Impact of Microbial Genomics Approaches for Novel Antibiotic Target

5

Hemant Joshi, Akanksha Verma, and Dharmendra Kumar Soni

Abstract

Infectious diseases are life-threatening and may lead to high mortality and morbidity rates. The existing danger of an increase and spread of multidrug resistance pathogens is a global concern. Therefore, the designing of novel antibiotics and vaccine to control and eliminate the disease is an utmost requirement. Traditional approaches for screening vaccine and drug targets are time-consuming and have been unsuccessful in controlling the spread of infectious diseases due to several reasons such as altered antigenic diversity, altered virulence potential, and antimicrobial resistance in the infectious agent population. To overcome this problem, there has been a paradigm shift from the conventional to microbial genomics approaches, as the availability of complete genome sequence of pathogenic microorganisms and multiple isolates of the same species provides a wealth of information on nearly all the potential drug targets. Microbial genomics approaches open up new avenues to pursuit novel antimicrobial agents that are highly conserved in a range of microbes, essential for the survival of pathogens and absent in humans. In this chapter, we present an overview of the microbial genomics approaches such as pan-genomics, comparative genomics, functional

Hemant Joshi and Akanksha Verma are both considered as first author.

H. Joshi
School of Biotechnology, Jawaharlal Nehru University, New Delhi, India

A. Verma
Department of Botany, MLKPG College, Balrampur, Uttar Pradesh, India

D. K. Soni (✉)
Department of Molecular Biology, Umeå Centre for Microbial Research (UCMR), Umeå University, Umeå, Sweden

Laboratory of Molecular Infection Medicine Sweden (MIMS), Umeå University, Umeå, Sweden
e-mail: dharmendra.soni@umu.se

V. Tripathi et al. (eds.), *Microbial Genomics in Sustainable Agroecosystems*,
https://doi.org/10.1007/978-981-32-9860-6_5

genomics, structural genomics, transcriptomics, and proteomics used in the discovery and development of novel antibiotics.

Keywords
Microbial genomics · Antibiotics and vaccine target · Pan-genomics · Structural genomics · Transcriptomics · Proteomics

5.1 Introduction

Earlier, the infections caused by microbes had been a massive problem, but in the year 1940s, it was resolved by the introduction of antibiotics. With this advancement in treatment protocol, in the 1960s, it was stated that the danger of microbial infection is no more a problem and the microbes could be successfully defeated. However, unfortunately over the last decades, the microbes acquired resistance toward antibiotics leading to a broader health concern. Therefore, nowadays microbial antibiotic resistance is an emergent and hazardous issue worldwide, and this necessitates novel antibiotics to combat microbial infection.

Development of effective antibiotics and vaccines against infectious disease has a major impact on health globally. The increasing antibiotic resistance and varied antigenic diversity among the pathogens are raising severe concern for the future pandemic. A recent report of the Centers for Disease Control and Prevention (CDC) on "antibiotic resistance: a global threat" showed that only in the USA, every year approximately 2 million people are infected by antibiotic-resistant strains, accounting for nearly 23,000 deaths (CDC 2018; https://www.cdc.gov/features/antibiotic-resistance-global/index.html). In this report, the negative impact of antimicrobial resistance on economy was also predicted with an expected loss of around $100 trillion by the year 2050. These estimation prioritize our action toward finding essential targets and mechanisms for the development of novel vaccines and drugs.

Conventional approaches have proven insufficient to study pathogens because of the complex mechanism of pathogenesis, varied antigenic diversity, as well as lack of a suitable animal model of infection. The arrival of the genomic era has a great impact on the development of vaccine and antibiotics. Microbial genomics data from genome, transcriptome, proteome, immunome, or structural genome provides a wealth of information about the different pathogens that seems to be sufficient for rapid development of novel vaccine and therapeutic and to limit the spread of infection. Therefore, the present chapter aims to provide a comprehensive overview of microbial genomics approaches and their significance in the development of novel vaccines and antibiotics.

5.2 Essential Criteria of Vaccines and Therapeutic Targets

The identification of drug and vaccine targets can be achieved by using various approaches such as the comparative and structural genome, transcriptome, proteome, and immunome. These approaches can be applied in several combinations based on nature of the pathogen under study. However, it is necessary to consider the following basic criteria while selecting the potential targets: (i) target should be specific and highly selective against the microbe rather than host and also active against a broad spectrum of pathogens, (ii) target should be essential for the growth and survival of pathogens at the time of infection, (iii) target should be expressed or easily accessible to the host immune system during the course of infection, and (iv) some prior information about the function of target is necessary so that high-throughput assays can be performed. Identification of new potential targets can be initiated by using the criteria mentioned above which would be helpful in finding out the successful targets.

