Computers in Biology and Medicine 167 (2023) 107594

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Risk assessment of cancer patients based on HLA-I alleles, neobinders and expression of cytokines

Anjali Dhall, Sumeet Patiyal, Harpreet Kaur, Gajendra P.S. Raghava [*,1]

Department of Computational Biology, Indraprastha Institute of Information Technology, Okhla Phase 3, New Delhi, 110020, India

ARTICLE INFO

Keywords:
HLA-Alleles
Neobinders
Cancer
Prognostic biomarkers
Cytokines

ABSTRACT

Advancements in cancer immunotherapy have shown significant outcomes in treating cancers. To design effective immunotherapy, it's important to understand immune response of a patient based on its genomic profile. However, analyses to do that requires proficiency in the bioinformatic methods. Swiftly growing sequencing technologies and statistical methods create a blockage for the scientists who want to find the biomarkers for different cancers but don't have detailed knowledge of coding or tool. Here, we are providing a web-based resource that gives scientists with no bioinformatics expertise, the ability to obtain the prognostic biomarkers for different cancer types at different levels. We computed prognostic biomarkers from 8346 cancer patients for twenty cancer types. These biomarkers were computed based on i) presence of 352 Human leukocyte antigen class-I, ii) 660959 tumor-specific HLA1 neobinders, and iii) expression profile of 153 cytokines. It was observed that survival risk of cancer patients depends on presence of certain type of HLA-I alleles; for example, liver hepatocellular carcinoma patients with HLA-A*03:01 are at lower risk. Our analysis indicates that neobinders of HLA-I alleles have high correlation with overall survival of certain type of cancer patients. For example, HLA-B*07:02 binders have 0.49 correlation with survival of lung squamous cell carcinoma and −0.77 with kidney chromophobe patients. Additionally, we computed prognostic biomarkers based on cytokine expressions. Higher expression of few cytokines is survival favorable like IL-2 for bladder urothelial carcinoma, whereas IL-5R is survival unfavorable for kidney chromophobe patients. Freely accessible to public, CancerHLA-I maintains raw and analysed data (https://webs.iiitd.edu.in/raghava/cancerhla1/).

1. Introduction

Cancer is one of the top causes of mortality worldwide; GLOBOCAN reported that 19.3 million new cancer cases and 10 million deaths were occurred in 2020 [1]. Several researchers have worked tirelessly over the last few decades to develop novel cures and treatments to fight against the deadly disease [2]. Traditional therapies like chemotherapy, radiation, and surgery are the most commonly used treatments [3]. These radiation-based therapies have negative consequences on the patient's health and survival [2,4,5]. To circumvent the drawbacks of conventional medicines, new treatment regimens have been developed, including targeted cancer therapies, adoptive T cell therapy, immune checkpoint inhibitor-based therapies, immunomodulators, interferons and oncolytic viruses [6–10]. Cancer immunotherapy have resulted in considerable outcomes and increase the life duration of many patients suffering from various solid tumours [11,12].

The central pillars of immunotherapy are immune checkpoint inhibitors and chimeric antigen receptor (CAR) T cells. These therapies are entirely dependent on T-lymphocytes (T cells), which recognize tumor-associated peptides displayed on the tumor cell surface by human leukocyte antigens (HLA) [13]. HLAs are the highly complex and polymorphic genes in the human genome, situated on chromosome 6. Class-I HLA alleles interact with the $CD8^+$ T cell receptors to activate T cells which further induce several immune responses to knock off the tumor cells from our system [14–17]. The immunotherapies are entirely based on the T cells, which identify tumor-associated peptides presented by human leukocytes antigens (HLA) on the infected cell surface. Recently, scientists majorly focused on HLA-dependent therapies, including $CD8^+$ T cell therapy, tumor-infiltrating lymphocytes (TILs) therapy, and TCR-engineered T cells (TCR-Ts) and neoantigen-based therapy [18,19] to treat cancer patients. HLA-dependent treatments are more effective

* Corresponding author. Department of Computational Biology, Indraprastha Institute of Information Technology, Delhi Okhla Industrial Estate, Phase III, Office: A-302 (R&D Block), 110020, New Delhi, India.
E-mail address: raghava@iiitd.ac.in (G.P.S. Raghava).
[1] http://webs.iiitd.edu.in/raghava/.

https://doi.org/10.1016/j.compbiomed.2023.107594
Received 21 December 2022; Received in revised form 6 October 2023; Accepted 17 October 2023
Available online 18 October 2023

List of abbreviations

BLCA	bladder urothelial carcinoma
BRCA	beast invasive carcinoma
CESC	cervical squamous cell carcinoma and endocervical adenocarcinoma
CHOL	cholangiocarcinoma
GBM	glioblastoma multiforme
HNSC	head and neck squamous cell carcinoma
KICH	kidney chromophobe
KIRC	kidney renal clear cell carcinoma
KIRP	kidney renal papillary cell carcinoma
LIHC	liver hepatocellular carcinoma
LUAD	lung adenocarcinoma
LUSC	lung squamous cell carcinoma
OV	ovarian serous cystadenocarcinoma
PAAD	pancreatic adenocarcinoma
PRAD	prostate adenocarcinoma
READ	rectum adenocarcinoma
SKCM	skin cutaneous melanoma
STAD	stomach adenocarcinoma
THCA	thyroid carcinoma
UCEC	uterine corpus endometrial carcinoma

and efficient than traditional chemotherapies. HLA-peptide binding is critical for determining the cancer immunogenicity.

In the past, number of repositories such as TCIA [20], TSNAdb [21], NEPdb [22], dbpepNeo [23], Ovirusdb [24], CancerTope [25] have been developed for designing immunotherapy. However, these databases lack the information of impact of class-I HLA-alleles, neobinders and cytokines on the survival of cancer patients. To construct patient-specific therapy, genetic information such as HLA-alleles, neoantigens, HLA-peptide binding affinity, and immune response need to be considered. Fortunately, it is now possible to detect patient-specific HLA-alleles with the help of latest technologies and the availability of sequencing data. In the pilot study, we have developed a resource which provides patient-specific information from public repositories such as (TCGA and TCIA) and analysed patient survival based on HLA-alleles, as well as the correlation of the amount of neobinders specific to HLA-alleles with the overall survival in various cancer types. Furthermore, we used correlational analysis to better understand the role of chemokines, cytokines, and their receptors in the prognosis of the cancer patients. We combine all aforementioned analysis of 20 different types of cancers onto a single user-friendly, and freely accessible platform named as "CancerHLA-I" (https://webs.iiitd.edu.in/raghava/cancerhla1/). CancerHLA-I is currently the most comprehensive web-based platform to explore the roles of HLA-I alleles, neobinders and cytokines expression on the survival of cancer patients for twenty different types of cancers.

