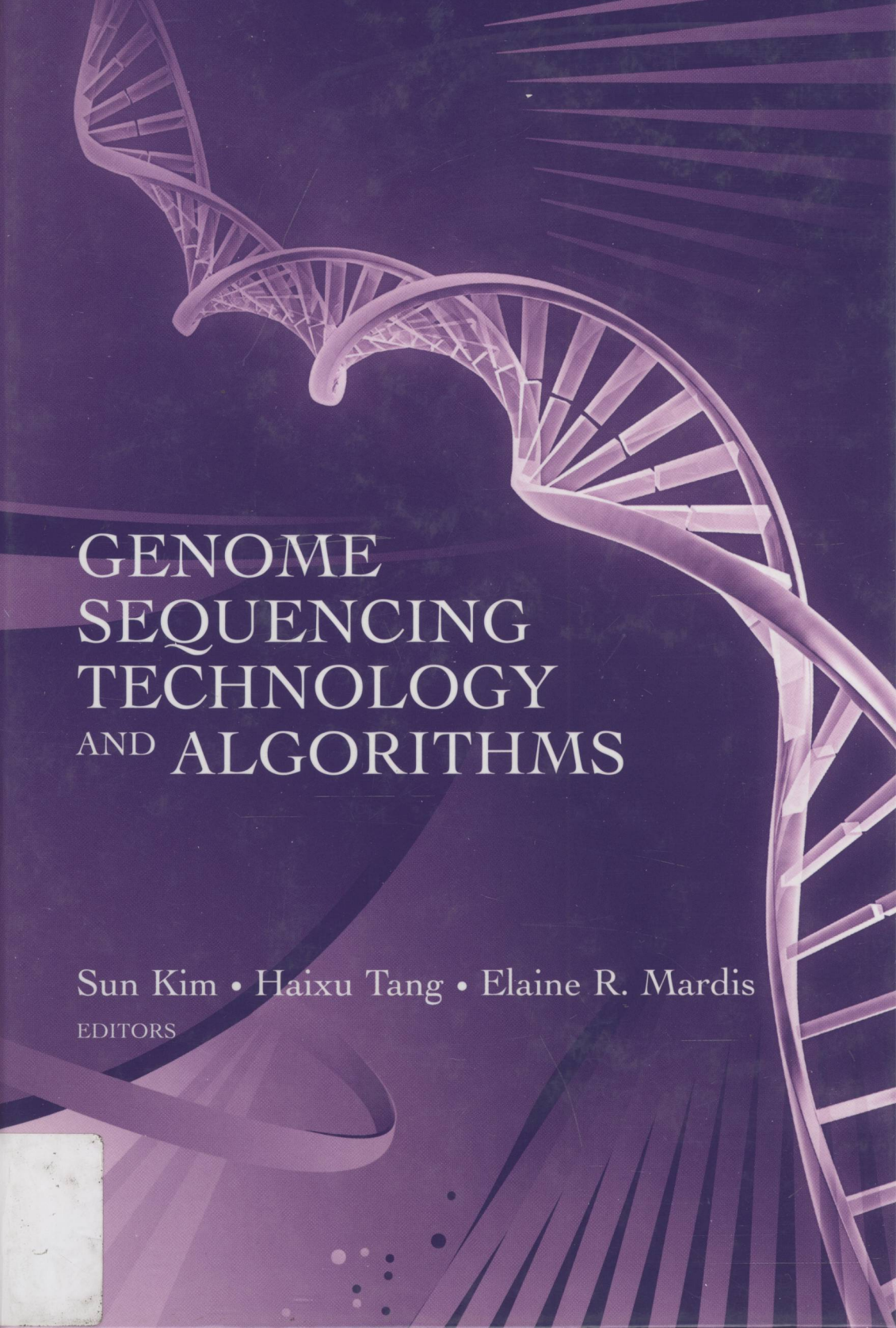


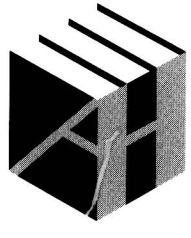
GENOME SEQUENCING TECHNOLOGY AND ALGORITHMS

Sun Kim • Haixu Tang • Elaine R. Mardis
EDITORS

Genome Sequencing Technology and Algorithms

Sun Kim
Haixu Tang
Elaine R. Mardis

Editors



ARTECH
HOUSE
BOSTON | LONDON
artechhouse.com

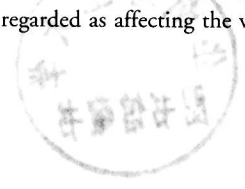
Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the U.S. Library of Congress.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library.

Cover design by Igor Valdman

ISBN 13: 978-1-59693-094-0

685 Canton Street
Norwood, MA 02062



10 9 8 7 6 5 4 3 2 1

Genome Sequencing Technology and Algorithms

For a listing of related Artech House titles,
turn to the back of this book.

List of Contributors

Chapter 1	Elaine R. Mardis, *Washington University*
Chapter 2	Baback Gharizadeh, Roxana Jalili, and Mostafa Ronaghi, *Stanford University*
Chapter 3	David Okou and Michael E. Zwick, *Emory University School of Medicine*
Chapter 4	Jay Shendure, *University of Washington* Gregory J. Porreca and George M. Church, *Harvard Medical School*
Chapter 5	Lewis J. Frey and Joyce A. Mitchell, *University of Utah* Victor Maojo, *Universidad Politecnica de Madrid*
Chapter 6	Sun Kim and Haixu Tang, *Indiana University*
Chapter 7	Sun Kim and Haixu Tang, *Indiana University*
Chapter 8	Jiacheng Chen and Steven Skiena, *Stony Brook University*
Chapter 9	Haixu Tang and Sun Kim, *Indiana University*
Chapter 10	Paola Bonizzoni and Gianluca Della Vedova, *University di Milano-Bicocca* Riccardo Dondi, *University of Bergamo* Jing Li, *Case Western Reserve University*
Chapter 11	Benjamin J. Raphael, *Brown University* Stas Volik, Colin C. Collins, *University of California at San Francisco*
Chapter 12	Curt Balch and Kenneth P. Nephew, *Indiana University* Tim H.-M. Huang, *The Ohio State University*
Chapter 13	Aleksandar Milosavljevic and Cristian Coarfa, *Baylor College of Medicine*

Contents

Part I
The New DNA Sequencing Technology

1

An Overview of New DNA Sequencing Technology

Elaine R. Mardis

1.1 An Overview

1.1.1 Background

The dideoxynucleotide termination DNA sequencing technology invented by Fred Sanger and colleagues, published in 1977, formed the basis for DNA sequencing from its inception through 2004 [1]. Originally based on radioactive labeling, the method was automated by the use of fluorescent labeling coupled with excitation and detection on dedicated instruments, with fragment separation by slab gel [2] and ultimately by capillary gel electrophoresis. A variety of molecular biology, chemistry, and enzymology-based improvements have brought Sanger's approach to its current state of the art. By virtue of economies of scale, high-throughput automation and reaction optimization, large sequencing centers have decreased the cost of a fluorescent Sanger sequencing reaction to around $0.30. However, it is likely that only incremental cost decreases will continue to be achieved for Sanger sequencing in its current manifestation. This fact, coupled with the ever-increasing need for DNA sequencing toward a variety of biomedical (and other) studies, has resulted in a rapid phase of technology development of so-called next generation or massively parallel sequencing technologies, that will revolutionize DNA sequencing as we now know it. Along with this revolution will come a significant and potentially unanticipated impact