

Colorectal cancer vaccines: in silico identification of tumor-specific antigens associated with frequent HLA-I alleles in the costan Rican Central Valley population

Vacunas contra el cáncer colorrectal: identificación en silico de antígenos específicos de tumores asociados con alelos HLA-I frecuentes en la población del Valle Central de Costa Rica

Diego Morazán-Fernández¹, José Arturo Molina-Mora²

Morazán-Fernández, D.; Molina-Mora, J.A. Colorectal cancer vaccines: in silico identification of tumor-specific antigens associated with frequent HLA-I alleles in the costan Rican Central Valley population *Tecnología en Marcha*. Vol. 35, special issue. November, 2022. International Work Conference on Bioinspired Intelligence. Pág. 83-92.

 <https://doi.org/10.18845/tm.v35i8.6458>

1 Laboratorio clínico, Hospital Nacional de Niños, San José. Costa Rica
2 Facultad de Microbiología, Universidad de Costa Rica, San José. Costa Rica.
Centro de Investigación de Enfermedades Tropicales and Facultad de Microbiología, Universidad de Costa Rica, San José. Costa Rica.
E-mail: jose.molinamora@ucr.ac.cr
 <https://orcid.org/0000-0001-9764-4192>

Keywords

Neopeptide; Colorectal cancer; Cancer vaccine; Single nucleotide variants; HLA.

Abstract

Colorectal cancer is a complex disease in which uncontrolled growth of abnormal cells occurs in the large intestine (colon or rectum). The study of tumor-specific antigens (neoantigens), molecules that interact with the immune system, has been extensively explored as a possible therapy called in silico cancer vaccine. Cancer vaccine studies have been triggered by the current high-throughput DNA sequencing technologies. However, there is no universal bioinformatic protocol to study tumor-antigens with DNA sequencing data.

We propose a bioinformatic protocol to detect tumor-specific antigens associated with single nucleotide variants (SNVs) or “mutations” in colorectal cancer and their interaction with frequent HLA alleles (complex that present antigens to immune cells) in the Costa Rican Central Valley population. We used public data of human exome (DNA regions that produce functional products, including proteins). A variant calling analysis was implemented to detect tumor-specific SNVs, in comparison to healthy tissue. We then predicted and analyzed the peptides (protein fragments, the tumor specific antigens) derived from these variants, in the context of its affinity with frequent alleles of HLA type I of the Costa Rican population.

We found 28 non-silent SNVs, present in 26 genes. The protocol yielded 23 strong binders peptides derived from the SNVs for frequent alleles (greater than 8%) for the Costa Rican population at the HLA-A, B and C loci. It is concluded that the standardized protocol was able to identify neoantigens and this can be considered a first step for the eventual design of a colorectal cancer vaccine for Costa Rican patients. To our knowledge, this is the first study of an in silico cancer vaccine using DNA sequencing data in the context of the Costa Rican HLA alleles.

Palabras clave

Neopéptido; cáncer colorrectal; vacuna contra el cáncer; variantes de un solo nucleótido; HLA.

Resumen

El cáncer colorrectal es una enfermedad compleja en la que se produce un crecimiento descontrolado de células anormales en el intestino grueso (colon o recto). El estudio de antígenos específicos de tumor (neoantígenos), moléculas que interactúan con el sistema inmunológico, se ha explorado ampliamente como una posible terapia llamada vacuna contra el cáncer en silico. Los estudios de vacunas contra el cáncer han sido impulsados por las tecnologías actuales de secuenciación de ADN de alto rendimiento. Sin embargo, no existe un protocolo bioinformático universal para estudiar antígenos tumorales con datos de secuenciación de ADN.

Proponemos un protocolo bioinformático para detectar antígenos específicos de tumores asociados con variantes de un solo nucleótido (SNV) o “mutaciones” en el cáncer colorrectal y su interacción con alelos HLA frecuentes (complejo que presenta antígenos a las células inmunes) en el Valle Central de Costa Rica. población. Usamos datos públicos del exoma humano (regiones de ADN que producen productos funcionales, incluidas proteínas). Se implementó un análisis de llamada variante para detectar SNV específicos del tumor, en comparación con el tejido sano. A continuación, predecimos y analizamos los péptidos (fragmentos de proteínas, los antígenos específicos del tumor) derivados de estas variantes, en el contexto de su afinidad con los alelos frecuentes de HLA tipo I de la población costarricense.

