Composite impact of genome-wide APOBEC3-mediated mutations and HLA haplotype on cancer immunogenicity has a sex-biased survival impact

Faezeh Borzooee[1,2], Alireza Heravi-Moussavi[1,2], Mani Larijani[1]*

1. Department of Molecular Biology and Biochemistry, Faculty of Science, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada

2. BC Cancer Research Center, Vancouver, British Columbia, V5Z 1G1, Canada

To whom correspondence may be addressed: mani_larijani@sfu.ca

Keywords: genome mutations, immunogenetics, anti-tumor immunity, cytotoxic T cells

Abstract

APOBEC3A and APOBEC3B genome mutator enzymes drive tumor evolution and drug resistance. However, their mutational activity can also generate neoepitopes that activate cytotoxic T cells (CTLs). Given the high polymorphism of Class I HLA, the CTL immunopeptidome is individual-specific. We used a genome-wide immunogenicity scanning pipeline to assess how APOBEC3A/B-induced mutations affect the immunogenicity of the entire human immunopeptidome, consisting of all possible 8-11mer peptides restricted by several thousand HLA class I alleles. We evaluated several billion APOBEC3-mediated mutations for their potential to alter peptide:MHC and T cell receptor binding, either increasing or decreasing immunogenicity. We then ranked HLA alleles based on the degree to which their restricted immunopeptidome lost or gained immunogenicity when mutated by APOBEC3A or APOBEC3B. We found that HLA class I alleles vary infinitely in the proportions of their immunopeptidome whose immunogenicity is diminished vs. enhanced by APOBEC3-mediated mutations, with mutations in APOBEC3B hotspots having the greatest potential for enhancement of immunogenicity. The cumulative potential of an individual's HLA haplotype's immunopeptidome to gain or lose immunogenicity upon APOBEC3-mediated mutation predicts survival in APOBEC3-mutated tumors and correlates with increased CD8+ T cell activation. Thus, HLA haplotype is a prognostic marker in APOBEC3-mutated tumors.

Significance statement

This work is the first to conduct a genome-wide scan to quantify the impact of A3A- and A3B-mediated mutations on the immunogenicity changes in the human immunopeptidome presented by several thousand HLA class I alleles. We found that the effects of APOBEC3-mediated mutations vary significantly across the immunopeptidomes of different HLA alleles, and that HLA genotype is a predictor of survival and immune activation status in APOBEC3-mutated tumors, reflecting each individual's unique HLA class I immunopeptidome's propensity for either increased or decreased immunogenicity when altered by A3A and/or A3B. The identification of HLA genotype as a biomarker in APOBEC3-mutated tumors presents a promising opportunity for developing personalized immunotherapy strategies and exploring potential A3 inhibition. This study highlights that the effects of APOBEC3-mediated mutations on immunogenicity are not random; rather, individual HLA class I polymorphisms determine whether these mutations enhance or suppress anti-tumor immunity, which has direct prognostic implications for solid tumors.

Introduction

The APOBEC3 cytidine deaminases are a major source of tumor genome mutations. APOBEC3s can mutate any cytidine, but each member has a favored hotspot. APOBEC3A (A3A) and APOBEC3B (A3B) which respectively prefer 5'YTCA3', 5'RTCA3' (R=purine, Y=pyrimidine) have emerged as the major mutators responsible for single base substitution (SBS) 2 and 13 signatures in solid tumors including lung, breast, bladder, head and neck and ovarian cancer (1-6). A3A- and A3B promote tumorigenesis, tumor clonal heterogeneity and drug resistance in cell culture and mouse models, and correlate with poor prognosis in human cancers (6-14)

Central to anti-tumor immunity are $CD8^+$ cytotoxic T cells (CTL) whose T cell receptor (TCR) recognizes 8-11mer peptide epitopes presented on tumour cells by MHC (encoded by HLA) class I. Established as A3A and A3B are as pro-tumorigenic factors, several findings suggest that they may also promote anti-tumor immunity: first, A3B-mediated generation of heteroclitic or neo-epitopes for CTL has been demonstrated in a mouse model (15). Second, higher expression of A3B and A3G in cervical cancer correlates with higher immune cell tumor infiltration (16,17). Third, a bioinformatics predictive study suggests that APOBEC3-mediated SBS2/13 mutations increase the hydrophobicity of CTL epitopes which promotes TCR and HLA binding (18,19).

HLA class I are highly polymorphic comprised of several thousand alleles. The HLA polymorphisms particularly affect their peptide epitope binding "anchor" residues, culminating in HLA alleles presenting different sets of peptides. Thus, each person's haplotype of six A, B, C alleles determine a unique CTL immunopeptidome. Accordingly, HLA allelic differences have been associated with anti-tumour immunity and immunotherapy outcomes (20,21).

In contrast to promotion of CTL epitope generation, APOBEC3-mediated mutations may also diminish peptide epitope immunogenicity leading to tumor CTL evasion. To gain insight into the balance of these opposing outcomes, we examined the predicted impact of A3A- or A3B-mediated mutations on HLA class I anchor and TCR contact amino acid positions for the CTL immunopeptidomes of several thousand HLA I alleles.

We report that A3A- and A3B-mediated genome mutations can indeed enhance or diminish CTL epitope immunogenicity. The relative probability of these two outcomes vary infinitely across the immunopeptidomes restricted to different HLA class I alleles, and the carriage of HLA alleles whose immunopeptidomes predominantly gain or lose immunogenicity due to A3A- and A3B-mediated mutations predicts survival in APOBEC3-mutated cancers. These results underscore the intricate relationship between APOBEC3-mediated mutations, HLA genotype and anti-tumor immunity, and highlight the importance of considering individual immunogenetic variations in APOBEC3-mutated tumors.

Materials and Methods

Scoring the genome-wide impact of APOBEC3-mediated mutations on the CTL immunopeptidome

The reference human HLA class I-binding immunopeptidome includes all possible 8-11mer peptides from ~ 21,000 protein including 71,000 isoform sequences, and ~46,000,000 unique peptides, whose HLA restriction includes 2,910 HLA class I alleles. This dataset (https://zenodo.org/record/1453418#.YoPo2ejMLUQ) which includes a total of 987,968,036 peptide:MHC (pMHC) combinations formed the initial basis of our AID/APOBEC3-mediated mutation simulation (Table 1) (22). tBlastn was used to find the corresponding encoding DNA coding sequence, using the following parameters: Database (Reference RNA sequences (refseq-rna)), Homo sapiens (taxid:9606), Max Target (100), Expected threshold (10,000), Word size (3), Low complexity filter (deselected). The encoding DNA sequence, including 2 nucleotides preceding the first codon were subject to mutation simulated at APOBEC3 hotspots (YTC (A3A), RTC (A3B), and WRC (AID), Y= purine, R= pyrimidine, W=A/T) (23) as follows: (AID: "TGC", "TGT"),(AID: "AGC", "AGT"),(AID: "TAC", "TAT"),(AID: "GCA", "ACA"),(AID:"GTA", "ATA"),(AID:"GCT", "ACT"),(A3A:"TTC", "TTT"),(A3A: "GAA", "AAA"),(A3A: "CTC", "CTT"),(A3A: "GAG", "AAG"),(A3B:"ATC", "ATT"),(A3B:"GAT", "AAT"),(A3B: "GTC", "GTT"), (A3B:"GAC", "AAC"). NetMHCpan (version 4.1) was used to predict pMHC binding affinities changes between the wildtype (WT) and APOBEC3-mutated (Mut) peptides, which were categorized as shown in Figure 1. In addition, we predicted the immunogenicity impact of changes in the TCR binding portion of each peptide, as shown in Figure 1. This combined immunogenicity score (CIS) was obtained by considering amino acid hydrophobicity, polarity changes at the terminal MHC-binding anchor residues (1,2,9,10) and TCR contact positions (residues 4 and 6 received a higher weight than 3,5, and 7, since they are typically the more dominant TCR contacts) (24,25). For each HLA allele, we tallied the counts of WT to Mut changes (CIS event count), as well as the magnitude of WT to Mut changes (CIS value) that were predicted to decrease immunogenicity ("-", pro-tumour) increase immunogenicity ("+", anti-tumour) and calculated a +/- ratio. The 2910 HLA alleles were ranked by their +/- ratios of pMHC binding, CIS event count, and CIS values.

HLA genotype- and APOBEC3-based solid tumor survival analysis

We used mutSignatures (version 2.1.5)(26) for extraction and analysis of mutational signatures in 8956 sequenced tumors from TCGA encompassing 30 cancer types and HLA OptiType as well as clinical data(26,27). The reliability of de novo extracted signatures was assessed by inspecting the silhouette plot at the end of the signature extraction process. The extracted signatures were then compared and matched with the COSMIC database to identify the cases bearing the COSMIC signatures 2 and 13 of APOBEC3s.

A scoring system was developed to evaluate the likelihood of an individual's HLA class I haplotype (six A, B, C alleles) to support the balance of pro-tumor or anti-tumor immunity because of APOBEC3-mediated mutations. If the anti-tumor (+)/pro-tumor (-) CIS ratio was > 1 or <1, an HLA allele was considered as anti- or pro-tumour, respectively. An HLA allele with a +/- CIS ratio of 1.6, has a restricted immunopeptidome that is 1.6 times more likely to result in an anti-tumor

than a pro-tumor outcome when mutated by A3A or A3B. Conversely, an HLA allele with a +/- CIS ratio of 0.1, has a 10-fold higher likelihood of supporting a pro-tumour decreased immunogenicity outcome when its immunopeptidome is mutated by APOBEC3, and would be assigned 1/CIS=10. The anti-tumor and pro-tumor values for each of the six HLA of an individual's haplotype were added to yield a cumulative +/- ratio for an individual, using the individual's cumulative HLA anti-tumor CIS values as the numerator, and cumulative HLA pro-tumor rank of CIS values as the denominator. This ratio (index) which reflects the cumulative +/- potential of each individual's six A, B and C HLA class I alleles was used to rank all patients, and correlations with immune parameters of tumors and survival data were examined using linear models, and the survival (version 3.3.1) and survminer (version 0.4.9) packages in R (24). The primary query of interest was overall survival, defined as the time from diagnosis to the last follow-up for which the Kaplan-Meier estimator was used. Survival curves were generated, and log-rank tests were conducted to compare survival curves between different groups.

Immune marker and differential gene expression analyses

To identify differentially expressed genes we utilized DESeq2 and immune marker analyses as previously described (27). TCGA sample count data were downloaded and used as the input. For A3A, A3B, as well as A3A+A3B combined we defined high and low index quantile groups. The count matrix was imported into R and converted into a DESeq2 dataset object. The APOEBC3 mutational signature and A3A and A3B TPM values were used as covariates in the final design model. P-values obtained from the differential expression analysis were adjusted for multiple testing using the Benjamini-Hochberg procedure to control the false discovery rate (FDR). Genes with an adjusted p-value (Padj) <0.05 and an absolute log2 fold change >1 were considered significantly differentially expressed.

Statistical analysis

For the predictive phase of the work on differences among the APOBEC3 enzymes, linear models were used to determine statistical significance. Wilcoxon test was used to compare means. For survival analyses, Kaplan Meyer log-rank tests were conducted to compare survival curves between different groups.

Data availability

Data are available at https://github.com/larijani-lab/

Results

Prediction of immunogenicity changes of the complete HLA class I immunopeptidome as a result of APOBEC3-mediated mutations

We simulated every possible mutation in hotspot motifs of A3A (YTC) and A3B (RTC) on both strands of the genomic sequences encoding each peptide in the CTL immunopeptidome which includes all possible 8-11mer peptides that can result from proteasomal degradation of the human proteome and potentially be presented by at least one of 2910 HLA class I. In contrast to A3A and A3B which are the main mutational culprits in solid tumors, AID's mutagenic activity is largely restricted to leukemia/lymphomas for which the CTL response is not a key factor. Thus, mutations in AID hotspots (WRC) were simulated as a control (Figure 1).

