The prevalence of HLA-I LOH in Chinese pan-cancer patients and genomic features of patients harboring HLA-I LOH

Jian Zhao¹, Xiaoxiong Xiao², Yue Li¹, Xuan Gao³, Xiuqin Zhang¹, Zeyi Liu¹, Yuting Yi³, Xiaorui Fu³, Han Wang³, Yanfang Guan³, Xuefeng Xia³, Weixing Zhang², and Jian'an Huang¹

¹First Affiliated Hospital of Soochow University ²Xiangya Hospital Central South University ³Geneplus-Beijing Institute

March 1, 2021

Abstract

Loss of heterozygosity in HLA-I (HLA-I LOH) may facilitate immune evasion. However, the large population study of HLA-I LOH in Chinese pan-cancer patients remains to be explored. In this study, analysis was performed in 1504 advanced pan-cancer patients and 134 early-stage NSCLC patients using a 1021-gene panel. The consistency between the 1021-gene panel and whole-exome sequencing (WES) was evaluated in 45 samples, where concordant results were obtained in 95.6% (43/45) of the samples. Analytical results revealed that the prevalence of HLA-I LOH presents considerable differences across cancer types. HLA-I LOH was relevant to genomic instability, reflected in higher tumor mutation burden (TMB) level. The incidence of HLA-I LOH in MSS samples was significantly higher than that in MSI-H samples. The alteration frequencies of p53 pathway, RTK/RAS pathway, Notch pathway, Hippo pathway, and Nrf2 pathway in HLA-I LOH group were significantly higher than that in HLA-I LOH group were significantly higher than that in HLA-I LOH group were significantly higher than that in HLA-I LOH group (p<0.0001, p<0.0001, p=0.032, p=0.013, p=0.003, respectively). In DNA damage response (DDR) pathways, alterations in CPF pathway and FA pathway are enriched in HLA-I LOH group (p<0.0001, p=0.023, respectively). Besides, HLA-I LOH was accompanied by higher mutation rates of several tumor suppressors, including TP53 and LRP1B. These results may shed light on follow-up research.

The prevalence of HLA-I LOH in Chinese pan-cancer patients and genomic features of patients harboring HLA-I LOH

Author list:

Jian Zhao^{1,2+}, Xiaoxiong Xiao³⁺, Yue Li^{1,2+}, Xuan Gao⁴, Xiuqin Zhang^{1,2}, Zeyi Liu^{1,2}, Yuting Yi⁴, Xiaorui Fu⁴, Han Wang⁴, Yanfang Guan⁴, Xuefeng Xia⁴, Weixing Zhang^{3*}, Jian'an Huang^{1,2*}

Affiliations:

- 1. Department of Pulmonary and Critical Care Medicine, The First Affiliated Hospital of Soochow University, Suzhou 215006, Jiangsu, China
- 2. Suzhou Key Laboratory for Respiratory Diseases, Suzhou 215006, Jiangsu, China
- 3. Thoracic Surgery Department, Xiangya Hospital Central South University, Changsha 410008, Hunan, China
- 4. Geneplus-Beijing, Beijing 102205, China

⁺Jian Zhao, Xiaoxiong Xiao, and Yue Li contributed equally to this work.

*Correspondence should be addressed to:

Dr. Weixing Zhang

Thoracic Surgery Department, Xiangya Hospital Central South University,

87 Xiangya Road, Changsha 410008, Hunan, China

E-mail:*zhangweixingdoctor@163.com*

Dr. Jian'an Huang

Department of Pulmonary and Critical Care Medicine, The First Affiliated Hospital of Soochow University,

899 Pinghai Road, Suzhou 215006, China

E-mail:

huang_jian_an@163.com

Abbreviation

HLA-I: human leukocyte antigen class I; HLA-I LOH: loss of heterozygosity in human leukocyte class I; WES: whole exome sequencing; TMB: tumor mutation burden; MSI: microsatellite instability; DDR: DNA damage response; CPF: check point factor; FA: Fanconi anemia; ICI: immune checkpoint inhibitor; PD-L1: programmed cell death-ligand 1; MHC-I: major histocompatibility complex class I; NK cell: natural killer cell; SqCC: squamous cell carcinoma; BRCA: breast invasive carcinoma; CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma; CRC: colorectal carcinoma; KIPAN: pan-kidney cohort; LIHC: liver hepatocellular carcinoma; LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma; NSCLC: non-small cell lung cancer; OV: ovarian serous cystadenocarcinoma; PAAD: pancreatic adenocarcinoma; PRAD: prostate adenocarcinoma; STES: stomach and esophageal carcinoma; UCEC: uterine corpus endometrial carcinoma.

Abstract

Loss of heterozygosity in HLA-I (HLA-I LOH) may facilitate immune evasion. However, the large population study of HLA-I LOH in Chinese pan-cancer patients remains to be explored. In this study, analysis was performed in 1504 advanced pan-cancer patients and 134 early-stage NSCLC patients using a 1021-gene panel. The consistency between the 1021-gene panel and whole-exome sequencing (WES) was evaluated in 45 samples, where concordant results were obtained in 95.6% (43/45) of the samples. Analytical results revealed that the prevalence of HLA-I LOH presents considerable differences across cancer types. HLA-I LOH was relevant to genomic instability, reflected in higher tumor mutation burden (TMB) level. The incidence of HLA-I LOH in MSS samples was significantly higher than that in MSI-H samples. The alteration frequencies of p53 pathway, RTK/RAS pathway, Notch pathway, Hippo pathway, and Nrf2 pathway in HLA-I LOH group were significantly higher than that in HLA-I stable group (p<0.0001, p=0.032, p=0.013, p=0.003, respectively). In DNA damage response (DDR) pathways, alterations in CPF pathway and FA pathway are enriched in HLA-I LOH group (p<0.0001, p=0.023, respectively). Besides, HLA-I LOH was accompanied by higher mutation rates of several tumor suppressors, including *TP53* and *LRP1B*. These results may shed light on follow-up research.

