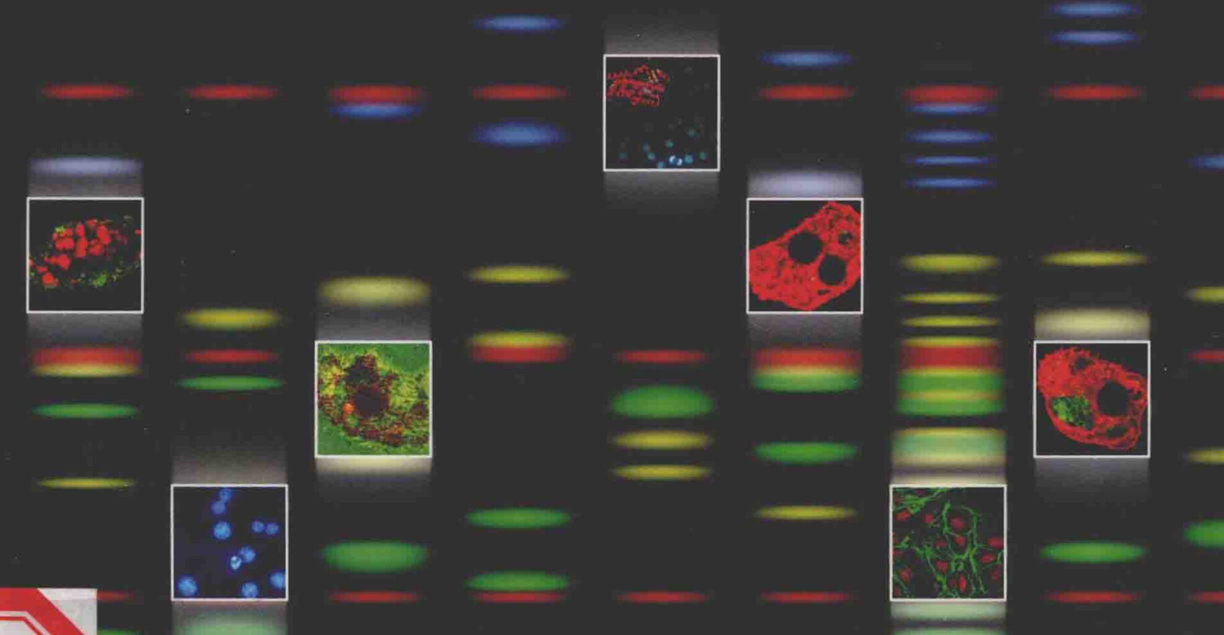
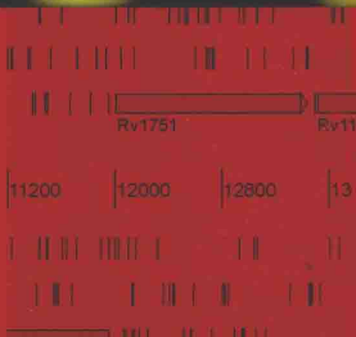
JEREMY W. DALE | MALCOLM VON SCHANTZ | NICK PLANT

FROM GENES TO GENOMES

CONCEPTS AND APPLICATIONS OF DNA TECHNOLOGY

THIRD EDITION





WILEY-BLACKWELL

From Genes to Genomes

Third Edition

Concepts and Applications of DNA Technology

Jeremy W. Dale, Malcolm von Schantz and **Nick Plant**

University of Surrey, UK

WILEY-BLACKWELL
A John Wiley & Sons, Ltd., Publication

From Genes to Genomes

Third Edition

This edition first published 2012

Wiley-Blackwell is an imprint of John Wiley & Sons, formed by the merger of Wiley's global Scientific, Technical and Medical business with Blackwell Publishing.

Registered office
John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial offices
9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK
111 River Street, Hoboken, NJ 07030-5774, USA

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

Library of Congress Cataloging-in-Publication Data

Dale, Jeremy, Professor.
From genes to genomes : concepts and applications of DNA technology / Jeremy W. Dale, Malcolm von Schantz, and Nick Plant. – 3rd ed.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-0-470-68386-6 (cloth) – ISBN 978-0-470-68385-9 (pbk.)
I. Schantz, Malcolm von. II. Plant, Nick. III. Title.
[DNLM: 1. Genetic Engineering. 2. Cloning, Molecular. 3. DNA, Recombinant. QU 450]
LC classification not assigned
660.6′5–dc23

2011030219

A catalogue record for this book is available from the British Library.

This book is published in the following electronic formats: ePDF 9781119953159; ePub 9781119954279; Mobi 9781119954286

Set in 10.5/13pt Times by Aptara Inc., New Delhi, India.
Printed in Singapore by Ho Printing Singapore Pte Ltd

First Impression 2012

Preface

The first edition of this book was published in 2002. By the time of the second edition (2007) the emphasis had moved away from just cloning genes, to embrace a wider range of technologies, especially genome sequencing, the polymerase chain reaction and microarray technology. The revolution has continued unabated, indeed even accelerating, not least with the advent of high-throughput genome sequencing. In this edition we have tried to introduce readers to the excitement engendered by the latest developments – but this poses a considerable challenge. Our aim has been to keep the book to an accessible size, so including newer technologies inevitably means discarding some of the older ones. Some might maintain that we could have gone further in that direction. Some methods that have been kept are no longer as important as they once were, and maybe there is an element of sentimentality in keeping them – but there is some virtue in retaining a balance so that we can maintain a degree of historical perspective. There is a need to understand, to some extent, how we got to the position we are now in, as well as trying to see where we are going.

