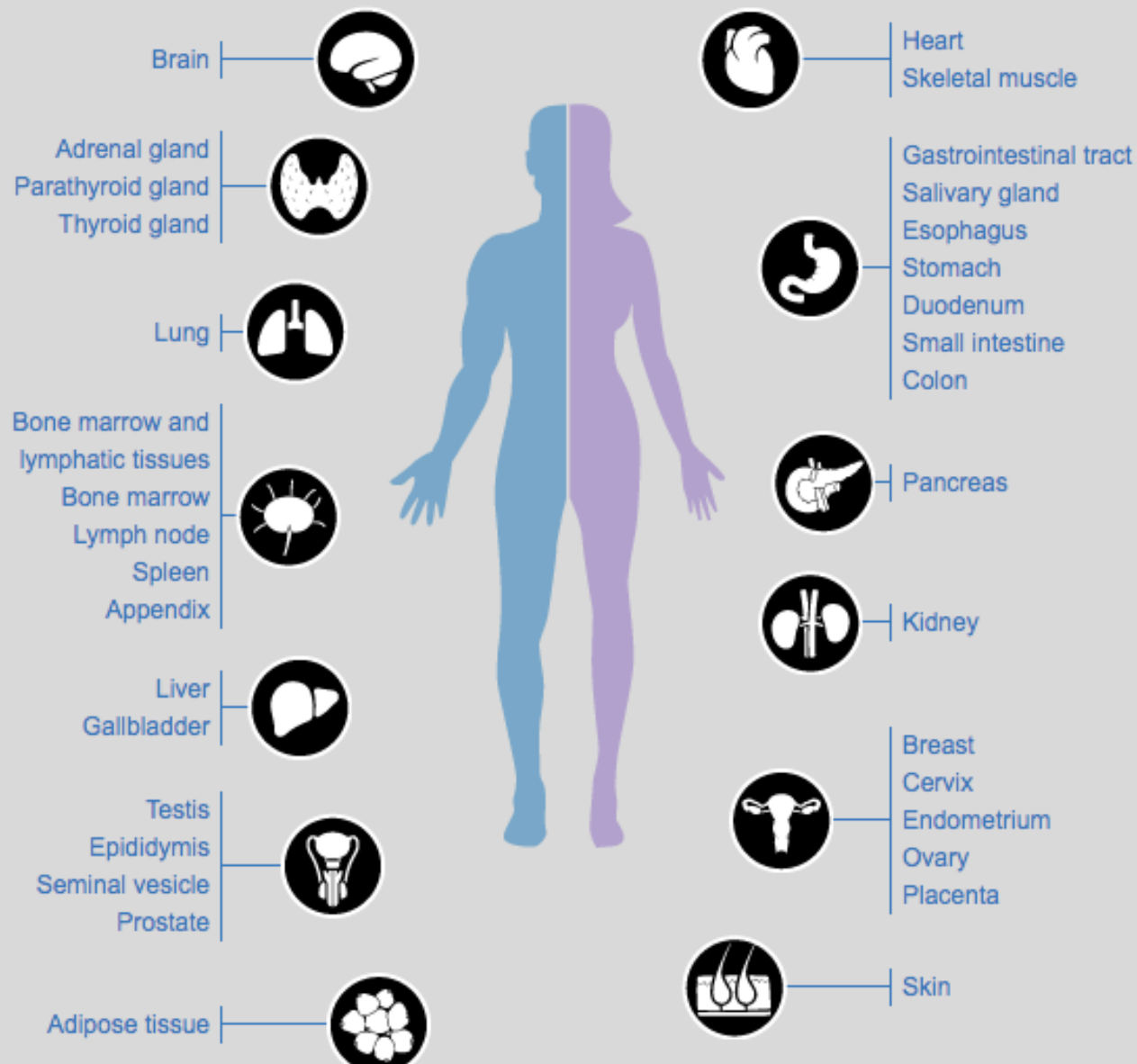


TRANSCRIPTOMICS

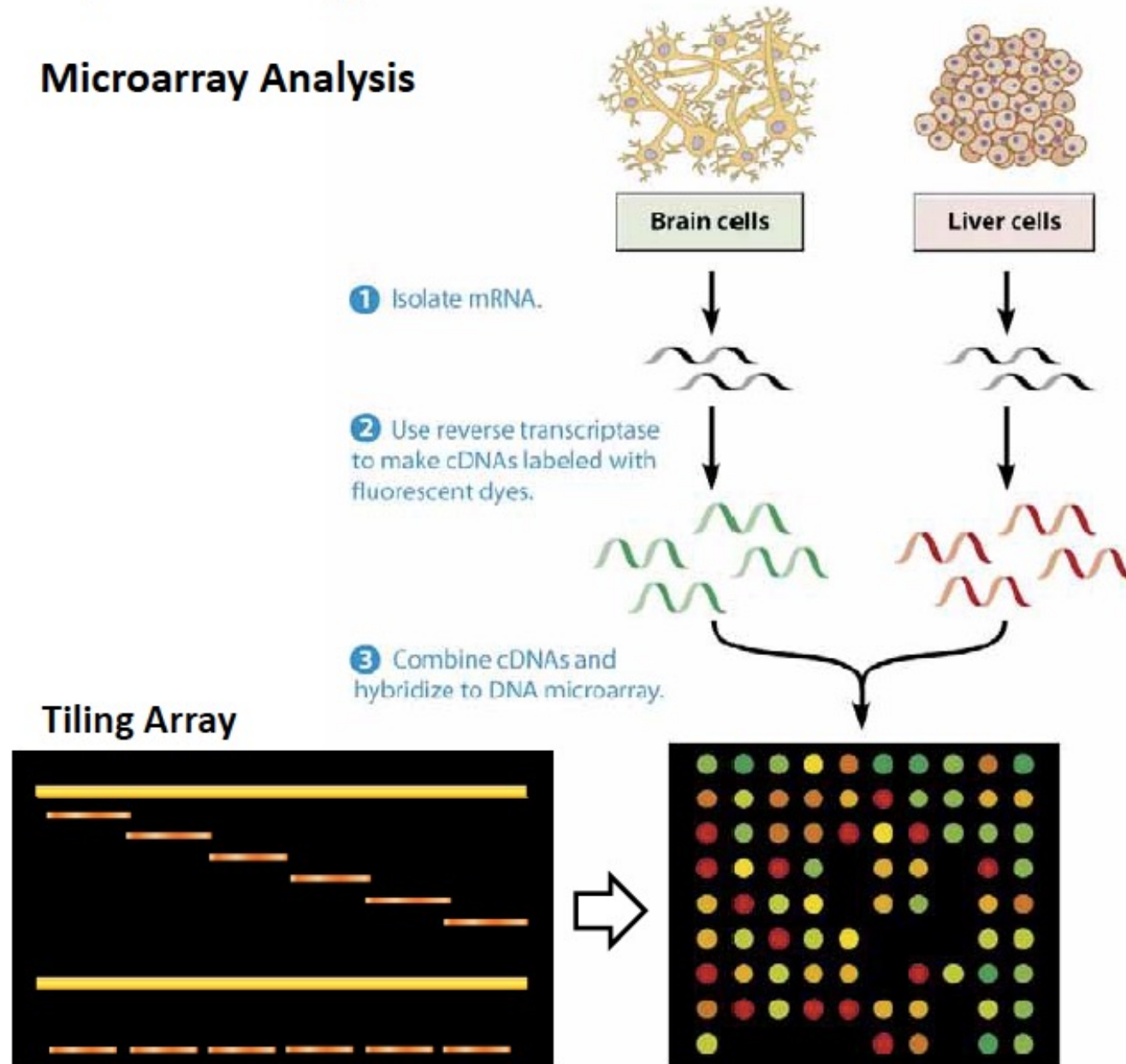
The study of the complete set of RNAs (**transcriptome**) encoded by the genome of a specific cell or organism at a specific time or under a specific set of conditions