As shown in table 1, most peptide-encoding genomic sequences (93%) contained at least one AID/APOBEC3 hotspot. In total, ~4.5 billion mutation events were simulated of which the majority (~2.6 billion, 62%) yielded non-synonymous (NS) amino acid changes. A3B-mediated mutations had a lower likelihood of generating NS changes relative to A3A and AID (NS/S ratio of 1.2, 1.5 and 1.9, respectively), and mutations in A3B hotspots yielded less mutated peptides relative to A3A (~571 vs 869 million) in total. The remaining epitopes devoid of APOBEC3 hotspots or in which mutations yielded synonymous codons may represent advantageous candidates for immunotherapy, being exempt from immune escape-inducing mutation by APOBEC3s.

Each APOBEC3-simulated mutation was analyzed for the change between the wildtype (WT) and the APOBEC3-mutated (Mut) pair's predicted binding affinity to the restricting HLA (pMHC formation). Each WT-Mut pair was classified as having one of four outcomes: lowering/loss of pMHC binding affinity considered as a pro-tumour or immune escape potential (designated "-"), gain of pMHC affinity considered as anti-tumour or enhanced CTL recognition potential (designated "+"), no change in binding affinity "NC", or synonymous mutations, the latter two being neutral events, neither pro- nor anti-tumor (supplementary tables 1-3).

Considering all WT-Mut pairs, A3A and AID were 4.6 and 3.7 fold more likely to generate "-" than "+" mutations, whereas this overall -/+ for A3B was a less pro-tumor biased 2.1-fold. Reflecting the lesser pro-tumor potential of A3B, the median of -/+ ratio across all HLA alleles for A3A, AID, and A3B were 20, 19 and 7.7, respectively (Figure 2A).

Thus, while both A3A and A3B mutations predominantly induce non-synonymous changes in the immunopeptidome, A3A and AID mutations are more likely to lead to immune escape events compared to A3B.

Variation among the immunopeptidomes of HLA class I alleles in the potential for immune escape or enhanced anti-tumour immunity after APOBEC3-mediation mutation

We ranked HLA alleles by +/- ratio reflecting how likely APOBEC3-mediated mutations of their restricted immunopeptidome was to generate enhanced "+" vs. diminished "-" immunogenicity

outcomes (Figure 2B, supplementary tables 4-6). Among HLA alleles, a remarkable variation of emerged in the +/- ratios (Figure 2B). The majority of HLA alleles fell at the extreme pro-tumor end as evident by two observations: first, A3A-, A3B- and AID-mediated mutations led to 100% pro-tumour "-" outcomes (i.e. +/- ratio=0) for ~15% of HLA alleles (428, 452, 411 alleles, respectively); second, A3A-, A3B-and AID-mediated mutations are 10-times more likely to yield pro- than anti-tumour pMHC outcomes (+/- ratio < 0.1) for 1820, 1330 and 1802 HLA alleles out of the total 2910 alleles, respectively. ~500 HLA alleles were 1000-times more likely to support a pro-tumour outcome for either A3A-, A3B-, or AID-mediated mutations (+/- ratio < 0.001).

At the other end of the spectrum, a +/- ratio >1 would indicate an HLA allele whose restricted immunopeptidome is more likely to have enhanced than diminished immunogenicity when mutated by APOBEC3s. The immunopeptidomes mutated by A3A and AID never yielded an HLA with a +/- ratio >1, with < 3% of all HLA alleles (86 and 57, respectively) supporting +/- ratio > 0.5, and the highest for an HLA allele being 0.9 and 0.7, respectively (Figure 2B). In contrast, a significantly higher proportion of HLA alleles (509 HLA alleles, 17% of all alleles) had a +/->0.5 by A3B-mediated mutation of their immunopeptidome, of which 115 HLA alleles had +/- ratio >1, with the highest reaching a +/- ratio ≈2 (Figure 2B).

The data revealed two main findings. First, there is significant variation among HLA alleles in the degree to which their restricted immunopeptidomes' immunogenicity is diminished or enhanced by APOBEC3-mediated mutations. Second, A3B-mediated mutations support the most anti-tumor outcomes (enhanced immunogenicity) relative to A3A and AID, both in the number of HLA alleles with higher +/- ratios and the higher value of the +/- ratios (supplementary tables 4-6). Thus, while most HLA alleles are skewed towards pro-tumor outcomes, particularly under A3A and AID mutations, a subset, especially under A3B influence, exhibits a higher likelihood of supporting anti-tumor responses.

Combined MHC binding and T cell receptor recognition score reveals that APOBEC3B-mediated mutations have the highest potential for enhancing CTL epitope immunogenicity

We then generated a comprehensive cumulative immunogenicity score (CIS) for each WT:Mut pair which also accounts for T cell recognition in addition to pMHC formation, by factoring in the position and nature of the amino acid changes in the TCR contact positions of the pMHC complex (Figure 1). For each HLA allele, +/- and -/+ ratios were calculated either by enumerating the event counts (mutational events with a - score / + score) or the CIS value (the sum of all - scores / + scores) (supplementary tables 7-9). As in the pMHC analysis, the overall A3A- and AID-mutated immunopeptidome produced a predominantly reduced immunogenicity "-" outcome reflected by event count -/+ of 1.9 and 3.3-fold and a CIS sum -/+ of 2.2 and 3.7-fold, respectively. In contrast, as in the pMHC binding analyses, A3B-mediated mutations produced a more anti-tumor profile, with event count and CIS sum -/+ of 0.9 and 1, respectively (Figure 3A). The median values of the +/- CIS count and sum for all HLAs were 0.6 and 0.2, for A3A and AID, respectively, while the median +/- CIS count and sum for A3B was 1.1, reflecting the significantly higher number of HLA alleles whose immunopeptidomes yielded a higher +/- CIS ratio when mutated by A3B (Figure 3A). The A3A-mutated immunopeptidome had 453 and 461 HLA alleles with +/- Cis event and sum ratios of >1, while 1140 and 1029 HLA alleles had +/- Cis event and sum ratios of >1 as a result of A3B-mediated mutations (Figure 3B). The AID-mutated immunopeptidome was the most

pro-tumor with only 70 and 98 alleles having a +/- Cis ratio of > 0.5, and none >1. The CIS event count and value sum trends exhibited a linear correlation across HLA alleles, as expected (Supplementary Figure 1), and the level of variation among HLA alleles due to APOBEC3-mediated mutations followed a pattern of "+">"-">NC/synonymous, with the two categories that are inert with respect to pMHC binding affinity changes (NC, and synonymous) showing the highest level of variation being in the relative proportion of the "+" category changes (Supplementary Figure 2).

Hydrophobic and non-polar amino acids are prevalent in HLA-binding anchor terminal peptide positions and in the TCR contact positions of immunogenic CTL epitopes(18,19). A3A-, A3B-, and AID-mediated mutations produced amino acid changes towards increased hydrophobicity and non-polarity in MHC anchor or TCR contacts, while A3A- and A3B-mediated mutations enhanced non-polarity compared to AID-mediated mutations at peptide anchor residues (Supplementary Figure 3). As a consequence of A3B-mediated mutations, 2 hydrophobic amino acids (L and M), 4 neutral amino acids (G, S, P, and H), and 2 hydrophilic amino acids (R and D) were modified to 6 hydrophobic (F, L, I, C, W, and Y), 3 neutral (K, N, and S), and 2 hydrophilic amino acids (E and G) (Supplementary Figure 4). A comparison of the pMHC and TCR components of CIS separately reveals that A3A- and A3B-mediated mutations generated a similar count and values of "+" TCR scores across the HLA alleles, but that A3B-mediated mutations led to significantly higher pMHC binding enhancements, approximately 3-fold better than A3A on average. (Supplementary Figure 5).

Thus, the combined analysis of MHC binding and TCR recognition reveals that A3B-mediated mutations have the highest potential to enhance CTL epitope immunogenicity compared to A3A and AID. This is reflected by a greater number of HLA alleles with positive CIS values for A3B, indicating a more favorable anti-tumor immune profile. APOBEC3-mediated mutations, especially from A3B, tend to increase hydrophobicity and non-polarity at MHC anchor and TCR contact positions in CTL epitopes. A3B mutations notably enhance pMHC binding more than A3A, potentially boosting immunogenicity and improving immune recognition of tumors.

HLA class I genotype is a predictor of survival and CD8+ T cell engagement in APOBEC3-mutated tumors

We generated a "triple index" to rank HLA alleles by combining each allele's +/- rank across pMHC binding, CIS event count, and CIS value sum to identify HLA alleles whose immunopeptidomes have the highest +/- potential when mutated by A3A or A3B individually, or the combined action of A3A+A3B (Table 2). We identified 332 HLA alleles for which the +/- ratio is >1 if mutated by A3A or A3B (Figure 4A, Table 2). Of these, 232 HLA alleles fall in the top 90th percentile of the triple index rank for both A3A and A3B. These represent HLA alleles that have immunopeptidomes with a higher likelihood of + than – category mutations and therefore the highest potential to support enhanced anti-tumour CTL immunity, if a tumour's genome is mutated by A3A, A3B or A3A+A3B. 55%, 54% and 60% of HLA alleles that supported the most anti-tumor "+" outcomes as a result of A3A-, A3B- and A3A+A3B-mediated mutations, respectively, are among the common, intermediate and well-documented (CIWD) list of HLA class I alleles

(Table 2)(28). Taken together, the pMHC and pMHC-TCR (CIS) binding predictions demonstrate two concepts: first, a remarkable variation in the potential of APOBEC3-mediated mutations to diminish or enhance the immunogenicity of the CTL immunopeptidome dependent on HLA class I genotype; second, that, relative to A3A, mutations in A3B hotspots have a higher potential to enhance epitope immunogenicity, in the number of HLA alleles and the magnitude of the +/- ratio.

We sought to examine whether the prediction that HLA class I alleles differentially channel APOBEC3-mediated mutations towards enhanced immunogenicity "+" or diminished immunogenicity "-" correlates with tumor immune markers and patient survival. We considered the TCGA subset with the APOBEC signature SBS2/13 which expectedly also exhibit a higher expression level of A3A and A3B relative to tumors that do not bear this mutational signature (Supplementary Figure 6). For each HLA class I allele, a combined CIS event count and CIS value +/- ratio for A3A+A3B was calculated (supplementary tables 10, 11) and a cumulative +/- ratio was determined for each patient reflecting the cumulative scores of their six HLA A, B and C alleles in their haplotype. Patients were ranked using this cumulative HLA genotype +/- index value, with "high" and "low" index denoting the anti-tumor and pro-tumor ends of the spectrum. Thus, "high index" patients have a HLA class I genotype with the highest potential to channel A3A+A3B-mediated mutations towards enhanced immunogenicity, whereas "low index" patients have an HLA class I genotype with the highest potential for diminished epitope immunogenicity upon A3A+A3B-mediated mutations. We observed a significant difference in survival in favor of the high index patients ($p=0.024$) while the same HLA class I-based stratification for patients whose tumors did not bear the APOBEC3 mutational signature revealed no survival difference (Figure 4B).

The pMHC and CIS data suggest that mutations in A3B-favored motifs potentiate a higher count and magnitude of "+" mutations than A3A. Thus, we stratified patients according to HLA class I haplotypes that are predicted to be at the anti- or pro-tumor ends of the spectrum for either A3A or A3B separately. We observed a significant correlation between enhanced patient survival and carriage of HLA class I's that are anti-tumor for A3B-mediated mutations ($p=0.0063$), but a non-significant trend in the case of A3A ($p=0.2$) (Figure 4B). When the survival data based on A3B were further stratified based on biological sex, we observed that the survival difference effect was significant in males ($p=0.031$), whereas non-significant trends were observed in females ($p=0.41$) and for both sexes for A3A-mediated mutations ($p>0.4$) (Figure 4B).

