

Contents

List of Figures	ix
Acknowledgements	xiii
1. Introduction: Bacterial Genomes and Gene Expression	1
2. Comparative Genomics in the Era of Sanger Sequencing	4
2.1 Introduction	4
2.2 The process of assembling and annotating bacterial genomes	6
2.2.1 Genome assembly and gap closure	7
2.2.2 Genome-scale computational identification of features	8
2.2.3 Annotating genes with functions	13
2.3 Case studies	16
2.3.1 The <i>Escherichia coli</i> complex and large-scale horizontal gene acquisition	16
2.3.2 Genome reduction in intracellular pathogens, endosymbionts and marine α -proteobacteria	20
2.3.3 The dynamic genomes of <i>Helicobacter pylori</i> and <i>Campylobacter jejuni</i>	26
2.4 Some lessons learnt from studying 2,000 bacterial genomes	27
2.4.1 Genome size	28
2.4.2 Coding density	29
2.4.3 Gene order conservation	30
2.4.4 Comparative genomics of gene functions: Systematic annotation	31
2.4.5 Comparative genomics of gene functions: Scaling laws	33
Summary	35
3. Studying Bacterial Genome Variation with Microarrays	36
3.1 Introduction	36
3.2 DNA microarrays: The concept	36

3.3	DNA microarrays: From fluorescence intensities to information	38
3.3.1	Background correction	39
3.3.2	Normalisation	41
3.3.3	Differences in signal from the same probe between two samples	44
3.4	Comparative genome hybridisation and bacterial phylogenomics	45
3.5	Case studies	46
3.5.1	Comparative genome hybridisation studies of <i>Escherichia coli</i>	47
3.5.2	Comparative genome hybridisation studies of <i>Staphylococcus aureus</i>	52
3.5.3	Comparative genome hybridisation studies of <i>Helicobacter pylori</i>	56
	Summary	59
4.	Studying Bacterial Genomes using Next-Generation Sequencing	60
4.1	Introduction	60
4.2	Next-generation sequencing technologies	61
4.2.1	Template preparation strategies	62
4.2.2	Sequencing strategies	64
4.3	Sequencing data processing for genome sequencing and re-sequencing	66
4.3.1	Genome assembly	66
4.3.2	Aligning short reads to long genomes	70
4.4	Case studies	73
4.4.1	Pyrosequencing-enabled complete genome sequence of <i>Acinetobacter baumannii</i>	74
4.4.2	On the track of pandemics: The genome of the aetiological agent of Black Death	75
4.4.3	From community genomes to complete genomes to single-cell genomes	80
4.4.4	Bacteria evolving in the laboratory	84
4.4.5	Bacteria evolving in their biotic hosts	86
	Summary	88

5	Genome-Scale Analysis of Gene Expression and its Regulation in Bacteria	89
5.1	Introduction	89
5.2	The process of transcription and the regulation of its initiation: An overview	90
5.3	Measuring gene expression on a genomic scale: Technologies	99
5.4	Next-generation sequencing for gene expression measurements: Data analysis	101
5.4.1	Transcriptome assembly	104
5.4.2	Measuring gene expression levels	110
5.5	Gene expression at high temporal resolution using fluorescent reporters	114
5.6	Constructing transcriptional regulatory networks: ChIP-chip and ChIP-seq	117
5.7	Case studies	122
5.7.1	Experimental annotation of bacterial genomes	122
5.7.2	Bioinformatic analysis of bacterial promoters	129
5.7.3	DNA topology and its interplay with gene expression	134
5.7.4	RNA polymerase occupancy and the σ -factors	145
5.7.5	Transcription factors and transcriptional regulatory networks	154
5.7.6	Transcriptional control by the small-molecule alarmone ppGpp	186
5.7.7	RNA chaperones and their regulons	188
	Summary	190
6	DNA Methylation in Bacteria: A Case for Bacterial Epigenetics	191
6.1	Introduction	191
6.2	DNA methyltransferases in bacteria: From restriction-modification systems	192
6.3	Identifying sites of DNA methylation on a genomic scale	195
6.3.1	Methylated DNA immunoprecipitation	195
6.3.2	Bisulphite sequencing	195
6.3.3	DNA cytosine methylation in laboratory <i>E. coli</i>	197

