Reverse vaccinology

Rino Rappuoli

Biochemical, serological and microbiological methods have been used to dissect pathogens and identify the components useful for vaccine development. Although successful in many cases, this approach is time-consuming and fails when the pathogens cannot be cultivated *in vitro*, or when the most abundant antigens are variable in sequence. Now genomic approaches allow prediction of all antigens, independent of their abundance and immunogenicity during infection, without the need to grow the pathogen *in vitro*. This allows vaccine development using non-conventional antigens and exploiting non-conventional arms of the immune system. Many vaccines impossible to develop so far will become a reality. Since the process of vaccine discovery starts in silico using the genetic information rather than the pathogen itself, this novel process can be named reverse vaccinology.

Addresses
IRIS, Chiron S.p.A., Via Fiorentina 1, 53100 Siena, Italy;
e-mail: Rino_Rappuoli@biocine.it

Current Opinion in Microbiology 2000, **3**:445–450

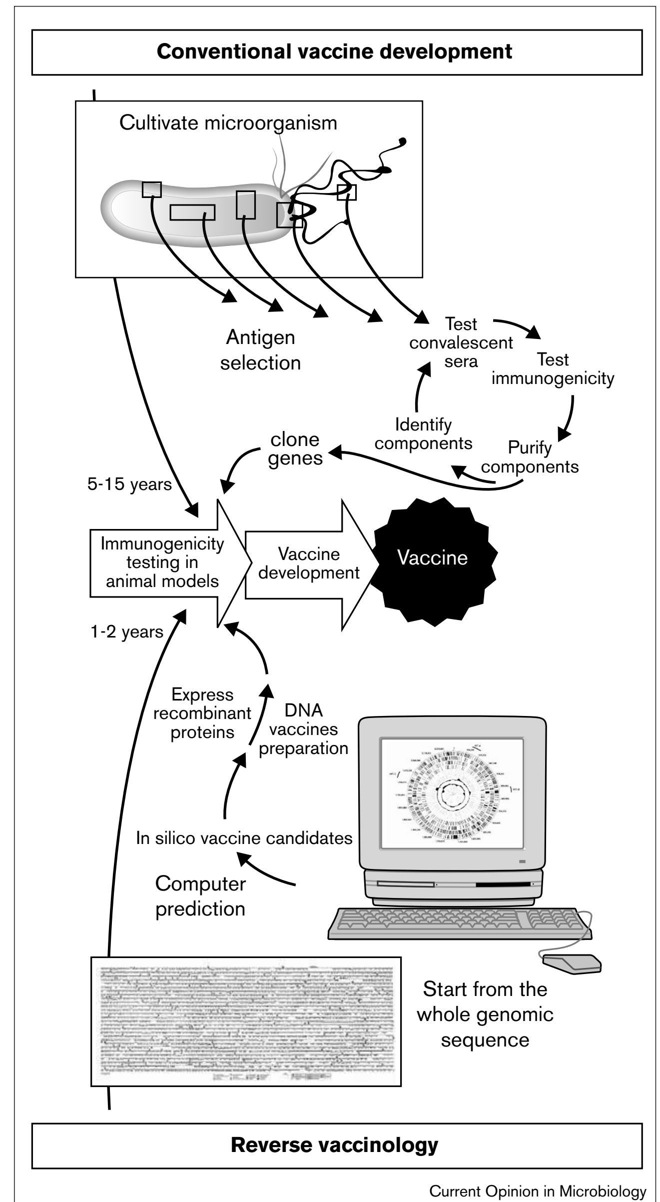
1369-5274/00/$ – see front matter
© 2000 Elsevier Science Ltd. All rights reserved.

