MOLECULAR BIOLOGY

LABFAX

EDITED BY
T.A. BROWN

MOLECULAR BIOLOGY

LABFAX

EDITED BY
T.A. BROWN
Department of Biochemistry and Applied Molecular Biology,
UMIST, Manchester M60 1QD, UK

BIOS SCIENTIFIC PUBLISHERS

BLACKWELL SCIENTIFIC PUBLICATIONS

BIOS Scientific Publishers Limited,
St Thomas House, Becket Street, Oxford OX1 1SJ

ISBN 1 872748 00 7

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

Published jointly in the United Kingdom by BIOS Scientific Publishers Ltd and Blackwell Scientific Publications Ltd.

Distributed in the United States, its territories and dependencies, and Canada exclusively by Academic Press, Inc., 1250 Sixth Avenue, San Diego, California 92101.

Distributed in Japan exclusively by Central Foreign Books Ltd (Chuo Yosho), 1–13 Jinbocho Kanda, Chiyoda-ku, Tokyo.

Distributed in the rest of the world exclusively by Blackwell Scientific Publications Ltd, Osney Mead, Oxford OX2 0EL.

Designed and typeset by the Opus Design Group Ltd, Oxford, UK
Printed by Information Press Ltd, Oxford, UK

The cover illustration shows, with permission, the pGEM-11Zf(–) vector available from Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711, USA

PREFACE

There is nothing new under the sun and so it would be foolhardy to suggest that *Molecular Biology Labfax* is an entirely new departure in scientific publishing. It is, however, different from the existing cloning manuals in that it is designed as a companion rather than a guide for molecular biology research. *Molecular Biology Labfax* does not contain procedures or methodology but instead is a detailed compendium of the essential information — on genotypes, reagents, enzymes, reaction conditions, cloning vectors and suchlike — that is needed to plan and carry out molecular biology research. Some of this information is already available in cloning manuals, catalogs and possibly on pieces of paper kept somewhere safe, but tracking down exactly what you need to know takes time and can be a frustrating experience. With molecular biology becoming an increasingly sophisticated science, an acute need has arisen for a databook to complement the traditional cloning manuals. *Molecular Biology Labfax* is intended to meet this need.

To be useful, the coverage of *Molecular Biology Labfax* has to be right. The scope of the book is of necessity a compromise between a desire to include everything and a need to keep within a reasonable size limit. The reader will expect to find extensive details of *Escherichia coli* genotypes and genetic markers, restriction enzymes, DNA and RNA modifying enzymes, chemicals and reagents, cloning vectors, restriction fragment patterns and suchlike. These topics are covered in as comprehensive a way as possible, so for instance in Chapter 4 all the known restriction enzymes are described along with reaction conditions for all the commercially available ones. A few items that might be expected are not included on the basis that they are of specialist interest, an example being cloning vectors for eukaryotes. Topics such as these will be covered by future editions in the Labfax series. Although experimental protocols are not given, certain key information is provided for subjects such as the growth of *E. coli* strains, use of radiochemicals, electrophoresis of nucleic acids, and hybridization analysis. These topics are of widespread importance in molecular biology procedures and so warrant sections of their own. Readers using other standard techniques will find their needs met by the data presented throughout the book. For instance, the DNA sequencer will find data on dideoxynucleotides and Maxam–Gilbert reagents (Chapter 2), radionuclectides and detection methods (Chapter 3), enzymes for chain termination sequencing (Chapter 5), M13 cloning vectors (Chapter 6), and electrophoresis systems (Chapter 8), as well as details of the genetic code and codon usages for interpretation of sequence information (Chapter 7).

A second essential requirement is accuracy. Wherever possible I have double-checked items that I have had doubts about, going back to the original publications if necessary. In a few cases the literature contains annoying contradictions that I have been unable to resolve, with *E. coli* genotypes providing some of the biggest headaches. The relevant entries carry a footnote or other warning to alert the reader and I welcome enlightenment if anyone knows any of the answers.

