Evaluation of tumor acceleration by HLA-I zygosity and identification of the immune characteristics in clear cell renal cell carcinoma

BeumJin Park

Department of Medical Science

The Graduate School, Yonsei University

Evaluation of tumor acceleration by HLA-I zygosity and identification of the immune characteristics in clear cell renal cell carcinoma

BeumJin Park

Department of Medical Science

The Graduate School, Yonsei University

Evaluation of tumor acceleration by HLA-I zygosity and identification of the immune characteristics in clear cell renal cell carcinoma

Directed by Professor Sangwoo Kim

The Master's Thesis
submitted to the Department of Medical Science,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree of
Master of Medical Science

BeumJin Park

June 2021

This certifies that the Master's Thesis of
BeumJin Park is approved.

Thesis Supervisor : Sangwoo Kim

Thesis Committee Member#1 : Inkyung Jung

Thesis Committee Member#2 : Han Sang Kim

The Graduate School

Yonsei University

June 2021

ACKNOWLEDGEMENTS

진로에 대해서 갈팡질팡하던 끝에 연구라는 도전을 해보겠다는 다짐으로 대학원에 진학하였고 시간이 빠르게 지나 졸업을 앞두었습니다. 열정으로 가득 찬 TGIL에서의 학위과정은 끊임없이 탐구하고 동시에 다방면으로 성장할 수 있는 시간이었습니다. 무엇보다도 학생들 개개인의 잠재력을 일깨워 독립적인 연구자로서 성장하도록 이끌어 주시는 지도교수 김상우 교수님의 가르침 속에서 많은 배움을 얻을 수 있었습니다. 또한, 이 논문이 완성되기까지 통계적인 가르침과 조언을 주신 정인경 교수님, 해박한 종양학적 지식으로 코멘트 주신 김한상 교수님 감사합니다. 그리고 논문의 모든 부분 열정을 가득 담아 지도해 주신 김상우 교수님께 감사의 말씀을 드립니다.

VFOS 프로젝트에서 여러 경험을 하게 해 주신 양인석 박사님. 항상 같이 상의해주고 조언해주었던 정신적 지주 미경 누나, 강형. 면역팀에서 저를 트레이닝 시켜준 소라 누나, 은영 누나. 서버관리에 힘써준 현주 형. 취미 부자 랩장 세영이 형. 부지런하고 닮고 싶은 것이 많은 혜인 누나. 친화력 갑 건호 형. 멋지고 긍정적인 유진이. 걸어 다니는 백과사전 택 쌤. 탐구력 끝판왕 신원 쌤. 자랑스러운 부사수 지윤이. 다재다능 정수 형. 항상 발전하는 승희. 면역팀의 미래를 밝혀줄 요한이, 아연 쌤, 동선 쌤. 해맑고 멘탈 강한 지호. ENFJ 판박이 알고리즘 고수 승석이 형. 노력하는 모습이 멋진 승호 형. 함께 지낸 시간이 짧아서 아쉬운 카리스마 의사 영수 쌤, 현구 쌤. 실력자 인턴 준한 쌤, 지수 쌤. TGIL의 모든 분들 덕분에 석사과정 동안 즐겁고 성실하게 연구할 수 있었습니다. 학구적인 분위기와 시스템을 만들어준 선배님들, 진심 어린 조언과 의견을 공유해준 연구실 소속원분들 모두 감사드립니다.

또한 고민이 있거나 지칠 때면 진심으로 위해주고 격려해주었던 친구들 고맙습니다. 그리고 저에게 가장 큰 영향을 주며 도전을 망설이지 않게 용기를 주는, 언제나 옆에서 힘이 되어준 지혜에게도 고맙습니다.

마지막으로 현재의 제가 있을 수 있는 이유인 양가 할머니, 할아버지 어머니, 아버지 형 감사합니다. 연구에 전념할 수 있도록 아낌없이 지원해주시고 잘할 거라는 믿음으로 지켜봐 주신 덕분에 학위과정을 잘 마칠 수 있었습니다. 은혜롭고 무한한 사랑에 항상 감사드리며 자랑스러운 아들이 되겠습니다. 그리고 소중한 내 삶의 멘토이며 동생을 끔찍하게 아끼는 진관이 형에게도 감사의 말을 전합니다.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Evaluation of tumor acceleration by HLA-I zygosity and identification of the immune characteristics in clear cell renal cell carcinoma

BeumJin Park

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor Sangwoo Kim)

Clear cell renal cell carcinoma (ccRCC) is a common histological subtype of renal cancer with a 107% increase in incidence over about 20 years and has distinct immunogenic characteristics. CcRCC has been reported to have an immune characteristic contrary to the conventional notion, such as short survival despite abundance of tumor-infiltrating lymphocytes. In tumor immunity, the HLA-I molecule leads to tumor suppression, which enables CD8+ T cells to recognize tumors by presenting neoantigen, a peptide containing mutations in tumor cells, on the cell surface. We investigated tumor acceleration by HLA-I zygosity through tumor occurrence analysis in clear cell renal cell carcinoma as well as pan-cancer. To evaluate the effectiveness of immune surveillance by T cells, we selected early tumors close to the time point of immune evasion, excluding tumor samples containing pathogenic factors that influence tumor development. The zygosity of HLA-I was classified into homozygous group and heterozygous group based on the

heterozygosity of all three classical HLA-I genes. The acceleration was assessed using an accelerated failure time (AFT) model by HLA zygosity. As a result, it was confirmed that heterozygous HLA-I delayed the tumor development in pan-cancer except ccRCC. In contrast, ccRCC was found to occur earlier in heterozygous HLA-I patients than homozygous HLA-I patients. To determine the cause of ccRCC acceleration in heterozygous HLA-I, we classified allele loss in major tumor suppressor genes, VHL and PBRM1. Tumor acceleration of heterozygous HLA-I was found to occur in tumors containing biallelic loss of VHL, according to the theory of the secondary hit hypothesis that the loss of both alleles results in a phenotypic change. The association of heterozygous HLA-I and VHL biallelic loss on tumor acceleration was validated in an independent ICGC clear cell renal cell carcinoma cohort.

Key words: Tumor acceleration, HLA, Pan-cancer analysis, clear cell renal cell carcinoma, AFT regression, Bioinformatics

Evaluation of tumor acceleration by HLA-I zygosity and identification of the immune characteristics in clear cell renal cell carcinoma

BeumJin Park

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor Sangwoo Kim)

I. INTRODUCTION

Cancer is a complex disease in which transformed cells proliferate uncontrolled, causing many deaths worldwide[1]. Advances in next-generation sequencing (NGS) technology and the success of cancer immunotherapy have focused more attention on tumor immunity. Although the understanding of tumor immunity has increased, many unclear mechanisms for tumors still exist. In particular, distinct immunogenic characteristic has been reported in clear cell renal cell carcinoma (ccRCC).

CcRCC is the most common subtype of kidney cancer with a 107% increase in incidence over about 20 years[2]. ccRCC is a hyper-vascular tumor due to dysregulation of HIF protein[3,4]. In 18 TCGA tumor types, ccRCC has a moderate

mutation burden compared to other tumors, but has the highest expression of the cytotoxic T cell genes GZMA and PRF1[5]. Reduced Human leukocyte antigen (HLA) expression is suggested as a major cancer immune evasion feature[6,7], but higher HLA expression has been determined in ccRCC compared to matched normal samples[8,9]. In addition, the abundance of total tumor-infiltrating lymphocytes (TIL) in ccRCC is associated with short survival[10]. However, the mechanism by which ccRCC progresses even with evidence of an active tumor immune response is still unclear.

Recently, evolutionary dynamics, such as selection of tumor cells by the immune system, have been studied to suppress tumors. Theory of immunoediting describes the process by which transformed cells become cancer through immune selection[11,12]. Diverse immune cells are involved in immunoediting process throughout elimination, equilibrium and escape[13]. The cancer immunoediting consists of three steps of elimination, equilibrium and escape, and T cell responses play an important role in tumor suppression at all steps. In particular, among the T cell subtypes, CD8 + T cell is adaptive immune cell that have efficient anti-tumor responses by recognizing tumor-specific antigen, neoantigen. HLA molecule is important for CD8+ T cells to recognize neoantigens.

