npj | precision oncology

Article

Published in partnership with The Hormel Institute, University of Minnesota

https://doi.org/10.1038/s41698-024-00665-z

Detecting HLA loss of heterozygosity within a standard diagnostic sequencing workflow for prognostic and therapeutic opportunities

Ariane Lozac'hmeur[1], Tyler Danek[1], Qidi Yang[1], Mario G. Rosasco[1], John S. Welch[2], William Y. Go[2], Eric W. Ng[2], Armen Mardiros[2], David G. Maloney[3], Edward B. Garon[4], Kedar Kirtane[5], Diane M. Simeone[6], Julian R. Molina[7], Ameen A. Salahudeen[8], Michelle M. Stein[1] ✉ & J. Randolph Hecht[4]

To enable interrogation of tumor HLA LOH as a clinical diagnostic for precision oncology, we developed and validated an assay that detects HLA LOH within the context of an FDA-approved clinical diagnostic test, Tempus xT CDx. Validation was conducted via: (1) analytical evaluation of 17 archival patient samples and 42 cell line admixtures and (2) independent clinical evaluation of LOH prevalence in the *HLA-A* gene (*HLA-A* LOH) across 10,982 patients. To evaluate the prognostic relevance of *HLA-A* LOH we assessed 256 immunotherapy-treated non-small cell lung cancer (NSCLC) patients. To determine the feasibility of prospectively identifying and enrolling *HLA-A* LOH patients into a clinical trial, we established BASECAMP-1 (NCT04981119). We observed a positive predictive agreement of 97% and a negative predictive agreement of 100% in samples with ≥ 40% tumor purity. We observed *HLA-A* LOH in 16.1% of patients (1771/10,982), comparable to previous reports. *HLA-A* LOH was associated with longer survival among NSCLC adenocarcinoma patients (HR = 0.60, 95% CI [0.37, 0.96], $p = 0.032$) with a trend towards shorter survival among squamous cell patients (HR = 1.64, 95% CI [0.80, 3.41], $p = 0.183$). In 20 months, we prospectively screened 1720 subjects using the Tempus AWARE program, identifying 26 *HLA-A*02* LOH patients at 8 sites, with 14 (54%) enrolled into BASECAMP-1. In conclusion, we developed and validated an investigational assay that detects tumor HLA LOH within an FDA-approved clinical diagnostic test, enabling HLA LOH utilization in diagnostic, prognostic, and therapeutic applications.

Precision oncology requires robust molecular assays that provide diagnostic insights, prognostic risk stratification, and enable patient accrual into targeted clinical trials. For molecular assays to be utilized in clinical care, they also must seamlessly integrate into routine diagnostic workflows while meeting applicable regulatory requirements. HLA loss of heterozygosity (LOH) is a type of somatic HLA defect that presents an opportunity for precision oncology, as it is a common genomic event that can be exploited therapeutically to differentially target tumors from normal cells[1]. However, to our knowledge, no clinical assay exists for HLA LOH detection in solid tumors.

HLA LOH is a mechanism of immune evasion by tumor cells that results in a decreased repertoire of tumor neoepitopes presented for immune surveillance[2–9]. HLA class I LOH has been associated with clonal selection and tumor cell survival, and has been detected in ~2–40% in the most common adult cancer types (and up to 73.8% in kidney chromophobe tumors)[10–15], and 15–20% in pediatric osteosarcoma and glioblastoma[14].

[1]Tempus AI Inc., Chicago, IL, USA. [2]A2 Biotherapeutics, Agoura Hills, CA, USA. [3]Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [4]David Geffen School of Medicine, University of California, Los Angeles, CA, USA. [5]Moffitt Cancer Center, Tampa, FL, USA. [6]Moores Cancer Center, UC San Diego Health, San Diego, CA, USA. [7]Mayo Clinic, Rochester, MN, USA. [8]University of Illinois at Chicago College of Medicine, Departments of Medicine, Biochemistry & Molecular Genetics, Chicago, IL, USA.
✉e-mail: publications@tempus.com

The Hormel Institute
University of Minnesota

Over the last decade, irreversible HLA LOH has been recognized as an important molecular marker of poor clinical outcomes. For example, in lung cancer, HLA LOH has been associated with worse overall survival as well as high mutational burden, tobacco signatures, *TP53* mutations, and metastatic disease[10]. In addition, a paired analysis of NSCLC primary and metastatic sites observed a nonsignificant trend toward increased prevalence of class I HLA LOH in metastatic sites (prevalence in brain metastasis versus primary sites, 43% vs 27%), suggesting an association with more aggressive disease[7]. Further, in thyroid cancer patients, a survey of the mechanisms of genetic immune escape revealed significant enrichment of immune escape in metastatic samples, driven by an increase in class I HLA LOH[13]. Parallel evolution with independent loss of the same HLA alleles has been observed as distinct events on separate branches of tumors' phylogenetic trees, suggesting recurrent immune pressure to silence specific HLA-restricted neoepitopes during subclonal evolution[7]. Therefore, HLA LOH represents a potential prognostic clinical marker of immune editing that may be clinically relevant for clonal tracking and disease monitoring or surveillance.

HLA LOH also represents a therapeutic opportunity, as allele-specific loss provides a distinction between tumor and normal cells allowing for selective therapeutic exploitation by logic-gated cell therapy[16–19]. The reverse is also therapeutically relevant; T-cell therapies and neoantigen vaccines require persistent expression of HLA allele-specific epitopes and HLA LOH could render the therapy ineffective. Given the association between HLA LOH and resistance to checkpoint immunotherapy, there is a need to accurately identify HLA LOH in routine clinical practice to implement relevant immunotherapy approaches.

Accurate and timely detection of HLA LOH for use in clinical practice, therefore, is of interest. Copy number-based detection algorithms can reliably detect large deletions that include the HLA region[12,14,20–22], but may have reduced sensitivity in cases where LOH results from a focal deletion, poor reference genome alignment, or variation in hybrid probe capture across different allele exons. Further, the performance of copy number variant-based detection tools is heavily influenced by tumor sample purity, and subclonal HLA LOH has been reported to be challenging to detect[7,10].

In this study, we expand on our prior experience to develop and characterize a robust HLA LOH detection assay[23]. We created an algorithm to identify LOH using sequencing data from intron 2 to intron 4 (including exons 2 and 3) of *HLA-A* within the gene panel of Tempus xT CDx, an FDA-approved in vitro diagnostic device that is routinely used in standard-of-care diagnostic evaluation of solid tumors in the United States[24]. We find that HLA LOH in submitted tumor samples is efficiently and seamlessly detected in parallel with the detection of tumor-associated variants. We further report the performance of this HLA LOH assay in the framework of an ongoing non-interventional, observational study, BASECAMP-1 (NCT04981119), providing experience with real-time integration of an HLA LOH test within a clinical workflow.

