THE EUKARYOTE GENOME IN DEVELOPMENT AND EVOLUTION

B. JOHN & G. MIKLOS

THE EUKARYOTE GENOME IN DEVELOPMENT AND EVOLUTION

Bernard John and George L. Gabor Miklos

Research School of Biological Sciences,
The Australian National University, Canberra

London
ALLEN & UNWIN
Boston Sydney Wellington

Allen & Unwin, the academic imprint of

Unwin Hyman Ltd

PO Box 18, Park Lane, Hemel Hempstead, Herts HP2 4TE, UK
40 Museum Street, London WC1A 1LU, UK
37/39 Queen Elizabeth Street, London SE1 2QB, UK

Allen & Unwin Inc.,
8 Winchester Place, Winchester, Mass. 01890, USA

Allen & Unwin (Australia) Ltd,
8 Napier Street, North Sydney, NSW 2060, Australia

Allen & Unwin (New Zealand) Ltd in association with the Port Nicholson Press Ltd,
60 Cambridge Terrace, Wellington, New Zealand

First published in 1988

British Library Cataloguing in Publication Data

John B.
The eukaryote genome in development and evolution
1. Genetics. 2. Eukaryotic cells
I. Title II. Miklos, George L. Gabor
574.87'322 QH430

ISBN 0-04-575032-7
ISBN 0-04-575033-5 Pbk

Library of Congress Cataloging-in-Publication Data

(applied for)

Typeset in 10 on 11 point Bembo by Oxford Print Associates Ltd
and printed in Great Britain by
Biddles, Guildford, Surrey

Preface

'The mind unlearns with difficulty what has long been impressed upon it.'

Seneca

Reductionism, is, without question, the most successful analytical approach available to the experimental scientist. With the advent of techniques for cloning and sequencing DNA, and the development of a variety of molecular probes for localizing macromolecules in cells and tissues, the biologist now has available the most powerful reductionist tools ever invented. The application of these new technologies has led to a veritable explosion of facts regarding the types and organization of nucleotide sequences present in the genomes of eukaryotes. These data offer a level of precision and predictability which is unparalleled in biology.

Recombinant DNA techniques were initially developed to gather information about the structure and organization of the DNA sequences within a genome. The power and potential of these techniques, however, extend far beyond simple data collection of this kind. In an attempt to use the new technology as a basis for analyzing development and evolution, attention was first focused on the topic of gene regulation, an approach that had proven so successful in prokaryotes. It is now clear that this has not been an adequate approach. Lewin (1984) has quoted Brenner as stating 'at the beginning it was said that the answer to the understanding of development was going to come from a knowledge of the molecular mechanisms of gene control. I doubt whether anyone believes this any more. The molecular mechanisms look boringly simple, and they don't tell us what we want to know.' Nor is this surprising. The role of genes in specifying the primary structure of protein molecules is usually far removed from the developmental end products to which those molecules contribute. Consequently, the solutions to developmental problems are much more likely to be found at the cellular and intercellular levels than at the level of transcription.

Although the approach was too limited, the philosophy was right. While we appreciate that there are aspects of development that transcend molecular biology, we believe it will not be possible to elucidate these aspects without an understanding of the molecular mechanisms upon which they depend. Thus while cell–cell interactions are a fundamental feature of all developmental programmes, the most direct approach to resolving this form of interaction must come from an understanding of gene products which, like the neural cell adhesion molecules, a series of membrane associated glycoproteins which control cell surface properties (Edelman 1985, reviewed in Rutishauser & Goridis 1986).

As far as evolutionary biology is concerned, Lewontin (1982) has argued that the deepest questions of evolution 'will never be resolved by molecular biology alone.' There is certainly no doubt that some occurrences of the deepest significance to evolution belong to this category. For example, virtually all the species of plants and animals that have existed on Earth are now extinct, so that extinction must rank as the most common evolutionary event. Major extinction events over the past 250 million years show a statistically significant periodicity (Raup & Sepkoski 1984). Although the precise causes of this periodicity are not known, there are grounds for arguing that the forcing agent was environmental rather than biological. If the impact of the physical environment did indeed result in non-random and short-lived mass extinctions, then the direction and type of evolutionary change may well have been uniquely and irrevocably altered, but molecular biology can have nothing to say about such events.