HLA molecule presents neoantigens to the cell surface according to affinity in new intracellular peptides which are containing amino acid altered by nonsynonymous mutation in tumors. Each type of HLA molecule has a different repertoire of peptide ligands[14]. Since each type of HLA molecule has a different repertoire of peptide ligands, negative selection that induce tumor cell killing by detecting randomly generated mutations is highly influenced by HLA[15]. Recently, it was reported that the HLA-I genotype is related to restriction of specific oncogenic mutations[16]. Across tumor types, lower mutation coverage group was diagnosed earlier than the higher group and vice versa. The diagnosis age showed a significant correlation with

the HLA-I coverage of the driver mutation, and the same aspect was also shown in the number of homozygous HLA-I genes[17]. These studies suggest that homozygous HLA-I is more disadvantageous for tumor immunosurveillance than heterozygous HLA-I. However, the association between age at diagnosis of ccRCC and HLA-I mutation coverage was not been assessed. Therefore, it was needed to investigate the immunological characteristics and acceleration of tumor development by HLA-I in ccRCC.

This study analyzed tumor resistance by controlling pathogenic germline mutations and viral infections that influence accelerate tumor formation. In addition, patients diagnosed at the time closest to immune escape were selected using clinical information and as patients corresponding to the onset time. For statistical analysis, more optimized accelerated failure time (AFT) model was applied to the onset comparison. Next, to investigate the cause of tumor acceleration by HLA-I zygosity in ccRCC, we classified allelic loss as a major tumor suppressor gene (TSG) and reported the distribution of loss status.

Our results report that high coverage HLA-I significantly accelerated tumor development than low coverage HLA-I in ccRCC. Contrary to aspect of ccRCC, high coverage HLA-I in other tumors was significant in tumor delay with the same results as in previous studies. Notably, it was confirmed that high coverage HLA-I is associated with tumor acceleration in patients with VHL bi-allele loss. These results were validated in an independent ICGC cohort.

II. MATERIALS AND METHODS

1. Acquisition of datasets

In this study, TCGA clinical information (TCGA-CDR) was downloaded from PanCanAtlas (https://gdc.cancer.gov/about-data/publications/pancanatlas). Tumor stage, age at diagnosis, gender, and race were used. In addition, for information on prior malignancy, clinical data provided by the GDC data portal was used. Most cancer types were classified by the AJCC pathologic staging system. Cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), Ovarian serous cystadenocarcinoma (OV), Uterine Corpus Endometrial Carcinoma (UCEC) and Uterine Carcinosarcoma (UCS) were staged by International Federation of Gynecology and Obstetrics (FIGO) staging system. For Lymphoid Neoplasm Diffuse Large B-cell Lymphoma (DLBC) and Thymoma (THYM), Ann Arbor and Masaoka staging systems were adopted, respectively. It also annotated germline mutations and viral infections that are considered pathogens.

In a recent study of 10,389 TCGA germline mutations, pathogenicity germline mutations that induce tumors were identified, and a list of pathogenicity annotation was available[18]. For viral infections, the TCGA Pan-cancer immune landscape study used RNA-seq to calculate the normalized read score of the viral sequence and determine the infection of HPV, EBV, HBV[19]. The data for calling somatic mutations using whole exome sequencing data[20] was also available in PanCanAtlas.

For validation, ICGC RNA-seq data was downloaded from Cancer Genome Collaboratory. Clinical information is available on the ICGC Data Portal. The histological type of RECA-EU was provided in clinical information data and

used to determine ccRCC. In addition, the result of analyzing the driver mutation of ccRCC could be obtained through DCC Data Releases.

2. HLA-I genotyping

HLA-I genotyping algorithms OptiType[21] and Polysolver[22] based on sequencing data were used for HLA-I allele typing in TCGA patients. Both tools outperformed in predicting HLA-I alleles. OptiType uses RNA-seq data, whole-exome sequencing (WES) and whole genome sequencing (WGS). The tool also shows high accuracy in case of using RNA-seq data. Polysolver uses WES data, and the homozygosity success rate was higher than that of OptiType as a benchmark result[22]. HLA-I typing results of both tools were available from PanCanAtlas supplementary files[19,23]. As the results were put together, 6,456 patients with predictive outcomes in both studies were 93.4% homozygosity. In shared patients, the results were determined by inference results of Polysolver, and for non-shared results, each tool result was used. The HLA-I allele in 11,167 TCGA patients was determined by the results of Polysolver and OptiType in 7,365 (66%) and 3,050 (27%) patients, respectively.

HLA-I allele types were determined using WGS and RNA-seq data from ICGC ccRCC patients (n = 66). For the memory efficiency of the prediction tool, pre-processing was performed in consideration of the highly polymorphic characteristic of the HLA-I gene located at chr6. RNA seq data (n=50) was converted into fastq format by extracting unmapped reads and mapped reads to chr6 region from bam file. RNA-seq data was predicted by OptiType, and genotyping was successful in 35 patients. The remaining RNA-seq data of 15 patients was typed by arcasHLA[24] For ccRCC patients without RNA data (n =

16), HLA Class I results predicted with ALPHLARD[25] using WGS in DCC data release were used (https://dcc.icgc.org/releases). The HLA genotype was unified as a 4-digit nomenclature spanning the allele group and specific HLA protein field.

3. Definition cohort at the time of onset and selection analysis dataset

The tumor onset cohort was defined using clinical information. To compare the difference in the period of immune surveillance, not only patients diagnosed at the time point closest to the point when immune escape occurred, but also samples containing pathogenic factors other than somatic mutations were excluded. Inferring the degree of tumor progression by the size of the tumor, the initial tumors, stage 0 and 1, were diagnosed closest to the escape point. In addition, factors affecting the probability and duration of immune selection occurring through elimination were considered for exclusion: prior malignancy, pathogenic mutation, no viral infection. Then, tumor types with a homozygous HLA-I group size less than 10 were excluded to statistically compare the HLA-I zygosity effects for each type. Finally, 11 tumor types were selected, including 1,573 patients (Table 1).

Table 1. Number of samples according to HLA-I zygosity for individual tumor types in the analyzed dataset

HLA-I zygosity[1]	THCA	LUAD	LUSC	KIRC	BRCA	KIRP	LIHC	TGCT	UCEC	SKCM	STAD	Total
Hetero zygous	203	162	148	142	117	107	89	72	72	52	24	1,188
Homo zygous	61	54	48	59	42	30	27	28	14	11	11	385

[1]Classification according to HLA-I zygosity: homozygous in at least one gene and heterozygous in all genes, respectively.

4. Examination of tumor acceleration according to HLA-I zygosity

Tumor acceleration was examined in the 1,573 patients, classified into homozygous and heterozygous groups. HLA-I zygosity was classified into at least one homozygous and all heterozygous in the three HLA class I genes (homozygous and heterozygous groups, respectively). First, the AFT model was used to investigate the tumor accelerated association between homozygous and heterozygous groups in the entire dataset. In addition, the relationship between HLA-I zygosity and onset acceleration was also confirmed in the dataset excluding ccRCC. Then, the effect of HLA-I zygosity in individual tumors was assessed by considering the heterogeneous tumor. For tumors with significant tumor incidence acceleration, tumor driver events were investigated to elucidate the cause.

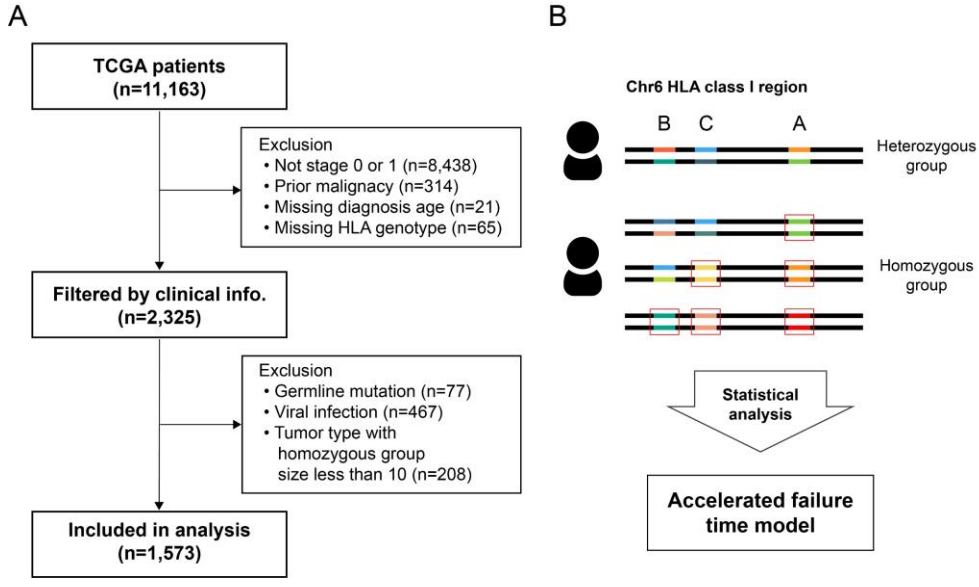

Figure 1. Overall workflow of evaluating the association between HLA-I zygosity and tumor acceleration. (A) The diagram shows the selection criteria by clinical information and annotated pathogenic factors. (B) The figure shows the analysis of tumor acceleration according to the HLA-I zygosity classification criteria.

