The landscapes of T cell epitope immunogenicity in sequence space

Authors: M. Ogishi,[1]* H. Yotsuyanagi[1]

Affiliations:

[1] Division of Infectious Diseases and Applied Immunology, The Institute of Medical Sciences Research Hospital, The University of Tokyo, Tokyo, 108-0071, Japan.

*Corresponding authors: M. Ogishi (oogishi-tky@umin.ac.jp)

One Sentence Summary: Immunogenicity prediction of MHC-I and MHC-II-restricted T cell epitopes from simulated TCR-peptide interaction profiles

Abstract: T cells immunity is triggered by epitopes bound to major histocompatibility complex (MHC). However, determinants of epitope immunogenicity remain poorly understood. Here, we comprehensively analyzed 21,162 and 31,693 MHC-I and MHC-II-restricted human peptides, respectively, through a computational framework mimicking the molecular scanning by T cell receptor (TCR) repertoire. Our analysis revealed that "immune transition," acquisition or loss of immunogenicity after single amino acid substitution, is a critical concept in untangling immunogenicity. With partially overlapping 32 and 27 features for MHC-I and MHC-II, respectively, our framework accurately predicted transitional status (AUC: 92% and 94% for MHC-I and MHC-II, respectively) and immunogenicity for non-transitional (AUC: 92% and 95%) and transitional (AUC: 85% and 93%) peptides. Furthermore, neoepitope burden curtailed by 76% according to immunogenicity-based dissimilarity-to-self criteria predicted survival after checkpoint blockade immunotherapy (N=197; $P = 7.7\times10^{-5}$). Collectively, our analysis delineates the landscapes of T cell epitopes, thus expediting translational efforts toward precision immunotherapy.

[Main Text]

Introduction

T cell epitopes bound to major histocompatibility complex [MHC; also called the human leukocyte antigen (HLA) in humans] molecules, when recognized by T cell receptor (TCR), activate T cells to initiate subsequent immunological orchestration (*1–3*). MHC class I (MHC-I) molecules typically present 8- to 11-aa peptides generated through proteasomal cleavage of intracellular proteins to activate $CD8^+$ cytotoxic T lymphocytes (CTLs), whereas MHC class II (MHC-II) molecules with an open-ended binding groove accommodate peptides of more variable length to activate $CD4^+$ T helper (Th) cells. Development of vaccines and other targeted immunotherapies necessarily involves exploration of potent T cell epitopes, which is accelerated by insights into the basis of molecular scanning of MHC-bound peptides by host TCR repertoire. However, in contrast to the success in sequence-based prediction of MHC-I and MHC-II binding (*4*, *5*), factors demarcating immunogenic epitopes from non-immunogenic MHC binders remain elusive. Although extensively studied, MHC binding affinity and stability most often do not predict immunogenicity (*6–8*). Instead, previous studies have attempted MHC-I immunogenicity prediction from position-weighted amino acid properties such as hydrophobicity and existence of aromatic side chains (*9*, *10*). In contrast to MHC-I, there is no validated methodology of immunogenicity prediction for MHC-II epitopes.

Why not all MHC binders subject to TCR recognition serve as epitopes? Since immunogenicity can be understood as a product of the amplitude of MHC-loaded peptides and the likelihood of recognition by T cells (*11*, *12*), the intermolecular contact potential profiles (CPPs) between the peptide-MHC complex and the host TCR repertoire are likely to be the critical determinants of immunogenicity (*13*, *14*). Moreover, it is plausible that, as a result of

evolutionary adaptation and thymic negative selection, public TCR clonotypes are more likely to recognize epitopes robustly immunogenic to a general population. Although molecular dynamics simulation is the most accurate method, it is virtually unrealistic to simulate the dynamics of every pair of peptides and TCRs. Instead, we hypothesized that a sequence-based framework that approximates CPPs using amino acid pairwise contact potential (AACP) scales is a feasible strategy. We made the following assumptions; First, only a small fraction of TCRs in the host TCR repertoire recognize a given MHC-loaded peptide; Second, complementarity-determining region 3 (CDR3) loops, which exhibit the highest degree of genetic variability, are sufficient in modeling the TCR-peptide interactions; Finally, only a small part of a CDR3 loop interacts with a small part of a given peptide. With those assumptions in mind, molecular scanning by TCR repertoire can be viewed as a parallelized pairwise sequence alignment problem between peptide sequences and CDR3 fragment sequences, with alignment scores calculated from AACP-derived custom substitution matrix as surrogate measurements of inter-fragment contact potentials. This strategy is independent of underlying MHC and therefore is expected to work both for MHC-I and -II epitopes.

Here, by implementing the aforementioned framework, we sought to identify the best predictive features that define immunogenicity of MHC-I and -II-restricted peptides. We found a partial overlap between the 32 and 27 most predictive features for MHC-I and MHC-II, respectively. The extremely randomized tree (ERT) models calculated the probability of immunogenicity in the testing subdatasets with the area under the receiver operating characteristic curve (AUC) of 79% and 73% for MHC-I and MHC-II, respectively. Further analysis led us to define “immune transition,” the acquisition or loss of immunogenicity after single amino acid substitutions. This is a higher-order concept encompassing various

terminologies such as escaping, cancer immunoediting, and epitope/neoepitope formation. Surprisingly, when focusing on non-transitional peptides, AUCs increased to 92% and 95% in MHC-I and MHC-II, respectively. Meanwhile, the same sets of features enabled accurate classification of transitional and non-transitional peptides, with AUCs being 92% and 94% in MHC-I and MHC-II, respectively. Moreover, mapping all possible mutants computationally generated from a given peptide in a neighbor network architecture enabled robust prediction of immunogenicity of transitional peptides, with AUCs being 85% and 93%, respectively. Finally, immunogenicity-based criteria of dissimilarity to corresponding wildtype peptides curtailed tumor neoepitope candidates by 76% while maintained the predictive power of neoepitope burden on patient survival after checkpoint blockade immunotherapy. Taken together, our findings define an integrated, high-dimensional architecture encompassing MHC-I and -II-restricted T cell epitope immunogenicity and provide a reference framework for systems biology of host-pathogen and host-tumor interactions.

Results

Immunogenicity prediction via TCR-peptide contact potential profiling

We collected a set of 21,162 8- to 11-aa HLA-I-restricted and 31,693 11- to 30-aa HLA-II-restricted peptide sequences from various sources (*9*, *10*, *15–21*). We regarded peptides with at least one positive T cell assay result as epitopes, and remaining ones as binders. The resultant metadataset contained 6,957 and 14,205 HLA-I epitopes and binders, and 16,642 and 15,051 HLA-II epitopes and binders, respectively. The immunogenicity prediction tool implemented in the Immune Epitope Database (*9*) showed only marginal predictive power (AUC: 54%) on this metadataset (fig. S1). Instead of further focusing on the peptide-MHC axis, we implemented a

framework termed “TCR-peptide contact potential profiling (CPP)” inspired by the hypothesis that public TCR repertoire shaped in favor of immunogenic peptides through germline-level adaptation and thymic selection could be used as a non-linear evaluation function to distinguish immunogenic and non-immunogenic peptides (Fig. 1A and Method). Iterative feature selections covering CPP-derived features and other peptide descriptors yielded the 32 and 27 most predictive features for MHC-I and MHC-II predictions, respectively (fig. S2). MHC binding prediction results from NetMHC4.0 (*4*) and NetMHCIIpan3.2 (*5*) to 12 and 6 representative HLA-I and HLA-II supertypes were also incorporated as features. We termed the mean probability estimates of immunogenicity from ERT models from 5 iterations as “immunogenicity scores.” The normalized score variances were distributed with medians (interquartile ranges) of 9.3% (6.4% to 13.6%) and 6.5% (4.3% to 9.4%) in MHC-I and MHC-II, respectively (fig. S3), showing that probability estimates were reasonably consistent between different iterations. Integration of predicted percentile ranks to HLA representatives improved the predictive power in MHC-I and less evidently in MHC-II (Figs. 1B and 1C). The optimal score thresholds were 0.36 and 0.51 in MHC-I and MHC-II, respectively. Under these thresholds, sensitivity and specificity were 69% and 74% in MHC-I, and 65% and 70% in MHC-II, respectively.

We next compared the predictive accuracies under different conditions. Notably, feature reduction only minimally undermined the predictive accuracy, suggesting that predictive rules defined by a tightly limited subset of the essential features apply across the entire datasets (fig. S4). In contrast, elimination of peptides with high sequence-level similarity led to a significant decrease in the predictive power particularly in MHC-I (fig. S5), suggesting that prediction was to some extent dependent on the sequence-level similarity of peptides. The difference between

MHC-I and MHC-II may arise from the fact that MHC-II accommodates peptides of a broader range of length for which sequence similarity-based clustering is less influential.

To investigate the generalizability of the predictive rules on the datasets from different species, we collected 8,756 and 411 peptides restricted on rodent and primate MHC-I, and 8,445 and 76 peptides restricted on rodent and primate MHC-II, respectively (fig. S6). We found that ERT models trained based on the human datasets had partial predictive power in MHC-II, whereas almost no predictive power was observed in MHC-I in rodents. Conversely, the immunogenicity of primate MHC-I-restricted peptides was to some extent predictable from human models, although no definitive conclusion can be made due to a dearth of data in primates.

Parameters utilized in the most predictive features have biological implications on the mechanisms of the molecular scanning by host TCR repertoire. Therefore, we next decomposed the CPP-derived features into a set of parameters, namely, fragment lengths, descriptive statistics, and AAIndex scales. Notably, features derived from short (*i.e.*, 3- and 4-aa) and long (*i.e.*, 8-aa and 11-aa in MHC-I and MHC-II, respectively) were more important (Fig. 2A). Meanwhile, as for descriptive statistics, features reflecting kurtosis and skewness were more critical (Fig. 2B). These trends were shared between MHC-I and MHC-II. In terms of AAIndex scales, the inverted version of BETM990101 (*22*) was the most important scale for both MHC-I and MHC-II (Fig. 2C). Also, both authentic and inverted versions of VENM980101 (*23*) were selected (Fig. 2C). The star plot shows that MHC-I prediction placed more considerable emphasis on the BETM990101-derived features, whereas MHC-II prediction utilized various scales more evenly (Fig. 2C). It is necessary to examine the biological relevance of the selected sets of parameters on TCR-peptide interactions and other receptor-ligand interactions in the future.

Collectively, the aforementioned observations strongly suggest that our framework recapitulates the high-dimensional architectures regulating CTL and Th cell activation to pave the way towards *in silico* prediction of T cell epitope immunogenicity.

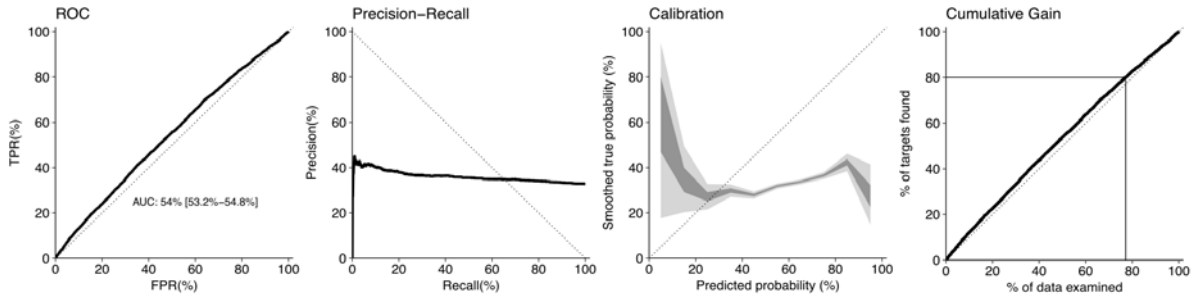

Fig. S1. MHC-I immunogenicity prediction by the IEDB Class I Immunogenicity tool. Immunogenicity was predicted using the tool with default settings. Note that peptides used for model training in the original paper(*9*) were not excluded in advance from prediction.

