INTERNATL MICROBIOL (1998) 1:255–258

REVIEW ARTICLE

The Renaissance of microbiology

Julian Davies

TerraGen Diversity, Inc. and Department of Microbiology and Immunology, University of British Columbia, Vancouver, B.C., Canada

Received 11 July 1998
Accepted 15 August 1998

Correspondence to:
Department of Microbiology and Immunology. University of British Columbia. 300–6174 University Boulevard. Vancouver, B.C. V67 1Z3. Canada.
Tel.: +1-604-2218896. Fax: +1-604-2218881.
E-mail: jed@unixg.ubc.ca

Summary Microbiology is finally occupying its true position as the pre-eminent field in life sciences. This is due to advances in molecular techniques that confirm the evolutionary significance of the biology of microbes. It is anticipated that the use of comparative genomics will provide information that will advance the understanding of mechanisms of pathogenesis and the importance of secondary metabolism in social microbiology. More emphasis on studies of microbial diversity will increase its value in both fundamental microbiology and its industrial applications.

Key words Genomics · Microbial diversity · Secondary metabolism · Microbial evolution · Antibiotic resistance

The science of microbiology (the study of the oldest living and most abundant organisms on this planet) is relatively young; it began seriously only in the mid-19th century with the work of Pasteur and Koch. Chemistry (or at least alchemy) and physics are more ancient scientific disciplines. All the sciences have developed rapidly in the past one hundred years, but in particular the explosion of new knowledge in life sciences that began some thirty years ago and continues to increase in an exponential manner is clearly beyond all expectation; the growth curve resembles that of a microbe! This expansion has its basis in advances in the techniques of genetics and molecular biology and, more recently, the advent of nucleic acid sequencing that has begun to reveal detailed information on gene structure and organization on a larger scale. During the early excitement of the Human Genome Project, the sequencing of bacterial genomes was given little attention, despite the fact that the first genomes completed in early sequencing efforts had been those of microbes, the bacteriophages ϕx174 and λ; now, however, the sequencing of prokaryotic genomes has moved to centre stage and the leading laboratories of eukaryotic sequence analysis are currently devoting much effort to the analysis of the genomes of microbes. It is reliably predicted that by the year 2000 there will be more than one hundred completed bacterial genome sequences [12]. While the prospect of all this knowledge is daunting, it should be considered appropriate in light of the fact that this represents only a tiny fraction of the microbial kingdom, 99.9% of which cannot be grown for laboratory study.

This emphasis is, of course, exactly as it should be! Microbes are the founder members of this planet and understanding bacterial function is a first priority in biology because of the critical role of microbes in the maintenance of all other forms of life. If the process of evolution is to be unravelled properly to elucidate the evolution of biosynthetic pathways and their regulation, complete bacterial genome sequences are obligatory to furnish the predictive information essential for functional genomic studies of more complex genomes, since gene identification and organization in higher organisms will be derived principally from comparative studies with simpler genomes.

At the time of writing, some twenty microbial genome sequences have been completed (although not necessarily annotated) and are in the process of more detailed analysis; clearly this work will keep microbiologists busy for some time. The nucleotide sequences of fifty other microbial genomes are well advanced, and the continuation of these efforts will result in the completion of the sequences of most of the major classes of bacteria, archaea and lower eukaryotes within 5 years. In addition, current commercial databases have been estimated to contain two- to three-fold more bacterial sequence information, which is not available to the scientific community. For much of this latter it is not known what organisms have been sequenced or the quality of the data. (It is regrettable that industry benefits from all of the publicly funded sequence information from academia [such as the yeast genome] but provides little in return.) Be that as it may, the publicly available sequences have provided new knowledge that whets the appetite for more, even though the analytical approaches (functional genomics) being used to obtain useful biological conclusions from these sequences are still fairly primitive.