However, in a strict sense, Lewontin's viewpoint is overly pessimistic since the deepest evolutionary questions which are resolvable will certainly require the aid of molecular biology. Indeed, there is no doubt that the findings of molecular biology impinge far more directly on the major problems of both development and evolution than any other approach to these problems. In the chapters that follow we shall try to show how, and why, they do so.

Specifically, we attempt to assess and interpret the facts concerning genome structure in terms of fundamental problems of genome function in relation to development and evolution. Our main aim is to evaluate the impact of the modern molecular work on long-standing biological problems.

B. John & G. L. G. Miklos
Canberra

Acknowledgements

Joan Madden, Marilyn Miklos, Val Rawlings, Jan Richie and Sylvia Stephens have typed and retyped the numerous drafts of this book and we thank them all for their contribution. Gary Brown prepared all the illustrations and our debt to him is obvious. Arnold Bendich, John Campbell, Mel Green, Ian Young and Brian Gunning read earlier drafts of this book while MasaToshi Yamamoto, John Langridge, Gabriel Dover and Ken Campbell drew our attention to relevant genetic, evolutionary and palaeontological works. Michael Rosbash and Steven Henikoff also provided helpful comments. We are grateful to all of them for their advice and comments. We also acknowledge the support of our respective wives, Ena and Marilyn, without whose perseverance this book would never have been written. Finally, we thank Linda Antoniw for her meticulous editorial work.

We are grateful to the following individuals and organizations who have kindly given permission for the reproduction of copyright material (figure numbers in parentheses):

IRL Press (1.5); Figure 1.6 reproduced by permission of Yelton & Scharff and *American Scientist*; Figure 1.9 reproduced by permission from *Nature*, Vol. 284, p. 215, Copyright © 1980 Macmillan Journals Limited; Figure 2.1 reproduced from *From egg to embryo* by permission of J. M. W. Slack and Cambridge University Press; Figure 2.2 reproduced from *Cell Interactions in Differentiation* by permission of Academic Press; Figure 2.4 reproduced from *Results and Problems in Cell Differentiation* **5**, by permission of Springer Verlag; Figure 2.5 reproduced by permission from *Nature*, Vol. 265, pp. 211–16, Copyright © 1977 Macmillan Journals Limited; Figure 2.7 reproduced from *37th Symp. Soc. Dev. Biol.* by permission of Academic Press; Figure 2.8 reproduced by permission from *Nature*, Vol. 287, pp. 795–801, Copyright © 1980 Macmillan Journals Limited; Figure 2.10 reproduced by permission from *American Zoologist*; Figure 2.12 reproduced by permission from Bender *et al.*, *Science*, Vol. 221, pp. 23–9, July 1983, Copyright 1983 the American Association for the Advancement of Science; Figure 2.13 reproduced by permission from *Nature*, Vol. 310, pp. 25–31, Copyright © 1984 Macmillan Journals Limited; Figure 2.13 reproduced by permission from *Nature*, Copyright © 1984 Macmillan Journals Limited; Figure 2.15 reproduced from *Dev. Biol.* **50** by permission of Academic Press; Figure 2.16 reproduced from *Genetics & Biology of Drosophila*, Vol. 2c by permission of Academic Press; Figure 2.18 reproduced from *Dev. Biol.* **89** by permission of Academic Press; Figure 2.20 reproduced from *31st Symp. Soc. Dev. Biol.* by permission of Academic Press; John Wiley & Sons'(2.19); *Genetics* (2.21, 3.24, 3.35, 3.43); Figures

