

Host–microbe interactions: bacteria Microbiology will never be the same again...

Editorial overview

Guy R Cornelis* and Rino Rappuoli†

Addresses

*Microbial Pathogenesis Unit, Université Catholique de Louvain and Christian de Duve Institute of Cellular Pathology, Avenue Hippocrate, 74 B1200 Brussels, Belgium; e-mail: cornelis@mipa.ucl.ac.be

†IRIS, Chiron SpA, Via Fiorentina 1, 53100 Siena, Italy; e-mail: rino_rappuoli@chiron.it

Current Opinion in Microbiology 2001, 4:13–15

1369-5274/01/\$ – see front matter

© 2001 Elsevier Science Ltd. All rights reserved.

Tremendous progress has been made recently in research in the field of microbial pathogenesis and especially in the subfield of cellular microbiology. For instance, there has been significant development in research involving well-known microorganisms such as *Salmonella*. However, this issue of *Current Opinion in Microbiology* deliberately ignores these well-known microorganisms in order to focus on microorganisms that may be less well understood, but for which a significant impact is expected from genomics. All the reviews appearing in this issue were written by experts with this in mind and, indeed, they show that genomics and proteomics will soon bring the state of understanding of the pathogenesis of these organisms up to that of *Salmonella*. In this overview, we shall provide a brief synopsis of the pathogens described in this issue, followed by a description of the areas in which genomics is expected to have a major impact.

The impact of genomics on the understanding of host–microbe interactions

Chlamydiae are microorganisms to which genomics is of tremendous help, as all the classical approaches are drastically restricted by the difficulty in growing these organisms and performing genetic studies on them. The complete genome sequence for *Chlamydia trachomatis* and *Chlamydia pneumoniae* immediately and comprehensively informed the biology of these human pathogens well beyond the scope of pre-genomic understanding (Stephens and Lammel, pp 16–20). Among the important outputs of genomic studies on these organisms was the discovery of families of polymorphic membrane proteins (Pmp) sometimes subjected to phase variation, and the discovery of a type III secretion system. There is an amazing paradox with *Chlamydia*: on the one hand, type III secretion was only detected by genomics, while on the other hand, the type III injectisome may well have been seen in the electron microscope long before the idea of type III secretion had been conceived!

Mycobacterium leprae is another organism for which research is severely hampered by the difficulty of *in vitro* culture. Not surprisingly, the genome appeared to be three times

smaller than that of the closely related *Mycobacterium tuberculosis*, implying the loss of about a thousand genes. Still, it is well adapted to intracellular parasitism and long term survival *in vivo*. Substantial progress has been recently made towards the understanding of the key issue of the microorganism's capacity to infect peripheral nerves (Rambukkana, pp 21–27). Recent studies have shown that a phenolic glycolipid of the *M. leprae* membrane interacts with laminin from the nerve basal lamina, which in turn interacts with α -dystroglycan on host Schwann cells.

The 4.4 Mb complete genome sequence of *M. tuberculosis* was published in 1998. Among other features, this genome contains an unusual but not unexpected abundance of genes involved in fatty acid metabolism. The completion of the sequence led to a marked intensification of research in the field of tuberculosis (Domenech, Barry and Cole, pp 28–34). In particular, it opened the way for genome comparison using DNA microarray technology. We now have a clear view of the diversity that exists among the various *M. tuberculosis* BCG strains used throughout the world. Several studies have already identified genes expressed within macrophages, and proteomics is also well on its way, although it does not deal so far with proteins specifically expressed in macrophages. The generation and analysis of knockout bacilli for several genes, such as anaerobic nitrate reductase, showed that oxygen-independent respiration plays an important role in the adaptation of *M. tuberculosis* to the intracellular environment. *M. tuberculosis* knockouts that are mutant in the synthesis of the siderophores, mycobactin, were also impaired for growth in macrophages.

The enteropathogen, *Campylobacter* (Linton, Karlyshev and Wren, pp 35–40), is one of the least understood pathogens. This may seem surprising if one considers that the other enteropathogens, such as *Salmonella*, *Shigella*, *Yersinia enterocolitica* and enteropathogenic *Escherichia coli* strains (EPECs), are among the best understood pathogens. This situation was the driving force for the *C. jejuni* genome sequencing project, and the results are not disappointing. The sequence first identified a group of genes with significant similarity to those involved in polysaccharide capsular biosynthesis in *E. coli* and *Neisseria* spp. Another major point is that a number of *Campylobacter* proteins, including flagellin, are glycosylated. The function of this glycosylation is still unclear but, because it occurs in a number of proteins, it is likely to play a role in pathogenesis. In addition, the sequence shows the presence of homopolymeric tracts in *C. jejuni* genes encoding surface structure, indicating phase variation. *C. jejuni* is thus a unique enteropathogen sharing this property with mucosal

pathogens like *Neisseria*, *Haemophilus* and *Helicobacter*. Finally, the sequence analysis confirms that *C. jejuni* lacks a type III secretion system, a finding that had been previously (perhaps prematurely) reported.

