SUPPLEMENTAL METHODS

RNA extraction, RNA sequencing, and quantitative real-time PCR

Total RNA was extracted from FFPE samples using the AllPrep DNA/RNA FFPE kit (Qiagen) per the manufacturer's guidelines. One microgram of extracted RNA was allocated for library preparation following the dUTP strand-specific protocol, with ribosome RNA removed using the Ribo-Zero rRNA Removal Kit (Human) (Illumina, San Diego, USA). 50bp single-end reads were generated on an Illumina HiSeq 4000 platform at The University of Chicago Functional Genomics Facility. The quality of raw paired-end (PE) reads was assessed by FastQC (v0.11.5).[1] Reads were aligned to human reference transcriptome (GRCh38) with Gencode gene annotation (v28)[2] using Kallisto (v0.43.1).[3] Transcript abundance was quantified at transcript level specifying strand-specific protocol by Kallisto and summarized into gene level by tximport (v1.6.0).[4] Lowly expressed genes were defined as those with CPM (counts per million of mapped reads) ≤ 2 in a sample and removed from further analysis. Genes with CPM > 2 in at least 6 samples were kept for normalization by TMM (trimmed mean of M values) method, followed by log2-transformation. Differentially expressed genes (DEGs) between *PD-L1* altered and *PD-L1* not altered groups were identified using Linear Models for Microarray Data (limma) voom method with precision weights (v3.34.5).[5] Significant DEGs were defined as those that passed filters of FDR-adjusted $P < 0.05$, and fold change ≥ 1.5 or ≤ -1.5.

GO (Gene Ontology)[6] biological processes and KEGG pathways [7] significantly enriched within the DEGs were identified using enrichGO and enrichKEGG function from Bioconductor package clusterProfiler (v3.6.0),[8] respectively. The prediction of upstream regulators in IPA was performed in accordance with previous studies.[9] Gene set enrichment analysis (GSEA)[10] was

performed using the javaGSEA desktop software to compare biologically relevant gene signatures in *PD-L1* gene-altered and not altered samples. Specifically, log2 transformed expression data was ranked according to the default Signal2Noise ranking metric, and GSEA was performed using datasets in the Hallmark,[11] Biocarta, and C2 gene sets from the Molecular Signature Database (MSigDB) (v6.2).[10] The statistical significance of the enrichment score was calculated by using default software settings, and permuting the phenotype labels 1000 times.[10,12]

Five hundred nanograms of extracted mRNA was also used to synthesize cDNA (Applied Biosystems high capacity cDNA reverse transcription kit) for quantitative real-time PCR of *PD-L1* transcripts. Primer pairs were generated across joining exons to avoid amplification of contaminating genomic DNA. Relative expression of *PD-L1* was assessed and normalized against a housekeeping gene (*β-actin*) using the Viia7 Applied Biosystem qPCR system.

DNA extraction and whole exome sequencing

DNA was isolated from FFPE samples using the AllPrep DNA/RNA FFPE kit (Qiagen) per the manufacturer's guidelines. Corresponding germline DNA was obtained from peripheral blood mononuclear cells using the DNeasy Blood and Tissue kit (Qiagen). Whole exomes and untranslated regions (UTR) were captured using the Agilent SureSelect Human All Exon V6 plus UTR kit (Agilent Technologies, Santa Clara, USA). 101bp PE reads were generated on an Illumina HiSeq 2500 instrument at Theragen Etex Bio Institute (Seoul, South Korea). Raw sequencing data were analyzed by an in-house pipeline constructed for WES analyses of paired or unpaired cancer genomes. The quality of raw reads was assessed by FastQC (v0.11.5),[1] and preprocessed to trim adaptors and low-quality bases using Trimmomatic (v0.36)[13] and merge 3' overlapping mates using FLASH (v1.2.11).[14] Reads were aligned to the decoy version of human

reference genome (GRCh38.d1.vd1) using BWA-MEM (v0.7.17)[15] with soft-clipping option activated by default. Read duplicates were marked using Sambamba (v0.6.7),[16] and alignments of mapping quality <30 were removed. Read alignment was further refined for insertions/deletions (indels) realignment and base quality score recalibration (BQSR) using GATK4 (v4.0.4.0).[17,18] Putative somatic mutations were identified using the GATK4-MuTect2 [19] following GATK's best practice protocol.[18] For the 19 tumors with matched normal tissues, variants were called using paired tumor-normal mode. For the 2 tumors without matched normal tissue available, we first generated a Panel of Normal (PON) using the normal samples of the other 19 tumors as recommended by GATK protocols, and then called somatic variants with this PON variant file as the control. The raw variant calls were post-processed by multiple filtering steps implemented in GATK4, followed by stringent downstream custom filters. The resulting high-quality variants were further filtered to remove potential germline variants identified as those at allele frequency (AF) $\geq$ 0.0001 in gnomMAD database (Genome Aggregation Database).[20]

Variants that passed all filters were carried on for annotation using ANNOVAR (April 2018 release).[21] Somatic mutation burden was calculated by the total number of somatic mutations that were predicted to cause a protein sequence change, including non-synonymous, stopgain, and stoploss SNVs (single-nucleotide variants), frameshift and non-frameshift indels, and variants that modify splicing sites. Putative neoantigens derived from protein-changing somatic mutations in each lymphoma specimen were predicted using netMHCpan (v4.0),[22] and filtered by gene expression from RNAseq data. Patients' MHC class I haplotypes were predicted from WES of germline DNA using Optitype (v1.3.1).[23] 9-mer mutant peptides were predicted using a sliding window approach centered at the mutated site of the protein sequence, and those

with a predicted binding affinity to HLA-A molecule <500 nM were selected as putative neoantigens.

External validation dataset

Preprocessed log2 ratios of hybridization probe intensities from Affymetrix SNP 6.0 arrays,[24] non-silent somatic mutations from whole exome sequencing in mutation annotation format (MAF),[25] gene expression FPKM (Fragments Per Kilobase of transcript per Million mapped reads)[26] values from RNAseq, and clinical data were downloaded for patients with DLBCL in a previously published study[27] as an external validation cohort from the National Cancer Institute's Genomic Data Commons data portal.[28] A total of 476 cases were included in the analysis.

Identification of *PD-L1* amplified cases in external dataset

The log2 ratios of hybridization probe intensities in the 476 DLBCL cases[27] were segmented using a circular binary segmentation (CBS) algorithm in DNAcopy (v1.56.0).[29] Discrete chromosomal regions with significant copy number variations (amplifications or deletions) were identified at a 99% confidence level in each sample using the Genomic Identification of Significant Targets in Cancer algorithm (GISTIC) (v2.0).[30] Patients with a GISTIC score of 2 at the *PD-L1* locus, corresponding to a log2 copy ratio ≥ 0.9 above baseline were classified as "amplified" (n = 21, 4.5%). Patients with a GISTIC score of 1 at the *PD-L1* locus, corresponding to a log2 copy ratio ≥ 0.1 above the baseline, were classified as "copy gain" (n = 46, 9.6%). Those with a GISTIC score of 0 or lower, corresponding to a log2 copy ratio < 0.1 above baseline, were classified as "non-amplified" (n = 409, 85.9%). The amplified group

(referred to hereafter as "*PD-L1* amplified") and non-amplified group were included in further analyses.

Clinical characterization and survival analysis

Clinical characteristics were extracted for the 476 DLBCL cases from Supplementary Data Appendix 2 of the published study,[27] matched to *PD-L1* amplification group assignments, and compared between *PD-L1* amplified and non-amplified cases by Fisher's exact test or Student's *t*-test (R v3.5.1).[31] Differences in the probability of overall survival (OS) and progression-free survival (PFS) between *PD-L1* amplified and non-amplified cases was tested by log-rank test using the R package survival (v2.43).[32] One and two year OS and PFS landmark analyses were performed using the R package landest (v1.0).[33]

Somatic mutation analysis

A total of 48745 non-silent somatic mutations identified in the 476 DLBCL cases were re-annotated using Oncotator (v1.9).[34] 24980 driver mutations that occurred at a significantly higher rate than expected background mutational rates (defined as q-score < 0.1) were identified using MutSig2CV (v3.11).[35] The frequency of driver mutations in each sample was compared between *PD-L1* amplified and non-amplified cases by Fisher's exact test (R v3.5.1). Benjamini and Hochberg (BH) FDR method was used for multiple testing correction of *P*-values.[36,37] Total tumor mutational burden (TMB) was assessed by the total number of non-silent protein-coding somatic mutations and compared between *PD-L1* amplified and non-amplified cases by non-parametric Wilcoxon rank-sum test (R v3.5.1).