The expression for all protein-coding genes in all major tissues and organs in the human body can be explored in this interactive database, including numerous catalogues of proteins expressed in a tissue-restricted manner.



Experimental Approaches Used to Characterize the Human Transcriptome

Microarray Analysis

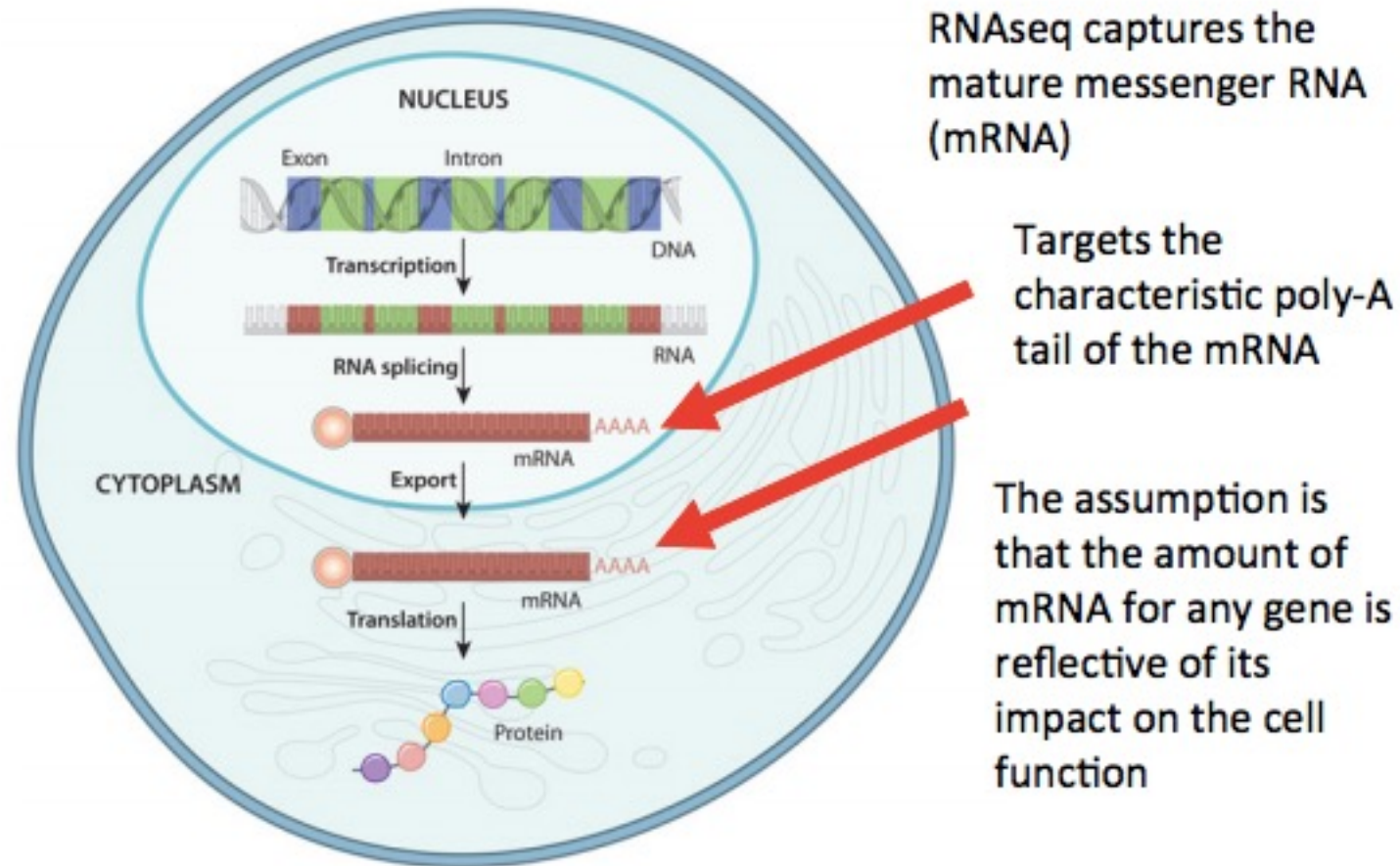


Why not sequence RNA?

