The complete genome sequence of *Mycobacterium tuberculosis*, along with novel genetic tools, provides the foundation for a new era of post-genomic research. The challenge is now to translate these opportunities into an improved understanding of the complex biology of tuberculosis infection.

A post-genomic perspective

DOUGLAS B. YOUNG

From being a gleam in the eye of Ron Davis at World Health Organization meetings in the mid-1980s, the mycobacterial genome program came of age with the publication of the complete sequence of *Mycobacterium tuberculosis* in June 1998 (ref.1). How do things stand in the third year of the post-genome era?

Genome sequencing has continued apace. The initial H37Rv sequence—from a strain maintained in laboratory culture for the best part of a century—can now be compared with that of an *M. tuberculosis* isolate involved in a recent US outbreak (http://www.tigr.org/tdb). Having provided the initial stimulus for much of mycobacterial genomics, the *Mycobacterium leprae* sequence is now also complete (http://www.sanger.ac.uk/Projects/M_leprae). Sequencing is at an advanced stage for an isolate of *M. bovis* (http://www.sanger.ac.uk/Projects/M_bovis), the agent of bovine tuberculosis, and its attenuated derivative, the bacillus Calmette-Guerin (BCG) vaccine; for the opportunistic pathogen *M. avium* (http://www.tigr.org/tdb); and for *M. smegmatis* (http://www.tigr.org/tdb), a representative of non-pathogenic environmental mycobacteria commonly used in laboratory cloning experiments. These sequences provide a wealth of opportunity for comparative genomics.

In terms of functional genomics, microarrays representing each of the 4,000 open reading frames of the H37Rv sequence have been used to screen for genotypic variation among mycobacterial isolates[2] and to characterize drug-induced changes in transcriptional activity[3]. Analysis of the mycobacterial proteome is underway[4], and *M. tuberculosis* has been adopted as the focus for structural genomics initiatives (http://www.doe-mbi.ucla.edu/TB). Techniques for targeted and random mutagenesis are now available, with a series of vectors and protocols for the stable introduction of genes into slow-growing mycobacteria[5].

Does this powerful post-genomic 'toolbox' presage the end of one of humankind's most persistent predators? Although the success in 'fast-tracking' *M. tuberculosis* to the forefront of bacterial genetics is a formidable achievement, considerable challenges remain in translating this into new insights into human disease. In the short term, the unprecedented post-genomic catalog of the 'nuts and bolts' of the tubercle bacillus provides clear opportunities for drug development. Many of the first-generation anti-mycobacterial agents have been shown to interfere with synthesis and assembly of components that make up the unique mycobacterial cell wall, and these should similarly provide reliable targets for new agents to stem the rising tide of multidrug-resistant isolates[6]. To go further than this—to develop interventions offering a substantial advance on those that have so far failed to control global tuberculosis—it will be necessary to dig deeper into the biology of the interaction between humans and *M. tuberculosis*. This provides a challenge to stretch the post-genomic imagination.

Comparative genomics: a window on evolution

M. tuberculosis is inextricably intertwined with humankind in terms of both physiology and social development. Almost certainly, *M. tuberculosis* or its earlier relatives were an important influence on the development of our immune system and, equally, the process of adaptation to living with the human immune system must be written in the genome of *M. tuberculosis*. Can we make use of comparative genomics to read this story? Extensive sequence comparison of selected genes identifies the *M. tuberculosis* complex (which includes human and bovine tubercle bacilli) as a highly related family of organisms. The near-total absence of single nucleotide polymorphisms indicates their derivation from a relatively recent evolutionary 'bottleneck' within the last 10,000–15,000 years[7]. More extensive diversity is demonstrated, however, by whole-genome comparisons. Set against the *M. tuberculosis* H37Rv sequence, the *M. bovis* genome shows a series of deletions resulting in the loss of around 100 genes[2,8]. Similarly, the attenuation of *M. bovis* during the derivation of BCG vaccine strains was associated with deletion events[2], and it is becoming evident that clinical isolates of *M. tuberculosis* also differ in the presence or absence of 5- to 10-kilobase 'chunks'[9]. What is the biological relevance of these genetic deletions? Do they represent an 'engine' for generation of biological diversity, or are they simply redundant fragments of information no longer essential to the current 'lifestyle' of the bacteria? Genome 'downsizing' associated with loss of redundant genes is exemplified by *M. leprae*, the related, highly specialized human pathogen that has scrambled or discarded almost half its genome[10]. Attempts to link comparative mycobacterial genomics with biological function are at the very early stage of research, but the combination of bacterial genetics with careful epidemiological and clinical characterization represents an exciting and relatively direct interface between post-genomics and human disease.