RNAseq gene expression analysis

Preprocessed normalized and log2-transformed FPKM values derived from RNA sequencing experiments for the 476 DLBCL cases[27] were used for the identification of differentially expressed genes (DEGs) and pathways between *PD-L1* amplified and non-amplified cases. Raw RNAseq FastQ files were not yet released at the time of this study. Out of 25066 genes total, 17981 genes with log2-FPKM > 5 in at least 10 samples were kept for further analysis. Significant DEGs were identified using limma (v3.38.3),[38] filtered by FDR-corrected $P < 0.05$, and expression fold change ≥ 1.5 or ≤ -1.5. Gene ontology (GO)[6] terms and Kyoto Encyclopedia of Genes and Genomes (KEGG)[7] pathways significantly enriched in genes of interest were identified using clusterProfiler (v3.10.1)[8] at FDR-corrected $P < 0.20$. Using the significant DEGs as downstream target molecules in Ingenuity Pathway Analysis (IPA, © Qiagen, accessed Feb 6, 2019) causal network analysis,[9] the upstream regulators were identified in accordance with previous studies.[6,7,9] Gene set enrichment analysis (GSEA) was performed using the javaGSEA desktop software to compare biologically relevant gene signatures between *PD-L1* amplified and non-amplified cases. In brief, log2-transformed gene expression data were ranked higher to lower according to the default Signal2Noise ranking metric, and GSEA was performed using Hallmark, Biocarta, and C2 curated gene sets from the Molecular Signatures Database (MSigDB v6.2).[10,11] The statistical significance of the enrichment score was accessed using default software settings with 1000 permutations.[10,12]

RESULTS: ANALYSIS OF EXTERNAL DATASET

Identification and characterization of *PD-L1* amplified cases

Using the gene-centric and broad peak results from GISTIC analysis, 21 samples (4.5%) were identified as *PD-L1* amplified (see Supplemental Methods). *JAK2*, which also resides on the 9p24.1 cytoband, was co-amplified in 19 (90.4%) of 21 amplified samples. *PD-L1* amplified cases were associated with significantly increased *PD-L1* expression (*P* = 2.3e-21, limma DEG analysis) (supplemental Figure 8), consistent with prior reports [39,40]. For the current analysis, we focused exclusively on the extreme phenotypes (21 *PD-L1* amplified cases *versus* 409 *PD-L1* non-amplified cases). *PD-L1* amplified cases were significantly associated with the activated B-cell (ABC) gene expression subtype, higher international prognostic index score, and higher LDH levels (supplemental Table 3). No strong association with *PD-L1* amplification was detected for age, sex, or genetic subgroup (as defined in Schmitz et al[27]). Overall survival (OS) was significantly inferior at one year in the *PD-L1* amplified group compared to the *PD-L1* non-amplified group (55.7% vs 83.7%, *P* = 0.04, log-rank test). No significant differences were detected in one-year progression-free survival (PFS) or two-year OS and PFS between the two groups.

PD-L1 amplifications and somatic mutations

No significant differences were detected in driver mutation frequency between *PD-L1* amplified and non-amplified cases after correction for multiple testing. No significant differences were detected in total tumor mutational burden (TMB) between the *PD-L1* amplified and non-amplified groups (supplemental Figure 10, *P* = 0.34, Wilcoxon rank-sum test).

PD-L1 amplifications and gene expression patterns

The *PD-L1* amplified group demonstrated significantly higher gene expression of *PD-L1*, *CD8A*, and *CD4* relative to the non-amplified group (supplemental Figure 8). 488 significantly differentially expressed genes (DEGs) were identified comparing the two groups at FDR-corrected p-value < 0.05, and fold change ≥ 1.5 or ≤ -1.5 (supplemental Figure 9A). Gene ontology analysis of the *PD-L1* amplified group's differentially expressed genes found the highest overlap in neutrophil-related pathways, followed by phagocytosis and T-cell activation (supplemental Figure 9B). IPA upstream regulator causal network analysis identified significant activation in the transcriptional programs of NF-κB (Z-score 2.93, P = 6.37e-6), TNF (Z-score 4.16, P = 3.06e-13) and IFNG (Z-score = 5.42, P = 4.35e-30) in the *PD-L1* amplified group (supplemental Figure 9C). GSEA found significant activation in the Hallmark TNF signaling via NF-κB pathway, as was seen in our primary tumor analysis (Figure 3E; supplemental Figure 9D).

To investigate the influence of activated B-cell (ABC) enrichment in the *PD-L1* amplified subgroup, we repeated the same analysis comparing *PD-L1* amplified (n = 21) to ABC-only non-amplified cases (n = 219). 1705 significant DEGs were identified using the same thresholds described above. IPA upstream regulator causal network analysis identified similar results as that of all non-amplified samples. For example, activation in NF-κB complex (Z-score = 5.267, P = 3.2e-10), TNF (Z-score = 6.325, P = 5.15e-32) and IFNG (Z-score = 6.752, P = 2.54e-41) pathways was observed within the *PD-L1* amplified group relative to ABC-only non-amplified cases. GSEA analysis also found significant activation in the Hallmark TNF signaling via NF-κB pathway within *PD-L1* amplified samples *(*P = 0.019). Collectively these results suggest that the differential gene expression seen between *PD-L1* amplified and non-amplified cases are independent of gene expression subgroup enrichment.

SUPPLEMENTAL TABLES

Supplemental Table 1. Treatment status of samples used for study experiments.

Subject ID	*PD-L1* Status	Disease Status	WES	RNA Seq	TCR Seq	CD4/CD8 IHC	HLA I/II IHC	PD-L1 IHC
1	Polysomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
2	Polysomic	Relapsed	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
3	Copy Gain	Treatment Naïve	Completed	Not Performed	Not Performed	Not Performed	Not Peformed	Completed
4	Copy Gain	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
5	Copy Gain	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Not Peformed	Completed
6	Copy Gain	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Completed
7	Copy Gain	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Completed
8	Copy Gain	Treatment Naïve	Completed	Completed	Not Performed	Completed	Not Peformed	Completed
9	Copy Gain	Treatment Naïve	Not Performed	Completed	Not Performed	Completed	Completed	Completed
10	Copy Gain	Treatment Naïve	Completed	Completed	Completed	Completed	Completed	Completed
11	Copy Gain	Relapsed	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
12	Copy Gain	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
13	Copy Gain	Treatment Naïve	Not Performed	Completed	Not Performed	Completed	Completed	Completed
14	Copy Gain	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Not Performed
15	Copy Gain	Treatment Naïve	Completed	Completed	Completed	Completed	Completed	Not Performed
16	Copy Gain	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
17	Copy Gain	Relapsed	Not Performed	Not Performed	Not Performed	Not Performed	Completed	Completed
18	Copy Gain	Treatment Naïve	Completed	Completed	Not Performed	Not Performed	Not Peformed	Completed
19	Copy Gain	Relapsed	Not Performed	Not Performed	Completed	Completed	Completed	Completed
20	Amplified	Treatment Naïve	Completed	Completed	Completed	Completed	Not Peformed	Completed
21	Amplified	Treatment Naïve	Completed	Not Performed	Completed	Completed	Completed	Completed
22	Amplified	Treatment Naïve	Completed	Completed	Completed	Completed	Completed	Completed
23	Amplified	Relapsed	Completed	Completed	Completed	Completed	Completed	Completed
24	Amplified	Treatment Naïve	Completed	Completed	Completed	Completed	Completed	Completed
25	Amplified	Treatment Naïve	Completed	Completed	Completed	Completed	Completed	Completed
26	Amplified	Treatment Naïve	Completed	Not Performed	Not Performed	Completed	Completed	Completed
27	Translocated	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Completed
28	Translocated	Treatment Naïve	Not Performed	Completed	Completed	Completed	Completed	Completed
29	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
30	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Completed
31	Disomic	Relapsed	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
32	Disomic	Relapsed	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
33	Disomic	Treatment Naïve	Completed	Completed	Not Performed	Not Performed	Not Peformed	Completed
34	Disomic	Relapsed	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
35	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
36	Disomic	Treatment Naïve	Not Performed	Completed	Not Performed	Completed	Completed	Completed
37	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
38	Disomic	Treatment Naïve	Completed	Completed	Completed	Completed	Completed	Completed
39	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
40	Disomic	Relapsed	Completed	Completed	Completed	Completed	Completed	Completed
41	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
42	Disomic	Treatment Naïve	Completed	Completed	Completed	Completed	Completed	Completed
43	Disomic	Relapsed	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
44	Disomic	Treatment Naïve	Not Performed	Completed	Completed	Completed	Completed	Completed
45	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Not Performed
46	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Not Performed
47	Disomic	Treatment Naïve	Completed	Completed	Completed	Not Performed	Not Peformed	Completed
48	Disomic	Relapsed	Not Performed	Not Performed	Not Performed	Not Performed	Completed	Completed
49	Disomic	Treatment Naïve	Completed	Completed	Completed	Completed	Completed	Completed
50	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Not Performed
51	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
52	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
53	Disomic	Relapsed	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed

