Developing vaccines in the era of genomics: a decade of Reverse Vaccinology

Kate L. Seib, Xin Zhao and Rino Rappuoli*

Novartis Vaccines and Diagnostics, Via Fiorentina 1, 53100, Siena, Italy

***Corresponding Author:** Rino Rappuoli. Novartis Vaccines and Diagnostics, Via Fiorentina 1, 53100, Siena, Italy. Phone: +39 0577243414. Fax: +39 0577278508.

Running title: Genomics and Reverse Vaccinology

Category: Review (CMI vaccine supplement article)

Abstract

Vaccines have a significant impact on public health, and vaccinology in the era of genomics is taking advantage of new technologies to tackle diseases for which vaccine development has so far been unsuccessful. Almost all existing vaccines were developed based on traditional vaccinology methods, which relied on empirical screening of a few candidates at a time based on known features of the pathogen. However, the ability to sequence a pathogen's genome provides access to its entire antigenic repertoire. As such, genomics has catalysed a shift in vaccine development towards sequence-based 'Reverse Vaccinology' approaches, which use high-throughput *in silico* screening of the entire genome of a pathogen to identify genes that encode proteins with the attributes of good vaccine targets. Furthermore, the increasing availability of genome sequences has led to the development and application of additional technologies to vaccine discovery, including comparative genomics, transcriptomics, proteomics, immunomics and structural genomics. Vaccine candidates identified from a pathogen's genome or proteome can then be expressed as recombinant proteins and tested in appropriate *in vitro* or *in vivo* models to assess immunogenicity and protection. The process of Reverse Vaccinology has been applied to several pathogens, including serogroup B *Neisseria meningitidis*, *Streptococcus agalactiae*, *Streptococcus pyogenes*, *Streptococcus pneumoniae* and pathogenic *Escherichia coli,* and has provided scores of new candidate antigens for pre-clinical and clinical investigation. As novel genome-based technologies continue to emerge, it is expected that new vaccines for unmet diseases will be within reach.

Key words: Reverse Vaccinology; genomics; proteomics; *Neisseria meningitidis; Streptococcus agalactiae*; *Streptococcus pyogenes*; *Streptococcus pneumoniae*; pathogenic *Escherichia coli.*

Introduction

Vaccination is one of the most successful and cost-effective ways of improving public health and has been used to save lives for at least two centuries Traditional vaccine development was largely based on Pasteur's principles, which are to "isolate, inactivate, and inject" the disease-causing agent thus producing vaccines composed of inactivated, live attenuated or purified subunits of a microorganism (1). However, sequencing of the first bacterial genome, *Haemophilus influenzae* in 1995, and the arrival of the era of microbial genomics have led to many changes in vaccinology. The continuing advances in genome sequencing technologies and bioinformatics now enable researchers to explore a microorganism's 'book of life', its genome, for antigen discovery. Microbial genomes contain the complete repertoire of possible antigens and are used as the starting point to capture information for vaccine development. This process, known as "Reverse Vaccinology" (2), was first proposed in 2000 based on the identification of novel meningococcal vaccine candidates from the genome sequence of a *Neisseria meningitidis* serogroup B strain. The use of Reverse Vaccinology over the last decade has changed the way vaccine development is approached and this new era of vaccine research has been referred to as a renaissance of vaccinology (Figure 1).

The original idea behind Reverse Vaccinology was to start *in silico* to screen the entire genome of a pathogen to identify genes that encode proteins with the attributes of good vaccine targets, for example, proteins that are predicted to be surface exposed and that are well conserved between strains. These selected proteins are then expressed in *Escherichia coli* and used to immunise mice to enable evaluation of immunogenicity and protection based on analysis of antisera by flow cytometry, and in serum bactericidal activity (SBA) and opsonophagocytosis assays or in animal models of infection. Over the past decade, the methodologies used in Reverse Vaccinology projects have greatly progressed beyond the original analysis of a single meningococcal genome (3). Improved sequencing technologies have greatly increased the number of genomes available and Reverse Vaccinology has evolved to include analysis of several genomes within a species (e.g. pan-genomic analysis of Group B Streptococcus (GBS)) (4), between closely related species (e.g. comparative-genomic analysis of GBS, Group A Streptococcus (GAS) and *Streptococcus pneumoniae*) or between pathogenic and commensal members of the same species (e.g. comparative/subtractive-genomic of *E. coli*)(5, 6). Furthermore, high throughput proteomic analyses are now possible, as proteins and peptides can be easily identified using data obtained by mass spectrometry to search genome databases, and can provide important information for vaccine development (e.g. analysis of the GAS surface proteome)(7, 8). Protein microarrays can also be used to identify proteins expressed during host infection (9). As other

innovative fields of research have emerged, such as immunoproteomics, structural biology, and systems biology, they have also been applied to vaccine development, and are helping researchers to overcome the limits of conventional approaches in the discovery and development of novel vaccines, especially against infectious diseases caused by diverse and variable microorganisms (10-12).

In this review, we summarise the progress of over a decade of genome and proteome-based vaccinology, focusing on several projects that have used Reverse Vaccinology as the main approach for antigen discovery (see Fig. 2), including vaccines against *N. meningitidis* serogroup B, for which an application has been submitted for licensure in Europe, and vaccines against GBS, GAS*, S. pneumoniae* and *E. coli* which are being analysed in preclinical studies.

Serogroup B *Neisseria meningitidis* - classic one-genome Reverse Vaccinology

The development of a serogroup B *Neisseria meningitidis* (MenB) vaccine represents the first example of the successful application of Reverse Vaccinology (3, 13-15). *N. meningitidis* is a major cause of septicaemia and meningitis leading to high mortality and morbidity worldwide, and MenB is the predominant serogroup responsible for meningococcal disease in most of the developed world (reviewed in 16). Conventional MenB vaccine approaches have been proven unsuccessful, largely due to the fact that the serogroup B capsular polysaccharide is similar to host cell components and is poorly immunogenic, as well as the high degree of variability of meningococcal major outer membrane proteins (17).

