Zhang *et al. BMC Bioinformatics* 2021, **22**(Suppl 8):40
https://doi.org/10.1186/s12859-021-03962-7

BMC Bioinformatics

SOFTWARE

Open Access

TANTIGEN 2.0: a knowledge base of tumor T cell antigens and epitopes

Guanglan Zhang[1*], Lou Chitkushev[1], Lars Rønn Olsen[2], Derin B. Keskin[3] and Vladimir Brusic[4]

From 3rd International Workshop on Computational Methods for the Immune System Function (CMISF 2019) San Diego, CA, USA. 18-21 November 2019

*Correspondence:
guanglan@bu.edu
[1] Metropolitan College, Boston University, Boston, USA
Full list of author information is available at the end of the article

Abstract

We previously developed TANTIGEN, a comprehensive online database cataloging more than 1000 T cell epitopes and HLA ligands from 292 tumor antigens. In TANTIGEN 2.0, we significantly expanded coverage in both immune response targets (T cell epitopes and HLA ligands) and tumor antigens. It catalogs 4,296 antigen variants from 403 unique tumor antigens and more than 1500 T cell epitopes and HLA ligands. We also included neoantigens, a class of tumor antigens generated through mutations resulting in new amino acid sequences in tumor antigens. TANTIGEN 2.0 contains validated TCR sequences specific for cognate T cell epitopes and tumor antigen gene/mRNA/protein expression information in major human cancers extracted by Human Pathology Atlas. TANTIGEN 2.0 is a rich data resource for tumor antigens and their associated epitopes and neoepitopes. It hosts a set of tailored data analytics tools tightly integrated with the data to form meaningful analysis workflows. It is freely available at http://projects.met-hilab.org/tadb.

Keywords: Neoepitope, Neoantigen, Tumor antigen, Cancer vaccine, Immunotherapy, T cell epitope prediction

Background

Advances in instrumentation and progress in immuno-oncology are driving a revolution in cancer care. New cancer treatment methods are emerging–targeted immunotherapies are among the most promising treatment options. Checkpoint blocking antibodies are currently providing stable cures to a subset of patients that could not be helped previously [1]. Chimeric antigen receptor (CAR) T cell/adoptive T cell therapies have been shown effective in some terminally ill patients [2]. Neoantigens are newly formed protein antigens that occur in individual patients, either from somatic mutations of genes or viral genes incorporated in the infected cell genome. These freshly emerged genes encode proteins that contain new T cell epitopes capable of inducing tumor-specific T cell recognition. Furthermore, it was demonstrated that the recognition of even a

small number of epitopes by T cells might be sufficient to reject tumors in terminally ill patients through adoptive T cell therapy [3, 4].

T cells recognize antigens through T cell receptor (TCR), a surface protein composed of multiple peptide chains. A TCR has such high antigen specificity that it can recognize its cognate targets at a level of single amino acid difference. A highly diverse repertoire of TCR sequences ensures the effectiveness of the adaptive immune system [5, 6]. It was shown that T cell repertoire could serve as a biomarker of immune responses in cancer patients [7, 8]. TCR sequences with validated cognate antigens are essential for cell-based vaccine design. Vaccines that target neoantigens and personalized cancer immunotherapies are considered as the current clinical and research frontier in immuno-oncology [9, 10].

Tumor derived antigens that induce productive antitumor immune responses are known as tumor antigens (TAs) [11]. TAs can be divided into two main groups, tumor-specific antigens (TSAs) and tumor-associated antigens (TAAs). TSAs may be mutated tumor neoantigens [12], cancer-testis antigens [13], or oncofetal proteins [14]. TSAs are exclusively found in tumors and are not expressed in normal tissues. TAAs are normal proteins that are overexpressed in tumor cells as compared to the expression level in healthy cells [15]. Neoantigens are TAs that are no longer self. They are not tolerated by T cell immunity and are exclusively tumor specific. TAs have been extensively studied and offer high promise for cancer therapeutics design and serve as cancer diagnosis targets [9, 10, 16, 17]. The therapeutic landscape of cancer has recently been transformed by the emergence of effective immunotherapies [18, 19]. Despite these advances, any one form of immunotherapy studied and used to date was shown to benefit only a subset of patients. These immunotherapies facilitate T cell mediated immunity against the tumors. However, we lack an understanding of what defines the specificity of protective T cell immune responses they generate. Correct identification and cataloging of T cell antigens that aid tumor rejection will allow the development of highly effective personalized cancer vaccine immunotherapies and understand the protection mechanism. Current epitope prediction algorithms only forecast HLA processing and presentation but cannot predict antigenicity. They produce large numbers of positives that are biochemically active but functionally inert [20]. Because not all HLA presented epitopes are recognized by T cells, their inclusion in the vaccines and immunotherapies may reduce immunotherapy effectiveness. Validated data sets of T cell epitopes with the associated HLA restriction may also allow the development of improved epitope prediction algorithms [20, 21]. TCR sequences specific for TAs can be utilized to generate antitumor T cells for adoptive T cell therapy or as templates for building models of TCR-antigen recognition for neoantigens that show high similarity to known immune response targets.