Encontramos 28 SNV no silenciosos, presentes en 26 genes. El protocolo produjo 23 péptidos ligantes fuertes derivados de los SNV para alelos frecuentes (más del 8%) para la población costarricense en los loci HLA-A, B y C. Se concluye que el protocolo estandarizado logró identificar neoantígenos y esto puede considerarse un primer paso para el eventual diseño de una vacuna contra el cáncer colorrectal para pacientes costarricenses. Hasta donde sabemos, este es el primer estudio de una vacuna contra el cáncer en silico que utiliza datos de secuenciación de ADN en el contexto de los alelos HLA de Costa Rica.

Introduction

Colorectal cancer is a complex disease in which uncontrolled growth of ab-normal cells occurs in the large intestine (colon or rectum). In 2018, colorectal cancer worldwide reached an incidence of 1 849 518 cases and 880 792 deaths, occupying the third and second cause of cancer incidence and deaths respectively [1]. Cancer vaccines and others immunotherapies have emerged as an alternative to treat this type of cancer. As summarized in Fig. 1, there is evidence that immune system cells, such as CD8+ cytotoxic lymphocytes, can recognize tumor-specific peptides or proteins fragments derived from gene mutations or variants in malignant cells. In the immune response, these neopeptides are presented by cells (cancer or normal) to CD8+ cells. The antigen presentation occurs with the participation of HLA, the human major histocompatibility complex in tumor (and normal) cells, and the TCR (T-cell receptor) in CD8+ cells. If the antigen is tumor-specific, the death of tumor cells is induced [2].

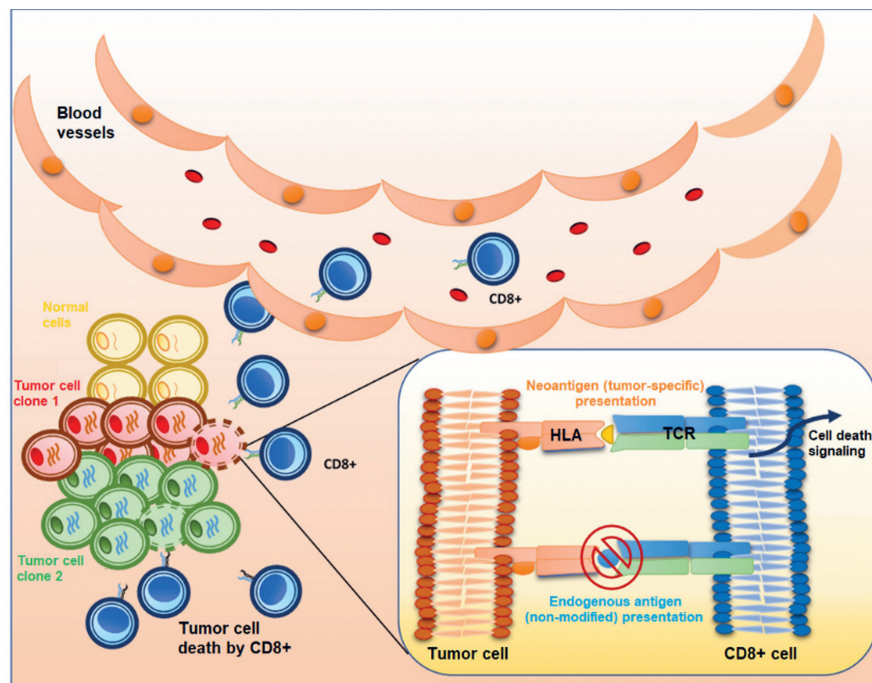


Figure 1. Conceptualization of the biological meaning of the neoantigens identification in cancer vaccines. The tumor-specific antigens (neoantigens) are originated from variants (mutations) in the genes of tumor cells, which are selected as candidate peptides for the vaccine. Peptides that are predicted to be presented in the HLA to cytotoxic CD8+ cells (HLA-antigen-TCR), are kept for the subsequent analyses. If the antigen is an endogenous antigen (non-modified and also present in normal cells), it is excluded. The eventual signaling in the immune response will induce death of tumor cells. Own elaboration.