Abbreviations
HCV hepatitis C virus
Men B group B meningococcus

Introduction: conventional vaccinology

The conventional approach to vaccine development uses two methods: first, attenuation of pathogens by serial passages *in vitro* to obtain live-attenuated strains to be used as vaccines, and second, identification of protective antigens to be used in non-living, subunit vaccines [1]. In this review, we focus on subunit vaccines. The conventional way to develop these vaccines is summarized in Figure 1. In order to identify the components of the pathogen suitable for vaccine development, the pathogen is grown in laboratory conditions and the components building the pathogen are first identified one at a time, by biochemical, serological or genetic methods. The identification of protective antigens that could be potential vaccine candidates involves separating each component of the pathogen one by one. This approach is time-consuming and allows the identification only of those antigens that can be purified in quantities suitable for vaccine testing. Since the most abundant proteins are most often not suitable vaccine candidates, and the genetic tools required to identify the less abundant components may be inadequate or not available at all, this approach can take years or decades. For the bacterial and parasitic pathogens studied to date, the maximum number of potential vaccine antigens identified during a century of vaccine development is usually less than ten. This conventional method also means that vaccine development is not possible when the pathogen cannot be grown in laboratory conditions. An exception to this has been the hepatitis B vaccine where the pathogen, although unable to grow *in vitro*, could be recovered in large quantities from the plasma of infected people [2].

Figure 1



Schematic representation of the essential steps of vaccine development by the conventional approach and by reverse vaccinology.

Once a suitable antigen is identified, it needs to be produced in large scale, often by growing the pathogen itself. Cloning of the gene coding for the antigen is often

necessary in order to better characterize and produce the identified antigen(s). Finally, the new molecule can enter vaccine development. Although successful in many cases, this approach took a long time to provide vaccines against those pathogens for which the solution was easy and failed to provide a solution for those bacteria and parasites that did not have obvious immunodominant protective antigens [3•].

Reverse vaccinology

The reverse approach to vaccine development takes advantage of the genome sequence of the pathogen. The genome sequence provides at once a catalog of virtually all protein antigens that the pathogen can express at any time. As shown in Figure 1, this approach starts from the genomic sequence and, by computer analysis, predicts those antigens that are most likely to be vaccine candidates. The approach can, therefore, be very naïve, and poses the question of whether any of the potential antigen candidates can provide protective immunity without knowing whether the antigen is abundant, immunogenic during infection or expressed *in vitro*. This approach allows not only the identification of all the antigens seen by the conventional methods, but also the discovery of novel antigens that work on a totally different paradigm. Therefore, this method allows the discovery of novel mechanisms of immune intervention. The feasibility of the approach relies heavily on the availability of a high-throughput system to screen protective immunity. When this is available, in theory all genes of a pathogen can be tested, without any bias of any type. Unfortunately, owing to our limited knowledge of vaccine immunology, good correlates of protection are rare and, therefore, screening for protective immunity is the rate-limiting step of reverse vaccinology. The other limit of this approach is the inability to identify non-protein antigens such as polysaccharides, which are important components of many successful vaccines, and the identification of CD1-restricted antigens such as glycolipids, which represent new promising vaccine candidates.

Applications of reverse vaccinology

The publication of the complete genome sequence of many bacteria, parasites and viruses means that the reverse approach to vaccine development can be put into practice. Below we discuss the different approaches that are being used or potentially could be used to develop novel and effective vaccines against a variety of pathogens.

Group B meningococcus

Group B meningococcus (MenB) represents the first example of the successful application of reverse vaccinology. The conventional approach to vaccine development against this pathogen had been struggling for four decades without progress. On the one hand, the capsular polysaccharide used to develop conventional and conjugate vaccines against all other pathogenic meningococci could not be used because the MenB capsule, which is chemically identical to an α2–8 linked polysialic acid present in many of our tissues, is poorly immunogenic and a potential cause of autoimmunity. On the other hand, the protein-based approach had identified as protective antigens the most abundant proteins of the outer membrane [4]. However, these abundant surface-exposed proteins usually contain many amphipathic domains, which span the outer membrane several times and assume a β-barrel conformation (Figure 2a). The protective epitopes in these proteins are located in the loops that are exposed on the external surface and are usually formed by the precise conformation of a few amino acids. Therefore, in order to induce protective immunity these antigens need to be folded within the outer membrane (recombinant proteins do not induce protection) and any change in one of the few amino acids of the loop will result in a different epitope. Vaccines based on outer membrane vesicles (OMV) and containing the major outer membrane proteins have been developed and shown to be efficacious in clinical trials; however, owing to the high sequence variability of the external loops in different MenB strains, protection is induced only against the immunizing strain. As a consequence, the conventional approach to vaccine development has failed to deliver a universal vaccine.