Subject ID	*PD-L1* Status	Disease Status	WES	RNA Seq	TCR Seq	CD4/CD8 IHC	HLA I/II IHC	PD-L1 IHC
54	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
55	Disomic	Treatment Naïve	Completed	Not Performed	Not Performed	Completed	Completed	Completed
56	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
57	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Not Performed
58	Disomic	Relapsed	Not Performed	Not Performed	Not Performed	Not Peformed	Completed	Completed
59	Disomic	Treatment Naïve	Completed	Completed	Completed	Completed	Completed	Completed
60	Disomic	Treatment Naïve	Completed	Completed	Completed	Not Performed	Completed	Completed
61	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
62	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
63	Disomic	Treatment Naïve	Not Performed	Completed	Completed	Completed	Completed	Completed
64	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Completed	Completed
65	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
66	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Completed	Completed
67	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
68	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
69	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Completed
70	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
71	Disomic	Relapsed	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
72	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Not Performed
73	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Completed
74	Disomic	Treatment Naïve	Not Performed	Completed	Completed	Completed	Completed	Completed
75	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
76	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
77	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
78	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
79	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Completed	Completed
80	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
81	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
82	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Completed
83	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Not Performed
84	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Not Performed
85	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Not Performed
86	Disomic	Relapsed	Not Performed	Not Performed	Not Performed	Not Performed	Completed	Completed
87	Disomic	Relapsed	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Not Performed
88	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
89	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
90	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Completed
91	Disomic	Relapsed	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Completed
92	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Completed
93	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Completed
94	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Completed	Completed
95	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
96	Disomic	Relapsed	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
97	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Completed
98	Disomic	Relapsed	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Completed
99	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Not Performed
100	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Performed	Not Peformed	Completed
101	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
102	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
103	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
104	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Completed	Completed	Completed
105	Disomic	Treatment Naïve	Not Performed	Not Performed	Not Performed	Not Peformed	Not Peformed	Completed

Supplemental Table 2. Differentially expressed genes between *PD-L1* altered and *PD-L1* not altered DLBCLs.

Gene	EnsemblGeneID	AveExpr	P.Value	PDL1alt.vs.PDL1notalt .adj.P.Val	PDL1alt.vs.PDL1notalt .FoldChange
CD274	ENSG00000120217	6.196242096	1.11E-06	0.008362921	8.586016997
NUP88	ENSG00000108559	5.800975158	1.38E-06	0.008362921	-1.866418188
SULT1C2	ENSG00000198203	4.972159193	1.78E-06	0.008362921	5.279783901
ACADM	ENSG00000117054	6.154588374	2.29E-06	0.008362921	-2.087329416
KNTC1	ENSG00000184445	7.885726221	2.91E-06	0.008498304	-1.956945716
AL132780.3	ENSG00000259132	-0.999823589	4.29E-06	0.010425213	12.49949752
B4GALNT1	ENSG00000135454	1.271986554	7.16E-06	0.012726308	4.18626551
GPD1L	ENSG00000152642	5.082129702	7.45E-06	0.012726308	-2.00978417
RAB13	ENSG00000143545	4.615473655	8.52E-06	0.012726308	2.575491214
KCNE3	ENSG00000175538	4.008157593	9.20E-06	0.012726308	2.048818103
RSBN1L	ENSG00000187257	6.748677893	9.89E-06	0.012726308	-1.648332486
EAF2	ENSG00000145088	4.498766756	1.05E-05	0.012726308	-2.971023001
ACO1	ENSG00000122729	5.61232566	1.24E-05	0.013887372	2.160340948
SLC9A7	ENSG00000065923	7.173800127	1.42E-05	0.013887372	-2.382487276
PDCD1LG2	ENSG00000197646	4.560056314	1.43E-05	0.013887372	2.407715601
NPFFR1	ENSG00000148734	4.4661185	1.80E-05	0.016442539	2.7487051
GAB1	ENSG00000109458	5.91078463	2.08E-05	0.016984647	-3.358209712
MAFG	ENSG00000197063	5.241187539	2.10E-05	0.016984647	1.644195333
STAP1	ENSG00000035720	4.017496811	2.38E-05	0.017701909	-4.858961987
MYO1G	ENSG00000136286	6.944730834	2.47E-05	0.017701909	2.164171724
HK3	ENSG00000160883	4.359156951	2.86E-05	0.018322721	4.557029933
PELI1	ENSG00000197329	6.49123533	3.05E-05	0.018322721	-2.198967414
RASGRP1	ENSG00000172575	7.240649249	3.08E-05	0.018322721	2.455894165
SLC2A6	ENSG00000160326	5.029002459	3.14E-05	0.018322721	2.495885487
TNFAIP2	ENSG00000185215	7.681332892	3.97E-05	0.021592901	2.633505325
LRP12	ENSG00000147650	4.148595801	4.00E-05	0.021592901	2.1456104
PLD1	ENSG00000075651	5.634969991	4.17E-05	0.021700504	1.984540682
METTL16	ENSG00000127804	5.96533389	4.60E-05	0.022780928	-1.58093691
FAS	ENSG00000026103	6.185125007	4.69E-05	0.022780928	2.149395287
FUT4	ENSG00000196371	5.008903861	4.85E-05	0.022797819	1.759305019
ACVR1B	ENSG00000135503	4.536680487	5.09E-05	0.022995759	2.093828179
MICALL2	ENSG00000164877	4.740938302	5.38E-05	0.022995759	2.714255487
HIVEP2	ENSG00000010818	7.367013588	5.40E-05	0.022995759	-1.698726434
ZNF860	ENSG00000197385	4.245318884	5.68E-05	0.022995759	-5.478437077
TNIP2	ENSG00000168884	4.774366543	5.74E-05	0.022995759	1.740585595
FCRL2	ENSG00000132704	6.185756083	5.84E-05	0.022995759	-9.130131036
KLHL14	ENSG00000197705	4.633618713	6.09E-05	0.022995759	-7.916045513
PYGO1	ENSG00000171016	4.596554393	6.23E-05	0.022995759	-6.851451786
G0S2	ENSG00000123689	2.037118144	6.31E-05	0.022995759	7.788386991
CSF3R	ENSG00000119535	5.215834223	7.04E-05	0.025037845	2.186832942
RUBCNL	ENSG00000102445	6.535755309	7.75E-05	0.026656514	-4.234934584
TMEM189	ENSG00000240849	5.412827006	7.88E-05	0.026656514	1.54014945
ZNRF1	ENSG00000186187	4.874819656	8.17E-05	0.026656514	2.526080519
PLEKHO2	ENSG00000241839	6.343334928	8.26E-05	0.026656514	1.816321757
PTAFR	ENSG00000169403	5.878947799	8.41E-05	0.026656514	2.110705846
TFEC	ENSG00000105967	7.461914119	8.80E-05	0.027291506	2.638147711
SPSB1	ENSG00000171621	3.72080779	9.21E-05	0.027759957	2.187710875
SNX33	ENSG00000173548	4.884239851	9.69E-05	0.027759957	2.068020348
NGLY1	ENSG00000151092	6.941479726	9.71E-05	0.027759957	-1.719400803

