Supplemental Information

Estimating the Binding of Sars-CoV-2 Peptides

to HLA Class I in Human Subpopulations

Using Artificial Neural Networks

Caterina A.M. La Porta and Stefano Zapperi
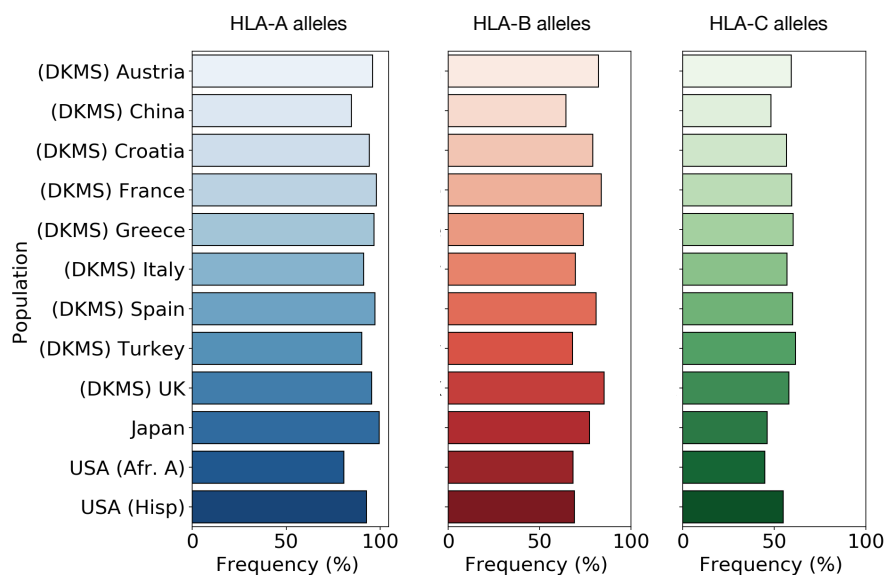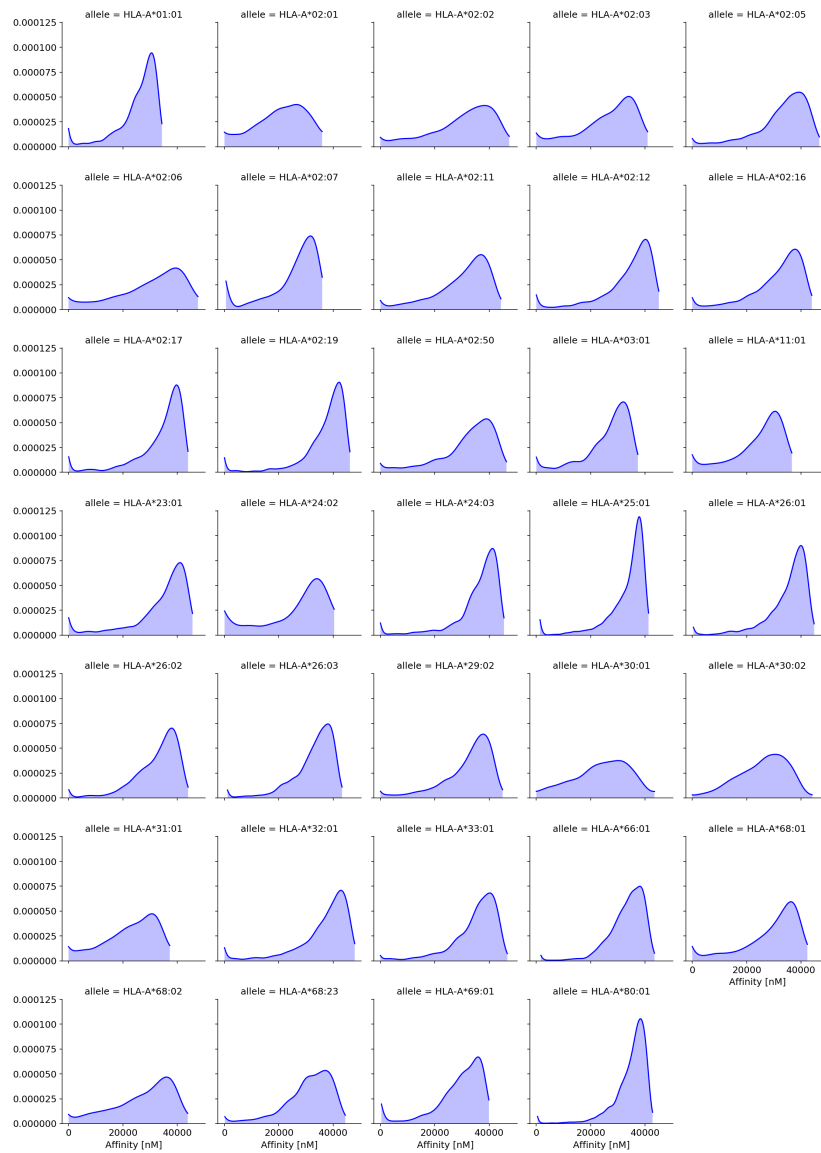
Supplementary Information

Figure S1: **Representation across human populations of the 79 HLA alleles studied in the present work. Related to Fig. 1** Percentage of the population that has one of the 79 a) HLA-A, b) HLA-B or c) HLA-C alleles.