Using reverse vaccinology, fragments of DNA were screened by computer analysis while the MenB nucleotide genome sequence was being determined [5•,6••]. Six hundred novel genes were predicted to code for surface-exposed or exported proteins. These were cloned and expressed in *Escherichia coli* as fusions to the glutatione transferase or to a histidine tag. Of these fusion proteins, 350 were successfully expressed, purified and used to immunize mice. The sera obtained were used to confirm the surface exposure of the proteins by ELISA and FACS analysis, and to test for the ability to induce complement-mediated *in vitro* killing of bacteria, a test that correlates with vaccine efficacy in humans. Within 18 months, while the nucleotide sequence was still being finalized, 85 novel surface-exposed proteins were discovered and 25 of these were shown to induce bactericidal antibodies [6••]. These numbers are impressive if one considers that during the past four decades no more than a dozen of such proteins had been identified. The surprising finding was not only the high number of the new proteins found but also the quality of the new proteins. In addition to the conventional outer membrane proteins with variable surface-exposed loops (as in Figure 2a), many of the new proteins were lipoproteins or other types of surface-associated proteins without membrane-spanning domains (Figure 2b). These were often conserved in sequence, and carried multiple protective epitopes conserved in most strains. These novel proteins provide an optimal basis for the development of a novel and effective vaccine against MenB [6••].

Malaria

Malaria, together with AIDS and tuberculosis, belongs to the triad of the most dangerous diseases that threaten

human health. The 500 million new infections each year and 2.5 million annual deaths indicate that all measures used so far to control the disease have failed [7]. Vaccination would be an effective way to control the spread of malaria, but vaccines are not available, despite many years of research [3•]. Approximately 20 antigens have been identified from the malaria parasite but none of them is good enough for a vaccine. The problem is further complicated by the different antigenic profiles expressed by sporozoites, merozoites and gametocytes, that the parasite assumes during its life cycle. The solution can only come from a genomic approach. The sequence of two of the 14 chromosomes of *Plasmodium falciparum* have been published [8•,9•] and provided the full set of genes contained in the two chromosomes. The complete sequence of the whole genome will soon provide information on the predicted 6000 genes. Analysis of the whole genome expression will show which genes are expressed by the sporozoite, liver and sexual life-stages of the parasite. Expression of genes predicted to be immunogenic as recombinant proteins delivered with adjuvants or as DNA vaccines will eventually provide the effective vaccine against malaria [10••,11].

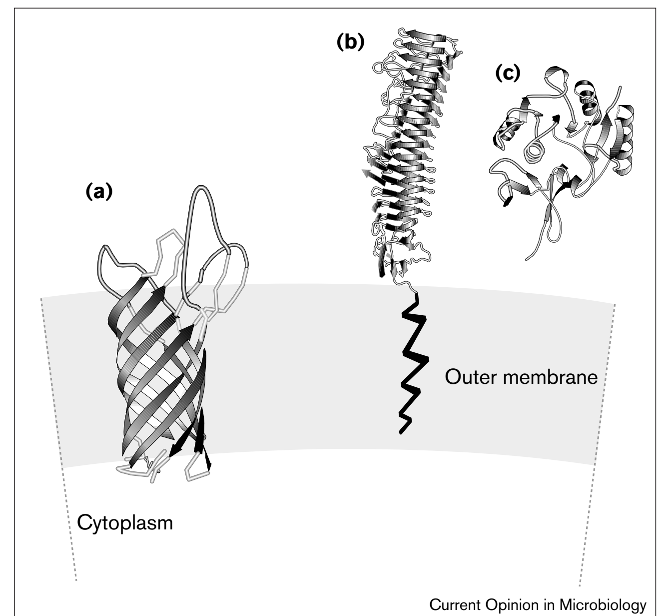
The task is a formidable challenge, however, it is doable. It is just a matter of resources and co-ordination. The most difficult task is the development of an *in vivo* or *in vitro* model that allows high-throughput screening of vaccine candidates.