What has been learned from this information so far? First and foremost is the finding that, depending on the organism,

as much as 30% of the open-reading-frame sequences are essentially uninterpretable, since they do not match the genes thus far identified in the available databases; this is both exciting and frustrating and underlines the need to develop the methods that can be used to predict protein function from DNA sequence information. Nonetheless, many interesting conclusions with respect to overall chromosome organization, biosynthetic pathways and their regulation have been revealed. (For examples, see references [3, 6, 11].) With the number of sequenced genomes increasing comparative studies have become more reliable, providing information on specific functions, as mentioned below. There are other constraints on the use of nucleotide sequence information. In most cases, in order to simplify the assignment and analysis of genetic organization, it has been considered necessary to restrict the minimum length of identified open reading frames to 50–100 amino acids; it is thus likely that many functions important to microbial life will not be identified during this first phase of the sequencing revolution. Small bacterial peptides and their derivatives are being overlooked, for example molecules of the microcin type, which are probably made by most bacteria and play roles in many aspects of the biology of microbial interactions [2]. The comparative analyses of the sequences of housekeeping genes are proving to be invaluable in establishing detailed phylogenetic trees, and initial work has uncovered several surprises that will influence our thinking on the evolution of genome function. It is evident that improved methods of protein domain identification and more sophisticated pattern tools will be needed if raw nucleotide sequences are to be interpreted to their full extent. For example, how does one predict the function of novel RNA genes?

At the present time, comparative (subtractive) analyses of closely related species are proving to be the most productive; for example, the identification and study of pathogenicity islands is an exciting area of investigation that is revealing new concepts of microbial behavior [14]. Simply described, since the difference in properties between a pathogen and a non-pathogen of the same bacterial species must lie in nucleotide sequence differences, subtractive analysis should identify those sequences involved in specific steps of pathogenesis; in addition, differences in G+C content, coding preferences or similar variations can provide evidence suggestive of horizontal gene transfer. The demonstration that genes responsible for virulence functions (toxins, adhesion and binding factors, specific secretion mechanisms, etc.) are found in discrete clusters on bacterial chromosomes and are clearly of foreign origin is of great significance in studying the evolution of pathogens and the molecular mechanisms of the process of host/pathogen interactions [8, 13, 22]. While the utility of the discovery of “pathogene” clusters to identify new targets for drug discovery remains untested at the time of writing, there is good reason to believe that future study of agents that can interfere with functions involved in host/pathogen interactions may lead to novel types of therapy. Possibilities include classes of therapeutic compounds that do not interfere with normal growth of the organism but rather turn off those functions that give the pathogen its foothold in infection (disarmament rather than killing), thus permitting the immune defences of the host to destroy a microbe that no longer has a survival advantage in the host. The work of Balaban et al. has given encouragement that this approach could be successful in some cases [1]. One obvious benefit of this type of agent would be that selection for antibiotic resistance might be less severe, although it is likely that the use of any agent capable of reducing the competitive survival (growth) of a microbe in a particular environment would still provide a selection for mechanisms of avoidance. Microbes are the ultimate survivors! However, should this approach prove successful, there is a significant operational difficulty: it will run counter to several generations of medical pharmaceutical schooling in the treatment of infectious disease, in particular the rationale for the use of broad-spectrum antimicrobial therapy with cidal agents. How readily would such a radical, new approach be accepted?

Comparative genomic analysis is likely to identify numerous other ‘islands’ or ‘islets’ of genetic information related to specific properties of the organism. For the study of antibiotic resistance, for instance, it should become possible to identify antibiotic resistance genes and their clusters on the basis of association with common elements such as attachment sites or integrases; these genes are most often found in linked and often co-regulated groups [15, 18]. Also, improvements in the process of gene-protein identification will provide clues concerning the origins and evolution of resistance and make it possible to predict ‘new’ antibiotic resistance genes directly from the scrutiny of nucleotide sequence data.

