

Fundamentals of Molecular Biology

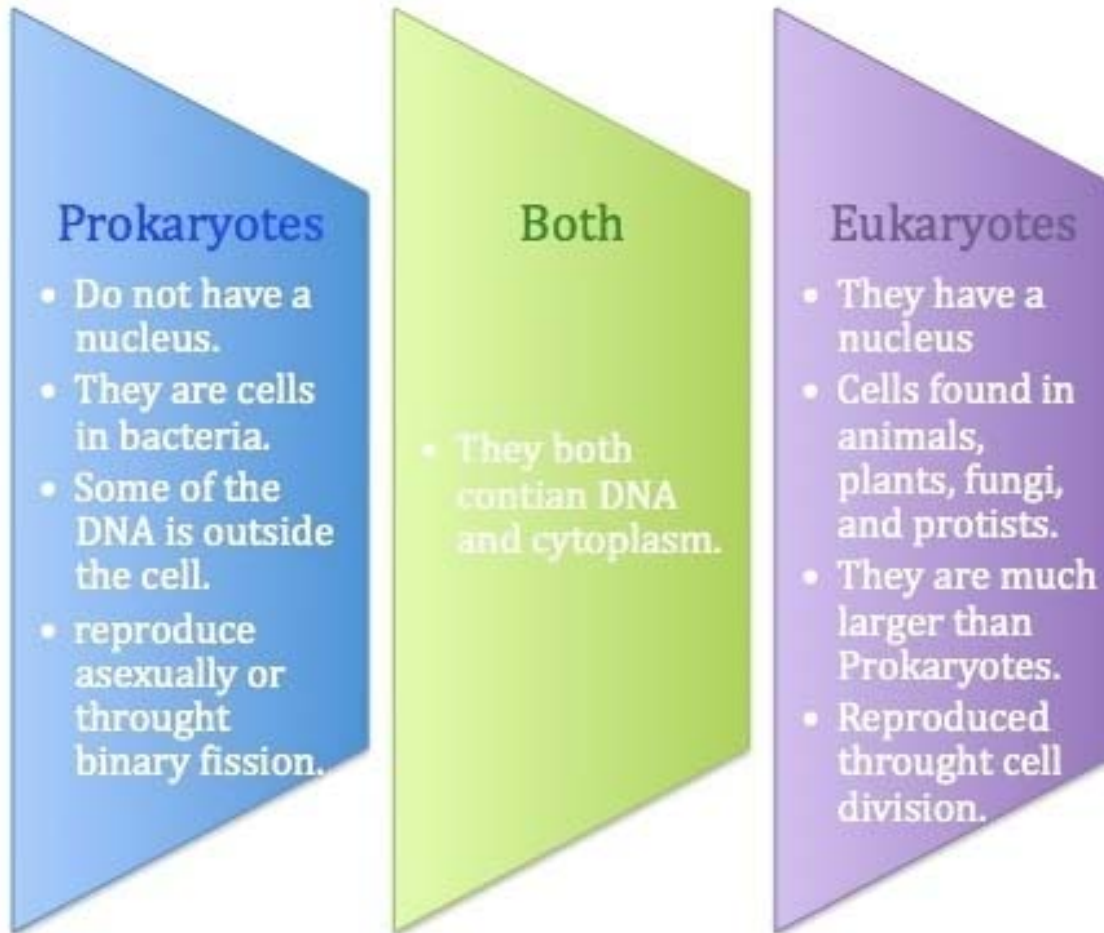
Malay Bhattacharyya
SRF, MIU, ISI Kolkata

Outline of the presentation

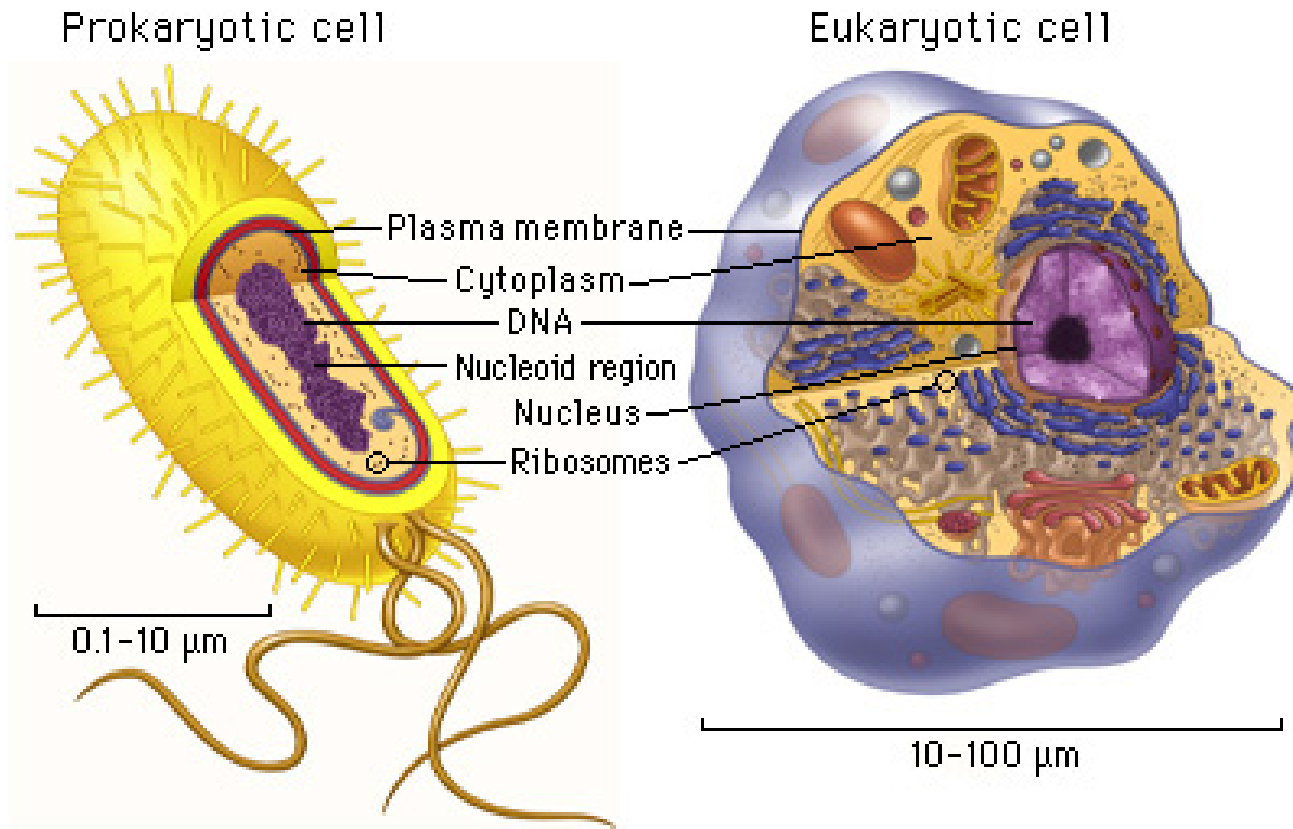
- ▶ Introduction
- ▶ Gene
- ▶ Central dogma of molecular biology
- ▶ Protein
- ▶ MicroRNA