Without the help of a number of people *Molecular Biology Labfax* would never have been completed. I am very grateful to Rich Roberts, Toshimichi Ikemura and Michael McClelland for their contributions, as well as GIBCO-BRL, Pharmacia, Promega, USB, Clontech, Stratagene

and FMC for providing artwork for the cloning vectors and electrophoresis sections. I would like to give a general thank-you to the various colleagues and friends who helped me out with points here and there. Half-way through the enterprise it became clear that you need to be slightly unbalanced to compile a databook: I became even more worried on the occasions when I thought I was actually enjoying the experience. Because of this I must thank my wife, Keri, who made sure I survived to tell the tale.

T.A. Brown

SAFETY NOTE

Safety is a critical aspect of laboratory work and all molecular biologists should be aware of the precautions needed to ensure personal safety and the safety of colleagues. National and local safety precautions must be followed at all times. All reputable suppliers of laboratory chemicals provide risk and safety information with those products that present a potential hazard, but several risk and safety classification schemes exist. This book cites risk and safety data for products where appropriate: full details of the classification schemes used, together with information on protection from microbiological and radiochemical hazards, are given in Chapter 11.

CONTRIBUTORS

T.A. BROWN

Department of Biochemistry and Applied Molecular Biology, UMIST, Manchester M60 1QD, UK

T. IKEMURA

National Institute of Genetics, Mishima, Shizuoka-ken 411, Japan

M. McCLELLAND

California Institute of Biological Research, 11099 North Torrey Pines Road, La Jolla, CA 92037, USA

R.J. ROBERTS

Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

ABBREVIATIONS

ATP	adenosine triphosphate
Bq	becquerel
bp	base pair
b.p.	boiling point
BSA	bovine serum albumin
c.p.m.	counts per minute
cccDNA	covalently-closed-circular DNA
Ci	Curie
DNase	deoxyribonuclease
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide
d.p.m.	disintegrations per minute
DTT	dithiothreitol
dsDNA	double-stranded DNA
dsRNA	double-stranded RNA
eV	electron volt
f.p.	flash point
g	acceleration due to gravity
h	hour
kb	kilobase
kd	kilodalton
m.p.	melting point
mRNA	messenger ribonucleic acid
min	minute
mol	mole
mol. wt	molecular weight
nt	nucleotide
ORF	open reading frame
pH	hydrogen-ion exponent
RNase	ribonuclease
sec	second
ssDNA	single-stranded DNA
ssRNA	single-stranded RNA
T_m	melting temperature
tRNA	transfer ribonucleic acid
u.v.	ultraviolet
(v/v)	volume/volume
(w/v)	weight/volume
(w/w)	weight/weight
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

CONTENTS

CHAPTER 1
BACTERIA AND BACTERIOPHAGES

1. *E. COLI* STRAINS USED IN RECOMBINANT DNA EXPERIMENTS

The *E. coli* strains routinely used in recombinant DNA experiments are listed with their genotypes in *Table 1*. The genotypes are described in accordance with the standard nomenclature as defined below.

1.1. Individual genes

(i) Each mutant locus is described by a three-letter abbreviation (e.g. *ara* = arabinose utilization). The abbreviations are defined in *Table 2*.
(ii) The capital letter following the locus refers to the individual gene that is mutated (e.g. *araD* = L-ribulosephosphate 4-epimerase). The genes are also described in *Table 2*.
(iii) Numbers following the gene designation refer to the specific allele involved (e.g. *araD139*).
(iv) A superscript '–' is generally not used, since, by convention, only mutated genes are listed in the genotype. A superscript '+' may be used to emphasize a locus or gene that is wild-type (e.g. lac^+ = no mutations in the genes involved in lactose utilization).
(v) A superscript 'q' indicates a constitutive mutation (e.g. $lacI^q$ = constitutive mutant for the *lac* repressor).
(vi) An amber mutation is denoted by 'am' following the gene designation (e.g. *malBam*).
(vii) If an antibiotic response is listed in the genotype then a superscript 'r' or 's' is used to denote resistance or sensitivity respectively (e.g. kan^r = kanamycin resistant).