5. Classification of somatic mutation and inactivated driver gene

The major driver genes of ccRCC were TSG, VHL and PBRM1. Notably, since VHL was found to follow the two-hit hypothesis[26], inactivation by alleles was classified to assess the association between driver mutations and HLA-I in tumor acceleration.

From the MC3 variant file, the somatic mutation call result of the TCGA dataset was obtained. For somatic mutation, the following filters were applied: Somatic mutation were extracted from the MC3 file (mc3.v0.2.8.CONTROLLED.maf) with the following filters: variant passes the calling filter, allele frequency > 0.1, read depth ≥ 10, nonsynonymous mutation (missense mutation, frame shift insertion, frame shift deletion, in frame insertion, in frame deletion, nonsense mutation, nonstop mutation, splice site). The copy count loss was inferred from the results of the copy count analysis for each sample provided by Firehose, using a threshold of -0.3 for the log2 ratio. In patients with copy number loss and mutation, it was determined that both alleles were inactivated.

The allelic inactivation of ccRCC of ICGC cohort was analyzed using the results of driver mutation provided by DCC Data Releases. For each sample driver gene, somatic mutation, copy number alteration (CNA) and structural variation (SV) were analyzed. Patients with two of the three events were determined to be inactive on both alleles.

6. Statistical Analysis

The difference of onset age was assessed using the Wilcoxon rank-sum test. Accelerated failure time (AFT) model was used to examine the Failure Rate (FR) with p-values and the 95% confidence interval. Distributions for the AFT model were chosen using the Akaike Information Criteria (AIC). The cumulative onset plot was produced through the log-lank test and Kaplan-Meier (K-M) estimator as well as the two groups were statistically compared. Comparison of tumor onset points in multiple groups classified by allelic loss was tested with Kruskal Wallis. Fisher's Exact Test was used to evaluate the categorical association of the significance of allelic loss status according to HLA-I zygosity. R software (version 3.6.3) was used for all statistical analysis. All statistical analysis results were described to be significant when the p value was less than 0.05, and statistical evaluation was performed two-sided.

III. RESULTS

1. Evaluation of tumor acceleration by HLA-I zygosity

A. Individual tumor types

In a total of 11 tumors, the AFT model was applied to measure tumor acceleration based on HLA-I zygosity. To measure the failure rate of HLA-I zygosity in individual tumors, the regression was calculated by adjusting the gender and race variables (Table 1).

Across tumor types, STAD showed the most significantly extreme failure rate as a result of AFT regression. The mean age onset of the two STAD groups classified by HLA-I zygosity was 66.54 in the homozygous group and 76.22 in the heterozygous group, which differed by more than 9 years. However, in STAD, the difference in onset age could have been overestimated due to the small sample size of 35. The failure rate was not significant in 7 tumor types, but the HLA-I homozygous group developed tumors earlier than the HLA-I heterozygous group. For tumors with failures lower than 1, the p value was calculated using the log rank test (Figure 2).

ccRCC showed significant acceleration and was occurred on average 4 years later when HLA-I zygosity was heterozygosity. The failure rate of ccRCC was 1.081 (95% confidence interval [CI]: 1.02,1.15), and the mean onset ages of the HLA-I heterozygous group and the homozygous group were 56.77 and 52.51, respectively. In addition to ccRCC, tumor incidence between HLA-I zygosity was compared with the log rank test for tumors with failure rates higher than 1 (Figure 3).

Table 2. Summary of the failure rate for HLA-I zygosity by tumor types

Tumor types	n	Failure Rate (95% CI)	*p* value	Heterozygous group age	Homozygous group age	Difference of age
ccRCC	201	1.081 (1.015,1.151)	0.015	52.51	56.77	-4.26
TGCT	100	1.037 (0.924,1.164)	0.540	29.17	30.24	-1.07
UCEC	86	1.018 (0.943,1.100)	0.645	65.04	66.23	-1.19
BRCA	159	0.986 (0.925,1.052)	0.680	56.40	55.64	0.76
SKCM	63	0.985 (0.817,1.188)	0.874	63.19	62.24	0.95
LUAD	216	0.982 (0.939,1.026)	0.414	61.98	60.85	1.13
LUSC	196	0.981 (0.943,1.020)	0.330	67.56	66.25	1.31
KIRP	137	0.974 (0.906,1.048)	0.485	62.13	60.54	1.59
THCA	264	0.965 (0.876,1.062)	0.464	34.28	33.07	1.21
LIHC	116	0.965 (0.896,1.040)	0.353	61.98	59.83	2.15
STAD	35	0.873 (0.791,0.963)	0.007	76.22	66.54	9.68

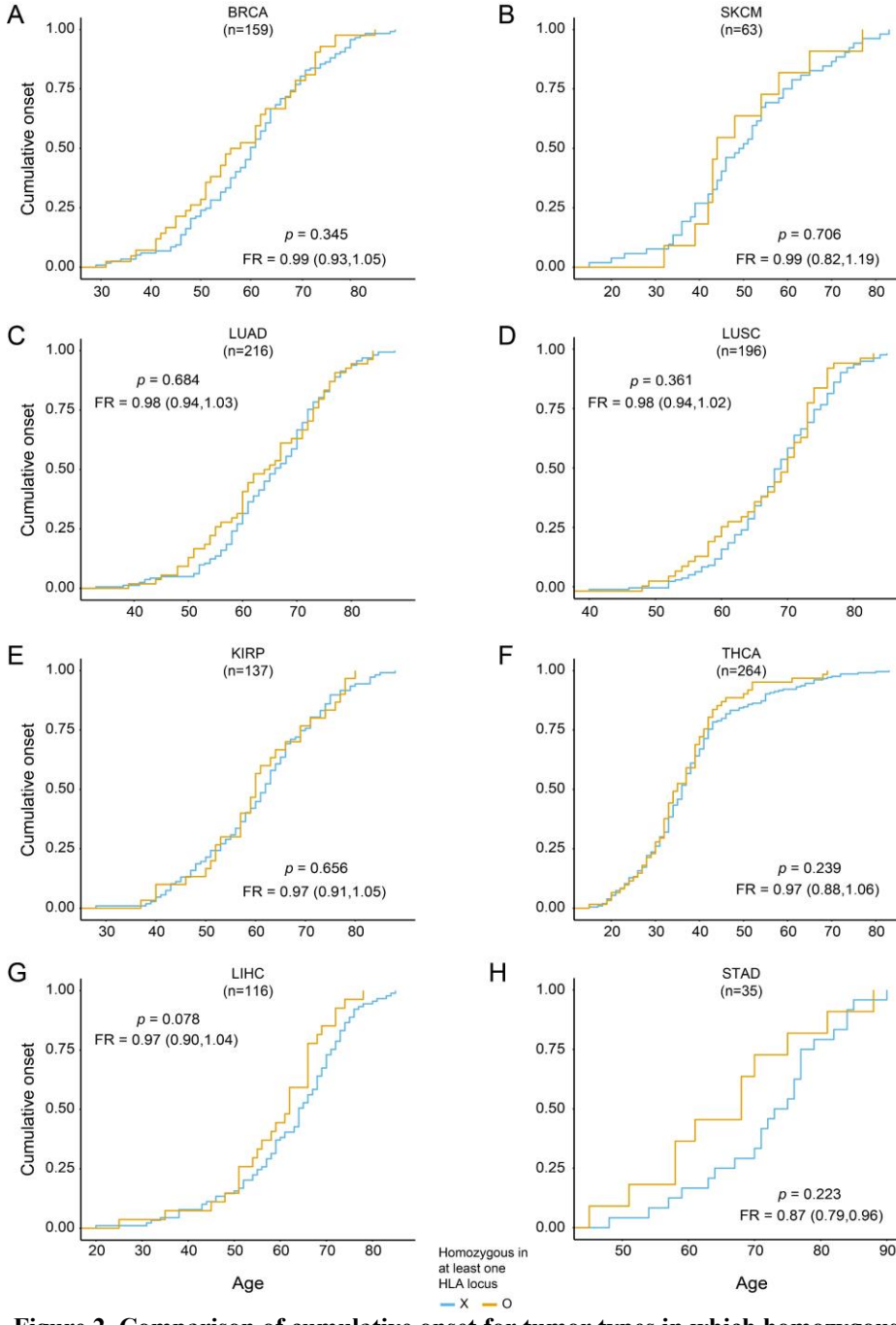

Figure 2. Comparison of cumulative onset for tumor types in which homozygous HLA-I accelerates tumors.