Introduction

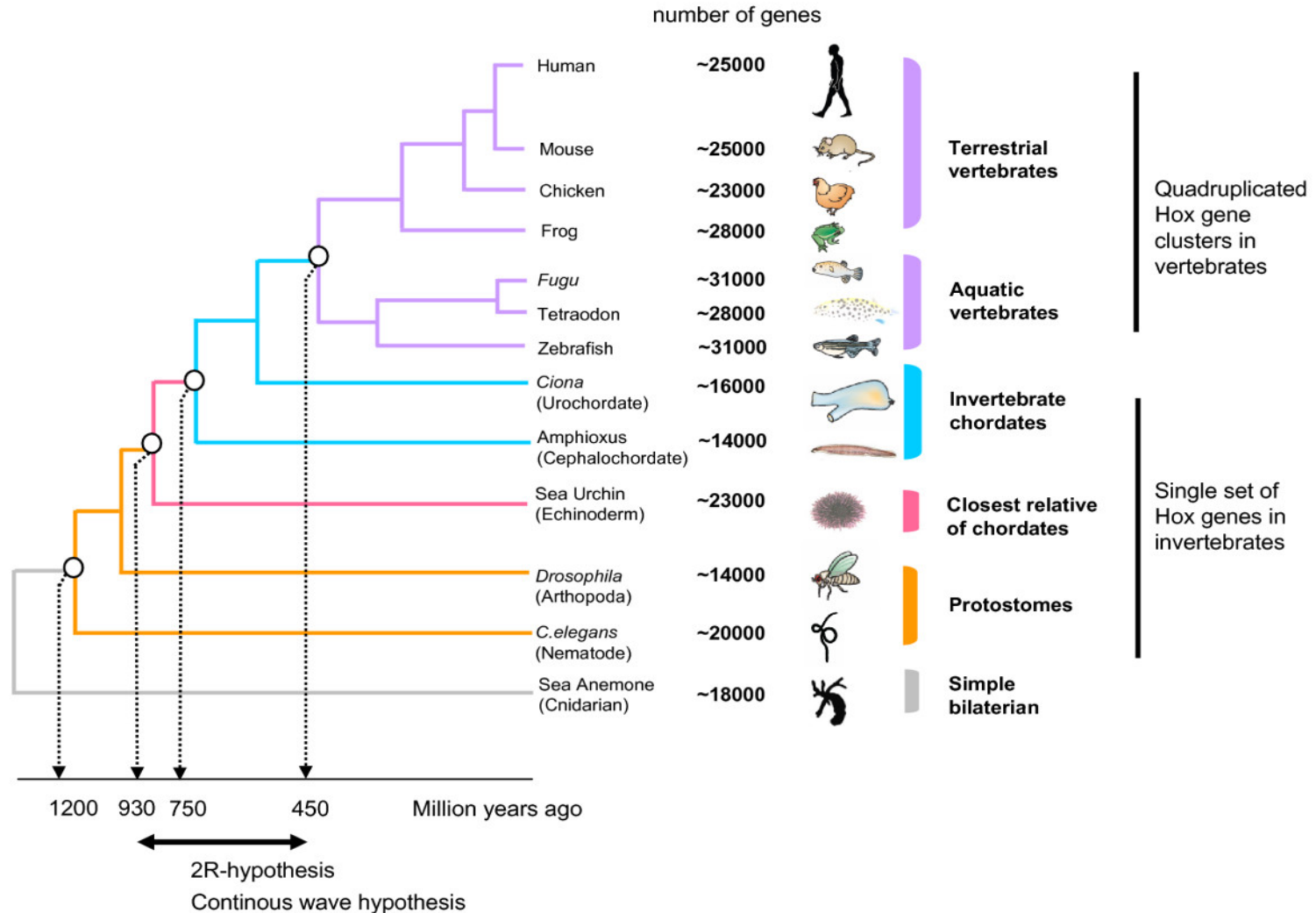
Organism



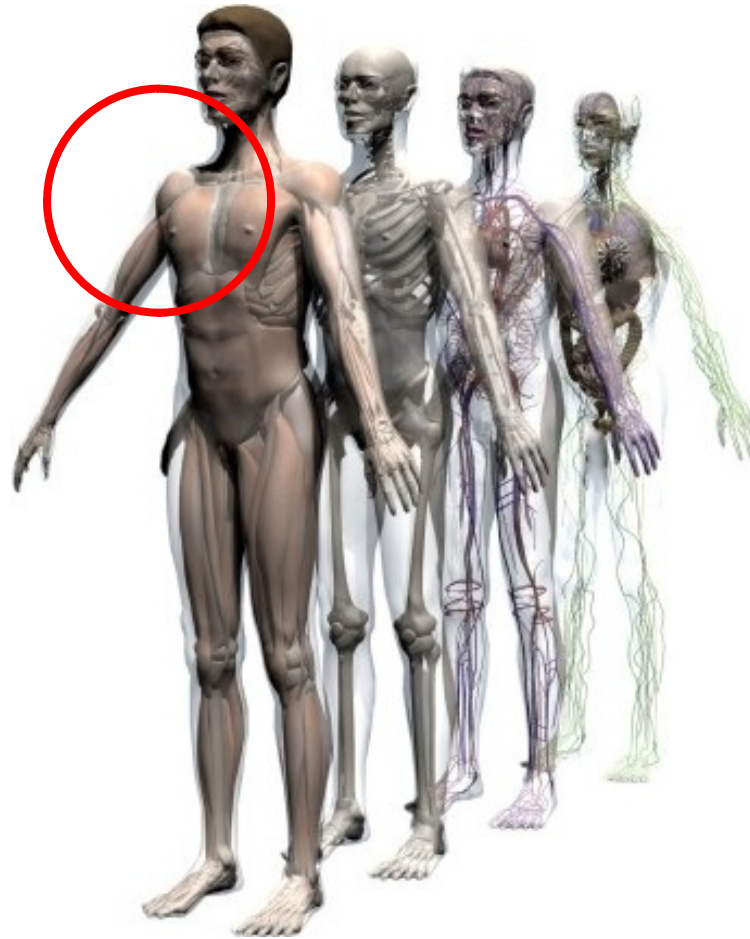
Eukaryote-prokaryote



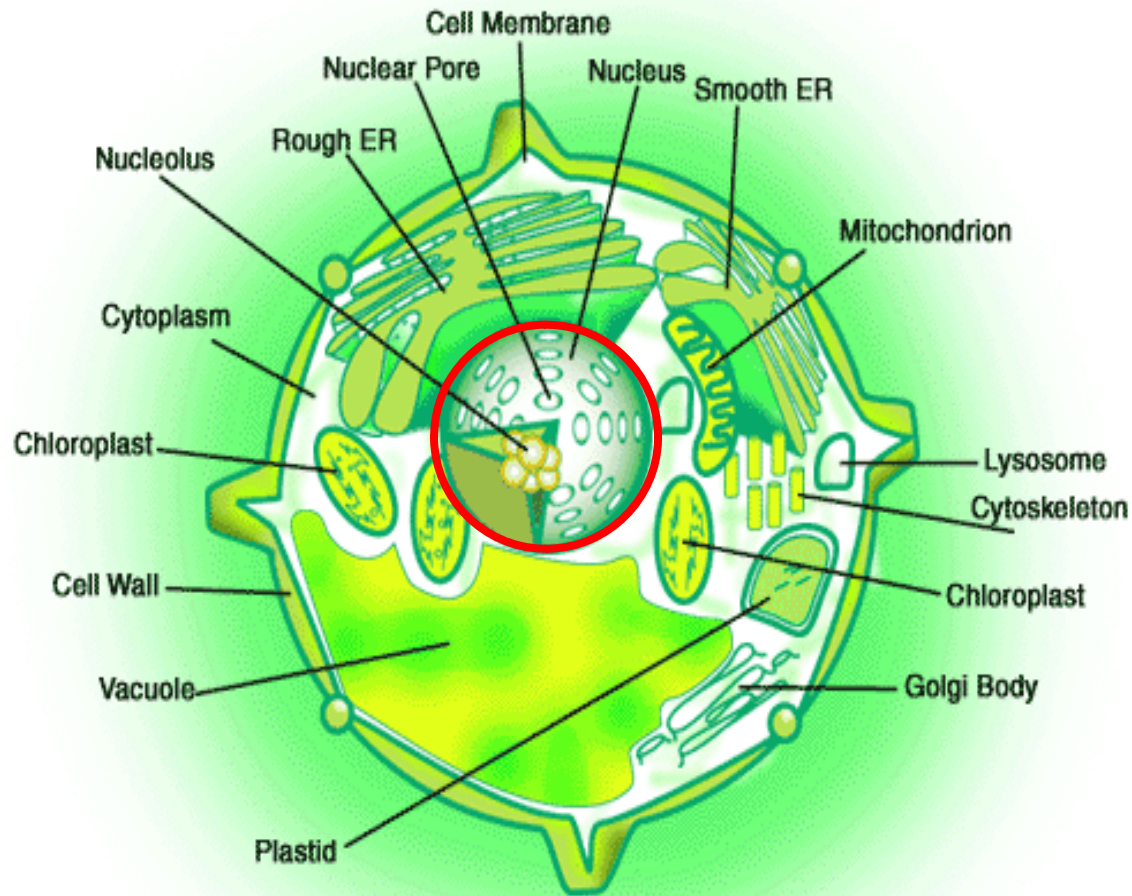
Phylogeny of organisms