With the advance of high-throughput DNA sequencing technologies, cancer vaccines and tumor/ immune system interactions cells have been extensively studied [3-5]. Cancer vaccine studies must consider HLA alleles (“gene versions”) that are frequent in the target population, the HLA type (type I or HLA-I for interaction with CD8+ cells), the peptide size (between 8-11 amino acids for HLA-I) and others [6]. However, the data complexity and the specific steps do not allow the application of universal protocols. Benchmark strategies are recommended for specific cases and populations [7-9]. Thus, the aim of this work was to implement a bioinformatic protocol of an in silico cancer vaccine to identify tumor-specific peptides in colorectal cancer considering the context of frequent HLA-I alleles from the Costa Rican Central Valley Population.

Methods

General workflow of the analysis is shown in Fig. 2.

Sequence data of tumor samples and pre-processing

Raw data from chromosome 1 of 7 exomes from human colorectal cancer (4 primary tumor and 3 from liver metastasis) and normal mucosa samples (all from same patient) were downloaded from NCBI Bioproject PRJNA271316 CHET9 experiment (sequenced by Illumina® technology, access: www.ncbi.nlm.nih.gov/bioproject/PRJNA271316). The exome reads were downloaded in Fastq format for subsequent analyzes. Quality control was done using FastQC, and trimming of reads was done with Trimmomatic, including filtering by quality and adapters removal [10].

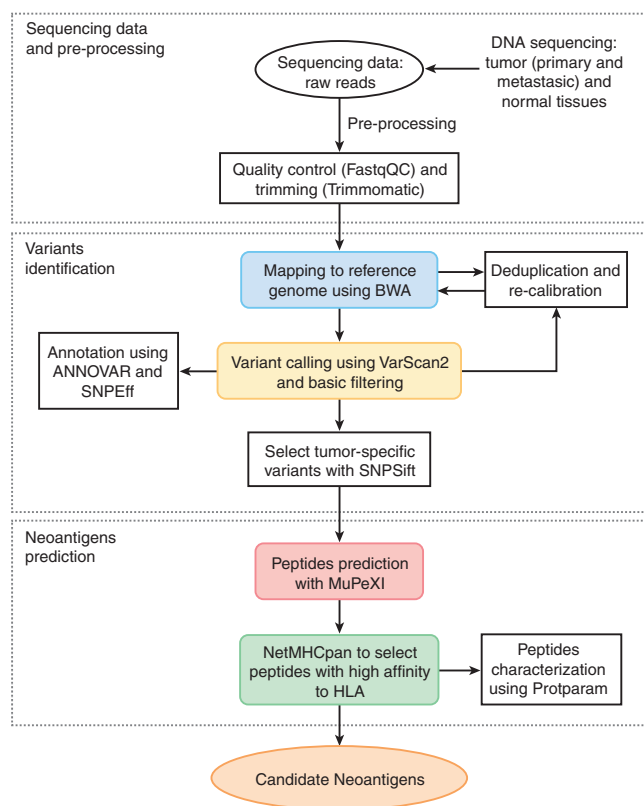


Figure 2. Bioinformatic pipeline to identify neoantigens in primary and metastatic tumor samples of colorectal cancer. The protocol includes three main steps: pre-processing, variant calling and neoantigens (peptides) identification.

Mapping to human genome and Variant Calling analysis

To map reads to the human reference genome (version hg19), the BWA program [11] was used with default parameters. Samtools [12] was used for extra data manipulation as usual. Picard (<http://broadinstitute.github.io/picard/>) was used to evaluate the quality of the mapping. VarScan2 program [13] was used to detect SNVs with quality Q higher than 20 in all samples; other parameters were used by default in the deduplication and re-calibration steps. In order to select mutations present in all the tumor samples, VCFlib (<https://github.com/vcflib/vcflib>) was used. Subsequently, the SnpEff program [14] was used to evaluate and compare the variants between samples. SnpSift [15] was used to filter all those SNVs that were in the normal mucosa sample to finally obtain the tumor exclusive variants. Finally, these mutations were annotated with ANNOVAR [16], obtaining the description of the SNVs and their possible effects on genes. The Atlas repository of colorectal cancer was used to see the involvement of the genes with selected SNVs and the possible metabolic or signaling pathways affected [17].