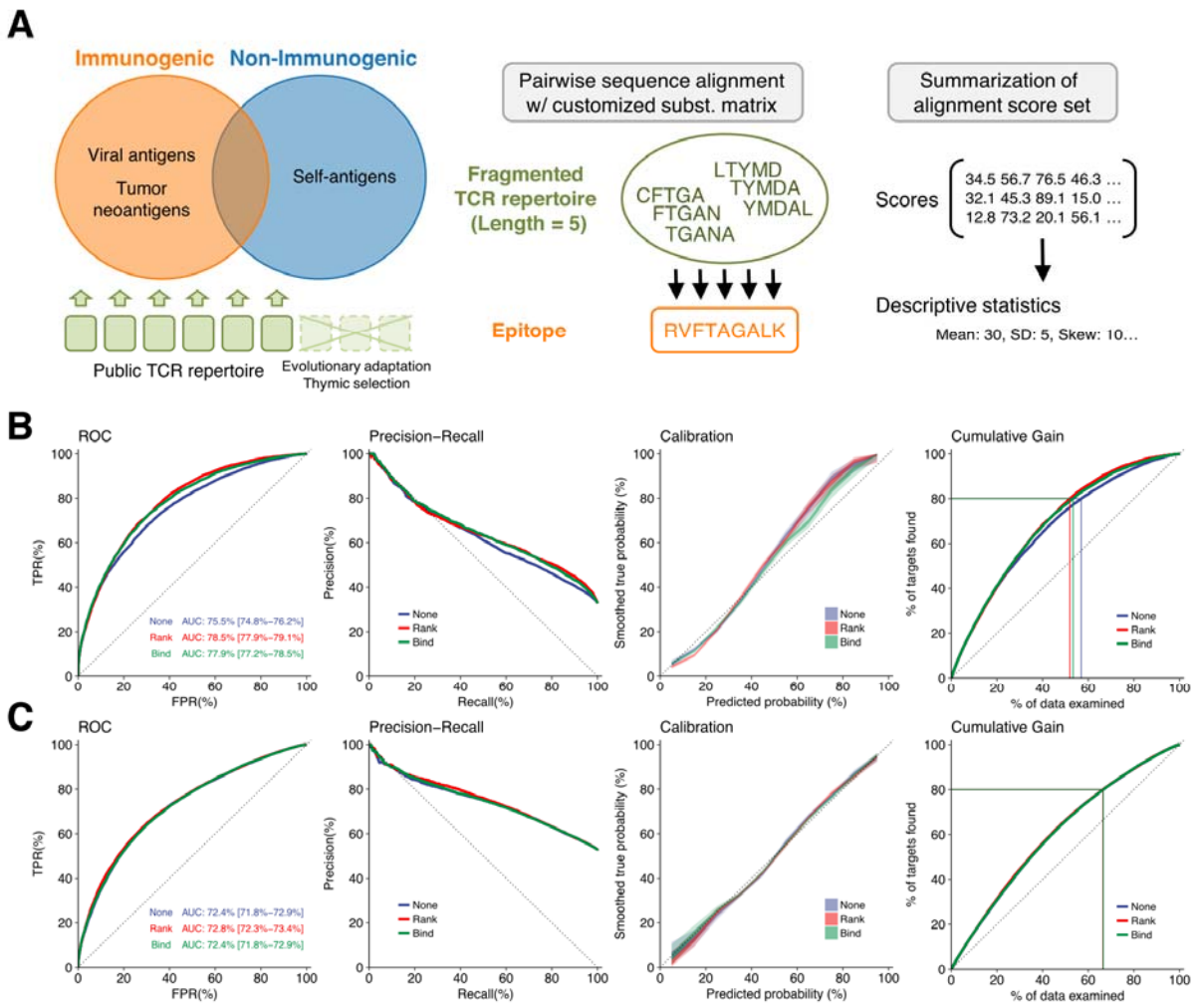

Fig. 1. Immunogenicity prediction from TCR-peptide contact potential profiling (CPP). (A) The concept of CPP. (B-C) Immunogenicity prediction of (B) HLA-I-restricted and (C) HLA-II-restricted peptides. Extremely randomized tree (ERT) models were trained using predefined sets of minimum features (fig. S2) with or without predicted binding to HLA representatives. None, ERT model training without predicted HLA binding data. Rank, with predicted percentile ranks to HLA representatives. Bind, with predicted binder categories to HLA representatives. Probabilities of immunogenicity were estimated for the testing subdataset. Averaged probability estimates from five iterations were termed “immunogenicity scores,” and diagnostic plots were

generated based on the scores. The AUC value was calculated as the average from ten-fold cross-validation, and the 95% confidence intervals were inferred from influence curves utilizing the method implemented in the R package *cvAUC* (*24*).

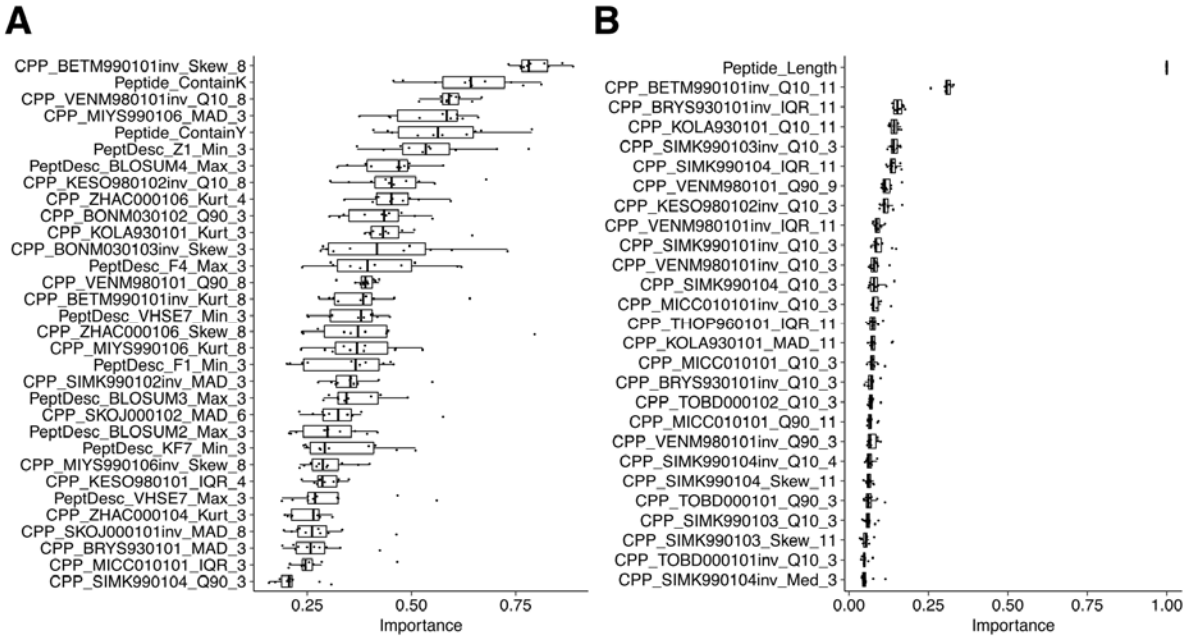

Fig. S2. Feature importance. The most predictive set of features for (A) MHC-I and (B) MHC-II immunogenicity prediction. Feature importance was estimated by constructing random forest (RF) models using a subset of the dataset, and the 100 most important features were selected. Features commonly selected from ten iterations were retained as the minimum essential feature set. The importance estimates from ten iterations are shown as box plots. For the nomenclature of features, see Method.

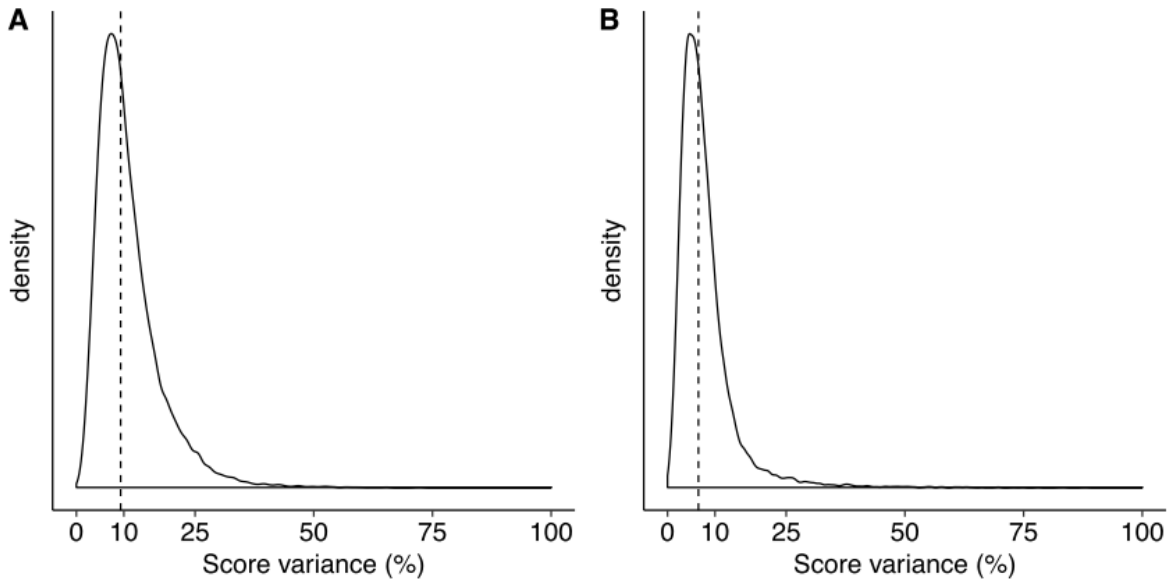

Fig. S3. Normalized score variances between iterations. Distributions of coefficients of variances (standard deviations normalized by means) of (A) MHC-I and (B) MHC-II probability estimates from five iterations are depicted.

146

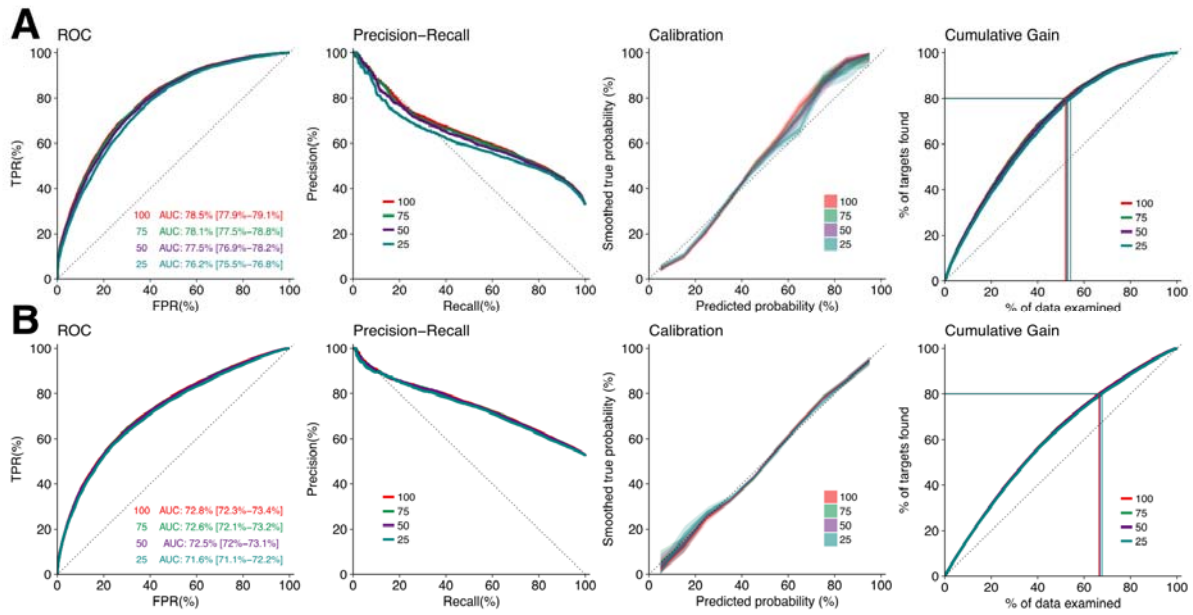

Fig. S4. Reduction of features only minimally affected the predictive power on immunogenicity. The cutoff of the number of features retained in each of the ten iterations was decreased from 100 to 25. (A) Immunogenicity prediction of HLA-I-restricted peptides. The actual numbers of features retained were 32, 26, 19, and 7, respectively. (B) Immunogenicity prediction of HLA-II-restricted peptides. The actual numbers of features retained were 27, 23, 21, and 12, respectively. Note that features derived from predicted percentile ranks to HLA representatives remained unchanged.

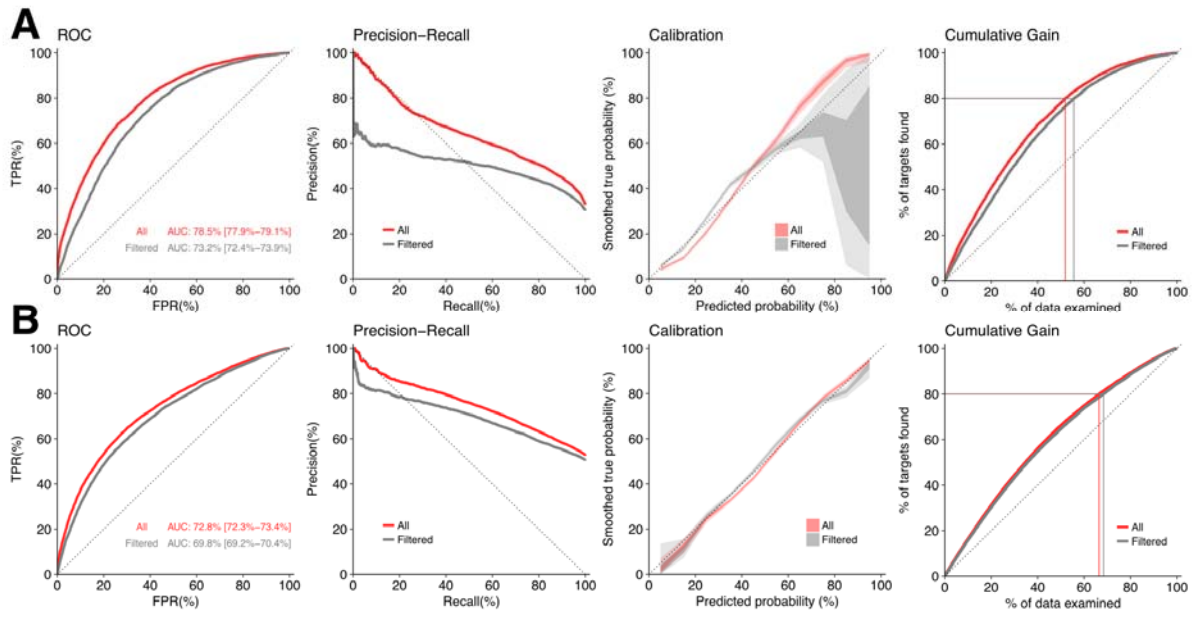

Fig. S5. Elimination of highly similar peptides significantly undermined the predictive power on immunogenicity. IEDB Epitope Cluster Analysis tool was used to cluster the peptides with the similarity threshold of 80%. All, all peptides. Filtered, single peptide randomly chosen from each cluster. (A) Immunogenicity prediction of HLA-I-restricted peptides. A total of 16,765 peptides were retained after clustering. (B) Immunogenicity prediction of HLA-II-restricted peptides. A total of 23,983 peptides were retained after clustering.

