Chapter 2
Bacterial Genomes and Vaccine Design

Valeria Cafardi, John L. Telford, and Davide Serruto

Abstract Since its introduction, vaccinology has been very effective in controlling and eliminating life-threatening infectious diseases. However, in several cases, the conventional approach to identifying protective antigens, based on biochemical, immunological, and microbiological methods, has failed to deliver successful vaccine candidates against major human pathogens. The availability of complete bacterial genome sequences has allowed scientists to change the paradigm and approach vaccine development starting from genomic information, a process named reverse vaccinology. This can be considered as one of the most powerful examples of how genomic information can be used to develop vaccines that were difficult or impossible to tackle with conventional approaches. The ever-growing genomic data, the new genome-based approaches and high-throughput sequencing technologies will help to complement reverse vaccinology to enable timely development of new vaccine antigens against emerging infectious diseases.

2.1 Introduction

Vaccines are currently available for infectious diseases caused by various viruses and bacteria and the prevention of disease and death by vaccination has profoundly improved the public health of many populations globally. Louis Pasteur, who developed the first vaccine against rabies, established in 1881 the basic paradigm for vaccine development, which included the isolation, inactivation, and injection of the causative microorganism. These basic principles have guided vaccine development during the twentieth century. All existing vaccines are based on killed or live-attenuated microorganisms or subunits purified from the microorganism such as toxins detoxified by chemical treatment, purified antigens or polysaccharide

J.L. Telford (✉)
Microbial Molecular Biology, Novartis Vaccines and Diagnostics, Siena, Italy
e-mail: john.telford@novartis.com

D.R. Flower and Y. Perrie (eds.), *Immunomic Discovery of Adjuvants and Candidate Subunit Vaccines*, Immunomics Reviews: 5, DOI 10.1007/978-1-4614-5070-2_2,

conjugated to proteins. At the end of the twentieth century, most of the vaccines that could be developed by these traditional technologies had been developed and they allowed the control and, in some cases, the eradication of many important infectious diseases [78]. Although very successful, in several instances these approaches were not able to deliver vaccines against certain pathogens and on other occasions the vaccines obtained with these classical approaches were no longer adequate due to safety concerns and low efficacy. Killed and attenuated vaccines, based on the whole organisms, may contain several factors that may have reactogenic activity and may induce undesirable inflammatory response. The attenuated vaccines could also revert to the virulent status and chemicals used for inactivating pathogens could be present as traces in the final composition. In addition, classical biochemical and microbiological methods used to identify protective subunits were hampered by the limited number of candidate antigens that could be identified as well as the time required for their identification. Remarkable progresses were recently made by the introduction of new technologies such as recombinant DNA and chemical conjugation of proteins to polysaccharides, as well as advances in the identification of novel adjuvants.

The genome era, initiated with the completion of the first bacterial genome, that of *Haemophilus influenzae* in 1995 [1], catalyzed a new revolution in vaccine development. Advances in sequencing technology and bioinformatics have resulted in an exponential growth of genome sequence information. The study of genomes by both computational and experimental approaches has significantly advanced our understanding of the physiology and pathogenicity of many microbes and has provided insights into the mechanisms of genome evolution as well as microbial population structures [2, 3].

Genomes and genome-based technologies have also the potential to help in the development of therapeutics and vaccines. The availability of whole-genome sequences has entirely changed the approach to vaccine development. The genome represents a list of virtually all the protein antigens that the pathogen can express at any time. It becomes possible to choose potentially surface-exposed proteins in a reverse manner, starting from the genome rather than from the microorganism with an approach called reverse vaccinology [4]. In this review we will describe how genomic information has been successful in the identification of novel protein antigens against various human pathogens. We will also focus on recent reports that have contributed to the discovery of novel vaccine candidates providing the proof of concept of genome-based approaches such as pan-genome investigation, subtractive reverse vaccinology, and DNA microarray analysis. A future view of how high-throughput sequencing methods might positively influence vaccine design will also be discussed.

2.2 Reverse Vaccinology: A Novel Genomic Approach to Antigen Identification

The recent genome revolution has extended the confines in vaccine research. Genome mining has revolutionized the approach to vaccine development and provided a new innovation to antigen selection and design. The approach starting from the genomic information leading to the identification of potential vaccine candidates is termed reverse vaccinology [4]. The availability of complete bacterial genome sequences offers a comprehensive catalogue of genes encoding all the potential proteins of a pathogen, with the potential to rationally select vaccine candidates rather than empirically test them one at a time. Furthermore, the prediction of antigens is independent of the need to culture the pathogen in vitro. On the basis of the concept that surface-exposed proteins are susceptible to antibody recognition and are therefore the most suitable vaccine antigens, a complete genome sequence can be screened using bioinformatics algorithms to select open reading frames (ORFs) encoding putative surface-exposed or secreted proteins. Putative surface proteins can be readily identified based on the combination of several features including the presence of signal peptide sequences, membrane spanning regions, lipoprotein signature, and motifs such as sortase attachment sites (LPTXG sites). Moreover, proteins with homology to known virulence factors or protective antigens from other pathogens can be selected based on homology. Several computational methods are available to search for surface-associated or secreted proteins: PSORT is used for the prediction of protein sorting signals and localization sites in amino acid sequences; SignalP predicts the presence and location of signal peptide cleavage sites in amino acid sequences from different organisms, Gram-positive and Gram-negative prokaryotes; TMpred program makes a prediction of membrane spanning regions and their orientation. Although much progress can be made in silico, the experimental approach is necessary to establish unambiguously the localization of the protein in living bacteria. Furthermore, screening for sequence homologies to human proteins for their exclusion in the selection process can help to avoid problems of autoimmunity.

After candidate surface antigens are identified in silico, they are produced as recombinant proteins and their immunogenicity is assayed to measure their potential as vaccine candidates. The feasibility of the reverse vaccinology approach relies on the availability of a high-throughput system for protective immunity screening and also on good correlate of protection. In the paragraphs below, we will describe how genomic information has been successful in the identification of novel potential vaccine candidates against various human pathogens, such as *Neisseria meningitidis* serogroup B, *Streptococcus agalactiae*, and pathogenic *Escherichia coli*.

2.2.1 The First Vaccine Obtained Through Reverse Vaccinology: The Serogroup B Meningococcus Vaccine

The concept of reverse vaccinology was applied for the first time to serogroup B *N. meningitidis* (MenB). *N. meningitidis* is the major cause of meningitis and sepsis, two devastating diseases that can kill children and young adults within hours, despite the availability of effective antibiotics. It is a Gram-negative bacterium that colonizes asymptomatically the upper nasopharynx tract of about 5–15 % of the human population. However, in a significant number of cases, the bacterium can traverse the epithelium and reach the bloodstream causing septicemia. From the blood meningococcus is able to cross the blood–brain barrier and infect the meninges, causing meningitis [5, 6].

N. meningitidis can be classified in 13 serogroups on the basis of the chemical composition of the capsule polysaccharide. However, more than 95 % of total cases of invasive disease are caused by five major serogroups: A, B, C, Y, and W135. Vaccines against serogroups A, C, Y, and W135 were developed in the 1960s by using the purified capsular polysaccharide as antigen. Second-generation, conjugated vaccines have now been introduced. The chemical composition of the polysaccharide of serogroup B, which resembles a molecule present in human tissues, makes a polysaccharide-based vaccine poorly immunogenic and a possible cause of autoimmunity.

In the last 40 years a lot of efforts have been directed to the identification of meningococcus B antigens as the basis of new vaccines. However, the high variability of these proteins among the different MenB strains represents a serious obstacle to the production of a globally effective anti-MenB vaccine [5]. As a consequence there are no effective vaccines available for the prevention of MenB disease, which is responsible for one third of meningococcal disease in the United States, and up to 80 % of cases in Europe.

