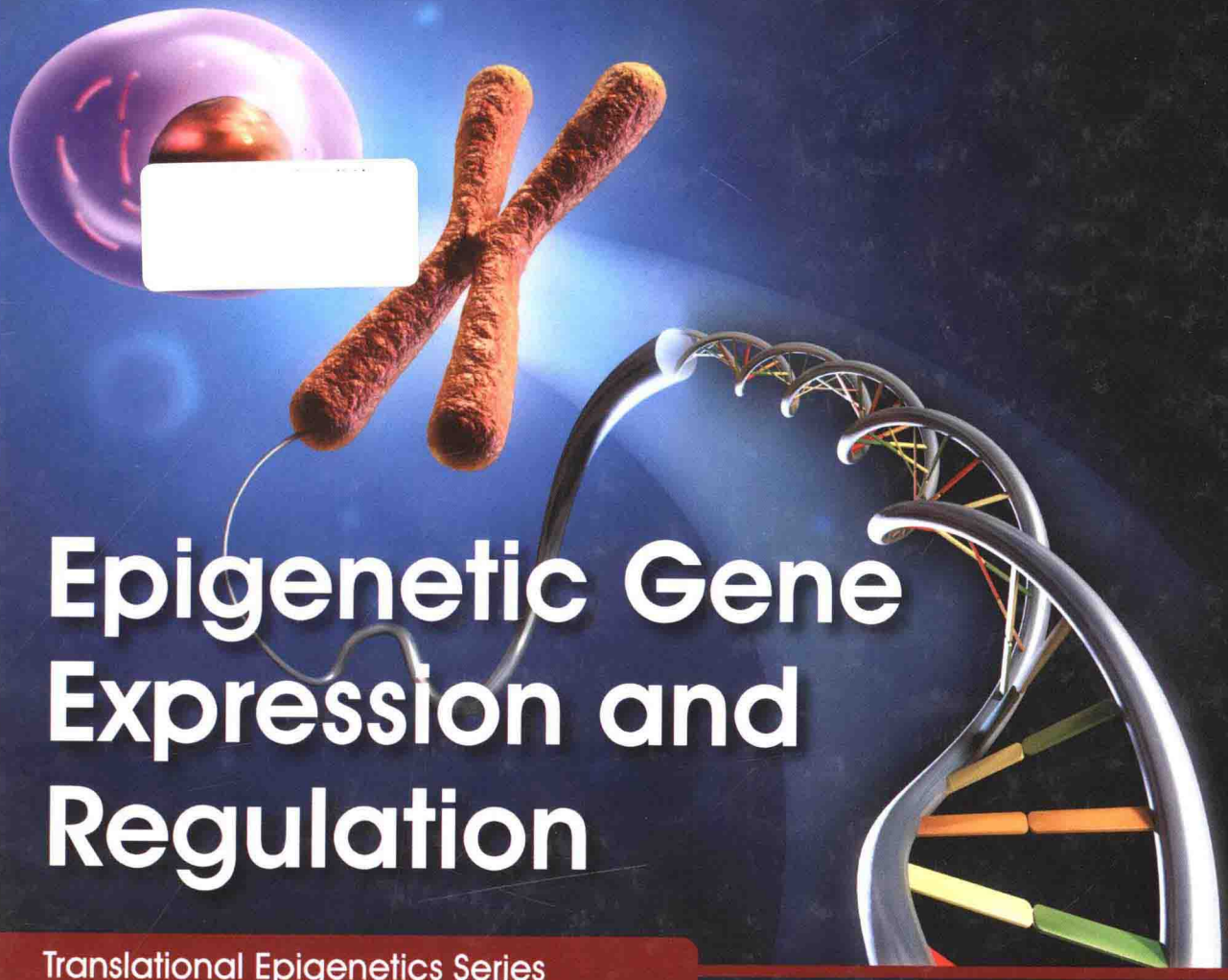

Epigenetic Gene Expression and Regulation

Translational Epigenetics Series

Edited by
Suming Huang
Michael D. Litt
C. Ann Blakey

AP

EPIGENETIC GENE EXPRESSION AND REGULATION

Edited by

SUMING HUANG
Department of Biochemistry and Molecular Biology
University of Florida, College of Medicine
Gainesville, FL, USA