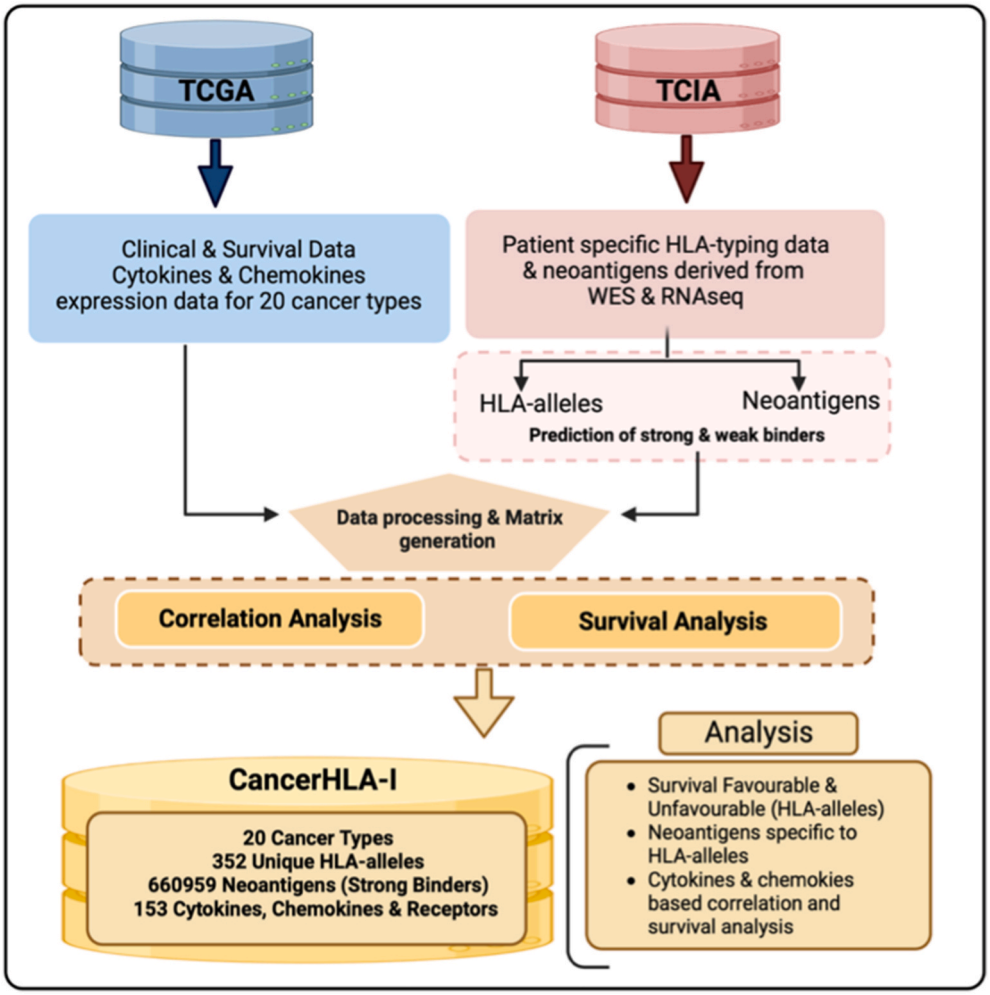

Fig. 1. Complete architecture of the study including dataset collection, analysis and website development.

2. Material & methods

2.1. Overall study design

The complete architecture of the study is depicted in Fig. 1.

2.2. Collection and pre-processing of datasets

In this study, we have obtained the cancer patients datasets from two major repositories, The Cancer Genome Atlas (TCGA) [20] and The Cancer Immunome Atlas (TCIA) [26]. We have considered those cancer types and patients, which are available in both databases. Therefore, we have a total of 8246 cancer patients with their HLA genotyping, gene expression, and clinical information. We gathered the genomic and clinical information for 8346 patients with 20 different types of cancer (BLCA, BRCA, CESC, CRC, GBM, HNSC, KICH, KIRC, KIRP, LIHC, LUAD, LUSC, OV, PAAD, PRAD, READ, SKCM, STAD, THCA, and UCEC). From TCIA, we obtained the control access [with the approval of dbGap (Project No. 17674)] patient-specific HLA-typing data and neoantigens data for 20 cancer types. Furthermore, RNA-seq expression profiles for 20 types of cancer were obtained using TCGA assembler 2.0. After that, we filtered out cytokines, chemokines, and their receptors gene expression for 20 types of cancer. The expression profiles were then transformed into log2 values after addition of 1.0 as a constant number to each of the expression value. The survival information covered vital status and overall survival time (OS).

2.3. Mean-overall and univariate survival analysis

For each cancer type, we first built a binary matrix based on the presence or absence of HLA-alleles. Each column represents HLA-alleles, and each row represents samples/patients. We used individual's survival data to calculate mean overall survival (MOS) based on the presence or absence of an HLA-allele. After that, we computed the difference in MOS (depending on presence/absence). Cox proportional hazard (Cox-PH) regression models were used in the current study to identify HLA-alleles linked with cancer patient survival. The "survival" package in R (V.3.5.1) was used for univariate analysis. The cox regression coefficient greater than 0 indicates that the presence of an HLA-allele affects survival (unfavorable), whereas less than 0 indicates that the presence of alleles increases survival (favorable). For each allele, we calculated the Hazard Ratio (HR) and 95 % CI (Confidence Interval). HR $>$ 1 denotes high-risk HLA-alleles, while HR $<$ 1 depicts low-risk alleles; however, HR $=$ 1 has no effect on survival. Furthermore, the log-rank test was conducted, and p-value was calculated to determine the significant distribution of low-risk and high-risk patients. To calculate the predictive performance of the models, we used the Concordance index (C). In this study, univariate survival analysis performed based on HLA-alleles, cytokines, and chemokines genes for 20 cancer type.

2.4. Correlation analysis

We extracted the strong binding neoantigens/epitopes corresponding to each HLA-allele for each cancer type using the MHCflurry 2.0 software [27]. We classify neoepitopes as strong or weak binders using the MHCflurry software's binding affinity (BA) percentile, where neoantigens with BA$<$2 were considered as strong binders, and BA$>$2 were taken as weak binders. Following that, we built a count matrix with the number of strong binders matching to each HLA-allele and cancer type. Further, we calculated the correlation coefficient between overall survival and the number of strong binders for each HLA-allele in order to understand the impact of number of binders on the survival using Pearson correlation test. The correlation coefficient and p-value ($<$0.05) demonstrated the significance of the number of HLA-binding neoepitopes on the survival of the cancer patient. Moreover, we conducted correlation analysis using the expression values of cytokines, chemokines genes for each cancer type. We used survival data and the expression of 153 cytokines, chemokines, and their receptors.

2.5. Database implementation

The web interface of CancerHLA-I (https://webs.iiitd.edu.in/raghava/cancerhla1/) was developed using MySQL and hosted on Linux based Apache server. In order to create the interactive user interface, we developed responsive framework made up of HTML, CSS, and JavaScript. To improve the data view, the user interface is responsive, which means that the web-interface recognizes the user device and modifies its structure and shape in accordance with the device resolution. With the help of this functionality, the interface is adaptable to a wide range of devices and browsers with various screen resolutions. The web-site can be searched using variety of devices (smartphones or tablets) and browsers (Google Chrome, Mozilla Firefox, and Safari).

3. Results

3.1. Statistical analysis of data of CancerHLA-I

CancerHLA-I incorporates data on cancer associated HLA-alleles, neoantigens, cytokines, chemokine, and their relationship with the overall survival for the 20 cancer types. The genomic and clinical information of 8346 cancer patients were downloaded from TCGA and TCIA repositories and processed to build this resource. In Fig. 2, we have provided the description of 20 types of cancer, with number of samples, HLA-alleles, total number neoantigens, strong and weak binders. As shown in Fig. 2B, HLA-B acquire maximum number of alleles i.e., 185 followed by HLA-A (97 alleles) and HLA-C (70 alleles). We observed that in the case of Uterine Corpus Endometrial Carcinoma (UCEC), highest number (i.e., 192 HLA-alleles) are reported, whereas Kidney Chromophobe (KICH) reported the lowest number i.e., 86 HLA-alleles (See Fig. 2C). In addition, we reported top-20 most frequent HLA-alleles whose frequency distribution is maximum among the cancer patients (See Fig. 2D). Moreover, we have obtained more than 90,000 strong binders i.e., neobinders corresponding to Skin Cutaneous Melanoma (SKCM) and Uterine Corpus Endometrial Carcinoma (UCEC) cancers. However, we have less than 5000 neobinders for Kidney chromophobe (KICH) and Thyroid Carcinoma (THCA) cancer types (See Fig. 2E). The complete data used in this analysis is stored in big MySQL table that can be searched and browsed by various categories defined in the database such as cancer type, HLA-allele, neoantigen, HR, p-value, cytokine/chemokine, correlation coefficient etc.

Of note, we have generated the UpSet plot to understand the distribution of HLA-alleles in different cancer types. These overlapping and exclusive HLA-alleles among each cancer type (See Fig. 3) are the ones that can serve the basis for explaining the molecular heterogeneity and similarity among the different cancer types. It also provides the potential insights into the progression of disease.