One observation from the comparative sequence analysis is that genes encoding antigens thought to be targets of protective immune responses show no diversity[11]. It would be expected that escape from immune pressure would confer some selective advantage on strains with variant sequences, and it is useful to consider why this is not the case. It may be that immune recognition is spread over a sufficient number of antigens that the advantage associated with modification of a single target is insufficient to drive selection. Alternatively, perhaps the antigens now being studied are not in fact essential for protection. The *M. tuberculosis* genome contains two extensive sets of genes of unknown function, called the PE and PPE families. These do show some degree of strain variation and have been discussed in terms of a possible contribution to 'antigenic variation'[1]. Members of the PE/PPE family have been identified as targets of T cell-mediated immunity[12], and may be involved in *in vivo* adaptation[13], but technical difficulties associated with their repetitive and high GC content has hindered progress in the detailed characterization of their genotypic and phenotypic diversity. Finally, perhaps immunogenicity has a positive selective advantage for *M. tuberculosis*. The very efficient aerosol transmission of *M. tuberculosis* is mediated by destruction of lung tissue through mechanisms thought to depend on immune activation. Although it guarantees a 'quieter life' within the infected host, loss of immunogenicity might therefore entail a loss of transmissibility. If comparative genomics can tell us what type of immune response suits *M. tuberculosis*, this may give us important clues as to what we might want to augment or reduce by new vaccination procedures.

In search of *in vivo* phenotypes

Central to the dilemma of tuberculosis control is the fact that most cases of disease arise years or decades after the first exposure to infection. For several weeks after initial infection, *M. tuberculosis* battles with the innate immune response of the host, leading to induction of an acquired immune response that in most individuals controls bacterial growth to an extent at least sufficient to prevent progression to clinical disease. Estimates based on skin testing indicate that around one-third of the global population is subject to this immunological stage of infection. The number of individuals who remain subject to microbiological infection—that is, continue to carry viable bacilli—is a question that taxes the hearts and minds of tuberculosis researchers. Evidence that a substantial proportion of immunologically infected individuals are also microbiologically infected comes from a series of autopsy studies in Europe and the United States in the pre-antibiotic era, using microbial culture and guinea pig infection protocols to detect viable *M. tuberculosis* in tissues from individuals with no overt signs of tuberculosis[14,15]. An intervention that targets this clinically latent infection would provide an essential addition to current control measures based on prophylactic vaccination and therapy. Can we use a post-genomic approach to identify the *in vivo* phenotype of mycobacteria that persist in the face of a mature immune response?

The first stage of such an analysis would seem relatively straightforward. From the genome sequence, it is apparent that *M. tuberculosis* is a very adaptable organism, replete with a range of transcriptional regulators, and several studies have demonstrated a change in the pattern of gene expression after transfer from laboratory medium to the intracellular environment offered by macrophage cultures. This is probably a reasonable model for the phenotype involved in the initial acute phase of infection, and its characterization using microarrays and proteomics may uncover new approaches to stimulate effective early immunity. An analogous approach to the characterization of the late-stage bacterial population expectorated in the sputum of tuberculosis patients may identify useful drug targets missed in simple *in vitro* screens. However, analysis of the persistent phenotype presents a more difficult challenge.