MICHAEL D. LITT
Center for Medical Education
Indiana University-Ball State University
Muncie, IN, USA

C. ANN BLAKEY
Department of Biology
Ball State University
Muncie, IN, USA

Amsterdam • Boston • Heidelberg • London
New York • Oxford • Paris • San Diego
San Francisco • Singapore • Sydney • Tokyo
Academic Press is an imprint of Elsevier

Academic Press is an imprint of Elsevier
125 London Wall, London EC2Y 5AS, UK
525 B Street, Suite 1800, San Diego, CA 92101-4495, USA
225 Wyman Street, Waltham, MA 02451, USA
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

ISBN: 978-0-12-799958-6

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

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

Typeset by TNQ Books and Journals
www.tnq.co.in

Printed and bound in the United States of America

EPIGENETIC GENE EXPRESSION AND REGULATION

Translational Epigenetics Series

Trygve O. Tollefsbol, Series Editor

Transgenerational Epigenetics
Edited by Trygve O. Tollefsbol, 2014

Personalized Epigenetics
Edited by Trygve O. Tollefsbol, 2015

Epigenetic Technological Applications
Edited by Y. George Zheng, 2015

Epigenetic Cancer Therapy
Edited by Steven G. Gray, 2015

DNA Methylation and Complex Human Disease
By Michel Neidhart, 2015

Epigenomics in Health and Disease
Edited by Mario F. Fraga and Agustin F. F Fernández

CONTRIBUTORS

Blake Atwood
Stem Cell Institute, Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, AL, USA

Aissa Benyoucef
The Sprott Center for Stem Cell Research, Regenerative Medicine Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada; Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON, Canada

C. Ann Blakey
Department of Biology, Ball State University, Muncie, IN, USA

Marjorie Brand
The Sprott Center for Stem Cell Research, Regenerative Medicine Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada; Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON, Canada

Jason O. Brant
Department of Biochemistry and Molecular Biology, Center for Epigenetics, Genetics Institute, University of Florida College of Medicine, Gainesville, FL, USA

Jörg Bungert
Department of Biochemistry and Molecular Biology, Center for Epigenetics, Genetics Institute, Powell Gene Therapy Center, College of Medicine, University of Florida, Gainesville, Florida, USA

Yun Chen
Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

Changwang Deng
Department of Biochemistry and Molecular Biology, University of Florida College of Medicine, Gainesville, FL, USA

Nora Engel
Fels Institute for Cancer Research, Temple University School of Medicine, Philadelphia, PA, USA

Alex Xiucheng Fan
Department of Biochemistry and Molecular Biology, Center for Epigenetics, Genetics Institute, Powell Gene Therapy Center, College of Medicine, University of Florida, Gainesville, Florida, USA

Ekaterina Gavrilova
Department of Biochemistry and Molecular Biology, Center for Epigenetics, Genetics Institute, Powell Gene Therapy Center, College of Medicine, University of Florida, Gainesville, Florida, USA

Shrestha Ghosh
Department of Biochemistry, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China

Keith E. Giles
Stem Cell Institute, Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, AL, USA

Suming Huang
Department of Biochemistry and Molecular Biology, University of Florida College of Medicine, Gainesville, FL, USA

Mir A. Hossain
Department of Biochemistry and Molecular Biology, Center for Epigenetics, Genetics Institute, Powell Gene Therapy Center, College of Medicine, University of Florida, Gainesville, Florida, USA

Gangqing Hu
Systems Biology Center, National Heart, Lung, and Blood Institute, NIH, Bethesda, MD, USA

Xin Hu
Edmond H. Fischer Signal Transduction Laboratory, School of Life Sciences, Jilin University, Changchun, China

Wei Jian
Department of Anatomy and Cell Biology, College of Medicine, University of Florida, Gainesville, Florida, USA