Results

Description of HLA genotype analysis, detection of allelic imbalances, clonal loss, and amplifications in tumors

HLA LOH determination requires the following inputs: germline HLA genotype of the patient, sequencing data from the tumor and a matching normal sample, and a pathologist-determined tumor purity estimate. The test determines candidate HLA reads by aligning HLA-mapped reads and unmapped reads to a reference containing all known HLA alleles (per the IMGT/HLA database)[25]. Reads mapping to at least one HLA allele are conserved and aligned to a custom HLA reference build using the known HLA genotype of the patient. The HLA alignment of the normal sample is used to cross-check the provided genotype and as quality control. After the alignment step, reads matching patient alleles exactly are retained and coverage within the tumor and matching normal sample are used to calculate features such as B-allele fraction (BAF) and log ratio coverage (Fig. 1a). The features generated are fed into two logistic regression models: the first determines the presence of allelic imbalance, and the second employs tumor purity to assess if the imbalance is clonal. If the probability of the presence of allelic imbalance is <50%, the loci are classified as LOH negative. If the probability of the presence of allelic imbalance is >55% and the probability of clonal allelic imbalance is <50%, the loci are classified as subclonal LOH. If both probabilities of the presence of allelic imbalance and of clonal allelic imbalance are >55% and 50%, respectively, the loci are classified as clonal allelic imbalance. To determine whether the allelic imbalance is due to a clonal LOH or amplification of the other allele, we use a proprietary genome-wide CNV calling algorithm based on normalized coverage and variant allelic fraction across the short arm of chromosome 6. Application of the CNV calling algorithm across the whole short arm smoothes out the noise associated with normalized coverage. To allow for greater interpretability, the test outputs a visualization of allele fraction and coverage in the HLA region and the whole genome (Fig. 1a), detecting both clonal and subclonal LOH, amplifications and copy neutral LOH (Supplementary Fig. 1a–e).

A critical step in the calculation of coverage and BAF features is the correction of hybrid capture bias in the HLA region (Fig. 1a, b). To characterize this bias, over 3000 germline samples were processed for each genotype (e.g., *HLA-A*02:01*/*HLA-A*03:01*) and the median coverage and BAF was computed (Fig. 1b). To correct the bias, a panel of normal controls was built using the germline samples. The median coverage and BAF values for each genotype are used to normalize the BAF observed in the tumor to adjust for the differences in hybrid capture. Applying this correction reduced the estimated false positive rate by 50% at lower tumor purities (Fig. 1c). A large number of germline samples used to build this panel of normals provides broad *HLA-A* coverage for each superpopulation (Supplementary Table 3), which allows our device to perform equally well across individuals from distinct ancestries.

Validation and performance evaluation for the detection of HLA LOH using cell line models

Due to a lack of comparable CLIA-validated methods for orthogonal testing, we conducted analytical validation on contrived controls. Contrived controls consisted of 198 admixtures of 42 well-characterized homozygous (HLA locus) cell lines at different ratios to simulate HLA LOH at various tumor purities, including germline and HLA LOH-negative samples (Supplementary Fig. 2a, see Supplementary Methods). Using admixture ratios to simulate a tumor purity of ≥40%, we demonstrated a positive and negative percent agreement of 97% and 100%, respectively (Fig. 2a, Supplementary Fig. 2b). The HLA LOH detection strategy employed is both repeatable and reproducible, with a precision of 100% on 88% of samples (Fig. 2b). The limit of detection (LOD) was established using a dilution study modeled after the probit approach (CLSI EP17-A2) and was determined to be at 40% tumor percentage where 100% (48/48) of positive samples were correctly classified (Fig. 2c)[26].

The sensitivity declined at tumor purities less than the LOD, with 71% and 27% of positive samples being detected at 30% and 20% tumor purity, respectively (Fig. 2c). Examples of decreased sensitivity below the LOD were observed in clinical cases, as shown in Fig. 2d where the HLA LOH status of the tumor sample was originally called negative at 30% tumor purity despite some evidence of HLA LOH, and then called positive when re-assessed on a second sample with 70% tumor purity. In contrast, false positive detection of LOH at the *HLA-A* locus was not influenced by low tumor purity (heterozygous *HLA-A* genotypes, Supplementary Fig. 2c). The minimal risk of false positive HLA LOH calls combined with the results from the limit of blank (LOB) study suggests that the test can be run at tumor purities as low as 20% (Supplementary Fig. 2c) with the risk of false negative results increasing below 40% tumor purity. The HLA LOH detection strategy employed is not impaired by a series of potential interfering substances (Supplementary Fig. 2d).

The HLA LOH assay accurately assesses LOH prevalence in large-scale analysis of multiple tumor types

Performance was assessed in a dataset of 10,982 patients from the Tempus Database. We determined the prevalence of clonal HLA LOH in several common cancer types and compared results to previous studies[10,12,13,15]. HLA LOH was common, with over 10% of tumors identified with LOH across