Prolonged asymptomatic carriage could reflect a static non-dividing population of bacteria, or a chronic infection during which active replication is countered by continuous immune killing. Non-dividing bacteria could be analogous to stationary-phase cultures or, at least conceptually, some more highly differentiated 'dormant' form of *M. tuberculosis*. Distinguishing between these possibilities is not simple, and different mechanisms need not be mutually exclusive. Given the difficulties associated with direct study of persistence in human, post-genomic analysis will require initial practice in model systems. Non-dividing cultures of *M. tuberculosis* are readily established by nutrient depletion *in vitro*, for example, with oxygen depletion possibly reflecting an important aspect of the environment within calcified lesions[16]. Persistence has been studied in mice using a protocol in which incomplete drug treatment is used to set up an infection that is undetectable in terms of bacterial culture from organ homogenates but can be reactivated by steroid-induced immune suppression[17]. This model may be particularly useful as a means of studying mycobacterial persistence during chemotherapy, but simpler models may be preferable for more general questions. Using a moderately low level of inoculum, the normal course of *M. tuberculosis* infection in untreated mice is characterized by an initial increase in bacterial load, followed by immune-mediated control and a subsequent chronic phase during which bacterial numbers remain constant

Post-genomic explorers set out for the Magic Mountain
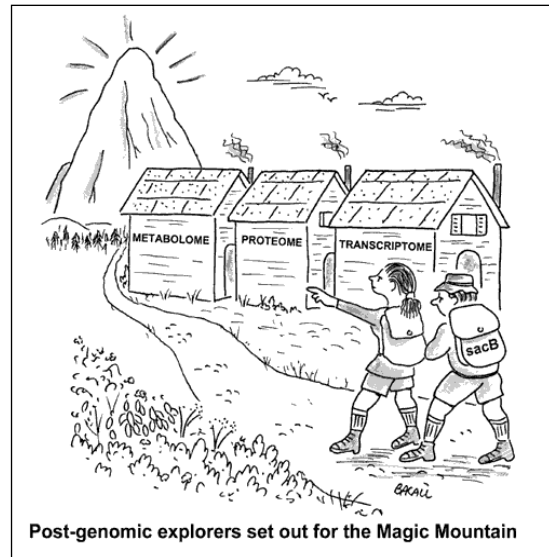
over a period of several months. The bacterial load during the chronic phase can be regulated by inoculum size or by prior BCG vaccination. This provides a useful model to study the protracted 'stalemate' between the forces of microbial virulence and host immunity that is the fundamental characteristic of tuberculosis infection. The bacterial population involved in the chronic infection is generally considered to be non-dividing—based on the acquisition of thermotolerance[18] and the absence of mycobacterial carcasses predicted to accumulate from steady-state replication and killing[19]—but this issue warrants more detailed analysis. An important attribute of this simple model is its accessibility to post-genomic analysis; two recent studies have exploited this model to identify genes required for chronic survival[20,21].

Whatever the model systems, direct application of post-genomic tools to human material will ultimately be essential to the understanding of persistence. It seems likely that this could best be accomplished by revisiting the autopsy approach used in the first half of the twentieth century; coordinating the skills of surgeons, pathologists and molecular biologists to facilitate detailed analysis of mycobacterial persistence in relevant biopsy or autopsy samples.

Post-genomic horizons

If post-genomics is envisaged as the road that leads from genome sequence through transcriptomes, proteomes and metabolomes to the land of biologically relevant phenotypes, then the ground underfoot at present would seem firm and reliable, although the phenotypes remain something of a shimmer on the far horizon. Although the considerations presented herein reflect the pathogen perspective, making sense of tuberculosis will undoubtedly rest on our ability to integrate mycobacterial post-genomics with human post-genomics.

