Neoantigen presentation arises as a result of tumor-specific mutations and is a critical component of immune surveillance that can be abrogated by somatic LOH of the human leukocyte antigen class I (HLA-I) locus. To understand the role of HLA-I LOH in oncogenesis and treatment, we utilized a pan-cancer genomic dataset of 83,644 patient samples, a small subset of which had treatment outcomes with immune checkpoint inhibitors (ICI). HLA-I LOH was common (17%) and unexpectedly had a nonlinear relationship with tumor mutational burden (TMB). HLA-I LOH was frequent at intermediate TMB, yet prevalence decreased above 30 mutations/megabase, suggesting highly mutated tumors require alternate immune evasion mechanisms. In ICI-treated patients with non–squamous non–small cell lung cancer, HLA-I LOH was a significant negative predictor of overall survival. Survival prediction improved when combined with TMB, suggesting TMB with HLA-I LOH may better identify patients likely to benefit from ICIs.

SIGNIFICANCE: This work shows the pan-cancer landscape of HLA-I LOH, revealing an unexpected “Goldilocks” relationship between HLA-I LOH and TMB, and demonstrates HLA-I LOH as a significant negative predictor of outcomes after ICI treatment. These data informed a combined predictor of outcomes after ICI and have implications for tumor vaccine development.

INTRODUCTION

Immune checkpoint inhibitors (ICI) have revolutionized current treatments for patients with advanced cancer by reinvigorating one’s own T cell–mediated immune response (1, 2). CD8 T cells are thought to recognize tumor cells via the presentation of tumor-specific mutant peptides (neoantigens) on human leukocyte antigen class I (HLA-I)–encoded Major Histocompatibility Complex (MHC) class I proteins (3). This hypothesis is supported by the efficacy of ICIs in diseases with a high tumor mutational burden (TMB), such as mismatch repair–deficient tumors, and the potential pan-cancer utility of TMB as a biomarker of ICI response (1, 2, 4). Yet, in prospective trials focused on non–small cell lung cancer (NSCLC),...
Pan-Cancer Analysis of HLA LOH as a Method of Immune Evasion

Recent findings suggest that HLA-I LOH is a frequent phenomenon in NSCLC and a potential mechanism for immune evasion (1, 5). However, widespread analysis of HLA-I LOH across varying tumor types is lacking. Here, we show that somatic HLA-I LOH is a poor prognostic OS in ICI-treated patients with nonsquamous NSCLC, highlighting the need for development of superior immunotherapy biomarkers and a deeper understanding of the mechanisms behind immune evasion (1, 5).

RESULTS

To assess the landscape of allele-specific HLA-I LOH, we developed a pipeline for tumor-only next-generation sequencing (NGS) of tissue biopsies that can detect LOH as well as germline homozygosity at the HLA-I locus (HLA-A, HLA-B, and HLA-C). We applied this pipeline to a Foundation Medicine dataset of 83,664 unique patient samples. Consistent with previous reports (6), nonsquamous NSCLCs with HLA-I LOH, classified as somatic LOH of at least one HLA-I allele, were enriched for samples with TMB≥10 mutations per megabase (mut/Mb; refs. 1, 7), PD-L1≥1% tumor proportion score, as well as APOBEC and tobacco mutational signatures (Fig. 1A). A subset of these data was used to investigate the impact of HLA-I LOH on ICI treatment by considering both TMB and HLA-I LOH biomarkers may better identify patients likely to benefit from ICI treatment.

TMB has failed to sufficiently predict a benefit in patients’ overall survival (OS), highlighting the need for development of superior immunotherapy biomarkers and a deeper understanding of the mechanisms behind immune evasion (1, 5).

Recently, LOH at the HLA-I locus has been described as a frequent phenomenon in NSCLC as well as nonsquamous NSCLCs with HLA-I LOH (blue, n = 10,471) and without evidence of HLA-I LOH (red, n = 2,769) and without evidence of HLA-I LOH (blue, n = 10,471). TMB high: ≥10 muts/Mb, TMB low: <10 muts/Mb).  