RNA: Ribonucleic Acid

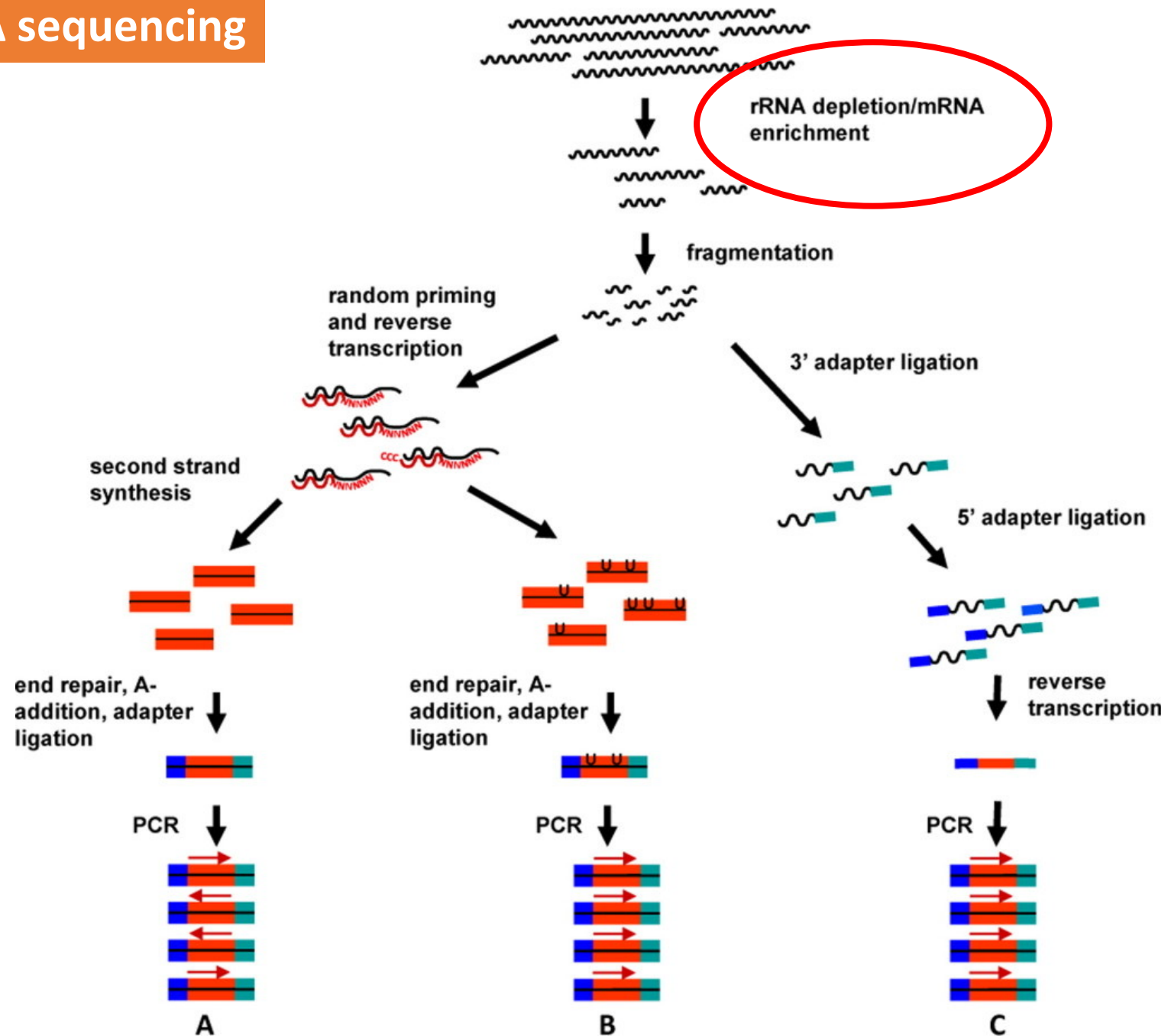
- Types of RNA
 - Ribosomal RNA (rRNA): catalytic component of ribosomes (about 80-85%)
 - Transfer RNA (tRNA): transfers amino acids to polypeptide chain at the ribosomal site of protein synthesis (about 15%)
 - Messenger RNA (mRNA): carries information about a protein sequence to the ribosomes (about 5%)
 - Other types
 - miRNA, siRNA, snRNA, dsRNA,...

The stage of gene expression we capture



RNA sequencing

RNA-seq

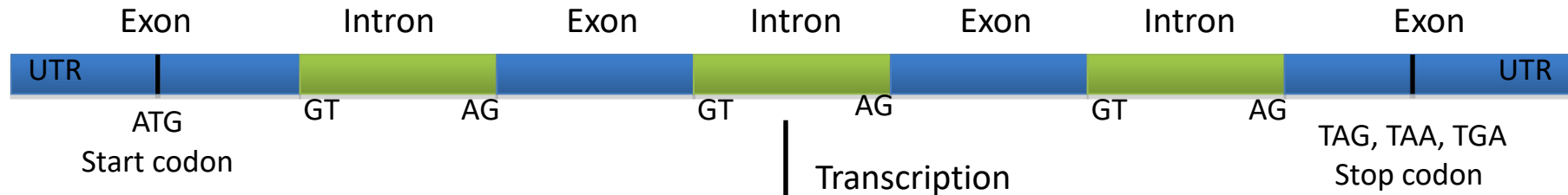


RNA sequencing

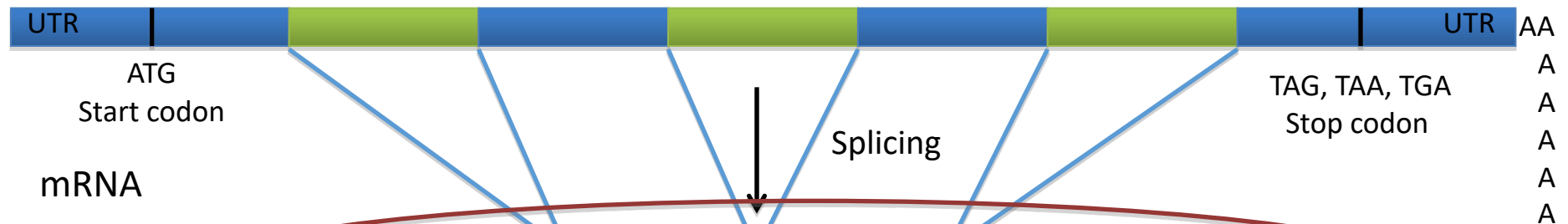
- Mapping-based transcriptomics (genome - based)
- Expression counts and differential expression
- (Alternative splicing)