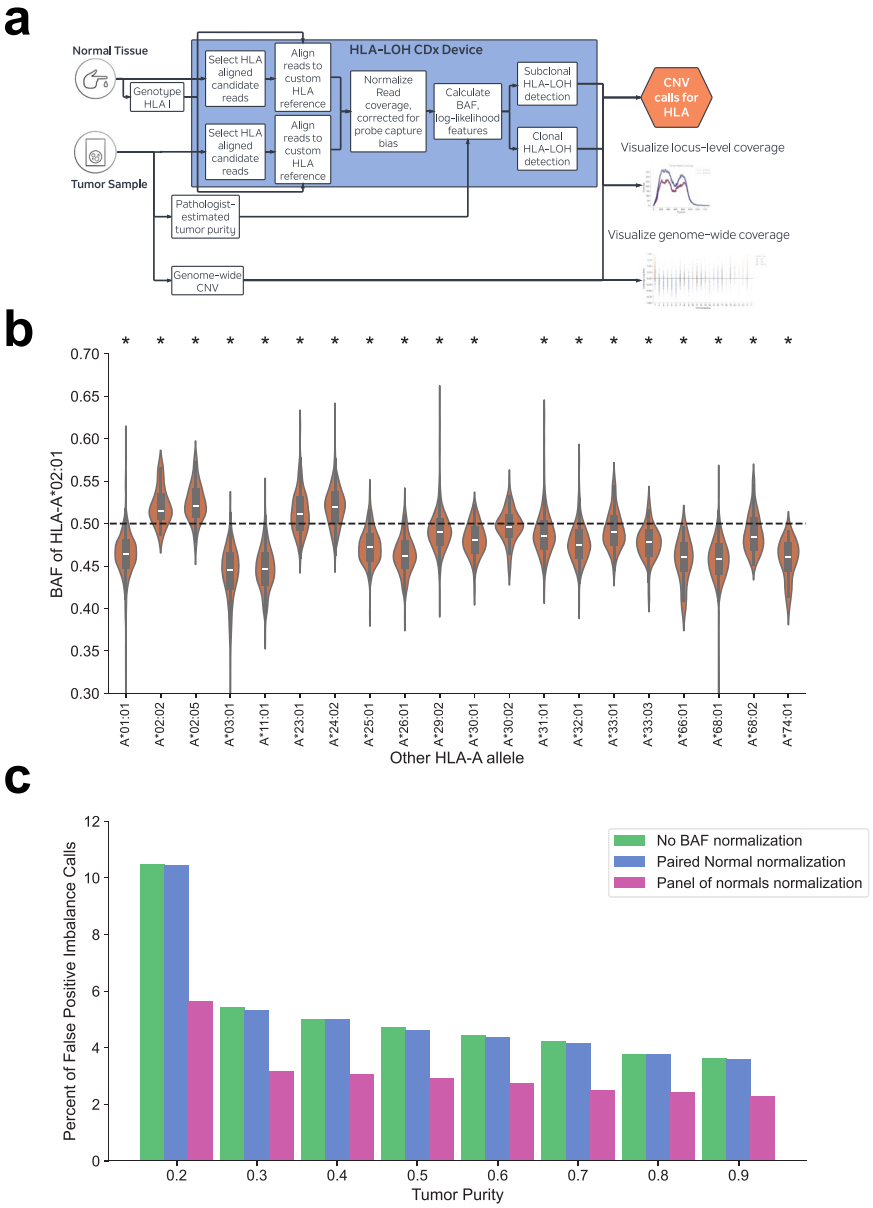


Fig. 1 | Description of the device employed for HLA genotype analysis. a Workflow describing the main steps of the HLA-LOH test. **b** B-allele fraction of the *HLA-A*02:01* allele in normal samples. Each violin plot only represents samples with a given genotype (eg. A*02:01/A*01:01). Most distributions significantly differ from the expected B-allele fraction (0.5), highlighting an experimental bias (caused most likely by the panel hybrid capture) that needs to be corrected. **c** False Positive Rate of the device on normal samples with different bias correction methods. The no B-allele fraction normalization uses the tumor sample B-allele fraction and coverage to determine HLA-LOH status. The paired normal normalization uses the patient's paired normal B-allele fraction and coverage to correct the observed bias. The panel of normals normalization uses all available normal samples with the same HLA genotype to correct the observed bias.

nearly all cancer types, with prevalence ranging from 3.3% in prostate cancer to 25.4% in head and neck squamous cell carcinoma (Table 1). These observations are consistent with results reported by The Cancer Genome Atlas Network (TCGA), Montesion et al., and the Martinez-Jimenez et al. cohorts, despite differences in cohort composition and HLA LOH detection methodology[10,13,15].

Real-world overall survival analysis

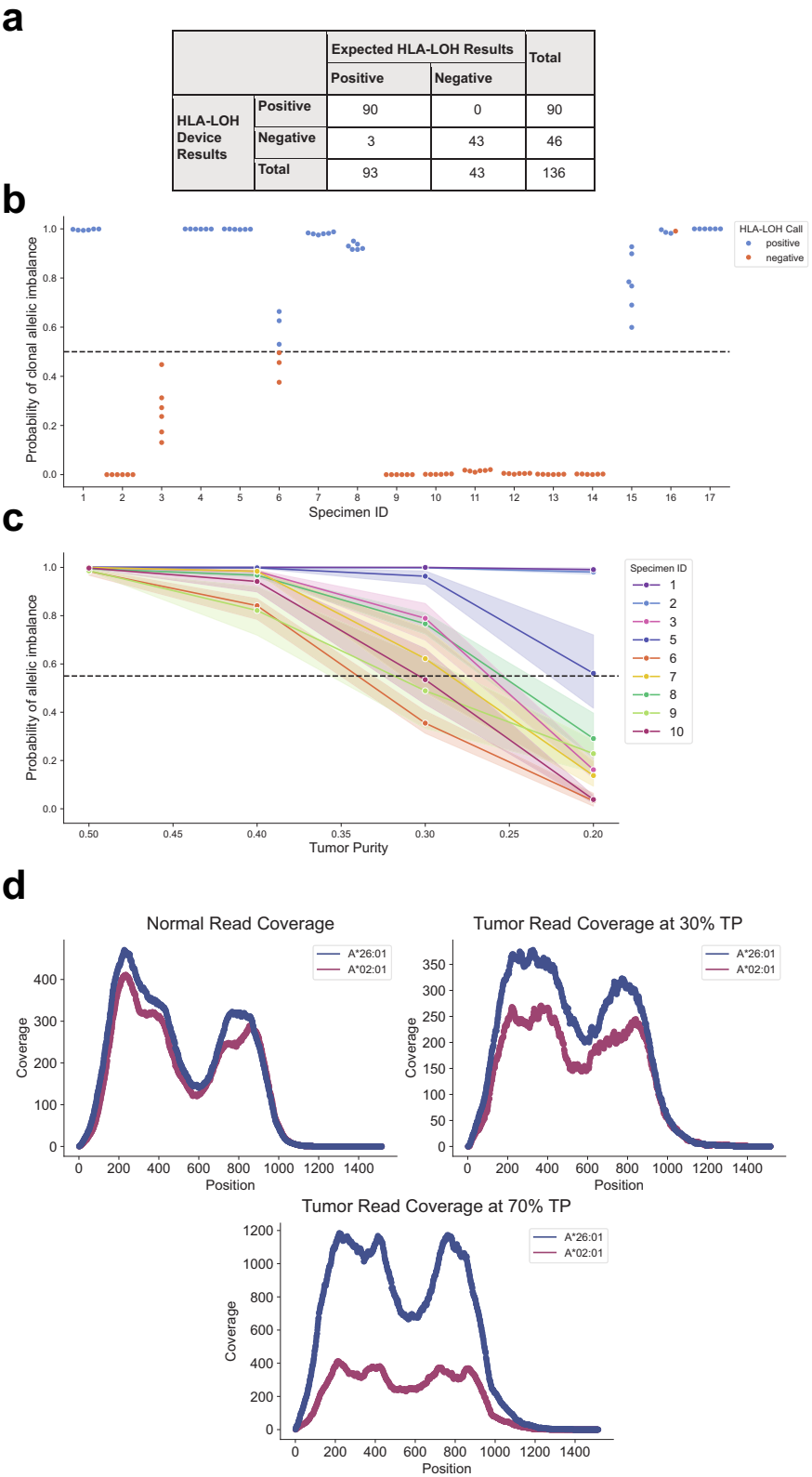
Having established that the LOH test accurately identified HLA LOH in tumors concordant with previous surveys of LOH prevalence, we next sought to determine the prognostic relevance of HLA LOH as a potential immune escape mechanism among patients receiving ICB. To do so, we leveraged the Tempus Database to study a cohort of 256 metastatic NSCLC patient samples (177 lung adenocarcinomas [LUAD] and 79 squamous cell carcinomas [LUSC]) whose tumors were collected prior to treatment with any FDA-approved regimen of ICB therapy (Supplementary Table 1).

Within the LUSC cohort, patients with clonal LOH at the *HLA-A* locus prior to ICB initiation tended to have shorter rwOS than patients with stable HLA loci (Fig. 3a and c), supporting the idea that antigen presentation via HLA-I is important for ICB response in this cohort[27–29]. We observed the opposite association within the LUAD cohort, finding that patients with *HLA-A* LOH had modestly longer rwOS following ICB treatment (Fig. 3b and c). To further understand immunological differences between these subgroups, we compared two biomarkers predictive of ICB response, TMB and PD-L1 protein expression measured by immunohistochemical staining, as well as immunological gene expression scores to assess the composition of the tumor-immune microenvironment (Supplementary Fig. 3). We found that LUAD patients with clonal HLA LOH had significantly higher TMB ($p = 0.024$, Wilcoxon), but found no significant differences in PD-L1 IHC staining or cytotoxicity-related gene expression scores (Supplementary Fig. 3). To confirm the association between HLA LOH and overall survival with a more stringent definition of HLA LOH, we extended the HLA LOH test as described for *HLA-A* to the *HLA-B* and *HLA-C* loci as well, and compared samples with consistent clonal LOH or HLA stable calls across all three loci (Supplementary Fig. 4). In this smaller set of samples, we again observed that LUAD samples with HLA LOH tended to have longer overall survival, whereas LUSC samples with HLA LOH were associated with shorter overall survival.

