CHAPTER TWENTY-ONE: Genomes & Their Evolution

21.2 *Scientists use bioinformatics to analyze genomes and their functions*

I. Centralized Resources for Analyzing Genome Sequences

- websites from arand the world for databases and software
- bioinformatics - application of computational methods to the storage and analysis of biological data
- ever expanding information
- sequences, bases, amino acids, etc
- yay Internet

II. Identifying Protein-Coding Genes & Understanding Their Functions

- reverse genetics: inferring phenotype from genotype
- gene annotation - identify all protein-coding genes in a sequence (+ their functions)
  - → software scans sequences for telltale signs of protein coding: codons, RNA splice sites, known mRNAs
  - → expressed sequence tags: short sequences collected from cDNA of known mRNA that are catalogued
- unknown sequences can be decoded sometimes if compared w/ known protein
- biochemical approach: determine 3D structure
- functional approach: block/disable gene (ie. RNAi)

III. Understanding Genes & Gene Expression at the Systems Level

- comparing genomes among species
- study of sets of genes
- how genes interact

} bioinformatics

- ENCODE project showed 90% of studied DNA region transcribed into RNA, only <2% coded for proteins
  - → led to experiments on fruit fly + soil nematode
- proteomics - study of the full sets of proteins of genomes

1. How Systems Are Studied: An Example
   - catalogs of genes + proteins
   - focus: functional integration in biological systems
   - systems bio approach: dynamic behavior of whole biological systems
     - → gene circuits + protein interaction networks
     - → double mutants compared to single mutants
     - → protein-protein interactions and fitness of an organism
2. Application of Systems Biology to Medicine
   - Cancer Genome Atlas → how changes in biological systems lead to cancer
   - common mutations in different cancers
     - → suspect genes + unknown genes
   - silicon + glass chips holding microarray of human genome
     - → analysis of gene expression patterns
     - → modest success
   - new: medical record of individual DNA sequence
     - → great potential
   - arrangement & interaction of the components of genetic systems
     - → important for deeper understanding of whole organisms

21.5 Duplication, rearrangement, and mutation of DNA contribute to genome evolution

I. Duplication of Entire Chromosome sets

- polyploidy: extra sets of chromosomes
  - → rarity: facilitates evolution of genes (if organism persists & reproduces)
    - ↳ genes w/new functions could evolve
  - → organism's phenotype could change if new gene encodes a protein acting in a novel way
- accumulation of mutations like this can lead to branch-off of new species
  - → mostly plants

II. Alterations of Chromosome structure

- humans: n = 24
- chimps: n = 23 (chromosome fused...)
- more genome info → compare chromosomes of different species → evolutionary processes
- chromosome 12+13 = chromosome 2
  ancestor   human
- comparison of human chr. 16 & mouse chromosomes 7, 8, 16, 17
  - → DNA sequences stayed together in mouse + human lineages from ancestor(s)
- evolutionary history of chromosomal re-arrangements in various species
  - → found duplications + inversions of large portions of chromosomes due to mistakes during meiotic recombination
- chromosomal rearrangements could lead to 2 populations that could not mate well
  - → led to two different species
- recombination "hot spots" - specific sites used a lot; not random

III. Duplication and Divergence of Gene-Sized Regions of DNA

- unequal crossing over: one chromosome w/ deletion, one w/ duplication of gene
  - → incorrect pairing of homologs leads to crossover at wrong location
- DNA replication slippage: segment deleted or duplicated
  - → template "slips" & part of it is skipped or used twice

1. Evolution of Genes w/ Related Functions: Human Globin Genes

- α-globin and β-globin
- evolved from common ancestral gene
  - → duplication & diverged from each other
- mutations must have occured that provided favorable outcomes without changing the protein product's actual function
- then came natural selection (Yay!)
- similarities of α and β globin support this

2. Evolution of Genes w/ Novel Functions

- duplicated gene can also alter so that a new function arises
- lysozyme: protects animals from infection by hydrolyzing bacterial cell walls
- α-lactalbumin: protein to do w/ milk production
  - → similar in amino sequence & 3D structure
- only lysozyme is in birds
  - → lysozyme gene ~~rep~~ duplicated in mammal lineage, not avian lineage
  - → lysozyme evolved into new gene (α-lactalbumin)

IV. Rearrangements of Parts of Genes: Exon Duplication & Exon Shuffling

- exon w/ a gene can be duplicated on one chromosome & deleted from the other
- duplication: protein w/ second copy of domain
  - → increasing stability
  - → enhancing binding to a ligand
  - → alteration of other property
  - → ex. collagen → repetitive exon pattern
- exon shuffling: mix & match of exons w/in a gene or between two separate genes
  - → new proteins, new functions

V. How Transposable Elements Contribute to Genome Evolution

- transposables can:
  - promote recombination
  - disrupt cellular genes or control elements
  - carry entire genes & individual exons to new place
  - provide homologous regions for cross over
    - → usually bad; sometimes advantageous
  - jump into protein-coding or regulatory sequence
    - → can stop normal transcript or change protein production; sometimes advantageous
  - move genes to new position in genome or allow tag-along of exons (new protein function)
    - → can be detrimental; sometimes advantageous
- these movements and mutations provide "valuable raw material for natural selection"