
Contents

LIST OF FIGURES	vii
LIST OF TABLES	xi
PREFACE	xv
ACKNOWLEDGMENTS	xix
CHAPTER 1 ■ Overview of Bioinformatics Databases	1
1.1 INTRODUCTION	1
1.2 SEQUENCE DATABASES	2
1.3 PHYLOGENETIC DATABASES	7
1.4 STRUCTURE AND PATHWAY DATABASES	12
1.5 MICROARRAY AND BOUTIQUE DATABASES	14
CHAPTER 2 ■ Biological Data Cleaning	17
2.1 INTRODUCTION	17
2.2 GENERAL DATA CLEANING	24
2.3 CASE STUDY IN BIOLOGICAL DATA CLEANING	34
CHAPTER 3 ■ Biological Data Integration	47
3.1 INTRODUCTION	47
3.2 GENERAL DATA INTEGRATION	48
3.3 TOPICS IN BIOLOGICAL DATA INTEGRATION	50
CHAPTER 4 ■ Biological Data Searching	57
4.1 INTRODUCTION	57
4.2 BIOLOGICAL DATA SEARCHING USING BLAST	58
4.3 BIOLOGICAL DATA SEARCHING USING THE UCSC GENOME BROWSER	58

4.4	CASE STUDY IN PHYLOGENETIC TREE DATABASE SEARCH	59
4.5	CASE STUDY IN RNA PSEUDOKNOT DATABASE SEARCH	73
CHAPTER 5 ■ Biological Data Mining		91
5.1	INTRODUCTION	91
5.2	GENERAL DATA MINING	93
5.3	BIOLOGICAL DATA MINING	94
5.4	CASE STUDY IN BIOLOGICAL MOTIF DISCOVERY	101
5.5	CASE STUDY IN BIOLOGICAL DATA MINING	111
CHAPTER 6 ■ Biological Network Inference		133
6.1	INTRODUCTION	133
6.2	GENE REGULATORY NETWORK INFERENCE	135
CHAPTER 7 ■ Cloud-Based Biological Data Processing		193
7.1	INTRODUCTION	193
7.2	DATA PROCESSING IN THE CLOUD	198
Bibliography		217
Index		261