

Immunogenicity Prediction by VaxiJen: A Ten Year Overview

Nevena Zaharieva¹, Ivan Dimitrov¹, Darren R Flower² and Iri Doytchinova^{1*}

¹Faculty of Pharmacy, Medical University of Sofia, Bulgaria

²School of Life & Health Sciences, Aston University, Aston Triangle, Birmingham, UK

Abstract

Vaccination stands alone as an unparalleled medical intervention able to save human lives numbering hundreds of millions. Modern licensed vaccines are either whole organism based or based on single proteins, together with carbohydrate epitope-based vaccines. Single protein or so-called subunit vaccines are prime targets for vaccine design and reverse vaccinology. Immunoinformatics is a branch of bioinformatics focusing on immunology and vaccinology, which includes database compilation, data mining, and epitope prediction. Immunoinformatics can aid in the discovery of subunit vaccines through algorithms for immunogenicity prediction. VaxiJen is the first and, ostensibly, the only widely-used method for distinguishing between immunogens and non-immunogens among proteins of bacterial, viral, parasite, fungal and tumour origin. In this review, we chart the applications of VaxiJen, placing these into context, and explore some of the future directions that research into identifying immunogens might take.

Keywords: VaxiJen; Subunit vaccine; Immunoinformatics; Immunogenicity prediction

Introduction

Vaccination is unquestionably one of the greatest achievements of human civilisation. The WHO estimates that today immunizations prevent between 2 and 3 million deaths annually and protect many more people from illness and disability [1]. Current vaccines are one of several types: live attenuated microorganisms (Hepatitis A virus, Influenza virus, Japanese encephalitis virus, Measles virus, Mumps virus, Poliovirus, Rabies virus, Vaccinia virus, Varicella zoster virus, Yellow fever virus, *M. tuberculosis*, *S. typhi*), microorganisms inactivated or killed by heat or chemical treatment (Influenza virus, Japanese encephalitis virus, Mumps virus, Poliovirus, *V. cholerae*, *B. pertussis*, *Y. pestis*), subunit vaccines (proteins from Influenza virus and Hepatitis B virus, toxoids of *B. pertussis*, *C. tetani* and *C. diphtheriae*, conjugates of *H. influenzae* type b, *N. meningitis* and *S. pneumoniae*), recombinant vaccines (Hepatitis B virus surface antigen and Human papillomavirus), or carbohydrate epitope-based vaccines (Pneumovax-23) [2].

The wide use of vaccines during the last 150 years has led to significant reduction (95-97%) in mortality from a wide range of previously deadly diseases, including diphtheria, tetanus, measles, mumps, rubella, pneumonia, hepatitis B and meningitis [3], as well as the total eradication of smallpox and the near eradication of polio. Yet, despite such tremendous success, a long list of deadly diseases still await efficient vaccines.

Vaccine development remains reliant on a gallimaufry of antiquated processes. The knowledge-based design of vaccines is still very much in an initial stage. Immunogenicity is the property of a molecule (protein, lipid or carbohydrate, or a combination thereof) or living organism (virus, bacterium, parasite or fungus) that induce humoral and/or cell-mediated response from the immune system. An immunogen triggering a protective immune response is a protective immunogen. Protective immunogens can make promising vaccine candidates.

Identification of protective immunogens among the proteome of a particular microorganism is the initial step in the process of vaccine design and development. This is the key step in what has come-to-be-known as Reverse Vaccinology [3]. Use of *in silico* methods in this initial stage can direct and thus greatly shorten subsequent

experimental work [4-7]. Moreover, the proper use of *in silico* methods can replace, reduce and refine the use of time-consuming and often misleading animal experimentation [8].

Immunoinformatics is a branch of bioinformatics focusing on immunology and vaccinology [5]. It includes strategies for databases compilation, data mining and analyses, *in silico* methods and algorithms for immunogenicity prediction, including tools for the prediction of B-cell and T-cell epitopes. The current state of immunoinformatics has been reviewed recently [9,10].

VaxiJen (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>); or, alternatively, just type vaxijen into Google) was the first server for prediction of protective antigens, tumour antigens and subunit vaccines [11]. It is also the first alignment-free bioinformatics tool for *in silico* identification of immunogens. VaxiJen uses Wold's z-scales [12] to describe the main physicochemical properties of the amino acids building the tested proteins, then converts the derived strings into uniform vectors by auto cross covariance (ACC) [13], selects the relevant variables by genetic algorithm (GA) [14] or stepwise regression and finally, classifies the proteins as protective antigens or non-antigens by partial least squares (PLS)-based discriminant analysis. Initially, the algorithm was trained to identify bacterial protective immunogens [15]. Later, models for viral and tumour immunogens were included [11] and VaxiJen was developed to give free access to the models. The last version of VaxiJen (VaxiJen 2.0) also includes models for identification of parasite and fungal immunogens [16].

Since its launch in 2007, VaxiJen has been widely used to identify candidate subunit vaccines among proteins of bacterial, viral, parasite,

***Corresponding author:** Iri Doytchinova, Faculty of Pharmacy, Medical University of Sofia, 2 Dunav st., Sofiq 1000, Bulgaria, Tel: +359-2-9236506; E-mail: idoytchinova@pharmfac.mu-sofia.bg

Received November 06, 2017; **Accepted** November 20, 2017; **Published** November 26, 2017

Citation: Zaharieva N, Dimitrov I, Flower DR, Doytchinova I (2017) Immunogenicity Prediction by VaxiJen: A Ten Year Overview. J Proteomics Bioinform 10: 298-310. doi: 10.4172/jpb.1000454

Copyright: © 2017 Zaharieva N, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

fungal and tumour origin. In the present review, we analyse the diverse applications of VaxiJen. Special attention is given to the experimentally validated predictions. We also take the opportunity to explore a few of the future directions that research into identifying immunogens might take.

Prediction of bacterial immunogens

Staphylococcus aureus (*S. aureus*) is one of the most important causes of nosocomial and community-acquired infections. Yet no vaccine against *S. aureus* exists. Fragments of the virulence proteins clumping factor A (ClfA), iron-regulated surface determinant (IsdB) and gamma hemolysin (Hlg) were selected and predicted as immunogenic by VaxiJen (scores 0.60, 0.47 and 0.58 at cutoff 0.4) [17]. A recombinant gene containing the three fragments and hydrophobic linkers between them was constructed and expressed in *E. coli* BL21 [18]. Immunisation of BALB/c mice with the recombinant protein evoked antigen-specific antibodies and increased survival following the intraperitoneal challenge with pathogenic *S. aureus*.

S. aureus clumping factor B was also identified by VaxiJen as a protective antigen (score 1.09) of *S. aureus* [19]. Shahbazi et al. [20] proposed a hexavalent subunit vaccine containing fragments of agglutinin-like sequence 9 (Als9) (VaxiJen score 0.93), ClfA (score 1.53), methicillin resistance determinant protein (FtmB) (score 1.15), immunoglobulin G-binding protein A (Spa) (score 1.55), serine-aspartate repeat-containing protein E (SdrE) (score 1.34) and biofilm-associated surface protein (Bap) (score 1.10). Hajjigharamani et al. [21] constructed a multi-epitope peptide vaccine containing B- and T-cell epitopes from alpha-enolase (Eno1), ClfA and IsdB. VaxiJen scored this candidate-vaccine as a moderately antigenic (score 0.50) [21].

Streptococcus pneumoniae (*S. pneumoniae*, pneumococcus) is the major pathogen causing pneumonia, meningitis and sepsis. Currently, two vaccines for prevention from pneumonia caused by *S. pneumoniae* are licensed for adults: 23-valent polysaccharide vaccine (PPV23) and 13-valent conjugate vaccine (PCV13) [22]. Both vaccines contain capsular polysaccharides eliciting weak immunogenicity in children, and are unable to induce effective immune memory [2]. The protein-containing PCV13 is more effective in children than PPV23; the high manufacturing costs limit widespread use in developing countries [23]. *S. pneumoniae* serotype 19F is among the main pneumococcal serotypes that cause invasive pneumonia in children under 5. Using bioinformatics analysis (including VaxiJen) on surface proteins from *S. pneumoniae* serotype 19F, Tarahomjoo [24] identified four putative candidate-vaccines: cell wall surface anchor family protein, D-alanyl-D-alanyl-carboxy peptidase, surface protein PspC and choline binding protein D. VaxiJen was also used to identify B-cell epitope regions in three surface proteins: autolysin, zinc binding lipoprotein and plasmid stabilization protein [25]. Additionally, two immunogenic conserved regions of choline binding protein A (CbpA) were predicted by VaxiJen [26].

Mycobacterium tuberculosis (*M. tuberculosis*) causes tuberculosis – the second leading cause of death worldwide [27]. BCG, the existing live attenuated vaccine against *M. tuberculosis*, protects newborns but does not prevent latent infection or reactivation of tuberculosis in adults [2]. VaxiJen was combined with other bioinformatics tools to identify several potential vaccine candidates: tyrosine phosphatase PtpA [28], proteins EsxL, PE26, PPE65, PE_PGERS49, PBP1 and Erp [29], Rv2031c protein [30], Rv3083 protein [31], hypothetical proteins Rv1904 and Rv2387 [32], Myt272-3 recombinant protein [33], and several methyltransferases [34].

The non-pathogenic, saprophytic *Mycobacterium indicus pranii* (MIP) provides protection against *M. tuberculosis* infection in mice [35,36]. It is used as an adjunct to chemotherapy in patients with type I and type II category tuberculosis [37]. VaxiJen identified proteins MIP0340, MIP5962 and MIP7697 as the most prominent putative antigens [38]. A comparative analysis of *M. tuberculosis* and MIP using VaxiJen highlighted the importance of the PE/PPE family in host immunomodulation, supporting further the likely potential of MIP as an effective vaccine against tuberculosis [39].

Strains of enterotoxigenic *Escherichia coli* (ETEC) are the most common cause of bacterial diarrhea in travelers to and children in developing countries [40]. ETECs adhere to the host small intestine by colonization factors (CFs) and produce enterotoxins [41]. ETEC secrete heat-stable enterotoxins (STs) and/or heat-labile enterotoxins (LTs). The design of subunit vaccines focus on CFs and toxins. A synthetic chimeric gene, encoding the colonization factors CFA/I, CS2 and CS3 and the B-subunit of LT was designed [42]. VaxiJen showed high antigenicity of the chimeric protein. However, there is no experimental validation of the immunogenicity of the chimeric protein. Similarly, a candidate subunit vaccine, composed of CFAB, CSSA, CSSB, and LTB, was designed, its immunogenicity predicted by VaxiJen, synthesized, and expressed successfully in a prokaryotic host but without subsequent experimental validation [40]. A recombinant protein containing CFAB, CFAE and LTB has been tested experimentally [41]. Mice have been immunized with this protein and the antibody titer and specificity of the sera analyzed by ELISA.

Mehla and Ramana applied another approach to identify novel immunogenic proteins from ETEC [43]. They identified proteins shared between virulent ETEC strains E24377A and H10407, but not by commensal *E. coli*. From this initial pool, human homologues were eliminated, the rest were predicted by VaxiJen, with only immunogenic proteins scrutinized for cellular location. Three novel probable immunogenic proteins were identified – putative membrane protein (uniprot ID: A7ZGR5), uncharacterized protein (uniprot ID: A7ZGK4) and O-antigen polymerase (uniprot ID: A7ZTH5) – and then analyzed for T- and B-cell epitopes.