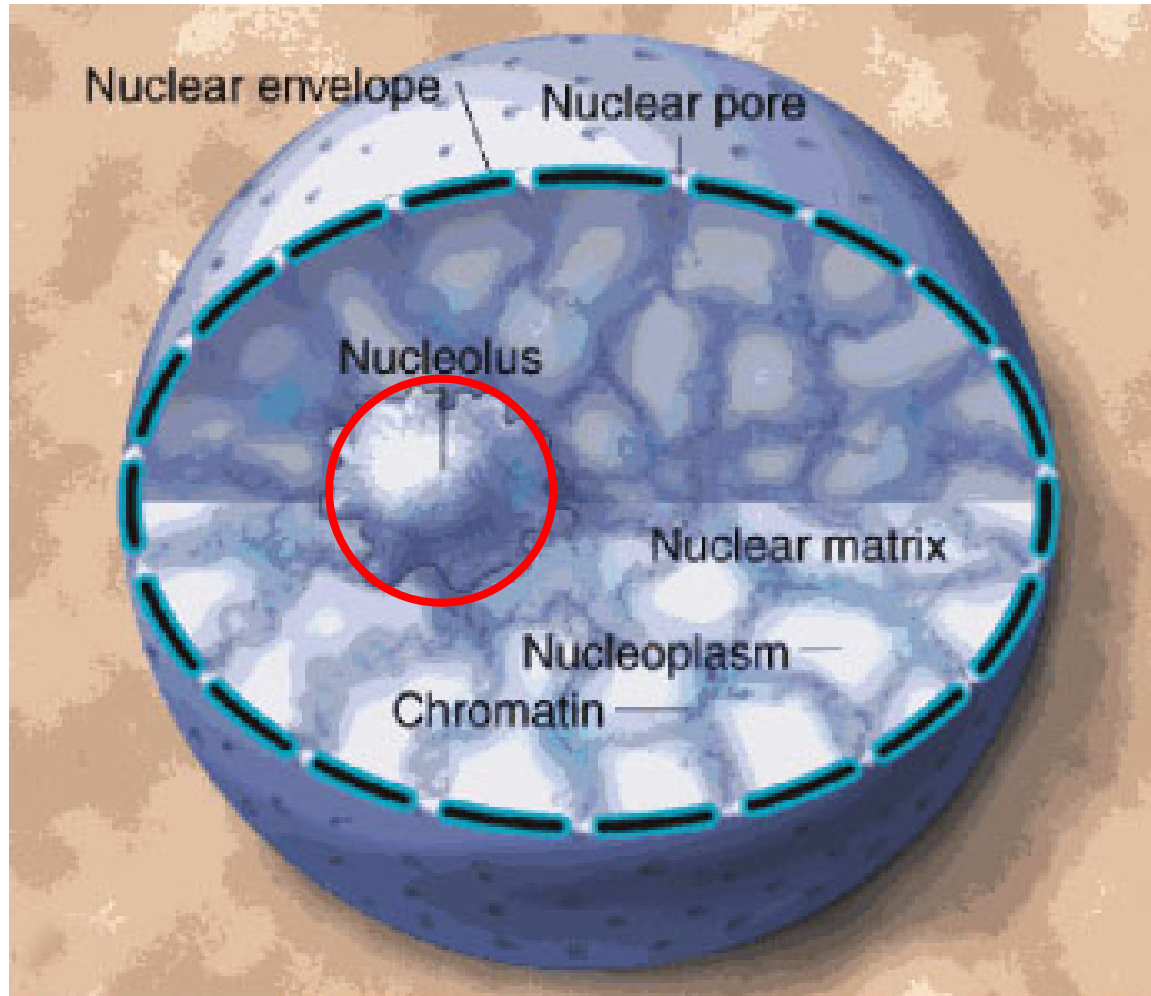
Body



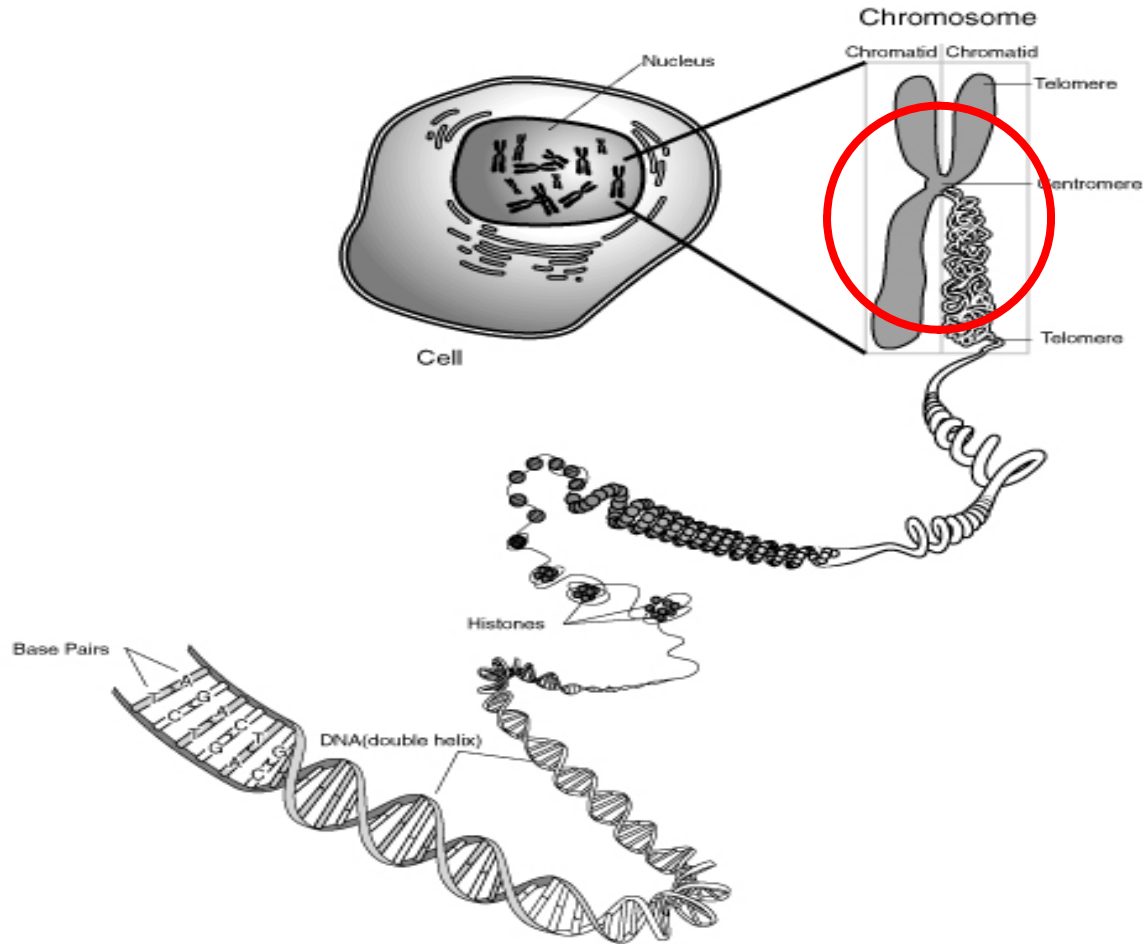
Cell



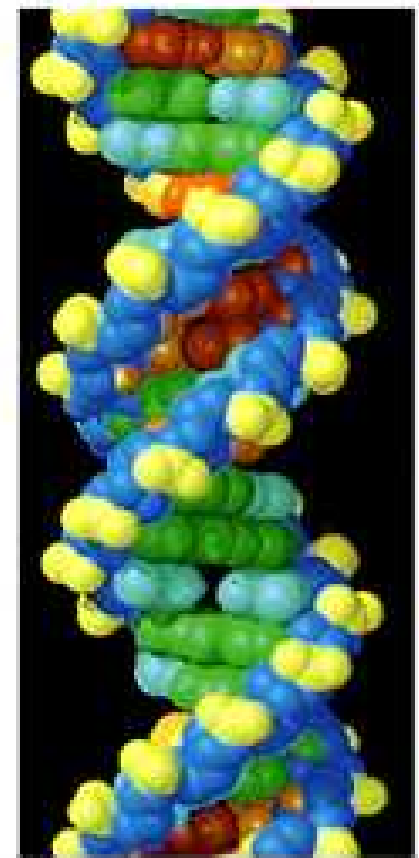
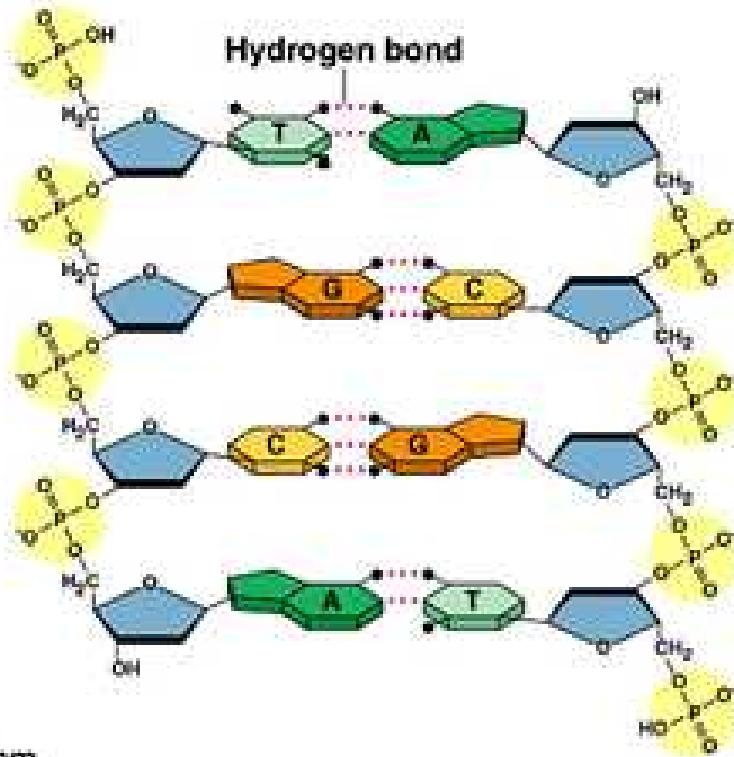
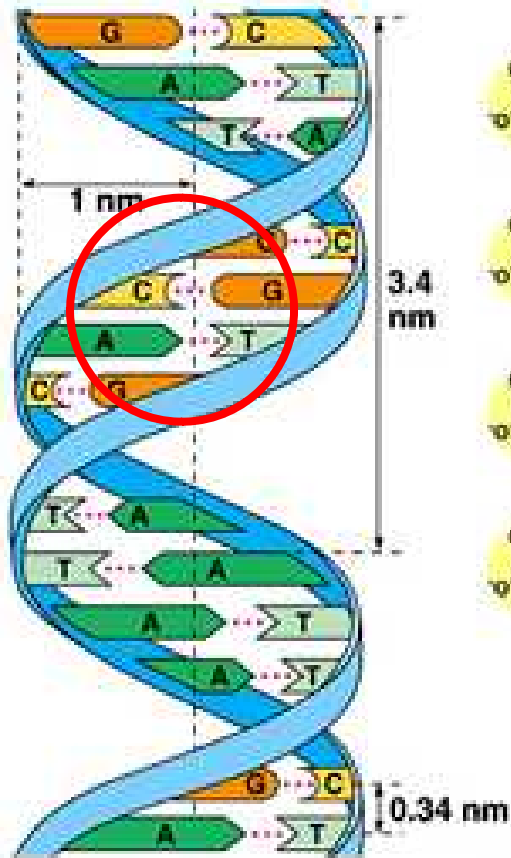
Nucleous



