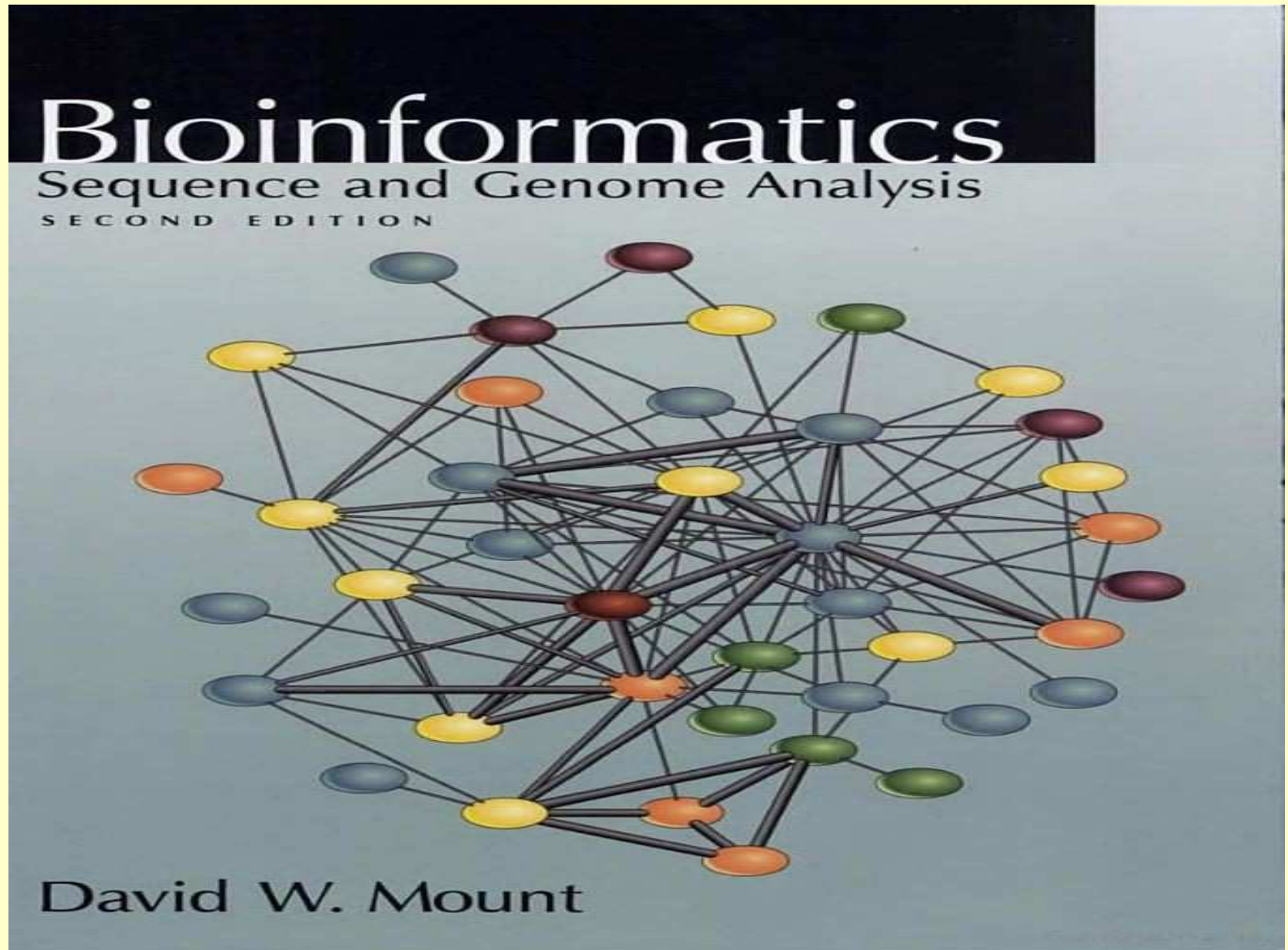


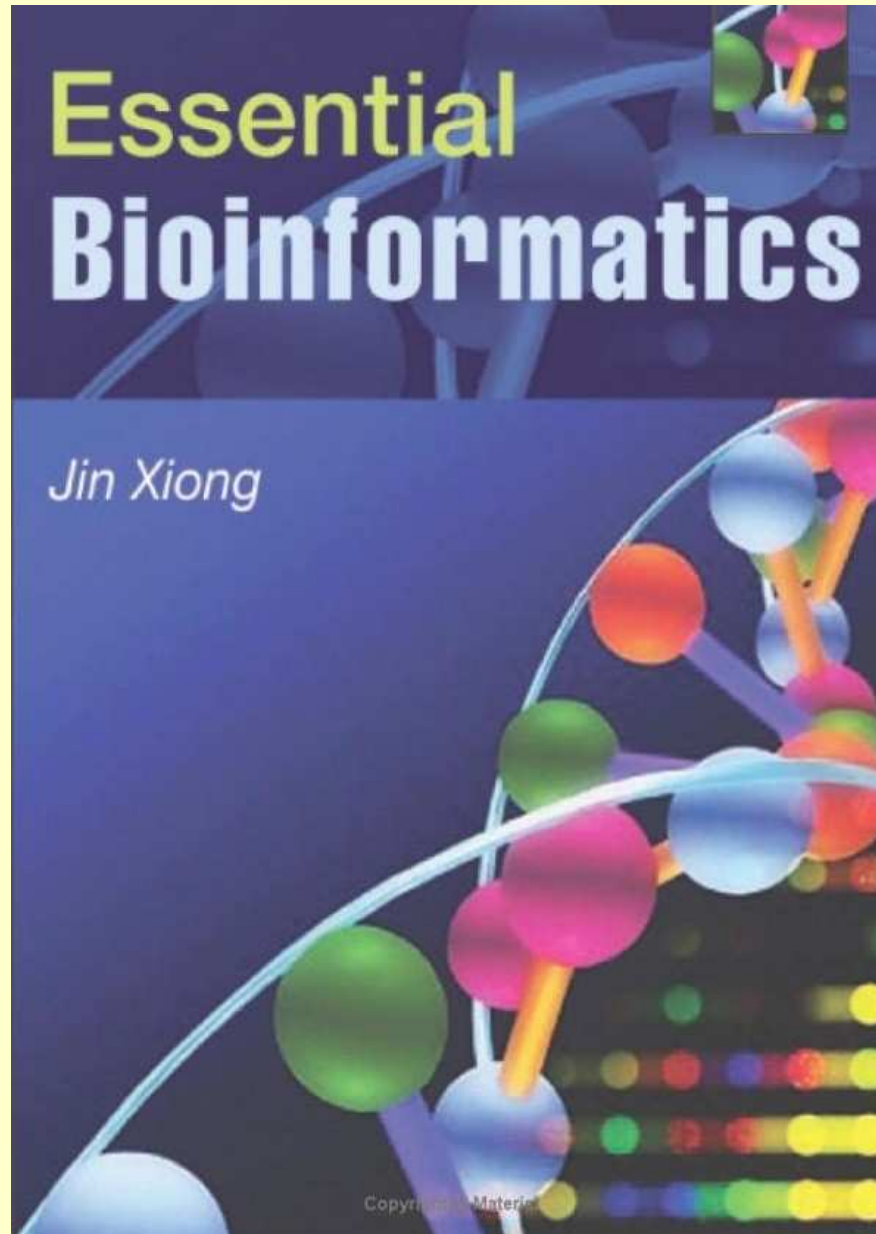
David Mount

Bioinformatics: Sequence and Genome Analysis 2nd Edition



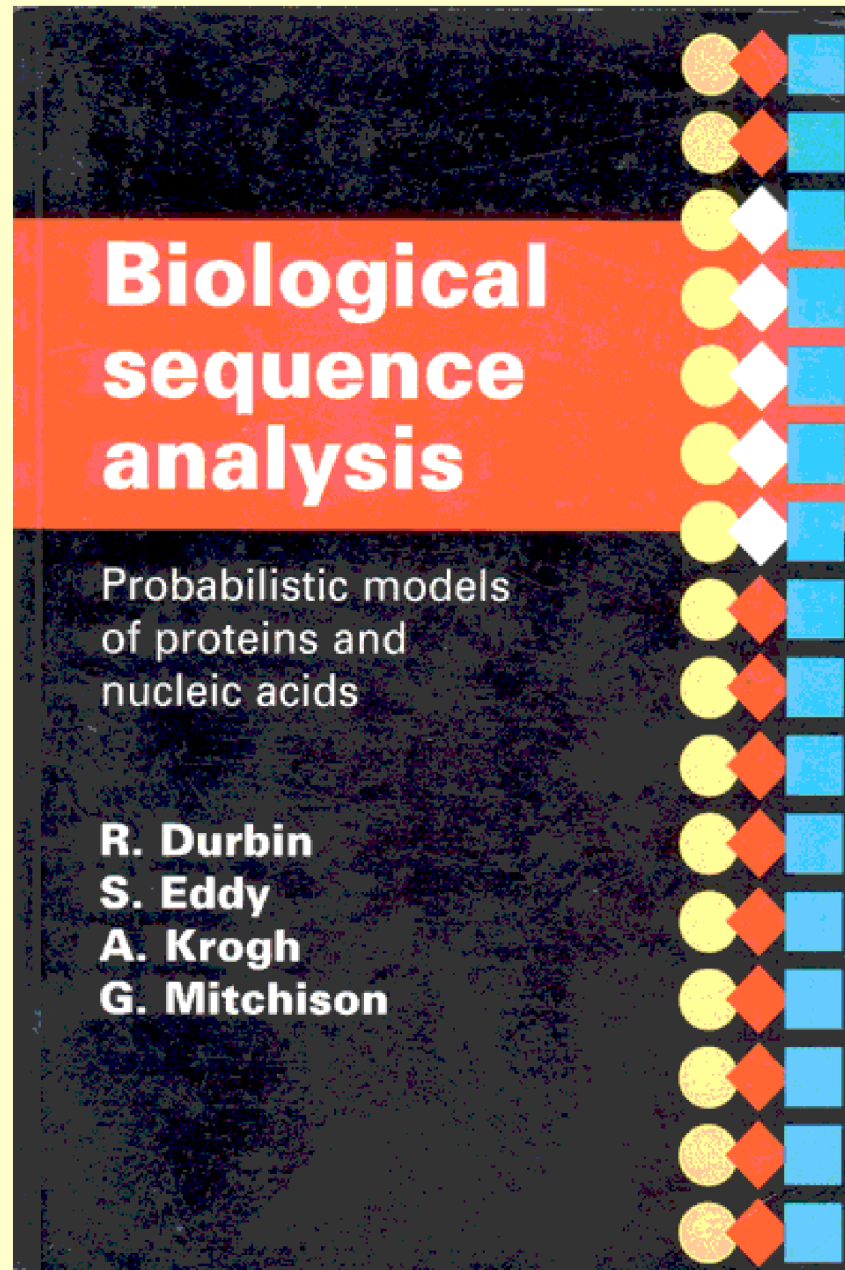
Jin Xiong

Essential Bioinformatics



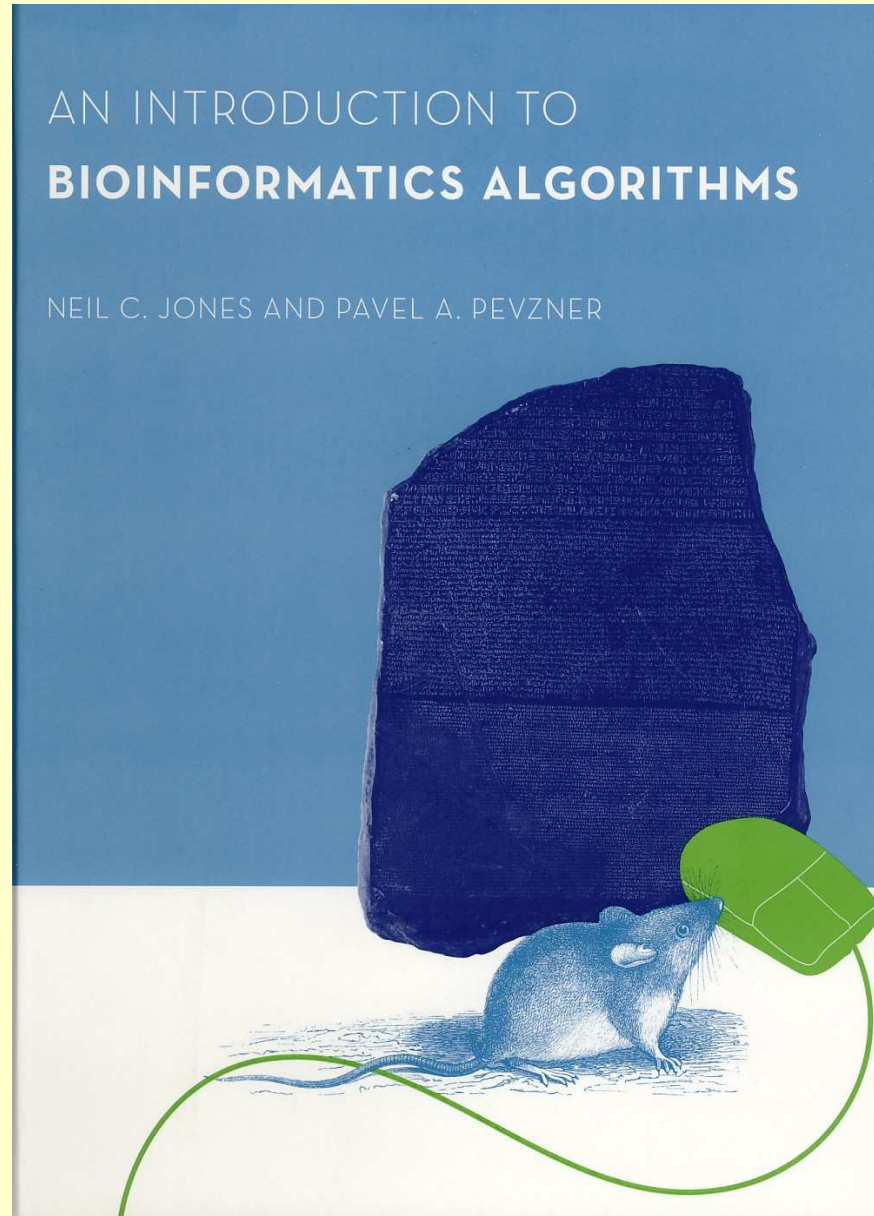
Richard Durbin *et al.*

