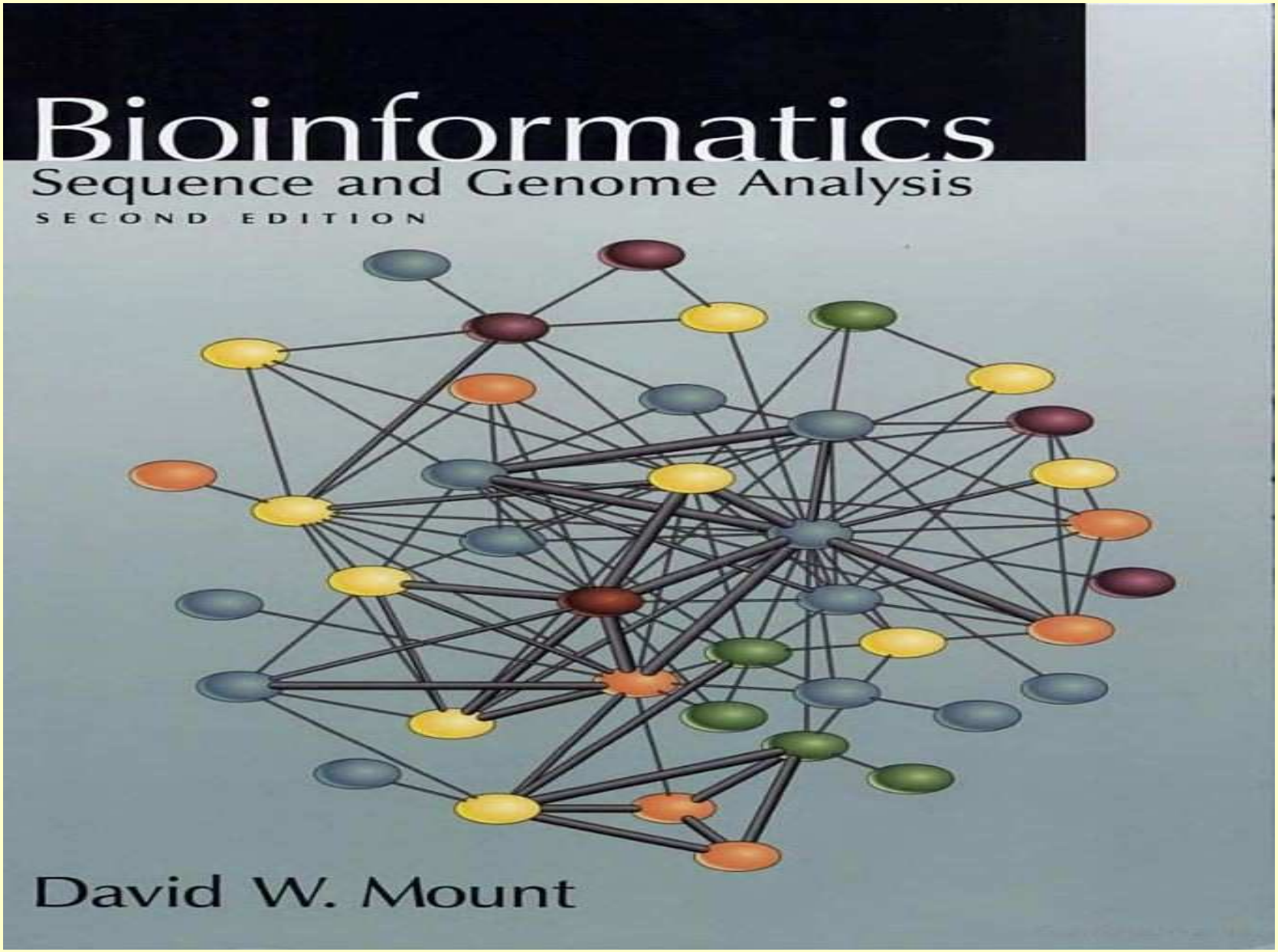
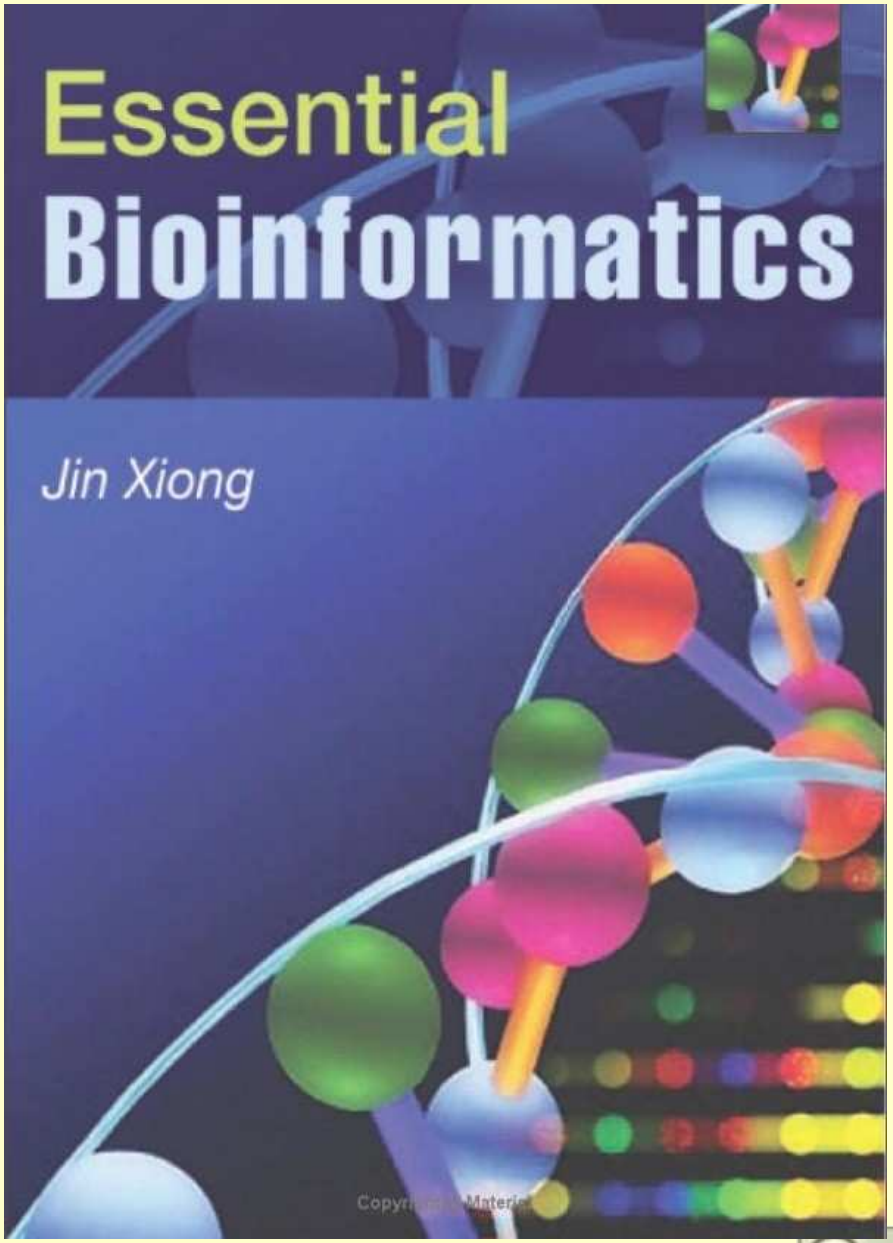
David Mount