RNA-seq

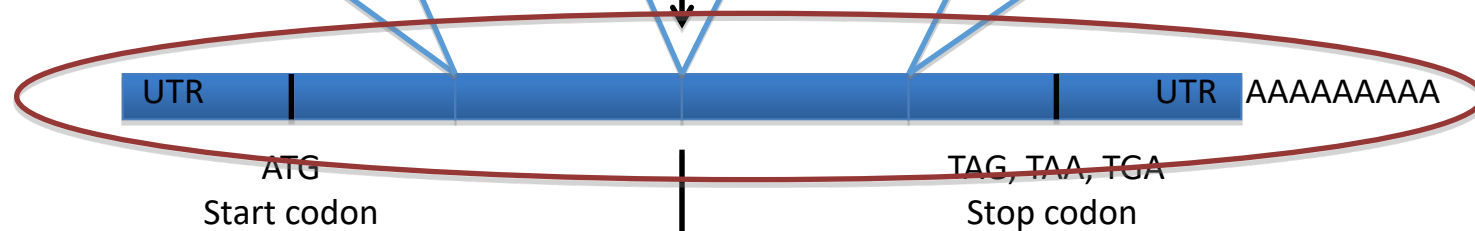
DNA



Pre-mRNA

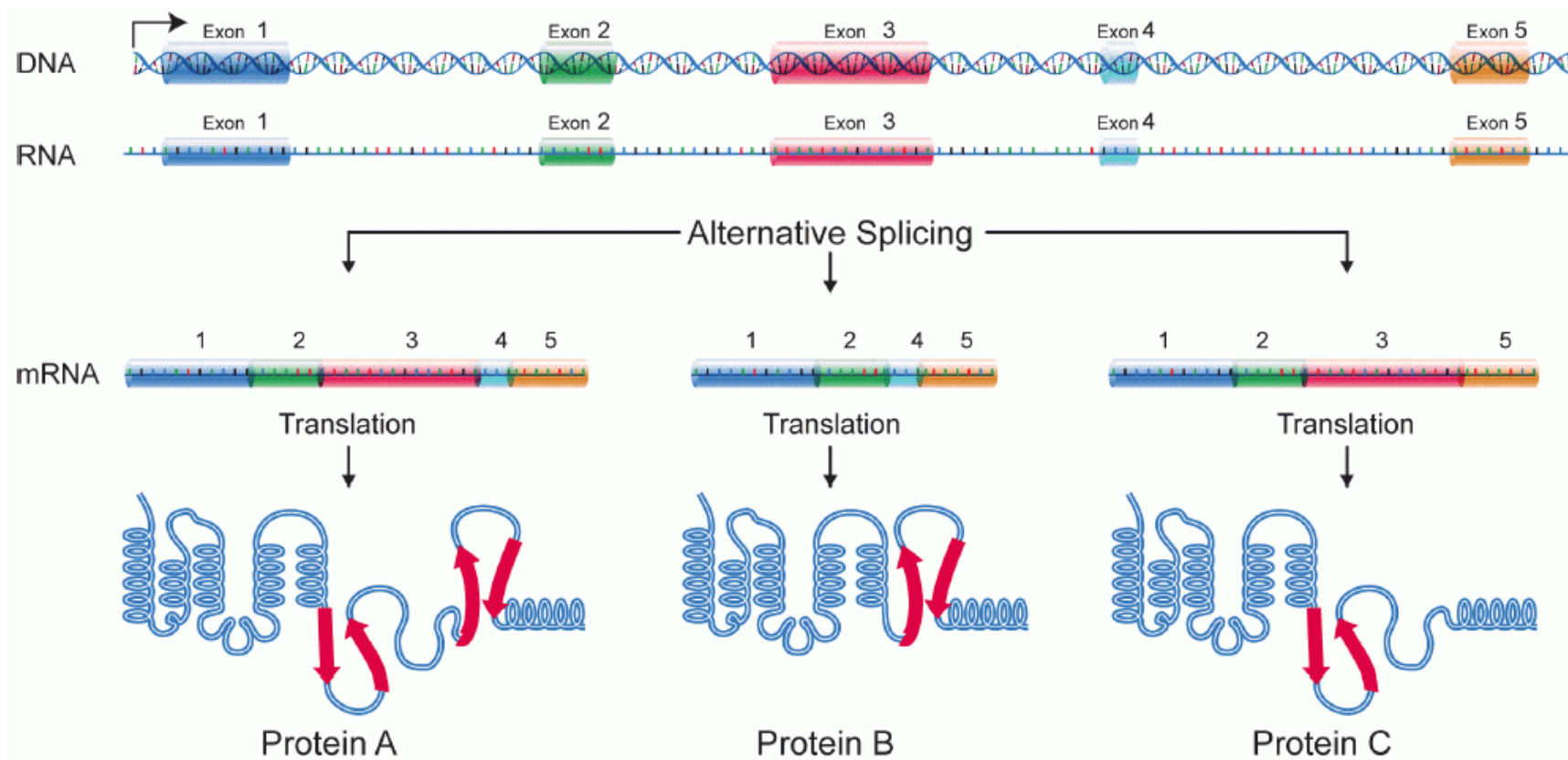


mRNA

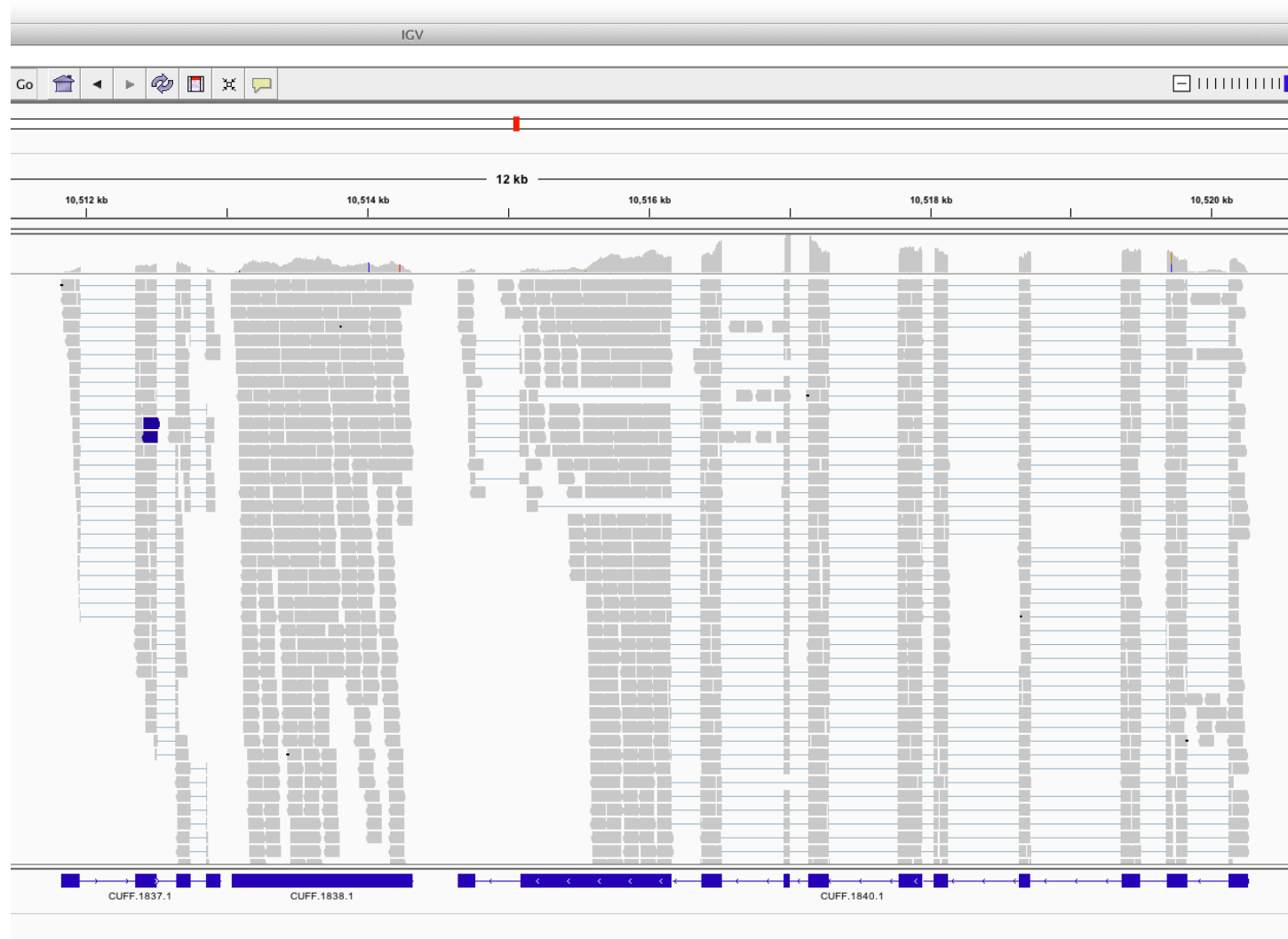


Translation

From DNA to Proteins



RNA-seq - Spliced reads



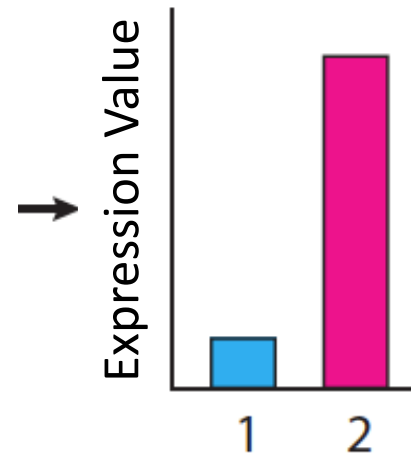
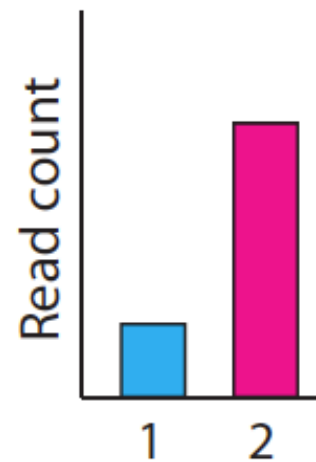
