Computational Molecular Biology and Bioinformatics

Basics of Molecular Biology

Malay Bhattacharyya

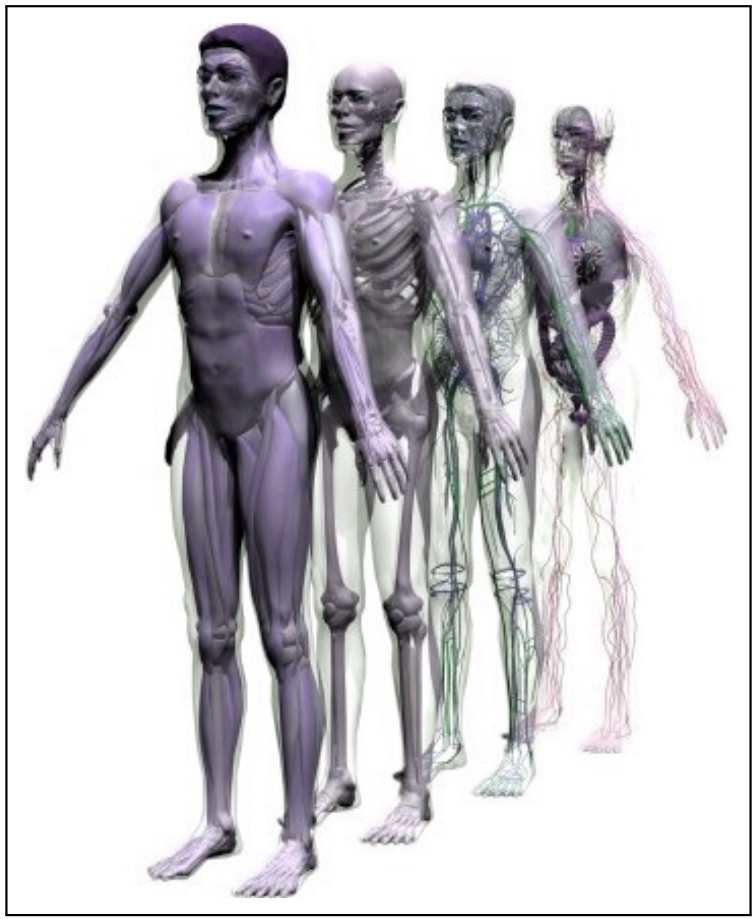
Associate Professor

Machine Intelligence Unit
Indian Statistical Institute, Kolkata