Keywords

Human leukocyte antigen complex; Loss of heterozygosity in HLA-I; Oncogenic signaling pathways; Genomic instability; DNA damage response system; Immunotherapy.

Introduction

Cancer treatment has been a global challenge. As part of the current standard treatment, immune checkpoint inhibitors (ICIs) therapy benefits the survival of a considerable number of patients across diverse cancer types (Havel et al., 2019; Hellmann et al., 2018; Samstein et al., 2019; Keenan et al., 2019). However, as the prognosis and response of ICI therapies vary greatly among patients, the need for effective biomarkers

to distinguish potential benefit population is urgent (Havel et al., 2019; Hellmann et al., 2018; Keenan et al., 2019; Chan et al., 2019). Current immunotherapy medication guidelines mainly focus on TMB level and PD-L1 expression, ignoring the role of antigen presentation in the activation of anti-tumor immune response (Hellmann et al., 2018; Samstein et al., 2019; Chan et al., 2019; Rizvi et al., 2018). The human leukocyte antigen (HLA) complex is the most polymorphic gene system in humans, which encode the heavy chain in the hetero-dimer of MHC-I (major histocompatibility complex class I) molecules (Parham et al., 1989; Klein & Sato, 2000). As essential components involved in tumor antigen presentation, MHC-I molecules present tumor neoantigens to the cell surface for recognition by T cell receptors and activate immune response (Bjorkman et al., 1987; Parham & Ohta, 1996; McGranahan et al., 2017). By interrupting tumor antigen recognition, tumor cells with defected antigen presentation may escape from anti-tumor immune clearance and survive, which in turn affects the efficacy of ICI therapies (Campoli & Ferrone, 2008; Hicklin et al., 1999; Hiraki et al., 2007).

Several studies have revealed the significance of HLA-I LOH in innate immunity. In particular, loss of alleles in HLA-C may inhibit immune clearance through the activation of NK cells (Thielens et al., 2012). In addition, multiple clinical researches have shown that HLA-I LOH is widespread in a variety of cancers and HLA-I LOH combined with traditional biomarkers could present better ICI therapies prognostic ability (Shim et al., 2020; Montesion et al., 2020). Shim et al. (2020) proposed that HLA-corrected TMB is a better predictor of ICI therapies than TMB alone. Montesion et al. (2020) pointed out that HLA-I LOH combined with TMB was a better factor for patient stratification in NSCLC. In addition, prior studies have revealed that the genotype of HLA-I alleles affects the response to immune checkpoint inhibitors (Chowell et al., 2018).

However, these previous studies were mainly based on western populations, while the incidence of HLA-I LOH in Chinese pan-cancer population remains unclear. In addition, the incidence of HLA-I LOH varies due to the differences in chromosomal instabilities among cancer types, highlighting the importance of HLA-I LOH studies in large population pan-cancer cohorts (Montesion et al., 2020). To illuminate this uncharted area, we analyzed the genomic variations in a 1638 Chinese pan-cancer cohort to delineate the HLA-I LOH landscape together with clinical and genomic features of those patients who harbored HLA-I LOH. In prior studies, the precise HLA-I LOH detection requires whole-exome sequencing (WES), which is not conducive in large-scale clinical applications due to the high cost of WES. Our study evaluated the reliability of HLA-I LOH analysis using the LOHHLA algorithm (McGranahan et al., 2017) based on a 1021-gene panel sequencing results. Exon 2, exon 3, and bilateral introns of HLA-A/B/C genes were well covered in the 1021-gene panel. We aimed to find the complex role of HLA-I LOH in tumor evolution which may provide valuable information for clinical practice and follow-up ICIs therapy research.

Results

Clinical characteristics of Chinese pan-cancer patients

Consisted of 1504 advanced pan-cancer patients and 134 early-stage NSCLC patients (Fig.1), this study included 1638 Chinese pan-cancer patients who had undergone tumor profiling with a 1021-gene panel and the prevalence of HLA-I LOH was analyzed. To assess the consistency of HLA-I LOH analysis between 1021-gene panel sequencing and WES, 45 samples were processed with both approaches, resulting in a consistency of 95.6% (43/45) (Fig.2A) which suggested that the 1021-gene panel sequencing is capable of detecting HLA-I genotype and HLA-I LOH status in this cohort. Among the 1638 patients, 24 (22 advanced patients and 2 early-stage patients) were excluded due to homozygosity at all three HLA-I loci (HLA-A, HLA-B, HLA-C), remaining with 1482 advanced pan-cancer patients and 132 early-stage NSCLC patients (Fig.1) for further analysis. The clinical characteristics of the study cohort are listed in Table 1. As shown in Table 1, patients in the HLA-I LOH group had a higher diagnosis age than those in the HLA-I stable group (median age 59 vs. 57, p=0.011), while no significant difference was found in gender (p=0.204). The distribution of patients among different cancer types and clinical data are shown in Supplementary Table 1.

