Journal of Biomedical Informatics xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Methodological Review

An overview of bioinformatics tools for epitope prediction: Implications on vaccine development

Ruth E. Soria-Guerra [a], Ricardo Nieto-Gomez [b], Dania O. Govea-Alonso [b], Sergio Rosales-Mendoza [b,*]

[a] Laboratorio de Ingeniería de Biorreactores, Facultad de Ciencias Químicas, Universidad Autónoma de San Luis Potosí, Av. Dr. Manuel Nava 6, SLP 78210, Mexico
[b] Laboratorio de Biofarmacéuticos Recombinantes, Facultad de Ciencias Químicas, Universidad Autónoma de San Luis Potosí, Av. Dr. Manuel Nava 6, SLP 78210, Mexico

ARTICLE INFO

Article history:
Received 25 February 2014
Accepted 3 November 2014
Available online xxxx

Keywords:
Antigenic determinants
Databases
Hypervariable pathogens
Immunological software
Immunoprotection

ABSTRACT

Exploitation of recombinant DNA and sequencing technologies has led to a new concept in vaccination in which isolated epitopes, capable of stimulating a specific immune response, have been identified and used to achieve advanced vaccine formulations; replacing those constituted by whole pathogen-formulations. In this context, bioinformatics approaches play a critical role on analyzing multiple genomes to select the protective epitopes *in silico*. It is conceived that cocktails of defined epitopes or chimeric protein arrangements, including the target epitopes, may provide a rationale design capable to elicit convenient humoral or cellular immune responses. This review presents a comprehensive compilation of the most advantageous online immunological software and searchable, in order to facilitate the design and development of vaccines. An outlook on how these tools are supporting vaccine development is presented. HIV and influenza have been taken as examples of promising developments on vaccination against hypervariable viruses. Perspectives in this field are also envisioned.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Exploitation of vaccination as a tool in fighting wide spread diseases has resulted in substantial strides in the combat against many infectious diseases such as influenza, smallpox, varicella, pertussis, diphtheria, tetanus, polio, hepatitis, and rotavirus [1,2]. The conventional vaccines, which include attenuated or killed agents, might take up to 15 years of development; this includes cultivation of the desired microorganism at a larger scale and under proper conditions as well as an effective inactivation with a subsequent evaluation of vaccine immunogenicity. Although this kind of vaccines has saved countless lives, it can have unfavorable consequences as adverse effects, induce the disease or, in some instances, even death [3–5].

Bioinformatics is a field of science in which several disciplines such as biology, computing, and information technology converge to organize and store large amounts of biological information driven by advances generated in genetics, molecular biology, and biotechnology [6]. One goal of bioinformatics is to streamline and interpret, effectively and timely, information from the genome, transcriptome, and/or proteome [7]. This discipline aims to promote health benefits including the area of vaccines.

The development of bioinformatics tools along with advances in recombinant DNA technology (rDNA) and the knowledge on the host immune response and the genetic background of the pathogen will lead to new vaccines against diseases that currently have few or no control measures in just 1 or 2 years through computer *in silico* predictions to define targets [8] see Fig. 1. The vaccines developed through rDNA technologies are designed to be safer, more efficacious, and/or less expensive than traditional vaccines. In order to achieve these aims, a thorough understanding of the disease agent, particularly, critical epitopes to induce the appropriate immunological reaction is required [9–11].

While the availability of the complete genome sequence permits the identification of all potential protein products, this information could be not sufficient to allow for the identification of the subset of proteins that are in fact expressed at any stage of the life

Abbreviations: HIV, Human Immunodeficiency Virus; rDNA, recombinant DNA technology; MHC, major histocompatibility complex; HLA, human leukocyte antigen; PSSM, Position Specific Scoring Matrices; ANN, Artificial Neuronal Networks; QM, quantitative matrices; KISS, Kernel-based Inter-allele peptide binding prediction SyStem; SVM, support vector machine; WAPP, Whole Antigen Processing Pathway; CTL, cytotoxic T cell; PI, Protrusion Index; gp120, envelop glycoprotein gp120; gag, structural polyprotein; mAb, monoclonal antibody; MCC, Matthews correlation coefficient; HA, hemagglutinin; NA, neuraminidase; TAP, transporter associated protein; APC, antigen presenting cells; SEPPA, Spatial Epitope Prediction of Protein Antigens; ATP, adenosine triphosphate; AIDS, acquired immunodeficiency syndrome; RV144, Thai HIV phase III prime/boost vaccine trial.

* Corresponding author. Fax: +52 444 826 2440.
E-mail address: rosales.s@fcq.uaslp.mx (S. Rosales-Mendoza).

http://dx.doi.org/10.1016/j.jbi.2014.11.003
1532-0464/© 2014 Elsevier Inc. All rights reserved.

of the pathogen. Proteomics also play an important role in this field as it can serve as a complementary strategy to the genomic-based approaches using immunomics techniques to identify and characterize immunogenic proteins. Vaccinomics, which consists on the characterization of host response to immunization, provides valuable information on pathogen–host cell interaction to validate candidate antigens. The information obtained from these disciplines also speeds the identification and characterization of new antigens [12].