1. Cole, S.T. *et al.* Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537–544 (1998).
2. Behr, M.A. *et al.* Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* **284**, 1520–1523 (1999).
3. Wilson, M. *et al.* Exploring drug-induced alterations in gene expression in *Mycobacterium tuberculosis* by microarray hybridization. *Proc. Natl. Acad. Sci. USA* **96**, 12833–12838 (1999).
4. Jungblut, P.R. *et al.* Comparative proteome analysis of *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG strains: towards functional genomics of microbial pathogens. *Mol. Microbiol.* **33**, 1103–1117 (1999).
5. Hatfull, G.F. & Jacobs, W.R. *Molecular Genetics of Mycobacteria* (ASM Press, Washington, DC, 2000).
6. Duncan, K. & Sacchettini, J.C. in *Molecular Genetics of Mycobacteria* (eds. Hatfull, G.F. & Jacobs, W.R.) 297–307 (ASM Press, Washington, DC, 2000).

7. Sreevatsan, S. *et al.* Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc. Natl. Acad. Sci. USA* **94**, 9869–9874 (1997).

8. Gordon, S.V. *et al.* Identification of variable regions in the genomes of tubercle bacilli using bacterial artificial chromosome arrays. *Mol. Microbiol.* **32**, 643–655 (1999).

9. Ho, T.B.L., Robertson, B.D., Taylor, G.M., Shaw, R.J. & Young, D.B. Comparison of *Mycobacterium tuberculosis* genomes reveals frequent deletions in a 20kb variable region in clinical isolates. *Yeast* **17**, 272–282 (2000).

10. Brosch, R., Gordon, S.V., Eiglmeier, K., Garnier, T. & Cole, S.T. Comparative genomics of the leprosy and tubercle bacilli. *Res. Microbiol.* **151**, 135–142 (2000).

11. Musser, J.M., Amin, A. & Ramaswamy, S. Negligible genetic diversity of *Mycobacterium tuberculosis* host immune system protein targets: evidence of limited selective pressure. *Genetics* **155**, 7–16 (2000).

12. Dillon, D.C. *et al.* Molecular characterization and human T-cell responses to a member of a novel *Mycobacterium tuberculosis* mtb39 gene family. *Infect. Immun.* **67**, 2941–2950 (1999).

13. Ramakrishnan, L., Federspiel, N.A. & Falkow, S. Granuloma-specific expression of Mycobacterium virulence proteins from the glycine-rich PE-PGRS family. *Science* **288**, 1436–1439 (2000).

14. Opie, E.L. & Aronson, J.D. Tubercle bacilli in latent tuberculous lesions and in lung tissue without tuberculous lesions. *Arch. Pathol.* **4**, 1 (1927).

15. Rich, A.R. *The Pathogenesis of Tuberculosis* (Blackwell Scientific Publications, Oxford, 1951).

16. Wayne, L.G. & Lin, K.Y. Glyoxylate metabolism and adaptation of *Mycobacterium tuberculosis* to survival under anaerobic conditions. *Infect. Immun.* **37**, 1042-1049 (1982).

17. McCune, R.M., Feldman, F.M., Lambert, H.P. & McDermott, W. Microbial persistence. I. The capacity of tubercle bacilli to survive sterilization in mouse tissues. *J. Exp. Med.* **123**, 445–468 (1966).

18. Wallace, J. G. 1961. The heat resistance of tubercle in the lungs of infected mice. *Am. Rev. Respir. Dis.* **83**, 866–871 (1961).

19. Rees, R.J.W. & Hart, P.D. Analysis of host-parasite equilibrium in chronic murine tuberculosis by total and viable bacillary counts. *Br. J. Exp. Pathol.* **42**, 83–88 (1961).

20. Glickman, M.S., Cox, J.S. & Jacobs, W.R.Jr. A novel mycolic acid cyclopropane synthetase is required for coding, persistence, and virulence of *Mycobacterium tuberculosis. Mol. Cell* **5**, 717–727 (2000).

