2

Evolution and Origin of Virulence Isolates

Voon Loong Chan, Philip M. Sherman, and Billy Bourke

Summary

Perhaps the most significant benefit of microbial genomic sequences is the knowledge gained on the molecular process of genome evolution in microbes derived from comparative genomic analysis. Genetic variations are the driving forces of evolution. These are generated not only through base substitutions, small deletions and insertions, major DNA rearrangements and deletions, but also through DNA acquisition by horizontal gene transfer. Pathogens are evolved from diverse bacterial species. The molecular mechanisms involved are diverse, and likely affected by the conditions of the microenvironment inhabited by the evolving bacterial species.

Key Words: Evolution; virulence; mutations; horizontal gene transfer.

1. Introduction

As a result of the ever increasing number of genomes being sequenced, we have a much better understanding of the molecular processes of evolution of prokaryotes, especially bacteria. Comparative genomic analysis between virulent and nonvirulent strains of the same species provide valuable insight regarding mechanisms of evolution of virulence, whereas comparative analysis of genomes of closely related species provide a much better understanding of how new species have evolved.

Genetic variations (mutations) by generating a large pool of diverse variants for natural selection are the driving forces of evolution. Such genetic variations can be generated through local base changes, deletions, and acquisition of new DNA through horizontal transfer and DNA rearrangements. With constant change in environmental conditions, some new variant strains may prove to be more fit for a new microenvironment and, thus, survive, replicate, and become the predominant clone.

In this chapter, the key elements of biological evolution are discussed and some critical information gained from comparative analyses of sequenced bacterial genomes on the evolution of virulent strains and species are presented.

2. Genetic Variation

A significant source of mutant genes and mutants is small mutations like base substitutions, addition, or deletion of a small number of bases (resulting in frameshift mutations). Regions of the genome with homopolymeric tracts are "hot spots" for frameshift mutations. These mutations could occur during DNA replication or repair. The rate of infidelity of DNA polymerases during DNA replication or repair is usually low, but it could be greatly enhanced by the presence of mutagens, generated either endogenously

From: *Bacterial Genomes and Infectious Diseases*
Edited by: V. L. Chan, P. M. Sherman, and B. Bourke © Humana Press Inc., Totowa, NJ

(e.g., free radicals generated through oxidative respiration), or from the environment. This source of mutations is paramount for those bacteria that have little opportunity to acquire new genes through horizontal gene transfer (HGT).

Mutations can also be generated through DNA rearrangement caused by intrachromosomal homologous recombination between repeated sequences, e.g., insertion sequence (IS) elements and *rrn* operons. An inversion results when the repeated sequences involved in the recombination are in reverse orientation on the chromosome. Deletions and tandem duplications are generated when the repeated sequences are in the same, or tandem, orientation. DNA rearrangement can generate a new gene through fusion of distantly linked domains. Some bacterial genomes have a great abundance of IS elements, while others have few or none (*see* Chapter 1). For example, *Bordetella pertussis* and *Bordetella parapertussis* genomes have 261 and 112 copies of different IS elements respectively, whereas *Bordetella bronchiseptica* RB50 strain has none *(1)*. *B. pertussis* and *B. parapertussis* are believed to have evolved from different *B. bronchiseptica*-like ancestors. The acquisition of IS elements and their proliferation in *B. pertussis* and *B. parapertussis* is a major contributor to the generation of a large number of pseudogenes in these two bacterial genomes. This process likely has played an important role in the evolution of these two species.

Perhaps the most important strategy for evolving new or variant genes is the acquisition of genetic material from distantly or closely related species by HGT. These acquired genes, which have evolved in the donor organisms, are likely functional and may provide enhanced ability for the recipient cells to survive and grow in new niches (e.g., in the presence of antibiotics, heavy metals, or iron-limitation). Certain bacteria are naturally competent and can acquire these DNA fragments through transformation. HGT can also be mediated by a virus (transduction) or a conjugative plasmid or conjugative transposon (conjugation). Analyses of bacterial genomes sequenced to date indicate a huge diversity in the amount of HGT genes that are accumulated ranging from 0% for some small genomes (*Rickettsia prowazekii, Borrelia burgdorferi* and *Mycoplasma genitalium*) to 17% in *Synechocystis* PCC6803, and *Escherichia coli* K12 with 12.8% *(2)*.