To support the development of rationally designed epitope-based cancer vaccines, we previously developed TANTIGEN, a comprehensive web-based database cataloging more than 1000 T cell epitopes and HLA ligands from 292 different tumor T cell antigens [22]. In the current build, TANTIGEN 2.0 [23], we extended the coverage of immune response targets in the original set of TAs, added more than 100 new TAs, and a selection of their T cell epitopes and ligands. All T cell epitopes and neoepitopes included in TANTIGEN 2.0 have been experimentally validated.

Implementation

Data collection, annotation, and organization

We assembled TANTIGEN 2.0 by compiling the new data from the Cancer Antigenic Peptide Database [24] and recent publications reporting neoepitopes and neoantigens. Previously we included mRNA expression information of TAs by providing EST profiles from UniGene. However, NCBI retired the UniGene web pages in July 2019. Large scale open-access efforts, such as the Cancer Genome Atlas (TCGA) and the Human Protein Atlas (HPA), provided data for genome-wide expression analysis of individual genes in different tissues and cancers [25, 26]. The Human Pathology Atlas that was created as a part of the Human Protein Atlas program analyzed and cataloged expression profiles at both RNA and protein levels. Protein expression data are shown in normal tissue and major human cancers [27]. In TANTIGEN 2.0, we utilized data from the HPA to enrich available information on TAs. TCR sequences were collected from McPAS-TCR, a TCR sequence database associated with various pathologies and antigens based on published literature [28]. It contains more than 5,000 sequences of TCRs associated with multiple pathologic conditions, including cancer and their respective antigens in humans and mice. To ensure data integrity and avoid the proliferation of data errors from external databases, we manually checked the data against the original publications. For example, in McPAS-TCR, an HLA-A*0201 restricted T cell epitope YLEPGPVTA (IEDB ID: 74,638) was mistakenly described as A*01. The error was corrected when data were integrated into TANTIGEN 2.0.

Bioinformatics tools

A set of bioinformatics tools has been integrated into TANTIGEN 2.0 to streamline data analysis. Users can look for antigens, epitopes, and HLA ligands using keyword searches. Sequence similarity searches can be performed using BLAST (Basic Local Alignment Search Tool) [29]. Sequence homology can be examined using multiple sequence alignment by MAFFT [30]. On-the-fly HLA binding prediction tools for 15 common HLA class I and class II alleles were integrated into TANTIGEN 2.0. They facilitate analysis of known immunogenicity in conjunction with predicted HLA binding and the prediction of additional potential epitopes. TANTIGEN 2.0 has a set of visualization tools that display the locations of peptides within their parent proteins. For each protein that contains point mutations, an interactive visualization tool shows a map of mutations in the tumor antigen sequences to provide a global view of all reported mutations for a given tumor antigen. For neoantigen entries, both the neoantigen fragment and the native sequence (called reference sequence) are included.

Webserver

The TANTIGEN webserver interface was constructed using the KB-builder framework, which streamlines the development and deployment of web-accessible immunological databases [31]. TANTIGEN 2.0 serves as an integrator of relevant information critical for the design of highly personalized cancer vaccines and immunotherapies.

Results

In TANTIGEN 2.0, we cataloged more than 1000 validated tumor T cell epitopes and over 500 HLA binding ligands (1,676 in total). The database hosts more than 4,000 antigen variants from 403 unique TAs reported in literature. Each record contains the information on tumor antigen sequence, variants (splice isoforms and mutation variants), known T-cell epitopes and HLA ligands, TCR information (if available), and literature references.