July, 2024

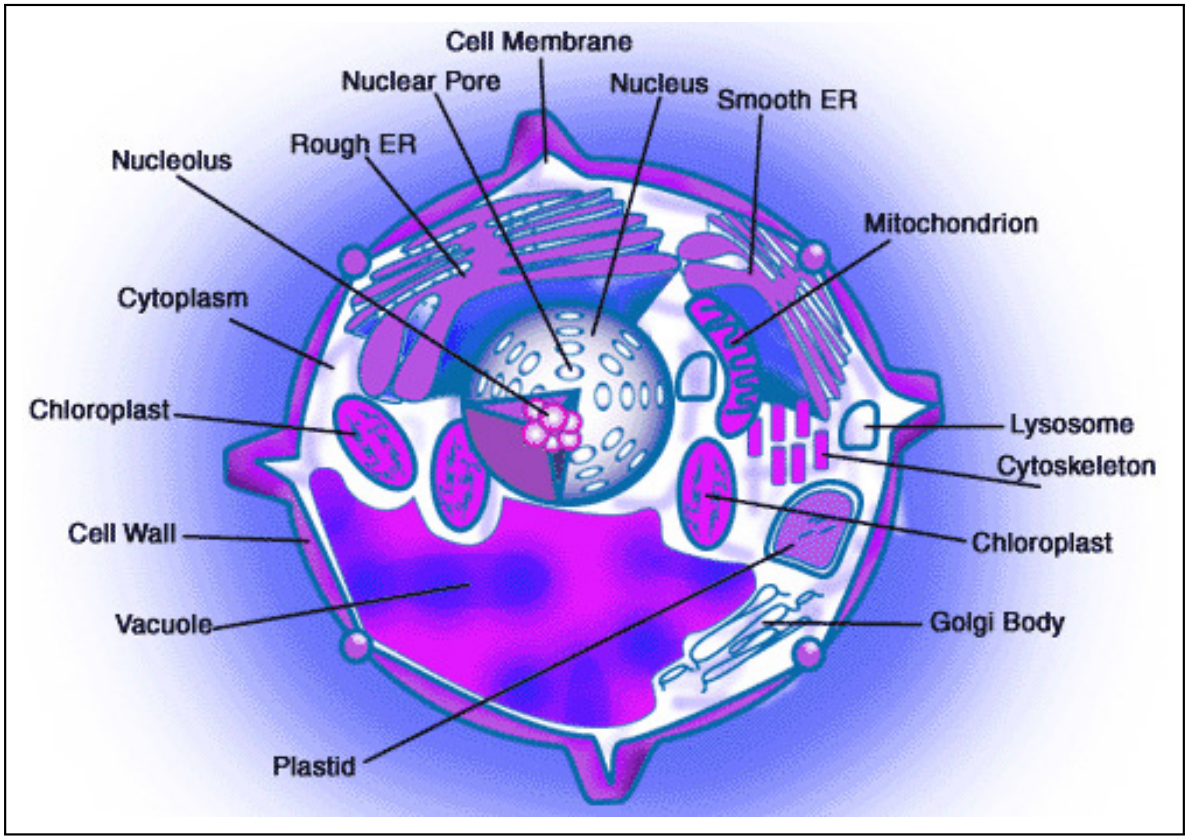
Imagine yourself having such eyes

Organism to cell



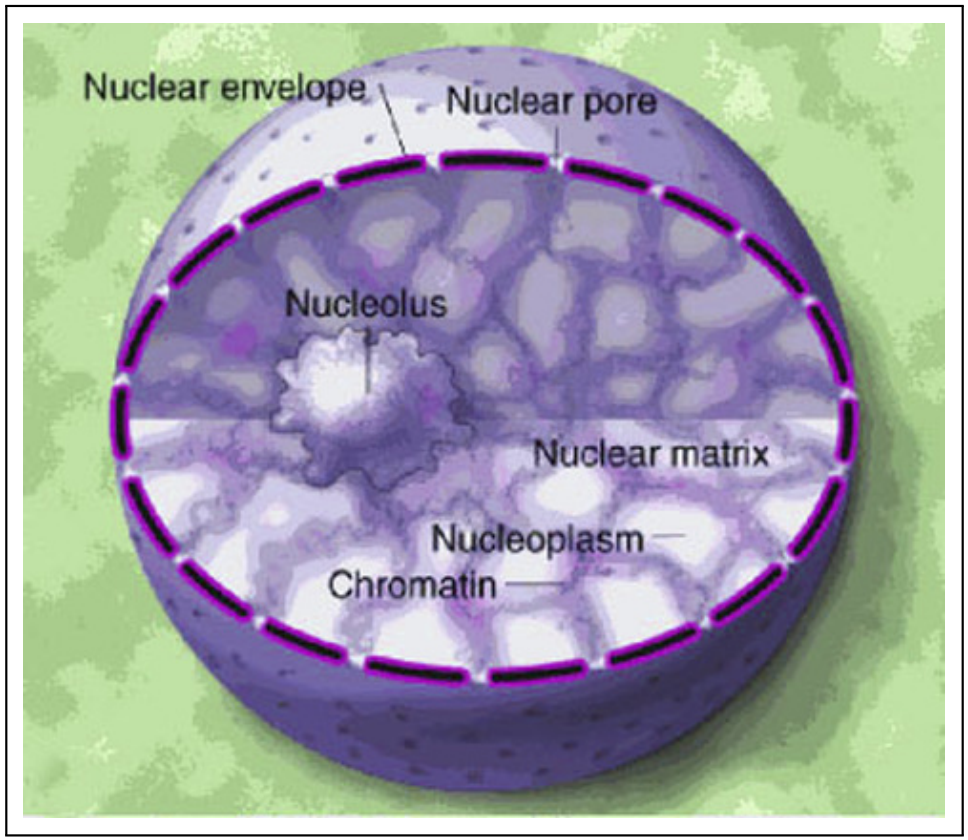
Zooming from the body of an organism to the cell