3.2. HLA-based prognostic biomarkers

At first, we have combined all the cancer types to perform the survival analysis in order to check the overall impact of presence/absence of HLA alleles on the survival of the cancer patients. As shown in the Supplementary Table S1, we have not observed any significant trend by combining all the cancer types due to the tumor heterogeneity. So, we have generated binary matrix corresponding to each cancer type based on the presence and absence of the HLA-alleles. With the utility of the survival package, we have performed survival analysis to identify the high risk and low risk HLA-alleles corresponding to each cancer type. In Table 1, we have reported only those HLA-alleles whose presence influences the survival of more than one type of cancer patients. For instance, presence of HLA-A*02:01 allele reduces the survival of kidney cancer (HR $=$ 5.46 with p-value $=$ 0.03) and skin cancer patients (HR $=$

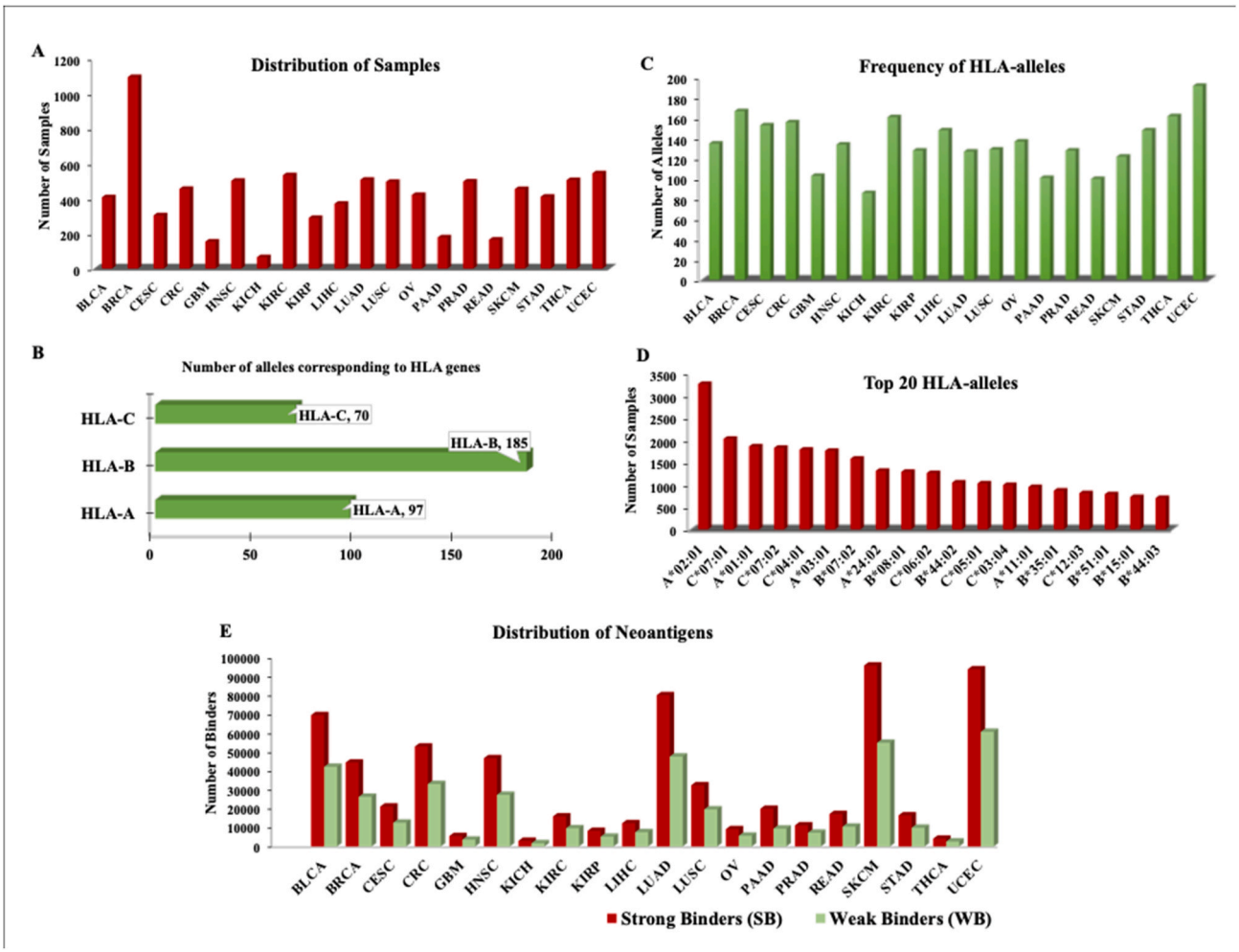

Fig. 2. Complete distribution of dataset in 20 cancer types with number of samples, frequency of alleles, HLA-alleles corresponding to class-I genes, most frequent HLA.

1.36 and p-value = 0.02). We observed that presence of HLA-A*02:01, HLA-A*68:01, HLA-B*52:01 and HLA-C*03:02 is significantly associated with the poor prognosis (with HR > 2 and p-value<0.05) in different types of cancer patients, as shown in Table 1. Moreover, the presence of certain HLA-alleles significantly improves the survival rate of cancer patients, for instance, HLA-C*14:02, HLA-C*12:03, HLA-A*03:01 significantly improve the survival rate and act as good prognostic markers in Bladder urothelial carcinoma (BLCA), Stomach Adenocarcinoma (STAD), Liver Hepatocellular Carcinoma (LIHC) and Glioblastoma Multiforme (GBM) (See Table 1).

3.3. HLA-I neobinders based correlation analysis

In this study, we have used MHCflurry 2.0 (O'Donnell et al., 2020) software for the prediction of neoepitopes having strong binding potential with the class-I HLA alleles. We have identified strong HLA-specific neobinders for each cancer type. In order to understand the impact of number of neobinders on the survival of the cancer patients, we performed Pearson correlation analysis. Fig. 4, shows the correlation values for the nine HLA-alleles (HLA-A*02:01, HLA-A*03:01, HLA-C*07:01, HLA-C*07:02, HLA-A*01:01, HLA-B*07:02, HLA-B*08:01, HLA-A*24:02, and HLA-B*44:02) present in most of the samples and all cancer types. At first, we have computed the correlation between the neobinders irrespective of the cancer types by combining all the data files. We observed that, the overall impact on combining the neobinders for all the cancer types is very less or negligible (See Fig. 4). On the other hand, some of the neobinders corresponding to the particular HLA-alleles have very high positive as well as negative correlation with the specific cancer types. For instance, HLA-B*07:02 neobinders have very high negative correlation (r = −0.77) with the survival of KICH patients, while on the opposite side it shows positive correlation coefficient of 0.49 with the LUSC patients. In the case of BLCA, LUSC, and OV, most of the nine alleles shows positive correlation with the overall survival. However, in case of KICH, some of the alleles showed positive association and some of them exhibited negative association with the survival. This signifies that, the binders corresponding to HLA-alleles have favorable and unfavorable impact based on the different cancer types.

3.4. Cytokines based prognostic biomarkers

In order to understand the prognostic role of cytokines and chemokines, we have performed univariate survival analysis using their expression profiles. In Fig. 5, we have reported those cytokine and chemokines, whose expression significantly impact the survival rate of the cancer patients. We observed that, the high expression of IL2, IFNB1, IFNA8, IL5 cytokines were having good impact on the survival of different cancer patients (HR < 0.4 and p-value <0.05). Whereas IL5RA, TGFBR3, CCR4, TGFB2, IL17A were highly associated with the poor survival rate in KICH, READ, and GBM patients (HR > 4 and p-value<0.05). The complete analysis for all the other cytokines and chemokines is available on CancerHLA-I server.

Moreover, we have done the correlation analysis by considering the

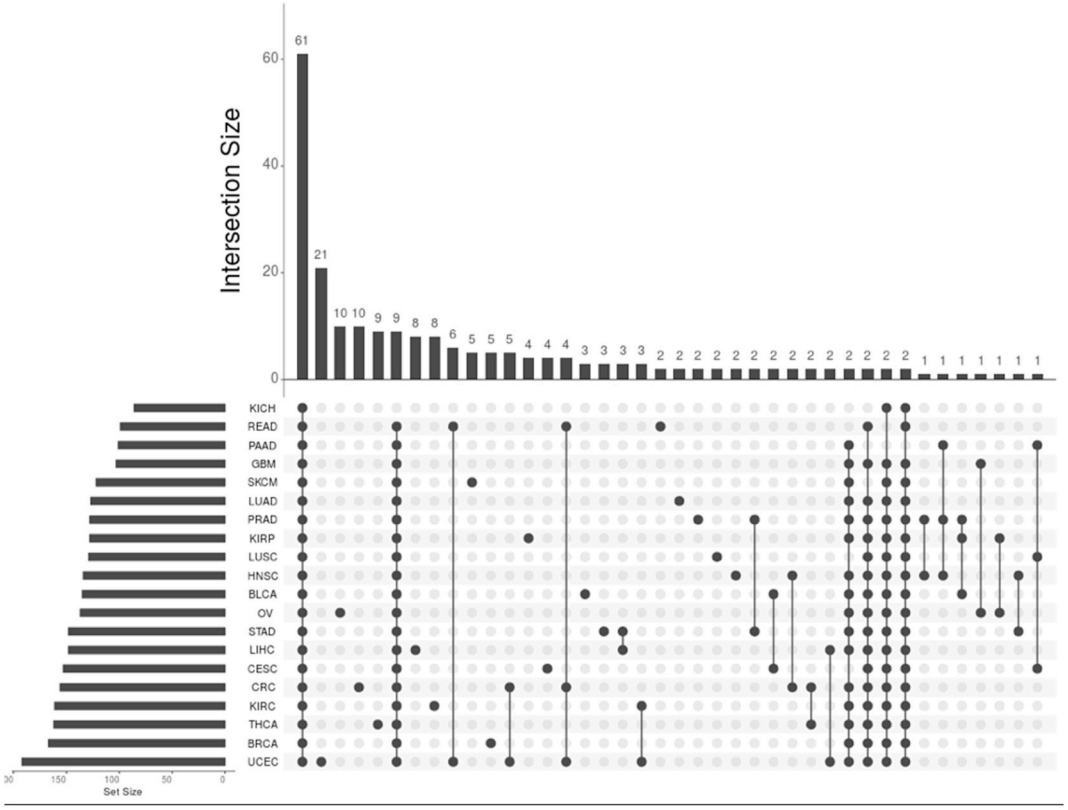

Fig. 3. UpSet plot of top 40 interactions including both common and unique HLA-alleles among each cancer type.

