Computers in Biology and Medicine xxx (xxxx) xxx-xxx

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: http://ees.elsevier.com

In-silico identification of subunit vaccine candidates against lung cancer-associated oncogenic viruses

Anjali Lathwal [a,1], Rajesh Kumar [b,1], Gajendra P.S. Raghava [a,*]

[a] Department of Computational Biology, Indraprastha Institute of Information Technology, New Delhi, India
[b] Bioinformatics Centre, Institute of Microbial Technology, Chandigarh, India

ARTICLE INFO

Keywords
Oncogenic viruses
Lung cancer
Antigenic epitopes
Promiscuous epitopes
Immune-stimulatory potential
Prediction pipeline
Subunit vaccine
Proteome
Immunoinformatics
Reverse vaccinology

ABSTRACT

Globally, ~20% of cancer malignancies are associated with virus infections. Lung cancer is the most prevalent cancer and has a 10% 5-year survival rate when diagnosed at stage IV. Cancer vaccines and oncolytic immunotherapy are promising treatment strategies for better clinical outcomes in advanced-stage cancer patients. Here, we used a reverse vaccinology approach to devise subunit vaccine candidates against lung cancer-causing oncogenic viruses. Protein components (945) from nine oncogenic virus species were systematically analyzed to identify epitope-based subunit vaccine candidates. Best vaccine candidates were identified based on their predicted ability to stimulate humoral and cell-mediated immunity and avoid self-tolerance. Using a rigorous integrative approach, we identified 125 best antigenic epitopes with predicted B-cell, T-cell, and/or MHC-binding capability and vaccine adjuvant potential. Thirty-two of these antigenic epitopes were predicted to have IL-4/IFN-gamma inducing potential and IL-10 non-inducing potential and were predicted to bind 15 MHC-type I and 49 MHC-type II alleles. All 32 epitopes were non-allergenic and 31 were non-toxic. The identified epitopes showed good conservancy and likely bind a broad class of human HLA alleles, indicating promiscuous potential. The majority of best antigenic epitopes were derived from *Human papillomavirus* and *Epstein-Barr virus* proteins. Of the 32 epitopes, 25 promiscuous epitopes were related to *E1* and *E6* envelope genes and were present in multiple viral strains/species, potentially providing heterologous immunity. Further validating our results, 38 antigenic epitopes were also present in the largest experimentally-validated epitope resource, Immune Epitope Database and Analysis Resource. We further narrowed the selection to 29 antigenic epitopes with the highest immunogenic/immune-boosting potential. These epitopes possess tremendous therapeutic potential as vaccines against lung cancer-causing viruses and should be validated in future experiments. All findings are available at https://webs.iiitd.edu.in/raghava/vlcvirus/.

1. Introduction

According to the World Health Organization (WHO), cancer is the second leading cause of death globally. Among all cancer types, lung cancer remains the leading cause of mortality in both men and women. Lung cancer is highly invasive and metastasizing in nature, causing more deaths than the other four leading cancers (pancreatic, colon, rectal, and breast) combined [1]. More than half of patients diagnosed with lung cancer die within a year and have a 5-year survival rate of 17.8% [2]. Risk factors contributing to lung cancer development include smoking, tobacco use, air pollutants, and oncogenic viruses [3]. Oncogenic viruses and other pathogenic agents cause nearly 20% of human cancers [4]. There is strong molecular evidence that oncogenic viruses such as *Human papillomavirus* (HPV), *Hepatitis B virus* (HBV), *Human T-lymphotropic virus* (HTLV), and *Rous sarcoma virus* (RSV) are the chief players involved in lung cancer pathogenesis [5]. These viruses are also involved in regulating gene expression and modulating drug responses [6]. Oncogenic viruses also benefit from the immunosuppressed tumor microenvironment, where they hijack the host replication machinery to synthesize oncoproteins and induce tumorigenesis [7,8].

Standard strategies such as radiotherapy, chemotherapy, targeted therapies, and resectable surgery are widely used to treat lung cancer. Commonly administered drugs include Axitinib, Carboplatin, Cediranib, CI-994, Cisplatin, Paclitaxel, Docetaxel, etc. The standard treatment strategies suffer from several limitations: toxicity, adverse side effects from radio/chemotherapy, and limited scope of resectable surgery for patients in advanced stages of lung cancer [1]. In this regard, immunotherapy is a widely employed advanced strategy for the treatment of cancer. It involves the administration of anti-CTLA 4 antibodies, anti-PD1 antibodies [9], and oncolytic viruses [10] to manage cancer. Immunotherapy employing vaccination is also the most widely used therapeutic strategy for the management of tumors. Several preventive

* Corresponding author. Department of Computational Biology, Indraprastha Institute of Information Technology, Delhi, Okhla Industrial Estate, Phase III, (Near Govind Puri Metro Station), Office: A-302 (R&D Block), New Delhi, 110020, India.
E-mail addresses: anjalil@iiitd.ac.in (A. Lathwal); b.rajesh@imtech.res.in (R. Kumar); raghava@iiitd.ac.in (G.P.S. Raghava)
[1] Contributed Equally.

https://doi.org/10.1016/j.compbiomed.2021.104215
Received 22 July 2020; Received in revised form 8 January 2021; Accepted 8 January 2021
Available online xxx
0010-4825/

vaccines, such as Cervarix, Gardasil, Heplisav-B, are in clinical trials for HPV related cancer types. Also, therapeutic vaccines such as Bacillus Calmette-Guerin (BCG) and Sipuleucel-T are approved for bladder and prostate cancer, respectively [11]. The advantage of vaccination is that it helps prepare the immune system for the selective eradication of tumor cells.

The past decade witnessed the evolution of an information-driven "reverse vaccinology" approach to identify vaccine candidates from the genomes and proteomes of pathogenic organisms. Several studies have used the reverse vaccinology approach to design subunit vaccine candidates against different diseases. Some specific examples include peptide-based subunit vaccines against Chagas disease [12], breast cancer [13], and prostate cancer [14]. In other studies, the reverse vaccinology approach has been utilized for identifying vaccine candidates for several emerging and pathogenic diseases. Successful examples include a proteome-based subunit vaccine candidates against emerging strains of mycobacterium [15], vaccines based on virulence factors and essential genes used against bacterial diseases [16], subunit vaccine candidates against the Zika virus [17], and many others against coronaviruses [18]. The advantages of using in-silico methods for identifying and predicting antigenic regions in the genome or proteome of pathogenic organisms are endless. Firstly, it not only brings down the cost of the experiment but also saves time and effort. Secondly, in-silico analysis can help find conserved regions of genomes or proteomes of pathogenic organism that aids in designing better and broad-acting therapies. Moreover, in-silico analysis can also help determine the minimum number of epitopes that need to be tested in the clinical laboratory. For example, predicted epitopes of SARS-Cov-2 viruses were recently used to design vaccine candidates [19].

Since oncogenic viruses are foreign to the body, an immune response can be generated against them. In the present study, we designed a computational pipeline that systematically identified subunit vaccine candidates against major lung cancer-causing viruses. We studied all the essential proteins of the viruses known to be involved in lung cancer pathogenesis in order to identify promising vaccine candidates. Oncogenic viruses evade the host immune response by producing ligands and proteins that resemble host molecules [20,21]. Thus, viral epitopes that matched peptides found in the human proteome were filtered out to avoid a self-immune reaction. Due to the technical limitations of clinical settings, most previous studies have considered either one or two viral proteins for epitope analysis [22]. To the best of our knowledge, no previous study has focused on identifying proteome-based epitope vaccine candidates for use against lung cancer-causing viruses. All the vaccine candidate epitopes identified in this study were stored and are available to the scientific community in the form of a freely accessible web resource at https://webs.iiitd.edu.in/raghava/vlcvirus/. Following assessment of patient genomic profiles for presence of oncogenic viruses, subunit vaccine candidates can be recommended to cancer patients in order to achieve more reliable, precise, and targeted treatments and better outcomes. We hope that the potential subunit vaccine candidates identified in this study will find their way into clinics and lead to improved lung cancer therapeutics.