Cell to nucleus



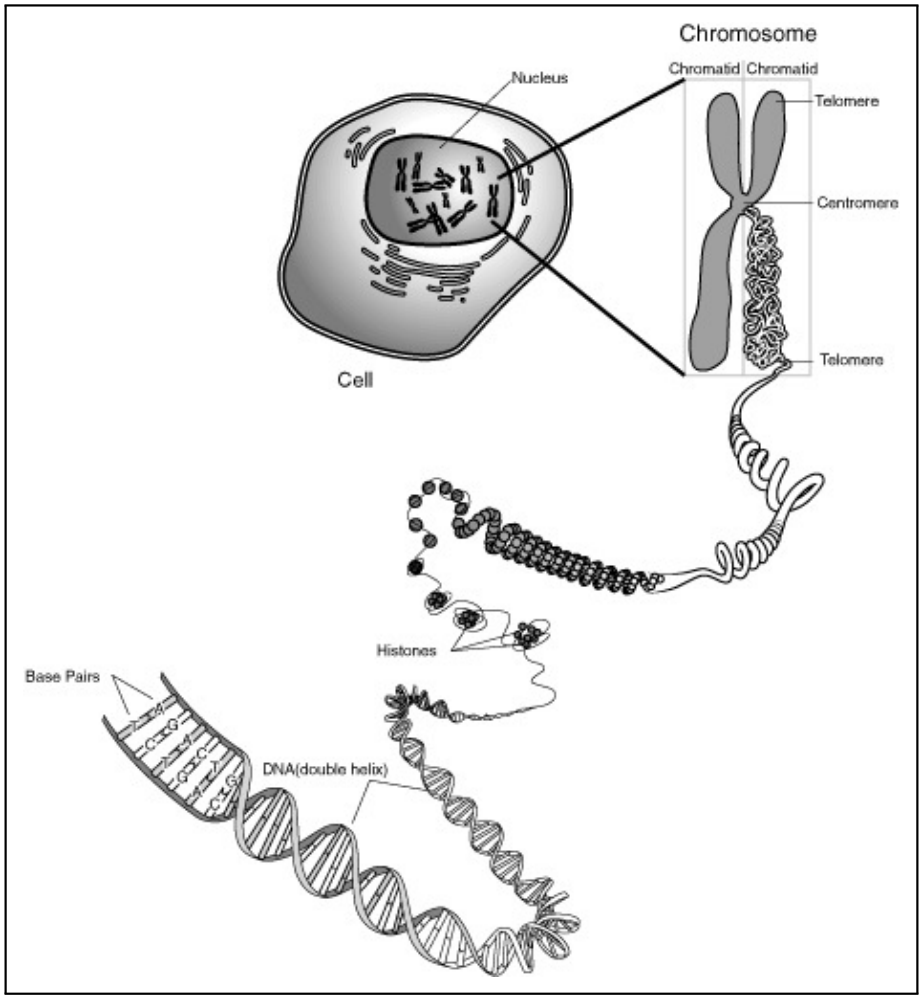
Zooming from cell into the nucleus

Nucleus to chromosome



Zooming from nucleus into the chromosome

Chromosome to DNA



Zooming from chromosome into the DNA

DNA to nucleotide

...	T	T	A	C	C	G	T	C	G	G	C	A	...
	|	|	|	|	|	|	|	|	|	|	|	|	
...	A	A	T	G	G	C	A	G	C	C	G	T	...

Note: There are 3 billion base pairs in each cell to fit into a space of approximately 6 μm across the human cells.