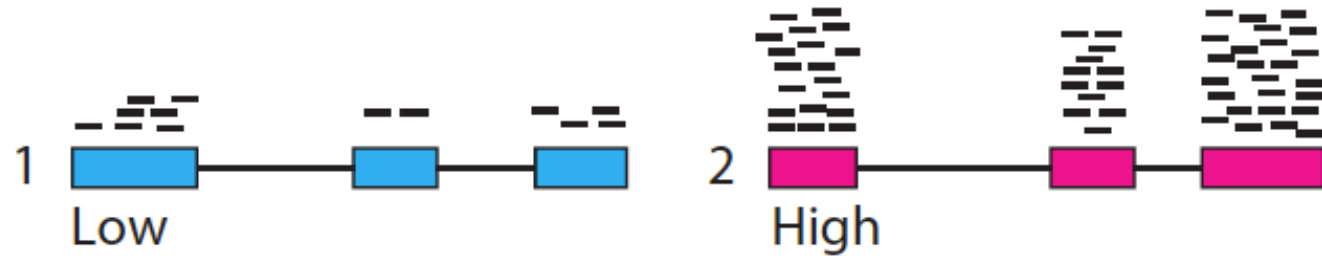
Differential Expression Analysis

1. Mapping & transcript reconstruction
2. Abundance estimation
3. Differential expression (gene level)
4. Enrichment analysis

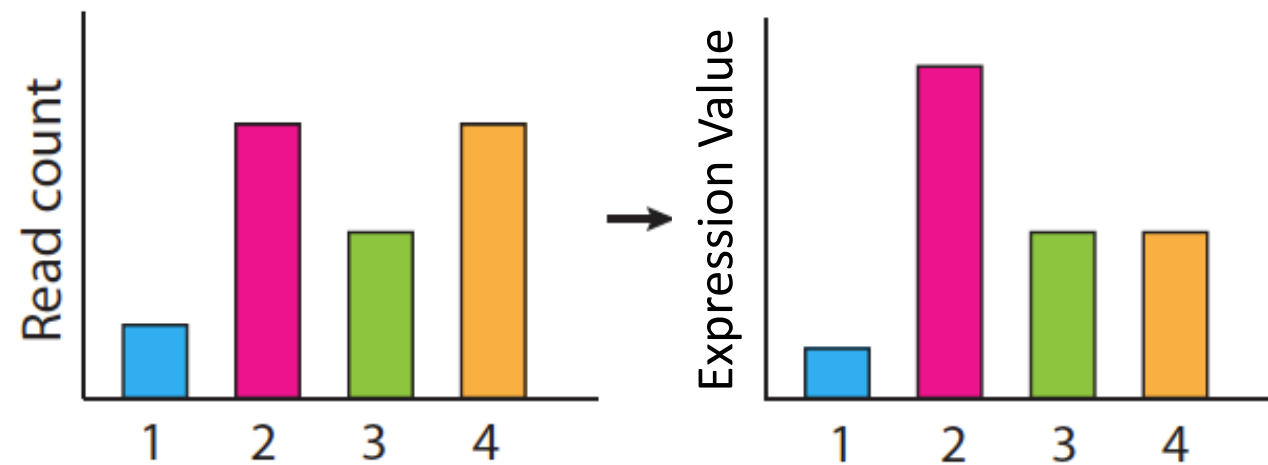
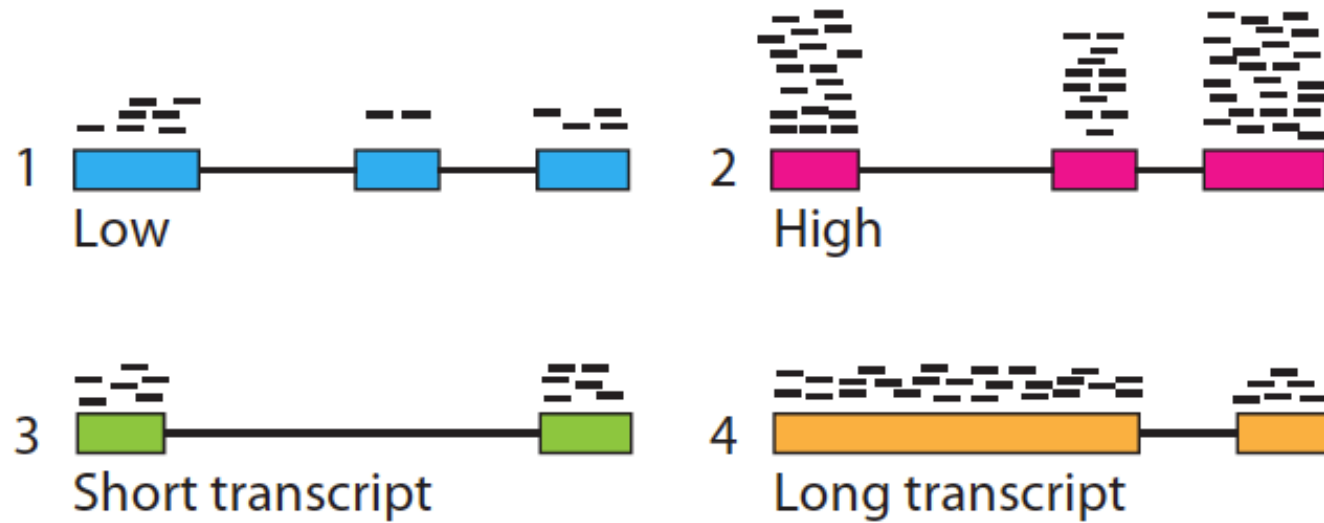
Abundance Estimation