Figure S2: **Distribution of binding affinities for SARS-Cov-2 peptides and HLA-A molecules. Related to Fig. 1** The probability distributions are estimated using using a Gaussian kernel.
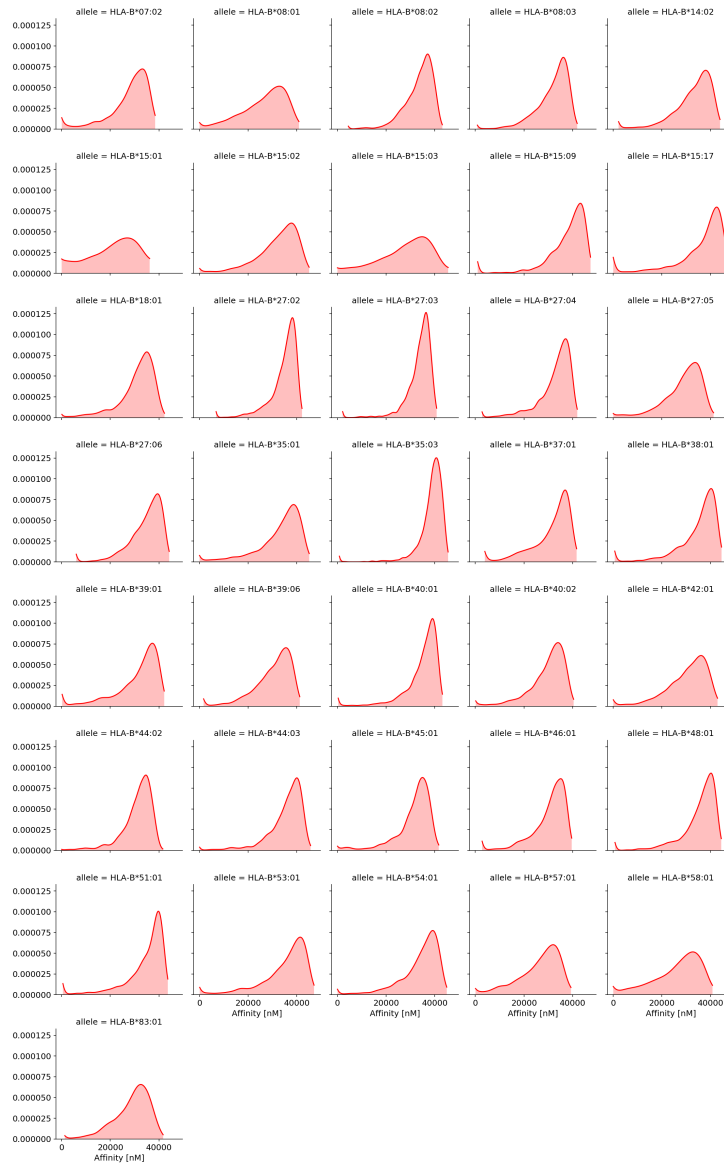
Figure S3: **Distribution of binding affinities for SARS-Cov-2 peptides and HLA-B molecules. Related to Fig. 1** The probability distributions are estimated using using a Gaussian kernel.