2. Material and methods

2.1. Data source and extraction

The present study focused on identifying epitope-based subunit vaccine candidates against major oncogenic lung cancer-causing virus species. Table 1 summarizes the complete data statistics in the context of viruses used in the present study.

The overall study only assessed viral species known to cause lung cancer by integrating their genome into the human genome. The Uniprot database was utilized to extract proteomic data by using the virus name as the search keyword. Similar to several previous studies that utilized proteomic data for vaccine design [13,23,24], the present study included only standard quality reference viral proteomic data.

Table 1
Brief description of oncogenic virus species, reference proteomes, and number of proteins utilized in the study.

Oncogenic virus species	Reference proteome	Proteins	Lung cancer type	References [PMIDs]
Human Papillomavirus (HPV)	89	621	Squamous cell carcinoma, Adenocarcinoma	27415011
Hepatitis B virus (HBV)	01	05	Squamous cell carcinoma, Adenocarcinoma	27415011
Rous Sarcoma virus (RSV)	01	04	Adenocarcinoma, Adenocarcinoma with lepidic spread, Stage IV adenocarcinoma, Squamous cell carcinoma	10885982
Simian T-cell Lymphotropic virus (STLV)	01	06	Squamous cell carcinoma, Adenocarcinoma	27415011
Bovine Leukemia virus (BLV)	01	06	Squamous cell carcinoma	27415011
Human T-cell Lymphotropic virus (HTLV)	01	06	Squamous cell carcinoma	27415011
Epstein-Barr virus (EBV)	02	171	Lung adenocarcinoma	31159203
Measles virus	03	22	Lung adenocarcinoma	19895323
JC virus	01	97	Squamous cell carcinoma, Adenocarcinoma	17534844

2.2. Target selection of vaccine candidate

One of the most critical aspects of designing subunit vaccines is identifying candidate epitopes that may elicit immune responses. Since viral proteins are integral for pathogenesis, targeting such proteins is an exciting research area for proteome-based virus vaccine research [25]. Table 2 briefly summarizes the major viral components utilized in our study.

Many ongoing clinical trials against various viral species suggest the importance of protein epitope-based subunit vaccine candidates (https://clinicaltrials.gov/ct2/results?cond=subunit+vaccine&term=virus). These studies mainly target only essential virus protein components such as non-structural proteins, glycoproteins, and receptor binding proteins for designing epitope-based subunit vaccine candidates, as these proteins are integral components for virus pathogenesis [26]. But components other than non-structural proteins can also potentially act as immunogens [27,28]. Thus, this study included essential protein components and all other virus components to identify epitope-based subunit vaccine candidates against lung cancer-causing viruses.

3. Epitope prediction methodology

3.1. Generation of ninemers

Literature evidence suggests that proteome-based subunit epitopes are attractive candidates for vaccination. It also suggests that peptides 8–10 residues long are enough to bind to MHC class I and II molecules. We generated all possible ninemer peptides (peptides containing nine

Table 2
Distribution of proteins among the oncogenic lung cancer virus species.

Oncogenic virus species	Protein descriptions
Human Papillomavirus (HPV)	NS1, VP1, 11 K, Replication protein E1, E7, E6, E4, E2, E8, Minor capsid protein L2, L1, E5A & E5B
Hepatitis B virus (HBV)	Large envelope protein, Capsid protein, Antigenic protein, Protein P & Protein X
Rous Sarcoma virus (RSV)	Gag-Pol polyprotein, Gag polyprotein, Envelope glycoprotein & Src tyrosine kinase
Simian T-cell Lymphotropic virus (STLV)	Tax protein, Gag polyprotein, Pro protein, Rex protein, Pol protein & Envelope glycoprotein
Bovine Leukemia virus (BLV)	RT-IN, Pr66, Gp60 SU, Gag polyprotein, p18 & p34
Human T-cell Lymphotropic virus (HTLV)	Gag-pro-pol, Protein Rex, Gag polyprotein, Envelope gp63 glycoprotein & Protein tax-2
Epstein-Barr virus (EBV)	BRRF2, BKRF4, BTRF1, BRRF1, BNLF2b, BZLF1, BLRF2, Capsid vertex 2, N protein, EBNA4, M protein, EBNA6, GP350, & BPLF1 protein
Measles virus	Phosphoprotein, L, N, M, H, protein C, Fusion glycoprotein F0 & Non-structural protein V
JC virus	DnaB-like helicase, Portal protein, Minor tail protein, Beta-lactamase protein, Capsid maturation, Integrase protein, DNA primase & Uncharacterized proteins

amino acid residues) from the given protein sequences of all nine oncogenic viruses using an overlapping window of nine amino acid residues length, which is a standard technique in the field of epitope-based vaccine candidates identification. An in-house python script was used for the generation of ninemers. These were continuous nine residue-long peptides derived from the proteins considered in the study. Further, to avoid redundancy, all duplicate ninemers were removed for further analysis.

3.2. Removal of self-epitopes

Self-tolerance must be considered in vaccine design, as a healthy immune system rarely acts against self-epitopes. All nonapeptides/ninemers that are also present in the healthy human body had to be removed to avoid self-tolerance [24]. All ninemers generated from the viral proteins were mapped onto the 1000 normal human proteomes, and an in-house python script was developed for the removal of self-epitopes. Ninemers that shared 100% amino acid identity with sequences from the 1000 normal human reference proteome were removed from further analysis. This methodology is well adopted in literature and has served as the basis for several critical findings related to immunotherapeutic development.

3.3. Predicting epitope properties in virus proteome

Our next objective was to select peptides that could activate the human immune system and generate memory cells. The engagement of an effective adaptive humoral immune system is necessary for the success of any vaccination strategy. Thus, it is necessary to identify epitopes that can elicit and induce a B-cell response. Since an adaptive T-cell response is also necessary for the defense against pathogenic organisms, it is also of utmost importance to identify epitopes that can elicit a T-cell response. Once antigenic proteins are recognized by antigen-presenting cells, they are broken down by digestive enzymes into several short fragments. Not all antigenic fragments can bind to class I and II MHC molecules. Only fragments that bind strongly to class I and II MHC molecules may be displayed by an MHC molecule on their cell surfaces and be subsequently recognized by T-lymphocytes. Thus, it is necessary to identify the regions in the antigenic proteins that can bind to class I or II MHC molecules and induce B and T-cell responses. Since antigenic fragments can be easily degraded in the body by proteolytic enzymes, it is also important that the vaccine molecules be delivered via immune adjuvants. In this regard, a computational pipeline was developed to predict different kinds of immune epitopes from the generated nonapeptides. The pipeline predicted (i) Linear B-cell epitopes using LBtope [29], (ii) MHC class I & II binders using ProPred1 and ProPred, respectively [30,31], (iii) T-cell epitopes using CTLpred [32], (iv) Vaccine adjuvant using VaxinPAD [33], and (v) Immune-boosting potential using IL4pred and IFNepitope [34,35]. Sequence-based methods for MHC I/II recognition were used rather than docking-based methods. It has been shown in previous studies that docking-based prediction methods performed poorly than sequence-based methods [31]. A large number of non-MHC binders can also fit into the groove of MHC molecules, so docking-based methods sometimes underperform compared to sequence-based methods. Besides, it is not computationally feasible to dock all peptides into a large number of MHC alleles. The complete workflow of the reverse vaccinology approach used to identify potential subunit vaccine candidates is shown in Fig. 1.