Biological Sequence Analysis



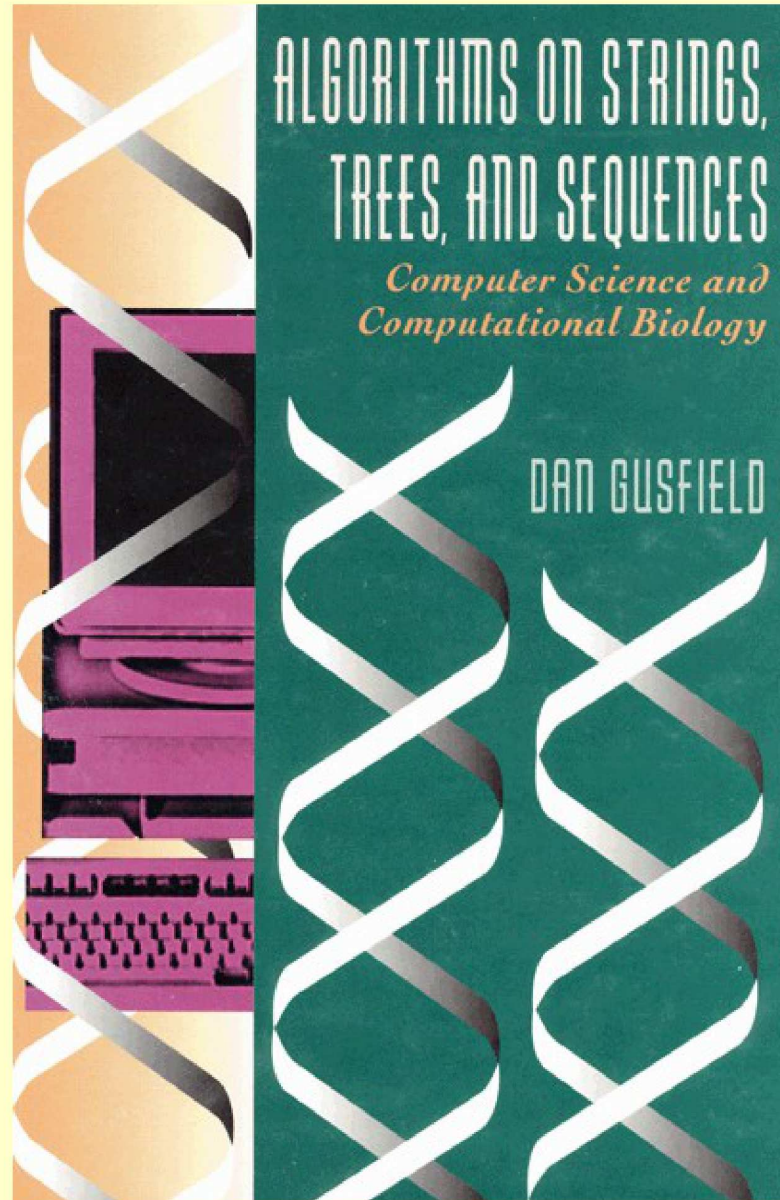
Jones & Pevzner

Bioinformatics Algorithms



Dan Gusfield

Algorithms on Strings, Trees & Sequences



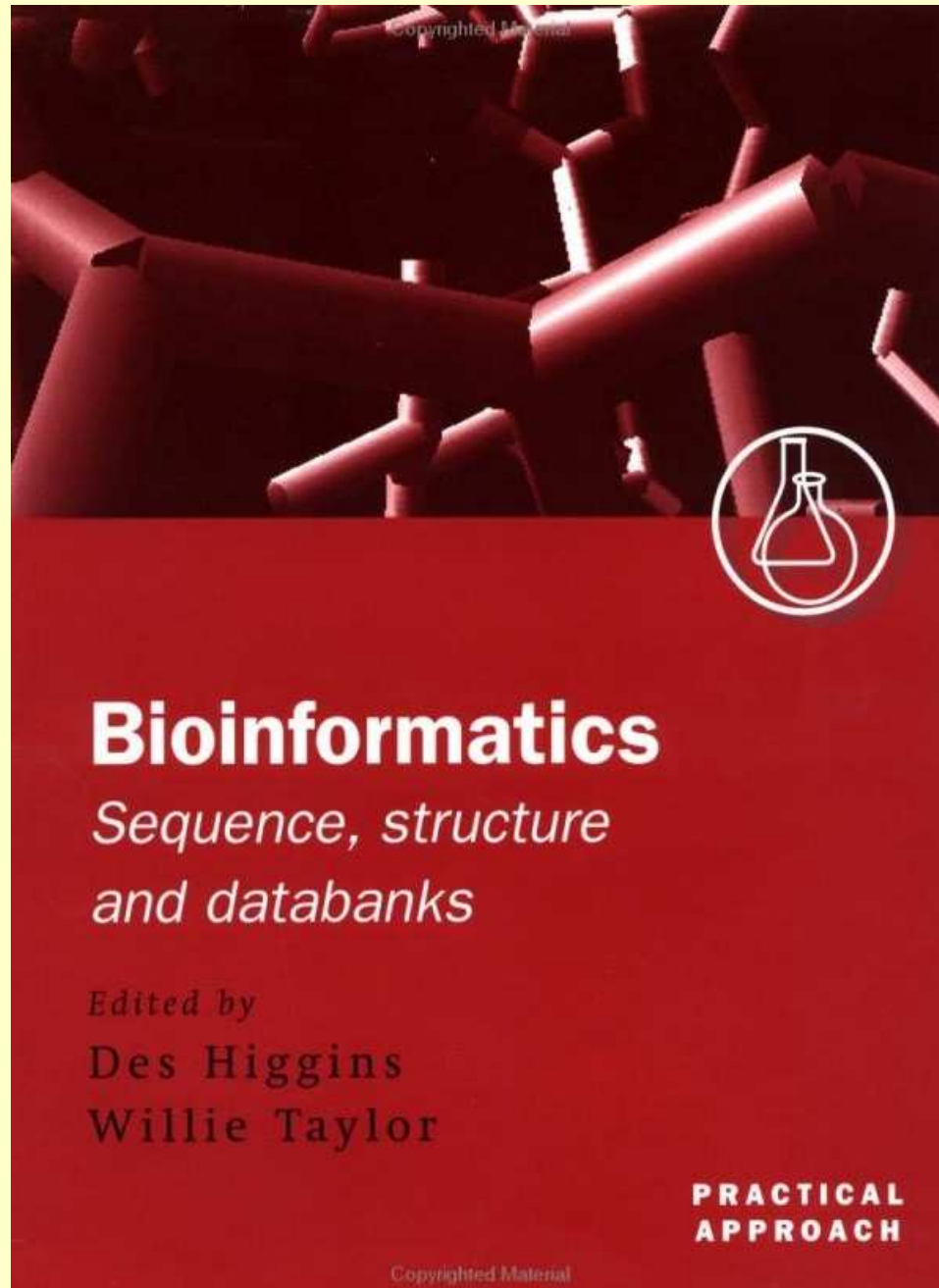
Baldi & Brunak

Bioinformatics: The Machine Learning Approach



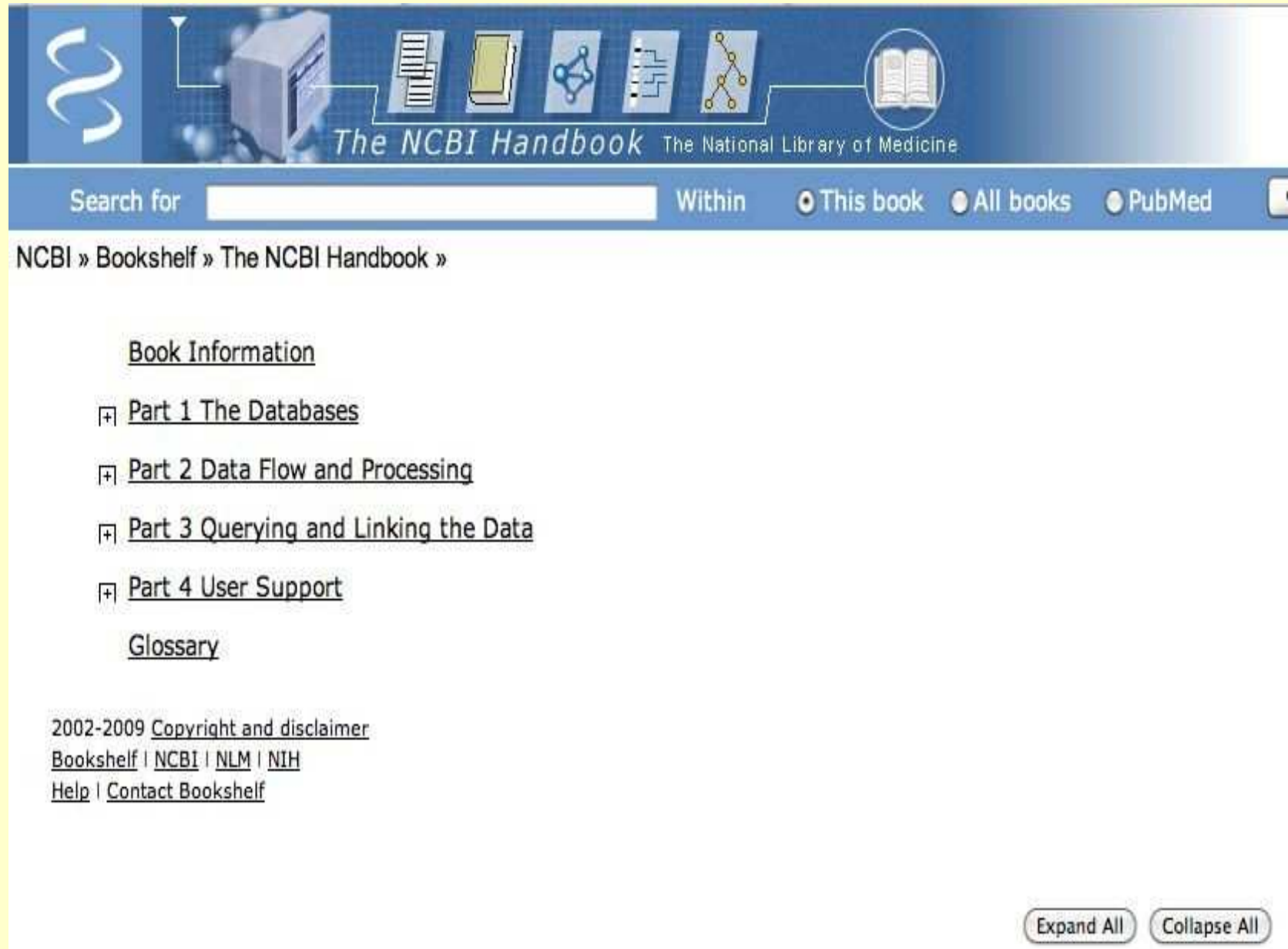

Higgins & Taylor

Bioinformatics: Sequence, Structure & Databanks



NCBI Handbook

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



The NCBI Handbook The National Library of Medicine

Search for Within ☒ This book ☐ All books ☐ PubMed

NCBI » Bookshelf » The NCBI Handbook »