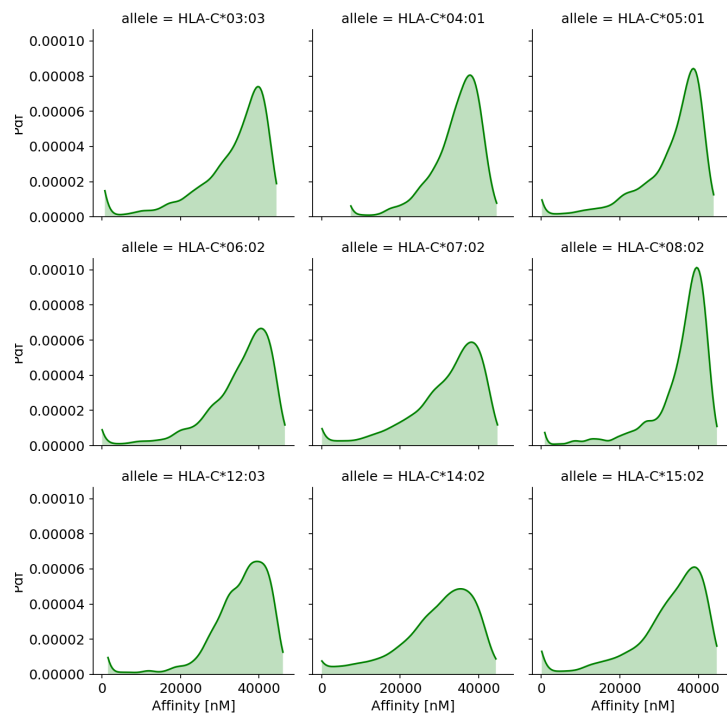
Figure S4: **Distribution of binding affinities for SARS-Cov-2 peptides and HLA-C molecules. Related to Fig. 1** The probability distributions are estimated using using a Gaussian kernel.
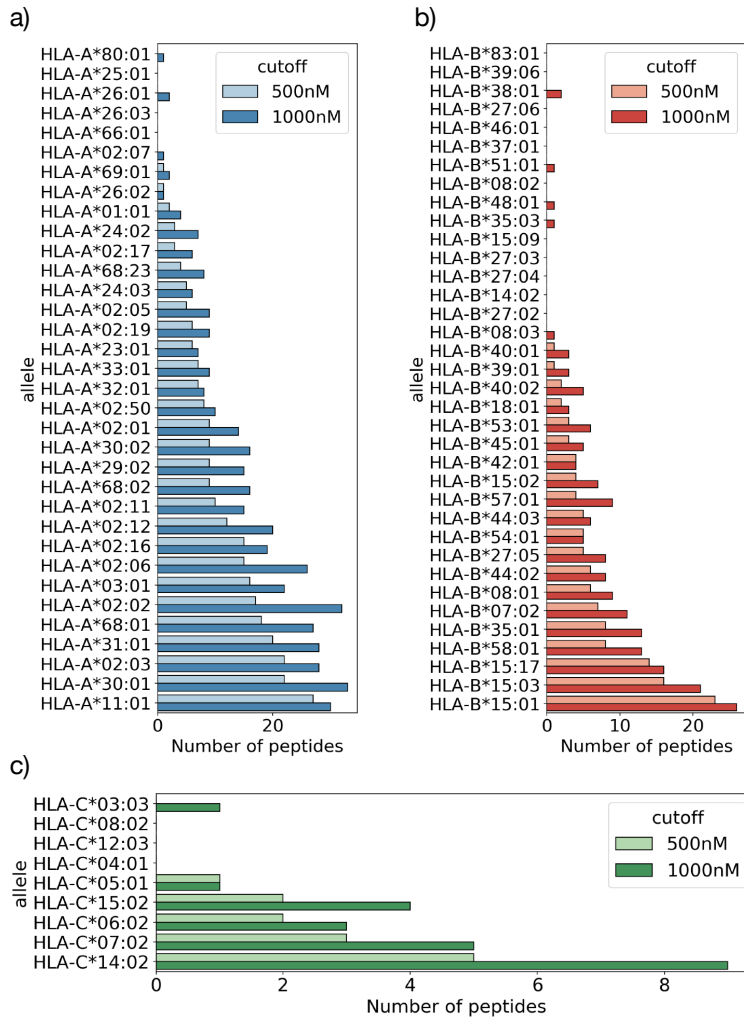
Figure S5: **Role of the cutoff in the number of strongly binding peptides. Related to Fig. 1** The number of strongly binding peptides for SARS-Cov-2 estimated for 79 Class I HLA alleles by combining predictions from netMHCpan and