Tuberculosis

Mycobacterium tuberculosis infects approximately two billion people worldwide and causes 1.5 million deaths annually [12]. The inability of AIDS patients to keep the infection under control and the appearance of multi-resistant strains make the disease an unrestrained danger. The available live-attenuated BCG vaccine is not a solution, because of the variable efficacy reported in the trials. Furthermore, subunit vaccines have not been developed because all the antigens identified by conventional vaccinology provide protection that in animal models is lower than that provided by BCG [12]. Also vaccine development and testing is complicated by the long time required for bacterial growth.

The sequence of the whole genome of *M. tuberculosis* [13•] has provided a list of all possible genes, which now can all be expressed as recombinant proteins or as DNA vaccines and tested for protective immunity [14]. The absence of a high-throughput screening for protective antigens makes the effort difficult but doable by means of a systematic approach. However, a number of genome- and proteome-based approaches are providing novel vaccine candidates, while at the same time the increased knowledge of this difficult bacterium makes it easier to approach [15,16,17••]. The combination of the genome and the use of the fast-growing *Mycobacterium marinum* is the winning combination to accelerate the discovery of an effective tuberculosis vaccine [18].

Figure 2



Examples of protective antigens of *Neisseria meningitidis*. **(a)** A schematic structure of a typical outer membrane protein that is mostly embedded within the membrane. These proteins contain one or two protective epitopes located at the tip of most external loops – change of one amino acid is enough to escape immunity. Outer membrane proteins represent the major part of the antigens of *N. meningitidis* identified by conventional vaccinology. Many novel outer membrane proteins have been identified by reverse vaccinology. **(b,c)** Schematic structures of membrane-anchored lipoproteins or secreted antigens from *N. meningitidis* that have been identified by reverse vaccinology.

Syphilis

During the past four centuries, syphilis has been a nightmare comparable to today's AIDS [19]. If untreated, this sexually-transmitted disease leads to neurological disorders, cardiovascular problems and death, but after the discovery of penicillin the disease became easy to control. However, today syphilis represents a new threat both in developed and developing countries because it causes genital ulcers, which facilitate the spread of HIV. There are approximately 9000 syphilis cases in the US and it is common in Africa and other developing countries, which are estimated to have 12.5 million cases [20].

Syphilis is caused by a bacterium, *Treponema pallidum*, that cannot be cultivated in the laboratory and, therefore, has been refractory to conventional approaches to vaccine development. Attempts to identify vaccine antigens using the bacterium grown in rabbits had identified approximately 20 different antigens. Once again, the sequence of the complete genome made available at once all the genes of the bacterium, which can all be expressed as recombinant proteins or as DNA vaccines [21•,22,23••]. Therefore, for the first time it is now possible to approach development of a syphilis vaccine in a systematic way. The absence of a high-throughput

Table 1

Comparison of conventional and genomic approaches to vaccine development.