In 1998, the research team at Novartis Vaccines embarked on a large-scale genome project. To develop a universal vaccine against serogroup B, the genome of a MenB isolate (MC58 strain) has been sequenced and used to discover novel antigens [7, 8].

The identification of new previously unidentified antigens was a process that took the research team 18 months to achieve. The sequence of the virulent strain was determined by the shotgun strategy and in order to identify novel vaccine antigens a strategy has been aimed to select, among the more than 2,000 predicted proteins, those that were predicted to be surface-exposed or secreted and their potential to induce protection against disease was tested. *N. meningitidis* is essentially an extracellular pathogen and the major protective response relies on circulating antibodies: complement-mediated bactericidal activity is, in fact, the accepted correlate for in vivo protection and as such is the surrogate endpoint in clinical trials of potential meningococcal vaccines. On the basis of this evidence, the group worked on the assumption that protective antigens are more likely to be found among surface-exposed or secreted proteins. Hence the initial selection of

candidates is based on computer predictions of secretion or surface location. Of the 2,158 predicted ORFs in the *N. meningitidis* genome, 570 were selected by these criteria and could therefore represent new potential vaccine candidates. The selected ORFs were amplified, cloned and analyzed for expression in a heterologous system as either C-terminal His-tag or N-terminal glutathione *S*-transferase fusion proteins. These two expression systems were chosen to achieve the highest level of expression and the easiest purification procedure by a single chromatography step. Of the 570, 350 ORFs were successfully cloned in *E. coli* and purified in a sufficient amount for mice immunizations. Most of the failures, both in cloning and in expression, were related to proteins with more than one transmembrane spanning region. This is likely to be due to toxicity for *E. coli* or to their intrinsic insolubility.

Each purified recombinant protein was used to immunize mice. Immune response was analyzed by Western blot analysis on total cell extracts to verify whether the protein was expressed and by enzyme-linked immunosorbent assay and flow cytometry on whole cells to verify whether the antigen was surface-exposed in meningococcus. Finally, the bactericidal assay was used to evaluate the complement-mediated killing activity of the antibodies (serum bactericidal activity), since this property correlates with vaccine efficacy in humans [9, 10].

Of the 91 proteins found to be positive in at least one of these assays, 28 were able to induce antibodies with bactericidal activity [8]. Several of the antigens previously identified using conventional approaches showed strain variability or were only expressed in some strains, and most of them are effective only against the homologous strains. Therefore, the potential vaccine candidates identified were evaluated for degree of sequence variability among multiple isolates and serogroups of *N. meningitidis*. Many of the newly identified serogroup B antigens included surface-exposed proteins or lipoproteins with a globular structure and without membrane spanning domains and many of them are not abundant on the bacterial surface. Reverse vaccinology has therefore proven to be a rapid and reliable approach to identifying vaccine candidates. In the case of serogroup B, these potential vaccine candidates, able to induce broad strain coverage, were subjected to further evaluation and characterization. The candidates were gradually funneled down, a process that took a further 24 months. Finally, the three most immunogenic antigens on the basis of their ability to induce bactericidal activity or in vivo passive protection were selected to be used in a multicomponent vaccine. They were NHBA [11], fHbp [12, 13], and NadA [14, 15]. Other two antigens (named GNA2091 and GNA1030) were also selected. To further enhance their immunogenicity and facilitate large-scale manufacturing of the vaccine, four of the selected antigens were combined into two fusion proteins so that the resulting protein vaccine contained three recombinant proteins. The antigen NHBA was fused to GNA1030 while GNA2091 was fused to fHbp. NadA was included as a single antigen as it did not perform well when fused to a partner. It is thought that this may be due to the fact that the protein may loose its native trimeric organization. Results showed that the two fusion proteins formulated with aluminum hydroxide induced immunes in both FACS and bactericidal assays. These antibodies were more potent than those induced by the individual antigens. Twenty

micrograms of each of the two fusion proteins and of the NadA antigen were adsorbed to an adjuvant suitable for human use aluminum hydroxide to make the vaccine formulation that was used in subsequent studies. The rationale behind combining antigen was to increase the spectrum of vaccine coverage, minimizing the possibility of bacterial evasion and development of selection mutants [16].

After successful preclinical studies, the MenB vaccine entered the long path of vaccine development that included testing safety and immunogenicity in adult volunteers [17], initial testing in infants [18, 19], and finally, a large-scale, phase III clinical trial that is the basis for a European license application (Fig. 2.1).

2.2.2 *From One to Multiple Genomes: Pan-Genome Reverse Vaccinology*

While the genome sequence of a single strain reveals many aspects of the biology of a species, it fails to address how genetic variability drives pathogenesis within a bacterial species and also limits genome-wide screens for vaccine candidates or for antimicrobial targets to a single strain. The availability of genome sequences for different isolates of a single species enables quantitative analyses of their genomic diversity through comparative genomic analyses. The higher the number of isolates and the broader the selection of strains, the better the estimate of whole species heterogeneity.

The advantage of multiple genome analysis in vaccine design is highlighted by the discovery of universal vaccine candidates against *S. agalactiae*, the group B *Streptococcus* (GBS). GBS is the leading cause of illness and death among newborn infants. Nine distinct capsular serotypes of GBS have been described; however, the major disease-causing isolates in Europe and the United States belong to only five serotypes: Ia, Ib, II, III, and V [20]. A comparative genomic hybridization analysis revealed that there was significant variation in gene content among different clinical isolates of GBS [21], supporting the idea that one genome sequence was not enough to fully capture the diversity of the species and to enable the identification of broadly protective vaccine candidates. For this reason, additional GBS strains belonging to the five major serotypes were sequenced and compared allowing the definition of the species pan-genome [22]. The pan-genome can be defined as the global gene repertoire pertaining to a species. In general, it can be divided in three parts: the core genome, which includes the set of genes invariably present and conserved in all the isolates; the "dispensable genome," comprising genes present in some but not all the strains, and the strain-specific genes, which are present only in one single isolate.

Maione et al. have applied the pan-genome concept to GBS vaccine discovery. Bioinformatic algorithms were used to select from the pan-genome genes that encode putative surface-associated and secreted proteins. After testing a consistent number of selected proteins in a mouse model of infection, four antigens were