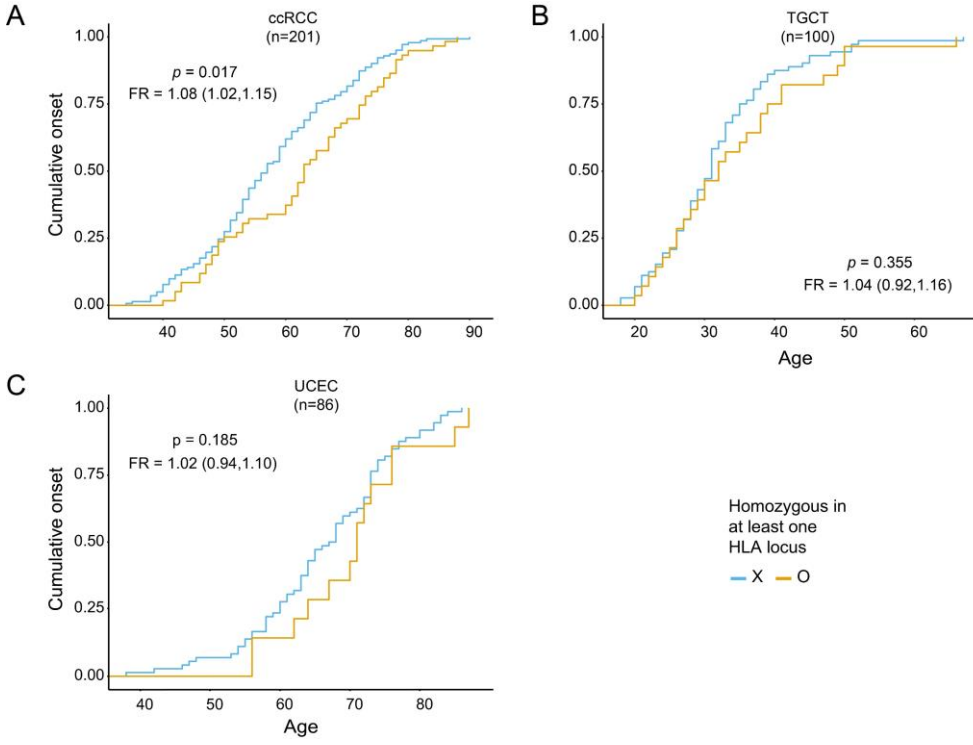

Figure 3. Comparison of cumulative onset for tumor types in which heterozygous HLA -I accelerates tumors.

B. Multiple tumor types

The failure rates were examined by AFT regression, which adjusted tumor type, gender and race with all tumor types as inputs (Table 3). HLA-I zygosity was not significant, but race showed a significant association with tumor acceleration. For race, most of the patients were white due to the characteristics of TCGA data, and non-white races were clustered for statistical analysis. White showed significantly later tumor development than non-white (p = 0.010, FR = 1.04, 95% CI 1.01 to 1.07).

Considering the significant association of heterozygous HLA-I with ccRCC acceleration, consistency AFT regression analysis was performed using data excluding ccRCC (Table 4). Race also showed a significant association with tumor acceleration. In the review of cancer susceptibility according to race, genetic diversity and evolutionary pressure due to infectious disease were suggested as causes. A review of cancer susceptibility by race suggested that blacks had an increased risk of malignancies than whites due to genetic diversity and evolutionary pressure by infectious disease[27]. The homozygous HLA-I group developed tumors earlier than the heterozygous group in tumors excluding ccRCC (p = 0.042, FR = 0.97, 95% CI 0.94 to 0.99). These results indicate that heterozygous HLA-I has the advantage of suppression in tumor, but in ccRCC this trend is reversed. In other words, it provides evidence for the possible effects of T cells in the tumorigenic process of ccRCC.

We comprehensively show the comparison of onset age and FR in the overall TCGA analysis (Figure 4).

Table 3. Multivariable AFT regression on total dataset (n = 1573)

	Failure Rate (95% CI)	*p* value
(Intercept)	57.222 (54.871,59.674)	< 0.001
Heterozygous HLA-I (reference)		
Homozygous HLA-I	0.988 (0.963,1.014)	0.359
BRCA (reference)		
KIRC	1.003 (0.956,1.053)	0.891
KIRP	1.046 (0.992,1.104)	0.099
LIHC	1.081 (1.023,1.143)	0.006
LUAD	1.112 (1.062,1.164)	< 0.001
LUSC	1.159 (1.103,1.218)	< 0.001
SKCM	0.834 (0.777,0.895)	< 0.001
STAD	1.204 (1.106,1.310)	< 0.001
TGCT	0.533 (0.501,0.568)	< 0.001
THCA	0.608 (0.581,0.638)	< 0.001
UCEC	1.138 (1.075,1.204)	< 0.001
Female(reference)		
MALE	0.977 (0.951,1.004)	0.095
Race cluster: Non-white(reference)		
Race cluster: white	1.039 (1.009,1.070)	0.010

Table 4. Multivariate AFT regression on datasets excluding ccRCC (n = 1372)

	Failure Rate (95% CI)	*p* value
(Intercept)	57.419 (54.374,60.635)	< 0.001
Heterozygous HLA-I (reference)		
Homozygous HLA-I	0.971 (0.944,0.999)	0.042
BRCA (reference)		
KIRP	1.036 (0.981,1.094)	0.198
LIHC	1.072 (1.014,1.134)	0.015
LUAD	1.105 (1.056,1.157)	< 0.001
LUSC	1.148 (1.092,1.208)	< 0.001
SKCM	0.825 (0.768,0.886)	< 0.001
STAD	1.195 (1.098,1.300)	< 0.001
TGCT	0.526 (0.494,0.561)	< 0.001
THCA	0.607 (0.579,0.636)	< 0.001
UCEC	1.137 (1.074,1.203)	< 0.001
Female(reference)		
MALE	0.990 (0.960,1.020)	0.496
Race cluster: Non-white(reference)		
Race cluster: white	1.040 (1.008,1.074)	0.014

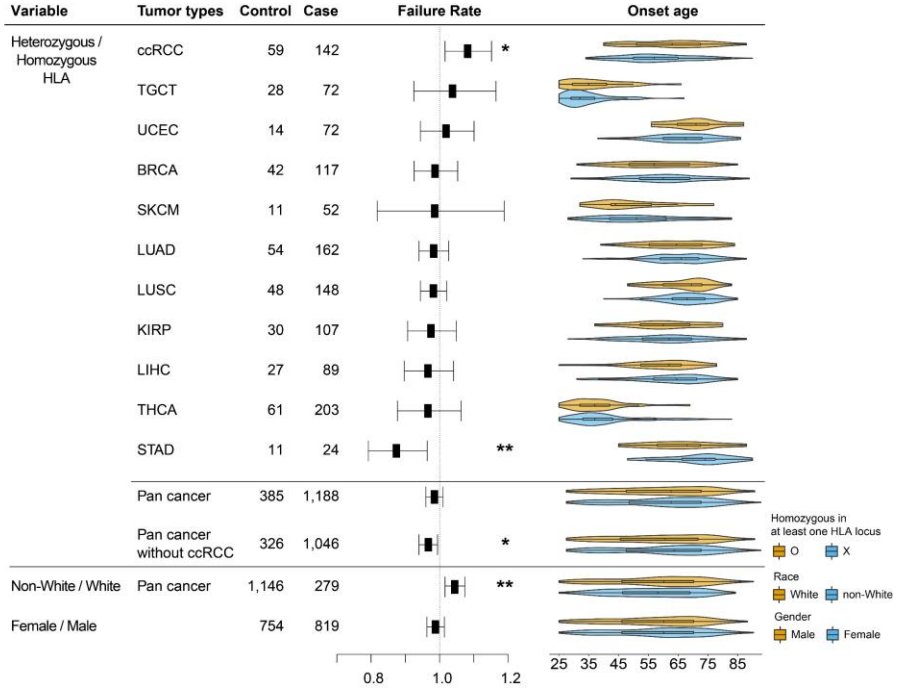

Figure 4. FR and onset age distribution by analysis dataset and variable. The top of the figure shows the acceleration for each tumor type and multiple tumor types by HLA-I zygosity. The lower part of the figure shows FR by race and gender in Pan-cancer. On the right side of the figure, onset age of the group classified by each variable is shown.