Fig. 2 | Performance of HLA-LOH test. a Summary of the accuracy study contingency results. **b** Swarm plot showing the results of the precision study. A total of 17 archival clinical tumor specimens with tumor percentages of 40% or 60% were selected and processed by multiple operators in replicates of 2 over 3 non-consecutive days. The dashed line denotes the threshold for detecting clonal allelic imbalance. **c** Line graph showing the LOD of the device at 40% tumor purity. The dashed line denotes the threshold for detecting allelic imbalance. **d** An example of LOD in real-world clinical cases. The two tumor specimens were extracted from the same sample, wherein the higher tumor purity specimen exhibited a stronger signal of *HLA-A*02:01* LOH.

a

		Expected HLA-LOH Results		Total
		Positive	Negative	
HLA-LOH Device Results	Positive	90	0	90
	Negative	3	43	46
	Total	93	43	136



Evaluating the HLA-A*02 LOH assay to identify patients for enrollment in prospective clinical studies

To evaluate the feasibility of this assay to identify cancer patients with tumor-associated *HLA-A*02* LOH, we established a prospective non-interventional, observational study (BASECAMP-1, clinicaltrials.gov NCT04981119). If eligible, patients are considered for treatment in a therapeutic study that uses a CAR T designed with a CEA-targeted activator paired with a blocker that recognizes HLA-A*02[30]; the blocker is included to prevent on-target, off-tumor toxicity in normal tissues while enabling the tumor-specific killing of cells with *HLA-A*02* LOH (EVEREST-1, A2B530, NCT05736731, (Supplementary Fig. 6)). In this ongoing study, patients are either directly tested for *HLA-A*02* LOH with the LOH test or identified

Table 1 | Prevalence of HLA class I LOH across cancer indications and datasets

	Tempus (n = 10982)		TCGA (n = 4693)		Montesion et al.[10] (n = 50139)		Martinez-Jimenez et al. (n = 953)	
Cancer Type	**No.**	**% LOH**	**No.**	**% LOH**	**No.**	**% LOH**	**No.**	**% LOH**
Colorectal Cancer	1854	15.6	615	9.6*	10682	15.3	42	16.7
Gastroesophageal Cancer	506	20.8	625	16.2	3174	22.2	137	24.8
Pancreatic Cancer	675	19.6	184	33.1*	4049	23.4*	284	33.8*
Prostate Cancer	1479	3.3	500	4.5	2774	5.8*	151	0.7
Ovarian, Fallopian tube, Primary peritoneal cancers	1383	17.5	579	17.0	4996	15.7	49	16.3
Non-Small Cell Lung Cancer (LUAD + LUSC)	2249	24.6	1017	18.4*	13240	23.0*	83	30.1
Lung Adenocarcinoma (LUAD)	1589	21.8	516	11.6*	NA	NA	NA	NA
Lung Squamous Cell Carcinoma (LUSC)	660	31.5	501	25.4	NA	NA	NA	NA
Breast Cancer	2426	12.2	1080	14.0	9686	13.2	161	11.2
H&N Squamous cell carcinoma	394	25.4	522	16.0*	1134	27.2	46	26.1
Mesothelioma	16	18.8	87	11.0	404	12.4*	NA	NA

*The prevalence significantly differs from Tempus prevalence (chi-square test Bonferroni corrected p-value < 0.0019).
Tempus LOH prevalence based on the prevalence of LOH at *HLA-A*.

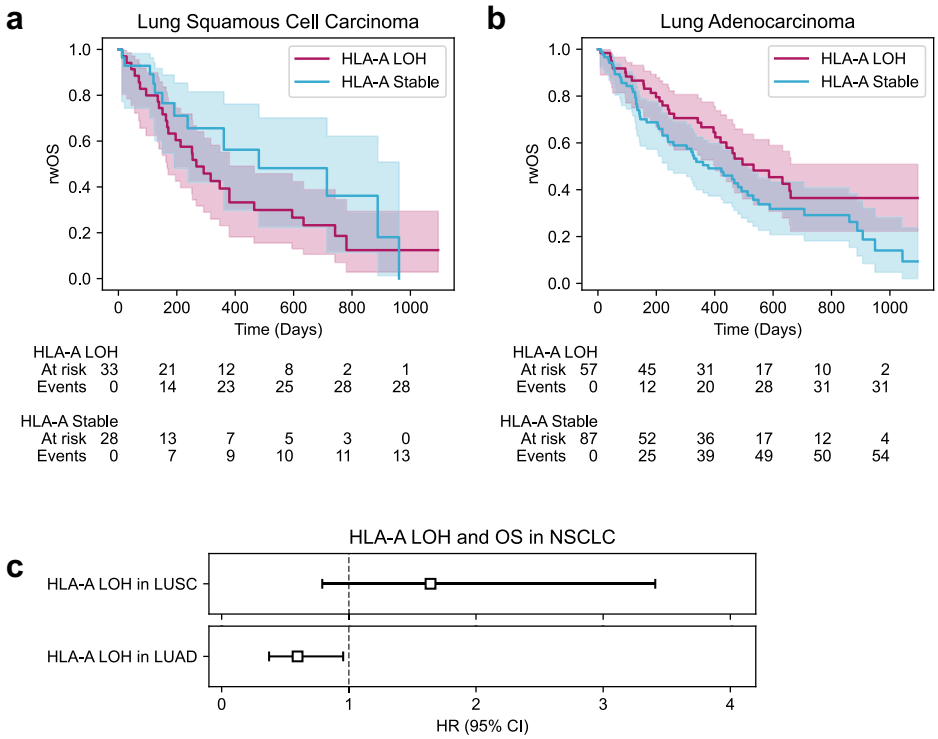


Fig. 3 | *HLA-A* LOH is associated with differential rwOS following ICB therapy in lung cancer. Kaplan-Meier analysis of rwOS for LUSC (**a**) and LUAD (**b**) patients with clonal *HLA-A* LOH (*HLA-A* LOH) and *HLA-A* stable patients. Median rwOS for LUSC *HLA-A* LOH patients was 267 days (95% CI [167, 380]), and median rwOS for LUSC *HLA-A* stable patients was 481 days (95% CI [191, 961]). In LUAD, median rwOS for *HLA-A* LOH patients was 533 days (95% CI [402, Inf]), whereas median survival for *HLA-A* stable patients was 378 days (95% CI [239, 513]). **c** Forest plot of univariable CoxPH models testing for associations between rwOS and *HLA-A* LOH within each of the LUSC and LUAD cohorts, stratified on treatment regimen type (ICB mono-therapy or ICB + chemotherapy) and line of therapy. In LUSC, patients with *HLA-A* LOH tended to have shorter rwOS (HR = 1.64, 95% CI [0.79, 3.41], p = 0.183), whereas in LUAD, *HLA-A* LOH is associated with longer rwOS (HR = 0.60, 95% CI [0.37, 0.96], p = 0.032).

