Intro to RNA-seq

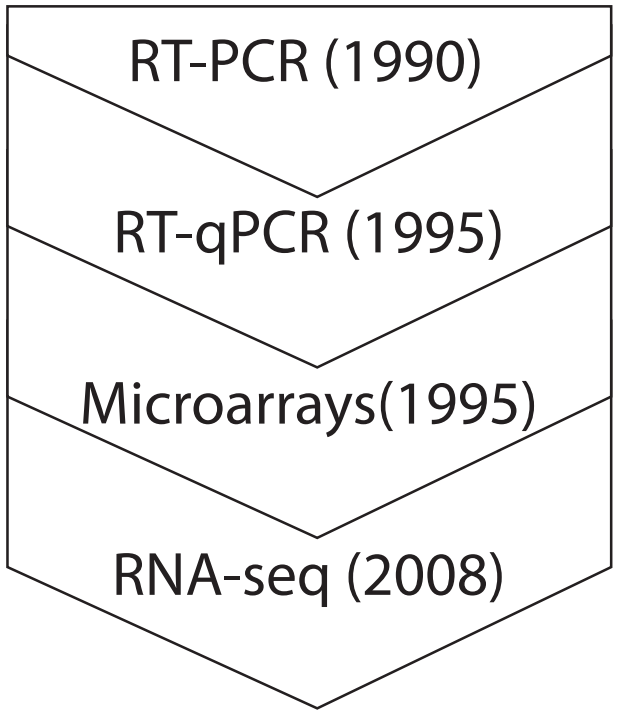
July 13, 2015

Goal of the course

- To be able to effectively design, and interpret genomic studies of gene expression.

- We will focus on RNA-seq, but the class will provide a foothold into other functional genomics assays.

Measuring Changes in Gene Expression



Throughput:	Quantitativeness:	Ease of use:
---	-	++
+	+++++	+
++	++	--
+++	+++	--

RNA-seq references:

1) Mortazavi et al, Nature Methods, 2008
2) Nagalakshmi et al, Science, 2008
3) Wilhelm et al, Nature, 2008

Why switch from arrays to sequencing?

Why switch from arrays to sequencing?

Technology	Tiling microarray	cDNA or EST sequencing	RNA-Seq
Technology specifications			
Principle	Hybridization	Sanger sequencing	High-throughput sequencing
Resolution	From several to 100 bp	Single base	Single base
Throughput	High	Low	High
Reliance on genomic sequence	Yes	No	In some cases
Background noise	High	Low	Low
Application			
Simultaneously map transcribed regions and gene expression	Yes	Limited for gene expression	Yes
Dynamic range to quantify gene expression level	Up to a few-hundredfold	Not practical	>8,000-fold
Ability to distinguish different isoforms	Limited	Yes	Yes
Ability to distinguish allelic expression	Limited	Yes	Yes
Practical issues			
Required amount of RNA	High	High	Low
Cost for mapping transcriptomes of large genomes	High	High	Relatively low

RNA-Seq: a revolutionary tool for transcriptomics. Zhong Wang, Mark Gerstein & Michael Snyder

Design of RNA-seq experiments

Key points:

1) What RNA molecules are important?
2) What aspects of the RNA molecules are important?

Answering these questions will inform:

1) What to sequence?
2) How to sequence?
3) How much to sequence?
4) Etc...

Parts of an RNA-seq experiment

1) Isolation of RNA molecules for sequencing.
   (Which molecules?)
3) Preparation of sequencing library
   (How to prepare?)
4) High-throughput sequencing
   (What, and how much?)
5) Analysis and interpretation
   (What question to ask, and how?)

Step 1: Isolating the right RNA.

Types of RNA in a cell

Ribosomal RNA (rRNA): translation

Messenger RNA (mRNA): protein coding

Transfer RNA (tRNA): translation

Long noncoding RNA (lncRNA): regulatory, structural

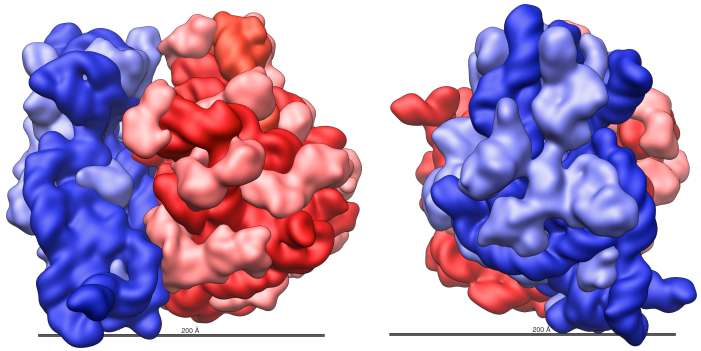
+ Currently a "hot topic"