For *Helicobacter pylori* (Censini, Stein and Covacci, pp 41–46), type IV secretion systems are now conceptually unified to type III secretion systems. In a fashion similar to the one by which EPECs deliver their receptor, Tir, to eukaryotes, *Helicobacter* delivers CagA by type IV secretion, and CagA becomes phosphorylated. This is followed by cortical actin polymerization. Interestingly, the pathogenicity island encoding CagA is genetically unstable, and this leads to frequent loss of CagA and new adaptation capacities.

For *Neisseria*, the pregenomic era had not revealed any major ‘virulence system’ except exceptional skills in adherence and invasion of eukaryotic cells, coupled to an extraordinary capacity to modify almost every surface-exposed component. At present, two strains of *Neisseria meningitidis* (Tinsley and Nassif, pp 47–52) have been sequenced, and the sequencing of one strain of *Neisseria gonorrhoeae* (Kooimey, pp 53–57) is nearing completion. In agreement with previous work that had shown many similarities between these two organisms, the two chromosomes appear to be 90% homologous. A close comparison of the two sequences will certainly lead to a more comprehensive understanding of the respective properties in pathogenesis. Genomics has also shown that more genes than expected are potentially phase variable, as they possess homopolymeric tracts or iterated oligonucleotide motifs. The next step in *Neisseria* research will be the comparative analysis of several *N. meningitidis* strains in order to determine why some strains of *N. meningitidis* are more virulent than others.

Although brucellosis remains a worldwide problem for cattle and a serious risk for the professionals in contact with infected cattle, our knowledge of *Brucella* pathogenesis is still lagging far behind that of other pathogens. Very recently, the 3.25 Mb genome sequence of *B. melitensis* was completed and the information gathered has begun to be exploited (Boschiroli, Foulongne and O’Callaghan, pp 58–64). As well as enabling classical transcriptome analysis and the engineering of a systematic knockout library, the sequence will also provide answers to a few immediate questions. For example, are the genes related to type III secretion really involved in type III secretion or do they belong to the flagellar system? What does the type IV secretion system deliver? What is the origin of the multiple replicons? What is the relation between *Brucella* and other α -proteobacteria?

A number of important contributions to the understanding of the pathogenesis of Group A streptococci (Graham *et al.*, pp 65–70) have appeared recently. The sequence of a serotype M1 strain is already available and a serotype M5 is being sequenced. These sequences contain the insertion elements and lysogenic bacteriophages that were expected. The sequences also contain numerous putative regulators

and two-component regulators, indicating that Group A streptococci can sense many environmental signals in addition to the ones that have been characterized in previous literature. Large scale comparative sequencing of one gene — the *sic* gene — also showed the emergence of variants in pharyngeal isolates in the population of infected people, prior to the onset of epidemic waves. In the same vein, comparative genome analyses using DNA microarrays has already begun to provide information on the differences present among serotypes with higher and lower virulence.

The pneumococcus (Hollingshead and Briles, pp 71–77) is one of the longest known pathogens. The study of its virulence for mice is famous not because it showed the role of a capsule in virulence, but because it has been instrumental to the discovery of the role of DNA in heredity. Now, the emergence of antibiotic-resistant strains makes the need for new antibiotics and broadly efficacious vaccines greater than ever. Genome information has already been used to identify new vaccine candidates, and it helped to identify new choline-binding proteins that seem to play a role in virulence. Several reports have already highlighted the role of such choline-binding proteins in pneumococcal growth, suggesting that they may constitute new drug targets. Finally, genomics and DNA arrays also help to unravel the mechanism of transformation, a mechanism that appears to be more sophisticated and controlled than may have previously been anticipated.

Sequencing of the two megaplasmsids pXO1 and pXO2, which encode the tripartite toxin and the capsule, respectively, was the first genomic step in *Bacillus anthracis* research (Baillie and Read, pp 78–81). Sequencing of the chromosome, which is still in progress, has so far confirmed the similarity of *B. anthracis* to *B. subtilis*, and has also indicated the presence of homologs to virulence genes in *B. cereus* and *B. thuringiensis*.

