

Contents

<i>Preface</i>	<i>page</i>	xi
1 Introduction		1
1.1 Original problems		2
1.2 Additional problems		5
1.3 Further reading		23
2 Pairwise alignment		24
2.1 Original problems		24
2.2 Additional problems and theory		43
2.2.1 Derivation of the amino acid substitution matrices (PAM series)		46
2.2.2 Distributions of similarity scores		57
2.2.3 Distribution of the length of the longest common word among several unrelated sequences		62
2.3 Further reading		65
3 Markov chains and hidden Markov models		67
3.1 Original problems		68
3.2 Additional problems and theory		77
3.2.1 Probabilistic models for sequences of symbols: selection of the model and parameter estimation		86
3.2.2 Bayesian approach to sequence composition analysis: the segmentation model by Liu and Lawrence		95
3.3 Further reading		102
4 Pairwise alignment using HMMs		104
4.1 Original problems		105
4.2 Additional problems		113
4.3 Further reading		125

viii	<i>Contents</i>	
5	Profile HMMs for sequence families	126
5.1	Original problems	127
5.2	Additional problems and theory	137
5.2.1	Discrimination function and maximum discrimination weights	150
5.3	Further reading	161
6	Multiple sequence alignment methods	162
6.1	Original problem	163
6.2	Additional problems and theory	163
6.2.1	Carrillo–Lipman multiple alignment algorithm	164
6.2.2	Progressive alignments: the Feng–Doolittle algorithm	171
6.2.3	Gibbs sampling algorithm for local multiple alignment	179
6.3	Further reading	181
7	Building phylogenetic trees	183
7.1	Original problems	183
7.2	Additional problems	211
7.3	Further reading	215
8	Probabilistic approaches to phylogeny	218
8.1	Original problems	219
8.1.1	Bayesian approach to finding the optimal tree and the Mau–Newton–Larget algorithm	235
8.2	Additional problems and theory	259
8.2.1	Relationship between sequence evolution models described by the Markov and the Poisson processes	264
8.2.2	Thorne–Kishino–Felsenstein model of sequence evolution with substitutions, insertions, and deletions	270
8.2.3	More on the rates of substitution	275
8.3	Further reading	277
9	Transformational grammars	279
9.1	Original problems	280
9.2	Further reading	290
10	RNA structure analysis	291
10.1	Original problems	292
10.2	Further reading	308

Cambridge University Press
978-0-521-61230-2 - Problems and Solutions in Biological Sequence Analysis
Mark Borodovsky and Svetlana Ekisheva
Table of Contents
[More information](#)

	<i>Contents</i>	ix
11	Background on probability	311
	11.1 Original problems	311
	11.2 Additional problem	326
	11.3 Further reading	327
	<i>References</i>	328
	<i>Index</i>	343