Figure 1. Somatic HLA-I LOH and TMB are independent and significant predictors of patient OS in ICI-treated NSCLC. A. Enrichment of sample attributes, including genomic driver alterations, in nonsquamous NSCLC samples with HLA-I LOH (red, n = 2,769) and without evidence of HLA-I LOH (blue, n = 10,471). TMB high: ≥10 muts/Mb, PD-L1≥1% tumor proportion score. Statistics conducted by Fisher exact test, and only significant (P < 0.01) associations are colored and labeled. The prevalence of each significant attribute within either the HLA-I–intact or HLA-I LOH cohort is depicted by the data point size. B, OS for patients with nonsquamous NSCLC in the real-world clinical-genomic cohort from start of second-line ICI monotherapy, stratified by HLA-I LOH status. mOS for HLA-I intact (n = 180) was 11.3 months (8.2–15.3) and for HLA-I LOH (n = 60) was 8.0 months (5.2–13.1). HR for HLA-I intact = 0.68 (0.49–0.95), P = 0.02. C, OS of patients with nonsquamous NSCLC in the real-world clinical-genomic cohort from start of second-line ICI monotherapy, stratified by HLA-I LOH and TMB status (TMB high: ≥10 muts/Mb, TMB low: <10 muts/Mb). TMB high, HLA-I intact (n = 82) mOS 14.1 months (9.0–21.1), TMB high, HLA-I LOH (n = 31) mOS 10.9 months (6.6–20.0), TMB low, HLA-I intact (n = 98) mOS 9.6 months (6.2–14.8), TMB low, HLA-I LOH (n = 29) mOS 4.8 months (2.9–12.6). HR for TMB high = 0.74 (0.54–0.99), P = 0.046. HR for HLA-I intact = 0.65 (0.47–0.91), P = 0.013. (continued on next page)
employing a real-world clinico-genomic-dataset (8) and analyzing a cohort of 240 patients with EGFR- and ALK-wild-type, nonsquamous NSCLC who were immunotherapy-naive and received second-line ICI monotherapy between July 2014 and February 2019. No significant differences in demographics, treatment history, or biopsy timing were observed when stratifying the cohort by HLA-I LOH (Supplementary Table S1). We chose a second-line ICI-treated cohort because patients were presumably treated irrespective of PD-L1 status.

Similar to other published cohorts (4), our clinical cohort had a 10.8-month median OS (mOS) from start of second-line ICI. Overall, 25% of patients exhibited HLA-I LOH, and this group had a significantly decreased mOS compared with the HLA-I-intact group [Fig. 1B; mOS (months) HLA-I LOH: 8.0 (5.2–13.1); HLA-I intact: 11.3 (8.2–15.3); HR HLA-I intact remained constant across all TMB cutoffs (Fig. 1D). The mOS for patients with nonsquamous NSCLC in the real-world clinico-genomic cohort from start of second-line ICI monotherapy, stratified by the most statistically significant TMB (muts/Mb) and HLA-I status combination. mOS for patients that were any TMB, HLA-I intact or TMB low, HLA-I intact (n = 13, HLA-I LOH (n ≥ 13, HLA-I LOH ≥ 10 muts/Mb, TMB low: 5.5 months (3.6–9.8). TMB high, HLA-I LOH (n = 20, mPFS = 2.8 months (1.7–5.4). TMB low, HLA-I LOH (n = 22) mPFS 2.1 months (0.51–1.0), P = 0.066. HR for HLA-I intact = 0.67 (0.46–0.99). P = 0.044. For B–F, significance is determined by log-rank test.

Figure 1. (Continued) D, OS of patients with nonsquamous NSCLC in the real-world clinico-genomic cohort from start of second-line ICI monotherapy, stratified by HLA-I LOH and TMB status across multiple TMB thresholds (1–20 mut/Mb). For each threshold, TMB high ≥ TMB threshold and TMB low < TMB threshold. The HR is derived from multivariate Cox proportional hazards models controlled for TMB at each TMB threshold. E, OS of patients with nonsquamous NSCLC in the real-world clinico-genomic cohort from start of second-line ICI monotherapy, stratified by the most statistically significant TMB (muts/Mb) and HLA-I status combination. mOS for patients that were any TMB, HLA-I intact or TMB ≥ 13, HLA-I LOH (n = 203) was 12.2 months (9.1–15.3). The mOS for patients who were TMB < 13, HLA-I LOH (n = 37) was 6.0 months (2.9–8.9). HR for any TMB, HLA-I intact; TMB ≥ 13, HLA-I intact = 0.45 (0.31–0.66), P = 0.00004. F, PFS of ICI-treated patients in the MSK nonsquamous NSCLC cohort, stratified by HLA-I LOH and TMB status (TMB high: ≥10 muts/Mb, TMB low: <10 muts/Mb). TMB high, HLA-I intact (n = 46) mPFS 5.5 months (3.6–9.8). TMB high, HLA-I LOH (n = 20) mPFS 2.8 months (1.8–8.3). TMB low, HLA-I intact (n = 92) mPFS 3.5 months (2.7–5.4). TMB low, HLA-I LOH (n = 22) mPFS 2.1 months (1.7–5.4). HR for TMB high = 0.72 (0.51–1.0), P = 0.066. HR for HLA-I intact = 0.67 (0.46–0.99). P = 0.044. For B–F, significance is determined by log-rank test.

