THE 3rd VIETNAM SCHOOL OF BIOLOGY (VSOB-3)

Bioinformatic Analysis For Bulk RNAseq Data

December 06th-08th, 2024, ICISE, Quy Nhon, Vietnam

Fundamental concepts of RNA sequencing experiments

NGUYEN Thuy Vy

Why do we care about RNA?

Genomics has revolutionized the way we see and understand organisms, but it is only the starting point.

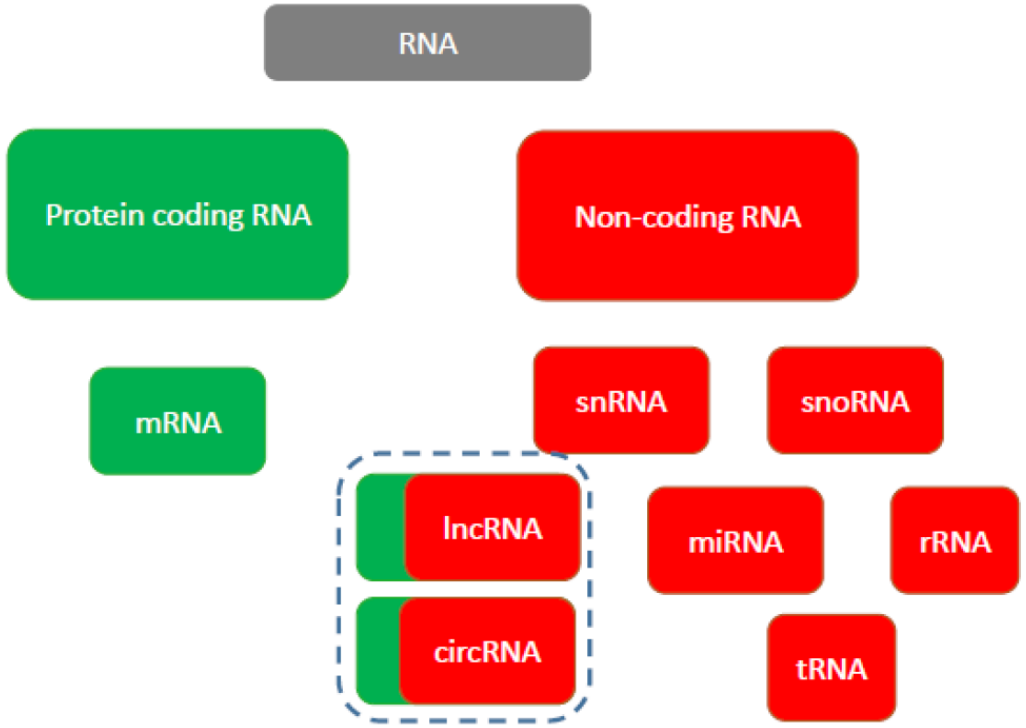
Just from genome sequence, we cannot know for sure how, when and where an organism expresses a gene, if at all.

Proteins are the effectors that carry out most of the molecular processes in the cell…

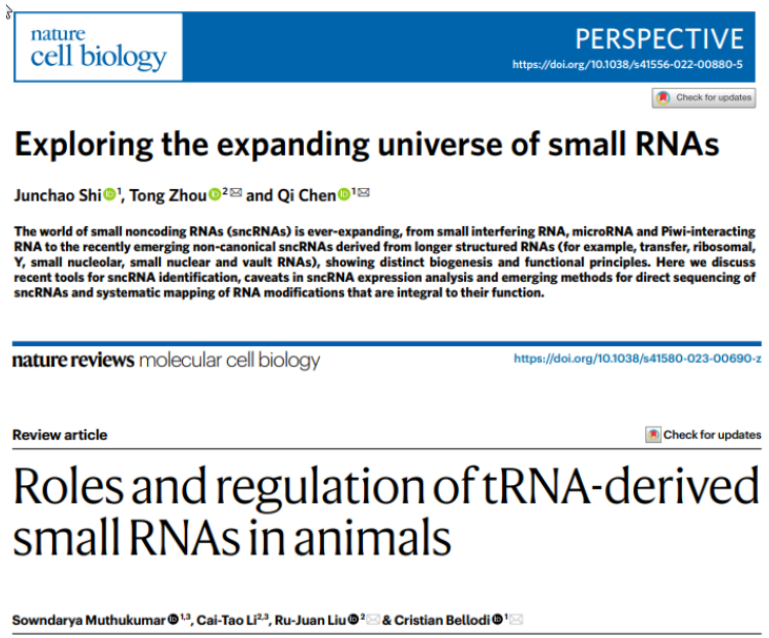
But proteomics is still more expensive than transcriptomics

