

Lecture Notes in Bioinformatics

8248

Edited by S. Istrail, P. Pevzner, and M. Waterman

Editorial Board: A. Apostolico S. Brunak M. Gelfand

T. Lengauer S. Miyano G. Myers M.-F. Sagot D. Sankoff

R. Shamir T. Speed M. Vingron W. Wong

Subseries of Lecture Notes in Computer Science

Wiesława Widlak

Molecular Biology

Not Only for Bioinformaticians



Springer

Author

Wiesława Widlak
Maria Skłodowska-Curie Memorial Cancer Center
and Institute of Oncology, Gliwice Branch
Wybrzeże Armii Krajowej 15, 44-101 Gliwice, Poland
E-mail: wwidlak@io.gliwice.pl

ISSN 0302-9743

e-ISSN 1611-3349

ISBN 978-3-642-45360-1

e-ISBN 978-3-642-45361-8

DOI 10.1007/978-3-642-45361-8

Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013955613

CR Subject Classification (1998): J.3

LNCS Sublibrary: SL 8 – Bioinformatics

© Springer-Verlag Berlin Heidelberg 2013

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The book is addressed to beginners in molecular biology, especially to computer scientists who would like to work as bioinformaticians (computational biologists). Bioinformatics, which could be defined as the application of computer science and information technology to the field of biology and medicine, has been rapidly developing over the past few decades. It generates new knowledge as well as the computational tools to create that knowledge. Certainly, understanding of the basic processes in living organisms is indispensable for bioinformaticians in the creation of knowledge. The book contains information about these processes presented in a condensed manner. Additionally, principles of several high-throughput technologies in molecular biology, which need the assistance of bioinformaticians, are explained from a biological point of view. Although some processes observed in living organisms are described here in a more detailed fashion, the book may be treated as an introduction to molecular biology. Thus it could be the first choice for people who are unfamiliar with the subject.

Contents

Preface	V
1 Cells and Viruses	1
1.1 Cell Structure	1
1.1.1 Prokaryotic Cell Structure	2
1.1.2 Eukaryotic Cell Structure	3
1.2 Biological Membranes and Their Lipid Component	4
1.2.1 Phospholipids	5
1.2.2 Cholesterol and Steroids	6
1.3 Nucleus	7
1.3.1 Chromosomes and Karyotype	7
1.4 Cytoplasmic Organelles	9
1.4.1 Mitochondria	9
1.4.2 Chloroplasts	10
1.4.3 Endoplasmic Reticulum and Golgi Complex	10
1.4.4 Peroxisomes, Lysosomes, Vacuoles, and Glyoxysomes	10
1.5 Viruses, Bacteriophages, and Virophages	11
1.5.1 Classification of Viruses	12
References	12
2 Protein Structure and Function	15
2.1 Amino Acids	15
2.2 Peptides and Proteins	17
2.3 Levels of Protein Structure	18
2.3.1 Hydrogen Bonds	19
2.3.2 Secondary Structure of Proteins	20
2.3.2.1 Alpha Helix	20
2.3.2.2 Beta Strand, Beta Sheet, and Beta Barrel ...	20
2.3.3 Tertiary and Quaternary Structure of Proteins	21

2.4	Proteins Functions	22
2.4.1	Enzymes	23
2.4.2	Membrane Proteins	25
2.4.2.1	Proteins Involved in Immune Response	26
2.4.2.2	Cell Surface Receptors	27
2.4.2.3	Ion Channels	27
2.4.3	Structural Proteins	28
	References	28
3	Nucleic Acids	31
3.1	Nucleotides	31
3.1.1	Pentoses	32
3.1.2	Nitrogenous Bases of Nucleotides	32
3.2	DNA Properties	34
3.2.1	Nucleotide Chain	34
3.2.2	DNA Structure	36
3.3	Structure and Function of RNA	37
3.3.1	Messenger RNA	38
3.3.2	Transfer RNA	39
3.3.3	Ribosomes and rRNA	39
3.3.4	Non Coding RNA	40
3.4	The Organization of Genetic Material in the Nucleus - Chromatin	42
3.4.1	Histones and Nucleosomes	42
3.4.2	Chromatin Structure	43
3.5	Genetic Code	44
3.6	Genes and Genome	45
3.7	Repetitive Sequences	46
	References	47
4	DNA Replication, Mutations, and Repair	49
4.1	DNA Replication Mechanism	49
4.1.1	DNA Polymerases	50
4.1.2	Polymerase Chain Reaction (PCR)	52
4.2	Telomerases and Cellular Senescence	53
4.3	Topoisomerases	54
4.4	Mutations and Their Consequences	56
4.4.1	Substitution	57
4.4.2	Deletion and Insertion	58
4.4.3	Exon Skipping	58
4.4.4	The Relationship between Mutations and Polymorphisms	59
4.5	Mutation Mechanisms	60
4.5.1	Mutations Caused by UV Light	61
4.5.2	Mutations Caused by Chemical Agents	62

