

## CS 262—Lecture 1 Notes

- 4-5 HWs, 3 late days
- (Optional) scribing: Due week after lecture; Replace 2 lowest problems from each problem set
- Optional textbook: *Biological Sequence Analysis* by Durbin
- Computational genomics started taking off around the time of the sequencing of the Human Genome in 2000
  
- Central Dogma of Molecular Biology
  - DNA in nucleus of eukaryotic cells encodes genes → transcription leads to RNA in cytoplasm → Translation leads to proteins, the building blocks for life
- Goal for 21<sup>st</sup> century: Convert biology from descriptive science to quantitative, predictive sciences
- Price of genome drastically dropping: \$30,000,000 in 2004 to \$1000 today

### Intro to Biology

- DNA in nucleus and each cell has copies of DNA
  - Although we have trillions of cells, an average somatic cell only has 30-60 differences from the “pure”, original genome
- DNA packed into chromosomes; DNA wrapped around histones, which are wrapped around nucleosomes, which form chromosomes
  - Multiple levels of packing
- 4 nucleotide bases: Adenine (A), Guanine (G), Thymine (T), Cytosine (C)
  - A binds with T and C binds with G
  - As such, DNA is double stranded
- DNA must be read from 5' to 3' end
- RNA usually single-stranded
  - Uracil replaces Thymine in RNA
- Gene transcription: Transcription factors recognize binding sites in DNA, recruits RNA polymerase
  - RNA polymerases actually transcribes the DNA strand
  - Presence or absence of transcription factor dictates whether transcription occurs, or in what amounts
- DNA → transcription to RNA → splicing out of introns leads to mRNA → translation leads to protein
  
- Proteins composed of one of 20 amino acids
- Specified start codon (AUG; Methionine) and multiple stop codons
- A non-coding RNA called tRNA folds into itself into a complex 3D-structure. It has an anti-codon that pairs with the reverse complement codon in the mRNA and releases the correct amino acid
  
- There can be errors, or mutations, in the translation process

- One amino acid mutation (substitution) could result in a different amino acid, or a stop codon
  - A deletion could result in a frame shift that affects subsequent amino acids
- Regulation can occur during transcription by the addition of enhancers or silencers
  - There are proteins that can bind to a promoter and either recruit or prevent binding of RNA polymerase