

# Pathogenomics

Genome Analysis of Pathogenic Microbes

*Edited by*

*Jörg Hacker and Ulrich Dobrindt*



WILEY-  
VCH

WILEY-VCH Verlag GmbH & Co. KGaA



**Pathogenomics**

*Edited by*

*Jörg Hacker and Ulrich Dobrindt*

## ***Related Titles***

Schumann, W.

### **Dynamics of the Bacterial Chromosome**

**Structure and Function**

2006

ISBN 3-527-30496-7

Black, J. G.

### **Microbiology**

**Principles and Explorations**

2005

ISBN 0-471-42084-0

Zhou, J., Thompson, D. K., Xu, Y., Tiedje, J. M.

### **Microbial Functional Genomics**

2004

ISBN 0-471-07190-0

Grandi, G. (ed.)

### **Genomics, Proteomics and Vaccines**

2004

ISBN 0-470-85616-5

Dale, J. W., Park, S.

### **Molecular Genetics of Bacteria**

2004

ISBN 0-470-85084-1

Streips, U. N., Yasbin, R. E. (eds.)

### **Modern Microbial Genetics**

2002

ISBN 0-471-38665-0

Gellissen, G. (ed.)

### **Hansenula polymorpha**

**Biology and Applications**

2002

ISBN 3-527-30341-3

Sensen, C. W. (ed.)

### **Essentials of Genomics and Bioinformatics**

2002

ISBN 3-527-30541-6

Singleton, P., Sainsbury, D.

### **Dictionary of Microbiology and Molecular Biology**

2001

ISBN 0-471-94150-6

Bahl, H., Dürre, P.

### **Clostridia**

**Biotechnology and Medical Applications**

2001

ISBN 3-527-30175-5

# Pathogenomics

Genome Analysis of Pathogenic Microbes

*Edited by*

*Jörg Hacker and Ulrich Dobrindt*



WILEY-  
VCH

WILEY-VCH Verlag GmbH & Co. KGaA

#### Series Editors

**Prof. Dr. Dr. h. c. mult. Jörg Hacker**

Universität Würzburg  
Institut für Molekulare Infektionsbiologie  
Röntgenring 11  
97070 Würzburg  
Germany

**Dr. Ulrich Dobrindt**

Universität Würzburg  
Institut für Molekulare Infektionsbiologie  
Röntgenring 11  
97070 Würzburg  
Germany

All books published by Wiley-VCH are carefully produced. Nevertheless, authors, editors, and publisher do not warrant the information contained in these books, including this book, to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

**Library of Congress Card No.:** applied for

**British Library Cataloguing-in-Publication Data**

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

**Bibliographic information published by**

**Die Deutsche Bibliothek**

Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <<http://dnb.ddb.de>>.

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

All rights reserved (including those of translation into other languages).

No part of this book may be reproduced in any form – nor transmitted or translated into machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Printed in the Federal Republic of Germany.  
Printed on acid-free paper.

**Typesetting** Kühn & Weyh, Satz und Medien, Freiburg

**Printing** betz-druck GmbH, Darmstadt

**Bookbinding** J. Schäffer GmbH i. G., Grünstadt

**ISBN-13:** 978-3-527-31265-8

**ISBN-10:** 3-527-31265-X

## Foreword

The determination of the genome sequences of many prokaryotic and eukaryotic organisms, together with the high-throughput techniques (transcriptomics, proteomics, metabolomics, interactomics, etc.) and the powerful tools of bioinformatics, have opened up to us all new perspectives for the deeper understanding of the basic mechanisms of life.

Because of the small size of bacterial genomes and the collinearity of their genetic information, prokaryotes are particularly suited to genome-based in-depth analysis of the essential processes which allow these microorganisms to survive and replicate in many different environments. Some of them can even multiply in those parts of the human body which are normally well protected by highly developed antimicrobial defense mechanisms. This latter property is the most outstanding evolutionary achievement of microbial pathogens that are capable of causing infectious diseases in humans.

The present monograph summarizes the state of genome-based research on some of the most important bacterial (and to a lesser extent viral and fungal) human pathogens today. The term pathogenomics has been coined for this new branch of microbiology. Not surprisingly, the major focus of pathogenomics was first on those human bacterial pathogens that (a) can cause major epidemics, especially in developing countries (e.g., *Shigella* spp. or *Vibrio cholerae*); (b) represent major health problems in almost all human societies in the form of food contaminants and/or agents of nosocomial infections (e.g., pathogenic *Enterobacteriaceae*, staphylococci, and streptococci); and (c), due to their frequent occurrence in humans (e.g., *Helicobacter pylori* and *Mycobacterium tuberculosis*) represent life-threatening problems not only for many people in developing countries, but also for immune-compromised and elderly persons in all the industrialized countries. These microorganisms therefore also represent the major objects described in this book.

A primary goal of pathogenomics is to use the new experimental tools in combination to unravel those genes – and gene arrangements – of pathogens that are essential for causing disease, and thereby to shed new light on the evolution, interbacterial gene transfer, and distribution of these virulence genes among bacterial populations. This area of pathogenomics, which is already highly advanced, makes up the largest part of this monograph.

The other, certainly not less important goal is to study the functional significance of the pathogen-specific genes in the infection process. In the years to come, the enormous amount of genetic information that has piled up in the last 10 years from the sequencing of the genomes of most of the important bacterial pathogens (and their closely related nonpathogenic environmental counterparts) will need to be functionally analyzed. It is anticipated that this functional pathogenomics will finally unravel the mechanisms of differential and coordinated regulation of the virulence genes, the structural, molecular, and physiological functions of their gene products, which will lead to a more comprehensive view on the pathogenic microorganisms. Some aspects of this most interesting future line of microbial research that can be expected to become the mainstream in future pathogenomics are already addressed in some of the chapters of this book.

Infections are the outcome of the encounter between the microbial pathogen and its host. Describing microbial infections in molecular terms therefore requires among other things a profound understanding of the host cell responses. The availability today of the genome sequences of man and some of the major model hosts in which microbial infections can be experimentally studied (the mouse in particular, but also alternative hosts such as amebae, the nematode *Caenorhabditis elegans*, *Drosophila melanogaster*, or the slime mold *Dictyostelium*) together with the new genetic and bioinformatics tools open up new avenues towards uncovering at least some of the basic host cell responses.

Whilst cellular microbiology has already delivered an enormous set of valuable host-response information at the cellular level, the new *in vivo* imaging techniques, siRNA technology, and the routine genetic manipulations of some of the abovementioned model hosts may now allow such molecular infection studies to be performed in real hosts. This monograph devotes several chapters to this important aspect of pathogenomics.