MenB Reverse Vaccinology became possible with the sequencing of the first genome of a serogroup B *N. meningitidis* strain (strain MC58) in 2000, as a result of the synergy of 3 groups, the J. Craig Venter Institute (formerly TIGR) USA, Novartis Vaccines and Diagnostics (formerly Chiron Corporation) Italy, and Oxford University UK (18). With the unravelling of the genome code of this pathogen, researchers were able to select nearly 600 potential vaccine candidates by *in silico* analysis. Following this, 350 recombinant proteins were successfully expressed in *E. coli* and used to immunise mice, from which 91 novel proteins were confirmed to be surface exposed by flow cytometry analysis, and 28 proteins were found to induce bactericidal antibodies (3). After further analysis, a final 4-component MenB (4CMenB) vaccine was formulated, which consists of 3 recombinant protein antigens (NHBA-GNA1030 and GNA2091-fHbp fusion proteins and NadA) and outer membrane vesicles (OMV) from the strain used for the New Zealand OMV vaccine (contains PorA 1.4) (13). Using the meningococcal genome, novel vaccine candidates were discovered that have since been identified as important virulence factors. Factor H binding protein (fHbp) binds factor H (fH), a key inhibitor of the complement alternative pathway, and enables the meningococcus to evade killing by

the innate immune system (19, 20). Neisserial heparin binding antigen (NHBA) also plays a role in serum resistance (21), and the Neisserial adhesin A (NadA) is important for bacterial-host cell interactions (22, 23).

Molecular epidemiological investigation revealed that the main antigens fHbp, NHBA and NadA, can be classified into 3, 14 and 5 variant groups, respectively (24). Furthermore, this analysis revealed that prediction of vaccine coverage was not possible on the basis of traditional meningococcal typing systems (24), such as multi-locus sequence typing (MLST) or capsule-based serogrouping (16, 25). In the serum bactericidal activity (SBA) assay, which is the correlate of protection used for meningococcal vaccines (26), cross-reactivity exists between NHBA variants (21, 27) and the 3 main NadA variants identified in hypervirulent strains (28), however little cross protection is seen between fHbp variant 1 (present in the 4CMenB vaccine) and the other two fHbp variant groups (29, 30). Furthermore, the level of expression of these antigens varies between strains (21, 29, 31). At this point, it became apparent that a new typing system was needed to predict the potential strain coverage of the 4CMenB vaccine, and the meningococcal antigen typing system (MATS) was developed. MATS is a vaccine antigen-specific sandwich ELISA that measures the amount of each antigen expressed by a strain and its immunologic cross-reactivity with the antigen present in the 4CMenB vaccine (32). MATS is correlated with killing of strains in the SBA assay, and MATS analysis of large strain panels from many countries predicted that adults and infants immunized with the 4CMenB vaccine would be protected from disease caused by 86% and 77% of the global MenB strains, respectively (32-34).

Clinical trial data from studies with more than 1,800 infants showed that 4CMenB (commercial name Bexsero) is safe and induces a robust immune response. Protective immune responses (human SBA titres ≥1:5) were seen against reference MenB strains containing the vaccine-antigens; at least 99% of participants produced antibodies that killed strains 44/76 (specific for fHbp) and 5/99 (specific for NadA), and 87% produced antibodies that killed strain NZ98/254 (specific for OMV PorA). MATS data together with a comprehensive clinical program with 4CMenB has served as the basis for submission for regulatory approval to the European Medicines Agency (14, 15, 35-39).

Streptococcus agalactiae - Pan-genome Reverse Vaccinology

Multigenome Reverse Vaccinology was applied to the Gram-positive pathogen *Streptococcus agalactiae* (group B streptococcus (GBS)), based on analysis of 8 GBS genomes to address the sequence diversity of this species (4, 40). GBS transmission mostly occurs from mother to newborn during delivery, causing sepsis, pneumonia and meningitis in the first month of life. GBS also causes severe invasive infections in the elderly, in pregnant women, and patients with underlying diseases

(reviewed in 41). There are 9 GBS serotypes, distinguished by their capsular polysaccharide (CPS), and serotypes Ia, Ib, II, III, and V account for the majority of disease in the USA (42). The risk of neonatal infection is inversely proportional to the maternal amounts of specific antibodies to CPS (43, 44) and CPS-conjugate vaccines are safe and immunogenic in clinical trials (45, 46). However, there are a number of non-typeable GBS isolates that would not be protected by these vaccines (46, 47), and efforts to develop vaccines based on conserved protein antigens are in progress (41).

Due to the high level of diversity between GBS strains, Reverse Vaccinology was applied to several genomes to enable the total gene repertoire or 'pan-genome' of the species to be screened for vaccine candidates. Analysis of 8 GBS genomes revealed a core-genome of approximately 1806 genes present in all strains, plus a dispensable-genome of 907 genes that are absent in one or more strains (40). *In silico* screening identified 589 predicted surface-exposed proteins, of which 396 were core genes and 193 were dispensable genes (4). Of these, 312 were successfully expressed in *E. coli* and evaluated for their ability to mediate protection in a mouse maternal immunization–neonatal pup challenge model. Four proteins were found to be protective: GBS322 from the core genome, as well as GBS67, GBS80 and GBS104, which are present in the dispensable genome and would not have been identified if only a single genome had been screened, however, in combination they proved to be protective against a large panel of strains (4).

