iScience, Volume *26*

Supplemental information

Multiple instance learning to predict immune checkpoint blockade efficacy using neoantigen candidates

Franziska Lang, Patrick Sorn, Barbara Schrörs, David Weber, Stefan Kramer, Ugur Sahin, and Martin Löwer

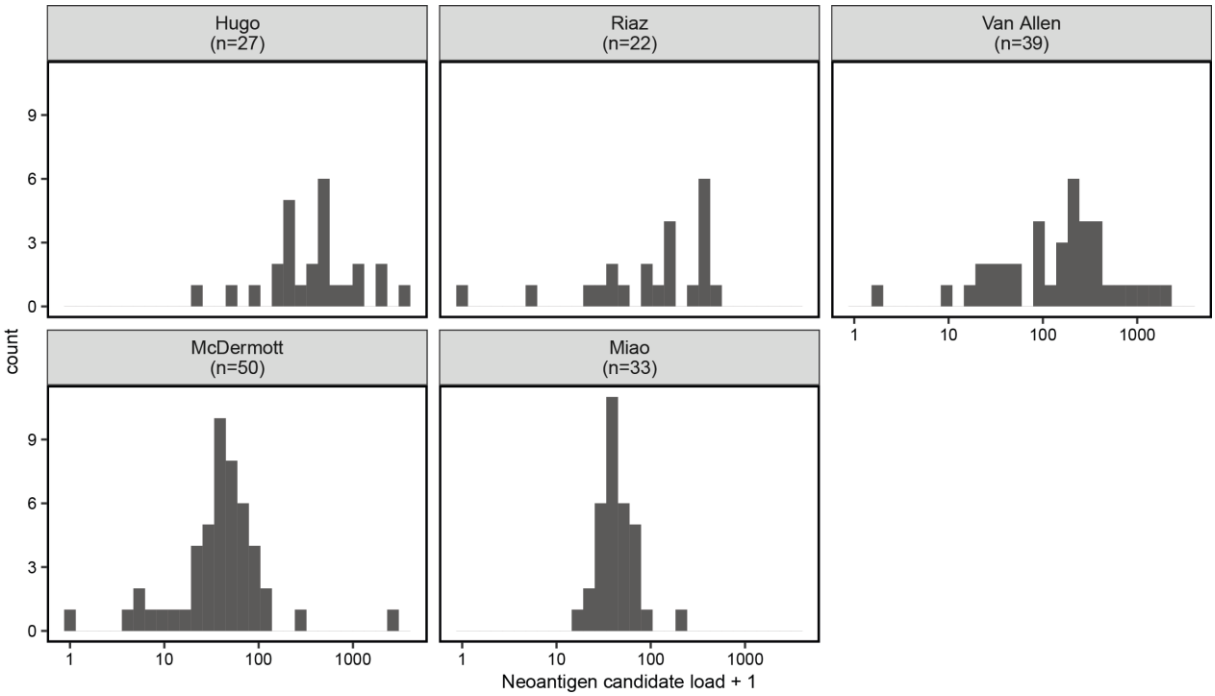

Figure S1: The distribution of the neoantigen candidate load from SNVs for each individual ICB cohort investigated in this study, related to Figure 2.