Science and, especially, the public and political worlds expect pathogenomics to provide novel ideas and strategies to combat infectious diseases through the rational design of better diagnostic tools and novel anti-infectives and vaccines. Indeed, some promising new developments deriving from pathogenomics are already visible and are in part outlined in some chapters of this monograph. These new approaches will most undoubtedly help in fighting the most dangerous infectious agents which still claim the highest death toll in mankind. However, scientists should be modest in their promises relating to the possibilities that arise from this exciting new field of science, since one major lesson which the attentive and critical reader of this monograph will quickly learn is the enormous genetic flexibility and adaptation potential of pathogenic microorganisms – and that means they will keep us busy for decades to come despite all the obvious successes of pathogenomics.

The book thus offers a representative view of the present state of the art in the new research area of pathogenomics. It is a valuable and reasonably comprehensive source of information for scientists and advanced students who wish to become acquainted with this most exciting field of modern microbiology.



## Contents

	<b>Foreword</b>	V
	<b>Preface</b>	XIX
	<b>List of Contributors</b>	XXI
	<b>Color Plates</b>	XXVII
I	<b>Methods</b>	1
1	<b>Bioinformatics: Data Mining Among Genome Sequences</b>	3
	<i>Susanne Kneitz and Thomas Dandekar</i>	
1.1	Systematic Genome Analysis of Pathogens as a Basis for Pharmacogenomic Strategies	3
1.2	Direct Sequence Annotation Tools for Functional Genomics	4
1.3	Identification of Protein Function	4
1.4	Obtaining Protein Information from a Domain Server	5
1.5	Pathway Analysis	6
1.6	Network Analysis	7
1.7	Adaptation in Time and to Stimuli	8
1.7.1	Experimental Design for Microarray Analysis	8
1.7.2	Data Analysis	9
1.8	Pathogen-Specific Challenges	13
1.9	Pathogen Adaptation Potential	14
1.10	The Fight Against Resistance	14
1.11	Drug Design and Antibiotics	15
1.12	Annotation Platforms Suitable for Pathogenomics	15
1.13	Conclusions	17

<b>2</b>	<b>Transcriptome Analysis: Towards a Comprehensive Understanding of Global Transcription Activity</b>	<b>21</b>
	<i>Ben Sidders and Neil Stoker</i>	
2.1	Introduction	21
2.2	Development of Transcriptomics	21
2.2.1	From Genomics to Functional Genomics	21
2.2.2	From Gene to Whole Genome	22
2.3	Introducing the Microarray	23
2.3.1	What Is a Microarray?	23
2.3.2	The Affymetrix Gene Chip	23
2.3.3	The Spotted Microarray	24
2.4	Microarray Methods	25
2.4.1	Experimental Design	25
2.4.1.1	Type of Experiment	25
2.4.1.2	Replicates	28
2.4.2	RNA Extraction	28
2.4.3	Labeling/Reverse Transcription	30
2.4.4	Hybridization	31
2.4.5	Scanning	31
2.5	Data Normalization and Analysis	32
2.5.1	Image Quantification	32
2.5.2	Data Processing	33
2.5.3	Data Analysis	34
2.5.3.1	Detection of Differential Expression	34
2.5.3.2	Pattern Recognition	35
2.5.3.3	Graphical Representations	35
2.5.6	Microarray Analysis Tools	36
2.5.7	Microarray Follow-Up	36
2.5.8	Data Storage and Reanalysis	37
2.6	Transcriptomics: Where We Are Now and What's to Come	37
<b>3</b>	<b>Physiological Proteomics of <i>Bacillus subtilis</i> and <i>Staphylococcus aureus</i>: Towards a Comprehensive Understanding of Cell Physiology and Pathogenicity</b>	<b>43</b>
	<i>Michael Hecker and Susanne Engelmann</i>	
3.1	Introduction	43
3.2	Proteomics of <i>Bacillus subtilis</i> : The Gram-positive Model Organism	46
3.2.1	The Vegetative Proteome	46
3.2.2	Proteomes of Nongrowing Cells: Proteomic Signatures of Stress/Starvation Stimuli	47
3.3	Physiological Proteomics of <i>Staphylococcus aureus</i>	53
3.3.1	The Postgenome Era of <i>S. aureus</i>	53
3.3.2	Proteomes of Growing and Nongrowing Cells	56

3.3.3	Extracellular Proteins and Pathogenicity Networks	62
3.4	Outlook: Second Generation Proteomics and New Fields in <i>S. aureus</i> Physiology and Infection Biology	65
<b>4</b>	<b>Impact of Genome Sequences on Mutational Analysis of Fungal and Bacterial Pathogens</b>	<b>69</b>
	<i>Vladimir Pelicic and Xavier Nassif</i>	69
4.1	The Long Road from Sequence to Function	69
4.2	Classical Genetics Still at the Forefront in the Postgenome Era	70
4.2.1	Reverse Genetics	70
4.2.2	Transposon Mutagenesis	71
4.3	Genome-Scale Mutational Analyses	72
4.3.1	<i>Saccharomyces cerevisiae</i>	73
4.3.2	Bacterial Workhorses: <i>E. coli</i> and <i>Bacillus subtilis</i>	75
4.3.3	Bacterial Pathogens	75
4.3.3.1	<i>Mycoplasma</i> Species	76
4.3.3.2	<i>Pseudomonas aeruginosa</i>	76
4.3.3.3	<i>Staphylococcus aureus</i>	77
4.3.3.4	<i>Neisseria meningitidis</i>	77
4.4	Conclusion	79
<b>II</b>	<b>Genomics of Pathogenic Bacteria</b>	<b>83</b>
<b>5</b>	<b>Pathogenomics of <i>Escherichia coli</i> and <i>Shigella</i> Species</b>	<b>85</b>
	<i>Ulrich Dobrindt and Jörg Hacker</i>	
5.1	Introduction	85
5.2	Comparative Genomics of <i>Shigella</i>	86
5.3	Comparative Genomics of <i>Escherichia coli</i>	92
5.3.1	Comparison of Complete Genome Sequences	92
5.3.2	Comparative Genomics Using DNA Arrays	94
5.3.3	Mobile Genetic Elements and Evolution of Pathogenic <i>E. coli</i>	95
5.3.4	Genomic Islands/Pathogenicity Islands	95
5.3.5	Plasmids and Bacteriophages	99
5.3.6	Genetic Diversity Among Extraintestinal Pathogenic <i>E. coli</i>	100
5.4	Conclusions	101