Book Information

- ☒ Part 1 The Databases
- ☒ Part 2 Data Flow and Processing
- ☒ Part 3 Querying and Linking the Data
- ☒ Part 4 User Support

Glossary

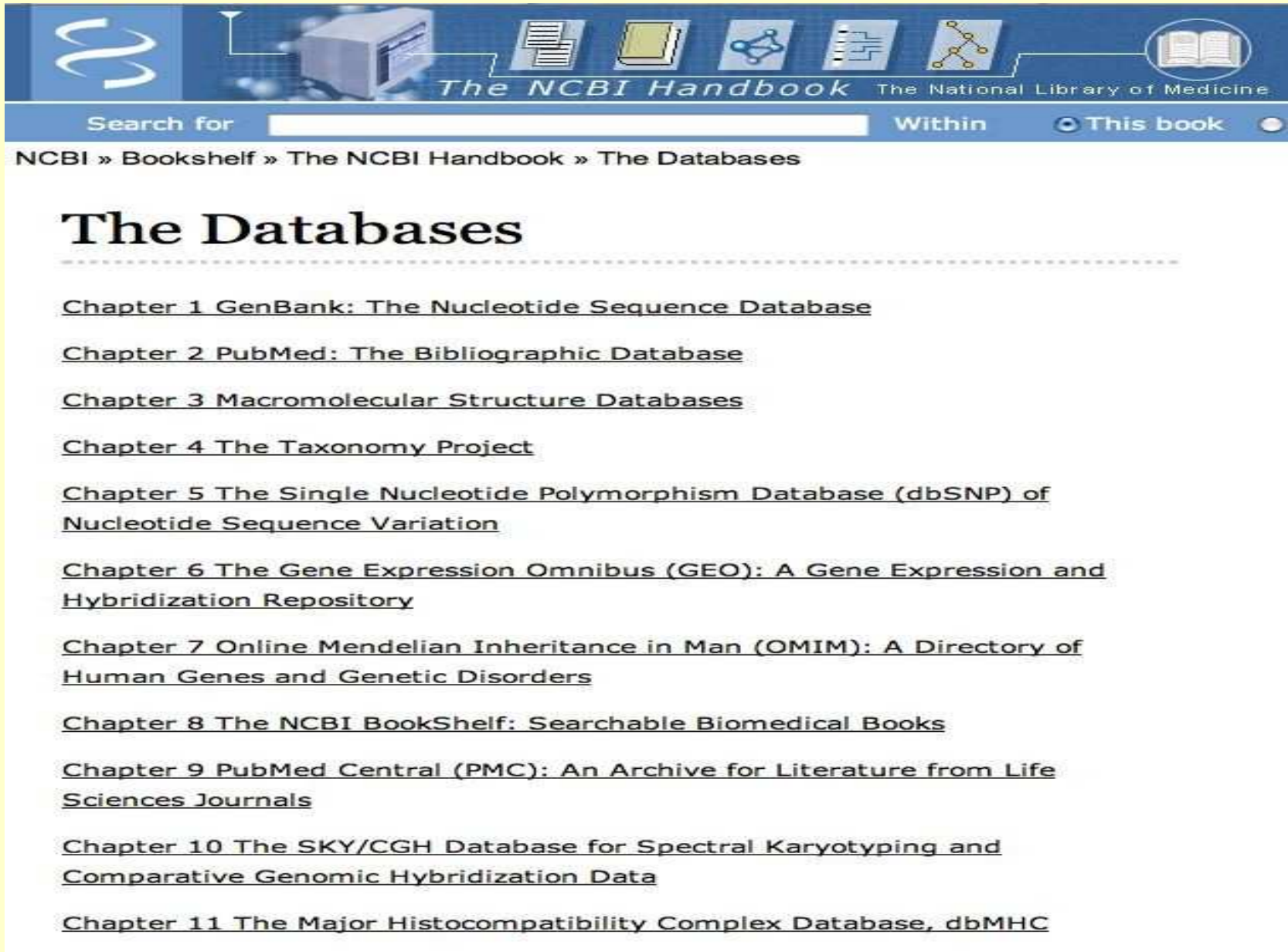

2002-2009 Copyright and disclaimer
Bookshelf | NCBI | NLM | NIH
Help | Contact Bookshelf

Expand All Collapse All



NCBI Handbook

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



The NCBI Handbook The National Library of Medicine

Search for Within ☒ This book ☐


NCBI » Bookshelf » The NCBI Handbook » The Databases



The Databases

- [Chapter 1 GenBank: The Nucleotide Sequence Database](#)
- [Chapter 2 PubMed: The Bibliographic Database](#)
- [Chapter 3 Macromolecular Structure Databases](#)
- [Chapter 4 The Taxonomy Project](#)
- [Chapter 5 The Single Nucleotide Polymorphism Database \(dbSNP\) of Nucleotide Sequence Variation](#)
- [Chapter 6 The Gene Expression Omnibus \(GEO\): A Gene Expression and Hybridization Repository](#)
- [Chapter 7 Online Mendelian Inheritance in Man \(OMIM\): A Directory of Human Genes and Genetic Disorders](#)
- [Chapter 8 The NCBI BookShelf: Searchable Biomedical Books](#)
- [Chapter 9 PubMed Central \(PMC\): An Archive for Literature from Life Sciences Journals](#)
- [Chapter 10 The SKY/CGH Database for Spectral Karyotyping and Comparative Genomic Hybridization Data](#)
- [Chapter 11 The Major Histocompatibility Complex Database, dbMHC](#)

EMBL-EBI Home Page

<http://www.ebi.ac.uk/>





EB-eye Search

All Databases

Enter Text Here

Go

Reset ?


Advanced Search

Give us feedback

DatabasesToolsEBI GroupsTrainingIndustryAbout UsHelpSite Index

Data Resources & Tools

<ul style="list-style-type: none">EMBL-BANKUniProtArrayExpressEnsemblInterProPDBe	<ul style="list-style-type: none">GenomesNucleotide SequencesProtein SequencesMacromolecular StructuresSmall Molecules	<ul style="list-style-type: none">Gene ExpressionMolecular InteractionsReactions & PathwaysProtein FamiliesEnzymes	<ul style="list-style-type: none">LiteratureTaxonomyOntologiesPatent Resources	<ul style="list-style-type: none">Sequence Similarity & AnalysisPattern & Motif SearchesStructure AnalysisText MiningDownloadsWeb Services
--	--	--	---	---



European Bioinformatics Institute


About the EBI

<ul style="list-style-type: none">ResearchPhD StudiesTrainingIndustry SupportGroup & Team LeadersEBI Funders	<ul style="list-style-type: none">User SupportEBI MissionPeopleEvents at the EBIGenome Campus EventsHow to Find us
---	---

EBI Hosted Project Websites

<ul style="list-style-type: none">1000 GenomesBioCatalogueBioSapiensE-MePEGAELIXIREMBRACEEMERALD	<ul style="list-style-type: none">ENFINFELICSIMPACTINSDCLRGSPINESYMBIOMatics
---	--

Latest News



New portal for plant genomics will support research into improved crops
08 October 2009

Today sees the launch of [Ensembl Plants](#) – a freely available web resource for plant genomics research – by EMBL-EBI, in partnership with the Cold Spring Harbor Laboratory, USA. Ensembl Plants allows researchers worldwide to access and visualise the results of genome-scale experiments in different plant species and will make it easier for scientists to improve the productivity and health of crops... [more](#)