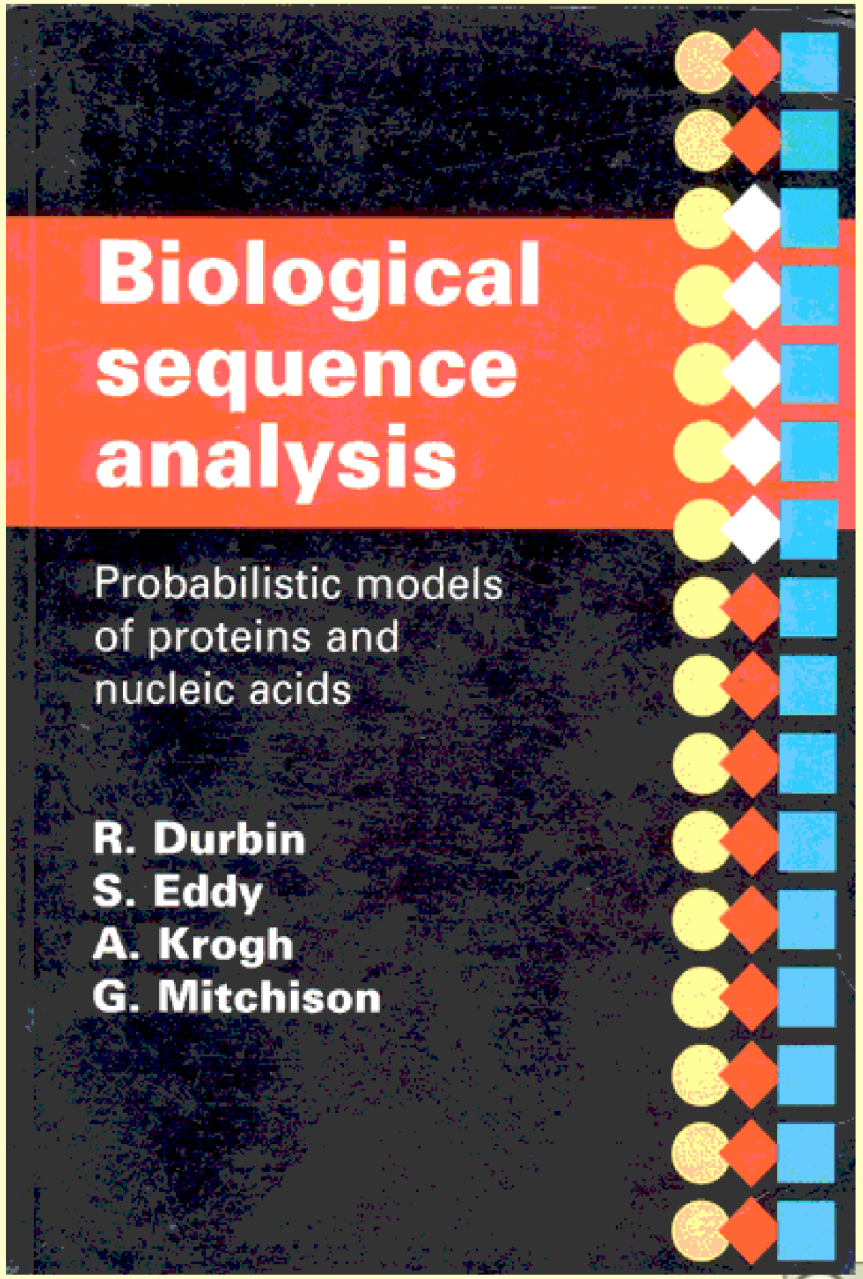
Bioinformatics: Sequence and Genome Analysis 2nd Edition