<b>6</b>	<b>Pathogenomics of <i>Salmonella</i> Species</b>	<b>109</b>
	<i>Helene Andrews-Polymeris and Andreas J. Bäuml</i>	
6.1	Introduction	109
6.2	<i>Salmonella</i> Signature Genes	109
6.3	Subspecies I Signature Genes	112
6.4	Host Restriction	115
<b>7</b>	<b>Pathogenomics of <i>Enterococcus faecalis</i></b>	<b>125</b>
	<i>Janet M. Manson and Michael S. Gilmore</i>	
7.1	Introduction	125
7.2	Enterococcal Pathogenesis	125
7.3	Genome Sequence of <i>E. faecalis</i>	126
7.3.1	Mobile Elements, Acquired DNA, and Antimicrobial Resistance	127
7.3.2	Environmental Adaptation and Stress Response	131
7.3.3	Survival <i>In Vivo</i>	133
7.3.4	Potential Virulence Factors	134
7.3.4.1	Hemolysins, Proteases, and other Enzymes	134
7.3.4.2	Cell-Wall-Associated Virulence Factors	136
7.3.5	Pathogenicity Island of <i>E. faecalis</i>	138
7.4	Conclusions and Future Perspectives	140
<b>8</b>	<b>Genomics of Streptococci</b>	<b>149</b>
	<i>Joseph J. Ferretti and W. Michael McShan</i>	
8.1	Introduction	149
8.2	Bacterial Genomes	152
8.2.1	Pyogenic Group	152
8.2.1.1	<i>Streptococcus pyogenes</i>	152
8.2.1.1.1	Virulence Factors	153
8.2.1.1.2	Horizontal Gene Transfer	154
8.2.2.1	<i>Streptococcus agalactiae</i>	155
8.2.2.2	Group C (GCS) and Group G Streptococci	156
8.2.2.3	<i>Streptococcus uberis</i>	157
8.2.2	Bovis Group	157
8.2.2.1	<i>Streptococcus bovis</i> and <i>Streptococcus suis</i>	157
8.2.3	Mitis Group	158
8.2.3.1	<i>Streptococcus pneumoniae</i>	158
8.2.3.2	<i>Streptococcus mitis</i> , <i>Streptococcus sanguis</i> , and <i>Streptococcus gordonii</i>	159
8.2.4	Anginosus and Salivarius Group	159
8.2.4.1	<i>Streptococcus salivarius</i>	159
8.2.4.2	<i>Streptococcus thermophilus</i>	160
8.2.5	Mutans Group	160
8.2.5.1	<i>Streptococcus mutans</i> and <i>Streptococcus sobrinus</i>	160

8.2.6	Other Organisms: <i>Enterococcus faecalis</i>	161
8.2.7	Comparative Genomics	161
8.3	Streptococcal Genomic Bacteriophages	162
8.3.1	Prophages and Streptococcal Genomes	162
8.3.2	GAS Genome Prophages	163
8.3.2.1	Prophages and Virulence Factors	163
8.3.2.2	Prophage Attachment Sites and Host Biology	165
8.3.2.3	Prophage Diversity	166
8.3.3	Prophages Associated with other Streptococcal Species	166
<b>9</b>	<b>Pathogenic Staphylococci: Lessons from Comparative Genomics</b>	<b>175</b>
	<i>Knut Ohlsen, Martin Eckart, Christian Hüttinger, and Wilma Ziebuhr</i>	
9.1	Introduction	175
9.2	Comparative Genomics of <i>S. aureus</i>	176
9.2.1	Overall Genome Structure	177
9.2.2	Core Genome	178
9.2.2.1	Metabolism	178
9.2.2.2	Information Pathways	180
9.2.2.3	Virulence Factors	181
9.2.3	Accessory Genome	184
9.2.3.1	Pathogenicity Islands	185
9.2.3.2	Staphylococcal Cassette Chromosome	190
9.2.3.3	Bacteriophages	192
9.2.3.4	Plasmids	194
9.3	<i>Staphylococcus epidermidis</i>	195
9.3.1	Genomic Islands	197
9.3.2	Phage SP $\beta$ and other <i>Bacillus</i> Genes	197
9.3.3	Virulence Factors	197
9.3.4	Staphylococcal Cassette Chromosome	198
9.3.5	Adherence and Biofilm Formation	199
9.3.6	Insertion Sequences	201
9.4	Concluding Remarks	202
<b>10</b>	<b>Pathogenomics: Insights into Tuberculosis and Related Mycobacterial Diseases</b>	<b>211</b>
	<i>Alexander S. Pym, Stephen V. Gordon, and Roland Brosch</i>	
10.1	Introduction	211
10.2	Molecular Basis of Pathogenicity	212
10.3	Evolution of the <i>M. tuberculosis</i> Complex	216
10.4	Some Metabolic Insight from the Genome Sequences	220
10.5	Other Major Mycobacterial Human Pathogens	222
10.5.1	<i>Mycobacterium leprae</i>	222
10.5.2	<i>Mycobacterium ulcerans</i>	223
10.6	Concluding Remarks	224

11	<p><b>Genomes of Pathogenic <i>Neisseria</i> Species</b> 231  <i>Christoph Schoen, Heike Claus, Ulrich Vogel, and Matthias Frosch</i></p> <p>11.1 Introduction 231</p> <p>11.2 Genomes of Pathogenic <i>Neisseria</i> Species 232</p> <p>11.2.1 The Flexible Genome Pool 233</p> <p>11.2.2 Repetitive DNA Sequence Elements Govern Neisserial Biology 236</p> <p>11.2.2.1 DNA Uptake Sequences, Horizontal Gene Transfer, and Antigenic Diversity 236</p> <p>11.2.2.2 Simple Sequence Repeats and Phase Variation 237</p> <p>11.2.2.3 Insertion Sequences and the Regulation of Gene Expression 238</p> <p>11.2.3 Genome-Wide Mutational Analyses 239</p> <p>11.2.4 Comparative Genomics 240</p> <p>11.2.5 Novel Virulence Factors of Meningococci Identified by Genomic Approaches 244</p> <p>11.3 Future Perspectives 249</p>
12	<p><b>Genomics of Pathogenic Clostridia and Bacilli</b> 257  <i>Armin Ehrenreich, Gerhard Gottschalk, and Holger Brüggemann</i></p> <p>12.1 Genomics of Pathogenic <i>Clostridia</i> spp. 257</p> <p>12.1.1 Introduction 257</p> <p>12.1.2 <i>C. perfringens</i> 258</p> <p>12.1.3 <i>C. tetani</i> 260</p> <p>12.1.4 <i>C. botulinum</i> 262</p> <p>12.1.5 <i>C. difficile</i> 263</p> <p>12.1.6 Conclusions and Perspectives 264</p> <p>12.2 Genomics of Pathogenic Bacilli 265</p> <p>12.2.1 Introduction 265</p> <p>12.2.2 Pathogenic Properties of Bacilli not Belonging to the <i>B. cereus</i> Group 266</p> <p>12.2.3 Pathogenicity of <i>B. cereus</i> 266</p> <p>12.2.4 Pathogenicity of <i>B. anthracis</i> 268</p> <p>12.2.4.1 Course of Anthrax 268</p> <p>12.2.4.2 Virulence Factors of <i>B. anthracis</i> 269</p> <p>12.2.5 Genome of <i>B. anthracis</i> 270</p> <p>12.2.5.1 Chromosomal Genes 270</p> <p>12.2.5.2 Genes Located on Plasmids pXO1 and pXO2 271</p> <p>12.2.5.3 Regulation of Virulence Genes 272</p> <p>12.2.5.4 Molecular Diversity in <i>B. anthracis</i> Genomes 272</p> <p>12.2.5.5 Genome of a Highly Virulent <i>B. cereus</i> Strain Resembling <i>B. anthracis</i> in Pathogenesis 273</p> <p>12.2.6 Comparison of <i>B. cereus</i> Group Genomes: How Did Pathogenicity Evolve? 273</p>