5.3 Microbial Genomics Approaches

Since the completion of the first bacterial genome sequence of *Haemophilus influenzae* in 1995, the idea for the development of vaccine and therapeutic approaches shifted from conventional approaches to microbial genome-based approaches. Several microbial genomics approaches such as genomics, pan-genomics, comparative genomics, functional genomics, structural genomics, transcriptomics, and proteomics have been utilized for this purpose. The schematic representation of several important microbial genomics approaches has been shown in Fig. 5.1. In summary, in silico screening of the entire genome sequence of the pathogen (genomics) provides complete information about the genetic repertoire of antigens and drug targets. Pan-genomics helps in the identification of conserved antigens and thereby in the identification of potential drug targets through the investigation of genetic material from numerous organisms of single species. Next, it is essential to compare the genetic material of pathogenic and nonpathogenic organisms of single species. This is crucial for identifying antigens or targets that are present in pathogenic strains but absent in nonpathogenic strains. Transcriptomics and proteomics aim to recognize the set of RNA transcripts and proteins expressed by an organism under a specified circumstance and in specific cellular location. Further, the analysis of genes and proteins array would help to understand the survival of an organism under a specific condition (functional genomics). Some interesting field of study emphasizes the identification of protein arrays or epitopes that interact with the host immune system and the possible mechanisms of their interaction (immunomics). Analysis of the three-dimensional structure of proteins of an organism and the process of interaction with antibody and therapeutics (structural genomics) can provide a clear idea about the biological phenomena and potentiality of a novel drug. Following this the vaccinomics approach enables the monitoring of the mode of response of the human immune system to a vaccine or drug. Finally, if the identified targets show

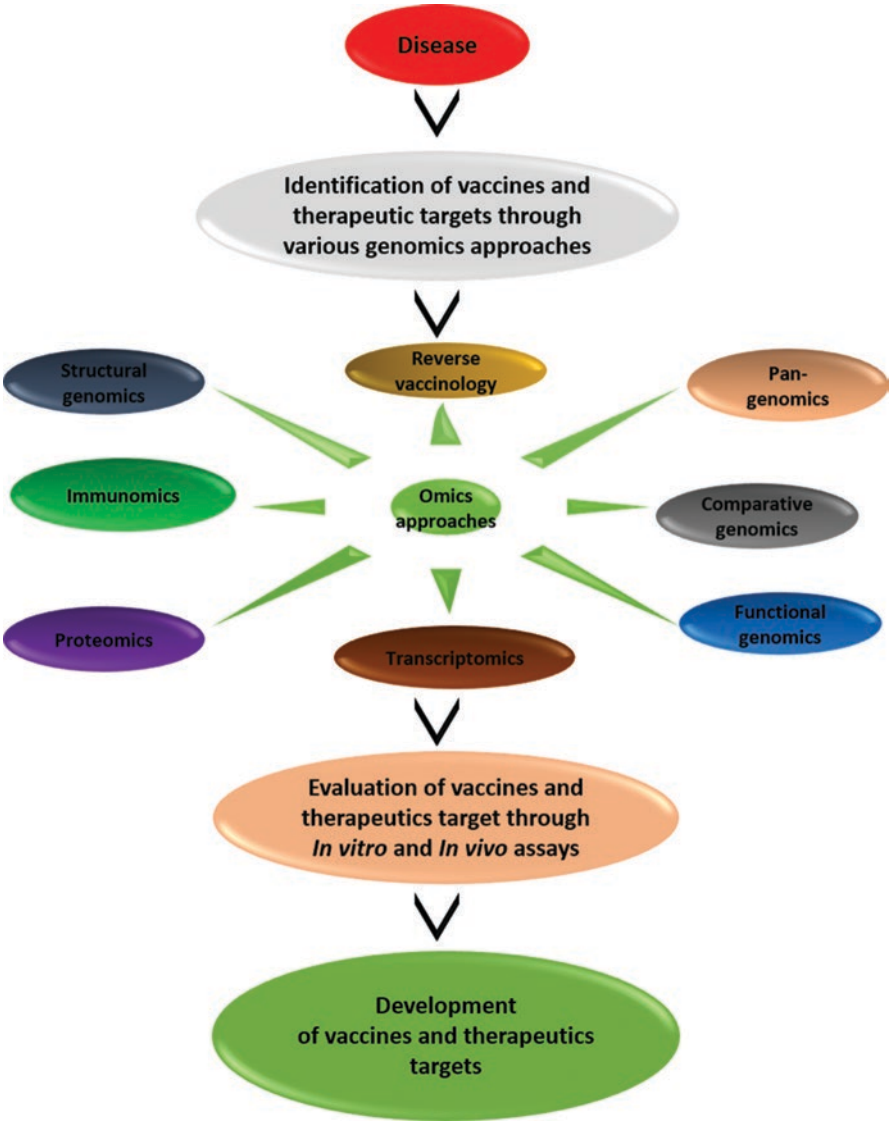

Fig. 5.1 Schematic representation of microbial genomics approaches for the development of vaccines and therapeutics

protection against disease and have low risk vs benefit ratio for humans, they are subjected for clinical trials, and then clinically tested vaccine and therapeutic targets can be licensed for use. In Table 5.1, a brief description of various microbial genomics approaches along with their limitations has been presented.