148

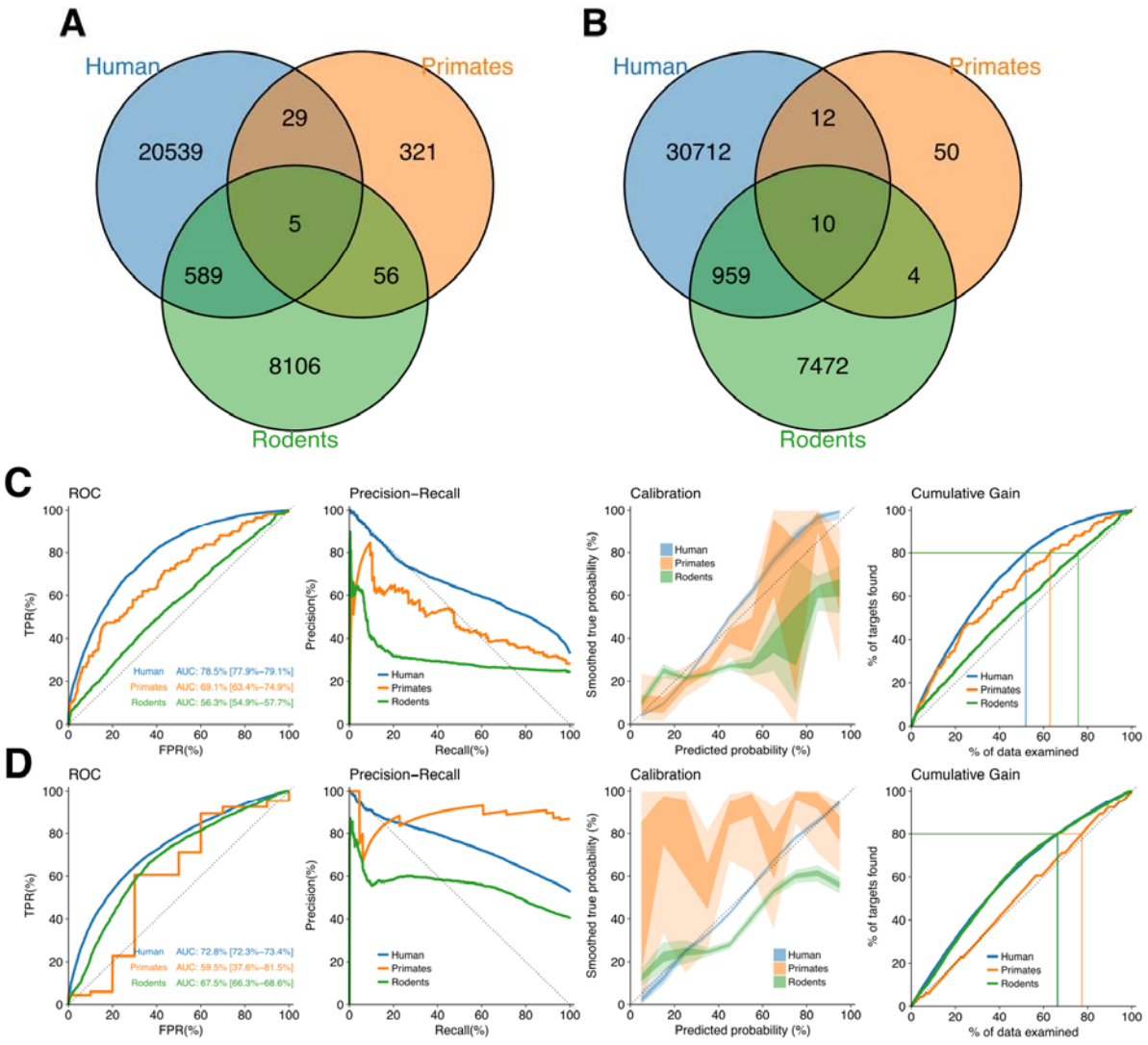

Fig. S6. Immunogenicity prediction on the peptide datasets from other species.

(A-B) Overlaps between (A) MHC-I-restricted and (B) MHC-II-restricted peptides. (C-D) Predictive performances of the probabilities of immunogenicity for (C) MHC-I-restricted and (D) MHC-II-restricted peptides. Probabilities were estimated from ERT models trained on the human peptide datasets.

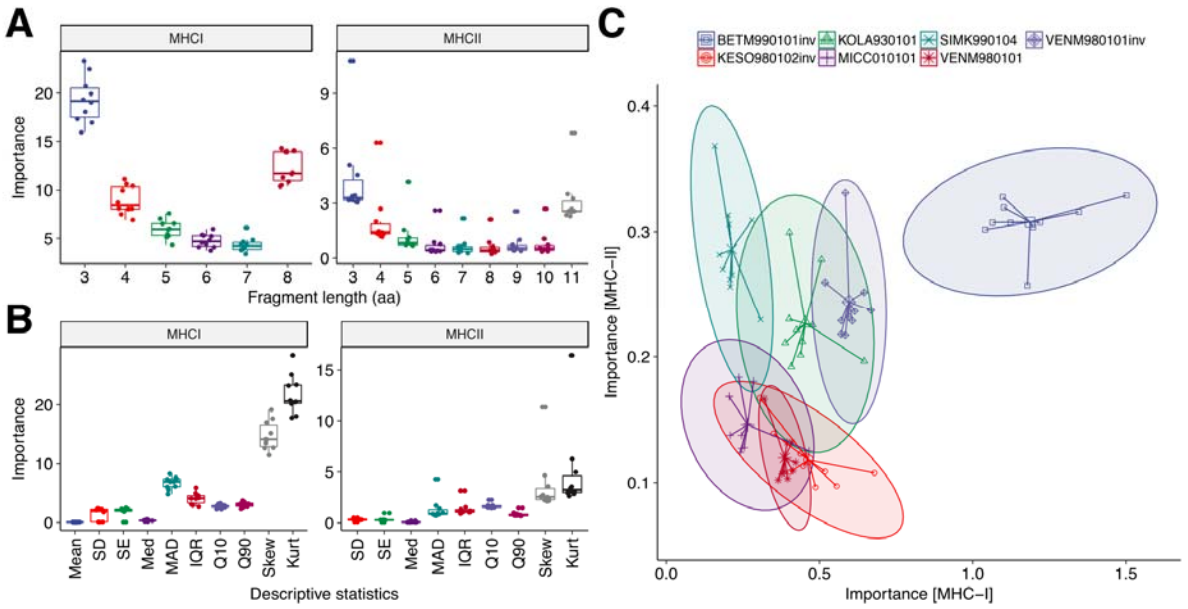

Fig. 2. Parameter importance analysis. (A-B) The sums of importance values of all CPP-derived features grouped by (A) fragment length and (B) descriptive statistics. The boxplot shows values from ten iterations. SD, standard deviation; SE, standard error of the mean; Med, median; MAD, median absolute deviation; IQR, interquartile range; Q10 and Q90, 10% and 90% quantiles, respectively; Skew, skewness; Kurt, kurtosis. (C) Importance stratified by AAIndex scales. Importance values of the most predictive 32 and 27 features for MHC-I and MHC-II, respectively, were aggregated by the AAIndex scales. Importance values of the scales utilized both in MHC-I and in MHC-II are depicted. Dots represent values from ten iterations.

Accurate immunogenicity prediction in non-transitional peptides

Single amino acid substitutions can drastically affect the immunogenicity of MHC-bound peptides. Mutations resulting in the loss of immunogenicity, also known as escaping mutations, have been extensively studied in the context of pathogen adaptation and cancer immunoediting

(*25–27*). Meanwhile, the recent advance in oncoimmunology has shed light on the roles of neoepitopes, epitopes engendered from somatic mutations that augment their immunogenicity, in the therapeutic effect of checkpoint blockade therapies and tumor evasion(*11*, *28*). However, acquisition and loss of immunogenicity are in principle the two sides of the same coin in sequence space, and understanding the bidirectional impacts of single amino acid changes in the landscapes of T cell epitope immunogenicity would provide valuable insights into the adaptation trajectories of pathogens and cancer.

To this end, we first screened our human epitope metadataset to obtain 6,128 and 3,413 MHC-I and -II-restricted peptides that have their neighbors, *i.e.*, counterparts with single amino acid differences. We termed the discordance of the annotated immunogenicity between neighbors as “immune transition.” Peptides are defined as “transitional” if there is at least one immune transition in the metadataset (Fig. 3A). In this manner, we found 660 and 700 transitional MHC-I epitopes and binders, and 500 and 476 transitional MHC-II epitopes and binders, respectively. Next, using the same feature sets and the same machine learning framework for immunogenicity prediction, we constructed classifiers predicting immune transition status. The probability estimates, which we termed “immune transition scores,” predicted immune transition status with AUCs of 92% and 94% in MHC-I and MHC-II, respectively (Figs. 3B-E).

Notably, the estimation of immunogenicity was near-perfect for non-transitional peptides (AUC: 92% and 95% in MHC-I and MHC-II, respectively), whereas transitional epitopes tended to have lower scores than transitional MHC binders (Figs. 3F-I). When we extrapolated the prediction of immune transitions to the entire peptide datasets (N=21,162 and N=31,693 for MHC-I and MHC-II, respectively), we found a monotonous decrease in the predictive accuracy

in terms of immunogenicity across the five strata based on the quintiles of immune transition scores (Fig. 4).

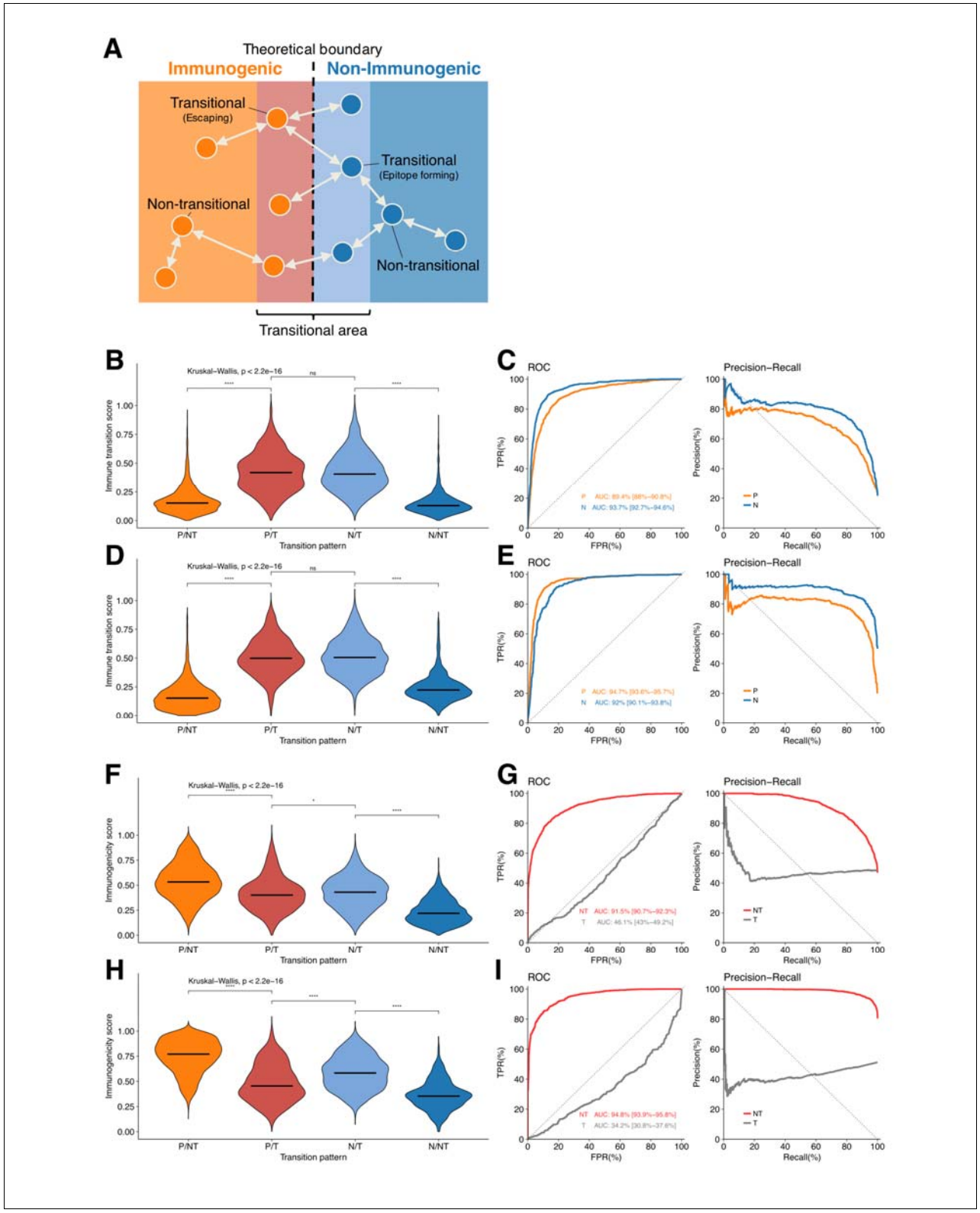

Fig. 3. Immunogenicity prediction stratified by immune transition status. (A) A schematic diagram of the concept of immune transition. Nodes represent peptides, and bidirectional arrows represent single amino acid mutations. (B, D) Distributions of immune transition scores stratified by transitional patterns in (B) MHC-I and (D) MHC-II. (C, E) Diagnostic plots of immune transition scores in (C) MHC-I and (E) MHC-II. (F, H) Distributions of immunogenicity scores stratified by transitional patterns in (F) MHC-I and (H) MHC-II. (G, I) Diagnostic plots of immunogenicity scores in (G) MHC-I and (I) MHC-II. P, positive (immunogenic); N, negative (non-immunogenic); T, transitional; NT, non-transitional. Black bars in violin plots represent medians. Bars represent medians. *P*-values were calculated by Kruskal-Wallis test followed by Wilcoxon rank sum test with false-discovery rate (FDR) adjustment. ns, not statistically significant; *, $P < 0.05$; ****, $P < 0.0001$.