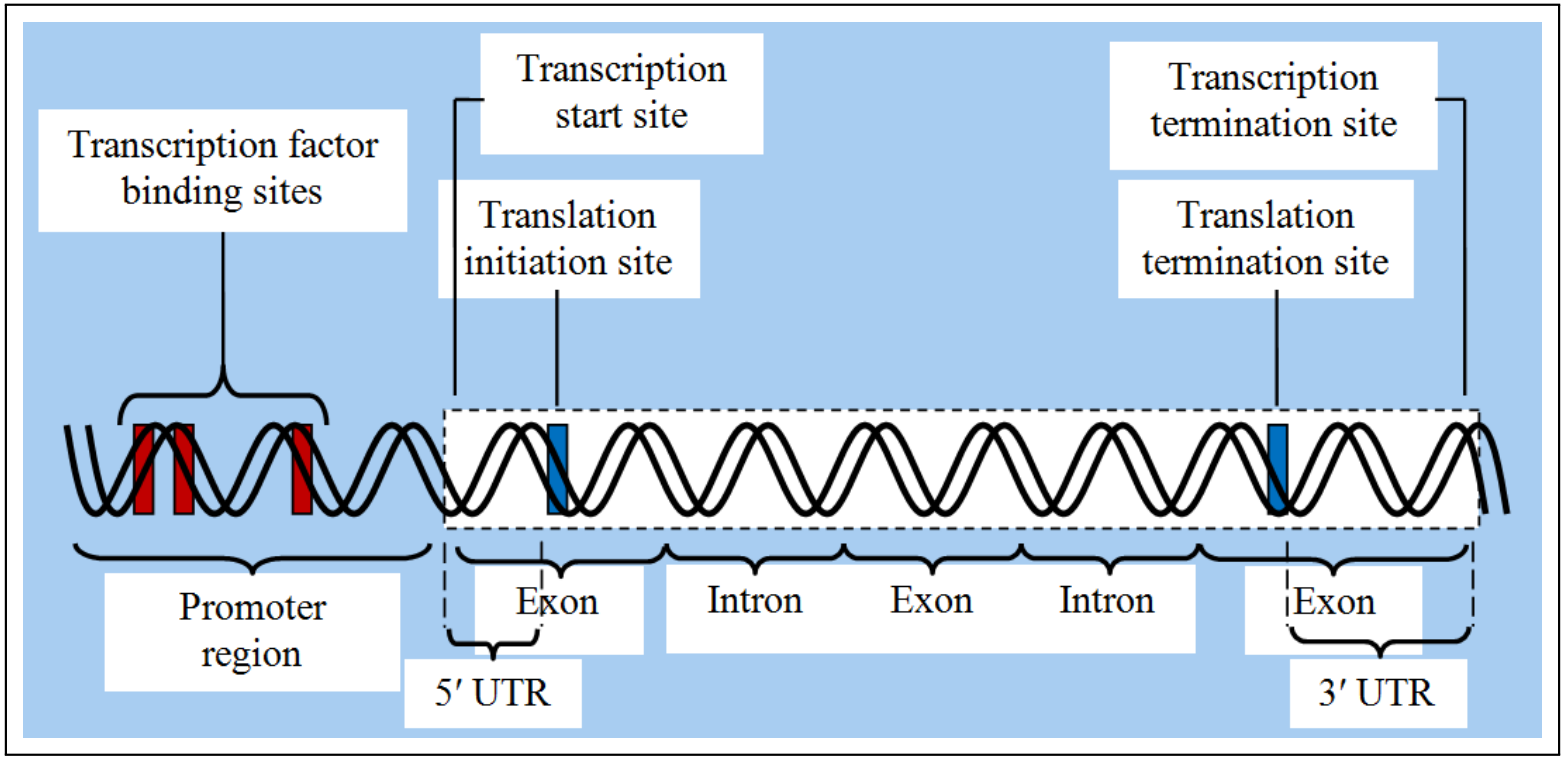
This is your haystack!!!

. . . AATTTCGGATTGCCAATTCGGAAAAATTTCGGATTGCCAA
TTCGGTTTCGGATTGCCAATTCGGAAAAGCCAATTCGGAAAAC
AAATTGCCAATTTTGCCAATTCGGAAAAATTTCGGATTGCCAA
TTCGGTTTCGGATTGCCAATTCGGAAAAGCCAATTCGGAAAAC
AAAAATTTCGGATTGCCAATTCGGAAAAATTTCGGATTGCCAA
TTCGGCGGATTATTGCCAATTCTTCAAAGCCAATTCGGAAAAC
CAAAATTTCGGATTGCCAATTCGGAAAAATTTCGGATTGCCAA
TTCGGTTTCGGATTGCCAATTCGGAAAAGCCAATTCGGAAAAC
TTCAATTTCGGATTGCCAATTCGGAAAAATTTCGGATTGCCAA
GGGAATTTCGGATTGCCAATTCGGAAAAATTTCGGATTGCCAA
ATCGGTTTCGGATTGCCAATTCGGAAAAGCCAATTCGGAAAAC
GAAAATTTCGGATTGCCAATTCGGAAAAATTTCGGATTGCCAA
ATCGGTTTCGGATTGCCAATTCGGAAAAGCCAATTCGGAA . . .

This is your haystack!!!