Gene	EnsemblGeneID	AveExpr	P.Value	PDL1alt.vs.PDL1notalt .adj.P.Val	PDL1alt.vs.PDL1notalt .FoldChange
SLC38A1	ENSG00000111371	9.131075031	0.000100958	0.028114701	-1.576380381
EPM2A	ENSG00000112425	5.152487031	0.0001042	0.028114701	-1.875921891
SH2D3C	ENSG00000095370	4.998470966	0.000104948	0.028114701	-2.299483239
SOCS3	ENSG00000184557	6.231280671	0.000109865	0.028114701	2.405864185
ARHGAP24	ENSG00000138639	6.172995338	0.000111744	0.028114701	-2.671569531
JUNB	ENSG00000171223	6.021125315	0.000113044	0.028114701	2.146588603
VPS37C	ENSG00000167987	4.174442586	0.000113754	0.028114701	1.627283527
FLNA	ENSG00000196924	9.737022925	0.000128265	0.029806731	1.923538048
PLXNB2	ENSG00000196576	7.220616093	0.000132829	0.029806731	1.932766779
TNFRSF17	ENSG00000048462	1.753473097	0.000133091	0.029806731	-6.947313337
ADGRE2	ENSG00000127507	5.650705462	0.000135179	0.029806731	2.22931729
ENO2	ENSG00000111674	4.806878074	0.000136417	0.029806731	2.277313811
RPA1	ENSG00000132383	6.379019492	0.000137089	0.029806731	-1.80211008
COL27A1	ENSG00000196739	5.15659974	0.000140177	0.029806731	2.4563606
OSBPL10	ENSG00000144645	6.497908278	0.000141013	0.029806731	-3.656707044
FAM69A	ENSG00000154511	4.249996383	0.000141041	0.029806731	-2.004323052
XKR6	ENSG00000171044	4.37399075	0.000145767	0.029974673	-2.245951504
PRMT7	ENSG00000132600	5.388294546	0.000147326	0.029974673	-1.530039552
SOCS1	ENSG00000185338	3.3410772	0.000148096	0.029974673	2.3549395
SRC	ENSG00000197122	4.987911061	0.000153092	0.029974673	1.988101833
DOCK5	ENSG00000147459	5.562024594	0.000153245	0.029974673	1.698873732
NUBPL	ENSG00000151413	5.164052652	0.00015417	0.029974673	-1.710604005
RTKN2	ENSG00000182010	5.242596782	0.000164198	0.031173576	-2.53101977
FCRL1	ENSG00000163534	5.282784114	0.000166497	0.031173576	-6.049482095
PPP1R18	ENSG00000146112	6.970800565	0.000166749	0.031173576	1.55909261
SUSD6	ENSG00000100647	7.143771107	0.000172536	0.031586636	1.640509131
BCAS4	ENSG00000124243	6.388014513	0.000173291	0.031586636	-2.416664687
TP53	ENSG00000141510	5.993229226	0.0001784	0.031834295	-1.688782145
CDCA7L	ENSG00000164649	6.427103281	0.000184388	0.031834295	-2.628307838
PPT2	ENSG00000221988	2.630075721	0.000184627	0.031834295	1.873377902
ARHGAP17	ENSG00000140750	7.384752049	0.000186732	0.031834295	-2.061197976
PRSS16	ENSG00000112812	1.087706152	0.000191001	0.031834295	-4.441661047
CYP27B1	ENSG00000111012	3.310196621	0.00019654	0.031834295	5.364356642
PRIM1	ENSG00000198056	4.243824722	0.000201986	0.031834295	-1.796701831
CYP26B1	ENSG00000003137	3.262766712	0.000202959	0.031834295	3.290794546
TMEM131L	ENSG00000121210	7.796621967	0.000204802	0.031834295	-1.871993524
DOCK4	ENSG00000128512	6.157145916	0.000205357	0.031834295	2.089392624
ETV5	ENSG00000244405	4.808663603	0.000205387	0.031834295	1.753862789
FCRLA	ENSG00000132185	5.866675911	0.000208209	0.031834295	-6.292200327
TXLNB	ENSG00000164440	4.068260317	0.000209436	0.031834295	-2.665568102
C11orf80	ENSG00000173715	4.308121357	0.000210342	0.031834295	-1.924818959
BRI3BP	ENSG00000184992	7.416630126	0.000216415	0.031834295	-1.987470366
TOP2B	ENSG00000077097	8.129077965	0.000216932	0.031834295	-1.680722385
IKZF2	ENSG00000030419	5.992219303	0.000217483	0.031834295	-2.616092524
LANCL1	ENSG00000115365	6.172995078	0.000218312	0.031834295	-1.537974209
ADAMTS14	ENSG00000138316	3.717792066	0.000229865	0.032646526	3.081413347
ZBED6CL	ENSG00000188707	4.67799075	0.000242603	0.034015753	2.01357447
IL10RB	ENSG00000243646	5.580469324	0.000245067	0.034033999	1.507490901
FCRL5	ENSG00000143297	8.396006944	0.000255106	0.035015437	-6.177235933
AGAP3	ENSG00000133612	4.985912618	0.000256937	0.035015437	1.879760005