3. Evolution Genes

Arber *(3)* postulated the existence of "evolution" genes. The products of these genes function to promote genetic variation or moderate the frequency of genetic variation for the benefit of the evolution of the micro-organism population. Genes involved in site-specific recombination are examples of evolution genes. Site-specific recombination systems are widespread in bacteria. In these systems, recombinases catalyze the inversion of DNA fragments that are flanked by consensus sequences. Genomes containing secondary consensus sequences at diverse sites provide the opportunity for a low level of aberrant novel inversion events, thereby leading to a pool of genetic variants. Transposase genes of IS elements are another example of evolution genes that promote genetic variation (*see* Chapter 1). Mismatched base pairs generated during DNA replication or DNA repair would also result in mutations if not corrected before cell division. The newly inserted noncomplementary nucleotide is generally removed by the methylation-dependent mismatch DNA repair enzymes *(4)*. Other DNA lesions can be repaired by base or nucleotide excision repair enzymes. These DNA repair genes are evolution genes that modulate the frequency of genetic variation.

Restriction-modification (RM) systems are composed of genes that encode a DNA restriction endonuclease and a modification methylase. A type II restriction enzyme recognizes a specific DNA sequence and cleaves it if the sequence is not methylated. The associated DNA methylase recognizes the same sequence and methylates a key residue within the specific sequence, thereby preventing cleavage of host DNA by its own restriction enzyme. Thus, RM systems are believed to play a role in reducing the uptake of foreign DNA to a tolerable level. However, DNA fragments which have been cleaved by restriction enzymes often have cohesive ends, and these DNA fragments are recombinogenic, thus facilitating recombination with host DNA. In this way RM systems can have a dual role in the evolution of prokaryotic genomes.

Analyses of sequenced prokaryotic genomes show that RM systems are widespread and some genomes have an abundance of RM genes. For example, 23 DNA methyltransferase genes were identified in the sequenced genome of *H. pylori* 26695 with the RM genes amounting to more than 1% of the genome *(5)*. Typically, genes encoding restriction enzymes are clustered with their cognate methyltransferase genes *(6)*. DNA methyltransferase genes are much easier to identify compared with genes for restriction enzymes because the former contain coding sequences for many characteristic protein sequence motifs *(7)*. Some RM gene complexes were shown to behave as selfish gene entities *(8)*. Close linkage of the restriction and modification genes of a RM genetic element favors mobility by HGT independent of the host chromosome. Furthermore, host cells that have lost the RM genetic element are killed because some of the target sites on the chromosome are no longer modified by the labile methylase and thus subjected to degradation by the stable restriction enzyme. There is abundant evidence that RM genetic elements are extensively transferred horizontally between distantly related genomes *(9)*. These genes are often linked with mobile genetic elements, including plasmids, prophages, transposons, and integrons.

Arber *(3)* suggested that bacterial viruses and plasmids also should be considered primarily as genetic elements with evolutionary functions because they are natural gene vectors and are involved in DNA acquisition. These elements promote genetic variation of host cells within a population and, therefore, enhance evolution.

4. Origin of Virulence Strains and Species

The first bacteria on this planet probably evolved about 3.8 billion yr ago. Virulence factors started to evolve about 1 billion years ago during the coevolution of bacteria and unicellular eukaryotic organisms *(10)*. Virulence factors are commonly defined as gene products that facilitate the bacterial interaction, subversion, and destruction of host cells, or the neutralization of host defense mechanisms. Bacterial toxins that are specific for highly conserved proteins like heterotrimeric G proteins, small G proteins, and actin are likely to have evolved in early time. However, some virulence factors evolved recently (since the arrival of higher eukaryotes). Other virulence factors of human-restricted bacterial pathogens like *Streptococcus pyogenes, Shigella* spp., and *Salmonella enterica* likely evolved in the past 1 million yr (the time frame of human evolution) *(10)*.

