Orthogonally and functionally validated algorithm for detecting HLA loss of heterozygosity

Rachel Marty Pyke*, Dattatreya Mellacheruvu*, Charles Abbott, Simo V. Zhang, Eric Levy, Steven Dea, Devayani Bhave, Manju Chinnappa, Gabor Bartha, John Lyle, John West, Richard Chen and Sean Michael Boyle
Personalis, Inc., Menlo Park, CA
*Co-first authors

Contacts:
Rachel.Pyke@personalis.com
Datta.Mellacheruvu@personalis.com
Sean.Boyle@personalis.com

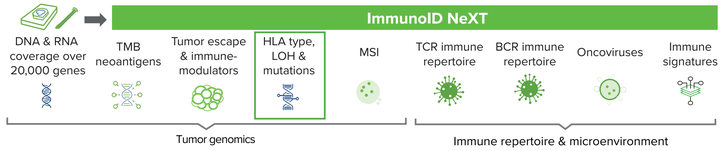
#73

I. Background

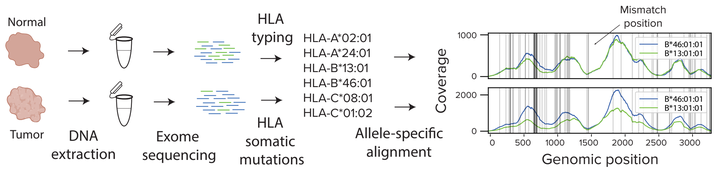
T cells can inhibit tumor growth and progression by recognizing major histocompatibility complex class I (MHC-I) -bound peptides derived from antigens that are mutated, inappropriately expressed, or overexpressed in the tumor cells (vs. healthy cells). There are many mechanisms by which tumors can avoid recognition or destruction by the immune system. Among them is HLA loss of heterozygosity (LOH), which occurs when a maternal or paternal HLA haplotype is lost, and can result in reduced tumor-specific antigen presentation. HLA-LOH has been associated impaired response to checkpoint blockade immunotherapy. Although HLA LOH is emerging as a key biomarker for response to immunotherapy, few tools exist to detect HLA LOH. Here, we describe DASH (Deletion of Allele-Specific HLAs), an algorithm to detect HLA LOH from exome sequencing data, and present a three-pronged validation approach to assess its performance.

II. Augmented exome capture with ImmunoID NeXT

The ImmunoID NeXT™ Platform provides joint tumor genomics and immune profiling from a paired tumor and healthy normal sample. Through augmented coverage of the HLA locus, ImmunoID NeXT also provides the data to accurately type HLA alleles, detect somatic mutations and probe copy number deletions in this highly polymorphic region.



III. Training DASH to detect HLA LOH

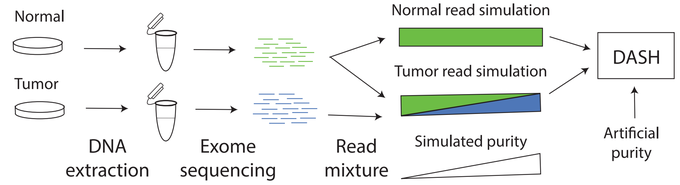


We performed exome sequencing with ImmunoID NeXT on tumor and normal samples from 279 patients to create a training dataset for our model. For each patient, we detected germline HLA types and somatic HLA mutations to build a patient-specific HLA reference. Then, we mapped their tumor and healthy normal reads to their patient-specific HLA reference and used this information as input for our machine learning features. Using six allele-specific and patient-specific features, we trained a gradient boosted regression algorithm on samples with manually annotated HLA LOH labels.

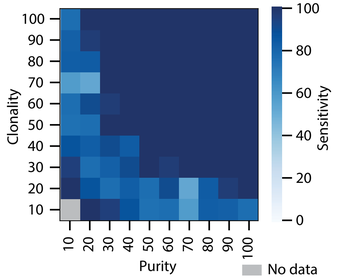
IV. Limit of detection using in silico cell line mixtures

In-silico evaluation of the limit of detection (LOD) of DASH was performed by deeply sequencing a tumor-normal paired cell line with HLA LOH and mixing reads at different proportions to simulate variable tumor purity and clonality. For fully clonal tumors, DASH had 100% sensitivity at all tumor purity levels above 8% and 100% specificity at tumor purity levels higher than 36%.

