Direct RNA sequencing reveals multi-intron splicing order and poly(A) tail lengths across subcellular compartments

Karine Choquet[1], Brendan M. Smalec[1], Robert Ietswaart[1], L. Stirling Churchman[1]

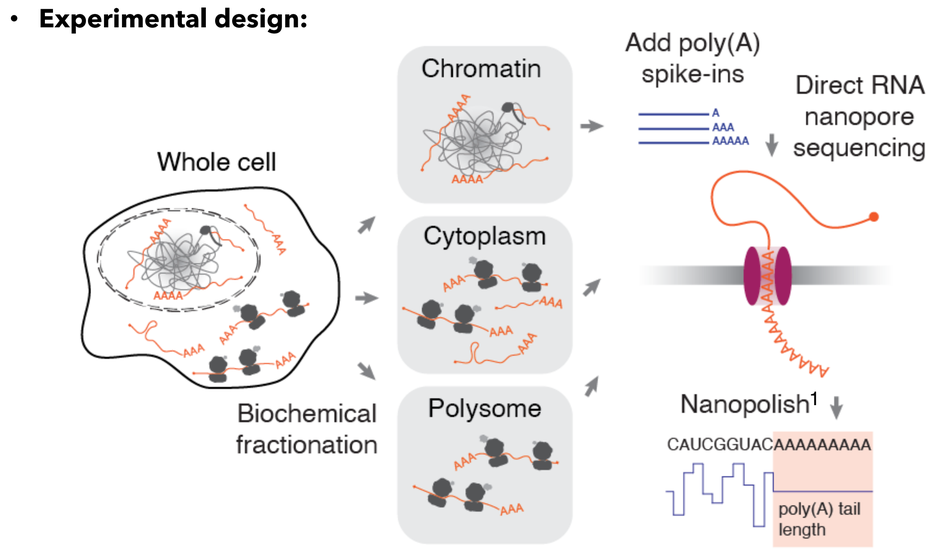
[1]Department of Genetics, Blavatnik Institute, Harvard Medical School, Boston, MA, USA; Contact: karine_choquet@hms.harvard.edu

Introduction

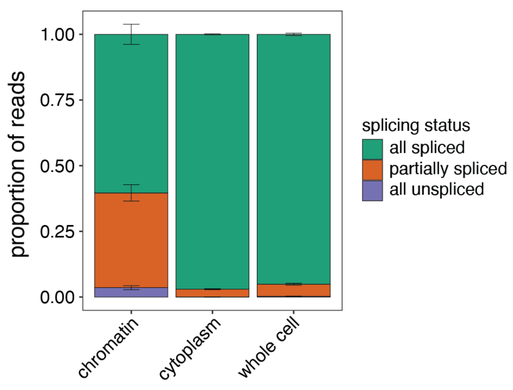
- Newly synthesized pre-mRNAs undergo **several processing steps on chromatin** prior to nuclear export. These include:
  - **Splicing** of the many long introns (co- or post-transcriptionally)
  - Addition of **poly(A) tails** (~200-250 A's)
- In the **cytoplasm, poly(A) tails are shortened** during the mRNA lifetime
- Several **outstanding questions** can be explored using **direct RNA nanopore sequencing:**
  - What is the **order of intron removal** across multi-intronic transcripts?
  - How **variable are poly(A) tail lengths** across subcellular compartments?
  - What is the relationship between **poly(A) tail length and the time an RNA spends** in each subcellular compartment?