<b>13</b>	<b>The Genomes of Pathogenic <i>Bartonella</i> Species</b>	<b>281</b>
	<i>Carolyn Frank, Eva Berglund, and Siv G. E. Andersson</i>	281
13.1	Introduction	281
13.1.1	<i>Bartonella</i> in a Phylogenetic Context	281
13.1.2	Hosts and Vectors for <i>Bartonella</i> Species	282
13.2	<i>Bartonella</i> Species and Pathogenicity	284
13.2.1	Infection of Reservoir and Incidental Host	284
13.2.2	<i>Bartonella</i> Species as Human Pathogens	285
13.3	The <i>Bartonella</i> Genomes	286
13.4	Genomic Islands and Phages	286
13.5	Genomic Islands and Phages in <i>Bartonella</i> Species	287
13.5.1	The <i>B. henselae</i> Prophage	288
13.5.2	<i>B. henselae</i> Genomic Islands and Islets	288
13.5.3	<i>B. quintana</i> Harbors Remnants of the <i>B. henselae</i> Islands	290
13.5.4	Role of Phages and Islands in the Evolution of <i>Bartonella</i>	290
13.6	The Chromosome II-Like Segment in <i>Bartonella</i>	291
13.6.1	Type IV Secretion Systems in <i>Bartonella</i> Species	292
13.6.2	The <i>virB-D4</i> Operon	293
13.6.3	The <i>trw</i> Operon	293
13.7	<i>B. quintana</i> 's Evolution into a Human Pathogen	294
13.8	Conclusions and Future Perspectives	295
<b>14</b>	<b>Pathogenomics of Gastric and Enterohepatic <i>Helicobacter</i> Species</b>	<b>301</b>
	<i>Sebastian Suerbaum, Sandra Schwarz, and Christine Josenhans</i>	
14.1	Introduction	301
14.2	<i>Helicobacter pylori</i>	302
14.2.1	Key Features of the <i>H. pylori</i> Genome Related to Pathogenesis	302
14.2.1.1	Colonization Factors: Urease and Motility	302
14.2.1.2	Phase Variation	302
14.2.1.3	The <i>H. pylori</i> Outer Membrane Protein Family	303
14.2.1.4	Intraspecies Variation of <i>H. pylori</i> Genomes	303
14.2.1.5	The <i>cag</i> Pathogenicity Island	304
14.2.1.6	Nucleotide Sequence Variation in <i>H. pylori</i>	305
14.3	<i>Helicobacter hepaticus</i>	305
14.3.1	The HHG11 Genomic Island	306
14.3.2	Other Putative <i>H. hepaticus</i> Virulence Factors	307
14.4	Genome Comparisons of Gastric and Enterohepatic <i>Helicobacter</i> Species with Related Bacteria	307
14.5	Outlook	308

<b>15</b>	<b>Genomics of the Opportunistic Pathogen <i>Legionella pneumophila</i></b>	<b>315</b>
	<i>Christel Cazalet and Carmen Buchrieser</i>	
15.1	The Genus <i>Legionella</i> : Epidemiology, Life Cycle, and Pathogenesis	315
15.2	Genomics of <i>Legionella pneumophila</i>	316
15.3	Specific Features of the <i>Legionella</i> Genomes	318
15.3.1	Eukaryotic-like Proteins in <i>Legionella pneumophila</i> : Modulation of Host Functions?	318
15.3.2	Secretion Machineries of <i>L. pneumophila</i> : Central to Its Life and to Pathogenesis	324
15.3.2.1	Type IV Secretion Systems in <i>Legionella</i>	325
15.3.2.1.1	The <i>dot/icm</i> Type IVB Secretion System	325
15.3.2.1.2	The <i>lvh</i> Type IVA Secretion System	326
15.3.2.2	A Putative Type I Secretion System in <i>Legionella</i>	326
15.3.2.3	A Type II Secretion System in <i>Legionella</i>	327
15.3.2.4	Secretion Across the Cytoplasmic Membrane	327
15.3.2.5	A Putative Type V Secretion System (Autotransporter) Specific to Strain Paris	328
15.3.3	Comparative Genomics: Diversity of the Species <i>L. pneumophila</i>	329
15.3.3.1	Genomic (Pathogenicity) Islands in the <i>L. pneumophila</i> Genomes	329
15.3.3.2	Plasmids and Genetic Diversity of <i>L. pneumophila</i>	331
15.4	Conclusions	333
<b>16</b>	<b>Genomics of <i>Listeria monocytogenes</i></b>	<b>339</b>
	<i>Michael Kuhn and Werner Goebel</i>	
16.1	Introduction: From Pregenomics to Postgenomics	339
16.2	<i>Listeria monocytogenes</i> : A Facultative Intracellular Pathogen	339
16.3	<i>Listeria monocytogenes</i> Genetics in the Pregenomic Era: Identification and Characterization of Important Virulence Factors	340
16.3.1	Internalins and the Invasion of Nonprofessional Phagocytic Cells	340
16.3.2	Listeriolysin O and Two Listerial Phospholipases Allow Escape from the Phagocytic Vacuole	342
16.3.3	Intracellular Motility and Cell-to-Cell Spread: The Surface Protein ActA	343
16.3.4	PrfA and the Regulation of Virulence Gene Expression	344
16.4	Genome Sequence of <i>L. monocytogenes</i> and Its Comparison with the Closely Related <i>L. innocua</i>	346
16.5	Genomic Approaches to Studying the other Members of the Genus <i>Listeria</i>	348
16.6	Evolutionary Aspects	349