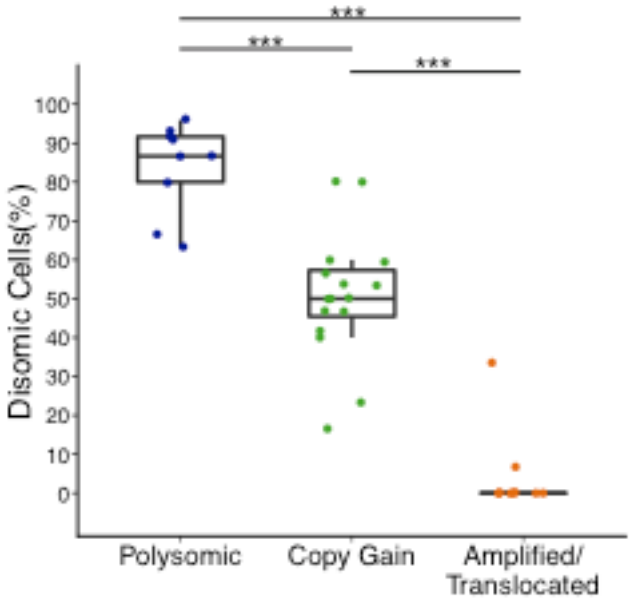
Gene	EnsemblGeneID	AveExpr	P.Value	PDL1alt.vs.PDL1notalt .adj.P.Val	PDL1alt.vs.PDL1notalt .FoldChange
INSIG2	ENSG00000125629	5.396106032	0.000263269	0.035220033	1.603433695
RFXAP	ENSG00000133111	4.200812851	0.000279874	0.037101116	-1.526179823
PFKFB3	ENSG00000170525	7.246433617	0.000295003	0.038754329	2.267788839
SPRY4	ENSG00000187678	3.848356053	0.000307968	0.039813781	2.262169566
FAM214B	ENSG00000005238	3.801462399	0.000313291	0.039813781	1.721327377
FLVCR2	ENSG00000119686	4.519951815	0.000318446	0.039813781	2.230507277
LRRC61	ENSG00000127399	3.42687456	0.000318567	0.039813781	2.189354714
HLA-DOB	ENSG00000241106	4.085408879	0.000318959	0.039813781	-2.541888391
GK	ENSG00000198814	6.326270947	0.000319449	0.039813781	2.08119694
HACD1	ENSG00000165996	2.642782314	0.000345766	0.042312119	-3.26392262
BLK	ENSG00000136573	6.098317942	0.000349809	0.042312119	-3.205553174
ALDH2	ENSG00000111275	6.429450044	0.000350628	0.042312119	2.574744984
TYMP	ENSG00000025708	6.739764267	0.000352467	0.042312119	2.185000975
SMAD4	ENSG00000141646	7.362415523	0.000356505	0.042312119	-1.601441225
TPRG1	ENSG00000188001	2.895823504	0.000359799	0.042312119	2.253827062
RFX2	ENSG00000087903	3.998613606	0.000368714	0.042312119	1.813821617
POU2F1	ENSG00000143190	7.454622669	0.000369647	0.042312119	-1.749402604
LRMP	ENSG00000118308	7.912642236	0.000371083	0.042312119	-2.686530635
ST18	ENSG00000147488	2.807463511	0.000371629	0.042312119	1.888400691
SPRED1	ENSG00000166068	6.976181004	0.000372209	0.042312119	1.590174152
MIPEP	ENSG00000027001	3.591098672	0.000374315	0.042312119	-1.502666564
SDC4	ENSG00000124145	4.720865197	0.000404107	0.04493853	2.6187
SLC16A3	ENSG00000141526	5.47175444	0.000406094	0.04493853	2.243502763
KLHL15	ENSG00000174010	4.970660361	0.000410799	0.04493853	-1.841870925
BCL2	ENSG00000171791	8.17861779	0.000413131	0.04493853	-3.248239818
KIAA0040	ENSG00000235750	6.70990791	0.000414363	0.04493853	-2.279944
CARMIL3	ENSG00000186648	3.110480163	0.00041604	0.04493853	4.160769055
FNDC3B	ENSG00000075420	7.094888833	0.000437544	0.04655589	1.719607513
GUCY1B3	ENSG00000061918	3.95662356	0.000438335	0.04655589	1.644890984
ABCC3	ENSG00000108846	4.890281953	0.000440592	0.04655589	1.950585516
CSF2RA	ENSG00000198223	4.816916285	0.000452477	0.047128672	1.612458259
IL10	ENSG00000136634	3.935129099	0.000466143	0.047868242	3.356417968
OXSM	ENSG00000151093	2.64637408	0.000469712	0.047897469	-1.61699319
RRAS2	ENSG00000133818	5.322479614	0.000478732	0.048170324	-2.455940526
IL17RC	ENSG00000163702	1.935315877	0.000478994	0.048170324	2.005844947
NECTIN2	ENSG00000130202	4.420067723	0.000489527	0.048892364	2.013836424
ARRB1	ENSG00000137486	5.080038922	0.000498103	0.049410458	1.846157832
ZNF519	ENSG00000175322	6.354557918	0.000506336	0.049887808	-1.691761599

Supplemental Table 3. Baseline characteristics of *PD-L1* amplified DLBCLs in an external dataset.

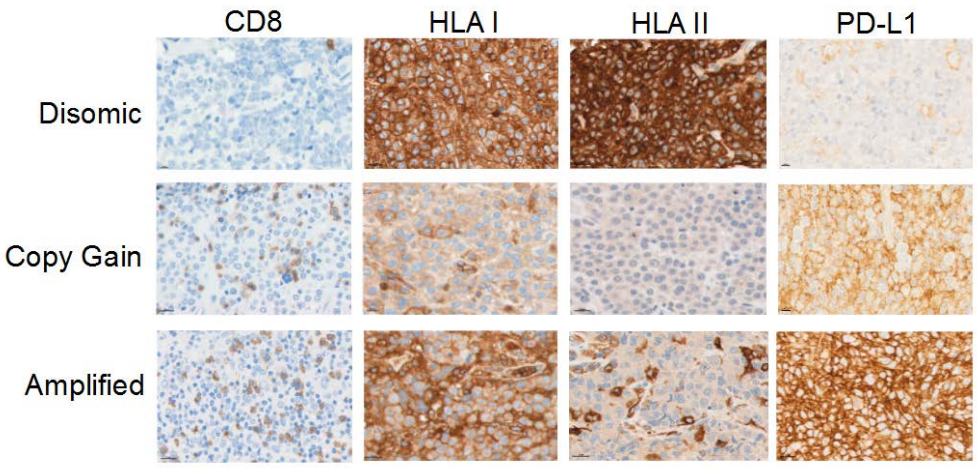
	***PD-L1* amplified (*n*=21)**	***PD-L1* non-amplified (*n*=409)**	***P*-value**
Mean age (SD)	58.8 (16.2)	61.2 (14.6)	NS
Female sex	7 (33%)	167 (41%)	NS
Cell of origin			
ABC	13 (62%)	202 (49%)	0.03
GCB	1 (5%)	119 (29%)	
IPI score			
High	2 (10%)	63 (15%)	0.03
Intermediate	14 (67%)	152 (37%)	
Low	1 (5%)	89 (22%)	
Stage			
Stage 1	0	47 (11%)	NS
Stage 2	7 (33%)	100 (24%)	
Stage 3	3 (14%)	102 (25%)	
Stage 4	8 (38%)	112 (27%)	
LDH elevated	13 (62%)	163 (40%)	0.02
Genetic subtype			
BN2	4 (19%)	74 (18%)	NS
EZB	1 (5%)	49 (12%)	

MCD	0	48 (12%)
N1	2 (10%)	13 (3%)

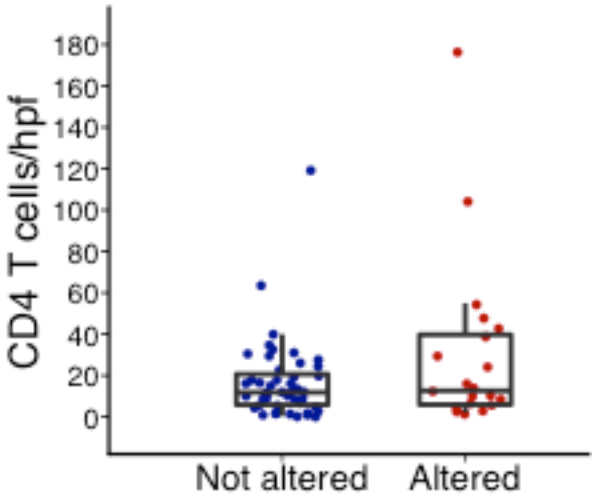
SUPPLEMENTAL FIGURES



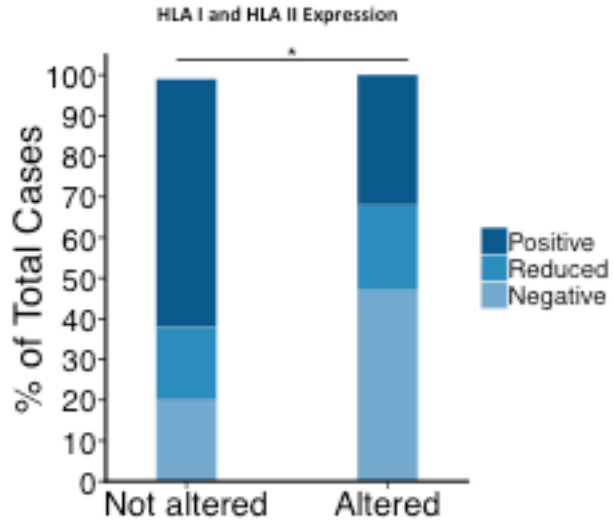
Supplemental Figure 1. Residual *PD-L1* disomic lymphoma cells in DLBCL specimens according to *PD-L1* locus status. Percentage of residual *PD-L1* disomic lymphoma cells in DLBCLs with the indicated mechanism of *PD-L1* gene alteration as assessed by FISH is shown (***$P < 0.001$, Mann-Whitney *U* test).