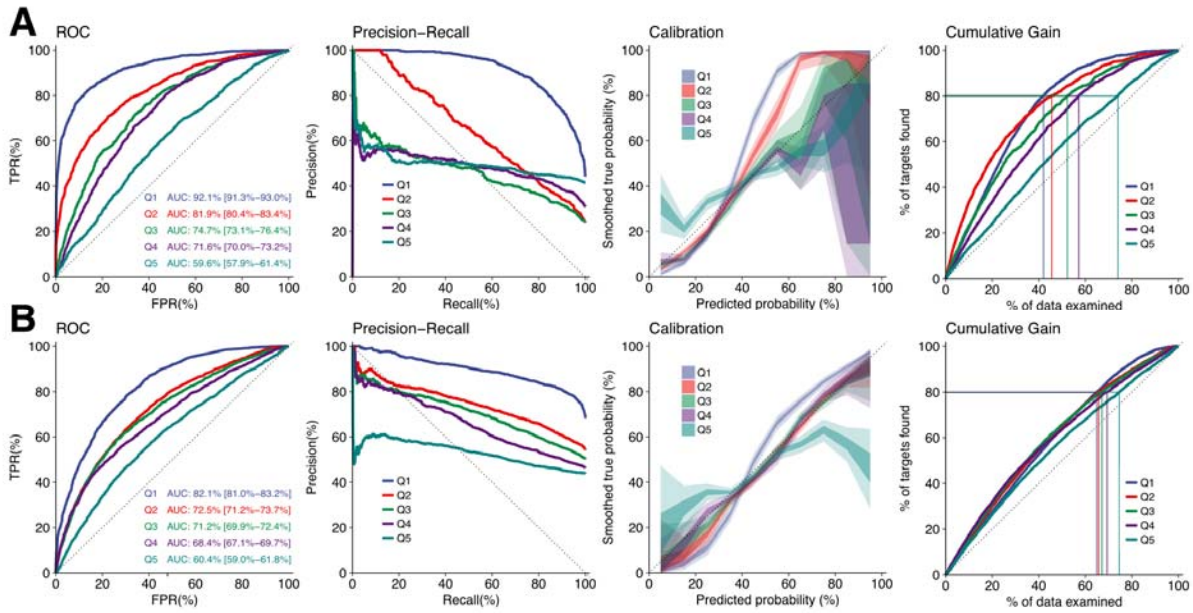

Fig. 4. Immunogenicity prediction was more reliable in peptides with lower immune transition likelihood. Diagnostic plots of (A) MHC-I and (B) MHC-II immunogenicity scores in the five strata based on the quintiles of immune transition scores are depicted.

Accurate immunogenicity predictions for transitional peptides from in silico mutant network architectures

To further characterize the immunogenicity landscape of transitional peptides, we examined the properties of neighbor networks, where each node represents a peptide, and peptides with single amino acid differences are connected with edge directions and weights defined from their immunogenicity scores. We first focused on the MHC-I epitope RRYQKSTEL (Fig. 5A). This epitope is transitional because 32 out of its 132 neighbors are non-immunogenic. We noted that all binders resided in Clusters 2 and 3, in which there were even epitopes that were predicted to be non-immunogenic. From these observations, we hypothesized that the cluster properties of neighbor networks constructed from computationally generated mutant sequences would be helpful for accurately predicting the immunogenicity of transitional peptides. As expected, mean immunogenicity scores of both all single amino acid mutants and the mutants in the corresponding cluster were significantly higher in epitopes (Figs. 5B and 5D). Using the per-cluster mean immunogenicity scores, immunogenicity was predictable with AUCs of 85% and 93% in MHC-I and MHC-II-restricted peptides, respectively (Figs. 5C and 5E).

Collectively, our framework provides a methodology for reliably quantifying the immunogenicity of MHC-bound peptides by taking into account their distance from a theoretical boundary of immunogenicity in sequence space.

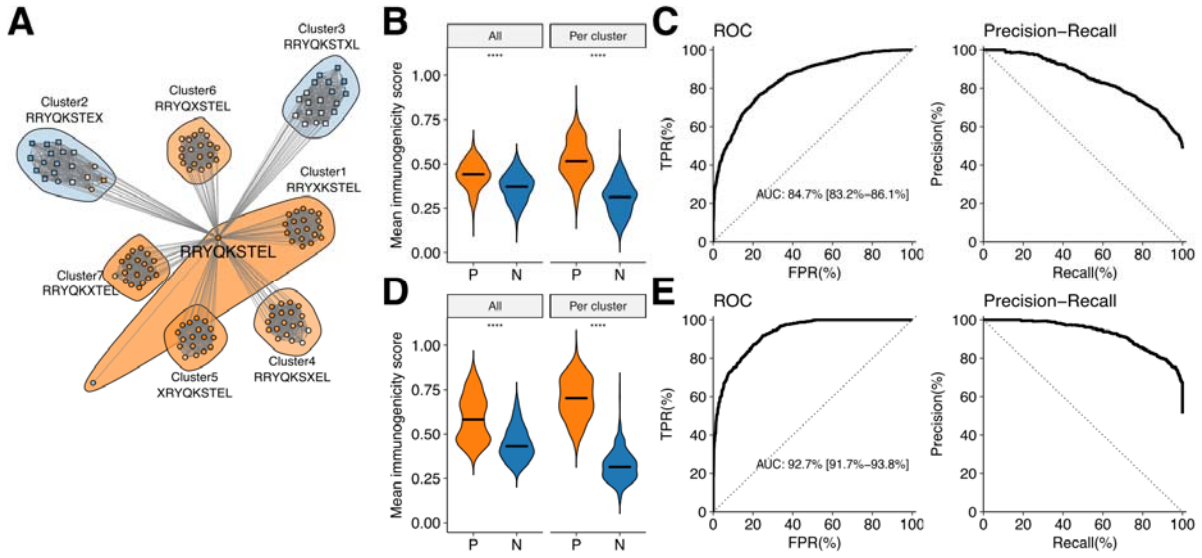

Fig. 5. Prediction of immunogenicity of transitional peptides from neighbor network clustering analysis. (A) A representative neighbor network of an HLA-I-restricted transitional epitope RRYQKSTEL and its 132 neighbors. There are 32 non-immunogenic neighbors. Peptides with single amino acid differences were linked. Edges were directed according to the differences in the immunogenicity scores between linked peptides. Edge distances were defined as the ratios of immunogenicity scores from smaller ones to larger ones. Peptides were clustered using a walktrap algorithm. The consensus sequence for each cluster was shown. X represents an arbitrary residue at the indicated position. The node shapes represent annotated immunogenicity (circle, immunogenic; square, non-immunogenic), and the node colors represent predicted immunogenicity, from blue (non-immunogenic) to orange (immunogenic). Each cluster was colorized by its mean immunogenicity score in the same color scheme. (B) The distributions of mean immunogenicity scores in transitional MHC-I-restricted peptides. P, positive (immunogenic); N, negative (non-immunogenic). (C) Prediction of immunogenicity of transitional MHC-I-restricted peptides from per-cluster mean immunogenicity scores. (D) The

distributions of mean immunogenicity scores in transitional MHC-II-restricted peptides. (E) Prediction of immunogenicity of transitional MHC-II-restricted peptides from per-cluster mean immunogenicity scores. Bars in violin plots represent medians. *P*-values were calculated by Wilcoxon rank sum test. ****, $P < 0.0001$.

The larger discrepancy of immunogenicity scores is a hallmark of transitional pairs of peptides

There are various occasions where comparison of a pair of peptides (*e.g.*, wildtype and mutant) are of interest. We hypothesized that the difference of immunogenicity scores is a hallmark of immunological dissimilarity between neighboring peptides. Indeed, we observed that immune transition between neighboring peptides was associated with the more substantial discrepancy between their immunogenicity scores (Fig. 6). We next examined the position-specific mutational effects. We focused on MHC-I because it has a closed-end groove to strictly constrain the anchor positions and orientation of presented peptides. Notably, while epitope-to-epitope non-transitional pairs consistently had smaller score discrepancies than epitope-to-binder transitional pairs, binder-to-epitope transitional pairs showed significantly larger discrepancies than binder-to-binder non-transitional pairs only when the mutation was introduced at positions 8 and 9 (fig. S7). Contrary to P9 that serves as an anchor in many MHC-I alleles, the role of P8 in MHC-I-restricted peptides has not yet been clarified, necessitating further investigation.

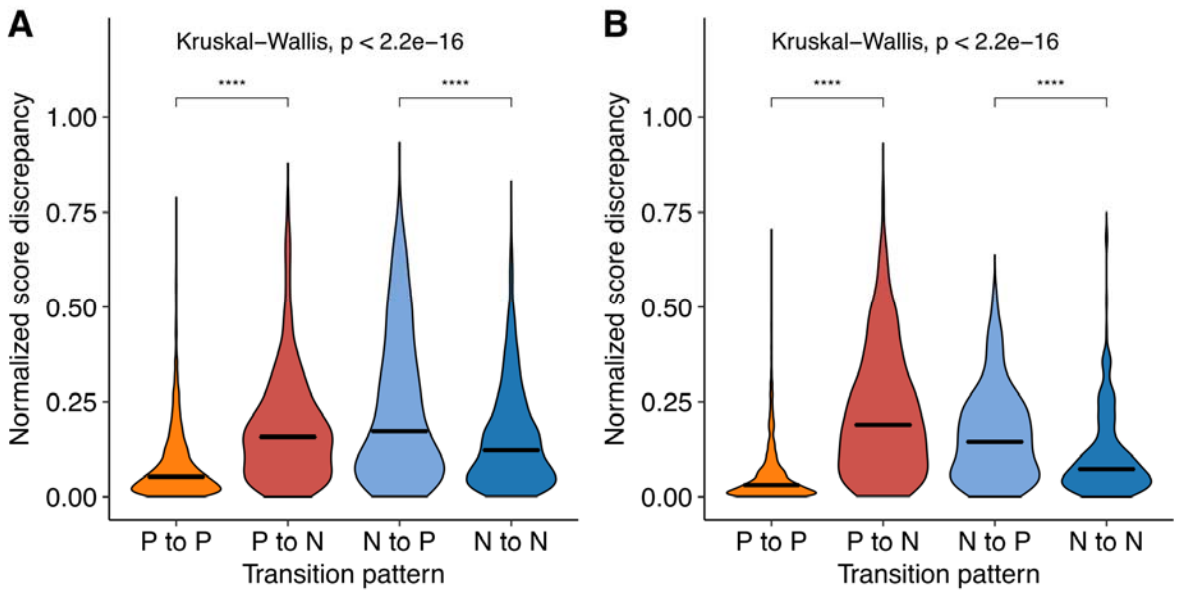

Fig. 6. Association of immune transition status and score discrepancies between neighboring peptides. Absolute differences in immunogenicity scores of neighboring peptides normalized by the sums of their scores are depicted. (A) MHC-I-restricted peptides. (B) MHC-II-restricted peptides. P, positive (immunogenic); N, negative (non-immunogenic). *P*-values were calculated by Kruskal-Wallis test followed by Wilcoxon rank sum test with FDR adjustment. ****, $P < 0.0001$.