	Conventional vaccinology	Reverse vaccinology
Essential features		
	Most abundant antigens during disease	All antigens immunogenic during disease
	Antigens immunogenic during disease	Antigens even if not immunogenic during disease
	Cultivable microorganism	Antigens even in non-cultivable microorganisms
	Animal models essential	Animal models essential
	Correlates of protection useful	Correlates of protection very important
		Correct folding in recombinant expression important
		High-throughput expression/analysis important
Advantages		
	Polysaccharides may be used as antigens	Fast access to virtually every single antigen
	Lipopolysaccharide-based vaccines are possible	Non-cultivable microorganisms can be approached
	Glycolipids and other CD1-restricted antigens can be used	Non abundant antigens can be identified
		Antigens that are not immunogenic during infection can be identified
		Antigens that are transiently expressed during infection can be identified
		Antigens not expressed *in vitro* can be identified
		Non-structural proteins can be used
Disadvantages		
	Long time required for antigen identification	Non proteic antigens cannot be used (polysaccharide, lipopolysaccharides, glycolipids and other CD1-restricted antigens)
	Antigenic variability of many of the identified antigens	
	Antigens not expressed *in vitro* cannot be identified	
	Only structural proteins are considered	

animal model again makes the problem difficult but not impossible to solve.

Hepatitis C virus

Hepatitis C virus (HCV) [24] is perhaps the best example of a vaccine being developed entirely by reverse vaccinology. In this case the virus that causes the disease has never been cultivated *in vitro* (it grows only in humans and chimpanzees [25]) and has never been visualized by electron microscopy, making it impossible to use any conventional approach to vaccine development. The cloning and sequencing of the HCV genome allowed the identification of the etiological agent [26], the recombinant expression of its proteins, and the immediate development of diagnostic tools, which prevents hundreds of new infections each day ever since. The availability of the genome sequence also allowed the prediction of the envelope proteins that normally are used to develop vaccines against enveloped viruses [27]. These proteins (E1 and E2) have been expressed in many hosts, but so far only mammalian cells have been able to express them in a form that induces production of antibodies able to interfere with the binding of E2 to the host receptor [28]. These recombinant proteins have been able to protect chimpanzees from infection with the homologous HCV virus [29].

While vaccine development using the E1 and E2 conventional vaccine targets is making progress, perhaps the most interesting questions are whether we can take advantage of the knowledge of the genome to design totally non-conventional vaccine targets and whether proteins never used in conventional vaccines (i.e. non-structural proteins) can become effective vaccines. These proteins should be able to confer protection mostly through cell-mediated immunity and not rely on antibody neutralization of viral infection. The encouraging results obtained with some early proteins such as Tat and Rev [30••–32••] in the case of HIV suggest that this may be a novel way to protect against viruses.

Other pathogens

The pathogens described above are perhaps some of the most representative among those that can be approached by reverse vaccinology. However, the list of the pathogens where the conventional approaches to vaccine development have failed or provided only partial solutions is extensive. Among these we can list bacteria such as *Chlamydia* [33•,34,35], pneumococcus [36–39], *Streptococcus*, *Staphylococcus*, pseudomonas, *Borrelia* [40,41••], *Escherichia coli*, gonococcus, typhoid, *Brucella*, *Ricksettia* [42•] and *Bartonella* (the genome sequences of most of these pathogens are about to be completed and available on the website http://www.tigr.org), and parasites such as *Leishmania* and many others.

Conclusions

Conventional approaches to vaccine development are time consuming, identify only abundant antigens that may or

may not provide immunity, and fail when the pathogen cannot be cultivated under laboratory conditions. Reverse vaccinology (i.e. genomic-based approaches to vaccine development) can overcome these problems (see Table 1) and allow researchers to identify novel antigen vaccine candidates. The sequencing of the complete genome of many pathogens, such as group B meningococcus, has allowed the successful application of reverse vaccinology where conventional approaches have failed. With the genome sequences of many other bacteria, parasites and viruses to be completed in the near future, reverse vaccinology means that many vaccines that were impossible to develop will become reality, and novel vaccines, using non-conventional antigens (i.e. non-structural proteins) can be developed.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Rappuoli R, Del Giudice G: **Identification of vaccine targets.** In *Vaccines: From Concept to Clinic.* Edited by Paoletti LC, McInnes PM. Boca Raton: CRC Press; 1999:1-17.