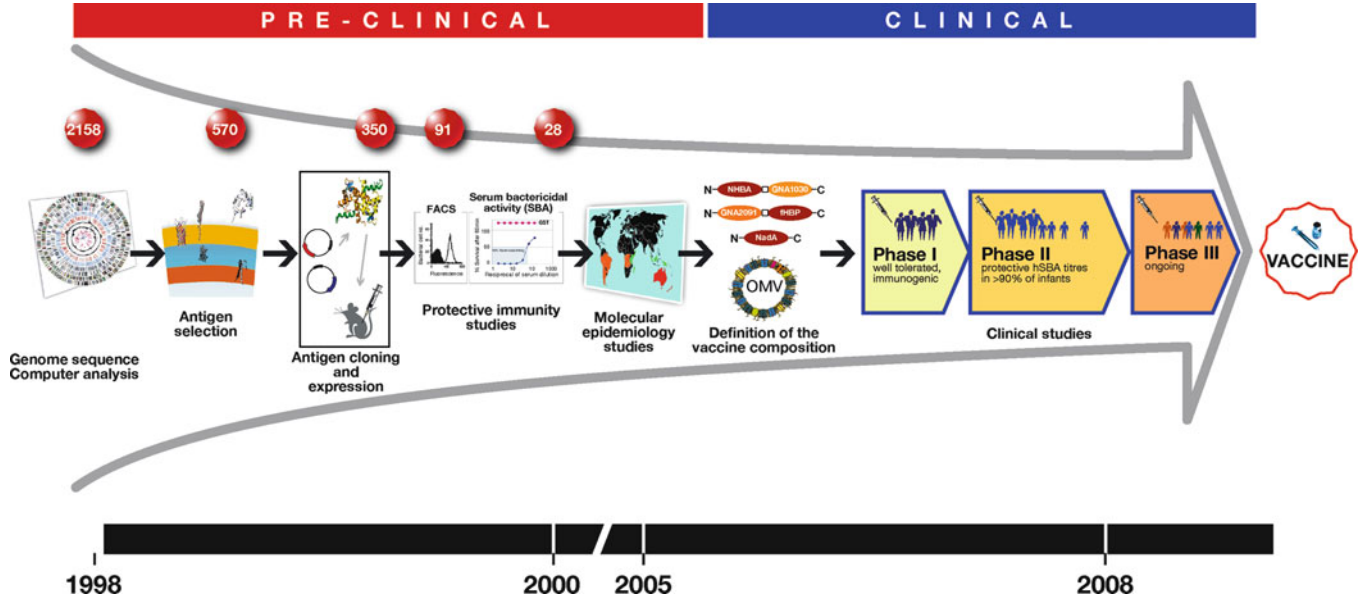

Fig. 2.1 Reverse vaccinology approach applied to *Neisseria meningitidis*. Based on the complete genome sequence of *N. meningitidis* strain MC58 (2158 ORFs), genetic sequences potentially encoding for novel surface exposed proteins were identified. DNA sequences encoding 570 potential surface-exposed antigens were amplified by PCR and cloned into an *Escherichia coli* expression vector. 350 recombinant proteins were successfully produced, purified, and used to immunize mice. The optimal recombinant protein candidates were then selected based on their surface expression and the ability to induce serum bactericidal antibodies: 91 new surface-exposed proteins and 28 novel protein antigens that were able to induce antibodies with bactericidal activity were identified. The antigens selected by reverse vaccinology were prioritized based on these criteria: (i) the protein has to be surface-exposed; (ii) the protein has to be conserved in sequence across a range of different MenB strains; (iii) the protein must induce a broad bactericidal antibody response. The three top antigens that met the prioritization criteria were NadA, fHBP, and NHBA; they gave high bactericidal titers and were bactericidal against most of the strains tested. The recombinant antigens were formulated with outer membrane vesicles and tested in preclinical studies and clinical trials in adults, adolescents, and infants.

identified that were capable of significantly increasing the survival rate among challenged mice. Unexpectedly, only one of these antigens was part of the core genome, the remaining three belonging to the dispensable genome. The final vaccine formulation comprises a combination of the four antigens, which provide overall almost universal strain coverage, with levels of protection similar to those seen when using capsular carbohydrate-based vaccines [23]. This example demonstrates the importance of having access to the genome sequence of multiple strains and performing upfront genome comparisons.

Characterization of the newly identified vaccine antigens revealed that three of them are able to form pilus-like structures extending from the bacterial surface [24]. Furthermore, three pilus islands have been discovered in GBS that encode structurally distinct pilus types [25]. Because of the limited variability of GBS pili, it has been suggested that a combination of three pilin subunits could lead to broad protective immunity against GBS [26].

2.2.3 *The Utility of Analyzing Genome Sequences of Commensal Bacteria: The Subtractive Reverse Vaccinology Approach*

While in the case of GBS, comparative genome analysis has been applied to disease-causing isolates, an alternative approach is the comparison between pathogenic and nonpathogenic strains of the same species. This kind of analysis can provide the information necessary for the identification of antigens that really make the difference in pathogenesis. An elegant example of comparative genome analysis is that described for uropathogenic strains of *E. coli*, whereby a comparison of UPEC genomes with other complete *E. coli* genome sequences was carried out to help identifying sets of "UPEC-specific" and strain-specific proteins, respectively, that may form the basis of their different individual phenotypes and uropathogenic potential. Comparative genome analysis revealed that genome evolution in these bacteria cannot be simply described by a "backbone and flexible gene pool" model, but must also be described by repeated insertions and deletions occurring in certain parts of the genome [27]. Similarly, whole-genome sequence analyses of *Listeria* and *Neisseria* spp. revealed that an extensive genome reduction had occurred in the nonpathogenic species with loss of genes, particularly those involved in virulence, host interaction, and metabolic pathways [28–30].

From a vaccine point of view, genes encoding for antigens conserved both in pathogenic and in nonpathogenic strains could be discarded during the selection, reducing the number of candidates to express and test in the animal model and, consequently, reducing the time for the delivery of a vaccine. This is mainly because the selection of antigens that are selectively present only in pathogenic strains might reduce the impact on the commensal flora, especially for microorganisms that possess both phenotypes, such as *E. coli*. The first example of a subtractive reverse vaccinology approach has been recently published.

Gomes Moriel et al. performed a comparative genome analysis between pathogenic extraintestinal *E. coli* (ExPEC) and commensal *E. coli* strains with the aim of identifying genes specific for pathogenic strains and therefore an ideal target for a vaccine that will not interfere with the natural commensal gut flora of the human host. Authors have selected 230 antigens that have been tested for protection in vivo. Nine antigens were protective in a mouse model of sepsis and the molecular epidemiology analysis showed that these nine antigens are conserved among several pathogenic *E. coli* strains suggesting that by combining some of these antigens a broadly cross-protective vaccine may be developed [31]. A protective antigen (named ECOK1_3385) was identified and active immunization with the antigen or passive protection with sera raised against it provided nearly complete protection from bacteremia and mortality. It is interesting to notice that this antigen was expressed by nonpathogenic strains but not secreted due to the absence of a type II secretion system involved in the surface localization of this antigen [31].

2.2.4 Reverse Vaccinology Explains Biology: Characterization of the Vaccine Antigens Identified

In addition to the discovery of many previously unknown antigens, which have led to successful vaccine development in several instances, reverse vaccinology has made possible studies on antigen function, leading to an understanding of the biology of the pathogen. The two most notable examples are the discovery of factor H binding protein in meningococcus and the discovery of pili in Gram-positive pathogens. In the case of meningococcus, following the publication of the protective antigen GNA1870, two independent laboratories found that this antigen binds the human complement regulator factor H (fH) [32, 33]. This discovery led to an understanding that meningococcus can grow in human blood by downregulating the alternative pathway of complement activation [13]. However, because the same protein was unable to bind Factor H from animal species such as mice and rats [79], the discovery also allowed the understanding of the species specificity of meningococcus and that failure to develop an animal model for meningococcus was because of the fact that the bacterium is unable to grow in the blood of mice and rats. Transgenic animals expressing human factor H may likely be the solution for a meningococcus animal model.

In the second case, while pili had been known for decades to be an essential component for the pathogenesis of Gram-negative bacteria, they were not known to be present in Gram-positive pathogens before the sequencing and analysis of the genomes of *Streptococcus pyogenes*, *S. agalactiae*, and *Streptococcus pneumoniae* [34]. Following screening for protective antigens by reverse vaccinology, it was found that a protective antigen of GBS was a component of a high molecular weight pilus [24]. This new line of research soon led to the identification of similar surface protrusions in *S. pyogenes* and in *S. pneumoniae* [35, 36], revealing a unique

mechanism of pathogenesis for these three important Gram-positive human pathogens.

2.3 Lessons Learned While Exploiting Reverse Vaccinology Approaches

The reverse vaccinology approach has been applied to other pathogenic bacteria including *Bacillus anthracis*, *Porphyromonas gingivalis*, *S. pneumoniae*, *Chlamydia pneumoniae*, and *Brucella melitensis* [37]. All the results obtained so far showed that genome mining allowed to increase the number of candidate vaccine antigens by several orders of magnitude. However, during the development of new vaccine antigens through reverse vaccinology, we understood that parallel genome-based studies can be applied in order to better characterize the candidates and understand their potentials.