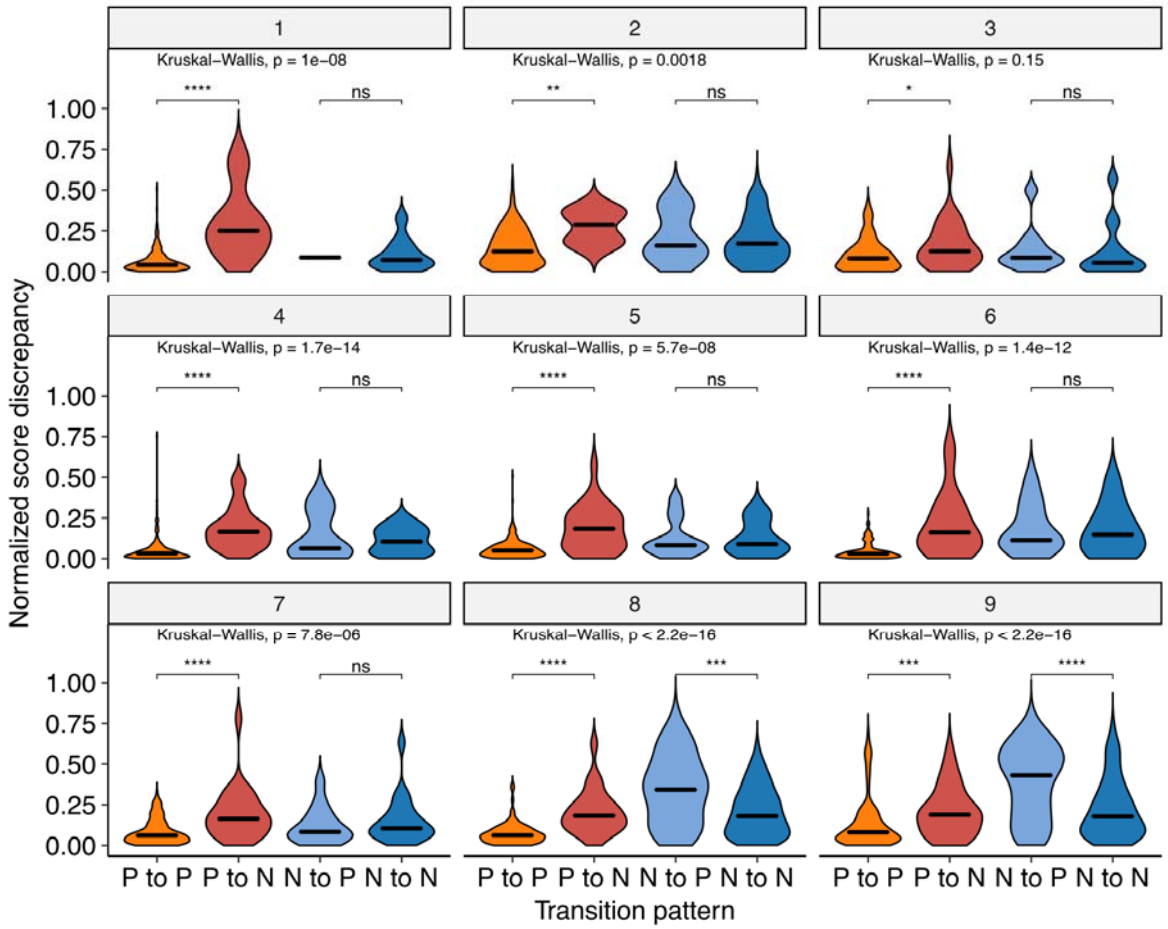

Fig. S7. Position-specific effects of mutations on the score discrepancies in MHC-I-restricted 9-aa peptides. Bars represent medians. *P*-values were calculated by Kruskal-Wallis test followed by Wilcoxon rank sum test with FDR adjustment. ns, not statistically significant; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$.

Neoepitopes reside in the middle of MHC binders and epitopes in sequence space

Accumulating evidence indicates that $CD8^+$ T cell immunity directed toward cancer neoepitopes is a hallmark of anti-tumor immunity (*28*, *29*). In contrast to microbial epitopes,

most neoepitopes still retain substantial homology to corresponding self-peptides. Indeed, our framework assigned intermediate levels of immunogenicity scores to the $CD8^+$ T cell neoepitopes of various tumor origins (Fig. 7), reflecting their homology to non-immunogenic self-peptides from an immunological perspective. Considering previous studies showing that homology of neoepitopes to microbial peptides is tightly associated with patient survival (*11*, *12*), it is reasonable to expect that more dissimilar-to-self neoepitopes are more immunogenic.

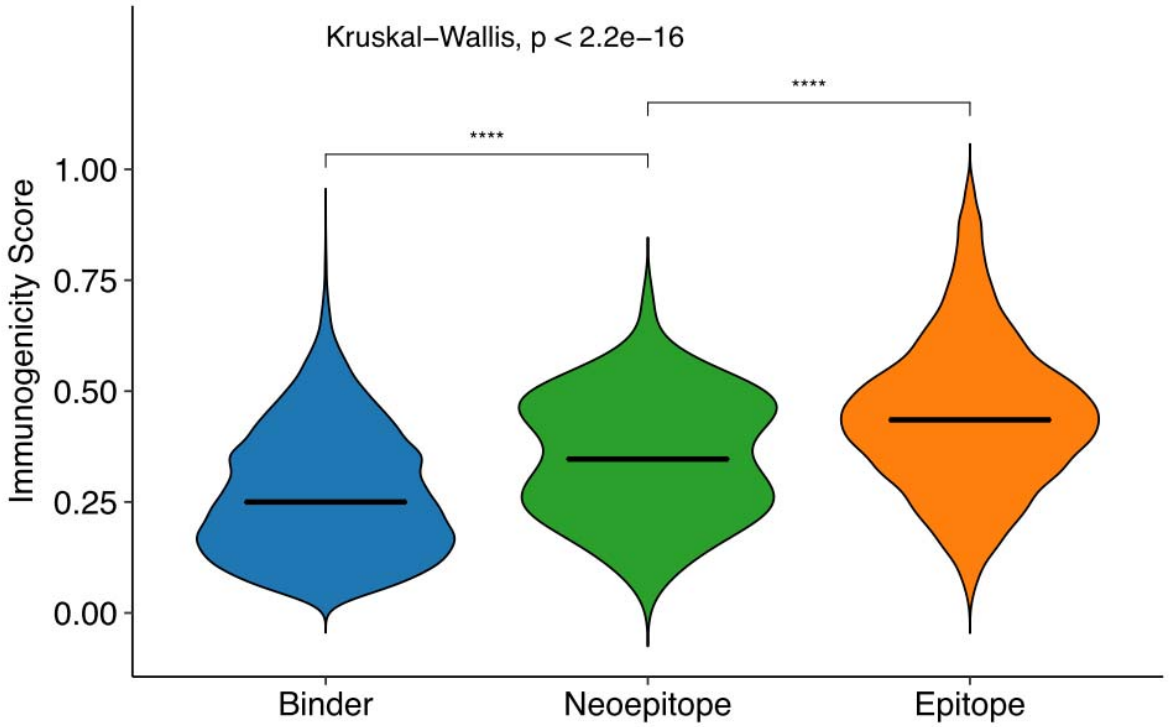

Fig. 7. Neoepitopes had intermediate levels of immunogenicity scores. Immunogenicity scores were calculated for a set of previously validated MHC-I-restricted neoepitopes (N=82) collected from the review by Durgeau *et al.* (*30*). Of the 82 neoepitopes, 45 were not included in our metadataset. Immunogenicity scores of MHC binders and epitopes in our metadataset are also depicted for comparison. Bars represent medians. *P*-values were calculated by Kruskal-

Wallis test followed by Wilcoxon rank sum test with FDR adjustment. ****, $P < 0.0001$.

Prediction of survival after checkpoint blockade immunotherapy from dissimilar-to-self neoepitope burden

Immune checkpoint blockade (ICB) therapy invigorates neoepitope-targeting tumor-infiltrating T cells to augment tumor rejection (*31*, *32*). The observations that high tumor neoepitope burden and microsatellite instability correlate with clinical benefits of ICB also support this notion (*29*, *33–35*). However, criteria for neoepitopes that elicit the most effective immune responses remain elusive, which dampens the development of cancer vaccines eliciting potent protective immunity (*36*). Indeed, it has been shown that the majority of predicted neoepitope candidates selected based on intracellular processing and MHC binding do not elicit T cell responses (*37*). Thus, strategies to "concentrate" potent neoepitope candidates are highly desired.

In this context, we asked whether the neoepitope burden curtailed by immunogenicity-based criteria still predicted patient survival after ICB therapy. We utilized the previously compiled HLA-I-restricted neoepitope candidate datasets (*11*) derived from two melanoma cohorts (*34*, *35*) and one non-small cell lung cancer (NSCLC) cohort (*29*). First of all, without immunogenicity-based filtering, patient overall survival was dichotomized (N=197; $P=8.6\times10^{-5}$ by the log-rank test) with the optimal cutoff of per-patient neoepitope burden being 83 (fig. S8A). The distributions of neoepitope burden were 38.2 ± 22.7 vs. 518.3 ± 603.8 in low- and high-burden groups, respectively (fig. S8B). Next, we attempted to filter out candidates by applying a cutoff to the immunogenicity scores of the mutated (MT) peptides. The optimal thresholds for immunogenicity score and per-patient neoepitope burden were 0.155 and 82, respectively, which

led to successful dichotomization (P=3.8×10^{-5}). However, thresholding with these parameters reduced the number of neoepitope candidates only by 3%, and the distributions of neoepitope burden were 38.0 ± 22.5 vs. 508.3 ± 587.4 in low- and high-burden groups, respectively (figs. S9). Finally, we utilized the normalized score discrepancy between WT and MT peptide pairs as a metric of immunological dissimilarity (see Fig. 6 and Methods). The underlying hypothesis is that neoepitopes more immunologically dissimilar from corresponding self-peptides are more likely to circumvent thymic and peripheral tolerance and thus to be recognized by highly reactive T cell repertoire. The optimal thresholds for score discrepancy and neoepitope burden were 0.0975 and 20, respectively. The criteria dichotomized survival (P=7.7×10^{-5}) and simultaneously curtailed neoepitope candidates by 76% so that the distributions of the modified neoepitope burdens were 8.2 ± 5.1 vs. 125.0 ± 152.7 in low- and high-burden groups, respectively (Fig. 8). Per-patient reduction rate was 76.7% ± 6.2%, and the reduction was more substantial in the low-burden group (fig. S10).

Overall, these findings place tumor neoepitope biology in a more generalized context of CTL immunity. Larger-scale cancer mutanome profiling study is warranted to validate and further sophisticate the prediction of the most promising neoepitope candidates, thus galvanizing research in precision oncoimmunology.

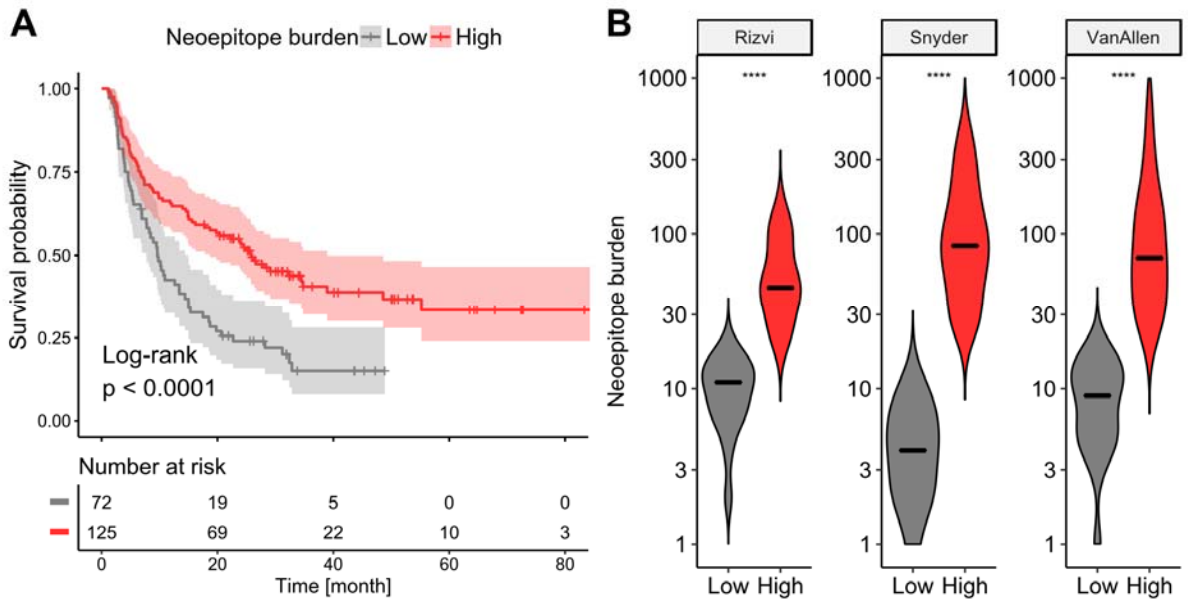

Fig. 8. Dissimilar-to-self neoepitope burden was associated with overall survival after checkpoint blockade immunotherapy. (A) Stratification of overall survival of patients treated with checkpoint blockade immunotherapy. The dataset compiled by Lukzsa *et al*. (*11*) derived from two melanoma cohorts (*34*, *35*) and one non-small cell lung cancer (NSCLC) cohort (*29*) contains HLA-I-restricted neoepitope candidates. The difference between the immunogenicity scores of wildtype (WT) and mutated (MT) peptides normalized by the sum of their scores was used as a metric of dissimilarity to self. The cutoff of normalized score discrepancy was 0.0975, and patients with 20 or larger neoepitopes were categorized as high neoepitope burden. *P*-value was calculated by the log-rank test. (B) Distributions of dissimilar-to-self neoepitope burden. Bars represent medians. *P*-values were calculated by Wilcoxon rank sum test. ****, $P < 0.0001$.

