

**SECOND EDITION** 

# Tene Control

**David S. Latchman** 

# **Gene Control**

#### **SECOND EDITION**

## David S. Latchman





Vice President: Denise Schanck Senior Editor: Elizabeth Owen Editorial Assistant: Deepa Divakaran

Production Editors: Ioana Moldovan and Deepa Divakaran

Illustrator: Oxford Designers & Illustrators

Layout: Techset Composition Cover Designer: Susan Schmidler

Copyeditor: Sally Livitt Proofreader: Susan Wood

©2015 by Garland Science, Taylor & Francis Group, LLC

Front cover shows the structure of the human Ago2 protein component of the RNA-induced silencing complex (RISC). Courtesy of Ian J MacRae, The Scripps Research Institute.

This book contains information obtained from authentic and highly regarded sources. Every effort has been made to trace copyright holders and to obtain their permission for the use of copyright material. Reprinted material is quoted with permission, and sources are indicated. A wide variety of references are listed. Reasonable efforts have been made to publish reliable data and information, but the author and the publisher cannot assume responsibility for the validity of all materials or for the consequences of their use. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or information storage and retrieval systems—without permission of the copyright holder.

ISBN 9780815345039

#### Library of Congress Cataloging-in-Publication Data

Latchman, David S., author.
Gene control / David S Latchman.
p.; cm.
ISBN 978-0-8153-4503-9 (alk. paper)
I. Title.
[DNLM: 1. Gene Expression Regulation--physiology. 2. Transcription, Genetic--physiology. QU 475]
QH450
572.8'65--dc23

2014044381

Published by Garland Science, Taylor & Francis Group, LLC, an informa business, 711 Third Avenue, New York, NY, 10017, USA, and 3 Park Square, Milton Park, Abingdon, OX14 4RN, UK.

Printed in the United States of America

15 14 13 12 11 10 9 8 7 6 5 4 3 2 1



Visit our web site at http://www.garlandscience.com

# **Gene Control**

**SECOND EDITION** 

To my mother and in memory of my father

#### **Preface**

It is now 25 years since the first edition of my book *Gene Regulation: A Eukaryotic Perspective* was published. Subsequently, four further editions appeared reflecting the enormous advances in the field over the ensuing years. By 2010, these advances allowed both a mechanistic understanding of the processes involved in gene regulation and a detailed analysis of the manner in which they operate in specific biological systems.

The work was therefore significantly expanded and revised to reflect these changes and given a new title, *Gene Control*. In the last few years, the field has developed further, necessitating this second and revised edition of *Gene Control*.

In particular, our understanding of two major themes in gene regulation has advanced considerably in the last few years and this is reflected in this new edition. The first of these is the regulation of chromatin structure and its role in gene regulation. Although a chapter was devoted to this topic in the first edition of Gene Regulation, its role was much less well understood at that time than, for example, that of DNA-binding transcription factors. However, over the last 25 years this topic has assumed a central position in our understanding of gene regulation. It is now clear, for example, that as well as the genomic DNA sequence, there exists an epigenome ('epi' is Greek for 'on') involving modifications of the DNA and of the proteins, such as histones, associated with it. Moreover, this epigenome varies between different tissues and cell types in contrast to the relative invariance of the genomic DNA sequence. This allows the epigenome to play a key role in the regulation of gene expression. To reflect this, all the sections dealing with this topic have been extensively revised and updated, while the title of Chapter 3 has been changed to 'The Epigenome: Role of Chromatin Structure in Gene Control.' Moreover, additional sections have been added on these modifications as a target of cellular signaling pathways (Section 8.5) and on the role of the epigenome in cancer (Section 11.4) and other human diseases (Section 12.2).