through the Tempus AWARE program, an enhanced clinical trial matching and patient eligibility delivery program within healthcare institutions that identifies patients with specific molecular alterations, in this case, *HLA-A*02* LOH, through Tempus xT testing as part of routine clinical care. Eligible subjects with LOH in the tumor may then consent for BASECAMP-1 enrollment, apheresis collection, and consideration of treatment in the CAR T therapeutic study. At an interim data cutoff of September 19, 2023 (20 months of screening), clinical data from 1720 subjects were screened for *HLA-A*02* LOH through the AWARE program across 8 sites. Germline A*02 heterozygous alleles were noted in 602 (35.0%), and 26 of those had LOH at the A*02 allele (4.32%). Of these, 14 (54%) have consented to BASECAMP-1, 1 has declined (4%), and 12 remain in screening at sites (46%), demonstrating the feasibility of identifying patients for potential recruitment into interventional precision oncology studies targeting HLA LOH.

Discussion

HLA LOH represents a recurrent immune escape mechanism observed in diverse cancer types[7,10,12,13]. Clinical utilization of HLA LOH detection requires a sensitive and specific assay that can be seamlessly integrated into clinical workflows, ideally within an approved in vitro diagnostic test used for solid tumor profiling. Tempus' xT CDx is an FDA-approved in vitro diagnostic device designed to detect and quantify diagnostically and therapeutically relevant variants in patient tumors during routine standard of

care. We have found that this platform can accommodate *HLA-A* LOH detection with high sensitivity and specificity (Fig. 1). Of note, we observe low frequencies of false-positive calls and a sensitivity that depends on tumor purity. These results suggest that this assay could be clinically relevant for detecting *HLA-A* LOH in patient samples with moderate to high tumor purity (i.e., > 40%). Analytical validation was leveraged to support designation of the HLA LOH assay as a Breakthrough Device by the FDA. We functionally validated this assay in a large cohort of 10,982 patients, where we recapitulated a similar prevalence of HLA LOH across tumor types compared to prior reports from TCGA and *post hoc* analysis of data from Foundation Medicine (Table 1).

Notably, we found associations between clonal *HLA-A* LOH and rwOS that vary depending on the histology of the tumor, indicating tumor-specific prognostic relevance of HLA LOH. Our findings differ from those of Montesion et al., who found that HLA LOH in non-squamous lung cancer was associated with shorter OS. In contrast, our findings are complementary to those of Zhang et al.[31], who observed that LOH was associated with shorter overall survival in LUSC but not LUAD patients. These discrepancies in the association between OS and HLA LOH may be due to differences in the baseline patient populations, including tumor mutational composition and racial makeup. Furthermore, Montesion et al. evaluated patients with non-squamous histology receiving second-line ICB monotherapy, whereas our and the Zhang et al. cohorts included patients receiving both ICB monotherapy as well as ICB in combination with chemotherapy.

Our observation that *HLA-A* LOH is associated with slightly longer rwOS in LUAD samples is surprising, given that samples with LOH are expected to promote less immune engagement during ICB treatment due to impaired antigen presentation ability. This finding was repeated in an alternate analysis using multivariable CoxPH models to attempt to identify and account for additional covariates which might impact this result (Supplementary Fig. 5, Supplementary Table 4). Interestingly, the association between LOH and rwOS in LUAD was less affected by additional covariates than in LUSC. These results suggest that mechanisms of immune engagement may play different roles in the tumor-immune microenvironment of different tumor types and in different treatment settings. We also note that TMB is higher in LUAD patients with clonal LOH compared to HLA stable patients, but not in LUSC patients (Supplementary Fig. 3), reproducing the results of McGranahan et al. Interestingly, we did not observe differences in cytotoxicity scores or PD-L1 comparable to those reported by McGranahan et al. This may be related to tumor stage differences between the two cohorts; the rwOS analysis cohort considered here was composed entirely of stage IV patients, whereas the McGranahan cohort represented early-stage disease. Nonetheless, further studies in larger independent cohorts is warranted to reconcile these differences.

HLA LOH is relevant for several therapeutic opportunities. The irreversible and clonal nature of HLA LOH at the DNA sequence level presents a fixed genetic alteration in tumor cells that definitively distinguishes them from nonmalignant cells[32–34]. CAR T logic-gated strategies[30,35] can exploit this difference using the co-expression of a modular activator receptor paired with a blocking receptor that recognizes a specific *HLA* allele. In contrast, T cell receptor (TCR)-restricted therapies and neoantigen vaccines target HLA-presented peptide antigens; patients with LOH of the *HLA* allele responsible for presenting this antigen will be resistant to therapy. Thus, HLA LOH analysis will be important to appropriately enroll patients in both logic-gated CAR T studies and TCR-restricted studies.

In conclusion, HLA LOH is observed in diverse tumors and has diagnostic, prognostic, and therapeutic relevance. We have established and validated an LOH detection test which was granted Breakthrough Device designation by the FDA, and demonstrated that subjects with HLA-specific LOH can be identified for prospective clinical studies by leveraging analysis within a routine clinical workflow (e.g., AWARE). This strategy could have implications for the application of diverse precision oncology therapies.

Methods

Tempus xT CDx sequencing

Tempus AI, Inc. ("Tempus") provides genomic testing for patients and physicians to guide treatment options and to identify patients eligible for clinical trials. The Tempus xT CDx assay is an FDA-approved (P210011) next-generation sequencing (NGS) in vitro diagnostic device that targets 648 actionable oncogenes[24]. The laboratory procedures are carried out in Tempus' College of American Pathologists (CAP)-accredited and Clinical Laboratory Improvement Amendments (CLIA)-certified laboratory.

The device uses solid tumor DNA extracted from formalin-fixed, paraffin-embedded (FFPE) tumors and matched normal specimens (saliva or blood) obtained from previously diagnosed patients with solid malignant neoplasms. Total nucleic acid is extracted from FFPE tumor slides and matched to normal blood or saliva. Slides are evaluated by pathologists and microdissected to meet a baseline requirement of 20% tumor cellularity. Nucleic acid is extracted (minimum extraction yield of 50 ng) using the Chemagic360 instrument, stored at -80°C, and quantified. Specimens are sonicated to achieve 200 base pair-sized fragments using a Covaris LE220 followed by library construction (KAPA Hyper Prep Kit (List KK8504)). If applicable, a bead-based size selection step is instituted to enrich target sequences at the library construction step. After elution from the magnetic bead cleanup, the control and study specimen libraries are evaluated. A minimum library yield of 175 ng is required as input into the hybridization step. Isolation of captured target sequences is performed using the xGen Hybridization and Wash Kit (List 1080577 or 1080584). The captured target sequences are amplified using the KAPA Library Amplification Kit (List KK2621 or KK2620). Specimens are then processed through a post-capture cleanup using the Axygen AxyPrep Mag PCR Clean-Up Beads (List MAGPCR-CL-250). A minimum molarity of 2 nM and a peak size of 200–800 base pairs of the post-capture library is required as input into sequencing. Specimens are sequenced on the Illumina® NovaSeq 6000 sequencing platform to > 500x median coverage of tumor samples with > 95% of exons at > 150x coverage and ≥98% of exons at ≥ 100x coverage. The quality of the sequencing data is evaluated post-sequencing to ensure minimum coverage and exclude poor-quality samples as well as samples with potential contamination or swaps. Sequence data are then processed using a customized analysis pipeline designed for use in the detection of substitutions (single-nucleotide variants [SNVs] and multi-nucleotide variants [MNVs]), and insertion and deletion alterations (INDELs).