265

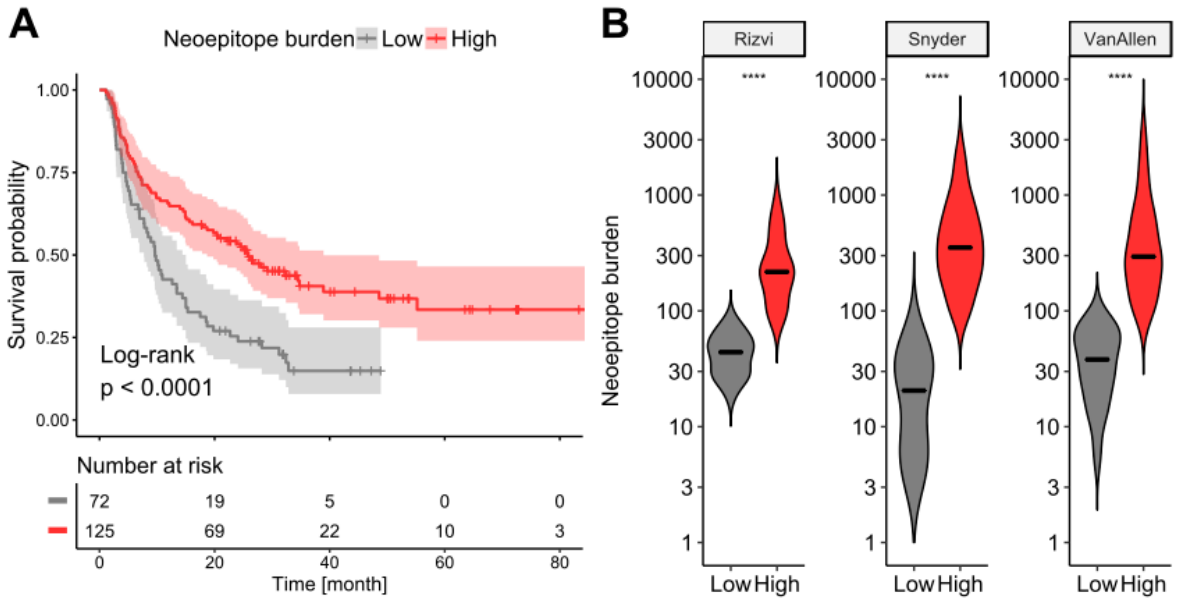

Fig. S8. Non-filtered predicted neoepitope burden was associated with overall survival after checkpoint blockade immunotherapy. See the main text and Fig. 8 for details.

266

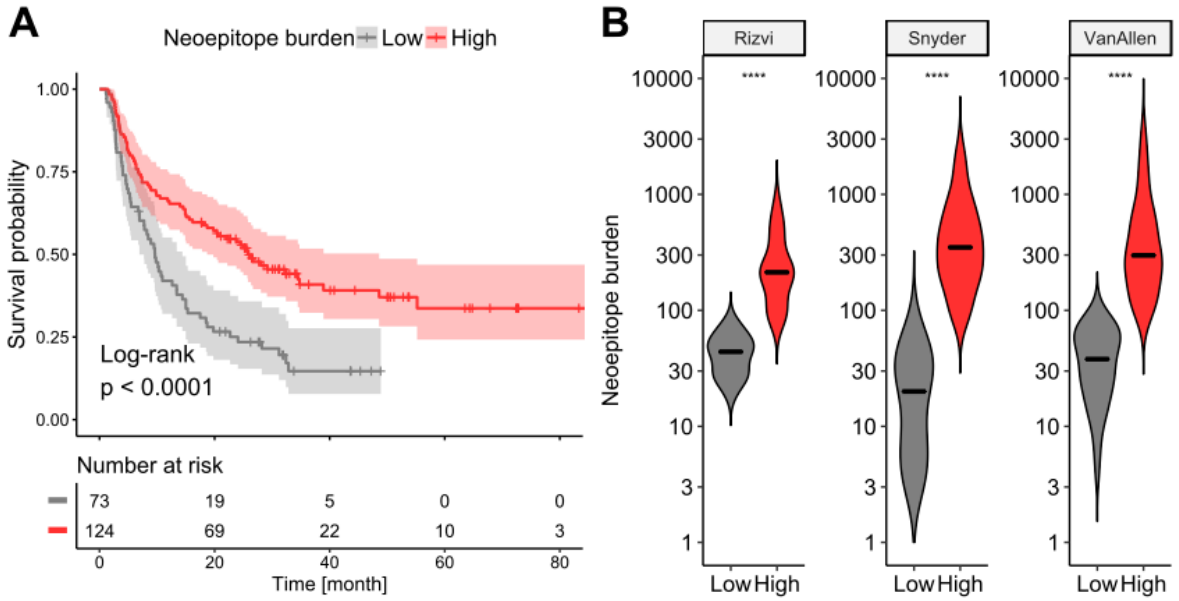

Fig. S9. Neoepitope burden filtered based on immunogenicity score was associated with overall survival after checkpoint blockade immunotherapy. See the main text and Fig. 8 for details.

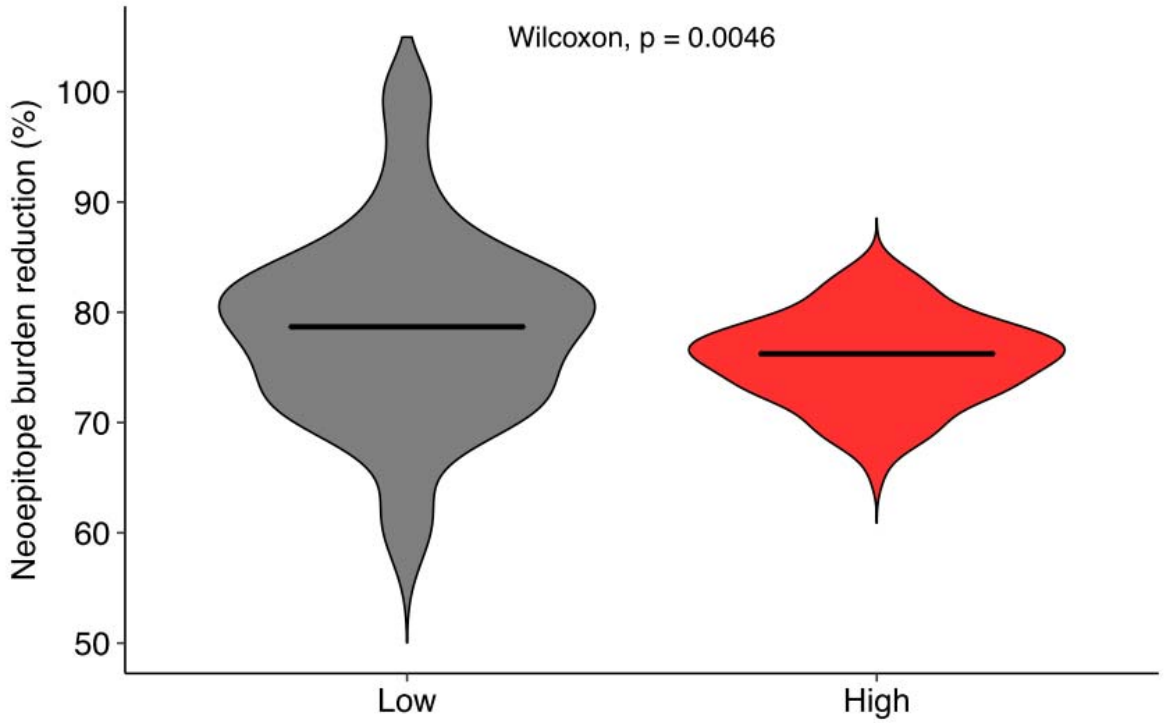

Fig. S10. Reduction of neoepitope burden per patient stratified by burden groups. Normalized score discrepancy was used as a metric for filtering neoepitope candidates. Per-patient neoepitope reduction rates are depicted. Red, high-burden group. Gray, low-burden group. Bars represent medians. *P*-values were calculated by Wilcoxon rank sum test.

Discussion

The ultimate goal of immunogenicity prediction is to predict which peptide best triggers sustained immunological response *in vivo*. In contrast to the complexity of the regulatory mechanisms *in vivo* including regulatory T cells (*38*), T cell exhaustion mediated through chronic immune checkpoint signals (*39*), and the immunosuppressive microenvironment engendered by solid tumors (*40*, *41*), T cell activation is primarily governed by the MHC-peptide-TCR interactions, with relatively minor assistance from coreceptors like CD8 (*42*). This asymmetry makes *in vitro* T cell activation assay results difficult to interpret. Indeed, efforts to establish predictive algorithms on immunogenicity remain mostly unsuccessful; the current method for MHC-I immunogenicity prediction (*9*) implemented in the Immune Epitope Database only yielded AUC of 54%, and there is no validated method for MHC-II to date.

In this work, we proposed a novel concept termed TCR-peptide contact potential profiling that may explain the immunogenicity of both MHC-I and -II-restricted epitopes in an integrated fashion. Basically, it utilizes TCR fragment libraries from public TCR repertoire as an evaluation function for measuring the homology of a peptide to a set of highly-immunogenic epitopes. In contrast to high-throughput *in vitro* T cell assays that only measure the immunogenicity of the peptides tested, our framework extracts the high-dimensional rules of immunogenicity applicable throughout the epitope sequence space. Indeed, it is intriguing that the machine-learned models for MHC-I and MHC-II utilized partially overlapped feature sets. Our findings may help deduce a unified and generalizable insight into MHC-peptide-TCR interactions. Furthermore, to our surprise, a simple concept which we termed immune transition effectively stratified both MHC-I and -II-restricted peptides to make unprecedentedly accurate immunogenicity prediction feasible. We speculate that transitional peptides are less stable than non-transitional ones in terms of immunogenicity due to the very biological limitation of TCR-

dependent self-nonself discriminating mechanisms. Additionally, the immunogenicity of transitional peptides may significantly be affected by the spatiotemporal dynamics of antigen-specific T cell clonotypes. However, one should bear in mind that misannotations owing to suboptimal assay conditions may also contribute to making immunogenicity prediction for transitional peptides even more complicated. Large-scale evaluation of the immunogenicity of neighboring peptides under comparable assay conditions would further elaborate the analysis on immune transition.

Tumor evasion from adaptive immunity can be mechanistically dissected into the following categories: (i) immunoediting to eliminate highly immunogenic neoepitopes; (ii) evasion from T cell surveillance due to the reduced level of antigen presentation; (iii) aberrant signaling pathways to prevent T cell migration and priming such as immune checkpoints (*27*, *32*, *43*, *44*). Meanwhile, tumor-extrinsic factors including gut microbiota also modulate anti-tumor immunity (*45*, *46*). Therefore, it is quintessential to take multiple biomarkers linked to different aspects of oncoimmunology into account to comprehensively understand the spectrum of clinical outcomes after ICB therapy. That being said, however, accumulating evidence suggests that ICB therapy mediates tumor rejection principally by potentiating neoepitope-specific T cell immunity (*28*, *31*, *32*). Therefore, it is of significant implications for translational immunology to explore criteria defining best-immunogenic neoepitopes which would enhance the efficacy of ICB therapy and serve as the targets of therapeutic cancer vaccines (*36*). An inspiring tumor fitness model proposed by Luksza *et al*. exploits sequence-level similarity of neoepitope candidates to known pathogen-derived epitopes as a metric of dissimilarity-to-self to predict response to ICB therapy (*11*). However, the scope of the fitness model is limited to MHC-I-restricted neoepitopes, and epitope-level immunogenicity prediction is not shown. Our finding that

dissimilar-to-self neoepitope burden was tightly associated with response to ICB therapy bolsters the notion that neoepitope “nonselfness” is the critical driving force of anti-tumor immunity. Immunogenicity score-guided *in silico* screening of neoepitopes with the highest therapeutic potential is an exciting next step.

Our study should be considered a starting point for the investigation of T cell epitope immunogenicity due to the following limitations. First, validation of the interaction profiles between MHC-bound peptides and TCR repertoire in a physiologically relevant condition would be an essential step to further sophisticate the algorithm of *in silico* TCR-peptide contact potential profiling, although it would be a challenging task given their extreme diversity. Second, the dearth of epitope data from non-microbial, non-tumorous sources may limit the applicability of our framework in other fields such as autoimmunity. Delineation of the immunogenicity landscape of self-epitopes, molecularly mimicking microbial epitopes, and their intermediates in sequence space in the context of epitope spreading would be an interesting topic to investigate. Finally, prospective identification of novel epitopes from pathogens and malignancies for which effective vaccines are not yet available would be quintessential to highlight the validity of our framework and warrant its clinical application.

Materials and Methods

Study design

The purpose of this retrospective study was to systematically analyze MHC-I and MHC-II-restricted peptide sequence datasets through a newly designed machine learning framework to identify the minimum sets of features that can classify immunogenic epitopes and non-immunogenic MHC binders and to establish a numerical scale for quantifying the immunogenicity of any given MHC-bound peptide.

All analysis is exploratory; no predetermined protocols were applied before initiating the entire study. All data inclusion and exclusion criteria were described in the following sections. Accounting for missing data values including imputation is not applicable.

Computational analysis

All computational analyses were conducted using R ver. 3.5.0 (*47*). The latest versions of R packages were consistently used. Scripts are available upon request.