Around 50% of the antigens (201 out of 403) have substitution mutations, with about 25% (95) having more than one. The antigens have more than 100 substitution mutations include TP53 (1,311 mutations), CDKN2A (239), EGFR (169), CTNNB1 (111), TYR

a

AgACC	Ag004463
Date	6-15-2020
Last updated	6-15-2020
Antigen Name	ERBB2
Common Name	HER2
Full Name	Receptor tyrosine-protein kinase erbB-2
Synonym	CD340 antigen; MLN 19; NEU proto-oncogene; Tyrosine kinase-type cell surface receptor HER2; c-erb B2; c-erbB2/neu protein; neuroblastoma/glioblastoma derived oncogene homolog; tyrosine kinase-type cell surface receptor; v-erb-b2 avian erythroblastic leukemia viral oncogene homolog; neuro/glioblastoma derived oncogene homolog; v-erb-b2 erythroblastic leukemia viral oncogene homolog; neuo/glioblastoma derived oncogene homolog (avain); c-erbB-2; EC 2.7.10.1; HER-2; HER-2/neu; HER2; NEU; NGL; TKR1; erb-2; herstatin; p185erbB2 another display format
Annotation	This is a fragment sequence.
Isoforms	Ag000001, Ag001732 Alignment of all Isoforms
Mutation entries	Ag001291, Ag001292, Ag001293, Ag001294, Ag001295, Ag001296, Ag001297, Ag001298, Ag001299, Ag001300, Ag001301, Ag001302, Ag001303, Ag001304, Ag001305, Ag001306, Ag001307, Ag001308, Ag001309, Ag001310, Ag001311, Ag001312, Ag001313, Ag001314, Ag002176, Ag002177, Ag002178, Ag002179, Ag002180, Ag002181, Ag004463 View mutation map
RNA/protein expression profile	RNA and protein Expression profile

T cell epitope	**Epitope sequence**	**Position**	**HLA allele**	**Reference**
	ALIHHNTYL	14-22	A*0201	26389673 30418433

Predicted HLA binders	Allele: HLA-A*02:01 predictions Peptide length: 9 Submit
Reference sequence	LGLRSLRELGSGLALIHHNTHLCFVHTVPWDQLFRNPHQAL
Antigen sequence	LGLRSLRELGSGLALIHHNTYLCFVHTVPWDQLFRNPHQAL

b

Mutation map of ERBB2

```
Reference sequence: Ag000001
          10        20        30      450       460       470
          |         |         |       |         |         |
MELAALCRWGLLLALLPPGAASTQVCTGTDMK •••  ISWLGLRSLRELGSGLALIHHNTHLC •••
................................      ..C.....................Y..
................................      ............................
```

Fig. 1 Example pages with neoantigen and neoepitope information: **a** HER2 neoantigen with a substitute mutation H473Y; **b** Part of the mutation Map; **c** HLA A*0201 restricted neoepitope; **d** On-the-fly HLA binding prediction with side-by-side comparison with the reference sequence. Predicted weak and strong binders are highlighted in green and magenta respectively; **e** TCR information

c

TACC	T001121
Epitope Sequence	ALIHHNTYL
HLA allele	A*0201
Reference	26389673 30418433
Experimental method	*in vivo*
Genomic Origin	neoepitope
TCR CD3 sequences	Click here

Ag004463 is the only antigen containing the HLA ligand. The HLA ligand is highlighted in yellow in the antigen s

Ag004463_ERBB2 LGLRSLRELGSGLALIHHNTYLCFVHTVPWDQLFRNPHQAL

d

RANK	Position	Peptide	IC50 (nM)	Experimental Status	Ref Peptide	Ref Pep IC50 (nM)	Ref Pep Experimental Status
1	1-9	LGLRSLREL	25184.06		LGLRSLREL	25184.06	
2	2-10	GLRSLRELG	34397.84		GLRSLRELG	34397.84	
3	3-11	LRSLRELGS	41518.95		LRSLRELGS	41518.95	
4	4-12	RSLRELGSG	39338.03		RSLRELGSG	39338.03	
5	5-13	SLRELGSGL	1929.23		SLRELGSGL	1929.23	
6	6-14	LRELGSGLA	39301.87		LRELGSGLA	39301.87	
7	7-15	RELGSGLAL	14772.45		RELGSGLAL	14772.45	
8	8-16	ELGSGLALI	1263.84		ELGSGLALI	1263.84	
9	9-17	LGSGLALIH	39846.52		LGSGLALIH	39846.52	
10	10-18	GSGLALIHH	41222.64		GSGLALIHH	41222.64	
11	11-19	SGLALIHHN	41754.11		SGLALIHHN	41754.11	
12	12-20	GLALIHHNT	1315.84		GLALIHHNT	1315.84	
13	13-21	LALIHHNTY	31959.98		LALIHHNTH	35749.43	
14	14-22	ALIHHNTYL	13.65	T-cell epitope	ALIHHNTHL	71.73	
15	15-23	LIHHNTYLC	9407.26		LIHHNTHLC	15154.51	