The HLA LOH assay and bioinformatics pipeline

An overview of the HLA LOH assay workflow is presented in Fig. 1a. This assay is an extension of the Tempus xT CDx functionality to detect LOH events in clinically actionable and biologically relevant HLA alleles for clinical trial enrollment and investigational use. The assay is capable of detecting LOH in *HLA-A*, *HLA-B*, and *HLA-C*; however, the validation study and downstream analyses presented here are exclusively in *HLA-A*. Following Tempus xT CDx sequencing, reads mapping to the larger HLA region and all unmapped reads are selected. A custom HLA reference is generated using the sample's germline genotype (provided as input), obtained from a matching blood or saliva sample. The sample's germline genotype is determined using the open-source tool Optitype[36], which displays an accuracy ≥95% both in benchmark studies and our own experiments. Candidate reads are aligned to the custom reference. The alignment of the germline to normal sample is used to verify the HLA genotype provided as input. If any variants with >40 supporting reads and a variant allele fraction superior to 75% are detected, the input genotype is considered to be incorrect and the test will return an error. In this validation study, however, there were no instances of incorrect input genotypes and thus no errors were returned. Once the genotype is confirmed, a strict read filtering is applied on both the tumor and germline alignments, where only correctly paired reads mapping exclusively to one allele with an edit distance of 0 are conserved. BEDTools[37] is then used to determine the read depth at every position for each allele. The panel of normals is applied by replacing the

germline read depth of the sample with the median read depth of samples with the same genotype in our panel. The features are derived from read depth at each genomic position and for each allele the log ratio of the read depth in the tumor and normal sample and the B-allele fraction (BAF) are calculated.

Only positions with >40X read depth in both alleles in both tumor and normal samples are sufficient and considered to be "high-coverage", which are used to generate the sample-level features. Samples with fewer than 300 high-coverage positions do not receive a classification due to low coverage. The sample-level features are the median BAF ratio (ratio of the BAF in the tumor to the BAF in the normal), the median of the difference between the log ratio of the read depth (logR diff), and the difference between logR diff to the expected value of that feature at the given tumor purity (ratio expected difference).

Once features are generated, two logistic regression models are applied. The subclonal LOH detection model takes the BAF of the non-targeted allele, the difference between the log ratios, and tumor purity as input. Samples with a probability of allelic imbalance inferior to 55% are classified as stable. The remaining samples are then fed into the clonal model. The clonal LOH detection model takes the BAF of the targeted allele, ratio expected difference, and tumor purity as input. Samples with a probability of clonal allelic imbalance inferior to 50% are classified as subclonal allelic imbalance; samples with a probability of clonal allelic imbalance superior to 50% are classified as clonal allelic imbalance. Both models were trained on manually labeled data. Finally, the algorithm determines whether a clonal allelic imbalance is an LOH or an amplification.

Once segmentation of the genome has been performed and copy number values have been assigned to every segment using our proprietary genome-wide copy number variation (CNV) calling algorithm, the segment overlapping with the HLA locus is selected. If this segment is determined to be a gain (major > 1) and does not present signs of LOH (minor > 0), the sample is classified as having an amplification of the HLA locus. Otherwise, the subclonal or clonal loss call is maintained. If no segment fully overlaps with the HLA locus, we use the gain/loss status of the segment either to the left or the right of the HLA locus to determine whether this locus shows signs of allelic imbalance (major =/= minor). If both segments to the right and left show no signs of allelic imbalance, the loss call is maintained.

Real-world cohort selection and feature generation to assess prognostic relevance of HLA LOH

De-identified records from a cohort of patients who received Tempus tissue-based NGS testing ($n = 256$) were selected from the Tempus Database (Tempus AI, Inc., Chicago, IL; Supplementary Table 1). This study was conducted on de-identified health information subject to an institutional review board-exempt determination (Advarra Pro00072742) and did not involve human subjects research. Patient records collected between 2016-2023 were selected for inclusion in the study based on the following criteria: (1) a diagnosis of stage IV non-small cell lung cancer (NSCLC) with either squamous cell carcinoma or adenocarcinoma histology; (2) received an FDA-approved immune checkpoint blockade (ICB)-containing regimen in the first, second, or third line of therapy; (3) regimen start date was on or after sample collection; (4) computationally assessed tumor purity ≥40%; (5) tumor sample did not have any somatic pathogenic mutations identified in *EGFR* or *ALK*; (6) a call of 'clonal loss' or 'stable' was made at the *HLA-A* locus. Patients with partial loss calls were excluded. The subset of these records with a consistent call of 'clonal loss' or 'stable' at all three classical class I HLA loci (*HLA-A*, *-B*, and *-C*) was also evaluated separately. Tumor mutational burden (TMB) was calculated for each sample as previously described[38].

RNAseq data generated using the Tempus xR whole-transcriptome RNA sequencing platform was used as the source of RNAseq data for this cohort[39]. Gene-level transcripts per million reads (TPM) values were used to compute cytotoxicity scores as previously described[40,41].

Real-world overall survival analysis

Survival analysis was performed in Python (v3.7.11) using the 'lifelines' package (v0.27.7[42]). Real-world overall survival (rwOS) was assessed for up to 3 years following first-line ICB initiation. Cox proportional hazards (CoxPH) or Kaplan-Meier (KM) models were fit to rwOS within each cohort with HLA status (LOH or stable) as the independent variable using the risk set adjustment method[43]. To account for differences in baseline hazards between treatment regimens and lines of therapy, stratified CoxPH models were used[40]. Regimens were grouped into classes of 'ICB monotherapy' or 'ICB + chemotherapy'. No differences in LOH prevalence were detected between patients who received ICB in the first line vs. later lines of therapy (Fisher's exact $p = 0.13$, Supplementary Table 2).

Study patients

BASECAMP-1 (NCT04981119) is an observational study to determine germline HLA genotypes and screen for tumor-associated LOH. Subjects screened at 8 BASECAMP-1 sites using the Tempus AWARE program were clinically tested using Tempus xT and met the following clinical parameters: germline sample available for LOH analysis, age ≥ 18, and unresectable/metastatic colorectal, lung, or ovarian cancer or any stage pancreatic cancer or mesothelioma. BASECAMP-1-participating patients with LOH of *HLA-A*02* can be considered for the Phase 1 clinical trial EVEREST-1 (NCT05736731). BASECAMP-1 and EVEREST-1 are approved by individual institutional review boards at each study site. Patients provided written informed consent for both studies and for AWARE participation as a part of testing.

