

Contents

1	Computers, Operating Systems and the Internet	1
1.1	Computers and Operating Systems	1
1.2	Internet and WWW	5
1.3	The Physical Connection to the Internet.	6
1.4	The Logical Connection to the Internet	8
1.5	Internet Services.	8
1.5.1	E-mail.	8
1.5.2	FTP	11
1.5.3	World Wide Web	16
1.6	Using UNIX.	18
1.7	The History of Bioinformatics	23
	WWW Links	29
2	The Biological Foundations of Bioinformatics	31
2.1	Nucleic Acids and Proteins	31
2.2	Structure of the Nucleic Acids DNA and RNA	31
2.3	The Storage of Genetic Information	33
2.4	The Structure of Proteins	36
2.4.1	Primary Structure.	36
2.4.2	Secondary Structure.	39
2.5	Tertiary and Quaternary Structure	41
	WWW Links	43

3	Biological Databases	45
3.1	Biological Knowledge is Stored in Global Databases	45
3.2	Primary Databases	46
3.2.1	Nucleotide Sequence Databases	46
3.2.2	Protein Sequence Databases	53
3.3	Secondary Databases	58
3.3.1	PROSITE	58
3.3.2	PRINTS	60
3.3.3	Pfam	61
3.3.4	Interpro	61
3.4	Genotype-Phenotype Databases	62
3.4.1	PhenomicDB	63
3.5	Molecular Structure Databases	64
3.5.1	Protein Data Bank (PDB)	64
3.5.2	SCOP	65
3.5.3	CATH	67
3.5.4	PubChem	67
	WWW Links	72
4	Sequence Comparisons and Sequence-Based Database Searches	75
4.1	Pairwise and Multiple Sequence Comparisons ...	75
4.2	Database Searches with Nucleotide and Protein Sequences	83
4.2.1	Important Algorithms for Database Searching	87
4.3	Software for Sequence Analysis	87
	WWW Links	92
5	The Decoding of Eukaryotic Genomes	95
5.1	The Sequencing of Complete Genomes	95
5.2	The Characterization of Genomes using STS and EST Sequences	96

5.2.1	Sequence Tagged Sites are Landmarks in the Human Genome.	96
5.2.2	Expressed Sequence Tags.	97
5.3	The Implementation of an EST Project.	99
5.4	The Identification of Unknown Genes.	103
5.5	The Discovery of Splice Variants	105
5.6	Genetic Causes for Individual Differences	108
5.6.1	Pharmacogenetics and Individual Medicine.	110
	WWW Links	115
6	Protein Structures and Structure-Based Rational Drug Design	117
6.1	Protein Structure	117
6.2	Signal Peptides	118
6.3	Transmembrane Proteins	122
6.4	Analyses of Protein Structures	123
6.4.1	Protein Modeling	124
6.4.2	The Determination of Protein Structures by High Throughput Methods	125
6.5	Structure-Based Rational Drug Design.	126
6.5.1	A Docking Example	128
6.5.2	Pharmacophore Modeling.	131
6.5.3	Successes of Structure-Based Rational Drug Design	133
	WWW Links	136
7	Systems Biology: The Functional Analysis of Genomes	139
7.1	The Identification of the Cellular Functions of Gene Products.	139
7.1.1	Transcriptomics	140
7.1.2	Proteomics.	153
7.1.3	Metabolomics	163
7.1.4	Phenomics	166

7.2	Systems Biology	170
	WWW Links	178
8	Comparative Genome Analyses	181
8.1	The Era of Genome Sequencing.....	181
8.2	Drug Research at the Target Protein	182
8.3	Comparative Genome Analyses Provide Information about the Biology of Organisms	185
8.3.1	Genome Structure	185
8.3.2	Coding Regions.....	186
8.3.3	Noncoding Regions	188
8.4	Comparative Metabolic Analyses.....	188
8.4.1	Kyoto Encyclopedia of Genes and Genomes.....	191
8.5	Groups of Orthologous Proteins	195
	WWW Links	199
9	Solutions to the Exercises	201
	Glossary	247
	Index.....	277