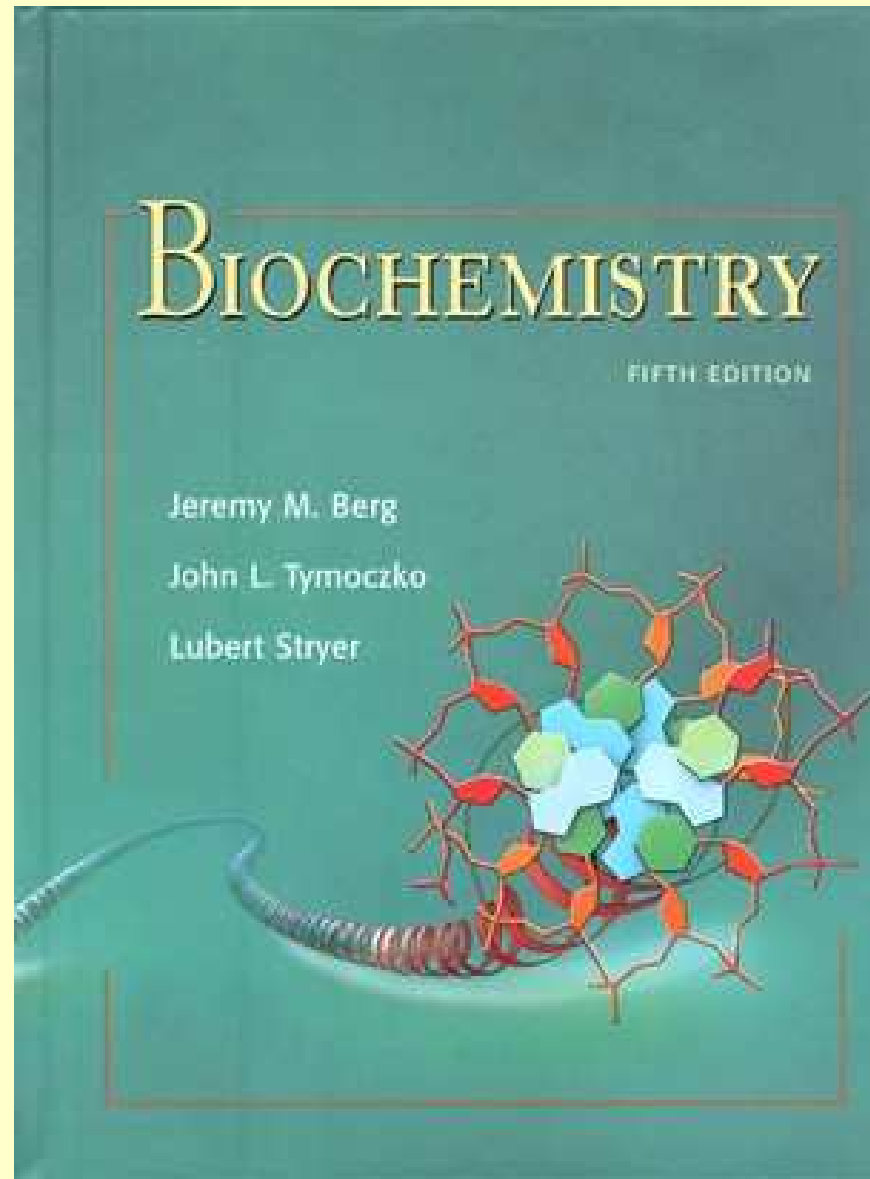
Research Highlights

EMBL-EBI articles are top of the list
20 November 2009

Articles on three resources hosted by EMBL-EBI ([PDBe](#), [Ensembl Genomes](#) and [Gene Expression Atlas](#)) are highlighted as featured articles in the latest Database issue of Nucleic Acids Research. Featured articles are selected by the journal's Executive Editors based upon their originality, significance and scientific excellence... [more](#)