200	6.4 Detecting DNA methylation by single-molecule real-time sequencing
202	6.4.1 DNA adenine methylation in pathogenic <i>E. coli</i> by SMRT sequencing
203	6.4.2 Insight into the epigenetic control of <i>Caulobacter crescentus</i> cell cycle from SMRT sequencing
205	Summary
207	Index
208	6.5 DNA Methylation in Bacteria: A Case for Bacterial Epigenetics
209	6.5.1 Introduction
210	6.5.2 DNA methylation in bacteria: from restriction modification systems to large-scale epigenetic control
211	6.5.3 Identifying sites of DNA methylation on a genomic scale
212	6.5.4 Methylated DNA immunoprecipitation
213	6.5.5 Bisulfite sequencing in bacteria
214	6.5.6 DNA methylation in <i>Escherichia coli</i>
215	6.5.7 RNA chaperones and their regions
216	6.5.8 RNA polymerase occupancy and the σ -factor
217	6.5.9 DNA topology and its interplay with gene expression
218	6.5.10 Bioinformatic analysis of bacterial genomes
219	6.5.11 Experimental annotation of bacterial genomes
220	6.5.12 Case studies
221	6.5.13 Summary
222	6.6 Case studies
223	6.6.1 Case studies
224	6.6.2 Case studies
225	6.6.3 Case studies
226	6.6.4 Case studies
227	6.6.5 Case studies
228	6.6.6 Case studies
229	6.6.7 Case studies
230	6.6.8 Case studies
231	6.6.9 Case studies
232	6.6.10 Case studies
233	6.6.11 Case studies
234	6.6.12 Case studies
235	6.6.13 Case studies
236	6.6.14 Case studies
237	6.6.15 Case studies
238	6.6.16 Case studies
239	6.6.17 Case studies
240	6.6.18 Case studies
241	6.6.19 Case studies
242	6.6.20 Case studies
243	6.6.21 Case studies
244	6.6.22 Case studies
245	6.6.23 Case studies
246	6.6.24 Case studies
247	6.6.25 Case studies
248	6.6.26 Case studies
249	6.6.27 Case studies
250	6.6.28 Case studies
251	6.6.29 Case studies
252	6.6.30 Case studies
253	6.6.31 Case studies
254	6.6.32 Case studies
255	6.6.33 Case studies
256	6.6.34 Case studies
257	6.6.35 Case studies
258	6.6.36 Case studies
259	6.6.37 Case studies
260	6.6.38 Case studies
261	6.6.39 Case studies
262	6.6.40 Case studies
263	6.6.41 Case studies
264	6.6.42 Case studies
265	6.6.43 Case studies
266	6.6.44 Case studies
267	6.6.45 Case studies
268	6.6.46 Case studies
269	6.6.47 Case studies
270	6.6.48 Case studies
271	6.6.49 Case studies
272	6.6.50 Case studies
273	6.6.51 Case studies
274	6.6.52 Case studies
275	6.6.53 Case studies
276	6.6.54 Case studies
277	6.6.55 Case studies
278	6.6.56 Case studies
279	6.6.57 Case studies
280	6.6.58 Case studies
281	6.6.59 Case studies
282	6.6.60 Case studies
283	6.6.61 Case studies
284	6.6.62 Case studies
285	6.6.63 Case studies
286	6.6.64 Case studies
287	6.6.65 Case studies
288	6.6.66 Case studies
289	6.6.67 Case studies
290	6.6.68 Case studies
291	6.6.69 Case studies
292	6.6.70 Case studies
293	6.6.71 Case studies
294	6.6.72 Case studies
295	6.6.73 Case studies
296	6.6.74 Case studies
297	6.6.75 Case studies
298	6.6.76 Case studies
299	6.6.77 Case studies
300	6.6.78 Case studies
301	6.6.79 Case studies
302	6.6.80 Case studies
303	6.6.81 Case studies
304	6.6.82 Case studies
305	6.6.83 Case studies
306	6.6.84 Case studies
307	6.6.85 Case studies
308	6.6.86 Case studies
309	6.6.87 Case studies
310	6.6.88 Case studies
311	6.6.89 Case studies
312	6.6.90 Case studies
313	6.6.91 Case studies
314	6.6.92 Case studies
315	6.6.93 Case studies
316	6.6.94 Case studies
317	6.6.95 Case studies
318	6.6.96 Case studies
319	6.6.97 Case studies
320	6.6.98 Case studies
321	6.6.99 Case studies
322	6.6.100 Case studies