e

TACC	T001121							
Epitope Sequence	ALIHHNTYL							
TCR Information	CD3 alpha AA	CD3 bata AA	TRAV	TRAJ	TRBV	TRBD	TRBJ	Reference
	CAVVLNSGYSTLTF	CASSDLQENQPQHF	TRAV8-3	NA	TRBV19	NA	NA	30418433
	CALSEAGYSSASKIIF	CSVVGTGSVITNEKLFF	TRAV19	NA	TRBV29-1	NA	NA	30418433

Fig. 1 continued

(111), and APC (103). The TA variants include 879 full-length sequences, 243 partial sequences, and 3,175 short fragments resulted from identified substitution mutations. Most defined short fragments are 41 aa long sequences with the mutation in the middle (position 21). Sometimes it is impossible to place the mutation position in the middle of the sequence. For example, in Ag000048, the mutation is at position 12 in the reference antigen. The length of these fragments is kept the same at 41. The reference sequence fragment is included for comparison with the mutated fragment; the mutation site is highlighted (Fig. 1a). The substitution mutation H473Y resulted in an A*0201 restricted neoepitope, as shown in Fig. 1b, c. The HLA binding prediction comparison in Fig. 1d shows that the peptide becomes a strong HLA-A*0201 binder instead of a weaker one due to the mutation. The TCR information for the neoepitope is shown in Fig. 1d.

Conclusions

TANTIGEN 2.0 provides a rich data resource for tumor associated epitope and neoepitope discovery studies. It hosts more than 4,000 antigen variants from 403 unique TAs (a 38% increase comparing to TANTIGEN) and 1,676 T cell epitopes and HLA ligands (a 46% increase). TCR information for some T cell epitopes is included too. Integrated computational analysis tools in TANTIGEN 2.0 enable users to combine data

and domain knowledge, use tailored bioinformatics tools, and simulate experiments. It represents a rich information resource for the study of cancer immunology and immunotherapy. All data and tools described here are available in TANTIGEN 2.0, an interactive open-access database (http://projects.met-hilab.org/tadb). The primary purpose of TANTIGEN2.0 is to support the design of neoantigen vaccine-based cancer immunotherapies. Immunological peptides from cancer-causing human viruses, such as Epstein-Barr virus (EBV), Human Papillomavirus (HPV), and Merkel Cell Polyomavirus (MCV), were not included in TANTIGEN. We developed EBVdb, HPVdb, and MCVdb to support studies on T cell immunology of EBV, HPV, and MCV [32–34].

Availability and requirements

Project name: TANTIGEN 2.0
Project home page: http://projects.met-hilab.org/tadb
Operating system(s): Platform independent.
Programming language: Perl and PHP.
Other requirements: None.
License: Not applicable.
Any restrictions to use by non-academics: None.

Abbreviations
BLAST: Basic local alignment search tool; CAR: Chimeric antigen receptor; EBV: Epstein-Barr virus; HLA: Human leukocyte antigen; HPA: Human protein atlas; HPV: Human papillomavirus; IEDB: Immune epitope database and analysis resource; MCV: Merkel cell polyomavirus; NCBI: National center for biotechnology information; TA: Tumore antigen; TAA: Tumor-associated antigen; TCGA: The cancer genome atlas; TCR: T cell receptor; TSA: Tumor-specific antigen.

Acknowledgments
Tantigen 2.0 was presented at the 2019 IEEE International Conference on Bioinformatics and Biomedicine.

About this supplement
This article has been published as part of BMC Bioinformatics Volume 22 Supplement 8 2021: Selected papers from the 3rd International Workshop on Computational Methods for the Immune System Function (CMISF 2019) – part 2. The full contents of the supplement are available at https://bmcbioinformatics.biomedcentral.com/articles/supplements/volume-22-supplement-8.