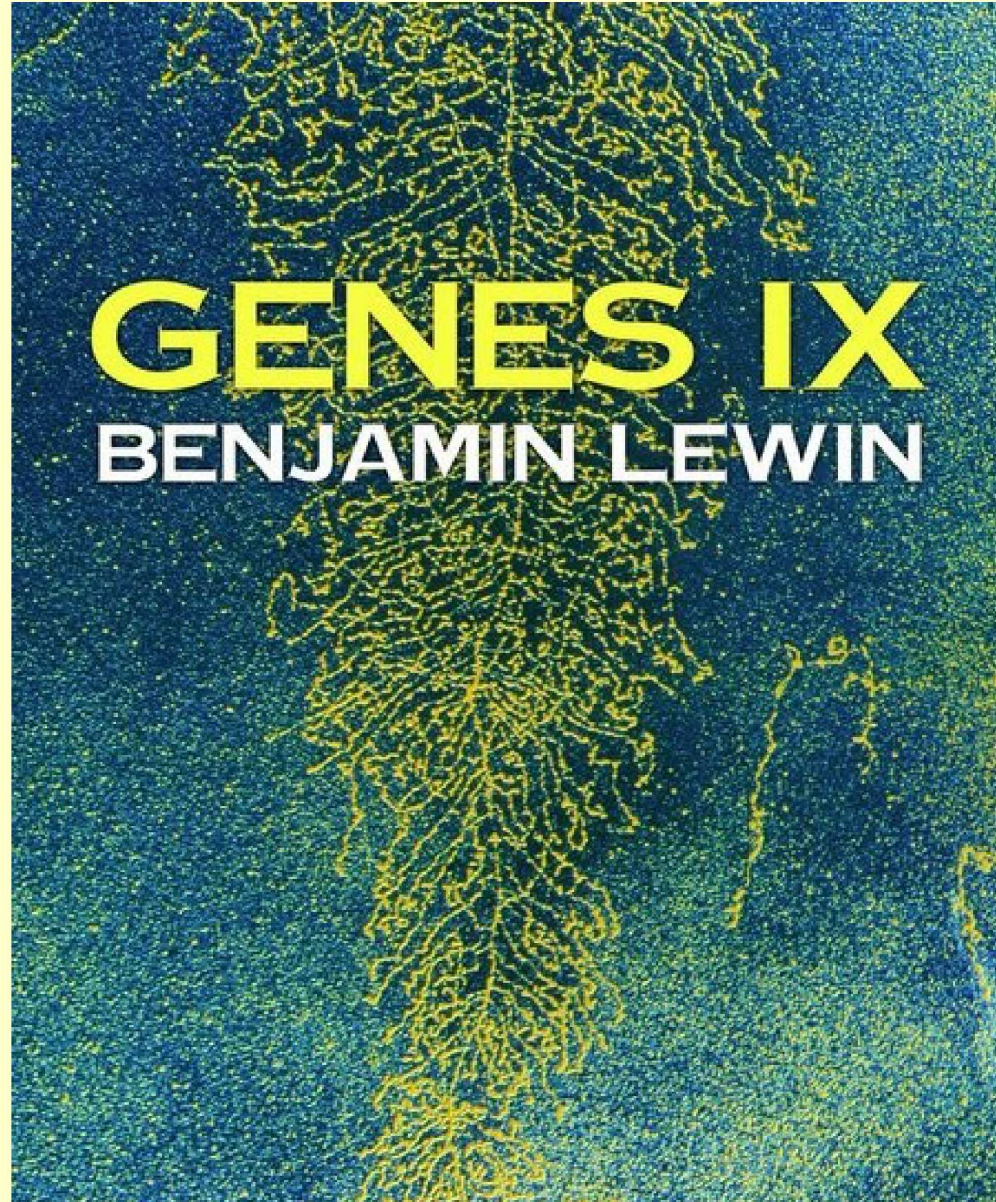
Berg, Tymoczko & Stryer

Biochemistry, Fifth Edition

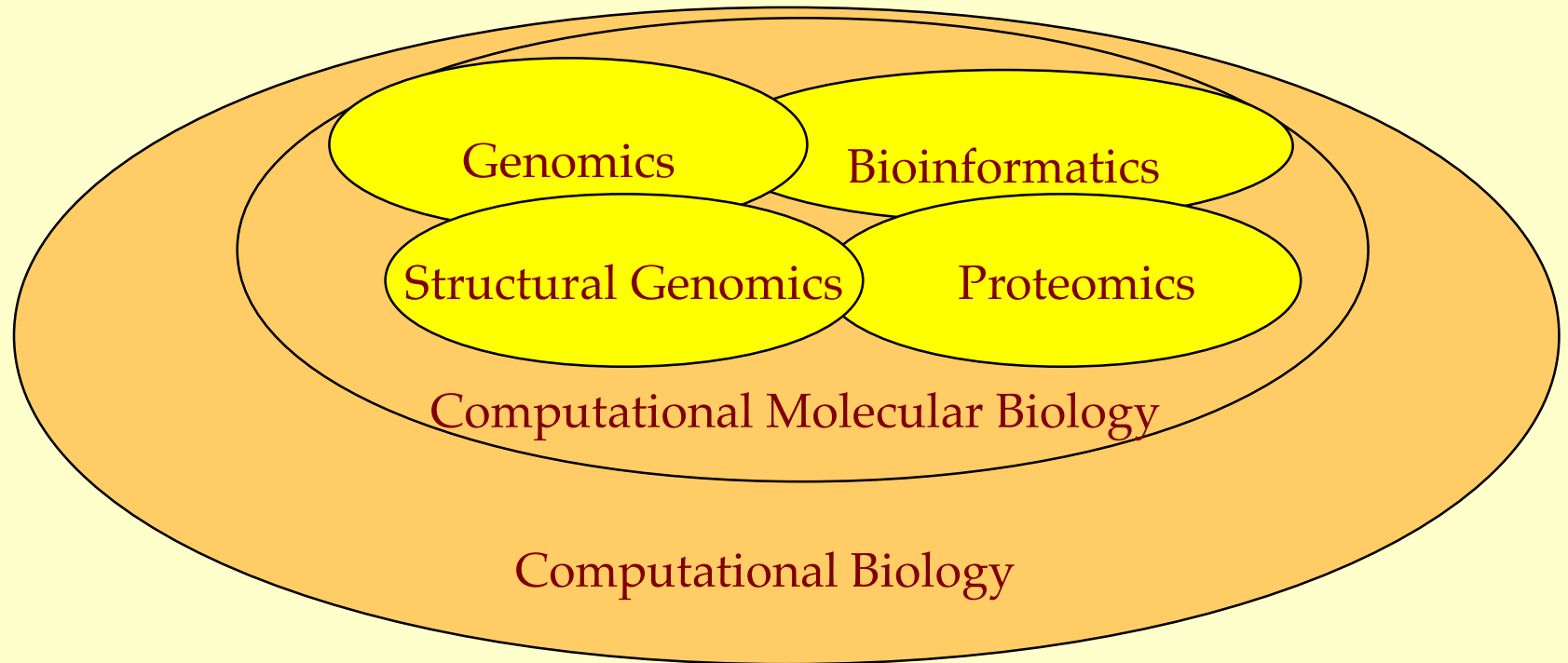


Benjamin Lewin

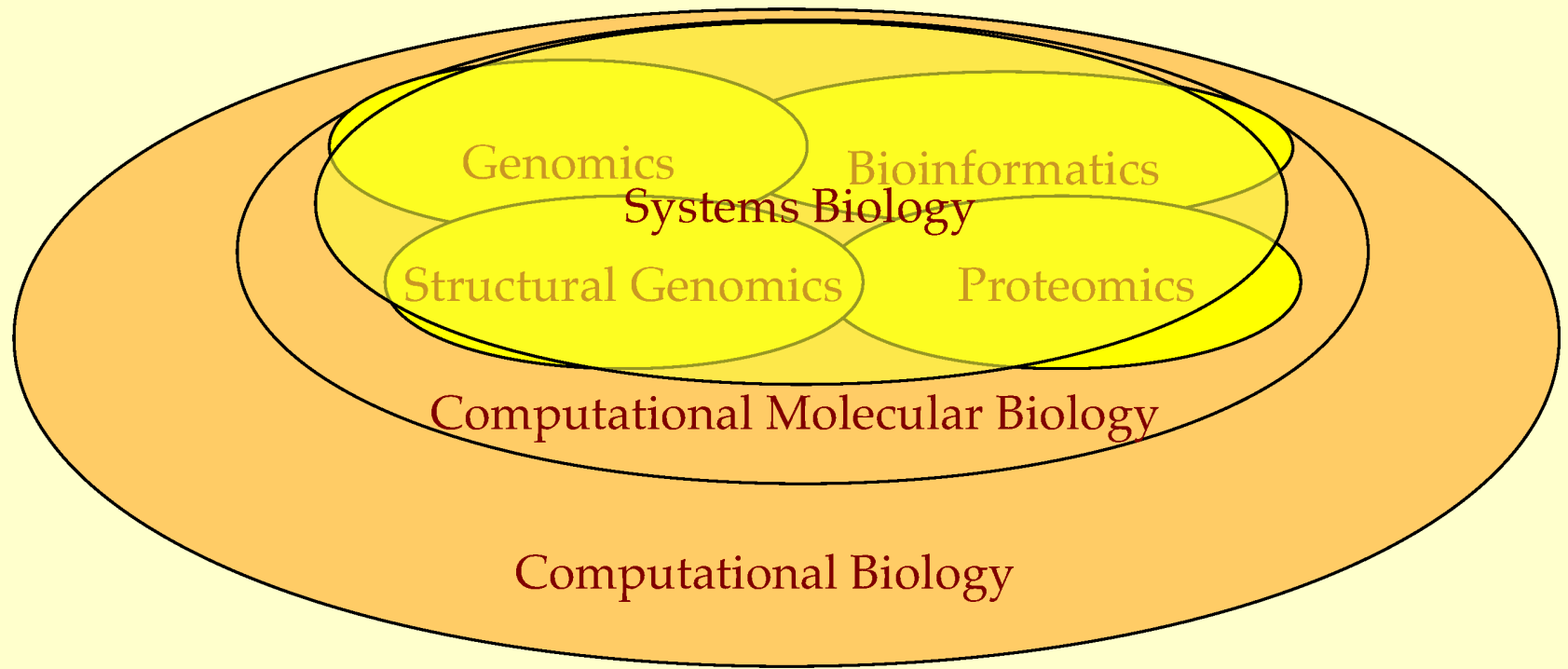
Genes IX



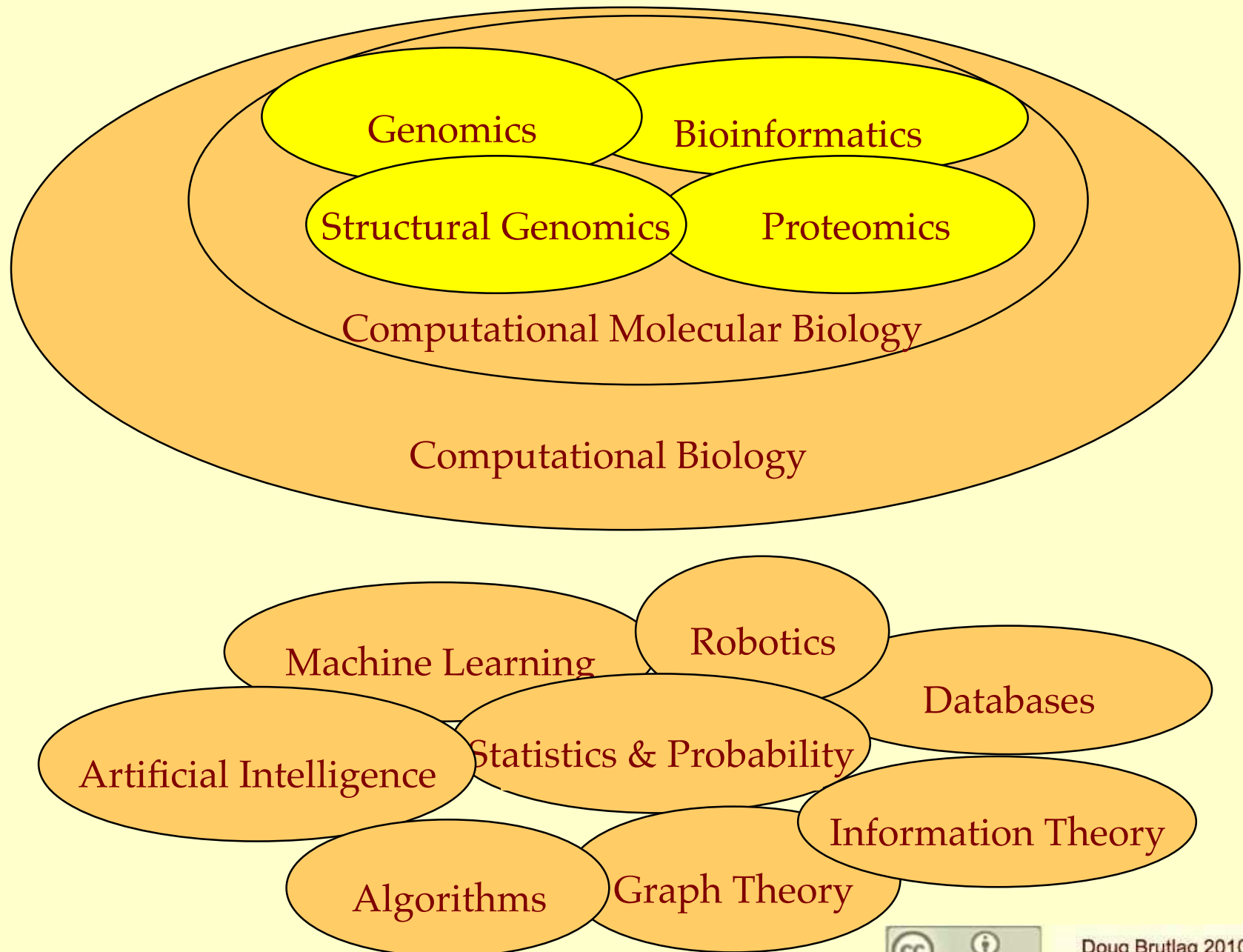
Genomics, Bioinformatics & Computational Biology