The observation that high-index patients exhibit improved survival compared to low-index patients due to the carriage of HLA class I genotypes that are more biased towards enhancing CD8+ T cell epitope immunogenicity would mean that APOBEC3-mediated mutations enhance CD8+ T cell response in high-index patients. To investigate this, we carried out immune marker and differential gene expression analyses on patients stratified based on high and low index HLA genotypes because of A3A+A3B- or A3B-mediated mutations, and whose tumors bear the APOBEC3 mutational signature. Indeed, we observed increased tumor-infiltrating CD8+ T cells (Figure 5A), and increased markers of CD8+ T cell activation, tumor killing, and T cell exhaustion in high-index patients relative to low-index patients (Figure 5B). These findings indicate that the enhanced survival of high-index patients due to the carriage of HLA class I alleles, whose immunopeptidomes become more immunogenic by APOBEC3-mediated mutations, indeed correlates with a more robust anti-tumor CD8+ T cell profile.

Discussion

Though the pro-tumorigenic roles of APOBEC3s' in cancer initiation and progression are established from clinical and mouse model studies, a few studies have noted positive correlations between APOBEC3 expression and improved clinical outcomes(4,5,7,15-17,29,30). In line with the known association between TMB and the success of immunotherapy, one anti-tumor activity of APOBEC3s could be neo-epitope generation. Such a context-dependent pro-tumor but also tumor-suppressive dual role including promotion of anti-tumor immunity has become increasingly established for A3B(15,30). A3B-induced mutations leads to a library of neo-epitopes, which increase the efficacy of immune inhibitor blockade therapy intermediated by CTL and NK cells. A3B activity also induces anti-tumor immune responses and $CD4^+$ T cell-mediated antigen-specific tumor growth inhibition(15). A3B-mediated mutagenesis sensitizes HER2-driven mammary tumors to anti-CTLA-4 monotherapy and CTLA-4/anti-PD1 combination therapy(30). Conversely, APOBEC3-mediated mutations could generate immune escape mutations. In clinical and mouse model studies, immune evasion mutations would be over-represented compared to those that boost immune recognition. The predictive phase of our work was motivated by the need to decipher the balance of APOBEC3's potential to generate immune escape vs. immune boosting mutations in an unbiased sandbox.

We report several observations in support of a clinically significant role for APOBEC3-mediated mutations in regulating the anti-tumor CTL response: first, that relative to A3A, A3B-mediated mutations have the highest potential to enhance the immunogenicity of the CTL immunopeptidome, in agreement with the aforementioned studies on the anti-tumor roles of A3B in neoepitope generation. Second, there is an infinite level of variation among HLA class I alleles: some alleles have immunopeptidomes that when mutated by APOBEC3s support entirely a pro-tumor immune escape consequence, while other alleles have immunopeptidomes with a higher potential for generation of neo-antigens than escape variants. Third, the carriage of the predicted pro- or anti-tumor HLA class I genotypes, as a cumulative function of one's HLA haplotype correlate with patient survival and CTL tumor infiltration. Fourth, the impact of APOBEC3-mediated mutations of the immunopeptidome on patient survival is not only HLA genotype dependent, but also sex-dependent. Relative to females, males generally display reduced neoantigen presentation, lower HLA expression, and reduced CTL effector properties(31). These differences may at least partially explain why the impact of A3B-mediated mutations in neo-epitope generation is a significant boost to CTL activation and survival in male patients.

Many of the HLA alleles whose immunopeptidomes generated the highest +/- values (anti-tumor potential of enhanced immunogenicity) when mutated by A3A and/or A3B, and the carriage of which correlates with survival, are among the CIWD HLA subset. Several belong to HLA-A03, HLA-A24 and HLA-A01A24 supertypes. HLA-A*03 is a negative biomarker for patients undergoing immune checkpoint inhibitors therapy(21). Based on our data, we speculate that the patients who carry HLA-A*03 with high APOBEC3 levels would have an abundance of neo-antigens leading to preferential elimination of APOBEC3+ tumors. Given the correlation between expression of APOBEC3 and immune checkpoint molecules, these would also be less sensitive to checkpoint inhibition therapy.

This work highlights the importance of considering an individual's HLA genotype in APOBEC3-mutated tumors. HLA genotype can serve as a biomarker to guide immunotherapy and future APOBEC3 inhibition therapeutic approaches(32). Study limitations include not assigning differential weights in mutation simulation for secondary DNA structure which may generate regions favored by each APOBEC3 enzymes. Including this level of refinement would present a challenge as these regions would differ across individual tumors. Second, it is possible that the loss of immunogenicity of an epitope due to APOBEC3-mediated mutations could coincide with gain of binding to another co-inherited HLA. In addition, within a given protein, an entirely new epitope can emerge because of APOBEC3-mediated mutations. These questions can be addressed at the individual genomic level.

References

1. Kim, Y.S., Sun, S., Yoon, J.S., Ko, Y.H., Won, H.S. and Kim, J.S. (2020) Clinical implications of APOBEC3A and 3B expression in patients with breast cancer. *PLoS One*, **15**, e0230261.
2. Mas-Ponte, D. and Supek, F. (2020) DNA mismatch repair promotes APOBEC3-mediated diffuse hypermutation in human cancers. *Nat Genet*, **52**, 958-968.
3. Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Campbell, P.J. and Stratton, M.R. (2013) Deciphering signatures of mutational processes operative in human cancer. *Cell Rep*, **3**, 246-259.
4. Burns, M.B., Temiz, N.A. and Harris, R.S. (2013) Evidence for APOBEC3B mutagenesis in multiple human cancers. *Nat Genet*, **45**, 977-983.
5. Roberts, S.A., Lawrence, M.S., Klimczak, L.J., Grimm, S.A., Fargo, D., Stojanov, P., Kiezun, A., Kryukov, G.V., Carter, S.L., Saksena, G. *et al.* (2013) An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nat Genet*, **45**, 970-976.
6. Petljak, M., Alexandrov, L.B., Brammeld, J.S., Price, S., Wedge, D.C., Grossmann, S., Dawson, K.J., Ju, Y.S., Iorio, F., Tubio, J.M.C. *et al.* (2019) Characterizing Mutational Signatures in Human Cancer Cell Lines Reveals Episodic APOBEC Mutagenesis. *Cell*, **176**, 1282-1294 e1220.
7. Swanton, C., McGranahan, N., Starrett, G.J. and Harris, R.S. (2015) APOBEC Enzymes: Mutagenic Fuel for Cancer Evolution and Heterogeneity. *Cancer Discov*, **5**, 704-712.
8. Isozaki, H., Sakhtemani, R., Abbasi, A., Nikpour, N., Stanzione, M., Oh, S., Langenbucher, A., Monroe, S., Su, W., Cabanos, H.F. *et al.* (2023) Therapy-induced APOBEC3A drives evolution of persistent cancer cells. *Nature*, **620**, 393-401.
9. Law, E.K., Sieuwerts, A.M., LaPara, K., Leonard, B., Starrett, G.J., Molan, A.M., Temiz, N.A., Vogel, R.I., Meijer-van Gelder, M.E., Sweep, F.C. *et al.* (2016) The DNA cytosine deaminase APOBEC3B promotes tamoxifen resistance in ER-positive breast cancer. *Sci Adv*, **2**, e1601737.
10. Alonso de la Vega, A., Temiz, N.A., Tasakis, R., Somogyi, K., Salgueiro, L., Zimmer, E., Ramos, M., Diaz-Jimenez, A., Chocarro, S., Fernandez-Vaquero, M. *et al.* (2023) Acute expression of human APOBEC3B in mice results in RNA editing and lethality. *Genome Biol*, **24**, 267.
11. Cui, Y., Chen, H., Xi, R., Cui, H., Zhao, Y., Xu, E., Yan, T., Lu, X., Huang, F., Kong, P. *et al.* (2020) Whole-genome sequencing of 508 patients identifies key molecular features associated with poor prognosis in esophageal squamous cell carcinoma. *Cell Res*, **30**, 902-913.
12. Law, E.K., Levin-Klein, R., Jarvis, M.C., Kim, H., Argyris, P.P., Carpenter, M.A., Starrett, G.J., Temiz, N.A., Larson, L.K., Durfee, C. *et al.* (2020) APOBEC3A catalyzes mutation and drives carcinogenesis in vivo. *J Exp Med*, **217**.
13. Caswell, D.R., Gui, P., Mayekar, M.K., Law, E.K., Pich, O., Bailey, C., Boumelha, J., Kerr, D.L., Blakely, C.M., Manabe, T. *et al.* (2024) The role of APOBEC3B in lung tumor evolution and targeted cancer therapy resistance. *Nat Genet*, **56**, 60-73.
14. Grasedieck, S., Panahi, A., Jarvis, M.C., Borzooee, F., Harris, R.S., Larijani, M., Avet-Loiseau, H., Samur, M., Munshi, N., Song, K. *et al.* (2024) Redefining high risk multiple myeloma with an APOBEC/Inflammation-based classifier. *Leukemia*.

15. Driscoll, C.B., Schuelke, M.R., Kottke, T., Thompson, J.M., Wongthida, P., Tonne, J.M., Huff, A.L., Miller, A., Shim, K.G., Molan, A. *et al.* (2020) APOBEC3B-mediated corruption of the tumor cell immunopeptidome induces heteroclitic neoepitopes for cancer immunotherapy. *Nat Commun*, **11**, 790.
16. Ruder, U., Denkert, C., Kunze, C.A., Jank, P., Lindner, J., Johrens, K., Kulbe, H., Sehouli, J., Dietel, M., Braicu, E. *et al.* (2019) APOBEC3B protein expression and mRNA analyses in patients with high-grade serous ovarian carcinoma. *Histol Histopathol*, **34**, 405-417.
17. Leonard, B., Starrett, G.J., Maurer, M.J., Oberg, A.L., Van Bockstal, M., Van Dorpe, J., De Wever, O., Helleman, J., Sieuwerts, A.M., Berns, E.M. *et al.* (2016) APOBEC3G Expression Correlates with T-Cell Infiltration and Improved Clinical Outcomes in High-grade Serous Ovarian Carcinoma. *Clin Cancer Res*, **22**, 4746-4755.
18. Boichard, A., Pham, T.V., Yeerna, H., Goodman, A., Tamayo, P., Lippman, S., Frampton, G.M., Tsigelny, I.F. and Kurzrock, R. (2019) APOBEC-related mutagenesis and neo-peptide hydrophobicity: implications for response to immunotherapy. *Oncoimmunology*, **8**, 1550341.
19. Chowell, D., Krishna, S., Becker, P.D., Cocita, C., Shu, J., Tan, X., Greenberg, P.D., Klavinskis, L.S., Blattman, J.N. and Anderson, K.S. (2015) TCR contact residue hydrophobicity is a hallmark of immunogenic CD8+ T cell epitopes. *Proc Natl Acad Sci U S A*, **112**, E1754-1762.
20. Sosman, J.A., Unger, J.M., Liu, P.Y., Flaherty, L.E., Park, M.S., Kempf, R.A., Thompson, J.A., Terasaki, P.I., Sondak, V.K. and Southwest Oncology, G. (2002) Adjuvant immunotherapy of resected, intermediate-thickness, node-negative melanoma with an allogeneic tumor vaccine: impact of HLA class I antigen expression on outcome. *J Clin Oncol*, **20**, 2067-2075.
21. Naranbhai, V., Viard, M., Dean, M., Groha, S., Braun, D.A., Labaki, C., Shukla, S.A., Yuki, Y., Shah, P., Chin, K. *et al.* (2022) HLA-A*03 and response to immune checkpoint blockade in cancer: an epidemiological biomarker study. *Lancet Oncol*, **23**, 172-184.
22. Brown, S.D. and Holt, R.A. (2019) Neoantigen characteristics in the context of the complete predicted MHC class I self-immunopeptidome. *Oncoimmunology*, **8**, 1556080.
23. Chan, K., Roberts, S.A., Klimczak, L.J., Sterling, J.F., Saini, N., Malc, E.P., Kim, J., Kwiatkowski, D.J., Fargo, D.C., Mieczkowski, P.A. *et al.* (2015) An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. *Nat Genet*, **47**, 1067-1072.
24. Thomas, C. and Tampe, R. (2017) Proofreading of Peptide-MHC Complexes through Dynamic Multivalent Interactions. *Front Immunol*, **8**, 65.
25. Kopicki, J.D., Saikia, A., Niebling, S., Gunther, C., Anjanappa, R., Garcia-Alai, M., Springer, S. and Uetrecht, C. (2022) Opening opportunities for K(d) determination and screening of MHC peptide complexes. *Commun Biol*, **5**, 488.
26. Fantini, D., Vidimar, V., Yu, Y., Condello, S. and Meeks, J.J. (2020) MutSignatures: an R package for extraction and analysis of cancer mutational signatures. *Sci Rep*, **10**, 18217.
27. Thorsson, V., Gibbs, D.L., Brown, S.D., Wolf, D., Bortone, D.S., Ou Yang, T.H., Porta-Pardo, E., Gao, G.F., Plaisier, C.L., Eddy, J.A. *et al.* (2018) The Immune Landscape of Cancer. *Immunity*, **48**, 812-830 e814.