An additional benefit of sequence comparisons of genomes will be the possibility to identify gene clusters encoding the pathways for catabolism of xenobiotic degradation and also pathways responsible for the biosynthesis of secondary metabolites. As the nucleic acid sequences of more of the complex bacteria (streptomycetes, myxococci, etc.) are determined, such comparative analyses should permit the identification of novel degradative and biosynthetic gene clusters. In most cases these might be expected to be present as genetic ‘islands’, distinguished from flanking DNA sequences in codon usage, regulatory organization, etc. Identification of such modules may then lead to the isolation (possibly by surrogate host expression) of novel bioactive secondary metabolites that have therapeutic value. As an indication of the power of this technique, one early result of the *Streptomyces coelicolor* genome project has been the complete structural characterisation of the biosynthetic cluster for undecylprodigiosin (the *red* cluster) [5]. As more of these clusters are picked out, the description of the complex regulatory processes involved in control of biosynthetic pathways for small molecules will aid in the rational engineering of secondary metabolite production levels for industrial uses.

The production of secondary metabolites is typical of many different microbial genera and species (perhaps all?), but their synthesis and functions are poorly characterized; the application of comparative genomics to the identification of the biosynthetic pathways, their potential products and their functions in the physiology of the producing organism is an exciting prospect. How easily will this be achieved? The gene clusters are highly specialized, but it should be possible to distinguish them from those encoding the housekeeping or essential cell functions using the same approaches that are being employed to detect other unique characteristics such as pathogenicity, as mentioned earlier. Obviously, dramatic advances in the ability to analyze nucleotide sequences within the context of entire genomes and predict their biochemical roles (functional genomics) will be required. This will not necessarily be an easy process since, if we take the microcins of *E.coli* as an example, the gene encoding the basic peptide structure would likely be missed under the search parameters currently used; the genes involved in post-translational modification would be difficult to identify unless particular types of gene organization are being sought. The identification of the previously unknown L-idonate metabolic pathway in *E. coli* is an interesting portent for the future of this type of study [20].

As differential genome sequence analysis becomes more sophisticated, there is little doubt that many examples of horizontal gene transfer at the inter- and intra-genic and species level will be uncovered. This is to be expected, since studies indicate that gene transfer is a significant factor in genome evolution for housekeeping and specialized functions [24]. For example, in the case of secondary metabolites, there is suggestive evidence for the horizontal transfer of the gene clusters required for β-lactam biosynthesis between bacteria and fungi [7]; it can be anticipated that this will be the case for many other biosynthetic pathway clusters. Such findings will engender questions as to what mechanisms of transfer are involved and what will be the organismal limits of the various types of gene exchange? Might there be novel types of transfer?

Since low-molecular-weight compounds are presumed to play important roles in maintaining microbial communities, their cognate biosynthetic pathways must be under the control of sensing mechanisms for cell density and a variety of stressful environmental changes. What role(s) will global regulatory systems play in these cases? To date, a number of peptide-signalling agents have been identified in microbial developmental cycles and also in the lifestyles of certain bacterial pathogens; it is not unreasonable to suggest that polyketides, nucleosides, sugars, and hybrids between these structures will also play roles in the maintenance of microbial community life [16]. A bacterial cytokine (protein) has recently been identified [19], and it will be important to unravel the processes involved in the function of informational molecules of this type, not only for what they reveal about the life-styles of single-celled organisms, but also as potential evolutionary precursors of the complex regulatory and controlling processes in higher eukaryotes, such as those of hematopoiesis and the immune system.