Chromosome



DNA



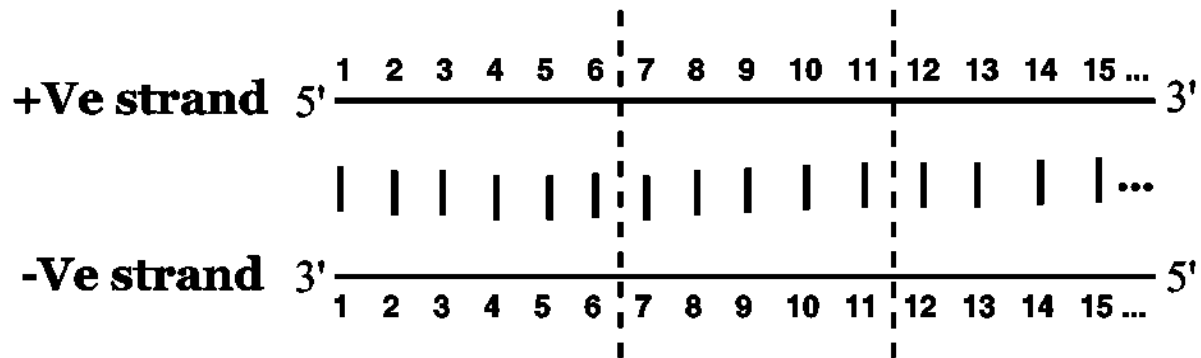
(a)

(b)

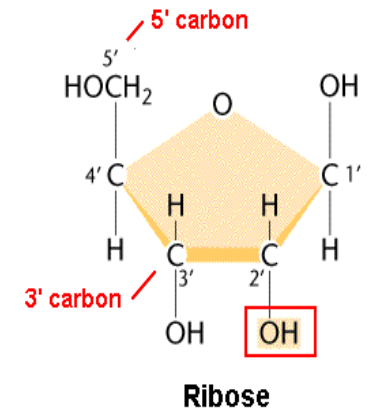
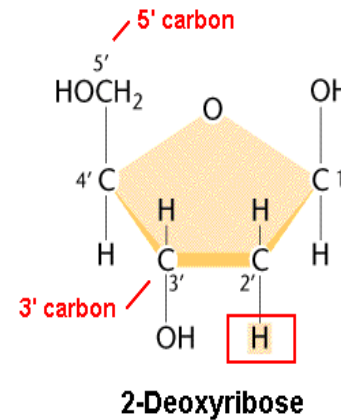
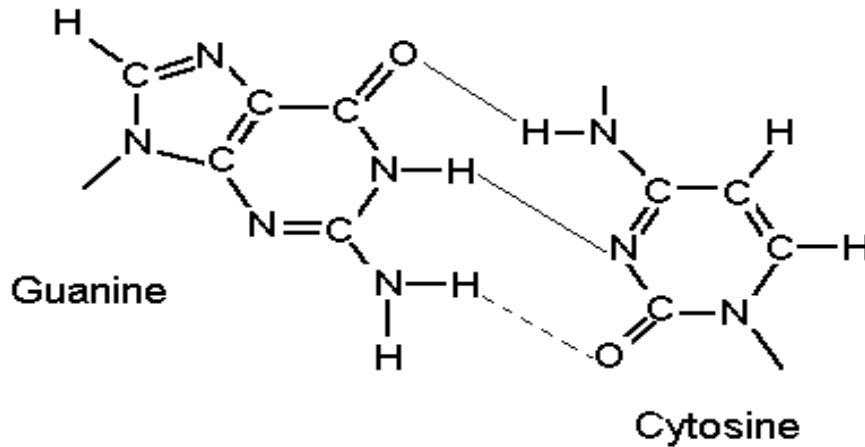
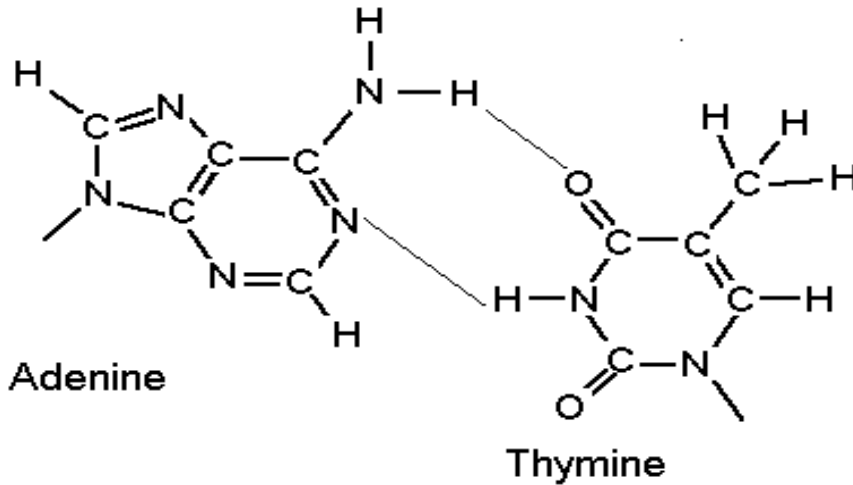
(c)

Copyright © Pearson Education, Inc., publishing as Benjamin Cummings.

Sequence data



Nucleotides & sugar phosphate backbone



(Klug & Cummings 1997)

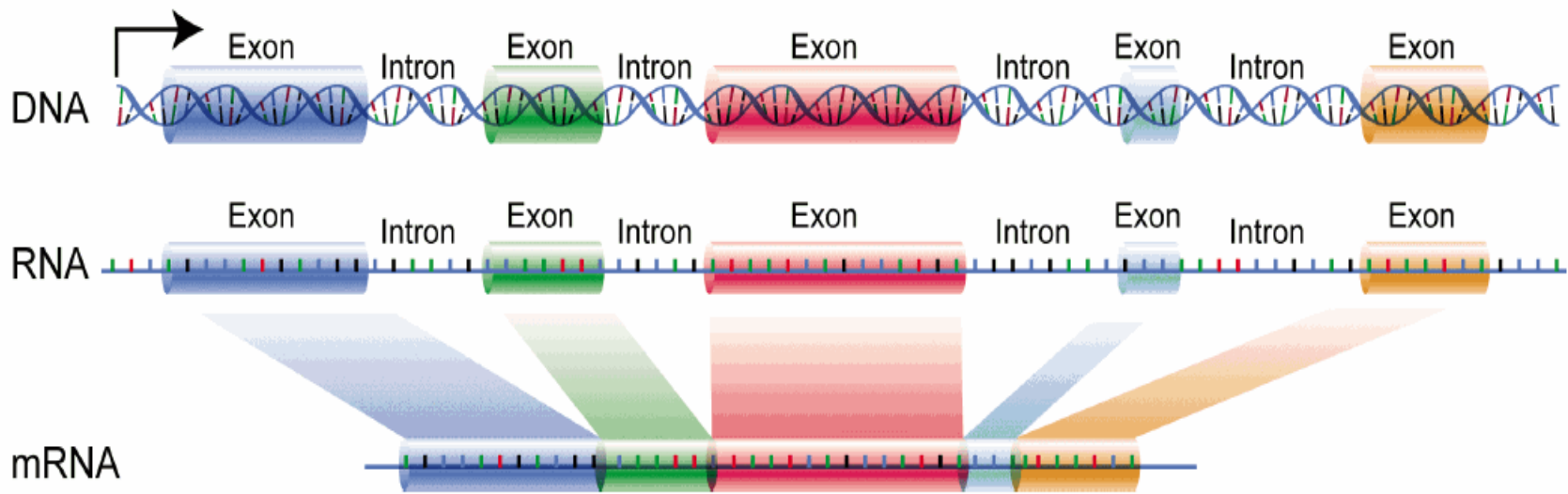
Open problems

- ▶ Aligning multiple DNA/RNA/Protein sequences (**multiple sequence alignment is an NP-hard problem**)
- ▶ Finding the conserved sequences for drug design (**closest substring finding is an NP-complete problem**)
- ▶ Sequence analysis of short length RNAs (**length margin**)