of 14.1 months (9.0–21.1), whereas the TMBlow, HLA-I LOH group had the shortest mOS of 4.8 months (2.9–12.6). To ensure the TMB cutoff had no effect on this observation, we tested all TMB thresholds between 1 and 20 mut/Mb. In agreement with published studies (7, 9), higher TMB cutoffs were associated with larger effects. However, the HR for HLA-I intact remained constant across all TMB cutoffs (Fig. 1D). Germline HLA-I zygosity did not affect mOS when stratifying by 6 versus <6 unique germline HLA-I alleles for either the entire cohort [Supplementary Fig. S1A; 6 allele HR = 1.0 (0.73–1.50), P = 0.8] or within the HLA-I-intact group [Supplementary Fig. S1B; 6 allele HR = 0.91 (0.60–1.40), P = 0.7]. To assess whether a combination biomarker could perform better than either HLA-I LOH or TMB alone, we stratified the cohort into two groups by setting different TMB thresholds depending on whether the sample was HLA-I intact or HLA-I LOH. The most significant difference was observed when combining all HLA-I-intact samples and TMB ≥ 13 mut/Mb for HLA-I LOH (Fig. 1E; HR = 0.45 (0.31–0.66), P = 0.00004), with 203 of 240 patients in the combination high group. Using TMB alone to create a predictor with a similar AUC yielded Area Under the Curve (AUC) = 0.73 (0.60–0.86), P = 0.00004.

Additional stratification by TMB showed TMB and HLA-I LOH were independent and statistically significant predictors of mOS in a multivariate Cox proportional hazards model [Fig. 1C; HR HLA-I intact: 0.65 (0.47–0.91), P = 0.013; HR TMBlow: 0.74 (0.54–0.99), P = 0.046]. The TMBhigh, HLA-I-intact group had the longest mOS
Pan-Cancer Analysis of HLA LOH as a Method of Immune Evasion

survival [mPFS; Fig. 1F, HR = significant effect of HLA-I LOH on median progression-free ≥ model with HLA-I LOH and TMB least one HLA-I gene. Similar to the real-world clinico-genomal panel (9). We defined HLA-I LOH as a shallow deletion of at treated patients with > LOH and TMB biomarker may better identify patients likely to benefit from ICI treatment.

We confirmed these observations in a cohort of 180 ICI-treated patients with EGRF– and ALK–wild-type, nonsquamous NSCLC profiled by the MSK-IMPACT–targeted NGS panel (9). We defined HLA-I LOH as a shallow deletion of at least one HLA-I gene. Similar to the real-world clinico-genomics cohort (Fig. 1C), a multivariate Cox proportional hazards model with HLA-I LOH and TMB ≥ 10 mut/Mb showed a significant effect of HLA-I LOH on median progression-free survival [mPFS; Fig. 1F, HR = 0.67 (0.46–0.99), P = 0.044]. When considered in isolation, HLA-I LOH trended toward significance for durable clinical benefit (DCB) and mPFS [Supplementary Fig. S2A and S2B; HLA-I LOH with DCB: 14% (8/56), without DCB: 28% (33/116), P = 0.055; mPFS HR = 0.71 (0.48–1.0), P = 0.077]. We also applied the combined predictor (Fig. 1E, all HLA-I intact; TMB ≥ 13 for HLA-I LOH) and found a statistically significant difference in mPFS [Supplementary Fig. S2C, HR = 0.61 (0.39–0.95), P = 0.040].