Putative tumor-specific peptides (neoantigens) identification

Selected tumor-specific variants were analyzed with MuPeXi package [18], to generate the peptides (8-11 amino acid residues) to the different alleles of HLA type I and the possible priorities of related mutant peptides. We chose those HLA type I alleles with frequencies greater than 8% for the Costa Rican Central Valley population [19], using Allele Frequency Net Database (<http://www.allelefrequencies.net/default.asp>). Due to HLA data had a resolution of two-digits, HLA supertypes (four-digits) were considered.

Mutant peptides were analyzed with the NetMHCpan program [20], to establish the affinity of these peptides to the respective alleles of HLA type I by means of the rank percentile (% rank). The 0.5% threshold was used for strong binders. Finally, the characterization of peptides was done using the ProtParam tool of the ExPasy server [21], which provides physicochemical properties.

Results and discussion

Identification of neoantigens or tumor-specific antigens was described in the end of the last century [22]. DNA sequencing technologies and bioinformatic protocols have triggered the understanding of the interactions of tumor cells with the immune system, including the design of cancer vaccines. This strategy is not suitable for all cancer types due to a relatively high mutation rate is expected to generate new antigens. In colorectal cancer, melanoma and other tumor types, a high number of neoantigens is associated with patient response to immune therapies, making them suitable for cancer vaccines.

In this context, to detect possible variants with potential to induce tumor-specific peptides in samples of colorectal tumors, a bioinformatic protocol was implemented with samples of human colorectal cancer for chromosome 1. We identified of 395 SNVs shared by all tumor samples (and absent in the normal mucosa sample).

Table 1. Strong binding peptides for most frequent HLA alleles of the Costa Rican population. Candidates are derived from the selected SNVs in the primary and metastatic tumor samples (pI: Isoelectric point, GRAVY: Grand average of hidropathy).

Alleles	Peptides	pI	Acidic nature	GRAVY	Lenght (AA)
A*02:01	KLVGFSLDV	5.84	acidic	1.133	9
	LLIHCDAYL	5.80	acidic	1.356	9
	YLHSPMYFFI	6.74	neutral	0.76	10
	ILLIHCDAYL	5.08	acidic	1.67	10
A*03:01	AMNILEINEK	4.53	acidic	-0.14	10
	GSSFYALEK	6.00	neutral	-0.256	9
A*24:02	AYLHSPMYFF	6.78	neutral	0.49	10
	AYLHSPMYFFI	6.78	neutral	0.855	11
	CDAYLHSPMYF	5.08	acidic	0.1	11
	DAYLHSPMYF	5.08	acidic	-0.14	10
	DAYLHSPMYFF	5.08	acidic	0.127	11
	RWYSTPSSYL	8.59	basic	-0.89	10
	YLHSPMYFF	6.74	neutral	0.344	9
B*07:02	GPLSSEKAAM	6.00	neutral	-0.17	10
B*35:01	DAYLHSPMY	5.08	acidic	-0.467	9
B*40:01	AEINILEI	3.80	acidic	1.075	8
	EEKLVGFSL	4.53	acidic	0.278	9
	LEEKLVGFSL	4.53	acidic	0.63	10
	RELTHLREKL	8.75	basic	-1.24	10
	SEKAAMNIL	5.52	acidic	0.233	9
B*44:02	AEINILEI	3.80	acidic	1.075	8
C*03:02	KAAMNILEI	6.00	neutral	0.822	9
C*04:01	AYLHSPMYF	6.78	neutral	0.233	9
	RWYSTPSSYL	8.59	basic	-0.89	10
	WYSTPSSYL	5.52	acidic	-0.489	9
	YLHSPMYFF	6.74	neutral	0.344	9
C*06:02	WYSTPSSYL	5.52	acidic	-0.489	9
C*07:01	WYSTPSSYL	5.52	acidic	-0.489	9

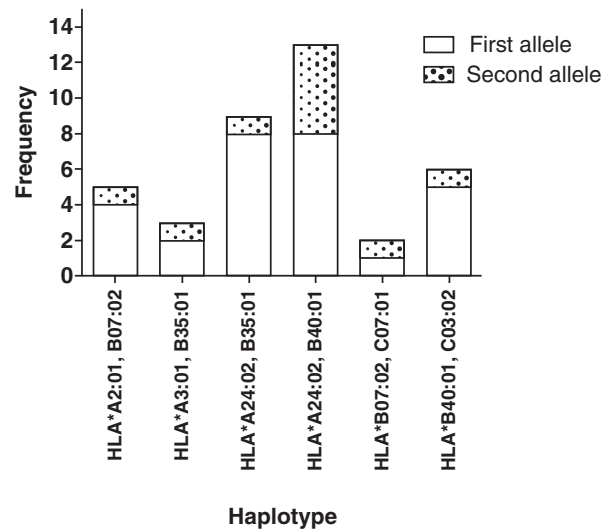


Figure 3. Number of strong binding peptides for different HLA I haplotypes from Costa Rican population. Peptides are derived from the selected SNVs in the primary and metastatic tumor samples of the study.

