A GENOMICS-BASED APPROACH TO BIODEFENCE PREPAREDNESS

Claire M. Fraser

The anthrax letter attacks in October 2001, followed by the SARS outbreak in early 2003, dramatically illustrated our vulnerability to both deliberate and natural outbreaks of infectious disease. The availability of pathogen genome sequences and high-throughput methods for studying the biology of both pathogens and their hosts have provided new insights into the mechanisms of pathogenesis and host defence. As infectious disease research expands to include major bioterror agents, genomics-based approaches will provide one of the cornerstones of efforts to develop more accurate diagnostics, new therapeutics and vaccines, and further capabilities for microbial forensics.

CATEGORY A AGENTS
Organisms that pose a risk to national security because they: can be easily transmitted from person to person; result in high mortality rates and have the potential for major public health impact; might cause public panic and social disruption; and require special action for public health preparedness.

CATEGORY B AGENTS
Organisms that are moderately easy to disseminate, result in moderate morbidity rates and low mortality rates. They require specific enhancements of the Center for Disease Control and Prevention's diagnostic capacity and enhanced disease surveillance.

The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, Maryland 20850, USA.
e-mail: cmfraser@tigr.org
doi:10.1038/nrg1245

Since the anthrax letter attacks in October 2001, the US federal government has placed a premium on homeland security, with research and development considered to be a key component of an integrated and unified biodefence strategy. For fiscal year 2003, President Bush proposed a $1.75 billion budget for biodefence research, to be administered primarily through the National Institutes of Allergy and Infectious Diseases (NIAID) at the National Institutes of Health (NIH). As summarized on the NIAID web site (see online links box), the proposed new biodefence research agenda will broadly focus on studies of CATEGORY A, CATEGORY B and CATEGORY C BIOTERROR AGENTS (for more information, see online links box). These potential bioterror agents share several features, including a high morbidity or mortality that is associated with infection, a high likelihood of person-to-person spread, the ability to be made into a weapon and spread by aerosols or food and water supplies, and the potential to cause widespread panic in the public if released into the environment. The goals of this accelerated research and development agenda are to: understand better the molecular mechanisms that are responsible for pathogen transmission, virulence and invasion; develop new animal models of disease that can be used in the study of potential bioterror agents; elucidate host responses to microbes to understand better the complex parameters of innate and acquired immunity, and how they might be manipulated after exposure to potential bioterror agents; develop new vaccines, antimicrobial and antiviral therapeutics, and diagnostic tools against the most threatening diseases, such as smallpox, anthrax, and plague; and to develop research resources such as BSL3 and BSL4 LABORATORY containment facilities to allow more work on category A–C agents than is possible at present. Ultimately, the expectation is that these new research activities will increase our level of preparedness so that we might better respond to, and protect the world's population against, bioterrorism.

Much needs to be done to realize the ambitious goals of the NIAID Strategic Plan for Biodefense Research — this includes not only an expanded research portfolio and the development of new tools and reagents, but also the creation of new research facilities and the recruitment of new investigators into the area of infectious disease research. On 4 September 2003, the US Department of Health and Human Services announced the creation of eight Regional Centers of Excellence for Biodefense and Emerging Diseases (RCEs). The aim of the RCEs is to bring together investigators from several disciplines and public and private institutions to tackle the most compelling research questions that are relevant to bioterror agents, particularly anthrax and smallpox, in large, comprehensive programmes (for more information on RCEs, see online links box). Collectively, what sets this new research agenda apart from previous efforts on the study of infectious disease agents is the

expectation that in a relatively short period of time these expanded research activities must achieve several key objectives that can provide immediate benefits in the public health arena. It is also anticipated that one of the other positive spin-offs of the new bioterrorism research agenda will most certainly be a better understanding of other more common and naturally occurring infectious agents. This is an important goal because we are at least as vulnerable to emerging infectious diseases as we are to deliberate attacks. The recent severe acute respiratory syndrome (SARS) epidemic that in 9 months killed over 900 people in 31 countries illustrates just how devastating a new disease can be, and how quickly it can spread around the world (for more information on SARS, see online links box).

Barry Bloom, the present Director of the Harvard School of Public Health, summarized the potential of genomics-based approaches to expand our understanding of the biology of pathogens in 1995 (REF. 1) after the publication of the second complete microbial genome sequence. He stated that "The power and cost effectiveness of modern genome sequencing technology mean that complete genome sequences of 25 of the major bacterial and parasitic pathogens could be available within 5 years. For about $100 million we could buy the sequence of every virulence determinant, every protein antigen, and every drug target. It would represent for each pathogen a one-time investment from which the information derived would be available to all scientists for all time. We could then think about a new post-genomic era of microbe biology." (REF. 1). As genomics, proteomics and bioinformatics are considered to be key enabling technologies in the development of new methods to deal with potential bioterror agents and emerging infectious diseases, this review summarizes the status of pathogen genome sequencing and analysis at present and discusses how these approaches might best be applied to biodefence preparedness.

Vulnerabilities and assessment of needs

The intentional spread of disease during war (biowarfare) has a long history that dates back to the ancient Greeks and Romans (BOX 1). Bioterrorism, by contrast, has a much shorter history. Recent events, in particular the anthrax letter attacks of 2001, have exposed our vulnerabilities in the area of BIOPREPAREDNESS against both natural and deliberate outbreaks of infectious disease. As there is a dearth of investigators who study biowarfare pathogens and only a few specialized facilities for carrying out such research, we only have a rudimentary understanding of the mechanisms of pathogenesis for most of these agents. We also lack rapid and precise diagnostic assays for identifying most of these species. The deliberate release of microbial pathogens can be done, for the most part, without being noticed, until significant numbers of symptomatic patients appear at

CATEGORY C AGENTS
These third highest priority agents include emerging pathogens that could be engineered for mass dissemination in the future because of their availability, ease of production and dissemination, and their potential for high morbidity and mortality rates and major health impact.

BSL3 LABORATORY
Laboratory facilities for work on biological agents that might cause serious or potentially lethal disease as a result of exposure by the inhalation route.

BSL4 LABORATORY
Laboratory facilities for work with dangerous and exotic biological agents that pose a high individual risk of aerosol-transmitted laboratory infections and life-threatening disease.

BIOPREPAREDNESS
A state of adequate preparation in case of bioterror attacks that will allow for a rapid and efficient response to contain the spread of the agent, minimize morbidity and mortality, and minimize the disruption to public infrastructure.

GLANDERS
A contagious disease of horses, mules and donkeys that is caused by the bacterium *Burkholderia mallei*. It can also be transmitted to humans who come into contact with infected animals.

Box 1 | **History of biowarfare and bioterrorism**

Biological warfare began with the ancient Greeks and Romans, who used toxic plants and human and animal corpses to poison drinking-water wells. More recently (in the fourteenth century) the Tartars catapulted bodies of plague victims over the city walls to infect the Genoese inhabitants of Caffa, a trade centre on the Black Sea. The Genoese took to their boats to escape, spreading the Black Death along shipping routes throughout Europe. In just a few years the plague killed approximately one-third of Europe's population.

The seventeenth and eighteenth centuries saw French and British soldiers intentionally infect native Americans with European diseases. For example, Sir Jeffrey Amherst ordered that smallpox be spread as a way "…to inoculate the Indians by means of blankets … to extirpate this execrable race…"[63].