Authors' contributions
VB, GLZ, and DBK designed the study. GLZ. LRO and LC implemented the database. All authors wrote the manuscript and approved the final manuscript.

Funding
DBK is supported by NIH/NCI R21 CA216772-01A1 and NCI-SPORE-2P50CA101942-11A1. Publication costs are funded by Boston University.

Availability of data and materials
All data are publicly available at http://projects.met-hilab.org/tadb. The organized datasets are available from the corresponding author on reasonable request.

Ethics approval and consent to participate
No ethics approval was required for the study.

Consent for publication
Not applicable.

Competing interests
DBK has previously advised and received consulting fees from Neon Therapeutics. DBK and GLZ own equity in Aduro Biotech, Agenus Inc., Armata pharmaceuticals, Breakbio Corp., Biomarin Pharmaceutical Inc., Bristol Myers Squibb Com., Celldex Therapeutics Inc., Editas Medicine Inc., Exelixis Inc., Gilead Sciences Inc., IMV Inc., Lexicon Pharmaceuticals Inc., Moderna Inc., Regeneron Pharmaceuticals, and Stemline Therapeutics Inc.

Author details
[1] Metropolitan College, Boston University, Boston, USA. [2] Department of Health Technology, Technical University of Denmark, Lyngby, Denmark. [3] Dana-Farber Cancer Institute, Harvard Medical School, Boston, USA. [4] School of Computer Science, University of Nottingham, Ningbo, China.