3.4. Identification of immune-boosting properties

Interleukin 4 (IL-4) and interferon-gamma (IFN-gamma) are cytokines that boost the immune response. The best antigenic subunit vaccine candidates that had all desired properties (i.e., MHC-binders, cell-mediated and humoral immunity inducing potential, and vaccine potential) were further investigated for the presence of IL-4 and IFN-gamma inducing potential.

3.5. Structural annotation of epitopes

Epitope structure plays an important role in determining its biological activity. With this in mind, the best antigenic epitope structures were predicted and provided in the VLCvirus resource. The information can be accessed by clicking the "Enquire - - Best Antigenic Epitope Properties" link under the Menu tab. The structure prediction algorithm PEPstrMOD [36] was employed to predict the tertiary structure of the candidate antigenic epitopes. PEPstrMOD can predict the structure of a given peptide or protein-containing natural and modified residues up to 25 amino acids in length.

3.6. Web resource interface development

All the results obtained from the analysis were stored and presented in the form of the web resource VLCvirus which is freely available for the scientific community at https://webs.iiitd.edu.in/raghava/vlcvirus/. The web resource was deployed using Linux-based Apache Server (LAMP). The website front-end was built using a bootstrap responsive HTML template that adjusts the screen by sensing the user's device. All the data-related queries were handled in the web resource by using MySQL.

4. Results and discussion

4.1. Protein sequence retrieval

The protein sequences from nine oncogenic lung cancer-causing virus species were retrieved from the Uniprot database. The Uniprot identifier of each virus species was as follows: *Hepatitis B virus* (UP000008591), *Bovine leukemia virus* (UP000202838), *Human-T cell lymphotropic virus 2* (UP000009254), *Simian T-cell lymphotropic virus* (UP000115666), *Rous sarcoma virus* (UP000159275), *Epstein-Barr virus* (UP000007639,UP000153037), *Measles virus* (UP000008699,UP000116098,UP000138367), *JC virus* (UP000008407) and *Human papil-*

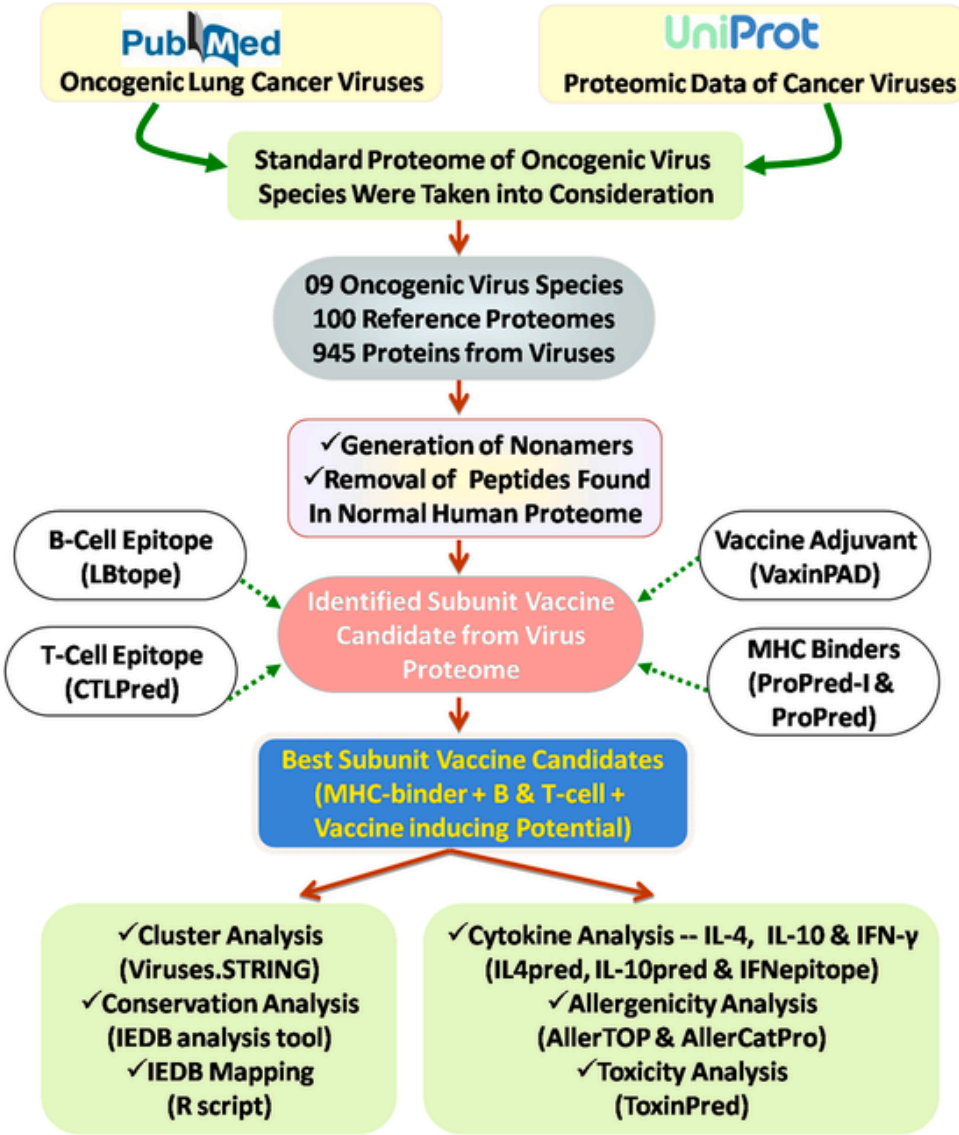

Fig. 1. Figure depicting the workflow of the reverse vaccinology approach used for identifying potential subunit vaccine candidates against lung cancer-causing virus species.

lomavirus (Supplementary S1 Table 1). Each viral species has a different set of core proteins necessary for their virulence (e.g., E2, L2, and capsid proteins). Each of the proteins (Table 2) present in each standard proteome was utilized to identify epitope-based subunit vaccine candidates. The present study employed 945 proteins from 100 reference standard proteomes from nine oncogenic virus species.

4.2. Identification of subunit vaccine candidates

After excluding peptides with 100% sequence identity with any protein from the reference human proteome, the list of viral protein nonapeptides was assessed to identify epitopes that could potentially exhibit different immune stimulation properties. Each nonapeptide of all virus proteins were utilized in the prediction pipeline. We integrated LBTope, CTLPred, VaxinPAD, ProPred, and ProPred-1 server in our pipeline, which identified and predicted B-cell epitopes, T-cell epitopes, vaccine adjuvant potential of the epitopes, and MHC-II and MHC-I binders. The integrated tools in the pipeline helped us find immunogenic regions of the candidate viral proteins. The high number of identified immunogenic nonapeptide epitopes was not practical for all to be tested in a clinical setting. Therefore, we applied a heuristic approach to arrive at a reasonable number that could be feasibly tested in clinical studies. We shortlisted epitopes that exhibited more than one property, including epitopes that were predicted to have MHC binder ability and subsequently might also possess B-cell and vaccine adjuvant potential. We termed these epitopes as multifunctional in our study. The results of this analysis are provided in Table 3.

Similarly, we further narrowed the epitope list by selecting those that were predicted to serve as T-cell epitopes with MHC binding ability, and vaccine adjuvant potential or to serve as B cell epitopes with MHC binding ability and vaccine adjuvant potential. Furthermore, epitopes with both T-cell and B-cell epitope potential, vaccine adjuvant potential, and MHC–I or II binding ability were also retrieved. These epitopes were deemed as the best antigenic subunit vaccine candidates since they were predicted to stimulate both arms of the immune system, possess MHC binding capability, and act as vaccine adjuvants. The results of this analysis are provided in Table 3. Many of these epitopes were considered promiscuous since they were predicted to have binding affinity towards more than one HLA class. Some of the best antigenic epitope candidates were predicted to bind to 45 different types of HLA molecules, which may account for most of the HLA allele class

Table 3
Table representing the number of epitopes predicted to exhibit different properties for each viral species after excluding epitopes that had 100% sequence identity with proteins from the human proteome.