Table 1
Results of univariate survival analysis based on the presence/absence of HLA-alleles in different type of cancers.

Cancer	HLA-allele	Allele Present (No. of Patients)	Allele Absent (No. of Patients)	Hazard Ratio (95%CI)	p-value	Concordance Index
KICH	HLA-A*02:01	26	39	5.462 (1.135–26.299)	0.034	0.72
SKCM		203	251	1.36 (1.038–1.781)	0.025	0.528
CRC	HLA-A*68:01	24	431	2.466 (1.237–4.918)	0.01	0.532
OV		21	399	1.936 (1.101–3.402)	0.022	0.51
GBM	HLA-B*44:03	16	138	1.805 (1.025–3.181)	0.041	0.528
LIHC		46	324	1.665 (1.041–2.663)	0.033	0.529
OV		31	389	1.583 (1.019–2.459)	0.041	0.512
THCA	HLA-B*52:01	25	480	4.057 (1.155–14.254)	0.029	0.62
SKCM		16	438	2.183 (1.154–4.132)	0.016	0.515
HNSC		18	483	2.113 (1.149–3.886)	0.016	0.514
LIHC	HLA-C*03:02	16	354	2.268 (1.104–4.658)	0.026	0.517
BRCA		96	997	1.659 (1.025–2.685)	0.039	0.52
LIHC		43	327	1.651 (1.034–2.636)	0.036	0.533
LUAD	HLA-C*07:01	140	367	1.366 (0.998–1.869)	0.051	0.538
LUSC		139	356	1.365 (1.018–1.829)	0.037	0.526
LUAD	HLA-C*12:03	51	456	1.668 (1.075–2.587)	0.022	0.518
BRCA		113	980	1.655 (1.042–2.627)	0.033	0.543
GBM		22	132	0.527 (0.29–0.936)	0.029	0.532
KIRP	HLA-A*03:01	74	215	1.884 (1.007–3.526)	0.044	0.531
LIHC		72	298	0.635 (0.397–1.015)	0.046	0.526
BLCA	HLA-C*14:02	15	392	0.140 (0.02–1.001)	0.048	0.517
STAD		20	390	0.315 (0.1–0.988)	0.048	0.516

gene expression of cytokine, chemokines and their receptors. The heatmap shows (Fig. 6) the correlation of overall survival with the expression of some of the cytokines and chemokines in 20 cancer types. The darker blue color shows the positive correlation, whereas light yellow color depicts the negative correlation. We observed that, cytokine IFNG have very high and significant positive correlation with the survival rate of GBM patients; higher expression of IL9 cytokine is associated with positive correlation in BLCA and OV cancer patients. Whereas, cytokine IL2, TNFA1P1, and TNF are associated with the negative correlations with the survival of KICH patients. In case of chemokines, we observed that CCL1 (CRC, KIRC and READ), CCL20 (BLCA, READ, and PAAD), CCL27 (GBM, KICH and THCA) had positive correlation with the overall survival. However, the over-expression of CCL18, CCL28, CCL4, CCL5 is associated with the negative correlation with most of the cancer types (See Fig. 6).

3.5. Utility of CancerHLA-I

CancerHLA-I can be interactively browsed and searched in a variety of different ways to satisfy the query of the user. The homepage of

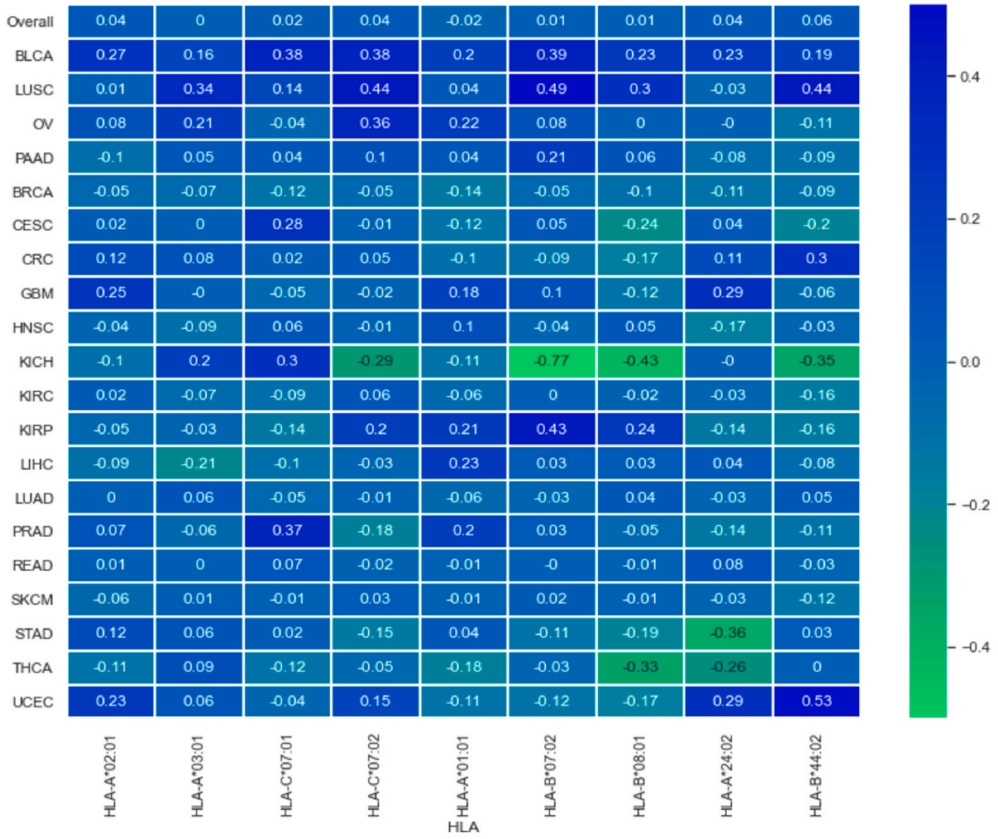

Fig. 4. Correlation analysis based on the number of HLA-I neobinders with the overall survival of cancer patients.

CancerHLA-I website provides a simple search page, where users can search the query in the database for specific cancer types, HLA-alleles, neoantigens, cytokines/chemokines, and its survival association (See Fig. 7). The advanced search page in the database provides the customized search facility for user defined query using Boolean expressions (AND/OR). We compiled the data in a tabular form corresponding to each cancer type for easy and efficient access. The advantage of the browsing facility is that users can quickly obtain all the results by clicking onto specific entry under concerned category. Moreover, 'Help' page on the website provides detailed visualization of the usage of the CancerHLA-I database. The data can be downloaded as a tab-delimited, comma separated, JSON, PDF and XLSX file formats.

3.6. Comparison with other resources

In the past, a number of databases have been created to store and provide the information related to various types of cancers such as TCGA, GEO, GDC etc. In order to explore the cancer biology, it is important to process and analyse these genomic datasets using different bioinformatics tools, pipelines and statistical approaches. However, to address the questions related to specific field of the cancer biology such as cancer immunotherapy; devoted resources are required. As shown in Table 2, we have compiled and compare various web-based resources with our database based on different parameters. TCIA [20], TSNAdb [21], NEPdb [22] dbPepNeo [23], TCLP [28], and CancerTope [25] databases provide information regarding the cancer specific peptides, and HLA-alleles. However, these resources lack the association of patient specific HLA-alleles, neoantigens and cytokine expressions with the overall survival of cancer patients in different cancer types. Of note, CancerHLA-I provides the impact of presence/absence of HLA-alleles, neoantigens and cytokines immune signatures on the survival of the 20 different kinds of cancer patients by implementing correlation and survival analysis.