In the development of a universal vaccine capable of inducing protection against virtually all circulating strains, accurate characterization of the selected vaccine candidates is highly recommended. The analysis of the sequence conservation of a given antigen is a fundamental aspect in its evaluation as vaccine candidate. Bacterial pathogens use a broad range of microevolutionary tools (like phase variation and antigenic variability) to escape the human immune response.

The analysis should be performed on a strain collection representing globally diverse geographic regions that take into account the target population of the vaccine. In the era of genomics, the availability of multiple genome sequences greatly facilitates the analysis of sequence conservation. However, the number of genomes available is usually not sufficient to cover the diversity of a species or its global geographic representation. Hence, large molecular epidemiology studies on a restricted number of genes encoding the best vaccine antigens are necessary. For example, in the case of MenB, a panel of strains was selected as representative of the meningococcus diversity. In order to evaluate the sequence variability the genes encoding vaccine antigens were sequenced and compared [38]. Performing these studies we learnt that the discovery of previously unknown protein antigens has generated new challenges because the presence and sequence variability of these antigens are not completely aligned with classical typing systems used to characterize the population structures of bacterial pathogens [38]. New antigens usually segregate in the bacterial population in a manner that is independent of conventional markers, such as those used to define serotypes, and genetic markers, such as those used in MLST (multilocus sequence typing), and so we need to identify new ways to type bacteria in order to evaluate the efficacy of new vaccines. In the case of MenB a new typing system, named MATS (meningococcal antigen typing system), has been developed as a basis for molecular epidemiology studies and evaluation of vaccine coverage [39]. Another critical aspect that must be considered during the development of a widely protective vaccine is the expression profile of the antigens

identified. Bacteria have different and complex mechanisms to regulate gene expression [40]. For vaccine discovery programs, it is of key importance to know what genes are expressed during host infection and also to understand the mechanism of regulation of the genes encoding the vaccine antigens. These notions will help to understand (i) whether these genes are expressed during infection so gene products can be the target of the antibodies generated by vaccination and (ii) whether they are expressed in the laboratory conditions used to evaluate antigen immunogenicity (e.g., serum bactericidal assay or opsonophagocitosis).

These lessons suggest that genome mining can be complemented by functional genomics approaches in order to obtain a comprehensive characterization of the antigens identified. For example an upfront comparison of whole-genome sequences from strains representative of the genetic diversity of a bacterial species can be a powerful tool for the selection of the most conserved antigens. Global genome profiling of gene expression in different conditions resembling infection or even better in infected hosts can be an approach to advance the study of genes involved in the pathogenesis and select those expressed in vivo.

2.4 Applications of Functional Genomics and High-Throughput Sequencing in Vaccine Design

Functional genomics methods, linking genotype to phenotype, empower the use of highly parallel methodologies that allow investigators to study all the genes or all the proteins of a pathogen in the context of a host or under various physiological states of interest [3]. These approaches are complementary to in silico antigen identification. These include the large-scale analysis of gene transcription by DNA microarray, the identification of the whole set of proteins encoded by a microorganism by two-dimensional gel electrophoresis and mass spectrometry, as well as the use of protein chips to analyze immunological responses in human sera. All these approaches have been widely applied to vaccine design, with the purpose of identifying sets of genes expressed during infection and involved in pathogenesis (Table 2.1).

In the past few years, unprecedented efforts have been made to develop and deploy new sequencing technologies [2, 41, 42]. The advent of new sequencing technologies that can produce sequence data much more cheaply and speedily than traditional methods has recently transformed the study of many fields of microbiology, from epidemiology to functional genomics.

With respect to traditional sequencers, new sequencing platforms can perform highly parallel sequencing of amplified DNA fragments without the need for cloning, thus generating, in the form of short reads, an output having a much higher order of magnitude. Three technologies have undergone a wide spread on the market: Roche 454 (Roche) [43], Genome Analyzer (Solexa/Illumina) [44], and SOLiDTM (Applied Biosystems) [45]. Their technical features are significantly different and have been described in several recent reviews [46, 47].

Table 2.1 Examples of bacterial genomes that have been explored for vaccine components using functional genomic approaches

Pathogen	Disease	Brief description of the approaches	References
Neisseria meningitidis B	Major cause of bacterial septicemia and meningitis	Reverse vaccinology, see text for details	[8]
		Microarray—Analysis of the MenB transcriptome during adhesion to host epithelial cells led to the identification of 189 genes with increased expression under conditions that mimicked in vivo host–pathogen interactions. Twelve of these genes were confirmed by FACS analysis to express surface proteins accessible to the immune system (with four of these being detected only after adhesion to epithelial cells), five of which induced bactericidal protective antibodies in mice	[50]
Streptococcus pneumoniae	Most common cause of fatal community-acquired pneumonia in the elderly and is also one of the most common causes of middle ear infections and meningitis in children	Reverse vaccinology: all the ORFs of the genome sequence of a clinical isolate of *S. pneumoniae* were evaluated to determine whether the gene products contained sequence motifs predictive of their localization on the surface of the bacterium. This led to the identification of 130 ORFs. Mice were immunized with 108 of these proteins and 6 were shown to confer protection against disseminated *S. pneumoniae* infection. All the six protective antigens were broadly distributed among several pneumococcus strains and showed immunogenicity during human infection	[51]
		Comparative genomics: the genome of an avirulent strain (R6) of *S. pneumoniae* has been sequenced. Comparative genome hybridization using DNA arrays revealed differences between the genomes of avirulent and virulent *S. pneumoniae*, which could contribute to differences in virulence and antigenicity. This comparison might lead to the	[52]

		identification of some specific proteins as potential target for vaccine development	
Staphylococcus aureus	Infects wounds and causes severe infections. Following acquisition of resistance to most available antibiotics has emerged as an important opportunistic pathogen	Genomic peptide libraries: *S. aureus* peptides were displayed on the surface of *Escherichia coli* via fusion to one or two outer membrane proteins (LamB and FhuA) and probed with sera selected for high antibody titers and opsonic activity. The exhaustive screening of these libraries by magnetic cell sorting determines the profile of antigens, which are expressed in vivo and elicit an immune response in humans. A total of 60 antigenic proteins were identified	[53]
		Serological proteome analysis: A surface protein preparation from *S. aureus* was resolved by 2D electrophoresis and analyzed by immunoblotting using two pools, each consisting of five sera coming from healthy donors or patients. Twenty-one spots were isolated and analyzed by mass spectrometry allowing the identification of 15 proteins including known and new vaccine candidates	[54]
Porphyromonas gingivalis	Periodontal pathogen that has been implicated in the etiology of chronic adult periodontitis	Reverse vaccinology: applying a series of bioinformatics tools 120 putative new antigens have been identified from the genome of *P. gingivalis*. The selected genes were cloned and expressed in *E. coli* and screened by Western blot using sera from human periodontitis patients. These candidates were reduced to a set of 40 proteins, which were purified and used to immunize mice that were subsequently challenged with live bacteria in a subcutaneous abscess model. Two antigens demonstrated protection in this model of infection and therefore could represent potential vaccine candidates	[55]

(continued)

Table 2.1 (continued)