The enterohemorrhagic *E. coli* (EHEC) strains are major human food-borne pathogens, responsible for bloody diarrhea and hemolytic-uremic syndrome worldwide [44]. The genomes of EHEC strains EDL933 and Sakai have been screened to identify common EHEC antigens absent in nonpathogenic *E. coli* strains. The analysis revealed 897 protein sequences. Applying immunoinformatics (including VaxiJen) their number was reduced to

65. Nine were tested for immunogenicity using a murine gastrointestinal infection model. Two vaccine candidates - Lom-like protein and a putative pilin subunit - significantly induced Th2 cytokines and production of sIgA, while the third (a fragment of the type III secretion structural protein EscC) reduced EHEC cecum colonization [44]. Chimeric proteins have been designed containing virulence factors of ETEC and EHEC [45] and ETEC, EHEC and Shigella [46].

Neisseria gonorrhoeae (*N. gonorrhoeae*), a Gram-negative diplococcus, causes one of the most common sexually transmitted diseases. It first adheres to the epithelium cells using pili and opa proteins, then penetrates, multiplying on basement membranes, producing lipopolysaccharide endotoxins [47]. Four essential membrane proteins - D-alanine-D-alanine ligase (ddl), sulfate transport permease protein C (cysW), competence lipoprotein (comL),

and type IV pilin protein (pilV) of *N. gonorrhoeae* virulent strain FA 1090 - previously identified as vaccine candidates [48], were studied for antigenicity using VaxiJen and then used to identify epitopes for both B- and T-cell mediated immune responses [49,50]. Bhairamadgi and Katti [46] screened the proteome of *N. gonorrhoeae* strain FA 1090 using immunoinformatic tools (including VaxiJen) and identified a novel immunogen - the hypothetical protein YP_208831.1 - as having the potential to induce both T- and B-cell mediated immunity. The proteomes of four *N. gonorrhoeae* strains FA 1090, TCDC_NG08107, NCCP11945 and MS11 were screened to identify probable secretory proteins [51]. From this initial pool, proteins with human homologues were removed, antigens selected by VaxiJen and B- and T-cell epitopes predicted. The final set consisted of 25 antigenic secretory proteins of *N. gonorrhoeae* which could be experimentally validated for vaccine development.

Helicobacter pylori (*H. pylori*) is a human gastric pathogen implicated as the major cause of peptic ulcer and second leading cause of gastric cancer around the world [52]. Based on the genome analysis of 39 *H. pylori* isolates, Ali et al. [52] selected 28 non-host homolog proteins as therapeutic targets. After selection by VaxiJen and epitope mapping, 3 highly conserved and 2 highly variable putative pathogenicity islands were revealed. Among them 5 potential vaccine candidates - vacA, babA, sabA, fecA and omp16 - were prioritized [53]. A chimeric gene containing four fragments of FlhD, UreB, VacA and CagL with a high density of B- and T-cell epitopes was designed and optimized in terms of solubility, antigenicity (by VaxiJen) and surface accessibility [54] for expression in *E. coli* BL21.

Campylobacter is one of the four key global causes of diarrheal diseases [55] and a major global cause of human gastroenteritis. Campylobacter infections are generally mild, but can be fatal among very young children, the elderly, and the immunosuppressed. Mehla and Ramana [56] screened the genome of *C. jejuni* pathogenic strain NCTC11168 and found 66 proteins non-homologous to the human proteome, 34 of them were deemed to be "drugable". As poultry constitutes the main animal reservoir, poultry vaccination is a promising way to reduce incidence of campylobacteriosis in humans [57]. Meunier et al. [57] screened the genome of the highly virulent Campylobacter jejuni subsp. jejuni 81-176 strain, selecting 24 extracellular and outer membrane proteins. Two of them were the known antigens FlaA and FlaB, the rest were newly identified antigens. After VaxiJen ranking, B- cell epitope mapping, and BLAST searching for conserved regions between *C. jejuni* and *C. coli*, 14 proteins were identified as potential vaccine candidates. The whole proteome of *C. jejuni* has been investigated for antigenicity, allergenicity, MHC binding, conservancy, and population coverage [58] and several conserved T-cell epitopes have been selected covering more than 80% of the human population.

Acinetobacter baumannii (*A. baumannii*) causes a variety of nosocomial infections of the respiratory tract, bloodstream, skin and soft-tissue infections [59]. Because of its biofilm-forming ability, it is resistant to most antibiotics [60]. The biofilm associated protein (Bap) of *A. baumannii* was searched for a conserved antigenic region using VaxiJen [61], identifying it as an antigen, with the potential to be a diagnostic. A comparative genome analysis was done to identify conserved proteins among five *A. baumannii* strains [62]. Three outer membrane proteins (outer membrane receptor protein, putative penicillin binding protein and glutamate synthase large chain precursor) with high VaxiJen scores were selected and B- and T-cell epitopes predicted. The protein *Baumannii acinetobactin* (BauA) was

investigated using bioinformatics tools (including VaxiJen) and two regions (26-191 of cork domain and 321-635 of barrel domain) selected as candidate vaccines [63]. Thirteen putative antigens were identified by a bioinformatic analysis of 30 *A. baumannii* strains [64]. The set included P pilus assembly protein, pili assembly chaperone, AdeK, PonA, OmpA, general secretion pathway protein D, FhuE receptor, Type VI secretion system OmpA/MotB, TonB dependent siderophore receptor, general secretion pathway protein D, outer membrane protein, peptidoglycan associated lipoprotein and peptidyl- prolyl cis-trans isomerase. The *A. baumannii* phospholipase D (pLD) was analyzed for antigenicity by VaxiJen, followed by B- and T-cell epitope mapping [65].

Vibrio cholerae (*V. cholerae*) is a noninvasive gram-negative bacterium causing water borne disease cholera [66]. Despite extant vaccines - inactivated whole organism and B subunit of cholera toxin [2] - development of novel vaccines remains appealing. Barn et al. [66] selected three candidate vaccines (OmpU, UppP and YajC) from the *V. cholerae* strain O395 using bioinformatics tools (including VaxiJen) and searched them for B- and T-cell epitopes. Nezafat et al. [67] have analyzed 6 known *V. cholerae* protective antigens (OmpW, OmpU, TcpA, TcpF and CTB) to identify promiscuous epitopes, binding to various HLA class II alleles and B-cell epitopes. The identified epitopes were linked together and the fused protein was predicted by VaxiJen to be antigenic. Two B-cell epitopes from Omp containing T-cell epitopes covering HLA class I and class II alleles have been designed as candidate epitope vaccines [68].

Salmonella Typhi (*S. typhi*) is a Gram-negative bacterium causing human typhoid fever. Prabhavathy et al. [69] proposed *S. typhi* proteins OmpLA and LsrC as suitable vaccine candidates based on antigenicity predicted by VaxiJen, followed by prediction of B- and T-cell epitopes, and identification of common epitopes for multiple pathogens. Toobak et al. [70] identified three major outer membrane proteins OmpC, OmpF and OmpA, which were amplified, cloned and expressed. The antigenicity of Omps was predicted by VaxiJen and confirmed by ELISA. Control and all mice immunized with a single Omp gene died within 24 hours of challenge with *S. typhi*. Mice immunized with two Omps survived for 48-50 hours. Mice immunized with three Omps survived 75 h. Generally, despite the high immunogenicity of Omps, they did not induce long-lasting protection in mice. As mice are not the natural hosts, results in human may well be different [70].

Leptospirosis caused by the pathogen *Leptospira* is one of the most widespread zoonotic diseases in the world [71]. Victor et al. [72] examined the evolutionary relationships of the outer membrane lipoprotein LipL41 taking 87 sequences from various *Leptospiral* serovars and strains followed by B-cell epitope mapping of conserved regions; identifying 8 LipL41 B-cell epitopes with high VaxiJen score. A similar systemic protein selection, antigenicity prediction and B- and T-cell mapping was applied to Omps of *Treponema pallidum* [73], *Shigella flexneri* [74], *Pseudomonas aeruginosa* [75] and *Clostridium botulinum* [76].

Prediction of viral immunogens

Influenza viruses are of four types: A, B, C and D. Type A infects a wide range of avian and mammalian species. Type B almost exclusively infects humans. Type C infects humans, dogs and pigs and causes mild respiratory illness, while type D only infects cattle. Among the four types, type A is the most virulent human pathogen and the most variable. It is classified into subtypes according to the serological reactivity of its surface glycoprotein antigens, haemagglutinin (HA)

and neuramidase (NA) [77]. Gupta et al. [77] collected 86 H1N1 HA protein sequences analyzing them for conservation. The analysis identified 15 conserved regions containing 13 HLA class I and 17 HLA class II epitopes. Their antigenicity was assessed by VaxiJen, with only 4 class I and 9 class II epitopes were proposed as a novel candidate epitope-based vaccine. Moattari et al. [78] compared 10 viral sequences collected from Iranian patients between 2010 and 2013 with 3 vaccine isolates: (A(H1N1)California/2009, A(H1N1) California X-157/2009 and A(H1N1)Brisbane/2007). This study detected several amino acids changes in HA which do not affect the epitope sites, antigenicity (predicted by VaxiJen), or secondary and tertiary structure of HA.

The Human immunodeficiency virus (HIV) causes acquired immunodeficiency syndrome (AIDS). HIV enters the host cell by forming a complex between the viral envelope glycoprotein (Env), the host receptor CD4, and chemokine co-receptors usually CCR5 or CXCR4 [79]. Antiretroviral peptides can inhibit the virus-coreceptor interaction by binding either virus envelope proteins or host proteins [80]. Rao et al. [80] compiled a dataset of 110 HIV antiviral peptides and analyzed them for conservancy, antigenicity by VaxiJen, hydrophobicity and antimicrobial activity. Fourteen peptides have been selected for promising AIDS treatment.

Hepatitis C virus (HCV) affects between 130 and 150 million people worldwide and cause chronic liver disease, liver cirrhosis and hepatocellular carcinoma [81]. *In silico* analyses and B- and T-cell epitope predictions have been made for the structural envelope glycoproteins 1 (E1) [82] and 2 (E2) [83,84] and the non-structural proteins NS3, NS4A, NS5A and NS5B [85]. All predicted epitopes are antigenic according to VaxiJen (VaxiJen score > 0.4).

Zika virus (ZIKV) is a mosquito-borne virus causing mild headache, cutaneous rash, fever, malaise, conjunctivitis and arthralgia [86]. It consists of 10 proteins: capsid, precursor of membrane, envelope and seven non-structural (NS) proteins [87]. The most investigated is the envelope glycoprotein protein. Several B- and T-cell epitopes have been identified by immunoinformatic tools (including VaxiJen) [88-91], but have not been tested experimentally. Dar et al. [92] analyzed 54 full length ZIKV polyprotein sequences identifying 23 HLA class I and 48 HLA class II binders. The most of them are localized in NS5, followed by envelope, NS1 and NS2.

Dengue virus (DENV) is a mosquito-borne virus causing life threatening hemorrhagic fever and shock syndrome [93]. There are four DENV serotypes and several genotypes [94]. The DENV genome encodes 3 structural (capsid C, precursor of membrane prM and envelope E) and 7 non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5). The DENV nucleocapsid is covered by a lipid bilayer containing envelope glycoprotein and membrane proteins [95]. The extramembrane location of these proteins makes them good candidates for vaccine development. The proteins E, prM, NS1 and NS3 have been identified by VaxiJen as probable antigens [96], and several B- and T-cell epitopes have been predicted [96-98]. The non-structural protein NS5 is the most conserved DENV protein making it a good vaccine target [99]. It has low intrinsic antigenicity but contains several potent and promiscuous T- and B-cell epitopes [98,99].

Ebola virus (EBOV) causes a fatal hemorrhagic fever with death rates up to 90% [100]. Fruit bats are the natural host, and disease transmission to human and other primates is mainly via bodily fluids (blood, secretions, semen) [101]. The EBOV proteome has been searched for antigenic proteins using VaxiJen: the L protein (UniprotKB ID: K4G1K7) was identified as the most antigenic (VaxiJen score

0.7024) [102]. Further, a highly promiscuous B- and T-cell epitope was identified as a probable epitope-based vaccine [102]. The glycoprotein 2 (GP2) and viral protein 24 (VP24) have also been analyzed for conservation, antigenicity using VaxiJen, and promiscuous B- and T-cell epitopes [103].