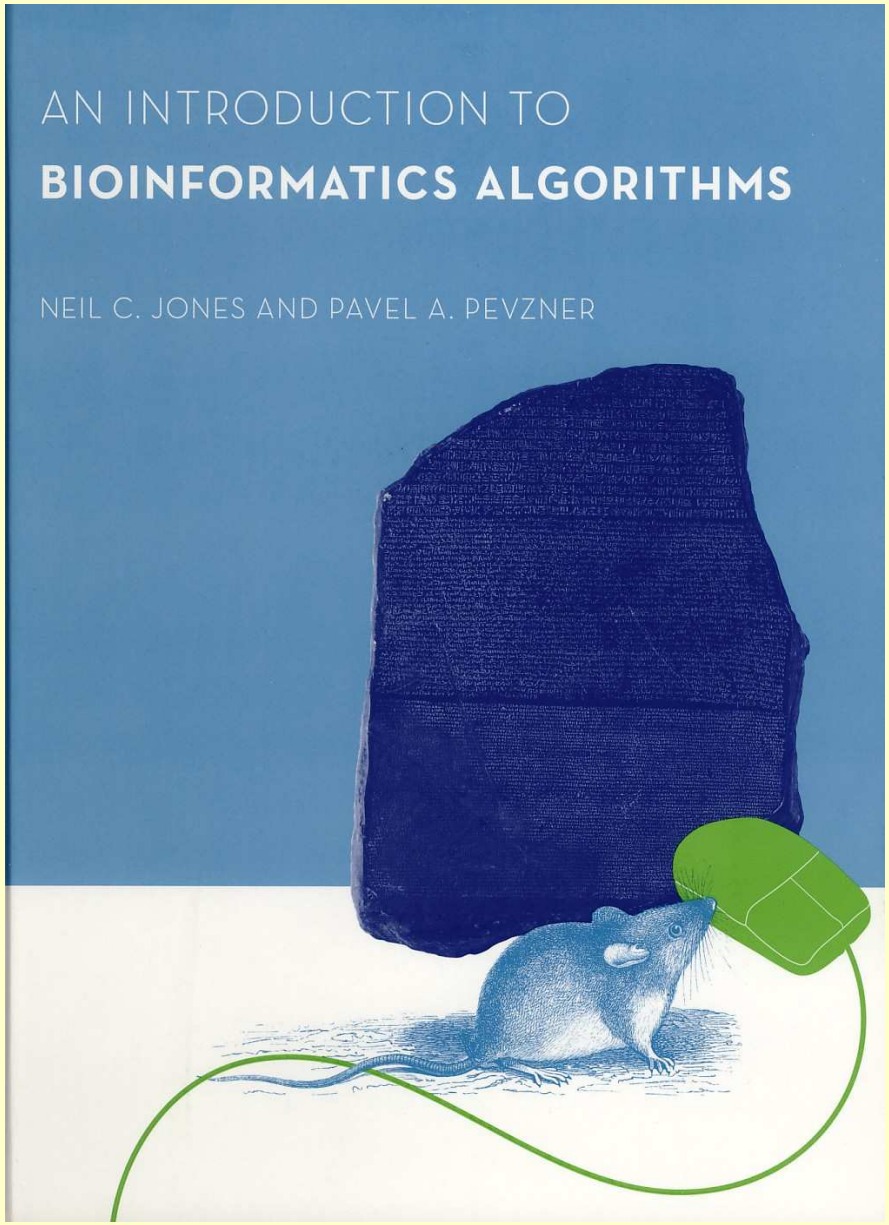
Jin Xiong
Essential Bioinformatics



Richard Durbin *et al.*
Biological Sequence Analysis

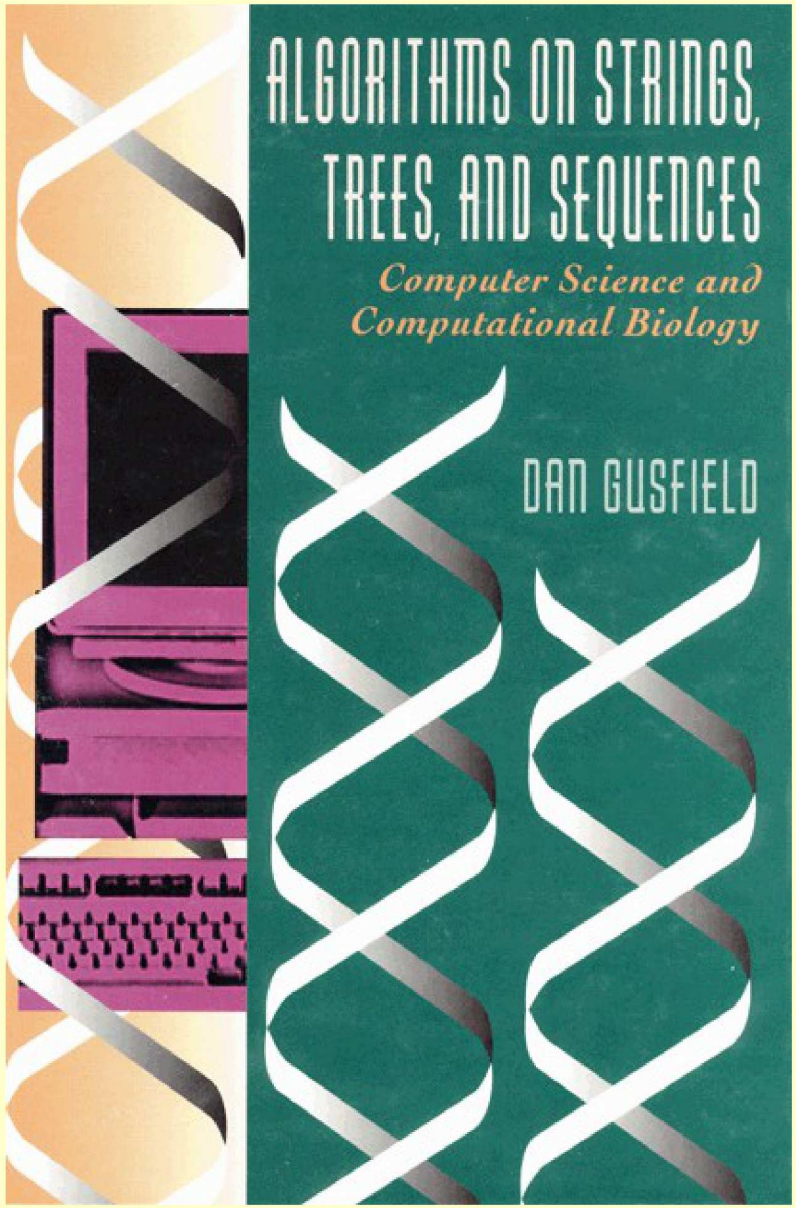


Jones & Pevzner
Bioinformatics Algorithms



Dan Gusfield

Algorithms on Strings, Trees & Sequences