Interestingly, three of the newly identified protective GBS antigens were seen to assemble into pili, which are long filamentous structures that extend from the bacterial surface and play an important role in virulence of many pathogens (48). Sequence analysis of a large panel of GBS isolates revealed the presence of 3 pilus islands, PI-1, PI-2a and PI-2b, and all strains characterized to date have at least 1, but often 2, of the islands (48-50). Each island encodes a pilus composed of three structural proteins containing a LPXTG motif, the backbone protein (BP) and two ancillary proteins (AP1 and AP2), of which BP and AP1 induce protective antibodies (4, 49, 50). Moreover, DNA sequence analysis has shown that the three subunits in strains carrying the same island are highly conserved and cross-protective, with the exception of BP-2a that has 6 non cross-protective variants (50). Based on these data, a vaccine containing 1 component of each pilus (BP-1, BP-2b and AP-2a) was tested and shown to elicit opsonophagocytic antibodies in an *in vitro* opsonophagocytosis assay and provide a high level of protection against a large panel of virulent strains from key serotypes in the maternal immunization/neonatal pup challenge model (50).

***Streptococcus pyogenes* – Comparative genomics and proteomic based Reverse Vaccinology**

Streptococcus pyogenes (Group A Streptococcus, GAS) is a Gram-positive pathogen that causes

disease ranging from uncomplicated pharyngitis to life-threatening necrotizing fasciitis, toxic shock, and rheumatic fever. Despite significant attempts, there is currently no vaccine available against GAS (51, 52).

Using comparative genomics and bioinformatics, five GAS genomes were analysed for typical pilus regions, which were identified by the presence of genes encoding proteins containing the cell-wall-attachment motif LPXTG that were also closely linked with sortase enzymes. Three to five genes coding for LPXTG-containing proteins were identified in the various genomes and a combination of 3 recombinant pilus proteins (Cpa, M1_128 and M1_130) was shown to confer protection in mice against mucosal challenge with virulent GAS strains (53). GAS pili are encoded by nine different gene clusters and are highly variable, however it was predicted that a vaccine combining 12 backbone variants could provide protection against over 90% of circulating GAS strains (54).

An innovative proteomic-based approach has also been applied to GAS to specifically isolate and identify surface proteins, termed the bacterial surface proteome or 'surfome'. Proteolytic enzymes were used to "shave" the bacterial surface, and the peptides released were analysed by mass spectrometry and identified using the available GAS genome sequences (8). 72 surface-exposed proteins were identified in strain M1_SF370, the majority of which were predicated by PSORT to be surface-associated and which included most of the protective GAS antigens described to date. Additional analysis of the highly virulent strain M23_DSM2071, led to the identification of 17 surface-exposed proteins, 14 of which were expressed in *E. coli* and tested to determine their ability to mediate protection in a mouse model. This approach led to the identification of one new protective antigen (Spy0416), which is highly conserved and a major cell surface component of >70% of the GAS strains (8).

***Streptococcus pneumoniae* – Comparative genomics and Reverse Vaccinology**

Several classical and genome-based approaches have been applied to *Streptococcus pneumoniae* for identification of virulence factors and vaccines antigens (55-57). *S. pneumoniae* is a Gram-positive bacterium that causes significant human disease, including sepsis, meningitis, pneumonia, otitis media and sinusitis, and accounts for approximately 11% of mortality worldwide in children under five years old (58). *S. pneumoniae* can be classified into more than 90 capsular serotypes, and multivalent capsule polysaccharide (Pneumovax 23) and polysaccharide protein conjugate vaccines (PCV; 7 or 13-valent Prevnar, 10-valent Synflorix) (59) are currently available. However, many pneumococcal proteins are being evaluated for use in a universal serotype-independent vaccine, due to increasing antibiotic resistance and limitations of polysaccharide vaccines, including poor immunogenicity of unconjugated

vaccines in infants, the complexity and cost of PCVs, the variable regional distribution of serotypes and the restricted coverage of available vaccines (59), as well as the occurrence of serotype replacement after vaccination (60).

The availability of multiple pneumococcal genome sequences combined with the increased understanding of pili in GBS and GAS led to the discovery of pili in *S. pneumoniae* and the investigation of pneumococcal pilus proteins as vaccine antigens (61, 62). Pneumococcal pili are involved in adherence and virulence as well as stimulating the host inflammatory response (61-64). Pilus 1 and pilus 2 are antigenically distinct and are encoded by two genomic pilus islets, PI-1 and PI-2, that are present in ~30% and 16% of pneumococcal clinical isolates, respectively (63, 65). Pilus 1 is encoded by the *rlrA* islet, and is composed of three subunits, a backbone protein (RrgB) and two ancillary proteins (RrgA and RrgC) (55, 61, 62, 66), while Pilus 2 consists of the backbone protein PitB (63). Active and passive immunization with recombinant pilus subunits RrgA, RrgB and RrgC showed protection against lethal challenge with strain TIGR4 (67). Further molecular epidemiology investigations have shown the presence of two variants of RrgA, both of which elicit cross-protection (68).

Pathogenic *Escherichia coli* – Subtractive genomics and Reverse Vaccinology

Pathogenic *E. coli* are responsible for a range of infections, and are classified as either extra-intestinal (ExPEC) or intestinal (InPEC) pathogenic *E. coli*. ExPEC strains are responsible for the majority of urinary tract infections (UPEC strains) and are the second most common cause of neonatal meningitis and sepsis (NMEC strains), whereas intestinal strains are the major cause of travellers' diarrhoea and can cause more serious haemolytic uremic syndrome (reviewed in 69).

