

Paul M. Selzer
Richard J. Marhöfer
Oliver Koch

Applied Bioinformatics

An Introduction

Second Edition

 Springer

Contents

1	The Biological Foundations of Bioinformatics	1
1.1	Nucleic Acids and Proteins.....	2
1.2	Structure of the Nucleic Acids DNA and RNA	2
1.3	The Storage of Genetic Information.....	2
1.4	The Structure of Proteins.....	7
1.4.1	Primary Structure.....	7
1.4.2	Secondary Structure	7
1.4.3	Tertiary and Quaternary Structure	10
1.5	Exercises.....	11
	References.....	12
2	Biological Databases	13
2.1	Biological Knowledge is Stored in Global Databases.....	14
2.2	Primary Databases.....	14
2.2.1	Nucleotide Sequence Databases.....	14
2.2.2	Protein Sequence Databases	20
2.3	Secondary Databases.....	23
2.3.1	Prosite.....	23
2.3.2	PRINTS.....	24
2.3.3	Pfam.....	25
2.3.4	Interpro	25
2.4	Genotype-Phenotype Databases	25
2.4.1	PhenomicDB.....	26
2.5	Molecular Structure Databases.....	27
2.5.1	Protein Data Bank	27
2.5.2	SCOP.....	29
2.5.3	CATH.....	29
2.5.4	PubChem.....	30
2.6	Exercises.....	31
	References.....	33
3	Sequence Comparisons and Sequence-Based Database Searches	35
3.1	Pairwise and Multiple Sequence Comparisons.....	36
3.2	Database Searches with Nucleotide and Protein Sequences	42
3.2.1	Important Algorithms for Database Searching.....	45
3.3	Software for Sequence Analysis	46
3.4	Exercises.....	48
	References.....	49

4	The Decoding of Eukaryotic Genomes	51
4.1	The Sequencing of Complete Genomes	52
4.2	Characterization of Genomes Using STS and EST Sequences	52
4.2.1	Sequence-Tagged Sites are Landmarks in the Human Genome.....	52
4.2.2	Expressed Sequence Tags.....	53
4.3	EST Project Implementation	55
4.4	Identification of Unknown Genes	56
4.5	The Discovery of Splice Variants	60
4.6	Genetic Causes for Individual Differences	61
4.6.1	Pharmacogenetics.....	63
4.6.2	Personalized Medicine and Biomarkers.....	65
4.6.3	Next-Generation Sequencing (NGS).....	67
4.6.4	Proteogenomics.....	68
4.7	Exercises	69
	References.....	71
5	Protein Structures and Structure-Based Rational Drug Design	73
5.1	Protein Structure	74
5.2	Signal Peptides	74
5.3	Transmembrane Proteins	77
5.4	Analyses of Protein Structures	78
5.4.1	Protein Modeling.....	78
5.4.2	Determination of Protein Structures by High-Throughput Methods.....	78
5.5	Structure-Based Rational Drug Design	79
5.5.1	A Docking Example Using DOCK.....	80
5.5.2	Docking Example Using GOLD.....	83
5.5.3	Pharmacophore Modeling and Searches.....	84
5.5.4	Successes of Structure-Based Rational Drug Design.....	85
5.6	Exercises	86
	References.....	88
6	The Functional Analysis of Genomes	91
6.1	The Identification of the Cellular Functions of Gene Products	92
6.1.1	Transcriptomics.....	93
6.1.2	Proteomics.....	102
6.1.3	Metabolomics.....	110
6.1.4	Phenomomics.....	112
6.2	Systems Biology	115
6.3	Exercises	118
	References.....	120
7	Comparative Genome Analyses	123
7.1	The Era of Genome Sequencing	124
7.2	Drug Research on the Target Protein	124

7.3	Comparative Genome Analyses Provide Information	
	About the Biology of Organisms	126
7.3.1	Genome Structure	126
7.3.2	Coding Regions	128
7.3.3	Noncoding Regions.....	128
7.4	Comparative Metabolic Analyses	129
7.4.1	Kyoto Encyclopedia of Genes and Genomes	133
7.5	Groups of Orthologous Proteins	135
7.6	Exercises	138
	References.....	139
	 Supplementary Information	
	Solutions to Exercises	142
	Glossary.....	164
	Index	179