The second major area to have advanced hugely in recent years is that of regulatory RNAs. At the time the first edition of *Gene Regulation* was published, it was believed that the regulation of gene expression was produced exclusively by regulatory proteins and the possibility of regulatory RNAs had not been put forward. It is now clear, however, that such regulatory RNAs are as important in the regulation of gene expression as regulatory proteins. Moreover, although small regulatory RNAs (20–30 bases in length) were discussed in the first edition of *Gene Control*, further studies have now indicated that longer regulatory RNAs (200 bases or more in length) are also of critical importance in the regulation of gene expression. Accordingly, the sections of this book dealing with this area have been updated and expanded to reflect the key role of all types of regulatory RNAs and to give equal emphasis to the different classes of such RNAs. Moreover, additional sections have been added on the role of regulatory RNAs in cancer (Section 11.6) and other human diseases (Section 12.4).

As with the first edition, the second edition of *Gene Control* is organized into two parts. The first part provides a detailed mechanistic analysis of the processes involved in controlling gene expression. After an introductory chapter, three pairs of chapters deal with the fundamental processes involved in gene regulation. In each pair of chapters, the first chapter deals with the basic process itself and the second deals with the manner in which it is involved in regulating gene expression. Thus, Chapters 2 and 3 deal with chromatin structure and its role in gene regulation, Chapters 4 and 5 deal with the process of transcription itself and the manner in which it is

regulated, while Chapters 6 and 7 deal with post-transcriptional processes and their role in the regulation of gene expression.

In contrast, the second part of the book deals with specific biological processes and the role played by gene control in their regulation. Thus, Chapter 8 deals with cellular signaling processes, Chapter 9 deals with the regulation of gene expression in development and Chapter 10 discusses the key role played by gene regulatory processes in the specification of individual differentiated cell types. Finally, Chapters 11 and 12 deal with the alterations in gene expression that can cause specific human diseases. Thus, Chapter 11 discusses the role of gene regulation in cancer while Chapter 12 deals with gene regulation in inherited and infectious diseases of humans, as well as discussing the manner in which advances in our understanding of gene regulatory processes may lead to improved therapies for human diseases.

Overall therefore it is hoped that the significant revision and updating of this book will allow it to continue to be of value to students, scientists, and clinicians interested in the topic of gene control, which is of such vital importance in both normal development and the proper functioning of the adult organism as well as playing a critical role in the development of different human diseases.

Finally, I should like to thank the staff of Garland Science who have produced this book so efficiently, particularly Elizabeth Owen who originally suggested the restructuring/expansion of this book, as well as being a valuable source of advice throughout the preparation of both editions. As always, I am indebted to Maruschka Malacos who has coped most efficiently with the need to introduce considerable new material as well as to rearrange and modify existing material in the preparation of the new edition.

David S. Latchman CBE

#### **Acknowledgments**

In writing this book I have benefited greatly from the advice of many geneticists, cell biologists, and biochemists. I would like to thank the following for their suggestions in preparing this edition.

John Acord, London South Bank University; Don Beitz, Iowa State University; Thomas Chen, University of Connecticut; Ferninando Chiaradonna, University of Milano-Bicocca; Tamara Davis, Bryn Mawr College; Paul Dyson, Swansea University; Richard Gomer, Texas A&M University; Barbara Guinn, University of Bedfordshire; Dean Jackson, University of Manchester; Catherine Kidner, University of Edinburgh; Darrell J Killian, Colorado College; Mairi MacKay, Thompson Rivers University; Phil Mitchell, University of Sheffield; John Rebers, Northern Michigan University; Susanna Cirera Salicio, University of Copenhagen; Paul Townsend, University of Southampton; Roland Valcke, University of Hasselt; Ann Vernallis, Aston University; Anders Virtanen, Uppsala University; Kenneth White, London Metropolitan University; Ian Wood, University of Leeds; Stephanie Wright, University of Leeds.

I would also like to thank all those colleagues who have given permission for material from their papers to be reproduced in this book and have provided images suitable for reproduction.

#### **Online Resources**

Accessible from www.garlandscience.com, the Student and Instructor Resource Websites provide learning and teaching tools created for Gene Control. The Student Resource Site is open to everyone, and users have the option to register in order to use book-marking and note-taking tools. The Instructor Resource Site requires registration and is available only to qualified instructors. To access the Instructor Resource Site, please contact your local sales representative or email science@garland.com. Below is an overview of the resources available for this book. On the Website, the resources may be browsed by individual chapters and there is a search engine. You can also access the resources available for other Garland Science titles.