The main title of the book, *From Genes to Genomes*, is derived from the progress of this revolution. It also indicates a recurrent theme within the book, in that the earlier chapters deal with analysis and investigation at the level of individual genes, and then later on we move towards genome-wide studies – ending up with a chapter directed at the whole organism.

Dealing only with the techniques, without the applications, would be rather dry. Some of the applications are obvious – recombinant product formation, genetic diagnosis, transgenic plants and animals, and so on – and we have attempted to introduce these to give you a flavour of the advances that continue to be made, but at the same time without burdening you with excessive detail. Equally important, possibly more so, are the contributions made to the advance of fundamental knowledge in areas such as developmental studies and molecular phylogeny.

The purpose of this book is to provide an introduction to the concepts and applications of this rapidly moving and fascinating field. In writing it, we had in mind its usefulness for undergraduate students in the biological and biomedical sciences (who we assume will have a basic grounding in molecular biology). However, it will also be relevant for many others, ranging from research workers and teachers who want to update their knowledge of related areas to anyone who would like to understand rather more of the background to current controversies about the applications of some of these techniques.

Jeremy W. Dale
Malcolm von Schantz
Nick Plant

Contents

1
From Genes to Genomes

1.1 Introduction

The classical approach to genetics starts with the identification of variants that have a specific *phenotype*, i.e., they differ from the *wildtype* in some way that can be seen (or detected in other ways) and defined. For Gregor Mendel, the father of modern genetics, this was the appearance of his peas (e.g., green versus yellow, or round versus wrinkled). One of the postulates he arrived at was that these characteristics assorted independently of one another. For example, when crossing one type of pea that produces yellow, wrinkled peas with another that produces green, round peas, the first generation (F_1) are all round and yellow (because round is dominant over wrinkled, and yellow is dominant over green). In the second (F_2) generation, there is a 3 : 1 mixture of round versus wrinkled peas, and independently a 3 : 1 mixture of yellow to green peas.

Of course Mendel did not know why this happened. We now know that if two genes are located on different chromosomes, which will segregate independently during meiosis, the genes will be distributed independently amongst the progeny. Independent assortment can also happen if the two genes are on the same chromosome, but only if they are so far apart that any recombination between the homologous chromosomes will be sufficient to reassort them independently. However, if they are quite close together, recombination is less likely, and they will therefore tend to remain associated during meiosis. They will therefore be inherited together. We refer to genes that do *not* segregate independently as *linked*; the closer they are, the greater the degree of linkage, i.e., the more likely they are to stay together during meiosis. Measuring the degree of linkage (*linkage analysis*) is a central tool in classical genetics, in that it provides a way of mapping genes, i.e., determining their relative position on the chromosome.

From Genes to Genomes: Concepts and Applications of DNA Technology, Third Edition.
Jeremy W. Dale, Malcolm von Schantz and Nick Plant.

Bacteria and yeasts provide much more convenient systems for genetic analysis, because they grow quickly, as unicellular organisms, on defined media. You can therefore use chemical or physical mutagens (such as ultraviolet irradiation) to produce a wide range of mutations, and can select specific mutations from very large pools of organisms – remembering that an overnight culture of *Escherichia coli* will contain some 10^9 bacteria per millilitre. So we can use genetic techniques to investigate detailed aspects of the physiology of such cells, including identifying the relevant genes by mapping the position of the mutations.

For multicellular organisms, the range of phenotypes is even greater, as there are then questions concerning the development of different parts of the organism, and how each individual part influences the development of others. However, animals have much longer generation times than bacteria, and using millions of animals (especially mammals) to identify the mutations you are interested in is logistically impossible, and ethically indefensible. Human genetics is even more difficult as you cannot use selected breeding to map genes; you have to rely on the analysis of real families, who have chosen to breed with no consideration for the needs of science. Nevertheless, classical genetics has contributed extensively to the study of developmental processes, notably in the fruit fly *Drosophila melanogaster*, where it is possible to study quite large numbers of animals, due to their relative ease of housing and short generation times, and to use mutagenic agents to enhance the rate of variation.

However, these methods suffered from a number of limitations. In particular, they could only be applied, in general, to mutations that gave rise to a phenotype that could be defined in some way, including shape, physiology, biochemical properties or behaviour. Furthermore, there was no easy way of characterizing the nature of the mutation. The situation changed radically in the 1970s with the development of techniques that enabled DNA to be cut precisely into specific fragments, and to be joined together, enzymatically –techniques that became known variously as genetic manipulation, genetic modification, genetic engineering or recombinant DNA technology. The term ‘gene cloning’ is also used, since joining a fragment of DNA with a vector such as a plasmid that can replicate in bacterial cells enabled the production of a bacterial strain (a clone) in which all the cells contained a copy of this specific piece of DNA. For the first time, it was possible to isolate and study specific genes. Since such techniques could be applied equally to human genes, the impact on human genetics was particularly marked.