We next assessed 59 different solid tumor types that comprised 83,664 unique patient samples (Supplementary Table S2), enriched for patients with advanced disease. HLA-I LOH was detected in 17% of samples with 85% of events involving all three HLA-I genes. Prevalence varied widely across tumor types (2%–42%; Fig. 2A). Between major tumor categories, the highest rate of HLA-I LOH was observed in squamous cell cancers (30%), followed by nonsquamous carcinomas (16%), neuroendocrine tumors (11%), sarcomas (11%), and nonsquamous skin cancers (6%). Pan-cancer, PD-L1+ samples were enriched for HLA-I LOH (Fig. 2B; HLA-I LOH: PD-L1+ 25% vs. PD-L1− 16%, P < 0.0001), and disease prevalences of PD-L1+ and HLA-I LOH were linearly and positively correlated (Fig. 2B, P < 0.0001). HLA-I LOH was observed in 21% of TMBhi samples (TMB ≥ 10) and 16% of TMBlo samples (Fig. 2C, P < 0.0001). However, this simple association failed to capture the complex interplay between TMB and HLA-I LOH. Overall, the prevalence of HLA-I LOH compared with TMB exhibited a “Goldilocks effect” whereby diseases with the lowest TMB (e.g., neuroendocrine tumors) and highest TMB (e.g., cutaneous melanoma) exhibited the lowest prevalence of HLA-I LOH, whereas tumors in between exhibited the highest prevalence of HLA-I LOH (Fig. 2C). In line with this observation, samples with microsatellite instability exhibited a similar or lower rate of HLA-I LOH as compared with microsatellite-stable samples (Supplementary Fig. S3). This relationship was also assessed on a per-sample basis, and the same “Goldilocks” trend was observed pan-cancer (Fig. 2D). Two outliers to the TMB and PD-L1 associations were pancreatic islet cell tumors and adrenocortical carcinomas, which had low percentages of TMBhi (5%–10%) and PD-L1+ (3%–7%) despite considerable HLA-I LOH (36%–38%). In both diseases, HLA-I LOH was associated with deleterious mutations in a nearby tumor suppressor, DAXX (Supplementary Fig. S4, both P < 0.01), suggesting HLA-I LOH is driven by LOH of a nearby tumor suppressor in pancreatic islet cell tumors and adrenocortical carcinomas.

Given the complex relationship between TMB and HLA-I LOH, we further assessed the link between tumor antigens and HLA-I LOH. Neoantigenic driver mutations present a unique subset of neoantigens that drive oncogenesis yet provoke an immune response. We explored six putative neoantigens with specific HLA-I restrictions in samples with HLA-I LOH (10–15). For each neoantigen, the presenting allele was lost in 74% to 94% of events, which was statistically significant (Fig. 3A, P < 0.05). We also assessed 127 recurrent driver neoantigens predicted using NetMHCpan (16) and found the presenting allele was more frequently lost for 98% (125/127)

Figure 2. Pan-cancer landscape of somatic HLA-I LOH. A, Prevalence of HLA-I LOH across 59 different solid tumor types in 83,664 unique patient samples. The number of patients within each tumor type is listed in Supplementary Table S2. (continued on next page)
of predicted driver neoantigens with 62% (77/125) reaching statistical significance (Fig. 3B, \( P < 0.05 \)). Viral infection can also drive oncogenesis and recognition by the immune system (17). In tumor types where viral infection mediates cell-intrinsic oncogenic transformation, the prevalence of HLA-I LOH was increased in all virally infected disease subsets, reaching statistical significance in three of four diseases (Fig. 3C; \( P < 0.05 \)). In contrast, hepatitis B virus, which induces cellular transformation through hepatitis and cirrhosis after chronic infection, was not associated with HLA-I LOH in hepatocellular carcinoma (Fig. 3C; \( P = 1.0 \)). These data implicate HLA-I LOH as a potential mechanism by which tumors abrogate neoantigen-mediated immune recognition.

Lastly, we investigated the enrichment patterns of frequent genomic alterations and mutational signatures with and without HLA-I LOH across all tumor types (Fig. 4A).
Pan-Cancer Analysis of HLA LOH as a Method of Immune Evasion

with known oncoviral associations. EBV, Epstein–Barr virus; HBV, hepatitis B virus; HPV, human papillomavirus. Number of samples (virus-positive, virus-

The percentage of times the predicted presenting allele was either lost or kept during an LOH event is shown. Only predicted neoantigens involved with

Neoantigen prediction of recurrent driver mutations was conducted by NetMHCpan. Predicted neoantigens are listed by the gene and protein effect.

Figure 3. Somatic HLA-I LOH is a potential mechanism for immune evasion in samples with tumor antigen presentation. A, Putative neoantigens are

listed by the gene, protein effect, and presenting HLA-I allele. The percentage of times the presenting allele was either lost or kept during an LOH event is

shown. All putative neoantigens were involved with >5 events. Statistics were conducted by Binomial Test, and significance was determined as P < 0.05.