Virulent species have evolved from numerous divergent clades of bacteria. Prevalence of pathogenic bacterial species has been estimated from various environmental sources, including the biota from animal surfaces and digestive tracts. These studies

showed that pathogens represent only a small percentage of microbial species *(11–14)*. Pathogenic strains and species are thought to be derived or evolved from either nonpathogenic siblings or closely related species.

A bacterial pathogen is defined as a microbe capable of causing host damage that results from either direct bacterial actions or from the host immune response to infection. There are many properties of a bacterium that can determine whether or not it will be a pathogen. As well as the production of a virulence factor, bacteria may need to infect the host or at least establish close contact and avoid host defenses. A pathogen may evolve from a commensal bacterium or a nonpathogen through various genetic mechanisms; such as the acquisition of virulence genes by HGT, major genomic deletions, genomic rearrangements, or point mutations.

The species of *E. coli* comprises not only nonpathogenic and commensal strains, but also different intestinal (intestinal pathogenic *E. coli* [IPEC]) and extraintestinal pathogenic (extraintestinal pathogenic *E. coli* [ExPEC]) strains that cause diseases in humans and some animals. To date, complete genome sequences have been published for five strains of *E. coli*, including nonpathogenic K12 strains MG1655 *(15)* and W3110 (http://ecoli/aist-nara.ac.jp/), two enterohemorrhagic *E. coli* (EHEC) O157:H7 strains *(16, 17)*, and one urosepsis (ExPEC) strain *(18)*. In addition, there are four pathogenic strains whose genomic sequences are at various stages of completion (*see* http://www.genomeonline.org). Comparison of the genome sequences of pathogenic and nonpathogenic *E. coli* strains identified a common core sequence of 4.1 Mb representing the backbone of the chromosome. This is interspersed with variable genomic islands (GEIs) containing strain-specific DNA sequences, some of which may contribute to virulence. Several of these GEIs in pathogenic *E. coli* strains represent pathogenicity islands (PAIs) or other mobile genetic elements that contain virulence genes *(16–19)*.

Genome analysis of the nonpathogenic probiotic *E. coli* strain Nissle 1917 identified four major GEIs (I–IV). The GEI II DNA region contains the *iuc, sat*, and *iha* genes that encode the aerobactin siderophore system, the serine protease Sat (autotransporter), and the putative adherence-conferring protein Iha, respectively *(20)*. These three genes are found and similarly organized in the *pheV*-associated GEI of the pathogenic (ExPEC) *E. coli* CFT073 strain, but the island of the pathogenic strain also contains the complete *hly* and *pap* gene clusters coding for the important virulence factors α-hemolysin and P fimbria, respectively. Interestingly, the *pap* operon in the GEI II of the probiotic strain is disrupted and partially deleted probably because of insertion of IS10 elements, and consecutive recombination events. These events, and the loss of the *hly* gene, were important steps in the evolution of the nonpathogenic Nissle 1917 strain *(20)*. Presumably the reverse could also occur, that is, a nonpathogenic *E. coli* strain, like Nissle 1917, on acquiring a wild type *hly* gene and a functional *pap* operon, could be transformed to a virulent strain.

Shigella species evolved from commensal *E. coli* strains in relatively recent evolutionary history. *Shigella* species and enteroinvasive *E. coli* (EIEC) both cause dysentery and have very similar phenotypic properties. Phylogenetic analysis indicates that *Shigella* and EIEC should be considered as a single pathovar of *E. coli (21)*. Evolution of pathogenic *Shigella* involved acquisition of a virulence plasmid, two PAIs (SHI-1 and SHI-2) *(2)* and the deletion of two virulence suppressor genes, *ompT* and *cadA (22,23)*.