microRNA (miRNA): regulatory

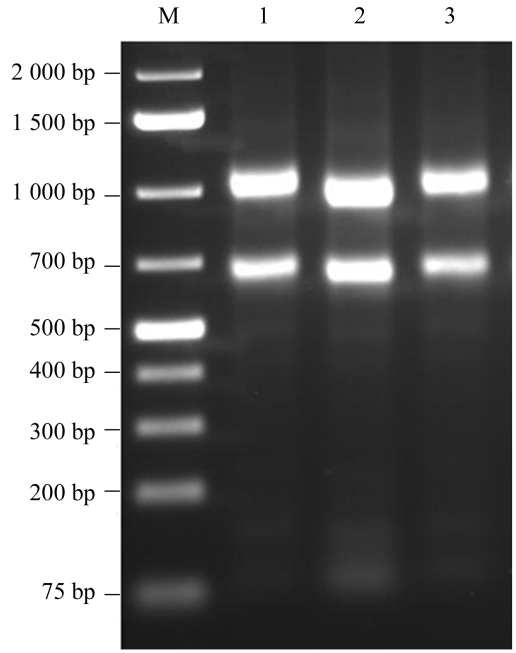
There are many others, but we will focus on these.

Ribosomal RNA

- Estimated to account for 90-95% of RNA in a human cell.
- Two predominant rRNAs in a human cell:
  - 18S rRNA (small component, blue)
  - 28s rRNA (large component, red)



source: wikipedia

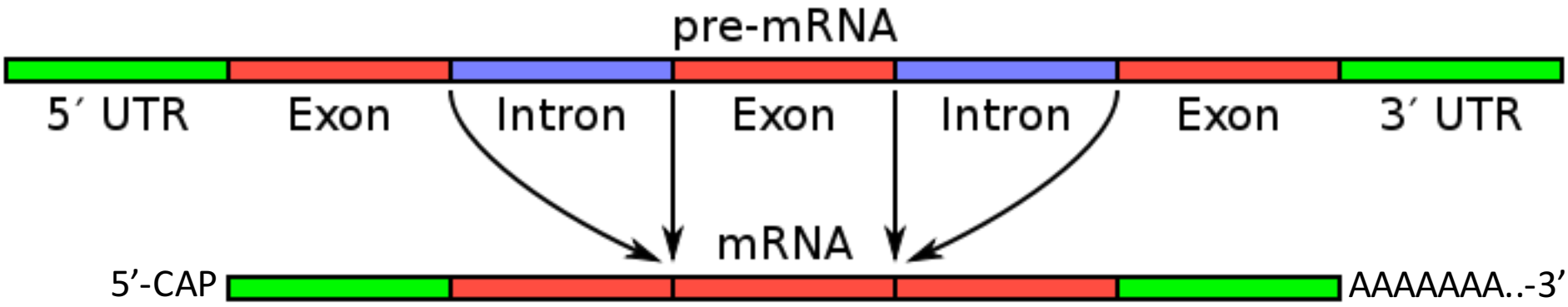


- Removing rRNA is a challenge

Gel image source: Cloning and Analysis of Gene Fragments Encoding *C4H* and *CAD* from *Miscanthus sinensis*