Despite this, only 28 non-silent SNPs in exonic regions were identified. They were selected to predict neoantigens. These variants were distributed in 26 genes, most of them related to cell cycle control. Interestingly, multiple of these genes, including NOTCH, OBSCN, HSPG2, and NBPF, have been demonstrated to be expressed in colorectal cancer [23-25]. Further analyses are required to assess the role of these specific genes in the development of the disease and the immune response, including studies of driving genes (important for the tumor survival), transcriptomic/proteomic profiling, epigenetic mechanisms or other genomic variants.

Regarding the SNVs identification, we used VarScan2 to call variants (mutations). This software has been shown to have a high sensitivity (81-100%) in the variant calling analysis, in particular with cancer data [26, 27]. In addition, the use of a normal tissue sample to filter variants has been advised by several studies to reduce false positives (such as germline mutations) [28, 29], as we did here. Confirmation strategies are required to validate identified variants (for example using Sanger technology to re-sequence genomic regions). Also, other approaches to call variants can be used to assess other DNA data, to deal with complexity, other sequencing technologies, and biological variability [30].

On the other hand, for HLA of the Costa Rican Central Valley population, information only had a two-digits resolution. Thus, we used the HLA supertypes (A*01, A*02, A*03, A*24 and B*07) or the representative allele of the group to predict interactions (four-digits resolution as required). The length of tumor-specific peptides was considered between 8-11 amino acids, which is the known peptide size that can be presented by HLA-I. It has been suggested that simultaneous testing of peptides between 8 and 11 amino acid residues is advisable in peptide prediction for putative cancer vaccines, due to the “core” or the common sequence (8 amino acids) may have greater versatility in different fragments [31]. This allowed to test several possible peptides of the same variant and to improve the searching for possible neoantigens.

Considering the above, 23 candidate peptides were found to have affinity against the most prevalent Costa Rican Central Valley HLA I alleles (Table 1). Some peptides were identified to be presented by different alleles of the HLA, such as AYLHSPMYF (two different loci HLA A*2402 and HLA * C0401), however, most peptides are predicted to be presented in only one allele.

Regarding the number of peptides and haplotypes, the HLA-A locus was the most significant, contributing to most of the high affinity peptides for the antigenic presentation (Fig. 3 and Table 1). For example, up to nine peptides can be presented in the case of haplotype A*2402/B*3501, in which the first allele is responsible for the presentation of nine peptides.

We also analyzed the physicochemical nature of the peptides (Table 1). Many of them are acidic peptides and, based on the hydrophobicity index, most of them have a slight hydrophobicity. Some “core” peptides are found in several peptides, such as AYLHSPMY which is present as part of seven candidate peptides of different HLA alleles (HLA-A*24:02, B*35:01 and C*04:01).

In the context of the selection of peptides, the algorithms use metrics such as IC50 value (concentration necessary to displace 50% of the classical peptide bound to an HLA allele) [32, 33]. However, it has been shown that it can yield a greater number of false positives in the detection of neopeptides, since there are some alleles that can bind a significant percentage of random peptides [20]. This motivated the present investigation to use only peptides with strong affinities (rank < 0.5%).

Other biologic factors should be considered to identify a possible neoantigens and not included in our study. For example, protein degradation is possible at the endoplasmic reticulum [34]. In addition, peptides affinity must be validated using *in vitro* assays with antigen-presenting cells of the Costa Rican population, for example the Enzyme-linked Immunospot Assay (ELISPOT) [35].

Conclusions

In summary, our approach was able to find 23 candidate neoantigens in all samples of colorectal cancer to eventually create a vaccine. The contribution of this study is the development of the first bioinformatic protocol to detect tumor-specific antigens in the context of the HLA alleles of the Costa Rican population. We hope that this study serve as a basis for future studies with therapeutic uses for cancer vaccines against colorectal tumors or other cancer types.