16.7	Identification of Listerial Virulence Factors in the Postgenomic Era	351
16.7.1	Internalins and Other Surface Proteins	351
16.7.2	Growth in the Host Cell Cytoplasm	352
16.7.3	Resistance to Bile	353
16.7.4	Two-component Systems and the Regulation of Virulence Gene Expression	354
16.7.5	Vitamin B <sub>12</sub> Biosynthesis and Anaerobic Use of Ethanolamine	354
16.8	Proteomics	355
16.9	Transcriptomics	356
16.10	Conclusions	358
<b>III</b>	<b>Genomics of Pathogens and Their Hosts: Applications</b>	<b>367</b>
<b>17</b>	<b>Genomics of Viruses</b>	<b>369</b>
	<i>Esteban Domingo, Alejandro Brun, José Ignacio Nuñez, Juan Cristina, Carlos Briones, and Cristina Escarmis</i>	
17.1	Introduction: Wide Scope of Virogenomics	369
17.2	Retrieving Information	371
17.3	Applications of Data Banks to Virology	374
17.4	Beyond Reference Strains: Towards a Second-Generation Virogenomics?	379
17.5	Virogenomics Through Microarrays	382
<b>18</b>	<b>Genomics of Pathogenic Fungi</b>	<b>389</b>
	<i>Gerwald A. Köhler, Alan Kuo, George Newport, and Nina Agabian</i>	
18.1	Introduction	389
18.2	Genomics of Primary Fungal Pathogens	390
18.2.1	<i>Histoplasma</i>	390
18.2.2	<i>Coccidioides</i>	396
18.2.3	<i>Blastomyces</i> and <i>Paracoccidioides</i>	396
18.3	Genomics of Opportunistic Fungal Pathogens	397
18.3.1	<i>Aspergillus</i>	397
18.3.2	<i>Cryptococcus</i>	398
18.3.3	<i>Pneumocystis</i>	401
18.3.4	Microsporidia	402
18.3.5	<i>Candida</i>	403
18.4	The Tool Box for Functional Genomics	406
18.4.1	Expression Analysis	407
18.4.2	Transformation and Mutagenesis	407
18.5	Fungal Virulence – From the Genomic Point of View	410
18.6	Conclusion	411

<b>19</b>	<b>Genomics of Pathogenic Parasites</b>	<b>417</b>
	<i>Gabriele Pradel and Thomas James Templeton</i>	
19.1	Exploring the Genomes of Pathogenic Protozoans	418
19.2	The Shaping of the Proteomes of the Pathogenic Protists	420
19.3	Role of Horizontal Gene Transfer in Protozoan Genome Plasticity	421
19.4	The Apicomplexa	422
19.4.1	<i>Plasmodium</i> , the Malaria Parasite	423
19.4.2	<i>Cryptosporidium</i>	427
19.4.3	<i>Toxoplasma</i>	428
19.5	The Pathogenic Kinetoplastids	429
19.5.1	<i>Trypanosoma</i>	430
19.5.2	<i>Leishmania</i>	431
19.6	The Pathogenic Diplomonad <i>Giardia</i> and the Parabasalid <i>Trichomonas</i>	431
19.7	Postgenomic Strategies and the Search for Cure	433
19.7.1	Gene Expression Analysis	434
19.7.2	Proteomics	434
19.7.3	Drug and Vaccine Development	435
19.7.4	Vector Genetics	436
<b>20</b>	<b>Model Host Systems: Tools for Comprehensive Analysis of Host–Pathogen Interactions</b>	<b>445</b>
	<i>Michael Steinert and Gernot Glöckner</i>	
20.1	Introduction	445
20.2	Host–Pathogen Interactions	446
20.3	<i>Arabidopsis thaliana</i> : A Plant as a Model for Human Disease	447
20.4	<i>Dictyostelium discoideum</i> : Perspectives from a Social Amoeba	448
20.5	<i>Caenorhabditis elegans</i> : Answers from a Worm	449
20.6	<i>Drosophila melanogaster</i> : A Fruitful Model	450
20.7	<i>Danio rerio</i> : Fishing for Knowledge	451
20.8	<i>Mus musculus</i> : Of Mice and Men	451
20.9	Clean Models and Dirty Reality	453
<b>21</b>	<b>Expression Analysis of Human Genes During Infection</b>	<b>457</b>
	<i>Erwin Bohn and Ingo B. Autenrieth</i>	
21.1	Introduction	457
21.2	Comparison of Gene Expression Profiles of Macrophages and Dendritic Cells In Vitro Upon Infection with Different Pathogens	458
21.3	Septicemia	460
21.4	Gene Expression in Epithelial Cells Modulated by Bacteria	461
21.4.1	<i>Helicobacter pylori</i>	462

21.4.2	<i>Yersinia enterocolitica</i>	464
21.4.3	<i>Pseudomonas aeruginosa</i>	466
21.4.4	<i>Bartonella henselae</i>	468
21.5	Common Signatures	469
21.6	Genetic Polymorphisms and Mutations Affect Gene Expression: Impact on Infection Susceptibility and Infection Course	471
21.7	Concluding Remarks	473
<b>22</b>	<b>Pathogenomics: Application and New Diagnostic Tools</b>	<b>481</b>
	<i>Sören Schubert and Jürgen Heesemann</i>	
22.1	Introduction: “In Our Hands”	481
22.2	Microbiological Diagnostics of Bacterial Pathogens: Aims, Tasks, and Current Limitations	482
22.3	The Pregenomic Era: Conventional and Molecular Methods in Microbiological Diagnostics	483
22.3.1	Conventional Culture-Based Methods in Microbiological Diagnostics	483
22.3.2	Molecular Microbiological Diagnostic Methods	484
22.3.2.1	Typing of Bacterial Isolates Using 16S-rRNA	484
22.3.2.2	Fluorescence <i>In Situ</i> Hybridization	484
22.3.2.3	PCR Methods for Microbial Diagnostics	485
22.4	The Postgenomic Era: Use of DNA Microarrays in the Diagnosis of Infectious Diseases in Humans and Animals	487
22.4.1	DNA Arrays: Platforms, Techniques and Targets	487
22.4.2	Detection and Typing of Microbial Pathogens	488
22.4.3	Pathoarrays	489
22.4.4	16S-/23S-rDNA Arrays	494
22.4.5	Detection of Antibiotic Resistance in Microbial Pathogens Using Microarray Technology	494
22.5	Microarray Technology in Bacteria: Further Areas of Applications	495
22.5.1	Gene Expression Microarrays and Host–Pathogen Interaction	495
22.5.2	DNA Microarray Technology in Food Technology	496
22.5.3	DNA Microarray Technology in Environmental Microbiology	496
22.5.4	Pathogenomic Tools (Microarrays) in the Diagnosis of Microbiologic Agents as Bioweapons	497
22.6	Current Limitations on the Use of DNA Microarrays in Diagnostics in Medical Microbiological Laboratories	498
22.7	Final Remarks	498