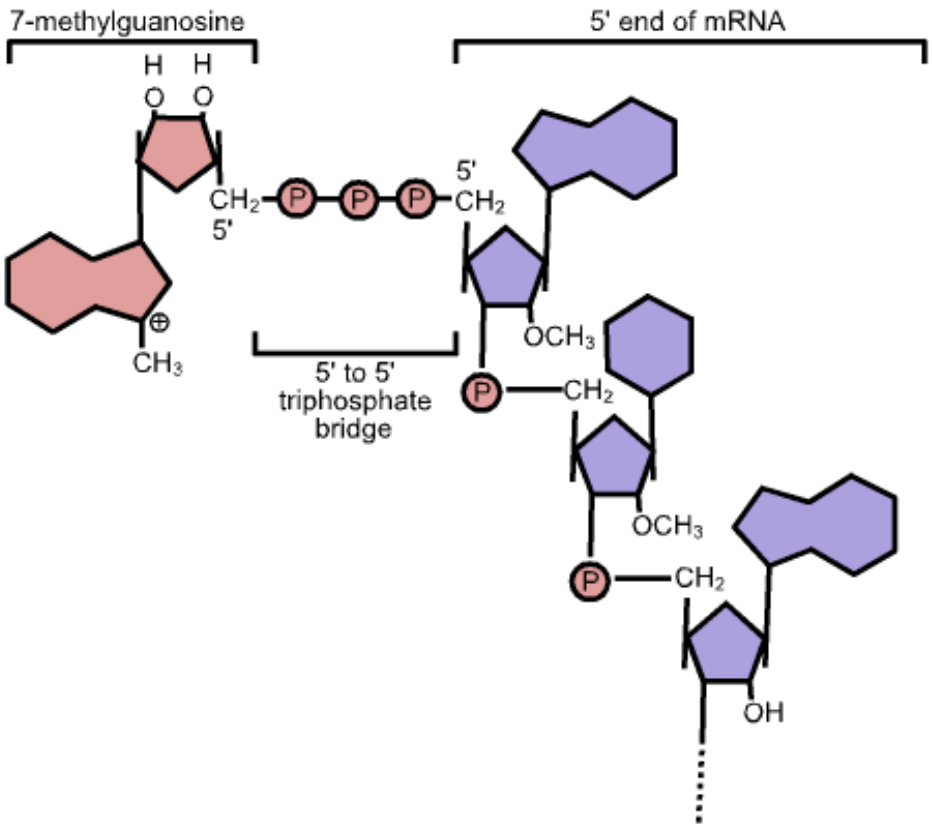
Messenger RNA

- The most commonly interrogated type of RNA
- Typical structure:
  - 5’ Cap
  - 5’ Untranslated region (5’ UTR)
  - Coding sequence (CDS)
  - 3’ Untranslated region (3’ UTR)
  - Poly-A tail



source: wikipedia

Messenger RNA, 5’ Cap



source: wikipedia

Messenger RNA

- 5’ Untranslated region (5’ UTR)
- Coding sequence (CDS)
- 3’ Untranslated region (3’ UTR)

What is the median size of an mRNA?

What is the median size of a CDS?

What is the median size of an hnRNA?

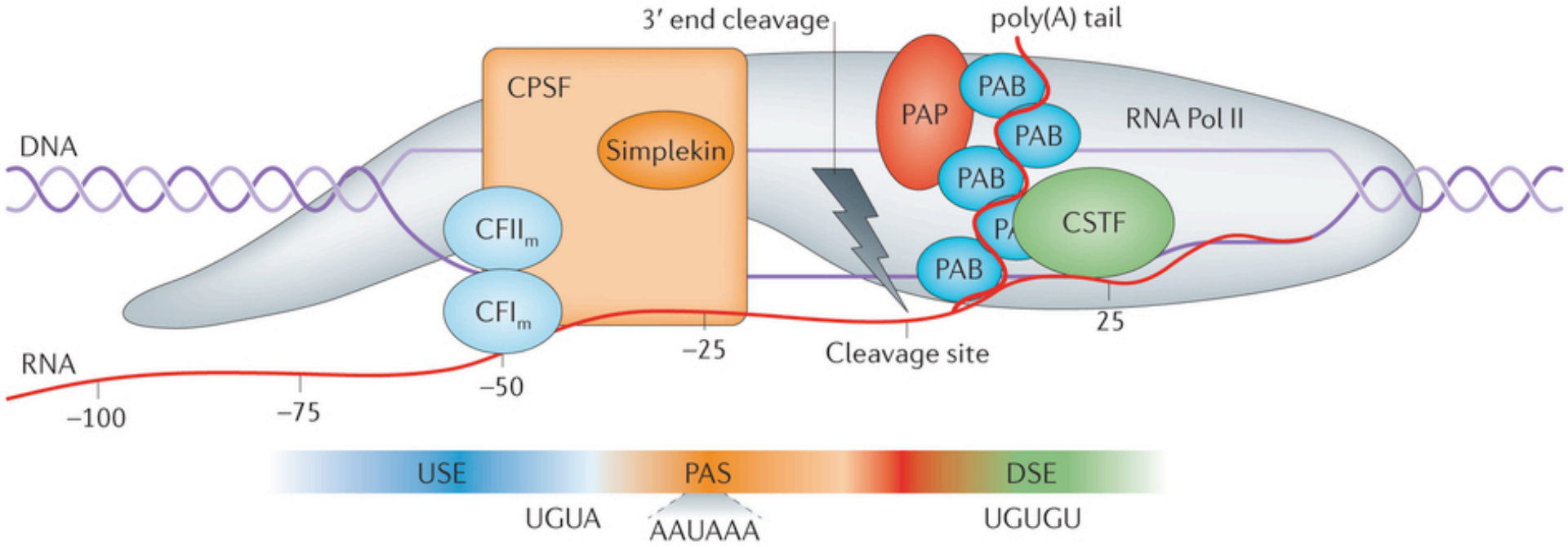
What is the median size of an exon?