(Computing Expression Values)

Calculating expression of genes and transcripts



Calculating expression of genes and transcripts



Normalized Expression Values

- Transcript-mapped read counts are normalized for both length of the transcript and total depth of sequencing.
- Reported as: Number of RNA-Seq **F**ragments
Per **K**ilobase of transcript
per total **M**illion fragments mapped

FPKM

Differential Expression Analysis Using RNA-Seq

Differential expression



Diff. Expression Analysis Involves

- Counting reads
- Statistical significance testing

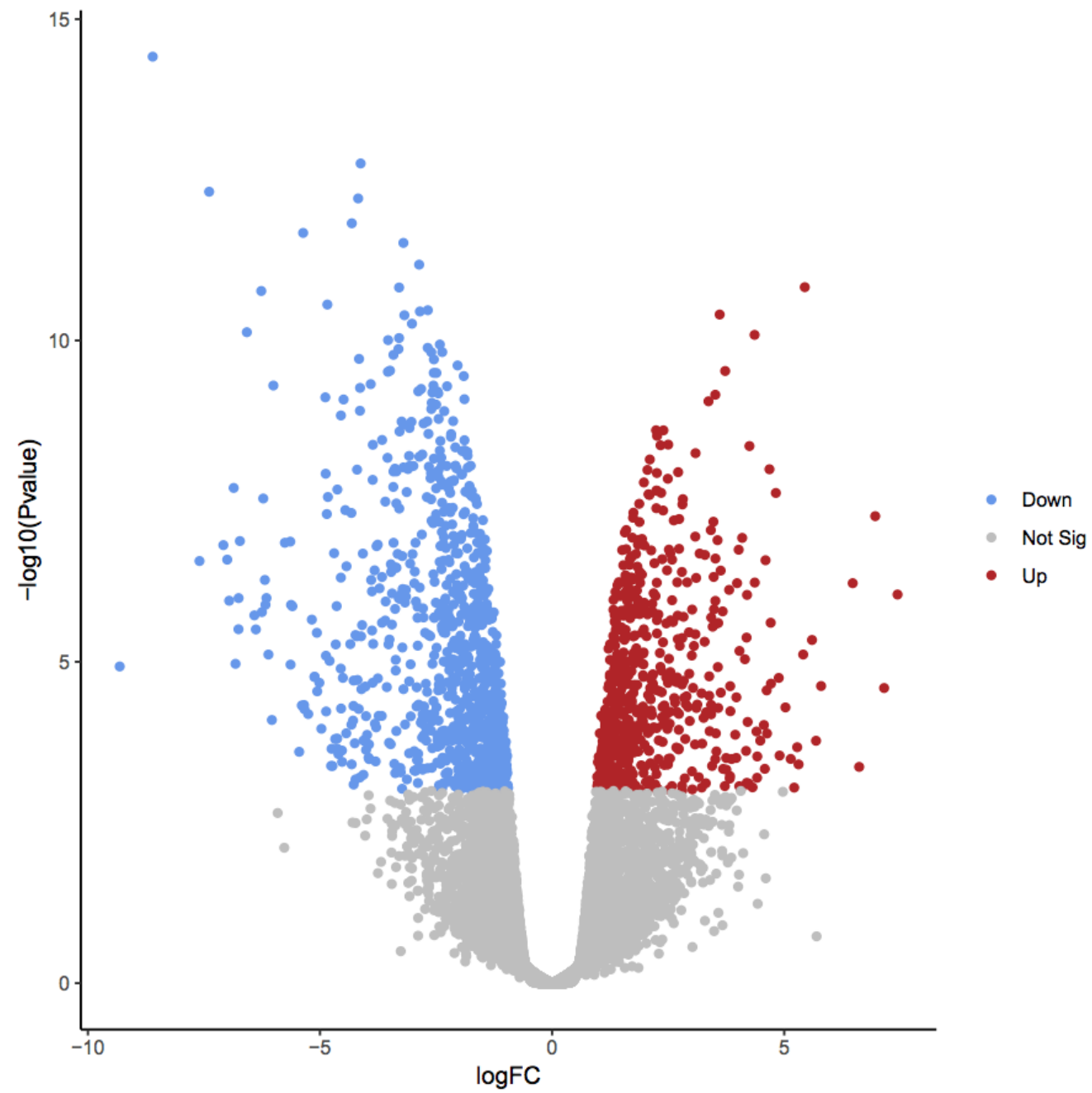
	Sample_A	Sample_B	Fold_Change	Significant?
Gene A	1	2	2-fold	No
Gene B	100	200	2-fold	Yes

More Counts = More Statistical Power

Example: 5000 total reads per sample.

Observed 2-fold differences in read counts.

	SampleA	Sample B	Fisher's Exact Test (P-value)
geneA	1	2	1.00
geneB	10	20	0.098
geneC	100	200	< 0.001



Analysis Workflow

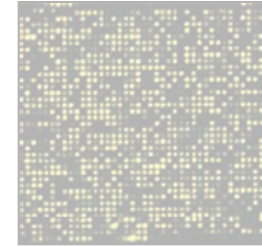
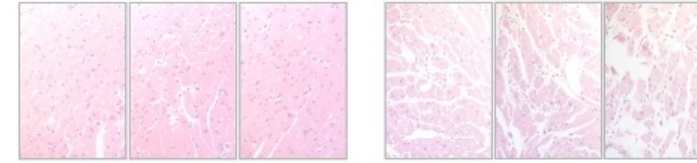
Define the **experimental design**

Collect the **biological samples**

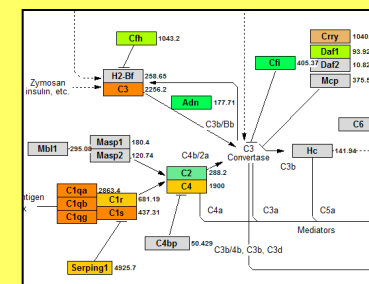
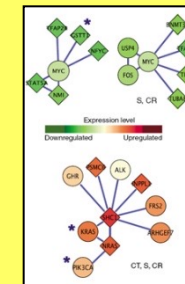
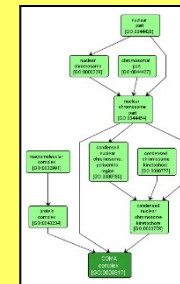
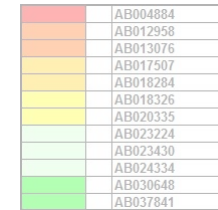
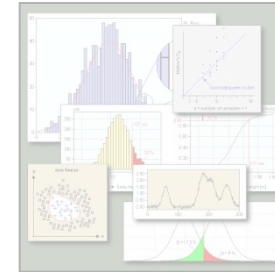
Generate the **expression data**