HLA-restricted peptide sequence datasets

HLA-I-restricted peptide sequences of 8-aa to 11-aa lengths with T cell assay results were collected from public databases [Immune Epitope Database (IEDB, as of May 7th, 2018) (*21*), the best-characterized CTL epitopes from Los Alamos National Laboratory (LANL) HIV Sequence Database (*17*), LANL HCV Sequence Database (*15*), EPIMHC (*19*), MHCBN (*16*), and TANTIGEN (*18*)] and previous publications (*9*, *10*, *20*). HLA-II-restricted peptide sequences of 11-aa to 30-aa lengths with T cell assay results were collected from the IEDB database (as of May 7th, 2018). As for peptides retrieved from the IEDB database, only those with valid evidence of T cell response were included (table S1). Peptides presented on non-

human MHC molecules were discarded, whereas those presented on HLA molecules in non-human hosts (*e.g.*, transgenic mice) were included. For the peptides having contradictory annotations on immunogenicity, peptides with at least one positive T cell assay result were considered immunogenic. In this manner, 21,162 and 31,693 HLA-I and HLA-II-restricted peptides were identified, respectively, of which 1,873 (8.9%) and 4,505 (14.2%) had contradictory annotations on immunogenicity, and 6,957 (32.9%) and 16,642 (52.5%) were considered immunogenic, respectively.

Peptide sequence datasets restricted on non-human MHC molecules

Peptides were collected from the IEDB database (as of as of May 7th, 2018) as described above. The following species were considered primate: bonobo, chimpanzee, gorilla, marmoset, and rhesus macaque. Meanwhile, mouse and rat were considered rodent. Eventually, 411 and 8,756 MHC-I-restricted peptides, and 76 and 8,445 MHC-II-restricted peptides were identified for primates and rodents, respectively.

TCR-peptide contact potential profiling (CPP)

Concept. T cell epitopes presented on the MHC molecules must be recognized by the TCRs of $CD8^+$ CTLs and $CD4^+$ Th cells with sufficient affinity to trigger subsequent immunological cascades (*14*). Apparently, the entire TCR repertoires do not necessarily recognize those peptides; instead, a quite limited set of epitope-specific high-affinity TCRs are required. Moreover, given the flexible conformational changes upon binding between TCR and peptide-MHC complex, it is plausible that only a restricted set of residue pairs between the peptide and the CDR3 loop initially act as a “seed” of molecular scanning that may endow the complex the time for conformational changes to further enhance the affinity and eventually activate T cells (*14*, *48*, *49*). In this view, the interacting surfaces could be approximated as a pair of linear

sequences, and by deconvoluting the TCR repertoire into a set of fragments of a defined amino acid length, the process of molecular scanning could be interpreted as a pairwise sequence alignment problem. One caution is that, although an ordinary sequence alignment algorithm involves a similarity matrix which gives higher values to biochemically similar residue pairs, higher scores must be given to more strongly interacting residue pairs rather than biochemically more similar residue pairs in order to interpret the alignment as the reflection of intermolecular interaction. For this purpose, amino acid pairwise contact potentials (AACPs) were adopted from the AAIndex database (*50*) (http://www.genome.jp/aaindex/AAindex/list_of_potentials) to generate custom substitution matrices. The Smith-Waterman local alignment algorithm maximizes the alignment score of a given TCR-derived sequence fragment against the whole sequence of a given peptide, and its alignment score can be interpreted as a correlate of the intermolecular contact potential. Consequently, the interactions between a given peptide and a given TCR repertoire can be profiled by summarizing the distribution of alignment scores.

Generation of TCR fragment library. We hypothesized that human public TCR clonotypes are biased in favor of "publicly" immunogenic peptides rather than non-immunogenic ones, and therefore could be utilized as a probe for immunogenicity. We focused on the CDR3 loops because they are primarily responsible for the interactions with MHC-loaded peptides, whereas MHC α-helices are typically contacted with more conserved CDR1 and CDR2 loops (*2*, *3*, *51*). We restricted our analysis to human TCRβ (TRB) repertoire. Due to the frequent lack of annotations of T cell sorting, we included TCR repertoires from all kinds of unsorted T cells. TCR repertoire datasets were collected from NCBI Sequence Read Archive (SRA) and a previous study led by Britanova *et al.* (*52*). SRA was searched with the following terms '(T Cell Receptor β) AND human.' The following BioProjects were included: PRJNA389805,

PRJNA329041, PRJNA298417, PRJNA273698, PRJNA258001, PRJNA229070, PRJNA79707, and PRJNA79435. Projects in which healthy donors were not involved (*e.g.*, cancer, autoimmune diseases) were omitted. Datasets derived from HIV-infected subjects in the BioProject PRJNA258001 were discarded. Meanwhile, PRJNA316572, involving the unique molecular identifier strategy, were not included in order to maintain consistency in the way of analysis. Fastq files were obtained using fastq-dump script with the following options: --gzip --skip-technical --readids --read-filter pass --dumpbase --split-files --clip --accession [SRA run number]. A total of 23,006,555 TRB CDR3 clonotypes were identified using MiXCR software (*53*), and a total of 191,326 clonotypes observed in at least 22 out of 220, *i.e.*, 10%, different datasets were retained as public clonotypes. Finally, the pooled CDR3 sequences were fragmented by a sliding window strategy to generate a fragment library. The sizes of fragments were from 3-aa to 8-aa in MHC-I cases, and from 3-aa to 11-aa in MHC-II cases. Moreover, since CDR3 loops could interact with peptides in either a forward parallel or antiparallel fashion, reversed CDR3 sequences were simultaneously fragmented and pooled. Resultant fragment pools were compressed to a defined library depth (*e.g.*, 10,000) by random down-sampling while preserving the original proportion of each fragment.

Pairwise sequence alignment. The *pairwiseAlignment* function implemented in the *Biostrings* package in *Bioconductor* (*54*) was used. This function seeks an optimal alignment which maximizes the overall alignment score defined as a sum of pairwise scores. Alignment type was set “global-local” to obtain an optimal alignment of a CDR3 fragment with consecutive subsequences of the peptide. Gaps were not allowed. A set of substitution matrices were generated from a set of thirty-five AACP scales retrieved from the AAIndex database (table S2) (*50*). AACP matrices were rescaled so that numerical indicators resided within the range between

0 and 1 for comparative purposes. For analytical purposes, we also prepared inversed AACP matrices (rescaled as well) to mimic the weakest interactions as well as the most robust interactions. A distribution of alignment scores was summarized by calculating representative statistics. Following statistics were calculated using the *describe* function implemented in the *psych* package: mean, standard deviation, standard error of the mean, median, median absolute deviation, interquartile range, 10% quantile, 90% quantile, skew, and, kurtosis. Predictive features were generated by combining the fragment length, the AACP scale, and the type of statistics. In this manner, 700 CPP features per one fragment length were generated for each of the peptides.

Peptide descriptors

Apart from CPP features, sequence-based estimates of physicochemical properties were also calculated. Each peptide sequence was converted into a set of consecutive fragments of a defined amino acid length, and peptide descriptors were calculated against each of the fragments using functions in the *Peptide* package. Following functions were utilized: *aIndex*, *blosumIndices*, *boman*, *charge*, *crucianProperties*, *fasgaiVectors*, *hmoment*, *hydrophobicity*, *instaIndex*, *kideraFactors*, *mswhimScores*, *pI*, *protFP*, *vhseScales*, and, *zScales*. The distributions of the values were summarized as described in the paragraph of CPP. In addition, a set of categorical features indicating whether the peptide of interest is free from a specific amino acid residue were included. Finally, the peptide length was used as a feature.

HLA binding prediction

We incorporated predicted binding strength against 12 HLA-I representatives (A1, A2, A3, A24, A26, B7, B8, B27, B39, B44, B58, and, B62) from NetMHC 4.0 (*4*) and 6 HLA-II representatives (DP, DQ, DRB1, DRB3, DRB4, DRB5) from NetMHCIIpan 3.2 (*5*) for MHC-I

and MHC-II immunogenicity prediction, respectively, because the stability of the peptide-MHC complex is a known correlate of immunogenicity (*13*). We decided not to use annotation-based HLA restriction data because experimentally validated HLA restrictions may be biased when the epitopes in question are relevant to specific diseases that have been associated with specific HLA alleles. In such cases, incorporating HLA annotation would cause overfitting in machine learning process. Moreover, this setting keeps the entire framework of immunogenicity prediction *in silico*. Predictions were made with default settings. Thresholds of percentile ranks for strong and weak binders were set as 0.5% and 2% in MHC-I, and 2% and 10% in MHC-II, respectively. For comparative purposes, binding to HLA representatives was predicted for peptides restricted to non-human MHC molecules.

Peptide filtering based on sequence-level homology

The Epitope Cluster Analysis tool implemented in the IEDB database was used to perform sequence similarity-based clustering. The similarity threshold was set as 0.80. One peptide per cluster was randomly chosen.

Feature selection

Preprocessing and feature selection steps were iterated ten times with different random seeds. Each iteration consisted of the following steps. First, the peptide set was randomly split in a ratio of 4:1 for the training and testing subdatasets. Second, numerical peptide descriptors other than peptide length and CPP features were centered and scaled. A peptide length and categorical variables representing existence or absence of each of the 20 residues were not preprocessed. Preprocessing functions defined using the training subdataset were also applied to the testing subdataset. Third, highly correlated features were removed, with the threshold of correlation coefficient being 0.75. Finally, features were filtered based on the importance values calculated

using the *generateFilterValuesData* function with the random forest (RF) method (*randomForestSRC.rfsrc*) implemented in the *mlr* package. The 100 most important features were retained unless otherwise stated. In the first five iterations, the names of the peptide descriptors and the CPP features were ordered in ascending order. In the latter five iterations, on the other hand, the names of those features were ordered in descending order. This strategy minimizes the inherent bias in the process of feature selection originated from the lexicographical order of the names of features. Features selected in all iterations were retained and utilized in the subsequent analyses.

Immunogenicity score

We chose an extremely randomized tree (ERT) algorithm implemented in the *extraTrees* package (*55*) from preliminary experiments. ERT is a derivative of RF and is more computationally efficient and at the same time less prone to overfitting. The following model-specific parameters were determined from preliminary experiments: mtry=35, and, numRandomCuts=2. Class weights were set to compensate imbalanced class distributions. The minimum set of features, with or without predicted binding to HLA representatives, were utilized. Machine learning was conducted in the leave-one-chunk-out fashion; the entire dataset was split into five chunks of approximately same sizes, and the ERT model trained using the training subdataset corresponding to four chunks was used to predict the probabilities of immunogenicity for the peptides in the held-out chunk. The prediction was conducted for every chunk. This strategy ensures that any peptide is subjected to predictions by models trained from a set of peptides that do not contain the target peptide. To confirm the robustness of the predictions, the model training and prediction steps for the entire dataset were iterated five times with different random seeds. The means of the probability estimates were termed “immunogenicity scores.”

Coefficients of variance defined as standard deviations divided by means were also calculated. For peptides which are not included in our epitope metadataset, immunogenicity scores are defined as the means of probability estimates from a total of 25 ERT models generated from the five chunks in each of the five iterations.

Immune transition

Single amino acid changes (substitutions, insertions, and deletions) have multifaceted effects on the immunogenicity of epitopes and MHC binders. Loss of immunogenicity is called "escape" or "immune evasion." Meanwhile, the term "immunoediting" is used in the field of cancer (*27*). In this work, we unify the terminology as "escape." For the sake of simplicity, we do not take into consideration epitope-extrinsic mechanisms such as impaired intracellular antigen processing, somatic loss of HLA heterozygosity (*43*, *56*), and checkpoint-mediated T cell exhaustion and apoptosis (*44*). In contrast, there is no appropriate terminology for the acquisition of immunogenicity. In this work, we refer to this phenomenon as "epitope/neoepitope formation." Escaping and epitope formation can be understood as two sides of the same coin in sequence space, and therefore, we propose a higher-order concept, "immune transition," which is defined as a change in immunogenicity between a pair of MHC-presented peptides with just one edit distance. By definition, immune transition is bidirectional. Note that synergistic effects of two or more mutations on immune transition are beyond the scope of the present study. As for peptide-level immune transition status, peptides with at least one immune transition in the neighbor network constructed from our epitope metadataset were defined as "transitional," and those without evidence of immune transition were defined as "non-transitional."

Immune transition score

ERT model training for binary classification of transitional vs. non-transitional peptides and estimation of the probability of immune transition status were conducted with the same feature sets except predicted percentile ranks to HLA representatives as described in the previous sections.

Neighbor network analysis

In a neighbor network, each node represents a peptide sequence. A pair of peptide sequences with only one edit distance were considered linked neighbors. Each edge was directed from the peptide with a lower immunogenicity score to the peptide with a higher score. Edge weight was defined as the ratio of the lower score to the higher score so that less immunogenically similar peptides were mapped more distant from each other in sequence space due to the smaller edge weight.

Neighbor network clustering was conducted using a walktrap algorithm implemented in the *igraph* package (*57*). A consensus sequence per cluster was generated from the multiple sequence alignment generated using the ClustalW algorithm implemented in the *msa* package (*58*) in *Bioconductor* (*54*).

***In silico* mutagenesis**

To gain more precise estimates of immunogenicity for transitional peptides, we computationally expand their neighbor networks by simulating all possible single amino acid substituted mutant peptides. Insertions and deletions were not taken into consideration. Note that since our primary attention is on exploring the effects of those mutations on immunogenicity, *i.e.,* presentation onto MHC and recognition by TCR, we ignored the possibility that these artificially introduced mutations could instead affect factors outside the MHC-peptide-TCR axis

such as proteolytic degradation. Immunogenicity scores of mutated peptides were calculated, and neighbor network construction and clustering were conducted as described. Per-cluster mean immunogenicity scores were calculated as a more robust metric of immunogenicity.