Pathogen	Disease	Brief description of the approaches	References
Streptococcus agalactiae (Group B streptococcus)	Leading cause of bacterial sepsis, pneumonia, and meningitis in neonates in the United States and Europe	Proteomics: Proteome analysis of the outer surface proteins of this pathogen allowed the discovery of novel surface proteins. Sera raised against some of these proteins were protective in a neonatal animal model against a lethal dose of the pathogen	[56]
		Pan genome reverse vaccinology—see text for details	[23]
Streptococcus pyogenes (Group A streptococcus)	Causes many human infections ranging from mild pharyngitis to severe diseases, including toxic shock syndrome, necrotizing fasciitis, and rheumatic fever	Comparative genomics: Analysis of the genome of four GAS strains led to the discovery of four new extracellular proteins. These proteins are very well conserved as observed applying sequencing and genetic population analysis. Western immunoblot confirmed that all four proteins are made during the course of distinct GAS infections and immunization with the purified form of one of these can confer protection in a murine model of infection	[57]
		Surface proteome—Surface digestion of live bacteria with different proteases allowed fast and consistent identification of proteins that are expressed on the bacteria surface and thus exposed to the immune system. The cell-surface peptide fragments generated after protease treatment of GAS strain SF370 were recovered, concentrated, and analyzed by tandem mass spectrometry and identified using bioinformatic examination. Seventy-two proteins were identified, of which only four were predicted by the PSORT algorithm to be cytoplasmic proteins indicating that the method was highly specific for surface-exposed proteins. Among the surface proteins identified it was proven that some induce protection in an animal model of infection	[58]

Chlamydia pneumoniae	Causes pneumonia and is also associated with atherosclerotic and cardiovascular disease	Reverse vaccinology, proteomic: As a result of in silico analysis of *C. pneumoniae* genome, 157 putative surface-exposed proteins have been identified. Recombinant forms were expressed in *E. coli*, purfied and used to immunize mice. Antisera were used to detect cell-surface localization by FACS analysis. 2D gel electrophoresis and mass spectrometry were used to confirm the expression of the FACS-positive antigens in the elementary body phase of development. The result of these systematic genome–proteome combined approach allowed the identification of 28 new vaccine candidate antigens	[59]
ExPEC	Extraintestinal pathogenic *E. coli* (ExPEC) are the cause of a diverse spectrum of invasive infections in humans and animals, leading to urinary tract infections, meningitis, or septicemia	Subtractive reverse vaccinology—see text for details	[31]
		Surface proteome of OMV: outer membrane vesicles (OMVs) produced by a tolR mutant of the pathogenic IHE3034 strain were analyzed by 2D electrophoresis coupled to mass spectrometry. The analysis led to the identification of 100 proteins, most of which are localized to the outer membrane and periplasmic spaces. Interestingly, seven of the identified proteins appear to be specific for pathogenic *E. coli* and therefore are potential targets for vaccine and drug development	[60]

The high amount of information that can result from high-throughput sequencing applications, thanks also to the development of computational tools able to manage so much sequence data, is improving our understanding of the complexity of bacterial genomes. One of the next challenges of vaccinology is to apply these technologies to help and improve the identification and characterization of vaccine antigens in the reverse vaccinology approach.

2.4.1 Studies on Antigen Variability and Population Structures

In the reverse vaccinology approaches conducted thus far, molecular epidemiology studies on antigen presence and variability have been conducted on a small number of antigens when they have been already selected and tested for expression and immunogenicity. On the contrary, it would be useful to start new reverse vaccinology projects by sequencing the genomes of a collection of strains representing globally diverse geographic areas that take into account the target population of the vaccine. The application of the new sequencing technologies will make this approach feasible. The large collection of genomic data generated will help upfront in the identification of the most conserved protein antigens. Moreover, genomic information will help to define the dimension of the pan-genome and to characterize the population structure, shedding light on the complexity of bacterial species. It will be possible to identify the major lineages inside the population and their geographical distribution and also evaluate how genetic exchange occurs between lineages, leading to a better understanding of the mechanisms involved in the evolutionary process of a bacterial species against which a vaccine has to be designed (Fig. 2.2a).

Recent works have demonstrated that new sequencing technologies have the potential to study the dynamics of bacterial pathogens. For example, studying a single strain of methicillin-resistant *Staphylococcus aureus*, Harris et al. demonstrated the potential of this approach in tracing transmission events over a short timescale, within a hospital environment or between hospitals, and in identifying the sources of outbreaks [48]. By comparing whole-genome sequences of many isolates of the same antibiotic-resistant strain of *S. pneumoniae*, Croucher et al. reconstructed its evolution and global spread. Moreover, they shed light on the effects that clinical practices as antibiotic use and vaccination can have on the evolution of this pathogen [49].

New sequencing technologies will also open up opportunities for monitoring pathogen vaccine escape by screening for evidence of immune selection in the genomes of pathogen populations before and after vaccination. By deep sequencing of microbial populations it will be possible to identify antigens under immune selection by monitoring the clustering of single nucleotide polymorphisms and other mutations that affect protein sequence.

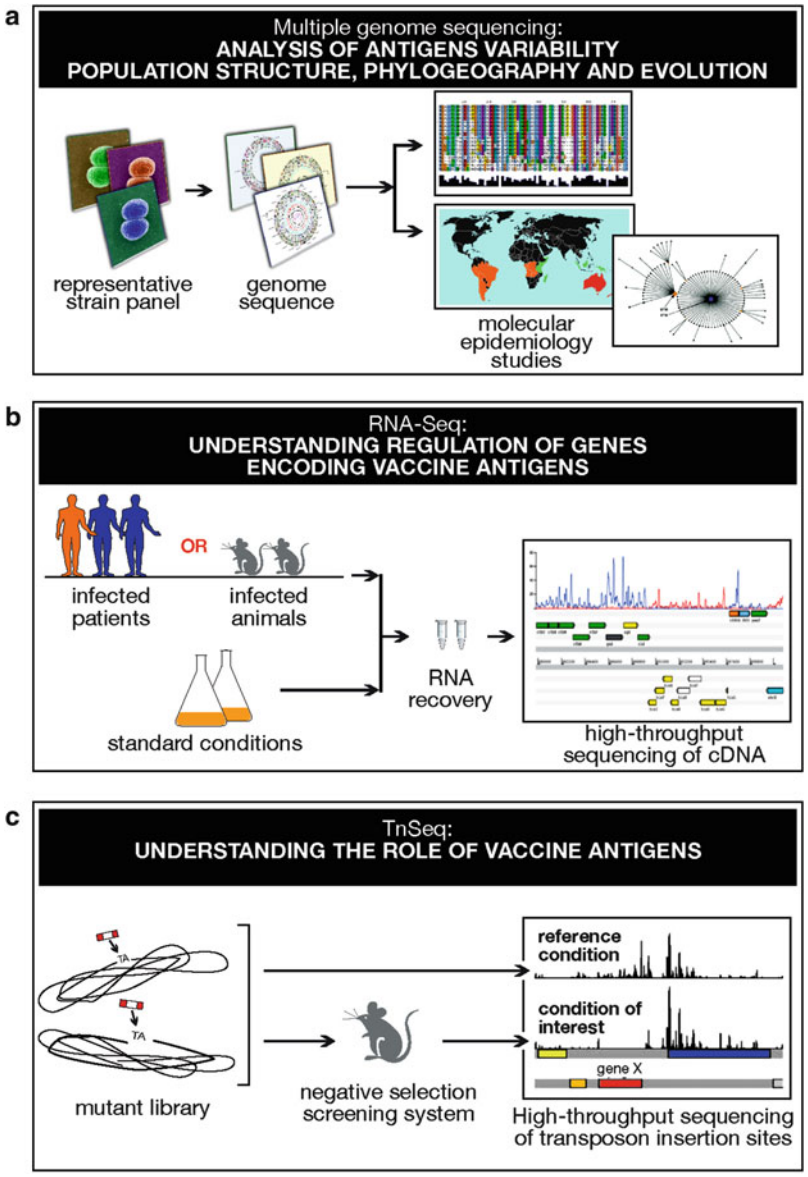

Fig. 2.2 Application of novel sequencing approaches to vaccine antigens identification. Three applications of next-generation sequencing technologies and their impact on vaccine design are schematically represented. (**a**) Analysis of antigen variability, distribution, and evolution by high-throughput sequencing of different isolates. (**b**) High-throughput transcriptional profiling of antigens by RNA-Seq in different in vitro, ex vivo, and in vivo conditions. (**c**) High-throughput functional characterization of vaccine antigens in different conditions mimicking infection by TnSeq.