2. Evaluation of tumor acceleration by driver events of ccRCC

To elucidate the effect of tumor acceleration in the mutation high coverage HLA-I of ccRCC, we focused on major driver genes. VHL is the most frequently mutated TSG in ccRCC and a recognizing substrate for the ubiquitin E3 ligase complex targeting HIF. The second most common mutant gene is PBRM1, and the PBRM1 molecule is a subunit of the SWI/SNF complex that remodels chromatin. Inactivation of VHL and PBRM1 was suggested as a major tumorigenic process for ccRCC through activation of the mTORC1 pathway[28].

The mutation landscape was generated using 150 tumors in the presence of a matched normal sample among 201 ccRCCs (Figure 5A). Mutations in VHL (47%) and PBRM1 (43%) were observed in most of the early ccRCCs and a loss of the 3p (79%) region was observed across VHL and PBRM1. To investigate the effects of VHL and PBRM1, both driver genes were classified into five categories based on wild type, monoallelic and biallelic loss: 19 wild type cohorts, 32 monoallelic loss of PBRM1, 36 monoallelic loss of VHL, 63 biallelic loss of PBRM1, 65 biallelic loss of VHL (Figure 5B).

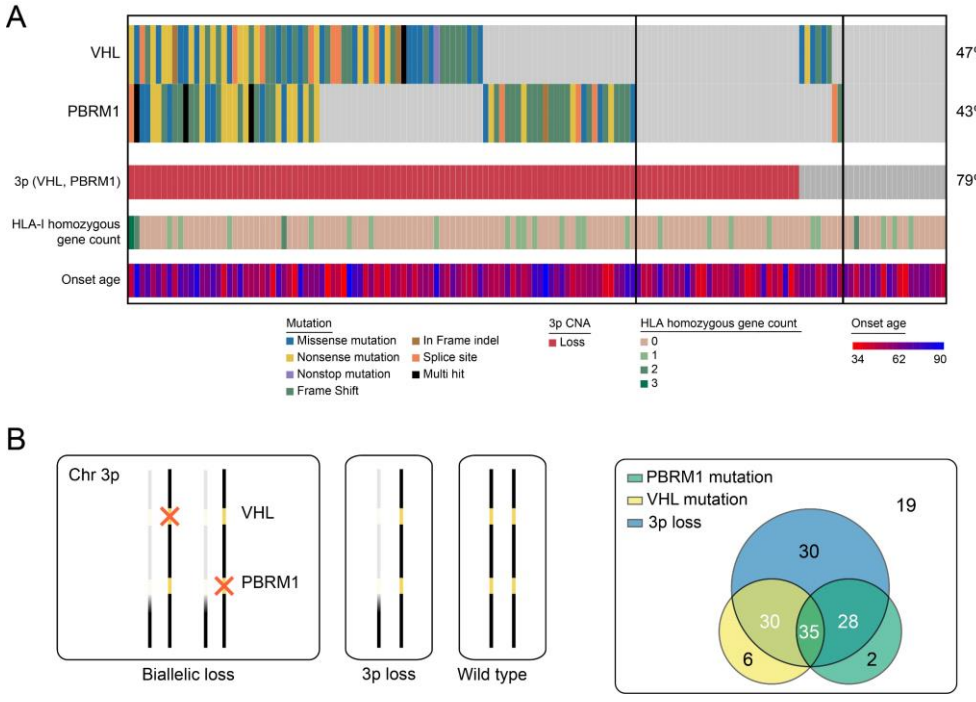

Figure 5. Driver events landscape and HLA-I zygosity status of ccRCC sample of 201 onset cohort. (A) The upper part shows the somatic mutations in the VHL and PBRM1 genes, which are the driver genes of ccRCC. The copy number loss for the region spanning the VHL and PBRM1 of the chromosome 3p arm is shown. The lower part shows the count status of the HLA-I zygosity gene and onset age. (B) A schematic of the allelic loss classification of VHL and PBRM1, and a Venn diagram of the sample distribution is shown.

3. Association of VHL inactivation and HLA-I zygosity with tumor acceleration

Next, the age of onset classified by the HLA-1 zygosity was compared by allelic loss in VHL and PBRM1 (Figure 6A). In a cohort of patients with VHL biallelic loss, HLA-I zygosity was significantly associated with onset age, but not PBRM1 and single allele loss. As a result of comparing the VHL and PBRM1 exclusive loss, onset age between the two groups by HLA-I zygosity was significantly different in cohort of the exclusive VHL biallelic loss (Figure 6B). In addition, we compared the two HLA-I groups for log rank tests in the cumulative onset Kaplan-Meier (Figure 7). In patients with VHL biallelic loss, HLA-I zygosity was significantly different in cumulative onset (p = 0.003), and failure rate was calculated by AFT regression using race and sex as adjusting variables (FR = 1.24, 95% CI 1.08 ~ 1.43).

The association between each allelic loss status and HLA-I zygosity was analyzed by Fisher's Exact Test and was not significant (p = 0.659). Kruskal Wallis was tested to compare the difference in tumor onset age for each allelic loss, and there was no significant (median age = 60.0, 53.5, 54.0, 60.0, 59.0: WT, $PBRM1_{WT/-}$, $VHL_{WT/-}$, $PBRM1_{-/-}$, $VHL_{-/-}$, respectively, p = 0.169).

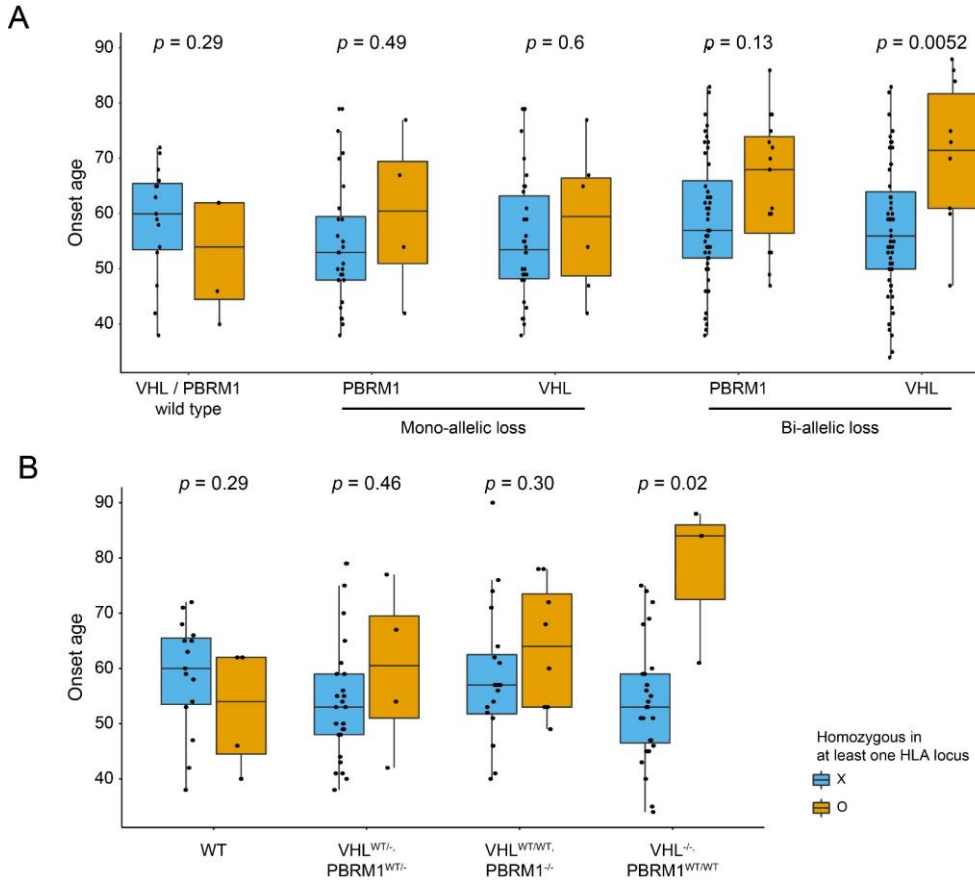

Figure 6. Comparison of tumor onset age by HLA-I zygosity and allelic loss of VHL and PBRM1 in TCGA ccRCC. (A) The figure is a boxplot comparing the onset age of HLA-I zygosity by allelic loss of PBRM1 and VHL. (B) The boxplot shows the onset age of HLA-I zygosity for the case of mutually exclusive allelic loss of VHL and PBRM1. WT = wild type, - = monoallelic loss.