Crimean-Congo hemorrhagic fever virus (CCHFV) is a tick-borne pathogen causing hemorrhagic fever with fatality rates between 15% and 70% [104]. A total of 80 envelope glycoproteins and 34 RNA-dependent RNA polymerase-L molecules from different CCHFV variants were analyzed for conservation, identifying 4 conserved regions [105]. Two regions were antigenic according to VaxiJen, and one nonamer epitope had the potential to interact with 8 HLA class I and 27 HLA class II alleles [105]. T-cell epitopes were also identified from the proteins GP1 and GP2 [106].

Human coronaviruses (HCoVs) cause mild to severe respiratory tract infections, and are named for the crown-like spike proteins on their surface. 56 outer membrane spike protein sequences from different variants belonging to five types (229E, NL63, HKU1, EMC, and OC43) were retrieved from UniProtKB and assessed for antigenicity using VaxiJen [107]. The spike protein (UniprotKB id: B2KKT9) with the highest VaxiJen score was searched for B- and T-cell epitopes, identifying two peptide sequences as conserved and promiscuous. Shi et al. [108] analyzed nucleocapsid (N) and spike (S) proteins of Middle East respiratory syndrome coronavirus (MERS-CoV) selecting 10 B- and T-cell epitopes as MERS vaccine candidates.

Chikungunya virus (CHIKV) is a mosquito-borne virus causing fever and severe joint pain [109]. The CHIKV proteome was analyzed by VaxiJen, identifying envelope protein 2 as the most immunogenic [110]. Kori et al. [111] have searched the proteomes of three different CHIKV strains for conserved regions and five B- and T-cell epitopes have been predicted from both structural and non-structural proteins.

VaxiJen, followed by B- and T-cell epitope predictions, was used to design epitope-based vaccines against several other viruses: Saint Louis encephalitis virus (SLEV) [112], Nipah virus (NiV) [113], Hantaan virus (HNTV) [114], human cytomegalovirus (HCMV) [115], Lassa virus [116], human enterovirus D-68 (EV-D68) [117], Cardamon mosaic virus (CdMV) [118], Henipavirus [119], Hantavirus [120], human rotavirus A [121], Lymphocytic choriomeningitis virus (LCMV) [122], human adenovirus E (HAdV-E) [34], and human papillomavirus type 16 (HPV) [123].

Prediction of parasite immunogens

Plasmodium falciparum (*P. falciparum*) causes malaria – the most important parasitic disease, killing one child under 5 year every 120 seconds [124]. *P. falciparum* has a large genome encoding over 5300 proteins, with stage-specific expression and variation within a single strain [125]. Singh et al. [126] used VaxiJen to identify 22 probable antigens, containing up to 15,000 predicted epitopes binding to HLA class I and II supertypes and covering 95% of the human population. The complete proteome of *P. falciparum*, after excluding human homologs, was subjected to subcellular localization prediction, with outer membrane proteins analyzed by VaxiJen [127]. Four membrane-associated hypothetical proteins were identified, containing B- and T-cell epitopes. Singh et al. [128] screened 32 extracellular secretory proteins of *P. falciparum* using VaxiJen, predicting 31 as antigenic and containing many epitopes binding to HLA-A, -B and -DR.

Leishmaniasis are a group of tropical diseases caused by protozoan parasites of genus *Leishmania* and transmitted to humans by

hematophagous sandflies [129]. The entire proteome of *Leishmania major* (*L. major*) comprising 8312 proteins was screened for signal peptides and GPI anchors [130]. A set of 151 proteins was selected and subjected to consensus antigenicity prediction using VaxiJen and ANTIGENpro [131], identifying 25 vaccine candidates. *Toxoplasma gondii* (*T. gondii*) is an obligate intracellular parasitic protozoan causing the disease toxoplasmosis [132]. The *T. gondii* major surface antigen, called SAG1 or p30, was screened for conservation and antigenicity using VaxiJen [133], identifying four conserved antigenic regions.

Prediction of fungal immunogens

Fusarium circinatum (*F. circinatum*) causes the economically important disease of pines pitch canker leading to devastating forestry industry losses [134]. The detection of the pathogen in plant growth media and in plant tissues during the early stages of infection is very important [135]. Maphosa et al. [135] conducted comparative genomic studies, identifying 24 unique antigenic proteins. The ORFs of 5 variants were selected and tested by PCR analyses and hybridization assays. The results have showed that three of the selected genes are common and unique to *F. circinatum* and are thus good candidates for rapid, in-the-field diagnostic assays specific to *F. circinatum*.

Prediction of cancer immunogens

Cancer immunotherapy aims to stimulate immune responses of B- and T-lymphocytes and prevent immune system suppression by IL-10, PGE2 and COX2, secreted by tumor cells [136]. Nezafet et al. [137] designed a multi-epitope polypeptide cancer vaccine containing two tumor-associated antigens E6 and E7 from human papillomavirus as cytotoxic T-cell epitopes, tetanus toxin fragment c (TTfrc) and pan-allelic DR epitope (PADRE) as helper T-cell epitopes, and the TLR4 agonist heparin-binding hemagglutinin (HBHA) as an adjuvant, epitopes separated by proteasome-sensitive linkers. This construct was analyzed for immunogenicity by VaxiJen and ANTIGENpro, and B-cell epitopes identified on the protein surface.

The receptor tyrosine kinase like orphan receptor 1 (ROR1) is a transmembrane protein overexpressed in several cancers, including gastric carcinoma and breast cancer [138]. A chimeric protein consisting of the extracellular domain of ROR1 and the powerful T-cell activator staphylococcal enterotoxin have been constructed as a potent vaccine for breast cancer and analyzed for antigenicity, allergenicity and the presence of B- and T-cell epitopes [139]. A multi-epitope vaccine against breast cancer was designed to include cytolytic T-cell epitopes from human epidermal growth factor receptor (HER2), mucin 1 protein and heparanase as well as helper T-cell epitopes from survivin and Por B from *Neisseria meningitidis* (TLR2 agonist) as an adjuvant [140].

Wilms' tumor gene WT1 is a zing finger transcription factor overexpressed in leukemias and solid tumors [141]. Khalili et al. [142] identified 44 novel epitopes in WT1 protein and constructed a DNA vaccine containing the predicted epitopes with acceptable population coverage (>65%).

Prediction of immunogens for diagnostic tools

Leptospirosis is a global zoonotic disease affecting humans and causing severe icteric Weil's disease, characterized by renal and liver failure [143]. It is caused by spirochetes of the genus *Leptospira*. The pathogenesis-associated leptospiral LigA protein expressed *in vivo*, has been evaluated as a diagnosis of the acute form of leptospirosis. The C-terminal sequence of LigA (LigA-C) was cloned into pET15b and

expressed in *E. coli* [144]. Two B- cell-specific immunogenic epitopes have been predicted and synthesized as peptides for evaluation along with recombinant LigA-C. Selected B-cell epitopes showed increased sensitivity over recombinant LigA-C in single and combination assays for IgM antibody detection, and may be useful in early diagnosis of leptospirosis.

Hepatitis B virus (HBV) and human T lymphotropic virus type I (HTLV-I) are blood-borne viruses. Since HBV and HTLV-I infections are asymptomatic for a long time, people are typically unaware of infection [145,146]. Recombinant proteins from these viruses are used as capture antigens in ELISA blood screening tests [147,148]. A chimeric antigen comprising antigenic fragments of HBV core protein and proteins gp46 and p16 of HTLV-I was constructed, predicted by VaxiJen as antigenic, expressed in *E. coli*, purified and tested using serum from patients infected with HBV and/or HTLV-I [149]. The antibodies were detected successfully by the chimeric protein.

Shawky et al. [150] amplified the region encoding proteins E1 and E2 (HCV-E) from hepatitis C virus (HCV) genotype 4a, cloned it into a plasmid, and used this to immunize mice. The DNA construct was immunogenic, as predicted by VaxiJen. This study also found that combining the HCV-E construct with extracts from *Echinacea purpurea* and *Nigella sativa* prior to immunizing mice significantly increased both humoral and cellular responses.

VaxiJen has also been used to predict the antigenicity of cystatin C developed as a diagnostic tool for accurate estimation of glomerular filtration rate (GFR) [151]. A stable fusion protein was constructed to include immunogenic fragments of several proteins (ApoB-100, hHSP60 and β -2-GPI) associated with atherosclerosis [152]. It could be used as a vaccine to prevent or modulate atherosclerosis. Bioinformatic tools (including VaxiJen) were used to locate a specific conserved region of ActA, a membrane protein from *Listeria monocytogenes* (*L. monocytogenes*) [153]. The region was used to design an antibody-antigen based diagnostic test for *L. monocytogenes*.

Prediction of immunogens for veterinary medicine

Brucellosis is a zoonotic illness transmitted from domestic animals to humans. It is caused by *Brucella* spp. Although a live attenuated vaccine against ovine brucellosis exists, investigations focus on developing a safer subunit vaccine. Several chimeric DNA vaccines have been designed to encode proteins omp19, omp31 and urease [154], BP26, omp31 and TF [155], and GroEL [156]. Based on systematic screening of the exoproteome and secretome of *B. melitensis*, Vishnu et al. [157] identified eight proteins as potential vaccine candidates, including LPS-assembly protein LptD, a polysaccharide export protein, a cell surface protein, heme transporter BhuA, flagellin FliC, 7-alpha-hydroxysteroid dehydrogenase, immunoglobulin-binding protein EIBE, and hemagglutinin. Several B- and T-cell epitopes were predicted, and also from the protein omp25 [158]. All these proteins were assessed to be antigenic by VaxiJen.

Histophilus somni (*H. somni*) is an opportunistic bacterial pathogen causing histophilosis in cattle and is associated with thrombotic meningoencephalitis and bovine respiratory disease [159]. The genome of 12 *H. somni* isolates were sequencing, protein coding regions predicted, and several programs (including VaxiJen) used to evaluate the antigenicity, surface exposure scores, and sequence conservation [160]. The first 20 ranked proteins have been analyzed in western blot with bovine serum, with 13 responding to bovine antibodies of *H. somni*.

Pasteurella multocida (*P. multocida*) is an opportunistic bacterial pathogen causing fowl cholera in poultry, haemorrhagic septicaemia in cattle and buffalo, pneumonia in lambs and goats, respiratory atrophic rhinitis in swine, and purulent rhinitis in rabbits [161]. Ragavendhar et al. [161] characterized the outer membrane antigen 87 (oma87) of *P. multocida* from sheep. B- and T-cell epitopes were predicted and it was found that the peptide 638-652 might form an epitope-based vaccine. A comparative genome analysis of four pathogenic *P. multocida* strains with their respective hosts (fowl, goat and buffalo) identified several outer membrane proteins responsible for the pathogenicity [162]. Among them the lipopolysaccharide (LPS) assembly outer membrane complex protein contained the most B- and T-cell epitopes and is a suitable target for vaccine development against the two economic devastating diseases: fowl cholera and hemorrhagic septicemia.

Pajaroellobacter abortibovis (*P. abortibovis*) is a tick-transmitted infection causing late-term abortion in cattle or birth of weak calves [163]. As this bacterium cannot be grown in culture, DNA and RNA was extracted from spleen tissue collected from experimentally-infected mice [164]. *In silico* prediction of vaccine candidates was performed and the top 10 candidate proteins, as ranked by VaxiJen, were tested using serum from *P. abortibovis* immunized mice. This confirmed the antigenicity of seven of the nine proteins.