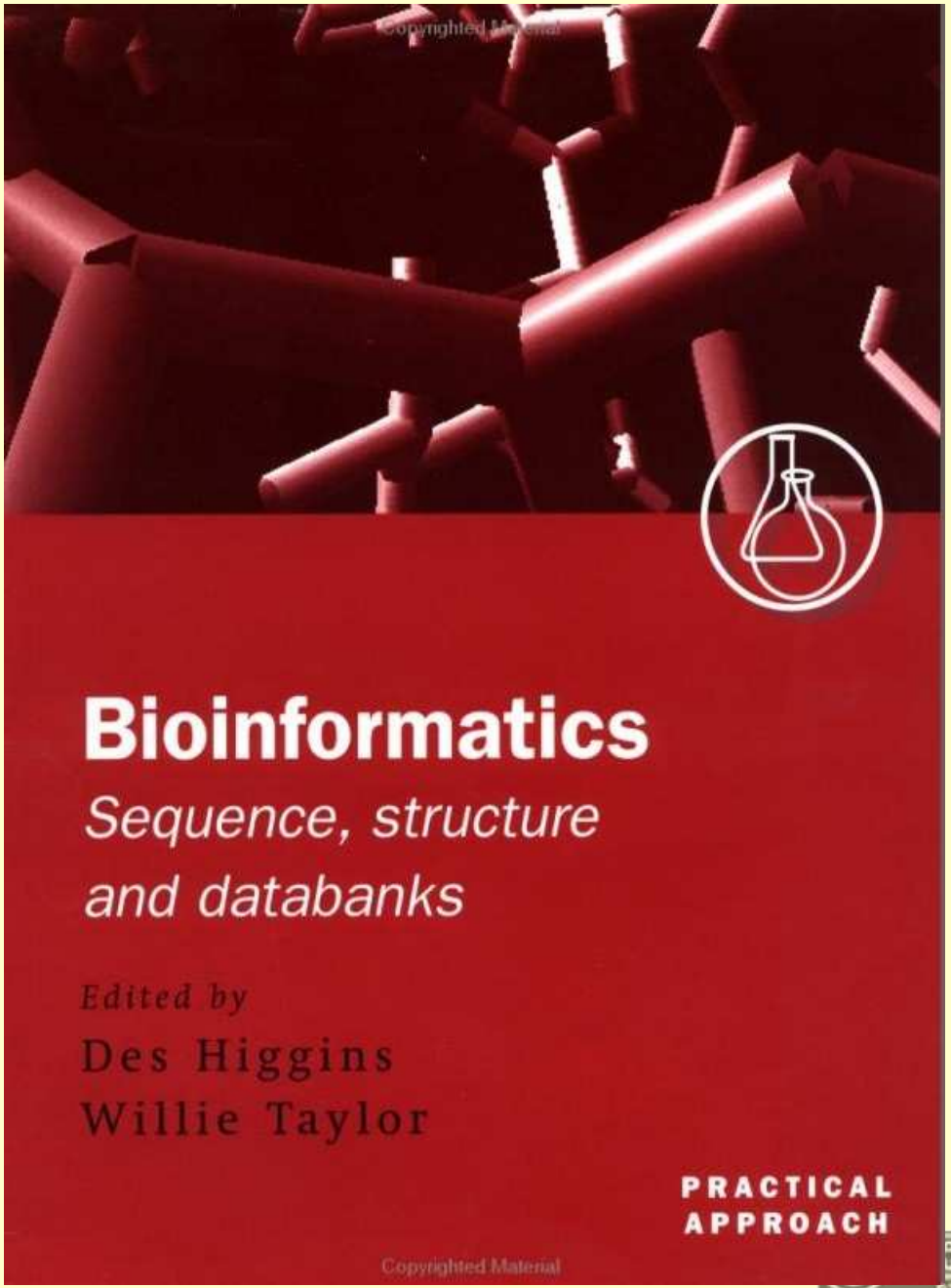


Baldi & Brunak

Bioinformatics: The Machine Learning Approach

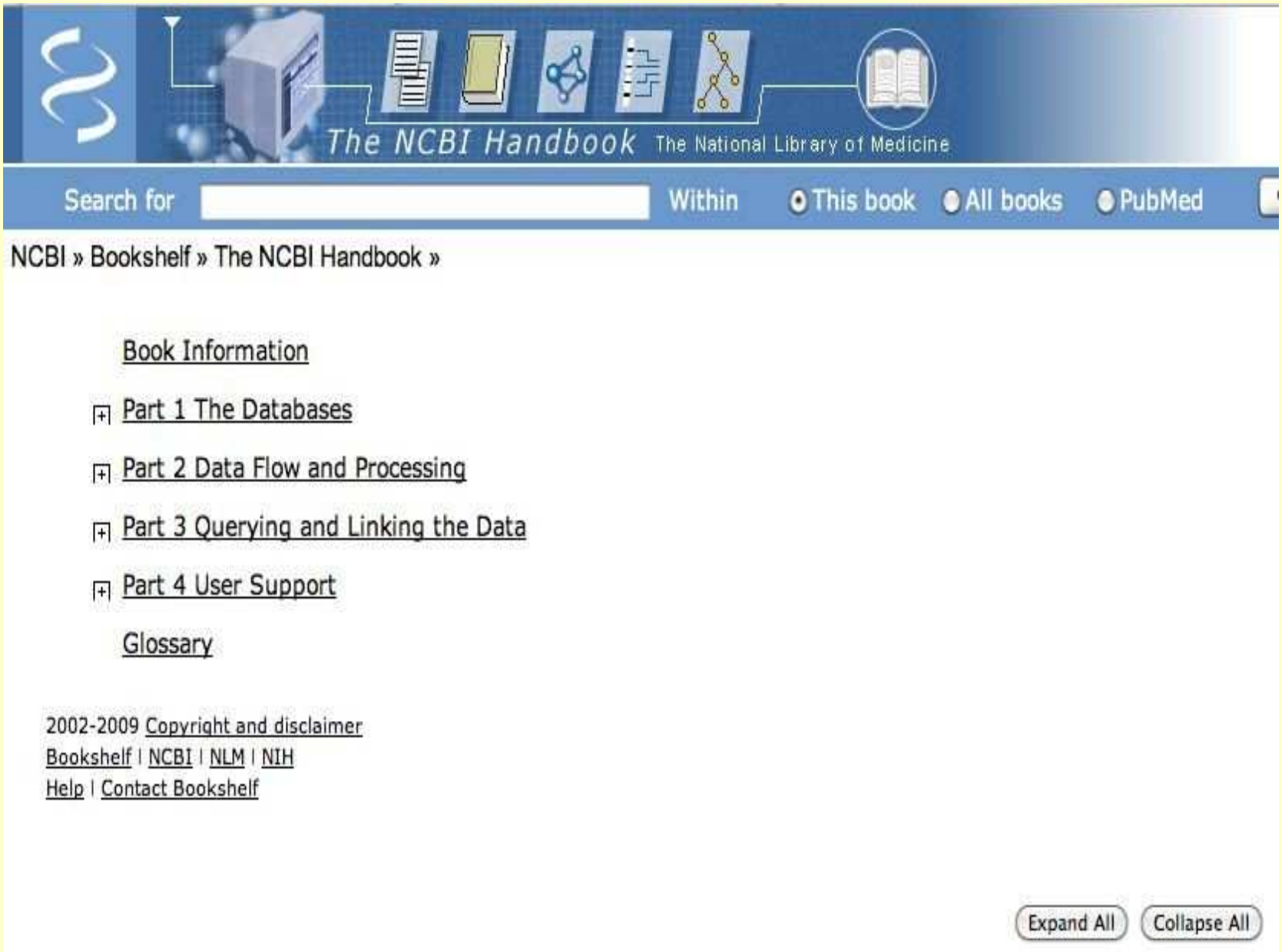
Higgins & Taylor

Bioinformatics: Sequence, Structure & Databanks



NCBI Handbook

http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=handbook