References

- [1] World Health Organization, Colorectal Cancer Statistics: 2020. Available from: https://gco.iarc.fr/today/data/factsheets/cancers/10_8_9-Colorectum-fact-sheet.pdf
- [2] Palucka AK, Coussens LM. The Basis of Oncoimmunology. *Cell* [Internet]. 2016;164(6):1233–47. Available from: <http://dx.doi.org/10.1016/j.cell.2016.01.049>
- [3] Hackl H, Charoentong P, Finotello F, Trajanoski Z. Computational genomics tools for dissecting tumour-immune cell interactions. *Nat Rev Genet*. 2016;17:441–58.
- [4] Spencer DH, Zhang B, Pfeifer J. Single Nucleotide Variant Detection Using Next Generation Sequencing. In: Kulkarni, S.; Pfeifer J, editor. *Clinical Genomics* [Internet]. Elsevier Inc.; 2014. p. 109–27. Available from: <http://dx.doi.org/10.1016/B978-0-12-404748-8.00008-3>
- [5] Ding L, Wendl MC, McMichael JF, Raphael BJ. Expanding the computational toolbox for mining cancer genomes. *Nat Rev Genet* [Internet]. 2014;15(8):556–70. Available from: <http://dx.doi.org/10.1038/nrg3767>
- [6] Liu XS, Mardis ER. Applications of Immunogenomics to Cancer. *Cell*. 2017;168(4):600–12.
- [7] Su Z, Ning B, Fang H, Hong H, Perkins R, Shi L. Next-generation sequencing and its application in molecular diagnostics. *Expert Rev Mol Diagn*. 2011;11(3):1–16.
- [8] Heather JM, Chain B. The sequence of sequencers: The history of sequencing DNA. *Genomics* [Internet]. 2016;107(1):1–8. Available from: <http://dx.doi.org/10.1016/j.ygeno.2015.11.003>
- [9] Rizzo JM, Buck MJ. Key principles and clinical applications of “next-generation” DNA sequencing. *Cancer Prev Res (Phila)* [Internet]. 2012 Jul 1 [cited 2018 Apr 25];5(7):887–900. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22617168>
- [10] Blankenberg D, Gordon A, Von Kuster G, Coraor N, Taylor J, Nekrutenko A, et al. Manipulation of FASTQ data with galaxy. *Bioinformatics*. 2010;26(14):1783–5.