During the First World War, Germany had plans to conduct covert operations in the United States with the intent to infect horses and cattle that were destined for service in Europe with GLANDERS and anthrax. During the Second World War, bioweapons programmes were initiated in every major participating country. The Japanese military practiced biowarfare on a large scale against the Chinese in Manchuria during the 1940s, infecting the population with various disease agents including anthrax, cholera and plague. The Cold War brought about an even greater escalation in the bioweapons programmes in both the United States and the Soviet Union, until President Nixon terminated the offensive bioweapons programme in the United States in 1969 and ordered that all stockpiled weapons be destroyed. In 1972, the United States and more than 100 nations signed the Biological and Toxin Weapons Convention, which was the first treaty to ban an entire class of weapons.

Bioweapons work continued in the Soviet Union — an outbreak of inhalation anthrax in Sverdlosk in 1979 was linked to secret weapons work in a nearby laboratory[64]. Soviet defectors in the 1980s confirmed that the Soviet Union was working on creating genetically engineered strains of pathogens that were resistant to antibiotics and vaccines. With the break-up of the Soviet Union in the mid 1980s, many of the Soviet scientists who carried out this work disappeared and resurfaced in countries such as Iraq, which launched its own bioweapons programme at around the same time.

Bioterrorism has a much shorter history. In 1984, followers of the guru Bhagwan Shree Rajneesh deliberately contaminated salad bars throughout one county in Oregon, USA, with *Salmonella* — the goal was to sicken a sufficient number of people to prevent them voting in an election[65]. More than 750 cases of food poisoning were reported, and it took a year to discover the source of the contamination. In 1995, the Aum Shinrikyo cult released sarin gas in a subway in Tokyo, Japan, killing 12 people and injuring thousands. Between 1993 and 1995, the same cult tried to spray botulinum toxin and anthrax in Kameido, a city near Tokyo, but was not successful[65,66]. Most recently, in October 2001, anthrax attacks were carried out in the United States using the US postal service as the delivery vehicle. This resulted in 5 deaths and 18 cases of anthrax infection, and shut down the postal service and the US Congressional offices for a time. The perpetrator of this crime is still at large.

medical facilities. Even then, the earliest symptoms that are associated with infection by biowarfare pathogens are nonspecific and can easily masquerade as something less serious. Once an attack is detected, the panic in the general population could overwhelm the public health-care system, making it more difficult to identify and treat those individuals who have actually been exposed.

Although recommendations for administering antimicrobial therapy after bioterror attacks with category A agents have been proposed by the Working Group on Civilian Biodefense[2–4], in many cases there is limited data available on the efficacy of such dose regimens. In some instances, three to eight weeks of antibiotic therapy is recommended; however, one of the dangers inherent in such a prolonged course of treatment is non-compliance and the emergence of antibiotic-resistant pathogens. With the continued emergence of antibiotic-resistant strains of more common pathogens as well, it is clear that we would greatly benefit from an expanded arsenal of new antibiotics.

Another alternative for combating infectious diseases is to develop new or better vaccines that confer long-term immunity or that could be used prophylactically after a bioterror attack. Such new generation vaccines must be able to protect all groups of civilians, including the large number of immuno-compromised patients who are present in the population. The recent reports of unexpected cardiac problems, including two deaths, which occured after vaccination of first-responders with the present smallpox vaccine[5,6], and numerous reports of adverse reactions to the present vaccine against anthrax[7–10] provide examples of how much extra work is needed.

Finally, there is an urgent need to understand better the genetic variation among, and the global distribution of, bacterial species and strains. The microbial genome is a dynamic entity that is shaped by numerous forces, including gene duplication and gene acquisition through lateral gene transfer, which influence the evolution of microbial species[11]. The plasticity of the microbial genome has important implications in epidemiologic studies, the spread of antibiotic resistance and pathogenicity, microbial forensics and the development of new therapeutics and vaccines.

Genomics efforts on pathogens

Despite the fact that the field of microbial genomics is still relatively young, tremendous progress has been made in a short amount of time. Essentially all of the 25–30 principal human bacterial pathogens, including many of those on the Center for Disease Control and Prevention (CDC) category A–C lists, have now been sequenced. Sequence data and associated genome annotation from these efforts are in the public domain (see the TIGR Microbial Database, online links box). In several cases, sequence data from more than one representative of the same species is available (for example, *Escherichia coli*, *Staphylococcus aureus*, *Mycobacterium tuberculosis*, *Bacillus anthracis*, *Helicobacter pylori* and *Chlamydia pneumoniae*) (TABLE 1), and this has proved extremely useful in providing new insights into the genetic variability that is present among strains.

SYNTENY
The conservation of gene order in chromosomes that are conserved over wide evolutionary distances.

Comparative genomics. Genome sequence data provide a powerful new starting point for follow-up investigations into the biology of microbial pathogens. Through the power of comparative genomics *in silico* it is possible to glean insights into new virulence genes, the molecular basis and evolution of pathogenicity, the diversity within closely related isolates, and to formulate hypotheses to be tested experimentally. For example, the sequencing of the *B. anthracis* genome resulted in the identification of several new putative virulence genes, some of which have recently been shown to encode functional toxins[12]. Several other *B. anthracis* genes have been identified that encode proteins that are related to proteins that are responsible for virulence in closely related *Bacillus* species that infect insects[13]. It is not clear whether these genes have the same role in *B. anthracis* as they do in the other species, nonetheless, they are obvious candidates for future experimental studies. The importance of such studies cannot be overemphasized, because bioinformatics approaches alone are not sufficient to unequivocally predict protein function. The identification of new virulence and pathogenicity genes in bioterror pathogens not only provides new insights into how these agents cause disease, but also potentially provides new targets for antimicrobial and vaccine development.

The comparison of the recently completed genome sequence of the intracellular pathogen and potential bioterrorism agent *Brucella suis*[14] with that of *Brucella melitensis*[15] defined a finite set of differences that could be responsible for the differences in virulence and host preference that are present between these closely related organisms. Analysis of the *B. suis* genome also revealed transport and metabolic abilities similar to those that are seen in soil/plant-associated bacteria[14], a similarity that is also evident in the extensive gene SYNTENY that is observed between *B. suis* chromosome 1 and the genome of the plant symbiont, *Mesorhizobium loti*[16]. These results illustrate the advantages that can come from taking a more broad view of the study of pathogenesis — new data on the biology of one pathogen could be directly applicable to the study of what might seem to be an unrelated pathogen.

Transcriptomics/proteomics. Although comparative *in silico* approaches can reveal the molecular differences that distinguish related species and identify potential virulence genes, they alone are not sufficient to uncover the complexities of the interaction between pathogen and host. However, the availability of large-scale approaches for transcriptome analysis are beginning to have an impact in infectious disease research by allowing investigators to move beyond the study of single genes to probe global changes in RNA expression[17]. Although this technology was only available in a small number of laboratories just a few years ago, it is now more widely available and is being used to study gene expression in both pathogen and host during different stages of infection. One of the most important advantages of this type of approach is that it allows the study of both known genes and new genes of unknown function.