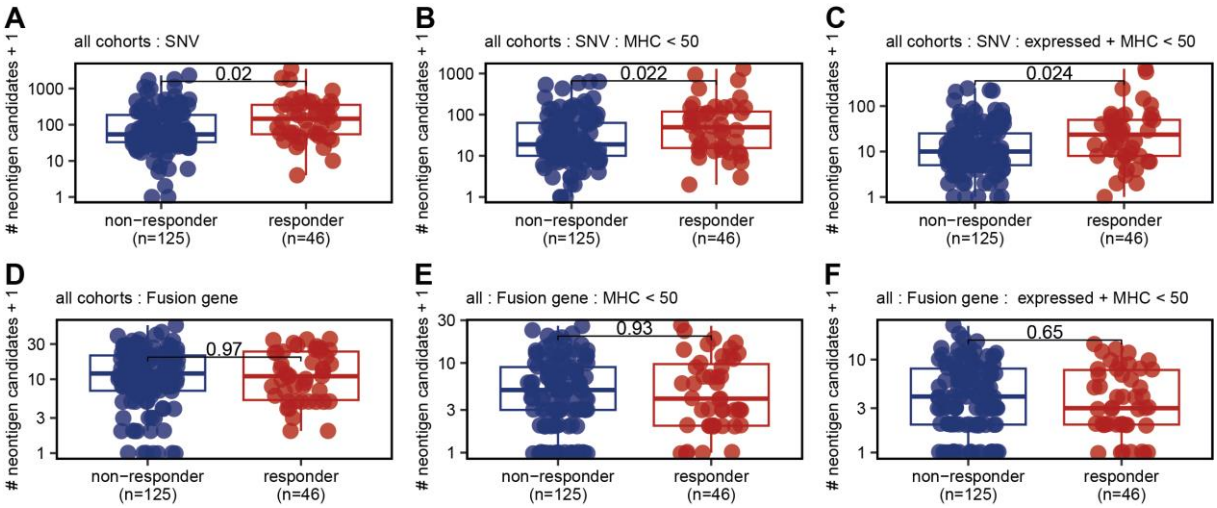

Figure S2: Comparison of the neoantigen candidate load between responders and non-responders to ICB therapy, related to Figure 3. (A-C) The SNV-derived neoantigen candidate burden was compared between responder and non-responder in a combined dataset of all ICB cohorts based on (**A**) all predicted neoantigen candidates, (**B**) candidates with MHC-I or MHC-II binding affinity < 50 nM and (**C**) expressed candidates with MHC-I or MHC-II binding affinity < 50 nM. (**D-F**) The fusion gene-derived neoantigen candidate burden was compared between responder and non-responder in a combined dataset of all ICB cohorts based on (**D**) all predicted neoantigen candidates, (**E**) candidates with MHC-I or MHC-II binding affinity < 50 nM and (**F**) expressed candidates with MHC-I or MHC-II binding affinity < 50 nM. Statistical testing was performed with Wilcoxon signed ranked test. P-values were corrected with multiple testing correction using the Benjamini Hochberg method. Statistical tests resulting in p-values < 0.05 after multiple testing correction were considered as significant.