Leptospira interrogans (*L. interrogans*) causes leptospirosis in dogs and other animal species, including humans [165]. Natarajaseenivasan et al. [166] have compared the humoral immune responses to 4 recombinant proteins (LipL32, LigA, LK73.5 and GroEL) and one lipopolysaccharide (LPS) of *L. interrogans* in dogs vaccinated by a commercial multivalent vaccine containing leptospiral whole cell lysates of two serovars. The proteins and the predicted B-cell epitopes have been assessed as antigenic by VaxiJen. Leptospiral whole cell lysates and LPS have elicited higher level of antibody response compared to the single proteins.

Bordetella bronchiseptica (*B. bronchiseptica*) causes acute and chronic respiratory infection in a variety of animals [167]. Currently, there is no vaccine to prevent these infections. Liu et al. [167] analyzed five *B. bronchiseptica* antigens, as defined by VaxiJen, including amino acid ATP-binding cassette transporter substrate-binding protein (ABC), lipoprotein (PL), outer membrane porin protein (PPP), leu/ile/val-binding protein (BPP), and conserved hypothetical protein (CHP). The murine immune responses to individual recombinant proteins were measured, with each tested protein inducing high antibody titers. PPP and PL showed protection against challenges with *B. bronchiseptica*, while the protection by ABC, BPP, and CHP was not been significantly different from controls. PPP and PL have been identified as candidates for a diagnostic test or vaccine for *B. bronchiseptica*.

Anaplasma marginale (*A. marginale*) is a tick-borne bacterium causing anaplasmosis in cattle. The major surface protein 1a (MSP1a) has been analyzed as a vaccine candidate containing 4 B-cell epitopes [168]. *Mycoplasma agalactiae* (*M. agalactiae*) causes agalactia in small ruminants [169]. The most important surface antigen P40 has been analyzed [169]. Three B- and three T-cell epitopes were identified as appropriate and could find use in developing recombinant vaccines.

Corynebacterium pseudotuberculosis (*C. pseudotuberculosis*) causes caseous lymphadenitis in sheep and goats, and is responsible for significant economic losses [170]. The putative virulence proteins SpaC, SodC, NanH, and PknG have been analyzed using immunoinformatics (including VaxiJen). It was found that SpaC, PknG and NanH presented better vaccine potential than SodC. *Aeromonas hydrophila*

(*A. hydrophila*) causes aeromoniasis in fish [171]. Rauta et al. [172] analyzed two antigenic outer membrane protein (Aha1) peptides as DNA vaccine candidates including them in a nanoparticle-based delivery system.

Bovine rotavirus and bovine coronavirus are the most important causes of diarrhea in newborn calves as well as pigs and sheep [173]. Rotavirus protein VP8 and coronavirus S2 spike glycoprotein are the major determinants of viral infectivity and neutralization. A chimeric VP8-S2 gene was designed computationally using VaxiJen [173], cloned and sub-cloned into vectors, and transferred into *E. coli*. The expressed protein was purified and used to immunize hens. The activity and specificity of the isolated and purified anti-VP8- S2 IgY was detected using several experimental assays. The specific anti-VP8-S2 IgY was suggested as a candidate for passive immunization against bovine rotavirus and bovine coronavirus.

Fowl adenovirus serotype 4 (FAV-4) causes hydropericardium syndrome in domestic fowl. The hexon gene was isolated from 3 virus isolates and then cloned. The vectors were analyzed using online bioinformatics tools [174]. All were predicted to be antigens according VaxiJen and to contain B-cell epitopes.

Newcastle disease virus (NDV) causes Newcastle disease, an extremely infectious viral disease affecting most bird species. Motamedi et al. [175] used an *in silico* approach, assembling potential and conserved epitopic regions of hemagglutinin-neuraminidase (HN) and fusion (F) glycoproteins of NDV to induce multi-epitopic responses against the virus. Epitope predictions have showed that the hypothetical synthetic construct could induce immature B and T cell epitopes. Most regions of the construct had a high antigenic propensity and surface accessibility.

The cattle tick, *Rhipicephalus microplus*, is an obligate hematophagous ectoparasite of cattle occurring in the tropical and subtropical regions of the world and is a vector of disease causing pathogens, such as *Babesia bovis*, *Babesia bigemina* and *Anaplasma marginale* [176]. A systematic approach using a combination of functional genomics (DNA microarrays) techniques and a pipeline of *in silico* predictions of subcellular localization and protective antigenicity using VaxiJen was used to identify novel anti-tick candidate vaccines [177]. 791 candidates were identified, of which 176 were membrane-associated and 86 secreted soluble proteins. Five predicted membrane-associated antigenic proteins were selected and synthetic peptides designed and then tested using polyclonal antisera from BALB/c mice immunized with a crude extract of tick midgut membrane proteins. Three showed antigenicity higher than that of Bm86 peptides. Moreover, 19 novel transmembrane proteins were identified using LS-MS/MS and suggested as putative tick targets [178]. Aguirre et al. [179] identified using immunoinformatics an antigenic peptide from ATAQ protein, a putative Bm86 homolog. The pure peptide and its conjugate with Keyhole Limpet Hemocyanin (KLH) were tested for immunogenicity in mice, rabbits and cattle. Between 35% and 47% of the animals developed a consistent immune response.

Dogs are the domestic reservoirs of the parasite *Leishmania infantum* (*L. infantum*) causing visceral leishmaniasis (VL) in humans and dogs [180]. The control of canine VL could reduce human infection rates. Agallou et al. [180] analyzed the total protein extract from late-log phase *L. infantum promastigotes* using two-dimensional western blots and probing with sera from asymptomatic and symptomatic dogs. Forty-two protein spots were found to react differentially with IgG from asymptomatic dogs. Of these, 21 were identified by mass

spectrometry and predicted with ANTIGENpro and VaxiJen for antigenicity. Six proteins were identified as novel candidate antigens able to be developed as vaccines or diagnostic tests.

Other methods for immunogenicity prediction

Following in the wake of VaxiJen, several groups have sort to emulate it's success or to develop alternative approaches for the identification of whole protein antigens from microbial life. NERVE (New Enhanced Reverse Vaccinology Environment) was the first *in silico* pipeline for identification of the best vaccine candidates from whole proteomes of bacterial pathogens [181]. Reverse vaccinology is a method build on genome-based antigen discovery [182] and it has been successfully applied to the development of an innovative vaccine against *N. meningitidis* serogroup B [183]. NERVE uses four filters for selection of immunogens: localization, topology, probability of being adhesin and similarity to human proteins. The threshold values have been tuned on 10 proteomes containing known immunogens. As a result, NERVE ranks the most probable vaccine candidates.

Vaxign is the first web-based vaccine design system that predicts vaccine targets based on genome sequences using the strategy of reverse vaccinology [184]. Vaxign predicts protein subcellular location, transmembrane helices, adhesion probability, conservation to human and/or mouse proteins, sequence exclusion from genome(s) of nonpathogenic strain(s), and epitope binding to MHC class I and class II. The precomputed Vaxign database contains prediction of vaccine targets for >70 genomes. Vaxign also performs dynamic vaccine target prediction based on input sequences.

ANTIGENpro is a two-stage architecture based on structural features as sequence length, molecular weight, absolute charge per residue, turn-forming residues fraction, hydrophathy, etc. and support vector machine (SVM) classifier [131]. The algorithm has been trained on two datasets: (i) antigens that elicit a strong antibody response in protected individuals but not in unprotected individuals, using human immunoglobulin reactivity data obtained from protein microarray analyses; and (ii) known protective antigens from the literature. ANTIGENpro has correctly classified 82% of the known protective antigens when trained using only the protein microarray datasets. The accuracy on the combined dataset was 76% by cross-validation experiments. ANTIGENpro has performed well also on an external pathogen proteomes.

Vaccine has been built around the concept of linked resources [185]. It consists of two parts: part A builds the proteome using gene predictors and similarity searches; and part B predicts protein characteristics such as localization, topology, number of transmembrane helices, binding to MHC class I and class II. The output is a list of vaccine candidates ranked according to the average probability of the classifiers used.

Jenner-Predict server has been developed for prediction of protein vaccine candidates (PVCs) from proteomes of bacterial pathogens [186]. The server targets host-pathogen interactions and pathogenesis by considering known functional domains from protein classes such as adhesin, virulence, invasin, porin, flagellin, colonization, toxin, choline-binding, penicillin-binding, transferring-binding, fibronectin-binding and solute-binding. It predicts non-cytosolic proteins containing these domains as PVCs. It also provides vaccine potential of PVCs in terms of their possible immunogenicity by comparing with experimentally known IEDB epitopes, absence of autoimmunity and conservation in different strains. Predicted PVCs are prioritized so that only few prospective PVCs could be validated experimentally. The performance of web server has been evaluated against known protective antigens from diverse classes of bacteria.

iVAX is an integrated set of tools for triaging candidate antigens, selecting immunogenic and conserved T cell epitopes, eliminating regulatory T cell epitopes, and optimizing antigens for immunogenicity and protection against disease [187]. iVAX has been applied to vaccine development programs for emerging infectious diseases, cancer antigens and biodefense targets. Several iVAX vaccine design projects have had success in pre-clinical studies in animal models and are progressing toward clinical studies.

VacSol is a high throughput *in silico* pipeline for vaccine candidate prediction from bacterial pathogens [188]. It consists of known tools like BLAST [189], PSORT [190], HMMTOP [191], ABCPred [192], ProPred [193] and databases DEG 10 [194], VFDB [195] and UniProt [196], integrated to work consequently. It is freely available to download from <https://sourceforge.net/projects/vacsol/>.

Protectome analysis is a bioinformatics tool for discovery of bacterial vaccine candidates based on "protective signatures" [197]. Authors have collected a database of all known protective antigens of 38 bacterial pathogens, have analysed them using BLAST [189], ClustalW [198], Smart [199] and Pfam [200] and have identified common structural features like function/biological role (toxins, iron-uptake systems, adhesins, etc.) and/or structural organization (multiple internal structural motifs). Support vector machine (SVM) classification has been applied to derive a model for discrimination between bacterial protective antigens and non-antigens [201]. Initially, the training set consisted of 136 antigens and 136 non-antigens (most of them taken from VaxiJen datasets), lately the training set was increased to 200 antigens and 200 non-antigens [202]. The validation data showed a better performance of the SVM model comparing to VaxiJen. The model is not accessible online.

Discussion

The role that immunoinformatics plays in vaccine design devolves in several ways: Helping to design transgenic whole-organism pathogens, which cannot grow or cause harm, through designing out one or more virulence factors and so rendering the organism effectively harmless or by inducing severely compromised reproductive capacity of a virus or other microorganism [203]; Helping to design epitope ensemble vaccines. Such efforts fall into two camps: un-validated prediction-only methods that predict supposedly high-binding and more modern approaches that use immunoinformatics to select rather than predict the best epitopes suitable for forming a vaccine [204,205]; Helping to identify immunogenic and potentially protective single proteins from the genome of a given pathogenic microorganism. Such methodologies come in two main guises: "pipelines" or networks of methods and algorithms that together are able to select appropriate proteins [206,207] and single methods that seek to predict immunogens. Vaxijen typifies such an approach. Vaxijen is now a decade old. While so-called pipelines for vaccine design abound, methods that address the more fundamental questions of immunogenicity and antigenicity do not. VaxiJen stands alone; almost. Why is this? The main answer lies with the difficulty of the task – were it easy to improve on extant result then solutions would abound, as they have in other areas of immunoinformatics and bioinformatics - and the complex perception of the accuracy and veracity of the outcome. While the research objective of VaxiJen is both obvious and profound, current implementation is certainly sub-optimal. We need to improve and to test VaxiJen rigorously.

More specifically, we at least need to address the following: First, we need much more "positive" data; that is, carefully curated and validated,

examples of protective antigens from pathogenic microorganisms and cancer. Databases of such do exist - AntigenDB [206] is, for example, a dedicated resource directly addressing this, while IEDB [206] also contains similar data, but as something of an afterthought, rather than being its focus. IEDB concentrates on epitopes and related information not antigens and protein immunogens. What we need is more and better data resources of this type upon which to draw upon. Thus there is a need for a concerted effort to enlarge, deepen, and broaden available data collations; much as has been done for epitopes [207].

