

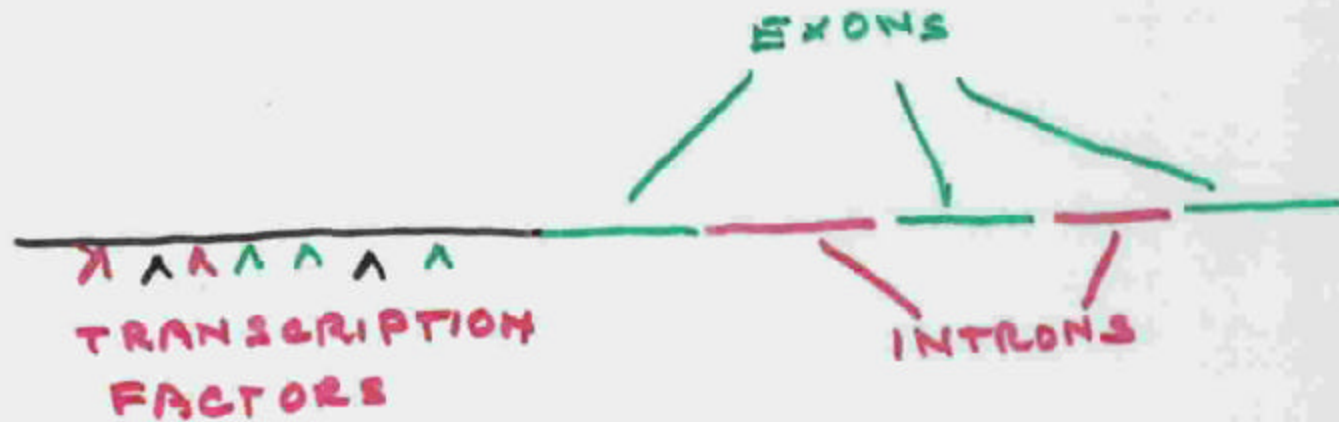
Genomics Made Easy

- **Genome:** set of DNA sequences present in every cell which carry the hereditary information and direct the production of proteins
- **Transcription of genes:** DNA to mRNA
- **Translation:** mRNA to protein, using universal triplet code

Proteins

- Sequences of amino acids
- Three-dimensional structures
- Structural elements of cells
- Catalysts of chemical reactions
- Carriers of intercellular communication
- Components of molecular machines
- Regulators of transcription and translation

Structure of a Gene



Computational Gene Finding

- Based on deterministic and stochastic features:
 - Predictable sequences at exon/intron boundaries
 - Constraints imposed by triplet code
 - Different nucleotide frequencies in exons, introns, intergenic regions
- Algorithm: stochastic finite-state automaton (Hidden Markov Model)

Structure of Regulatory Regions

- Transcription factor (TF): protein that enhances or inhibits transcription of a gene by binding to nearby DNA
- Binding site: site where a TF binds
- Binding sites for a given TF can be described by a “probabilistic word”

Regulatory Module

- Region containing binding sites for several TFs. Depending on which TFs bind, the module may issue an excitatory or inhibitory signal.
- The level of expression of the gene is determined by the signals from the modules

Analysis of Transcriptional Regulation

- Recognize binding sites for given TF
 - Stochastic model
 - Parameters: distribution of nucleotides in each position
 - Tendency for highly conserved positions to cluster

Architecture of Transcriptional Regulation

- Recognize regulatory modules
- Determine how the signals from the modules determine the expression level
- Find common features of the organization of transcriptional regulation in a set of related organisms

Uses of Phylogenetic Trees

- Given phylogenetic tree for a set of related organisms can study critical evolutionary events that have influenced:
 - Complex traits (organization of eye)
 - Developmental program
 - Genes and their functions
 - Mechanisms of gene regulation
 - Function and structure of proteins

Evolutionary Genomics

- Organisms are best understood in terms of their evolutionary history
- Conserved genomic regions among a set of related organisms tend to have functional significance.
- Example: genes tend to be conserved among man, mouse, rat and chimpanzee