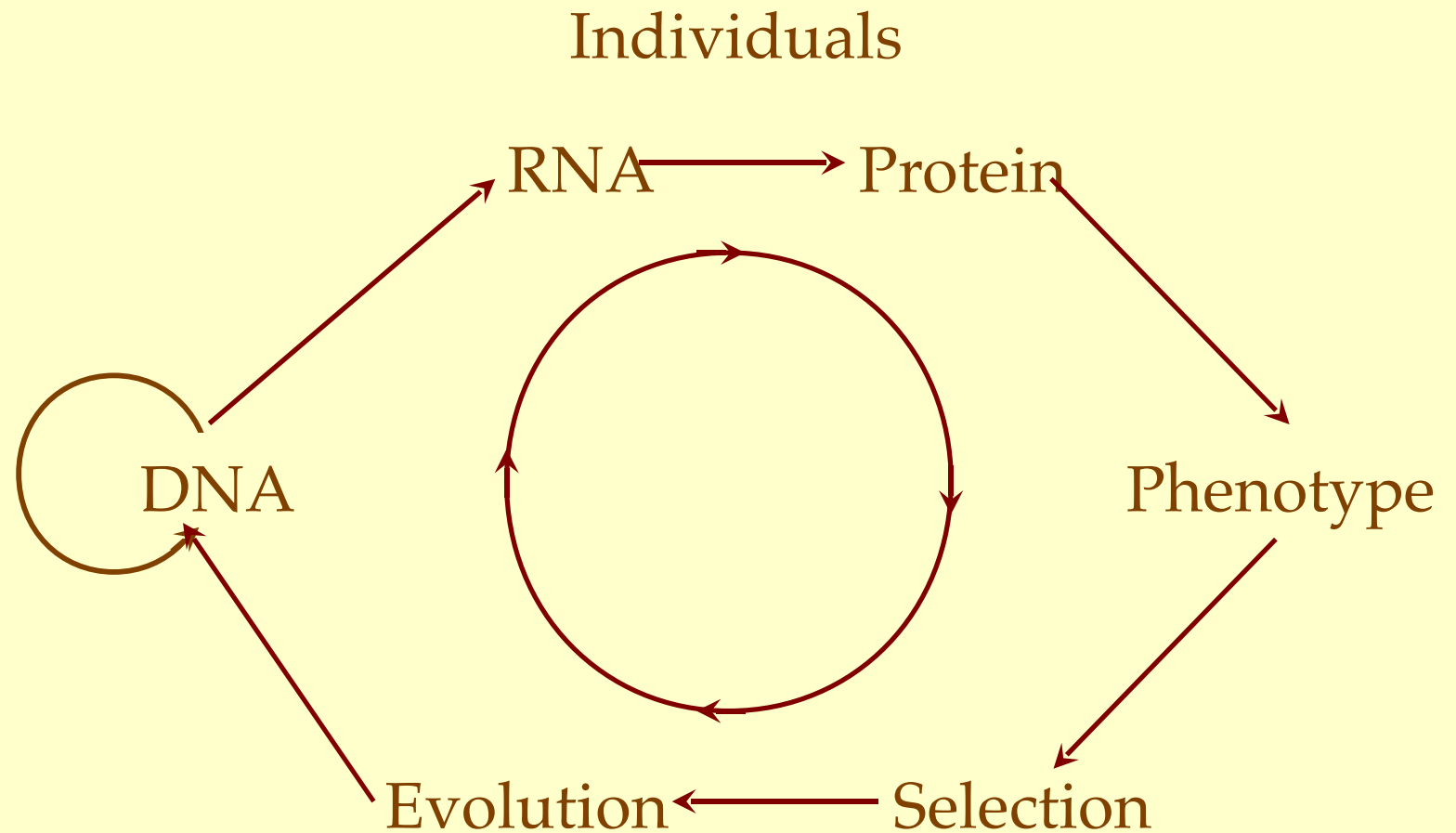
Genomics, Bioinformatics & Computational Biology



Genomics, Bioinformatics & Computational Biology



What is Bioinformatics?