B, Neoantigen prediction of recurrent driver mutations was conducted by NetMHCpan. Predicted neoantigens are listed by the gene and protein effect. The percentage of times the predicted presenting allele was either lost or kept during an LOH event is shown. Only predicted neoantigens involved with >5 events are included. Statistics were conducted by Binomial Test, and significance was determined as P < 0.05. C, Prevalence of HLA-I LOH in tumor types with known oncoviral associations. EBV, Epstein–Barr virus; HBV, hepatitis B virus; HPV, human papillomavirus. Number of samples (virus-positive, virus-negative): head and neck squamous cell carcinoma (SqCC, n = 363, n = 771), cervical (n = 141, n = 121), gastric (n = 189, n = 1018), nasopharyngeal (n = 50, n = 38), hepatocellular (n = 64, n = 506). Statistics conducted by Fisher exact test, and significant (P < 0.05) associations are labeled with an asterisk.
Figure 4. Enrichment of genomic alterations in samples with somatic HLA-I LOH. A. Enrichment of genomic alterations in tumor types with HLA-I LOH (red) and without evidence of HLA-I LOH (blue). Tumor types in the top quartile overall in prevalence of APOBEC mutational signature (purple), tobacco smoking mutational signature (brown), and UV mutational signature (yellow) are shown. Only genes with enrichment in at least six different tumor types are included. B–E, Tumor type enrichment in samples with select genomic mutations, stratified by HLA-I LOH status. In A–E, statistics were conducted by Fisher exact test, and only significant (P < 0.05) associations are shown. In B–E, significant disease associations are colored by their respective major tumor type category.
Although we observed a significant association between metastatic samples and HLA-I status for seven tumor types, these results do not support a pan-tumor correlation between HLA-I LOH and tumor metastasis (Fig. 4B). Several genes were frequently associated with HLA-I LOH including TP53 in 14 tumor types, CDKN2A in 16 tumor types, and Pik3ca in 5 tumor types, of which were squamous cell carcinomas (Fig. 4C–E, P < 0.05).

DISCUSSION

In this work, we describe the landscape of HLA-I LOH across 59 different tumor types totaling 83,664 samples. HLA-I LOH is mechanistically connected to neoantigen presentation, which was supported by the preferential loss of the HLA-I allele presenting driver neoantigens in this work, and has also been associated with high TMB in NSCLC (6). However, our large pan-cancer dataset revealed that prevalence of HLA-I LOH appears to follow a “Goldilocks” pattern whereby tumors with high and low TMB exhibit low HLA-I LOH with higher rates in between. We hypothesize 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 a natural killer (NK)-cell response. In contrast, highly mutated cancers such as Merkel cell carcinoma have been shown to lose expression of all MHC-I from the cell surface, a mechanism that could potentiate an NK-cell response (18). We also found an association between HLA-I LOH and deleterious alterations in TP53, raising a possible link between genomic instability and the likelihood of an HLA-I LOH event occurring such that selection could act on it. Combined, these results suggest a surprising finesse to tumor evolution whereby tumors lose the proper amount of HLA-I presentation ability to avoid an immune response.

HLA-I LOH has the potential to refine TMB as a biomarker of checkpoint inhibitor responses based on a better understanding of neoantigen presentation by the tumor. In noninflamed tumors, the incidence of HLA-I LOH is low, and thus high TMB in these tumors may enrich for patients with superior responses to checkpoint inhibitors. This may explain the responses to monotherapy ICI seen in the pan-tumor trial investigating pembrolizumab efficacy in patients representing noninflamed cancers with high TMB (19). In contrast, high TMB was not predictive of OS in late-phase trials of NSCLC (1, 5). There has been active investigation into combining TMB with other biomarkers in late-phase trials of NSCLC (1, 5). There has been active investigation into designing therapeutic modalities that involve CD8 T-cell presentation (in a rational combination may lead to better patient stratification in NSCLC). Beyond ICI, assessing HLA-I LOH will be an important consideration for designing therapeutic modalities that involve CD8 T-cell responses such as neoantigen vaccines.

METHODS

Genomic Profiling

Genomic data were collected as part of routine clinical care for 83,664 patients using a targeted comprehensive genomic profiling assay in a Clinical Laboratory Improvement Amendments (CLIA)–certified, College of American Pathologists (CAP)–accredited, New York State–approved laboratory (FoundationOne), as previously described (22). DNA was extracted, and adaptor-ligated hybridization capture for all coding exons of 315 genes plus 28 introns frequently rearranged in cancer was performed. Libraries were sequenced to a median unique coverage depth of >500X. Analysis for genomic alterations, including short variant alterations (base substitutions, insertions, and deletions), copy-number alterations (amplifications and homozygous deletions), as well as gene rearrangements, was performed as previously described (22, 23). TMB was defined as the number of nondriver somatic coding mut/Mb of genome sequenced. Mutational signatures were determined in samples with ≥20 nondriver somatic mutations, including silent and noncoding alterations. Signatures were assigned using the COSMIC signatures of mutational processes in human cancer, as previously described (24). A positive status was determined if a sample had ≥40% fit to a mutational process. Viral DNA detection was performed through Velvet de novo assembly of sequencing reads left unmapped to the human reference genome (hg19). Assembled contigs were competitively mapped by BLASTN (BLAST+ v2.6.0) to the NCBI database of ≥3 million viral nucleotide sequences, and a positive viral status was determined by contigs ≥80 nucleotides in length with ≥97% identity to the BLAST sequence.