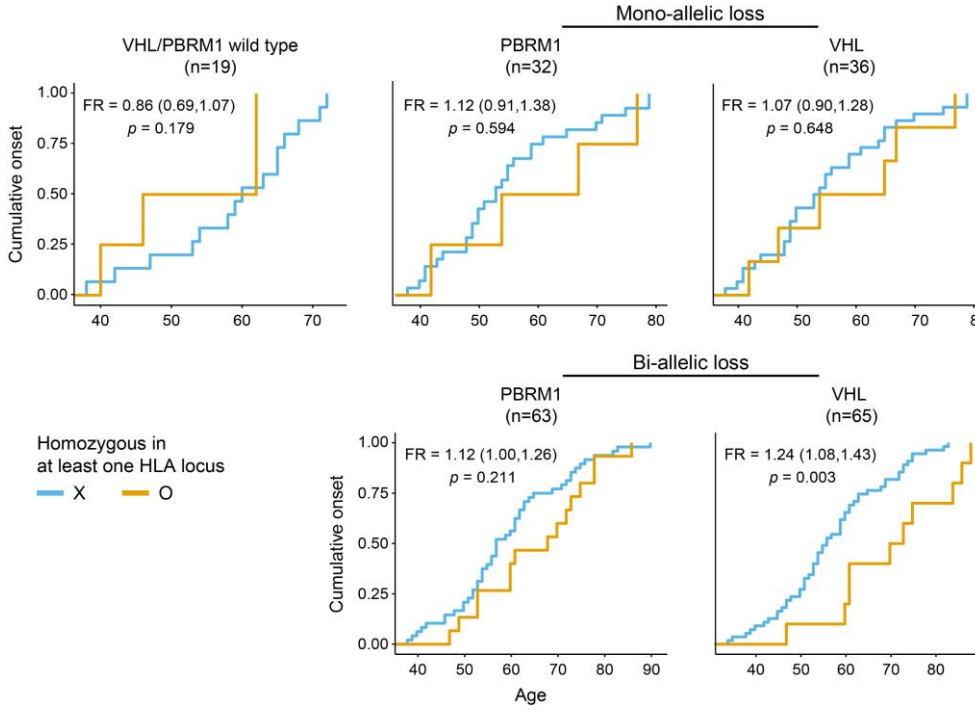

Figure 7. Comparison of cumulative onset differences in HLA-I zygosity by the loss of each allele of VHL and PBRM1 in TCGA ccRCC. Kaplan-Meier plots were generated by comparing the cumulative incidence differences of HLA-I zygosity through log rank and including the FRs obtained using AFT regression.

4. Validation in independent dataset

Independent ICGC ccRCC were used to validate tumor acceleration of heterozygous HLA-I in VHL biallelic loss. In the VHL biallelic loss cohort, heterozygous HLA-I occurred tumors earlier than homozygous HLA-I (p = 0.058, Figure 8A). The comparison of cumulative onset by HLA-I zygosity was significant in VHL biallelic loss (p=0.037, log-rank test), and tumor acceleration was evaluated through gender-adjusted AFT regression (FR = 1.19, 95% CI 1.04

to 1.37, Figure 8B). Taken together, the tumor acceleration of heterozygous HLA-I in ccRCC is associated with VHL biallelic loss.

Loss of tumor suppressor gene VHL induces hypoxia through dysregulation of HIF-1α[29]. As a result, hypoxia-inducible genes are transcriptionally activated. Analysis of multi-region sequencing revealed early occurrence of 3p loss in ccRCC patients. In patients with biallelic loss of VHL, the second VHL inactivation is estimated to have occurred 15 to 30 years prior to cancer diagnosis[30]. These results suggest that the VHL loss influence during the long dormancy state. During immunosurveillance, heterozygous HLA-I leads to more diverse neoantigens to be presented than homozygous HLA-I, resulting in a more frequent T cell response per unit time in heterozygous HLA-I tumors. Consequently, this study suggests that in transformed cells containing VHL biallelic, CD8+ T cell responses influence the progression of ccRCC rather than tumor cell killing.

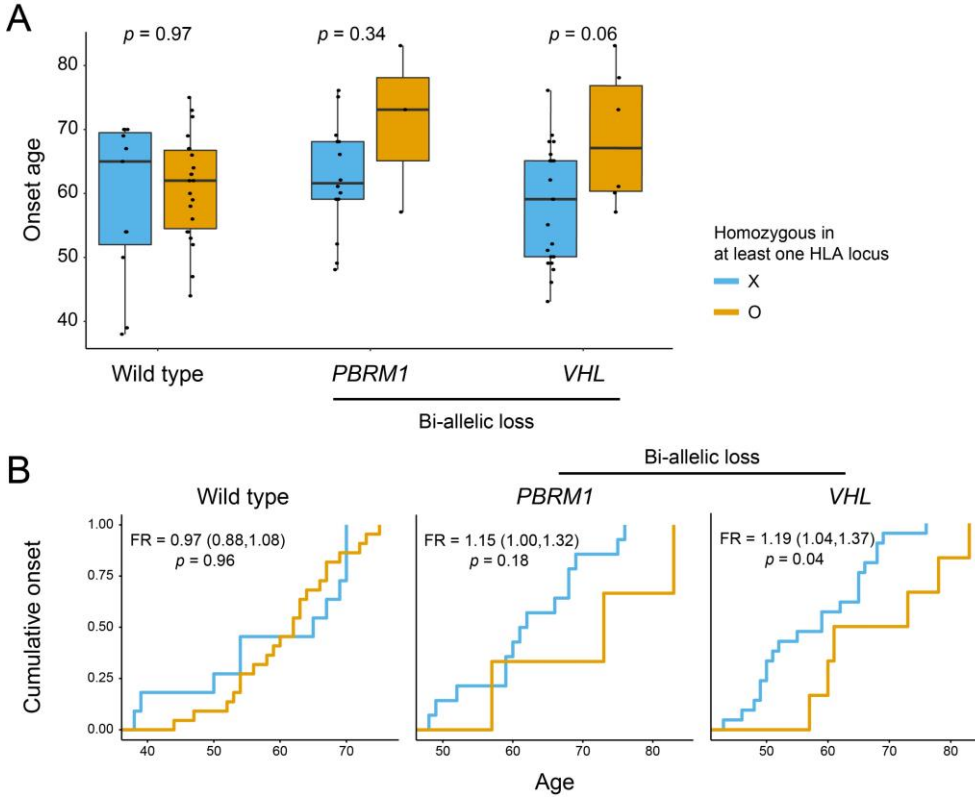

Figure 8. Validation of association between HLA-I zygosity and acceleration of tumorigenesis due to loss of each allele of VHL and PBRM1 in ICGC ccRCC. (A) The figure is a boxplot comparing the onset age of HLA-I zygosity by biallelic loss of PBRM1 and VHL. (B) Kaplan-Meier plots were generated by comparing the cumulative incidence differences of HLA-I zygosity through log rank and including the FRs obtained using AFT regression.

IV. DISCUSSION

AFT analysis for pan-cancers excluding ccRCC showed that tumors were delayed in the heterozygous group, which could present more diverse neoantigens. These results indicated that heterozygous HLA-I induces tumor cell death by increasing the likelihood of presenting antigens against random mutations. In contrast, tumor onset in ccRCC was accelerated in the heterozygous group. CcRCC with distinct immunological characteristics required investigation into the causes of these results.

VHL and PBRM1, the main TSGs of ccRCC, were selected and classified losses by allele. The tumor acceleration of heterozygous HLA-I was significantly correlated in the cohort with VHL biallelic loss. Inactivation of the VHL gene induce HIF1 dysregulation followed by the activation of STAT3, leading to the expression of angiogenesis-related genes[31]. VHL biallelic loss was found to be an early event long before ccRCC was diagnosed, indicating that transformed cells, including biallelic loss, become cancerous after a long dormancy state[30]. We suggest that the cause of ccRCC acceleration of heterozygous HLA-I is mTORC1 activation as a downstream event of VHL inactivation (Figure 9A). Among the products of the CD8+ T cell response, IFN-γ can influence chronic HIF-1 activation and accelerate tumors through the hypoxia pathway and mTORC1 activation. IFN-γ induces activation of immune cells through signal transduction, but has been reported to play a dual role in tumor suppression and progression during immune editing[32,33]. In ccRCC, IFN-γ activates STAT3 and upregulates HIF-1 expression, and STAT3 and HIF-1 interact with each other to amplify[34-36]. In addition, IFN-γ stimulates mTORC1 by activating the PI3K/AKT pathway[37]. T cell response leads to accelerate tumor by activating mTORC1 in VHL-inactivation transformed cells through paracrine of IFN-γ with frequent T cell responses (Figure 9B).