28. Hurley, C.K., Kempenich, J., Wadsworth, K., Sauter, J., Hofmann, J.A., Schefzyk, D., Schmidt, A.H., Galarza, P., Cardozo, M.B.R., Dudkiewicz, M. *et al.* (2020) Common, intermediate and well-documented HLA alleles in world populations: CIWD version 3.0.0. *HLA*, **95**, 516-531.
29. Xia, S., Gu, Y., Zhang, H., Fei, Y., Cao, Y., Fang, H., Wang, J., Lin, C., Zhang, H., Li, H. *et al.* (2021) Immune inactivation by APOBEC3B enrichment predicts response to chemotherapy and survival in gastric cancer. *Oncoimmunology*, **10**, 1975386.
30. DiMarco, A.V., Qin, X., McKinney, B.J., Garcia, N.M.G., Van Alsten, S.C., Mendes, E.A., Force, J., Hanks, B.A., Troester, M.A., Owzar, K. *et al.* (2022) APOBEC Mutagenesis Inhibits Breast Cancer Growth through Induction of T cell-Mediated Antitumor Immune Responses. *Cancer Immunol Res*, **10**, 70-86.
31. Haupt, S., Caramia, F., Klein, S.L., Rubin, J.B. and Haupt, Y. (2021) Sex disparities matter in cancer development and therapy. *Nat Rev Cancer*, **21**, 393-407.
32. Hata, A.N. and Larijani, M. (2024) Targeting APOBECs in cancer: It's about timing. *Cancer Cell*, **42**, 497-501.

Starting wildtype human immunopeptidome	The HLA class I immunopeptidome includes all unique 8-11mer peptides (~ 46 million) of which 28% (13,684,528) have a functional affinity (< 500 nM) for at least one of 2910 HLA class I alleles = ~ 1 billion peptide:MHC (pMHC) combinations		
Mutations simulated (C >T, G>A)	APOBEC3A (CTC, GAG, TTC, GAA)		
	APOBEC3B (ATC, GAT, GTC, GAC)		
	AID (AGC, TAC, TGC, GCT, GTA, GCA)		
APOBEC3-mutable immunopeptidome	12,749,500 unique peptides=914,059,070 pMHC (93% of HLA class I immunopeptidome)		
Simulated mutation events	APOBEC3A (1,430,782,669)		
	APOBEC3B (1057937974)		
	AID (2,013,086,955)		
	Total= 4,245,047,967		
Genome coverage	381,534,562 nts (~ 12% of the genome)		
Total mutated pMHC combinations	6,009,308,016		
Amino acid consequence of simulated mutations (pMHC)	APOBEC3A	non-synonymous	869,066,666 (61%)
		synonymous	561,716,003
	APOBEC3B	non-synonymous	571,472,553 (54%)
		synonymous	486,465,421
	AID	non-synonymous	1,320,733,062 (66 %)
		synonymous	692,353,893

Table 1. APOBEC3-mediated mutation simulation and prediction of immunogenicity consequences

APOBEC3A							
Top 10% T-index				Top 10% CIS +/- ratio			
A03-01	A11-23	A24-130	A29-21	A01-19	A23-15	A24-56	A31-35
A03-02	A11-24	A24-133	A29-22	A03-01	A23-16	A24-57	A31-36
A03-04	A11-25	A24-135	A30-07	A03-04	A23-17	A24-58	A31-37
A03-05	A11-26	A24-136	A30-09	A03-05	A23-18	A24-59	A32-07
A03-06	A11-27	A24-137	A30-17	A03-06	A23-20	A24-61	A32-13
A03-07	A11-29	A24-138	A30-26	A03-07	A23-21	A24-62	A32-15
A03-09	A11-30	A24-139	A31-01	A03-09	A23-22	A24-63	A33-08
A03-10	A11-31	A24-141	A31-02	A03-12	A23-23	A24-66	A33-09
A03-12	A11-32	A24-142	A31-03	A03-13	A23-24	A24-68	A33-21
A03-13	A11-33	A24-144	A31-04	A03-14	A23-25	A24-69	A36-02
A03-14	A11-34	A24-15	A31-05	A03-15	A23-26	A24-70	A80-02
A03-15	A11-35	A24-18	A31-06	A03-16	A24-02	A24-71	B08-49
A03-16	A11-37	A24-20	A31-07	A03-17	A24-03	A24-72	B15-197
A03-17	A11-41	A24-22	A31-08	A03-19	A24-04	A24-73	B15-44
A03-20	A11-42	A24-23	A31-09	A03-20	A24-05	A24-74	B15-45
A03-22	A11-43	A24-25	A31-10	A03-22	A24-06	A24-75	B27-18
A03-25	A11-46	A24-26	A31-11	A03-25	A24-08	A24-76	B35-135
A03-26	A11-47	A24-27	A31-12	A03-26	A24-100	A24-77	B46-09
A03-27	A11-48	A24-29	A31-13	A03-27	A24-101	A24-78	B48-08
A03-28	A11-49	A24-33	A31-15	A03-28	A24-102	A24-79	B48-14
A03-29	A11-51	A24-34	A31-16	A03-29	A24-103	A24-80	B48-15
A03-30	A11-53	A24-35	A31-17	A03-30	A24-105	A24-81	C03-39
A03-33	A11-54	A24-37	A31-18	A03-33	A24-106	A24-82	C03-67
A03-34	A11-55	A24-38	A31-19	A03-34	A24-107	A24-85	C07-29
A03-35	A11-56	A24-39	A31-20	A03-35	A24-109	A24-91	
A03-37	A11-57	A24-43	A31-22	A03-37	A24-110	A24-92	
A03-38	A11-58	A24-47	A31-23	A03-38	A24-111	A24-93	
A03-39	A11-59	A24-49	A31-24	A03-39	A24-113	A24-94	
A03-40	A11-61	A24-51	A31-25	A03-40	A24-114	A24-95	
A03-42	A11-62	A24-53	A31-26	A03-42	A24-115	A24-97	
A03-44	A11-63	A24-54	A31-27	A03-44	A24-116	A24-98	
A03-45	A11-64	A24-57	A31-28	A03-45	A24-117	A24-99	
A03-46	A23-01	A24-58	A31-29	A03-46	A24-118	A29-01	
A03-47	A23-03	A24-61	A31-30	A03-47	A24-119	A29-02	
A03-48	A23-04	A24-63	A31-31	A03-48	A24-120	A29-03	
A03-49	A23-05	A24-66	A31-32	A03-49	A24-121	A29-04	
A03-50	A23-06	A24-68	A31-33	A03-50	A24-122	A29-05	
A03-51	A23-10	A24-69	A31-34	A03-51	A24-123	A29-06	
A03-52	A23-12	A24-70	A31-35	A03-52	A24-124	A29-07	
A03-53	A23-13	A24-71	A31-36	A03-53	A24-125	A29-09	
A03-54	A23-14	A24-72	A31-37	A03-54	A24-126	A29-10	
A03-55	A23-15	A24-74	A32-07	A03-55	A24-127	A29-11	
A03-56	A23-16	A24-75	A33-03	A03-56	A24-128	A29-12	
A03-58	A23-17	A24-76	A33-06	A03-58	A24-13	A29-14	
A03-59	A23-18	A24-78	A33-08	A03-59	A24-130	A29-15	
A03-60	A23-20	A24-79	A33-09	A03-60	A24-133	A29-16	
A03-61	A23-21	A24-80	A33-12	A03-61	A24-134	A29-17	
A03-62	A23-22	A24-81	A33-14	A03-62	A24-135	A29-18	
A03-63	A23-23	A24-85	A33-15	A03-63	A24-136	A29-19	
A03-64	A23-24	A24-91	A33-18	A03-64	A24-137	A29-20	
A03-65	A23-25	A24-92	A33-20	A03-65	A24-138	A29-21	
A03-67	A23-26	A24-94	A33-21	A03-67	A24-139	A30-29	
A03-70	A24-02	A24-95	A33-23	A03-70	A24-140	A31-01	
A03-71	A24-03	A24-97	A33-25	A03-71	A24-141	A31-02	
A03-72	A24-05	A24-98	A33-26	A03-72	A24-142	A31-05	
A03-73	A24-10	A24-99	A33-29	A03-74	A24-144	A31-07	
A03-74	A24-100	A29-01	A33-30	A03-75	A24-17	A31-08	
A03-75	A24-101	A29-02	A33-31	A03-77	A24-18	A31-09	
A03-76	A24-102	A29-03	A68-10	A03-78	A24-19	A31-10	
A03-77	A24-103	A29-04	A68-13	A03-79	A24-20	A31-11	
A03-78	A24-105	A29-05	A68-14	A03-80	A24-21	A31-13	
A03-79	A24-107	A29-06	A68-30	A03-81	A24-23	A31-16	
A03-80	A24-110	A29-07	B15-68	A03-82	A24-25	A31-17	
A03-81	A24-111	A29-09	C14-15	A11-04	A24-26	A31-18	
A11-01	A24-113	A29-10		A11-25	A24-27	A31-19	
A11-02	A24-114	A29-11		A11-27	A24-29	A31-20	
A11-03	A24-115	A29-12		A11-36	A24-32	A31-22	
A11-04	A24-116	A29-13		A11-43	A24-33	A31-23	
A11-05	A24-117	A29-14		A23-01	A24-34	A31-24	
A11-07	A24-118	A29-15		A23-02	A24-35	A31-25	
A11-09	A24-120	A29-11		A23-03	A24-37	A31-26	
A11-12	A24-122	A29-12		A23-04	A24-38	A31-27	
A11-13	A24-123	A29-13		A23-05	A24-39	A31-28	
A11-14	A24-124	A29-14		A23-06	A24-41	A31-29	
A11-15	A24-125	A29-15		A23-09	A24-42	A31-30	
A11-16	A24-126	A29-16		A23-10	A24-43	A31-31	
A11-19	A24-127	A29-17		A23-12	A24-47	A31-32	
A11-20	A24-128	A29-18		A23-13	A24-49	A31-33	
A11-22	A24-13	A29-20		A23-14	A24-54	A31-34	