Yersinia pestis, the causative agent of plague, also evolved recently in evolutionary time (1500–20,000 yr ago). It evolved from *Yersinia pseudotuberculosis (24)*. Two fully virulent *Y. pestis* strains, CO92 (Orientalis strain) *(25)* and KIM (Mediaevalis strain), *(26)* were sequenced in 2001 and 2002, respectively, followed by a nonvirulent strain 91001 *(27)* in 2004. The genome of *Y. pseudotuberculosis* IP32953, an enteropathogen of humans and animals, was recently sequenced *(28)*. All three sequenced *Y. pestis* strains and *Y. pseudotuberculosis* possess a virulence plasmid pYV (70.3 kb) (pCD1 in *Y. pseudotuberculosis*) that encodes a type III secretion system, which is responsible for injecting a number of cytotoxins and effectors into host cells (*29*; Chapter 12). *Y. pestis* possesses two additional plasmids, pPCP1 (9.6 kb) and pMT1 (102 kb), which encode plasminogen activator, murine toxin, and capsule-like antigen, respectively. The chromosome of *Y. pestis* CO92 (4.65 Mb) and *Y. pestis* KIM (4.60 Mb) are scattered with 21 GEIs including a cluster encoding a type-III secretion system that shows similar gene content and order to the Spi2 type III secretion system of *S. enterica* serovar Typhimurium *(30)*. Comparative analysis of the genomes of *Y. pseudotuberculosis* IP32953 and *Y. pestis* CO92 and KIM, identified 112 genes that are found in the two *Y. pestis* strains, but not in IP32953 *(28)*. Further analysis of 19 *Yersinia* strains showed that 32 of the 112 genes, located in six clusters, are unique to *Y. pestis (28)*. These six clusters include many genes that encode phage-related proteins, membrane proteins, and proteins with no known function. None of these are known virulence factors, but their role in pathogenesis deserves to be explored.

Evolution of *Y. pestis* from *Y. pseudotuberculosis* also involved major accumulation of pseudogenes, many of which result from insertion of IS elements. At least nine of the pseudogenes are mutated regulatory genes, which likely have contributed to the phenotype of *Y. pestis,* including its virulence *(28)*.

Vibrio cholerae is the etiological agent of cholera, a severe diarrheal disease endemic in many areas of Southern Asia and the Indian subcontinent. The emergence of pathogenic strains from nonpathogenic environmental strains involved the acquisition of at least three essential elements: (1) the PAI VPI-1, which encodes the type IV toxin coregulated pilus, an essential colonization factor and the receptor for phage CTXφ, (2) CTXφ that contains the *ctx AB* genes that encode cholera toxin (CT) *(31)*, and (3) transcriptional regulator genes (*toxRS*) and (*toxT* and *tcpPH*) encoded on the core chromosome and on the VPI-1, respectively (*32,33*; *see* Chapter 13). Determinants for resistance to trimethoprim and aminoglycosides are encoded by super integrons, and integrative and conjugative elements encode resistance to sulfamethoxazole, trimethoprim, chloramphenicol, and streptomycin *(34)*.

Unlike the evolution of *E. coli* pathogens, the evolution of obligate intracellular pathogens of both mammals and arthropods, the Rickettsiales (*Rickettsia*, *Anaplasma*, and *Wolbachia*), involves major genome reduction. It is estimated that a few thousand genes were lost at the early stage of evolution prior to the divergence of *Rickettsia* and *Wolbachia* spp. *(35)*. *R. prowazekii*, the causative agent of typhus, has the smallest genome of the group with 1.1 Mb. Genomes of *Rickettsia* spp. have characteristic low levels of coding content and a high number of pseudogenes *(36)*. *R. prowazekii* is the extreme with 24% of the genome as noncoding DNA *(37)*. Noncoding DNA and pseudogenes are intermediate stages of genome reduction. It appears that elimination of genes

was most pronounced at an early stage of the transition to the intracellular environment *(36)*, but the process is ongoing in the modern species *(38)*.