2. Buynak EB, Roehm RR, Tytell AA, Bertland AU II, Lampson GP, Hilleman MR: **Vaccine against human hepatitis B.** *JAMA* 1976, **235**:2832-2834.

3. • National Institutes of Health: **Jordan Report 2000. Accelerated Development of Vaccines.** 2000:i-173 [http://www.nih.gov].

A comprehensive report on the state-of-the-art of most vaccines in development.

4. Zollinger WD: **New and improved vaccines against meningococcal disease.** In *New Generation Vaccines.* Edited by Levine MM, Woodrow GC, Kaper JB, Cobon GS. New York: Marcel Dekker Inc; 1997:469-488.

5. • Tettelin H, Saunders NJ, Heidelberg J, Jeffries AC, Nelson KE, Eisen JA, Ketchum KA, Hood DW, Peden JF, Dodson RJ *et al.*: **Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.** *Science* 2000, **287**:1809-1815.

The genome sequence of serogroup B meningococcus is reported.

6. •• Pizza M, Scarlato V, Masignani V, Giuliani MM, Aricò B, Comanducci M, Jennings GT, Baldi L, Bartolini E, Capecchi B *et al.*: **Whole genome sequencing to identify vaccine candidates against serogroup B meningococcus.** *Science* 2000, **287**:1816-1820.

This paper provides the first example of a rational approach to vaccine development using reverse vaccinology. The whole genome sequence of *N. meningitidis* is screened by computer to identify vaccine candidates. The antigens identified in silico are expressed in *E. coli* that are used to immunize mice and the sera tested for bactericidal activity *in vitro*. Many novel vaccine candidates were discovered that had been missed by all other technologies used so far.

7. World Health Organization: **World malaria situation in 1994.** *Wkly Epidemiol Rec* 1997, **72**:269-274.

8. • Gardner MJ, Tettelin H, Carucci DJ, Cummings LM, Aravind L, Koonin EV, Shallom S, Mason T, Yu K, Fujii C *et al.*: **Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum*.** *Science* 1998, **282**:1126-1132.

The complete sequence of chromosome 2 of *Plasmodium falciparum*.

9. • Bowman S, Lawson D, Basham D, Brown D, Chillingworth T, Churcher CM, Craig A, Davies RM, Devlin K, Feltwell T *et al.*: **The complete nucleotide sequence of chromosome 3 of *Plasmodium falciparum*.** *Nature* 1999, **4000**:532-538.

The complete sequence of chromosome 3 of *Plasmodium falciparum*.

10. •• Hoffman SL, Rogers WO, Carucci DJ, Venter JC: **From genomics to vaccines: malaria as a model system.** *Nat Med* 1998, **4**:1351-1353.

The possibility of using reverse vaccinology against malaria is discussed. Malaria, for which a vaccine has been refractory to any other approach, may be an ideal target for reverse vaccinology, which may succeed where everything else failed.

11. Wang R, Doolan DL, Le TP, Hedstrom RC, Coonan KM, Charoenvit Y, Jones TR, Hobart P, Margalith M, Ng J *et al.*: **Induction of antigen-specific cytotoxic T lymphocytes in humans by a malaria DNA vaccine.** *Science* 1998, **282**:476-480.

12. Ridzon R, Hannan M: **Tuberculosis vaccines.** *Science* 1999, **286**:1298-1300.

13. • Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE III *et al.*: **Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence.** *Nature* 1998, **393**:537-544.

The complete sequence of *Mycobacterium tuberculosis*.

14. Doherty TM, Andersen P: **Tuberculosis vaccines: development work and future.** *Curr Opin Pulm Med* 2000, **6**:203-208.

15. Brosch R, Philipp WJ, Stravropoulos E, Colston MJ, Cole ST, Gordon SV: **Genomic analysis reveals variation between *Mycobacterium tuberculosis* H3 and the attenuated *M. tuberculosis* H37Ra strain.** *Infect Immun* 1999, **67**:5768-5774.