Comparative genome analysis and subtractive Reverse Vaccinology, using the genomes of 2 NMEC, 5 UPEC strains, 1 pathogenic avian and 1 non-pathogenic *E. coli* K12 strain, led to the identification of 230 antigens present in ExPEC but absent (or poorly conserved) in non-pathogenic strains (5). This approach was aided by the earlier characterisation of outer membrane and secreted proteins by proteomic analysis of OMVs from pathogenic *E. coli* strains (70-72). Further analysis of candidates led to the identification of nine potential vaccine antigens that, following immunization with the recombinant protein, were able to induce protection in a mouse challenge sepsis model. For the most protective antigen, ECOK1_3385, both active immunization with the recombinant protein, or passive immunization with sera raised against this antigen, provided nearly complete protection from bacteraemia and mortality in the mouse model of sepsis. ECOK1_3385 is widely distributed among *E. coli* pathotypes with amino acid sequence identity ranging from 86% to 100%. Some of the novel

protective antigens are also present in intestinal pathogenic *E. coli*, suggesting the possibility of a broadly protective vaccine against pathogenic *E. coli* (5).

Conclusions and Perspectives

Significant progress has been made in the field of vaccinology during the era of genomics, and vaccines are set to have an increasing impact on public health in the coming years. A decade of genome and proteome based Reverse Vaccinology, targeting bacterial pathogens including MenB, GBS, GAS, *S. pneumoniae* and pathogenic *E. coli*, has resulted in the discovery of numerous candidate vaccine antigens that are undergoing preclinical development, in addition to the clinical development of the 4CMenB vaccine (Figure 2). The continual advancements in genome sequencing technologies and the approaches used to screen the genome and proteome of a bacterial pathogen have greatly improved the efficiency and time needed for antigen discovery. These advancements have also improved antigen discovery and characterization for eukaryotic pathogens, which often encode more than 10,000 genes (73). Genomics is also increasingly used to characterize variability between viral pathogens, which typically contain less than 10 genes (74).

Reverse vaccinology is now being applied to many bacterial, viral and eukaryotic pathogens and has been successful in all cases in providing novel antigens for the design of new vaccines (10-12). Furthermore, the ability to improve candidate antigens by rational design and structural vaccinology can provide increased protection against antigenically variable pathogens. For example, a structural vaccinology approach has been used to address the sequence variability of the MenB vaccine antigen fHbp, where a single antigen has been designed, which could elicit broadly protective immunity against strains expressing any fHbp variant, by grafting multiple immunodominant epitopes onto a single fHbp variant 1 molecule (75). Similarly, a synthetic GBS pilus antigen has been constructed that comprises the protective domain of each of six variants of the backbone protein, BP-2a, to facilitate the development of a broadly protective pilus-based vaccine against GBS (76). In addition, formulation of selected antigens with appropriate adjuvants may also expand the protection provided against variable pathogens. In the coming years, genomics will continue to contribute to vaccine development as human genomics and systems biology address questions aimed at understanding the types of immune responses needed to control certain diseases, the ways in which adjuvants direct the immune system and the different outcome of vaccination in the population.

Acknowledgements

We would like to thank Giorgio Corsi for artwork.

References

1. Plotkin SA. Vaccines: the fourth century. Clin Vaccine Immunol. 2009 Dec;16(12):1709-19.
2. Rappuoli R. Reverse vaccinology. Curr Opin Microbiol. 2000 Oct;3(5):445-50.
3. Pizza M, Scarlato V, Masignani V, Giuliani MM, Arico B, Comanducci M, et al. Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. Science. 2000 Mar 10;287(5459):1816-20.
4. Maione D, Margarit I, Rinaudo CD, Masignani V, Mora M, Scarselli M, et al. Identification of a universal Group B streptococcus vaccine by multiple genome screen. Science. 2005 Jul 1;309(5731):148-50.
5. Moriel DG, Bertoldi I, Spagnuolo A, Marchi S, Rosini R, Nesta B, et al. Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic Escherichia coli. Proc Natl Acad Sci U S A. 2010 May 18;107(20):9072-7.
6. Rasko DA, Rosovitz MJ, Myers GS, Mongodin EF, Fricke WF, Gajer P, et al. The pangenome structure of Escherichia coli: comparative genomic analysis of E. coli commensal and pathogenic isolates. J Bacteriol. 2008 Oct;190(20):6881-93.
7. Walters MS, Mobley HL. Bacterial proteomics and identification of potential vaccine targets. Expert Rev Proteomics. 2010 Apr;7(2):181-4.
8. Rodriguez-Ortega MJ, Norais N, Bensi G, Liberatori S, Capo S, Mora M, et al. Characterization and identification of vaccine candidate proteins through analysis of the group A Streptococcus surface proteome. Nature biotechnology. 2006 Feb;24(2):191-7.
9. Grandi G. Genomics and proteomics in reverse vaccines. Methods Biochem Anal. 2006;49:379-93.
10. Bagnoli F, Baudner B, Mishra RP, Bartolini E, Fiaschi L, Mariotti P, et al. Designing the next generation of vaccines for global public health. OMICS. 2011 Sep;15(9):545-66.
11. Rinaudo CD, Telford JL, Rappuoli R, Seib KL. Vaccinology in the genome era. J Clin Invest. 2009 Sep;119(9):2515-25.
12. Buonaguro L, Pulendran B. Immunogenomics and systems biology of vaccines. Immunol Rev. 2011 Jan;239(1):197-208.
13. Giuliani MM, Adu-Bobie J, Comanducci M, Arico B, Savino S, Santini L, et al. A universal vaccine for serogroup B meningococcus. Proc Natl Acad Sci U S A. 2006 Jul 18;103(29):10834-9.

14. Toneatto D, Ismaili S, Ypma E, Vienken K, Oster P, Dull P. The first use of an investigational multicomponent meningococcal serogroup B vaccine (4CMenB) in humans. Hum Vaccin. 2011 Jun 1;7(6):646-53.

15. Toneatto D, Oster P, Deboer AC, Emerson A, Santos GF, Ypma E, et al. Early clinical experience with a candidate meningococcal B recombinant vaccine (rMenB) in healthy adults. Hum Vaccin. 2011 Jul 1;7(7):781-91.

16. Harrison LH, Trotter CL, Ramsay ME. Global epidemiology of meningococcal disease. Vaccine. 2009 Jun 24;27 Suppl 2:B51-63.