4. Discussion

Class-I HLA molecules are essential for immunosurveillance and cancer immunotherapy [29,30]. Several studies report that, there is a significant impact of HLA-alleles with the survival of cancer patients. For instance, Naranbhai et al., shows the presence of HLA-A*03 allele in kidney cancer patients reduces the survival rate and associated with the poor response against the immune checkpoint blockade (ICB) therapy [31]. HLA-B*55 and HLA-A*01 significantly improves the survival, while HLA-B*50 allele reduces the survival rate in the melanoma patients [32]. Moreover, due to the mutations at the genetic level, loss of heterozygosity occurs in HLA genes at chromosome 6 in the colorectal cancer and non-small cell lung carcinoma patients [30,33]. However, the loss of the functions of class-I HLA molecules exhibits the escape mechanism by different cancer types [30,34,35]. Moreover, mutations in type-I and II interferon pathway genes also effect the survival of cancer patients. Interleukins such as IL-6, IL-11, IL-1, and TGF-β induces cancer cell proliferation and progression [36]. The relationship between cytokines and cancer types is highly complicated and variable. The outcome or effect of certain types of cytokines, such as IL-2, IL-5, TNF, etc., varies with the different types of cancer. There could be a number of reasons associated with the favorable or unfavorable effect of cytokines with the different types of cancers, such as tumor microenvironment [37–39] the stage of cancer [40–42], heterogeneous nature of cancer [43,44], and interaction with the other factors, like other cytokines, immune cells, and tumor antigens [45].

To understand the overall impact of patient-specific HLA-alleles, neobinders and cytokines/chemokines expression profiles, we have conducted a pan-cancer analysis using 20 cancer types. Here, we have performed survival and correlation analysis using the genomic information of cancer patients. Furthermore, we investigated the relationship between cytokine expression with the overall survival of the cancer patients. Our analysis reveals that HLA-C*14:01 and HLA-C*12:03

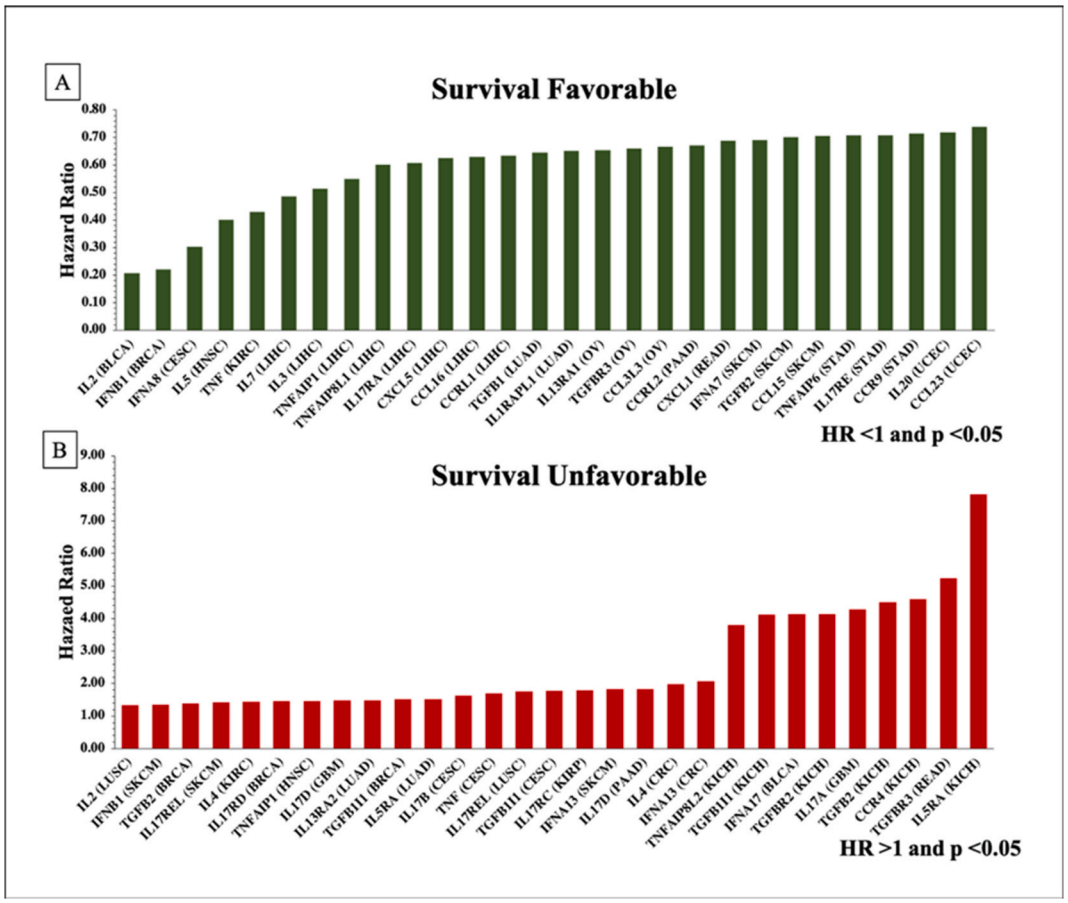

Fig. 5. Univariate survival analysis shows survival favorable (HR < 1) and unfavorable (HR > 1) cytokines, chemokines and their receptors in different cancer.

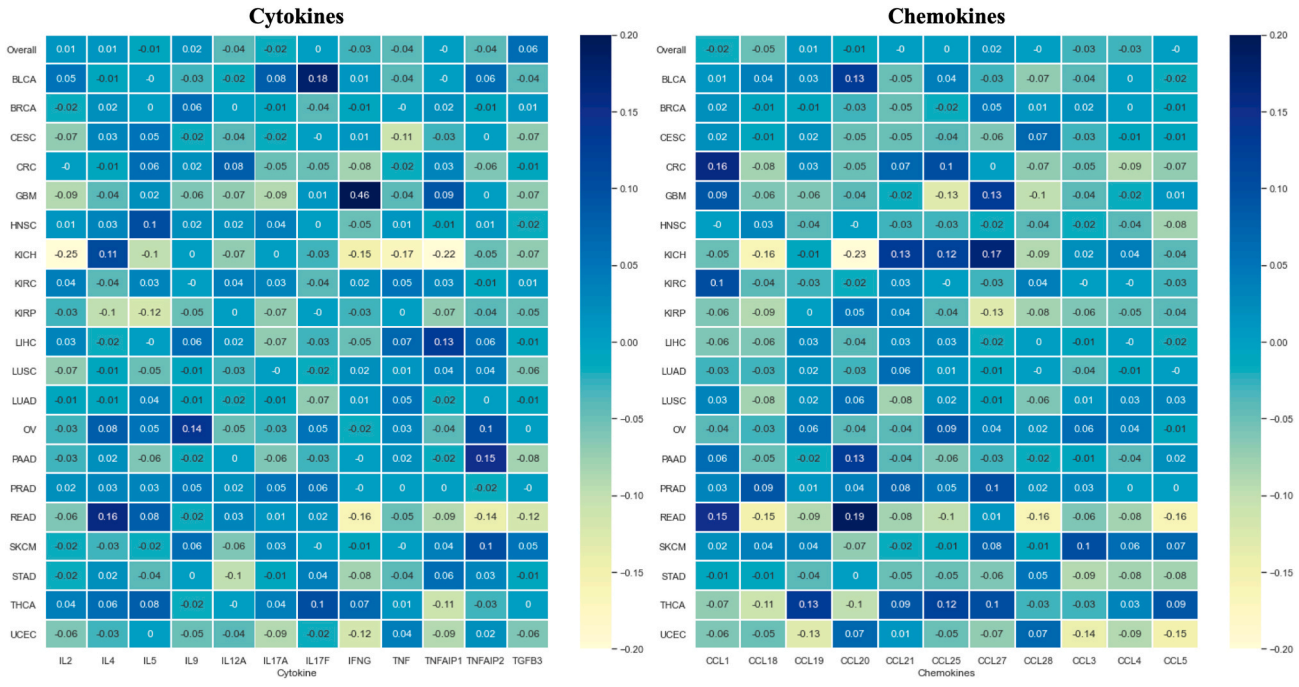

Fig. 6. Correlation analysis based on the expression profile of cytokines & chemokines with the overall survival of cancer patients.