1.2. Deletions

Deletions are denoted by 'Δ' with the deleted gene or genes listed in brackets, possibly followed by an allele designation outside of the brackets [e.g. *Δ(gal–uvrB)40* = deletion of the region from *gal* to *uvrB*].

1.3. Fusions

(i) A fusion is denoted in the same way as a deletion, except that the symbol 'Φ' is used.
(ii) A prime (') is used to designate that the fused gene is incomplete (e.g. *'lacZ* indicates that the *lacZ* gene involved in the fusion is deleted in the 5' region; *lacZ'* indicates a deletion in the 3' region).
(iii) A superscript '+' (e.g. $lacZ^+$) denotes that the fusion involves an operon rather than a single gene.

1.4. Insertions

(i) An insertion is denoted by '::', preceded by the position of the insertion and followed by the inserted DNA (e.g. *trpC22*::Tn*10* = insertion of Tn*10* into the *trpC* gene, allele 22).
(ii) If the insertion does not occur within a known gene then the map position is denoted by a three-letter code. The first letter is always *z*, followed by *a–g* to indicate a 10 min interval, and *a–i* to indicate a 1 min interval (e.g. *zgi* = 79 min).

1.5. Phages and plasmids

A plasmid or lysogenic phage carried by the bacterium is listed at the end of the genotype in brackets [e.g. (pMC9) = carries plasmid pMC9].

1.6. Fertility status

(i) Strains are assumed to be F^- unless the status is given.
(ii) F^+ and Hfr strains are denoted by the relevant symbol at the start of the genotype.
(iii) When the strain is listed as F', the genes carried on the episome are listed in square brackets. The F' status is usually placed at the end of the genotype.

The full restriction and modification status of individual strains is not given in *Table 1*. The genotypes are correct for *hsdR*, *hsdS* and *hsdM*, but *mcrA*, *mcrB* and *mrr* are not included. This is because for many strains the *mcrA*, *mcrB* and *mrr* genotypes have not yet been determined. Full descriptions of the restriction and modification status, as far as is known, of important strains are given in *Table 3*. *Table 4* classifies strains according to their specific application(s) in recombinant DNA experiments.

Table 1. Genotypes of *E. coli* strains used in recombinant DNA experiments

Strain	Genotype	References
594	*rpsL*	1
1101	F^+ *supE*	2
71/18	*supE thi* Δ*(lac–proAB)* F'[*proAB*$^+$ *lacI*^{q} *lacZ*Δ*M15*]	3–6
χ1776[a]	*tonA53 dapD8 minA1 supE42* Δ*(gal–uvrB)40 minB2 rfb-2 gyrA25 thyA142 oms-2 metC65 oms-1* Δ*(bioH–asd)29 tte-1 cycB2 cycA1 hsdR2*	7, 8
AG1[b]	*recA1 endA1 gyrA96 thi hsdR17 supE44 relA1*	9
BB4	*supF58 supE44 hsdR514 galK2 galT22 trpR55 metB1 tonA* Δ*(lac)U169* F'[*proAB*$^+$ *lacI*^{q} *lacZ*Δ*M15* Tn*10(tet*^{r}*)*]	1
BHB2600	*supE supF* (λCH616)	10
BHB2688	N205 *recA* (λ*imm*434 cIts *b2 red3 E*am4 *S*am7)/λ	11, 12
BHB2690	N205 *recA* (λ*imm*434 cIts *b2 red3 D*am15 *S*am7)/λ	11, 12
BJ5183	*endA sbcB recBC galK met str*^{r} *thi-1 bioT hsdR*	13
BL21(DE3)[c]	*hsdS gal* (λcIts857 *ind*1 *S*am7 *nin*5 *lacUV5*-T7 gene 1)	14
BNN93	see C600	
BNN102[d]	*supE44 hsdR thi-1 thr-1 leuB6 lacY1 tonA21 hflA150*[*chr*::Tn*10(tet*^{r}*)*]	15, 16
C-1A	wild-type	17, 18
C600[e]	*supE44 hsdR*? *thi-1 thr-1 leuB6 lacY1 tonA21*	13, 15, 16, 19–21
C600*galK*	see C600	
C600*hflA*	see BNN102	
CES200	*sbcB15 recB21 recC22 hsdR*	23
CES201	*recA sbcB15 recB21 recC22 hsdR*	24
CJ236	*dut1 ung1 thi-1 relA1* (pCJ105[*cam*^{r}F'])	25
CR34	see C600	