#### For instructors

#### **Figures**

The images in this book are available in two convenient formats: PowerPoint® and JPEG. They have been optimized for display on a computer.

PowerPoint is a registered trademark of Microsoft Corporation in the United States and/or other countries.

## **Contents in Brief**

1	Levels of Gene Control	
2	Structure of Chromatin	35
3	The Epigenome: Role of Chromatin Structure in Gene Control	67
4	The Process of Transcription	115
5	Transcription Factors and Transcriptional Control	159
6	Post-transcriptional Processes	205
7	Post-transcriptional Regulation	233
8	Gene Control and Cellular Signaling Pathways	283
9	Gene Control In Embryonic Development	317
10	Control of Cell Type-Specific Gene Expression	355
11	Gene Regulation and Cancer	389
12	Gene Regulation and Human Disease	425
13	Conclusions and Future Prospects	457

## **Contents in Detail**

Preface		vii	Transcriptional control can operate	
Acknowledgments		viii	at the level of chromatin structure and at the level of production of the	
Online resources		ix	primary RNA transcript	20
1 L	evels of Gene Control	1	1.5 Regulatory RNAs and the regulation of gene expression	21
	oduction	1	RNA molecules of different lengths are	- 2.1
1.1	The protein content of different cell types is different	1	involved in the control of gene expression MicroRNAs are processed from a	1 21
	Specific methods can be used to study the expression of individual proteins in tissues and cells	1	single-stranded precursor which folds to form a double-stranded hairpin loop	21
	General methods can be used for studying the overall protein composition of tissues and cells	2	Many small interfering RNAs are processed from a double-stranded precursor	24
1.2	The mRNA content of different cell types is different	5	Piwi-interacting RNAs are processed in a Dicer-independent manner	26
	Specific methods can be used to study the expression of individual mRNAs in different tissues and cells	5	Small RNAs can move from cell to cell and from one part of an organism to another	28
	General methods can be used to study the overall population of mRNAs		Long noncoding RNAs function by a variety of mechanisms	28
	expressed in different tissues and cells	7	Conclusions	32
1.3	The DNA content of different cell types is generally the same	9	Key concepts Further reading	33 33
	Specific methods can be used to study individual genes in different tissues and cells	9	2 Structure of Chromatin Introduction	<b>35</b>
	General methods can be used to study the total DNA in different tissues and cells	10	Regulation of transcription in eukaryotes is much more complex than in prokaryotes	
	Exceptional cases do exist in which changes to the DNA occur in specific tissues or cell types	13	2.1 Commitment to the differentiated state and its stability	36
1.4	Transcriptional or post- transcriptional control?	15	Cells can remain committed to a particular differentiated state even in the absence of its phenotypic	36
	Pulse labeling studies directly demonstrate transcriptional control	16	characteristics  Cells can become committed to a	30
	Nuclear run-on assays allow transcriptional control to be demonstrated for a wide range of genes	17	particular differentiated state before actual phenotypic differentiation  2.2 The nucleosome	37 <b>39</b>
		1 /		3)
	Polytene chromosomes provide further		The nucleosome is the basic unit of	