Table 1a | **Pathogen genome sequences completed**

Species	Disease	Genome size (Mb)
Alpha proteobacteria		
Brucella melitensis 16M*	Brucellosis	3.29
Brucella melitensis biovar *suis**	Brucellosis	3.31
Rickettsia conorii	Mediterranean spotted fever	1.27
Rickettsia prowazekii Madrid E*	Typhus	1.11
Beta proteobacteria		
Bordetella bronchiseptica RB50 NCTC-13252	Respiratory infections (animals)	5.34
Bordetella parapertussis 12822 NCTC-13253	Whooping cough	4.77
Bordetella pertussis Tohama I NCTC-13251	Whooping cough	4.09
Neisseria meningitidis MC58 (serogroup B)	Meningitis	2.27
Neisseria meningitidis Z2491 (serogroup A)	Meningitis	2.18
Chlamydias		
Chlamydia trachomatis D/UW-3/CX	Sexually transmitted disease	1.04
Chlamydia trachomatis MoPn/Nigg	Pneumonia (mice)	1.06
Chlamydophila caviae GPIC	Conjunctivitis (guinea pigs)	1.17
Chlamydophila pneumoniae AR39	Pneumonia, bronchitis	1.23
Chlamydophila pneumoniae CWL029	Pneumonia, bronchitis	1.23
Chlamydophila pneumoniae J138	Pneumonia, bronchitis	1.23
Chlamydophila pneumoniae TW-183	Pneumonia, bronchitis	1.23
Epsilon proteobacteria		
Camplyobacter jejuni subspecies *jejuni* NCTC11168	Food-borne gastroenteritis	1.64
Helicobacter hepaticus ATCC51449	Hepatitis/liver cancer (mice)	1.80
Helicobacter pylori 26695	Stomach ulcer, gastric cancer	1.67
Helicobacter pylori J99	Stomach ulcer, gastric cancer	1.64
Fusobacteria		
Fusobacterium nucleatum	Periodontal disease (facilitator)	2.17
Gamma-proteobacteria		
Coxiella burnetii RSA 493*	Q fever	2.10
Escherichia coli 0157:H7 *Sakai**	Food-borne gastroenteritis	5.5B
Escherichia coli UPEC-CFT073	Urinary tract infections	5.23
Haemophilus ducreyi 35000HP	Chancre	1.69
Haemophilus influenzae Rd KW20	Otitis media	1.83
Pasteurella multocida Pm70	Multiple diseases and species	2.25
Pseuodomonas aeruginosa PA01	Opportunistic infections	6.26
Salmonella enterica serovar *Typhi* CT18*	Typhoid fever	4.81
Salmonella enterica Typhi T2	Typhoid fever	4.79
Salmonella typhimurium LT2 SGSC1412	Gastroenteritis	4.86
Shigella flexneri serotype 2a 2457T	Dysentery/diarrhoea	4.59
Shigella flexneri serotype 2a 301	Dysentery/diarrhoea	4.61
Vibrio cholerae serotype O1, biotype El Tor, strain N16961*	Cholera	4.00
Vibrio parahaemolyticus	Food-borne gastroenteritis	5.16
Yersinia pestis CO-92 (biovar *orientalis*)‡	Plague	4.65
Yersinia pestis KIM5 P12 (biovar *mediaevalis*)‡	Plague	4.60

*CDC category B agent. ‡CDC category A agent. CDC, Center for Disease Control and Prevention; Mb, megabase pairs.

Recent studies of *Neisseria meningitidis*, which is a causative agent of septicaemia and meningococcal meningitis, provide an excellent example of how useful transcriptome analysis can be[18,19]. These studies showed that distinct sets of genes were differentially regulated during two key steps in the meningococcal infection of human cells — the initial interaction with epithelial cells in the respiratory tract and the later interaction with endothelial cells in the blood–brain barrier. These differentially regulated genes — which include those that encode membrane transporters, transcription factors, general metabolic pathways and several hypothetical

Table 1b | **Pathogen genome sequences completed**

Species	Disease	Genome size (Mb)
Gram-positive		
Mycobacterium bovis AF2122/97	Tuberculosis (man and animals)	4.35
Mycobacterium leprae TN	Leprosy	3.28
Mycobacterium tuberculosis CDC 1551	Tuberculosis	4.50
Mycobacterium tuberculosis 37 Rv	Tuberculosis	4.41
Tropheryma whipplei TW08/27	Whipple's disease	0.925
Tropheryma whipplei Twist	Whipple's disease	0.927
Low GC Gram-positive		
Bacillus anthracis Ames‡	Anthrax	5.23
Bacillus cereus ATCC 14579	Opportunisitic infections	5.41
Clostridium perfringens 13§	Gas gangrene	3.03
Clostridium tetani Massachusetts E88	Tetanus	2.80
Enterococcus faecalis	Opportunisitic/hospital-acquired infections	3.21
Listeria monocytogenes EGD-e	Listeriosis/food-borne gastroenteritis	2.94
Mycoplasma genitalium G-37	Urethritis, arthritis	0.58
Mycoplasma penetrans HF-2	Urogential/respiratory infections	1.36
Mycoplasma pneumoniae M129	Pneumonia	0.816
Mycoplasma pulmonis UAB CTIP	Respiratory infections (mice)	0.963
Staphylococcus aureus Mu50 (VRSA)§	Community/hospital-acquired infections	2.88
Staphylococcus aureus N315 (MRSA)	Community/hospital-acquired infections	2.81
Staphylococcus aureus subspecies *Aureus*	Community/hospital-acquired infections	2.82
Staphylococcus epidermidis ATCC 12228	Community/hospital-acquired infections	2.49
Streptococcus agalactiae 2603V/R	Neonatal infections	2.16
Streptococcus agalactiae NEM316	Neonatal infections	2.21
Streptococcus mutans UA159	Dental caries	2.03
Streptococcus pneumoniae R6	Pneumonia/meningitis/bacteraemia	2.04
Streptococcus pneumoniae TIGR4	Pneumonia/meningitis/bacteraemia	2.16
Streptococcus pyogenes M1 GAS SF370	Suppurative infections	1.85
Streptococcus pyogenes M18 MGAS8232	Suppurative infections	1.89
Streptococcus pyogenes M3 (SSI-1)	Suppurative infections	1.89
Streptococcus pyogenes M3 MGAS315	Suppurative infections	1.90
Ureaplasma urealyticum (parvum) serovar 3	Mucosal pathogen	0.751
Spirochetes		
Borrelia burgdorferi B31	Lyme disease	1.50
Leptospira interrogans serovar *Lai*	Leptospirosis	4.69
Treponema pallidum subspecies *pallidum Nichols*	Syphilis	1.14

‡CDC category A agent. §The toxin is a CDC category B agent. CDC, Center for Disease Control and Prevention; GC, guanine plus cytosine; Mb, megabase pairs; MRSA, methicillin-resistant *Staphylococcus aureus*; VRSA, vancomycin-resistant *Staphylococcus aureus*.

proteins — are obvious candidates for further studies, which in turn could lead to new approaches to preventing diseases that are caused by *N. meningitidis*. Studies like these, which identify suites of genes that are differentially expressed at different stages of infection, could similarly lead to new biodefence strategies if applied to bioterror agents such as *B. anthracis*.

In parallel with transcriptome studies, highly sensitive two-dimensional gel electrophoresis and protein identification by MASS SPECTROMETRY are being used to explore the proteomes of several pathogens. Such analyses can be used to identify differences in protein expression between virulent and avirulent strains[20] and potential vaccine candidates[21]. Although proteomics approaches are still limited by factors such as sensitivity and scalability, they are particularly well suited to identifying where proteins are localized in the pathogen (cytoplasmic versus membrane versus secreted proteins), which might be a key step towards understanding the pathogenesis of numerous potential bioterror agents. Of particular interest with regard to the study of bioterror pathogens are the results of a recent study, in which proteomic analysis was used to study the spore coat proteins of *B. anthracis* and a related but non-pathogenic species, *Bacillus subtilis*[22]. In response to external stress, both of these bacilli form spores that

MASS SPECTROMETRY
Analysis using an analytical instrument that provides accurate information about the molecular mass and structure of complex molecules. This technique can identify and quantify extremely small amounts of peptide by their mass-fragment spectrum.

allow them to persist in the environment under extremes of temperature, dessication and time[23]. In the case of *B. anthracis*, the interaction of the spore with the host is essential for infection[24]. Proteome analysis uncovered a set of conserved spore coat proteins and sets of new spore coat proteins that are unique to each species of *Bacillus*[22]. The identification of unique spore coat proteins in *B. anthracis* could accelerate efforts to develop new methods for the detection of this organism in the environment, as well as provide new targets through which the interaction of the pathogen with host macrophages might be disrupted.