It is of interest to note that, in recent years, many so-called 'eukaryotic' processes (such as tyrosine- and serine-threonine kinases and their associated phosphatases as well as calmodulin-based regulation, to name just a few) have been identified in prokaryotes [17, 25] and have been shown to play critical roles in cell development and host/pathogen interactions [4, 10]. Conversely, the histidine-aspartyl-kinase two-component signalling processes (typically 'prokaryotic') have been identified in eukaryotes [17]. If truth be known, all such biochemical processes are 'prokaryote-like'. Suggesting that these functions are specific to any genus or species is non-productive, since they are all related through cellular evolution. It is to be expected that common post-translational processes will exist throughout biology and an evolutionary approach will be essential to provide an understanding of cellular biochemistry and behavior; as increasing nucleotide sequence becomes available, it can be predicted that more universal functions will be identified. Jacques Monod once proposed that when we understand *E. coli* we will understand elephants; this prophecy, in a general sense, now seems plausible. The notion that a 'pool' of progenotes (rather than a single universal progenitor) was the likely basis for cellular evolution, coupled with evidence of extensive horizontal gene transfer during the evolution of genera and species indicates that the evolution of living cells has the characteristics of a chaotic process and needs to be considered in this way.

Finally, returning once more to the subject of secondary metabolism, the function of these processes and their many products in microbial ecology remains largely a mystery and is now ripe for more detailed investigation. As mentioned previously, new studies of non-cultivable organisms and comparative genomic analyses will probably unmask new classes and types of products. While there is definite evidence for a 'true' antibiotic role in certain circumstances (especially in plant/bacterial relationships and protection against infection), the enormous number and variety of microbial secondary metabolites produced by microbes provoke many questions and call for further analysis of their natural functions. Many activities and roles for secondary metabolism (in one sense the most creative form of cellular metabolism) have been suggested [9], but few convincing experiments have been carried out; obviously the experiments are not easy! Given the fact that microbes always grow in communities—a certainty that is becoming increasingly apparent to present-day microbiologists—it is reasonable to assume that these communities, such as biofilms, soils and other macro- and micro-ecosystems, must sustain their structural and metabolic stability through microbe/microbe interactions involving (principally) low-molecular-weight compounds as the intermediates for inter- and intra-cellular cross-talk and regulation [21]. More attention to the subject of 'social microbiology' should reveal novel

Table 1 Why study microbial diversity?

1. Explore the environmental limits of life.
2. Establish the organismal requirements for biosphere sustainability.
3. Provide new resources for biotechnology, pharmaceuticals, energy production, waste treatment.
4. Monitor and predict environmental change, understand global chemical cycling.
5. Broaden conservation biology and bioremediation.
6. Test community ecology principles.
7. Obtain records of early eras—evolutionary relationships and mechanisms (e.g., horizontal gene transfer) and their role in the evolution and maintenance of the biosphere.
8. Study gene transfer in nature: environmental safety, fate of released DNA, survival of introduced microbes. Examine the extent of natural gene exchange.
9. Define and prevent harmful microbial activities such as corrosion, biofouling, fertilizer degradation.
10. Expand knowledge of the gene pool.
11. Identify the nature of pathogenicity and other host-parasite relationships.
12. Understand the sociobiology of microbial communities and their maintenance.

aspects of the chemical interactions (hypothesized to be mostly synergistic) between microbes in nature.

This is a very good time to be a microbiologist. The ability to grow single bacterial colonies on solid medium, developed by Robert Koch in 1881, was a critical discovery in the development of microbiological science, but one which, for more than a century, has focused the efforts of the majority of microbiologists on studying microbes as single colony entities and not as complex communities. The vast array of powerful technical methods available for the study of all aspects of microbial growth, coupled with the benefits of complete microbial genome sequences, has now put microbiology in its "place as the fundamental discipline of all biology" [23]. Employing and interpreting the accumulating information and applying the derived knowledge to the broader questions of evolution, disease, and the functions of microbial diversity in all aspects of life on earth will entail an enormous amount of work. Perhaps more than anything, concerted efforts to understand the full nature of microbial diversity (Table 1), its genetics, biochemistry and control, will earn the most substantial dividends for the future of biology.

Acknowledgments This paper is dedicated to Sagrario Mochales, whose passion for science and discovery is matched only by her friendship, benevolence and love of life. The author wishes to thank Dorothy Davies for her assistance in preparing the manuscript.

