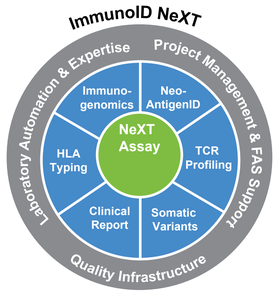
Improving Neoantigen Identification for Therapeutic and Diagnostic Use in Immuno-oncology Using Mass Spectrometry and Machine Learning

Sean Michael Boyle, Eric Levy, Nick Phillips, Datta M, Gabor Bartha, Jason Harris, Rena McClory, John West, Richard Chen
Personalis, Inc. 1330 O'Brien Drive, Menlo Park, CA 94025

Background

Neoantigens are increasingly critical in immuno-oncology as therapeutic targets for neoantigen-based personalized cancer vaccines (PCVs) and as potential biomarkers 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 PCVs and enabling the potential use of neoantigens as a biomarker in immunotherapy. Current MHC binding prediction algorithms are primarily trained using in vitro MHC binding data, which does not encompass proteasomal cleavage and transport, important factors for neoantigen presentation. Recent advances in immuno-affinity purification and mass spec technology makes it possible to identify processed cell surface MHC bound peptides in an in vivo setting, providing the opportunity for development of improved neoantigen prediction pipelines.

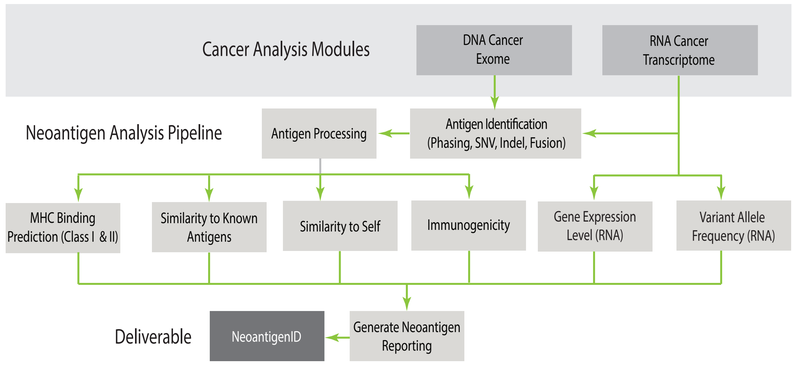
Introducing ImmunoID NeXT a Comprehensive Platform For IO



We have developed a neoantigen detection pipeline built upon our analytically validated ImmunoID NeXT exome and transcriptome sequencing platform and somatic variant calling pipeline through combined DNA and RNA analysis. Our newest generation of ImmunoID platform both augments difficult to sequence regions that are missed by traditional exomes and improves HLA Typing, TCR Profiling, and somatic mutation calling. This is particularly important for Immuno-oncology as immunogenic antigens have the potential to arise from mutations

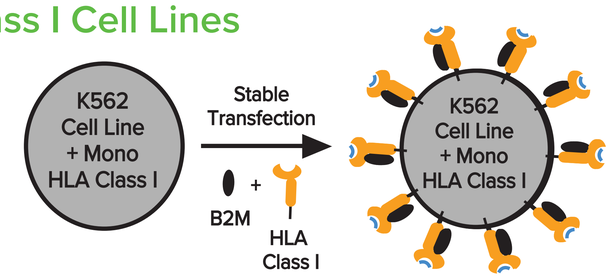
Developing a Neoantigen Analytics Engine

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. Collectively, our ImmunoID product provides a comprehensive assessment of features that may be used for identifying and ranking potentially immunogenic neoantigens.



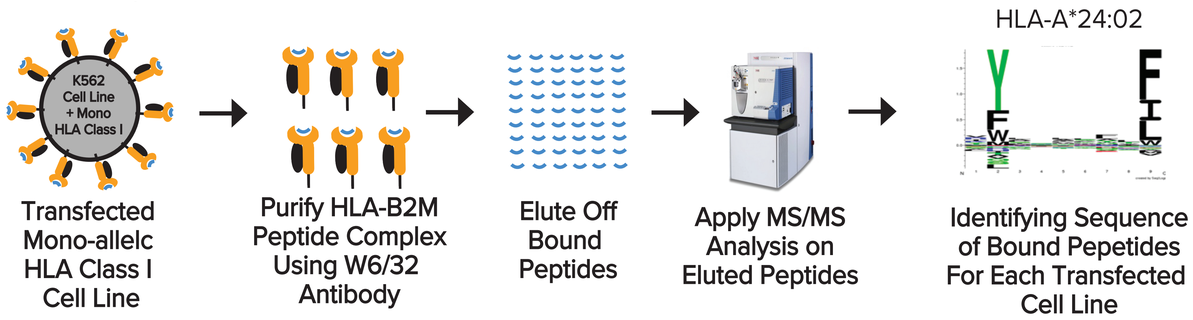
Generating Mono-Allelic HLA Class I Cell Lines

Mono-allelic HLA class I cell lines were generated by transfecting individual class I HLA alleles into the HLA class I null cell line K562, prioritizing HLA class I alleles that were of high abundance in different populations.