Alta Johnson
Department of Anatomy and Cell Biology, College of Medicine, University of Florida, Gainesville, Florida, USA

Priscilla Nga Ieng Lau
Leukemia and Stem Cell Biology Group, Department of Haematological Medicine, King's College London, London, UK

Michael D. Litt
Center for Medical Education, Indiana University-Ball State University, Muncie, IN, USA

Xuehui Li
Department of Anatomy and Cell Biology, College of Medicine, University of Florida, Gainesville, Florida, USA

Yangqiu Li
Institute of Hematology, Jinan University Medical College, Guangzhou, China

Yuanyuan Li
Department of Medicine, Division of Hematology and Oncology, University of Alabama at Birmingham, Birmingham, AL, USA; Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, AL, USA; Nutrition Obesity Research Center, University of Alabama at Birmingham, Birmingham, AL, USA

Xiumei Lin
Department of Hematology, Guangzhou First People's Hospital, Guangzhou, China

Jianrong Lu
Department of Biochemistry and Molecular Biology, College of Medicine, University of Florida, Gainesville, FL, USA

Huacheng Luo
Department of Biochemistry and Molecular Biology, College of Medicine, University of Florida, Gainesville, FL, USA

Christine M. McBride
Department of Biology, The University of Alabama at Birmingham, Birmingham, AL, USA

Benjamin B. Mills
Department of Biology, The University of Alabama at Birmingham, Birmingham, AL, USA

Bhavita Patel
Department of Medicine, University of Florida College of Medicine, Gainesville, FL, USA

Yi Qiu
Department of Anatomy and Cell Biology, College of Medicine, University of Florida, Gainesville, Florida, USA

Félix Recillas-Targa
Instituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, Ciudad de México, México

Nicole C. Riddle
Department of Biology, The University of Alabama at Birmingham, Birmingham, AL, USA

Chi Wai Eric So
Leukemia and Stem Cell Biology Group, Department of Haematological Medicine, King's College London, London, UK

Jared Stees
Department of Biochemistry and Molecular Biology, Center for Epigenetics, Genetics Institute, Powell Gene Therapy Center, College of Medicine, University of Florida, Gainesville, Florida, USA

Ming Tang
Department of Biochemistry and Molecular Biology, College of Medicine, University of Florida, Gainesville, FL, USA

Jessica L. Woolnough
Stem Cell Institute, Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, AL, USA

Bowen Yan
Department of Anatomy and Cell Biology, College of Medicine, University of Florida, Gainesville, Florida, USA

Thomas P. Yang
Department of Biochemistry and Molecular Biology, Center for Epigenetics, Genetics Institute, University of Florida College of Medicine, Gainesville, FL, USA

Yurong Yang
Department of Anatomy and Cell Biology, College of Medicine, University of Florida, Gainesville, Florida, USA

Keji Zhao
Systems Biology Center, National Heart, Lung, and Blood Institute, NIH, Bethesda, MD, USA

Zhizhuang Joe Zhao
Edmond H. Fischer Signal Transduction Laboratory, School of Life Sciences, Jilin University, Changchun, China; Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

Lei Zhou
Department of Molecular Genetics and Microbiology, UF Health Cancer Center, College of Medicine, University of Florida, Gainesville, Florida, USA

Zhongjun Zhou
Department of Biochemistry, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China; Shenzhen Institute of Research and Innovation, The University of Hong Kong, Shenzhen, China

PREFACE

The definition of epigenetics has changed a number of times over the past decades, reflecting our increasingly detailed understanding of regulatory mechanisms in eukaryotes. The chapters of this book effectively point the way to a contemporary definition. Epigenetics, as defined in recent years, concerns itself with the transmission through cell division, and in some cases through the germline, of phenotypic information that is not contained in the DNA sequence itself. The search for biochemical mechanisms that could implement an epigenetic program has, reasonably, focused on reactions that lead to covalent modifications of DNA, and on the identity and biochemistry of proteins and nucleic acids that might bind to DNA tightly enough to remain attached through cell division. That search initially focused on DNA methylation, and more recently on the nucleosome, the octamer histone complex that binds tightly to DNA and packages virtually all of the genome.