significantly improves the overall survival of cancer patients, whereas HLA-A*02:01, HLA-A*68:01, HLA-B*52:01, HLA-C*03:02 are associated with the poor survival (See Table 1). Some previous studies also support our findings; for example, Andersson et al., reported that HLA-A*02 positive epithelial ovarian cancer patients possess poor overall survival rate [46]. Alifu et al., shows that HLA-A*68:01 associated with

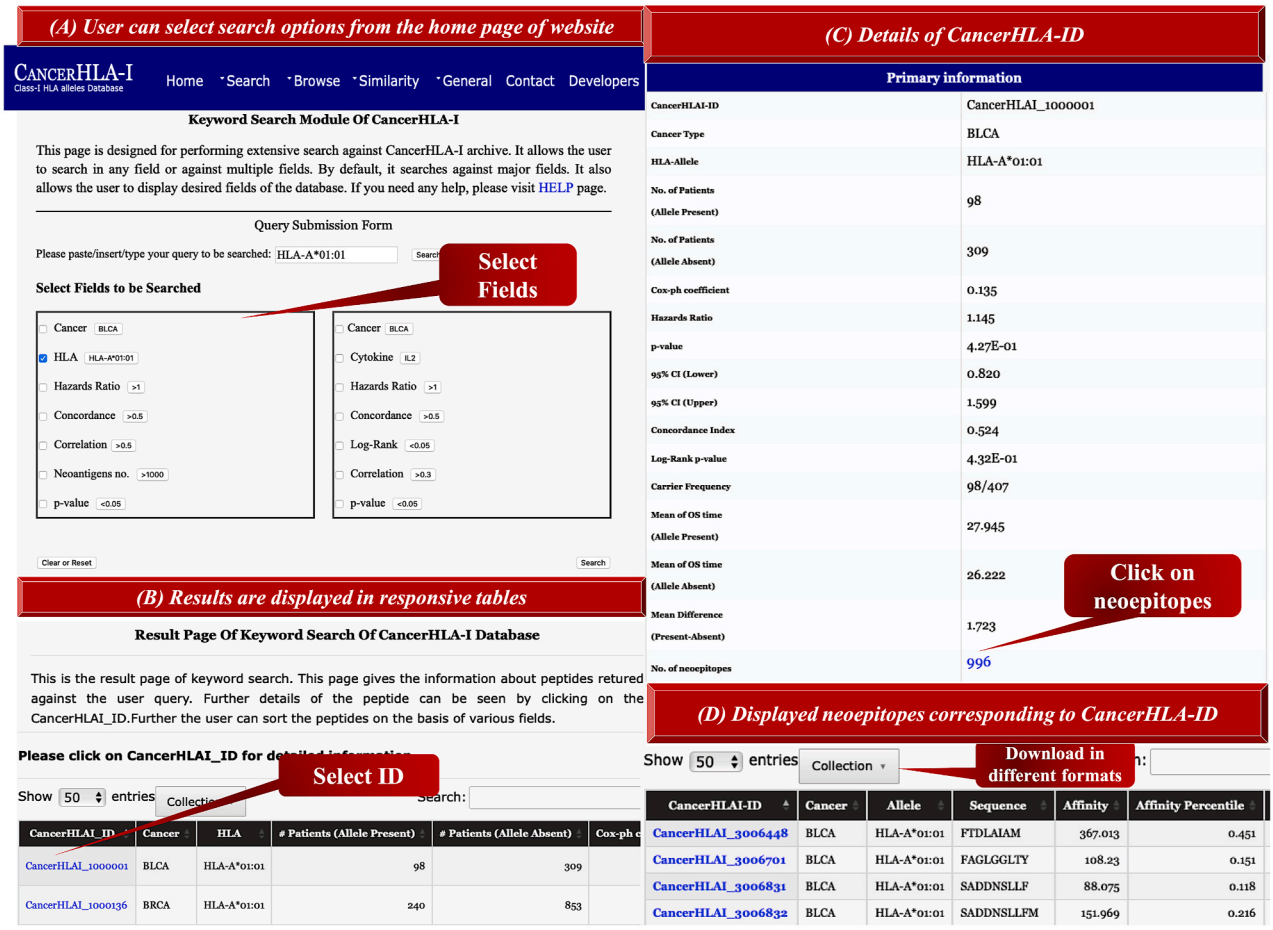

Fig. 7. Schematic illustration of the utility of Cancer HLA-I (A). Simple search page of the database Cancer HLA-I, where users can select fields for querying against the.

Table 2
Shows major features integrated in our repository CancerHLA-I and in existing major repositories.

Major features integrated in repositories	Different Repositories/Databases				
	CancerHLA1	dbPepNeo	TCIA	TSNAdb	NEPdb
Type of cancers covered	20	14	20	16	23
Search of neopeptides	Yes	Yes	No	Yes	Yes
Similarity search with neopeptides	Yes	No	No	No	Yes
Browse on type of cancer	Yes	Yes	Yes	Yes	Yes
HLA allele wise categorization	Yes	Yes	Yes	Yes	Yes
Correlation (overall survival vs neobinders of HLA alleles)	Yes	No	No	No	No
Correlation (overall survival vs expression of cytokines)	Yes	No	No	No	No
Prognostic value of HLA alleles	Yes	No	No	No	No
Prognostic value of cytokines	Yes	No	Yes	No	No

the poor survival rate in advanced squamous cell cervical cancer patients [47]. A study report that presence of HLA-B*52:01 in Japanese women with cervical cancer significantly associated with poor survival rate with HR = 1.6 and $p < 0.05$ [48]. In addition, HLA-B*52 allele is also having strong association with the development of takayasu arteritis (TAK) [49,50]. Moreover, correlation analysis revealed the positive and negative association of neobinders with the survival rate of the cancer patients (See Supplementary Figs. S1–S9).

Torrenté et al., reported that it is not necessary that the assumption of gene expression transcriptomic data follows a specific kind of distribution, such as, normal distribution always holds true [33371881]. Therefore, we have implanted the Fitter library of Python to explore the distributions of cytokine expression. Our results correspond with their observation that less that 50 % of the cytokine expression data follows the normal distribution, whereas rest of them follows other distributions such as Lognormal, Cauchy, and Gamma (See Supplementary Fig. S10). Further, we identified that, higher expression of some of the cytokines, such as, IL2, IFNG, IFNB, TNF significantly improves the survival of cancer patients, while some of the cytokines like IL5, IL17, CCR4, TGF reduces the survival rate significantly. IL2 elicit anti-tumor immune response and acts as a key cytokine in our immune system, FDA approved IL2-based immunotherapy for the treatment of metastatic renal cell carcinoma and metastatic melanoma. Studies also revealed that higher expression of IL2 is associated with improved survival rate [51,52]. Moreover, high expression of interferon-gamma improved the progression-free survival rate of non-small cell lung cancer patients treated with nivolumab; and melanoma patients treated with

pembrolizumab [53]. Wang et al., report that chemokines receptor (CCR4) acts as negative regulator with survival and its high expression levels CCR4 in oral tongue carcinoma patients significantly associated with poorer survival rate [54]. While high serum levels of IL17 cytokine associated with poor prognosis in non-small cell lung carcinoma, HBV-related hepatocellular carcinoma, and gastric cancer [55]. These studies indicate that presence/absence of HLA-alleles and cytokines expression rate are significantly impact the survival rate of different cancer patients. In order to aid the clinicians and scientific community we attempted to develop a highly interactive web-based platform for the analysis and identification of cancer-specific biomarkers which shows the overall impact of HLA and cytokines with the overall survival rate. We have integrated user-friendly browsing, searching, and analysis modules in our resource "CancerHLA-I" (https://webs.iiitd.edu.in/raghava/cancerhla1/) for easy data retrieval, data comparison, and examination.

5. Conclusion

Understanding the prognostic importance of HLA-alleles, binding peptides, and cytokines is crucial to examining the impact and efficacy of cancer immunotherapy. In this study, we have developed a repository "CancerHLA-I" which supports the correlation and survival analysis based on the genomic profiles of 20 types of cancer patients. CancerHLA-I provides a platform to clinicians and researchers to identify the associations between 352 Class-I HLA alleles with the survival of cancer patients. Moreover, it also provides the correlation and statistical significance values of cytokines/chemokines with the survival of different type of cancer patients. We anticipate that our research yields promising novel HLA and cytokines-based biomarkers for improved cancer immunotherapy and treatment.

Authors contributions

AD and GPSR collected and processed the datasets. AD, HK, SP and GPSR conceived the idea and analysed the results. AD and SP created the back-end of the web server the front-end user interface. AD, SP and GPSR penned the manuscript. GPSR conceived and coordinated the project. All authors have read and approved the final manuscript.