Overview of HLA-I LOH events in Chinese pan-cancer patients

Among 1482 advanced pan-cancer patients, 45.6% (676/1482) were identified with LOH in at least one HLA-I locus (Fig.2B). The prevalence of HLA-I LOH presented considerable differences among cancer types, ranging from 12.2% (prostate adenocarcinoma) to 68.0% (cervical squamous cell carcinoma) (Fig.2C, Table 2), which was in accordance with previous studies that HLA-I LOH was more enriched in squamous cell carcinoma (68.0% in cervical squamous cell carcinoma, 62.2% in lung squamous cell carcinoma, Table 2) (Montesion et al., 2020). In NSCLC patients, there was an enrichment of HLA-I LOH in both early-stage (I-IIIa) LUSC patients (24/88 [27.3%] of LUAD vs. 26/42 [61.9%] of LUSC, p<0.001, Fig.2D) and advanced (IIIb-IV) LUSC patients (162/306 [52.9%] of LUAD vs. 156/251 [62.2%] of LUSC, p=0.029, Fig.2D), which was consistent with the results that HLA mutations occurred more frequently in LUSC (Shukla et al., 2015; McGranahan et al., 2017). However, in contrast to the enrichment of HLA-I LOH in advanced LUAD (24/88 [27.3%] of early-stage LUAD vs. 162/306 [52.9%] of advanced LUAD, p<0.001, Fig.2D), the incidence of HLA-I LOH between early-stage and advanced LUSC was equivalent (26/42 [61.9%] of early-stage LUSC vs. 156/251 [62.2%] of advanced LUSC, p=0.976, Fig.2D). In 43.3% (293/676) patients harboring HLA-I LOH, LOH co-occurred in HLA-A, B, and C (Fig.2E). The HLA-I gene status of HLA-I LOH patients in each cancer type was shown in Supplementary Figure 1.

The occurrence of HLA-I LOH is relevant to genomic instability

Consistent with previous studies, the occurrence of HLA-I LOH was relevant to an elevated TMB level (median TMB 6.72 vs. 5.76, p<0.0001, Fig.3A) (Anagnostou et al., 2020; McGranahan et al., 2017; Shim et al., 2020). In addition, the TMB level of patients with HLA-I LOH at all three loci was significantly higher than those with HLA-I LOH at only one locus (median TMB 7.68 vs. 6.72, p=0.027, Fig.3B). When distinguished by cancer types, significant differences were observed in Non-SqCC NSCLC, LUSC, STES and BRCA (p=0.009, p=0.021, p=0.029, p=0.049, respectively, Supplementary Figure 2). However, despite that MSI-H samples presented the highest incidence of TMB-H (identified as TMB[?]10.56, Supplementary Figure 3C), the incidence of HLA-I LOH decreased in these samples (Supplementary Figure 3C), which suggested that the incidence of HLA-I LOH might have a non-linear relationship with TMB. In colorectal cancer, the HLA-I LOH incidence of MSS samples was significantly higher than that of MSI-H samples (61/130 [46.9%] of MSS vs. 3/17 [17.6%] of MSI-H, p=0.022, Fig.3C). Conformably, in patients with other cancer types (Non-SqCC NSCLC, LUSC, BRCA, STES, OV, KIPAN, PAAD, PRAD, UCEC, CESC, LIHC), the ratio was 485/967 [50.2%] of MSS vs. 5/20 [25%] of MSI-H, p=0.026, Fig.3D). The MSI status of the study cohort was listed in Supplementary Table 2.

Alterations of multiple oncogenic signaling pathways are enriched in HLA-I LOH patients

We further analyzed the alteration frequencies of DDR pathways and 10 canonical oncogenic signaling pathways in advanced pan-cancer cohort. A signaling pathway was considered as altered if one or more genes in this pathway have non-synonymous mutations (Sanchez-Vega et al., 2018). The alteration frequencies of CPF pathway, FA pathway, p53 pathway, RTK/RAS pathway, Notch pathway, Hippo pathway and Nrf2 pathway in the HLA-I LOH group were significantly higher than that in HLA-I stable group (p<0.0001, p=0.023, p<0.0001, p<0.0001, p=0.032, p=0.013, p=0.003, respectively). Among these signaling pathways, the CPF pathway, p53 pathway and RTK/RAS pathway were most frequently mutated in the study cohort, in which the differences between the HLA-I LOH group and HLA-I group were the most significant (Fig.4A, Fig.4B). Consistent result was obtained in the NSCLC cohort, where the most frequently altered pathways had the most significant differences between groups (p53 pathway in LUSC and RTK/RAS pathway in Non-SqCC NSCLC, Supplementary Figure 4A, Supplementary Figure 4B).

Genomic features of patients harboring HLA-I LOH

Based on SNV data, we further investigated mutations in 1021 cancer genes within HLA-I LOH subtypes. Results showed that TP53, LRP1B, EGFR, KRAS, and PIK3CA were the most frequently altered genes in advanced pan-cancer cohort (Fig.4D). Among these genes, the alteration frequencies of TP53, EGFR, and LRP1B in the HLA-I LOH group were significantly higher than that in HLA-I stable group (p<0.001 for TP53, p<0.01 for EGFR and p<0.05 for LRP1B, Fig.4D). However, as the distribution of EGFR

mutation across cancer types biased in NSCLC patients $(220/242 \ [90.9\%])$ while the incidence of HLA-I LOH in NSCLC cohort differed from the average incidence in pan-cancer cohort (Fig.2C), the difference in pan-cancer *EGFR* mutation between HLA-I LOH and HLA-I stable group might not reflect the relationship between *EGFR* alteration and HLA-I LOH. In contrast, there was no significant difference in *EGFR* alteration frequency between the HLA-I LOH and HLA-I stable group in NSCLC cohort (Supplementary Figure 4C).

Discussion

In this study, we evaluated the prevalence of HLA-I LOH in Chinese cancer patients using the 1021-gene panel and characterized the genomic features of patients harboring HLA-I LOH.