4.5.3	Mutations Caused by Replication Errors	63
4.5.4	DNA Methylation and CpG (CG) Islands	63
4.6	Mechanisms of DNA Repair	65
4.6.1	Base Excision Repair (BER)	66
4.6.2	Nucleotide Excision Repair (NER)	67
4.6.3	DNA Mismatch Repair (MMR)	68
	References	69
5	Transcription and Posttranscriptional Processes	71
5.1	Gene Expression Overview	71
5.2	Gene Transcription and RNA Polymerases	72
5.2.1	RNA Polymerases	73
5.3	DNA Regulatory Elements	75
5.3.1	Response Elements	76
5.4	Structural Motifs of Transcription Factors	77
5.5	Transcription of Protein-Coding Genes	78
5.6	Transcription of RNA Genes	80
5.7	The Role of Transcription Factors and Chromatin Structure in Regulation of Transcription	81
5.7.1	Mechanism of Gene Bookmarking	82
5.7.2	Insulators	83
5.8	RNA Processing	84
5.8.1	Capping Process	85
5.8.2	Polyadenylation	85
5.8.3	Processing of Pre-rRNA and Pre-tRNA	87
5.8.4	RNA Splicing	87
5.8.5	RNA Editing	88
5.9	Reverse Transcription	89
	References	91
6	Synthesis and Posttranslational Modifications of Proteins	93
6.1	Protein Synthesis	93
6.1.1	The Initiation Signal for Protein Synthesis	93
6.1.2	Role of tRNA in Translation	94
6.1.3	Process of Protein Synthesis	96
6.1.3.1	Initiation	96
6.1.3.2	Elongation	97
6.1.3.3	Termination	98
6.1.4	Reading Frame Shift	98
6.1.5	Overlapping Genes	99
6.2	Posttranslational Modifications and Protein Sorting	99
6.2.1	Targeting of Proteins to the Rough Endoplasmic Reticulum	100

6.2.2	Transport of Proteins into and across the Golgi Apparatus	102
6.2.3	Transport of Proteins into Lysosomes	103
6.2.4	Nuclear Transport	103
6.2.4.1	Nuclear Localization Signal and Nuclear Export Signal	103
6.2.4.2	Mechanism of Importins and Exportins Action	104
6.3	Protein Degradation	105
	References	106
7	Cell Division	109
7.1	Cell Cycle	109
7.1.1	CDK and Cyclins	110
7.1.2	Mechanism of Initiation and Termination of the S Phase	111
7.2	Mitosis	111
7.2.1	Initiation of Mitosis	113
7.2.2	Mechanism of Mitosis Termination	113
7.3	Meiosis	114
7.3.1	Independent Chromosome Segregation	114
7.4	DNA Recombination	116
7.4.1	Homologous Recombination	116
7.4.2	Site-specific Recombination	117
7.4.3	Transposition	118
	References	120
8	Cell Signaling Pathways	121
8.1	Signaling through Hydrophobic Molecules	121
8.2	Signaling through Ion Channel Linked Receptors	123
8.3	Signal Transduction via Receptors Associated with G Protein	125
8.3.1	G Protein-Coupled Receptors	125
8.3.2	G Proteins	126
8.3.3	Effectors and Secondary Messengers	128
8.4	Signal Transduction through Enzyme Linked Receptors	130
8.4.1	Receptor Tyrosine Kinases	130
8.4.2	Receptor Serine/Threonine Kinases	132
8.5	Intracellular Signaling Pathways	132
8.5.1	cAMP-dependent Signaling Pathway	132
8.5.2	NF-kappaB Signaling Pathway	133
8.5.3	JAK-STAT Signaling Pathway	134
8.5.4	MAPK Signaling Pathway	135
8.6	Apoptosis	136
	References	138

9	High-Throughput Technologies in Molecular Biology	139
9.1	DNA Sequencing	139
9.1.1	Chain Termination Sequencing	139
9.1.2	Full Genome Sequencing	141
9.2	Analysis of Gene Expression at RNA Level	143
9.2.1	DNA Microarrays	143
9.2.1.1	Two-Color Microarrays	145
9.2.1.2	One-Color Microarrays	146
9.3	Studies of Gene Expression Regulation by Chromatin Immunoprecipitation	147
9.4	Analysis of Protein Expression	149
9.4.1	MALDI Mass Spectrometry	149
	References	153