- [11] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(15):1754–60.
- [12] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–9.
- [13] Koboldt DC, Zhang Q, Larson DE, Shen D, Mclellan MD, Lin L, et al. VarScan 2 : Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res*. 2012;22:568–76.
- [14] Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin)*. 2012;6(2):80–92.
- [15] Cingolani P, Patel VM, Coon M, Nguyen T, Land SJ, Ruden DM. Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program , SnpSift. *Front Genet*. 2012;3(315):1–9.
- [16] Wang K, Li M, Hakonarson H. ANNOVAR : functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):1–7.
- [17] Chisanga D, Keerthikumar S, Pathan M, Ariyaratne D, Kalra H, Boukouris S, et al. Colorectal cancer atlas : An integrative resource for genomic and proteomic annotations from colorectal cancer cell lines and tissues. *Nucleic Acids Res*. 2016;44(D1):969–74.
- [18] Mette A, Morten B, Hadrup SR. MuPeXI : prediction of neo - epitopes from tumor sequencing data. *Cancer Immunol Immunother*. 2017;66(9):1123–30.
- [19] Arrieta-Bolaños E, Maldonado-Torres H, Dimitriu O, Hoddinott MA, Fowles F, Shah A, et al. HLA-A, -B, -C, -DQB1, and -DRB1,3,4,5 allele and haplotype frequencies in the Costa Rica Central Valley Population and its relationship to worldwide populations. *Hum Immunol*. 2011;72(1):80–6.
- [20] Nielsen M, Andreatta M. NetMHCpan-3.0; improved prediction of binding to MHC class I molecules integrating information from multiple receptor and peptide length datasets. *Genome Med [Internet]*. 2016;8(1):1–9. Available from: <http://dx.doi.org/10.1186/s13073-016-0288-x>
- [21] Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein Identification and Analysis Tools on the ExPASy Server. In: *Wlaker JM, editor. The Proteomics Protocols Handbook*. Human Press; 2005. p. 571–608.
- [22] Emens LA, Ascierto PA, Darcy PK, Demaria S, Eggermont AMM, Redmond WL, et al. Cancer immunotherapy: Opportunities and challenges in the rapidly evolving clinical landscape [Internet]. Vol. 81, *European Journal of Cancer*. 2017 [cited 2018 May 4]. p. 116–29. Available from: https://ac.els-cdn.com/S0959804917309188/1-s2.0-S0959804917309188-main.pdf?_tid=852cff99-60b9-4e00-9a8c-eca592b4c0be&acdnat=1525459824_17fe72594087a4b1c451941055acd504
- [23] Andries V, Vandepoele K, Staes K, Bex G, Bogaert P, Isterdael G, et al. NBPF1, a tumor suppressor candidate in neuroblastoma, exerts growth inhibitory effects by inducing a G1 cell cycle arrest. *BMC Cancer*. 2015;15(1):1–25.
- [24] Shriver M, Stroka KM, Vitolo MI, Martin S, DL Huso DL, Konstantopoulos K, et al. Loss of giant obscurins from breast epithelium promotes epithelial-to-mesenchymal transition, tumorigenicity and metastasis. *Oncogene*. 2015;34(32):4248–59.
- [25] Liao W, Li G, You Y, Wan H, Wu Q, Wang C, et al. Antitumor activity of Notch-1 inhibition in human colorectal carcinoma cells. *Oncol Rep*. 2018;39(3):1063–71.
- [26] Spencer DH, Tyagi M, Vallania F, Bredemeyer AJ, Pfeifer JD, Mitra RD, et al. Performance of common analysis methods for detecting low-frequency single nucleotide variants in targeted next-generation sequence data. *J Mol Diagnostics [Internet]*. 2014;16(1):75–88. Available from: <http://dx.doi.org/10.1016/j.jmoldx.2013.09.003>
- [27] Xu C. A review of somatic single nucleotide variant calling algorithms for next-generation sequencing data. *Comput Struct Biotechnol J [Internet]*. 2018;16:15–24. Available from: <https://doi.org/10.1016/j.csbj.2018.01.003>
- [28] Bae JM, Kim JH, Kang GH. Molecular subtypes of colorectal cancer and their clinicopathologic features, with an emphasis on the serrated neoplasia pathway. *Arch Pathol Lab Med*. 2016;140(5):406–12.
- [29] Wang Y, Liping GUO, Feng L, Zhang W, Xiao T, Xuebing DI, et al. Single nucleotide variant profiles of viable single circulating tumour cells reveal CTC behaviours in breast cancer. *Oncol Rep*. 2018;39(5):2147–59.
- [30] Liu ZK, Shang YK, Chen ZN, Bian H. A three-caller pipeline for variant analysis of cancer whole-exome sequencing data. *Mol Med Rep*. 2017;15(5):2489–94.
- [31] Luo H, Ye H, Ng HW, Shi L, Tong W, Mattes W, et al. Understanding and predicting binding between human leukocyte antigens (HLAs) and peptides by network analysis. *BMC Bioinformatics [Internet]*. 2015 [cited 2018 May 16];16(13). Available from: <http://www.biomedcentral.com/1471-2105/16/S13/S9>

- [32] Giannakis M, Mu XJ, Shukla SA, Qian ZR, Cohen O, Nishihara R, et al. Genomic Correlates of Immune-Cell Infiltrates in Colorectal Carcinoma. *Cell Rep.* 2016;15(4):857–65.
- [33] Karasaki T, Nagayama K, Kuwano H, Nitadori J ichi, Sato M, Anraku M, et al. An Immunogram for the Cancer-Immunity Cycle: Towards Personalized Immunotherapy of Lung Cancer. *J Thorac Oncol* [Internet]. 2017;12(5):791–803. Available from: <http://dx.doi.org/10.1016/j.jtho.2017.01.005>
- [34] Fleri W, Paul S, Dhanda SK, Mahajan S, Xu X, Peters B, et al. The immune epitope database and analysis resource in epitope discovery and synthetic vaccine design. *Front Immunol.* 2017;8(278):1–16.
- [35] Kato T, Matsuda T, Ikeda Y, Park J-H, Leisegang M, Yoshimura S, et al. Effective screening of T cells recognizing neoantigens and construction of T-cell receptor-engineered T cells. *Oncotarget.* 2018;9(13):11009–19.