Mycoplasma and *Ureaplasma* are two genera members of the taxonomic class of Mollicutes. *Mycoplasmas*, like all mollicutes, are bacteria that do not have a cell wall. *Mycoplasma* spp are all obligate parasites of various hosts and have a small genome, generally between 0.6 and 1.4 Mb. At least five species of *Mycoplasma* (*M. pneumoniae, M. genitalium, M. hominis, M. fermentans, M. penetrans*) and one *Ureaplasma* spp. (*U. urealyticum*) are human pathogens (*see* Chapter 10). Their small genomes are thought to be the result of reductive evolution from a low guanine + cytosine Gram-positive bacterial ancestor that is common with *Clostridium* spp. and *Bacillus* spp *(39)*. Comparative analysis of the sequenced genomes of mycoplasma and ureaplasma species showed high levels of divergence and little conservation of gene order, except between *M. genitalium* and *M. pneumoniae* (*40*; Chapter 10). *Mycoplasmas* evolved different mechanisms of varying the antigenic structure of lipoproteins or adhesions, which they encode, to evade the host immune system (*41–45*; Chapter 10).

Three closely related *Bordetella* species, *B. pertussis, B. parapertussis*, and *B. bronchiseptica*, were recently sequenced *(46)*. These Gram-negative β-proteobacteria colonize the respiratory tract of mammals. *B. pertussis* is a strict human pathogen and is the causative agent of whooping cough. *B. parapertussis* infects both humans and sheep, and also can cause whooping cough in human infants. *B. bronchiseptica* has a broader host range causing respiratory infections in a wide range of animals and only occasionally humans. Comparative analysis of the three genomes indicates that *B. pertussis* and *B. parapertussis* evolved independently from a *B. bronchiseptica*-like ancestor *(46)*. Loss of genes and gene function play a major role in the evolution process *(1)*. *B. pertussis* Tohama I (4.1 Mb) and *B. parapertussis* strain 12822 (4.8 Mb) are not only significantly smaller than *B. bronchiseptica* RB50 (5.3 Mb), but also have a higher percentage of pseudogenes present in their genomes, 9.4 and 5.0 vs 0.4%, respectively.

Staphylococcus aureus, the causative agent of a wide range of human diseases, including carbuncles, food poisoning, bacteremia, necrotizing pneumonia, toxic shock syndrome, and endocarditis, is an important nosocomial and community-acquired pathogen. *S. aureus* encodes a large number of virulence factors that promote adhesion, colonization, cell–cell interactions, immune-system evasion, and tissue damage (*see* Chapter 11). Multilocus sequence typing of a large population of clinical isolates showed that the population structure of *S. aureus* is highly clonal *(47)*. Five *S. aureus* strains have been sequenced, including hospital-acquired methicillin-resistant strains (MRSAs), methilicillin-sensitive strains (MSSA), and vancomycin intermediate susceptible strains *(48–50)*. The main differences between the strains are in accessory genetic elements, including Staphylococcal chromosome cassette (SCC) elements, PAIs, GEIs, transposons, prophages, plasmids, and insertion elements *(50)*. Virulence factors are encoded by the PAIs and lysogenic phages; therefore, it would appear as though evolution of virulence involved acquiring these mobile genetic elements (*50,51*; Chapter 11).

5. Conclusions

Comparative analyses of genomic sequences of virulent and nonvirulent strains and species provide great insight regarding the mechanisms of evolution of bacterial pathogens. Many bacterial lineages, especially those that occupy microenvironments that are