Analytical results indicated that HLA-I LOH was widespread in Chinese cancer patients while varied greatly across cancer types, with an average incidence of 45.6%. When compared with a western population cohort containing 83644 samples, Chinese cancer patients shown a higher incidence of HLA-I LOH (45.6% vs. 17.0%) (Montesion et al., 2020). Since HLA-I LOH is a common immune evasion mechanism later in tumor evolution, we speculate that this difference may be due to the different tumor progression stages of the patients in the study cohort (McGranahan et al., 2017). The percentage of LOH co-occurrence at all three HLA-I loci in HLA-I LOH patients is 43.3% (293/676) in Chinese cohort and 85% in the western population cohort, respectively. (Montesion et al., 2020). Although there is a difference in the frequency of HLA-I LOH, the proportion of patients with LOH at all three sites concurrently is roughly equivalent (19.7% in our cohort vs. 14.5% in western cohort) (Montesion et al., 2020). As confirmed by previous studies, HLA-I LOH is enriched in squamous cell carcinoma (Montesion et al., 2020). In our cohort, the highest incidence of HLA-I LOH occurs in cervical squamous cell carcinoma and lung squamous cell carcinoma. Consistent with prior researches, the occurrence of HLA-I LOH was accompanied by elevated tumor mutation burden (TMB), suggesting the genomic instability in patients harboring HLA-I LOH (Anagnostou et al., 2020; McGranahan et al., 2017). When focusing on patients harboring HLA-I LOH, the TMB level of patients with LOH at a single site was lower than with LOH at all three sites. As one of the important indicators of immunotherapy, high TMB level indicated better immunotherapy prognosis (Hellmann et al., 2018; Samstein et al., 2019; Chan et al., 2019; Rizvi et al., 2018). Considering that the occurrence of multi-loci HLA-I LOH may impair antigen binding and presentation, the survival benefits of high TMB levels could be diminished, in which circumstances patients might have a poor response to immunotherapy. Thus, in clinical practice, the TMB level should be evaluated along with the occurrence of HLA-I LOH and number of LOH loci when predicting the response to immunotherapy.

However, when dividing TMB more finely, the relationship between HLA-I LOH and TMB is more in line with the "Goldilocks" pattern that tumors with low or high TMB levels show a low incidence of HLA-I LOH, while higher incidence is presented in between (Supplementary Figure 3A) (Montesion et al., 2020). Further analysis find that MSI-H patients have significantly higher TMB levels but lower incidence of HLA-I LOH compared with MSS patients, which led to the "Goldilocks" phenomenon in our cohort (Supplementary Figure 3C). Montesion et al. (2020) hypothesized "HLA-I LOH is selected for in tumors with enough neoantigens to elicit an immune response but few enough that HLA-I LOH abrogates immune recognition without eliciting an NK cell response". We believe that the reasons behind this phenomenon need to be further explored by analyzing the genomic differences of those patients. In addition, the higher incidence of HLA-I LOH in MSS patients suggests that HLA-I LOH may be one of the reasons for the poor response to immunotherapy in some MSS patients, prompting the underlying significance of HLA-I LOH testing in predicting potential benefit population in MSS patients. Since MSI-H group mainly consist of patients with gastrointestinal tumors (59.5% [22/37]) in our cohort, conclusions above might be specifically reflected in these cancer types.

Another important discovery is that patients harboring HLA-I LOH have unique genomic features, such as increased alteration frequencies of several signaling pathways and oncogenes. Since the occurrence of HLA-I LOH is closely related to genomic instability, we first analyzed the differences in the alterations in DDR pathways (McGranahan et al., 2017). Results revealed that alteration frequencies of CPF pathway and FA pathway in the HLA-I LOH group was significantly higher than that of the HLA-I stable group.

In addition, the occurrence of HLA-I LOH was accompanied by alterations in several important oncogenic signaling pathways, including p53 pathway, RTK/RAS pathway, Hippo pathway, Notch pathway and Nrf2 pathway, indicating that HLA-I LOH might involve in tumor progression by influencing the normal signaling transduction. As a signaling cascade pathway initiated by activation of RTKs, RTK/RAS pathway regulates cell proliferation and survival and is frequently altered in many cancer types (Malumbres & Barbacid, 2003; Sanchez-Vega et al., 2018). For several oncogenic genes, such as TP53 and LRP1B, the occurrence of HLA-I LOH is related to higher mutation frequencies. As the central gene of p53 pathway and CPF pathway, TP53 regulates apoptosis, cell cycle arrest, senescence, and DNA repair (Wade et al., 2013; Sanchez-Vega et al., 2018). The somatic mutations of this important tumor suppressor may cause genomic instability and replication stress, which is closely related to tumorigenesis in various cancer types (Nigro et al., 1989; Baker et al., 1989; Stracquadanio et al., 2016; Malekzadeh et al., 2019). LRP1Bmay play an essential role in tumorigenesis of NSCLCs, and mutations of this gene are associated with a higher TMB level (Lan et al., 2019; Wang et al., 2021). These factors may be involved in positive selection during tumor evolution and promote the occurrence of HLA-I LOH. Therefore, the mechanism underlying these observations and the causal relationship between HLA-I LOH and mutations of certain genes need further investigation.

One of the major limitations of this study is the insufficient collection of clinical information, such as smoking history and clinical outcome, which precluded us to set a cohort to analyze the clinical significance of our findings. Another shortcoming is that the number of samples used for analysis is relatively small compared with other large population studies. In our follow-up study, after enrolling a sufficient number of patients and collecting adequate clinical data, these limitations will be overcome. Therefore, deeper mechanism mining and clinical verification will be achieved.

In general, we confirmed the validity of 1021-gene panel for HLA-I LOH detection through comparison with whole-exome sequencing (WES). The 1021-gene panel could be applied for HLA-I LOH analysis, provided that the relevant regions are well captured. We further depicted the landscape of HLA-I LOH events in Chinese cancer patients and revealed the genomic features of patients harboring HLA-I LOH. These insights may provide valuable information for clinical practice and follow-up immunotherapy research.

Materials and methods

Study cohort and tumor samples

The study cohort in this research included 1638 Chinese cancer patients across 11 cancer types, who had undergone tumor profiling using target capture next generation sequencing (NGS) of 1021 cancer genes in tumor DNA and paired germline DNA as part of the clinical care. This study was approved by the Academic Advisory Board of Soochow University and all participants signed a written informed consent.