APOBEC3B							
Top 10% T- index				Top 10% CIS +/- ratio			
A03-07	A24-94	A68-16	B44-46	A11-27	B40-104	B44-09	B44-99
A03-12	A29-01	A68-19	B44-89	A11-43	B40-105	B44-100	B45-01
A03-42	A29-02	A68-21	B44-91	A23-01	B40-107	B44-101	B45-02
A03-50	A29-03	A68-22	B45-09	A23-03	B40-108	B44-102	B45-03
A03-64	A29-04	A68-24	B48-02	A23-04	B40-113	B44-103	B45-04
A03-67	A29-06	A68-25		A23-05	B40-120	B44-104	B45-05
A03-72	A29-09	A68-29		A23-06	B40-122	B44-105	B45-07
A11-01	A29-10	A68-30		A23-10	B40-123	B44-106	B45-08
A11-02	A29-11	A68-32		A23-14	B40-125	B44-107	B45-09
A11-03	A29-12	A68-33		A23-15	B40-126	B44-109	B45-10
A11-04	A29-14	A68-35		A23-16	B40-128	B44-11	B45-11
A11-05	A29-15	A68-37		A23-17	B40-130	B44-110	B45-12
A11-07	A29-16	A68-38		A23-18	B40-134	B44-12	B49-02
A11-09	A29-17	A68-39		A23-20	B40-135	B44-13	
A11-12	A29-18	A68-41		A23-21	B40-138	B44-14	
A11-13	A29-21	A68-43		A23-22	B40-139	B44-15	
A11-14	A29-22	A68-45		A23-23	B40-140	B44-16	
A11-15	A30-17	A68-47		A23-24	B40-141	B44-17	
A11-16	A31-01	A68-50		A23-25	B40-146	B44-20	
A11-20	A31-02	A68-52		A23-26	B40-147	B44-21	
A11-23	A31-03	B13-10		A24-03	B40-18	B44-22	
A11-24	A31-04	B15-123		A24-10	B40-20	B44-24	
A11-25	A31-05	B15-131		A24-125	B40-24	B44-26	
A11-26	A31-06	B15-151		A24-13	B40-25	B44-27	
A11-27	A31-07	B15-185		A24-23	B40-27	B44-28	
A11-29	A31-08	B15-91		A24-33	B40-28	B44-29	
A11-30	A31-09	B18-12		A24-34	B40-29	B44-30	
A11-31	A31-10	B35-28		A24-57	B40-30	B44-32	
A11-32	A31-11	B35-63		A24-92	B40-33	B44-33	
A11-33	A31-12	B40-02		A24-94	B40-34	B44-34	
A11-34	A31-13	B40-03		A29-01	B40-35	B44-35	
A11-35	A31-15	B40-04		A29-02	B40-38	B44-36	
A11-37	A31-16	B40-09		A29-03	B40-39	B44-38	
A11-41	A31-17	B40-104		A29-04	B40-42	B44-40	
A11-42	A31-18	B40-105		A29-06	B40-45	B44-41	
A11-43	A31-19	B40-107		A29-09	B40-47	B44-42	
A11-46	A31-20	B40-11		A29-10	B40-48	B44-43	
A11-47	A31-22	B40-111		A29-11	B40-51	B44-44	
A11-48	A31-23	B40-114		A29-12	B40-52	B44-45	
A11-49	A31-24	B40-115		A29-14	B40-54	B44-46	
A11-51	A31-25	B40-119		A29-15	B40-55	B44-48	
A11-53	A31-26	B40-120		A29-16	B40-56	B44-50	
A11-54	A31-27	B40-121		A29-17	B40-57	B44-51	
A11-55	A31-28	B40-122		A29-18	B40-59	B44-53	
A11-56	A31-29	B40-14		A29-20	B40-60	B44-54	
A11-57	A31-30	B40-145		A29-21	B40-62	B44-55	
A11-58	A31-31	B40-18		A31-07	B40-63	B44-59	
A11-60	A31-32	B40-20		A31-18	B40-64	B44-63	
A11-61	A31-34	B40-24		A31-30	B40-65	B44-64	
A11-62	A31-35	B40-27		A33-03	B40-66	B44-65	
A11-63	A31-36	B40-28		A33-06	B40-67	B44-66	
A11-64	A31-37	B40-29		A33-12	B40-69	B44-67	
A23-01	A32-07	B40-35		A33-14	B40-72	B44-68	
A23-03	A33-03	B40-37		A33-15	B40-74	B44-69	
A23-04	A33-06	B40-38		A33-18	B40-76	B44-70	
A23-05	A33-08	B40-39		A33-20	B40-78	B44-71	
A23-06	A33-09	B40-45		A33-23	B40-79	B44-72	
A23-10	A33-12	B40-51		A33-25	B40-81	B44-73	
A23-14	A33-14	B40-52		A33-26	B40-82	B44-74	
A23-15	A33-15	B40-56		A33-29	B40-84	B44-75	
A23-16	A33-18	B40-57		A33-30	B40-87	B44-76	
A23-17	A33-20	B40-59		A33-31	B40-88	B44-77	
A23-18	A33-21	B40-60		A68-45	B40-89	B44-78	
A23-20	A33-22	B40-64		B15-91	B40-91	B44-79	
A23-21	A33-23	B40-78		B18-03	B40-92	B44-80	
A23-22	A33-25	B40-82		B18-06	B40-94	B44-81	
A23-23	A33-26	B40-85		B18-12	B40-97	B44-82	
A23-24	A33-29	B40-89		B18-13	B40-99	B44-83	
A23-25	A33-30	B40-90		B18-14	B41-01	B44-84	
A23-26	A33-31	B40-91		B18-21	B41-05	B44-85	
A24-03	A34-02	B40-94		B18-30	B41-06	B44-86	
A24-10	A34-03	B40-97		B40-01	B41-07	B44-87	
A24-125	A34-04	B40-99		B40-02	B41-09	B44-88	
A24-13	A68-01	B44-09		B40-03	B41-12	B44-89	
A24-23	A68-09	B44-17		B40-04	B44-02	B44-92	
A24-33	A68-10	B44-29		B40-09	B44-03	B44-93	
A24-51	A68-12	B44-42		B40-100	B44-04	B44-94	
A24-57	A68-13	B44-43		B40-101	B44-05	B44-95	
A24-92	A68-14	B44-44		B40-102	B44-07	B44-98	

APOBEC3A & APOBEC3B			
Top 10% T-index		Top 20% CIS +/-	
A03-02	A24-92	A01-19	A31-30
A03-07	A24-94	A03-12	A31-31
A03-10	A29-01	A03-67	A31-32
A03-12	A29-02	A03-72	A31-34
A03-42	A29-03	A11-04	A31-35
A03-50	A29-04	A11-27	A31-36
A03-64	A29-06	A11-43	A31-37
A03-67	A29-09	A23-01	A33-09
A03-72	A29-10	A23-03	A33-21
A03-73	A29-11	A23-04	
A03-76	A29-12	A23-05	
A11-01	A29-14	A23-06	
A11-02	A29-15	A23-10	
A11-03	A29-16	A23-13	
A11-04	A29-17	A23-14	
A11-05	A29-18	A23-15	
A11-07	A29-21	A23-16	
A11-09	A30-17	A23-17	
A11-12	A31-01	A23-18	
A11-13	A31-02	A23-20	
A11-14	A31-03	A23-21	
A11-15	A31-04	A23-22	
A11-16	A31-05	A23-23	
A11-20	A31-06	A23-24	
A11-23	A31-07	A23-25	
A11-24	A31-08	A23-26	
A11-25	A31-09	A24-03	
A11-27	A31-10	A24-106	
A11-29	A31-11	A24-125	
A11-30	A31-12	A24-13	
A11-32	A31-13	A24-138	
A11-33	A31-15	A24-18	
A11-34	A31-16	A24-23	
A11-35	A31-17	A24-33	
A11-37	A31-18	A24-34	
A11-41	A31-19	A24-92	
A11-42	A31-20	A24-94	
A11-43	A31-22	A29-01	
A11-46	A31-23	A29-02	
A11-47	A31-24	A29-03	
A11-48	A31-25	A29-04	
A11-49	A31-26	A29-05	
A11-51	A31-27	A29-06	
A11-53	A31-28	A29-09	
A11-54	A31-29	A29-10	
A11-55	A31-30	A29-11	
A11-56	A31-31	A29-12	
A11-57	A31-32	A29-14	
A11-58	A31-34	A29-15	
A11-61	A31-35	A29-16	
A11-62	A31-36	A29-17	
A11-63	A31-37	A29-18	
A11-64	A32-07	A29-20	
A23-01	A33-08	A29-21	
A23-03	A33-09	A30-29	
A23-04	A33-18	A31-01	
A23-05	A33-21	A31-02	
A23-06	A68-10	A31-03	
A23-10	A68-13	A31-04	
A23-14	A68-14	A31-05	
A23-15	A68-30	A31-06	
A23-16		A31-07	
A23-17		A31-08	
A23-18		A31-09	
A23-20		A31-10	
A23-21		A31-11	
A23-22		A31-13	
A23-23		A31-16	
A23-24		A31-17	
A23-25		A31-18	
A23-26		A31-19	
A24-03		A31-20	
A24-10		A31-22	
A24-125		A31-23	
A24-13		A31-25	
A24-23		A31-26	
A24-33		A31-27	
A24-51		A31-28	
A24-57		A31-29	

Table 2. List of HLA alleles with the highest anti-tumor scores as a result of APOBEC3-mediated mutations of their immunopeptidome, according to the combined immunogenicity score (CIS), or triple index (TI) rankings. The HLA alleles colored green are shared across the top CIS and TI rankings, and alleles highlighted in gray are classified as common, intermediate or well documented (CIWD).

Figure 1. Prediction of the impact of APOBEC3-mediated mutations on the immunogenicity of the human immunopeptidome. The human MHC (HLA) class I immunopeptidome is comprised of millions of 8-11mer peptides that can potentially be bound and presented by several thousand human HLA class I molecules, making several billion peptide:MHC (pMHC) combinations that can potentially be presented for recognition by T cells. For each peptide epitope in the immunopeptidome, we recovered the genomic coding sequence and simulated mutations at favored hotspot motifs of APOBEC3A or APOBEC3B, and as control, AID. Following mutation simulation (scale described in table 1), we evaluated the predicted impact of each mutation event by comparing each wild-type (WT): mutant (MUT) pair, for the impact of the mutation on pMHC binding affinity (lower left), and pMHC + T cell receptor (TCR) binding affinity (lower right). For the pMHC analyses (lower left), we evaluated whether the percentile rank pMHC affinity was altered using NetMHC 4. We used two criteria to account for small inter-category movements: a percentile rank value change > 0.3 for a 1-category shift and a significant shift between the two extremes of a category (separated by distribution Z score > 0.5). Mutational events were thus binned into the following four categories "-": lost or decreased immunogenicity=pro-tumour, "+": gained immunogenicity=anti-tumour, NC: no change in immunogenicity, Synonymous: silent. The +/- ratio of +(enhanced immunogenicity=anti-tumor mutational events) / -(lost or diminished immunogenicity=pro-tumor mutational events) was thus calculated for the total mutations induced by each APOBEC3 enzyme, as well as for the immunopeptidome of each HLA allele. For TCR binding of the pMHC complex (lower right), we took into account the amino acid property changes in each WT:MUT pair which are predicted to impact TCR binding, in addition to the pMHC affinity changes. Changes in the hydrophobicity and polarity of amino acid substitutions at each position in each WT:MUT pair, from WT to MUT were determined using the Kyte-Doolittle and Grantham scales. To illustrate the magnitude of the changes, an index based on the quartiles of hydrophobicity and polarity substitution at each position was employed. Peptide positions 4 and 6 received a higher weight in score than positions 3,5,7, since the former are more dominant TCR contacts. In this manner, each mutational event was given a score that combines pMHC binding and TCR recognition. For the immunopeptidome of each HLA allele, the number of simulated mutation events in the pool of its restricted epitopes that resulted in a change in immunogenicity towards either a potential "+" (enhanced immunogenicity=anti-tumor mutational events) or "-" (lost or diminished immunogenicity=pro-tumor mutational events) was tallied to obtain the probability of an APOBEC3-mediated mutation to instigate either an anti-tumor or pro-tumor event. The +/- ratio of + (anti-tumor mutational events)/- (pro-tumor mutational events) was thus calculated, considering changes in HLA and TCR binding affinities. In parallel to the mutation count of + and - events, the numerical scores which reflect the magnitude of changes in immunogenicity for pMHC and TCR binding for each mutational event were used to assess the magnitude of each change of immunogenicity event in either the "+" anti-tumor or "-" pro-tumor direction. For each HLA allele, we then calculated a cumulative immunogenicity score (CIS) which is the sum of the individual mutational event's score values for the HLA and TCR immunogenicity changes, in either the anti-tumor (+ values) or pro-tumor (- values) direction. The +/- ratio of these values indicates the anti-tumor/pro-tumor potential of each HLA allele, considering the number of mutational events that cause an immunogenicity change in either the anti- or pro-tumor directions, and also the magnitude of each

immunogenicity change event as determined by the type and position of amino acid changes in the epitope.