Host responses. Another important research area for biodefence is the study of host responses that are triggered by infections of potential bioterror pathogens. The availability of the human genome sequence and human DNA arrays means that the types of studies that are required are now possible. For example, one recent study[25] found marked differences in the gene-expression profiles of gastric epithelial cells that were exposed to strains of *H. pylori* (a gastric pathogen[26]) with a 40-kilobase PATHOGENICITY ISLAND[27–29] compared with cells that were exposed to strains that lacked this island. Such studies highlight the genes that might be involved in the observed differences in disease outcome after infection. Another study[30] used several approaches, including microarray analysis, to identify 22 genes in murine macrophages that are upregulated in response to infection of cells with *Yersinia pestis*, which is the causative agent of plague. Several of these genes have previously been shown to be involved in apoptosis. These data are consistent with the fact that many pathogens, such as *Y. pestis*, are known to circumvent the immune response of the host by interfering with the normal function of circulating macrophages, in this case by triggering cell death, and they indicate that there are possible interventional strategies that might be pursued in the development of new therapeutics to treat this disease.

DNA arrays have also been used to investigate the response of human peripheral blood mononuclear cells (PBMCs) after exposure to various bacterial pathogens[31]. These cells play a vital surveillance role for detecting exposure to infectious agents and can be considered as the immune system's 'canary in the coal mine'. The exposure of PBCMs from numerous donors to many Gram-positive and Gram-negative bacterial pathogens showed that there is a consistent programme of gene expression that is characterized by the upregulation of many cytokines and chemokines. These results could provide the basis for new diagnostic tests for pathogenic microbes, perhaps even before overt symptoms develop. Such diagnostic tests could be extremely important as biological 'early warning' systems for biodefence purposes.

Genetic characterization of pathogen strains. DNA microarray technology can also be used to assess molecular differences between closely related pathogen strains for which complete genome sequences are available[32–34]. This comparative genome hybridization technology can rapidly be applied to the study of large numbers of isolates, identify subsets of genes that might be responsible for phenotypic differences in pathogenesis and virulence, and advance our understanding of microbial evolution and genome heterogeneity. However, one of the limitations of microarray-based analysis of gene content is that the arrays can only provide information on those genes that are represented on an array. Given the genetic diversity that has been observed in some species of bacteria, such as the 26% difference in gene content between the avirulent *E. coli* K12 strain and the enteropathogenic *E. coli* 0157:H7 strain[35], caution is necessary when interpreting comparative genome hybridization data in the absence of comprehensive information about strain variability.

One of the most exciting applications of comparative genome hybridization approaches might come from the development of more rapid methods for the detection of biological agents, even when they have been genetically engineered. It is not at all unreasonable to contemplate the fabrication of a DNA chip that contains all of the predicted coding sequences from numerous isolates of the most important human, animal and plant pathogens as a first step in the development of new detection technologies. The read-out from such a 'detector' could provide information on the full genetic complement of any bioterror agent, even if it contains genes or plasmids from other species, any unusual properties that are related to virulence or antibiotic-resistance, or is a synthetic organism that has been built from component genes. The ability to quickly identify and characterize a potential bioterror agent in a single assay would greatly reduce the delays that are inherent in present methods of detection. Progress towards this goal has begun with the development of a multi-pathogen identification microarray (MPID) that has been used to identify 18 pathogenic prokaryotes, bacteria and viruses with a high degree of specificity[36], and an oligonucleotide array for the detection of closely related orthopox viruses including variola, which is the causative agent of smallpox[37].

Antimicrobial and vaccine development. There has been an alarming increase in resistance to multiple antibiotics among pathogens such as *S. aureus*, *Streptococcus pneumoniae*, *Enterococcus faecalis* and *M. tuberculosis* in both community and hospital settings over the past 15 years[38–41]. Until quite recently, the primary approach to antibiotic development within the pharmaceutical industry was to make incremental improvements to existing antibiotic classes to stay one step ahead of resistant microbes[42]. At present antibiotics target three cellular processes in the bacterial cell: DNA/RNA synthesis, protein synthesis and cell-wall synthesis. The reason that these three pathways make excellent targets for antimicrobial compounds is that they represent essential cellular functions. By analogy, any other protein that is essential for cell viability becomes a possible target for the development of entirely new classes of antibiotics.

PATHOGENICITY ISLAND
The genetic element within the genome of an organism that is responsible for its capacity to cause disease (its pathogenicity). The virulence of the organism is modulated by the genes that are harboured by this island.

