Unbiased Genome-wide Discovery Using TScan Reveals Shared Immunodominant CD8+ T Cell Epitopes in SARS-CoV-2

Andrew P. Ferretti[1], Tomasz Kula[1], Yifan Wang[1], Dalena M.V. Nguyen[1], Adam Weinheimer[1], Garrett S. Dunlap[1], Qikai Xu[1], Nancy Nabilsi[1], Candace R. Perullo[1], Alexander W. Cristofaro[1], Holly J. Whitton[1], Amy Virbasius[1], Kenneth J. Olivier Jr[1], Lyndsey R. Buckner[2], Angela T. Alistar[3], Eric D. Whitman[3], Sarah A. Bertino[1], Shrikanta Chattopadhyay[1], Gavin MacBeath[1]

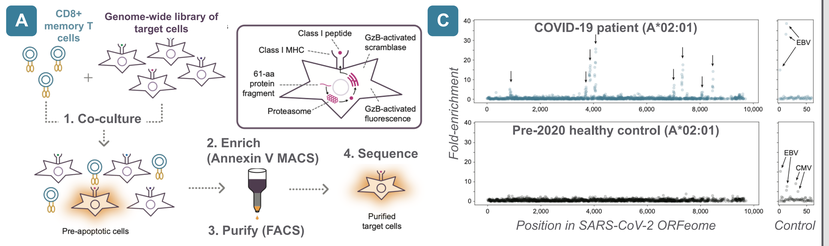
[1]TScan Therapeutics, Waltham, MA, [2]Ochsner Medical Center, New Orleans, LA; [3]Atlantic Health System, Morristown, NJ

See also https://www.medrxiv.org/content/10.1101/2020.07.24.20161653v2

Abstract

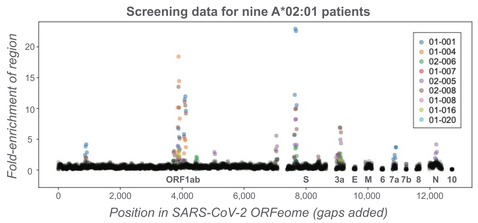
Developing effective strategies to prevent or treat COVID-19 requires a detailed understanding of the natural immune response to SARS-CoV-2, including how CD8+ T cells recognize the virus. We used an unbiased, genome-wide screening technology to determine the global landscape and exact peptide sequences recognized by the memory CD8+ T cells of 25 COVID-19 convalescent patients, focusing on the six most prevalent human leukocyte antigen (HLA) types. Overall, we find that there are 3–8 immunodominant (shared) epitopes for each HLA type, and that these antigens reside in regions of the virus that are not subject to mutational variation. In contrast to recent studies of CD4+ T cells, we find that CD8+ T cells generally do not cross-react with epitopes in the four endemic coronaviruses that cause the common cold, arguing that pre-existing immunity to other coronaviruses does not significantly shape CD8+ T cell responses to SARS-CoV-2. Notably, we find that only ~10% of immunodominant epitopes (3 of 29) reside in the S protein, supporting the development of next-generation vaccines that better recapitulate natural CD8+ T cell immunity to SARS-CoV-2.

TScan approach for comprehensive mapping of the CD8+ T cell response to SARS-CoV-2



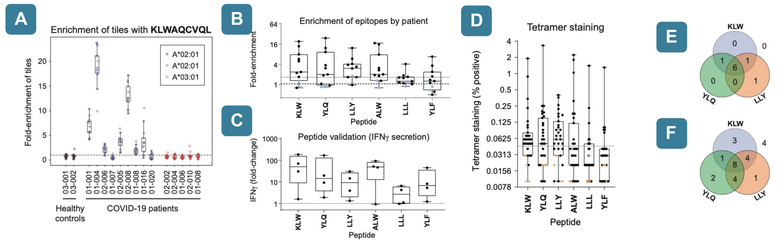
(A) Overview of the T-Scan antigen discovery screen. (B) Design of the ORFeome-wide SARS-CoV-2 antigen library. (C) Example SARS-CoV-2 ORFeome-wide T-Scan screen data for a convalescent COVID-19 patient (top panel) and healthy control (bottom panel). Each circle represents a single 61-aa SARS-CoV-2 protein fragment, with the x-axis showing the position of each fragment in the concatenated SARS-CoV-2 ORFeome. The y-axis shows the performance of the fragment in the screen, calculated as the ratio of sorted target cells expressing the protein fragment relative to the unsorted target library. The right panels show the performance of the 60 positive control protein fragments derived from CMV, EBV, and Influenza.