Properties	BLV	HBV	HPV	HTLV	RSV	STLV	EBV	JCV	MV
T-Cell Epitope	717	427	48347	665	729	500	17559	3429	3396
B-cell Epitope	228	251	17985	445	297	229	7208	1102	1397
Vaccine Adjuvant	324	256	25723	293	401	234	7312	1466	1833
MHC_I	10943	6332	665437	11384	9769	8293	270008	48869	49913
MHC_II	4241	3043	272264	4835	3736	3297	117045	16236	24343
Multifunctional Epitopes Identified From Oncogenic Virus Proteomes									
MHC_I_II	292	228	21472	289	282	298	5265	1615	776
B_MHC_I	113	137	9771	197	141	145	2203	767	369
B_MHC_II	19	19	1815	41	21	35	335	103	86
B_MHC_I_II	18	16	1443	34	18	30	272	82	66
T_MHC_I	509	357	38753	468	486	490	9057	3243	1311
T_MHC_II	97	71	6628	82	83	88	1554	466	236
T_MHC_I_II	97	69	6480	82	81	87	1535	455	228
BV_MHC_I	12	27	1130	15	13	8	191	70	52
BV_MHC_II	2	2	329	4	4	2	59	11	14
BV_MHC_I_II	2	2	272	4	4	2	50	8	11
TV_MHC_I	86	55	6743	60	75	51	1246	437	221
TV_MHC_II	17	9	1578	8	17	9	311	96	58
TV_MHC_I_II	17	9	1539	8	17	8	306	94	54
TBV (Tcell, Bcell & Vaccine)	5	11	494	4	5	2	75	34	26
Antigenic Epitopes Identified From Oncogenic Virus Proteomes									
TBV_MHC_I	5	10	473	4	5	2	72	34	26
TBV_MHC_II	1	0	99	1	0	1	18	3	5
TBV_MHC_I_II	1	0	97	1	0	1	17	3	5

types present throughout humans. Thus, the predicted promiscuous epitopes could potentially be used towards developing prophylactic immunity throughout the human population.

4.3. Residue preference of the best antigenic subunit vaccine candidates

After narrowing the selection of epitope vaccine candidates to epitopes predicted to bind MHC I or II and stimulate T-cells and/or B-cells, the remaining epitopes were accessed for preferred residues. Fig. 2 demonstrates the probability of preferred residues of the ninemeric epitopes we identified against the best antigenic epitopes found in three different classes. Basic amino acids were preferred in most positions of the best antigenic epitopes. Since predicted epitope residues can significantly increase the success ratio of immunotherapy, this information will help researchers design antibodies against them in in-vitro conditions [37].

4.4. Identification of promiscuous nature of best antigenic subunit vaccine candidates

Earlier studies have suggested that epitope-based immunization is capable of eliciting a protective immune response. Yet, experimental evidence suggests that promiscuous epitopes should be considered to further improve epitope-based vaccine design efficacy [38]. Promiscuous epitopes pose great therapeutic benefits by exhibiting binding affinities with a large number of HLA molecules. In our study, we also identified several promiscuous epitopes that may be able to bind to a large number of HLA molecules. These epitopes are required for T-cell stimulation and priming and boosting the immune response. The results of this analysis are available in the download section of the VLCvirus web resource. The most frequently expressed alleles in human populations are HLA-A1, HLA-A2, DRB1, and DRB4 [39]. Out of the 125 best

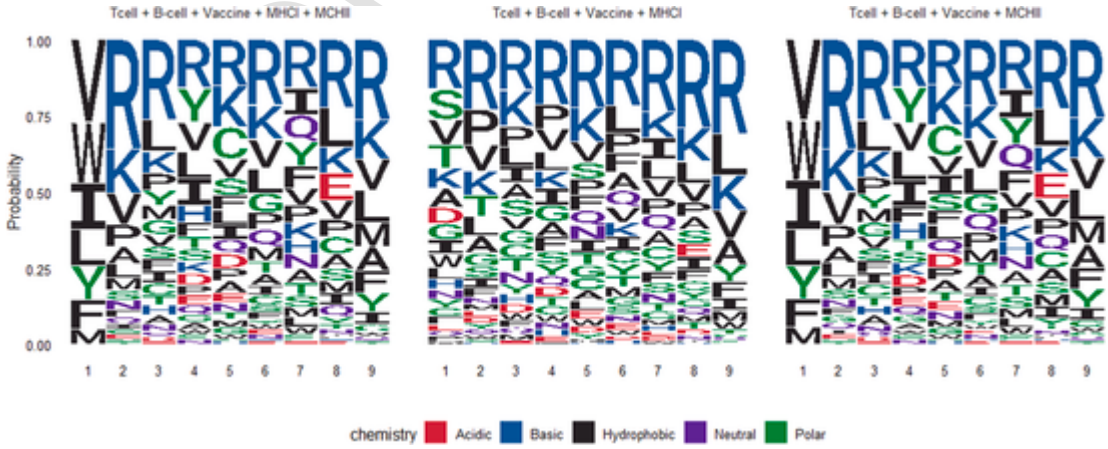

Fig. 2. Sequence logo of preferred residue at each position for the best antigenic epitopes. The probability of preferred residues of ninemeric epitopes were compared against the best antigenic epitopes in three different classes. Taller letters signify the higher number of the particular residues at the respective position of the peptide sequences. Residues were scaled according to their frequency.

antigenic epitopes, some were predicted to bind to nearly 15 MHC-type I ("VMFVSRVPV") and 49 MHC-type II ("LRRFMVALI") alleles. Thus, the identified epitopes likely exhibit broad coverage of the human population and have the potential to be utilized in clinics for providing mass scale immunogenicity.

4.5. Identification of immune-dominant conserved epitopes across strains

Infection dynamics can vary across different strains of pathogenic species [40]. Therefore, it is important to strive for cross-strain immunity and design epitopes that provide protective immunity across different viral strains. When designing epitope-based vaccines, it is likely more prudent to explore conserved epitopes that may offer broader protection across multiple strains, or even species, than epitopes derived from highly variable genome regions [41]. We identified the epitopes present in multiple strains/species of oncogenic viruses that exhibit various immunogenic properties (Fig. 3). Supplementary information on detailed parameters such as HLA binding frequencies and associated genes/proteins of the promiscuous epitopes are available in the download section of the VLCvirus resource. The identified conserved epitopes across the strains may be useful for providing heterologous immunity and offer an alternative strategy to conventional approaches of epitope-based vaccine design [42].

4.6. Identification of cytokine-inducing properties

The most promising subunit vaccine candidates were further analyzed for their IL-4 and IFN- gamma inducing potential. The foundation behind this approach is that for successful anticancer therapy, immunostimulants should be present in high amounts near the tumor microenvironment. Keeping this idea in mind, we extended our methodology to predict the IL-4 and IFN-gamma inducing potential of the epitopes and further narrow our selection of antigenic peptides for testing in clinical settings. The majority of the identified best antigenic epitopes were also predicted to induce IL-4 and IFN-gamma cytokines; 103 epitopes were predicted to induce IL-4, 38 epitopes were predicted to induce IFN-gamma, and 32 epitopes were predicted to induce both (refer to downloads section of VLCvirus). We also assessed other properties such as IL-10 inducing potential, allergenicity, and toxicity of the antigenic peptides, the results of which can be found in Supplementary S1 Table 4. These cytokine inducers are attractive agents for use during cancer immunotherapies since the co-treatment of various cytokines with chemotherapy has been successfully tested in many clinical trials. We also analyzed the parent genes of the 32 promiscuous epitopes that were predicted to induce both IL-4 and IFN-gamma. We found that most of the epitopes belonged to *E1* and *E6* genes (Fig. 4). Several previous studies have also determined that envelope proteins are important for designing effective epitope-based vaccines against pathogenic/virus strains [43,44].