At the same time, we also need much better and much more carefully constructed negative training sets and learning protocols. We need to balance the selection of negative test sets so that any signal present in the analysis reflects antigenicity and no other quality. We need to select similar protein lengths, similar origin species, similar subcellular locations, and similar functions: the list goes on. Imagine we wanted to separate a particular evolutionarily-conserved family of membrane protein, a fair test would be other membrane proteins of similar length from the same species; not soluble or fibrous proteins selected from a distant branch of the tree-of-life.

Better representations of the sequence data: currently VaxiJen employs Wold's z-scales to characterise proteins using ACC transform. This works, but it is not clear that this is optimal. Other descriptors are available. Single descriptors characterising the whole sequence [208] and other multivariate descriptors of sequences. One could envisage a phase space of disjoint descriptor variables from which we could use variable selection protocols to select a compact and near-optimal choice of indicative variables.

Better algorithms: Artificial Intelligence or AI now is focused on the development of deep-learning protocols. Powerful machine learning toolkits, such as Weka [209], are already available, and these are more than capable of delivering robust and extensible methods provided the data and the data representation are adequate. Nonetheless, as new algorithms do appear we must be open and embrace them. Avoiding complacency must be our mantra.

More validation: better protocols for establishing the immunogenicity and recall of identified vaccines. This is the world of experimentalist, and here what is needed is a fast, straightforward methodology which can be used to give much more consistent and much more accurate estimates of individual proteins.

More generally, as with most computational studies of real world problems, there exists a pressing need for experimental validation. The publication of an ever increasing stockpile of papers relating to the *in silico* analyses of pathogen genomes and virtual proteomes, have generated many potential vaccine candidates. Such papers typically use methodology largely embodied in web-servers; operating such systems is facile, and the resulting analysis straightforward. We need to highlight the severe limitations of all such studies in the absence of proper experimental validation. Although many papers are technically sound, their utility is hard to quantify and their significance questionable. Likewise, more examples of experimentally validated Vaxijen predictions are needed. A majority of papers citing the use of Vaxijen do not contain any such verification. Publishing unverified papers ultimately becomes counterproductive. Other studies give credence to their computational results [18,41,44,149,150,160,164,166,167,173,179] by combining vaccine design with experimental validation in animal models. Even in the days of AI hysteria, prediction without validation exerts little influence and convinces few.

Reference

1. WHO (2015) Global vaccination targets 'off-track' warns.
2. Arnon R (2011) Overview of vaccine strategies. In: Vaccine design. Innovative approaches and novel strategies. Caister Academic Press, UK.
3. Rappuoli R, Miller HI, Falkow S (2002) Medicine. The intangible value of vaccination. *Science* 297: 937-939.
4. Davies MN, Flower DR (2007) Harnessing bioinformatics to discover new vaccines. *Drug Discov Today* 12: 389-395.
5. Vivona S, Gardy JL, Ramachandran S, Brinkman FSL, Raghava GPS, et al. (2008) Computer-aided biotechnology: from immune-informatics to reverse vaccinology. *Trends Biotechnol* 26: 190-200.
6. Liljeroos L, Malito E, Ferlenghi I, Bottomley MJ (2015) Structural and computational biology in the design of immunogenic vaccine antigens. *J Immunol Res* 156241.
7. Pappalardo F, Flower D, Russo G, Pennisi M, Motta S (2015) Computational modelling approaches to vaccinology. *Pharmacol Res* 92: 40-45.
8. Tanner R, McShane H (2017) Replacing, reducing and refining the use of animals in tuberculosis vaccine research. *Altex* 34: 157-166.
9. Flower DR, Davies MN, Doytchinova IA (2015) Identification of candidate vaccine antigens in silico. Springer, New York pp: 39-71.
10. Tomar N, De RK (2014) Immunoinformatics: a brief review. In: Immunoinformatics. *Methods Mol Biol* 1184: 23-55.
11. Doytchinova IA, Flower DR (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* 8: 4.
12. Hellberg S, Sjöström M, Skagerberg B, Wold S (1987) Peptide quantitative structure-activity relationships, a multivariate approach. *J Med Chem* 30: 1126-1135.
13. Wold S, Jonsson J, Sjöström M, Sandberg M, Rännar S (1993) DNA and peptide sequences and chemical processes multivariately modelled by principal component analysis and partial least squares projections to latent structures. *Anal Chim Acta* 277: 239-253.
14. Leardi R, Boggia R, Terrile M (1992) Genetic algorithms as a strategy for feature selection. *J Chemom* 6: 267-281.
15. Doytchinova IA, Flower DR (2007) Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties. *Vaccine* 25: 856-866.
16. Doytchinova IA, Flower DR (2008) Bioinformatic approach for identifying parasite and fungal candidate subunit vaccines. *Open Vaccine J* 1: 22-26.
17. Delfani S, Imani Fooladi AA, Mobarez AM, Emameini M, Amani J, et al. (2015) In silico analysis for identifying potential vaccine candidates against *Staphylococcus aureus*. *Clin Exp Vaccine Res* 4: 99-106.
18. Delfani S, Mobarez AM, Imani Fooladi AA, Amani J, Emameini M (2016) Protection of mice against *Staphylococcus aureus* infection by a recombinant protein C1fA-IsdB-Hlg as a vaccine candidate. *Med Microbiol Immunol* 205: 47-55.
19. Shilpa MD, Anil Kumar SK, Sulochana MB (2016) In silico analysis of clumping factors C1fA and C1fB of multidrug-resistant *Staphylococcus aureus* MRSA 252 strain. *Int J Adv Biotech Res* 7: 1103-1110.
20. Shahbazi M, Haghkhah M, Rahbar MR, Nezafat N, Ghasemi Y (2016) In silico sub-unit hexavalent peptide vaccine against an *Staphylococcus aureus* biofilm-related infection. *Int J Pept Res Ther* 22: 101-117.
21. Hajighahramani N, Nezafat N, Eslami M, Negahdaripour M, Rahmatabadi SS, et al. (2017) Immunoinformatics analysis and in silico designing of a novel multi-epitope peptide vaccine against *Staphylococcus aureus*. *Infect Genet Evol* 48: 83-94.
22. Falkenhorst G, Remschmidt C, Harder T, Hummers-Pradier E, Wichmann O, et al. (2017) Effectiveness of the 23-valent pneumococcal polysaccharide vaccine (PPV23) against pneumococcal disease in the elderly: systematic review and meta-analysis. *PLoS One* 12: e0169368.
23. Gentile A, Bazán V (2011) Prevention of pneumococcal disease through vaccination. *Vaccine* 29S: C15- C25.

24. Tarahomjoo S (2014) Bioinformatic analysis of surface proteins of *Streptococcus pneumoniae* serotype 19F for identification of vaccine candidates. *Am J Microbiol Res* 2: 174-177.
25. Tarahomjoo S (2015) Identification of B-cell Epitope Regions in Cell Surface Proteins of *Streptococcus pneumoniae* Serotype 19F Using Bioinformatic Tools. *Am Sci Res J Eng Technol Sci* 14: 107-117.
26. Abry MF, Kimenyi KM, Osowo FO, Odhiambo WO, Sewe SO, et al. (2015) Genetic diversity of the pneumococcal CbpA: implications for next-generation vaccine development. *Hum Vaccin Immunother* 11: 1261-1267.
27. Junqueira-Kipnis AP, Marques Neto LM, Kipnis A (2014) Role of fused *Mycobacterium tuberculosis* immunogens and adjuvants in modern tuberculosis vaccines. *Front Immunol* 5: 188.
28. Merly DP, Anil Kumar S, Mohan R, Tejaswini B, Darley SP, et al. (2012) In-silico analysis of PtpA—an antigenic protein of *Mycobacterium tuberculosis*. *Recent Res Sci Technol* 4: 23-26.
29. Monterrubio-López GP, González-Y-Merchand JA, Ribas-Aparicio RM (2015) Identification of novel potential vaccine candidates against tuberculosis based on reverse vaccinology. *BioMed Res Int* 2015: 483150.
30. Mushtaq K, Sheikh JA, Amir M, Khan N, Singh B, et al. (2015) Rv2031c of *Mycobacterium tuberculosis*: a master regulator of Rv2028–Rv2031 (HspX) operon. *Front Microbiol* 6: 351.
31. Saraav I, Kirti Pandey K, Sharma M, Singh S, Dutta P, et al. (2016) Predicting promiscuous antigenic T cell epitopes of *Mycobacterium tuberculosis* mymA operon proteins binding to MHC Class I and Class II molecules. *Infect Genet Evol* 44: 182-189.
32. Gazi A, Kibria MG, Mahfuz M, Rezaul I, Ghosh P, et al. (2016) Functional, structural and epitopic prediction of hypothetical proteins of *Mycobacterium tuberculosis* H37Rv: an in silico approach for prioritizing the targets. *Gene* 591: 442-455.
33. Usman MM, Salmah Ismail S, Teoh TC (2016) Protein expression of Myt272-3 recombinant clone and in silico prediction of a possible vaccine candidate against *Mycobacterium tuberculosis*. *Trop J Pharm Res* 15: 1393-1400.
34. Patwary NI, Islam MS, Sohel M, Ara I, Sikder MO, et al. (2016) In silico structure analysis and epitope prediction of E3 CR1-beta protein of Human Adenovirus E for vaccine design. *Biomed J* 39: 382-390.
35. Singh IG, Mukherjee R, Talwar GP, Kaufmann SH (1992) In vitro characterization of T cells from *Mycobacterium w*-vaccinated mice. *Infect Immun* 60: 257-263.
36. Gupta A, Geetha N, Mani J, Upadhyay P, Katoch VM, et al. (2009) Immunogenicity and protective efficacy of "Mycobacterium w" against *Mycobacterium tuberculosis* in mice immunized with live versus heat-killed M. w by the aerosol or parenteral route. *Infect Immun* 77: 223-231.
37. Nyasulu PS (2011) The role of adjunctive *Mycobacterium w* immunotherapy for tuberculosis. *J Exp Clin Med* 2: 124-129.
38. Saini V, Raghuvanshi S, Khurana JP, Ahmed N, Hasnain SE, et al. (2012) Massive gene acquisitions in *Mycobacterium indicus pranii* provide a perspective on mycobacterial evolution. *Nucl Acids Res* 40: 10832-10850.
39. Singh Y, Kohli S, Sowpati DT, Rahman SA, Tyagi AK, et al. (2014) Gene cooption in *Mycobacteria* and search for virulence attributes: Comparative proteomic analyses of *Mycobacterium tuberculosis*, *Mycobacterium indicus pranii* and other mycobacteria. *Int J Med Microbiol* 304: 742-748.
40. Zeinalzadeh N, Salmanian AH, Ahangari G, Sadeghi M, Amani J, et al. (2014) Design and characterization of a chimeric multi-epitope construct containing CfaB, heat-stable toxoid, Cssa, CcssB, and heat-labile toxin subunit B of enterotoxigenic *Escherichia coli*: a bioinformatic approach. *Biotech Appl Biochem* 61: 517-527.
41. Hayat SMG, Gargari SLM, Nazarian S (2016) Construction and immunogenic properties of a chimeric protein comprising CfaE, CfaB and LTb against Enterotoxigenic *Escherichia coli*. *Biologicals* 44: 503-510.
42. Nazarian S, Gargari SLM, Rasooli I, Amani J, Bagheri S, et al. (2012) An in silico chimeric multi subunit vaccine targeting virulence factors of enterotoxigenic *Escherichia coli* (ETEC) with its bacterial inbuilt adjuvant. *J Microbiol Methods* 90: 36-45.
43. Mehla K, Ramana J (2016) Identification of epitope-based peptide vaccine candidates against enterotoxigenic *Escherichia coli*: a comparative genomics and immunoinformatics approach. *Mol BioSyst* 12: 890-901.
44. García-Angulo VA, Kalita A, Kalita M, Lozano L, Torres AG (2014) Comparative Genomics and Immunoinformatics Approach for the Identification of Vaccine Candidates for Enterohemorrhagic *Escherichia coli* O157:H7. *Infect Immun* 82: 2016-2026.
45. Jeshvaghani FS, Rahjerdi AK, Amani J, Rad I, Jafari M, et al. (2016) Designing and structure evaluation of multi-epitope vaccine against ETEC and EHEC, an in silico approach. *Protein Pept Lett* 23: 33-42.
46. Hajizade A, Ebrahimi F, Amani J, Arpanaei A, Salmanian AH (2016) Design and in silico analysis of pentavalent chimeric antigen against three enteropathogenic bacteria: enterotoxigenic *E. coli*, enterohemorrhagic *E. coli* and *Shigella*. *Biosci Biotech Res Comm* 9: 225-239.
47. Bhairamadgi RN, Katti AKS (2013) In silico identification and sequence annotations of potential vaccine candidate in *N. gonorrhoeae*. *Int J Adv Biotech Res* 4: 404-414.
48. Barh D, Kumar A (2009) In silico identification of candidate drug and vaccine targets from various pathways in *Neisseria gonorrhoeae*. *In Silico Biol* 9: 225-231.
49. Merly DP, Kumar SA, Reddy KM, Tejaswini B, Darley SP, et al. (2012) In-silico analysis of PtpA - an antigenic protein of *Mycobacterium tuberculosis*. *Rec Res Sci Technol* 4: 23-26.
50. Barh D, Misra AN, Kumar A (2010) In silico identification of dual ability of *N. gonorrhoeae* ddl for developing drug and vaccine against pathogenic *Neisseria* and other human pathogens. *J Proteomics Bioinform* 3: 82-90.
51. Ravi Jain R, Sonkar SC, Chaudhry U, Bala M, Saluja D (2013) In-silico hierarchical approach for the identification of potential universal vaccine candidates (PUVCs) from *Neisseria gonorrhoeae*. *J Theor Biol* 410: 36-43.
52. Ali A, Naz A, Soares SC, Bakhtiar M, Tiwari S, et al. (2015) Pan-genome analysis of human gastric pathogen *H. pylori*: comparative genomics and pathogenomics approaches to identify regions associated with pathogenicity and prediction of potential core therapeutic targets. *BioMed Res Int* 2015: 139580.
53. Naz A, Awan FM, Obaid A, Muhammad SA, Paracha RZ, et al. (2015) Identification of putative vaccine candidates against *Helicobacter pylori* exploiting exoproteome and secretome: A reverse vaccinology based approach. *Infect Genet Evol* 32: 280-291.
54. Mohammad N, Karsabet MT, Amani J, Ardjmand A, Zadeh MR, et al. (2016) In Silico design of a chimeric protein containing antigenic fragments of *Helicobacter pylori*: a bioinformatic approach. *Open Microbiol J* 10: 97-112.
55. WHO (2016) *Campylobacter*. Fact sheet.
56. Mehla K, Ramana J (2015) Novel drug targets for food-borne pathogen *Campylobacter jejuni*: an integrated subtractive genomics and comparative metabolic pathway study. *J Integr Biol* 19: 393-406.
57. Meunier M, Nicodème MG, Hirschaud E, Parra A, Chemaly M, et al. (2016) Identification of novel vaccine candidates against *Campylobacter* through reverse vaccinology. *J Immunol Res* 2016: 5715790.
58. Mehla K, Ramana J (2017) Surface proteome mining for identification of potential vaccine candidates against *Campylobacter jejuni*: an in silico approach. *Funct Integr Genomics* 17: 27-37.
59. Fournier P, Richet H (2006) The epidemiology and control of *Acinetobacter baumannii* in health care facilities. *Clin Infect Dis* 42: 692-699.
60. Jain R, Danziger LH (2004) Multidrug-resistant *Acinetobacter* infections: an emerging challenge to clinicians. *Ann Pharmacother* 38: 1449-1459.
61. Rahbar MR, Rasooli I, Gargari SLM, Sandstrom G, Amani J, et al. (2012) A potential in silico antibody-antigen based diagnostic test for precise identification of *Acinetobacter baumannii*. *J Theor Biol* 294: 29-39.
62. Taiwo AA, Fallilat AJ, Ezemuel YS (2014) Computational design of peptide vaccine against *Acinetobacter baumannii* infection using comparative genomic approach. *Comput Biol Bioinf* 2: 13-18.
63. Sefid F, Iraj Rasooli I, Jahangiri A, Bazmara H (2015) Functional exposed amino acids of BauA as potential immunogen against *Acinetobacter baumannii*. *Acta Biotheor* 63: 129-149.
64. Hassan A, Naz A, Obaid A, Paracha RZ, Naz K, et al. (2016) Pangenome and immuno-proteomics analysis of *Acinetobacter baumannii* strains revealed the core peptide vaccine targets. *BMC Genomics* 17: 732.