And it is sequence dependent, whereas transcriptomics provides the sequence by itself

The Growing RNA World



for this



nature cell biology

PERSPECTIVE

https://doi.org/10.1038/s41556-022-00880-5

Check for updates

Exploring the expanding universe of small RNAs

Junchao Shi[1], Tong Zhou[2] ✉ and Qi Chen[1] ✉

The world of small noncoding RNAs (sncRNAs) is ever-expanding, from small interfering RNA, microRNA and Piwi-interacting RNA to the recently emerging non-canonical sncRNAs derived from longer structured RNAs (for example, transfer, ribosomal, Y, small nucleolar, small nuclear and vault RNAs), showing distinct biogenesis and functional principles. Here we discuss recent tools for sncRNA identification, caveats in sncRNA expression analysis and emerging methods for direct sequencing of sncRNAs and systematic mapping of RNA modifications that are integral to their function.

nature reviews molecular cell biology

https://doi.org/10.1038/s41580-023-00690-z

Review article

Check for updates

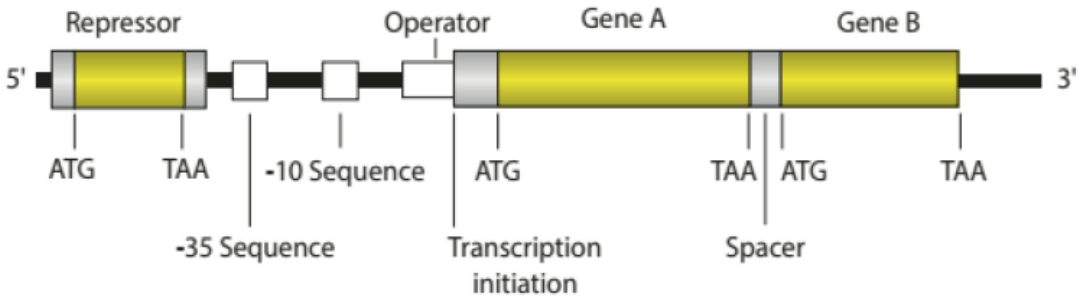
Roles and regulation of tRNA-derived small RNAs in animals

Sowndarya Muthukumar[1,3], Cai-Tao Li[2,3], Ru-Juan Liu[2] ✉ & Cristian Bellodi[1] ✉