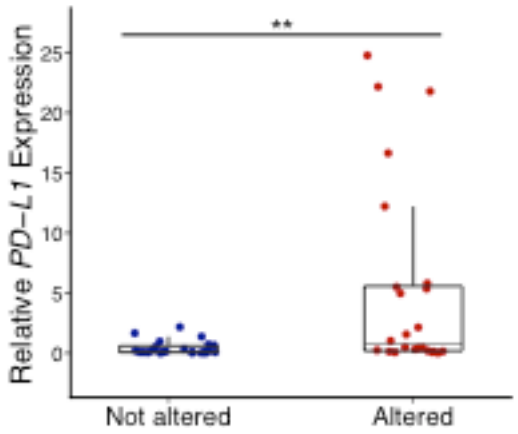
Supplemental Figure 2. Representative IHC images for PD-L1, CD8, and HLA class I and II in *PD-L1* gene-altered and *PD-L1* not altered cases. IHC images depicting $CD8^+$ T cell infiltration and expression of HLA I/II and PD-L1 in representative DLBCLs with *PD-L1* disomy, copy gain or amplification. The *PD-L1* disomic DLBCL specimen (top row) was largely devoid of $CD8^+$ T cells, exhibited normal HLA class I and II staining, and was PD-L1 negative. The DLBCL case harboring a relative *PD-L1* copy gain (middle row) contained numerous $CD8^+$ T cells, was negative for class I and II HLA staining (note positive HLA class I staining of infiltrating non-malignant immune cells), and displayed moderately-intense PD-L1 expression. Finally, the *PD-L1* gene-amplified DLBCL specimen (bottom row) was robustly infiltrated by $CD8^+$ T cells, had reduced HLA class I expression compared to surrounding non-malignant cells, was negative for HLA class II, and exhibited very intense and diffuse PD-L1 expression.



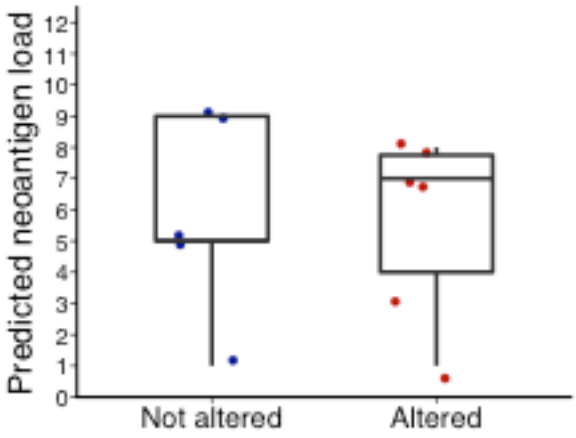
Supplemental Figure 3. CD4$^+$ T cell infiltration in *PD-L1* altered versus not altered DLBCLs. Numbers of CD4$^+$ T cells/hpf in *PD-L1* altered and *PD-L1* not altered DLBCLs as assessed by IHC (P = 0.30, Mann-Whitney U test).



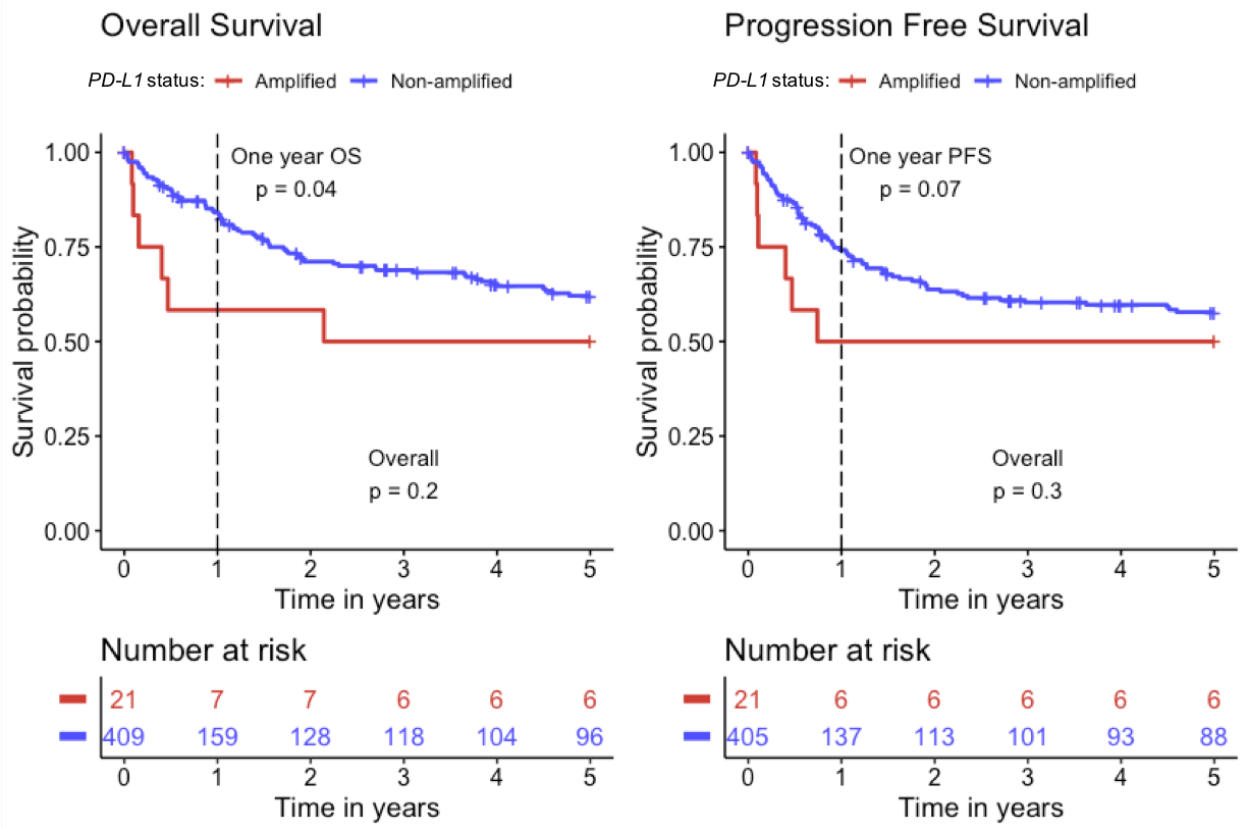
Supplemental Figure 4. Combined HLA class I and II cell surface expression according to *PD-L1* locus status. Percentage of DLBCLs with normal, reduced, or absent HLA class I and/or II expression as assessed by IHC according to *PD-L1* locus status (*P = 0.028, Fisher's exact test).