<b>23</b>	<b>The Search for New Antibiotics</b>	<b>505</b>
	<i>Harald Labischinski, Christoph Freiberg, and Heike Brötz-Oesterhelt</i>	
23.1	The Need for Novel Antibiotics	505
23.2	Where Will the New Antibiotics Come From?	507
23.2.1	The Past	507
23.2.2	The Present	509
23.2.3	Future Directions	511
23.3	Contributions of Genomic Technologies to Antibacterial Research	513
23.3.1	Target Identification and Validation	513
23.3.2	Target Prioritization	517
23.3.3	Genetic Tools for Drug Screening and Mode-of-Action Determination	517
23.3.4	Genome-Wide Expression Profiling for Mode-of-Action Characterization	520
23.3.5	Outlook for Genomic Technologies for Antibiotic Drug Discovery	521
23.4	Alternative Approaches in Antibacterial Drug Discovery	521
23.4.1	Targeting the Resistance Mechanism	522
23.4.2	Extremely Narrow-Spectrum Drugs	523
23.4.3	Phage Therapies and other Bacteriolytic Approaches	524
23.4.4	Strategies for Reducing Virulence and/or Influencing Pathogenesis	525
<b>24</b>	<b>Reverse Vaccinology: Revolutionizing the Approach to Vaccine Design</b>	<b>533</b>
	<i>Laura Serino, Mariagrazia Pizza, and Rino Rappuoli</i>	
24.1	Impact of Genomics on Vaccine Design	533
24.2	MenB Vaccine Approach by Reverse Vaccinology	535
24.3	Following the MenB Experience: Other Pathogens	538
24.4	Functional Genomics	539
24.5	Gene Expression <i>In Vivo</i> : IVET and STM	540
24.6	Transcriptome Analysis and Comparative Genomics	541
24.7	Proteomics and Vaccine Design	546
24.8	Conclusions	546
	<b>Index</b>	<b>555</b>

## Preface

In the year 1995, the first full genome sequence of a free-living-organism, the bacterium *Haemophilus influenzae* strain Rd, was published. This publication, which appeared in the journal *Science*, represented the starting point for a new field in molecular biology called genomics. Today, 10 years later, complete genome sequences of almost all the major pathogenic microbes have been determined. As a consequence, a new discipline has arisen, which has been named “pathogenomics.” As the name implies, pathogenomics is the analysis at the genomic level of the processes involved in bacterial pathogenesis caused by the interaction of pathogenic microbes and their hosts.

The present volume is the first handbook to be entirely devoted to the newly established discipline of pathogenomics. We are very grateful to our colleagues for their input, especially those associated with the German Pathogenomics competence network, established by the German Federal Ministry of Science and Education to analyze pathogenic microbes at the genomic level. Werner Goebel in particular – the network’s speaker, who contributed the Foreword to this book’s preface – has influenced the entire discipline with his spirit and his vision.

Our thanks are due to the staff at Wiley-VCH, most notably Andrea Pillmann, who encouraged us to put this book together. We are also very grateful to all our authors for their contributions to this important project.

Würzburg, September 2005

*Ulrich Dobrindt, Jörg Hacker*



## List of Contributors

### ***Nina Agabian***

University of California San Francisco  
Department of Cell and Tissue Biology  
521 Parnassus, Box 0640  
San Francisco, CA 94143-0640  
USA

### ***Siv Andersson***

Uppsala University  
Evolutionary Biology Center  
Department of Molecular Evolution  
Norbyvägen 18C  
752 36 Uppsala University  
Sweden

### ***Helene Andrews-Polymeris***

Department of Medical Microbiology  
and Immunology  
College of Medicine  
Texas A&M University HSC  
407 Reynolds Medical Building  
College Station, TX 77843-1114  
USA

### ***Ingo B. Autenrieth***

Universitätsklinikum Tübingen  
Institut für Medizinische Mikrobiologie  
und Hygiene  
Elfriede-Aulhorn-Str. 6  
72076 Tübingen  
Germany

### ***Andreas Bäumlér***

Department of Medical Microbiology  
and Immunology  
School of Medicine  
University of California at Davis  
One Shields Ave.  
Davis, CA 95616-8645  
USA

### ***Eva Berglund***

Uppsala University  
Evolutionary Biology Center  
Department of Molecular Evolution  
Norbyvägen 18C  
752 36 Uppsala University  
Sweden

### ***Erwin Bohn***

Universitätsklinikum Tübingen  
Institut für Medizinische Mikrobiologie  
und Hygiene  
Elfriede-Aulhorn-Str. 6  
72076 Tübingen  
Germany

### ***Carlos Briones***

Centro de Astrobiología (CSIC-INTA)  
Carretera de Ajalvir, Km. 4  
Torrejón de Ardoz  
28850 Madrid  
Spain

**Roland Brosch**

Institut Pasteur  
Unité de Génétique Moléculaire  
Bactérienne  
28 Rue du Dr. Roux  
75724 Paris Cedex 15  
France

**Heike Brötz-Oesterhelt**

Bayer Healthcare AG  
Pharma Research Center  
Anti-Infective Research  
Gebäude 405  
42096 Wuppertal  
Germany

**Holger Brüggemann**

Institut Pasteur  
Unité de Génomique des  
Microorganismes Pathogènes (GMP)  
28, Rue du Dr Roux  
75724 Paris, Cedex 15  
France

**Alejandro Brun**

CISA-INIA  
Valdeolmos  
28130 Madrid  
Spain

**Carmen Buchrieser**

Laboratoire de Génomique des  
Microorganismes Pathogènes  
Institut Pasteur  
25 Rue du Dr. Roux  
75724 Paris, Cedex 15  
France

**Christel Cazalet**

Laboratoire de Génomique des  
Microorganismes Pathogènes  
Institut Pasteur  
25 Rue du Dr. Roux  
75724 Paris, Cedex 15  
France

**Heike Claus**

Institut für Hygiene und Mikrobiologie  
Universität Würzburg  
Josef-Schneider-Str. 2  
97080 Würzburg  
Germany

**Juan Cristina**

Centro de Biología Molecular  
“Severo Ochoa” (CSIC-UAM)  
Universidad Autónoma de Madrid  
Cantoblanco  
28049 Madrid  
Spain

**Thomas Dandekar**

University of Würzburg  
Chair of Bioinformatics  
Biocenter  
Am Hubland  
97074 Würzburg  
Germany

**Ulrich Dobrindt**

Institut für Molekulare  
Infektionsbiologie  
Universität Würzburg  
Röntgenring 11  
97070 Würzburg  
Germany

**Esteban Domingo**

Centro de Biología Molecular  
“Severo Ochoa” (CSIC-UAM)  
Universidad Autónoma de Madrid  
Cantoblanco  
28049 Madrid  
Spain

**Martin Eckart**

Institut für Molekulare  
Infektionsbiologie  
Röntgenring 11  
97070 Würzburg  
Germany



**Armin Ehrenreich**

Göttingen Genomics Laboratory  
Georg-August-Universität Göttingen  
Grisebachstraße 8  
37077 Göttingen  
Germany