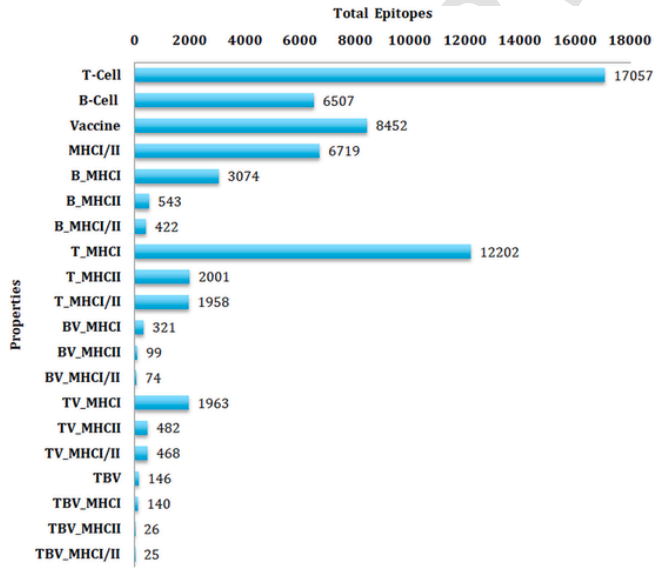

Fig. 3. Promiscuous epitopes across two or more viral strains/species against different properties.

4.7. IEDB mapping of epitopes

In order to partly validate our list of 125 best antigenic epitopes vaccine candidates, we mapped the epitopes to data available from the Immune Epitope Database and Analysis Resource (IEDB) version 3 that houses the largest number of experimentally validated immune epitopes [45]. The scope of IEDB data can be categorized into four broad disease categories, including allergenic diseases, autoimmune disorders, transplantation, and infectious diseases. The epitopes present in the IEDB data can be used to provide protective immunity against the disease categories mentioned above. We implemented continuous substring matching criteria (ninemers as a continuous subsequence of IEDB epitope) using an R script while identifying matching epitopes since the peptides in IEBD are varied in length. We found 38 matched records against our list of 125 best antigenic epitopes and 29,193 records for the epitopes (193,437) that had any of the following computed properties: T-cell epitopes, B-cell epitopes, vaccine adjuvant potential, MHC I binders, and MHC II binders. Thus, we hypothesize that our predicted epitopes that were validated by the IEDB data can also be used both for promoting immunity against lung cancer-causing virus species and the different disease classes that were previously mentioned. In this regard, we can say that these identified epitopes may be best suited for providing heterologous immunity among the same virus type and unrelated strains of different pathogenic organisms [45,46]. IEDB mapping results are available in Supplementary S1 Table 2 and 3. We also provided the IEDB mapping results of the best antigenic epitopes on the web resource under the Enquire section. This analysis further supports the reliability and credibility of our identified epitopes.

4.8. Antigenic epitope conservancy analysis

Examining epitope conservancy is a crucial step in designing epitope-based vaccines. Conservancy analysis signifies the degree of distribution of epitopes in a homologous set of proteins. We analyzed peptide conservancy using the IEDB conservancy analysis tool. The best antigenic epitopes found in this study were analyzed against respective

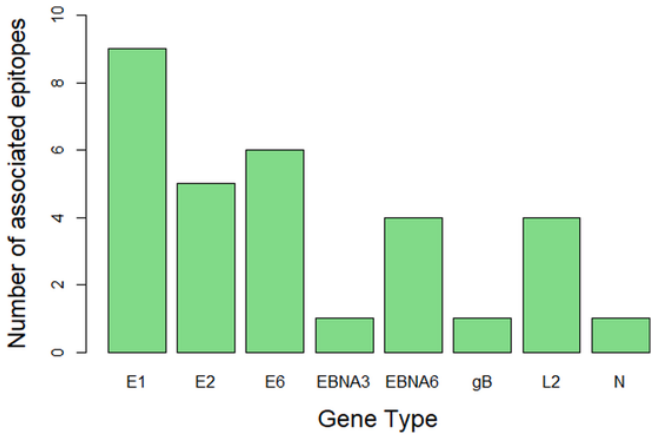

Fig. 4. Gene-wide distribution of 32 promiscuous epitopes with IL-4 and IFN-gamma immunostimulatory properties.

viral species from which they were derived, and some of them showed roughly 50% conservancy (Supplementary S1 Table 4).

4.9. Toxicity and allergenicity epitope analysis

Toxicity assessment of peptide-based therapeutics is also incredibly important in clinical settings. Peptide-based drugs should be free from toxicity such as hemotoxicity and liver toxicity. In order to identify the best antigenic epitopes for testing in clinical settings, we used the ToxinPred server [48] to identify the toxicity of the best antigenic epitopes (Supplementary S1 Table 4). The result suggests that out of the 125 best antigenic epitopes, only five epitopes were predicted to elicit toxic responses. These epitopes were YVRECITMV, IRCQECYAF, VKMYKRYEM, VRGHWRNCC, and YNSDKTCCR. We also identified the allergenicity the 125 best antigenic epitopes using the AllerCatPro server [49]. We found that none of the identified epitopes exhibited allergic properties. Further, various physicochemical properties of the epitopes were evaluated (Supplementary S1 Table 4).

4.10. Epitope-host protein interaction and cluster analyses

Better understanding virus-host protein interactions may help advance vaccines and therapeutic regimen design and development. In this regard, we retrieved all the virus-human protein interactions using the Viruses. STRING [50] server for the viruses used in the current study. However, we could find the data for only six viral species. *Epstein-Barr virus* had the highest number of protein interactions with human proteins among all the viruses. We then mapped the protein-coding genes to identify their role in lung cancer pathogenesis using the DisGeNET resource (Supplementary S1 Table 5). The epitopes were generally similar in sequence, which complicated the epitope-associated analysis. We performed cluster analysis using the IEDB clustering tool [51] by selecting the "cluster-breaking algorithm" to find consensus sequence driven sub-clusters (Supplementary S1 Table 6). The total identified clusters were 97, of which 15 were consensus type and 82 were singletons.

4.11. Web resource implementation and utility

All the identified immunogenic ninemeric epitopes were compiled and provided in the form of a web resource VLCvirus (https://webs.iiitd.edu.in/raghava/vlcvirus/) for the advancement of peptide-based subunit vaccine design against lung cancer-causing virus species. The web resource is built on a LAMP server (version 2.4.7). The front-end of VLCvirus was created using a bootstrap responsive HTML5.0 template, which adjusts the screen ratio by sensing the user device. The back-end data related queries were managed by MySQL server of version 5.5.55-0ubuntu0.14.04.1 - (Ubuntu). VLCvirus has various data browsing categories, which helps users get responses against queries in a time-efficient manner. In order to ease data visualization, a user-defined query is displayed in a responsive tabular form. The data browsing facility of VLCvirus has three different categories: I) Multifunctional Epitopes, II) Best antigenic Epitopes, and III) Promiscuous Epitopes across strains. Multifunctional browsing helps the user query the data against all virus species that have multiple immune-stimulating properties such as epitopes predicted to have B-cell inducing potential, vaccine adjuvant potential, and/or MHC-binding ability. The best antigenic module allows users to query against the epitopes with maximum identified multiple properties, HLA frequency, and their cytokine (i.e., IL-4 and IFN-gamma) inducing potentials. The promiscuous epitope module facilitates the user to query the ninemers epitopes across the different viral strains/species present in the study. This module may be of high importance for researchers who wish to design a subunit vaccine candidate that can effectively vaccinate against the different strains of the species. The Enquire module's multi-query provides the user with the ability to get customized results based on four properties: virus name, property, and MHC class I/II types. Moreover, the VLCvirus web resource is also equipped with several data analysis tools such as BLAST and epitope identification via protein query. The epitope identification tool allows users to map the epitopes against IEDB, MHCBN, and BCIPEP databases. The implemented BLAST tool allows the user to find related sequences to the user-defined query protein.