The revolution also depended on the development of a variety of other molecular techniques. The earliest of these (actually predating gene cloning) was *hybridization*, which enabled the identification of specific DNA sequences on the basis of their sequence similarity. Later on came methods for determining the sequence of these DNA fragments, and the polymerase chain reaction (PCR), which provided a powerful way of amplifying specific DNA sequences. Combining those advances with automation, plus the concurrent advance in computer power, led to the determination of the full genome sequence of many organisms, including the human genome, and thence to enormous advances in understanding the roles of genes and their products. In recent years, sequencing technology has advanced to a stage where it is now a routine matter to sequence the full genome of many individuals, and thus attempt to pinpoint the causes of the differences between them, including some genetic diseases.

Furthermore, since these techniques enabled the cloning and expression of genes from any one organism (including humans) into a more amenable host, such as a bacterium, they allowed the use of genetically modified bacteria (or other hosts) for the production of human gene products, such as hormones, for therapeutic use. This principle was subsequently extended to the genetic modification of plants and animals – both by inserting foreign genes and by knocking out existing ones – to produce plants and animals with novel properties.

As is well known, the construction and use of genetically modified organisms (GMOs) is not without controversy. In the early days, there was a lot of concern that the introduction of foreign DNA into *E. coli* would generate bacteria with dangerous properties. Fortunately, this is one fear that has been shown to be unfounded. Due to careful design, genetically modified bacteria are, generally, not well able to cope with life outside the laboratory, and hence any GM bacterium released into the environment (deliberately or accidentally) is unlikely to survive for long. In addition, one must recognize that nature is quite capable of producing pathogenic organisms without our assistance – which history, unfortunately, has repeatedly demonstrated through disease outbreaks.

The debate on GMOs has now largely moved on to issues relating to genetically modified plants and animals. It is important to distinguish the *genetic modification* of plants and animals from *cloning* of plants and animals. The latter simply involves the production of genetically identical individuals; it does not involve any genetic modification whatsoever. (The two technologies can be used in tandem, but that is another matter.) There are ethical issues to be considered, but cloning plants and animals is not the subject of this book.

Currently, the debate on genetic modification can be envisaged as largely revolving around two factors: food safety and environmental impact. The first thing to be clear about is that there is no imaginable reason why genetic modification, per se, should make a foodstuff hazardous in any way. There is no reason to suppose that cheese made with rennet from a genetically modified bacterium is any more dangerous than similar cheese made with 'natural' rennet. It is possible to imagine a risk associated with some genetically modified foodstuffs, due to unintended stimulation of the production of natural toxins – remembering, for example, that potatoes are related to deadly nightshade. But this can happen equally well (or perhaps is even more likely) with conventional cross-breeding procedures for developing new strains, which are not always subject to the same degree of rigorous safety testing as GM plants.

The potential environmental impact is more difficult to assess. The main issue here is the use of genetic modification to make plants resistant to herbicides or to insect attack. When such plants are grown on a large scale, it is difficult to be certain that the gene in question will not spread to related wild plants in the vicinity (although measures can be taken to reduce this possibility), or the knock-on effect that such resistance may have on the ecosystem – if all the insects are killed, what will small birds and animals eat? But these concerns may be exaggerated. As with the bacterial example above, these genes will not spread significantly unless there is an evolutionary pressure favouring them. So we would not expect widespread resistance to weedkillers unless the plants are being sprayed with those weedkillers. There might be an advantage in becoming resistant to insect attack, but the insects concerned have been around for a long time, so the wild plants have had plenty of time to develop natural resistance anyway. In addition, targeted resistance in a group of plants may arguably have less environmental impact than the less targeted spraying of insecticides. We have to balance the use of genetically modified plants against the use of chemicals. If genetic modification of the plants means a reduction in the use of environmentally damaging chemicals, then that is a tangible benefit that could outweigh any theoretical risk.

The purpose of this book is to provide an introduction to the exciting developments that have resulted in an explosion of our knowledge of the genetics and molecular biology of all forms of life, from viruses and bacteria to plants and mammals, including of course ourselves – developments that continue as we write. We hope that it will convey some of the wonder and intellectual stimulation that this science brings to its practitioners.

1.2 Basic molecular biology

In this book, we assume you already have a working knowledge of the basic concepts of molecular and cellular biology. This section serves as a reminder of the key aspects that are especially relevant to this book.

1.2.1 The DNA backbone

Manipulation of nucleic acids in the laboratory is based on their physical and chemical properties, which in turn are reflected in their biological function. Intrinsically, DNA is a remarkably stable molecule. Indeed, DNA of sufficiently high quality to be analysed has been recovered from frozen