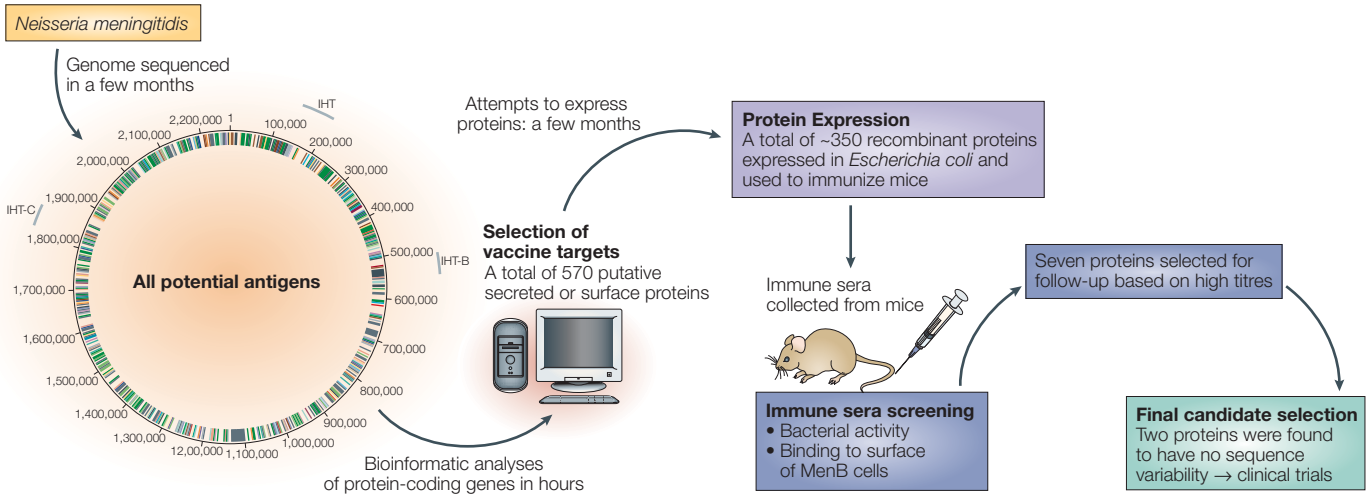


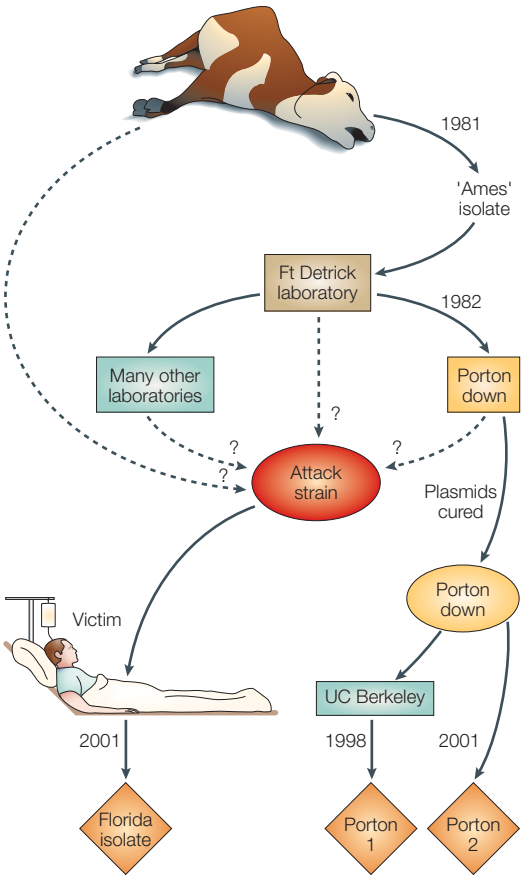
Figure 1 | **Reverse vaccinology: a genomics-enabled approach to vaccine development.** The reverse vaccinology strategy was first successfully used in the search for vaccine candidates against serogroup B *Neisseria meningitidis* (MenB)[53]. The first step in this process is the completion of the genome sequence of the pathogen of interest. In the case of MenB, this took approximately 18 months from the time when it was undertaken in 1998–1999. Today, the complete genome sequence of a pathogen can be obtained in a matter of days to weeks. Several algorithms are used to identify putative cell-surface or secreted proteins that could potentially elicit antibody responses in a human host. For MenB, 570 potential vaccine candidates were identified by bioinformatics approaches. The next step in the process was to produce recombinant proteins in *Escherichia coli*; approximately 350 proteins were expressed at high levels, purified and used as immunogens in mice. Immune sera were collected and assayed for their ability to bind to the surface of MenB cells and for their bactericidal activity *in vitro*. Seven proteins had high titres in all of the assays that were carried out and were taken into the final stage of evaluation, which assessed the extent of protein sequence variability in these proteins across large numbers of MenB isolates. From this large-scale screening process, two new vaccine candidates emerged that met all of the criteria. These vaccine candidates are now in Phase I clinical trials. Modified with permission from REF. 67 © Macmillan Magazines Ltd (2000).

The availability of pathogen genome sequence data has catalysed research on the identification of new targets for antimicrobial compound, by providing a complete catalogue of genes across a wide range of organisms, which can be compared at various levels[43]. The four attributes of an ideal antimicrobial target are that it is: essential for viability in the microbial pathogen; absent or significantly different in the human host (a parameter that is now much easier to assess given the availability of the human genome sequence); conserved across the appropriate range of organisms; and expressed and relevant to the infectious process. *In silico* approaches that compare pathogenic organisms and their non-pathogenic relatives can provide a first-pass list of potential genes that are required for colonization, invasion and virulence. Other methods such as transposon mutagenesis[44–46] and *in vivo* expression technology (IVET)[47], for example, have allowed large-scale screening for essential proteins *in vitro* and *in vivo*, even when the biological function of a protein is not known. New targets that are identified through such screening programmes can in turn be used in high-throughput screening assays with combinatorial chemistry libraries to identify potential small-molecule inhibitors of protein function. This kind of approach makes full use of the information that is in genome databases by allowing the assay system to drive the identification of new targets[48]. It is not limited by *a priori* knowledge of the function of a specific protein. Genomics-based methods have identified several new targets and pathways for antimicrobial development, including aminoacyl-tRNA synthetases, polypeptide deformylase, fatty-acid biosynthesis, DNA replication, protein secretion, cell division, peptidoglycan biosynthesis, cell signalling and amino-acid biosynthesis[42,49–51]. Although there are pros and cons to developing new broad-spectrum antibiotics versus antibiotics that are highly specific for a particular infectious agent, it seems that it might be possible to achieve success on both fronts by using genomic information in the drug-discovery process.

Genomics-based approaches have also greatly accelerated the search for new vaccine candidates, particularly for those pathogens that induce antibody formation in humans and/or experimental animals. As described above, transcriptome and proteome analyses of various pathogens have identified subsets of genes that are important in infectivity and, therefore, possible targets for vaccine development. An alternative strategy that is known as reverse vaccinology has used bioinformatics methods to identify potential cell-surface proteins, followed by the large-scale expression and evaluation of these candidate antigens in animal models[52]. So far, this approach has been successfully applied to *N. meningitidis*[53] and *S. pneumoniae*[54], and a new *N. meningitidis* vaccine candidate that was discovered through reverse vaccinology has entered Phase I clinical trials (R., Rappuoli, personal communication; FIG. 1). Also, the use of immune sera together with a cell-surface display of *S. aureus* peptides has identified several potential vaccine candidates that patients infected

Box 2 | **Comparative genome sequencing: the search for polymorphisms**

VARIABLE NUMBER OF TANDEM REPEAT (VNTR) analysis of *Bacillus anthracis* spores that were used in the anthrax letter attacks in October 2001 indicated that this strain was related to one that was originally isolated from a dead cow in Texas, USA, in 1981 and designated as the Ames strain[57,68]. This strain of *B. anthracis* was subsequently sent to the US Army Medical Research Institute (USAMRIID) in Fort Detrick, Maryland, where it was used in the US defensive biological weapons programme and distributed to laboratories around the world. At the time of the anthrax letter attacks, scientists at The Institite for Genomic Research (TIGR) were completing the genome sequence of another isolate of *B. anthracis* Ames that had been obtained from Porton Down in the United Kingdom (Porton strain 1 and 2). Before growth of this bacterial culture, the virulence plasmids had been cured by heat treatment. Because VNTR analysis only examines limited regions within a genome, a project was initiated to evaluate whether large-scale DNA sequence analysis could provide further information on polymorphisms between the different Ames isolates. A DNA sample was obtained from a culture of *B. anthracis* taken from the first patient to die of inhalational anthrax in Florida, USA (Florida isolate). DNA libraries were made and the Florida isolate was sequenced to eightfold coverage. The sequences of the Florida and Porton isolates were compared and putative polymorphisms were identified and validated as described in REF. 57. The suggested relationship of multiple Ames isolates is illustrated in the figure. Known direct transfers of the isolates (and hence generations of growth) are shown as full arrows. Diamond boxes represent the sources of DNA that were used for genome sequencing. Hypothetical lines of descent are shown as dotted lines. Porton1 and Porton2 are different cultures of the Porton Ames isolate. The figure is modified with permission from REF. 57 © American Association for the Advancement of Science (2002).



with this pathogen can immunologically recognize[55]. Although it might still be a bit too early to fully assess the impact of genomics on vaccine development, it has been estimated that the availability of genome sequence data, together with the application of large-scale approaches, has reduced the time that is required to identify new vaccine candidates by several years. An added benefit is that the initial screening can be done in a comprehensive way, evaluating all of the potential antigens in a pathogen's genome, and therefore increases the likelihood that the most promising candidates will emerge. Unfortunately, as yet there has not been as much progress using genomics-based approaches in the development of vaccines against pathogens for which T-cell-mediated immunity is most important. This can be largely attributed to the difficulty of predicting T-cell epitopes from protein sequence data alone.