Sample processing and DNA extraction

All fresh samples were stored in -80 refrigerators immediately till all sample collection was finished. Peripheral blood was collected in EDTA Vacutainer tubes (BD Diagnostics, Franklin Lakes, NJ, USA) and processed within 4 h. Peripheral blood lymphocytes (PBLs) from the first centrifugation were used for the extraction of germline genomic DNA. The genomic DNA of frozen tumor tissue and matched peripheral blood lymphocytes samples were isolated using a commercially available DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany). The genomic DNA of FFPE samples were isolated using Maxwell® RSC DNA FFPE Kit (Promega, Madison, WI, USA). The DNA concentration was measured using a Qubit fluorometer and the Qubit dsDNA HS (High Sensitivity) Assay Kit (Invitrogen, Carlsbad, CA, USA). The total DNA yield must be greater than or equal to 1 μ g, while 260/280 and 260/230 is greater than or equal to 1.8 and 2, respectively.

Sequencing library construction and target enrichment

Before library construction, 1 µg each of genomic DNA extracted from PBL and tumor specimen was sheared to 300-bp fragments with a Covaris S2 ultrasonicator (Covaris, Woburn, MA, USA). A volume of 20–80 ng DNA from plasma were used for library construction. Indexed Illumina next-generation sequencing (NGS)

libraries were prepared from PBL DNA and tumor DNA using the KAPA Library Preparation Kit (Kapa Biosystems, Wilmington, MA, USA).

Target enrichment was performed with a custom SeqCap EZ Library (Roche NimbleGen, Madison, WI, USA). To explore the comprehensive genetic properties of Chinese pan-cancer patients, the capture probe was designed based on ~1.04 Mb genomic regions of 1021 genes frequently mutated in common solid tumors. Capture hybridization was carried out according to the manufacturer's protocol. Following hybrid selection, the captured DNA fragments were amplified and then pooled to generate several multiplex libraries.

NGS sequencing

Sequencing was carried out using Illumina 2×75 bp paired-end reads on an Illumina HiSeq 3000 instrument according to the manufacturer's recommendations using TruSeq PE Cluster Generation Kit v3 and the TruSeq SBS Kit v3 (Illumina, San Diego, CA, USA).

Sequence data analysis

After removal of terminal adaptor sequences and low-quality data, reads were mapped to the reference human genome (hg19) and aligned using BWA (0.7.12-r1039). MuTect2 (Cibulskis et al., 2013) (3.4–46-gbc02625) was employed to call somatic small insertions and deletions (InDels) and single nucleotide variants (SNVs). Mutations were considered as a candidate somatic mutation only when (i) the mutation had at least five high-quality reads (Phred score [?]30, mapping quality [?]30, and without paired-end reads bias) containing the particular base; (ii) the mutation was not presented in >1% of population in the 1,000 Genomes Project or dbSNP databases (The Single Nucleotide Polymorphism Database); and (iii) the mutation was not present in an on-house database of normal samples. For somatic tumor mutations, a mutant allele must be present in [?]3% of reads. Somatic non-synonymous mutations per megabase of the panel region were used in tumor mutation burden (TMB) analysis.

Contra (Li et al., 2012) (2.0.8) was used to detect copy number variations and LOH HLA algorithm was used to identify LOH based on informative SNPs. For structure variations (SV), baits were designed to capture selected exons and introns of RET, ALK, ROS1, and NTRK1 oncogenes based on previously reported SVs in these genes and an in-house algorithm was used to identified split-read and discordant read-pair to identify SVs. All final candidate variants were manually verified with the integrative genomics viewer browser.

Statistical analysis

Mann-Whitney test was used to compare TMB between groups. The Chi-square test or Fisher's exact test was used to test frequency between groups. All statistical analysis were carried out using GraphPad Prism software (v. 8.0.2; GraphPad Software, La Jolla, CA, USA) and R 4.0.2 software (R Foundation for Statistical Computing, Vienna, Austria). Statistical significance was defined as a two-sided p-value of <0.05).

Acknowledgement

We sincerely thank Dr. Jian'an Huang, Dr. Weixing Zhang, Dr. Xiaoxiong Xiao, Dr. Xuan Gao, Dr. Han Wang, Dr. Xuefeng Xia, and Dr. Yanfang Guan for their constructive discussions. We thank the patients, families, nurses and investigators who participated in this study.

References

Anagnostou, V., Niknafs, N., Marrone, K., Bruhm, D. C., White, J. R., Naidoo, J., Hummelink, K., Monkhorst, K., Lalezari, F., Lanis, M., Rosner, S., Reuss, J. E., Smith, K. N., Adleff, V., Rodgers, K., Belcaid, Z., Rhymee, L., Levy, B., Feliciano, J., ... Velculescu, V. E. (2020). Multimodal genomic features predict outcome of immune checkpoint blockade in non-small-cell lung cancer. Nature Cancer, 1(1). https://doi.org/10.1038/s43018-019-0008-8

Baker, S. J., Fearon, E. R., Nigro, J. M., Hamilton, S. R., Preisinger, A. C., Jessup, J. M., vanTuinen, P., Ledbetter, D. H., Barker, D. F., Nakamura, Y., White, R., & Vogelstein, B. (1989). Chromosome

17 deletions and p53 gene mutations in colorectal carcinomas. Science (New York, N.Y.), 244(4901). https://doi.org/10.1126/science.2649981

Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L., & Wiley, D. C. (1987). The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. Nature, 329(6139). https://doi.org/10.1038/329512a0

Campoli, M., & Ferrone, S. (2008). HLA antigen changes in malignant cells: Epigenetic mechanisms and biologic significance. Oncogene, 27(45). https://doi.org/10.1038/onc.2008.273

Chan, T. A., Yarchoan, M., Jaffee, E., Swanton, C., Quezada, S. A., Stenzinger, A., & Peters, S. (2019). Development of tumor mutation burden as an immunotherapy biomarker: Utility for the oncology clinic. Annals of Oncology, 30(1), 44–56. https://doi.org/10.1093/annonc/mdy495