Discovery of immunodominant SARS-CoV-2 epitopes for HLA-A*02:01 presented on HLA-A*0201



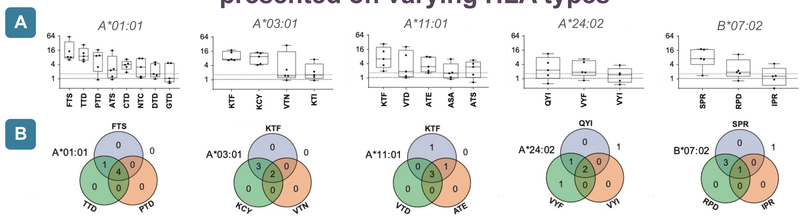
CD8+ T cell reactivity across the SARS-CoV-2 ORFeome for nine HLA-A*02:01 COVID-19 patients. In keeping with other viruses, specific fragments of SARS-CoV-2 are recurrently recognized by the T cells of multiple patients (1). Each circle corresponds to a 20-aa segment of the SARS-CoV-2 ORFeome, with the x-axis indicating the position of the segment in the ORFeome (gaps added for display purposes). Results for each patient are denoted with different colors.

Validation of immunodominant SARS-CoV-2 epitopes



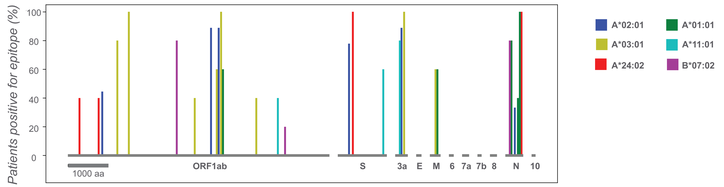
Specific fragments of SARS-CoV-2 are recurrently recognized by the T cells of multiple patients (i.e., are immunodominant). (A) For example, ORF1ab aa 3881-3900 is recognized by 7 of 9 A*02:01 patients, but not by healthy controls or A*03:01 patients. (B) In total, six regions are targeted by CD8+ T cells from at least three patients. In each case, the exact peptide epitope in the region was identified using NetMHC4.0 and validated by peptide-dependent T-cell activation as determined by interferon-γ secretion (C). As further validation, we constructed MHC tetramers with the six peptides and used them to stain the memory CD8+ T cells of all nine A*02:01 patients, as well as an additional test-set of 18 A*02:01 patients that had not been screened (D). The top three epitopes are recognized by most screened patients (E) and test-set patients (F).

Discovery and validation of immunodominant epitopes presented on varying HLA types



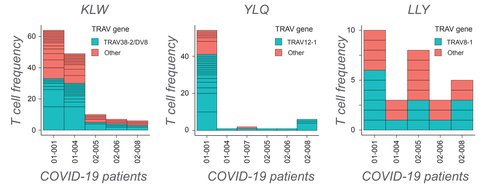
As CD8+ T cell responses are profoundly shaped by host MHC alleles, we next mapped memory CD8+ T cell reactivities for five additional MHC alleles representing 90% of the U.S. population. We again observed recurrent recognition of specific protein fragments by most or all patients for each allele, indicating a narrow set of immunodominant responses. (A) Collapsed T-Scan screen data and (B) recognition of the three most common epitopes for each prioritized MHC, as described above.