2.4.2 Transcriptomics

To rapidly respond to environmental changes, bacteria have evolved different and complex mechanisms to regulate gene expression. While the genome carries all the necessary genes, bacteria can save considerable energy by expressing only those that are required in a particular stage of their life cycle [40]. Transcriptional regulation is an important aspect that needs to be considered when designing a new vaccine. During the phase of antigen selection, transcriptional profiling of bacteria in conditions mimicking infection can lead to the identification of novel virulence factors that can be considered as vaccine antigens. Moreover, once a panel of candidates has been selected, antigens of interest can be better characterized by transcriptional analysis, to understand which are the conditions influencing their expression.

From a vaccine coverage point of view, an important aspect to take into consideration is that discrepancy between the sequence conservation of an antigen and its expression can occur. Even when gene conservation is high, variability can occur in terms of expression. There may be several reasons for such discrepancy. First, the genomic organization can vary from strain to strain thus affecting gene expression. Second, activity of transcriptional regulators can also vary, and this is not predictable from the nucleotide sequence of the selected antigens. Therefore transcriptional studies must be carried out, as an approach complementary to genome mining, in order to select antigens that are expressed in the conditions of interest used to assess the functionality of an antigen. For example, in the case of MenB we use an in vitro serum bactericidal assay to assess the ability of antibodies to induce complement-mediated killing. It is fundamental that the target antigen under evaluation be expressed in the particular in vitro condition otherwise we might underestimate the role of the antigen as a vaccine candidate. If sufficient information is known about the mechanism of regulation of the gene encoding an antigen, we may think to modify the in vitro assay performing it in the conditions where it is expressed. An example is represented by the NadA antigen, which is highly regulated and repressed during standard in vitro growth conditions {[15] #642;[80] #29037}.

Last but not least, the analysis of the transcriptional profile of pathogens during infection or in conditions mimicking infection has considerably improved our understanding of the mechanisms implicated in host–pathogen interaction.

In the last few years, DNA microarrays have been extensively used to simultaneously monitor the expression levels of all annotated genes in a bacterial population (Table 2.1). Gene expression can be studied by growing pathogens in an appropriate in vivo or ex vivo model of infection (e.g., cell and tissue cultures, animal models) and, after recovering the bacteria for RNA preparation and labeling, the gene activity is analyzed and compared to the expression of the genes under in vitro conditions [61].

The need for a more precise and comprehensive identification of microbial transcriptomes, independently of any annotated sequence feature, recently led to

the development of tiling microarrays, which generally represent both strands and the intergenic regions of the genome. These new arrays have revealed an unexpectedly high degree of transcriptome complexity in bacterial species, allowing detection of unannotated genes, noncoding RNAs, alternative and antisense transcripts, and untranslated regions [62]. However, hybridization-based approaches present several limitations. In particular, they rely upon existing knowledge about genome sequence and the dynamic range of detection is limited by cross-hybridization and saturation signals.

Following progress in DNA sequencing and development of new computational tools, genome-wide studies on bacterial gene expression have begun to shift from microarray technology to a new method termed RNA sequencing (RNA-seq), in which high-throughput sequencing of cDNA is carried out.

Over preexisting approaches, RNA-Seq presents several advantages. Following sequencing, the resulting reads are either aligned to a reference genome or assembled de novo to produce a transcription map. This also means that transcripts that do not correspond to existing genomic sequences can be detected. RNA-Seq provides a far more precise measurement of the levels of transcripts than other methods, having very low if any background signal and no upper limit for quantification, which correlates with the number of sequences obtained. Transcripts can be precisely located to a single-base resolution. Moreover, it can help genome sequencing in revealing sequence variations in the transcribed regions [63].

In the last few years, RNA-seq has been applied to the transcriptome analysis of several pathogens. Yoder-Himes and coworkers used RNA-Seq to examine the transcriptional response of two closely related strains of *Burkholderia cenocepacia*, one isolated from a cystic fibrosis patient and the other from soil, in conditions mimicking human sputum and soil. Despite the high degree of DNA sequence similarity of the two strains, the authors were able to detect a large number of regulatory differences between them, which may represent specific adaptations to the niches from which they were isolated [64]. Similarly, the RNA-seq approach has been successfully applied to the transcriptome analysis of other species, such as *Salmonella enterica* serovar Typhi [65], *Listeria monocytogenes* [66], *B. anthracis* [67], *Acinetobacter baumannii* [68], *Chlamydia thracomatis* [69], and *Helicobacter pylori* [70].

We can predict that in the near future this technique, enabling the simultaneous transcriptomic analysis of all the antigens of interest more speedily and precisely than it was previously possible, will become a helpful tool for antigen identification and characterization. Moreover, without any need to accurately design probed arrays, RNA-Seq will allow comparison of the expression profiles of multiple strains. We can also imagine that RNA-seq will be used in some applications that aim to characterize host–pathogen interactions without any need to separate the two RNA populations and allowing to understand not only how the pathogen adapts to cause an infection but also how the host (or an animal model) responds to the infection. For example, RNA could be isolated from infected patients and both the pathogen and the host transcriptional responses analyzed (Fig. 2.2b).

Knowledge of strains and conditions in which an antigen of interest is expressed can provide fundamental complementary information regarding the potential protective capability of each antigen that will help in the prioritization of the antigens. Moreover, analyzing expression levels on a whole-transcriptome scale and in a high-throughput fashion can lead in a short time to the identification of in vitro and in vivo conditions in which an antigen is expressed. These will be the conditions in which that antigen can be immunologically characterized and its function investigated in depth.

2.4.3 Identification of Genes Involved in Pathogenesis

While RNA expression can provide indication on the expression of an antigen in a certain condition related to pathogenesis, other approaches are needed to demonstrate its role in the establishment of the disease or the survival of the pathogen in the host. In order to identify bacterial genes involved in survival and pathogenesis, a variety of methods have been set up in the past few years. All these approaches rely on random mutagenesis and require a detection system that allows simultaneous screening of a pool of mutants.

A strategy that has been extensively used to evaluate gene expression in vivo is IVET (in vivo expression technology), which identifies promoters that are active in vivo but not in vitro using a library of random genomic fragments ligated to a promoterless reporter gene [71]. Another methodology that has been used is STM (signature-tagged mutagenesis). In STM, unique hybridization tags are introduced into each mutant by transposon mutagenesis [72]. All the mutants are pooled and used to infect an animal or alternatively are exposed to a selected condition mimicking a stage of infection. "Negative selection" of the clones unable to survive is then carried out by comparing hybridization of the input pool (the entire mutant library) with that of the output pool (a subset of the library containing all the survived mutants). This approach allowed the identification of *Salmonella typhimurium* mutants showing an attenuated virulence in a murine model of typhoid fever [72]; similarly, *N. meningitidis* genes involved in the establishment of systemic infection could be selected in an infant rat model [73] (Table 2.1).

These methods allow selection of virulence factors on a large scale, providing useful information for both identification of new vaccine candidates and their prioritization.

Taking advantage of next-generation sequencing technology, a new method has recently been developed. In this method, called Tn-Seq, a saturated transposon insertion library is generated. Upon growth under a test condition of interest (in vitro or in vivo) insertion mutants with a lower fitness decrease in the population, while other mutants remain the same or increase in frequency. Changes in frequency are determined by high-throughput sequencing of the transposon flanking regions [74, 75].