Microbial forensics. Microbial forensics — the use of molecular variation between closely related strains to trace relationships and study population structure[56] — has also benefited tremendously from the availability of comparative genomic data that can discriminate between two samples at the level of a single base pair of DNA. For example, the comparative genome sequencing of a reference Ames strain of *B. anthracis* and the Ames isolate from the first patient to die of inhalation anthrax in the attacks of 2001 allowed new polymorphic loci that distinguished multiple Ames isolates to be identified[57] (BOX 2). Further sequencing of other *B. anthracis* strains, as well as representatives of the closely related *Bacillus cereus* and *Bacillus thuringiensis* is underway, with the goal being the development of a polymorphism database for the *B. anthracis* group that will provide important information for tracking the origin and history of particular isolates (C.M.F., unpublished data). The technology is available to quickly expand these efforts to include other category A–C agents in addition to *B. anthracis*.

Where do we go from here?

Genomics-based approaches to the study of microbial pathogens and their hosts have had a profound impact on the way in which we approach the study and treatment of infectious disease. One of the most important challenges that is facing us today is how to best exploit

VARIABLE NUMBER OF TANDEM REPEATS (VNTRs). A linear arrangement of multiple copies of short repeated DNA sequences that vary in length and are highly polymorphic, which makes them useful as markers.

these large-scale technologies in the biodefence arena. There is no reason we cannot begin to use DNA sequence analysis and DNA microarray technology to collect information on natural variability in large numbers of isolates of the most important pathogens. *B. anthracis* has been the focus of the first large-scale efforts to catalogue such variability and create a forensics database. However, these efforts should be expanded to include all of the principal human, animal and plant pathogens if we are to effectively track naturally occurring and emerging infectious diseases.

It is likely that transcriptome and proteome analyses that are applied to the study of bioterror pathogens will quickly find subsets of genes that have a role in pathogenesis and host–pathogen interactions. This information could potentially provide leads for the development of new antibiotics and vaccines; however, as has already been shown, this information might not always be required, nor does it guarantee success. One of the biggest challenges before us is the fact that only a limited number of animal model systems are available for the study of the most threatening bioterror agents. Moreover, any clinical trials to evaluate the efficacy of new antibiotics and vaccines against this group of infectious agents will be limited by the the inability to do this testing in human subjects — both because it would be considered highly unethical to deliberately expose volunteers to these infectious agents and because natural outbreaks of these diseases are rare or non-existent. So, although it might be tempting to focus considerable research efforts on a small number of bioterror agents, there is much to be gained by taking a broader comparative approach to the study of a large number of pathogens because common mechanisms might emerge.

Anthony Fauci, Director of the NIAID at the NIH recently stated that "The goal of developing 'universal' antibiotics, antivirals and antitoxins … is not unattainable." (REF. 58). It is still too early to tell whether or not this vision is realistic, however, it has set the tone for the infectious disease research agenda for the coming years. Although the consensus is that a large infusion of US federal government funds into biodefence research is essential if these ambitious goals are to be met, increased funding alone is no guarantee that new diagnostics, antibiotics and vaccines will emerge. However, there are several things that can be done to increase the likelihood of success.

Funding of research. There must be a sustained and sufficient commitment of funds to enable the basic research that is beginning today to be translated into deliverable achievements that will have an impact on public health. Increased funding for two to three years is not enough, given the nature of scientific research and development.

Incentives must be created to attract a sufficient number of outstanding new investigators to infectious disease research to carry out an expanded research agenda. Given the nature of the scientific enterprise, the creation of long-term funding programmes that allow investigators to focus on research, and not on grant writing, should be one option that is considered. It should be recognized that there is a need for both individual investigator-initiated research programmes on potential bioterror pathogens as well as larger, more comprehensive, multi-investigator research programmes, such as the recently created RCEs. Although research programmes that focus on developing new diagnostics, antibiotics and vaccines using well-proven methods are essential, further funding initiatives that reward innovative thinking should be encouraged.

Incentives, such as Project BioShield (for more information, see online links box), that will enable the US government to purchase vaccines and drugs in large amounts, must be created to provide incentives for the pharmaceutical and biotechnology industry to rapidly translate basic research results into new products. It is very disturbing that, at a time when microbial genomics has delivered hundreds of potential new targets for antimicrobial compounds, the pharmaceutical industry is scaling back its collective efforts in antibiotic development because the profit margin is not sufficient[43]. In some cases, the key achievements of these research programmes might be drugs or vaccines that are stockpiled for use only in the event of an emergency. Without a continued and sustained market for new products, it is all but guaranteed that these products will not be developed. The US government must extend its efforts in the biodefence arena to become a full partner with industry to increase the level of preparedness for the next natural or deliberate outbreak of disease.

Research priorities. If we have learned anything from the collective efforts in genomics over the past several years it is how little we know about the biology of organisms. Some individuals might be inclined to debate whether we are most in need of 'universal' antibiotics and antivirals versus ones that are more organism-specific, but there are potential uses for both and both should be vigorously pursued if the science that underlies the research is sound.

Although it is essential that a great deal of emphasis be placed on the study of the most important bioterror agents, such as *B. anthracis*, *Y. pestis* and *Variola major*, it is also important to also take a broader view of the microbial world, given that so many features that are relevant to microbial pathogenesis are shared among phylogenetically diverse species. Related to this point is the need to remember that collectively we have focused a great deal of attention on human pathogens, but the social and economic devastation that could result from an attack with a virulent plant or animal pathogen, such as foot and mouth disease, could be much greater than what was seen with the anthrax attacks.

Information sharing. Another topic that is relevant to biodefence research is the need to strike an appropriate balance for the sharing of information and research results on bioterror agents. On the one hand, we have already seen that providing information for downstream work is crucial, particularly for the development of new antibiotics and vaccines. On the other hand,

there is a concern that those individuals with malicious intent could use this information inappropriately[59]. The risks that are posed by the potential misuse of this information have been difficult to quantify, thereby complicating many of these discussions. In January 2003, several scientific journals issued a statement on scientific publication and security to the effect that they will now implement a policy to screen and possibly reject manuscripts that have been submitted for publication if "…an editor … concludes that the potential harm of publication outweighs the societal benefits." (REF. 60; see also REF. 61 and references therein). This is certainly an important statement of responsibility, but it in no way abrogates scientific investigators of their responsibility in this area.

On 8 October 2003, the National Research Council of the National Academies of Science released a report entitled 'Biotechnology Research in an Age of Terrorism' that contained a new set of recommendations for dealing with such dual-use research[62]. The report noted that there was a need for: the increased education of the scientific community about the dual-use dilemma; an Institutional Biosafety Committee review of plans for experiments that fall into potential areas of concern; a review of publications for potential national security risks at the publication stage; the creation of a National Science Advisory Board for Biodefense as a resource for ongoing discussions between the scientific and security communities; and the creation of an International Forum on Biosecurity to promote discussions about all of the relevant issues around the world. This last point is perhaps one of the most important to emerge from this report. Technologies that could potentially be misused are available in many countries around the world, and unless there is international consensus about the appropriate way to deal with dual-use data, the implementation of new policies within the United States alone could potentially frustrate good scientists without increasing the overall level of national security. This objective might not be so easy to achieve, because there are many who believe that the perceived threat that is posed by this kind of information is not universally agreed on by scientists around the globe.

Conclusions

It is clear that genome-enabled science will have a central role in the new biodefence research agenda. On the basis of successes so far, it is highly likely that these approaches will lead to important new breakthroughs on several fronts. As the new biodefence programmes get underway, it will be vitally important to monitor progress and set new priorities as dictated by the most promising research results.