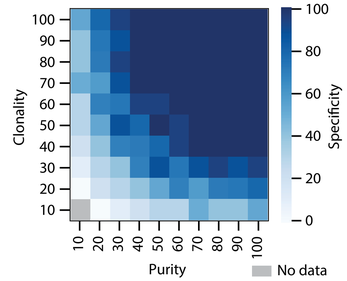
A. Experimental design



B. Sensitivity of DASH



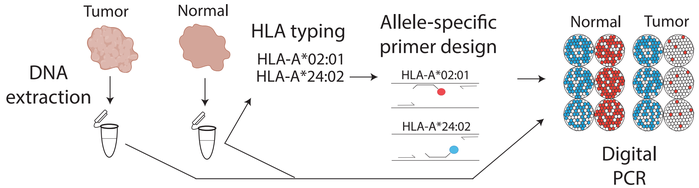
C. Specificity of DASH



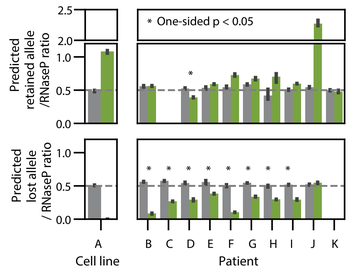
V. Genomic validation using patient-specific digital PCR

Direct genomic validation was performed using digital PCR (dPCR) with allele-specific primers targeting both predicted retained and lost alleles in ten patient samples and one cell line. In addition to having to target the alleles unique to each patient, the probes were also designed to avoid targets on other patient-specific alleles. Probes for RNaseP were multiplexed as the positive control for all samples

A. Experimental design



B. Digital PCR results

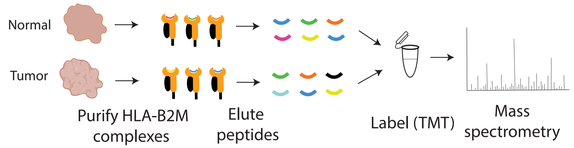


We investigated the specificity of the allele-specific primers and found that 20 of the 22 primers were highly specific (one primer from patient C and patient K were excluded for a lack of specificity). Next, we compared the allele to RNaseP ratios in the tumor samples to the ratios in the normal samples. We found a significant depletion in only one of the nine predicted retained alleles and a significant depletion in eight of the nine predicted lost alleles in the tumor sample. Several of these samples had low tumor content, confirming the accuracy of DASH across variable tumor purities.

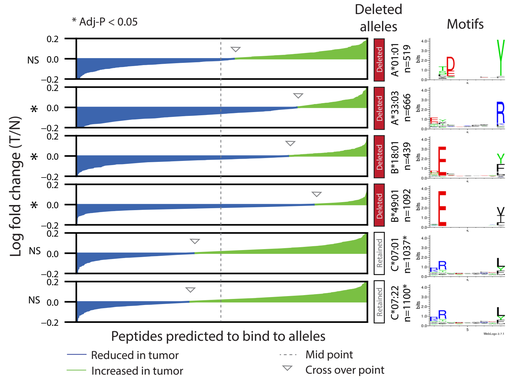
VI. Functional assessment of HLA LOH with quantitative immunopeptidomics

Mechanistically, HLA LOH is expected to effectively reduce the neoantigen load by eliminating surface presentation of neoantigens that would bind to specific HLA alleles. Quantitative immunopeptidomics was performed on six samples (two without HLA LOH and four with HLA LOH) to compare peptides presented by HLA alleles in tumor cells and adjacent normal cells. In the samples without any predicted HLA LOH, we found minimal differences in surface peptide intensities between the tumor and normal samples. In contrast, the samples with predicted HLA LOH showed twice as much variability in peptide presentation between the tumor and normal samples, with more variation in higher tumor purity samples. In a moderate purity sample (shown here), we found that three of four deleted alleles (predicted bi-allelic deletions) had significantly fewer predicted binding peptides in the tumor sample than in the adjacent normal sample.

A. Experimental design



B. Allele-specific enrichment in tumor peptides



VII. Conclusion

HLA LOH detection methods need to be rigorously validated in order to be used as a clinical biomarker. Here, we introduced three methods to assess performance, demonstrated the strong predictive power of DASH, and highlighted the need to consider tumor purity in such assessments.

Personalis®