Using a saturated mariner transposon insertion library, van Opijnen et al. have been able to evaluate fitness of *S. pneumoniae* genes in standard laboratory conditions [74]. Langridge et al. have used a similar approach, named TraDIS (transposon-directed insertion-site sequencing), to assay simultaneously every gene of *S. enterica* serovar Typhi for both essentiality in standard laboratory growth conditions and contribution toward bile tolerance. They clearly show how the possibility to sequence at high resolution large numbers of mutants allows the screening of very rich libraries. Moreover, the semi-quantitative nature of the assay allows the identification not only of essential and dispensable genes, but also advantageous and disadvantageous genes during growth [76]. Using mouse as a model of pulmonary infection Gawronski et al have identified genes required by *H. influenzae* to resist host defenses during lung pathogenesis [77]. This study shows the potential of this Tn-seq approach to provide new vaccine targets for prevention of *H. influenzae* pulmonary infection.

Providing a link between genotype and phenotype in a high-throughput fashion, Tn-Seq is likely to become the tool of choice for the characterization of bacterial pathogenesis (Fig. 2.2c). Given a selected condition, the entire gene content of a pathogen can be screened speedily and at unprecedented resolution. Moreover, negative selection can be carried out in multiple conditions, mimicking different stages of infection, thus revealing which genes are most important to establish successful infection. Tn-Seq can thus become very helpful in screening vaccine candidates on a whole-genome scale on the basis of their role in vivo.

2.5 Conclusions

Genomics has introduced a new paradigm to bacterial pathogenesis. Instead of dissecting bacterial components in vitro, the new approach starts with the complete information on the genome sequence and then identifies the important factors in virulence. Moreover, the availability of complete genome sequence information of many pathogens has led to a new paradigm in vaccine development. If a suitable assay is available, every protein synthesized by the pathogen can be tested as a vaccine candidate without any prior selection based on incomplete knowledge of the pathogenicity and immunogenicity of the organism. Several years into the reverse vaccinology approach, we have a previously unbelievable insight into a pathogen's genome and its use to identify new vaccine candidates. However, the approaches used for vaccine development are continually being refined based on improved understanding of microbial molecular epidemiology, evolution, virulence, host–pathogen interactions and increased understanding of the complexity of microbial communities as well as improved genome-based technologies. One, or a combination, of the approaches described in this chapter typically drives vaccine discovery projects, with the approach used being heavily dependent on the characteristics of the target pathogen and the vaccine antigens to be identified.

Finally, the progress made in the genomic era has finally put the realization of vaccines for many pathogens within reach as demonstrated by the new MenB vaccine recently developed. Many diseases still cannot be controlled by vaccination and more infectious diseases are expected to emerge or reemerge. In this new era of vaccine development, we would expect the evolution of genome-based approaches that will provide a number of candidate antigens for new vaccines to control and eradicate microbial infections.

References

1. Fleischmann RD et al (1995) Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science 269(5223):496–512
2. Medini D et al (2008) Microbiology in the post-genomic era. Nat Rev Microbiol 6(6):419–430
3. Raskin DM et al (2006) Bacterial genomics and pathogen evolution. Cell 124(4):703–714
4. Rappuoli R (2001) Reverse vaccinology, a genome-based approach to vaccine development. Vaccine 19(17–19):2688–2691
5. Stephens DS, Greenwood B, Brandtzaeg P (2007) Epidemic meningitis, meningococcaemia, and Neisseria meningitidis. Lancet 369(9580):2196–2210
6. Caugant DA, Maiden MC (2009) Meningococcal carriage and disease—population biology and evolution. Vaccine 27(suppl 2):B64–B70
7. Tettelin H et al (2000) Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. Science 287(5459):1809–1815
8. Pizza M et al (2000) Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. Science 287(5459):1816–1820
9. Borrow R, Balmer P, Miller E (2005) Meningococcal surrogates of protection–serum bactericidal antibody activity. Vaccine 23(17–18):2222–2227
10. Goldschneider I, Gotschlich EC, Artenstein MS (1969) *Human immunity to the meningococcus*. I. The role of humoral antibodies. J Exp Med 129(6):1307–1326
11. Serruto D et al (2010) Neisseria meningitidis GNA2132, a heparin-binding protein that induces protective immunity in humans. Proc Natl Acad Sci USA 107(8):3770–3775
12. Masignani V et al (2003) Vaccination against Neisseria meningitidis using three variants of the lipoprotein GNA1870. J Exp Med 197(6):789–799
13. Seib KL et al (2009) Factor H-binding protein is important for meningococcal survival in human whole blood and serum and in the presence of the antimicrobial peptide LL-37. Infect Immun 77(1):292–299
14. Capecchi B et al (2005) Neisseria meningitidis NadA is a new invasin which promotes bacterial adhesion to and penetration into human epithelial cells. Mol Microbiol 55 (3):687–698
15. Comanducci M et al (2002) NadA, a novel vaccine candidate of Neisseria meningitidis. J Exp Med 195(11):1445–1454
16. Giuliani MM et al (2006) A universal vaccine for serogroup B meningococcus. Proc Natl Acad Sci USA 103(29):10834–10839
17. Kimura A et al (2011) Immunogenicity and safety of a multicomponent meningococcal serogroup B vaccine and a quadrivalent meningococcal CRM197 conjugate vaccine against serogroups A, C, W-135, and Y in adults who are at increased risk for occupational exposure to meningococcal isolates. Clin Vaccine Immunol 18(3):483–486
18. Findlow J et al (2010) Multicenter, open-label, randomized phase II controlled trial of an investigational recombinant Meningococcal serogroup B vaccine with and without outer membrane vesicles, administered in infancy. Clin Infect Dis 51(10):1127–1137

19. Snape MD et al (2010) Immunogenicity and reactogenicity of a 13-valent-pneumococcal conjugate vaccine administered at 2, 4, and 12 months of age: a double-blind randomized active-controlled trial. Pediatr Infect Dis J 29(12):e80–e90
20. Johri AK et al (2006) Group B Streptococcus: global incidence and vaccine development. Nat Rev Microbiol 4(12):932–942
21. Tettelin H et al (2002) Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V Streptococcus agalactiae. Proc Natl Acad Sci USA 99 (19):12391–12396
22. Tettelin H et al (2005) Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pan-genome". Proc Natl Acad Sci USA 102 (39):13950–13955
23. Maione D et al (2005) Identification of a universal Group B streptococcus vaccine by multiple genome screen. Science 309(5731):148–150
24. Lauer P et al (2005) Genome analysis reveals pili in Group B Streptococcus. Science 309 (5731):105
25. Rosini R et al (2006) Identification of novel genomic islands coding for antigenic pilus-like structures in Streptococcus agalactiae. Mol Microbiol 61(1):126–141
26. Margarit I et al (2009) Preventing bacterial infections with pilus-based vaccines: the group B streptococcus paradigm. J Infect Dis 199(1):108–115
27. Brzuszkiewicz E et al (2006) How to become a uropathogen: comparative genomic analysis of extraintestinal pathogenic Escherichia coli strains. Proc Natl Acad Sci USA 103 (34):12879–12884
28. Hain T et al (2007) Pathogenomics of Listeria spp. Int J Med Microbiol 297(7–8):541–557
29. Perrin A et al (2002) Comparative genomics identifies the genetic islands that distinguish Neisseria meningitidis, the agent of cerebrospinal meningitis, from other Neisseria species. Infect Immun 70(12):7063–7072
30. Perrin A, Nassif X, Tinsley C (1999) Identification of regions of the chromosome of Neisseria meningitidis and Neisseria gonorrhoeae which are specific to the pathogenic Neisseria species. Infect Immun 67(11):6119–6129
31. Moriel DG et al (2010) Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic Escherichia coli. Proc Natl Acad Sci USA 107 (20):9072–9077
32. Madico G et al (2007) Factor H binding and function in sialylated pathogenic neisseriae is influenced by gonococcal, but not meningococcal, porin. J Immunol 178(7):4489–4497
33. Schneider MC et al (2007) Interactions between Neisseria meningitidis and the complement system. Trends Microbiol 15(5):233–240
34. Telford JL et al (2006) Pili in gram-positive pathogens. Nat Rev Microbiol 4(7):509–519
35. Barocchi MA et al (2006) A pneumococcal pilus influences virulence and host inflammatory responses. Proc Natl Acad Sci USA 103(8):2857–2862
36. Mora M et al (2005) Group A Streptococcus produce pilus-like structures containing protective antigens and Lancefield T antigens. Proc Natl Acad Sci USA 102(43):15641–15646
37. Adu-Bobie J et al (2003) Two years into reverse vaccinology. Vaccine 21(7–8):605–610
38. Bambini S et al (2009) Distribution and genetic variability of three vaccine components in a panel of strains representative of the diversity of serogroup B meningococcus. Vaccine 27 (21):2794–2803
39. Donnelly J et al (2010) Qualitative and quantitative assessment of meningococcal antigens to evaluate the potential strain coverage of protein-based vaccines. Proc Natl Acad Sci USA 107 (45):19490–19495
40. Telford JL (2008) Bacterial genome variability and its impact on vaccine design. Cell Host Microbe 3(6):408–416
41. Hall N (2007) Advanced sequencing technologies and their wider impact in microbiology. J Exp Biol 210(Pt 9):1518–1525