Figure 2. Impact of simulated APOBEC3-mediated mutations on changing the MHC binding affinity of the immunopeptidome. A. The left panel shows the total mutation events of each enzyme as broken down into the four categories of "-": decreased pMHC binding = pro-tumour, "+": enhanced pMHC binding = anti-tumour, "NC": no change in pMHC binding affinity, and "synonymous": silent mutations. The y-axis shows the raw count of mutational events. The middle panel shows the four category relative proportions normalized (1 = 100%) for each enzyme. The right panel shows the +/total and -/total mutational counts for each enzyme, as an average or median of the value determined for the immunopeptidomes of each of the 2910 HLA alleles. The "+" or "-" mutational event counts was normalized to the total mutational count, for each of the 2910 HLA alleles, and the average and medians of this relative proportion across all HLA alleles is plotted. **B.** The top panel shows the proportion of the 4 categories for all 2910 HLA alleles, to demonstrate the variation among alleles. The bottom panel shows the 2910 HLA alleles ranked by +/- ratio of their immunopeptidomes as mutated by A3A, A3B and AID. For each HLA allele, the relative count of "+" (anti-tumour)/total MUT events was divided by "-" (pro-tumour)/total MUT events to obtain a +/- ratio which is plotted for each of the 2910 HLA alleles.

Figure 3. APOBEC3-mediated mutations alter peptide epitope immunogenicity by impacting pMHC binding and T cell receptor recognition. A. After ascribing a combined immunogenicity score (CIS) which accounts for changes in both pMHC binding affinity as well as TCR contact residues, for each mutational event, we tallied the count of events that produced a "-"(diminished immunogenicity=pro-tumour potential) vs. "+" (enhanced immunogenicity=anti-tumour potential) in the immunopeptidome of each HLA allele. The anti-tumour event counts were compared to the pro-tumour event counts for the total immunopeptidome as mutated by each enzyme (top left). The +/- ratio averages, medians and highest values of CIS even counts across the 2910 HLA alleles were derived (top right). The CIS values of all "+" vs "-" events for each HLA allele were added and the total magnitude of the CIS value in the "+" anti-tumour direction was compared to the total CIS value in the "-" pro-tumour direction, reflecting both the event count and the magnitude of immunogenicity changes (bottom left). The +/- ratio averages, medians, and highest CIS values across the 2910 HLA alleles were derived (bottom right). **B.** For each of the 2910 HLA alleles, we tallied the count of CIS events (top row) or the total CIS value (bottom row) in the "+ " and "-" directions and calculated a +/- ratio, based on which the HLA alleles were ranked.

Figure 4. The HLA class I-dependence of the impact of APOBEC3-mediated mutations on the immunopeptidome's immunogenicity is a predictor of survival in APOBEC3-mutated tumors. A. The left panels shows the distribution of the 2910 HLA alleles based on CIS value +/- ratios of their A3A-mutated vs. A3B-mutated immuopeptidomes. The right panel shows the distribution of HLA alleles based on the triple index values of their A3A-mutated vs. A3B-mutated immunopeptidomes. The triple index integrates the +/- rank of each HLA allele relative to the average of all alleles, across the three measures of pMHC binding affinity, CIS event count, and CIS value. **B.** Using the CIS event count and CIS value +/- ratios of each HLA allele, a combined +/- score for each individual in the TCGA dataset was devised which takes into account the +/- scores of each of the six HLA class I A, B and C alleles in an individual's HLA haplotype. Patients were then ranked based on the cumulative +/- index of their HLA haplotype. The "High index" quantile represents patients that carry an HLA class I haplotype at the highest end of the cumulative +/- scores and therefore most potential for enhanced immunogenicity of the HLA class I immunopeptidome after APOBEC3-mediated mutations. At the other end of the spectrum, patients in the "Low index" quantile carry an HLA class I haplotype that falls at the lowest range of +/- scores, and therefore the most potential for lost or reduced immunogenicity and immune escape after APOBEC3-mediated mutations. Survival was then examined based for the "High index" vs. "Low index" patients for tumors that carry an APOBEC3 mutation signature (top row), and, as control, for those that do not carry an APOBEC3 mutational signature (bottom row). The left, middle and right panels denote different "High index" and "Low index" HLA haplotypes, based on the combined +/- scores of A3A+A3B, or A3B or A3A, alone. **C.** The data shown in panel B was further stratified based on biological sex in addition to the carriage of the "High index" and "Low index" HLA haplotypes.

Figure 5. CD8+ T cell marker analyses of tumors stratified by HLA haplotype, APOBEC3 mutational signature, and survival. **A.** Tumors were stratified based on the carriage of High-index vs. Low-index HLA class I haplotype for A3A+A3B (left panel), A3B (middle panel), or A3A (right panel), and analyzed for CD8+ T cell infiltration according to immune landscape data. **B.** Tumors were stratified based on the carriage of High-index vs. Low-index HLA class I haplotype for A3A+A3B, as well as the presence of APOBEC3 signature mutations, followed by differential gene expression analyses to examine expression levels of genes that are markers of CD8+ T cell infiltration (surface markers of CD8+T cells), activation (cytokines, transcription factors), tumor killing by CD8+ T cells (granzyme, perforin, Fas Ligand) and T cell exhaustion (PD Ligand, Lag3, Icos). Differentially expressed CD8+ T cell markers with Padj <0.05 were plotted. The left panel shows a comparison of differentially expressed markers in either the High index quantile relative to all samples (green), or the Low index quantile relative to all samples (red). The right panel shows the enrichment ratios of the High index relative to the Low index quantile.

Figure 1

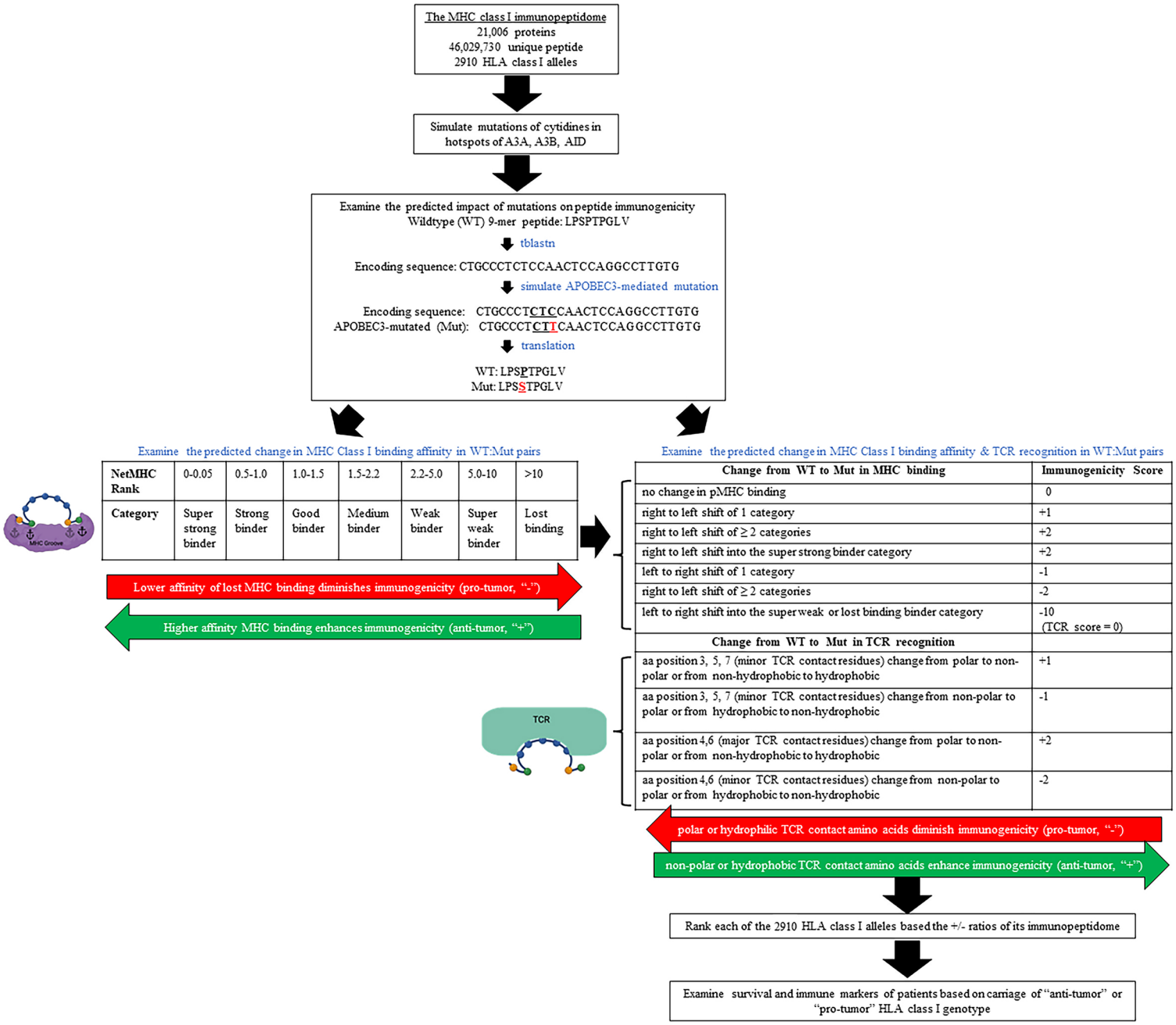
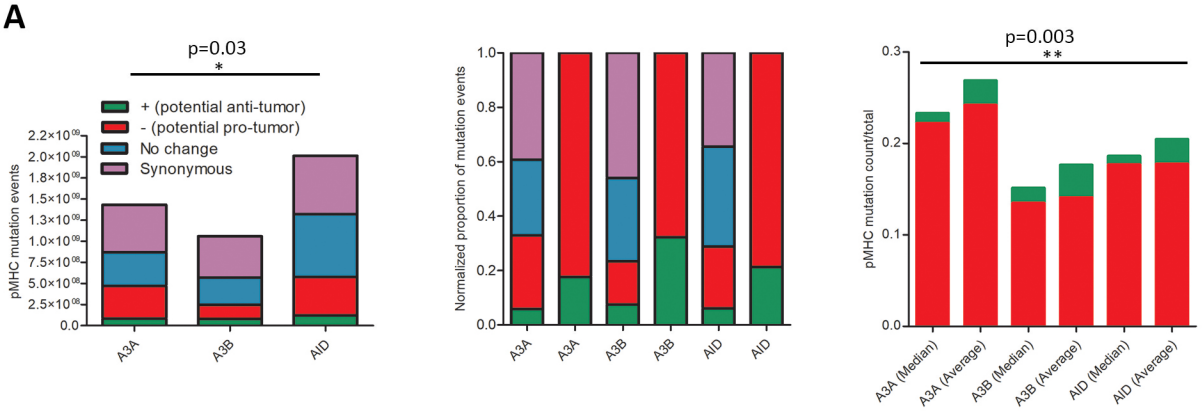