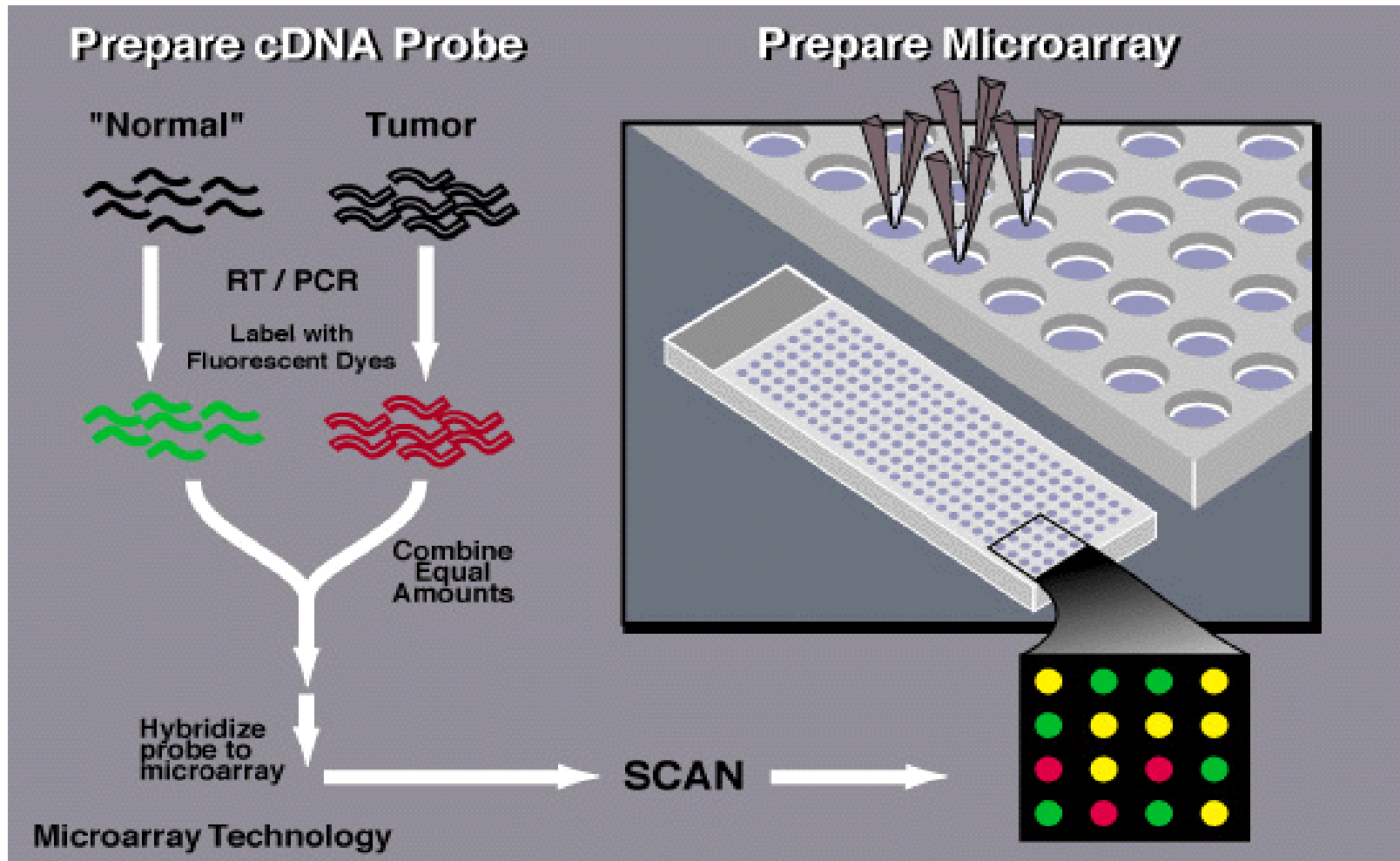
Gene



Gene



Microarray profiling



Snapshot of expression data

Gene	ID	t1	t2	t3	t4	t5	t6	t7
G1	...	1.2	1.9	2.4	3.2	1.1	5.7	7.4
G2	...	3.2	3.9	4.4	5.3	3	7.8	9.5
G3	...	1	2.1	3.2	6.2	7.3	8.5	3.7
...
G1000	...	2.2	3.1	6.3	5.3	8.2	2.5	4.3

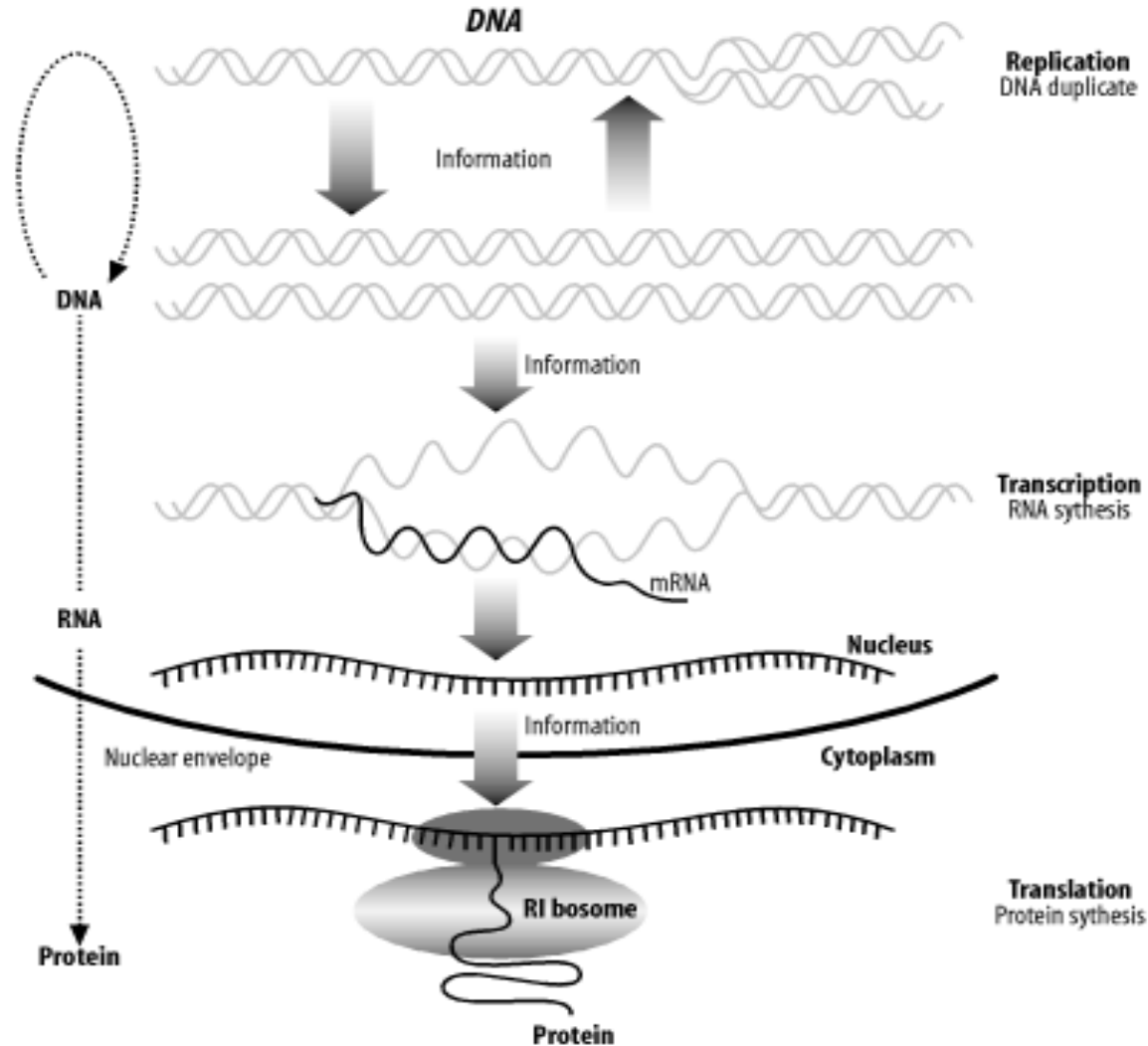
Gene	ID	c1	c2	c3	c4	n1	n2	n3	n4	n5	n6
G1	...	1.2	1.6	1.8	1.1	1	2	1.3	4	2	1.1
G2	...	1.1	1.5	1.3	1.8	2.1	1.1	1.1	1.1	2.3	1.5
G3	...	1.2	1.7	1.8	1.1	2	1.1	2.1	0.8	1.1	1.9
...
G100	...	2.1	1.6	1.7	1.4	2.3	2.7	2.8	2.9	1.3	2.1