Histology

PD-L1 staining was performed at Foundation Medicine. PD-L1 status was determined through IHC performed on formalin-fixed
paraffin-embedded tissue sections, with the use of the commercially available antibody clones 22C3 (Dako/Agilent) or SP142 (Ventana). A board-certified pathologist determined the percentage of tumor cells with expression (0%–100%) and the intensity of expression (0, 1+, 2+). PD-L1 expression was reported as a continuous variable with the percentage of tumor cells staining with 2+ intensity. PD-L1 expression for each sample was also summarized as negative (<1% tumor cells) or positive (≥1% tumor cells). The pathology laboratory established performance characteristics for this assay per the requirements of CLIA ’88 and in accordance with CAP checklist requirements and guidance.

**HLA LOH Determination**

HLA-I zygosity was determined using the investigational somatic-germline-zygosity (SGZ) algorithm, a computational method for zygosity prediction from NGS results of mixed tumor–normal samples (20%–95% tumor), from pipeline v3.1.3 (23). SGZ models zygosity by taking into account the tumor purity, tumor ploidy, minor allele frequency (MAF), and local copy number of each genomic segment. The MAF of each HLA-I gene (HLA-A, HLA-B, and HLA-C) was calculated separately. HLA-I genotyping by FoundationOne sequencing results was performed by OptiType (25) v1.3.1 to a four-digit resolution. HLA-I reference sequences that matched the germline alleles for each sample were obtained from the IPD-IMGT/HLA database. Only germline heterozygous alleles were assessed for LOH, and samples identified as being germline homozygous at all three loci were not used in this study.

Sequencing reads that aligned to the HLA region of the human reference genome (hg19, 6p21-22) as well as all unmapped reads were extracted using Samtools v1.5. Picard v1.56 was used to remove all PCR and optical duplicates. Reads were competitively realigned to the HLA-I reference sequences specific to each sample using BWA v0.7.17. Samtools v1.5 was used to keep only uniquely aligned reads and to remove all unpaired mates. A local alignment was performed between each germline homologous allele using BLAST+ v2.6.0. The BTOP function was used to identify mismatch positions between each homologous allele. Samtools v1.5 was used to collect all unique reads that aligned to each mismatch position, and the allele frequency (AF) for each allele was calculated as the number of reads that uniquely aligned to one homologous allele divided by the total number of uniquely aligned reads to both homologous alleles.

We used a baited sequencing method to isolate regions of interest. For the HLA-I locus, we observed that homologous HLA pairs had a consistent AF skew (Supplementary Fig. S5A). To account for effects of hybridization on baiting efficiency due to HLA type, we modeled the observed AF (obsAF) as a function of AF and an association constant of an HLA type for the baits.

\[
\text{obsAF}_{ij} = \frac{k_i \times AF_{ij}}{k_i \times AF_{ij} + k_j \times AF_{ij}},
\]

where \(\text{obsAF}_{ij}\) is the observed AF of HLA type \(i\) in the pair \(ij\), \(k_i\) is the association constant of HLA type \(i\), and \(AF_{ij}\) is the AF of HLA type \(i\) in the sample with pair \(ij\). Note \(AF_{ij} = 1 - AF_{ij}\).

To fit the association constants, we assumed allelic balance (\(AF_{ij} = AF_{ij}\)). Given that most samples are not under LOH, we used the median AF of all samples with the same pair of homologous HLA types. We also added the constraint that 50 samples were required to provide a representative AF. Thus,

\[
\text{obsAF}_{ij} = \frac{k_i}{k_i + k_j}
\]

Combining these equations yields:

\[
\frac{\text{obsAF}_{ij} - k_i}{\text{obsAF}_{ij} - k_i} = \frac{1}{1 + \frac{AF_{ij}}{AF_{ij}^*}}
\]

To determine the best \(k\) values for each HLA type, we used least squares fitting:

\[
\min \sum \left( \frac{\text{obsAF}_{ij} - k_i \times \text{obsAF}_{ij}}{k_i \times \text{obsAF}_{ij} - k_j \times \text{obsAF}_{ij}} \right)^2 \text{ for all pairs}
\]

With the \(k\) values determined, the input AF can be determined from the observed AF after rearrangement of the model equation:

\[
AF_{ij} = \frac{k_i \times \text{obsAF}_{ij}}{k_i \times \text{obsAF}_{ij} + k_j \times \text{obsAF}_{ij}}
\]

We assessed how the association constants \((k)\) values mapped to sequence diversity in HLA-A. A dendrogram of HLA-A two-digite sequences exhibited two major branches with one branch having \(k\) values > 1 and the other having \(k\) values < 1, supporting our hypothesis that sequence-driven hybridization effects were the underlying cause of AF skewing (Supplementary Fig. S5B).