Table 1. Continued

Strain	Genotype	References
CSH18	*supE thi* Δ*(lac–pro)* F'[*proAB*$^+$ *lacZ*$^-$]	26, 27
D1210	*supE44 hsdS20 recA13 ara-14 proA2 lacI*^{q} *galK2 rpsL20 xyl-5 mtl-1*	28
DH1	*supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	13, 29
DH5	*supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	13, 29
DH5α	*supE44* Δ*lacU169* (ϕ80 *lacZ*Δ*M15*) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	13
DH20	*supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1* F'[*lacI*^{q} *lacZ*$^+$ *proAB*$^+$]	13
DH21	*supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1* F'[*lacI*^{q} *lacZ*$^+$ *proAB*$^+$]	13
DP50*supF*[f,g]	*supE44 supF58 hsdS3 dapD8 lacY1 glnV4* Δ*(gal–uvrB)47 tyrT58 gyrA29 tonA53* Δ*(thyA)57*	30
ED8654	*supE supF hsdR metB lacY gal trpR*	18, 31
ED8767[g]	*supE44 supF58 hsdS3 recA56 galK2 galT22 metB1*	31
ER1398[h,i]	*supE44 hsdR endA1 thi*	20
ER1451[j]	*supE44 endA1 hsdR17 gyrA96 relA1 thi* Δ*(lac–proAB)* F'[*traD36 proAB*$^+$ *lacI*^{q} *lacZ*Δ*M15*]	20
ER1647*	*serB28* Δ*(mrr–hsdRMS–mcrB) mcrA1272 recD*	32
ER1648*	*serB28* Δ*(mrr–hsdRMS–mcrB) mcrA1272*	32
GM48	*thr leu thi lacY galK galT ara tonA tsx dam dcm supE44*	3
GM2163	*hsdR dam dcm supE*	20, 33
HB101[g,k]	*supE44 hsdS20 recA13 ara-14 proA2 lacY1 galK2 rpsL20 xyl-5 mtl-1*	13, 20, 34, 35
HMS174	*recA1 hsdR rif*^{r}	14, 36
JC7623	*recB1 recC22 sbcB15*	37
JC9956	*recA99 thr-1 leu-6 thi-1 lacY1 galK2 ara-14 xyl-5 mtl-1 proA2 his-4 argE3 str-31 tsx-33*	38
JM83	*ara* Δ*(lac–proAB) strA thi-1* (ϕ80 *lacZ*Δ*M15*)	3, 39
JM101	*supE thi-1* Δ*(lac–proAB)* F'[*traD36 proAB*$^+$ *lacI*^{q} *lacZ*Δ*M15*]	3, 20, 40
JM103[l]	*supE thi-1 endA1 hsdR4 sbcB15 strA* Δ*(lac–proAB)* F'[*traD36 proAB*$^+$ *lacI*^{q} *lacZ*Δ*M15*]	13, 41
JM103Y	see JM103	
JM105[m]	*endA sbcB15 hsdR4 rpsL thi* Δ*(lac–proAB)* F'[*traD36 proAB*$^+$ *lacI*^{q} *lacZ*Δ*M15*]	3
JM106	*supE44 endA1 hsdR17 gyrA96 relA1 thi* Δ*(lac–proAB)*	3
JM107	*supE44 endA1 hsdR17 gyrA96 relA1 thi* Δ*(lac–proAB)* F'[*traD36 proAB*$^+$ *lacI*^{q} *lacZ*Δ*M15*]	3, 20
JM108	*recA1 supE44 endA1 hsdR17 gyrA96 relA1 thi* Δ*(lac–proAB)*	3
JM109	*recA1 supE44 endA1 hsdR17 gyrA96 relA1 thi* Δ*(lac–proAB)* F'[*traD36 proAB*$^+$ *lacI*^{q} *lacZ*Δ*M15*]	3

* See note on page 24.