**Susanne Engelmann**

Ernst-Moritz-Arndt-Universität  
Institut für Mikrobiologie  
Friedrich-Ludwig-Jahn-Straße 15  
17487 Greifswald  
Germany

**Cristina Escarmís**

Centro de Biología Molecular  
“Severo Ochoa” (CSIC-UAM)  
Universidad Autónoma de Madrid  
Cantoblanco  
28049 Madrid  
Spain

**Joseph J. Ferretti**

Department of Microbiology and  
Immunology  
University of Oklahoma Health  
Sciences Center  
Oklahoma City, OK 73104  
USA

**Carolin Frank**

Uppsala University  
Evolutionary Biology Center  
Department of Molecular Evolution  
Norbyvägen 18C  
752 36 Uppsala University  
Sweden

**Christoph Freiberg**

Bayer Healthcare AG  
Pharma Research Center  
Anti-Infective Research  
Gebäude 405  
42096 Wuppertal  
Germany

**Matthias Frosch**

Institut für Hygiene und Mikrobiologie  
Universität Würzburg  
Josef-Schneider-Str. 2  
97080 Würzburg  
Germany

**Michael S. Gilmore**

The Schepens Eye Research Institute  
and Harvard Medical School  
Department of Ophthalmology  
20 Staniford St.  
Boston, MA 02114  
USA

**Gernot Glöckner**

Leibnitz-Institut für Altersforschung  
Fritz-Lipmann-Institut  
Beutenbergstr. 11  
07745 Jena  
Germany

**Werner Goebel**

University of Würzburg  
Theodor-Boveri-Institute for  
Biological Sciences  
Chair of Microbiology  
Am Hubland  
97074 Würzburg  
Germany

**Stephen V. Gordon**

TB Research Group  
Veterinary Laboratories Agency  
Weybridge  
Woodham Lane, New Haw  
Addlestone, Surrey  
KT15 3NB  
United Kingdom

**Gerhard Gottschalk**

Göttingen Genomics Laboratory  
Georg-August-Universität Göttingen  
Grisebachstraße 8  
37077 Göttingen  
Germany

**Jörg Hacker**

Institut für Molekulare  
Infektionsbiologie  
Universität Würzburg  
Röntgenring 11  
97070 Würzburg  
Germany

**Michael Hecker**

Ernst-Moritz-Arndt-Universität  
Institut für Mikrobiologie  
Friedrich-Ludwig-Jahn-Straße 15  
17487 Greifswald  
Germany

**Jürgen Heesemann**

Max-von-Pettenkofer-Institut für  
Hygiene und Medizinische  
Mikrobiologie  
Universität München  
Pettenkoferstraße 9a  
80336 München  
Germany

**Christian Hüttinger**

Institut für Molekulare  
Infektionsbiologie  
Röntgenring 11  
97070 Würzburg  
Germany

**Christine Josenhans**

Hannover Medical School  
Institute for Medical Microbiology  
OE5210  
Carl-Neuberg-Straße 1  
30625 Hannover  
Germany

**Susanne Kneitz**

IZKF, Labor für Microarray-  
Anwendungen  
Institut für Virologie und  
Immunbiologie  
97078 Würzburg  
Germany

**Gerwald A. Köhler**

University of California San Francisco  
Department of Cell and Tissue Biology  
521 Parnassus, Box 0640  
San Francisco, CA 94143-0640  
USA

**Michael Kuhn**

University of Würzburg  
Competence Network PathoGenoMik  
Biocenter  
Am Hubland  
97074 Würzburg  
Germany

**Alan Kuo**

University of California San Francisco  
Department of Cell and Tissue Biology  
521 Parnassus, Box 0640  
San Francisco, CA 94143-0640  
USA

**Harald Labischinski**

Combinature Biopharm AG  
Robert-Rössle-Str. 10,  
Gebäude 80  
13125 Berlin  
Germany

**Janet M. Manson**

The Schepens Eye Research Institute  
and Harvard Medical School  
Department of Ophthalmology  
20 Staniford St.  
Boston, MA 02114  
USA

**W. Michael McShan**

Department of Pharmaceutical Sciences  
P.O. Box 26901, CPB 307  
University of Oklahoma Health  
Sciences Center  
Oklahoma City, OK 73190  
USA

**Xavier Nassif**

INSERM U570  
Faculté René Descartes Paris 5  
Site Necker-Enfants Malades  
156 rue de Vaugirard  
75730 Paris Cedex 15  
France

**George Newport**

University of California San Francisco  
Department of Cell and Tissue Biology  
521 Parnassus, Box 0640  
San Francisco, CA 94143-0640  
USA

**José Ignacio Nuñez**

CISA-INIA  
Valdeolmos  
28130 Madrid  
Spain

**Knut Ohlsen**

Institut für Molekulare  
Infektionsbiologie  
Röntgenring 11  
97070 Würzburg  
Germany

**Vladimir Pelicic**

INSERM U570  
Faculté René Descartes Paris 5  
Site Necker-Enfants Malades  
156 rue de Vaugirard  
75730 Paris Cedex 15  
France

**Mariagrazia Pizza**

Chiron Vaccines  
Via Fiorentina, 1  
53100 Siena  
Italy

**Gabriele Pradel**

University of Würzburg  
Research Center for Infections Diseases  
Röntgenring 11  
97070 Würzburg  
Germany

**Alexander S. Pym**

Unit for Clinical and Biomedical TB  
Research  
South African MRC  
491 Ridge Road  
PO Box 70380  
Overport, 4067  
Durban  
South Africa

**Rino Rappuoli**

Chiron Vaccines  
Via Fiorentina, 1  
53100 Siena  
Italy

**Christoph Schoen**

Institut für Hygiene und Mikrobiologie  
Universität Würzburg  
Josef-Schneider-Str. 2  
97080 Würzburg  
Germany

**Sören Schubert**

Max von Pettenkofer-Institut für  
Hygiene und Medizinische  
Mikrobiologie  
Universität München  
Pettenkoferstr. 9a  
80336 München  
Germany

**Sandra Schwarz**

Institute for Medical Microbiology  
OE5210  
Hannover Medical School  
Carl-Neuberg-Straße 1  
30625 Hannover  
Germany

**Laura Serino**

Chiron Vaccines  
Via Fiorentina, 1  
53100 Siena  
Italy

**Ben Sidders**

The Royal Veterinary College  
Pathology and Infectious Diseases  
Royal College Street  
London, NW1 0TU  
United Kingdom

**Michael Steinert**

Universität Würzburg  
Institut für Molekulare  
Infektionsbiologie  
Röntgenring 11  
97070 Würzburg  
Germany