Using this model, adjusted MAF, representing the true MAF in the sample, was calculated from the observed MAF. The adjusted MAF was then used in the SGZ algorithm described above. At loci identified as having HLA-I LOH, the allele with the lower AF was determined to be the allele under LOH.

**Neoantigen Prediction**

End-to-end processing and MHC-I-binding predictions were calculated using NetMHCpan-4.0 and the IEDB API for all wild-type and mutant peptides (16). The API produces proteasomal cleavage scores, TAP transport scores, and MHC-I-binding affinities, as well as a total score which combines these values in an HLA-to-peptide-specific manner. Total scores of at least -0.8 and MHC-I-binding affinities of at most 500 nM/L, were used to dichotomize each peptide as a binder or a nonbinder in each sample. Upon identification of binders, binder mutant peptides were filtered against their wild-type counterpart.

**Clinico-Genomics Cohort and Survival Analysis**

The retrospective clinical analysis utilized the Foundation Medicine–Flatiron Health real-world clinico-genomic dataset (data collected through June 30, 2019) which includes electronic health record (EHR)–derived deidentified data for patients in the Flatiron Health database who underwent comprehensive genomic profiling by Foundation Medicine (8). The deidentified patient-level clinical data from the EHR included structured data (e.g., medication orders and administrations, lab tests, diagnostic codes) in addition to unstructured data (e.g., smoking status, histology) collected via technology-enabled chart abstraction from physicians’ notes by trained medical record abstractors who followed prespecified, standardized policies and procedures. Deidentified patient-level genomic data included specimen (e.g., TMB, tumor purity) and genomic (e.g., gene altered, alteration type) data reported by the FoundationOne comprehensive genomic profiling test.

The patients included in the clinical analysis were diagnosed with nonsquamous NSCLC and negative for EGFR and ALK alterations, as tested by FoundationOne. A variety of second-line ICI monotherapies were received, specifically nivolumab, pembrolizumab, durvalumab, and atezolizumab. The primary clinical end point studied was mOS from start of second-line ICI regimen until death or end of follow-up. In the OS analyses, to account for left truncation, patients were treated as at risk of death only after the later of their first sequencing report date and their second visit in the Flatiron
Health network on or after January 1, 2011, as both are requirements for inclusion in the cohort. For the Kaplan–Meier analyses, the log-rank test was used to compare groups. Significance of survival outcomes in this cohort was not affected when adjusted for race/ethnicity, age at the start of second-line ICI, first-line therapy received, and medical practice type, in a multivariate analysis. Statistics on patient demographics were conducted by a two-sided Fisher exact test except for the age comparisons, which were conducted by a two-sided Wilcoxon Rank-Sum Test. Analyses were performed on the R software version 3.6.0.

**MSK Cohort and Survival Analysis**

The MSK cohort analysis utilized publicly available data from ICI-treated patients with NSCLC who were profiled by the MSK-IMPACT targeted NGS panel (9). To better mirror the clinico-genomics cohort, we excluded all patients with squamous cell carcinoma as well as EGFR or ALK alterations. A list of alterations used for exclusion is included in Supplementary Table S3. We inferred HLA-I LOH–positive samples as those with a shallow deletion (log, copy number < -0.3) of a chromosomal segment that encompassed at least one HLA-I allele. One caveat to this inference method is that it will not account for copy-neutral HLA-I LOH and therefore is expected to have lower sensitivity than the HLA-I LOH determination method described above. For the Kaplan–Meier analyses, the log-rank test was used to compare groups, and analyses were performed on the R software version 3.6.0.

**Patient Consent and Data Availability**

Approval for this study, including a waiver of informed consent and a HIPAA waiver of authorization, was obtained from the Western Institutional Review Board (Protocol No. 20152817). Consented data that can be released are included in the article and its supplementary files. Patients were not consented for the release of underlying sequence data. Academic researchers can gain access to Foundation Medicine data in this study by contacting the corresponding author and filling out a study review committee form. You and your institution will be required to sign a data transfer agreement.