Identify the **Differential Genes**

Identify the **Functional Groups**



	A	B	C	D	E
1	Unique id	2009w0	2009w2	2009w4	2009LFU
2	AB004884	8.294597	8.086672	8.457752	8.367051
3	AB012958	7.648038	7.69157	7.844407	7.558314
4	AB013076	14.94777	14.73669	14.85634	14.7109
5	AB017507	11.61236	11.98187	12.13949	11.71435
6	AB018284	6.090057	6.566097	6.227057	5.927692
7	AB018326	8.823744	9.30355	8.94253	8.987581
8	AB020335	8.760232	8.922606	8.827492	8.819615
9	AB023224	14.39539	14.07356	14.11936	13.73096
10	AB023430	11.21844	11.28691	11.39944	11.29315
11	AB024334	13.62434	13.67584	13.29712	12.69672
12	AB030648	10.99952	10.46016	10.48638	10.72427
13	AB037841	10.08971	9.880636	9.743547	9.721693
14	AB040961	12.53133	11.83537	12.24728	12.62928

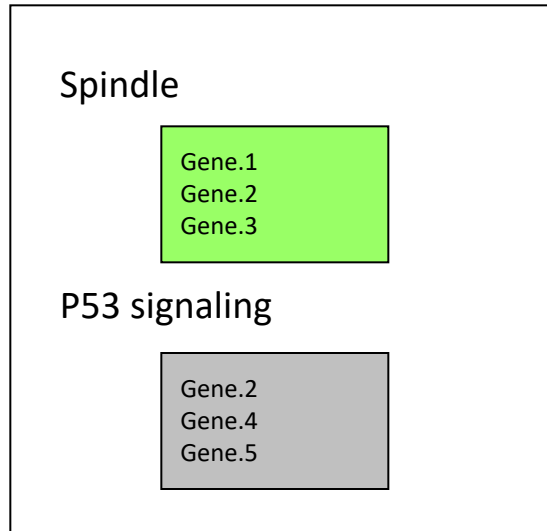


Basic Assumption

- Genes involved in the same biological processes, functions, or localizations **present correlated behaviors** in terms of expression levels, signal intensities, allele occurrences, and so on.
-
- We can therefore apply statistical tests to find perturbed pathways.

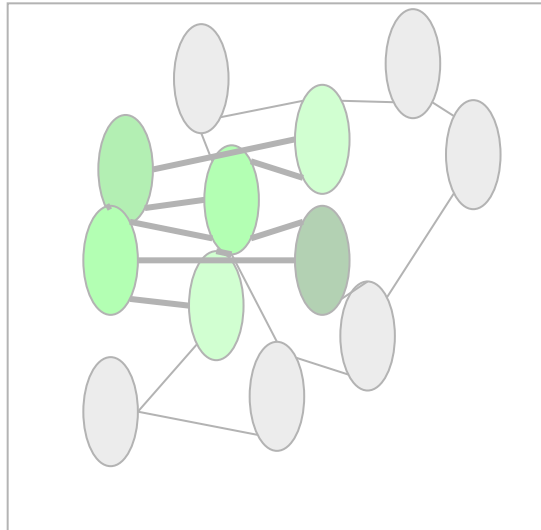
Identification of Functional Groups

GENE SETS



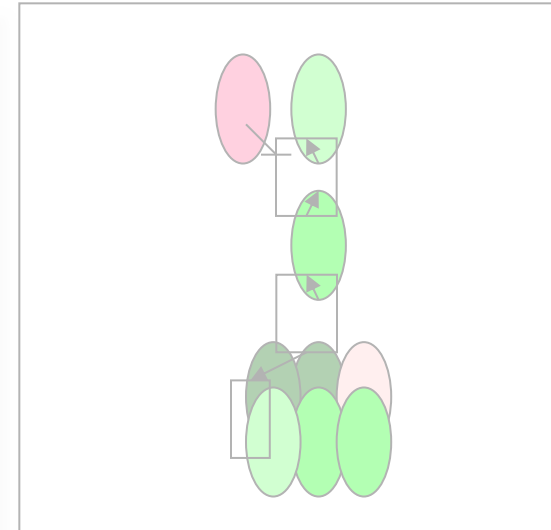
Score the set depending on the gene expression of its member genes

NETWORKS



Just visual, or Identify modules satisfying some joint gene expression and topology requirement

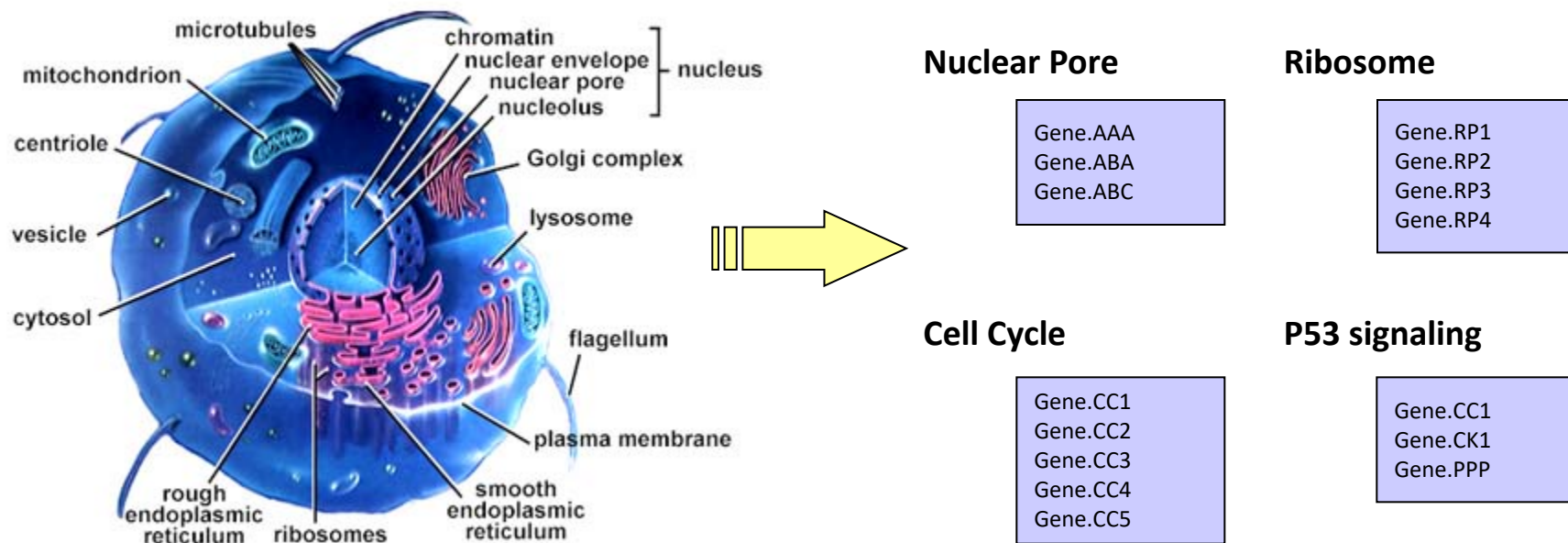
PATHWAYS



Just visual, or Score the pathways exploiting gene expression and topology

What's Gene-set Enrichment Analysis?