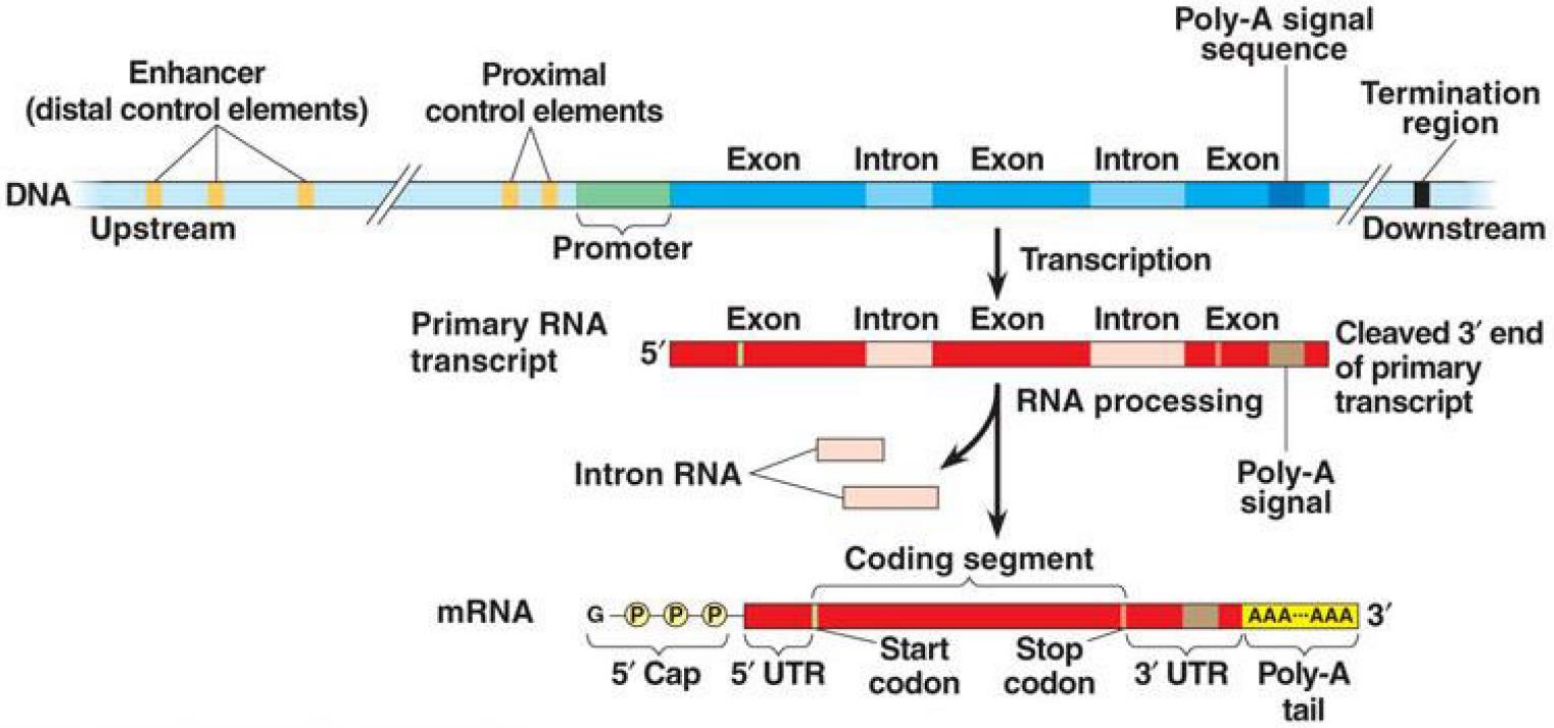
GeneCore EMBL

Gene structure

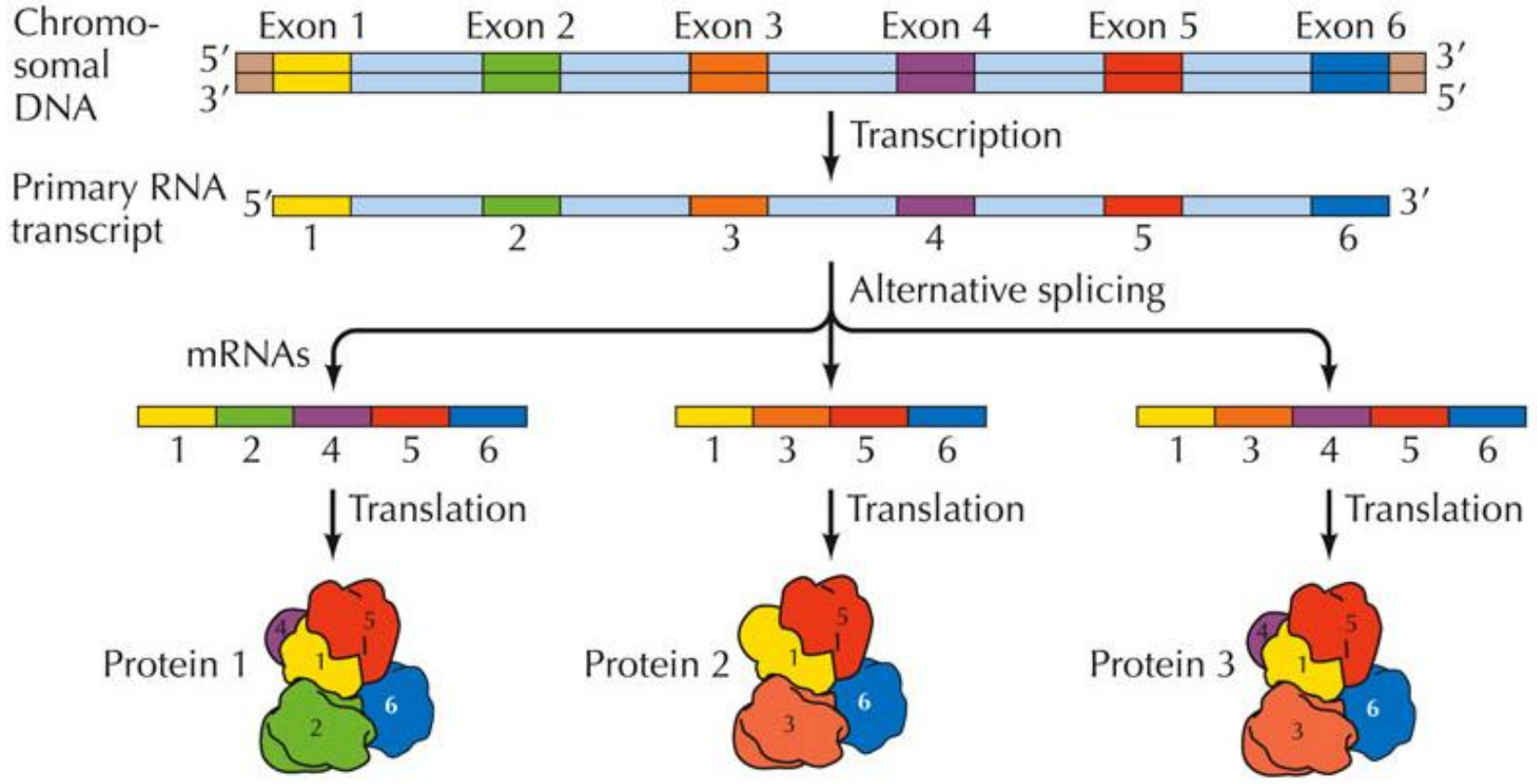
Prokaryotes



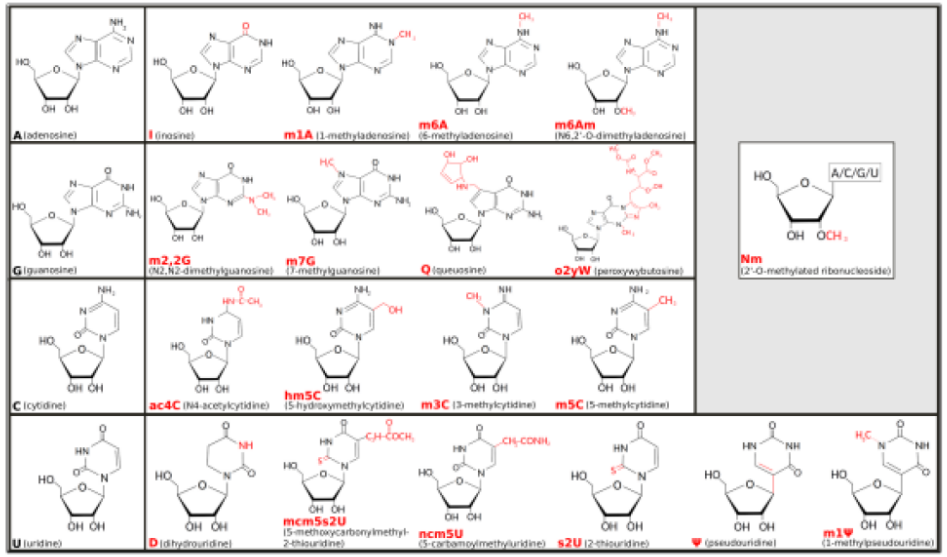
Eukaryotes

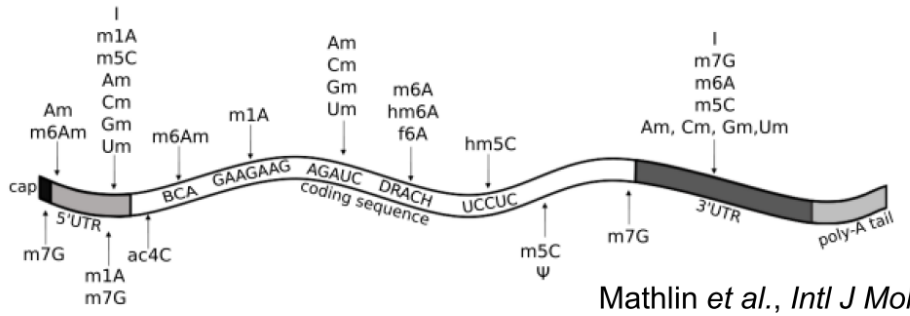


Alternative splicing- gives rise to different isoforms



Epitranscriptome a.k.a. RNA modifications





Mathlin *et al.*, *Intl J Molecular Sciences* (2020)