What is the median size of an intron?

How many exons per mRNA?

mRNA polyA tail

How long is the polyA tail added to mRNAs?



Alternative cleavage and polyadenylation: extent, regulation and function. Elkon, Ugalde, Agami

Long noncoding RNAs

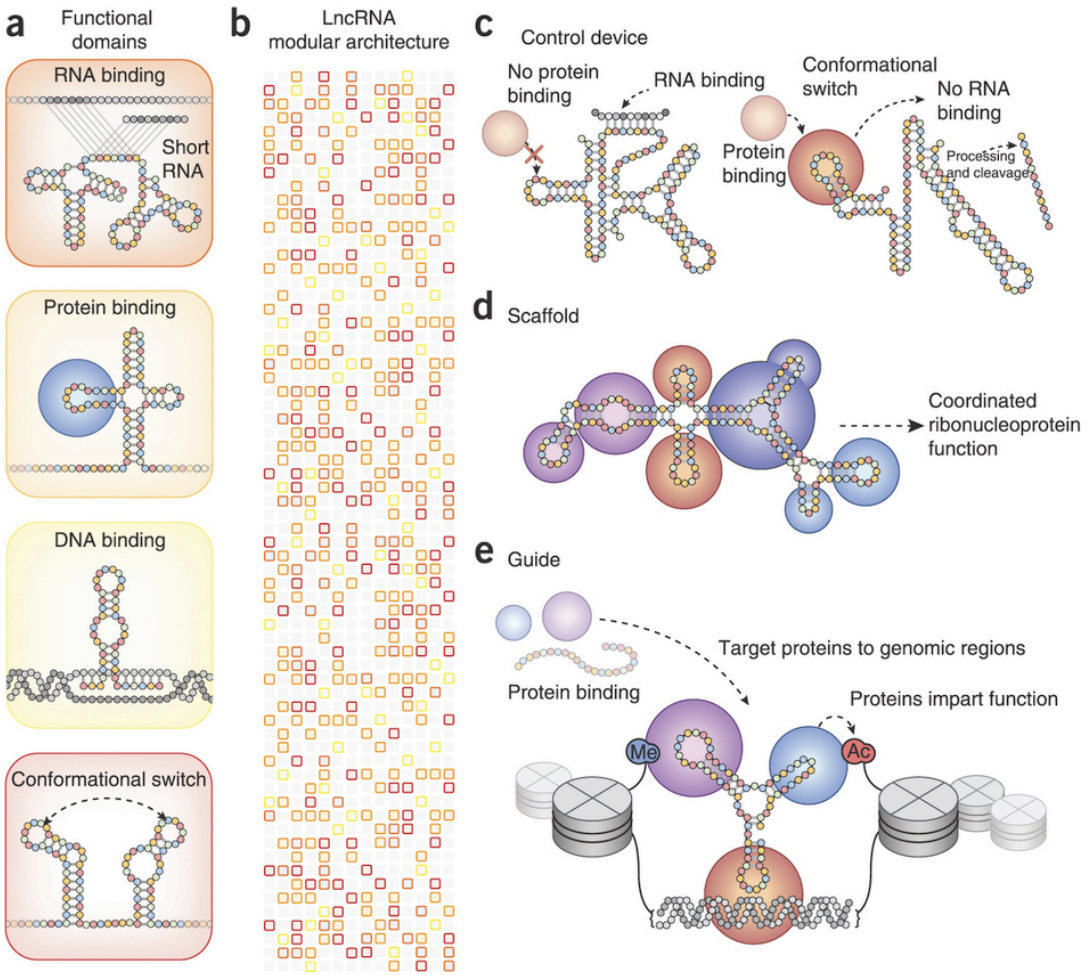
- Typically kilobases in size
- Lowly expressed
- Polyadenylated

Classic Examples:

XIST (X silencing)

HOTAIR (HOXD)