Proteomic experiments conducted on bacterial species have not only been verified with data obtained from genome sequencing and bioinformatic analyses but have also lead to the discovery of new proteins, which could be potential new vaccine candidates [12–14]. An extensive review focused on the impact of proteomics on the development of antibacterial and antiviral vaccines was published by Adamczyk-Poplawska [12].

It is important to note that the standard bioinformatics web tools presented here, are not enough for detailed analysis of whole genomes. The immunoinformatics is a discipline whose main objective is to convert large-scale immunological data, using computational and mathematical approaches, to understand and organize these large scale data to obtain immunologically meaningful interpretations [15,16]. The tools in this field are based on statistical and machine learning system and are used for studies in modeling molecular interactions (such as antigen processing and presentation) and also plays a role in defining new hypotheses related to understand the immune system mechanisms [17,18].

This review is intended to provide a gateway to some of the most useful online immunological softwares and searchable databases for genomes analyses, based in our own experience and a laborious search in the literature and web databases, providing an outlook on how these tools have aided on the vaccine development field particularly on the development of epitope-based vaccines.

2. Epitope-based vaccines

Epitopes are of particular interest to both clinical and basic biomedical researchers as they hold huge potential for vaccine design, disease prevention, diagnosis, and treatment. Using rDNA technologies, we can isolate specific epitopes which can replace the whole pathogen in a vaccine. However, within the diversity of epitopes in a pathogen, it is important to notice that not all of the epitopes, even those that seem to be dominant, are equal in their ability to elicit antibody production [19–21].

Besides producing particular immunogens instead whole pathogens, rDNA has allowed for a rational vaccine design comprising the production of chimeric proteins that opens a wide number of possibilities for immunogen design; including the conception of multiepitopic vaccines having advantages such as: several immunoprotective epitopes are included in a single molecule, immunodominant but non-protective epitopes are discarded, and epitopes exerting adjuvant effects such as promiscuous T cell epitopes can be included to enhance immunogenicity [22]. These features offer the possibility of designing multitarget, highly efficient vaccines. However a requisite for the design of such immunogens consists on the discovery of the immunoprotective epitopes and the variants when genetic variability is of relevance for a particular pathogen.

The epitope-driven vaccine is an attractive concept that is being successfully pursued in a large number of research groups, especially to the development of vaccines targeting conserved epitopes in variable or rapidly mutating pathogens [23,24].

The selected epitopes in a vaccine should ideally be conserved across different stages of the pathogen and its variants. Furthermore it should be taken into consideration the desired immune response. Cytotoxic T cell-mediated response is elicited by a pathway comprising intracellular antigen processing with linear epitopes as predominant targets [21]. In this regard, the epitopes selected for a vaccine must have binding affinity with more than one major histocompatibility complex (MHC) allele and must cover a major population [25,26].

The proteins that contain many epitopes recognized by the common MHC alleles are known as promiscuous binders [26]. The human leukocyte antigen (HLA) supertype refers to a set of HLA alleles with overlapping peptide binding specificities. The alleles in the given HLA super type often represent the same epitope, which refers to the region on the surface of an antigen capable of eliciting an immune response for T cell recognition [25,27].

On the other hand, elicitation of humoral responses relies on the recognition of linear epitopes and conformational epitopes. The latter constitute a challenge for chimeric vaccine design as they must retain their native conformation to be functional [28]. Therefore, knowledge on the whole antigen structure is necessary to aid on the rational design of vaccines targeting conformational B cell epitopes [27].

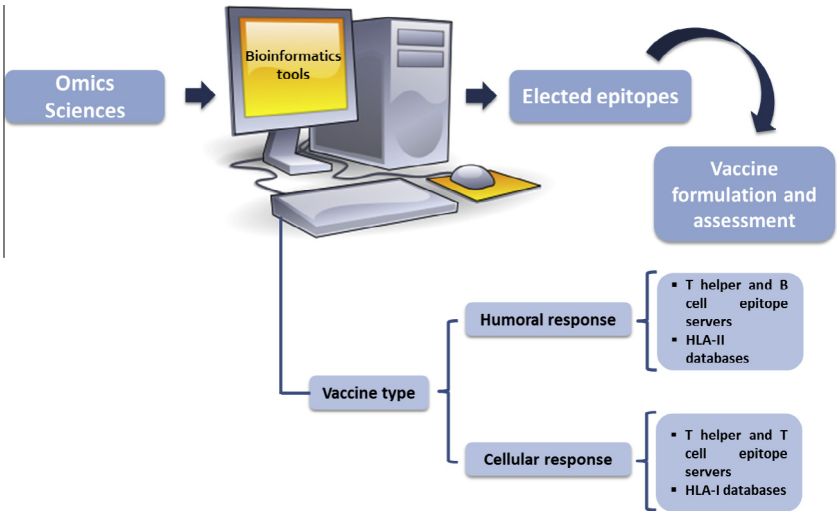

Fig. 1. Schematic representation of the workflow to identify epitopes for vaccine development.

Please cite this article in press as: Soria-Guerra RE et al. An overview of bioinformatics tools for epitope prediction: Implications on vaccine development. J Biomed Inform (2014), http://dx.doi.org/10.1016/j.jbi.2014.11.003

Download English Version:

https://daneshyari.com/en/article/6928300

Download Persian Version:

https://daneshyari.com/article/6928300

Daneshyari.com