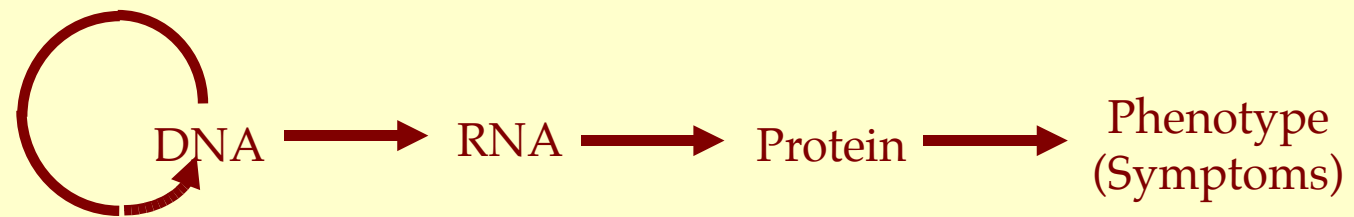


Biological Information

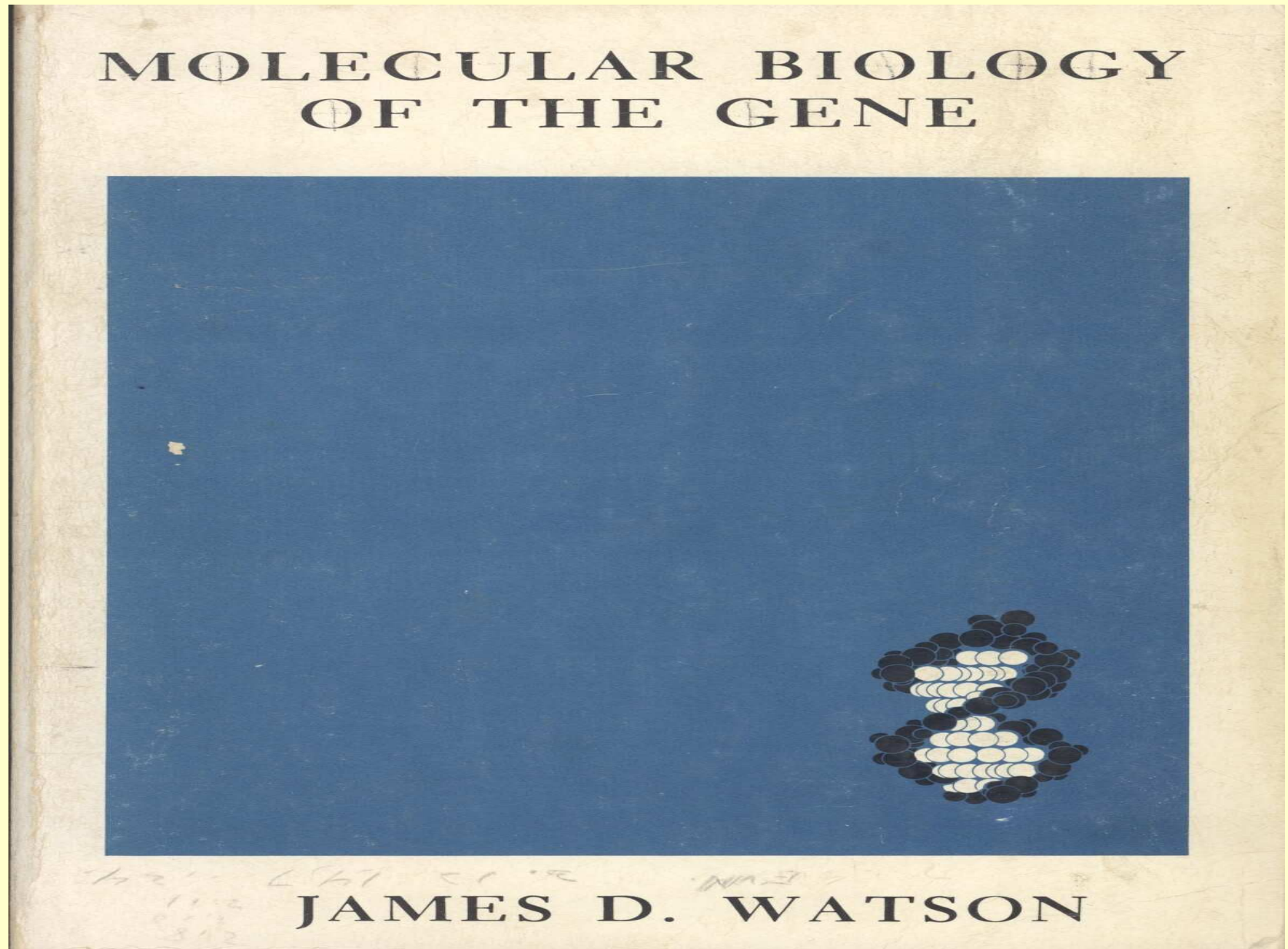
Computational Goals of Bioinformatics

- 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

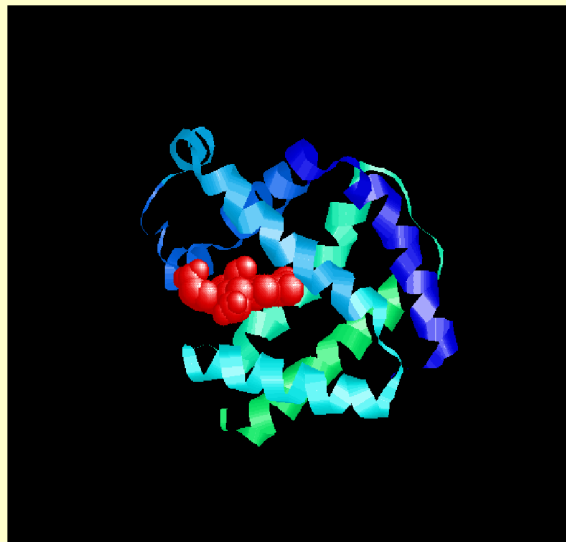


Central Paradigm of Bioinformatics

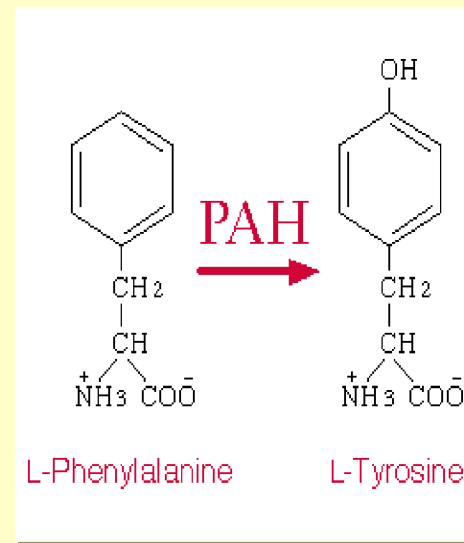
Genetic
Information

MVHLTPEEKT
AVNALWGKVN
VDAVGGEALG
RLLVVYPWTQ
RFFESFGDLS
SPDAVMGNPK
VKAHGKKVLG
AFSDGLAHLD
NLKGTFSQLS
ELHCDKLHVD
PENFRLLGNV
LVCVLARNFG
KEFTPQMCAA
YQKVAVGVAN
ALAHKYH

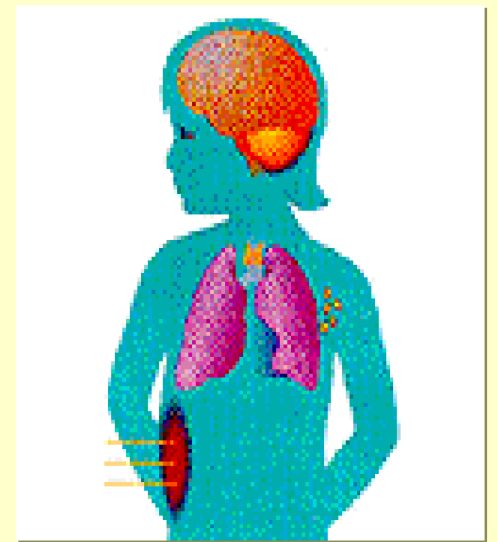
Molecular
Structure



Biochemical
Function



Phenotype
(Symptoms)

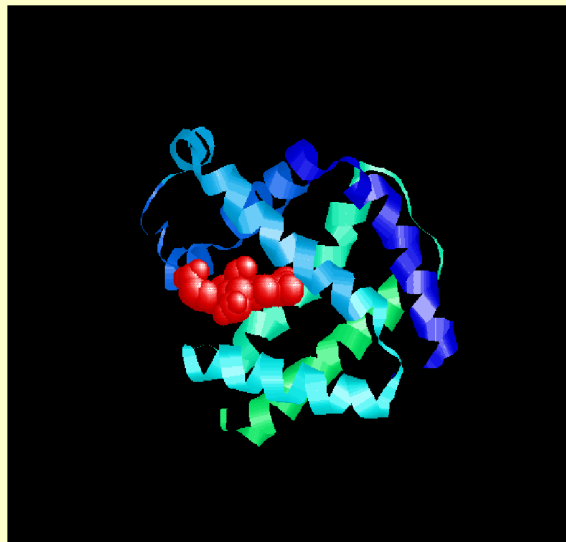


Central Paradigm of Bioinformatics

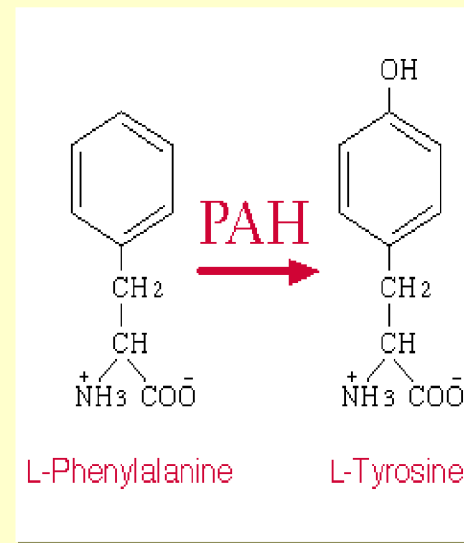
Genetic
Information

MVHLTPEEK
TAVNALWGK
VAVDVGGEAL
GRLLVVYPWT
QRFFESFGDLS
SPDAVMGNPK
VKAHGKKVLG
AFSDGLAHLD
NLKGTFSQLS
ELHCDKLHVD
PENFRLLGNV
LVCVLARNFG
KEFTPQMCAA
YQKVVAGVAN
ALAHKYH

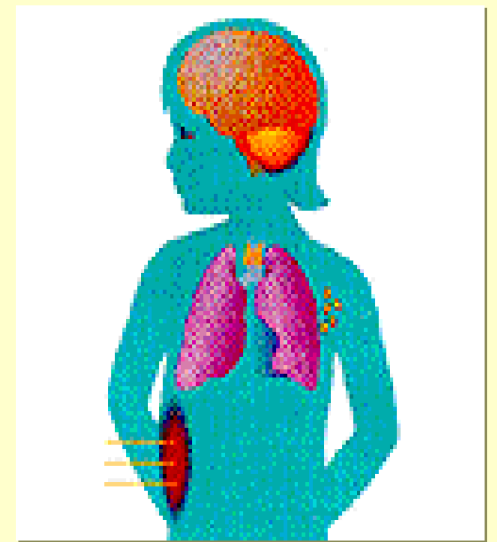
Molecular
Structure



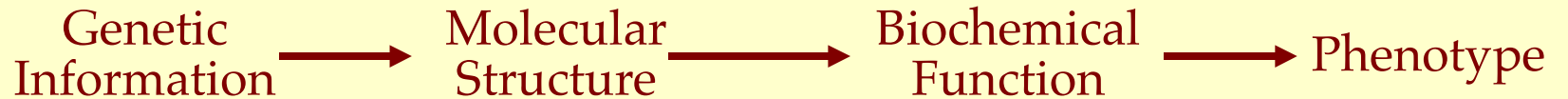
Biochemical
Function



Phenotype
(Symptoms)



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

