github.com/YalamanchiliLab/PolyAMiner-Bulk>

APalyzer: <https://bioconductor.org/packages/release/bioc/html/APalyzer.html>

APAtrap: <https://sourceforge.net/projects/apatrap/>

DaPars: <https://github.com/ZhengXia/dapars>

QAPA: <https://github.com/morrislab/qapa>

ROAR: <https://bioconductor.org/packages/release/bioc/html/roar.html>

TAPAS: <https://github.com/arefeen/TAPAS>

### References:

Wang, R., et al. APalyzer: a bioinformatics package for analysis of alternative polyadenylation isoforms. *Bioinformatics* 2020; 36(12):3907-3309

Arefeen, A., et al. TAPAS: tool for alternative polyadenylation site analysis. *Bioinformatics* 2018;34(15):2521-2529.

Grassi, E., et al. Roar: detecting alternative polyadenylation with standard mRNA sequencing libraries. *BMC bioinformatics* 2016;17(1):423.

Ha, K.C., Blencowe, B.J. and Morris, Q. QAPA: a new method for the systematic analysis of alternative polyadenylation from RNA-seq data. *Genome biology* 2018;19(1):45.

Xia, Z., et al. Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3'-UTR landscape across seven tumour types. *Nature communications* 2014;5(1):1-13.

Ye, C., et al. APAtrap: identification and quantification of alternative polyadenylation sites from RNA-seq data. *Bioinformatics* 2018;34(11):1841-1849.