The MHC class I immunopeptidome
21,006 proteins
46,029,730 unique peptide
2910 HLA class I alleles
Simulate mutations of cytidines in hotspots of A3A, A3B, AID
Examine the predicted impact of mutations on peptide immunogenicity
Wildtype (WT) 9-mer peptide: LPSPTPGLV
tblastn
Encoding sequence: CTGCCCTCTCCAACTCCAGGCCTTGTG
simulate APOBEC3-mediated mutation
Encoding sequence: CTGCCCTCTCCAACTCCAGGCCTTGTG
APOBEC3-mutated (Mut): CTGCCCTCTTCAACTCCAGGCCTTGTG
translation
WT: LPSPTPGLV
Mut: LPSSTPGLV
Examine the predicted change in MHC Class I binding affinity in WT:Mut pairs
NetMHC Rank | 0-0.05 | 0.5-1.0 | 1.0-1.5 | 1.5-2.2 | 2.2-5.0 | 5.0-10 | >10
Category | Super strong binder | Strong binder | Good binder | Medium binder | Weak binder | Super weak binder | Lost binding
MHC Groove
Lower affinity of lost MHC binding diminishes immunogenicity (pro-tumor, "-")
Higher affinity MHC binding enhances immunogenicity (anti-tumor, "+")
Examine the predicted change in MHC Class I binding affinity & TCR recognition in WT:Mut pairs
Change from WT to Mut in MHC binding | Immunogenicity Score
no change in pMHC binding | 0
right to left shift of 1 category | +1
right to left shift of ≥ 2 categories | +2
right to left shift into the super strong binder category | +2
left to right shift of 1 category | -1
right to left shift of ≥ 2 categories | -2
left to right shift into the super weak or lost binding binder category | -10 (TCR score = 0)
Change from WT to Mut in TCR recognition
aa position 3, 5, 7 (minor TCR contact residues) change from polar to non-polar or from non-hydrophobic to hydrophobic | +1
aa position 3, 5, 7 (minor TCR contact residues) change from non-polar to polar or from hydrophobic to non-hydrophobic | -1
aa position 4,6 (major TCR contact residues) change from polar to non-polar or from non-hydrophobic to hydrophobic | +2
aa position 4,6 (minor TCR contact residues) change from non-polar to polar or from hydrophobic to non-hydrophobic | -2
TCR
polar or hydrophilic TCR contact amino acids diminish immunogenicity (pro-tumor, "-")
non-polar or hydrophobic TCR contact amino acids enhance immunogenicity (anti-tumor, "+")
Rank each of the 2910 HLA class I alleles based the +/- ratios of its immunopeptidome
Examine survival and immune markers of patients based on carriage of "anti-tumor" or "pro-tumor" HLA class I genotype

Figure 2

A


p=0.03
*
+ (potential anti-tumor)
- (potential pro-tumor)
No change
Synonymous
pMHC mutation events
A3A
A3B
AID
Normalized proportion of mutation events
p=0.003
**
pMHC mutation count/total
A3A (Median)
A3A (Average)
A3B (Median)
A3B (Average)
AID (Median)
AID (Average)

B

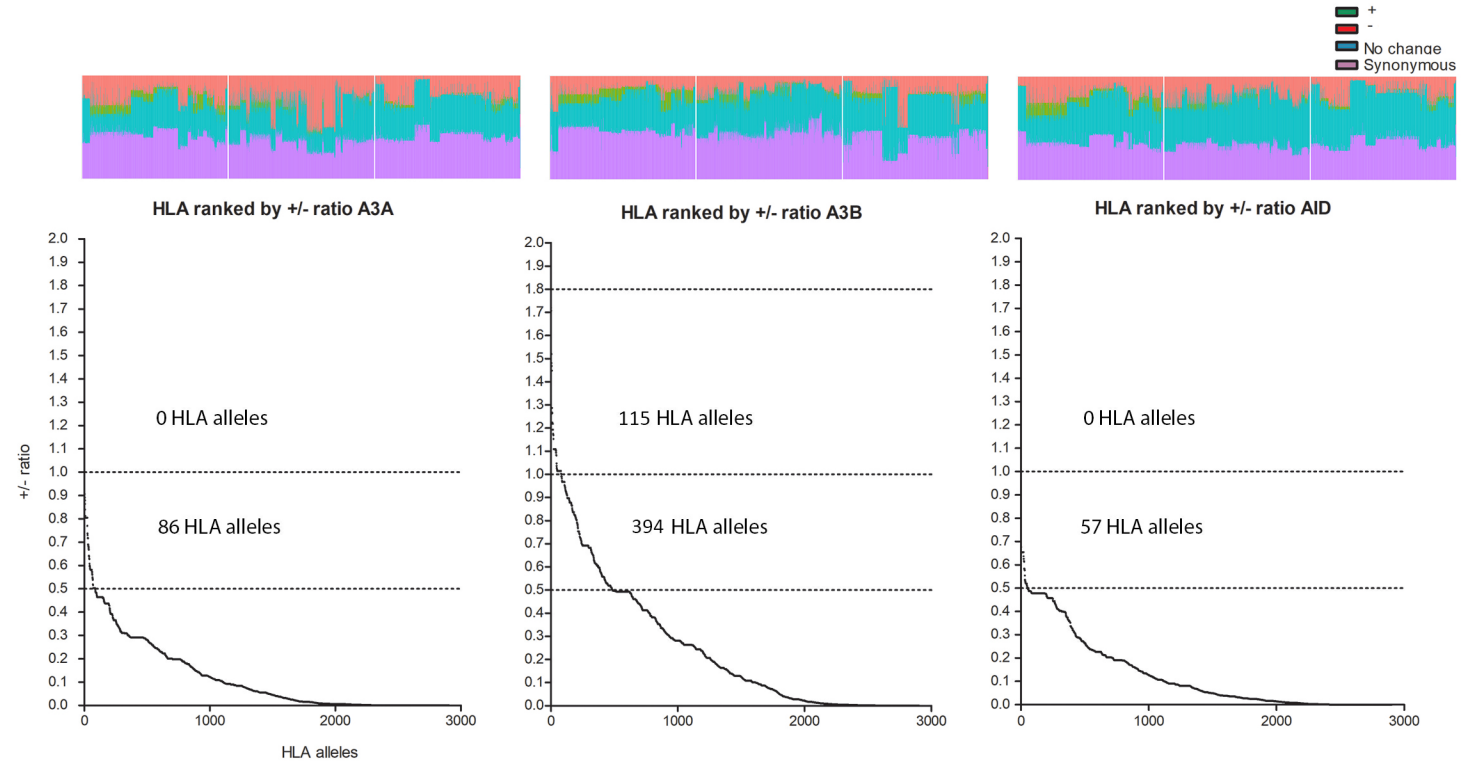

+
-
No change
Synonymous
HLA ranked by +/- ratio A3A
HLA ranked by +/- ratio A3B
HLA ranked by +/- ratio AID
+/- ratio
0 HLA alleles
86 HLA alleles
115 HLA alleles
394 HLA alleles
0 HLA alleles
57 HLA alleles
HLA alleles

Figure 3

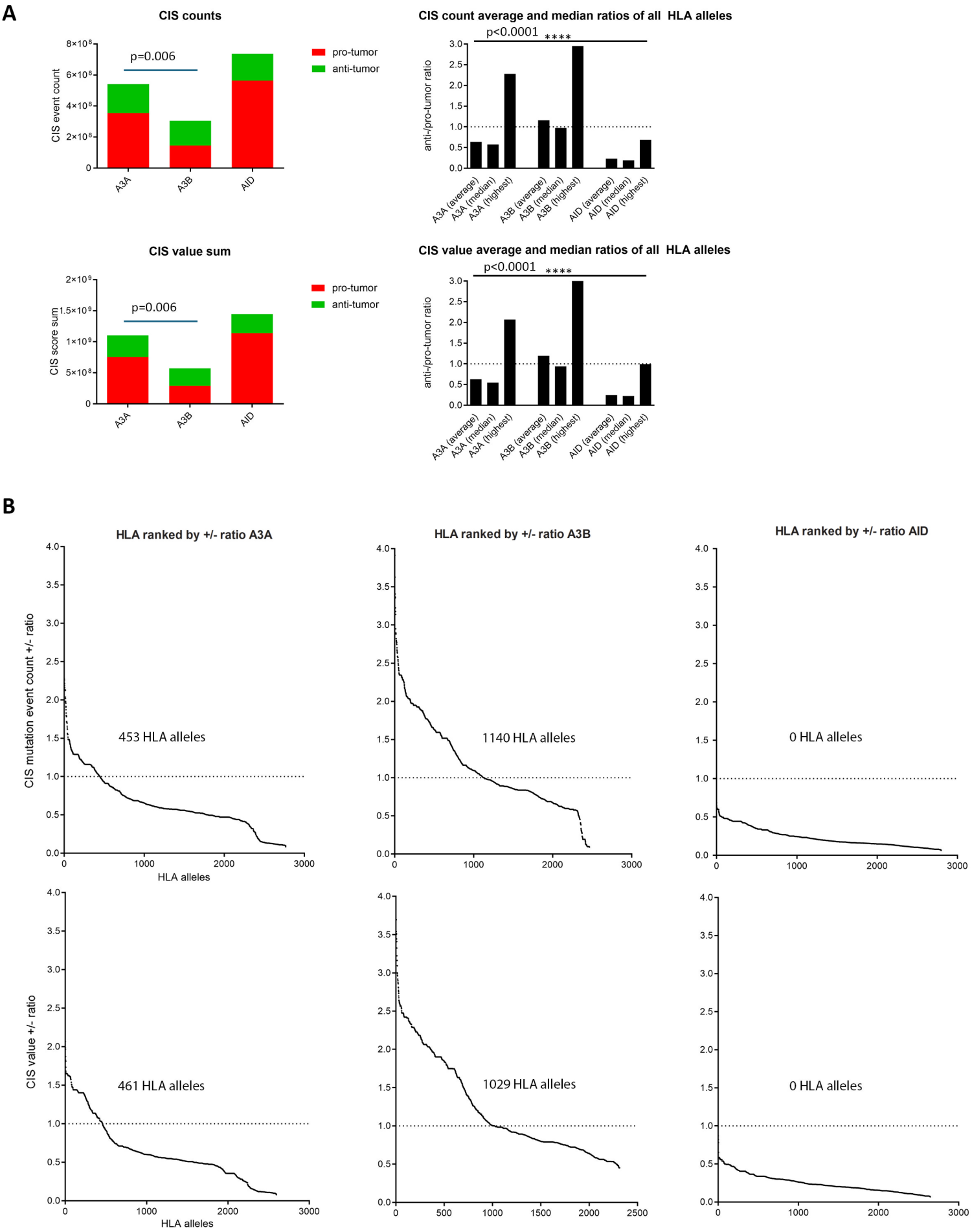
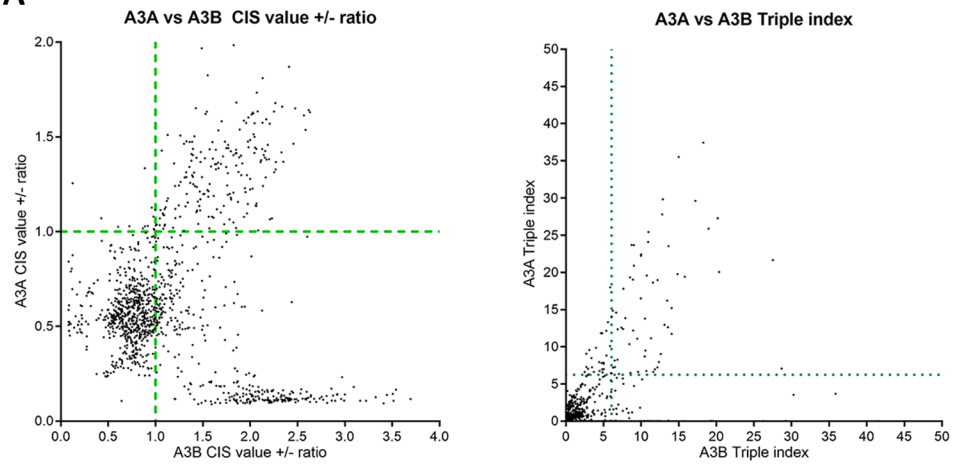
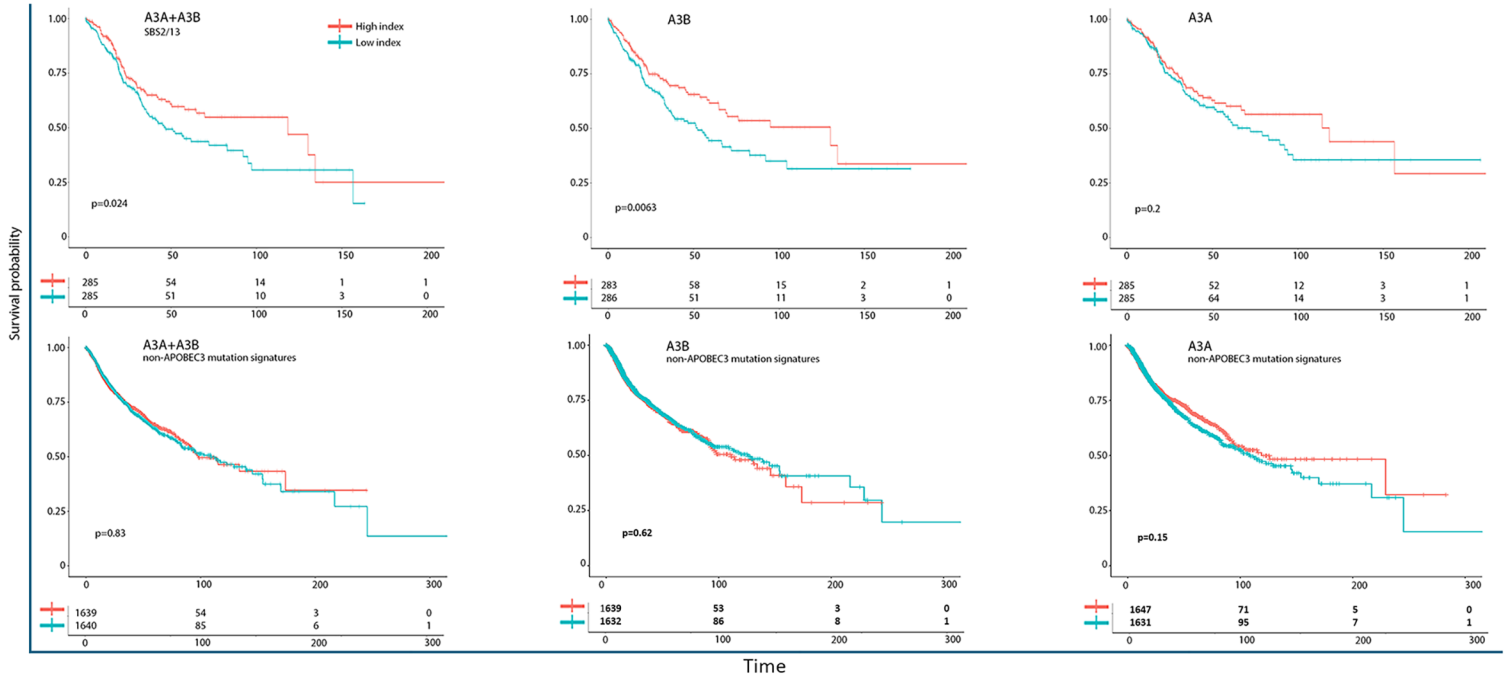