Chowell, D., Morris, L. G. T., Grigg, C. M., Weber, J. K., Samstein, R. M., Makarov, V., Kuo, F., Kendall, S. M., Requena, D., Riaz, N., Greenbaum, B., Carroll, J., Garon, E., Hyman, D. M., Zehir, A., Solit, D., Berger, M., Zhou, R., Rizvi, N. A., & Chan, T. A. (2018). Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. Science, 359(6375), 582–587. https://doi.org/10.1126/science.aao4572

Cibulskis, K., Lawrence, M. S., Carter, S. L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E. S., & Getz, G. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nature Biotechnology, 31(3). https://doi.org/10.1038/nbt.2514

Havel, J. J., Chowell, D., & Chan, T. A. (2019). The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. Nature Reviews Cancer, 19(3), 133–150. https://doi.org/10.1038/s41568-019-0116-x

Hellmann, M. D., Nathanson, T., Rizvi, H., Creelan, B. C., Sanchez-Vega, F., Ahuja, A., Ni, A., Novik, J. B., Mangarin, L. M. B., Abu-Akeel, M., Liu, C., Sauter, J. L., Rekhtman, N., Chang, E., Callahan, M. K., Chaft, J. E., Voss, M. H., Tenet, M., Li, X.-M., ... Wolchok, J. D. (2018). Genomic Features of Response to Combination Immunotherapy in Patients with Advanced Non-Small-Cell Lung Cancer. Cancer Cell, 33(5), 843-852.e4. https://doi.org/10.1016/j.ccell.2018.03.018

Hicklin, D. J., Marincola, F. M., & Ferrone, S. (1999). HLA class I antigen downregulation in human cancers: T-cell immunotherapy revives an old story. MOLECULAR MEDICINE TODAY, 5, 9.

Hiraki, A., Fujii, N., Murakami, T., Kiura, K., Aoe, K., Yamane, H., Masuda, K., Maeda, T., Sugi, K., Darzynkiewicz, Z., Tanimoto, M., & Harada, M. (2004). High Frequency of Allele-specific Down-regulation of HLA class I Expression in Lung Cancer Cell Lines. ANTICANCER RESEARCH, 4.

Keenan, T. E., Burke, K. P., & Van Allen, E. M. (2019). Genomic correlates of response to immune checkpoint blockade. Nature Medicine, 25(3), 389–402. https://doi.org/10.1038/s41591-019-0382-x

Klein, J., & Sato, A. (2000). The HLA system. First of two parts. The New England Journal of Medicine, 343(10). https://doi.org/10.1056/NEJM200009073431006

Lan, S., Li, H., Liu, Y., Ma, L., Liu, X., Liu, Y., Yan, S., & Cheng, Y. (2019). Somatic mutation of *LRP1B* is associated with tumor mutational burden in patients with lung cancer. Lung Cancer (Amsterdam, Netherlands), 132. https://doi.org/10.1016/j.lungcan.2019.04.025

Li, J., Lupat, R., Amarasinghe, K. C., Thompson, E. R., Doyle, M. A., Ryland, G. L., Tothill, R. W., Halgamuge, S. K., Campbell, I. G., & Gorringe, K. L. (2012). CONTRA: copy number analysis for targeted resequencing. Bioinformatics (Oxford, England), 28(10). https://doi.org/10.1093/bioinformatics/bts146

Malekzadeh, P., Pasetto, A., Robbins, P. F., Parkhurst, M. R., Paria, B. C., Jia, L., Gartner, J. J., Hill, V., Yu, Z., Restifo, N. P., Sachs, A., Tran, E., Lo, W., Somerville, R. P., Rosenberg, S. A., & Deniger, D. C. (2019). Neoantigen screening identifies broad *TP53* mutant immunogenicity in patients with epithelial cancers. The Journal of Clinical Investigation, 129(3). https://doi.org/10.1172/JCI123791

Malumbres, M., & Barbacid, M. (2003). RAS oncogenes: The first 30 years. Nature Reviews. Cancer, 3(6). https://doi.org/10.1038/nrc1097

McGranahan, N., Rosenthal, R., Hiley, C. T., Rowan, A. J., Watkins, T. B. K., Wilson, G. A., Birkbak, N. J., Veeriah, S., Van Loo, P., Herrero, J., Swanton, C., Swanton, C., Jamal-Hanjani, M., Veeriah, S., Shafi, S., Czyzewska-Khan, J., Johnson, D., Laycock, J., Bosshard-Carter, L., ... Dessimoz, C. (2017). Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. Cell, 171(6), 1259-1271.e11. https://doi.org/10.1016/j.cell.2017.10.001

Mehta, A. M., Jordanova, E. S., Kenter, G. G., Ferrone, S., & Fleuren, G.-J. (2007). Association of antigen processing machinery and HLA class I defects with clinicopathological outcome in cervical carcinoma. Cancer Immunology, Immunotherapy, 57(2), 197–206. https://doi.org/10.1007/s00262-007-0362-8

Montesion, M., Murugesan, K., Jin, D. X., Sharaf, R., Sanchez, N., Guria, A., Minker, M., Li, G., Fisher, V., Sokol, E. S., Pavlick, D. C., Moore, J. A., Braly, A., Singal, G., Fabrizio, D., Comment, L. A., Rizvi, N. A., Alexander, B. M., Frampton, G. M., ... Albacker, L. A. (2020). Somatic HLA class I loss is a widespread mechanism of immune evasion which refines the use of tumor mutational burden as a biomarker of checkpoint inhibitor response. Cancer Discovery, CD-20-0672. https://doi.org/10.1158/2159-8290.CD-20-0672

Nigro, J. M., Baker, S. J., Preisinger, A. C., Jessup, J. M., Hostetter, R., Cleary, K., Bigner, S. H., Davidson, N., Baylin, S., & Devilee, P. (1989). Mutations in the p53 gene occur in diverse human tumour types. Nature, 342(6250). https://doi.org/10.1038/342705a0