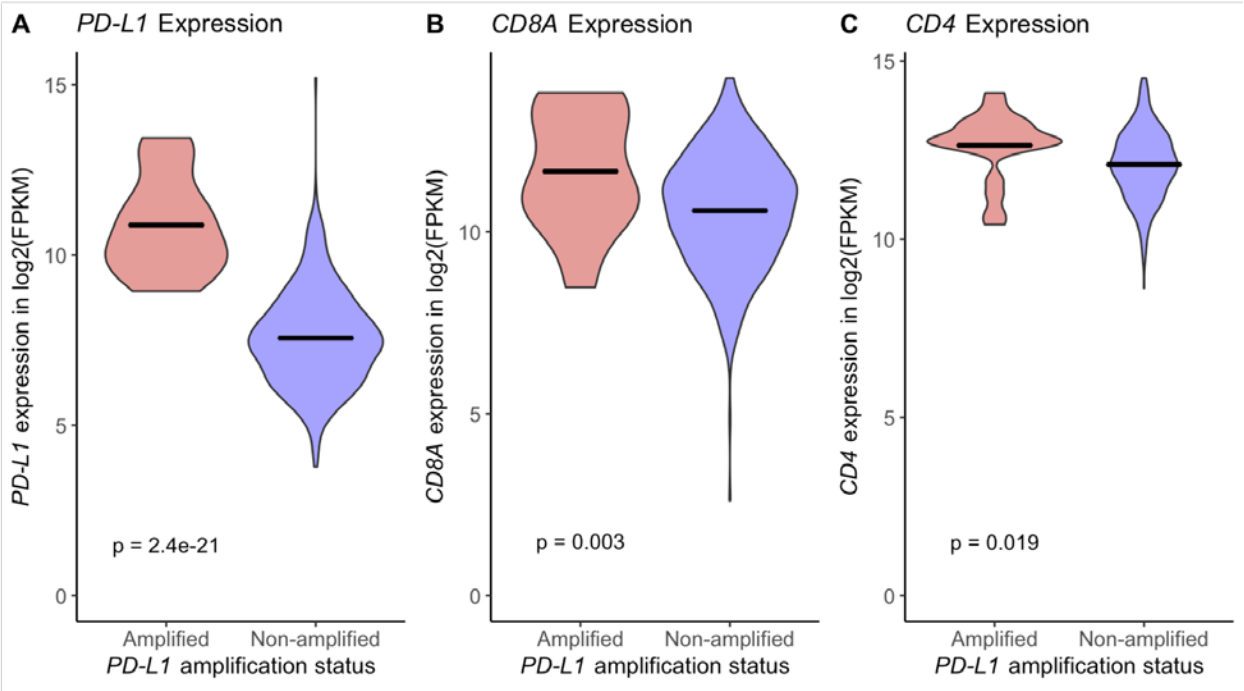
Supplemental Figure 5. *PD-L1* mRNA expression according to *PD-L1* locus status. Relative *PD-L1* mRNA expression compared to a housekeeping gene (*ACTB)* was assessed by quantitative PCR in *PD-L1* gene-altered and not altered DLBCLs (n=47). Experiments were performed in triplicate with the average being depicted (****P* = 0.004, Mann-Whitney *U* test).



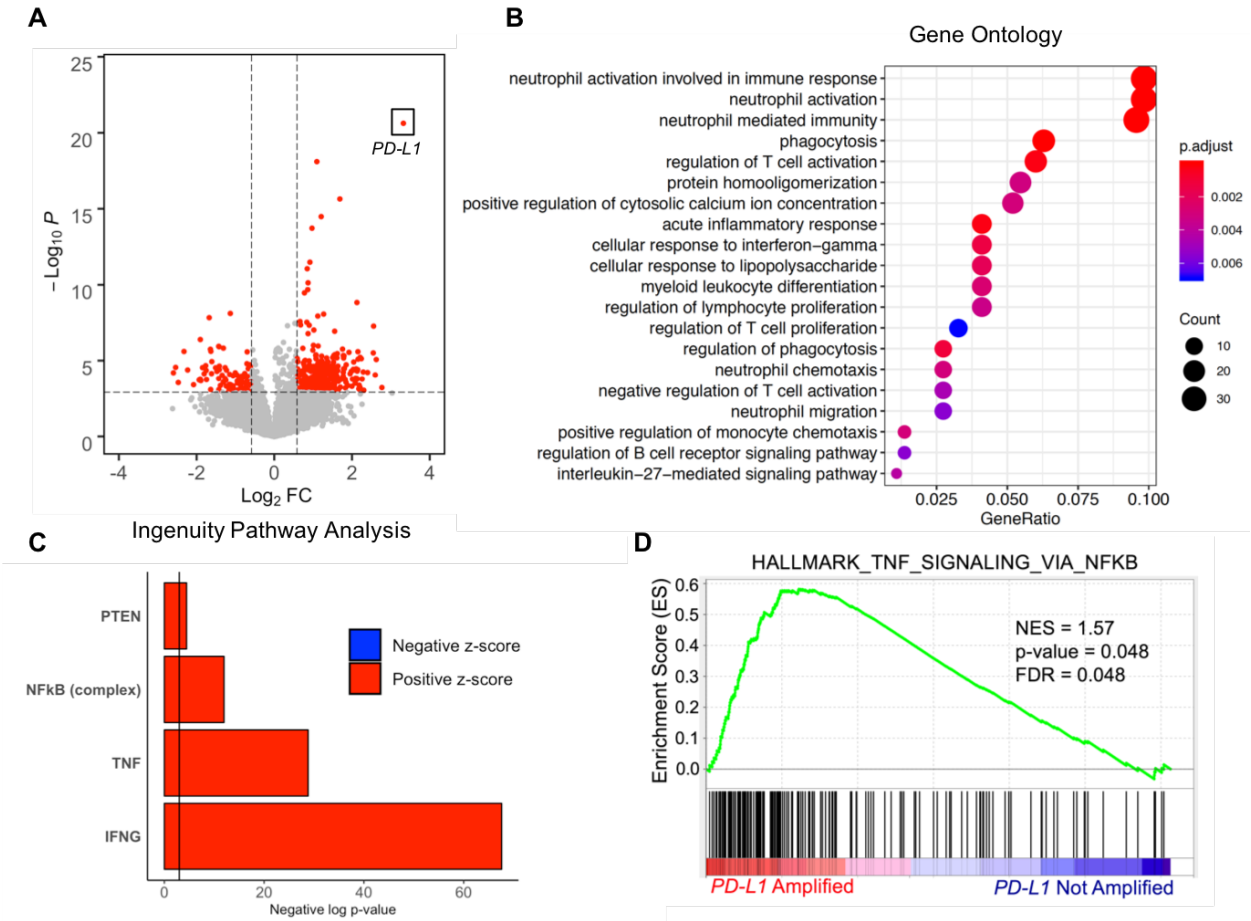
Supplemental Figure 6. Predicted HLA-A neo-antigen burden in *PD-L1* gene-altered versus not altered DLBCL (*P* = 0. 853, Mann-Whitney *U* test).



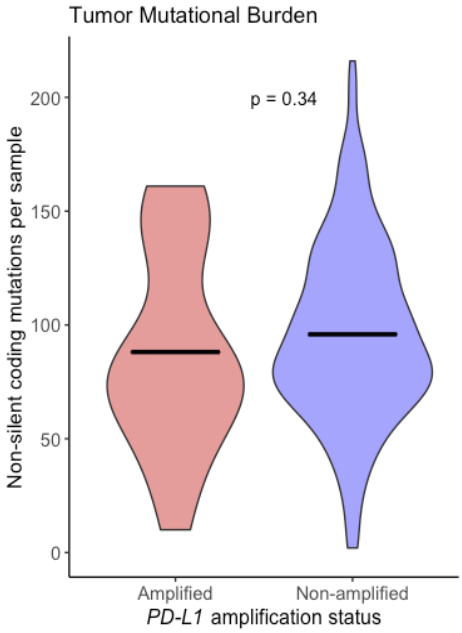
Supplemental Figure 7. One-year OS is significantly worse in the *PD-L1* amplified group, with non-significant observed differences in PFS in a validation dataset. Overall survival and progression free survival were not significantly different between the *PD-L1* amplified and non-amplified groups over the entire course of follow-up (OS *P* = 0.2, PFS, *P* = 0.3 with log-rank test). *P*-value from landmark survival analysis at one year using the Kaplan-Meier estimator is shown for OS *(P* = 0.04, log-rank test) and PFS (*P* = 0.07, log-rank test).