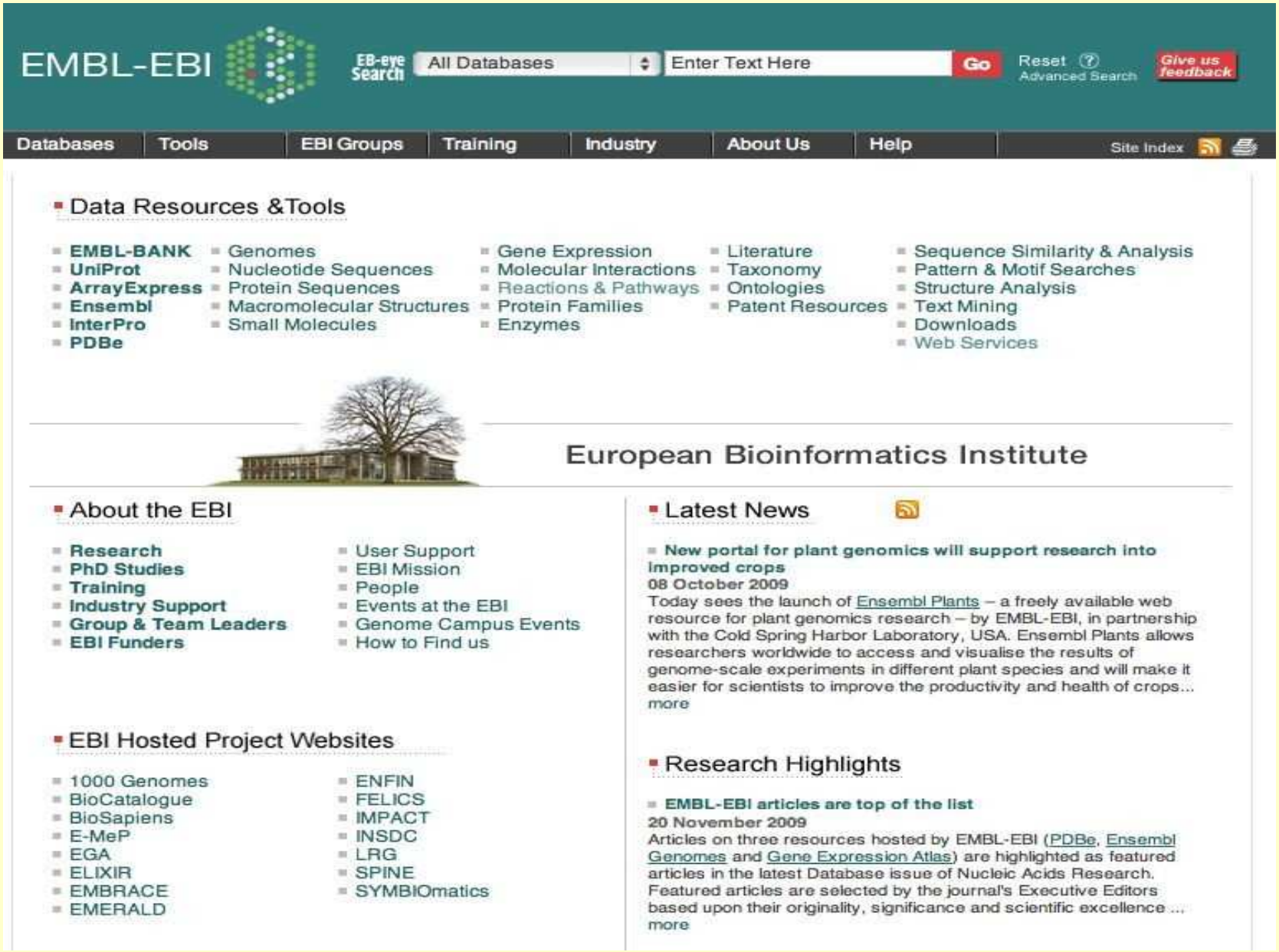
NCBI Handbook

http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=handbook

NCBI » Bookshelf » The NCBI Handbook » The Databases

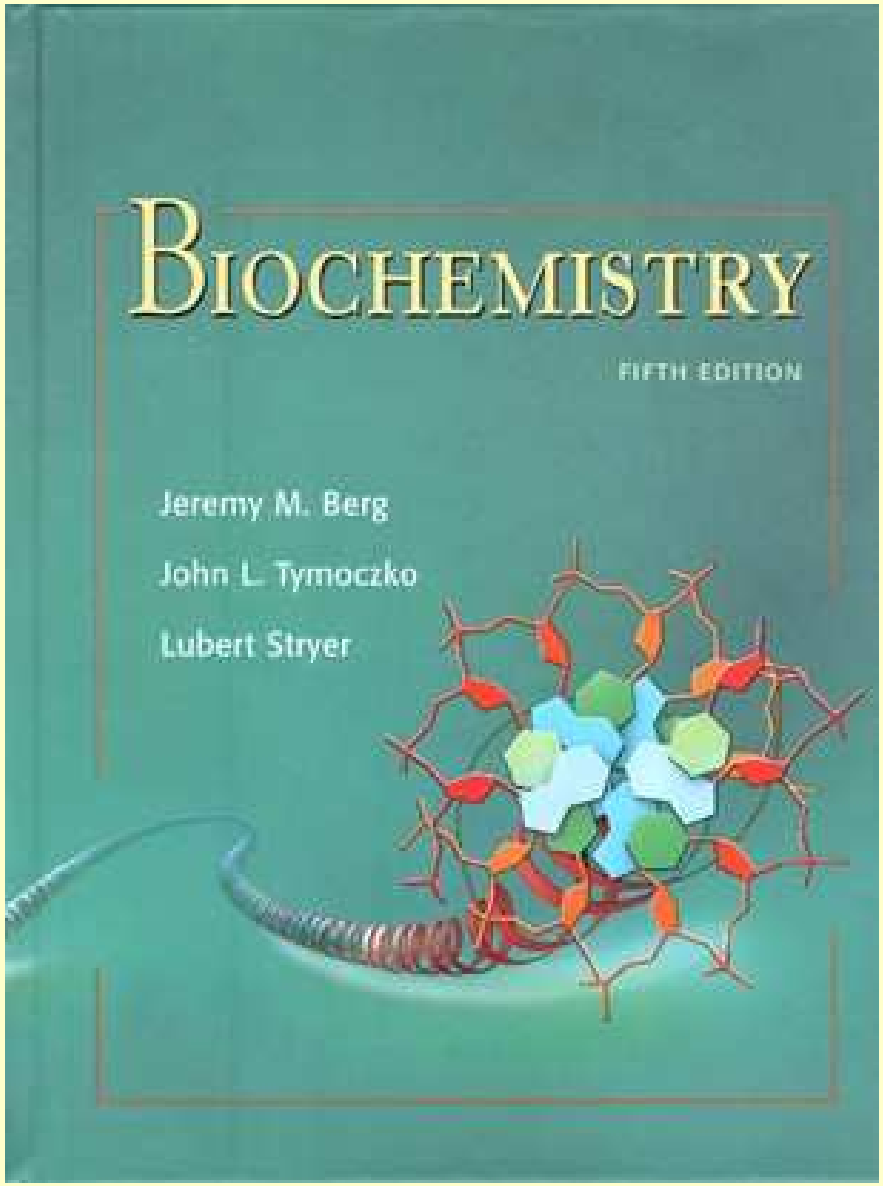
The Databases

EMBL-EBI Home Page
http://www.ebi.ac.uk/