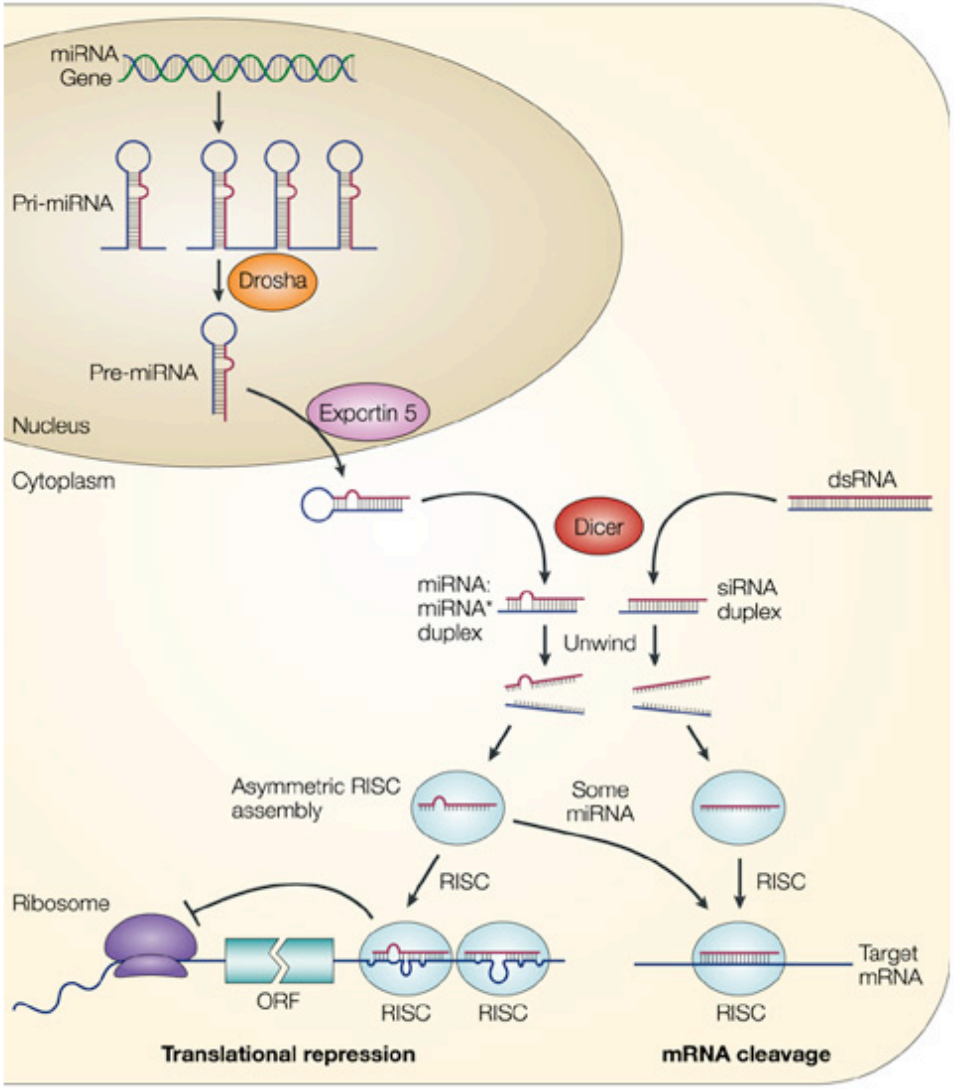
KCNQ1OT1



Mercer and Mattick, Nat. Struct. Mol. Biol., 2013

miRNAs

- Act via RNAi pathway
- Pre-mRNA (70bp)
- miRNA (21-25bp)
- polyA- once processed



Lin He & Gregory J. Hannon

Nature Reviews | Genetics

Isolation of RNA molecules

Three overall approaches are commonly used:

- Organic Extraction / Precipitation
  - Phenol:Chloroform:Isoamyl Alcohol
  - Alcohol Precipitation
- Column purification:
  - Several Manufacturers
- Bead-based purification:
  - SPRI beads
  - Oligo-dT beads for polyA selection
  - 5’ Cap capture

Considerations when isolating RNA

- Size exclusion:
  - Most beads and columns do not collect RNA < 70-100bp
  - Small RNA molecules are also difficult to precipitate
- RNA integrity:
  - PolyA+ selection does not work for degraded RNA
- Amount of RNA:
  - Special techniques are employed for low-abundance samples
- Downstream processing:
  - Phenol carryover can interfere with downstream steps and can bias quantitation.

Typical prep approaches for RNA-seq

- mRNA/lncRNA selection:
  - Purification of “total RNA” with columns
  - Selection of polyA+ RNA with oligo-dT beads

- mRNA from low-quality samples:
  - Purification of “total RNA” with columns
  - Removal of ribosomes with beads

- miRNA/tRNA/other small RNAs:
  - Precipitation of total RNA, or specially designed columns
  - Gel-based size selection of small molecules

*extra steps are generally included to remove genomic DNA