Application of the work: The development, application, and study of epitope-focused immunogens are conceptually attractive for managing diseases where conventional approaches to vaccination face obstacles. One notable attribute of epitope-based vaccines is their ability to elicit antibodies toward sequences that otherwise might not be immunogenic, which may happen with some epitopes found in vaccines made with attenuated or inactivated pathogens. Indeed, employing various filtration steps, including the removal of peptides present in the healthy human proteome, in order to avoid initiating the self-tolerance mechanism. In order to select the best antigenic epitopes, we further screened for epitopes that were promiscuous. The selected epitopes may also facilitate the generation of a preventive immune response against oncogenic lung cancer viruses since they were predicted to elicit both B- and T-cell responses.

As previously mentioned, lung cancer is one of the leading causes of morbidity among cancer patients. Several treatment options, such as Paclitaxel, Carboplatin, and Cisplatin, have improved patient lives, but even today, there is no definitive cure for lung cancer. Additionally, chemotherapy has several side effects including toxicity, immune suppression, and many others. In contrast to chemotherapy, vaccines can initiate a host immune response against the pathogen and can offer multifaceted management of the infection. Lately, researchers around the globe have identified several potential vaccine candidates against different pathogenic and virus species and have even successfully developed vaccines such as DPT, BCG, etc. As such peptide-based subunit vaccine drug development is an exciting area of research[47].

Previous studies have typically identified several peptide-based vaccine candidates using computational and molecular docking analysis. For example, one such study assessed peptide MHC binding conservation score to identify vaccine candidates against eukaryotic pathogens [52]. In another study, the HBV polymerase protein was used as a target for the identification of vaccine candidates [53]. Another study utilized glycoproteins as a target protein to identify a multisubunit vaccine against the Kaposi sarcoma [54]. Similarly, another study utilized spike proteins of MERS-CoV viruses to predict the effective epitopes in order to develop prophylactic measures [18]. Lastly, nucleotide data of tumor antigens was also used to identify the subunit vaccine for prostate cancer [14]. To the best of the our knowledge, no previous consensus study has analyzed the viruses and their proteins known to be involved in lung cancer pathogenesis to identify epitope-based subunit vaccine candidates. Moreover, no other previous study has provided comprehensive epitopes information and results as a user-friendly resource. We integrated various tools in our pipeline to identify protein-based subunit vaccine candidates against lung cancer-causing viruses. All the tools integrated into our pipeline have been widely used in literature to identify and design potential vaccine candidates. We implemented sequence-based methods in our pipeline and did not utilize structure/docking-based algorithms because of their intrinsic limitations. The number of vaccine candidates identified by each method was in thousands, too large to be feasibly tested in an experimental/clinically setting. Therefore, we further screened the epitopes using a heuristic approach that extracted promiscuous epitopes via multiple properties. These epitopes may have the capability of evoking different arms of the immune system. Moreover, we implemented several user-friendly analysis tools, such as identifying epitopes in the query sequence, in our web resource in order to facilitate data analysis with a single platform.

As we know by now, peptides are an attractive target for immunotherapy but are less immunogenic when used alone. Peptide therapeutics need a corresponding immunologic adjuvant in order to enhance their therapeutic potential. With this in mind, we predicted the vaccine adjuvant potential as well as the cytokine inducing potential of all the epitopes from the different virus species in order to identity epitopes that may improve immune stimulation of less immunogenic molecules. In addition, the identified promiscuous epitopes could provide prophylactic immunity. Since the identified promiscuous epitopes were conserved across different strains/species of oncogenic viruses, they could alternatively be utilized in diagnostic and disease monitoring settings [41]. Moreover, the identified most promising epitopes did not have glycosylation residues, indicating that these epitopes could be used to elicit robust neutralizing antibody responses against viruses involved in lung cancer pathogenesis. Previous studies have suggested that conformational epitopes lead to better vaccine design, but designing such epitopes is very time-consuming and unreliable since conformational epitopes can change with time. As such, we only predicted the immunogenic potential of linear epitopes for easy synthesis and greater stability. Our study identified 125 subunit vaccine candidates that can activate the both arms of the immune system, including 32 candidates that were predicted to have IL-4 and IFN-gamma inducing potential.

The vaccine candidates identified in this study could also lead to the generation of preventive therapy against lung cancer. The best antigenic epitopes identified in this study were predicted to have IL-4 and IFN-gamma inducing potential but were not predicted to induce anti-inflammatory cytokines such as IL-10. As a consequence of our analysis, we selected 29 epitopes that outperformed the remaining epitopes in terms of immunogenic potential and lacking allergenicity and toxicity (Supplementary S1 Table 7). The target binding sites of these 29 epitopes are provided in Supplementary S2. While considering the peptide-based therapeutics, the possibility of stimulating a cytokine storm should be considered carefully. However, in order to avoid a cytokine storm, immune system has to be manipulated (e.g., use of adjuvant, depletion of regulatory T lymphocytes, inhibition of anti-inflammatory cytokine, etc.) [55]. Further, immune-modulation and prevention of a cytokine storm can be achieved by targeting the COX-2 and cholinergic anti-inflammatory pathways.

5. Future directions

The area of immune-modulator research is still in its infancy; however, in the coming decade, more research will lead to advancements in reliable immune modulators that aid in disease therapy management like GTS-21 and PAF-AH. In addition to the above-mentioned strategies, several computational tools such as PIP-EL [56] and ProIn-Fuse [57] can also be used to assess the proinflammatory cytokine inducing potential of a given peptide sequence and be used to avoid initiating cytokine storms. Cautionary steps should be adopted while suggesting therapeutic options, in order to avoid triggering cytokine storms. Genomic analysis of epitopes identified in this study would also impart greater accuracy and allow one to trace virus pathogenicity, but we did not do this analysis since it was outside the scope of this study. Since each prediction algorithm used in this study has its own limitations, we suggest that the epitopes identified in this study be verified experimentally in a clinical setup before it is used as a therapeutic agent. Despite such limitations, we hope that the data stored in the VLCvirus will facilitate the vaccine designing process against lung cancer. Moreover, the integrative pipeline may pave new ways to design vaccine candidates against new bacterial, viral, fungal, or other pathogenic organisms.

Author's contribution

Conception and design: Anjali Lathwal, Rajesh Kumar and GPS Raghava, Development of methodology: Anjali Lathwal, Rajesh Kumar and GPS Raghava, Acquisition of data: Anjali Lathwal and Rajesh Kumar, Analysis and interpretation of data and results: Anjali Lathwal, Rajesh Kumar and GPS Raghava, Writing, reviewing, and revision of the manuscript: Anjali Lathwal, Rajesh Kumar and GPS Raghava.

Funding

There is no funding available for this article.

Data availability

All the data used in the study is freely available on UniProt.

Declaration of competing interest

All authors declare that there is no potential financial and personal conflict of interest.

Acknowledgment

The authors are thankful for the funding agencies, the Council of Scientific and Industrial Research (CSIR), University Grants Commission (UGC), and Indraprastha Institute of Information Technology (IIIT-D) Delhi for providing the necessary facility and infrastructure to carry out this research work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2021.104215.