Received: 6 January 2021 Accepted: 8 January 2021
Published: 14 April 2021

References

1. Paschen A, Schadendorf D. The era of checkpoint inhibition: lessons learned from melanoma. Recent Results Cancer Res. 2020;214:169–87.
2. Dave H, Jerkins L, Hanley PJ, Bollard CM, Jacobsohn D. Driving the CAR to the bone marrow transplant program. Curr Hematol Malignancy Rep. 2019;14(6):1–9.
3. Rosenberg SA. Raising the bar: the curative potential of human cancer immunotherapy. Sci Transl Med. 2021;4(127):8.
4. Zacharakis N, Chinnasamy H, Black M, Xu H, Lu YC, Zheng Z, et al. Immune recognition of somatic mutations leading to complete durable regression in metastatic breast cancer. Nat Med. 2018;24(6):724–30.
5. Robins HS, Srivastava SK, Campregher PV, Turtle CJ, Andriesen J, Riddell SR, et al. Overlap and effective size of the human CD8+ T cell receptor repertoire. Sci Transl Med. 2010;2(47):ra 64.
6. Laydon DJ, Bangham CR, Asquith B. Estimating T-cell repertoire diversity: limitations of classical estimators and a new approach. Philos Trans R Soc B Biol Sci. 2015;370(1675):20140291.
7. Cui JH, Lin KR, Yuan SH, Jin YB, Chen XP, Su XK, et al. TCR repertoire as a novel indicator for immune monitoring and prognosis assessment of patients with cervical cancer. Front Immunol. 2018;9:2729.
8. Hopkins AC, Yarchoan M, Durham JN, Yusko EC, Rytlewski JA, Robins HS, et al. T cell receptor repertoire features associated with survival in immunotherapy-treated pancreatic ductal adenocarcinoma. JCI insight 2018;3(13).
9. Ott PA, Hu Z, Keskin DB, Shukla SA, Sun J, Bozym DJ, et al. An immunogenic personal neoantigen vaccine for patients with melanoma. Nature. 2017;547(7662):217–21.
10. Keskin DB, Anandappa AJ, Sun J, Tirosh I, Mathewson ND, Li S, et al. Neoantigen vaccine generates intratumoral T cell responses in phase Ib glioblastoma trial. Nature. 2019;565(7738):234–9.
11. Rosenberg SA, Yang JC, Restifo NP. Cancer immunotherapy: moving beyond current vaccines. Nat Med. 2004;10(9):909–15.
12. Hacohen N, Fritsch EF, Carter TA, Lander ES, Wu CJ. Getting personal with neoantigen-based therapeutic cancer vaccines. Cancer Immunol Res. 2013;1(1):11–5.
13. Boon T, Cerottini JC, Van den Eynde B, van der Bruggen P, Van Pel A. Tumor antigens recognized by T lymphocytes. Annu Rev Immunol. 1994;12(1):337–65.
14. Coggin JH Jr, Barsoum AL, Rohrer JW. 37 kiloDalton oncofetal antigen protein and immature laminin receptor protein are identical, universal T-cell inducing immunogens on primary rodent and human cancers. Anticancer Res. 1999;19(6C):5535.
15. Andersen MH, Pedersen LØ, Becker JC, thor Straten P. Identification of a cytotoxic T lymphocyte response to the apoptosis inhibitor protein survivin in cancer patients. Cancer Research. 2001;61(3):869–72.
16. Sahin U, Derhovanessian E, Miller M, Kloke BP, Simon P, Löwer M, et al. Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. Nature. 2017;547(7662):222–6.
17. Hilf N, Kuttruff-Coqui S, Frenzel K, Bukur V, Stevanović S, Gouttefangeas C, et al. Actively personalized vaccination trial for newly diagnosed glioblastoma. Nature. 2019;565(7738):240–5.
18. Brahmer JR, Tykodi SS, Chow LQ, Hwu WJ, Topalian SL, Hwu P, et al. Safety and activity of anti–PD-L1 antibody in patients with advanced cancer. N Engl J Med. 2012;366(26):2455–65.
19. Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, et al. Safety, activity, and immune correlates of anti–PD-1 antibody in cancer. N Engl J Med. 2012;366(26):2443–54.
20. Abelin JG, Keskin DB, Sarkizova S, Hartigan CR, Zhang W, Sidney J, et al. Mass spectrometry profiling of HLA-associated peptidomes in mono-allelic cells enables more accurate epitope prediction. Immunity. 2017;46(2):315–26.
21. Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, Nielsen M. NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8–11. Nucl Acids Res. 2008;36(2):W509-12.
22. Olsen LR, Tongchusak S, Lin H, Reinherz EL, Brusic V, Zhang GL. TANTIGEN: a comprehensive database of tumor T cell antigens. Cancer Immunol Immunother. 2017;66(6):731–5.
23. Zhang G, Chitkushev L, Keskin DB, Brusic V. TANTIGEN 2.0: an online database and analysis platform for tumor T cell antigens. In 2019 IEEE international conference on bioinformatics and biomedicine (BIBM) 2019;2228–2231.
24. Vigneron N, Stroobant V, Van den Eynde BJ, van der Bruggen P. Database of T cell-defined human tumor antigens: the 2013 update. Cancer Immunity Archive 2013;13(3).
25. Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, et al. The cancer genome atlas pan-cancer analysis project. Nat Genet. 2013;45(10):1113.
26. Uhlen M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Tissue-based map of the human proteome. Science. 2015;347:6220.
27. Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhori G, et al. A pathology atlas of the human cancer transcriptome. Science. 2017;357:6352.
28. Tickotsky N, Sagiv T, Prilusky J, Shifrut E, Friedman N. McPAS-TCR: a manually curated catalogue of pathology-associated T cell receptor sequences. Bioinformatics. 2017;33(18):2924–9.
29. Mount DW. Using the basic local alignment search tool (BLAST). Cold Spring Harbor Protocols. 2007;2007(7):17.

30. Katoh K, Toh H. Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinform. 2008;9(4):286–98.
31. Zhang GL, Sun J, Chitkushev L, Brusic V. Big data analytics in immunology: a knowledge-based approach. BioMed Res Int 2014;2014.
32. Zhang GL, Keskin D, Chitkushev L, Reinherz EL, Brusic V. EBVdb: a data repository and analysis platform for knowledge discovery in Epstein-Barr virus with applications in T cell immunotherapy. ICSI3 2015, July 17–18, Taomina, Italy.
33. Zhang GL, Riemer AB, Keskin DB, Chitkushev L, Reinherz EL, Brusic V. HPVdb: a data mining system for knowledge discovery in human papillomavirus with applications in T cell immunology and vaccinology. Database. 2014;2014.
34. Zhang GL, Keskin DB, DeCaprio JA, Wu CJ, Chitkushev L, Brusic V. MCVdb: A database for knowledge discovery in Merkel cell polyomavirus with applications in T cell immunology and vaccinology. In 2017 IEEE international conference on bioinformatics and biomedicine (BIBM) 2017;1483–1488.

Publisher's Note