Table 1. Continued

Strain	Genotype	References
JM109(DE3)[c]	*recA1 supE44 endA1 hsdR17 gyrA96 relA1 thi* Δ*(lac–proAB)* F'[*traD36 proAB*$^+$ *lacI*^{q} *lacZ*Δ*M15*] (λcIts857 *ind*1 Sam7 *nin*5 *lacUV5*-T7 gene 1)	43
JM110[n]	*dam dcm supE44 thi leu rpsL lacY galK galT ara tonA thr tsx* Δ*(lac–proAB)* F'[*traD36 proAB*$^+$ *lacI*^{q} *lacZ*Δ*M15*]	3
K802	*supE hsdR gal metB*	20, 44
KK2186	see JM103	45
KLF41	*leuB6 hisG1 recA1 argG6 metB1 lacY1 gal-6 xyl-7 mtl-2 malA1 rpsL104 tonA tsx supE44* F'141	1
LE30	*mutD5 rpsL azi galU95*	1
LE292	HfrH *argEam rpoB galT*::[λΔ(*int–*FII)]	1
LE392	*supE44 supF58 hsdR514 galK2 galT22 metB1 trpR55 lacY1*	13, 18, 20, 31
LE392.23	*supE44 supF58 hsdR514 galK2 galT22 metB1 trpR55 lacY1* Δ*(argF–lac)U169*	1
LG90	Δ*(lac–proAB)*	46
M5219	*lacZ trpA rpsL* (λ*bio*252 *c*Its857 *H*1)	47, 48
MAL103	Δ*(gpt–proAB–argF–lac)XIII rpsL* [Mudl (*lac*, *Ap*)] (Mucts62)	1
MB100	Δ*(argF–lac)U169 rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR leuABCD*::Tn*10*	1
MB101	*araCam araD* Δ*(argF–lac)U169 trpam malBam rpsL relA thi supF* Φ*(araBA'–lacZ*$^+$*)101* [λ*p*1(209)]	1
MBM7007	*araCam araD* Δ*(argF–lac)U169 trpam malBam rpsL relA thi*	1
MBM7014	*araCam araD* Δ*(argF–lac)U169 trpam malBam rpsL relA thi supF*	1
MBM7014.5	*hsdR2 zjj202*::Tn*10(tet*^{r}*) araD139 araCU25am* Δ*(lac)U169*	20
MBM7060	*araCam araD* Δ*(argF–lac)U169 trpam malBam rpsL relA thi supF* (λ*p*1048)	1
MC1000	*araD139* Δ*(araABC–leu)7679 galU galK* Δ*(lac)X74 rpsL thi*	1
MC1061	*hsdR araD139* Δ*(araABC–leu)7679* Δ*(lac)X74 galU galK rpsL thi*	16, 49, 50
MC4100	*araD139* Δ*(argF–lac)U169 rpsL150 relA1 flbB3501 deoC1 ptsF25 rbsR*	1
MH225	*araD139* Δ*(argF–lac)U169 rpsL150 relA1 flbB3501 deoC1 ptsF25 rbsR* Δ*(ompC'–lacZ*$^+$*)10–25* [λ*p*1(209)]	1
MH513	Δ*(argF–lac)U169 rpsL150 relA1 flbB3501 deoC1 ptsF25 rbsR* Φ*(ompF'–lacZ*$^+$*)16–23* [λ*p*1(209)]	1