References

[1] H. Lemjabbar-Alaoui, O.U.I. Hassan, Y.W. Yang, P. Buchanan, Lung cancer: biology and treatment options, Biochim. Biophys. Acta Rev. Canc 1856 (2015) 189–210, doi:10.1016/j.bbcan.2015.08.002.
[2] C. Zappa, S.A. Mousa, Non-small cell lung cancer: current treatment and future advances, Transl. Lung Cancer Res. 5 (2016) 288–300, doi:10.21037/tlcr.2016.06.07.
[3] Y.H. Wang, X. Di Shen, Human immunodeficiency virus infection and mortality risk among lung cancer patients, Med. (United States). 97 (2018), doi:10.1097/MD.0000000000010361.
[4] Y. Kim, C.M. Pierce, L.A. Robinson, Impact of viral presence in tumor on gene expression in non-small cell lung cancer, BMC Canc. 18 (2018), doi:10.1186/s12885-018-4748-0.
[5] L.A. Robinson, C.J. Jaing, C. Pierce Campbell, A. Magliocco, Y. Xiong, G. Magliocco, J.B. Thissen, S. Antonia, Molecular evidence of viral DNA in non-small cell lung cancer and non-neoplastic lung, Br. J. Canc. 115 (2016) 497–504, doi:10.1038/bjc.2016.213.
[6] M. Pogorzelski, S. Ting, T.C. Gauler, F. Breitenbuecher, I. Vossebein, S. Hoffarth, J. Markowetz, S. Lang, C. Bergmann, S. Brandau, J.A. Jawad, K.W. Schmid, M. Schuler, S. Kasper, Impact of human papilloma virus infection on the response of head and neck cancers to anti-epidermal growth factor receptor antibody therapy, Cell Death Dis. 5 (2014), doi:10.1038/cddis.2014.62.
[7] C.C. Spriggs, M.C. Harwood, B. Tsai, How non-enveloped viruses hijack host machineries to cause infection, Adv. Virus Res., Academic Press Inc., 2019, pp. 97–122, doi:10.1016/bs.aivir.2019.05.002.
[8] S.K. Thaker, J. Ch'ng, H.R. Christofk, Viral hijacking of cellular metabolism, BMC Biol. 17 (2019) 59, doi:10.1186/s12915-019-0678-9.
[9] C. Sanghera, R. Sanghera, Immunotherapy – strategies for expanding its role in the treatment of all major tumor sites, Cureus 11 (2019), doi:10.7759/cureus.5938.
[10] A. Lathwal, R. Kumar, G.P.S. Raghava, OvirusTdb: a database of oncolytic viruses for the advancement of therapeutics in cancer, Virology 548 (2020) 109–116, doi:10.1016/j.virol.2020.05.016.
[11] R.E. Hollingsworth, K. Jansen, Turning the corner on therapeutic cancer vaccines, Npj Vaccines 4 (2019), doi:10.1038/s41541-019-0103-y.
[12] L. Michel-Todó, P.A. Reche, P. Bigey, M.J. Pinazo, J. Gascón, J. Alonso-Padilla, In silico design of an epitope-based vaccine ensemble for Chagas disease, Front. Immunol. 10 (2019), doi:10.3389/fimmu.2019.02698.
[13] A.A. Imani Fooladi, H.M. Hosseini, J. Amani, An in silico chimeric vaccine targeting breast cancer containing inherent adjuvant, Int. J. Canc. Manag. 8 (2015), doi:10.17795/ijcp2326.
[14] P. Patra, M. Bhattacharya, A.R. Sharma, P. Ghosh, G. Sharma, B.C. Patra, B. Mallick, S.S. Lee, C. Chakraborty, Identification and design of a next-generation multi epitopes bases peptide vaccine candidate against prostate cancer: an in silico approach, Cell Biochem. Biophys. (2020), doi:10.1007/s12013-020-00912-7.
[15] S.K. Dhanda, P. Vir, D. Singla, S. Gupta, S. Kumar, G.P.S. Raghava, A Web-Based platform for designing vaccines against existing and emerging strains of