17. Tan LK, Carlone GM, Borrow R. Advances in the development of vaccines against Neisseria meningitidis. N Engl J Med. 2010 Apr 22;362(16):1511-20.

18. Tettelin H, Saunders NJ, Heidelberg J, Jeffries AC, Nelson KE, Eisen JA, et al. Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. Science. 2000 Mar 10;287(5459):1809-15.

19. Madico G, Welsch JA, Lewis LA, McNaughton A, Perlman DH, Costello CE, et al. The meningococcal vaccine candidate GNA1870 binds the complement regulatory protein factor H and enhances serum resistance. J Immunol. 2006 Jul 1;177(1):501-10.

20. Schneider MC, Exley RM, Chan H, Feavers I, Kang YH, Sim RB, et al. Functional significance of factor H binding to Neisseria meningitidis. J Immunol. 2006 Jun 15;176(12):7566-75.

21. Serruto D, Spadafina T, Ciucchi L, Lewis LA, Ram S, Tontini M, et al. Neisseria meningitidis GNA2132, a heparin-binding protein that induces protective immunity in humans. Proc Natl Acad Sci U S A. 2010 Feb 23;107(8):3770-5.

22. Capecchi B, Adu-Bobie J, Di Marcello F, Ciucchi L, Masignani V, Taddei A, et al. Neisseria meningitidis NadA is a new invasin which promotes bacterial adhesion to and penetration into human epithelial cells. Mol Microbiol. 2005 Feb;55(3):687-98.

23. Comanducci M, Bambini S, Brunelli B, Adu-Bobie J, Arico B, Capecchi B, et al. NadA, a novel vaccine candidate of Neisseria meningitidis. J Exp Med. 2002 Jun 3;195(11):1445-54.

24. Bambini S, Muzzi A, Olcen P, Rappuoli R, Pizza M, Comanducci M. Distribution and genetic variability of three vaccine components in a panel of strains representative of the diversity of serogroup B meningococcus. Vaccine. 2009;27(21):2794-803.

25. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci U S A. 1998 Mar 17;95(6):3140-5.

26. Borrow R, Carlone GM, Rosenstein N, Blake M, Feavers I, Martin D, et al. Neisseria meningitidis group B correlates of protection and assay standardization--international meeting report Emory University, Atlanta, Georgia, United States, 16-17 March 2005. Vaccine. 2006 Jun 12;24(24):5093-107.

27. Serruto D, Biolchi A, Santini L, Brunelli B, Del Tordello E, Bambini S, et al., editors. Evaluation of the immunological properties of the Neisserial heparin binding antigen (NHBA). The 11th meeting of The European Meningococcal Disease Society (EMGM); 2011 18-28 May 2011; Ljubljana.

28. Comanducci M, Bambini S, Caugant DA, Mora M, Brunelli B, Capecchi B, et al. NadA diversity and carriage in Neisseria meningitidis. Infect Immun. 2004 Jul;72(7):4217-23.

29. Masignani V, Comanducci M, Giuliani MM, Bambini S, Adu-Bobie J, Arico B, et al. Vaccination against Neisseria meningitidis using three variants of the lipoprotein GNA1870. J Exp Med. 2003 Mar 17;197(6):789-99.

30. Brunelli B, Del Tordello E, Palumbo E, Biolchi A, Bambini S, Comanducci M, et al. Influence of sequence variability on bactericidal activity sera induced by Factor H binding protein variant 1.1. Vaccine. 2011 Jan 29;29(5):1072-81.

31. Metruccio MM, Pigozzi E, Roncarati D, Berlanda Scorza F, Norais N, Hill SA, et al. A novel phase variation mechanism in the meningococcus driven by a ligand-responsive repressor and differential spacing of distal promoter elements. PLoS Pathog. 2009 Dec;5(12):e1000710.

32. Donnelly J, Medini D, Boccadifuoco G, Biolchi A, Ward J, Frasch C, et al. Qualitative and quantitative assessment of meningococcal antigens to evaluate the potential strain coverage of protein-based vaccines. Proc Natl Acad Sci U S A. 2010 Nov 9;107(45):19490-5.

33. Rappuoli R. The challenge of developing universal vaccines. F1000 Med Rep. 2011;3:16.

34. Donnelly J, Medini D, Giuliani MM, Boccadifuoco G, Stella M, Frosi G, et al., editors. Estimating the potential strain coverage in Europe of a multilcomponent vaccine targeting serogroup B meningococci. The 11th meeting of The European Meningococcal Disease Society (EMGM); 2011 18-28 May 2011; Ljubljana.

35. Vesikari T, Prymula R, Liese J, Kollaritsch H, Bona G, Kimura A, et al., editors. Booster dose at 12 months of an investigational meningococcal serogroup B vaccine (4CMenB) in toddlers previously primed at 2,4,6 months 29th Annual Meeting of the European Society for Paediatric infectious Deseases; 2011 June 7-11,2011 The Hague.

36. Santolaya ME, O´Ryan ML, Valenzuela MT, Prado V, Vergara R, Muñoz A, et al., editors. Immunogenicity and tolerability of an investigational multicomponent meningococcal serogroupB

(4CMenB) vaccine in healthy adolescents 29th Annual Meeting of the European Society for Paediatric infectious Deseases; 2011 June 7-11, 2011; The Hague.

37. Prymula R, Vesikari T, Esposito S, Kohl I, Ypma E, Toneatto D, et al., editors. Catch-up vaccination of healthy toddlers with an investigational multicomponent meningococcal serogroup B vaccine(4BmenB)-exploration of a two-dose schedule 29th Annual Meeting of the European Society for Paediatric infectious Deseases; 2011 June 7-11, 2011; The Hague.