65. Hosseingholi EZ, Rasooli I, Gargari SLM (2014) In silico analysis of *Acinetobacter baumannii* phospholipase D as a subunit vaccine candidate. *Acta Biotheor* 62: 455-478.
66. Barh D, Barve N, Gupta K, Chandra S, Jain N, et al. (2013) Exoproteome and secretome derived broad spectrum novel drug and vaccine candidates in *Vibrio cholera* targeted by piper betel derived compounds. *PLoS One* 8: e52773.
67. Nezafata N, Karimia Z, Eslamia M, Mohkama M, Zandiana S, et al. (2016) Designing an efficient multi-epitope peptide vaccine against *Vibrio cholerae* via combined immunoinformatics and protein interaction based approaches. *Comput Biol Chem* 62: 82-95.
68. Rauta PR, Ashe S, Nayak D, Nayak B (2016) In silico identification of outer membrane protein (Omp) and subunit vaccine design against pathogenic *Vibrio cholera*. *Comput Biol Chem* 65: 61-68.
69. Prabhavathy K, Perumal P, Baalaji NS (2011) In silico identification of B- and T-cell epitopes on OMPLA and LsrC from *Salmonella typhi* for peptide-based subunit vaccine design. *Ind J Biotechnol* 10: 440-451.
70. Toobak H, Rasooli I, Talei D, Jahangiri A, Owlia P, et al. (2013) Immune response variations to *Salmonella enterica* serovar Typhi recombinant porin proteins in mice. *Biologicals* 41: 224-230.
71. Levett PN, Branch SL, Whittington CU, Edwards CN, Paxton H (2001) Two methods for rapid serological diagnosis of acute leptospirosis. *Clin Diagn Lab Immunol* 8: 349-351.
72. Victor AAR, Abraham S, Chinniah MN, Madathiparambil MG, Tennyson J (2014) Phylogenetic characterization and threading based-epitope mapping of Leptospiral outer membrane lipoprotein LipL41. *J Proteomics Bioinform* 7: 222-231.
73. Malipatil V, Madagi S, Bhattacharjee B (2011) Subtractive genomics approach for in silico identification of novel drug targets and epitopes for vaccine design in *Treponema pallidum* subsp. *pallidum* str. Nichols. *Int J Pharm Sci Res* 2: 1855-1859.
74. Sharma D, Patel S, Padh H, Desai P (2016) Immunoinformatic identification of potential epitopes against shigellosis. *Int J Pept Res Ther* 22: 481-495.
75. Chirani AS, Majidzadeh R, Dabiri H, Rezaei J, Esmaili A, et al. (2016) Immunological study on integrated PilQ and disulphide loop region of PilA against acute *Pseudomonas aeruginosa* infection: In silico analysis and in vitro production. *J Acute Dis* 5: 131-142.
76. Mishra P, Velavan S (2013) Prediction of peptide vaccine candidate for dairy products borne human botulism. *Int J Pharm Sci Rev Res* 21: 158-162.
77. Gupta SK, Srivastava M, Akhoun BA, Smita S, Schmitz U, et al. (2011) Identification of immunogenic consensus T-cell epitopes in globally distributed influenza-A H1N1 neuraminidase. *Infect Genet Evol* 11: 308-319.
78. Moattari A, Dehghani B, Khodadad N, Tavakoli F (2015) In silico functional and structural characterization of H1N1 influenza A viruses hemagglutinin, 2010–2013, Shiraz, Iran. *Acta Biotheor* 63: 183-202.
79. Xiao X, Wu L, Stantchev TS, Feng YR, Ugolini S, et al. (1999) Constitutive cell surface association between CD4 and CCR5. *Proc Natl Acad Sci USA* 96: 7496-7501.
80. Rao BS, Gupta KK, Kumari S, Gupta A, Pujitha K (2013) Conserved HIV wide spectrum antipeptides – A hope for HIV treatment. *Adv Tech Biol Med* 1: 1000102.
81. WHO (2016) Hepatitis C. Fact sheet.
82. Idrees S, Ashfaq UA (2013) Structural analysis and epitope prediction of HCV E1 protein isolated in Pakistan: an in-silico approach. *Viro J* 10: 113.
83. Ikram A, Anjum S, Tahir M (2014) In silico identification and conservation analysis of B-cell and T-cell epitopes of hepatitis C virus 3a genotype enveloped glycoprotein 2 from Pakistan: a step towards heterologous vaccine design. *Hepat Mon* 14: e9832.
84. Afzal S, Idrees M, Hussain M (2014) De Novo modeling of Envelope 2 protein of HCV isolated from Pakistani patient and epitopes prediction for vaccine development. *J Transl Med* 12: 115.
85. Ikram A, Obaid A, Awan FM, Hanif R, Naz A, et al. (2017) Identification of drug resistance and immune-driven variations in hepatitis C virus (HCV) NS3/4A, NS5A and NS5B regions reveals a new approach toward personalized medicine. *Antivir Res* 137: 112-124.
86. Dick GW, Kitchen SF, Hadow AJ (1952) Zika virus. I. Isolations and serological specificity. *Trans R Soc Trop Med Hyg* 46: 509-520.
87. Faye O, Freire CCM, Iamarino A, Faye O, Oliveira JVC, et al. (2014) Molecular evolution of Zika virus during its emergence in the 20th century. *PLoS Negl Trop Dis* 8: e2636.
88. Shawan MAM, Mahmud HA, Hasan MM, Parvin A, Rahman MN, et al. (2014) In silico modeling and immunoinformatics probing disclose the epitope based peptide vaccine against Zika virus envelope glycoprotein. *Indian J Pharm Biol Res* 2: 44-57.
89. Ashfaq UA, Ahmed B (2016) De novo structural modeling and conserved epitopes prediction of Zika virus envelop protein for vaccine development. *Viral Immunol* 29: 436-443.
90. Alam A, Ali S, Ahamad S, Malik MZ, Ishrat R (2016) From ZikV genome to vaccine: in silico approach for the epitope based peptide vaccine against Zika virus envelope glycoprotein. *Immunology* 149: 386-399.
91. Janahia EM, Dhasmanab A, Srivastavad V, Sarangie AN, Razac S, et al. (2017) In silico CD4+, CD8+ T-cell and B-cell immunity associated immunogenic epitope prediction and HLA distribution analysis of Zika virus. *EXCLI J* 16: 63-72.
92. Dar H, Zaheer T, Rehman MT, Ali A, Javed A, et al. (2016) Prediction of promiscuous T-cell epitopes in the Zika virus polyprotein: An in silico approach. *Asian Pac J Trop Med* 9: 844-850.
93. Norrby R (2014) Outlook for a dengue vaccine. *Clin Microbiol Infect* 20: 92-94.
94. Fatima Z, Idrees M, Bajwa MA, Tahir Z, Ullah O, et al. (2011) Serotype and genotype analysis of dengue virus by sequencing followed by phylogenetic analysis using samples from three mini outbreaks- 2007-2009 in Pakistan. *BMC Microbiol* 11: 200.
95. Kuno G, Chang G-JJ, Tsuchiya KR, Karabatsos N, Cropp CB (1998) Phylogeny of the genus *Flavivirus*. *J Virol* 72: 73-83.
96. Baskar V, Madhan R, Srinivasan G, Selvakumar K, Radha M, et al. (2011) Identifying the potential tetravalent vaccine candidate for Dengue virus using in silico approach. *Insight Bioinf* 1: 6-10.
97. Hussain M, Idrees M, Afzal S (2015) Development of global consensus of Dengue virus envelope glycoprotein for epitopes based vaccine design. *Curr Comp Aided Drug Des* 11: 84-97.
98. Venkatachalam R, Subramanian V (2015) Antigenicity of envelop and non-structural proteins of dengue serotypes and their potentiality to elicit specific antibody. *Asian Pac J Trop Des* 5: S47-S50.
99. Shi J, Sun J, Wu M, Hu N, Li J, et al. (2015) Inferring protective CD8+ T-cell epitopes for NS5 protein of four serotypes of Dengue virus Chinese isolates based on HLA-A, -B and -C allelic distribution: implications for epitope-based universal vaccine design. *PLoS ONE* 10: e0138729.
100. Feldmann H, Geisbert TW (2011) Ebola haemorrhagic fever. *Lancet* 377: 849-862.
101. Leroy EM, Kumulungui B, Pourrut X, Rouquet P, Hassanin A, et al. (2005) Fruit bats as reservoirs of Ebola virus. *Nature* 438: 575-576.
102. Khan MA, Hossain MU, Rakib-Uz-Zaman SM, Morshed MN (2015) Epitope-based peptide vaccine design and target site depiction against Ebola viruses: an immunoinformatics study. *Scand J Immunol* 82: 25-34.
103. Srivastava PN, Jain R, Dubey SD, Bhatnagar S, Ahmad N (2016) Prediction of epitope-based peptides for vaccine development from coat proteins GP2 and VP24 of Ebola virus using immunoinformatics. *Int J Pept Res Ther* 22: 119-133.
104. Lacy MD, Smego R (1995) Viral hemorrhagic fevers. *Adv Pediatr Infect Dis* 12: 21-53.
105. Oany AR, Ahmad SAI, Hossain MU, Jyoti TP (2015) Identification of highly conserved regions in L-segment of Crimean–Congo hemorrhagic fever virus and immunoinformatic prediction about potential novel vaccine. *Adv Appl Bioinform Chem* 8: 1-10.
106. Tipu HN (2016) Immunoinformatic analysis of Crimean Congo hemorrhagic fever virus glycoproteins and epitope prediction for synthetic peptide vaccine. *J Coll Physicians Surg Pak* 26: 108-112.
107. Oany AR, Emran AA, Jyoti TP (2014) Design of an epitope-based peptide vaccine against spike protein of human coronavirus: an in silico approach. *Drug Des Devel Ther* 8: 1139-1149.

