

CONTENTS

Preface

Acknowledgements

ix

1. History, Scope and Importance	1
1.1. Important Contributions	1
1.2. Sequencing Development	5
1.3. Aims and Tasks of Bioinformatics	7
1.4. Applications of Bioinformatics	9
1.4.1. Sequence Homology Analysis	9
1.4.2. Drug Design	9
1.4.3. Predictive Functions	9
1.4.4. Medical Areas	10
1.4.5. Intellectual Property Rights	10
1.5. Challenges and Opportunities	11
Study Questions	12
2. Computers, Internet, World Wide Web and NCBI	13
2.1. Computers and Programs	13
2.2. Internet	15
2.3. World Wide Web	17
2.4. Browsers	18
2.5. EMBnet and SRS	19
2.6. NCBI	21
Study Questions	24
3. DNA, RNA and Proteins	25
3.1. Background	25
3.2. DNA	28
3.3. RNA	32
3.4. Transcription and Translation	37
3.5. Proteins	40
Study Questions	44
4. DNA and Protein Sequencing and Analysis	45
4.1. Genomics and Proteomics	45
4.2. Genome Mapping	48
4.3. DNA Sequencing Methods	49
4.4. Open Reading Frame	52
4.5. Determining Sequence of a Clone	53
4.6. Expressed Sequence Tags	54

4.7.	Protein Sequencing	57
4.8.	Gene and Protein expression analysis	57
4.8.1.	DNA Microarrays	58
4.8.2.	Protein Expression Analysis	62
4.8.3.	Gene Discovery	64
4.9.	Human Genome Project	66
	Study Questions	70
5.	Databases, Tools and their Uses	71
5.1.	Importance of Databases	71
5.2.	Nucleic acid Sequence Databases	75
5.3.	Protein Sequence Databases	78
5.4.	Structure Databases	80
5.5.	Bibliographic Databases and Virtual Library	86
5.6.	Specialized Analysis Packages	87
5.7.	Use of Databases	91
	Study Questions	91
6.	Sequence Alignment	92
6.1.	Algorithm	92
6.2.	Goals and Types of Alignment	93
6.3.	Study of Similarities	94
6.4.	Scoring Mutations, Deletions and Substitutions	97
6.5.	Sequence Alignment Methods	100
6.6.	Pairwise Alignment	101
6.7.	Multiple Sequence Alignment	105
6.8.	Algorithms for Identifying Domains within a Protein Structure	109
6.9.	Algorithms for Structural Comparison	109
6.10.	Carrying out a Sequence Search	110
	Study Questions	113
7.	Predictive Methods using DNA and Protein Sequences	114
7.1.	Gene Prediction Strategies	114
7.1.1.	Gene Prediction Programs	116
7.2.	Protein Prediction Strategies	117
7.2.1.	Secondary Structure Prediction	118
7.2.2.	Intrinsic Tendency of Amino acids to form β -turns	121
7.2.3.	Rotamer Libraries	121
7.2.4.	Three-Dimensional Structure Prediction	121
7.2.5.	Comparative Modeling	122
7.2.6.	Threading	123

7.2.7. Energy-based Prediction of Protein Structure	123
7.2.8. Protein Function Prediction	124
7.3. Protein Prediction Programs	125
7.4. Molecular Visualization	126
Study Questions	127
8. Homology, Phylogeny and Evolutionary Trees	129
8.1. Homology and Similarity	129
8.2. Phylogeny and Relationships	131
8.2.1. Approaches used in Phylogenetic Analysis	132
8.2.2. Phylogenetic Trees	135
8.2.3. Tree-building Methods	136
8.3. Molecular Approaches to Phylogeny	139
8.4. Phylogenetic Analysis Databases	141
Study Questions	141
9. Drug Discovery and Pharmainformatics	142
9.1. Discovering a Drug	142
9.1.1. Target Identification and Validation	142
9.1.2. Identifying the Lead Compound	144
9.1.3. Optimization of Lead Compound	144
9.2. Pharmainformatics	146
9.2.1. Chemical Libraries	147
9.3. Search Programs	147
Study Questions	152
Appendix: List of Special Websites and Web Addresses	153
References	170
Glossary	173
Index	187