Core concepts

- Gene: sequence of nucleotides that encodes a functional molecule (mRNA, tRNA, rRNA, among others)
- Genome: the entire genetic information an organism (or virus) possesses
- Transcript: RNA molecule derived from a gene
- Transcriptome: The entire repertoire of transcripts present in an organism at a given time
- Isoform: transcript that originates from alternative transcription start sites, or alternative transcription termination sites, or by alternative splicing sites, or any combination of the above

Technologies to measure RNA

- Initially RNA was measured by hybridisation (mid 70s), then, retro-transcription was coupled to sequencing (80s-00s), then quantitative PCR was used, and then, high throughput hybridisation (00s-present).
- All of these methodologies are great on their own, but were rather incomplete

Victor Flores. 2024

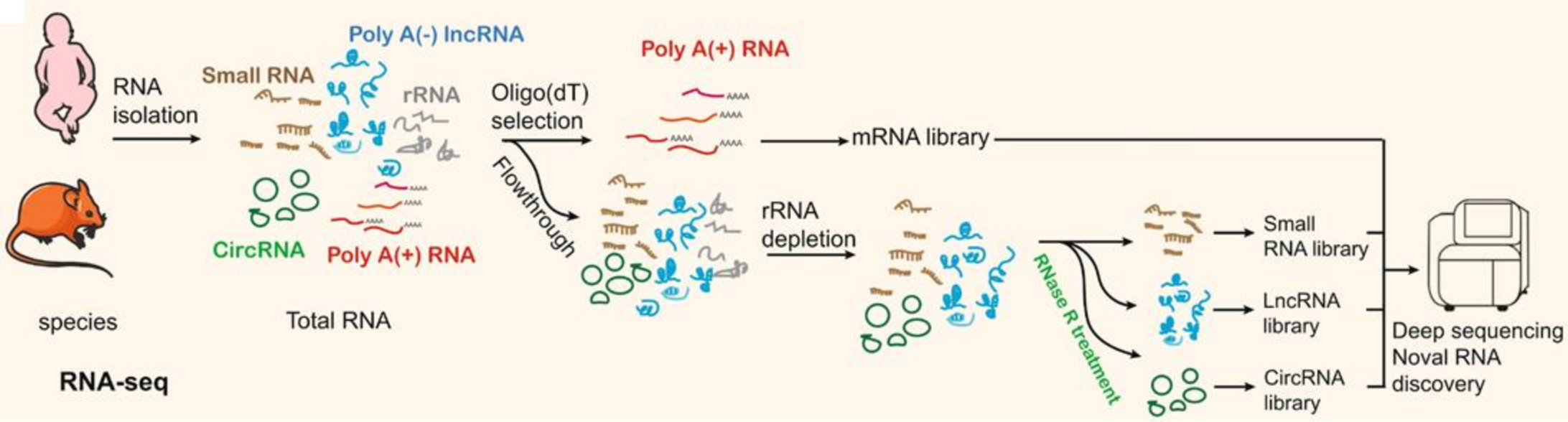
Technologies to measure RNA

- Northern blot and qRT-PCR are highly specific, but they are time consuming and yield very low throughput
- ESTs sequencing and SAGE can have high throughput, but are time consuming, expensive and rarely capture the whole transcriptome
- Microarrays have high throughput, high specificity, moderate sensitivity but fail to detect novel transcripts

Victor Flores. 2024

Enter RNAseq

- Once massive sequencing platforms started to become available, they were applied to sequence RNA, thus greatly impacting the way we conduct RNA research and applications



Sun. 2020. doi: 10.1186/s13045-020-00945-8

RNAseq

- Just as the previous technologies are not perfect, neither it is RNAseq, however, the advantages of performing RNAseq outweigh its limitations
  - It is genome independent
  - It yields high throughput
  - It can be used to discover novel transcripts/isoforms/genes
  - It can be easily parallelized for statistical analyses
  - There are multiple programs, algorithms and frameworks to process the data

We are happy to help you!