Funding source

The current work has received grant from the Department of Bio-Technology (DBT), Govt. of India, India.

Declaration of competing interest

The authors declare no competing financial and non-financial interests.

Acknowledgements

Authors are thankful to the Department of Bio-Technology (DBT) and Department of Science and Technology (DST-INSPIRE) for fellowships and the financial support and Department of Computational Biology, IIITD New Delhi for infrastructure and facilities.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2023.107594.

References

[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA A Cancer J. Clin. 71 (2021) 209–249.
[2] C. Pucci, C. Martinelli, G. Ciofani, Innovative approaches for cancer treatment: current perspectives and new challenges, E Cancer Med.Sci. 13 (2019) 961.
[3] M. Arruebo, N. Vilaboa, B. Saez-Gutierrez, J. Lambea, A. Tres, M. Valladares, A. Gonzalez-Fernandez, Assessment of the evolution of cancer treatment therapies, Cancers 3 (2011) 3279–3330.
[4] V. Dilalla, G. Chaput, T. Williams, K. Sultanem, Radiotherapy side effects: integrating a survivorship clinical lens to better serve patients, Curr. Oncol. 27 (2020) 107–112.
[5] I. Altun, A. Sonkaya, The most common side effects experienced by patients were receiving first cycle of chemotherapy, Iran, J. Public Health 47 (2018) 1218–1219.
[6] V.V. Padma, An overview of targeted cancer therapy, Biomedicine 5 (2015) 19.
[7] K. Esfahani, L. Roudaia, N. Buhlaiga, S.V. Del Rincon, N. Papneja, W.H. Miller Jr., A review of cancer immunotherapy: from the past, to the present, to the future, Curr. Oncol. 27 (2020) S87–S97.
[8] J. Dine, R. Gordon, Y. Shames, M.K. Kasler, M. Barton-Burke, Immune checkpoint inhibitors: an innovation in immunotherapy for the treatment and management of patients with cancer, asia pac, J. Oncol. Nurs. 4 (2017) 127–135.
[9] R. Franzin, G.S. Netti, F. Spadaccino, C. Porta, L. Gesualdo, G. Stallone, G. Castellano, E. Ranieri, The use of immune checkpoint inhibitors in oncology and the occurrence of AKI: where do we stand? Front. Immunol. 11 (2020), 574271.
[10] O. Hemminki, J.M. Dos Santos, A. Hemminki, Oncolytic viruses for cancer immunotherapy, J. Hematol. Oncol. 13 (2020) 84.
[11] A. Ruiz-Patino, O. Arrieta, A.F. Cardona, C. Martin, L.E. Raez, Z.L. Zatarain-Barron, F. Barron, L. Ricaurte, M.A. Bravo-Garzon, L. Mas, L. Corrales, L. Rojas, L. Lupinacci, F. Perazzo, C. Bas, O. Carranza, C. Puparelli, M. Rizzo, R. Ruiz, C. Rolfo, P. Archila, J. Rodriguez, C. Sotelo, C. Vargas, H. Carranza, J. Otero, L. E. Pino, C. Ortiz, P. Laguado, R. Rosell, ClicaP, Immunotherapy at any line of treatment improves survival in patients with advanced metastatic non-small cell lung cancer (NSCLC) compared with chemotherapy (Quijote-CLICaP), Thorac. Cancer 11 (2020) 353–361.
[12] S. Amin, M.J. Baine, J.L. Meza, C. Lin, Association of immunotherapy with survival among patients with brain metastases whose cancer was managed with definitive surgery of the primary tumor, JAMA Netw. Open 3 (2020), e2015444.
[13] A.D. Waldman, J.M. Fritz, M.J. Lenardo, A guide to cancer immunotherapy: from T cell basic science to clinical practice, Nat. Rev. Immunol. 20 (2020) 651–668.
[14] K.F. Chan, B.S. Gully, S. Gras, D.X. Beringer, L. Kjer-Nielsen, J. Cebon, J. McCluskey, W. Chen, J. Rossjohn, Divergent T-cell receptor recognition modes of a HLA-I restricted extended tumour-associated peptide, Nat. Commun. 9 (2018), 1026.
[15] Q. He, X. Jiang, X. Zhou, J. Weng, Targeting cancers through TCR-peptide/MHC interactions, J. Hematol. Oncol. 12 (2019) 139.
[16] J.D. Buhrman, J.E. Slansky, Improving T cell responses to modified peptides in tumor vaccines, Immunol. Res. 55 (2013) 34–47.
[17] B. Engels, V.H. Engelhard, J. Sidney, A. Sette, D.C. Binder, R.B. Liu, D.M. Kranz, S. C. Meredith, D.A. Rowley, H. Schreiber, Relapse or eradication of cancer is predicted by peptide-major histocompatibility complex affinity, Cancer Cell 23 (2013) 516–526.
[18] Y. Sun, F. Li, H. Sonnemann, K.R. Jackson, A.H. Talukder, A.S. Katailiha, G. Lizee, Evolution of CD8(+) T cell receptor (TCR) engineered therapies for the treatment of cancer, Cells (2021) 10.
[19] M. Yarmarkovich, Q.F. Marshall, J.M. Warrington, R. Premaratne, A. Farrel, D. Groff, W. Li, M. di Marco, E. Runbeck, H. Truong, J.S. Toor, S. Tripathi, S. Nguyen, H. Shen, T. Noel, N.L. Church, A. Weiner, N. Kendsersky, D. Martinez, R. Weisberg, M. Christie, L. Eisenlohr, K.R. Bosse, D.S. Dimitrov, S. Stevanovic, N. G. Sgourakis, B.R. Kiefel, J.M. Maris, Cross-HLA targeting of intracellular oncoproteins with peptide-centric CARs, Nature 599 (2021) 477–484.
[20] P. Charoentong, F. Finotello, M. Angelova, C. Mayer, M. Efremova, D. Rieder, H. Hackl, Z. Trajanoski, Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade, Cell Rep. 18 (2017) 248–262.
[21] J. Wu, W. Zhao, B. Zhou, Z. Su, X. Gu, Z. Zhou, S. Chen, TSNAdb: a database for tumor-specific neoantigens from immunogenomics data analysis, Dev. Reprod. Biol. 16 (2018) 276–282.
[22] J. Xia, P. Bai, W. Fan, Q. Li, Y. Li, D. Wang, L. Yin, Y. Zhou, NEPdb: a database of T-cell experimentally-validated neoantigens and pan-cancer predicted neoepitopes for cancer immunotherapy, Front. Immunol. 12 (2021), 644637.
[23] X. Tan, D. Li, P. Huang, X. Jian, H. Wan, G. Wang, Y. Li, J. Ouyang, Y. Lin, L. Xie, dbPepNeo: a Manually Curated Database for Human Tumor Neoantigen Peptides, Database, Oxford), 2020, 2020.
[24] 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.
[25] S. Gupta, K. Chaudhary, S.K. Dhanda, R. Kumar, S. Kumar, M. Sehgal, G. Nagpal, G. P. Raghava, A platform for designing genome-based personalized immunotherapy or vaccine against cancer, PLoS One 11 (2016), e0166372.
[26] K. Tomczak, P. Czerwinska, M. Wiznerowicz, The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge, Contemp. Oncol. 19 (2015) A68–A77.
[27] T.J. O'Donnell, A. Rubinsteyn, U. Laserson, MHCflurry 2.0: improved pan-allele prediction of MHC class I-presented peptides by incorporating antigen processing, Cell Syst. 11 (2020) 42–48 e47.
[28] J. Scholtalbers, S. Boegel, T. Bukur, M. Byl, S. Goerges, P. Sorn, M. Loewer, U. Sahin, J.C. Castle, TCLP: an online cancer cell line catalogue integrating HLA type, predicted neo-epitopes, virus and gene expression, Genome Med. 7 (2015) 118.