Here, we are summarizing genomics, pan-genomics, comparative genomics, functional genomics, structural genomics, transcriptomics, and proteomics-based

Table 5.1 Overview of microbial genomics approaches for the development of vaccines and therapeutics

Genomic approaches	Description	Limitations
Reverse vaccinology/ genomics	Identification of surface-exposed proteins as vaccine/therapeutic targets that targets organisms has the potential to be express	Nonprotein antigens such as lipids (glycolipid, phospholipid) and polysaccharides and posttranslational modification (glycosylation, methylation, mannosylation) cannot be identified (Seib et al. 2009)
		Requirement of high-throughput cloning and protein expression
Pan-genomics	Identification of conserved targets through analysis of genetic material of several organisms of single species	Requirement of the genomic sequence of several strains of same species
		Same limitation as described above
Comparative genomics	Identification of genes that are present in pathogenic strain but absent in nonpathogenic strain through analysis of genetic material of different strains of the same species	Same limitations as described for the above two approaches
Transcriptomics	Identification of global changes in gene expression through analysis of RNA transcripts level of an organisms under specified conditions	Large quantity of mRNA is required for in vivo studies
		Difficulty to make probes because bacterial mRNA is unstable
		There is no direct correlation between transcription product (i.e., mRNA) and translational product (i.e., protein)
Functional genomics	Evaluate the function of genes and proteins to identify the genes that are essential for survival and pathogenesis of an organism under specific conditions	Pathogen should be naturally competent; otherwise, it will not be able to accept the transposon (Seib et al. 2009)
Proteomics	Identification of the entire set of proteins expressed by an organism under specified conditions	Proteins should not be in low abundance
		Those proteins, which are expressed in in vivo conditions but not capable to express in in vitro conditions, may not be able to detect
		Low-solubility proteins also may not be able to identify
Immunomics	Evaluate the complete set of proteins that interact with the host immune system to identify the B-cell and T-cell epitopes	Difficult to predict the B-cell antigenic determinant
		Only configurational epitopes can be detected. Chances to detect conformational epitopes is very low (Seib et al. 2009)
		Potential bias against non-displayed sequences
Structural genomics	Identification of the three-dimensional structure of proteins expressed by an organism and how they interact with drugs or antibodies	Inadequate apprehension of antigenic determinants of immunogenicity (Seib et al. 2009)
		Poor understanding of the structure-function relationship

microbial genomics approaches in the context of identification and characterization of potential targets as a drug or vaccine candidate.

5.3.1 Reverse Vaccinology/Genomics

Reverse vaccinology is the in silico screening of the pathogen genome to find out the repertoire of antigens/drug targets that are expressed by the organism. By using various bioinformatics tools, it is possible to predict the ORFs of all the genes that are exposed or secreted on the surface of pathogen. Genes which are uniquely present in a certain pathogen can be selected for in vitro and in vivo studies. This involves a few critical experimental steps like gene cloning and expression, protein purification, and then selection of the potential candidate (Grandi 2001). One of the best examples of reverse vaccinology approach is serogroup B *Neisseria meningitidis* (MenB) project. In this project, numerous novel vaccine candidates were determined in a period of 18 months, and it outnumbered the discovery made in 40 years of conventional vaccinology (Pizza et al. 2000). In the analysis of MC58 strain genome (belongs to MenB), 570 ORFs out of 2158 ORFs were predicted to encode either surface exposed or secreted (Pizza et al. 2000). Antigen sorting was continued based on handful criteria which include the identification of the ability of antigens to be cloned and expressed in *Escherichia coli* as recombinant proteins (350 candidates) followed by the validation of the antigen exposed on the cell surface (91 candidates) by ELISA and flow cytometry. To confer protective immunity, the ability of induced antibodies (28 candidates) was measured by serum bactericidal assay or passive protection in infant rat. Further, screening was performed to identify the conservation of potential antigens in a panel of diverse meningococcal strains especially pathogenic strains of MenB (Rappuoli 2008; Giuliani et al. 2006). Using this methodology it was possible to identify five antigens, (i) genome-derived *Neisseria* antigens 1870 (GNA1870; which is factor H-binding protein [fHBP]), (ii) GNA1994 (which is NadA), (iii) GNA213, (iv) GNA1030, and (v) GNA2091. It also enabled the classification of outer membrane vesicles (OMV) from the New Zealand MeNZB vaccine strain that contains the immunogen PorA (Martin et al. 2006) and has been combined to form the Novartis MenB vaccine which entered the phase III clinical trials in 2008 (Rappuoli 2008; Giuliani et al. 2006).

5.3.2 Comparative Genomics

This approach is used to compare the pathogenic and nonpathogenic strains of the same species in order to identify the unique genes that are only present in pathogenic strains but absent in nonpathogenic strains. Those unique genes that are involved in pathogenesis and virulence of organisms might be the potential target for the development of vaccine and therapeutics (Bhagwat and Bhagwat 2008). Rasko et al. (2008) identified some genes that are present only in pathogenic strains of *E. coli* but absent in commensal strains during comparison of up

to 17 commensal and pathogenic strains of *E. coli*. With the rapid advancement in sequencing technology and bioinformatics, an exponential growth in genome sequence information has been achieved. Studying the genome sequence information of various pathogens to find out the genes conserved among the bacteria enables the identification of potential targets for the development of broad-spectrum antibiotics, while unique genes specific to particular species of bacteria might be an ideal target for narrow-spectrum antibiotics. For example, 26 genes in *E. coli* out of which most of them were conserved in various species such as *B. subtilis*, *M. genitalium*, *H. influenzae*, *H. pylori*, *Streptococcus pneumoniae*, and *Borrelia burgdorferi* genomes were identified by Arigoni F and colleagues (Arigoni et al. 1998). To potentially select the target, it is crucial to compare the genome sequence of the pathogen and the eukaryote so that the bacterial target proteins that are conserved among the mammalian proteins could simply be avoided to reduce the chances of human toxicity (Tatusov et al. 1997). For example, a previous report indicated significant sequence similarity between the broadly conserved proteins (15 out of 26) across the bacterial species and that of *Saccharomyces cerevisiae* (Arigoni et al. 1998).