Contrived sample method and QC for cell line simulated samples in an accuracy study

Cell lines were purchased from the Fred Hutchinson International Histocompatibility Working Group Cell and Gene Bank as purified genomic DNA and detailed in Supplementary Table 5. The *HLA-A*02:01* cell lines were represented by IHW09287, IHW09046, IHW09031, IHW09039, IHW09004, IHW09056, IHW09062, IHW09059, IHW09084, IHW09036, IHW09058, IHW09070, IHW09068, and IHW09052. Each of the *HLA-A*02:01* cell lines (cell line 1) was mixed with two different non-*HLA-A*02:01* cell lines (cell line 2) separately to make two cell line pairs per *HLA-A*02:01* cell line. Each cell line was mixed at a 1:1 mass ratio to simulate an expected HLA stable specimen (i.e., negative control). These cell lines were then mixed at five different mass ratios to simulate LOH signal consistent with clinical specimens across tumor percentages ranging from 20%-90%. These specimens were mixed such that the relative percentage of DNA input for the *HLA-A*02:01* cell line and the non-*HLA-A*02:01* cell lines, respectively, were 44.4% and 55.6% to simulate LOH of a tumor specimen with 20% purity; 37.5% and 62.5% to simulate LOH of a tumor with 40% purity; 28.6% and 71.4% to simulate LOH of a tumor with 60% purity; 16.7% and 83.3% to simulate LOH of a tumor with 80% purity; and 9.1% and 90.9% to simulate LOH of a tumor with 90% purity.

To perform cell line contrived quality control (QC) calculations, germline variants (determined by the Tempus xT CDx device) shared among samples contrived under different mixture ratios were selected. Variants with a VAF standard deviation of <5% were removed because they were homozygous or heterozygous in both cell lines and not informative in determining the intended proportion of each cell line.

For each remaining variant, a line of best fit was calculated using the target proportion of cell line 1 (p) as the independent variable and the observed VAF as the dependent variable. The slope was used to determine the variant's zygosity status in both cell lines. Variants with a slope within 0.15 of the expected slope were labeled as one of the four categories: heterozygous in cell line 1 and homozygous in cell line 2 (VAF = 1–0.5p, expected slope = -0.5), homozygous in cell line 1 and heterozygous in cell line 2 (VAF = 0.5 + 0.5p, expected slope = 0.5), only homozygous in cell line 1 (VAF = p, expected slope = 1), only homozygous in cell line 2 (VAF = 1-p,

expected slope = -1). Variants with other zygosity combinations could not be annotated as germline in the Tempus xO variant calling pipeline and thus were not considered.

After determining the variant's zygosity, the observed proportion was calculated using the observed VAF and the corresponding equation based on the variant's category. The observed proportion of cell line 1 in the admixture was calculated as the mean of observed proportions from all identified variants. Samples passing this contrived sample QC had an observed proportion within 5% of the target proportion.

28 samples were excluded using this method. Of those 11 were true positive samples, 9 were true negative samples, 5 were false negative samples, 1 was a false positive sample and 2 were excluded from the study for other quality reasons.

Data availability

De-identified data used in the research was collected in a real-world healthcare setting and is subject to controlled access for privacy and proprietary reasons. Ethics committee and/or informed consent does not allow for public availability. When possible, derived data supporting the findings of this study have been made available within the Article, its Supplementary Figures/Tables, and Source Data. Tempus may make access to further data available pending a signed data use agreement; requests for access should be sent to publication.inquiry@tempus.com. TCGA data used in this study are available from the cBioPortal, data from Montesion et al. is available in the supplementary data tables from the publication[10].

Code availability

Requests for code used to generate the figures are subject to a non-commercial data software agreement and review by Tempus and can be directed to publications@tempus.com.

Received: 19 December 2023; Accepted: 24 July 2024;
Published online: 05 August 2024

References

1. Garrido, F., Cabrera, T. & Aptsiauri, N. 'Hard' and 'soft' lesions underlying the HLA class I alterations in cancer cells: implications for immunotherapy. *Int. J. Cancer* **127**, 249–256 (2010).
2. Smith, M. E., Marsh, S. G., Bodmer, J. G., Gelsthorpe, K. & Bodmer, W. F. Loss of HLA-A,B,C allele products and lymphocyte function-associated antigen 3 in colorectal neoplasia. *Proc. Natl Acad. Sci. USA* **86**, 5557–5561 (1989).
3. Marincola, F. M. et al. Loss of HLA haplotype and B locus down-regulation in melanoma cell lines. *J. Immunol.* **153**, 1225–1237 (1994).
4. Feenstra, M. et al. HLA class I expression and chromosomal deletions at 6p and 15q in head and neck squamous cell carcinomas. *Tissue Antigens* **54**, 235–245 (1999).
5. Jiménez, P. et al. Chromosome loss is the most frequent mechanism contributing to HLA haplotype loss in human tumors. *Int. J. Cancer* **83**, 91–97 (1999).
6. Maleno, I., López-Nevot, M. A., Cabrera, T., Salinero, J. & Garrido, F. Multiple mechanisms generate HLA class I altered phenotypes in laryngeal carcinomas: high frequency of HLA haplotype loss associated with loss of heterozygosity in chromosome region 6p21. *Cancer Immunol. Immunother.* **51**, 389–396 (2002).
7. McGranahan, N. et al. Allele-specific HLA loss and immune escape in lung cancer evolution. *Cell* **171**, 1259–1271.e11 (2017).
8. Garrido, M. A. et al. HLA class I alterations in breast carcinoma are associated with a high frequency of the loss of heterozygosity at chromosomes 6 and 15. *Immunogenetics* **70**, 647–659 (2018).
9. Krishna, C. et al. An immunogenetic basis for lung cancer risk. *Science* **383**, eadi3808 (2024).
10. Montesion, M. et al. Somatic HLA class I loss Is a widespread mechanism of immune evasion which refines the use of tumor mutational burden as a biomarker of checkpoint inhibitor response. *Cancer Discov.* **11**, 282–292 (2021).
11. Zhao, J. et al. The prevalence of HLA-I LOH in Chinese pan-cancer patients and genomic features of patients harboring HLA-I LOH. *Hum. Mutat.* **42**, 1254–1264 (2021).
12. Pyke, R. M. et al. A machine learning algorithm with subclonal sensitivity reveals widespread pan-cancer human leukocyte antigen loss of heterozygosity. *Nat. Commun.* **13**, 1925 (2022).
13. Martínez-Jiménez, F. et al. Genetic immune escape landscape in primary and metastatic cancer. *Nat. Genet.* **55**, 820–831 (2023).
14. Lim, W. C. et al. Divergent HLA variations and heterogeneous expression but recurrent HLA loss-of- heterozygosity and common HLA-B and TAP transcriptional silencing across advanced pediatric solid cancers. *Front. Immunol.* **14**, 1265469 (2023).
15. The Cancer Genome Atlas program (TCGA). CCG- *National Cancer Institute* https://www.cancer.gov/tcga (2022).
16. Simeonov, D. R. & Marson, A. CRISPR-based tools in immunity. *Annu. Rev. Immunol.* **37**, 571–597 (2019).
17. Clubb, J. D., Gao, T. A. & Chen, Y. Y. Synthetic biology in the engineering of CAR-T and CAR-NK cell therapies: facts and hopes. *Clin. Cancer Res.* **29**, 1390–1402 (2023).
18. DiAndreth, B., Hamburger, A. E., Xu, H. & Kamb, A. The Tmod cellular logic gate as a solution for tumor-selective immunotherapy. *Clin. Immunol.* **241**, 109030 (2022).
19. Cable, J. et al. Synthetic biology: at the crossroads of genetic engineering and human therapeutics-a Keystone Symposia report. *Ann. N. Y. Acad. Sci.* **1506**, 98–117 (2021).
20. Van Loo, P. et al. Allele-specific copy number analysis of tumors. *Proc. Natl Acad. Sci. USA.* **107**, 16910–16915 (2010).
21. Koboldt, D. C. et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012).
22. Favero, F. et al. Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Ann. Oncol.* **26**, 64–70 (2015).
23. Larsen, B. M. et al. A pan-cancer organoid platform for precision medicine. *Cell Rep.* **36**, 109429 (2021).
24. U.S. Food and Drug Administration. *Tempus xT CDx FDA Label*. https://www.accessdata.fda.gov/cdrh_docs/pdf21/P210011C.pdf (2023).
25. Robinson, J. et al. IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. *Nucleic Acids Res.* **31**, 311–314 (2003).
26. Pierson-Perry, J. F. et al. *Evaluation of Detection Capability for Clinical Laboratory Measurement Procedures; Approved Guideline,* Vol. 68 (Clinical Laboratory Standards Institute, 2012).
27. Sade-Feldman, M. et al. Resistance to checkpoint blockade therapy through inactivation of antigen presentation. *Nat. Commun.* **8**, 1136 (2017).
28. Gettinger, S. et al. Impaired HLA Class I antigen processing and presentation as a mechanism of acquired resistance to immune checkpoint inhibitors in lung cancer. *Cancer Cancer Discov.* **7**, 1420–1435 (2017).
29. Zaretsky, J. M. et al. Mutations associated with acquired resistance to PD-1 blockade in melanoma. *N. Engl. J. Med.* **375**, 819–829 (2016).
30. Sandberg, M. L. et al. A carcinoembryonic antigen-specific cell therapy selectively targets tumor cells with HLA loss of heterozygosity in vitro and in vivo. *Sci. Transl. Med.* **14**, eabm0306 (2022).
31. Zhang, X. et al. Integrated investigation of the prognostic role of HLA LOH in advanced lung cancer patients with immunotherapy. *Front. Genet.* **13**, 1066636 (2022).
32. Tokatlian, T. et al. Mesothelin-specific CAR-T cell therapy that incorporates an HLA-gated safety mechanism selectively kills tumor cells. *J. Immunother. Cancer* **10**, e003826 (2022).
33. Hamburger, A. E. et al. Engineered T cells directed at tumors with defined allelic loss. *Mol. Immunol.* **128**, 298–310 (2020).