42. Schuster SC (2008) Next-generation sequencing transforms today's biology. Nat Methods 5 (1):16–18
43. Margulies M et al (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437(7057):376–380
44. Bennett ST et al (2005) Toward the 1,000 dollars human genome. Pharmacogenomics 6 (4):373–382
45. Shendure J et al (2005) Accurate multiplex polony sequencing of an evolved bacterial genome. Science 309(5741):1728–1732
46. Ansorge WJ (2009) Next-generation DNA sequencing techniques. N Biotechnol 25 (4):195–203
47. Metzker ML (2010) Sequencing technologies—the next generation. Nat Rev Genet 11 (1):31–46
48. Harris SR et al (2010) Evolution of MRSA during hospital transmission and intercontinental spread. Science 327(5964):469–474
49. Croucher NJ et al (2011) Rapid pneumococcal evolution in response to clinical interventions. Science 331(6016):430–434
50. Grifantini R et al (2002) Previously unrecognized vaccine candidates against group B meningococcus identified by DNA microarrays. Nat Biotechnol 20(9):914–921
51. Wizemann TM et al (2001) Use of a whole genome approach to identify vaccine molecules affording protection against Streptococcus pneumoniae infection. Infect Immun 69 (3):1593–1598
52. Hoskins J et al (2001) Genome of the bacterium Streptococcus pneumoniae strain R6. J Bacteriol 183(19):5709–5717
53. Etz H et al (2002) Identification of in vivo expressed vaccine candidate antigens from Staphylococcus aureus. Proc Natl Acad Sci USA 99(10):6573–6578
54. Vytvytska O et al (2002) Identification of vaccine candidate antigens of Staphylococcus aureus by serological proteome analysis. Proteomics 2(5):580–590
55. Ross BC et al (2001) Identification of vaccine candidate antigens from a genomic analysis of Porphyromonas gingivalis. Vaccine 19(30):4135–4142
56. Hughes MJ et al (2002) Identification of major outer surface proteins of Streptococcus agalactiae. Infect Immun 70(3):1254–1259
57. Reid SD et al (2002) Postgenomic analysis of four novel antigens of group a streptococcus: growth phase-dependent gene transcription and human serologic response. J Bacteriol 184 (22):6316–6324
58. Rodriguez-Ortega MJ et al (2006) Characterization and identification of vaccine candidate proteins through analysis of the group A Streptococcus surface proteome. Nat Biotechnol 24 (2):191–197
59. Montigiani S et al (2002) Genomic approach for analysis of surface proteins in Chlamydia pneumoniae. Infect Immun 70(1):368–379
60. Berlanda Scorza F et al (2008) Proteomics characterization of outer membrane vesicles from the extraintestinal pathogenic Escherichia coli DeltatolR IHE3034 mutant. Mol Cell Proteomics 7(3):473–485
61. Serruto D et al (2004) Biotechnology and vaccines: application of functional genomics to Neisseria meningitidis and other bacterial pathogens. J Biotechnol 113(1–3):15–32
62. Sorek R, Cossart P (2010) Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. Nat Rev Genet 11(1):9–16
63. Croucher NJ, Thomson NR (2010) Studying bacterial transcriptomes using RNA-seq. Curr Opin Microbiol 13(5):619–624
64. Yoder-Himes DR et al (2009) Mapping the Burkholderia cenocepacia niche response via high-throughput sequencing. Proc Natl Acad Sci USA 106(10):3976–3981
65. Perkins TT et al (2009) A strand-specific RNA-Seq analysis of the transcriptome of the typhoid bacillus Salmonella typhi. PLoS Genet 5(7):e1000569

66. Mraheil MA et al (2011) The intracellular sRNA transcriptome of Listeria monocytogenes during growth in macrophages. Nucleic Acids Res 39(10):4235–4248
67. Passalacqua KD et al (2009) Structure and complexity of a bacterial transcriptome. J Bacteriol 191(10):3203–3211
68. Camarena L et al (2010) Molecular mechanisms of ethanol-induced pathogenesis revealed by RNA-sequencing. PLoS Pathog 6(4):e1000834
69. Albrecht M et al (2010) Deep sequencing-based discovery of the Chlamydia trachomatis transcriptome. Nucleic Acids Res 38(3):868–877
70. Sharma CM et al (2010) The primary transcriptome of the major human pathogen Helicobacter pylori. Nature 464(7286):250–255
71. Mahan MJ, Slauch JM, Mekalanos JJ (1993) Selection of bacterial virulence genes that are specifically induced in host tissues. Science 259(5095):686–688
72. Hensel M et al (1995) Simultaneous identification of bacterial virulence genes by negative selection. Science 269(5222):400–403
73. Sun YH et al (2000) Functional genomics of Neisseria meningitidis pathogenesis. Nat Med 6 (11):1269–1273
74. van Opijnen T, Bodi KL, Camilli A (2009) Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. Nat Methods 6(10):767–772
75. van Opijnen, T. and A. Camilli, Genome-wide fitness and genetic interactions determined by Tn-seq, a high-throughput massively parallel sequencing method for microorganisms. Curr Protoc Microbiol, 2010. Chapter 1: p. Unit1E 3.
76. Langridge GC et al (2009) Simultaneous assay of every Salmonella Typhi gene using one million transposon mutants. Genome Res 19(12):2308–2316
77. Gawronski JD et al (2009) Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for Haemophilus genes required in the lung. Proc Natl Acad Sci USA 106(38):16422–16427
78. Rappuoli R (2004) From Pasteur to genomics: progress and challenges in infectious diseases. Nat Med 10(11): 1177–85
79. Granoff DM et al. (2009) Binding of complement factor H (fH) to Neisseria meningitidis is specific for human fH and inhibits complement activation by rat and rabbit sera. Infect Immun 77(2): 764–9
80. Metruccio MM et al (2009) A novel phase variation mechanism in the meningococcus driven by a ligand-responsive repressor and differential spacing of distal promoter elements. PLoS Pathog 5(12): e1000710

http://www.springer.com/978-1-4614-5069-6

Immunomic Discovery of Adjuvants and Candidate Subunit Vaccines

Flower, D.R.; Perrie, Y. (Eds.)
2013, X, 314 p., Hardcover
ISBN: 978-1-4614-5069-6