38. Gossger N, Snape M, Finn A, Heath P, Collinson A, Bona G, et al., editors. Immunogenicity of an investigational multicomponent meningococcal serogroup B vaccine (4CMenB) administered with or without routine infant vaccinations in different schedules. 29th Annual Meeting of the European Society for Paediatric infectious Deseases; 2011 June 7-11, 2011; The Hague.

39. Beeretz I, Snape M, Finn A, Heath P, Collinson A, Bona G, et al., editors. Reactogenicity and safety of multicomponent meningococcal serogroup B vaccine (4CMenB) administered with or without routine infant vaccinations in different schedules 29th Annual Meeting of the European Society for Paediatric infectious Deseases; 2011 June 7-11, 2011; The Hague.

40. Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, et al. Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pan-genome". Proc Natl Acad Sci U S A. 2005 Sep 27;102(39):13950-5.

41. Johri AK, Paoletti LC, Glaser P, Dua M, Sharma PK, Grandi G, et al. Group B Streptococcus: global incidence and vaccine development. Nat Rev Microbiol. 2006 Dec;4(12):932-42.

42. Phares CR, Lynfield R, Farley MM, Mohle-Boetani J, Harrison LH, Petit S, et al. Epidemiology of invasive group B streptococcal disease in the United States, 1999-2005. JAMA. 2008 May 7;299(17):2056-65.

43. Lin FY, Weisman LE, Azimi PH, Philips JB, 3rd, Clark P, Regan J, et al. Level of maternal IgG anti-group B streptococcus type III antibody correlated with protection of neonates against early-onset disease caused by this pathogen. J Infect Dis. 2004 Sep 1;190(5):928-34.

44. Baker CJ, Edwards MS, Kasper DL. Role of antibody to native type III polysaccharide of group B Streptococcus in infant infection. Pediatrics. 1981 Oct;68(4):544-9.

45. Paoletti LC, Madoff LC. Vaccines to prevent neonatal GBS infection. Semin Neonatol. 2002 Aug;7(4):315-23.

46. Baker CJ, Edwards MS. Group B streptococcal conjugate vaccines. Arch Dis Child. 2003 May;88(5):375-8.

47. Ippolito DL, James WA, Tinnemore D, Huang RR, Dehart MJ, Williams J, et al. Group B streptococcus serotype prevalence in reproductive-age women at a tertiary care military medical center relative to global serotype distribution. BMC Infect Dis. 2010;10:336.

48. Lauer P, Rinaudo CD, Soriani M, Margarit I, Maione D, Rosini R, et al. Genome analysis reveals pili in Group B Streptococcus. Science. 2005 Jul 1;309(5731):105.

49. Rosini R, Rinaudo CD, Soriani M, Lauer P, Mora M, Maione D, et al. Identification of novel genomic islands coding for antigenic pilus-like structures in Streptococcus agalactiae. Mol Microbiol. 2006 Jul;61(1):126-41.

50. Margarit I, Rinaudo CD, Galeotti CL, Maione D, Ghezzo C, Buttazzoni E, et al. Preventing bacterial infections with pilus-based vaccines: the group B streptococcus paradigm. J Infect Dis. 2009 Jan 1;199(1):108-15.

51. Pandey M, Sekuloski S, Batzloff MR. Novel strategies for controlling Streptococcus pyogenes infection and associated diseases: from potential peptide vaccines to antibody immunotherapy. Immunol Cell Biol. 2009 Jul;87(5):391-9.

52. Bisno AL, Rubin FA, Cleary PP, Dale JB. Prospects for a group A streptococcal vaccine: rationale, feasibility, and obstacles--report of a National Institute of Allergy and Infectious Diseases workshop. Clin Infect Dis. 2005 Oct 15;41(8):1150-6.

53. Mora M, Bensi G, Capo S, Falugi F, Zingaretti C, Manetti AG, et al. Group A Streptococcus produce pilus-like structures containing protective antigens and Lancefield T antigens. Proc Natl Acad Sci U S A. 2005 Oct 25;102(43):15641-6.

54. Falugi F, Zingaretti C, Pinto V, Mariani M, Amodeo L, Manetti AG, et al. Sequence Variation in Group A Streptococcus Pili and Association of Pilus Backbone Types with Lancefield T Serotypes. J Infect Dis. 2008 Oct 17;198(12):1834-41.

55. Hava DL, Camilli A. Large-scale identification of serotype 4 Streptococcus pneumoniae virulence factors. Mol Microbiol. 2002 Sep;45(5):1389-406.

56. Paton JC, Giammarinaro P. Genome-based analysis of pneumococcal virulence factors: the quest for novel vaccine antigens and drug targets. Trends Microbiol. 2001 Nov;9(11):515-8.

57. Wizemann TM, Heinrichs JH, Adamou JE, Erwin AL, Kunsch C, Choi GH, et al. Use of a whole genome approach to identify vaccine molecules affording protection against Streptococcus pneumoniae infection. Infect Immun. 2001 Mar;69(3):1593-8.

58. O'Brien KL, Wolfson LJ, Watt JP, Henkle E, Deloria-Knoll M, McCall N, et al. Burden of disease caused by Streptococcus pneumoniae in children younger than 5 years: global estimates. Lancet. 2009 Sep 12;374(9693):893-902.

59. Moffitt KL, Malley R. Next generation pneumococcal vaccines. Curr Opin Immunol. 2011 Jun;23(3):407-13.

60. Hsu HE, Shutt KA, Moore MR, Beall BW, Bennett NM, Craig AS, et al. Effect of pneumococcal conjugate vaccine on pneumococcal meningitis. N Engl J Med. 2009 Jan 15;360(3):244-56.

61. Barocchi MA, Ries J, Zogaj X, Hemsley C, Albiger B, Kanth A, et al. A pneumococcal pilus influences virulence and host inflammatory responses. Proc Natl Acad Sci U S A. 2006 Feb 21;103(8):2857-62.