mycobacterium tuberculosis, PloS One 11 (2016), doi:10.1371/journal.pone.0153771.
[16] G. Nagpal, S.S. Usmani, G.P.S. Raghava, A web resource for designing subunit vaccine against major pathogenic species of bacteria, Front. Immunol. 9 (2018), doi:10.3389/fimmu.2018.02280.
[17] A.K. Gupta, K. Kaur, A. Rajput, S.K. Dhanda, M. Sehgal, M.S. Khan, I. Monga, S.A. Dar, S. Singh, G. Nagpal, S.S. Usmani, A. Thakur, G. Kaur, S. Sharma, A. Bhardwaj, A. Qureshi, G.P.S. Raghava, M. Kumar, ZikaVR: an integrated Zika virus resource for genomics, proteomics, phylogenetic and therapeutic analysis, Sci. Rep. 6 (2016), doi:10.1038/srep32713.
[18] M. Tahir Ul Qamar, S. Saleem, U.A. Ashfaq, A. Bari, F. Anwar, S. Alqahtani, Epitope-based peptide vaccine design and target site depiction against Middle East Respiratory Syndrome Coronavirus: an immune-informatics study, J. Transl. Med. 17 (2019), doi:10.1186/s12967-019-2116-8.
[19] A. Grifoni, D. Weiskopf, S.I. Ramirez, J. Mateus, J.M. Dan, C.R. Moderbacher, S.A. Rawlings, A. Sutherland, L. Premkumar, R.S. Jadi, D. Marrama, A.M. de Silva, A. Frazier, A.F. Carlin, J.A. Greenbaum, B. Peters, F. Krammer, D.M. Smith, S. Crotty, A. Sette, Targets of T Cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals, Cell 181 (2020), doi:10.1016/j.cell.2020.05.015.
[20] A. Christiaansen, S.M. Varga, J.V. Spencer, Viral manipulation of the host immune response, Curr. Opin. Immunol. 36 (2015) 54–60, doi:10.1016/j.coi.2015.06.012.
[21] A. Lathwal, R. Kumar, G.P.S. Raghava, Computer-aided designing of oncolytic viruses for overcoming translational challenges of cancer immunotherapy, Drug Discov. Today (2020), doi:10.1016/j.drudis.2020.04.008.
[22] M.M. Tarek, A.E. Shafei, M.A. Ali, M.M. Mansour, Computational prediction of vaccine potential epitopes and 3-dimensional structure of XAGE-1b for non-small cell lung cancer immunotherapy, Biomed. J. 41 (2018) 118–128, doi:10.1016/j.bj.2018.04.002.
[23] R.A. Shey, S.M. Ghogomu, K.K. Esoh, N.D. Nebangwa, C.M. Shintouo, N.F. Nongley, B.F. Asa, F.N. Ngale, L. Vanhamme, J. Souopgui, In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases, Sci. Rep. 9 (2019), doi:10.1038/s41598-019-40833-x.
[24] A. Zawawi, R. Forman, H. Smith, I. Mair, M. Jibril, M.H. Albaqshi, A. Brass, J.P. Derrick, K.J. Else, In silico design of a T-cell epitope vaccine candidate for parasitic helminth infection, PLoS Pathog. 16 (2020), doi:10.1371/journal.ppat.1008243.
[25] N. Wang, J. Shang, S. Jiang, L. Du, Subunit vaccines against emerging pathogenic human coronaviruses, Front. Microbiol. 11 (2020), doi:10.3389/fmicb.2020.00298.
[26] A.C. Galassie, A.J. Link, Proteomic contributions to our understanding of vaccine and immune responses, Proteomics - Clin. Appl. 9 (2015) 972–989, doi:10.1002/prca.201500054.
[27] S.J.R. da Silva, C.T.A. da Silva, R.P.G. Mendes, L. Pena, Role of nonstructural proteins in the pathogenesis of SARS-CoV-2, J. Med. Virol. (2020), doi:10.1002/jmv.25858.
[28] C. Klemm, Y. Boergeling, S. Ludwig, C. Ehrhardt, Immunomodulatory nonstructural proteins of influenza A viruses, Trends Microbiol. 26 (2018) 624–636, doi:10.1016/j.tim.2017.12.006.
[29] H. Singh, H.R. Ansari, G.P.S. Raghava, Improved method for linear B-cell epitope prediction using antigen's primary sequence, PloS One 8 (2013), doi:10.1371/journal.pone.0062216.
[30] ProPred: prediction of HLA-DR binding sites - PubMed, (n.d.)accessedhttps://pubmed.ncbi.nlm.nih.gov/11751237/s 21 July 2020
[31] ProPred1: prediction of promiscuous MHC Class-I binding sites - PubMed, (n.d.)accessedhttps://pubmed.ncbi.nlm.nih.gov/12761064/s 21 July 2020
[32] M. Bhasin, G.P.S. Raghava, Prediction of CTL epitopes using QM, SVM and ANN techniques, Vaccine 22 (2004) 3195–3204, doi:10.1016/j.vaccine.2004.02.005.
[33] G. Nagpal, K. Chaudhary, P. Agrawal, G.P.S. Raghava, Computer-aided prediction of antigen presenting cell modulators for designing peptide-based vaccine adjuvants, J. Transl. Med. 16 (2018), doi:10.1186/s12967-018-1560-1.
[34] Prediction of IL4 inducing peptides - PubMed, (n.d.)accessedhttps://pubmed.ncbi.nlm.nih.gov/24489573/s 21 July 2020
[35] S.K. Dhanda, P. Vir, G.P.S. Raghava, Designing of interferon-gamma inducing MHC class-II binders, Biol. Direct 8 (2013), doi:10.1186/1745-6150-8-30.
[36] S. Singh, H. Singh, A. Tuknait, K. Chaudhary, B. Singh, S. Kumaran, G.P.S. Raghava, PEPstrMOD: structure prediction of peptides containing natural, non-natural and modified residues, Biol. Direct 10 (2015), doi:10.1186/s13062-015-0103-4.
[37] S. S, K. D, S. H, K. M, H. N, Use of amino acid composition to predict epitope residues of individual antibodies, Protein Eng. Des. Sel. 23 (2010), doi:10.1093/PROTEIN/GZQ014.
[38] S. Ebrahimi, H. Mohabatkar, M. Behbahani, Predicting promiscuous T cell epitopes for designing a vaccine against Streptococcus pyogenes, Appl. Biochem. Biotechnol. 187 (2019) 90–100, doi:10.1007/s12010-018-2804-5.
[39] M. Oprea, F. Antohe, Reverse-vaccinology strategy for designing T-cell epitope candidates for Staphylococcus aureus endocarditis vaccine, Biologicals 41 (2013) 148–153, doi:10.1016/j.biologicals.2013.03.001.
[40] C.J. Mitran, S.K. Yanow, The case for exploiting cross-species epitopes in malaria vaccine design, Front. Immunol. 11 (2020), doi:10.3389/fimmu.2020.00335.
[41] H.H. Bui, J. Sidney, W. Li, N. Fusseder, A. Sette, Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines, BMC Bioinf. 8 (2007), doi:10.1186/1471-2105-8-361.
[42] K.W. Wu, C.Y. Chien, S.W. Li, C.C. King, C.H. Chang, Highly conserved influenza A virus epitope sequences as candidates of H3N2 flu vaccine targets, Genomics 100 (2012) 102–109, doi:10.1016/j.ygeno.2012.06.003.
[43] Y. Lei, F. Zhao, J. Shao, Y. Li, S. Li, H. Chang, Y. Zhang, Application of built-in adjuvants for epitope-based vaccines, PeerJ (2019) 2019, doi:10.7717/peerj.6185.
[44] S.N. Sabah, M.A. Gazi, R.A. Sthity, A.B. Husain, S.A. Quyyum, M. Rahman, M.R. Islam, Designing of epitope-focused vaccine by targeting E6 and E7 conserved protein sequences: an immuno-informatics approach in human papillomavirus 58 isolates, Interdiscipl. Sci. Comput. Life Sci. 10 (2018) 251–260, doi:10.1007/s12539-016-0184-5.
[45] The immune epitope database (IEDB) 3.0 - PubMed, (n.d.)accessedhttps://pubmed.ncbi.nlm.nih.gov/25300482/s 21 July 2020
[46] R.M. Welsh, L.K. Selin, No one is naive: the significance of heterologous T-cell immunity, Nat. Rev. Immunol. 2 (2002) 417–426, doi:10.1038/nri820.
[47] R.M. Welsh, R.S. Fujinami, Pathogenic epitopes, heterologous immunity and vaccine design, Nat. Rev. Microbiol. 5 (2007) 555–563, doi:10.1038/nrmicro1709.
[48] S. Gupta, P. Kapoor, K. Chaudhary, A. Gautam, R. Kumar, G.P.S. Raghava, Silico approach for predicting toxicity of peptides and proteins, PloS One 8 (2013), doi:10.1371/journal.pone.0073957.
[49] S. Maurer-Stroh, N.L. Krutz, P.S. Kern, V. Gunalan, M.N. Nguyen, V. Limviphuvadh, F. Eisenhaber, G. Frank Gerberick, AllerCatPro-prediction of protein allergenicity potential from the protein sequence, Bioinformatics 35 (2019) 3020–3027, doi:10.1093/bioinformatics/btz029.
[50] H.V. Cook, N.T. Doncheva, D. Szklarczyk, C. von Mering, L.J. Jensen, Viruses.STRING: a virus-host protein-protein interaction database, Viruses 10 (2018), doi:10.3390/v10100519.
[51] S.K. Dhanda, K. Vaughan, V. Schulten, A. Grifoni, D. Weiskopf, J. Sidney, B. Peters, A. Sette, Development of a novel clustering tool for linear peptide sequences, Immunology 155 (2018) 331–345, doi:10.1111/imm.12984.
[52] S.J. Goodswen, P.J. Kennedy, J.T. Ellis, Enhancing in silico protein-based vaccine discovery for eukaryotic pathogens using predicted peptide-MHC binding and peptide conservation scores, PloS One 9 (2014), doi:10.1371/journal.pone.0115745.
[53] J. Zheng, X. Lin, X. Wang, L. Zheng, S. Lan, S. Jin, Z. Ou, J. Wu, Silico analysis of epitope-based vaccine candidates against hepatitis B virus polymerase protein, Viruses 9 (2017), doi:10.3390/v9050112.
[54] V. Chauhan, T. Rungta, K. Goyal, M.P. Singh, Designing a multi-epitope based vaccine to combat Kaposi Sarcoma utilizing immunoinformatics approach, Sci. Rep. 9 (2019), doi:10.1038/s41598-019-39299-8.
[55] H. Kitamura, C. Sedlik, A. Jacquet, B. Zaragoza, M. Dusseaux, V. Premel, X. Sastre-Garau, O. Lantz, Long peptide vaccination can lead to lethality through CD4 + T cell-mediated cytokine storm, J. Immunol. 185 (2010) 892–901, doi:10.4049/jimmunol.1000933.
[56] B. Manavalan, T.H. Shin, M.O. Kim, G. Lee, E.L. Pip, A new ensemble learning method for improved proinflammatory peptide predictions, Front. Immunol. 9 (2018) 1783, doi:10.3389/fimmu.2018.01783.
[57] M.S. Khatun, M.M. Hasan, W. Shoombuatong, H. Kurata, ProIn-Fuse: improved and robust prediction of proinflammatory peptides by fusing of multiple feature representations, J. Comput. Aided Mol. Des. (2020), doi:10.1007/s10822-020-00343-9.