Parham, P., Benjamin, R. J., Chen, B. P., Clayberger, C., Ennis, P. D., Krensky, A. M., Lawlor, D. A., Littman, D. R., Norment, A. M., & Orr, H. T. (1989). Diversity of class I HLA molecules: Functional and evolutionary interactions with T cells. Cold Spring Harbor Symposia on Quantitative Biology. https://doi.org/10.1101/sqb.1989.054.01.063

Parham, P., & Ohta, T. (1996). Population biology of antigen presentation by MHC class I molecules. Science (New York, N.Y.), 272(5258). https://doi.org/10.1126/science.272.5258.67

Rizvi, H., Sanchez-Vega, F., La, K., Chatila, W., Jonsson, P., Halpenny, D., Plodkowski, A., Long, N., Sauter, J. L., Rekhtman, N., Hollmann, T., Schalper, K. A., Gainor, J. F., Shen, R., Ni, A., Arbour, K. C., Merghoub, T., Wolchok, J., Snyder, A., ... Hellmann, M. D. (2018). Molecular Determinants of Response to Anti–Programmed Cell Death (PD)-1 and Anti–Programmed Death-Ligand 1 (PD-L1) Blockade in Patients With Non–Small-Cell Lung Cancer Profiled With Targeted Next-Generation Sequencing. Journal of Clinical Oncology, 36(7), 633–641. https://doi.org/10.1200/JCO.2017.75.3384

Samstein, R. M., Lee, C.-H., Shoushtari, A. N., Hellmann, M. D., Shen, R., Janjigian, Y. Y., Barron, D. A., Zehir, A., Jordan, E. J., Omuro, A., Kaley, T. J., Kendall, S. M., Motzer, R. J., Hakimi, A. A., Voss, M. H., Russo, P., Rosenberg, J., Iyer, G., Bochner, B. H., ... Morris, L. G. T. (2019). Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nature Genetics, 51(2), 202–206. https://doi.org/10.1038/s41588-018-0312-8

Sanchez-Vega, F., Mina, M., Armenia, J., Chatila, W. K., Luna, A., La, K. C., Dimitriadoy, S., Liu, D. L., Kantheti, H. S., Saghafinia, S., Chakravarty, D., Daian, F., Gao, Q., Bailey, M. H., Liang, W.-W., Foltz, S. M., Shmulevich, I., Ding, L., Heins, Z., ... Schultz, N. (2018). Oncogenic Signaling Pathways in The Cancer Genome Atlas. Cell, 173(2). https://doi.org/10.1016/j.cell.2018.03.035

Shim, J. H., Kim, H. S., Cha, H., Kim, S., Kim, T. M., Anagnostou, V., Choi, Y.-L., Jung, H. A., Sun, J.-M., Ahn, J. S., Ahn, M.-J., Park, K., Park, W.-Y., & Lee, S.-H. (2020). HLA-corrected tumor mutation burden and homologous recombination deficiency for the prediction of response to PD-(L)1 blockade in advanced non-small-cell lung cancer patients. Annals of Oncology, 31(7), 902–911. https://doi.org/10.1016/j.annonc.2020.04.004

Shukla, S. A., Rooney, M. S., Rajasagi, M., Tiao, G., Dixon, P. M., Lawrence, M. S., Stevens, J., Lane, W. J., Dellagatta, J. L., Steelman, S., Sougnez, C., Cibulskis, K., Kiezun, A., Hacohen, N., Brusic, V., Wu, C.

J., & Getz, G. (2015). Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. Nature Biotechnology, 33(11). https://doi.org/10.1038/nbt.3344

Stracquadanio, G., Wang, X., Wallace, M. D., Grawenda, A. M., Zhang, P., Hewitt, J., Zeron-Medina, J., Castro-Giner, F., Tomlinson, I. P., Goding, C. R., Cygan, K. J., Fairbrother, W. G., Thomas, L. F., Saetrom, P., Gemignani, F., Landi, S., Schuster-Bockler, B., Bell, D. A., & Bond, G. L. (2016). The importance of p53 pathway genetics in inherited and somatic cancer genomes. Nature Reviews. Cancer, 16(4). https://doi.org/10.1038/nrc.2016.15

Thielens, A., Vivier, E., & Romagne, F. (2012). NK cell MHC class I specific receptors (KIR): From biology to clinical intervention. Current Opinion in Immunology, 24(2), 239–245. https://doi.org/10.1016/j.coi.2012.01.001

Wade, M., Li, Y.-C., & Wahl, G. M. (2013). MDM2, MDMX and p53 in oncogenesis and cancer therapy. Nature Reviews Cancer, 13(2), 83–96. https://doi.org/10.1038/nrc3430

Wang, L., Yan, K., He, X., Zhu, H., Song, J., Chen, S., Cai, S., Zhao, Y., & Wang, L. (2021). *LRP1B* or mutations are associated with higher tumor mutational burden and worse survival in hepatocellular carcinoma. Journal of Cancer, 12(1). https://doi.org/10.7150/jca.48983

Funding

This work was supported by Jiangsu Provincial Key Medical Discipline Laboratory (No. ZDXKB2016007), The Science and Technology Development Projects of Suzhou (No. SYS2020008), Jiangsu Provincial Medical Youth Talent (No. QNRC2016746), Gusu Youth Health Talent of Suzhou (No. 2020-082), and Suzhou Key Laboratory for Respiratory Medicine (No. SZS201617).