Ultimately, if the United States is committed to developing a comprehensive programme in biodefence, the most important thing that can be done is to ensure that sufficient funding is provided for a sufficient period of time for the ambitious goals of this new initiative to be realized. If successful, there is likely to be a dual payoff. Not only will a new generation of diagnostics, antimicrobial and antiviral compounds be delivered to protect citizens around the world, but the increased level of biopreparedness that will result could probably also be a strong deterrent to any future deliberate bioterror attack.

1. Bloom, B. R. Genome sequences. A microbial minimalist. *Nature* **378**, 236 (1995).
2. Inglesby, T. V. *et al.* Anthrax as a biological weapon: medical and public health management. Working Group on Civilian Biodefense. *JAMA* **281**, 1735–1745 (1999).
3. Dennis, D. T. *et al.* Tularemia as a biological weapon: medical and public health management. *JAMA* **285**, 2763–2773 (2001).
4. Inglesby, T. V. *et al.* Plague as a biological weapon: medical and public health management. Working Group on Civilian Biodefense. *JAMA* **283**, 2281–2290 (2000).
5. Sullivan, P. Cardiac cases trigger advisory about smallpox vaccination. *CMAJ* **168**, 1167 (2003).
6. Bartlett, J. *et al.* Smallpox vaccination in 2003: key information for clinicians. *Clin. Infect. Dis*. **36**, 883–902 (2003).
7. Hayes, S. C. & World, M. J. Adverse reactions to anthrax immunization in a military field hospital. *J. R. Med. Corps* **146**, 191–195 (2000).
8. Hotopf, M. *et al.* Role of vaccinations as risk factors for ill health in veterans of the Gulf War: cross sectional study. *Br. Med. J.* **320**, 1363–1367 (2000).
9. Swanson-Biearman, B. & Krenzelok, E. P. Delayed life-threatening reaction to anthrax vaccine. *J. Toxicol. Clin. Toxicol*. **39**, 81–84 (2001).
10. Asa, P. B., Wilson, R. B., & Garry, R. F. Antibodies to squalene in recipients of anthrax vaccine. *Exp. Mol. Pathol*. **73**, 19–27 (2002).
11. Boucher, Y., Nesbo, C. L, & Doolittle, W. F. Microbial genomes: dealing with diversity. *Curr. Opin. Microbiol*. **4**, 285–289.
Comparison of the genome sequences of several pairs of closely related pathogens has revealed tremendous diversity, with some strains of the same species differing by more than 20% in their gene content. This review summarizes these findings and suggests new conceptual and methodological approaches for further analysis of microbial genome diversity.
12. Shannon, J. G., Ross, C. L., Koehler, T. M. & Rest, R. F. Characterization of anthrolysin O, the *Bacillus anthracis* cholersterol-dependent cytolysin. *Infect. Immun.* **71**, 3183–3189 (2003).
13. Brousseau, R. & Masson, L. *Bacillus thuringiensis* insecticidal crystal toxins: gene structure and mode of action. *Biotechnol. Adv*. **6**, 697–724 (1988).
14. Paulsen, I. T. *et al.* The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts. *Proc. Natl Acad. Sci. USA* **99**,13148–13153 (2002).
This paper reports on the sequence of the bioterror pathogen *Brucella suis*, with a discussion of the similarities between this organism and other animal and plant pathogens.
15. Del Vecchio, V. G. *et al.* The genome sequence of the facultative intracellular pathogen *Brucella melitensis*. *Proc. Natl Acad. Sci. USA* **99**, 443–448 (2002).
16. Kaneko, T. *et al.* Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res*. **7**, 331–338 (2000).
17. Diehn, M. & Relman, D. A. Comparing functional genomic datasets: lessons from DNA microarray analyses of host–pathogen interactions. *Curr. Opin. Microbiol*. **4**, 95–101 (2001).
18. Grifanti *et al.* Previously unrecognized vaccine candidates against group B meningococcus identified by DNA microarrays. *Nature Biotechnol*. **20**, 914–921 (2002).
19. Dietrich, G. *et al.* Transcriptome analysis of *Neisseria meningitidis* during infection. *J. Bacteriol*. **185**, 155–164 (2003).
Analysis of the global changes in gene expression that are induced in *Neisseria meningitidis* after two key steps in meningococcal infection. These data provide new insights into pathogenicity mechanisms of *N. meningitidis* and reveal differences in gene expression that could possibly be exploited in the development of new therapeutics to treat this infection.
20. He, X. Y., Zhuang, Y. H., Zhang, X. G. & Li, G. L. Comparative proteome analysis of culture supernatant proteins of *Mycobacterium tuberculosis* H37Rv and H37Ra. *Microbes Infect*. **5**, 851–856 (2003).
21. Lei, B., Mackie, S., Lukomski, S. & Musser, J. M. Identification and immunogenicity of Group A *Streptococcus* culture supernatant proteins. *Infect. Immun.* **68**, 6807–6818 (2000).
22. Lai, E.-M. *et al.* Proteomic analysis of the spore coats of *Bacillus subtilis* and *Bacillus anthracis*. *J. Bacteriol.* **185**, 1443–1454 (2003).
23. Sonenshein, A. L. in *Prokaryotic Development* (eds Brun, Y. V. & Shimkets, L. J.) 133–150 (ASM Press, Washington DC, 2000).
24. Dixon, T. C., Meselson, M., Guillemin, J., & Hanna, P. C. Anthrax. *N. Engl. J. Med*. **341**, 815–826 (1999).
25. Cox, J. M. *et al.* cDNA array analysis of *cag* pathogenicity island-associated *Helicobacter pylori* epithelial cell response genes. *Infect. Immun.* **69**, 6970–6980 (2001).
26. Cover T. L. & Blaser, M. J. *Helicobacter pylori* infections, a paradigm for chronic mucosal inflammation: pathogenesis and implications for eradication and therapy. *Adv. Intern. Med*. **41**, 85–117 (1996).
27. Webb, P. M., Crabtree, J. E. & Forman, D. Gastric cancer, cytotoxin-associated gene A positive *Helicobacter pylori*, and serum pepsinogens: an international study. The Eurogst Study Group. *Gastroenterology* **116**, 269–276 (1999).
28. Crabtree, J. E. *et al.* Mucosal IgA recognition of *Helicobacter pylori* 120 kDa protein, peptic ulceration, and gastric pathology. *Lancet* **338**, 332–225 (1991).
29. Parsonnet, J., Friedman, G. D., Orentreich, N. & Vogelman, H. Risk for gastric cancer in people with CagA positive or CagA negative *Helicobacter pylori* infection. *Gut* **40**, 297–301 (1997).
30. Ng, L. C., Forslund, O., Koh, S., Kuoppa, K. & Sjostedt, A. The response of murine macrophages to infection with *Yersinia pestis* as revealed by DNA microarray analysis. *Adv. Exp. Med. Biol.* **529**, 155–160 (2003).