62. LeMieux J, Hava DL, Basset A, Camilli A. RrgA and RrgB are components of a multisubunit pilus encoded by the Streptococcus pneumoniae rlrA pathogenicity islet. Infect Immun. 2006 Apr;74(4):2453-6.

63. Bagnoli F, Moschioni M, Donati C, Dimitrovska V, Ferlenghi I, Facciotti C, et al. A second pilus type in Streptococcus pneumoniae is prevalent in emerging serotypes and mediates adhesion to host cells. J Bacteriol. 2008 Aug;190(15):5480-92.

64. Telford JL, Barocchi MA, Margarit I, Rappuoli R, Grandi G. Pili in gram-positive pathogens. Nat Rev Microbiol. 2006 Jul;4(7):509-19.

65. Moschioni M, Donati C, Muzzi A, Masignani V, Censini S, Hanage WP, et al. Streptococcus pneumoniae contains 3 rlrA pilus variants that are clonally related. J Infect Dis. 2008 Mar 15;197(6):888-96.

66. Hilleringmann M, Giusti F, Baudner BC, Masignani V, Covacci A, Rappuoli R, et al. Pneumococcal pili are composed of protofilaments exposing adhesive clusters of Rrg A. PLoS Pathog. 2008 Mar;4(3):e1000026.

67. Gianfaldoni C, Censini S, Hilleringmann M, Moschioni M, Facciotti C, Pansegrau W, et al. Streptococcus pneumoniae pilus subunits protect mice against lethal challenge. Infect Immun. 2007 Feb;75(2):1059-62.

68. Moschioni M, Emolo C, Biagini M, Maccari S, Pansegrau W, Donati C, et al. The two variants of the Streptococcus pneumoniae pilus 1 RrgA adhesin retain the same function and elicit cross-protection in vivo. Infect Immun. 2010 Dec;78(12):5033-42.

69. Kaper JB, Nataro JP, Mobley HL. Pathogenic Escherichia coli. Nat Rev Microbiol. 2004 Feb;2(2):123-40.

70. Berlanda Scorza F, Doro F, Rodriguez-Ortega MJ, Stella M, Liberatori S, Taddei AR, et al. Proteomics characterization of outer membrane vesicles from the extraintestinal pathogenic Escherichia coli DeltatolR IHE3034 mutant. Mol Cell Proteomics. 2008 Mar;7(3):473-85.

71. Alteri CJ, Mobley HL. Quantitative profile of the uropathogenic Escherichia coli outer membrane proteome during growth in human urine. Infect Immun. 2007 Jun;75(6):2679-88.

72. Molloy MP, Herbert BR, Slade MB, Rabilloud T, Nouwens AS, Williams KL, et al. Proteomic analysis of the Escherichia coli outer membrane. Eur J Biochem. 2000 May;267(10):2871-81.

73. Aurrecoechea C, Heiges M, Wang H, Wang Z, Fischer S, Rhodes P, et al. ApiDB: integrated resources for the apicomplexan bioinformatics resource center. Nucleic Acids Res. 2007 Jan;35(Database issue):D427-30.

74. Hirahata M, Abe T, Tanaka N, Kuwana Y, Shigemoto Y, Miyazaki S, et al. Genome Information Broker for Viruses (GIB-V): database for comparative analysis of virus genomes. Nucleic Acids Res. 2007 Jan;35(Database issue):D339-42.

75. Scarselli M, Arico B, Brunelli B, Savino S, Di Marcello F, Palumbo E, et al. Rational design of a meningococcal antigen inducing broad protective immunity. Sci Transl Med. 2011 Jul 13;3(91):91ra62.

76. Nuccitelli A, Cozzi R, Gourlay LJ, Donnarumma D, Necchi F, Norais N, et al. Structure-based approach to rationally design a chimeric protein for an effective vaccine against Group B Streptococcus infections. Proc Natl Acad Sci U S A. 2011 Jun 21;108(25):10278-83.