5.3.3 Pan-Genomics

This is an advanced future of comparative genomics which aims at understanding the content, organization, and evolution of genomes and explains genotype-phenotype relationships. Availability of multiple genome sequences for a single species highlighted the importance of pan-genomics approach in identifying vaccine candidates in antigenically diverse species (Muzzi et al. 2007). The analysis of variation in genome sequences of pathogenic and its nonpathogenic strain leads to the rapid identification of genes involved in virulence. Pan-genomics focused on the variation in genomic sequence of different strains of same species which indicates that single genome sequence may not be enough or may not provide the complete understanding of intraspecies genetic variability (Fitzgerald et al. 2001; Dorrell et al. 2001; Fukiya et al. 2004; Obert et al. 2006). In pan-genomics approach, open reading frames (ORFs) are selected by screening of multiple genomes either by comparative genomics hybridization or by direct sequencing (Muzzi et al. 2007). These studies suggest that a potential vaccine and antimicrobial targets have to be conserved across all strains of the same species and are involved in the pathogenesis of bacterial pathogens. One of the best examples of genetic diversity studied through pan-genomics approach is seen in *Streptococcus agalactiae* (also known as group B streptococcus), a multiserotype bacterial pathogen that causes life-threatening disease in newborns. Genome sequence analysis of eight different strains of *S. agalactiae* predicted genetic variability and the extended collection of genes of the species (Tettelin et al. 2005). It can be classified into three parts: genes that are present only in one strain (strain specific genes), genes present in some strains but not in all strains (dispensable genome), and set of genes that are present in all strains (core genome). The bioinformatics screening predicted 589 genes as surface-exposed or

secreted proteins in the *S. agalactiae* genome. Among them, 396 and 193 genes are from the core and dispensable genome, respectively. Further, screening of these genes revealed four proteins that elicited protection in mice against all strains of *S. agalactiae* (Maione et al., 2005). Interestingly, it was found that a combination of four proteins GBS322, GBS104, GBS67, and GBS80 can act as a universal vaccine. However, only one of these proteins belonged to the core genome, while the rest of the three are from the dispensable genome of *S. agalactiae.* Therefore, the authors suggested that it is not the only conserved protein which essentially provide broad-spectrum protection (Kaushik and Sehgal 2008).

5.3.4 Transcriptomics

This genomics approach can be used for analysis of global changes in bacterial gene expression under a specific condition. Thus, genes which are essential for survival and pathogenesis of microorganisms in the host can be identified by the transcriptomic approach. The highly expressed genes can be selected for further analyses as they are crucial for microbial pathogenesis. On the contrary, low-expressed genes in host environment should be considered less important for a potential target. It is reported that targeting such genes which are shown to be essential for survival and expressed in virulence-induced condition has a higher potentiality to be drug target (Moir et al. 1999). Information about such essential genes that are also expressed in the animal model would indicate the importance of such genes in infection as well. There are commonly two types of methods for gene expression: first, cDNA-based microarray (cDNA derived from the RNA transcripts by using reverse transcription under specific condition) and second, ultra-high-throughput sequencing technologies that allow rapid sequencing and direct quantification of cDNA.

Identification of potential vaccine and therapeutic targets under experimental conditions by mimicking host-pathogen interaction is a good way. For example, in a study using microarray-based transcriptional profiling, it was found that adhesion to epithelial cells altered the expression of 350 genes by more than twofold, in which 189 genes were upregulated, 151 downregulated, and 7 genes either up- or downregulated depending on the time point in kinetics (Grifantini et al. 2002). They identified five new adhesion-induced proteins (NMB0315, NMB1119, NMB0995, NMB0652, and NMB1876) capable of inducing bactericidal antibodies in mice (Grifantini et al. 2002). However, there are some major limitations of this approach. Firstly, there was no direct correlation between the levels of proteins and mRNA. Secondly, in vivo studies require large amounts of mRNA; amplification of mRNA further creates additional technical challenges. Thirdly, they failed to establish a correlation between animal or cell-culture systems and the human host. Some other examples of microarray-based transcriptional profiling are (i) *Mycobacterium tuberculosis* genes encoding proteins that could be targeted for vaccine development, which are expressed during host infection (Talaat and Stemke-Hale 2005), and (ii) transcriptional profiling of *Vibrio cholerae* genes that are expressed during human infection (Merrell et al. 2002).

In addition to these techniques, several alternative techniques (in vitro expression technology (IVET), in vivo induced antigen technology (IVAT), and expression library immunization) are also developed for the study of bacterial gene expression globally (Angelichio and Camilli 2002; Talaat and Stemke-Hale 2005). Besides these techniques, signature-tagged mutagenesis (STM), genome analysis and mapping by in vitro transposition, and transposon site hybridization (TraSH) techniques are also developed with the special emphasis on the bacterial genes whose expression is dependent on host-pathogen interaction. The idea behind the development of such high-throughput techniques is to find out number of vaccine and therapeutic targets from bacterial species (Merrell et al., 2002; Moxon and Rappuoli 2002; Scarselli et al. 2005).