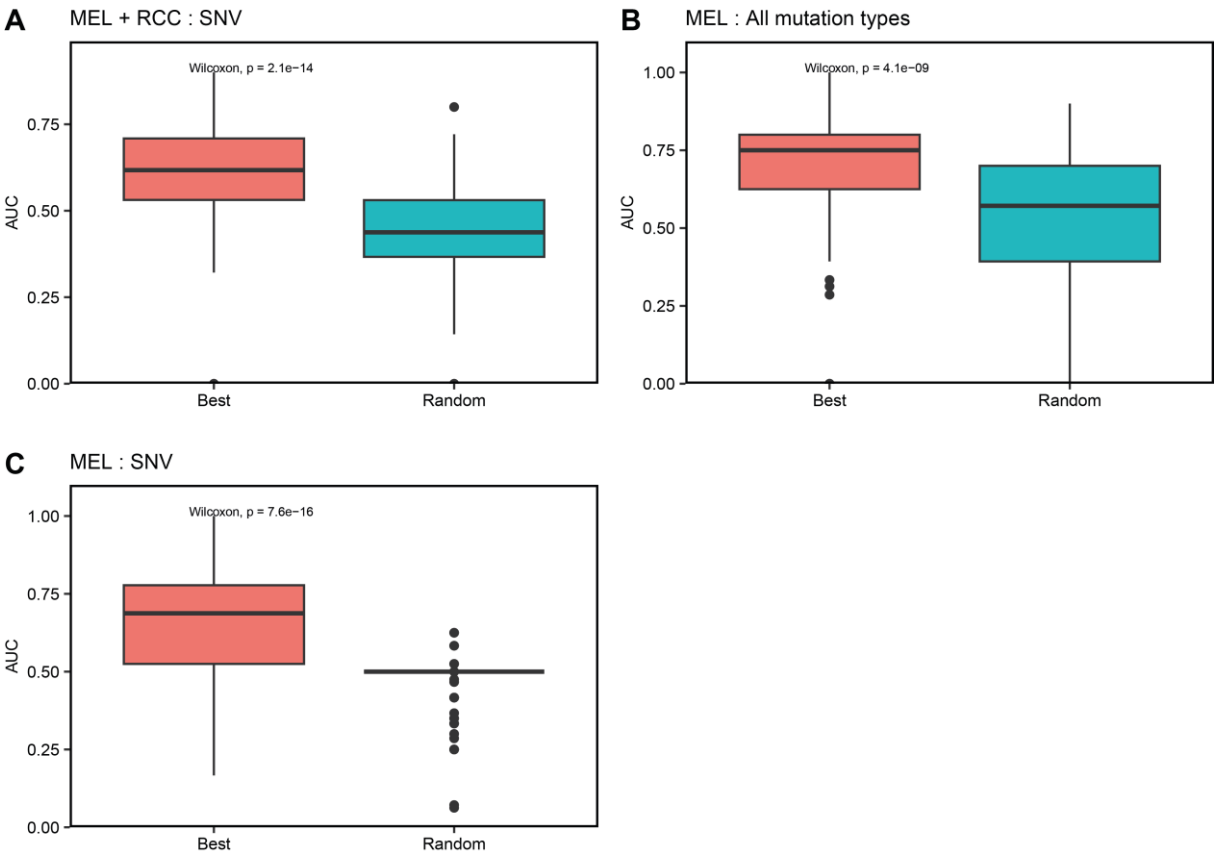

Figure S3: Performance of MILES on randomized datasets, related to Figure 4. (A-C) The optimal hyperparameter sets were used to train and evaluate the performance of MILES on randomized datasets of **(A)** all SNV-derived neoantigen candidates from MEL+ RCC cohort, of **(B)** all neoantigen candidates from MEL cohort and of **(C)** all SNV-derived neoantigen candidates from MEL cohort. Randomized refers to the randomization of neoantigen candidates across patients while keeping the original number of neoantigen candidates per patient.

