Validation of an Expanded Neoantigen Identification Platform for Therapeutic and Diagnostic Use in Immuno-oncology

289

Sean Michael Boyle, Jason Harris, Gabor Bartha, Ravi Alla, Patrick Jonganeel, Mirian Karbelashvili, Scott Kirk, Steve Chervitz, Eric Levy, Craig Rowell, Robert McCord, Shujun Luo, Rena McCloy, John West, Richard Chen

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

Contact:
sean.boyle@personalis.com

Background

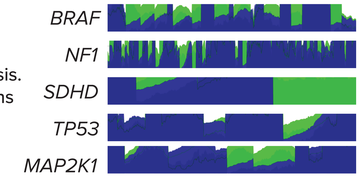
Neoantigen identification is increasingly critical for clinical immuno-oncology applications including predicting immunotherapy response and neoantigen-based personalized cancer vaccines. Although standard research pipelines have been developed to aid neoantigen identification, building a robust, validated neoantigen identification platform suitable for clinical applications has been challenging due to the complex processes involved.

Methods

To improve neoantigen identification, we extended standard sequencing and informatics methods. We developed an Accuracy and Content Enhanced (ACE) exome sequenced at 200X to increase sensitivity to SNPs and indels used for neoantigen identification as well as HLA performance. To accurately identify fusions and variants from RNA, we optimized our ACE transcriptome for FFPE tissue. To improve neoantigen pipelines based on MHC binding algorithms, we developed peptide phasing, high accuracy HLA typing, immunogenicity, and transcript isoform estimation tools to detect neoantigens from indel and fusion events. We performed comprehensive analytical validation of the platform including the ACE exome, somatic SNV/indel calls, RNA based variant and fusion calls, and HLA typing. This was followed by an overall in silico validation of neoantigen identification using 23 experimentally validated immunogenic neoepitopes spiked into exome data.

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 difficult 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.

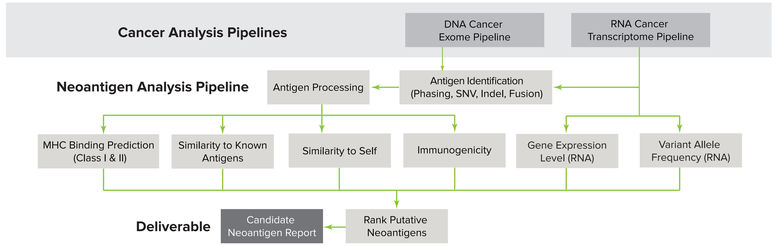


Accurately Detecting Somatic Variants for Neoantigen Assessment

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.

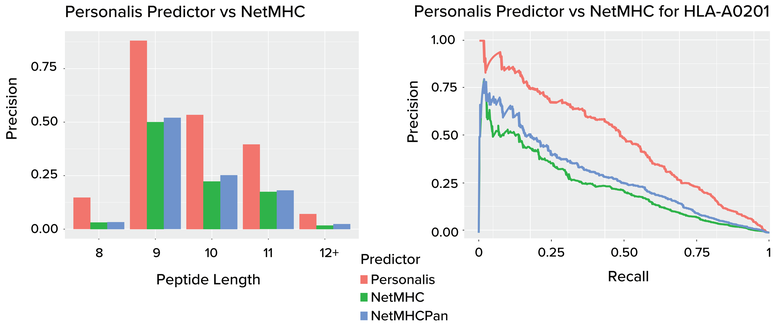
	Sensitivity	LOD >10% (PPA)	LOD >10% (PPV)
DNA (Small Vars)	> 98%	>97%	>98%
RNA (Small Vars)	> 97%	> 97%	
Fusions	> 99%		

Developing a Comprehensive Neoantigen Pipeline



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



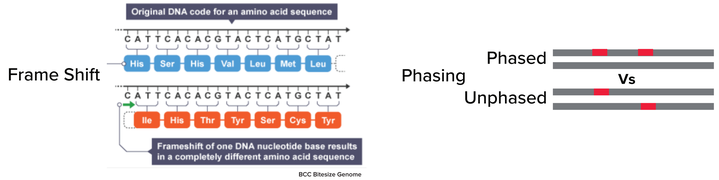
HLA binding prediction is arguably the most important component of neoantigen assessment. Recent advances in training data generation, made possible by mass spec, provide the opportunity to accrue large numbers of peptide binders and non-binders for individual HLA alleles. Further this new binding data 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 cutoffs.

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 validated an extremely accurate tool built upon our ACE platform. We performed a blinded clinical validation with 20 orthogonally typed samples involving > 300 HLA alleles, achieving high concordance.

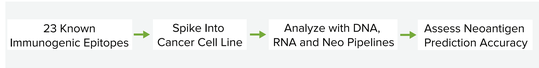
HLA Loci	# Calls	# Agree	Concordance
Class I	120	117	98%
Class II	236	224	95%
Class I + II	356	341	96%

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 effectiveness 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
Pioneering Genome-Guided Medicine