**Neil Stoker**

The Royal Veterinary College  
Pathology and Infectious Diseases  
Royal College Street  
London, NW1 0TU  
United Kingdom

**Sebastian Suerbaum**

Hannover Medical School  
Institute for Medical Microbiology  
OE5210  
Carl-Neuberg-Straße 1  
30625 Hannover  
Germany

**Thomas J. Templeton**

Weill Medical College of Cornell  
University  
Department of Microbiology and  
Immunology  
1300 York Avenue, Box 62  
New York, NY 10021  
USA

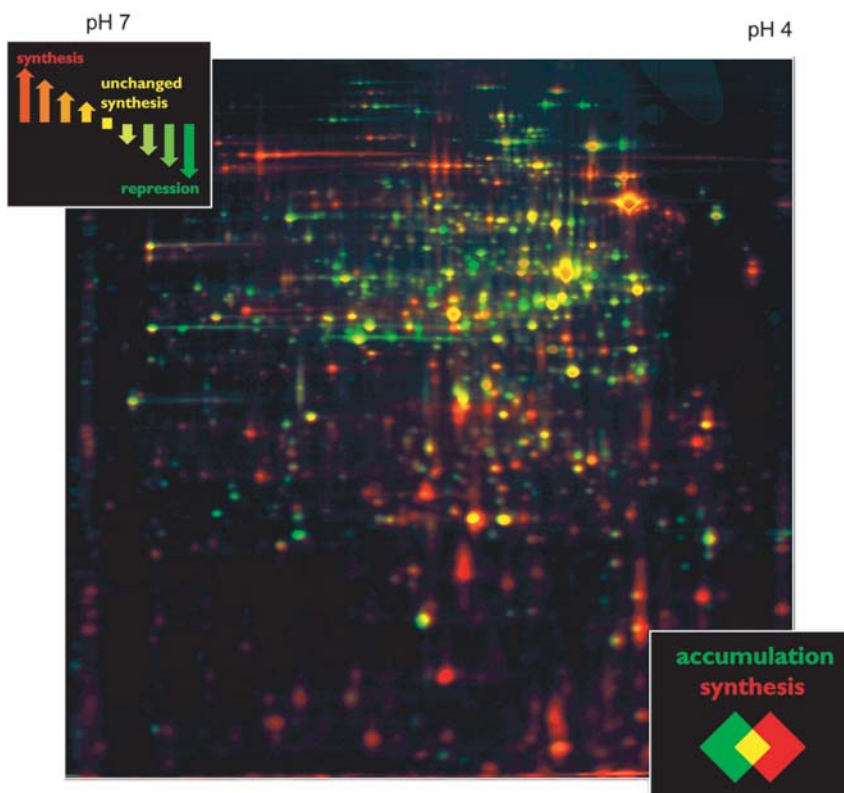
**Ulrich Vogel**

Institut für Hygiene und Mikrobiologie  
Universität Würzburg  
Josef-Schneider-Str. 2  
97080 Würzburg  
Germany

**Wilma Ziebuhr**

Institut für Molekulare  
Infektionsbiologie  
Röntgenring 11  
97070 Würzburg  
Germany

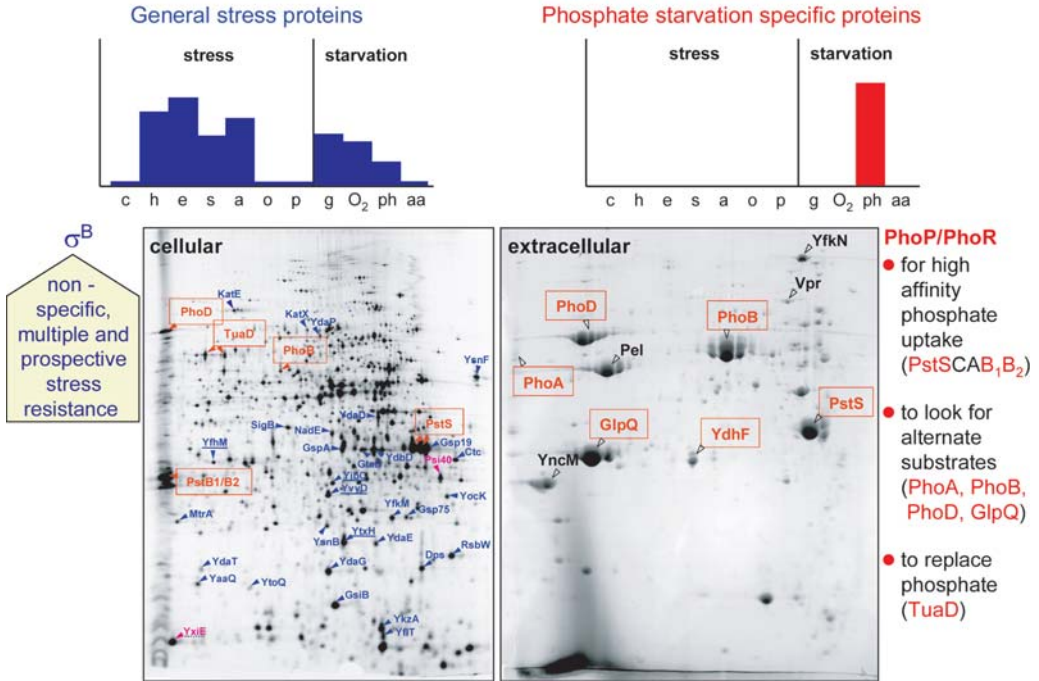
## Color Plates



**Fig. 3.2** Protein pattern of heat-shocked *B. subtilis* cells. During exposure to heat shock (37–to 48 °C), *B. subtilis* cells were pulse-labeled with  $^{35}\text{S}$ -L-methionine. After the separation of crude protein extract by 2-D gels, the proteins were visualized by silver staining and the protein synthesis rate was

determined by phosphoimaging. The two images were overlaid with the aid of the Delta-2D software package (dual-channel imaging technique [11]; Decodon, Greifswald, Germany). The red-labeled proteins form the heat stress stimulon. (This figure also appears on page 48.)

**Phosphate starvation stimulon in *B. subtilis***



**Fig. 3.4** The phosphate starvation stimulon of *B. subtilis* consists of both general stress proteins also induced by phosphate starvation ( $\sigma^B$  regulon,  $\sigma^B$ -dependent proteins marked in blue) and starvation-specific proteins (PhoR regulon, marked in red). The stress/starvation induction profile of both groups and the functions of the proteins

are indicated. C, control; h, heat stress; e, ethanol stress; s, NaCl stress; a, pH 5.5; o, oxidative stress (H<sub>2</sub>O<sub>2</sub>); p, puromycin; g, glucose starvation; O<sub>2</sub>, oxygen starvation; ph, phosphate starvation; aa, amino acids starvation. (This figure also appears on page 50.)