2.22 & 5.2 reproduced from *Dahlem Konferenzen* 1982 by permission of Springer Verlag; Figure 2.26 reproduced from *37th Symp. Soc. Dev. Biol.* by permission of Academic Press; Figure 2.27 reproduced from *Mobile Genetic Elements* by permission of Academic Press; Figure 2.28 reproduced from *C.S.H.S.Q.B.* **45**, Copyright 1981 by permission of Cold Spring Harbor Laboratory; Figure 2.29 reproduced by permission from *Nature*, Vol. 290, pp. 625–7, Copyright © 1981 Macmillan Journals Limited; Figures 2.30, 2.31, 2.35, 6.5 reproduced from *Scientific American* by permission of W. H. Freeman; Figure 2.32 reproduced from *Cell* **37**, Copyright M.I.T.; Figure 2.34 reproduced from *Chromosoma* **45** by permission of Springer Verlag; Figure 2.37 reproduced from Varmus, *Mobile genetic elements* by permission of Academic Press; P.N.A.S (3.2, 3.8, 3.21, 3.27, 3.39, 4.13); Figure 3.3 reproduced from *Nature*, Vol. 311, pp. 163–5, Copyright © 1984 Macmillan Journals Limited; Figure 3.4 reproduced from *Cell* **38**, Copyright M.I.T.; Figure 3.5 reproduced from *Cell* **32**, Copyright M.I.T.; Figure 3.6 reproduced from *Trends in Genetics* by permission of Elsevier; Figure 3.7 reproduced from *Nature*, Vol. 297, pp. 365–71, Copyright © 1982 Macmillan Journals Limited; Figure 3.9 reproduced from Rosenfield *et al.*, *Science*, Vol. 225, pp. 1315–20, Copyright 1984 the American Association for the Advancement of Science; Figure 3.15 reproduced from Hall *et al.*, *J. Mol. Biol.*, Vol. 169 by permission of Academic Press; Figure 3.22 reproduced from *Cell* **7**, Copyright M.I.T.; Harvard University Press (3.23); Figure 3.25 reproduced from *Chromosoma* **62** by permission of Springer Verlag; Figure 3.27 reproduced from *Cell* **26**, Copyright M.I.T.; Figure 3.28 reproduced from *C.S.H.S.Q.B.* **42**, Copyright 1978, by permission of Cold Spring Harbor Laboratory; Figure 3.30 reproduced from *Cell* **19**, Copyright M.I.T.; Figure 3.36 reproduced from *Experimental Cell Research* **118**, by permission of Academic Press; Figure 3.37 reproduced from *Cell* **27**, Copyright M.I.T.; Figure 3.38 reproduced from *Cell* **29**, Copyright M.I.T.; Figure 3.40 reproduced from *Cell* **34**, Copyright M.I.T.; Figure 4.1 reproduced by permission from *Heredity* **43**; Figures 4.2 & 4.3 reproduced from R. V. Short, *Reproduction in Mammals* **6**, by permission of Cambridge University Press; Figure 4.8 reproduced from *Genetics & Biology of Drosophila*, Vol. 3B, by permission of Academic Press; Figure 4.9 reproduced from *Chromosoma* **85** by permission of Springer Verlag; Figure 4.10 reproduced by permission from *Genetica* **52/53**; Figure 4.11 reproduced from *Cell* **32**, Copyright M.I.T.; Figure 4.12 reproduced from the *Ann. Rev. Genet.*, Vol. 14, © 1980 by permission of Annual Reviews Inc.; Alan R. Liss Inc. (4.14); Figure 4.16 reproduced from *Chromosomes Today*, Vol. 6, by permission of Elsevier; Figure 5.1 reproduced from *Metamorphosis* by permission of Plenum Press; Figure 5.3 reproduced by permission from *Evolution* **36**; Figure 5.4 reproduced from the Royal Society's Philosophical Transactions by permission of the Royal Society; Figure 6.1 reproduced from *Cell* **45**, Copyright Cell Press; Figure 6.2 reproduced from *Mol. Gen. Genet.* **198** by permission of Springer Verlag; Figure 6.4 reproduced from *Molecular Evolutionary Genetics* by permission of Plenum Press; Figure 6.5 reproduced from *Scientific American* by permission of W. H. Freeman.

Contents

List of tables

1

General molecular organization of genomes

'Far more critical than what we want to know or do not know is what we do not want to know.'

Eric Hoffer

1.1 DISSECTING GENOMES

DNA is the most important component of the eukaryote genome in the sense that it ultimately provides the essential coding information necessary for specifying the production of all other molecules within the eukaryote cell. Added to this, it is now the most tractable component in a technological sense. Much of this technology is now so well known as to need no introduction (Weinberg 1985). In this book we have chosen to highlight the fly, *Drosophila melanogaster*, which is, without question, the most thoroughly studied eukaryote in genetical, developmental and molecular terms. For this reason we begin by outlining the four principal techniques currently in use for dissecting the genome of *Drosophila melanogaster*, for isolating specific genes within that genome and to which we refer in the other sections of this book.