108. Shi J, Zhang J, Li S, Sun J, Teng Y, et al. (2015) Epitope-based vaccine target screening against highly pathogenic MERS-CoV: an in silico approach applied to emerging infectious diseases. *PLoS ONE* 10: e0144475.
109. Thiboutot MM, Kannan S, Kawalekar OU, Shedlock DJ, Khan AS, et al. (2010) Chikungunya: a potentially emerging epidemic? *PLoS Negl Trop Dis* 4: e623.
110. Hasan A, Khan A, Datta A, Md. Mazumder HH, Hossain MU (2015) A comprehensive immunoinformatics and target site study revealed the cornerstone toward Chikungunya virus treatment. *Mol Immunol* 65: 189-204.
111. Priyanka Kori P, Sajjan SS, Madagi SB (2015) In silico prediction of epitopes for Chikungunya viral strains. *J Pharm Invest* 45: 579-591.
112. Hasan A, Hossain M, Alam J (2013) A computational assay to design an epitope-based peptide vaccine against Saint Louis encephalitis virus. *Bioinf Biol Insights* 7: 347-355.
113. Sakib MS, Islam R, Mahbub Hasan AKM, Nurun Nabi AHM (2014) Prediction of epitope-based peptides for the utility of vaccine development from fusion and glycoprotein of Nipah virus using in silico approach. *Adv Bioinf* 2014: 402492.
114. Rahman M, Parvege M (2014) In silico structural analysis of Hantaan virus glycoprotein G2 and conserved epitope prediction for vaccine development. *J Appl Virol* 3: 62-77.
115. Sabbaghian E, Roodbari F, Rafiei A, Amani J (2014) In Silico design of a multimeric polytope as a highly immunogenic DNA vaccine against human cytomegalovirus. *J Appl Biotech Rep* 1: 143-153.
116. Verma SK, Yadav S, Kumar A (2015) In silico prediction of B- and T-cell epitope on Lassa virus proteins for peptide based subunit vaccine design. *Adv Biomed Res* 4: 201.
117. Yasmin T (2015) Prediction of B and T cell epitope-based peptide vaccines from highly conserved regions in Enterovirus D68 capsid protein VP1: a computational approach. *J Innov Pharm Biol Sci* 2: 426-438.
118. Kumar V, Damodharan S, Pandaranayaka EP, Madathiparambil MG, Tennyson J (2016) Molecular modeling and in silico engineering of Cardamon mosaic virus coat protein for the presentation of immunogenic epitopes of Leptospira LipL32. *J Biomol Struct Dyn* 34: 42-56.
119. Parvege M, Rahman M, Nibir YM, Hossain MS (2016) Two highly similar LAEDDTNAQKT and LTDKIGTEI epitopes in G glycoprotein may be useful for effective epitope based vaccine design against pathogenic Henipavirus. *Comput Biol Chem* 61: 270-280.
120. Mostafa SM, Islam ABMMK (2016) An in silico approach predicted potential therapeutics that can confer protection from maximum pathogenic Hantaviruses. *Fut Virol* 11: 411-428.
121. Hossain MU, Hashem A, Keya CA, Salimullah M (2016) Therapeutics insight with inclusive immunopharmacology explication of Human Rotavirus A for the treatment of diarrhea. *Front Pharmacol* 7: 153.
122. Sahrawat TR, Kaur A (2016) In-silico design of an epitope-based peptide vaccine: A computational biology approach. *Int J Comput Biol* 5: 1-5.
123. Lissabet JFB (2016) Computational prediction of linear B-cell epitopes in the E5 oncoprotein of the human papillomavirus type 16 using several bioinformatics tools. *Vacunas* 17: 18-26.
124. WHO (2016) Malaria. Fact sheet: World Malaria Report.
125. Doolan DL, Aguiar JC, Weiss WR, Sette A, Felgner PL, et al. (2003) Utilization of genomic sequence information to develop malaria vaccines. *J Exp Biol* 206: 3789-3802.
126. Singh SP, Khan F, Mishra BN (2010) Computational characterization of Plasmodium falciparum proteomic data for screening of potential vaccine candidates. *Hum Immunol* 71: 136-143.
127. Ajao AT, Adekeye AO, Yakubu SE (2014) Application of reverse vaccinology in designing a vaccine for malaria. *Am J Bioinf Res* 4: 6-10.
128. Singh SP, Verma V, Mishra BN (2015) Characterization of Plasmodium falciparum proteome at asexual blood stages for screening of effective vaccine candidates: an immunoinformatics approach. *Immunol Immunogen Insights* 7: 21-30.
129. Dumonteil E (2007) DNA vaccines against protozoan parasites: advances and challenges. *J Biomed Biotech* 2007: 90520.
130. Singh SP, Roopendra K, Mishra BN (2015) Genome-wide prediction of vaccine candidates for Leishmania major: an integrated approach. *J Trop Med* 2015: 709216.
131. Magnan CN, Zeller M, Kayala MA, Vigil A, Randall A, et al. (2010) High-throughput prediction of protein antigenicity using protein microarray data. *Bioinformatics* 26: 2936-2943.
132. Hill D, Dubey J (2002) Toxoplasma gondii: transmission, diagnosis and prevention. *Clin Microbiol Infect* 8: 634-640.
133. Hasheminasab SS, Maghsood H, Khalili S (2016) Immunoinformatic prediction about potential novel vaccine in surface antigen fragment protein of Toxoplasma gondii. *Int J Health* 4: 1-5.
134. Wingfield M, Hammerbacher A, Ganley R, Steenkamp E, Gordon T, et al. (2008) Pitch canker caused by Fusarium circinatum: a growing threat to pine plantations and forests worldwide. *Australas Plant Pathol* 37: 319-334.
135. Maphosa MN, Steenkamp ET, Wingfield BD (2016) Genome-based selection and characterization of Fusarium circinatum-specific sequences. *G3 (Bethesda)* 6: 631-639.
136. Mellman I, Coukos G, Dranoff G (2011) Cancer immunotherapy comes of age. *Nature* 480: 480-489.
137. Nezafat N, Ghasemi Y, Javadi G, Khoshnoud MJ, Omidinia E (2014) A novel multi-epitope peptide vaccine against cancer: An in silico approach. *J Theor Biol* 349: 121-134.
138. Zhang S, Chen L, Cui B, Chuang HY, Yu J, et al. (2012) ROR1 is expressed in human breast cancer and associated with enhanced tumor-cell growth. *PLoS One* 7: e31127.
139. Fooladi AAI, Hosseini HM, Amani J (2015) An In silico chimeric vaccine targeting breast cancer containing inherent adjuvant. *Iran J Cancer Preven* 8: e2326.
140. Mahmoodi S, Nezafat N, Berzegar A, Negahdaripour M, Nikanfar AR, et al. (2016) Harnessing bioinformatics for designing a novel multi-epitope peptide vaccine against breast cancer. *Curr Pharm Biotechnol* 17: 1100-1114.
141. Sugiyama H (2010) WT1 (Wilms'tumorgene1): biology and cancer immunotherapy. *Jpn J Clin Oncol* 40: 377-387.
142. Khalili S, Rahbar MR, Dezfulian MH, Jahangiri A (2015) In silico analyses of Wilms' tumor protein to designing a novel multi-epitope DNA vaccine against cancer. *J Theor Biol* 379: 66-78.
143. WHO (2012) Leptospirosis. Fact sheet.
144. Kanagavel M, Shanmughapriya S, Anbarasu K, Natarajaseenivasana K (2014) B-cell-specific peptides of Leptospira interrogans LigA for diagnosis of patients with acute leptospirosis. *Clin Vaccine Immunol* 21: 354-359.
145. WHO (2017) Hepatitis B. Fact sheet.
146. Proietti FA, Carneiro-Proietti ABF, Catalan-Soares BC, Murphy EL (2005) Global epidemiology of HTLV-I infection and associated diseases. *Oncogene* 24: 6058-6068.
147. Gallo D, Penning LM, Hanson C (1991) Detection and differentiation of antibodies to human T-cell lymphotropic virus types I and II by the immunofluorescence method. *J Clin Microbiol* 29: 2345-2347.
148. Wang JT, Lee CZ, Chen PJ, Wang TH, Chen DS (2002) Transfusion transmitted HBV infection in an endemic area: the necessity of more sensitive screening for HBV carriers. *Transfusion* 42: 1592-1597.
149. Zarnagh HH, Ravanshad M, Pourfatollah AA, Rasaei MJ (2015) Expression and purification of a novel computationally designed antigen for simultaneously detection of HTLV-1 and HBV antibodies. *Iran J Allergy Asthma Immunol* 14: 168-178.
150. Shawky H, Maghraby AS, Solliman MED, El-Mokadem MT, Sherif MM, et al. (2015) Expression, immunogenicity and diagnostic value of envelope proteins from an Egyptian hepatitis C virus isolate. *Arch Virol* 160: 945-958.
151. Abdollah Z, Karami A, Amani J, Khodi S, Gheybi E (2015) In silico study of cystatin C protein to develop an ELISA kit using computational tools and servers. *Bull Env Pharma Life Sci* 4: 87-94.
152. Karkhah A, Amani J (2016) A potent multivalent vaccine for modulation of immune system in atherosclerosis: an in silico approach. *Clin Exp Vaccine Res* 5: 50-59.