Direct nanopore sequencing of subcellular RNA

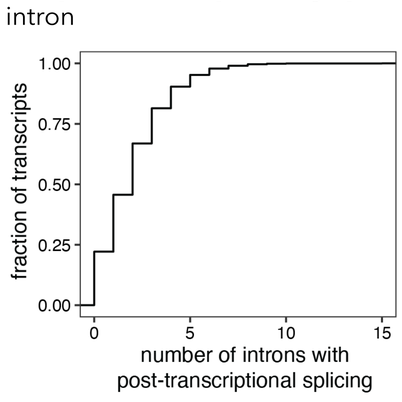
- **Experimental design:**



- **Chromatin-associated RNA** is enriched for **partially spliced RNAs**

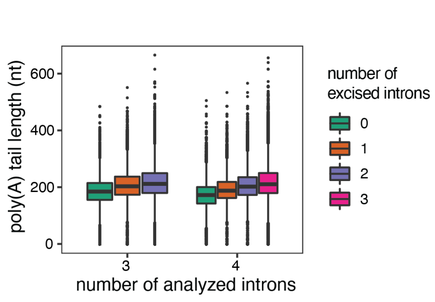


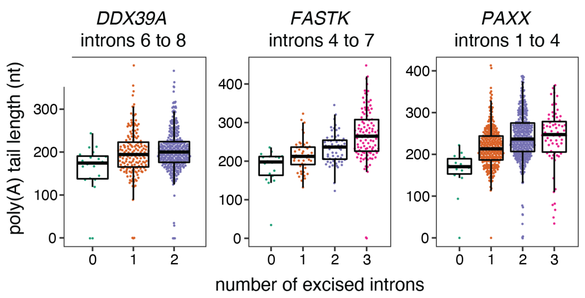
- **>75% of transcripts** contain at least one **post-transcriptionally spliced** intron



Splicing and polyadenylation progress in parallel

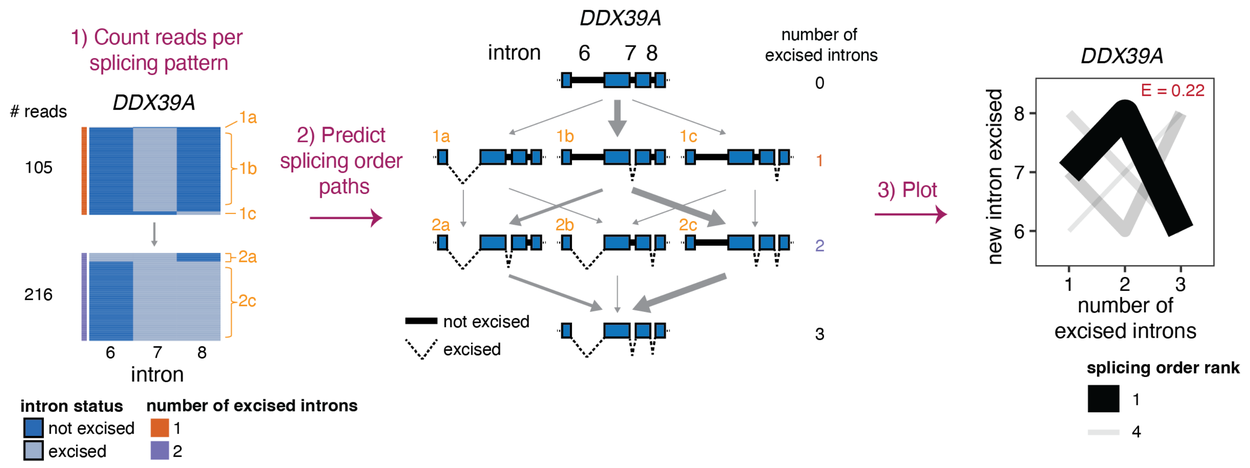
- **Chromatin poly(A) tails are longer** on intermediate isoforms that have **more excised introns**



