Extensively validated HLA LOH algorithm demonstrates an association between HLA LOH and genomic instability

#79

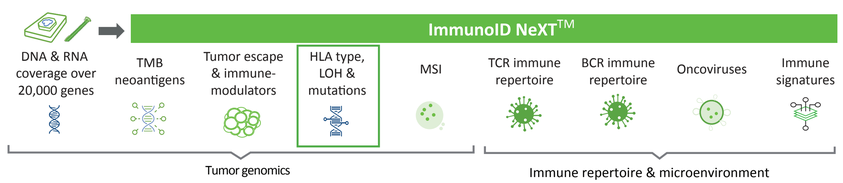
Rachel Marty Pyke[1], Steven Dea[1], Dattatreya Mellacheruvu[1], Charles W. Abbott[1], Simo V. Zhang[1], Lee McDaniel[1], Eric Levy[1], Gabor Bartha[1], John West[1], Michael P. Synder[2], Richard Chen[1] and Sean Michael Boyle[1]
[1]Personalis, Inc., Menlo Park, CA; [2]Stanford, Palo Alto, CA

I. Background

Human Leukocyte Antigen (HLA) genes are critical for the presentation of neoantigens to the immune system by cancer cells. Deletion of HLA alleles, known as HLA loss of heterozygosity (LOH), has been highlighted as a key immune escape mechanism. Validated algorithms to detect HLA LOH from sequencing data are critical for exploring the biological impact of HLA LOH and assessing its utility as a clinical biomarker. To address this need, we developed a machine learning algorithm to detect HLA LOH, Deletion of Allele-Specific HLAs (DASH™), extensively validated it with several approaches and applied it to a large cancer cohort to understand its frequency of occurrence and biological impact.

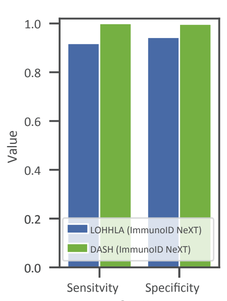
II. Augmented exome capture with ImmunoID NeXT™

The ImmunoID NeXT Platform® provides joint tumor genomics and immune profiling from a single tumor/normal sample. Through augmenting 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.



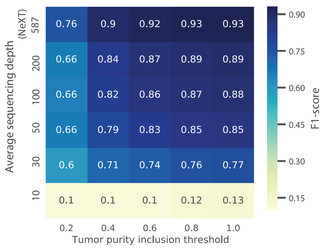
III. Approach and performance on held out data

We performed exome sequencing with ImmunoID NeXT on tumor and normal samples from 279 patients to create a training dataset for DASH. For each patient, we detected germline HLA types and somatic HLA mutations. Then, we mapped their tumor and normal reads to their patient-specific HLA reference and engineered seven features: adjusted b-allele frequency, allele-specific tumor-normal ratio, total sequencing depth ratio, consistency of sequencing depth, tumor purity, tumor ploidy and deletion of flanking regions. To label our dataset, we visualized each allele pair and manually annotated LOH. Then, we trained DASH with an XGBoost model to discriminate between genes with and without HLA LOH. We benchmarked our performance on our held out dataset against LOHHLA, a publicly available algorithm, and demonstrated improved sensitivity and specificity. Both algorithms were tested using the ImmunoID NeXT Platform.

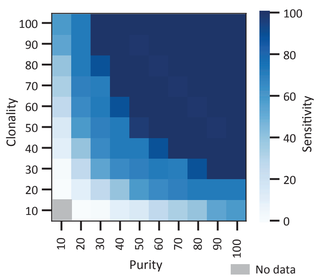


IV. Performance with decreasing sequencing depth

The algorithms were tested on ImmunoID NeXT, suggesting that performance may decrease with other exome platforms. ImmunoID NeXT boosts sequencing in the HLA loci, ~1600x prior to strigent post-processing and ~600x afterward, as compared to ~150x on other exome platforms. Thus, we evaluated DASH and found that performance drops with lower sequencing depth and tumor purity.



