51 A COMBINATION OF ANTIGEN PRESENTATION AND T-CELL RECOGNITION FEATURES IMPROVES NEOANTIGEN IMMUNOGENICITY PREDICTIONS

Neeraja Ravi*, Hima Anbunathan, Rachel Marty Pyke, Steven Dea, Prateek Tandon, Richard O Chen, Sean Michael Boyle. *Personalis, San Francisco, CA, USA*

Background The assessment of tumor neoantigen burden has been shown to outperform tumor mutational burden in predicting patient response to checkpoint inhibitor immunotherapy by better capturing the biological mechanism underlying response.[1]. However, immune recognition of neoantigens by T-cells requires more than antigen presentation, which has been the focus of tumor neoantigen burden assessment to date. Here, we extend the existing SHERPA® MHC-presentation framework.[2] To include a model for the prediction of neoantigen immunogenicity.

Methods For feature engineering, training and validation, we utilized two datasets containing peptides experimentally validated for immunogenicity. The first dataset, curated by Schmidt et al.,[3] aggregates experiments from 17 different sources, identifying 1282 immunogenic peptides across 67 MHC alleles. While the diversity of this dataset enables generalizability, a lack of associated sequencing data limits the features that can be investigated. The second dataset, curated by the TESLA consortium, contains 37 immunogenic peptides across 13 MHC alleles and patient-specific exome and transcriptome sequencing data, broadening the potential feature landscape.[4] Using both datasets, we developed and validated features associated with antigen availability, processing, presentation and recognition. To inform the assessment of antigen availability, we measured gene expression level and variant allele fraction. We built a cleavage probability predictor from immunopeptidomics data for antigen processing, while SHERPA MHC binding probability was used to quantify antigen presentation. Finally, we included measures to predict T-cell recognition based on antigen hydrophobicity, agreotopicity, dissimilarity to self antigens and similarity to known foreign antigens. We utilized a two-tiered machine learning model that selectively learns the weights of features from the dataset that is most informative and least biased.

Results The Schmidt et al. dataset was used in the first tier of the model to develop an immunogenicity score using peptide-derived features. The first tier score distinguished immunogenic peptides with an area under the precision recall curve (AUPRC) of 0.74, far greater than SHERPA or NetMHCpan-4.1 alone (0.48 and 0.39 respectively). The second tier of the model was trained on the TESLA dataset and used the first tier score as a feature along with other patient-specific features. Cross validation yielded a 37% fold increase in AUPRC over the method developed by the TESLA consortium.

Conclusions By combining antigen presentation and T-cell recognition features in a two-tiered model, we can better predict immunogenic neoantigens and make progress towards using neoantigens as biomarkers to assess checkpoint inhibitor efficacy.

REFERENCES

1. Abbott, C.W. *et al*. Prediction of Immunotherapy Response in Melanoma through Combined Modeling of Neoantigen Burden and Immune-Related Resistance Mechanisms. *Clin Cancer Res*. 2021; **27**(15):4265–4276.
2. Pyke, R.M. *et al*. Precision Neoantigen Discovery Using Large-scale Immunopeptidomes and Composite Modeling of MHC Peptide Presentation. *Mol Cell Proteomics*. 2021;**20**:100111.
3. Schmidt, J. *et al*. Prediction of neo-epitope immunogenicity reveals TCR recognition determinants and provides insight into immunoediting. *Cell Rep Med*. 2021;**2** (2):100194.
4. Wells, D. *et al*. Key Parameters of Tumor Epitope Immunogenicity Revealed Through a Consortium Approach Improves Neoantigen Prediction. *Cell*. 2020 Oct 29;**183**(3):818–834.e13.

http://dx.doi.org/10.1136/jitc-2022-SITC2022.0051