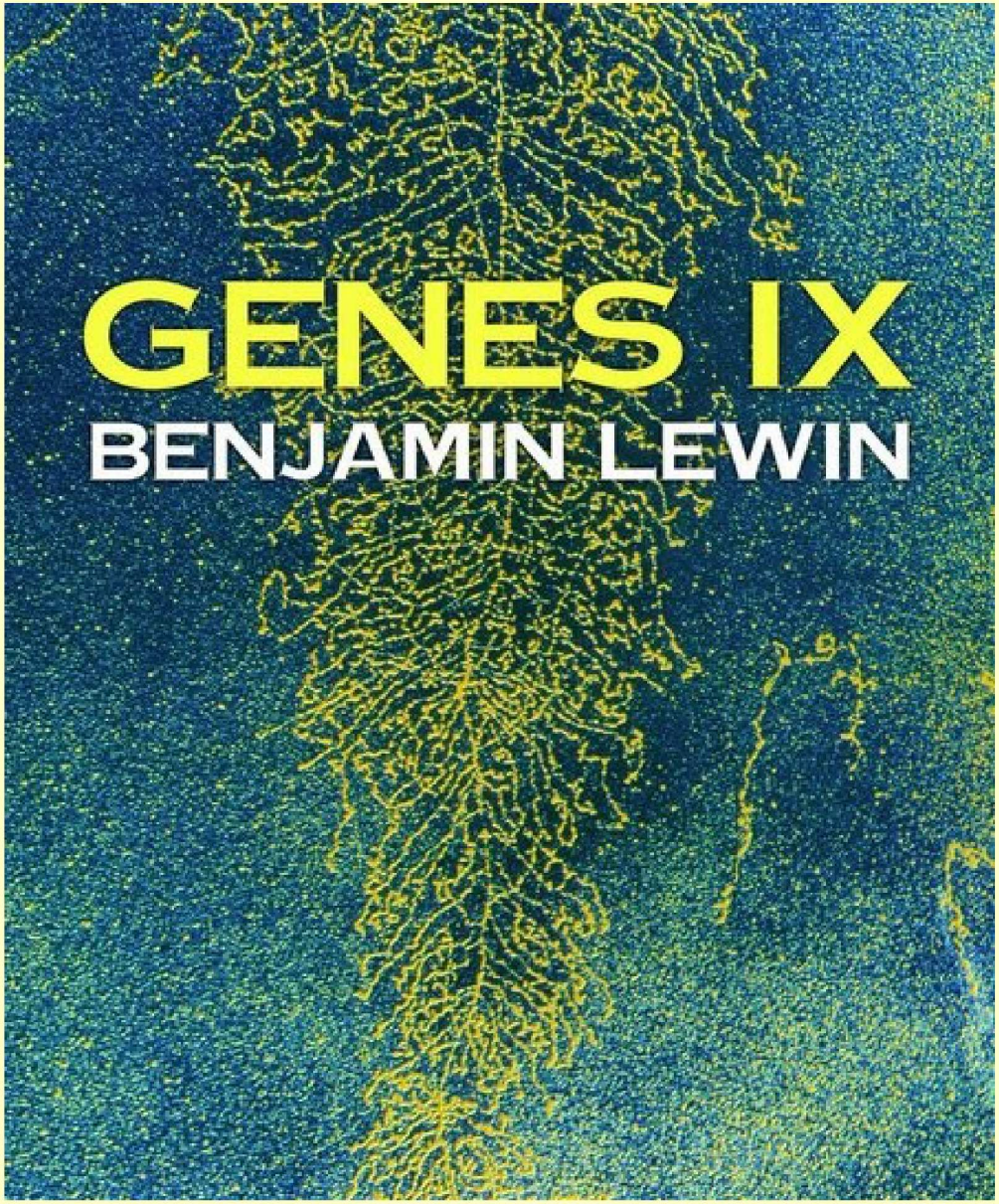
Berg, Tymoczko & Stryer

Biochemistry, Fifth Edition

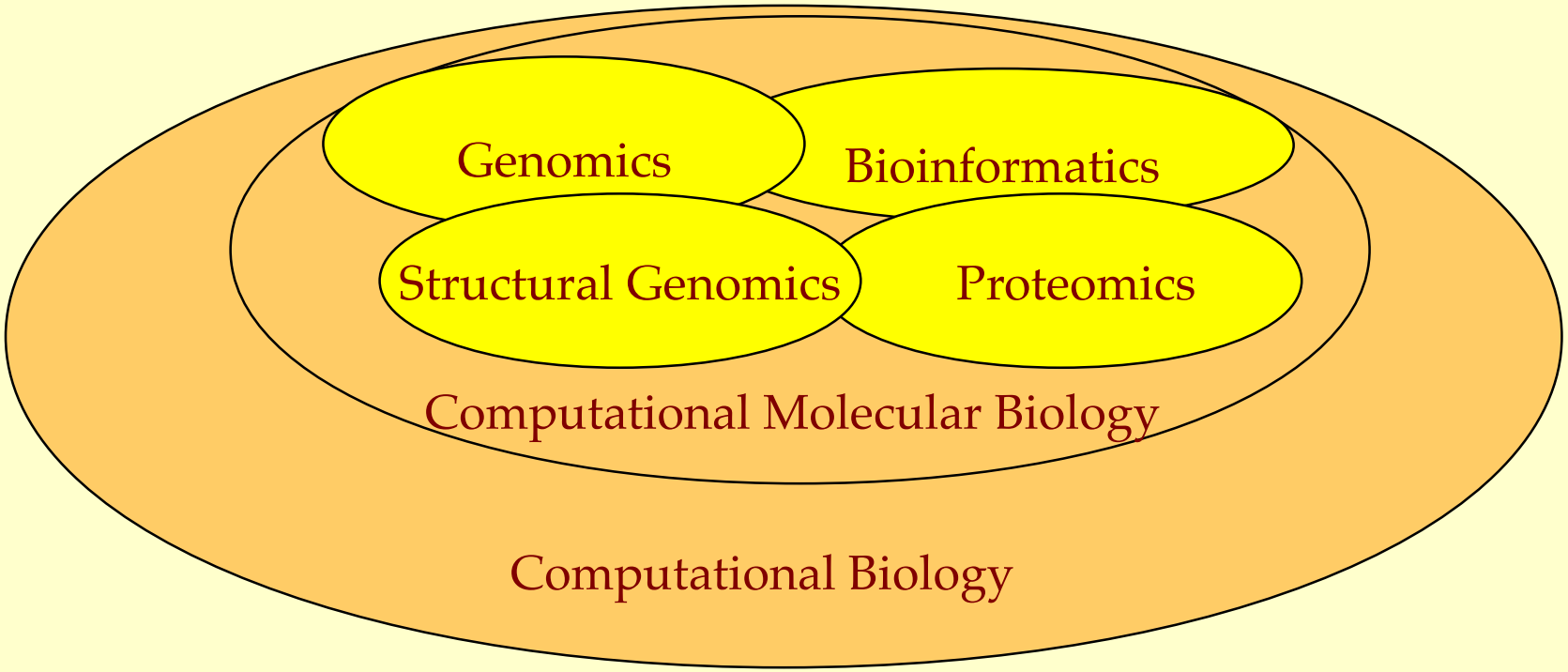


Benjamin Lewin

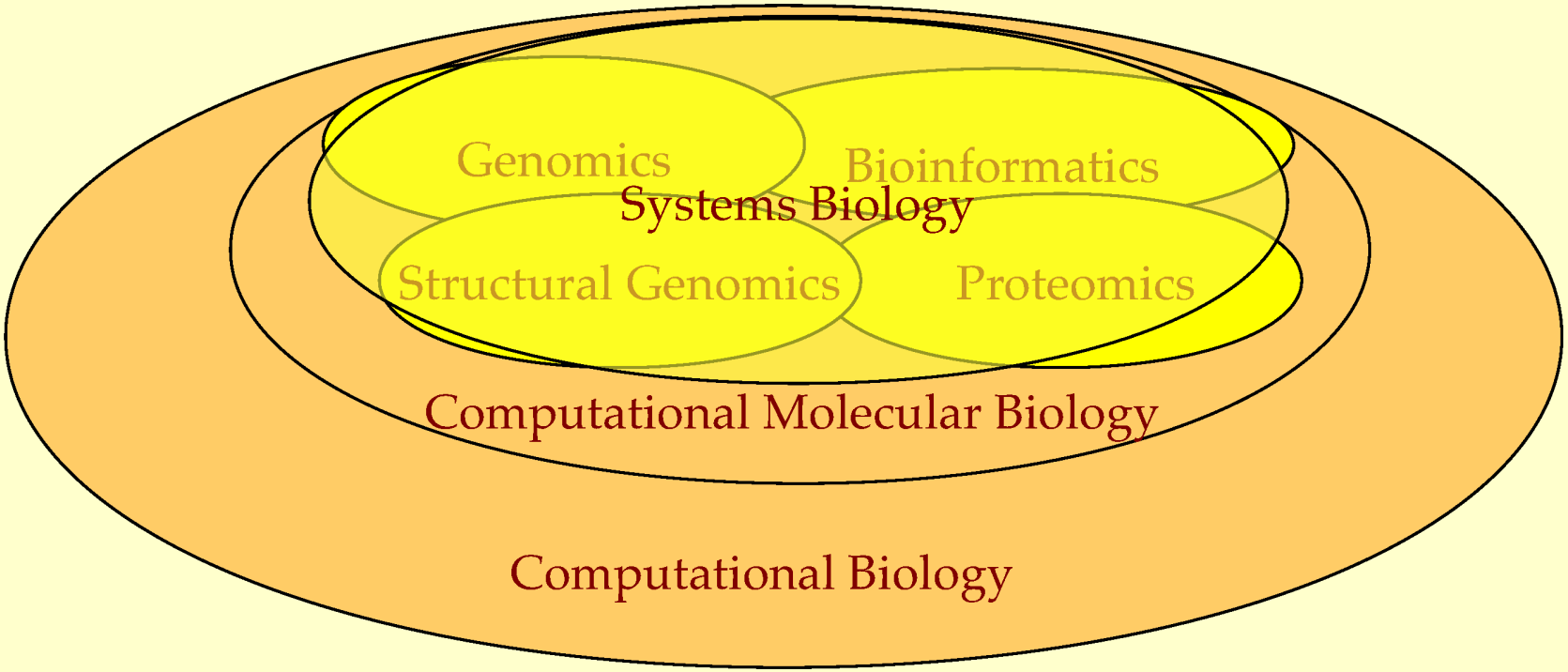
Genes IX



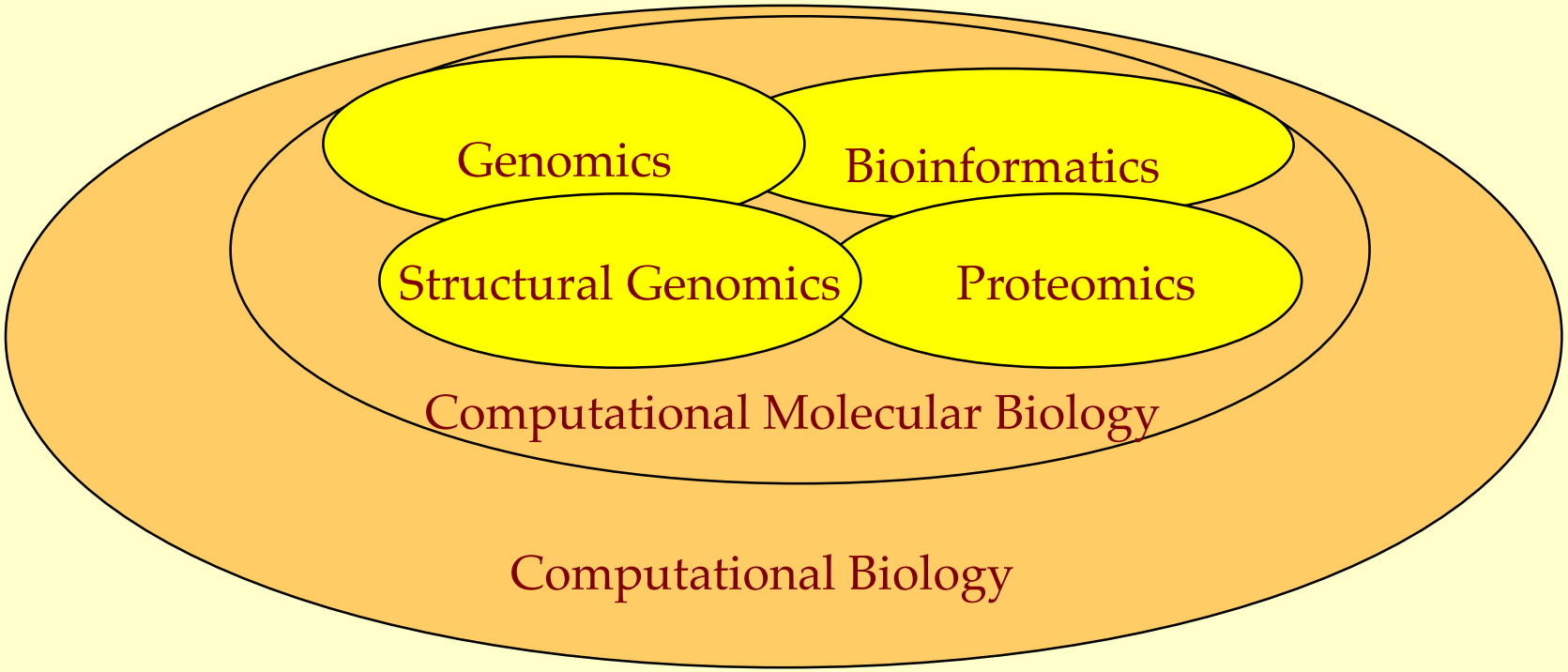
Genomics, Bioinformatics & Computational Biology