1.1.1 Microdissection and microcloning

One technique that has enormously speeded up gene isolation in *Drosophila melanogaster*, and at the same time made the screening of large DNA libraries effectively obsolete, is termed microcloning (Pirrotta 1984). This involves the molecular cloning of picogram quantities of DNA. Although this technique has now been applied to mammals, it was pioneered in *Drosophila melanogaster*. Here giant polytene chromosomes are present in a variety of larval tissues. The remarkable size of these is a consequence of up to ten rounds of chromosome replication without either chromatid separation or cell division. Consequently, the DNA helices are duplicated more than 1000 times. In this polytenization process the different DNA sequences do not behave uniformly. The highly repeated sequences that make up the heterochromatic regions of the mitotic chromosomes replicate

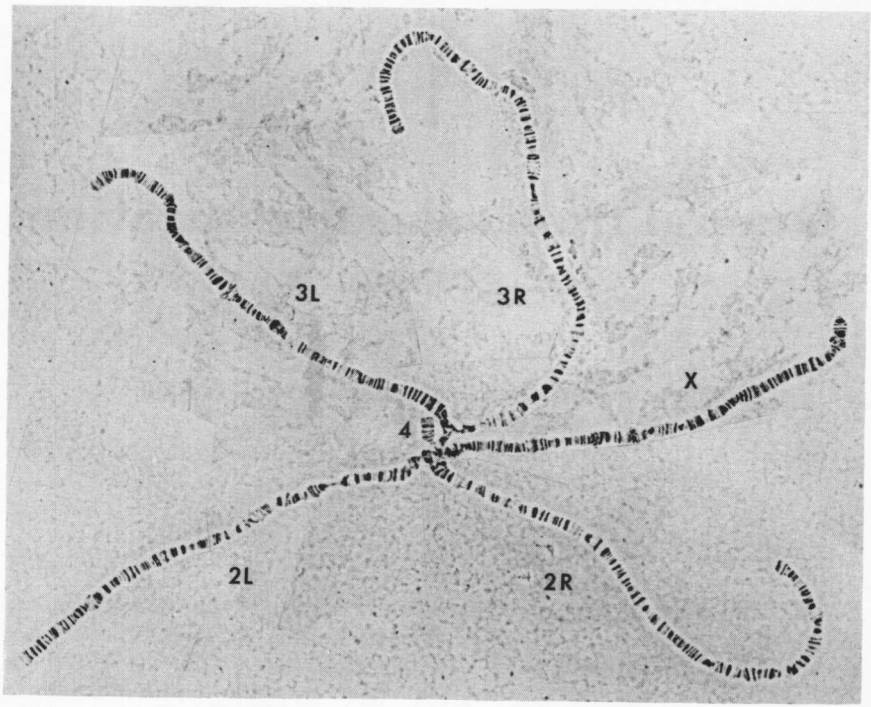

Plate 1 The polytene chromosomes of the larval salivary gland of *Drosophila melanogaster* (photograph kindly supplied by Professor George Lefevre).

little, if at all. These regions coalesce in the polytene nucleus to form a chromocentre, a region consisting of numerous attenuated strands which join the base of each chromosome arm. The result is that the polytene chromosomes are made up almost entirely of the euchromatic portions of the mitotic chromosomes which constitute some 70% of the total DNA. These are laterally duplicated by a factor of 1000 times or more, and the chromosome arms are partitioned into a series of transverse chromatic bands (Plate 1).

These polytene chromosomes provide a cytological map of the genome, and individual polytene bands furnish specific landmarks which allow for the localization of both genes and structural rearrangements. With over 5000 definable bands, the genome is well partitioned even at the level of the light microscope. Consequently, *in situ* hybridization with DNA or RNA probes has a large effective target size even when unique sequences are employed. Figure 1.1 illustrates the chromosomal zip code employed in referring to individual bands. The euchromatin of each major chromosome arm is divided into exactly 20 numbered divisions in the following way: the X-chromosome (1) = 1–20; the left arm chromosome 2 (2L) = 21–40; the right arm of chromosome 2 (2R) = 41–60; the left arm of chromosome 3 (3L) = 61–80; the right arm of chromosome 3 (3R) = 81–100; with two