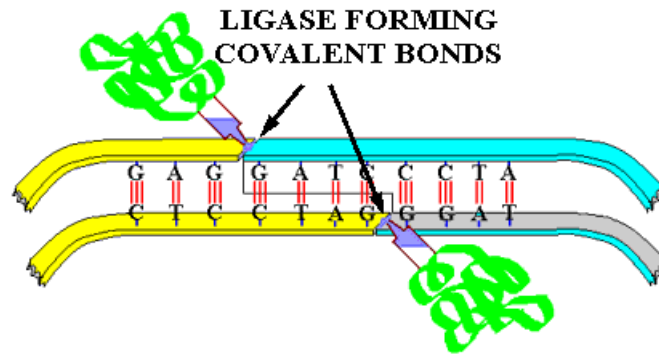
- Break down cellular function into gene sets
 - Every set of genes is associated to a specific cellular function, process, component or pathway



Gene Ontology (GO)

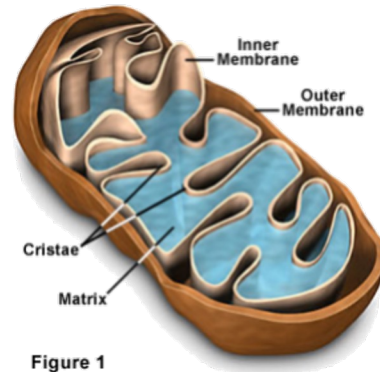
- Gene Ontology is:
 - a hierarchically-structured,
 - *Functional categories are organized hierarchically, i.e. a system of inter-related sets with increasing scope specificity*
 - *(parent-child relations)*
 - controlled vocabulary
 - *Functional categories are defined by experts, and then must be used consistently for annotation*
 - for gene product function annotation
 - *Gene products (i.e. proteins) are annotated using GO functional categories (“terms”)*
 - It is general for all species

**MOLECULAR
FUNCTION**



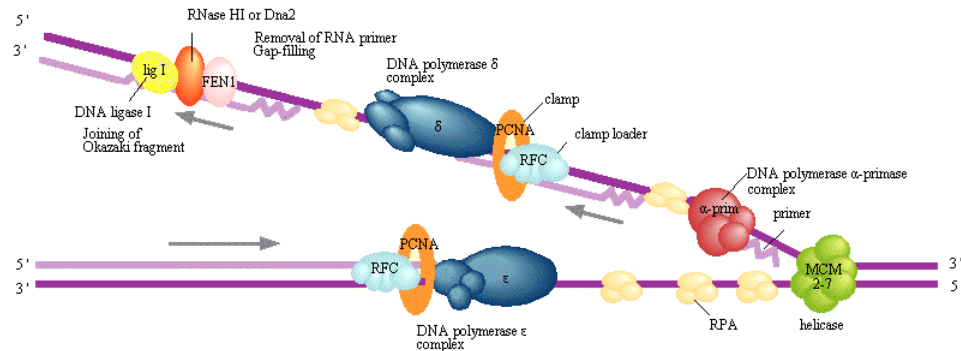
Ligase Activity

CELLULAR COMPONENT



Mitochondrion

**BIOLOGICAL
PROCESS**



NA Replication

Gene Ontology: Partitions

- GO has three independent partitions, which are not interconnected:
 - **Molecular Function**
 - Describes biochemical activities, in-vitro binding specificities, etc...
 - Example: *Ligase Activity, Kinase Activity, DNA Binding*
 - **Cellular Component**
 - Describes parts of the cell
 - Example: *Mitochondrion, Spindle Microtubule*
 - **Biological Process**
 - Describes processes at the intra-cellular and organism level
 - Example: *DNA Replication, Apoptosis, Development*

Gene Ontology: Partitions

First-level children (list)

MOLECULAR FUNCTION

- I [GO:0045735](#) nutrient reservoir activity
- I [GO:0010860](#) proteasome regulator activity
- I [GO:0003824](#) catalytic activity
- I [GO:0030528](#) transcription regulator activity
- I [GO:0042056](#) chemoattractant activity
- I [GO:0045499](#) chemorepellent activity
- I [GO:0031386](#) protein tag
- I [GO:0030234](#) enzyme regulator activity
- I [GO:0005198](#) structural molecule activity
- I [GO:0005215](#) transporter activity
- I [GO:0008369](#) obsolete molecular function
- I [GO:0045182](#) translation regulator activity
- I [GO:0005488](#) binding
- I [GO:0016209](#) antioxidant activity
- I [GO:0016247](#) channel regulator activity
- I [GO:0060089](#) molecular transducer activity
- I [GO:0016530](#) metallochaperone activity
- I [GO:0009055](#) electron carrier activity

CELLULAR COMPONENT

- I [GO:0045202](#) synapse
- I [GO:0019012](#) virion
- I [GO:0005623](#) cell
- I [GO:0031974](#) membrane-enclosed lumen
- I [GO:0008370](#) obsolete cellular component
- I [GO:0005576](#) extracellular region
- I [GO:0043226](#) organelle
- I [GO:0055044](#) symplast
- I [GO:0032991](#) macromolecular complex
- I [GO:0044421](#) extracellular region part
- I [GO:0044422](#) organelle part
- I [GO:0044423](#) virion part
- I [GO:0044456](#) synapse part
- I [GO:0044464](#) cell part

BIOLOGICAL PROCESS

- I [GO:0002376](#) immune system process
- I [GO:0040007](#) growth
- I [GO:0051234](#) establishment of localization
- I [GO:0000003](#) reproduction
- I [GO:0043473](#) pigmentation
- I [GO:0016032](#) viral reproduction
- I [GO:0050896](#) response to stimulus
- I [GO:0008371](#) obsolete biological process
- I [GO:0051704](#) multi-organism process
- I [GO:0009987](#) cellular process
- I [GO:0008152](#) metabolic process
- I [GO:0016265](#) death
- I [GO:0022414](#) reproductive process
- I [GO:0022610](#) biological adhesion
- PR [GO:0048518](#) positive regulation of biological process
- NR [GO:0048519](#) negative regulation of biological process
- I [GO:0023046](#) signaling process
- I [GO:0023052](#) signaling
- I [GO:0032501](#) multicellular organismal process
- I [GO:0032502](#) developmental process
- I [GO:0040011](#) locomotion
- R [GO:0050789](#) regulation of biological process
- I [GO:0048511](#) rhythmic process
- I [GO:0071554](#) cell wall organization or biogenesis
- I [GO:0051179](#) localization
- I [GO:0065007](#) biological regulation
- I [GO:0001906](#) cell killing
- I [GO:0016043](#) cellular component organization
- I [GO:0044085](#) cellular component biogenesis