16. Jungblut PR, Schaible UE, Mollenkopf HJ, Zimny-Arndt U, Raupach B, Mattow J, Halada P, Lamer S, Hagens K, Kaufmann SH: **Comparative proteome analysis of *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG strains: towards functional genomics of microbial pathogens.** *Mol Microbiol* 1999, **33**:1103-1117.

17. •• Behr MA, Wilson MA, Gill WP, Salamon H, Schoolnik GK, Rane S, Small PM: **Comparative genomics of BCG vaccines by whole-genome DNA microarray.** *Science* 1999, **284**:1520-1523.

Whole genome comparison using microchips is used to identify the differences between the live-attenuated strains of BCG, which are used as vaccines against tuberculosis, and the wild type, disease-causing strains. Several surprises (unexpected differences) emerge from this analysis.

18. Ramakrishnan L, Federspiel NA, Falkow S: **Granuloma-specific expression of mycobacterium virulence proteins from the glycine-rich PE-PGRS family.** *Science* 2000, **288**:1436-1439.

19. Singh AE, Romanowski B: **Syphilis: review with emphasis on clinical, epidemiologic, and some biologic features.** *Clin Microbiol Rev* **12**:187-209.

20. St Louis ME, Wasserheit JN: **Elimination of syphilis in the United States.** *Science* 1998, **281**:353-354.

21. • Fraser CM, Norris SJ, Weinstock GM, White O, Sutton GG, Dodson R, Gwinn M, Hickey EK, Clayton R, Ketchum KA *et al.*: **Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.** *Science* 1998, **281**:375-388.

The complete sequence of *Treponema pallidum*, the causative agent of syphilis.

22. Pennisi E: **Genome reveals wiles and weak points of syphilis.** *Science* 1998, **281**:324-325.

23. •• Norris SJ, Weinstock GM: **The genome sequence of *Treponema pallidum*, the syphilis spirochete: will clinicians benefit?** *Curr Opin Infect Dis* 2000, **13**:29-36.

The authors discuss how the knowledge of the genome may help improve the approach to syphilis research, including vaccine development.

24. Sarbah SA, Younossi ZM: **Hepatitis C: an update on the silent epidemic.** *J Clin Gastroenterol* 2000, **30**:125-143.

25. Bradley DW, McCaustland KA, Cook EH, Ebert JW, McCaustland KA, Schable CA, Fields HA: **Posttrasfusion non-A non-B hepatitis in chimpanzees: physicochemical evidence that the tubule-forming agent is a small enveloped virus.** *Gastroenterol* 1985, **88**:773-779.

26. Choo Q-L, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M: **Isolation of a cDNA clone derived from a blood-borne non-A and non-B viral hepatitis genome.** *Science* 1989, **244**:359-362.

27. Reed KE, Rice CM: **Overview of hepatitis C virus genome structure, polyprotein processing, and protein properties.** *Curr Top Microbiol Immunol* 2000, **242**:55-84.

28. Rosa D, Campagnoli S, Moretto C, Guenzi E, Cousens L, Chin M, Dong C, Weiner AJ, Lau JY, Choo QL *et al.*: **A quantitative test to estimate neutralizing antibodies to the hepatitis C virus: cytofluorimetric assessment of envelope glycoprotein 2 binding to target cells.** *Proc Natl Acad Sci USA* 1996, **93**:1759-1763.

29. Choo QL, Kuo G, Ralston R, Weiner A, Chien D, Van Nest G, Han J, Berger K, Thudium K, Kuo C *et al.*: **Vaccination of chimpanzees against infection by the hepatitis C virus.** *Proc Natl Acad Sci USA* 1994, **91**:1294-1298.

30. Pauza CD, Trivedi P, Wallace M, Ruckwardt TJ, Le Baunce H, Lu W,
•• Bizzini B, Burny A, Zagury D, Gallo RC: **Vaccination with tat toxoid attenuates disease in simian/HIV-challenged macaques.** *Proc Natl Acad Sci USA* 2000, **97**:3515-3519.