In the pan-cancer analysis, the filter criteria such as selecting tumors at the early stage made the sample size of each tumor type different. Statistical analysis in the sample size of 35 patients may have overestimate the degree of tumor acceleration, and further analysis is needed to evaluate aspects in various tumors using the data obtained.

Taken together, our results provide evidence for short survival in ccRCC patients despite abundant TIL and evidence for HLA-I expression in ccRCC than in matched normal samples. The distinct immunogenic characteristic of ccRCC require further study based on the presence of VHL biallelic loss. The mechanism by which the CD8+ T cell response accelerates tumors in VHL-inactivated transformed cells will enable understanding of heterogeneous clinical outcomes for immune checkpoint blockade therapy in ccRCC.

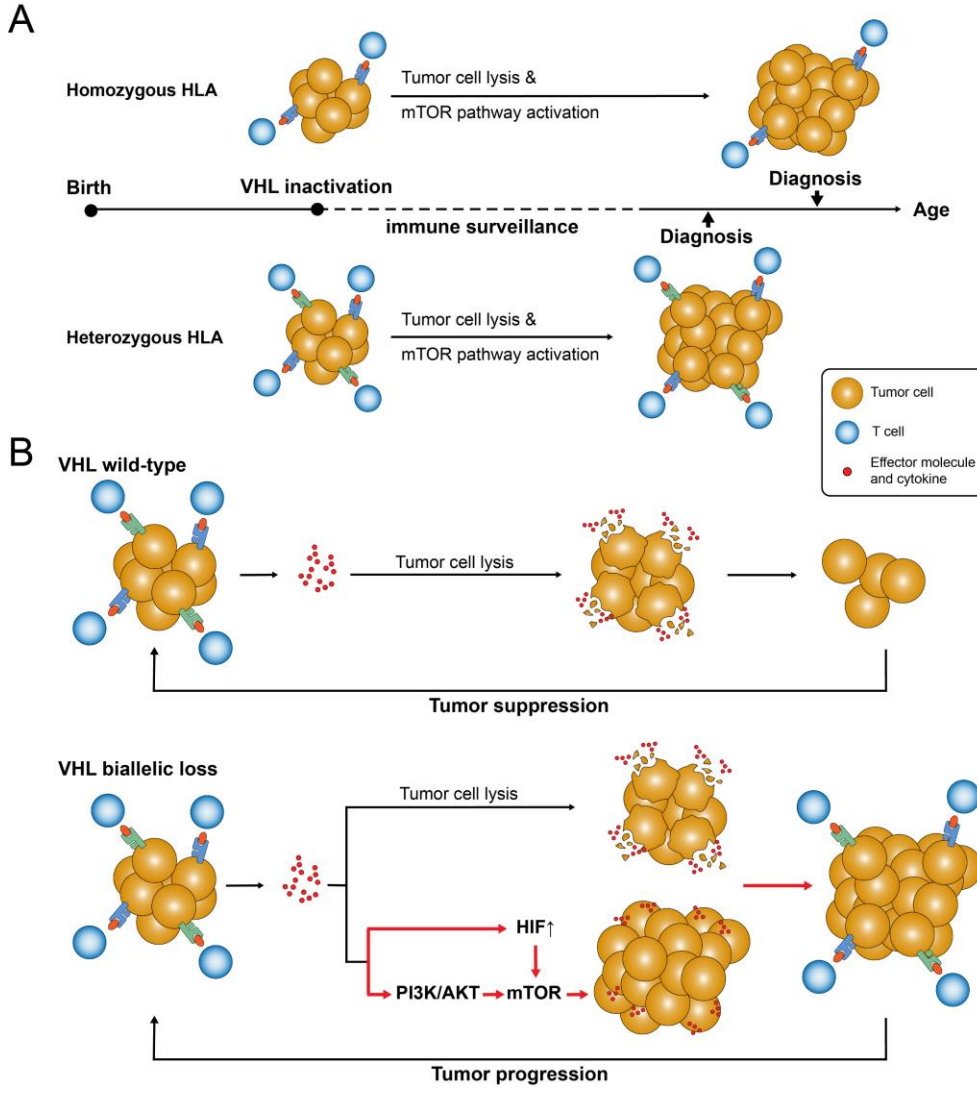

Figure 9. Hypothesis model for the relationship between VHL inactivation and T cell response. (A) The illustration shows that after VHL inactivation, heterozygous HLA accelerates tumors with more frequent T cell responses. (B) In transformed cells with VHL biallelic loss, T cell response induces tumor proliferation through mTOR activation rather than tumor cell lysis.

V. CONCLUSION

In this study, in order to compare the role of CD8+ and tumor suppression during immunoediting by HLA-I zygosity, patients close to the immune escape point were selected and tumor acceleration was evaluated using the AFT model. In the TCGA dataset excluding ccRCC, it was found that heterozygous HLA-I, which enable presenting neoantigens against a more diverse mutations, was earlier in tumor than homozygous HLA-I. In contrast, ccRCC had early tumor onset in heterozygous HLA-I.

The immunogenic characteristics of ccRCC using TSG were investigated and found that heterozygous HLA-I was significantly associated with tumor acceleration in patients with VHL biallelic loss. In an independent ICGC cohort, consistent results of the association between VHL biallelic loss and HLA zygosity were validated.

VHL inactivation is an early event in life, found as a clonal mutation in many ccRCC patients after prolonged dormancy and has a significant impact on the formation of ccRCC immunity. The findings that VHL inactivation influences tumor acceleration of CD8+ T cell responses is provided as evidence for uncovering ccRCC immune characteristics.

REFERENCES

1. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Pineros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer 2019;144:1941-53.
2. Global Burden of Disease Cancer C, Fitzmaurice C, Dicker D, Pain A, Hamavid H, Moradi-Lakeh M, et al. The Global Burden of Cancer 2013. JAMA Oncol 2015;1:505-27.
3. Hsieh JJ, Purdue MP, Signoretti S, Swanton C, Albiges L, Schmidinger M, et al. Renal cell carcinoma. Nat Rev Dis Primers 2017;3:17009.
4. Liu XD, Hoang A, Zhou L, Kalra S, Yetil A, Sun M, et al. Resistance to Antiangiogenic Therapy Is Associated with an Immunosuppressive Tumor Microenvironment in Metastatic Renal Cell Carcinoma. Cancer Immunol Res 2015;3:1017-29.
5. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell 2015;160:48-61.
6. McGranahan N, Rosenthal R, Hiley CT, Rowan AJ, Watkins TBK, Wilson GA, et al. Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. Cell 2017;171:1259-71 e11.
7. Kaneko K, Ishigami S, Kijima Y, Funasako Y, Hirata M, Okumura H, et al. Clinical implication of HLA class I expression in breast cancer. BMC Cancer 2011;11:454.
8. Stickel JS, Stickel N, Hennenlotter J, Klingel K, Stenzl A, Rammensee HG, et al. Quantification of HLA class I molecules on renal cell carcinoma using Edman degradation. BMC Urol 2011;11:1.
9. Saenz-Lopez P, Gouttefangeas C, Hennenlotter J, Concha A, Maleno I, Ruiz-Cabello F, et al. Higher HLA class I expression in renal cell carcinoma than in autologous normal tissue. Tissue Antigens 2010;75:110-8.
10. Danaher P, Warren S, Dennis L, D'Amico L, White A, Disis ML, et al. Gene expression markers of Tumor Infiltrating Leukocytes. J Immunother Cancer 2017;5:18.
11. Dunn GP, Old LJ, Schreiber RD. The immunobiology of cancer immunosurveillance and immunoediting. Immunity 2004;21:137-48.
12. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. Science 2011;331:1565-70.
13. O'Donnell JS, Teng MWL, Smyth MJ. Cancer immunoediting and resistance to T cell-based immunotherapy. Nat Rev Clin Oncol 2019;16:151-67.
14. Chowell D, Morris LGT, Grigg CM, Weber JK, Samstein RM, Makarov V, et al. Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. Science 2018;359:582-7.
15. Lakatos E, Williams MJ, Schenck RO, Cross WCH, Household J, Zapata L,