Open problems

- ▶ Study on intronic regions (lack of data)
- ▶ Methylation pattern analysis of genes (scarcity of large-scale data)
- ▶ Proper framework for co-expression, contra-expression, differential co-expression, co-expression dynamics (lack of mathematical analysis)

Central dogma of molecular biology

Gene → RNA → protein



Transcription

Introducing: Transcription

DNA



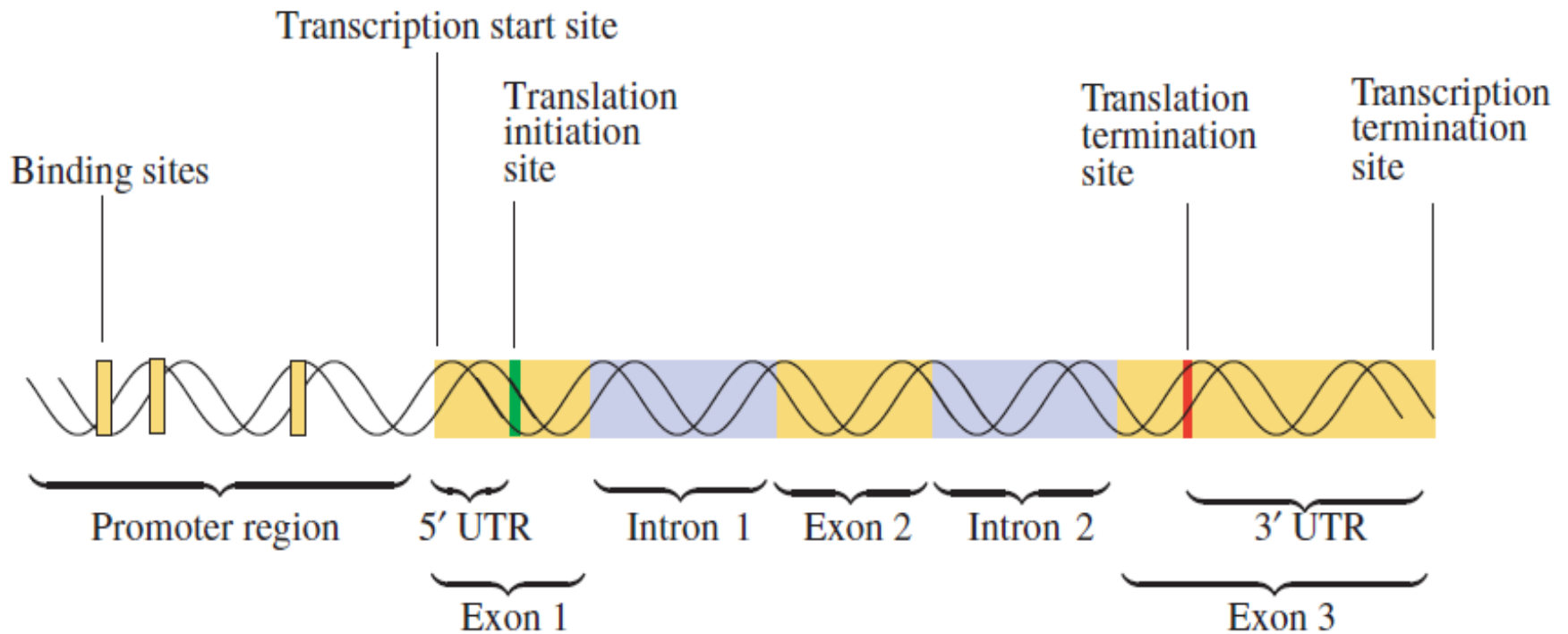
Translation

Introducing: Translation

mRNA



Promoter



Types of promoters

▶ Core promoter

- ▶ RNA polymerase binding site (within 1 kb from the upstream)
 - Transcription start site
 - Pol I transcribes genes encoding rRNA
 - Pol II transcribes genes encoding mRNA, miRNA, etc.
 - Pol III transcribes genes encoding tRNA, short RNAs, etc.

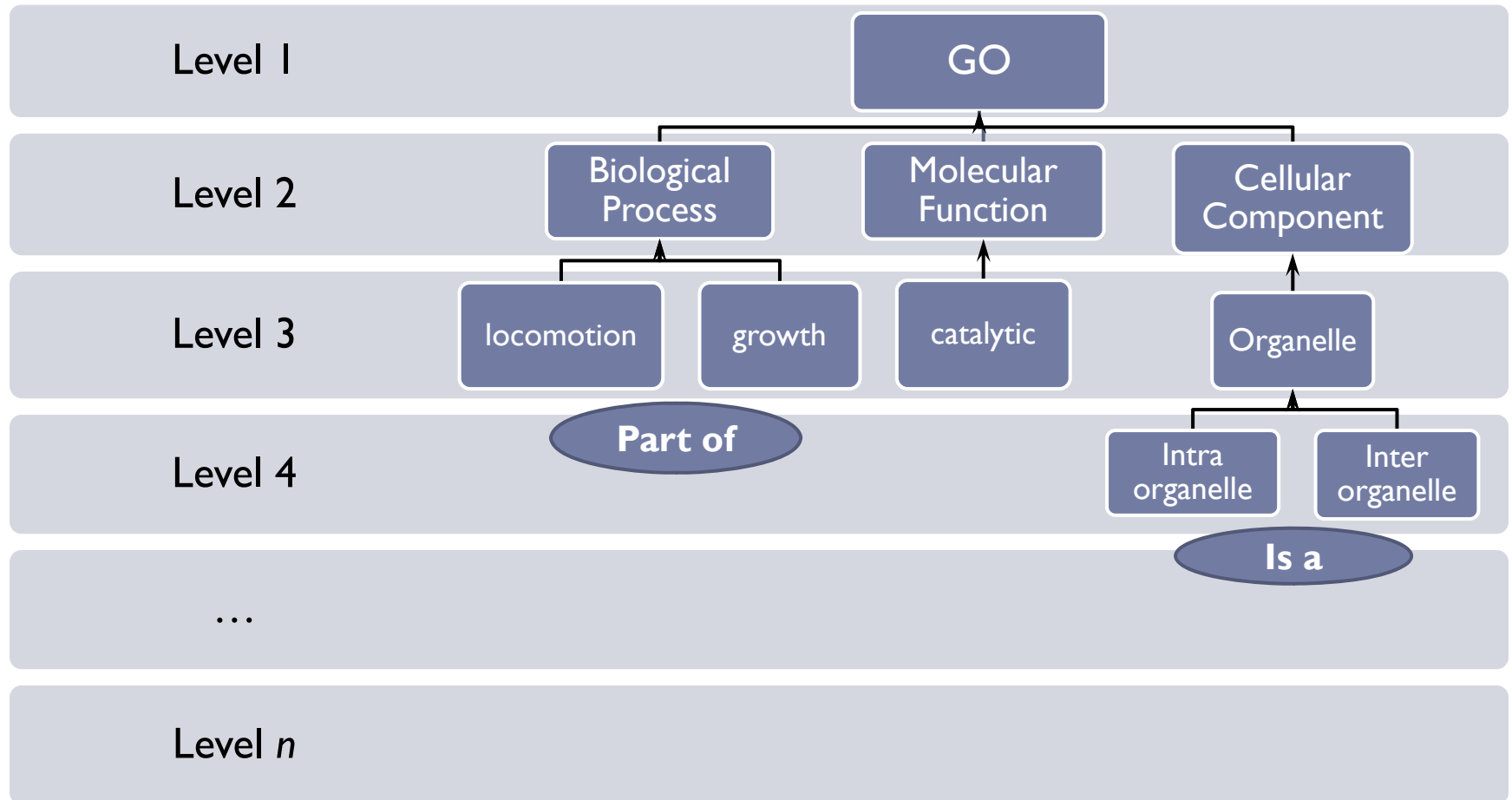
▶ Proximal promoter

- ▶ Transcription factor binding site (within 2-3 kb from the upstream)