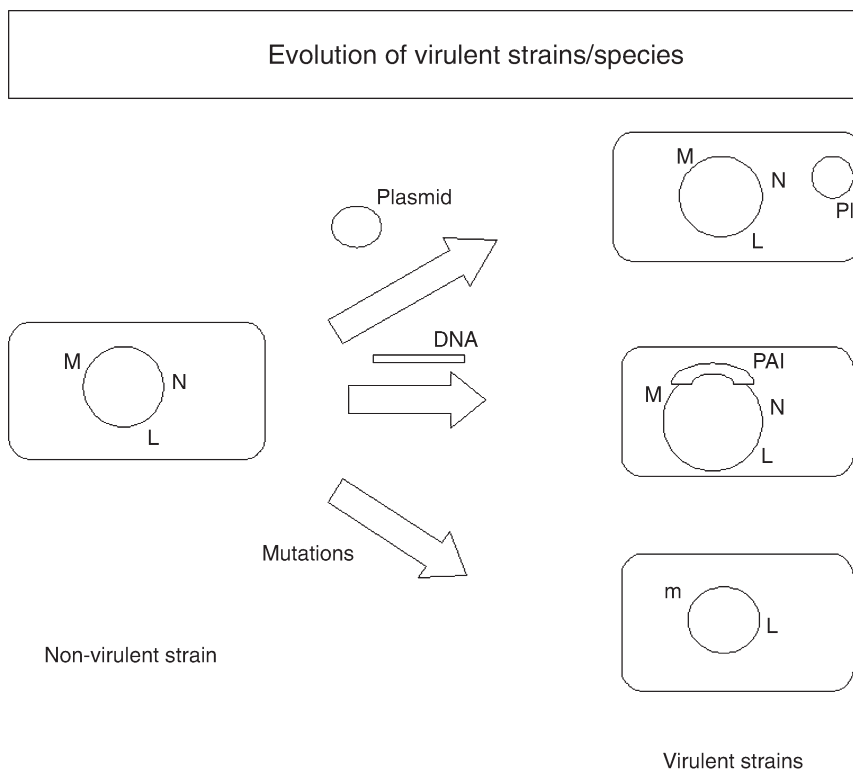

Fig. 1. Mechanisms of evolution of virulent strains and species. Mechanisms are not exclusive. Evolution of virulent strains from a nonvirulent strain may include one or more mechanisms, whereas evolution of a virulent species generally would involve multiple mechanisms. The letters M, N, and L are different genes, and m is a mutant gene of M. PAI, pathogenicity island; pl, plasmid.

rich in micro-organisms and phages, acquired their virulence genes from another species through HGT. However, intracellular nonpathogenic bacterial lineages evolved to virulent strains or species through endogenous mutations, including deletions, insertions, gene duplications, gene fusions, and rearrangements (Fig. 1).

Acknowledgments

The authors are grateful for the insightful and thoughtful comments on the chapter by Drs. Penny Chan and Gavin Clark. Assistance on preparation of the figure by Chee Chan is much appreciated.

References

1. Preston, A., Parkhill, J., and Maskell, D. J. (2004) The bordetellae: lessons from genomics. *Nat. Rev. Microbiol.* **2,** 379–390.
2. Ochman, H., Lawrence, J. G., and Grolsman, E. A. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* **405,** 299–304.
3. Arber, W. (2000) Genetic variation: mechanisms and impact on microbial evolution. *FEMS Microbiol. Rev.* **24,** 1–7.
4. Schofield, M. J. and Hsieh, P. (2003). DNA mismatch repair: molecular mechanisms and biological function. *Annu. Rev. Microbiol.* **57,** 579–608.