DNA methylation provides potentially the most straightforward mechanism for epigenetic transmission of information. As first pointed out over 40 years ago, cytosine methylation at CpG sites can be propagated during replication by an enzyme that recognizes the hemimethylated site and methylates the CpG on the newly synthesized strand. Subsequent studies, notably at imprinted loci, have shown that DNA methylation can affect, directly or indirectly, the binding of transcription factors. The DNA methylation model, in fact, provided the impetus for the present definition of epigenetics, because it appeared to represent the first plausible pathway for transmission through DNA replication of information not encoded in DNA sequence. Since then, however, discovery of the role of chromatin structure in regulation of gene expression, and, during the last twenty years, of the role of noncoding RNAs, has greatly expanded the scope of the epigenetic enterprise.

A large number of articles in this book examine the now extensive literature on the relationship between chromatin structure and gene expression. At the level of individual nucleosomes, the obvious first question is whether nucleosome stability and placement play an important regulatory role. One major research direction has addressed the identity, biochemical consequences, and biological function of histone covalent modifications. The identified distinct modifications include acetylation, mono-, di-, or trimethylation, phosphorylation, ubiquitylation, sumoylation, ADP ribosylation, and more. Each modification targets specific residues on the individual histones of the nucleosome, but the effects, it is now clear, cannot be described in terms of a single code. As with other regulatory mechanisms in eukaryotes, organisms appear to take advantage of the combinatorial versatility afforded by the wide variety of modifications. As described in several chapters,

these modifications are an integral part of mechanisms that control initiation of transcription, transcriptional pausing, transcript elongation, termination, and RNA splicing. Other modifications are involved in DNA replication and DNA damage repair. Some marks are predominantly associated with transcriptionally active or inactive chromatin. Certain combinations of marks tend to appear at promoters, others at enhancers. In pluripotent stem cells, some promoters carry both activating and silencing modifications, in principle making the cells ready for rapid changes in expression, depending on the developmental path they take. Histone modifications also play a critical role in transcriptional elongation. When RNA polymerase II transcribes a gene, it carries with it histone modifying enzymes, signaling the formation of a chromatin structure that discourages inappropriate initiation of transcription from within the gene coding region.

All of these reactions are carried out by an array of enzyme complexes specialized to add or remove the individual modifications at particular sites on the histones. Typically, these complexes carry, in addition to the subunit containing the active site, a variety of other protein cofactors that can confer further target specificity. Notable among these modifying complexes are the Polycomb and Trithorax groups, associated respectively with transcriptional silencing and activation. Other complexes are devoted to the task of removing modifications: enzymes and pathways have been identified not only for deacetylation, but for removal of methyl groups from histone lysine and arginine residues. Modifications are thus dynamically regulated in response to regulatory signals and are an essential part of the mechanism by which genetic information contained in DNA is selectively expressed.

How are these modifications used to transmit that information? Modified histones can, in some cases, recruit transcription factor complexes to promoters and enhancers. Other large families of protein complexes use ATP to move nucleosomes away from important DNA regulatory sites or to position them so that they cover these sites. Some histone modifications can alter the strength of interaction between histones and DNA. Certain modified histones can recruit enzymes involved in DNA methylation, and 5-methylcytosine can recruit histone modifying enzymes, thus coupling the two kinds of potentially epigenetic signals. In every case, the protein complexes involved can vary in details of subunit content and specificity, making it necessary to study each regulatory pathway as a separate problem. This has become especially clear as we have begun to understand the role of noncoding RNAs (ncRNAs) in regulatory processes and especially in epigenetic mechanisms involving chromatin. In particular, ncRNAs help to deliver Polycomb or Trithorax group complexes to specific sites, thus guiding the delivery of silencing or activating marks. Other ncRNAs function as part of regulatory protein complexes, and are essential for their activity.