xiii

3.6	Other situations in which chromatin structure is regulated	103		Polymerase pausing is widespread and represents a potential control	
	In female mammals one of the two X chromosomes is inactivated	103		point for the regulation of transcription	130
	The active and inactive X chromosomes have a different chromatin structure	104		Termination of transcription occurs downstream of the polyadenylation signal	131
	The XIST regulatory RNA is		4.3	The gene promoter	134
	specifically transcribed on the inactive X chromosome	104		The 70 kDa heat-shock protein gene contains a typical promoter for RNA polymerase II	135
	Genomic imprinting involves the specific inactivation of either the maternally or paternally inherited copy of specific genes	106		The <i>hsp70</i> gene promoter contains several DNA sequence motifs which are found in a variety of other gene promoters	135
	Imprinting involves changes in chromatin structure	107		The heat-shock element is found	
Con	clusions	110		only in heat-inducible genes	136
Кеу	concepts	111		Other response elements are found in the promoters of genes with	
Fur	ther reading	112		different patterns of expression	138
4 T	the Process of Transcription	115		The proteins binding to short DNA sequence elements can be characterized by a variety	
Intr	oduction	115		of techniques	140
4.1	Transcription by RNA polymerases	115		The use of high-throughput methods such ChIP-chip and ChIP-seq allows	
	Transcription by RNA polymerase I is relatively simple	116		the analysis of DNA-binding sites and epigenetic modifications on a whole	1 4 4
	Transcription by RNA polymerase III is more complex than for RNA polymerase I	117		genome basis  Promoter regulatory elements act by binding factors which either affect	144
	Transcription by RNA polymerase II is much more complex than			chromatin structure and/or influence transcription directly	145
	transcription by RNA polymerases I and III	119	4.4	Enhancers and silencers	146
	Plants contain two additional RNA polymerases which are related to RNA polymerase II and play a key role			Enhancers are regulatory sequences that act at a distance to increase gene expression	146
	in gene silencing by small RNAs	122		Many enhancers have cell type- or tissue-specific activity	148
	Transcription by RNA polymerases I, II, and III has a number of common features	123		Enhancers influence transcription both by recruiting DNA-binding transcription factors and by altering	
	Transcription takes place in defined regions of the nucleus	127		chromatin structure The "at a distance" action of	150
4.2	Transcriptional elongation and termination	129		enhancers involves both DNA-binding transcription factors and alteration of chromatin structure	152
	Transcriptional elongation requires further phosphorylation of RNA polymerase II	129		Many enhancers are transcribed to produce long noncoding RNAs	153

	Silencers can act at a distance to inhibit gene expression	154	Several different classes of activation domain exist	183
Cor	nclusions	155	Activators can interact with the TAFs	
Key concepts		156	within the TFIID complex	184
Fur	ther reading	157	Activators can interact with co-activators	s 186
C 1	Françorintian Factors and		Activators can interact with modulators of chromatin structure	187
	ranscription Factors and ranscriptional Control	159	Activators have a multitude of targets	188
	oduction	159	5.3 Repression of transcription	190
	DNA binding by transcription factors	162	Repressors can act indirectly by inhibiting the positive effect of activators	191
	The helix-turn-helix motif is found in a number of transcription factors which regulate gene expression during embryonic development	162	Repressors can act directly by inhibiting the assembly or activity of the basal transcriptional complex	191
	The helix-turn-helix domain found in homeodomain proteins is a DNA-	102	5.4 Regulation at transcriptional elongation	194
	binding domain In the POU domain transcription	164	Regulation of transcription can occur at the elongation stage, as well as at initiation	194
	factors, the homeodomain forms part of a larger DNA-binding motif	166	Factors which regulate transcriptional elongation target the C-terminal	
	The two-cysteine-two-histidine zinc finger is found in multiple copies in many transcription factors	168	domain of the RPB1 subunit of RNA polymerase II	197
	The nuclear receptors contain two copies of a multi-cysteine zinc finger distinct from the two-cysteine-two-histidine zinc finger	100	5.5 Regulation of transcription by RNA polymerases I and III	199
		170	Transcription by RNA polymerases I and III can be regulated by altering the expression or activity of components	
	The leucine zipper is a dimerization domain which allows DNA binding	175	of their basal transcriptional complexes Regulation of transcription by RNA	199
	In some transcription factors, the basic DNA-binding domain is found associated with a helix-loop-helix		polymerase III can involve specific transcription factors binding to RNA as well as to DNA	200
	dimerization domain  Dimerization between factors provides an additional level of regulation	176	Transcription by RNA polymerases I and III can be regulated by alterations in chromatin structure	201
		177	Conclusions	202
	Other domains can also mediate DNA binding			202 203
	The DNA sequence bound by a particular transcription factor and the effect of such binding is affected by the presence or absence of other regulatory factors	180		205 205
5.2	Activation of transcription	181		205
	Activation domains can be identified by "domain-swap" experiments	181	The capping process modifies the	205