31. Boldrick, J. C. *et al.* Stereotyped and specific gene expression programs in human innate immune responses to bacteria. *Proc. Natl Acad. Sci. USA* **99**, 972–977 (2002).
This work uses DNA microarray technology to investigate the gene expression programme that is triggered in human PBMCs after exposure to bacteria and bacterial products. It revealed a remarkable conserved set of changes in gene expression under different conditions and that cell communication within the immune system was an important component of this response.
32. Perrin, A. *et al.* Comparative genomics identified the genetic islands that distinguish *Neisseria meningitidis*, the agent of cerebrospinal meningitis, from other *Neisseria* species. *Infect. Immun.* **70**, 7063–7072 (2002).
33. Tettelin, H. *et al.* Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V *Streptococcus agalactiae*. *Proc. Natl Acad. Sci. USA* **99**, 12391–12396 (2002)
34. Read T. D. *et al.* The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. *Nature* **423**, 81–86 (2003).
35. Perna N. *et al.* Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409**, 529–533 (2001).
36. Wilson, W. J. *et al.* Sequence-specific identification of 18 pathogenic microorganisms using microarray technology. *Mol. Cell Probes* **16**, 119–127 (2002).
37. Laassri, M. *et al.* Detection and discrimination of orthopoxviruses using microarrays of immobilized oligonucleotides. *J. Virol. Methods* **112**, 67–78 (2003).
38. Jones, R. N. *et al.* Epidemiological trends in nosocomial and community-acquired infections due to antibiotic-resistant Gram-positive bacteria. *Diagn. Microbiol. Infect. Dis*. **33**, 101–112 (1999).
39. Archer, G. L. *Staphylococcus aureus*: a well armed pathogen. *Clin. Infect. Dis.* **26**, 1179–1181 (1998).
40. Pfaller M. A. *et al.* Bacterial pathogens isolated from patients with blood stream infection: frequency of occurrence and antimicrobial susceptibility patterns. *Antimicrob. Agents Chemother.* **42**, 1762–1770 (1998).
41. World Health Organization. *Anti-tuberculosis drug resistance in the world: the WHO/IUATLD Global Project on Anti-tuberculosis Drug Resistance Surveillance, 1994–1997.* Report No. WHO/TB/97.229 (WHO Global Tuberculosis Programme Geneva, Switzerland, 1997).
42. McDevitt, D. & Rosenberg, M. Exploiting genomics to discover new antibiotics. *Trends Microbiol*. **9**, 611–617, 2001.
This review summarizes many of the new genome-enabled approaches that are being used in the discovery and validation of new antibiotic targets.
43. Hughes, D. Exploiting genomics, genetics and chemistry to combat antibiotic resistance. *Nature Rev. Genet.* **4**, 432–441 (2003).
44. Hutchison, C. A. *et al.* Global transposon mutagenesis and a minimal *Mycoplasma* genome. *Science* **286**, 2165–2169 (1999).
45. Lamichhane, G. *et al.* A postgenomic method for predicting essential genes at subsaturation levels of mutagenesis: application to *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. USA* **100**, 7213–7218 (2003).
46. Akerley, B. J. *et al.* A genome-scale analysis for identification of genes required for growth or survival of *Haemophilus influenzae*. *Proc. Natl Acad. Sci. USA* **99**, 966–971 (2002).
47. Mahan, M. J. *et al.* Antibiotic-based selection for bacterial genes that are specifically induced during infection of a host. *Proc. Natl Acad. Sci. USA* **92**, 669–673 (1995).
48. Schmid, M. B. Novel approaches to the discovery of antimicrobial agents. *Curr. Opin. Chem. Biol*. **2**, 529–234 (1998).
49. Alksne, L. *et al.* Identification and analysis of bacterial secretion inhibitors utilizing a SecA–LacZ reporter fusion system. *Antimicrob. Agents Chemother*. **44**, 1418–1427 (2000).
50. Barrett, J. F. *et al.* Antimicrobial agents that inhibit two component signal transduction systems. *Proc. Natl Acad. Sci. USA* **95**, 5317–5322 (1998).
51. Payne D. J. *et al.* The impact of genomics on novel antibacterial targets. *Curr. Opin. Drug Discov. Devel.* **3**, 177–190 (2000).
52. Rappuoli, R. Reverse vaccinology, a genome-based approach to vaccine development. *Vaccine* **19**, 2688–2691 (2001).
This review summarizes many of the new genome-enabled approaches that are being used in the discovery of new vaccine candidates.
53. Pizza *et al.* Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. *Science* **287**, 1816–1820 (2000).
54. Wizemann, T. M., *et al.* Use of a whole genome approach to identify vaccine molecules affording protection against *Streptococcus pneumoniae* infection. *Infect. Immun*. **69**, 1593–1598 (2001).
55. Etz, H. *et al.* Identification of *in vivo* expressed vaccine candidate antigens from *Staphylococcus aureus*. *Proc. Natl Acad. Sci. USA* **99**, 6573–6578 (2002).
56. Cummings, C. A. & Relman, D. A. Microbial forensics — cross-examining pathogens. *Science* **296**, 1976–1978 (2002).
57. Read, T. D. *et al.* Comparative genome sequencing for the discovery of novel polymorphisms in *Bacillus anthracis*. *Science* **296**, 2028–2033 (2002).
58. Fauci, A. S. Biodefence on the research agenda. *Nature* **421**, 787 (2003).
This commentary summarizes the new biodefence research agenda of the NIAID at the NIH.
59. Epstein, G. L. Controlling biological warfare threats: resolving potential tensions among the research community, industry, and the national security community. *Crit. Rev. Microbiol.* **27**, 321–354 (2001).
60. Statement on the consideration of biodefence and biosecurity. *Nature* **421**, 771 (2003).
61. Atlas, R. M. Bioterrorism and biodefence research: changing the focus of microbiology. *Nature Rev. Microbiol.* **1**, 70–74 (2003).
62. Committee on Research Standards and Practices to Prevent the Destructive Application of Biotechnology, Development, Security, and Cooperation, National Research Council of the National Academies of Science. *Biotechnology Research in an Age of Terrorism: Confronting the "Dual Use" Dilemma*, [online] < http://www.nap.edu/books/0309089778/html/> (National Academies Press, Washington DC, 2003).
This report summarizes the present rules and regulations that govern research on pathogens and potentially dangerous biotechnology research. It recommends changes in these practices that could further prevent misuse of dual-use research. This is a seminal report that is likely to have a profound impact on how dual-use research is conducted in the years to come.
63. Amherst, J. Letter, British Manuscript Project, Library of Congress, Washington DC (Amherst to Bouquet, 16 July 1763).
64. Meselson, M. *et al.* The Sverdlosk anthrax outbreak of 1979. *Science* **266**, 1202–1208 (1994).
65. Smithson, A. E. in *Ataxia: The Chemical and Biological Terrorism Threat and the US Response.* (eds Smithson, A. E. & Levy, L. E.) Stimson Center Report No. 35, 71–111 (The Henry L. Stimson Center, Washington DC, 2000).
66. Keim P. *et al.* Molecular investigation of the Aum Shinrikyo anthrax release in Kameido, Japan. *J. Clin. Microbiol*. **39**, 4566–4567, 2001.
67. Fraser C. *et al.* Microbial genome sequencing. *Nature* **406**, 799–803 (2000).
68. Keim, P. *et al.* Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. *J. Bacteriol*. **182**, 2928–2936 (2000).

Acknowledgments
The author gratefully acknowledges the many contributions of her TIGR colleagues to the field of microbial genomics.

Competing interests statement
The author declares that she has no competing financial interests.

Online links

FURTHER INFORMATION
NIAID Biodefense Research: www.niaid.nih.gov/biodefense
The CDC Category Listing of Potential Bioterrorism Agents: http://www.scav.org/Category%20Agent%20List.pdf
World Health Organization Information on the SARS outbreak: http://www.who.int/csr/sars/country/2003_08_15/en/
Description of NIAID-funded Regional Centers of Excellence: http://www.niaid.nih.gov/newsroom/releases/HHS_RCE.htm
TIGR Microbial Database: http://www.tigr.org/tdb/mdb/mdbcomplete.html
The Project Bioshield Initiative: www.whitehouse.gov/news/releases/2003/01/20030128-19.html
Access to this interactive links box is free online.