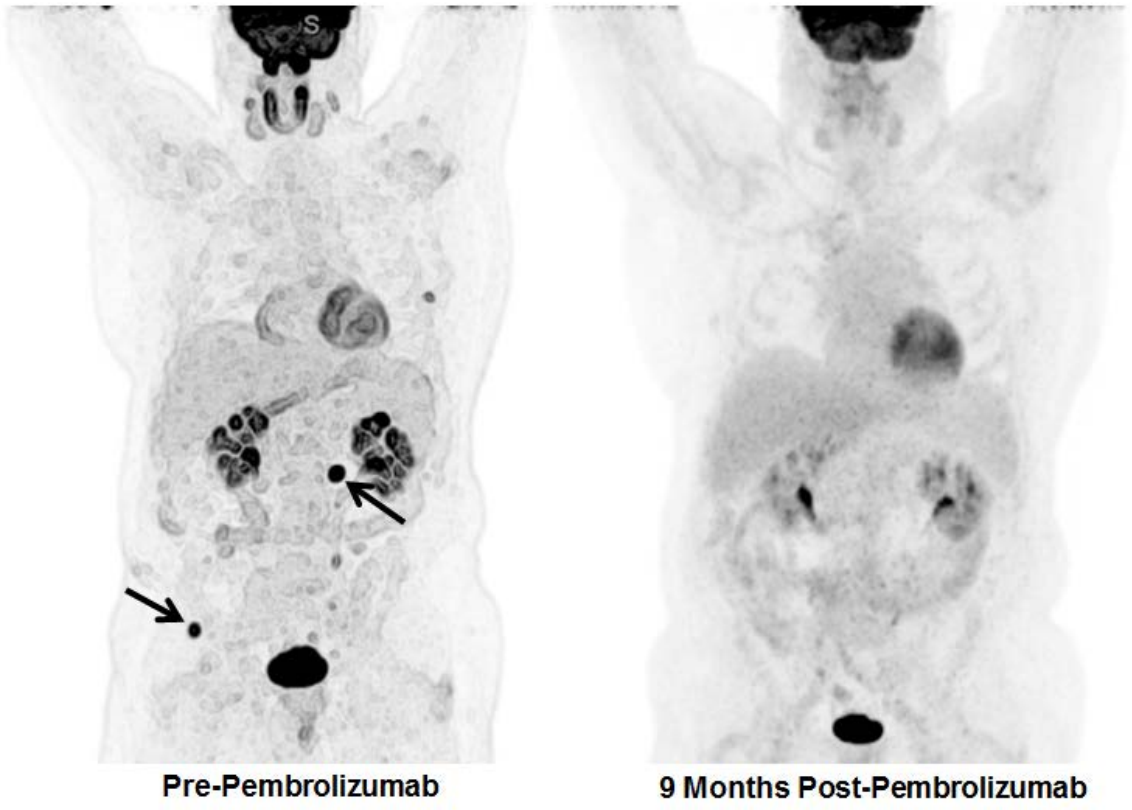
Supplemental Figure 8. *PD-L1* gene amplifications are associated with increased expression of PD-L1, CD8A and CD4 in a validation dataset. Gene expression of *PD-L1*, *CD8A*, and *CD4* was compared between *PD-L1* amplified (n = 21) and non-amplified groups (n = 409). Violin plots of (**A**) *PD-L1*, (**B**) *CD8A*, (**C**) *CD4* gene expression in *PD-L1* amplified versus non-amplified DLBCLs. The mean is marked with a horizontal bar, and width of the violin plot corresponds to the number of samples with that y-value. Significance testing was performed in the entire gene expression cohort with limma and Ebayes as described above, reported here for the individual genes.



Supplemental Figure 9. RNA sequencing identifies differentially-expressed genes in *PD-L1* amplified relative to *PD-L1* non-amplified DLBCLs in a validation dataset. (**A**) Volcano plot summarizing 488 differentially expressed genes (DEG) in DLBCLs with vs without *PD-L1* gene amplifications (FDR-adjusted $P < 0.05$, fold change ≥1.5 or ≤ -1.5) in 430 previously-published RNAseq profiles.[27] (**B**) GO analysis performed on the 488 DEGs. GO terms significantly enriched in DLBCLs with *PD-L1* gene amplifications are denoted by red dots. (**C**) Ingenuity Pathway Analysis revealing predicted upstream regulators of gene expression activated (positive z-score) or inhibited (negative z-score) in DLBCLs with *PD-L1* amplifications relative to those without. Vertical line indicates the position of $P = 0.05$ on the x-axis. (**D**) GSEA demonstrating enrichment of NF-κB-regulated genes in *PD-L1* amplified compared to *PD-L1* non-amplified DLBCLs.



Supplemental Figure 10. Tumor mutational burden in *PD-L1* gene-amplified and non-amplified DLBCLs in a validation dataset. Total tumor mutational burden (TMB), defined as non-silent protein-changing somatic mutations per sample, was compared between *PD-L1* amplified (n = 21) and non-amplified samples (n = 409) (P = 0.34, Wilcoxon rank sum test). The mean is marked with a horizontal bar, and the width of the violin plot corresponds to the number of samples with a given y-value.



Supplemental Figure 11. Coronal PET images before and after pembrolizumab therapy for patient included in Figure 5C.

SUPPLEMENTAL REFERENCES

1. Andrews S. FastQC: A quality control application for high throughput sequence data. *Babraham Institute Project page: http://www. bioinformatics. bbsrc. ac. uk/projects/fastqc*. 2012;
2. Derrien T, Johnson R, Bussotti G, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 2012;22(9):1775–1789.
3. Bray NL, Pimentel H, Melsted P, Pachter L. Erratum: Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* 2016;34(8):888.
4. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res.* 2015;4:1521.
5. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 2014;15(2):R29.
6. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 2000;25(1):25–29.
7. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference

resource for gene and protein annotation. *Nucleic Acids Res.* 2016;44(D1):D457–62.
8. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS: A Journal of Integrative Biology*. 2012;16(5):284–287.
9. Kramer A, Green J, Pollard J Jr, Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics*. 2014;30(4):523–530.
10. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 2005;102(43):15545–15550.
11. Liberzon A, Subramanian A, Pinchback R, et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011;27(12):1739–1740.
12. Mootha VK, Lindgren CM, Eriksson K-F, et al. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* 2003;34:267.
13. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–2120.
14. Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*. 2011;27(21):2957–2963.
15. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv [q-bio.GN]*. 2013;
16. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics*. 2015;31(12):2032–2034.
17. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 2011;43(5):491–498.
18. Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics*. 2013;43:11.10.1–33.
19. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* 2013;31(3):213–219.
20. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536(7616):285–291.
21. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38(16):e164–e164.
22. Jurtz V, Paul S, Andreatta M, et al. NetMHCpan-4.0: Improved Peptide–MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. *The Journal of Immunology*. 2017;ji1700893.
23. Szolek A, Schubert B, Mohr C, et al. OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics*. 2014;30(23):3310–3316.
24. Use of Affymetrix Mapping Arrays in the Diagnosis of Gene Copy Number Variation. *Current Protocols in Human Genetics*. 2001;79:500.
25. GDC MAF Format v1.0.0. *Genomic Data Commons Documentation*. .
26. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12:323.
27. Schmitz R, Wright GW, Huang DW, et al. Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. *N. Engl. J. Med.* 2018;378(15):1396–1407.
28. Jensen MA, Ferretti V, Grossman RL, Staudt LM. The NCI Genomic Data Commons as an

engine for precision medicine. *Blood*. 2017;130(4):453.
29. Seshan VE, Olshen A. DNAcopy: DNA copy number data analysis. R package version 1.56.0. 2018.
30. Mermel CH, Schumacher SE, Hill B, et al. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* 2011;12(4):R41.
31. Team RC, Others. R: A language and environment for statistical computing. 2018;
32. Therneau TM, Lumley T. Package "survival." *R Top Doc*. 2015;128.:
33. Xie J, Liu C. Adjusted Kaplan-Meier estimator and log-rank test with inverse probability of treatment weighting for survival data. *Stat. Med.* 2005;24(20):3089–3110.
34. Ramos AH, Lichtenstein L, Gupta M, et al. Oncotator: cancer variant annotation tool. *Hum. Mutat.* 2015;36(4):E2423–9.
35. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013;499(7457):214–218.
36. Glickman ME, Rao SR, Schultz MR. False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. *J. Clin. Epidemiol.* 2014;67(8):850–857.
37. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Series B Stat. Methodol.* 1995;57(1):289–300.
38. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47.
39. Georgiou K, Chen L, Berglund M, et al. Genetic basis of PD-L1 overexpression in diffuse large B-cell lymphomas. *Blood*. 2016;127(24):3026–3034.
40. Roemer MG, Advani RH, Ligon AH, et al. PD-L1 and PD-L2 Genetic Alterations Define Classical Hodgkin Lymphoma and Predict Outcome. *J. Clin. Oncol.* 2016;34(23):2690–2697.