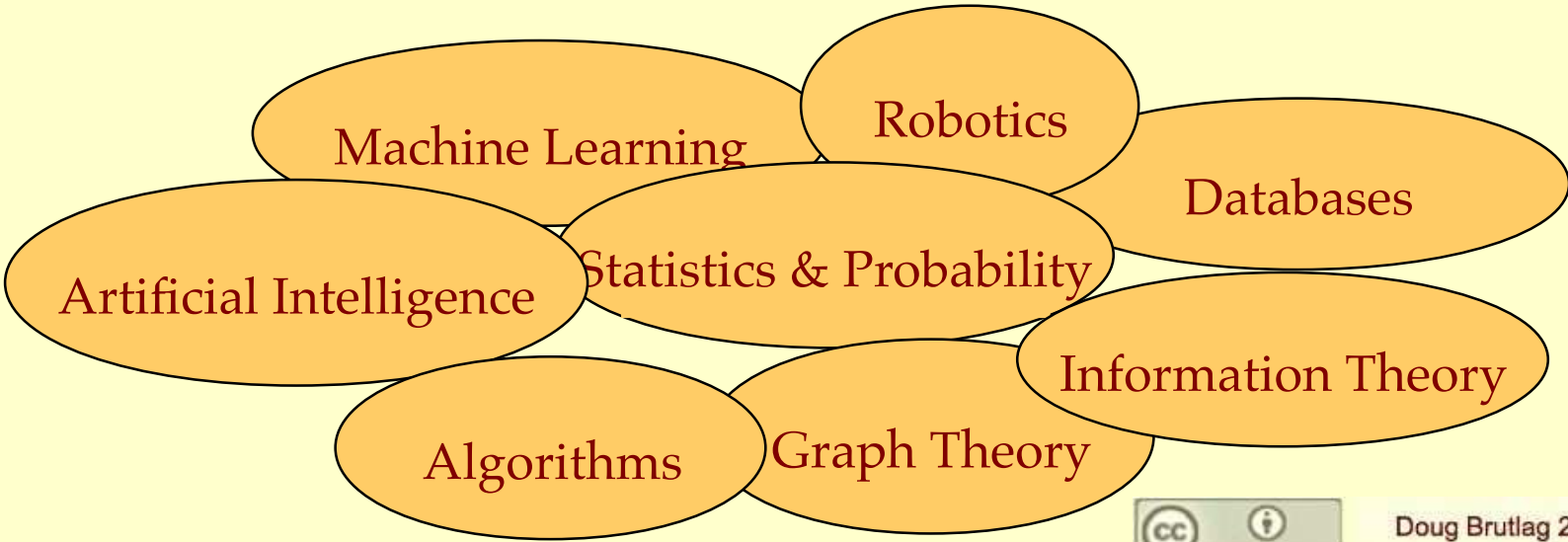


Genomics, Bioinformatics & Computational Biology

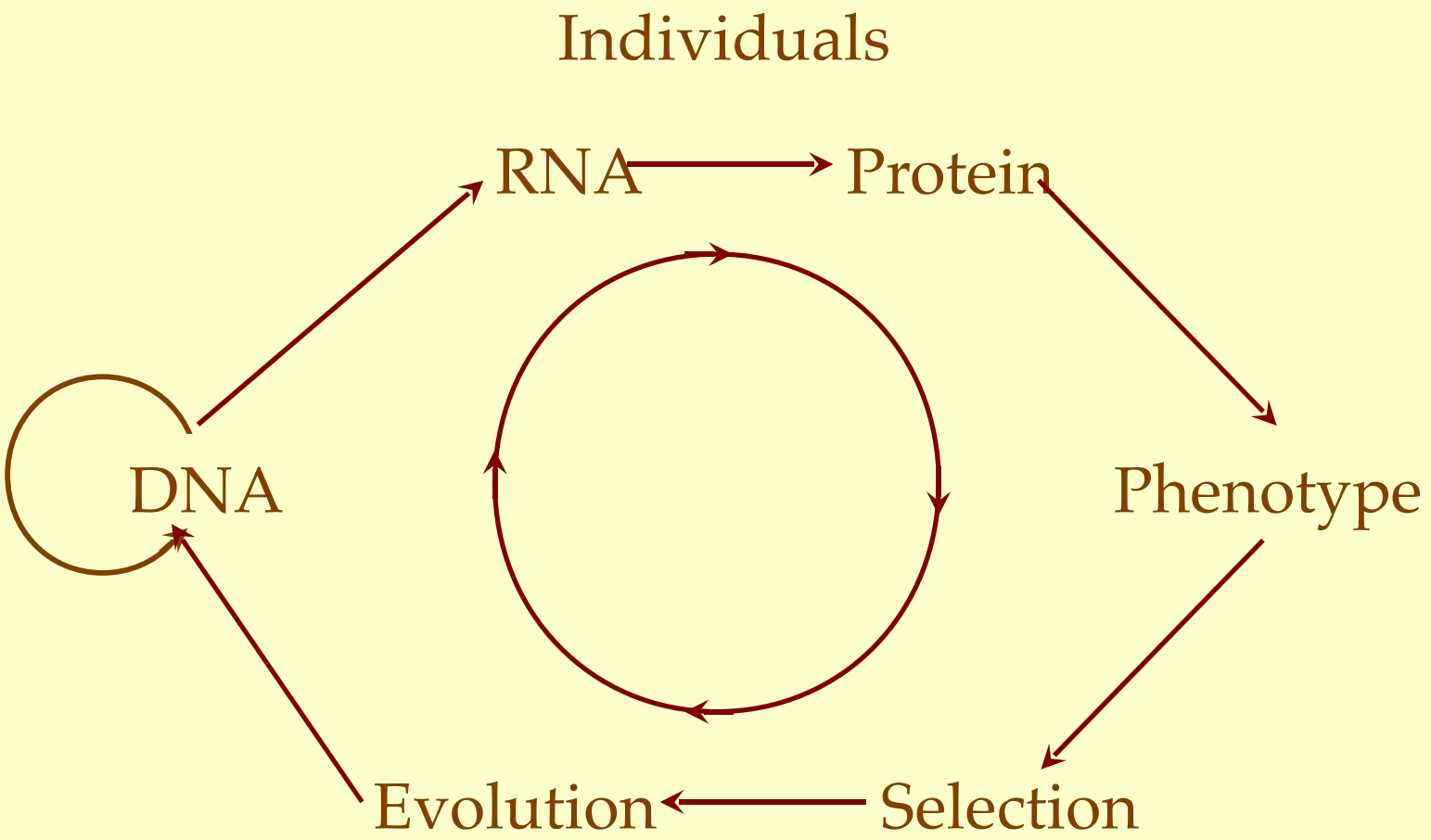


Genomics, Bioinformatics & Computational Biology





What is Bioinformatics?

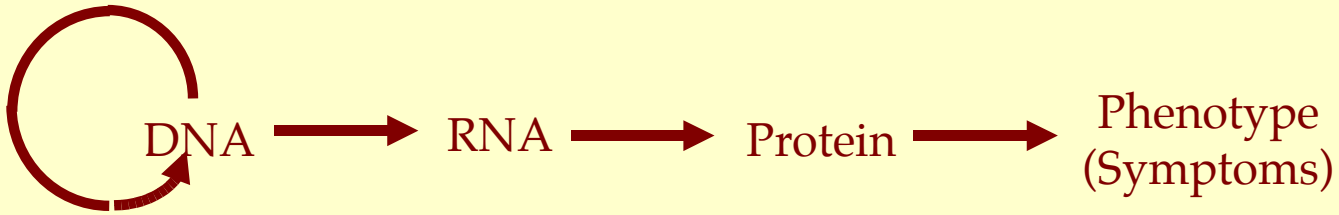


Biological Information

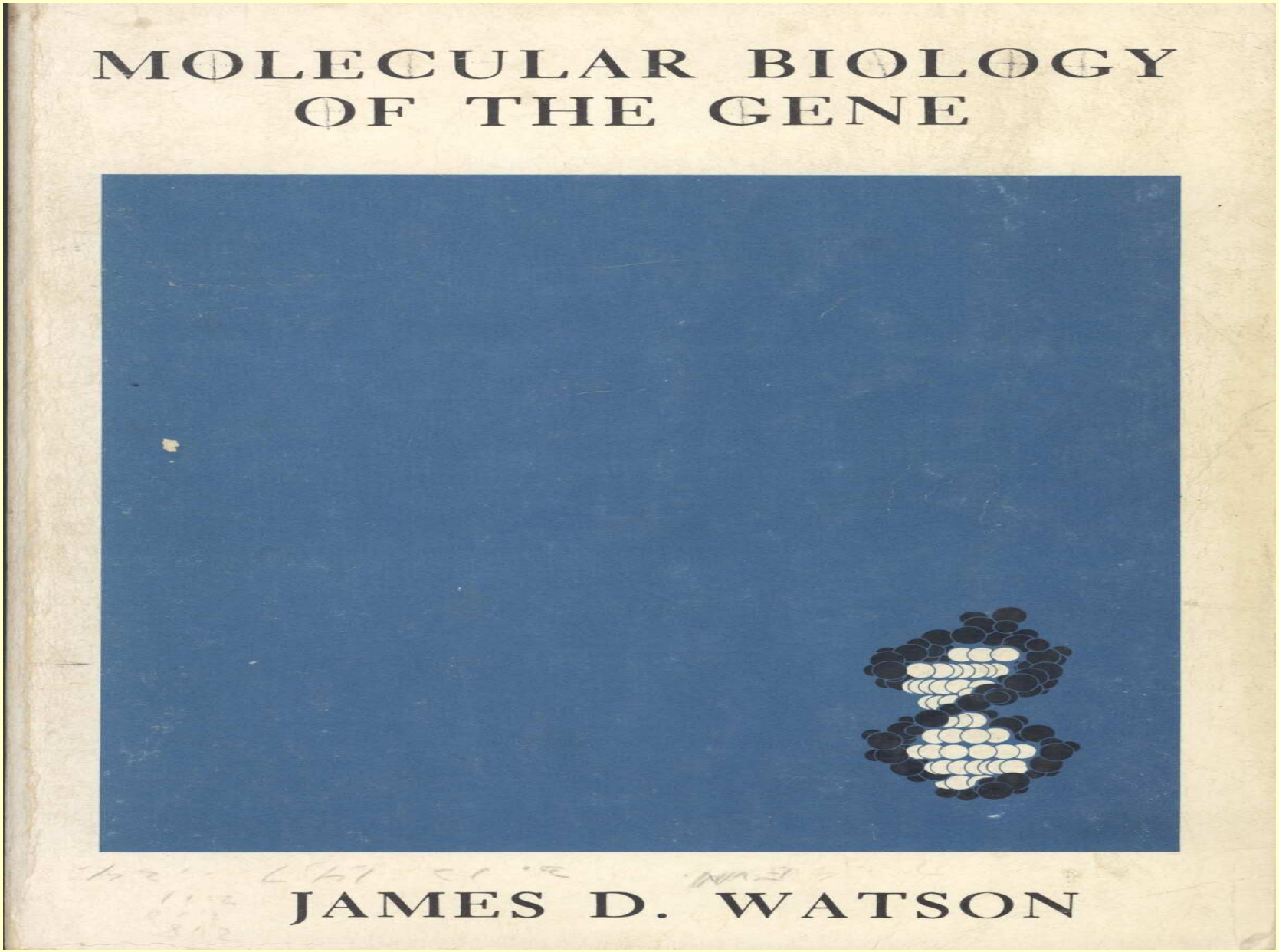
Computational Goals of Bioinformatics

- Learn & Generalize: Discover conserved patterns (models) of sequences, structures, interactions, metabolism & chemistries from well-studied examples.
- Prediction: Infer function or structure of newly sequenced genes, genomes, proteins or proteomes from these generalizations.
- Organize & Integrate: Develop a systematic and genomic approach to molecular interactions, metabolism, cell signaling, gene expression…
- Simulate: Model gene expression, gene regulation, protein folding, protein-protein interaction, protein-ligand binding, catalytic function, metabolism…
- Engineer: Construct novel organisms or novel functions or novel regulation of genes and proteins.
- Gene Therapy: Target specific genes, or mutations, RNAi to change a disease phenotype.