5.3.5 Proteomics

Proteomics refers to analyzing a set of proteins expressed under specified conditions or in specific cellular location. Using this approach, the potential vaccine and therapeutic targets could be predicted by obtaining an overall view of the pathogen proteome and the host's immune response after infection. High-throughput proteomic analysis can also be performed by using several techniques such as mass spectrometry, chromatographic techniques, and protein microarrays (Grandi 2006). One of the chromatographic techniques like 2D-PAGE separates proteins that appear as fine spot on the gels; these are then isolated and subjected to further analysis by mass spectrometry. Mass spectrometric techniques such as MALDI-TOF (matrix-assisted laser desorption ionization-time of flight) and MS/MS (tandem mass spectrometry) are used for peptide mass and sequence analysis of protein spots on a gel (Patterson and Aebersold 2003; Zhu et al. 2003). One of the common examples of proteomics-based approach is the identification of 27 outer surface proteins of *S. agalactiae*, first by 2D-gel electrophoresis and then by peptide sequencing. Out of these, six proteins were cloned, expressed, purified, and then utilized for mice immunization experiments. Two potential candidates were found to be protective against a lethal dose of bacteria in a neonatal mouse model (Hughes et al. 2002). Grandi (2006) also analyzed the surface proteome of *Streptococcus pyogenes* to identify novel vaccine and therapeutic targets (Rodriguez-Ortega et al. 2006). This novel proteome-based approach is used to identify novel proteins in several organisms such as *Bacillus anthracis* (Ariel et al., 2003), *Streptococcus pneumoniae* (Ling et al., 2004), *Streptococcus iniae* (Shin et al., 2007), *Bartonella quintana* (Boonjakuakul et al., 2007), and *Mycobacterium tuberculosis* (Malen et al., 2008).

5.3.6 Immunomics

Immunomics is the analysis of a set of proteins and epitopes of the pathogen that interact with the host immune system. The proteome of bacteria can also be screened

to identify immunome of that bacterium by in silico and in vitro techniques. In silico techniques can be used to predict pathogen epitopes that can be recognized by B-cell and T-cell. Large-scale screening for B-cell and T-cell epitopes in pathogens including HIV, *B. anthracis*, *M. tuberculosis*, *F. tularensis*, *Yersinia pestis*, flaviviruses, and influenza is currently under process (Sette et al. 2005; De Groot et al. 2008a). Although epitope prediction may serve as a steer for further biological evaluation, T-cell epitopes are recognized by MHC/HLA complex on the surface of antigen-presenting cells (B-cell, macrophages, and dendritic cells), which differ considerably between hosts, confounding the task of functional epitope prediction. Furthermore, B-cell epitopes can be both linear and conformational. Finally, the rationale behind the study was to create a single peptide which could represent defined epitope combinations from a protein or organism and overcome the genetic variability of both pathogen and host (De Groot et al. 2008b).

Antibodies present in host serum upon exposure to a pathogen can be used to identify vaccine candidates. There are several established techniques which allow the high-throughput display of pathogen proteins and the subsequent screening for proteins that interact with antibodies present in host serum (Seib et al. 2009). Immunogenic surface proteins of various organisms have been identified in several studies, including *Staphylococcus aureus* using 2D-PAGE, membrane blotting, and MS (Vytvytska et al. 2002); *Streptococcus agalactiae*, *Streptococcus pyogenes*, and *S. pneumoniae* using phage- or *E. coli*-based comprehensive genomic peptide expression libraries (Meinke et al. 2005; Giefing et al. 2008); and *Francisella tularensis* (Eyles et al., 2007) and *Vibrio cholerae* using protein microarray chips (Rolfs et al. 2008). Characterization of protein-drug interactions, as well as other protein-protein, protein-nucleic acid, ligand-receptor, and enzyme-substrate interactions, can also be done by using protein microarray (Stoevesandt et al. 2009).

5.3.7 Structural Genomics

Structural genomics mainly focuses on the three-dimensional structure of an organism's proteins and how they interact with antibodies and therapeutics. NMR (nuclear magnetic resonance) and crystallography techniques are used to determine the structure of proteins and the conformational changes that occur during the interaction of proteins with antibodies and therapeutics. This approach is quite useful to engineer antibodies and inhibitors against specific proteins by using their structure-based design to find out the residues involved in the active site of that protein. High-resolution techniques for protein structure determination are mainly focused on understanding and analyzing the structural basis of immune-dominant and recessive antigens as well as active sites and potential drug binding sites of proteins (Dormitzer et al. 2008; Nicola and Abagyan 2009). Several methods have been developed for high-throughput characterization of proteins on the basis of their genome information (Todd et al. 2005). For example, structural characterization of two HIV

envelope proteins gp120 (glycoprotein 120) and gp41 (glycoprotein41) have shown mechanisms used by the virus to evade host antibody responses due to hypervariability in immunodominant epitopes (Zhou et al. 2007; Prabakaran et al. 2007). However, there are some limitations to this approach such as poor understanding of determinants of immunogenicity, immunodominance, and structure-function (Seib et al. 2009). Nevertheless, this approach is very important for high-throughput modification of proteins and their screening for immunogenicity and interaction with antimicrobials to develop some novel vaccine and therapeutics (Dormitzer et al. 2008).