	The cap enhances translation of the mRNA by the ribosome	206	Key con	ncepts reading	231 231
6.2	Polyadenylation	208	Further	reaunig	231
	The polyadenylation process modifies the 3' end of the RNA transcript	208		t-transcriptional ulation	233
	Polyadenylation enhances the stability of the mRNA	208	Introdu		233
6.3	RNA splicing	210	7.1 Alt	ernative RNA splicing	234
	RNA splicing removes intervening sequences and joins exons together	210	Alte	A splicing can be regulated ernative splicing represents a	234
	Specific RNAs and proteins catalyze the process of RNA splicing within		sup	jor regulatory process which plements transcriptional control	234
6.4	the spliceosome  Coupling of transcription and RNA processing within the nucleus	210 <b>215</b>	spe	ernative RNA splicing involves cific splicing factors that promote nhibit the use of specific splice sites	235
	Transcriptional initiation and elongation are coupled to post-transcriptional processes	215	hav	tors regulating alternative splicing been identified by genetic and chemical methods	237
	Post-transcriptional processes can interact with one another	218	alte	e processes of transcription and ernative splicing interact with one other	241
6.5	RNA transport	218	Alte	ernative RNA splicing is a	
	RNA transport is coupled to other post-transcriptional processes	218	ver	y widely used method of plementing transcriptional control	243
6.6	Translation	220	7.2 Reg	gulation of polyadenylation	245
	Translation of the mRNA takes place on cytoplasmic ribosomes	220		h the amount and the site of yadenylation can be regulated	245
	Translational initiation involves initiation factors binding to the cap	222		ernative polyadenylation can lead Alternative splicing	246
	Translation involves base pairing of		7.3 RN	A editing	249
	triplet codons in the mRNA with tRNA anticodons	222		ne cases of RNA editing involve nange from a C to a U residue	249
	Translational elongation involves tRNAs located in two distinct sites within the ribosome	225		ny cases of RNA editing involve nange from A to an I residue	250
	Translational termination occurs		7.4 Reg	gulation of RNA transport	252
	at specific stop codons	226	Spe	ecific proteins can regulate the	
6.7	RNA degradation	228		nsport of individual mRNAs from	252
	RNA degradation occurs in both the nucleus and the cytoplasm	228	RNA	cleus to cytoplasm  A transport processes can also	252
	RNA degradation in the cytoplasm involves prior deadenylation and	222	mR	ulate the location of individual  NAs within the cytoplasm	254
	decapping of the mRNA Abnormal mRNAs are degraded in	229		A localization can be linked to er stages of gene expression	255
	the cytoplasm by a nonsense-mediated RNA decay process	229		gulation of RNA stability	256
Con	clusions	231		ne regulation can involve erations in RNA stability	256

XVII

8.3	Regulation of transcription factor activity by other post-translational modifications	298	Conclusions Key concepts	311 314
8.4	Regulation of transcription factor activity by signals which		Further reading	314
	regulate precursor processing	301	9 Gene Control In Embryonic Development	317
	Transcription factors can be activated by cleavage of a precursor which contains an inhibitory region	301	Introduction	317
	Transcription factors can be activated	501	Regulation of mRNA translation occurs following fertilization	317
	by cleavage of a membrane-bound precursor	301	Transcriptional control processes activate the embryonic genome	318
	Cleavage of a transcription factor can convert it from an activator to a repressor	302	The Oct4 and Cdx2 transcription factors regulate the differentiation of inner cell mass and trophectoderm cells	320
	Cleavage of a lipid link can be used to activate a transcription factor	303	9.1 Regulation of gene expression in pluripotent embryonic stem cells	321
8.5.	Regulation of histone modification and chromatin structure by cellular signaling pathways	304	Embryonic stem cells can differentiate into a wide variety of cell types	321
	Cellular signaling pathways can induce the post-translational modification of histones	304	Several transcription factors which are expressed specifically in ES cells can together reprogram differentiated cells to an ES cell-like phenotype	322
	Histone modifications link intermediary metabolism and chromatin structure/gene expression	304	ES cell-specific transcription factors can activate or repress the expression of their target genes	324
8.6	Regulation of post-transcriptional processes by cellular signaling pathways	305	ES cell-specific transcription factors regulate genes encoding chromatin-modifying enzymes and miRNAs	326
	The PI3-kinase/Akt system plays a key role in regulating gene expression in response to growth factors or insulin	305	ES cells have an unusual pattern of histone methylation	328
	Akt regulates RNA splicing by phosphorylating splicing factors	306	The polycomb complex regulates histone methylation in ES cells	329
	Akt can differentially regulate microRNA activity by phosphorylating	307	Polycomb protein complexes regulate the expression of microRNA genes in ES cells	331
	Akt regulates mRNA translation via	307	Chromatin structure in ES cells is regulated by multiple effects on histones	332
	the TOR kinase which phosphorylates proteins involved in translation  Akt/TOR can also stimulate mRNA	307	ES cells show very specific patterns of chromosome looping which bring together transcriptionally active loci	332
	translation by enhancing the transcription of genes encoding RNAs and proteins involved in protein synthesis	on 308	The regulation of chromatin structure in ES cells plays a critical role in maintaining their undifferentiated natur	L
	A variety of kinases inhibit translation by phosphorylating eIF2	309	and in their ability to differentiate into a variety of cell types	333
	Individual kinases can produce multi-level regulation of gene expression	310	Gene expression in ES cells is regulated at the post-transcriptional level as well as at transcription	333