5. Tomb, J. F., White, O., Kerlavage, A. R., et al. (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* **388,** 539–547.
6. Bickle, T. A. and Kruger, D. H. (1993) Biology of DNA restriction. *Microbiol. Review* **57,** 434–450.
7. Roberts, R. J. (1998) Bioinformatics: a new world of restriction and modification enzymes. *NEB Transcript* **9,** 1–4.
8. Nakayama, Y. and Kobayashi, I. (1998) Restriction-modification gene complexes as selfish gene entities: Roles of regulatory system in their estabishment, maintenance, and apoptotic mutual exclusion. *Proc. Natl. Acad. Sci. USA* **95,** 6442–6447.
9. Kobayashi, I. (2001) Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. *Nucleic Acids Res.* **29,** 3742–3756.
10. Brussow, H., Canchaya, C., and Hardt, W. D. (2004) Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev.* **68,** 560–602.
11. Kroes, I., Lepp, P. W., and Relman, D. A. (1999) Bacterial diversity within the human subgingival crevice. *Proc. Natl. Acad. Sci. USA* **96,** 14,547–14,552.
12. DeLong, E. F. (2002) Microbial population genomics and ecology. *Curr. Opin. Microbiol.* **5,** 520–524.
13. Hugenholtz, P. (2002) Exploring prokaryotic diversity in the genomic era. *Genome Biol.* **3,** REVIEWS0003.
14. Torsvik, V. and Ovreas, L. (2002) Microbial diversity and function in soil: from genes to ecosystems. *Curr. Opin. Microbiol.* **5,** 240–245.
15. Blattner, F. R., Plunkett, G. 3rd, Bloch, C. A., et al. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277,** 1453–1474.
16. Perna, N. T., Plunkett, G. 3rd, Burland, V., et al. (2001) Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409,** 529–533.
17. Hayashi, T., Makino, K., Ohnishi, M., et al. (2001) Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res.* **8,** 11–22.
18. Welch, R. A., Burland, V., Plunkett, G. 3rd, et al. (2002) Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **99,** 17,020–17,024.
19. Dobrindt, U., Hochhut, B., Hentschel, U., and Hacker, J. (2004) Genomic islands in pathogenic and environmental microorganisms. *Nat. Rev. Microbiol.* **2,** 414–424.
20. Grozdanov, L., Raasch, C., Schulze, J., et al. (2004) Analysis of the genome structure of the nonpathogenic probiotic *Escherichia coli* strain Nissle 1917. *J. Bacteriol.* **186,** 5432–5441.
21. Lan, R., Alles, M. C., Donohoe, K., Martinez, M. B., and Reeves, P. R. (2004) Molecular evolutionary relationships of enteroinvasive *Escherichia coli* and *Shigella* spp. *Infect. Immun.* **72,** 5080–5088.
22. Nakata, N., Tobe, T., Fukuda, I., et al. (1993) The absence of a surface protease, OmpT, determines the intercellular spreading ability of *Shigella*: the relationship between the *ompT* and *kcpA* loci. *Mol. Microbiol.* **9,** 459–468.
23. Maurelli, A. T., Fernandez, R. E., Bloch, C. A., Rode, C. K., and Fasano, A. (1998) "Black holes" and bacterial pathogenicity: a large genomic deletion that enhances the virulence of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **95,** 3943–3948.
24. Achtman, M., Zurth, K., Morelli, G., Torrea, G., Guiyoule, A., and Carniel, E. (1999) *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis*. *Proc. Natl. Acad. Sci. USA* **96,** 14,043–14,048.