et al. Evolutionary dynamics of neoantigens in growing tumors. Nat Genet 2020.
16. Marty R, Kaabinejadian S, Rossell D, Slifker MJ, van de Haar J, Engin HB, et al. MHC-I Genotype Restricts the Oncogenic Mutational Landscape. Cell 2017;171:1272-83 e15.
17. Marty Pyke R, Thompson WK, Salem RM, Font-Burgada J, Zanetti M, Carter H. Evolutionary Pressure against MHC Class II Binding Cancer Mutations. Cell 2018;175:416-28 e13.
18. Huang KL, Mashl RJ, Wu Y, Ritter DI, Wang J, Oh C, et al. Pathogenic Germline Variants in 10,389 Adult Cancers. Cell 2018;173:355-70 e14.
19. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The Immune Landscape of Cancer. Immunity 2018;48:812-30 e14.
20. Ellrott K, Bailey MH, Saksena G, Covington KR, Kandoth C, Stewart C, et al. Scalable Open Science Approach for Mutation Calling of Tumor Exomes Using Multiple Genomic Pipelines. Cell Syst 2018;6:271-81 e7.
21. Szolek A, Schubert B, Mohr C, Sturm M, Feldhahn M, Kohlbacher O. OptiType: precision HLA typing from next-generation sequencing data. Bioinformatics 2014;30:3310-6.
22. Shukla SA, Rooney MS, Rajasagi M, Tiao G, Dixon PM, Lawrence MS, et al. Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. Nat Biotechnol 2015;33:1152-8.
23. Kahles A, Lehmann KV, Toussaint NC, Huser M, Stark SG, Sachsenberg T, et al. Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. Cancer Cell 2018;34:211-24 e6.
24. Orenbuch R, Filip I, Comito D, Shaman J, Pe'er I, Rabadan R. arcasHLA: high-resolution HLA typing from RNAseq. Bioinformatics 2020;36:33-40.
25. Hayashi S, Yamaguchi R, Mizuno S, Komura M, Miyano S, Nakagawa H, et al. ALPHLARD: a Bayesian method for analyzing HLA genes from whole genome sequence data. BMC Genomics 2018;19:790.
26. McNeill A, Rattenberry E, Barber R, Killick P, MacDonald F, Maher ER. Genotype-phenotype correlations in VHL exon deletions. Am J Med Genet A 2009;149A:2147-51.
27. Ozdemir BC, Dotto GP. Racial Differences in Cancer Susceptibility and Survival: More Than the Color of the Skin? Trends Cancer 2017;3:181-97.
28. Nargund AM, Pham CG, Dong Y, Wang PI, Osmangeyoglu HU, Xie Y, et al. The SWI/SNF Protein PBRM1 Restrains VHL-Loss-Driven Clear Cell Renal Cell Carcinoma. Cell Rep 2017;18:2893-906.
29. Semenza GL. HIF-1 mediates metabolic responses to intratumoral hypoxia and oncogenic mutations. J Clin Invest 2013;123:3664-71.
30. Mitchell TJ, Turajlic S, Rowan A, Nicol D, Farmery JHR, O'Brien T, et al. Timing the Landmark Events in the Evolution of Clear Cell Renal Cell Cancer: TRACERx Renal. Cell 2018;173:611-23 e17.
31. Jung JE, Lee HG, Cho IH, Chung DH, Yoon SH, Yang YM, et al. STAT3 is

a potential modulator of HIF-1-mediated VEGF expression in human renal carcinoma cells. FASEB J 2005;19:1296-8.

32. Schurch C, Riether C, Amrein MA, Ochsenbein AF. Cytotoxic T cells induce proliferation of chronic myeloid leukemia stem cells by secreting interferon-gamma. J Exp Med 2013;210:605-21.
33. Zou Q, Jin J, Xiao Y, Zhou X, Hu H, Cheng X, et al. T Cell Intrinsic USP15 Deficiency Promotes Excessive IFN-gamma Production and an Immunosuppressive Tumor Microenvironment in MCA-Induced Fibrosarcoma. Cell Rep 2015;13:2470-9.
34. Wen Z, Zhong Z, Darnell JE, Jr. Maximal activation of transcription by Stat1 and Stat3 requires both tyrosine and serine phosphorylation. Cell 1995;82:241-50.
35. Qing Y, Stark GR. Alternative activation of STAT1 and STAT3 in response to interferon-gamma. J Biol Chem 2004;279:41679-85.
36. Yeh YH, Hsiao HF, Yeh YC, Chen TW, Li TK. Inflammatory interferon activates HIF-1alpha-mediated epithelial-to-mesenchymal transition via PI3K/AKT/mTOR pathway. J Exp Clin Cancer Res 2018;37:70.
37. Lekmine F, Sassano A, Uddin S, Smith J, Majchrzak B, Brachmann SM, et al. Interferon-gamma engages the p70 S6 kinase to regulate phosphorylation of the 40S S6 ribosomal protein. Exp Cell Res 2004;295:173-82.

ABSTRACT (IN KOREAN)

HLA-I 접합성에 따른 종양 가속화 평가 및 투명 세포 신장 세포 암에서 면역 특징의 발견

< 지도교수 김상우 >

연세대학교 대학원 의과학과

박범진

투명 세포 신장 세포 암은 약 20년 동안 발병률이 107% 증가하는 신장암의 하위유형 중 가장 흔한 암이며 뚜렷한 면역원성 특징을 갖는다. 특히 투명 세포 신장 세포 암은 종양 침윤 T 림프구가 많이 존재할 때 짧은 생존을 보이는 등 기존의 개념과는 반대되는 면역 특징이 보고되어져 왔다. 종양 발생과정 중 면역반응을 설명하는 면역편집 이론은 최근 활발히 연구되고 있으며, 면역편집 동안 CD8+ T 세포는 종양세포를 인지하고 종양세포 특이적 사멸을 유발하는 반응을 한다. CD8+ T 세포가 종양세포를 인식하는 반응에는 HLA-I 분자가 중요한 작용을 하게 되는데, HLA-I는 종양 특이적 항원인 신항원(neoantigen)을 세포표면으로 표지하여 CD8+ T 세포가 non-self 항원인 neoantigen을 인지하여 반응을 유발할 수 있게 한다. 본 연구는 HLA-I 접합성에 의한 종양 발병가속도를

조사하기 위해 TCGA의 대규모의 암 데이터를 이용하여 투명 세포 신장 세포 암 뿐만 아니라 범암(pan-cancer) 단위의 분석을 진행하였다. 병원성 요인을 포함한 종양을 제외하여 T 세포에 의한 면역감시효과를 주요하게 평가하였으며 면역회피가 일어난 시점에서 가장 가까운 초기 단계의 종양을 선정하였다. HLA-I의 접합성은 HLA-I 유전자(HLA-A, B, C) 접합성에 따라 동형 접합, 이형 접합 HLA-I으로 분류되었고 HLA-I 접합성의 종양 가속도를 측정하기 위해 accelerated failure time 모델이 사용되었다. 그 결과 투명 세포 신장 세포 암을 제외한 pan-cancer에서 이형 접합 HLA-I은 종양 발병이 늦춰지는 것과 연관성이 있었다. 대조적으로 투명 세포 신장 세포 암에서는 이형 접합 HLA-I 환자에서 동형 접합 HLA-I 환자보다 일찍 종양이 발병하는 것이 밝혀졌다. 이형 접합 HLA-I에서 종양 가속화가 일어나는 원인을 규명하기 위해 투명 세포 신장 세포 암에서 가장 빈번하게 돌연변이가 발생하는 종양 억제 유전자 VHL과 PBRM1를 선별하였다. VHL 유전자는 두 대립 유전자가 모두 손실되면 표현형 변화가 생긴다는 two-hit hypothesis를 따르는 것이 보고되어 왔으며, 이형 접합 HLA-I은 VHL 이중 대립 유전자 손실을 포함하는 경우에서 유의한 종양 가속화를 나타냈다. 종양 가속화에 대한 이형 접합 HLA-I와 VHL 이중 대립 유전자 손실의 연관성은 독립적인 ICGC 투명 세포 신장 세포 암 집단에서 검증되었다.

핵심되는 말: 종양 가속화, 조직적합성항원, 투명 세포 신장 세포암, 대규모 암 분석

PUBLICATION LIST