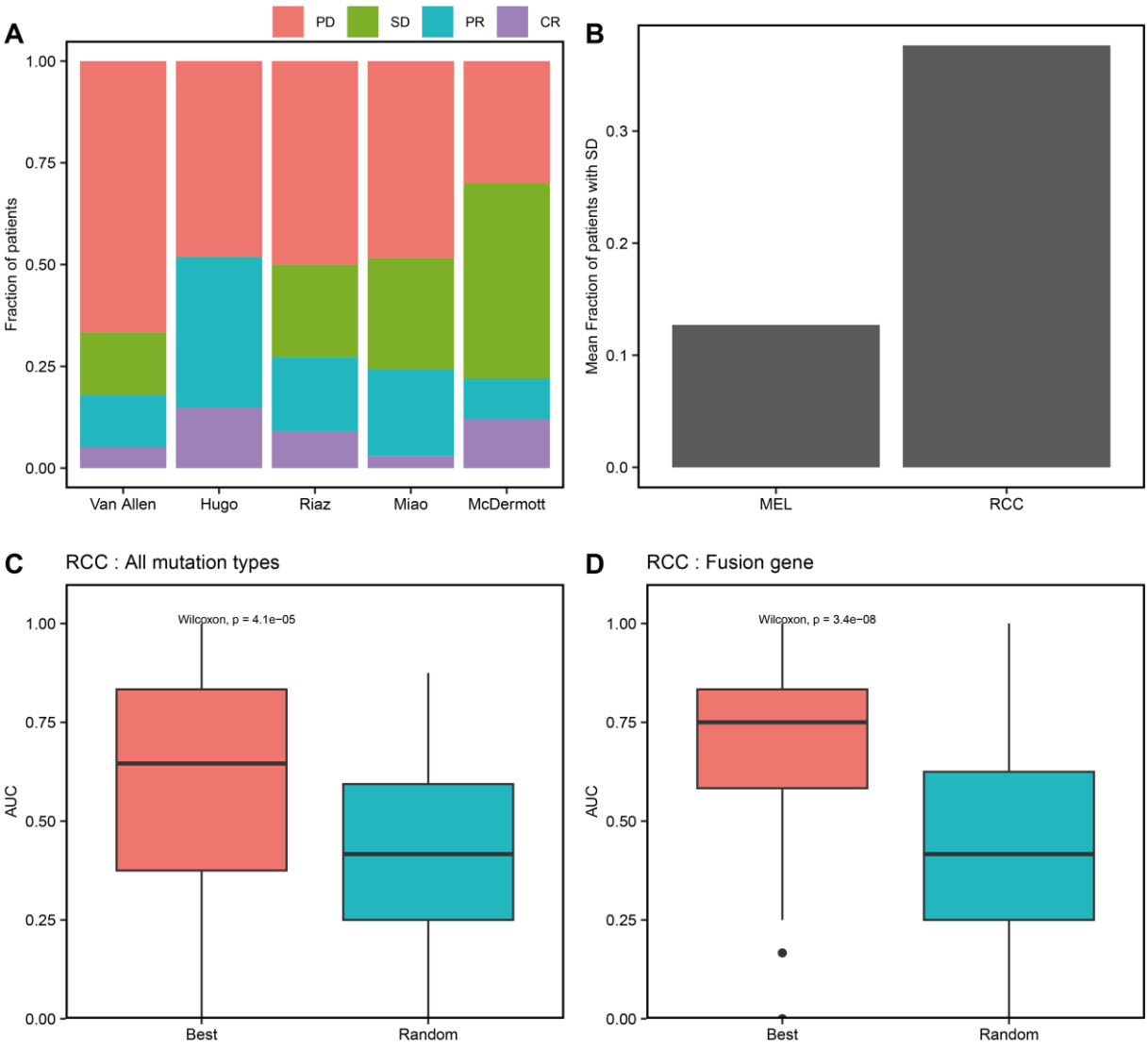

Figure S4: MILES on a dataset excluding patients with stable disease, related to Figure 4. (A) Distribution of response categories within the individual cohorts. PD: progressive disease, SD: stable disease, PR: partial response, CR: complete response. **(B)** Fraction of patients with stable disease in the melanoma (MEL) and renal cell carcinoma cohort. **(C-D)** The optimal hyper parameter set were used to train and evaluate the performance of MILEs on a randomized dataset without patients with stable disease for the RCC cohort of **(C)** neoantigen candidates from all mutation types and of **(D)** fusion gene-derived neoantigen candidates. Randomized refers to the randomization of neoantigen candidates across patients while keeping the original number of neoantigen candidates per patient in the respective dataset.