MHCflurry. The results are obtained using two different cutoff values ($IC_{50} < 1000$nM and $IC_{50} < 500$nM) for a) HLA-A, b) HLA-B and c) HLA-C molecules.
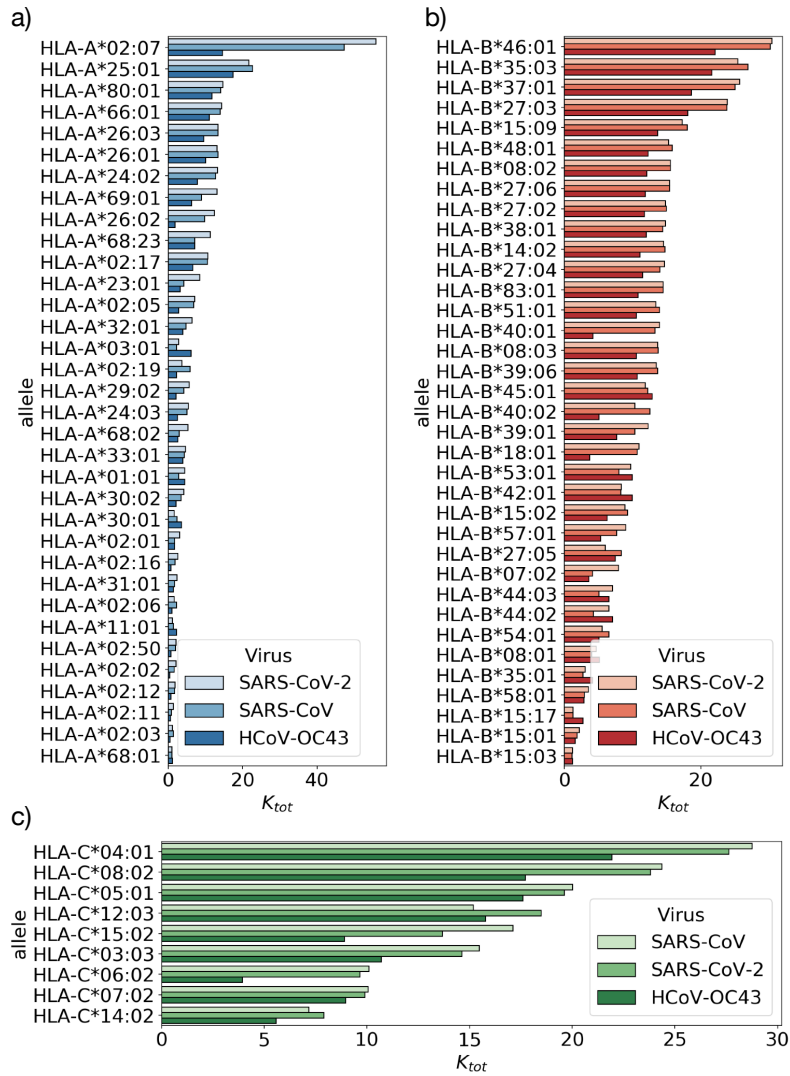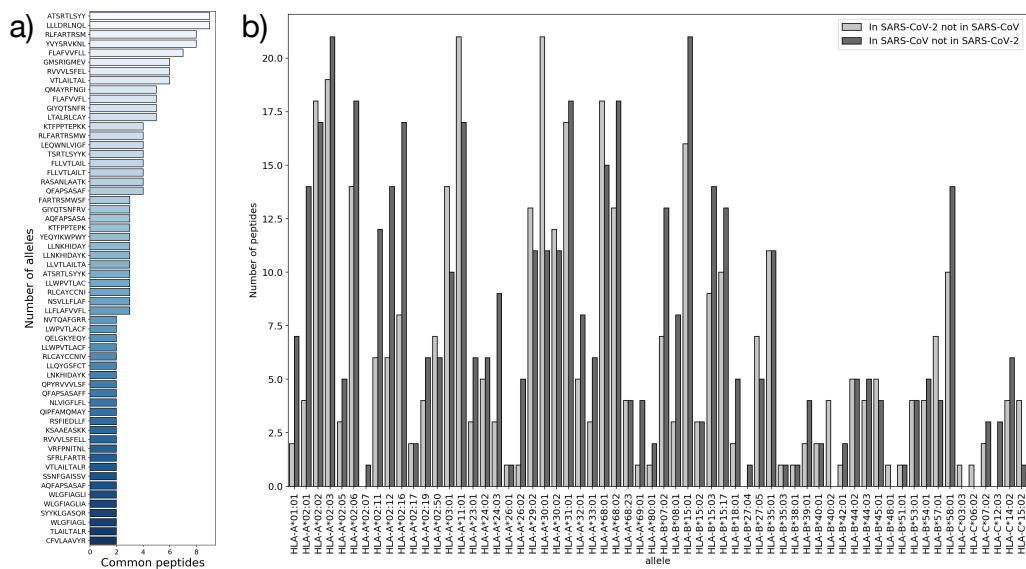
Figure S6: **Total binding affinity for each HLA molecule. Related to Fig. 1** The total binding affinity for SARS-Cov-2, SARS-Cov and HCOV-OC43 peptides estimated for 79 Class I HLA alleles by combining predictions from netMHCpan and MHCflurry. Results for a) HLA-A, b) HLA-B and c) HLA-C molecules.

Figure S7: **Strongly binding peptides in SARS-CoV-2 and SARS-CoV. Related to Fig. 1** a) The list of peptides that bind strongly to multiple HLA molecules for both SARS-CoV-2 and SARS-CoV (affinity less than 1000 nM). Peptides are ranked according to the number of common HLA molecules to which they bind strongly. b) The number of peptides that bind strongly only for either one between SARS-CoV-2 and SARS-CoV are reported for each of 79 alleles studied.