153. Jahangiri A, Rasooli I, Rahbar MR, Khalili S, Amani J, et al. (2012) Precise detection of *L. monocytogenes* hitting its highly conserved region possessing several specific antibody binding sites. *J Theor Biol* 305: 15-23.
154. Abkar M, Lotfi AS, Amani J, Ghorashi SA, Brujeni GN, et al. (2014) Design of a chimeric DNA vaccine against *Brucella* spp. *Minerva Biotec* 26: 223-233.
155. Ghasemi A, Ranjbar R, Amani J (2014) In silico analysis of chimeric TF, Omp31 and BP26 fragments of *Brucella melitensis* for development of a multi subunit vaccine candidate. *Iran J Basic Med Sci* 17: 172-180.
156. Sekhavati MH, Heravi RM, Tahmoorespur M, Yousefi S, Abbassi-Dalooi T, et al. (2015) Cloning, molecular analysis and epitopes prediction of a new chaperone GroEL *Brucella melitensis* antigen. *Iran J Basic Med Sci* 18: 499-505.
157. Vishnu US, Sankarasubramanian J, Gunasekaran P, Rajendhran J (2015) Novel vaccine candidates against *Brucella melitensis* identified through reverse vaccinology approach. *OMICS* 19: 722-729.
158. Yousefi S, Tahmoorespur M, Sekhavati MH (2015) B and T-cell epitope prediction of the OMP25 antigen for developing *Brucella melitensis* vaccines for sheep. *Iran J Appl Anim Sci* 5: 629-638.
159. Corbeil LB, Widders PR, Gogolewski R, Arthur J, Inzana TJ, et al. (1986) *Haemophilus somnus*: bovine reproductive and respiratory disease. *Can Vet J* 27: 90-93.
160. Madampage CA, Rawlyk N, Crockford G, Wang Y, White AP, et al. (2015) Reverse vaccinology as an approach for developing *Histophilus somni* vaccine candidates. *Biologicals* 43: 444-451.
161. Ragavendhar K, Prabhakar P, Thangavelu A, Kirubaharan JJ, Chandran NDJ. (2015) Sequence analysis, 3- dimensional protein modelling and epitope prediction of Oma87 gene originating from *Pasteurella multocida* isolate of sheep. *J Inf Mol Biol* 3: 1-12.
162. Ghaffar A, Tariq A. In-silico analysis of *Pasteurella multocida* to identify common epitopes between fowl, goat and buffalo. (2016) *Gene* 580: 58-66.
163. Howarth JA, Moulton JE, Frazier LM9 (1956) Epizootic bovine abortion characterized by fetal hepatopathy. *J Am Vet Med Assoc* 128: 441-44.
164. Welly BT, Miller MR, Stott JL, Blanchard MT, Islas-Trejo AD, et al. (2017) Genome report: Identification and validation of antigenic proteins from *Pajaroellobacter abortibovis* using de novo genome sequence assembly and reverse vaccinology. *G3 (Bethesda)* 7: 321-331.
165. Faine S. *Leptospira* and *Leptospirosis* (1994) CRC Press, Boca Raton,
166. Natarajaseenivasan K, Priya CL, Vanithamani S, Shanmughapriya S, Anandhagiri S, et al. (2012) Humoral immune response of inactivated bivalent *Leptospira* vaccine among dogs in Tiruchirappalli, Tamilnadu, India. *World J Vac*. 85-90.
167. Liu Y, Chen H, Wei Q, Xiao C, Ji Q, et al. (2015) Immune efficacy of five novel recombinant *Bordetella bronchiseptica* proteins. *BMC Vet Res* 11: 173.
168. Cabezas-Cruz A, Passos LMF, Lis K, Kenneil R, Valdés JJ, et al. (2013) Functional and immunological relevance of *Anaplasma marginale* major surface protein 1a sequence and structural analysis. *PLoS ONE* 8: e65243.
169. Forouharmehr A, Nassiry MR (2015) B and T-cell epitopes prediction of the P40 antigen for developing *Mycoplasma Agalactiae* vaccine using bioinformatic tools. *G3M* 13: 3954-3961.
170. Santana- Jorge KTO, Santos TM, Tartaglia NR, Aguiar EL, Souza RFS, et al. (2016) Putative virulence factors of *Corynebacterium pseudotuberculosis* FRC41: vaccine potential and protein expression. *Microb Cell Fact* 15: 83.
171. Austin B, Austin DA, Dalsgaard I, Gudmundsdottir BK, Høie S, et al. (1998) Characterization of atypical *Aeromonas salmonicida* by different methods. *Syst Appl Microbiol* 21: 50-64.
172. Rauta PR, Nayak B, Monteiro GA, Mateus M (2017) Design and characterization of plasmids encoding antigenic peptides of Aha1 from *Aeromonas hydrophila* as prospective fish vaccines. *J Biotechnol* 241: 116-126.
173. Nasiri K, Nassiri MR, Tahmoorespur M, Haghparast A, Zibae S (2016) Production and characterization of egg yolk antibody (IgY) against recombinant VP8-S2 antigen. *Pol J Vet Sci* 19: 271-279.
174. Asthana M, Singh VK, Kumar R, Chandra R (2011) Isolation, Cloning and in silico study of hexon gene of fowl adenovirus 4 (FAV4) isolates associated with hydropericardium syndrome in domestic fowl. *J Proteomics Bioinform* 4: 190-195.
175. Motamedi MJ, Amani J, Shahsavandi S, Salmanian AH (2014) In silico design of multimeric HN-F antigen as a highly immunogenic peptide vaccine against Newcastle disease virus. *Int J Pept Res Ther* 20: 179- 194.
176. de Castro JJ (1997) Sustainable tick and tickborne disease control in livestock improvement in developing countries. *Vet Parasitol* 71: 77-97.
177. Maritz-Olivier C, van Zyl W, Stutzer C (2012) A systematic, functional genomics, and reverse vaccinology approach to the identification of vaccine candidates in the cattle tick, *Rhipicephalus microplus*. *Tick Tick Borne Dis* 3: 179-187.
178. Richards SA, Stutzer C, Bosman AM, Maritz-Oliviera C (2006) Transmembrane proteins – Mining the cattle tick transcriptome. *Ticks Tick Borne Dis* 6: 695-710.
179. Aguirrea AAR, Lobo FP, Cunha RC, Garcia MV, Andreotti R (2006) Design of the ATAQ peptide and its evaluation as an immunogen to develop a *Rhipicephalus* vaccine. *Vet Parasitol* 221: 30-38.
180. Agallou M, Athanasiou E, Samiotaki M, Panayotou G, Karagouni E (2016) Identification of immunoreactive *Leishmania infantum* protein antigens to asymptomatic dog sera through combined immunoproteomics and bioinformatics analysis. *PLoS ONE* 11: e0149894.
181. Vivona S, Bernante F, Filippini F (2006) NERVE: New Enhanced Reverse Vaccinology Environment. *BMC Biotechnol* 6: 35.
182. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, et al. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269:496-512.
183. Pizza M, Scarlato V, Masignani V, Giuliani MM, Aricó B, et al. (2000) Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. *Science* 287: 1816-1820.
184. He Y, Xiang Z, Mobley HLT (2010) Vaxign: The first web-based vaccine design program for reverse vaccinology and applications for vaccine development. *J Biomed Biotechnol* 2010: 297505.
185. Goodswen SJ, Kennedy PJ, Ellis JT (2014) Vaxceed: a high-throughput in silico vaccine candidate discovery pipeline for eukaryotic pathogens based on reverse vaccinology. *Bioinformatics* 30: 2381-2383.
186. Jaiswal V, Chanumolu SK, Gupta A, Chauhan RS, Rout C (2013) Jenner-predict server: prediction of protein vaccine candidates (PVCs) in bacteria based on host-pathogen interactions. *BMC Bioinformatics* 14: 211.
187. Moise L, Gutierrez A, Kibria F, Martin R, Tassone R, et al. (2015) iVAX: An integrated toolkit for the selection and optimization of antigens and the design of epitope-driven vaccines. *Hum Vac Immunother* 11: 2312-2321.
188. Rizwan M, Naz A, Ahmad J, Naz K, Obaid A, et al. (2017) VacSol: a high throughput in silico pipeline to predict potential therapeutic targets in prokaryotic pathogens using subtractive reverse vaccinology. *BMC Bioinformatics* 18: 106.
189. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, et al. (2009) BLAST+: architecture and applications. *BMC Bioinformatics* 10: 421.
190. Gardy JL, Laird MR, Chen F, Rey S, Walsh CJ, et al. (2005) PSORTb v. 2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. *Bioinformatics* 21: 617-623.
191. Tusnady GE, Simon I (2001) The HMMTOP transmembrane topology prediction server. *Bioinformatics* 17: 849-850.
192. Saha S, Raghava G (2006) Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* 65: 40-48.
193. Singh H, Raghava G (2003) ProPred1: prediction of promiscuous MHC Class-I binding sites. *Bioinformatics* 19: 1009-1014.
194. Luo H, Lin Y, Gao F, Zhang CT, Zhang R, et al. (2013) DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements. *Nucleic Acids Res* 42: D574- D580.
195. Chen L, Yang J, Yu J, Yao Z, Sun L, et al. (2005) VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res* 33: D325-D328.
196. Consortium U. The universal protein resource (UniProt) (2008) *Nucleic Acids Res* 36: D190-D195.
197. Altindis E, Cozzi R, Di Palo B, Necchi F, Mishra RP, et al. (2015) Protectome analysis: a new selective bioinformatics tool for bacterial vaccine candidate discovery. *Mol Cell Proteomics* 14: 418- 429.

198. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680.
199. Letunic I, Doerks T, Bork P (2012) SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res* 40: D302-D305.
200. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, et al. (2014) Pfam: the protein families database. *Nucleic Acids Res.* 42: D222-D230.
201. Bowman BN, McAdam PR, Vivona S, Zhang JX, Luong T, et al (2011) Improving reverse vaccinology with a machine learning approach. *Vaccine* 29: 8156-8164.
202. Heinson AI, Gunawardana Y, Moesker B, Hume CCD, Vataga E, et al. (2017) Enhancing the biological relevance of machine learning classifiers for reverse vaccinology. *Int J Mol Sci* 18: E312.
203. Lim KI, Lang T, Lam V, Yin J (2006) Model-based design of growth-attenuated viruses. *PLoS Comput Biol* 2: e116.
204. Molero-Abraham M, Lafuente EM, Flower DR, Reche PA (2013) Selection of conserved epitopes from hepatitis C virus for pan-population stimulation of T-cell responses. *Clin Dev Immunol* 2013: 601943.
205. Sheikh QM, Gatherer D, Reche PA, Flower DR (2016) Towards the knowledge-based design of universal influenza epitope ensemble vaccines. *Bioinformatics* 32: 3233-3239.
206. Ansari HR, Flower DR, Raghava GP (2010) AntigenDB: an immunoinformatics database of pathogen antigens. *Nucleic Acids Res* 38: D847-D853.
207. Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, et al. (2015) The immune epitope database (IEDB) 3.0. *Nucleic Acids Res* 43: D405- D412.
208. Chattopadhyay AK, Nasiev D, Flower DR (2015) A statistical physics perspective on alignment-independent protein sequence comparison. *Bioinformatics* 31: 2469-2474.
209. Frank E, Hall MA, Witten IH (2006) The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Kaufmann M, Ed, Fourth Edition, 2016.