Supplemental materials and methods

DNA and RNA extraction

DNA was extracted from frozen biopsy samples using the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's recommendations. For the cases analyzed via whole-exome sequencing (WES), the matched germline DNA was extracted from peripheral blood (n = 42). A total amount of 0.6 μg genomic tumor DNA per sample was used as input material for the DNA sample preparation.

Total RNA was extracted using the RNeasy Kit (Qiagen). RNA purity was checked using the NanoPhotometer® spectrophotometer (Implen GmbH, Munich, Germany) and RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA). A total amount of 3 μg RNA per sample was used as input material for the RNA sample preparations.

Whole-exome sequencing (WES)

For the cases analyzed by WES, DNA libraries were generated using the Agilent SureSelect Human All Exon Kit (Agilent Technologies) following the manufacturer's recommendations and then sequenced on the Illumina HiSeq platform. Quality control steps were as follows: (1) Discard paired reads if either one read contains adapter contamination (>10 nucleotides aligned to the adapter, allowing ≤ 10% mismatches); (2) Discard paired reads if > 10% of bases are uncertain in either one read; (3) Discard paired reads if the proportion of low quality bases (Phred quality <5) is > 50% in

either one read. Valid sequencing data was mapped to the reference human genome (UCSC hg19) by Burrows-Wheeler Aligner software for variations detection. Single nucleotide variants (SNVs) were identified by MuTect version 1.1.4. and copy number variations were identified by Contra version 2.0.4.

Targeted deep sequencing (TDS)

For the cases analyzed by TDS, target-specific primers were designed to flank sites of interest. Reads were aligned to UCSC hg19. Identified mutations were processed using in-house scripts to remove low-confidence and likely false positive mutations. To remove potential germline polymorphisms, the identified SNPs and Indels were filtered using the following steps: 1). SNPs annotated as synonymous were removed; 2). SNPs or Indels with a mutation allele frequency (MAF) ≥1% in databases of 1000 genome all, 1000 genome East Asia, or Esp6500, or with a MAF≥1% in databases of ExAC all or ExAC East Asia were removed. 3). SNPs or Indels defined as benign in the ClinVar database were filtered out; 4). SNPs or Indels detected in tumor samples, which were also detected in 42 matching peripheral blood DNA with MAF ≥1% were filtered out; 5). SNPs or Indels detected in >50% of samples in our cohort with MAF ≥10% were discarded; 6). All remaining SNPs and Indels with MAF between 10%-90% were kept.

RNA sequencing and gene set enrichment analysis (GSEA)

For the cases analyzed by RNA sequencing, the libraries obtained were sequenced on

an Illumina platform using 150 bp paired-end reads. Raw data (raw reads) of fastq format were firstly processed through in-house perl scripts for obtaining clean data (clean reads) with high quality. The reference genome and gene model annotation files were downloaded directly from the genome website. Then paired-end clean reads were aligned with the reference genome using Hisat2 v2.0.5.

Differential expression genes (fold change) were calculated on raw expression counts using the R package DESeq2. GSEA analysis was further performed using the clusterProfiler R package and significance threshold was set at |*normalized enrichment score*|>1, *p value* < 0.05, *false discovery rate q value* < 0.05. Three publicly available gene expression datasets (GSE117556, N = 928; GSE31312, N = 498; and GSE147986, N = 111) were downloaded from the NCBI GEO database (https://www.ncbi.nlm.nih.gov/geo/), as validation cohorts.

Quantitative real-time polymerase chain reaction

To characterize gene expression at the transcriptional level, samples harboring target gene variations were selected for the absolute and/or relative quantitative real-time polymerase chain reaction (PCR) assay. RNA samples were converted to cDNA using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA, USA) according to the manufacturer’s instructions. cDNA was further amplified by real-time PCR using the 7900HT Fast Real-Time PCR System (Applied Biosystems) with three technical repeats. For absolute quantitative expression analysis, gene expression was calculated from their Ct values and compared with the standard

curve. [1] The GAPDH gene was used as a reference for relative quantitative expression analysis. [2] Three samples of patients with lymphoproliferative diseases were used as a control.

Immunohistochemistry

Formalin-fixed paraffin-embedded (FFPE) DLBCL tissues from patients with PD-L1/2 and CD73/A2aR genetic variations were constructed to visualize protein expression by immunohistochemistry. Samples were deemed positive for protein expression when membranous or cytoplasmic staining was detected.

Multiplex immunofluorescence staining, multispectral imaging

In total, 4-5 μm FFPE whole tissue sections were deparaffinized and rehydrated prior to antigen recovery in heated citric acid buffer (pH 6.0) for 15 min using microwave incubation. Sections were blocked with blocking buffer (Dako, Carpinteria, CA, USA) for 10 min followed by overnight incubation at 4 ℃ with primary antibodies and the corresponding horseradish peroxidase-conjugated secondary antibody (PerkinElmer, Waltham, MA, USA). Each horseradish peroxidase-conjugated polymer mediated the covalent binding of a different fluorophore using tyramide signal amplification. This covalent reaction was followed by additional antigen recovery in heated citric acid buffer (pH 6.0) for 15 min to remove bound antibodies before the next step. The process was repeated in sequence for the following antibodies/ fluorescent dyes (listed in order): anti-CD8/Opal 520, anti-A2aR/Opal 620, and anti-PD-1/Opal 570. After the

Zhang T, *et al. J Immunother Cancer* 2022; 10:e004114. doi: 10.1136/jitc-2021-004114

three sequential reactions, sections were counterstained with DAPI for nucleolus staining (Life Technologies, Carlsbad, CA, USA) and mounted with VECTASHIELD hard set medium, and then stored in a light-proof box at 4 ℃ prior to imaging. Multiplex stained sections were imaged using the Mantra System (PerkinElmer), where 20 fields of view in each section were randomly acquired at 200× multispectral images for digital quantitative analysis. Image analysis and phenotyping was performed using InForm software (PerkinElmer). [3][4] The spectrally unmixed and segmented images were subjected to a proprietary active learning phenotyping algorithm. Specifically, each cell phenotype-specific algorithm was based upon an iterative training/test process, whereby a few cells (training phase) were manually selected as being most representative of each phenotype of interest; the algorithm then predicted the phenotype for all remaining cells (testing phase). The decisions made by the software could be overruled to improve accuracy until phenotyping was optimized. This allowed for the individual identification of each DAPI-stained cell according to their pattern of fluorophore expression and cell morphological features. InForm automatically derives the maps of cell membranes and fluorescent pixel intensity for CD8-, PD-1-, and A2aR-phenotyped cells within the image. Thresholds for "positive" staining and the accuracy of the phenotypic algorithms were confirmed by the pathologist for each case. The cells positive for CD8, PD-1, and A2aR expression were obtained according to the threshold. The average number of positive cells per mm^2 were calculated for 20 fields of view.

References

[1] Dhana Sekaran S, Doherty TM, Kenneth J, et al. Comparison of different standards for real-time PCR-based absolute quantification. *J Immunol Methods* 2010;354(1-2):34–9.

[2] Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 2001;29(9):e45.

[3] Carstens JL, Pedro CDS, Yang D, et al. Spatial computation of intratumoral T cells correlates with survival of patients with pancreatic cancer. *Nat Commun* 2017;8:1–13.

[4] Carey C D, Gusenleitner D, Lipschitz M, et al. Topological analysis reveals a PD-L1 associated microenvironmental niche for Reed-Sternberg cells in Hodgkin lymphoma. *Blood* 2017;130(22):2420–30.