HLA-I tumor neoepitopes

A total of 82 mutation-derived neoepitopes recognized by CTLs were collected from the review by Durgeau *et al.* (*30*). Immunogenicity scores were calculated as described in the previous sections.

Tumor neoepitope burden and survival after checkpoint blockade

The compiled dataset of predicted tumor neoepitopes and clinical annotations by Luksza *et al*. (*11*) was utilized. These data were derived from two melanoma cohorts and one non-small cell lung cancer (NSCLC) cohort treated either with anti-PD-1 and anti-CTLA-4 checkpoint inhibitors (*29*, *34*, *35*). Immunogenicity scores of wildtype (WT) and mutant (MT) peptides were calculated as described in the previous sections. Two immunogenicity-based metrics were attempted: the immunogenicity scores of MT peptides, and the absolute differences of scores normalized by the sums of WT and MT scores. Neoepitope candidates were filtered based on these metrics with an arbitrary threshold, and neoepitope burden per patient was estimated. Again, neoepitope burden was dichotomized with an arbitrary threshold. The significance of the stratification of overall survival after checkpoint blockade immunotherapy was calculated by the log-rank test. In this manner, the optimal combination of thresholds for neoepitope filtering and dichotomization that minimized the *P*-value by the log-rank test was determined.

Statistical analyses

When calculating the descriptive statistics in the CPP analysis, the *describe* function implemented in the *psych* package was utilized. Non-parametric hypothesis testing methods were consistently employed unless otherwise stated. For continuous variables, Kruskal-Wallis test and Wilcoxon rank sum test were employed, with FDR adjustment according to the method proposed by Benjamini and Hochberg (*59*). When reporting measures of central tendency and dispersion, data were presented as mean ± SD unless otherwise stated. AUC values were calculated based on ten-fold cross-validations, and their 95% confidence intervals were inferred from influence curves using the method implemented in the *cvAUC* package (*24*). Survival period analysis was conducted using the *survival* package. Survival distributions between groups were compared by the log-rank test using the *survdiff* function implemented in the *survival* package. All statistical analyses were conducted in R (*47*).

Supplementary Materials

Fig. S1. MHC-I immunogenicity prediction by the IEDB Class I Immunogenicity tool.

Fig. S2. Feature importance.

Fig. S3. Normalized score variances between iterations.

Fig. S4. Reduction of features only minimally affected the predictive power on immunogenicity.

Fig. S5. Elimination of highly similar peptides significantly undermined the predictive power on immunogenicity.

Fig. S6. Immunogenicity prediction on the peptide datasets from other species.

Fig. S7. Evaluation of various immunogenicity-based criteria to curtail neoepitope burden in checkpoint blockade immunotherapy cohorts.

Fig. S8. Non-filtered predicted neoepitope burden was associated with overall survival after checkpoint blockade immunotherapy.

Fig. S9. Neoepitope burden filtered based on immunogenicity score was associated with overall survival after checkpoint blockade immunotherapy.

Fig. S10. Reduction of neoepitope burden per patient stratified by burden groups.

Table S1. A list of evidence of T cell response for the peptide dataset derived from the Immune Epitope Database.

Table S2. Amino acid pairwise contact potential scales utilized in this study.

References and Notes:

1. A. Grakoui *et al.*, The immunological synapse: a molecular machine controlling T cell activation. *Science*. **285**, 221–7 (1999).

2. M. G. Rudolph, R. L. Stanfield, I. A. Wilson, How TCRs bind MHCs, peptides, and coreceptors. *Annu. Rev. Immunol.* **24**, 419–66 (2006).

3. K. C. Garcia, L. Teyton, I. A. Wilson, Structural basis of T cell recognition. *Annu. Rev. Immunol.* **17**, 369–97 (1999).

4. M. Andreatta, M. Nielsen, Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics*. **32**, 511–517 (2016).

5. K. K. Jensen *et al.*, Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* (2018), doi:10.1111/imm.12889.

6. A. Sette *et al.*, The relationship between class I binding affinity and immunogenicity of potential cytotoxic T cell epitopes. *J. Immunol*. **153**, 5586–92 (1994).

7. J. Schmidt *et al.*, In silico and cell-based analyses reveal strong divergence between prediction and observation of T-cell-recognized tumor antigen T-cell epitopes. *J. Biol. Chem.* **292**, 11840–11849 (2017).

8. R. Mazor, C.-H. Tai, B. Lee, I. Pastan, Poor correlation between T-cell activation assays and HLA-DR binding prediction algorithms in an immunogenic fragment of Pseudomonas exotoxin A. *J. Immunol. Methods*. **425**, 10–20 (2015).

9. J. J. A. Calis *et al.*, Properties of MHC Class I Presented Peptides That Enhance Immunogenicity. *PLoS Comput. Biol.* **9** (2013), doi:10.1371/journal.pcbi.1003266.

10. D. Chowell *et al.*, TCR contact residue hydrophobicity is a hallmark of immunogenic CD8 + T cell epitopes. *Proc. Natl. Acad. Sci.* **112**, E1754–E1762 (2015).

11. M. Luksza *et al.*, A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. *Nature*. **551**, 517–520 (2017).

12. V. P. Balachandran *et al.*, Identification of unique neoantigen qualities in long-term survivors of pancreatic cancer. *Nature*. **551**, S12–S16 (2017).

13. M. Rasmussen *et al.*, Pan-Specific Prediction of Peptide–MHC Class I Complex Stability, a Correlate of T Cell Immunogenicity. *J. Immunol.* **197**, 1517–1524 (2016).

14. S. Tian, R. Maile, E. J. Collins, J. A. Frelinger, CD8+ T Cell Activation Is Governed by TCR-Peptide/MHC Affinity, Not Dissociation Rate. *J. Immunol.* **179**, 2952–2960 (2007).

15. C. Kuiken, K. Yusim, L. Boykin, R. Richardson, The Los Alamos hepatitis C sequence database. *Bioinformatics*. **21**, 379–384 (2005).

16. S. Lata, M. Bhasin, G. P. S. Raghava, MHCBN 4.0: A database of MHC/TAP binding peptides and T-cell epitopes. *BMC Res. Notes*. **2**, 61 (2009).

17. A. Llano, A. Williams, A. Olvera, S. Silva-Arrieta, C. Brander, Best-Characterized HIV-1 CTL Epitopes: The 2013 Update. *HIV Mol. Immunol. 2013*, 3–25 (2013).

18. L. R. Olsen *et al.*, TANTIGEN: a comprehensive database of tumor T cell antigens. *Cancer Immunol. Immunother*. **66**, 731–735 (2017).

19. P. A. Reche, H. Zhang, J. P. Glutting, E. L. Reinherz, EPIMHC: A curated database of MHC-binding peptides for customized computational vaccinology. *Bioinformatics*. **21**, 2140–2141 (2005).

20. C. W. Tung *et al.*, POPISK: T-cell reactivity prediction using support vector machines and string kernels. *BMC Bioinformatics*. **12**, 446 (2011).

21. W. Fleri *et al.*, The immune epitope database and analysis resource in epitope discovery and synthetic vaccine design. *Front. Immunol.* **8**, 1–16 (2017).

22. M. R. Betancourt, D. Thirumalai, Pair potentials for protein folding: choice of reference states and sensitivity of predicted native states to variations in the interaction schemes. *Protein Sci.* **8**, 361–369 (1999).

23. M. Vendruscolo, E. Domany, Pairwise contact potentials are unsuitable for protein folding. *J. Chem. Phys.* **109**, 11101 (1998).

24. E. Ledell, M. Petersen, M. Van Der Laan, Computationally efficient confidence intervals for cross-validated area under the ROC curve estimates. *Electron. J. Stat.* **9**, 1583–1607 (2015).

25. D. G. Bowen, C. M. Walker, Adaptive immune responses in acute and chronic hepatitis C virus infection. *Nature*. **436**, 946–952 (2005).

26. Y. Kawashima *et al.*, Adaptation of HIV-1 to human leukocyte antigen class I. *Nature*. **458**, 641–645 (2009).

27. M. Efremova *et al.*, Targeting immune checkpoints potentiates immunoediting and changes the dynamics of tumor evolution. *Nat. Commun.* **9** (2018), doi:10.1038/s41467-017-02424-0.

28. N. McGranahan *et al.*, Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science*. **351**, 1463–9 (2016).

29. N. A. Rizvi *et al.*, Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science (80-.).* **348**, 124–128 (2015).

30. A. Durgeau, Y. Virk, S. Corgnac, F. Mami-Chouaib, Recent advances in targeting CD8 T-cell immunity for more effective cancer immunotherapy. *Front. Immunol.* **9** (2018), doi:10.3389/fimmu.2018.00014.

31. M. M. Gubin *et al.*, Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature*. **515**, 577–581 (2014).

32. E. M. E. Verdegaal *et al.*, Neoantigen landscape dynamics during human melanoma–T cell interactions. *Nature*. **536**, 91–95 (2016).

33. D. T. Le *et al.*, Mismatch-repair deficiency predicts response of solid tumors to PD-1 blockade. *Science*. **6733**, 1–11 (2017).

34. A. Snyder *et al.*, Genetic Basis for Clinical Response to CTLA-4 Blockade in Melanoma. *N. Engl. J. Med.*, 2189–2199 (2014).

35. E. M. Van Allen *et al.*, Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science (80-.).* **350**, 207–211 (2015).

36. P. A. Ott *et al.*, An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature*. **547**, 217–221 (2017).

37. A. M. Bjerregaard *et al.*, An analysis of natural T cell responses to predicted tumor neoepitopes. *Front. Immunol.* **8**, 1–9 (2017).

38. L. Wojciech *et al.*, The same self-peptide selects conventional and regulatory CD4+ T cells with identical antigen receptors. *Nat. Commun.* **5**, 5061 (2014).

39. J. Attanasio, E. J. Wherry, Costimulatory and Coinhibitory Receptor Pathways in Infectious Disease. *Immunity*. **44**, 1052–1068 (2016).

40. R. Eil *et al.*, Ionic immune suppression within the tumour microenvironment limits T cell effector function. *Nature*. **537**, 539–543 (2016).

41. C. H. Chang *et al.*, Metabolic Competition in the Tumor Microenvironment Is a Driver of Cancer Progression. *Cell*. **162**, 1229–1241 (2015).

42. I. F. Luescher *et al.*, CD8 modulation of T-cell antigen receptor-ligand interactions on living cytotoxic T lymphocytes. *Nature*. **373**, 353–6 (1995).

43. N. McGranahan *et al.*, Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. *Cell*, 1259–1271 (2017).

44. D. M. Pardoll, The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer*. **12**, 252–264 (2012).

45. J. M. Pitt *et al.*, Resistance Mechanisms to Immune-Checkpoint Blockade in Cancer: Tumor-Intrinsic and -Extrinsic Factors. *Immunity*. **44**, 1255–69 (2016).

46. L. Zitvogel, M. Ayyoub, B. Routy, G. Kroemer, Microbiome and Anticancer Immunosurveillance. *Cell*. **165**, 276–287 (2016).

47. R Core Team, R: A Language and Environment for Statistical Computing (2018), , doi:ISBN 3-900051-07-0.

48. J. J. Boniface, Z. Reich, D. S. Lyons, M. M. Davis, Thermodynamics of T cell receptor binding to peptide-MHC: evidence for a general mechanism of molecular scanning. *Proc Natl Acad Sci U S A*. **96**, 11446–11451 (1999).

49. S. Zhong *et al.*, T-cell receptor affinity and avidity defines antitumor response and autoimmunity in T-cell immunotherapy. *Proc. Natl. Acad. Sci.* **110**, 6973–6978 (2013).

50. S. Kawashima *et al.*, AAindex: Amino acid index database, progress report 2008. *Nucleic Acids Res.* **36**, 202–205 (2008).

51. D. K. Cole *et al.*, T-cell Receptor (TCR)-peptide specificity overrides affinity-enhancing TCR-major histocompatibility complex interactions. *J. Biol. Chem.* **289**, 628–638 (2014).

52. O. V. Britanova *et al.*, Age-Related Decrease in TCR Repertoire Diversity Measured with Deep and Normalized Sequence Profiling. *J. Immunol.* **192**, 2689–2698 (2014).

53. D. A. Bolotin *et al.*, MiXCR: software for comprehensive adaptive immunity profiling. *Nat. Methods*. **12**, 380–381 (2015).

54. R. C. Gentleman *et al.*, Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80 (2004).

55. P. Geurts, D. Ernst, L. Wehenkel, Extremely randomized trees. *Mach. Learn.* **63**, 3–42 (2006).

56. D. Chowell *et al.*, Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. *Science (80-.).* **4572**, eaao4572 (2017).

57. G. Csárdi, T. Nepusz, The igraph software package for complex network research. *InterJournal*. **Complex Sy**, 1695 (2006).

58. U. Bodenhofer, E. Bonatesta, C. Horejš-Kainrath, S. Hochreiter, msa: an R package for multiple sequence alignment. *Bioinformatics*. **31**, btv494 (2015).

59. Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B.* **57**, 289–300 (1995).

Acknowledgments: We thank Dr. Couture-Cossette and Ueno for helpful advice. **Funding**: This work is not supported by any external funding sources. **Author contributions**: M.O. conceived the idea, designed the study, performed computational analyses, and drafted the manuscript; M.O. and H.Y. wrote the manuscript. **Competing interests**: The authors declare no competing interests. **Data availability**: Compiled datasets and essential in-house functions are available as the R package *Repitope* on GitHub (https://github.com/masato-ogishi/Repitope/). Other analytical scripts are available upon request.