Although much of this work has focused on epigenetic effects at the level of individual genes or gene families, recent studies have revealed the importance of large scale organization within the nucleus in the control of gene expression. It has long been

known that enhancers can in some instances act over very great distances to activate specific promoters. It has not been clear how the enhancer chooses its target. Chromatin conformation capture technology (3C, 4C, 5C, and Hi-C) has revealed the presence of large scale organization of the genome into loop domains, with interactions within loops favored over those between loops. How such loops are established and maintained through cell division is a question that remains to be explored and may well be an important part of the epigenetic machinery.

One question raised by all these mechanisms is the extent to which they fit the present definition of epigenetics. There is, at least in principle, a way for DNA methylation patterns to be transmitted through cell division, as described above. Patterns of histone modifications could also be preserved: for example, the Polycomb complex, PRC2, which methylates lysine 27 on histone H3, can be recruited specifically to histones carrying that mark, and could then modify adjacent histones newly deposited at the replication fork. Similarly, one could propose that some part of the loop domain organization of the genome is maintained, although it is known that many features of this higher order structure are not preserved during mitosis. Demonstrating the actual use of such mechanisms in vivo has, however, been difficult, particularly for those mechanisms involving chromatin structure. To qualify according to the strict definition, chromatin structure at a gene should be preserved through cell division and should be determinative of the state of activity of that gene in the daughter cells. From one point of view, changes in histone modifications or nucleosome position could be viewed simply as a part of the mechanism of gene expression–a consequence, rather than a cause of the activity state. For example, an active gene could be maintained in that state because relevant transcription factors remain at high concentrations through mitosis. Afterward, these factors might be sufficient to reestablish activity, and appropriate histone modifications could be regenerated.

Because it is difficult to distinguish between cause and effect, there has been resistance to the idea that chromatin structure and histone modifications convey epigenetic information. Given their evident significance for cell function, this is perhaps not such an important objection. As has been suggested before, it may be more appropriate to revise the definition of epigenetics, which in earlier times referred to the developmental processes that led from a single fertilized egg to a complete organism. There can be no doubt about the role of histones, chromatin, and DNA methylation in those processes. At the same time, it should be clear that mechanisms do exist for inheritance of information that is not carried in the DNA sequence. Position effect variegation in *Drosophila* provided the first evidence for a change in phenotype that was connected with location of a gene within a chromosome, rather than with changes in the gene itself. A clear example exists even in *Saccharomyces*, where genes located near telomeres can be maintained in a silent state through many generations, then switch to an active state, which is likewise stable through many cell divisions. In neither case is there any change in the DNA sequence itself; changes in chromatin structure provide the necessary signals.

Although other examples of transmission through cell division exist, epigenetic inheritance through the germline has been more difficult to demonstrate. A clear example is provided by the agouti mouse, in which coat color, determined by DNA methylation patterns of a retrotransposon inserted near the *agouti* gene, is transmitted to offspring. Other examples exist in plants. There is also evidence for multigenerational epigenetic transmission of phenotypes in flies, mice, and humans. It should be noted that the fidelity of transmission is not as high as for genetic inheritance. However to the extent that these phenotypic changes help to stabilize certain mutations, they may contribute to more permanent changes marked directly in the genome. Because DNA methylation and probably chromatin chemistry can be influenced by environment (e.g., diet), epigenetically controlled phenotypes can reflect environmental signals. The old arguments about heredity vs. environment can now be seen in a new light. But whatever the contribution of these epigenetic signals to inherited phenotype, their importance transcends definitions. They are the manifestations of the array of biochemical pathways that modulate all DNA function in eukaryotes and make possible the enormous variety of cell behaviors required for the success of multicellular organisms.

Gary Felsenfeld
Laboratory of Molecular Biology
National Institute of Diabetes and Digestive and Kidney Diseases
National Institutes of Health, Bethesda, MD, USA

ACKNOWLEDGMENT

This work was supported by the intramural research program of the NIH, National Institute of Diabetes and Digestive and Kidney Diseases.

CONTENTS