References

1. Balaban N, Goldkorn T, Nhan RT, Dang LB, Scott S, Ridgley RM, Rasooly A, Wright SC, Larrick JW, Rasooly R, Carlson JR (1998) Autoinducer of virulence as a target for vaccine and therapy against *Staphylococcus aureus*. Science 280:438–440
2. Baquero F, Moreno F (1984) The microcins. FEMS Microbiol Lett 23:117–124
3. Blattner FR, Plunkett III G, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y (1997) The complete genome sequence of *Escherichia coli* K-12. Science 277:1453–1462
4. Bliska JB (1995) Crystal structure of the *Yersinia* tyrosine phosphatase. Trends Microbiol 3:125–127
5. Cerdeño Terraga AM, Bibb M (1998) Antibiotic undecylprodigiosin production by *Streptomyces coelicolor* A3(2). Abstract, 8th International Symposium on the Genetics of Industrial Microorganisms, Jerusalem, June 28–July 2, 1998
6. Clayton RA, White O, Ketchum KA, Venter JC (1997) The first genome from the third domain of life. Nature 387:459–462
7. Cohen G, Shiffman D, Mevarech M, Aharonowitz Y (1990) Microbial isopenicillin N synthase genes: structure, function, diversity and evolution. Trends Biotechnol 8:105–111
8. Covacci A, Falkow S, Berg DE, Rappuoli R (1997) Did the inheritance of a pathogenicity island modify the virulence of *Helicobacter pylori*? Trends Microbiol 5:205–208
9. Davies J (1992) Introduction. In: Secondary metabolites: Their function and evolution. John Wiley & Sons, Chichester. Ciba Foundation Symposium 171:1–2
10. Finlay BB, Cossart P (1997) Exploitation of mammalian host cell functions by bacterial pathogens. Science 276:718–725
11. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldman H, Galibert F, Hoheisel JD, Jacq D, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG (1996) Life with 6000 genes. Science 274:546–567
12. Goodwin P (1998) Microbial genome sequencing – where are we now? SGM Quarterly 25:52–53
13. Groisman EA, Ochman H (1997) How *Salmonella* became a pathogen. Trends Microbiol 5:343–349
14. Hacker J, Blum-Oehler G, Mühldorfer I, Tschäpe H (1997) Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. Mol Microbiol 23:1089–1097
15. Hall RM (1997) Mobile gene cassettes and integrons: moving antibiotic resistance genes in Gram-negative bacteria. In: Antibiotic Resistance: Origins, Evolution, Selection and Spread. John Wiley & Sons, Chichester. Ciba Foundation Symposium 207:192–205
16. Kell DB, Kaprelyants AS, Grafen A (1995) Pheromones, social behaviour and the functions of secondary metabolism in bacteria. TREE 10:126–129
17. Kennelly PJ, Potts M (1996) Fancy meeting you here! A fresh look at "prokaryotic" protein phosphorylation. J Bacteriol 178:4759–4764
18. Mazel D, Dychinco B, Webb VA, Davies J (1998) A distinctive class of integron in the *Vibrio cholerae* genome. Science 280:605–607
19. Mukamolova GV, Kaprelyants AS, Young DI, Young M, Kell DB (1998) A bacterial cytokine. Proc Natl Acad Sci USA 95:8916–8921
20. Peekhaus N, Conway T (1998) What's for dinner?: Entner-Doudoroff metabolism in *Escherichia coli*. J Bacteriol 180:3495–3502
21. Shapiro JA, Dworkin M (eds) (1997) Bacteria as Multicellular Organisms. New York: Oxford University Press
22. Strauss EJ, Falkow S (1997) Microbial pathogenesis: genomics and beyond. Science 276:707–712
23. Woese CR (1994) There must be a prokaryote somewhere: microbiology's search for itself. Microbiol Rev 58:1–9
24. Woese C (1998) The universal ancestor. Proc Natl Acad Sci USA 95:6854–6859
25. Zhang C-C (1996) Bacterial signalling involving eukaryotic-type protein kinases. Mol Microbiol 20:9–15