5.4 Conclusions

In this chapter, we review the impact of microbial genomics approach on the development of novel vaccine and therapeutics. This chapter covers several microbial genomics approaches that have emerged to identify the potential candidates for vaccine and therapeutic design from our better understanding of the human genome. Genomic and proteomic approaches have been used to identify the surface proteins during host-pathogen interaction. Furthermore, transcriptomics tells us about the expression level of RNA transcripts during infection, which is useful to dig out the essential target for the development of vaccine and antibiotic targets. All these approaches are useful, but there still remain some challenges such as understanding of molecular nature of B-cell and T-cell antigenic determinants of immunogenicity, mechanisms of different adjuvants, and structure-function relationship of proteins. These challenges can be fulfilled by improvement of structural studies of antigenic determinants, immunogenicity, and B-cell and T-cell epitope prediction. Identification of novel vaccine and therapeutic targets through genome-based approaches has to be subjected to confirmation and validation by in vitro (e.g., bactericidal assay) and in vivo assays (e.g., animal protection experiments). Unavailability of valid models to measure efficacy and protection against disease is still a major issue of animal protection experiments. In spite of that, a wealth of information about the microbial pathogenesis obtained through genome-based approaches can be useful in sorting out this issue. Several effective vaccines and therapeutic candidates have to pass through confirmatory tests including stepwise series of pre-licensure clinical trials (phases I, II, and III) before being introduced into the market. However, preclinical trials that are required to check the safety, efficacy, and immunogenicity of potential vaccine and therapeutic targets are time-consuming and costly. We therefore believe that with advancements in the field of technology, we can expect to witness more effective and specific vaccine and therapeutic targets against a disease in the near future.

References

Angelichio MJ, Camilli A (2002) In vivo expression technology. Infect Immun 70:6518–6523

Ariel N, Zvi A, Makarova KS, Chitlaru T, Elhanany E, Velan B, Cohen S, Friedlander AM, Shafferman A (2003) Genome-based bioinformatics selection of chromosomal *Bacillus anthracis* putative vaccine candidates coupled with proteomic identification of surface-associated antigens. Infect Immun 71:4563–4579

Arigoni F, Talabot F, Peitsch M, Edgerton MD, Meldrum E, Allet E, Fish R, Jamotte T, Curchod ML, Loferer H (1998) A genome-based approach for the identification of essential bacterial genes. Nat Biotechnol 16:851–856

Bhagwat AA, Bhagwat M (2008) Methods and tools for comparative genomics of foodborne pathogens. Foodborne Pathog Dis 5:487–497

Boonjakuakul JK, Gerns HL, Chen YT, Hicks LD, Minnick MF, Dixon SE, Hall SC, Koehler JE (2007) Proteomic and immunoblot analyses of *Bartonella quintana* total membrane proteins identify antigens recognized by sera from infected patients. Infect Immun 75:2548–2561

Centers for Disease Control and Prevention (CDC) (2018) Antibiotic resistance: a global threat. https://www.cdc.gov/features/antibiotic-resistance-global/index.html

De Groot AS, Rivera DS, McMurry JA, Buus S, Martin W (2008a) Identification of immunogenic HLA-B7 "Achilles' heel" epitopes within highly conserved regions of HIV. Vaccine 26:3059–3071

De Groot AS, Moise L, McMurry JA, Martin W (2008b) Epitope-based immunome derived vaccines: a strategy for improved design and safety. In: Falus A (ed) Clinical applications of immunomics. Springer, New York, pp 39–69

Dormitzer PR, Ulmer JB, Rappuoli R (2008) Structure-based antigen design: a strategy for next generation vaccines. Trends Biotechnol 26:659–667

Dorrell N, Mangan JA, Laing KG, Hinds J, Linton D, Al-Ghusein H, Barrell BG, Parkhill J, Stoker NG, Karlyshev AV, Butcher PD, Wren BW (2001) Whole genome comparison of *Campylobacter jejuni* human isolates using a low-cost microarray reveals extensive genetic diversity. Genome Res 11:1706–1715

Eyles JE, Unal B, Hartley MG, Newstead SL, Flick-Smith H, Prior JL, Oyston PC, Randall A, Mu Y, Hirst S, Molina DM, Davies DH, Milne T, Griffin KF, Baldi P, Titball RW, Felgner PL (2007) Immunodominant *Francisella tularensis* antigens identified using proteome microarray. Proteomics 7:2172–2183

Fitzgerald JR, Sturdevant DE, Mackie SM, Gill SR, Musser JM (2001) Evolutionary genomics of *Staphylococcus aureus*: insights into the origin of methicillin-resistant strains and the toxic shock syndrome epidemic. Proc Natl Acad Sci U S A 98:8821–8826

Fukiya S, Mizoguchi H, Tobe T, Mori H (2004) Extensive genomic diversity in pathogenic *Escherichia coli* and *Shigella* strains revealed by comparative genomic hybridization microarray. J Bacteriol 186:3911–3921