A
CIS counts
p=0.006
pro-tumor
anti-tumor
CIS event count
A3A
A3B
AID
CIS count average and median ratios of all HLA alleles
p<0.0001
****
anti-/pro-tumor ratio
A3A (average)
A3A (median)
A3A (highest)
A3B (average)
A3B (median)
A3B (highest)
AID (average)
AID (median)
AID (highest)
CIS value sum
p=0.006
pro-tumor
anti-tumor
CIS score sum
CIS value average and median ratios of all HLA alleles
p<0.0001
****
B
HLA ranked by +/- ratio A3A
HLA ranked by +/- ratio A3B
HLA ranked by +/- ratio AID
CIS mutation event count +/- ratio
453 HLA alleles
1140 HLA alleles
0 HLA alleles
HLA alleles
CIS value +/- ratio
461 HLA alleles
1029 HLA alleles
0 HLA alleles
HLA alleles

Figure 4

A


A3A vs A3B CIS value +/- ratio
A3A CIS value +/- ratio
A3B CIS value +/- ratio
A3A vs A3B Triple index
A3A Triple index
A3B Triple index

B


A3A+A3B
SBS2/13
High index
Low index
p=0.024
A3B
p=0.0063
A3A
p=0.2
A3A+A3B
non-APOBEC3 mutation signatures
p=0.83
A3B
non-APOBEC3 mutation signatures
p=0.62
A3A
non-APOBEC3 mutation signatures
p=0.15
Survival probability
Time

C

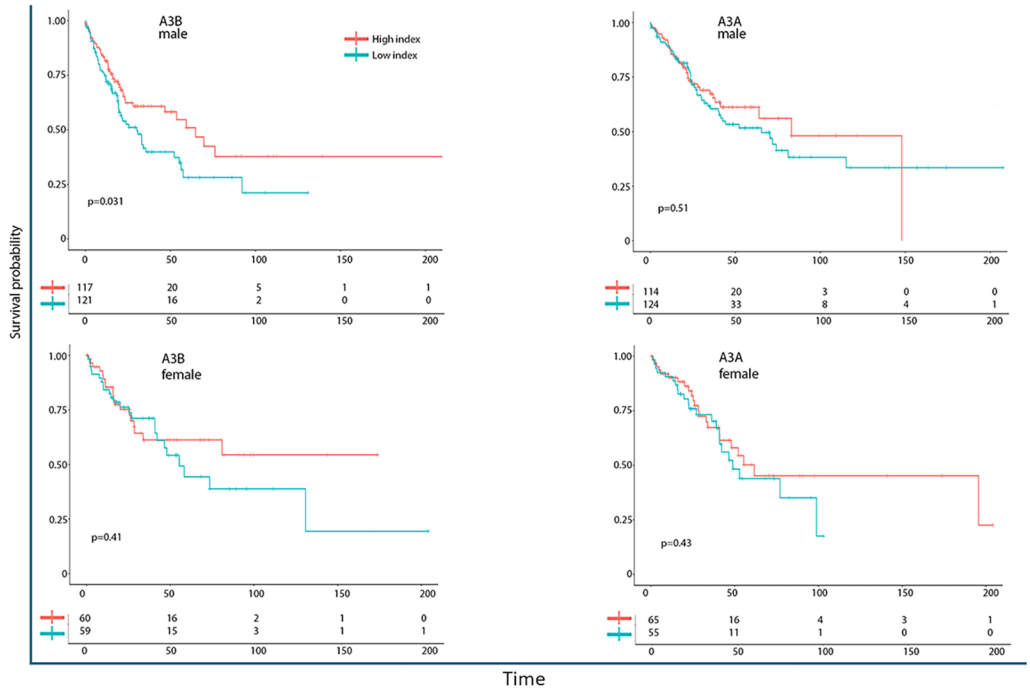
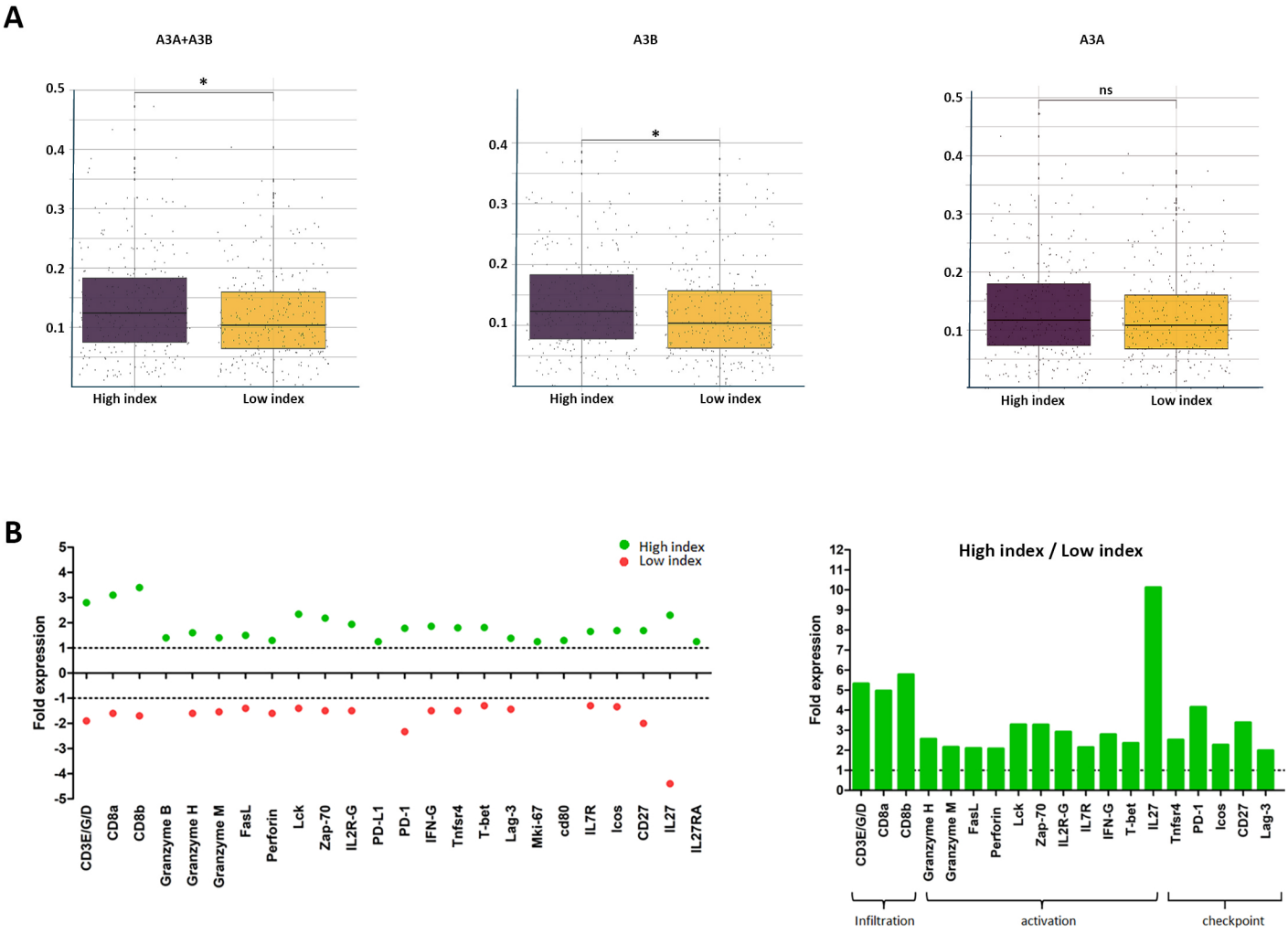

A3B
male
High index
Low index
p=0.031
A3A
male
p=0.51
A3B
female
p=0.41
A3A
female
p=0.43
Survival probability
Time

Figure 5


A
A3A+A3B
A3B
A3A
*
*
ns
High index
Low index
B
High index
Low index
Fold expression
CD3E/G/D
CD8a
CD8b
Granzyme B
Granzyme H
Granzyme M
FasL
Perforin
Lck
Zap-70
IL2R-G
PD-L1
PD-1
IFN-G
Tnfsr4
T-bet
Lag-3
Mki-67
cd80
IL7R
Icos
CD27
IL27
IL27RA
High index / Low index
Infiltration
activation
checkpoint