[29] F. Sabbatino, L. Liguori, G. Polcaro, I. Salvato, G. Caramori, F.A. Salzano, V. Casolaro, C. Stellato, J.D. Col, S. Pepe, Role of human leukocyte antigen system as A predictive biomarker for checkpoint-based immunotherapy in cancer patients, Int. J. Mol. Sci. 21 (2020).
[30] A. Hazini, K. Fisher, L. Seymour, Deregulation of HLA-I in cancer and its central importance for immunotherapy, J. Immunother Cancer 9 (2021).
[31] V. Naranbhai, M. Viard, M. Dean, S. Groha, D.A. Braun, C. Labaki, S.A. Shukla, Y. Yuki, P. Shah, K. Chin, M. Wind-Rotolo, X.J. Mu, P.B. Robbins, A. Gusev, T. K. Choueiri, J.L. Gulley, M. Carrington, HLA-A*03 and response to immune checkpoint blockade in cancer: an epidemiological biomarker study, Lancet Oncol. 23 (2022) 172–184.
[32] A. Dhall, S. Patiyal, H. Kaur, S. Bhalla, C. Arora, G.P.S. Raghava, Computing skin cutaneous melanoma outcome from the HLA-alleles and clinical characteristics, Front. Genet. 11 (2020) 221.
[33] X. Zhang, T. Sjoblom, Targeting loss of heterozygosity: a novel paradigm for cancer therapy, Pharmaceuticals (Basel) 14 (2021).
[34] F. Garrido, N. Aptsiauri, E.M. Doorduijn, A.M. Garcia Lora, T. van Hall, The urgent need to recover MHC class I in cancers for effective immunotherapy, Curr. Opin. Immunol. 39 (2016) 44–51.
[35] K. Dhatchinamoorthy, J.D. Colbert, K.L. Rock, Cancer immune evasion through loss of MHC class I antigen presentation, Front. Immunol. 12 (2021), 636568.
[36] M. Esquivel-Velazquez, P. Ostoa-Saloma, M.I. Palacios-Arreola, K.E. Nava-Castro, J.I. Castro, J. Morales-Montor, The role of cytokines in breast cancer development and progression, J. Interferon Cytokine Res. 35 (2015) 1–16.
[37] M. Akdis, S. Burgler, R. Crameri, T. Eiwegger, H. Fujita, E. Gomez, S. Klunker, N. Meyer, L. O'Mahony, O. Palomares, C. Rhyner, N. Ouaked, A. Schaffartzik, W. Van De Veen, S. Zeller, M. Zimmermann, C.A. Akdis, Interleukins, from 1 to 37, and interferon-gamma: receptors, functions, and roles in diseases, J. Allergy Clin. Immunol. 127 (2011) 701–721, e701-770.
[38] G.A. Rabinovich, D. Gabrilovich, E.M. Sotomayor, Immunosuppressive strategies that are mediated by tumor cells, Annu. Rev. Immunol. 25 (2007) 267–296.
[39] L. Gorelik, R.A. Flavell, Immune-mediated eradication of tumors through the blockade of transforming growth factor-beta signaling in T cells, Nat. Med. 7 (2001) 1118–1122.
[40] P.M. Voorhees, R.Z. Orlowski, The proteasome and proteasome inhibitors in cancer therapy, Annu. Rev. Pharmacol. Toxicol. 46 (2006) 189–213.
[41] S.I. Grivennikov, M. Karin, Inflammation and oncogenesis: a vicious connection, Curr. Opin. Genet. Dev. 20 (2010) 65–71.
[42] F. Balkwill, Tumour necrosis factor and cancer, Nat. Rev. Cancer 9 (2009) 361–371.
[43] W. Ma, B.M. Gilligan, J. Yuan, T. Li, Current status and perspectives in translational biomarker research for PD-1/PD-L1 immune checkpoint blockade therapy, J. Hematol. Oncol. 9 (2016) 47.
[44] P. Sharma, S. Hu-Lieskovan, J.A. Wargo, A. Ribas, Primary, adaptive, and acquired resistance to cancer immunotherapy, Cell 168 (2017) 707–723.
[45] M.D. Rosenblum, I.K. Gratz, J.S. Paw, K. Lee, A. Marshak-Rothstein, A.K. Abbas, Response to self antigen imprints regulatory memory in tissues, Nature 480 (2011) 538–542.
[46] E. Andersson, L. Villabona, K. Bergfeldt, J.W. Carlson, S. Ferrone, R. Kiessling, B. Seliger, G.V. Masucci, Correlation of HLA-A02* genotype and HLA class I antigen down-regulation with the prognosis of epithelial ovarian cancer, Cancer Immunol. Immunother. 61 (2012) 1243–1253.
[47] X.C. Mayinuer Alifu1, Gulina Kuerban, Yaning Feng, Xuan Yao, Yanchun Peng, Yunhui Hu, Tao Dong, Ruozheng Wang, Distribution of HLA-A Alleles and its Relation to Clinical Outcome in Uyghur and Han Patients with Advanced Squamous Cell Cervical Cancer in Xinjiang, TCR, China, 2018.
[48] T. Masuda, H. Ito, J. Hirata, S. Sakaue, Y. Ueda, T. Kimura, F. Takeuchi, Y. Murakami, K. Matsuda, K. Matsuo, Y. Okada, Fine mapping of the major histocompatibility complex region and association of the HLA-B*52:01 allele with cervical cancer in Japanese women, JAMA Netw. Open 3 (2020), e2023248.
[49] Z. Sahin, M. Bicakcigil, K. Aksu, S. Kamali, S. Akar, F. Onen, O. Karadag, Z. Ozbalkan, A. Ates, H.T. Ozer, V. Yilmaz, E. Seyahi, M.A. Ozturk, A. Cefle, V. Cobankara, A.M. Onat, E. Tunc, N. Duzgun, S.Z. Aydin, N. Yilmaz, I. Fresko, Y. Karaaslan, S. Kiraz, N. Akkoc, M. Inanc, G. Keser, F.A. Uyar, H. Direskeneli, G. Saruhan-Direskeneli, G. Turkish Takayasu Study, Takayasu's arteritis is associated with HLA-B*52, but not with HLA-B*51, in Turkey, Arthritis Res. Ther. 14 (2012) R27.
[50] C. Terao, H. Yoshifuji, T. Matsumura, T.K. Naruse, T. Ishii, Y. Nakaoka, Y. Kirino, K. Matsuo, T. Origuchi, M. Shimizu, Y. Maejima, E. Amiya, N. Tamura, T. Kawaguchi, M. Takahashi, K. Setoh, K. Ohmura, R. Watanabe, T. Horita, T. Atsumi, M. Matsukura, T. Miyata, Y. Kochi, T. Suda, K. Tanemoto, A. Meguro, Y. Okada, A. Ogimoto, M. Yamamoto, H. Takahashi, S. Nakayamada, K. Saito, M. Kuwana, N. Mizuki, Y. Tabara, A. Ueda, I. Komuro, A. Kimura, M. Isobe, T. Mimori, F. Matsuda, Genetic determinants and an epistasis of LILRA3 and HLA-B*52 in Takayasu arteritis, Proc. Natl. Acad. Sci. U. S. A. 115 (2018) 13045–13050.
[51] M. Fishman, J.P. Dutcher, J.I. Clark, A. Alva, G.P. Miletello, B. Curti, N. Agarwal, R. Hauke, K.M. Mahoney, H. Moon, J. Treisman, S.S. Tykodi, G. Daniels, M. A. Morse, M.K.K. Wong, H. Kaufman, N. Gregory, D.F. McDermott, Overall survival by clinical risk category for high dose interleukin-2 (HD IL-2) treated patients with metastatic renal cell cancer (mRCC): data from the PROCLAIM(SM) registry, J. Immunother Cancer 7 (2019) 84.
[52] T. Jiang, C. Zhou, S. Ren, Role of IL-2 in cancer immunotherapy, OncoImmunology 5 (2016), e1163462.
[53] N. Karachaliou, M. Gonzalez-Cao, G. Crespo, A. Drozdowskyj, E. Aldeguer, A. Gimenez-Capitan, C. Teixido, M.A. Molina-Vila, S. Viteri, M. De Los Llanos Gil, S.M. Algarra, E. Perez-Ruiz, I. Marquez-Rodas, D. Rodriguez-Abreu, R. Blanco, T. Puertolas, M.A. Royo, R. Rosell, Interferon gamma, an important marker of response to immune checkpoint blockade in non-small cell lung cancer and melanoma patients, Ther. Adv. Med. Oncol. 10 (2018), 1758834017749748.
[54] L. Wang, M. Zhang, Y. Zhu, X. Zhang, Y. Yang, C. Wang, CCR4 expression is associated with poor prognosis in patients with early stage (pN0) oral tongue cancer, J. Oral Maxillofac. Surg. 77 (2019) 426–432.
[55] S. Punt, J.M. Langenhoff, H. Putter, G.J. Fleuren, A. Gorter, E.S. Jordanova, The correlations between IL-17 vs. Th17 cells and cancer patient survival: a systematic review, OncoImmunology 4 (2015), e984547.