**Authors’ Disclosures**

M. Montesion reports personal fees from Foundation Medicine and other from Roche Holding AG outside the submitted work; in addition, M. Montesion has a patent for HLA-I LOH Algorithm pending. K. Murugesan reports personal fees from Foundation Medicine and other from Roche Holding AG outside the submitted work. D.X. Jin reports personal fees from Foundation Medicine and other from Roche Holding AG outside the submitted work. R. Sharaf reports personal fees from Foundation Medicine Inc. during the conduct of the study and personal fees from Foundation Medicine Inc. outside the submitted work. N. Sanchez reports Employee at Foundation Medicine at time of manuscript preparation. A. Guria reports personal fees from Foundation Medicine and other from Roche Holding AG outside the submitted work. M. Minker reports personal fees from Foundation Medicine and other from Roche Holding AG outside the submitted work. G. Li reports personal fees from Foundation Medicine outside the submitted work and stock ownership in F. Hoffmann-La Roche. E.S. Sokol reports personal fees from Foundation Medicine and other from Roche Holding AG outside the submitted work. D.C. Pavlick reports personal fees from Foundation Medicine, Inc. and personal fees from Roche Holding AG outside the submitted work. J.A. Moore reports personal fees from Foundation Medicine and other from Roche Holding AG outside the submitted work. A. Braly reports personal fees from Foundation Medicine and other from Roche Holding AG outside the submitted work. G. Singal reports personal fees from Foundation Medicine outside the submitted work. L.A. Comment reports personal fees from Foundation Medicine and other from Roche Holding AG outside the submitted work. N.A. Rizvi reports other from Gritstone Oncology outside the submitted work; in addition, N.A. Rizvi has a patent for PCT/US2015/062208 pending to PGDX. B.M. Alexander reports personal fees from Foundation Medicine and other from Roche during the conduct of the study; personal fees from Foundation Medicine and other from Roche outside the submitted work. P.S. Hegde reports personal fees from Foundation Medicine/Roche outside the submitted work. L.A. Albacker reports personal fees from Foundation Medicine and other from Roche Holding AG outside the submitted work; in addition, L.A. Albacker has a patent for HLA-I LOH algorithm pending. No disclosures were reported by the other authors.

**Authors’ Contributions**

**M. Montesion:** Conceptualization, data curation, software, formal analysis, validation, investigation, visualization, methodology, writing-original draft, writing-review and editing.

**K. Murugesan:** Data curation, software, formal analysis, validation, investigation, visualization, methodology, writing-review and editing.

**D.X. Jin:** Data curation, software, formal analysis, validation, methodology, writing-review and editing.

**R. Sharaf:** Conceptualization, data curation, formal analysis, validation.

**N. Sanchez:** Data curation, project administration.

**A. Guria:** Data curation.

**M. Minker:** Data curation.

**G. Li:** Data curation, validation.

**E.S. Sokol:** Data curation, software, formal analysis, validation, methodology.

**D.C. Pavlick:** Data curation, validation.

**L.A. Albacker:** Data curation, validation.

**B.M. Alexander:** Supervision.

**G.M. Frampton:** Conceptualization, data curation, supervision.

**P.S. Hegde:** Conceptualization, supervision, writing-original draft.

**L.A. Albacker:** Conceptualization, data curation, software, formal analysis, supervision, visualization, methodology, writing-original draft, writing-review and editing.

**Acknowledgments**

The authors thank Ericka Ebot, Jeremy Snider, and Emily Castelanos for scientific discussions and thoughtful comments on the text. This study was self-funded by Foundation Medicine, Inc.

Received May 18, 2020; revised September 14, 2020; accepted October 27, 2020; published first October 30, 2020.

**REFERENCES**


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

Meagan Montesion, Karthikeyan Murugesan, Dexter X. Jin, et al.


**Updated version**
Access the most recent version of this article at:
doi:10.1158/2159-8290.CD-20-0672

**Supplementary Material**
Access the most recent supplemental material at:
http://cancerdiscovery.aacrjournals.org/content/suppl/2020/10/30/2159-8290.CD-20-0672.DC1

**Cited articles**
This article cites 25 articles, 5 of which you can access for free at:
http://cancerdiscovery.aacrjournals.org/content/11/2/282.full#ref-list-1

**Citing articles**
This article has been cited by 1 HighWire-hosted articles. Access the articles at:
http://cancerdiscovery.aacrjournals.org/content/11/2/282.full#related-urls

**E-mail alerts**
Sign up to receive free email-alerts related to this article or journal.

**Reprints and Subscriptions**
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

**Permissions**
To request permission to re-use all or part of this article, use this link
http://cancerdiscovery.aacrjournals.org/content/11/2/282.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.