Giefing C, Meinke AL, Hanner M, Henics T, Bui MD, Gelbmann D, Lundberg U, Senn BM, Schunn M, Habel A, Henriques-Normark B, Ortqvist A, Kalin M, von Gabain A, Nagy E (2008) Discovery of a novel class of highly conserved vaccine antigens using genomic scale antigenic fingerprinting of pneumococcus with human antibodies. J Exp Med 205:117–131

Giuliani MM, Adu-Bobie J, Comanducci M, Aricò B, Savino S, Santini L, Brunelli B, Bambini S, Biolchi A, Capecchi B, Cartocci E, Ciucchi L, Di Marcello F, Ferlicca F, Galli B, Luzzi E, Masignani V, Serruto D, Veggi D, Contorni M, Morandi M, Bartalesi A, Cinotti V, Mannucci D, Titta F, Ovidi E, Welsch JA, Granoff D, Rappuoli R, Pizza M (2006) A universal vaccine for serogroup B meningococcus. Proc Natl Acad Sci U S A 103:10834–10839

Grandi G (2001) Antibacterial vaccine design using genomics and proteomics. Trends Biotechnol 19(5):181–188

Grandi G (2006) Genomics and proteomics in reverse vaccines. Methods Biochem Anal 49:379–393

Grifantini R, Bartolini E, Muzzi A, Draghi M, Frigimelica E, Berger J, Ratti G, Petracca R, Galli G, Agnusdei M, Giuliani MM, Santini L, Brunelli B, Tettelin H, Rappuoli R, Randazzo F,

Grandi G (2002) Previously unrecognized vaccine candidates against group B meningococcus identified by DNA microarrays. Nat Biotechnol 20:914–921

Hughes MJ, Moore JC, Lane JD, Wilson R, Pribul PK, Younes ZN, Dobson RJ, Everest P, Reason AJ, Redfern JM, Greer FM, Paxton T, Panico M, Morris HR, Feldman RG, Santangelo JD (2002) Identification of major outer surface proteins of *Streptococcus agalactiae*. Infect Immun 70:1254–1259

Kaushik DK, Sehgal D (2008) Developing antibacterial vaccines in genomic and proteomic era. Scand J Immunol 67:544–552

Ling E, Feldman G, Portnoi M, Dagan R, Overweg K, Mulholland F, Chalifa-Caspi V, Wells J, Mizrachi-Nebenzahl Y (2004) Glycolytic enzymes associated with the cell surface of *Streptococcus pneumoniae* are antigenic in humans and elicit protective immune responses in the mouse. Clin Exp Immunol 138:290–298

Maione D, Margarit I, Rinaudo CD, Masignani V, Mora M, Scarselli M, Tettelin H, Brettoni C, Iacobini ET, Rosini R, D'Agostino N, Miorin L, Buccato S, Mariani M, Galli G, Nogarotto R, Nardi-Dei V, Vegni F, Fraser C, Mancuso G, Teti G, Madoff LC, Paoletti LC, Rappuoli R, Kasper DL, Telford JL, Grandi G (2005) Identification of a universal group B streptococcus vaccine by multiple genome screen. Science 309:148–150

Målen H, Søfteland T, Wiker HG (2008) Antigen analysis of *Mycobacterium tuberculosis* H37Rv culture filtrate proteins. Scand J Immunol 67:245–252

Martin DR, Ruijne N, McCallum L, O'hallahan J, Oster P (2006) The VR2 epitope on the PorA P1. 7-2, 4 protein is the major target for the immune response elicited by the strain-specific group B meningococcal vaccine MeNZB. Clin Vaccine Immunol 13(4):486–491

Meinke A, Henics T, Hanner M, Minh DB, Nagy E (2005) Antigenome technology: a novel approach for the selection of bacterial vaccine candidate antigens. Vaccine 23:2035–2041

Merrell DS, Butler SM, Qadri F, Dolganov NA, Alam A, Cohen MB, Calderwood SB, Schoolnik GK, Camilli A (2002) Host-induced epidemic spread of the cholera bacterium. Nature 417:642–645

Moir DT, Shaw KJ, Hare RS, Vovis GF (1999) Genomics and antimicrobial drug discovery. Antimicrob Agents Chemother 43(3):439–446

Moxon R, Rappuoli R (2002) Bacterial pathogen genomics and vaccines. Br Med Bull 62:45–58

Muzzi A, Masignani V, Rappuoli R (2007) The pan-genome: towards a knowledge-based discovery of novel targets for vaccines and antibacterials. Drug Discov Today 12:429–439

Nicola G, Abagyan R (2009) Structure-based approaches to antibiotic drug discovery. Curr Protoc Microbiol; Chapter 17:Unit 17.2

Obert C, Sublett J, Kaushal D, Hinojosa E, Barton T, Tuomanen EI, Orihuela CJ (2006) Identification of a candidate *Streptococcus pneumoniae* core genome and regions of diversity correlated with invasive pneumococcal disease. Infect Immun 74:4766–4777

Patterson SD, Aebersold RH (2003) Proteomics: the first decade and beyond. Nat Genet 33(Suppl):311–323