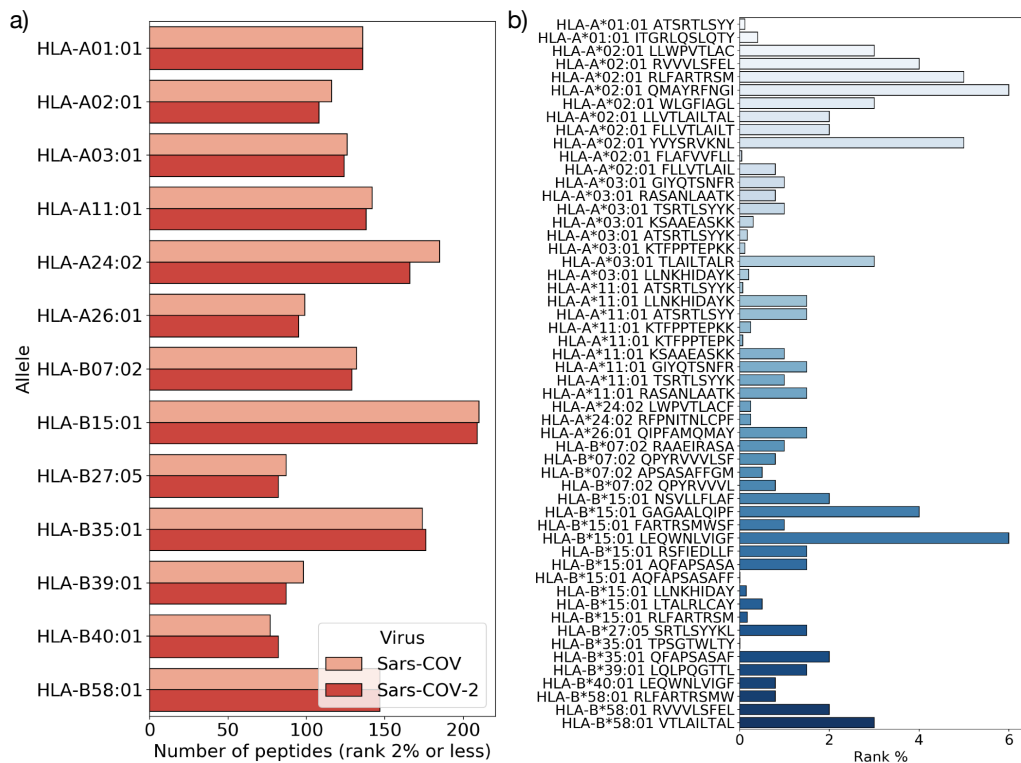
Figure S8: **Potential T-cell epitopes are shared between SARS-CoV-2 and SARS-CoV. Related to Fig. 1** a) The number of potential T-cell epitopes for SARS-CoV-2 and SARS-CoV peptides estimated with netTepi (see Methods). b) Highly ranked peptides for different HLA alleles that are in common for SARS-CoV-2 and SARS-CoV.
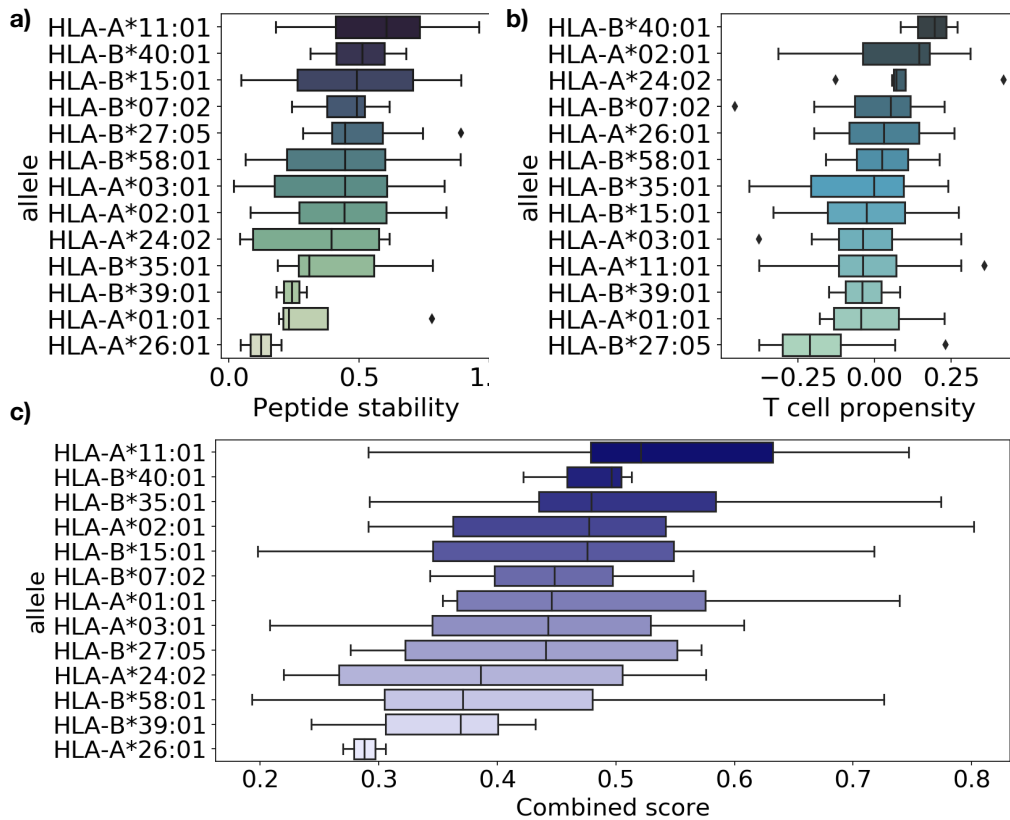
Figure S9: **T-cell epitopes for SARS-CoV-2. Related to Fig. 1** The distribution of a) peptide stability, b) T cell propensity and c) combined T-cell epitope score computed by netTepi (see Methods) for different HLA alleles and all the peptides from SARS-CoV-2 structural proteins.