			CONTENTS	XIX
9.2	Role of gene regulation in the development of <i>Drosophila</i> melanogaster	337	Key concepts Further reading	352 352
	A gradient in expression of the Bicoid transcription factor plays a key role in defining the anterior–posterior axis in the early <i>Drosophila embryo</i>	337	10 Control of Cell Type-Specific Gene Expression Introduction	355 355
	Bicoid activates a cascade of genes encoding other transcription factors producing a segmented pattern of		10.1 Regulation of gene expression in skeletal muscle cells	357
	Eve gene expression	338	The MyoD protein can induce muscle cell differentiation	357
	The Bicoid system involves both transcriptional and post-transcriptional regulation	339	MyoD is a basic helix-loop-helix transcription factor which is able to regulate gene expression	358
	Homeodomain transcription factors specify segment identity in the <i>Drosophila embryo</i>	340	MyoD is regulated by controlling both its synthesis and its activity	360
	Protein–protein interactions control the effect of homeodomain-containing transcription factors on gene expression		Other muscle-specific transcription factors can induce muscle cell differentiation	361
9.3	Role of homeodomain factors in mammalian development	341	MEF2 proteins are downstream regulators of muscle-cell specific gene transcription	363
	Homeodomain transcription factors are also found in mammals	341	10.2 Regulation of gene expression in neuronal cells	366
	Mammalian <i>Hox</i> genes are expressed in specific regions of the developing embryo	342	Basic helix-loop-helix transcription factors are also involved in neuronal differentiation	366
	Transcription of individual <i>Hox</i> genes is regulated by gene-specific regulatory regions	342	The REST transcription factor represses the expression of neuronal genes	369
	Hox gene transcription is also dependent on the position of the gene	244	Neuronal cells express specific alternative splicing factors	371
	changes in chromatin structure play a	344	Translational control plays a key role in synaptic plasticity in neuronal cells	373
	key role in activating the expression of Hox genes and in their spatial/ temporal pattern of gene expression	344	miRNAs play a key role in the regulation of neuronal gene expression	375
	Hox gene expression is also regulated at the post-transcriptional level by specific miRNAs		10.3 Regulation of yeast mating type	377
		345	Yeast cells can be ${f a}$ or ${f lpha}$ in mating type	377
	Differential regulation of different Hox genes by Sonic Hedgehog controls	J40	Mating-type switching is controlled by regulating the transcription of the <i>HO</i> gene	378
	the differentiation of cells in the neural tube  Regulation of <i>Hox</i> gene expression by	346	The SBF transcription factor activates HO transcription only in the G1 phase of the cell cycle	378
	Sonic Hedgehog is also involved in limb formation	348	The Ash-1 transcription factor represses HO transcription in	_, 0
Con	clusions	349	daughter cells	379