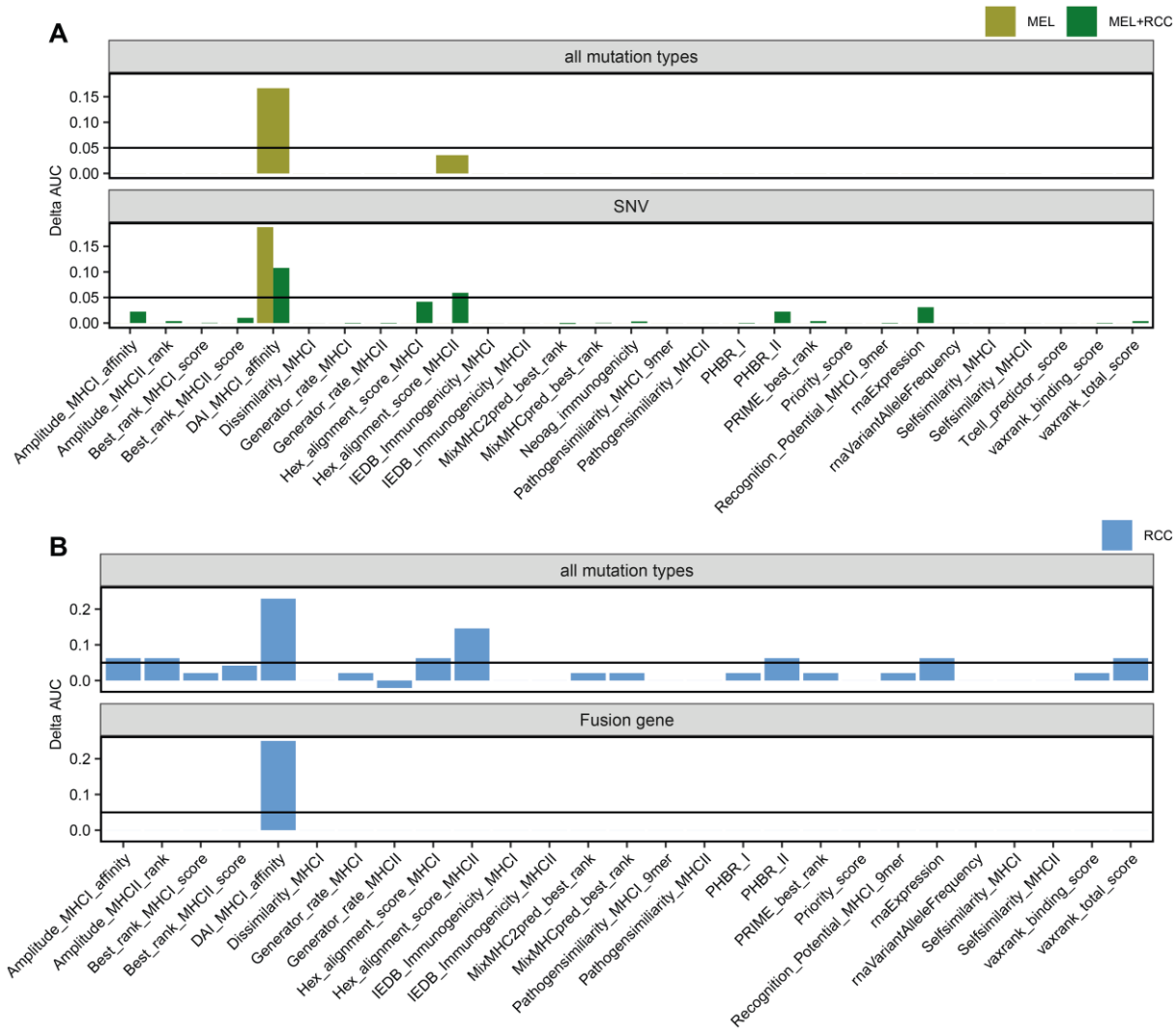

Figure S5: Feature importance analysis, related to Figure 4: To estimate the importance of each neoantigen feature, the nested CV approach was 50x repeated for the best hyperparameter setting on a dataset in which the neoantigen feature of interest was permutated. Feature importance was approximated for approaches with median AUROC > 0.6 by the delta AUROC of learning method on the original data and the learning method on the data with permutated feature. **(A**) Feature importance for MILES on neoantigen candidates from all RCC and MEL or MEL-only cohorts. (**B**) Feature importance for MILES on neoantigen candidates from renal cell carcinoma cohorts, excluding patients with stable disease