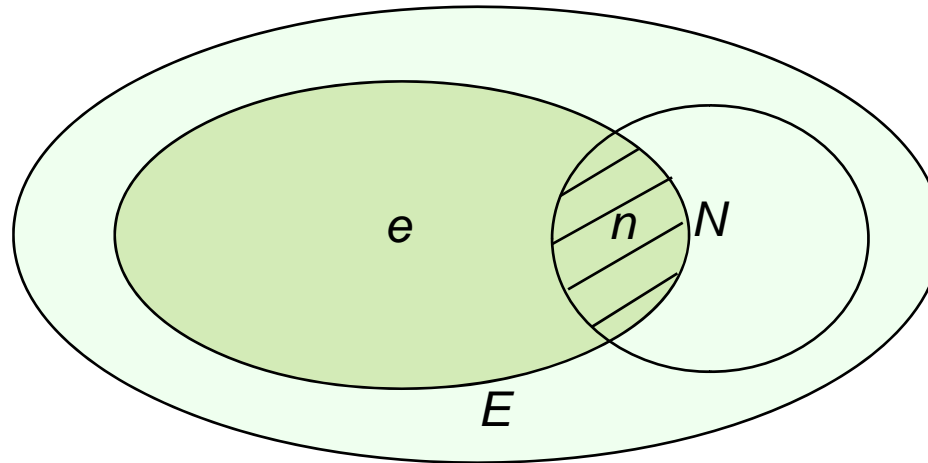
▶ Distal promoter

- ▶ Specific transcription factor binding site (within 10kb from the upstream)

Gene ontology



p -value



$$p\text{-value} = \sum_{i=n}^{\min(e, N)} \frac{\binom{e}{i} \binom{E-e}{N-i}}{\binom{E}{N}}$$

Open problems

- ▶ Transcription termination protocol (lack of information)
- ▶ Quantifying the biological significance of a gene cluster (proper mathematical framework is lacking)

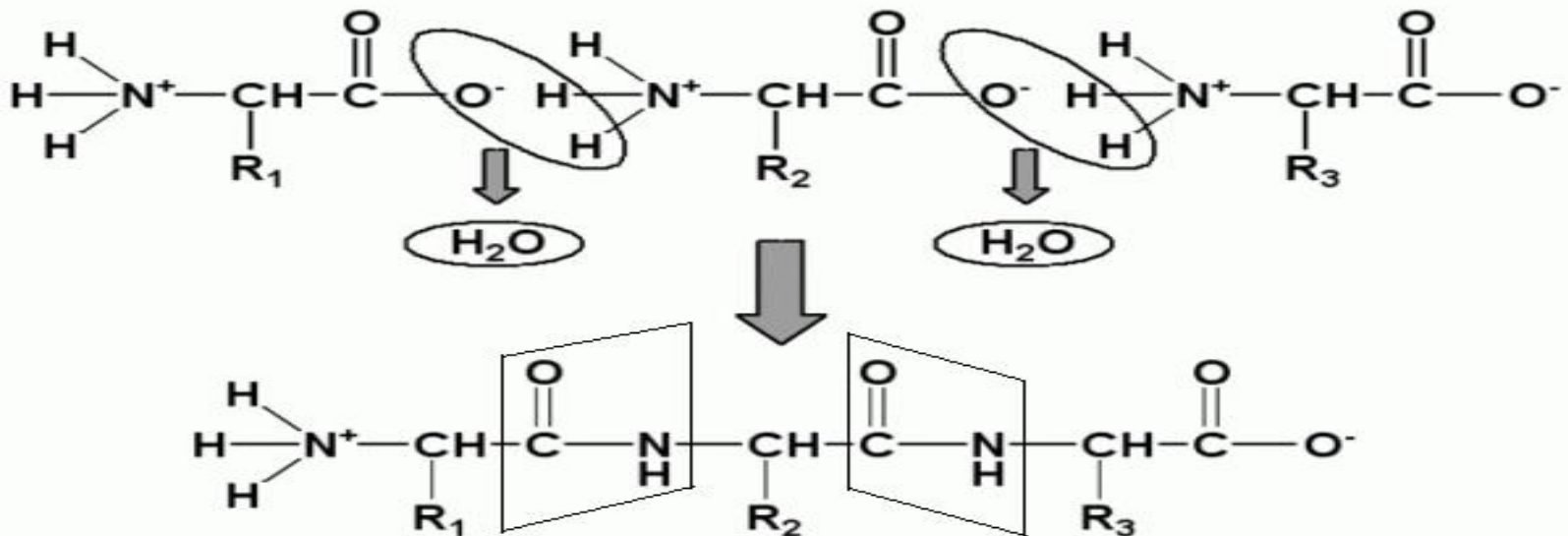
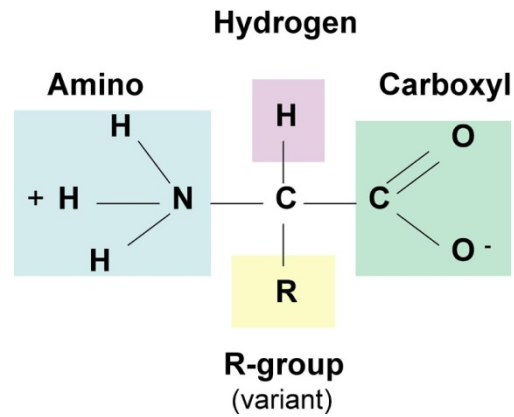
Protein

Types of proteins

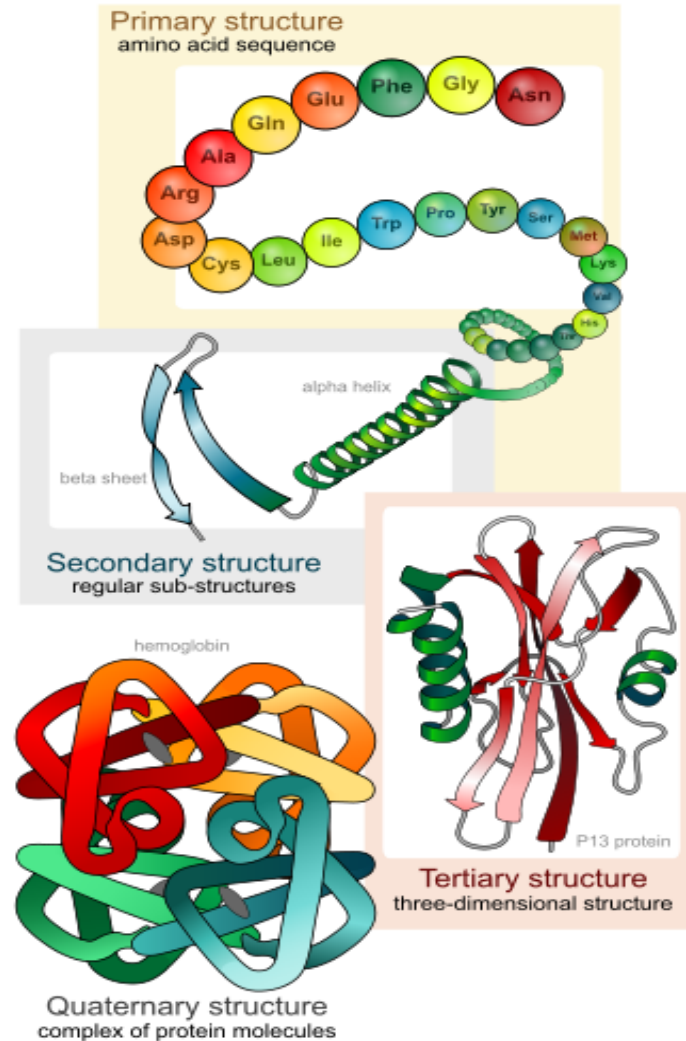
- ▶ **Structural**
 - ▶ Keratin, Actin, Fibrillin, etc.
- ▶ **Functional**
 - ▶ Ciliary neurotrophic factor, Erythropoietin, etc.
- ▶ **Enzymes**
 - ▶ Oxidoreductases, Transferases, Hydrolases, etc.