Immunodominant epitopes span the SARS-CoV-2 ORFeome



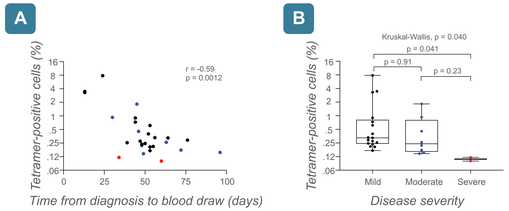
Distribution of immunodominant CD8+ T cell epitopes across the SARS-CoV-2 genome. We observed broad reactivity to many SARS-CoV-2 proteins, including ORF1ab, S, N, M, and ORF3a. Notably, only three of the 29 epitopes are located in the S protein, with most located in ORF1ab and the highest density of epitopes located in the N protein. Each bar represents one validated epitope, with the color indicating its MHC restriction and the height of the bar indicating the percentage of MHC-matched patients recognizing the epitope.

Dominant epitopes are recognized by TCRs with shared features



We next examined the features of the TCRs recognizing the identified immunodominant epitopes. We isolated epitope-specific T cells and used single cell sequencing to determine the paired TCR α and β chains expressed by these T cells. We observed oligoclonal recognition of each epitope and found that the same TCR α variable (TRAV) genes were used repeatedly across patients. In the figure, the height of each box corresponds to the number of T cells within each clonotype.

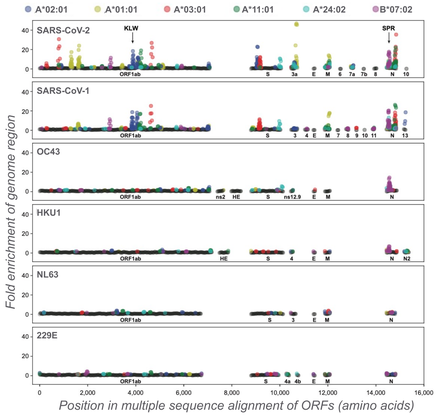
Correlations of T cell response with clinical characteristics



(A) Magnitude of detected memory CD8+ T cell response correlates negatively with time from diagnosis to blood draw. (B) Magnitude of response correlates negatively with disease severity. For both panels, the magnitude of response indicates the fraction of memory CD8+ T cells that stain positive with tetramers for one of the six identified HLA-A*02:01 epitopes (KLW, YLQ, LLY, ALW, LLL, YLF) for 27 HLA-A*02:01 patients.

Minimal cross-reactivity of SARS-CoV-2-reactive memory CD8+ T cells with other coronaviruses

We examined T-cell reactivity to SARS-CoV-2, SARS-CoV-1, and all four endemic coronaviruses (OC43, HKU1, NL63, or 229E) in the 34 genome-wide screens that we conducted across all patients and MHC alleles. We observed broad reactivity to the corresponding epitopes in SARS-CoV-1 in over half of cases, consistent with a recent study reporting the existence of long-lasting memory T cells cross-reactive to SARS-CoV-2 in patients that had been infected by SARS-CoV-1 during the 2002/2003 SARS outbreak (2). In contrast, we detected almost no reactivity to the endemic coronaviruses, suggesting that prior exposure to these viruses is unlikely to provide CD8+ T cell-based protection from SARS-CoV-2. Each panel shows the collective reactivity to one coronavirus genome.



Conclusion

Although studies using peptide pools have uncovered CD8+ T cell responses in COVID-19 patients, it has not been clear which specific epitopes are being recognized. Our genome-wide screens provide an unbiased view of the overall landscape of T cell reactivity and reveal that memory CD8+ T cells generally recognize a relatively small set of immunodominant epitopes: 3-8 for each of the six most common MHC alleles. Notably, only ~10% of these epitopes reside in the Spike protein, highlighting the need for second-generation vaccines that more fully recapitulate the natural CD8+ T cell response to SARS-CoV-2 infection.

References: (1) Yewdell, 2006, *Immunity*, **25**, 533; (2) Le Bert et al., 2020, *Nature*, **584**, 457.