34. Savanur, M. A., Weinstein-Marom, H. & Gross, G. Implementing logic gates for safer immunotherapy of cancer. *Front. Immunol.* **12**, 780399 (2021).
35. Hwang, M. S. et al. Targeting loss of heterozygosity for cancer-specific immunotherapy. *Proc. Natl Acad. Sci. USA* **118**, e2022410118 (2021).
36. Szolek, A. et al. OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics* **30**, 3310–3316 (2014).
37. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
38. Aggarwal, C. et al. Assessment of tumor mutational burden and outcomes in patients with diverse advanced cancers created with immunotherapy. *JAMA Netw. Open* **6**, e2311181 (2023).
39. Tempus xT & xR. *Tempus* https://www.tempus.com/oncology/genomic-profiling/xt-xr/ (2023).
40. Lau, D. et al. Integration of tumor extrinsic and intrinsic features associates with immunotherapy response in non-small cell lung cancer. *Nat. Commun.* **13**, 4053 (2022).
41. Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **160**, 48–61 (2015).
42. Davidson-Pilon, C. lifelines: survival analysis in python. *J. Open Source Softw.* **4**, 1317 (2019).
43. Tsai, W.-Y., Jewell, N. P. & Wang, M.-C. A note on the product-limit estimator under right censoring and left truncation. *Biometrika* **74**, 883–886 (1987).

Acknowledgements

The authors thank Amrita A. Iyer, Ph.D. and Matthew Kase, M.S. of Tempus AI, Inc., Chicago, IL, USA for providing medical writing and editorial support, which was funded by Tempus AI, Inc., Chicago, IL, USA in accordance with Good Publication Practice (GPP3) guidelines.

Author contributions

Methods development and design: A.L., Q.Y., J.S.W., T.D. Generation and analysis of data: A.L., Q.Y., M.G.R., T.D. Provision of patients: E.B.G, K.K, D.M.S, J.R.M., J.R.H. Interpretation of data: all authors. Drafting of manuscript: A.L., T.D., Q.Y., M.G.R., J.S.W., A.A.S., M.M.S. Review, edits, and final approval of completed manuscript: all authors.

Competing interests

A.L., T.D., Q.Y., M.G.R., and M.M.S. are employees and own restricted stock units of Tempus AI, Inc. A.S. owns restricted stock units of Tempus AI, Inc. A.L. and W.Y.G. hold patents related to the HLA LOH test (A.L. US20230154563A1; W.Y.G MX2020007543A). J.S.W., W.Y.G., E.W.N., and A.M. are employees of A2 Biotherapeutics and hold A2 Biotherapeutics stock options. D.G.M. holds stock options in A2 Biotherapeutics and Navan Technologies, and research funding from Kite Pharma, Juno Therapeutics, Celgene, Legend Biotech, and Bristol-Myers Squibb. E.B.G. holds grants from ABL-Bio; AstraZeneca; Bristol Myers Squibb; Dynavax Technologies; Eli Lilly; EMD Serono; Genentech; Iovance Biotherapeutics; Merck; Mirati Therapeutics; Neon; and Novartis, and is a consultant or advisor for: Abbvie; ABL-Bio; AstraZeneca, Boehringer-Ingelheim; Bristol Myers Squibb; Dracen Pharmaceuticals; EMD Serono; Eisai; Eli Lilly; Gilead; GlaxoSmithKline; Merck; Natera; Novartis; Personalis; Regeneron; Sanofi; Shionogi; and Xilio. Joel R. Hecht owns stock options of Actym, Triumvirate and Rafael. His research is funded by A2 Biotherapeutics, Revolution Medicine, Regeneron, Seagen, Pfizer, Janssen, Gilead, Exelixis, Affin-T, Neogene, Lyell, Cardiff, IGM, NGM, Tizona, Camurus, Crinetics, Gritstone, Mirati, Agenus. Advisor for BMS, Astellas, Novartis, BeiGene. K.K. receives travel support from A2 Biotherapeutics. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41698-024-00665-z.

Correspondence and requests for materials should be addressed to Michelle M. Stein.