25. Parkhill, J., Wren, B. W., Thomson, N. R., et al. (2001) Genome sequence of *Yersinia pestis*, the causative agent of plague. *Nature* **413,** 523–527.
26. Deng, W., Burland, V., Plunkett, G. 3rd, et al. (2002) Genome sequence of *Yersinia pestis* KIM *J. Bacteriol.* **184,** 4601–4611.
27. Song, Y., Tong, Z., Wang, J., et al. (2004) Complete genome sequence of *Yersinia pestis* strain 91001, an isolate avirulent to humans. *DNA Res.* **11,** 179–197.
28. Chain, P. S., Carniel, E., Larimer, F. W., et al. (2004) Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*. *Proc. Natl. Acad. Sci. USA* **101,** 13,826–13,831.
29. Cornelis, G. R. (2002) *Yersinia* type III secretion: send in the effectors. *J. Cell Biol.* **158,** 401–408.
30. Wren, B. W. (2003) The yersiniae—a model genus to study the rapid evolution of bacterial pathogens. *Nat. Rev. Microbiol.* **1,** 55–64.
31. Waldor, M. K. and Mekalanos, J. J. (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272,** 1910–1914.
32. Faruque, S. M., Asadulghani, Saha, M. N., et al. (1998) Analysis of clinical and environmental strains of nontoxigenic *Vibrio cholerae* for susceptibility to CTXPhi: molecular basis for origination of new strains with epidemic potential. *Infect Immun.* **66,** 5819–5825.
33. Dziejman, M., Balon, E., Boyd, D., Fraser, C. M., Heidelberg, J. F., and Mekalanos, J. J. (2002) Comparative genomic analysis of *Vibrio cholerae*: genes that correlate with cholera endemic and pandemic disease. *Proc. Natl. Acad. Sci. USA* **99,** 1556–1561.
34. Hochhut, B. and Waldor, M. K. (1999) Site-specific integration of the conjugal *Vibrio cholerae* SXT element into *prfC*. *Mol. Microbiol.* **32,** 99–110.
35. Batut, J., Andersson, S. G., and O'Callaghan, D. (2004) The evolution of chronic infection strategies in the alpha-proteobacteria. *Nat. Rev. Microbiol.* **2,** 933–945.
36. Boussau, B., Karlberg, E. O., Frank, A. C., Legault, B. A., and Andersson, S. G. (2004) Computational inference of scenarios for alpha-proteobacterial genome evolution. *Proc. Natl. Acad. Sci. USA* **101,** 9722–9727.
37. Andersson, S. G., Zomorodipour, A., Andersson, J. O., et al. (1998) The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* **396,** 133–140.
38. Amiri, H., Davids, W., and Andersson, S. G. (2003) Birth and death of orphan genes in *Rickettsia*. *Mol. Biol. Evol.* **20,** 1575–1587.
39. Maniloff, J. (2002) Phylogeny and evolution, in *Molecular Biology and Pathogenicity of Mycoplasmas* (Rasin, S. and Herrmann, R., eds.). Kluwer Academic/Plenum Publishers, New York, pp. 31–44.
40. Sasaki, Y., Ishikawa, J., Yamashita, A., et al. (2002) The complete genomic sequence of *Mycoplasma penetrans*, an intracellular bacterial pathogen in humans. *Nucleic Acids Res.* **30,** 5293–5300.
41. Neyrolles, O., Chambaud, I., Ferris, S., et al. (1999) Phase variations of the *Mycoplasma penetrans* main surface lipoprotein increase antigenic diversity. *Infect. Immun.* **67,** 1569–1578.
42. Glew, M. D., Papazisi, L., Poumarat, F., Bergonier, D., Rosengarten, R., and Citti, C. (2000) Characterization of a multigene family undergoing high-frequency DNA rearrangements and coding for abundant variable surface proteins in *Mycoplasma agalactiae*. *Infect. Immun.* **68,** 4539–4548.
43. Roske, K., Blanchard, A., Chambaud, I., et al. (2001) Phase variation among major surface antigens of *Mycoplasma penetrans*. *Infect. Immun.* **69,** 7642–7651.
44. Liu, L., Panangala, V. S., and Dybvig, K. (2002) Trinucleotide GAA repeats dictate *pMGA* gene expression in *Mycoplasma gallisepticum* by affecting spacing between flanking regions. *J. Bacteriol.* **184,** 1335–1339.

45. Winner, F., Markova, I., Much, P., et al. (2003) Phenotypic switching in *Mycoplasma gallisepticum* hemadsorption is governed by a high-frequency, reversible point mutation. *Infect. Immun.* **71,** 1265–1273.
46. Parkhill, J., Sebaihia, M., Preston, A., et al. (2003) Comparative analysis of the genome sequences of *Bordetella pertussis, Bordetella parapertussis* and *Bordetella bronchiseptica*. *Nat. Genet*. **35,** 32–40.
47. Feil, E. J., Cooper, J. E., Grundmann, H., et al. (2003) How clonal is *Staphylococcus aureus*? *J. Bacteriol.* **185,** 3307–3316.
48. Kuroda, M., Ohta, T., Uchiyama, I., et al. (2001) Whole genome sequencing of meticillin-resistant *Staphylococcus aureus*. *Lancet* **357,** 1225–1240.
49. Baba, T., Takeuchi, F., Kuroda, M., et al. (2002) Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet* **359,** 1819–1827.
50. Holden, M. T., Feil, E. J., Lindsay, J. A., et al. (2004) Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. *Proc. Natl. Acad. Sci. USA* **101,** 9786–9791.
51. Lindsay, J. A. and Holden, M.T. (2004) *Staphylococcus aureus*: superbug, super genome? *Trends Microbiol.* **12,** 378–385.

http://www.springer.com/978-1-58829-496-8

Bacterial Genomes and Infectious Diseases
Chan, R.V.L.; Sherman, P.M.; Bourke, B. (Eds.)
2006, XIII, 270 p., Hardcover
ISBN: 978-1-58829-496-8
A product of Humana Press