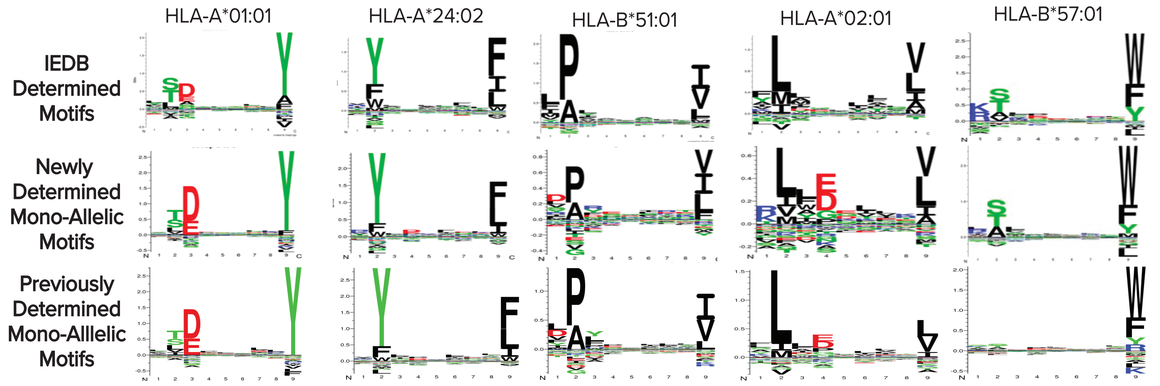


Identifying HLA Class I Bound Peptides Using Immunopeptidomics

Immunopeptidomics was applied to isolate and identify HLA Class I bound peptides. As all cells in each stably transfected population expressed the only 1 HLA gene, all immuno-precipitated peptides from a pool of cells were known to be bound to the same HLA allele.

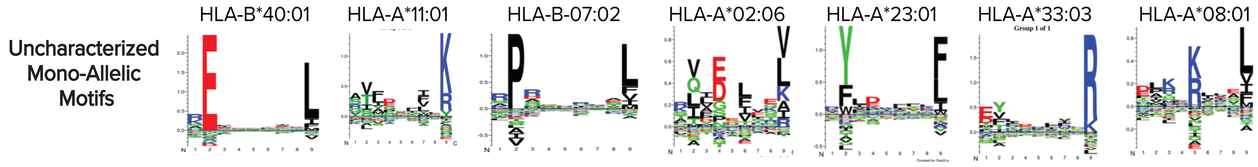


Comparing Sequence Motifs for Alleles With Known Motifs



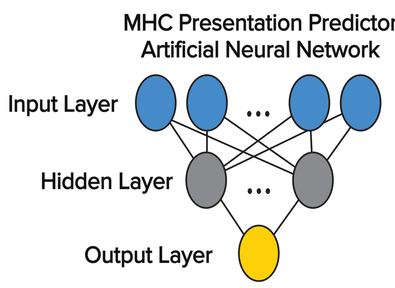
For the 5 HLA class I alleles with previsouly published mono-allelic derived sequence motifs, we identified similarirty and distinctions between both IEDB and mono-allelic derived published motifs. For example, our derived A*02:01 peptides cotntain a similar signature at position 4 to mono-allelic motifs. However, we retain the stronger VL patter at position 9 indicating a more conserved binding preference. As peptides derived from mono-allelic lines do not require deconvolution and include proteosomal processing, we ave increased confidence in our newly identifed peptide motifs over multi-allelic results.

Identifiyng Motifs For Alleles Previously Uncharacterized Using Mono-Allelic Lines

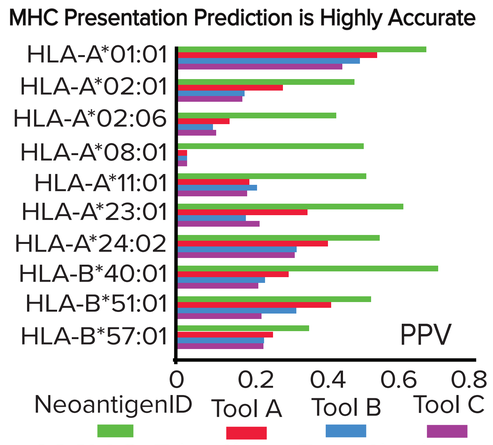


For 7 HLA alleles we have determined mono-allelic derived peptide motifs for the first time. As anticipated we observe strong signatures for the majority of alleles at positions 9 and / or 2 in each motif, suggesting strong receptor facing amino acid preferences. We also observe the solution facing amino acids for each allele are more diverse, with the exception of alleles A*08:01 and A*02:06, which have a conserved preference at position 5 and 4/6, respectively.

Developing a Class I MHC Presentation Predictor Using Mono-Alleic Data

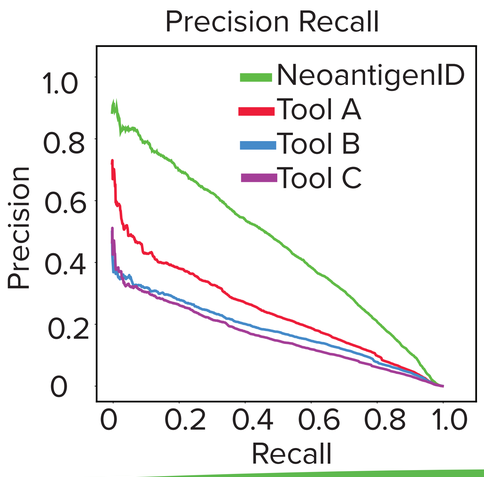


We developed and trained neural networks to predict MHC class I presentation for each assayed HLA class I allele.



Our neoantigen prediction algorithm, trained on our own in vivo peptide data, consistently achieves a higher overall sensitivity and specificity than other commerically avialble tools based on either in vitro MHC binding or immuno-peptidomics derived data, when tested on the same HLA class I alleles.

Improved Performance of MHC Binding Prediction



HLA binding prediction is a critical important component of neoantigen assessment. Recent advances in training data generation, including our mono-allelic MHC class I cell line generation and immuno-peptidcomics, provide the opportunity to accurately detect large numbers of naturally processed peptide binders and non-binders for individual HLA alleles. 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, generating highly accurate binding data and developing a brand new MHC binding prediction algorithm which outperforms publically available predictors across the range of prediction score cutoffs.

Personalis®
Pioneering Genome-Guided Medicine