Protein

Amino Acid Structure



Protein structure

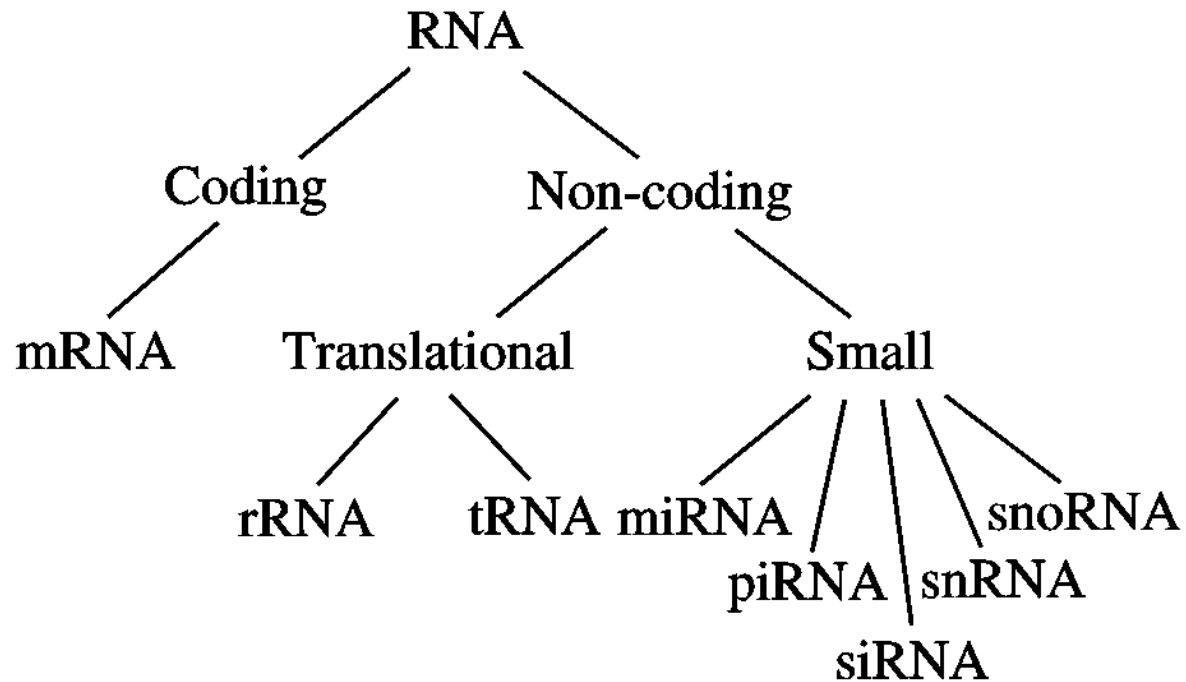


Open problems

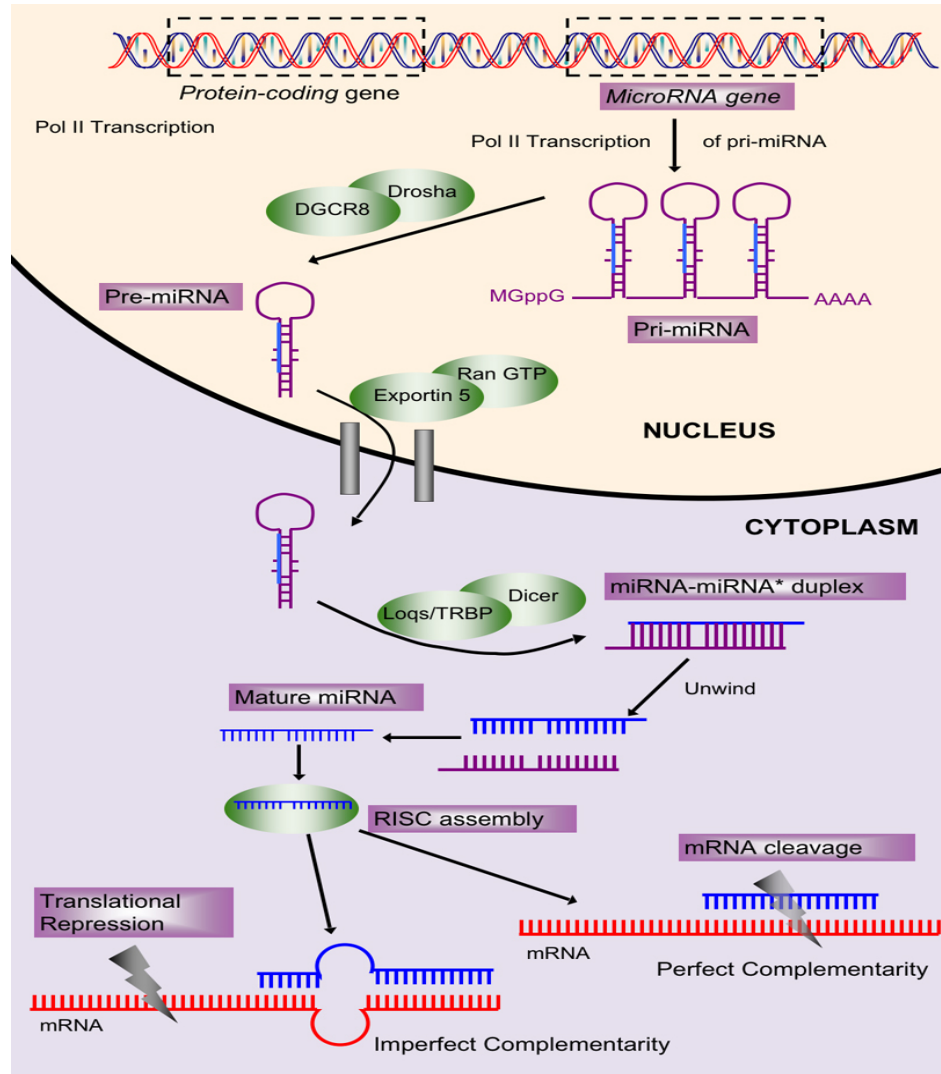
- ▶ Drug design (the only approach is hypothetical analysis)
- ▶ Host-viral protein interaction network analysis for diseases (lack of data for many diseases)

MicroRNA

RNA tree

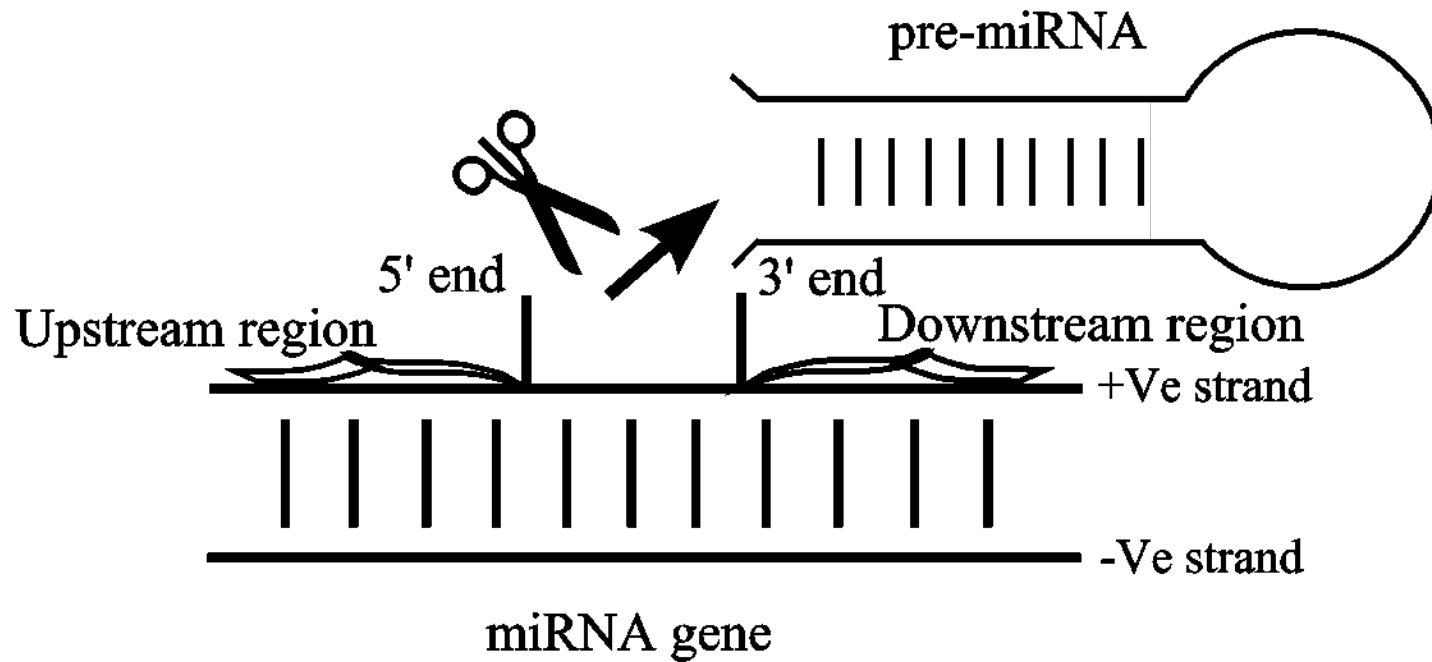


Biogenesis of miRNAs



[Youtube](#)

Regulation



Open problems

- ▶ Regulation of miRNAs by transcription factors (**lack of proper information fusion**)
- ▶ The transcriptional protocol for different categories of miRNAs (**scarcity of data**)

References

1. J. Setubal and J. Meidanis, *Introduction to Computational Molecular Biology*, PWS publishing company, Boston, 1999.
2. M. S. Waterman, *Introduction to Computational Biology: Maps, Sequences and Genomes*, Chapman & Hall, London, 1995.
3. B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, *Molecular Biology of the Cell*, Fourth Edition, Taylor & Francis Group, 2002.

Thank you