Pizza M, Scarlato V, Masignani V, Giuliani MM, Aricò B, Comanducci M, Jennings GT, Baldi L, Bartolini E, Capecchi B, Galeotti CL, Luzzi E, Manetti R, Marchetti E, Mora M, Nuti S, Ratti G, Santini L, Savino S, Scarselli M, Storni E, Zuo P, Broeker M, Hundt E, Knapp B, Blair E, Mason T, Tettelin H, Hood DW, Jeffries AC, Saunders NJ, Granoff DM, Venter JC, Moxon ER, Grandi G, Rappuoli R (2000) Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. Science 287:1816–1820

Prabakaran P, Dimitrov AS, Fouts TR, Dimitrov DS (2007) Structure and function of the HIV envelope glycoprotein as entry mediator, vaccine immunogen, and target for inhibitors. Adv Pharmacol 55:33–97

Rappuoli R (2008) The application of reverse vaccinology, Novartis MenB vaccine developed by design. 16th International Pathogenic Neisseria Conference, Rotterdam, The Netherlands. http://www.IPNC2008.org. Abstract, 81 p

Rasko DA, Rosovitz MJ, Myers GS, Mongodin EF, Fricke WF, Gajer P, Crabtree J, Sebaihia M, Thomson NR, Chaudhuri R, Henderson IR, Sperandio V, Ravel J (2008) The pangenome struc-

ture of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. J Bacteriol 190:6881–6893

Rodríguez-Ortega MJ, Norais N, Bensi G, Liberatori S, Capo S, Mora M, Scarselli M, Doro F, Ferrari G, Garaguso I, Maggi T, Neumann A, Covre A, Telford JL, Grandi G (2006) Characterization and identification of vaccine candidate proteins through analysis of the group A Streptococcus surface proteome. Nat Biotechnol 24:191–197

Rolfs A, Montor WR, Yoon SS, Hu Y, Bhullar B, Kelley F, McCarron S, Jepson DA, Shen B, Taycher E, Mohr SE, Zuo D, Williamson J, Mekalanos J, Labaer J (2008) Production and sequence validation of a complete full length ORF collection for the pathogenic bacterium *Vibrio cholerae*. Proc Natl Acad Sci U S A 105:4364–4369

Scarselli M, Giuliani MM, Adu-Bobie J, Pizza M, Rappuoli R (2005) The impact of genomics on vaccine design. Trends Biotechnol 23:84–91

Seib KL, Dougan G, Rappuoli R (2009) The key role of genomics in modern vaccine and drug design for emerging infectious diseases. PLoS Genet 5(10):e1000612

Sette A, Fleri W, Peters B, Sathiamurthy M, Bui HH, Wilson S (2005) A roadmap for the immunomics of category A-C pathogens. Immunity 22:155–161

Shin GW, Palaksha KJ, Kim YR, Nho SW, Kim S, Heo GJ, Park SC, Jung TS (2007) Application of immunoproteomics in developing a *Streptococcus iniae* vaccine for olive flounder (*Paralichthys olivaceus*). J Chromatogr B Analyt Technol Biomed Life Sci 849:315–322

Stoevesandt O, Taussig MJ, He M (2009) Protein microarrays: high-throughput tools for proteomics. Expert Rev Proteomics 6:145–157

Talaat AM, Stemke-Hale K (2005) Expression library immunization: a road map for discovery of vaccines against infectious diseases. Infect Immun 73:7089–7098

Tatusov RL, Koonin EV, Lipman DJ (1997) A genomic perspective on protein families. Science 278:631–637

Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, Angiuoli SV, Crabtree J, Jones AL, Durkin AS, Deboy RT, Davidsen TM, Mora M, Scarselli M, Margarit y Ros I, Peterson JD, Hauser CR, Sundaram JP, Nelson WC, Madupu R, Brinkac LM, Dodson RJ, Rosovitz MJ, Sullivan SA, Daugherty SC, Haft DH, Selengut J, Gwinn ML, Zhou L, Zafar N, Khouri H, Radune D, Dimitrov G, Watkins K, O'Connor KJ, Smith S, Utterback TR, White O, Rubens CE, Grandi G, Madoff LC, Kasper DL, Telford JL, Wessels MR, Rappuoli R, Fraser CM (2005) Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". Proc Natl Acad Sci U S A 102:13950–13955

Todd AE, Marsden RL, Thornton JM, Orengo CA (2005) Progress of structural genomics initiatives: an analysis of solved target structures. J Mol Biol 348:1235–1260

Vytvytska O, Nagy E, Blüggel M, Meyer HE, Kurzbauer R, Huber LA, Klade CS (2002) Identification of vaccine candidate antigens of *Staphylococcus aureus* by serological proteome analysis. Proteomics 2:580–590

Zhou T, Xu L, Dey B, Hessell AJ, Van Ryk D, Xiang SH, Yang X, Zhang MY, Zwick MB, Arthos J, Burton DR, Dimitrov DS, Sodroski J, Wyatt R, Nabel GJ, Kwong PD (2007) Structural definition of a conserved neutralization epitope on HIV-1 gp120. Nature 445:732–737

Zhu H, Bilgin M, Snyder M (2003) Proteomics. Annu Rev Biochem 72:783–812