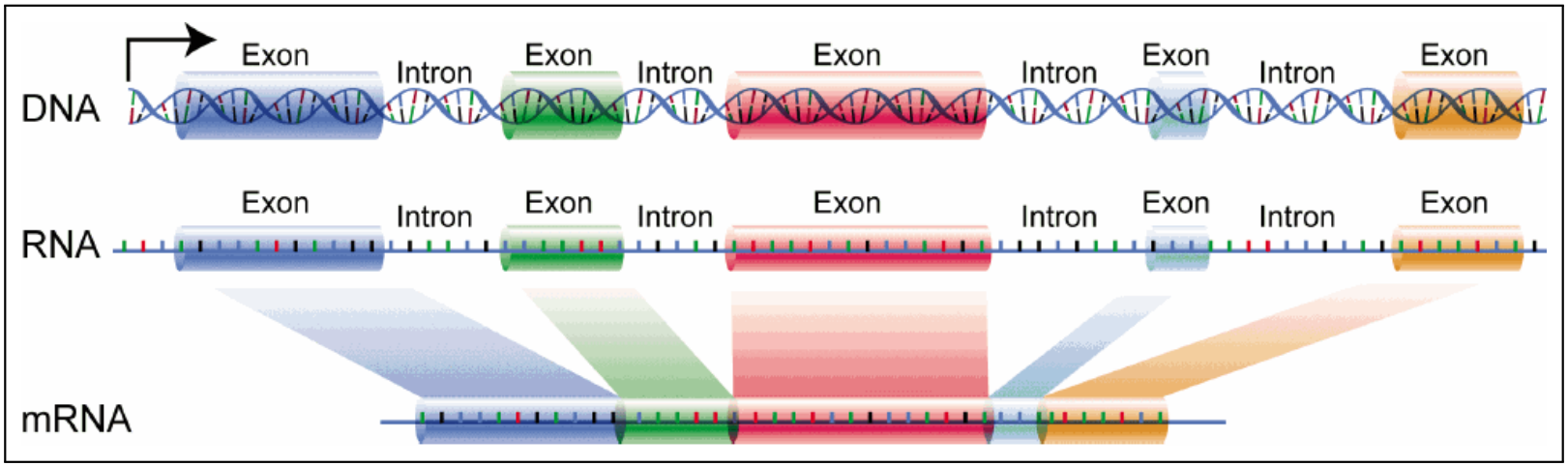
. . . AATTTCGGATTGCCAATTCGGAAAAATTTCGGATTGCCAA
TTCGGTTTCGGATTGCCAATTCGGAAAAGCCAATTCGGAAAAC
AAATTGCCAATTTTGCCAATTCGGAAAAATTTCGGATTGCCAA
TTCGGTTTCGGATTGCCAATTCGGAAAAGCCAATTCGGAAAAC
AAAAATTTCGGATTGCCAATTCGGAAAAATTTCGGATTGCCAA
TTCGGCGGATTATTGCCAATTCTTCAAAGCCAATTCGGAAAAC
CAAAATTTCGGATTGCCAATTCGGAAAAATTTCGGATTGCCAA
TTCGGTTTC Blue eyes??? CGGAAAAGCCAATTCGGAAAAC
TTCAATTTCGGATTGCCAATTCGGAAAAATTTCGGATTGCCAA
GGGAATTTCGGATTGCCAATTCGGAAAAATTTCGGATTGCCAA
ATCGGTTTCGGATTGCCAATTCGGAAAAGCCAATTCGGAAAAC
GAAAATTTCGGATTGCCAATTCGGAAAAATTTCGGATTGCCAA
ATCGGTTTCGGATTGCCAATTCGGAAAAGCCAATTCGGAA . . .