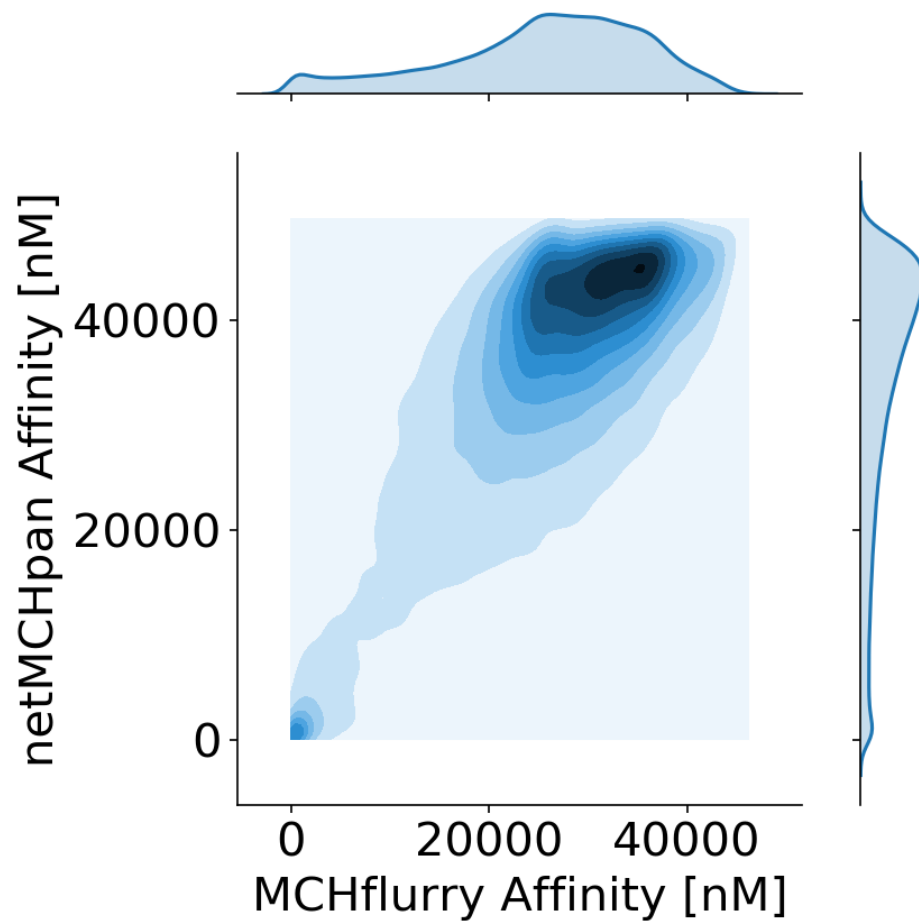
Figure S10: **Comparison of binding affinity predictions. Related to Fig. 1** Joint density estimates for the binding affinity predictions of netMHCpan and MHCflurry.