Author Contributions

Jian'an Huang and Weixing Zhang conceived the overall study. Xuefeng Xia, Yanfang Guan, and Yuting Yi designed the study, Xiaorui Fu constructed and managed the sequencing and variant-calling pipelines. Xiaoxiong Xiao, Xiuqin Zhang, Zeyi Liu and Yue Li collected the samples and clinical data. Jian Zhao analyzed the data. Xuan Gao aided in statistical analyses. Jian Zhao wrote the manuscript. Xuan Gao and Han Wang participated the optimization of the manuscript. All authors contributed to writing the manuscript. All authors reviewed and approved the final draft. Jian'an Huang and Weixing Zhang jointly supervised this work.

Competing Interests

Xuan Gao, Yuting Yi, Xiaorui Fu, Han Wang, Yanfang Guan, and Xuefeng Xia are current employees of Geneplus-Bejing. The remaining authors declare no conflict of interest.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Figure legends

Figure 1. Study scheme of patient cohort. Patients across 12 cancer types went through targeted sequencing as a part of clinical care from May 2017 to October 2019 were included for further analysis. NGS, next generation sequencing; NSCLC, non-small cell lung cancer; HLA-I LOH, loss of heterozygosity in human leukocyte antigen class I locus.

Figure 2. Overview of the occurrence of HLA-I LOH in Chinese cancer patients. A. The consistency in HLA-I LOH analysis of 1021-gene panel and whole-exome sequencing (WES).B. Landscape of HLA-I LOH in Chinese patients with advanced cancer. C. The incidence of HLA-I LOH across cancer types. There is an enrichment for HLA-I LOH in squamous cell carcinoma.D. The occurrence of HLA-I LOH in NSCLC cohort. There is a significant difference in the incidence of HLA-I LOH between LUAD

and LUSC in both early-stage NSCLC (24/88 [27.3%] of LUAD vs. 26/42 [61.9%] of LUSC, p<0.001) and advanced patients (162/306 [52.9%] of LUAD vs. 156/251 [62.2%] of LUSC, p=0.029). E. Status of HLA-I locus in advanced patients harboring HLA-I LOH. In 43.3% (293/676) patients harboring HLA-I LOH, LOH occurred simultaneously at all three HLA class-I loci (HLA-A, B, C).

Figure 3. The relationship between HLA-I LOH and genomic instability. A. The tumor mutation burden (TMB) level of HLA-I LOH group is significantly higher than that of HLA stable group (median TMB 6.72 vs. 5.76, p<0.0001, Mann-Whitney test), indicating genomic instability in patients harboring HLA-I LOH. B. The TMB level of patients with LOH at one single site is significantly lower than that of patients with LOH co-occurrence at all three sites (median TMB 6.72 vs. 7.68, p=0.027, Mann-Whitney test). C. The proportion of HLA-I LOH in MSS samples is significantly higher than that in MSI-H samples in advanced CRC patients (61/130 [46.9%] vs. 3/17 [17.6%], p=0.022). D. Consistent results were obtained in patients with other cancer types (485/967 [50.2%] vs. 5/20 [25%], p=0.026).

Figure 4. The occurrence of HLA-I LOH is related to alterations in several oncogenic signaling pathways and genes. A. The significance histogram of DDR pathways in advanced pan-cancer cohort. The alteration frequencies of CPF pathway and FA pathway in HLA-I LOH group are significantly higher than that of HLA-I stable patients. (p<0.0001, p=0.023, respectively). B. The significance histogram of 10 oncogenic signaling pathways in advanced pan-cancer cohort. Statistical results revealed that the alteration frequencies of p53 pathway, RTK/RAS pathway, Notch pathway, Hippo pathway and Nrf2 pathway in HLA LOH group are significantly higher than HLA stable group (p<0.0001, p=0.032, p=0.013, p=0.003, respectively). C. Overview of alterations of 10 oncogenic signaling pathways in advanced pancancer cohort. D. Waterfall chart showing mutation details of patients in advanced pan-cancer cohort.

Supplementary Figure 1. Status of HLA-I locus in advanced patients harboring HLA-I LOH across different cancer types.

Supplementary Figure 2. Comparison of TMB levels between HLA-I LOH group and HLA-I stable group across different cancer types.

Supplementary Figure 3. A. When dividing TMB more finely, the relationship between HLA-I LOH and TMB is more in line with the "Goldilocks" pattern. **B.** Comparison of TMB levels between MSI-H group and MSS group. **C.** There is a decrease in the incidence of HLA-I LOH in MSI-H samples, suggesting that the incidence of HLA-I LOH may not linear correlated with TMB.

Supplementary Figure 4. A. The alteration frequencies of ten oncogenic pathways in advanced non-squamous NSCLC cohort. **B.** The alteration frequencies of ten oncogenic pathways in advanced LUSC cohort. **C.** Waterfall chart showing mutation details of patients in advanced NSCLC (LUSC and Non-SqCC NSCLC) cohort.

Main tables

Table 1. Clinical characteristics of the study cohort.

Table 2. The occurrence of HLA-I LOH across different cancer types.

Supplementary information

Supplementary Table 1. Distribution of patients with various cancers and clinical characteristics.

Supplementary Table 2. MSI status of patients across different cancer types.

Supplementary data

Supplementary data 1. Clinical information of 1638 pan-cancer patients.

Supplementary data 2. 1021 genes in the NGS panel.



Landscape of HLA-I LOH in Chinese advanced pan-cancer patients (n=1482)



The prevalence of HLA-I LOH across cancer types (%)







HLA-I status 🔄 HLA-I LOH 🖮 HLA-I stable







Advanced CRC patients(n=147)



Advanced patients across 11 cancer types(n=987)









Hosted file

Table 1.pdf available at https://authorea.com/users/398779/articles/511388-the-prevalenceof-hla-i-loh-in-chinese-pan-cancer-patients-and-genomic-features-of-patients-harboringhla-i-loh

Hosted file

Table 2.pdf available at https://authorea.com/users/398779/articles/511388-the-prevalenceof-hla-i-loh-in-chinese-pan-cancer-patients-and-genomic-features-of-patients-harboringhla-i-loh