Methods of Improving Accuracy of Neoantigen Identification for Therapeutic and Diagnostic Use in Immuno-oncology

#1292

Sean Michael Boyle, Ravi Alla, Ryan Wang, Eric Levy, Gabor Bartha, Jason Harris, Rena McClory, John West, Richard Chen

Personalis, Inc. 1330 O'Brien Drive, Menlo Park, CA 94025

Contact: sean.boyle@personalis.com

Background

Neoantigens are increasingly critical in immuno-oncology as a therapeutic target for neoantigen-based personalized cancer vaccines and as a potential biomarker for immunotherapy response. However, the methods for identifying which neoepitopes are more likely to provoke an immune response remains an important challenge for improving both the effectiveness of neoantigen-based vaccines and enabling the potential use of neoantigens as a biomarker in immunotherapy.

Methods

We sought to improve overall neoantigen identification performance by systematically improving critical components of our ACE ImmunoID assays and neoantigen pipeline. Personalis' Accuracy and Content Enhanced (ACE) technology was developed to fill critical gaps in conventional exome and transcriptome sequencing that can lead to missed neoantigens. To improve MHC-epitope binding prediction, we trained neural networks on mass spectrometry derived MHC-epitope binding data. This is in contrast to other MHC binding algorithms that have been primarily trained using in vitro competitive binding data, which sufer from having not been processed, loaded, nor shuttled natively into the HLA binding domain. HLA typing, a key input into the neoantigen prediction algorithms, was improved through exome augmentation of the HLA region with an optimized HLA typing algorithm. Other enhancements include RNA based somatic variant calling, peptide phasing, transcript isoform estimation, and identification of indel and fusion derived neoepitopes.

The Importance of an Augmented Exome (ACE)

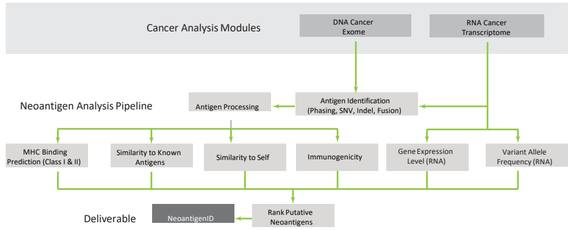
We have developed a neoantigen detection pipeline built upon our analytically validated ACE exome and transcriptome sequencing platform and somatic variant calling pipeline through combined DNA and RNA analysis. Our ACE platform augments digcult to sequence regions (green) that are missed by traditional exomes (blue). This is particularly important for Immuno-oncology as immunogenic antigens have the potential to arise from mutations anywhere in the exome.

Accurately Detecting Somatic Variants for Neoantigen Assessment

Accurately detecting variants directly from genomic regions and expressed transcripts is incredibly important for neoantigen prediction. Expressed variants are far more likely to be translated into proteins which are in turn cleaved, transported, and presented on MHC class I. We took this importance into consideration when developing and validating a highly accurate method for somatic variant calling in both DNA and RNA.

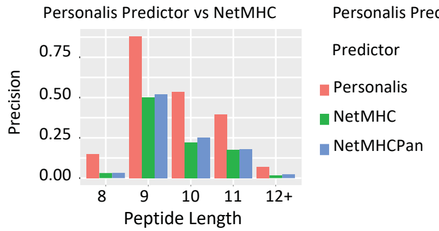
	Sensitivity	LOD >10% (PPA)	LOD >10% (PPV)
DNA (Small Vars)	> 98%	>97%	>98%
RNA (Small Vars)	> 97%	> 97%	
Fusions	> 99%		

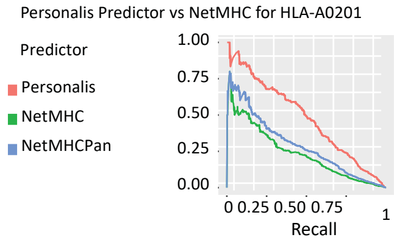
Developing a Neoantigen Analytics Engine



Within our neoantigen pipeline, variants that are detected by our DNA and RNA cancer analysis pipelines are processed for antigen identification, including SNVs, indels, and fusion events. Importantly, both in-frame and out-of-frame events are accurately considered by transcript, allowing for detection of a wealth of candidate neoantigens. Our pipeline includes assessment of important immunologic components including HLA prediction, MHC binding (class I and II), immunogenicity, similarity to self, and similarity to known antigens. Additionally, peptides are evaluated for variant allele frequency in both the RNA and DNA of the tumor sample and gene expression level is considered. Collectively, our ImmunoID product provides a comprehensive assessment of features that may be used for identifying and ranking potentially immunogenic neoantigens.

Improving MHC Binding Prediction





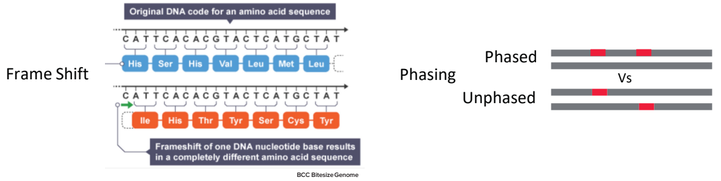
HLA binding prediction is arguably the most important component of neoantigen assessment. Recent advances in training data generation, made possible by mass spec, provide the opportunity to accrue large numbers of peptide binders and non-binders for individual HLA alleles. Further this new binding data takes two important additional components into consideration: cleavage and transportation, which are critically important for presentation assessment. We leverage this advancement, developing a brand new MHC binding prediction algorithm which outperforms both NetMHC predictors across the range of prediction score cutofs.

Correct HLA Typing For Accurate Neoantigen Prediction

HLA typing is a cornerstone of neoantigen prediction. If the typing isn't correct, then all downstream binding predictions will be wrong. In order to provide the most accruate HLA typing, we have developed and valdiated an extremely accurate tool built upon our ACE platform. We performed a blinded clinical validation with > 60 orthogonally typed samples, achieving high concordance.

HLA Loci	# Calls	# Agree	Concordance
Class I	318	311	98%
Class II	378	355	94%
Class I + II	696	666	96%

Phasing and Frame-Shift Variant Peptide Creation



Frame-shift variants can result in entirely distinct downstream protein sequences. Correctly assessing these mutations at the transcript level produces a wealth of peptides. Proper phasing allows for assignment of correct peptides. As each amino acid matters when assessing MHC binding or immunogenicty, phasing can have a major impact on neoantigen predictions.

Accurately Detecting Known Immunogenic Neoantigens

Neoantigen Prediction
96% Sensitivity (22/23 Epitopes)

To assess the efectiveness of this pipeline in predicting immunogenic neoantigens, we assembled a gold-set of 23 known, previously experimentally-validated immunogenic neoantigens from the literature. We spiked in these neoepitopes into exome data and assessed the ability of our neoantigen pipeline to find and rank these immunogenic known neoantigens. Preliminary results show our neoantigen pipeline is able to accurately identify 22 out of 23 (~96%) of the spiked in neoantigens as being potentially immunogenic.

Conclusions

We developed sequencing and informatics improvements to standard approaches that can enhance neoantigen identification, including advancements in HLA typing and MHC binding prediction. When coupled with comprehensive validations these may support neoantigen use in future clinical settings.

Personalis