Transcription unit and promoter



Organization of the transcription unit and promoter region

DNA-RNA-mRNA

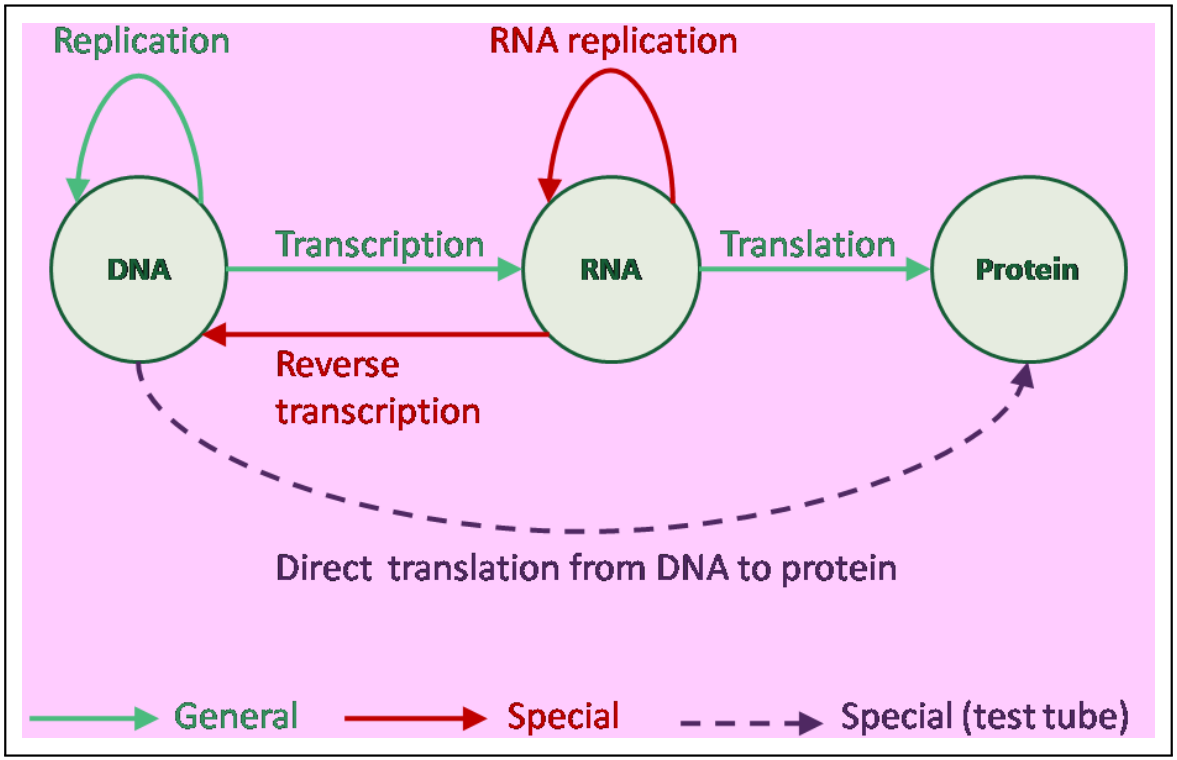


The formation of DNA-RNA-mRNA

Note: The nucleotides A, T, C and G constitutes the DNA, whereas in RNA T is replaced by U.

Central dogma of molecular biology

The central dogma of molecular biology comprises transcription followed by translation.



Scope of research

"Biology is the only science in which multiplication is the same thing as division"

– Jupiter Scientific.

Molecular biology, which serves as a link between the genetics and biochemistry, covers the following areas:

- Sequence analysis
- Expression analysis
- Genetic analysis
- Epigenetic analysis
- System-level analysis
- Pathway analysis

Resources

Books:

1. C. Setubal and J. Meidanis: Introduction to Computational Molecular Biology, PWS Publishing Company, Boston, 1997.
2. P. A. Pevzner: Computational Molecular Biology – An Algorithmic Approach, MIT Press, 2000.
3. R. Durbin, S. R. Eddy, A. Krogh and G. Mitchison: Biological Sequence Analysis - Probabilistic Models of Proteins and Nucleic Acids, Cambridge University Press, 1998.
4. D. Gusfield: Algorithms on Strings, Trees, and Sequences, Cambridge University Press, USA, 1997.
5. H. Lodish, A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore and J. Darnell: Molecular Cell Biology, W. H. Freeman, USA, 2000.
6. C.-I. Branden, J. Tooze: Introduction to Protein Structure, Garland Publishing, 1998.
7. A. Kowald, C.hristoph Wierling, E. Klipp, and W. Liebermeister: Systems Biology, Wiley-VCH, 2016.
8. B.O. Palsson: Systems Biology – Constraint based Reconstruction and Analysis, Cambridge Univer- sity Press, 2015.

Resources

Journals:

1. WIREs Computational Molecular Science, Wiley.
2. Briefings in Bioinformatics, Oxford University Press.
3. Bioinformatics, Oxford University Press.
4. PLoS Computational Biology, PLoS.
5. IEEE/ACM Transactions on Computational Biology and Bioinformatics, IEEE-ACM.

Conferences:

1. ISMB/ECCB
2. RECOMB

Resources

Other similar courses:

1. Serafim Batzoglou at Stanford – `https://web.stanford.edu/class/cs262`
2. Manolis Kellis at MIT – `https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-047-computational-biology-fall-2015`
3. Christopher Burge, David Gifford and Ernest Fraenkel at MIT – `https://ocw.mit.edu/courses/biology/7-91j-foundations-of-computational-and-systems-biology-spring-2014`
4. Curtis Huttenhower at Harvard – `https://canvas.harvard.edu/courses/71191`

Evaluation criteria

1. **End-semester Evaluation:**
   - Written examination (50 marks)
2. **Internal Evaluation:**
   - Mid-semester examination (30 marks)
   - Assignment submission (10 marks)
   - Project work (10 marks)

Web: https://www.isical.ac.in/ malaybhattacharyya/Courses/CMBB/Fall2024