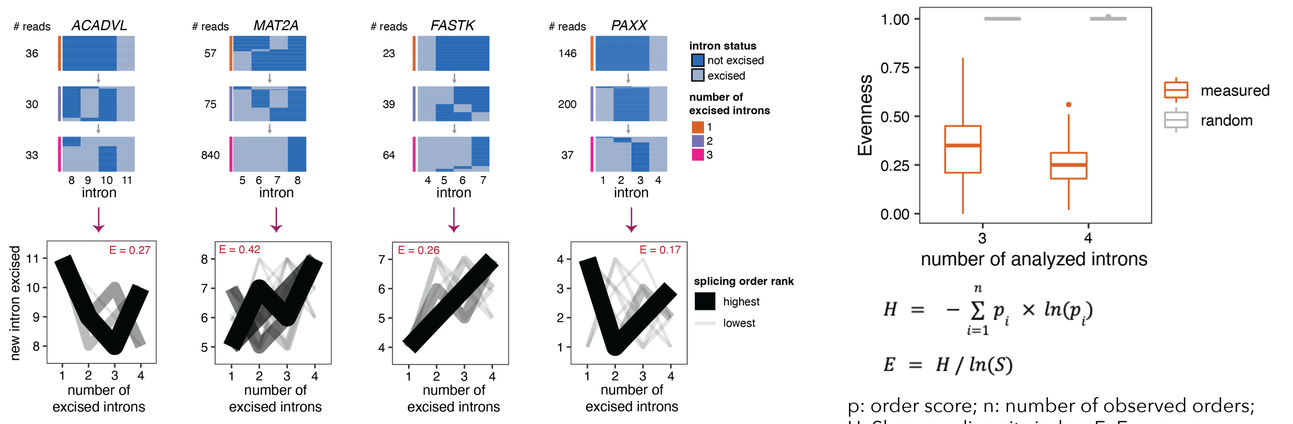


Multi-intron splicing order is defined in human cells

- Multi-intron **splicing orders** are computed based on the frequencies of different **intermediate isoforms** (partially spliced reads) for **groups of 3-4 proximal introns**



- Most intron groups follow only **1-2 predominant splicing orders** and have a low evenness (E)



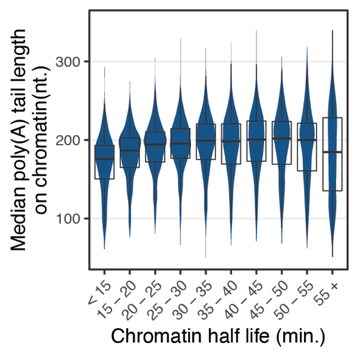
$$H = -\sum_{i=1}^{n} p_i \times ln(p_i)$$

$$E = H / ln(S)$$

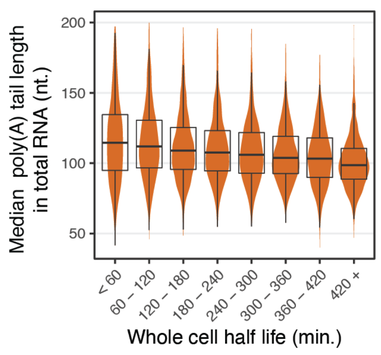
p: order score; n: number of observed orders;
H: Shannon diversity index; E: Evenness;
S: total number of possible orders

Poly(A) tails reflect the time spent in different subcellular compartments

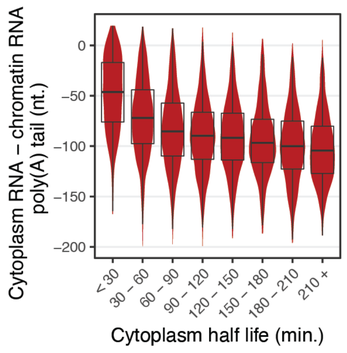
- Subcellular TimeLapse-seq provides **subcellular compartment-specific RNA half-lives**[2]
- Poly(A) tail lengths **increase with chromatin half-life**

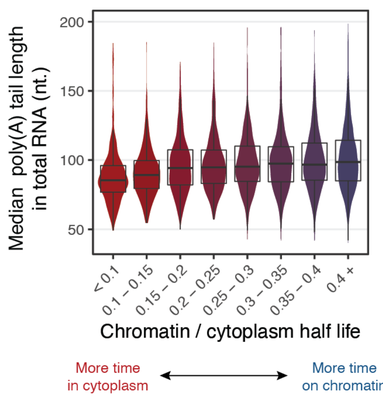


- Poly(A) tail lengths **decrease** as **whole cell and cytoplasmic** (not shown) **half-lives increase**