The possibility of producing large quantities of recombinant non-structural proteins (which would be impossible by growing the virus using conventional vaccinology) allows researchers to test novel paradigms for vaccination. Here, along with [31••,32••], early viral proteins are used to vaccinate against HIV, trying to exploit cellular-mediated immunity against these antigens to protect from infection. This paradigm is novel in vaccinology.

31. Cafaro A, Caputo A, Fracasso C, Maggiorella MT, Goletti D,
•• Baroncelli S, Pace M, Sernicola L, Koanga-Mogtomo ML *et al.*: **Control of SHIV-89.6P-infection of cynomolgus monkeys by HIV-1 Tat pro vaccine.** *Nat Med* 1999, **5**:643-650.

See annotation for [30••].

32. Osterhaus AD, van Baalen CA, Gruters RA, Schutten M,
•• Siebelink CH, Hulskotte EG, Tijhaar EJ, Randall RE, van Amerongen G, Fleuchaus A *et al.*: **Vaccination with Rev and Tat against AIDS.** *Vaccine* 1999, **17**:2713-2714.

See annotation for [30••].

33. Stephens RS, Kalman S, Lammel C, Fan J, Marathe R, Aravind L,
• Mitchell W, Olinger L, Tatusov RL, Zhao Q *et al.*: **Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*.** 1998, **282**:754-759.

The complete sequence of *Chlamydia trachomatis*.

34. Read TD, Brunham RC, Shen C, Gill SR, Heidelberg JF, White O, Hickey EK, Peterson J, Utterback T, Berry K *et al.*: **Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39.** *Nucleic Acids Res* 2000, **28**:1397-1406.

35. Kalman S, Mitchell W, Marathe R, Lammel C, Fan J, Hyman RW, Olinger L, Grimwood J, Davis RW, Stephens RS: **Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*.** *Nat Genet* 1999, **21**:385-399.

36. Klein DL, Ellis RW: **Conjugate vaccines against *Streptococcus pneumoniae*.** In *New Generation Vaccines*. Edited by Levine MM, Woodrow GC, Kaper JB, Cobon GS. New York: Marcel Dekker Inc; 1997:504-525.

37. Toumanen E: **Molecular and cellular biology of pneumococcal infection.** *Curr Opin Microbiol* 1999, **2**:35-39.

38. Baltz RH, Norris FH, Matsushima P, DeHoff BS, Rockey P, Porter G, Burgett S, Peery R, Hoskins J, Braverman L *et al.*: **DNA sequence sampling of the *Streptococcus pneumoniae* genome to identify novel targets for antibiotic development.** *Microb Drug Resist* 1998, **4**:1-9.

39. Polissi A, Pontiggia A, Feger G, Altieri M, Mottl H, Ferrari L, Simon D: **Large-scale identification of virulence genes from *Streptococcus pneumoniae*.** *Infect Immun* 1998, **66**:5620-5629.

40. Fraser CM, Casjens S, Huang WM, Sutton GG, Clayton R, Lathigra R, White O, Ketchum KA, Dodson R, Hickey EK *et al.*: **Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*.** *Nature* 1997, **390**:580-586.

41. Nordstrand A, Barbour AG, Bergström S: ***Borrelia* pathogenesis
•• research in the post-genomic and post-vaccine era.** *Curr Opin Microbiol* 2000, **3**:86-92.

Although a vaccine based on OspA had been developed against Lyme disease before the genome sequence of the bacterium had been determined, the efficacy of this vaccine is limited to the East Coast of the USA. The authors discuss how sequencing the genome changed the approach to the problem and how it may help improve the vaccine.

42. Andersson SG, Zomorodipour A, Andersson JO,
• Sicheritz-Ponten T, Alsmark UC, Podowski RM, Naslund AK, Eriksson AS, Winkler HH, Kurland CG: **The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria.** *Nature* 1998, **396**:133-140.

The complete sequence of the *Rickettsia prowazekii*.