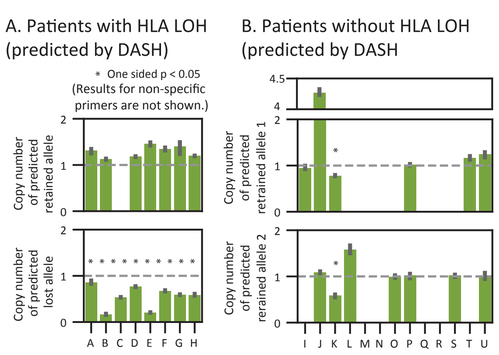
V. Limit of detection with in silico mixtures



To understand the limit of detection of DASH, we profiled over 30 paired tumor-normal cell lines on the ImmunoID NeXT Platform and identified four cell lines with HLA LOH. Using in silico mixtures for three of these cell lines, DASH demonstrates greater than 99% specificity across all tumor purity and sub-clonality levels (not shown) and greater than 98% sensitivity for above 27% tumor purity (shown to the right). In comparison, LOHHLA's sensitivity dropped dramatically across all tumor purities for events that were less than 80% clonal.

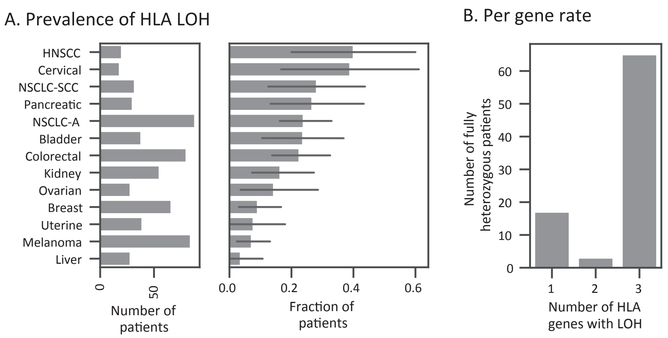
VI. Digital PCR validation with tumor samples

We designed a digital PCR (dPCR) assay using patient-specific and allele-specific primers that target a single HLA allele while avoiding all other HLA alleles. We tested 8 patients predicted to have HLA LOH by DASH and 13 patients predicted not to have HLA LOH. Excluding one ambiguous patient (K), DASH demonstrated 100% sensitivity and specificity in dPCR experiments across 21 tumor samples with stable controls.



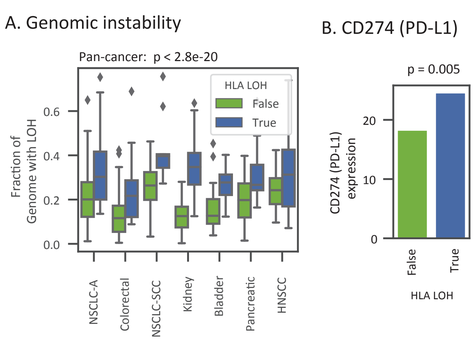
VII. Prevalence of HLA LOH across tumor types

We applied DASH to a large pan-cancer cohort and found that 18% of patients pan-cancer had HLA LOH, ranging from 40% in head and neck squamous cell carcinoma to only 4% in liver cancer. Moreover, patients preferentially lost all three genes.



VIII. Relationship to other tumor features

Next, we explored the cohort to understand how HLA LOH related to other tumor features. We identified strong associations between HLA LOH and genomic instability. Moreover, we demonstrated relationships between HLA LOH and markers of immune pressure, such as a correlation with CD274 (PD-L1) expression and allele-specific neoantigen enrichment for deleted HLA alleles.



IV. Conclusion

DASH, a highly sensitive HLA LOH algorithm that has been extensively validated using cross validation, in silico downsampling, cell line mixtures and dPCR, has demonstrated the widespread impact of HLA LOH in a large pan-cancer cohort.

Contact:
Rachel.Pyke@personalis.com
Sean.Boyle@personalis.com

Personalis®