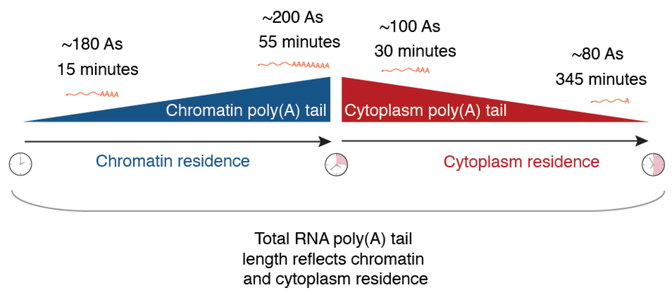


- mRNAs are **not all deadenylated to the same extent** before degradation



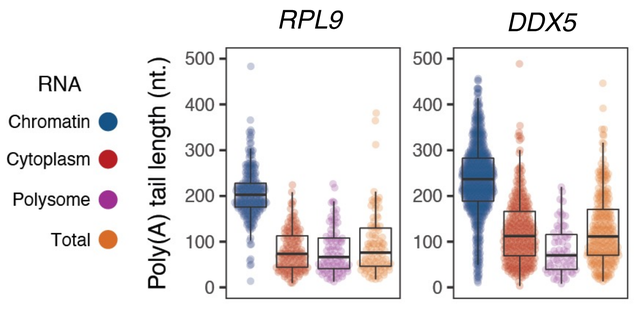


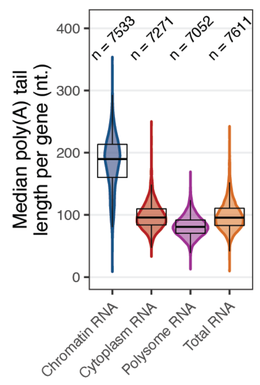
- **Whole cell tail length** reflects the **relative time** spent in each subcellular compartment

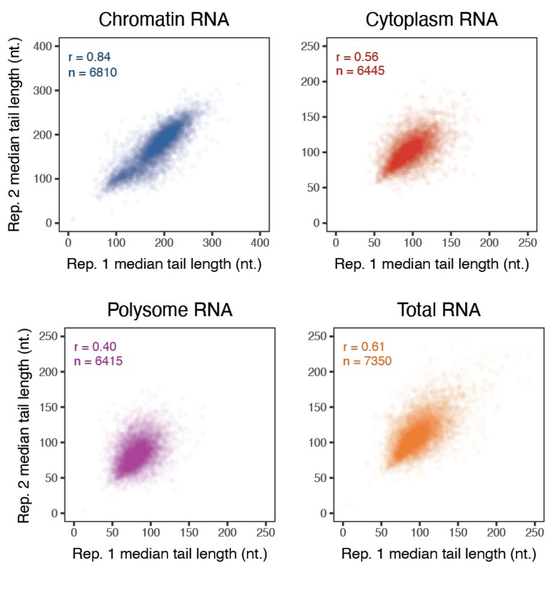


Poly(A) tails are dynamic across and within subcellular compartments

- **Chromatin: longest poly(A) tails** and best correlation between replicates
- **Polysome: shortest poly(A) tails** and lowest correlation between replicates
- **Whole cell and cytoplasm** have very **similar** poly(A) tail lengths
- **Heterogeneity** between and within genes in each subcellular compartment







Conclusions

- **Subcellular direct RNA nanopore sequencing** reveals the dynamics of **splicing, polyadenylation and deadenylation** as RNAs flow through the cell.
- On chromatin, multi-intron **splicing order converges on a few predominant paths.**
- Analysis of poly(A) tail lengths reveals **extensive heterogeneity** between genes and subcellular compartments.
- Poly(A) tails can be used as a proxy for **the time spent by an RNA** in each subcellular compartment.

Acknowledgements

This work was supported by the NIH (R01-HG010538, R01-GM136794 and R21-HG011682 to L.S.C.). K.C. received post-doctoral fellowships from the Fonds de Recherche du Québec - Santé and the Canadian Institutes of Health Research.

References

1) Workman *et al.* (2019) *Nat Methods* 16, 1297-1305.
2) Smalec, Ietswaart *et al.* bioRxiv (2022), https://doi.org/10.1101/2022.08.21.504696.