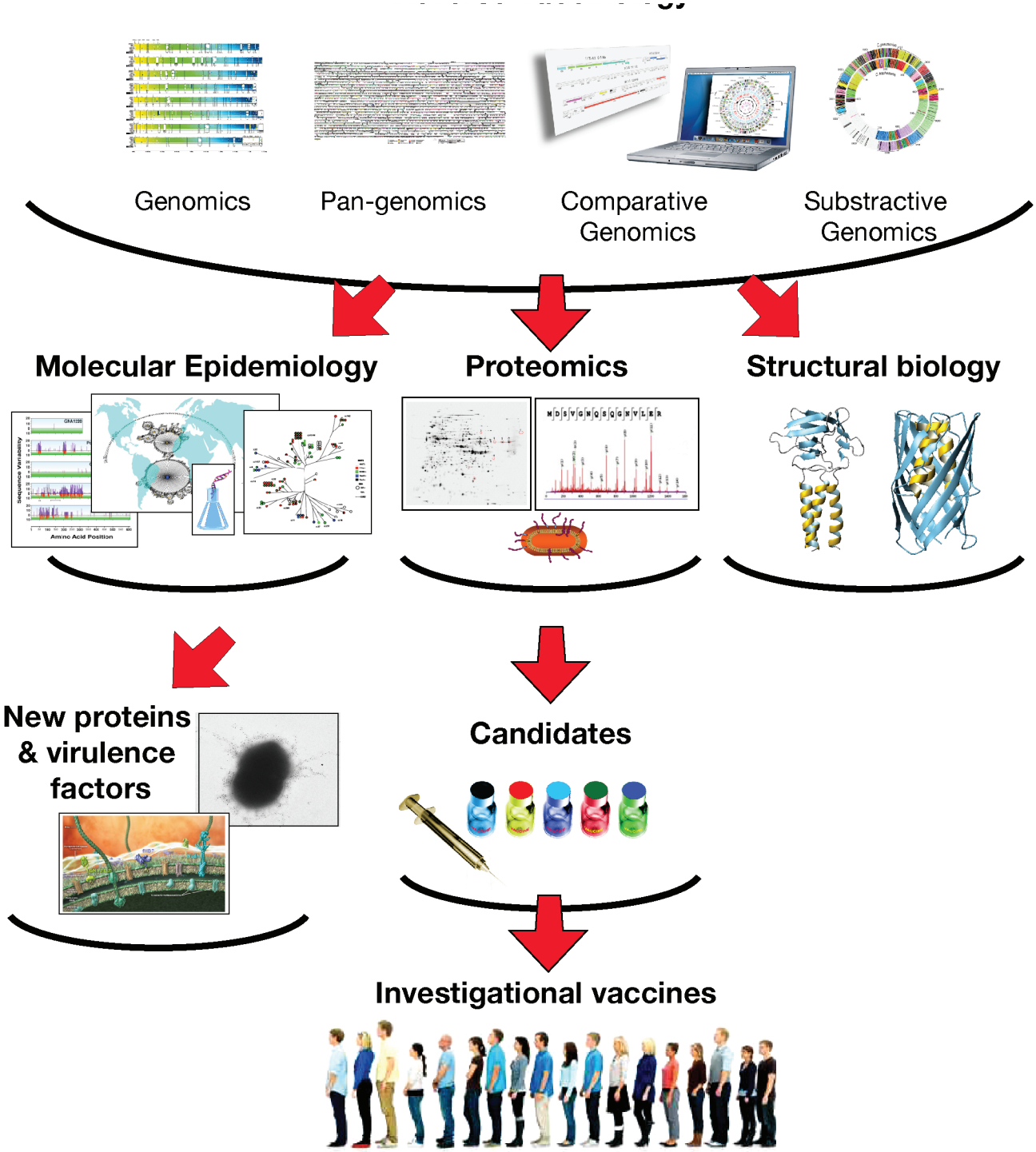

Figure 1. Schematic overview of Reverse Vaccinology. Reverse Vaccinology started with the analysis of a single genome of a Meningococcus serogroup B strain, and pan-genomics, comparative genomics and substractive genomics were subsequently used for various projects, with the incorporation of proteomics, molecular epidemiology and structural biology to further refine the process of antigen discovery. Furthermore, during the process of vaccine development several new virulence factors have been identified, including the *Neisseria meningitidis* factor H binding protein (fHbp) and *Neisserial* heparin binding antigen (NHBA), as well as pili of Group B streptococcus, Group A streptococcus and *Streptococcus pneumoniae*.

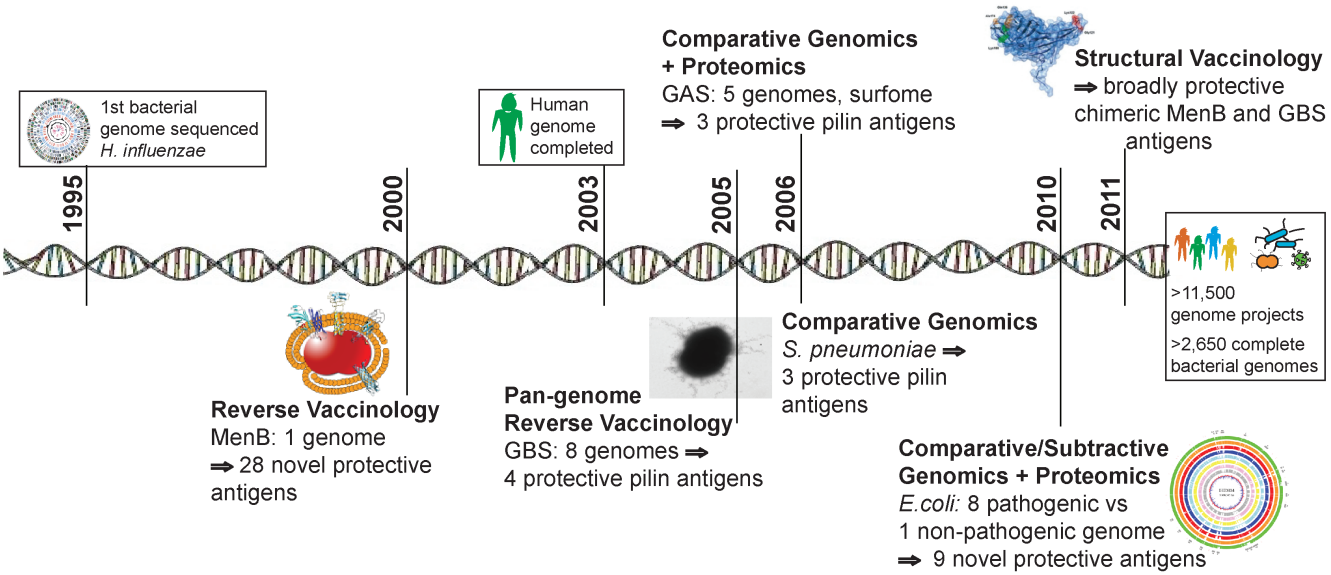

Figure 2. Timeline of a decade of Reverse Vacacinology. Vaccine projects that have used genomic and proteomic based Reverse Vaccinology approaches over the last decade are shown, along with their key outcomes: MenB, *Nesseria meningitis* serogroup B; GBS, Group B Streptococcus; GAS, Group A Streptococcus, *Streptococcus pneumoniae*; and pathogenic *Escherichia coli*. The rapid advancement of genome sequencing is also highlighted, with only 1 bacterial genome available in 1995 and the human genome being completed in 2003, however there are now more than 11,500 genome sequencing projects underway and at least 2,650 bacterial genome sequences available (http://genomesonline.org/cgi-bin/GOLD/bin/gold.cgi).