

Jerome K. Percus

Cambridge Studies in Mathematical Biology

Mathematics of Genome Analysis



Contents

<i>Preface</i>	<i>page ix</i>
1 Decomposing DNA	1
1.1 DNA Sequences	1
1.2 Restriction Fragments	6
1.3 Clone Libraries	9
Assignment 1	9
2 Recomposing DNA	13
2.1 Fingerprint Assembly	13
2.2 Anchoring	18
2.3 Restriction-Fragment-Length Polymorphism (RFLP) Analysis	23
Assignment 2	27
2.4 Pooling	28
Assignment 3	35
2.5 Reprise	36
3 Sequence Statistics	42
3.1 Local Properties of DNA	42
3.2 Long-Range Properties of DNA	49
3.2.1 Longest Repeat	50
Assignment 4	53
3.2.2 Displaced Correlations	53
3.2.3 Nucleotide-Level Criteria	54
3.2.4 Batch-Level Criteria	61
3.2.5 Statistical Models	65
Assignment 5	73
3.3 Other Measures of Significance	73

3.3.1 Spectral Analysis	73
3.3.2 Entropic Criteria	78
4 Sequence Comparison	81
4.1 Basic Matching	81
4.1.1 Mutual-Exclusion Model	82
4.1.2 Independence Model	85
4.1.3 Direct Asymptotic Evaluation	87
4.1.4 Extreme-Value Technique	89
Assignment 6	96
4.2 Matching with Imperfections	92
4.2.1 Score Distribution	92
4.2.2 Penalty-Free Limit	97
4.2.3 Effect of Indel Penalty	99
4.2.4 Score Acquisition	100
4.3 Multisequence Comparison	103
4.3.1 Locating a Common Pattern	103
4.3.2 Assessing Significance	105
Assignment 7	107
4.3.3 Category Analysis	108
4.3.4 Adaptive Techniques	111
5 Spatial Structure and Dynamics of DNA	118
5.1 Thermal Behavior	118
5.2 Dynamics	123
5.3 Effect of Heterogeneity	127
Assignment 8	127
<i>Bibliography</i>	129
<i>Index</i>	137