Central Paradigm of Molecular Biology



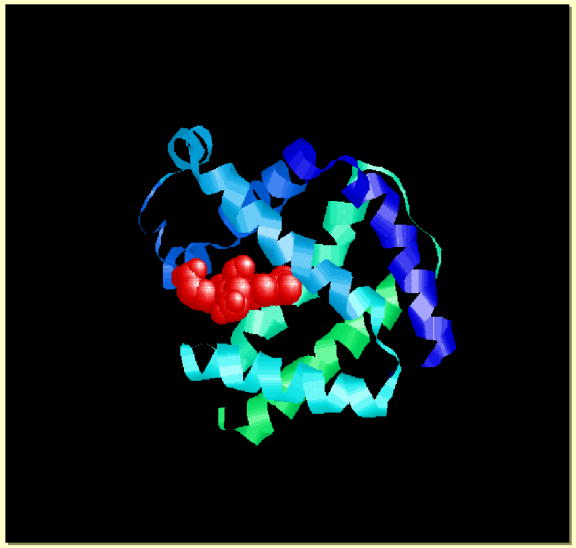
Molecular Biology of the Gene 1965

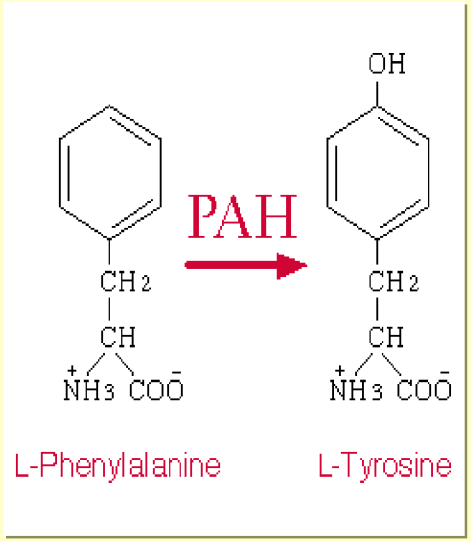


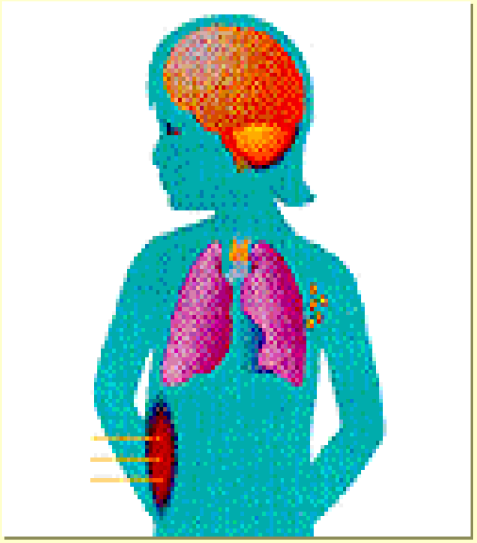
Central Paradigm of Bioinformatics

Genetic Information → Molecular Structure → Biochemical Function → Phenotype (Symptoms)

MVHLTPEEKT
AVNALWGKVN
VDAVGGEALG
RLLVVYPWTQ
RFFESFGDLS
SPDAVMGNPK
VKAHGKKVLG
AFSDGLAHLD
NLKGTFSQLS
ELHCDKLHVD
PENFRLLGNV
LVCVLARNFG
KEFTPQMQAA
YQKVVAGVAN
ALAHKYH



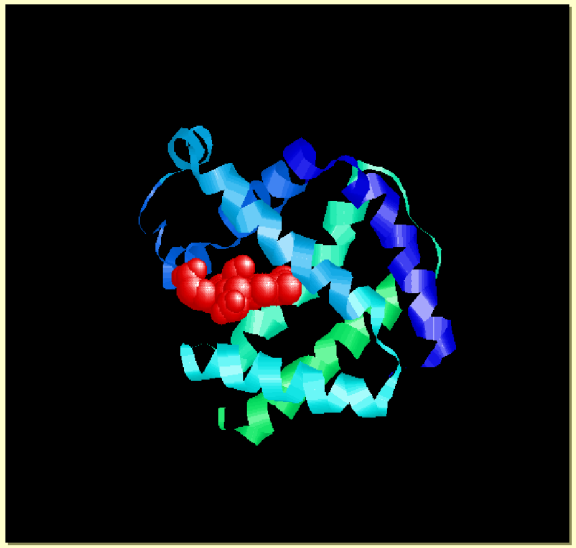


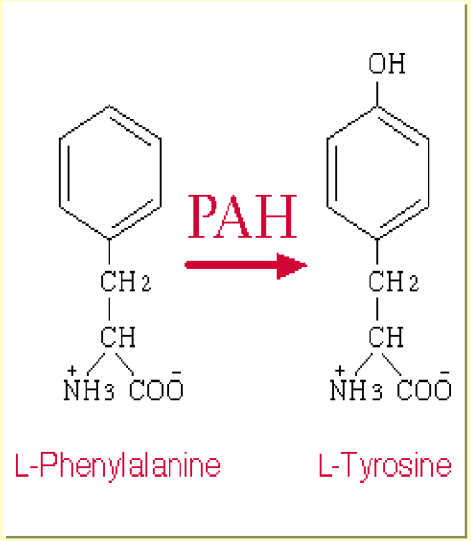


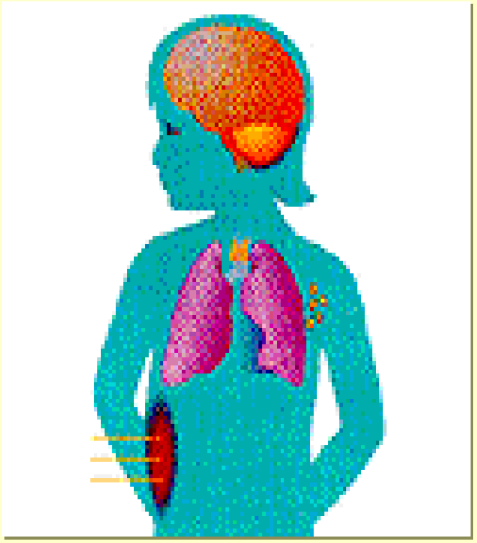
Central Paradigm of Bioinformatics

Genetic Information → Molecular Structure → Biochemical Function → Phenotype (Symptoms)

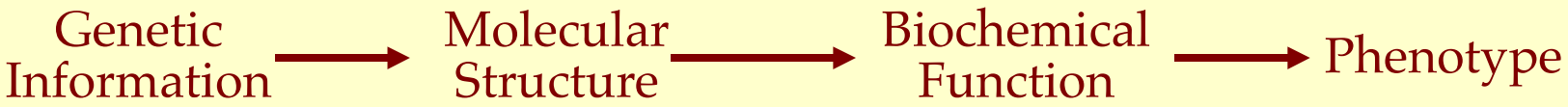
MVHLTPEEKT
AVNALWGKVN
VDAVGGEALG
RLLVVYPWTQ
RFFESFGDLS
SPDAVMGNPK
VKAHGKKVLG
AFSDGLAHLD
NLKGTFSQLS
ELHCDKLHVD
PENFRLLGNV
LVCVLARNFG
KEFTPQMQAA
YQKVVAGVAN
ALAHKYH







Challenges Understanding Genetic Information



- Genetic information is redundant
- Structural information is redundant
- Genes and proteins are meta-stable
- Single genes have multiple functions
- Genes are one dimensional but function depends on three-dimensional structure

Redundancy in Genomic & Protein Sequences

- DNA is double-stranded
- Genetic code
- Acceptable amino-acid replacements
- Intron-exon variation
- Alternative splicing
- Strain variations (SNPs)
- Sequencing errors