Initially considered to be a toxin producer, *Bordetella pertussis* (Locht, Antoine and Jacob-Dubuisson, pp 82–89) is now recognized to be one of the most complex pathogens that interacts in multiple ways with its host. It is a pathogen well endowed with an array of different adhesins, with type I, type III and type IV secretion systems, with two autotransporters, and an iron acquisition system. Sequencing of the *B. pertussis* genome, which is still in progress, has so far revealed putative additional virulence factors and an impressive number of regulators, including 13 sigma factors, suggesting the existence of an as yet uncharacterized sensing and regulatory network.

Finally, this issue will also revisit the multifaceted *E. coli*, with its various pathogenicity islands and virulence plasmids (Dougan *et al.*, pp 90–94).

Global prospects

The determination of the nucleotide sequences of the genomes of each of the bacteria reviewed in this issue is a

fundamental milestone that can advance the knowledge of the most difficult bacteria almost to the same level of that of the best-studied microorganisms. In practice, there are no more secrets to discover in these genomes, and thus they can be the starting points from which the answers to important questions can be found. These include the following questions. How are the genes really used *in vivo*? What is the relationship between bacteria and variability within a bacterial population worldwide? How do bacteria change their gene expression when interacting with host cells or other microorganisms? The areas in which the genomes are expected to have major impact are described below.

Whole-genome assays

Microarrays containing the whole genomes of pathogens are becoming available and will soon be used routinely. This will be a revolution for microbiology laboratories, which, instead of investigating one area at a time, will be able to perform experiments on multiple problems at the same time. Microarrays can be used in diagnostics and to compare the whole genomes of different bacterial isolates. Using the same effort required to test by PCR or Southern blotting methods whether or not, for instance, a gene is present in different isolates, we can compare the whole genomes of these isolates and obtain global information. Mutants generated in the laboratory can be tested not only for the presence or absence of a gene of interest, but their whole genomes can be checked by DNA microarray experiments. Similarly, microarrays will allow the testing of changes in gene expression in the whole genome during infection. Diehn and Relman (pp 95–101) compared two recent studies on the responses of human respiratory epithelial cells to the two Gram-negative respiratory tract pathogens, *Pseudomonas aeruginosa* and *Bordetella pertussis*. Although these two studies used different platforms and experimental conditions, they revealed two sets of genes that are activated by both pathogens. The first set encodes proteins that are involved in nuclear factor κ -B (NF κ B) activation. The second set of co-induced genes is, interestingly, involved in the clotting cascade. Perhaps more surprising is the observation that, in both cases, genes encoding antagonistic players were simultaneously induced. After comparing the two studies, Diehn and Relman pinpoint a number of caveats that researchers using DNA microarray technology should be mindful of.

Clearly, the microarray technology will be at the forefront of cellular microbiology. Global studies of gene expression will be possible for pathogens interacting with host cells *in vitro*, and also for pathogens grown in animal models or directly from infected tissues and biopsies.

New targets for vaccines

So far, vaccines have been developed against bacteria whose protective virulence factors (for example, toxins, capsular polysaccharides and adhesins) can be easily identified and purified in large quantities by growing the pathogen *in vitro*. However, vaccine development has been difficult or impossible for bacteria that cannot be grown *in vitro*, or when virulence factors are variable in sequence or difficult to identify. The whole genomic sequence provides a catalogue of all the antigens that a pathogen can express at any time, and eliminates the time-consuming antigen discovery phase and the need to grow the pathogen *in vitro*. Now, vaccine development can begin with computer prediction of possible antigens, and their expression in a recombinant system. Using this approach, called reverse vaccinology, novel vaccine candidates have already been discovered for serogroup B *N. meningitidis*, and will soon be discovered for many other pathogens, including those not cultivable *in vitro*.

New targets for antimicrobials

The successful antibiotic control of many bacterial diseases relies on the use of molecules that target no more than five metabolic pathways in the bacterial cell. Genome sequencing can aid this method of control by providing a catalogue of all possible targets for antimicrobials. The new targets can be expressed in a recombinant host and used for high-throughput screening of new molecules. The availability of antimicrobial agents that target different pathways and the use of pathogen-specific molecules will enable better control of antimicrobial resistance.

In conclusion, genomics leads to a quantum leap in the speed of acquisition of new information and the understanding of microbial pathogenesis. New vaccines are in sight and the targets for new therapeutics can be identified. This issue summarizes the situation for a dozen bacterial pathogens that remain or will become again major health problems.