21. McKinney, J.D. *et al.* Persistence of *Mycobacterium tuberculosis* in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. *Nature* **406**, 735–738 (2000).

Imperial College, London W2 1PG, UK
Email: d.young@ic.ac.uk

The World Health Organization estimates that tobacco will become the largest single health problem by 2020, causing an estimated 8.4 million deaths annually. But the smoking burden will not be distributed evenly across the globe; deaths in developed nations are set to rise 50% to 2.4 million while those in Asia will soar fourfold to an estimated 4.2 million in 2020. In the face of such discrepancy, Martin Raw, Honorary Lecturer in evidence-based treatment at Guys, Kings and St Thomas' School of Medicine, London, explains why attention can not be focused solely on Asia and why efforts are still needed to stop smoking in Europe.

Fighting tobacco dependence in Europe

Martin Raw

Following the publication of clinical guidelines on smoking cessation, in 1996 in the United States and 1998 in the United Kingdom, and the development of government funded treatment services in Britain for dependent smokers, the World Health Organization (WHO) has developed a project to reduce the effects of tobacco dependence in its 51 member state European region.

The project is called the WHO European Partnership Project to Reduce Tobacco Dependence. It is an ambitious endeavor, the overall goal of which is to reduce the burden of disease caused by tobacco dependence in Europe, by promoting smoking cessation and related activities. These include supporting the Framework Convention on Tobacco Control, a treaty which UN member states will sign and which aims to control global aspects of the tobacco problem (like advertising and smuggling).

In the past tobacco control has tended to lose out to more urgent or high profile health problems, such as tuberculosis and AIDS, partly because it was thought to be a lifestyle choice rather than an addiction—a perception strongly pushed by the tobacco industry. What is different now is that public health professionals have joined forces with private partners to acknowledge the addictive nature of tobacco and the health problems caused by smoking. For the first time ever, resources are available that enable us to tackle the problem on a slightly more equal footing. This is no small consideration when you are fighting an industry whose annual profits are greater than the gross national products of many countries.

The Partnership Project is enabling work to take place that simply would not have happened otherwise. This is certainly true in my field—smoking cessation—where it is extremely difficult to find funds to work internationally. The project has brought together a broad coalition of public and private partners and is open to any appropriate partners, public or private, commercial or non-commercial. It has been funded so far mainly by GlaxoWellcome, Novartis, Pharmacia, and SmithKline Beecham, each of which manufactures a smoking cessation product such as a nicotine patch, for example. It is managed and led by WHO, and supported also by representatives of a number of tobacco control advocates and pan-European organizations including consumers' organizations and health-professional bodies. Partners also include national government representatives of its initial target countries. The project includes five main activity areas: monitoring smoking status, intention to change, and impact of policy changes; regulation of tobacco and of treatment products; promoting smoke free places; promoting evidence based treatment; and communicating the message.

According to the World Bank, tobacco consumption is only falling in developed countries, while it is rising steadily in developing countries. Though annual per capita cigarette consumption has fallen slightly in the developed world, from about 2,700 to 2,600 since 1970, over the same period it has risen from about 800 to 1,400 in developing countries[1]. So, why is the Partnership Project only in Europe at the moment, when tobacco consumption has been rising in low- and middle-income countries (in Asia for example) since the 1970s? The answer is that this was the region in which it was felt that maximum benefit could be gained by sharing the experiences of countries that have made good progress in tobacco control, but which still suffer high smoking prevalence and disease rates and in some cases, are still influenced by an extremely and insidiously powerful tobacco industry.

Just as with China, the former communist states in the East face huge problems and costs in re-construction and some of them will in the next decades be spending more and more treating smoking-related diseases. Life expectancy in Eastern Europe has fallen disturbingly in the last decade. Smoking is exceedingly pervasive in Germany where there are no services to help smokers stop and no political interest at all in the issue.