Molecular profiling reveals immunogenic cues in anaplastic large cell lymphomas with *DUSP22* rearrangements

Luchtel RA et al.

SUPPLEMENTAL MATERIAL

Table of Contents

SUPPLEMENTAL METHODS

Gene set enrichment analysis (GSEA)

Analyses were performed using GSEA software (Broad Institute) configured to use the "GSEAPreranked" method:

http://software.broadinstitute.org/cancer/software/genepattern/modules/docs/GSEAPreranked/1

Enrichment of the cancer-testis antigen (CTA) gene set was evaluated using CTpedia (http://www.cta.lncc.br/modelo.php) as the gene set in GSEA.

Reverse-phase protein array (RPPA) analysis

Proteins were extracted from frozen tissue samples as previously described[1] and RPPA analysis was performed at the M.D. Anderson Cancer Center RPPA Core Facility using STAT3 and pSTAT3^{Y705}antibodies validated by Western blot as published.[2] Protein intensity data were normalized and standardized as previously described.[1]

RNA sequencing from FFPE tissue

RNA sequencing (RNAseq) was performed from FFPE tissue as previously described.[3] Briefly, RNA was extracted using the AllPrep DNA/RNA FFPE kit (Qiagen). Sequencing libraries were prepared using TruSeq RNA Access kit (Illumina) and analyzed on a HiSeq 4000 sequencer (Illumina). Sequenced reads were aligned to the hg38 reference using the previously published MAP-RSeq pipeline[4] slightly modified to use the STAR aligner.[5] Gene-level read counts based on Ensembl version 78 were transformed into RPKMs and resulting expression data were quantile-normalized to remove batch effects. Normalized expression data for signature genes

derived from the frozen discovery set described above were utilized for clustering with statistical analysis as described above.

Immunohistochemistry

Immunohistochemistry was performed on 4-micron FFPE sections as previously published[6,7] using antibodies to pSTAT3^{Y705} (clone D3A7, Cell Signaling; 1:400), PD-L1 (clone SP263, Ventana; prediluted), PD-1 (clone NAT105, Abcam; 1:300), or HLA-DR (clone LN3, BioLegend; 1:12,800). Details of ALCLs used in immunohistochemistry studies are summarized in supplemental Table 2. Stains were scored in a blinded fashion. For pSTAT3, scoring was based on percentage of tumor cell nuclei staining; sections with no internal positive control staining (e.g., endothelial cells) were excluded. PD-L1 and HLA-DR were scored as percent positive tumor cells. PD-1 was scored as the average number of positive infiltrating non-neoplastic cells per high power field. Statistical differences among groups were assessed using the Wilcoxon test. Photomicrographs were taken using an Olympus DP71 camera, Olympus BX51 microscope, and Olympus cellSens image acquisition software at the original magnifications indicated.

DNA methylation analysis

DNA methylation analysis was performed on the 31-sample frozen ALCL discovery set, a validation set of 71 FFPE ALCL samples, and 6 ALCL cell lines treated with either decitabine or vehicle. DNA was extracted from frozen ALCL samples and cell lines as published[8] and reduced representation bisulfite sequencing (RRBS) was performed with base-pair resolution as previously described.[9] CpG methylation ratios were segmented into 200-bp regions and

differentially methylated regions (DMRs) between genetic subtypes were identified using methylKit software.[10] DMRs with Q-values ≤0.01 and absolute delta methylation (Δβ) differences of ≥10% were considered significant. DMRs were annotated according to various types of genomic regions of interest: promoters (-1500 ≤ TSS ≤ 500; TSS=transcription start site); CpG islands (CpGi); CpG shores (-2000 ≤ CpGi ≤ 2000), gene body (i.e. exons and introns), 3'UTR, 5'UTR, SINE repeat regions and LINE repeat regions. The number of DMRs in each genomic region was tallied and compared across genetic subtypes using chi-square tests configured to assume a null hypothesis of uniform distribution of hypomethylated and hypermethylated DMRs in each region type of interest across different genetic types.

For FFPE ALCL samples, DNA was extracted using the QIAamp DNA FFPE Tissue Kit (Qiagen) following the manufacturer's recommendations. Extracted DNA was bisulfite converted, amplified, fragmented, and hybridized to Infinium MethylationEPIC BeadChip arrays (Illumina) at the University of Minnesota Genomics Center (Minneapolis, MN) following the manufacturer's recommendations. Raw data were processed using minfi software using default configuration parameters for processing Illumina HumanMethylationEPIC array data.[11] β values representing the methylation status at each CpG locus were calculated as previously published.[12] Differentially methylated CpG probes (DMCs) were evaluated using ANOVA tests to compare percent methylation values. DMCs with absolute Δβ ≥ 10% and a corrected *P*-value of ≤0.05 were considered significant.

Cell lines and decitabine treatment

Cell lines were obtained and maintained as previously described[13] in RPMI 1640 media (Invitrogen) supplemented with 10% fetal bovine serum (HyClone; except SR-786, 15%). Cells

were treated for 96 h with 10 μM decitabine or vehicle (phosphate-buffered saline). DNA methylation was measured by RRBS and RNA sequencing as described above. Gene expression response was measured by RNA sequencing as previously described.[14] Genes with <50 reads in all samples were considered inevaluable.

To examine the relationship between decitabine-induced gene expression and gene expression in ALCLs with and without *DUSP22* rearrangements, genes expressed in ≥4 cell lines were ranked for GSEA as described above. DUSP22-associated genes were defined as genes overexpressed in ALCLs with *DUSP22* rearrangements in the frozen discovery set with a $\log_2$ fold change ≥5 and an adjusted *P* value ≤0.05. GSEA then was used to assess enrichment of DUSP22-associated genes among decitabine-induced genes. Genes down-regulated in *DUSP22*-rearranged ALCLs ($\log_2$ fold change ≤-5 and adjusted *P* value ≤0.05) were used as a negative control.

Statistical analysis

All statistical analyses were performed either using JMP Pro 10 (SAS Institute) or in the SPSS or R statistical environment. Survival analyses were conducted using the log-rank test and plotted using the Kaplan-Meier method. Modeling was performed using the Cox proportional hazards methods. Other statistical tests were used as noted. *P*-values ≤0.05 were considered statistically significant except where indicated.

SUPPLEMENTAL TABLES

Supplemental Table 1. Summary of ALCLs in frozen discovery set[6]

Case	Age/Sex	WHO Diagnosis	Genetic Subtype
1	46/M	ALCL, ALK negative	DUSP22
2	40/M	ALCL, ALK negative	Other
3	18/F	ALCL, ALK positive	ALK
4	74/M	ALCL, ALK positive	ALK
5	39/F	ALCL, ALK negative	Other
6	76/M	ALCL, cutaneous	DUSP22
7	65/M	ALCL, cutaneous	DUSP22
8	59/F	ALCL, cutaneous	DUSP22
9	75/M	ALCL, ALK negative	Other
10	51/M	ALCL, ALK negative	DUSP22
11	43/M	ALCL, ALK negative	Other
12	69/F	ALCL, cutaneous	Other
13	50/F	ALCL, ALK negative	Other
14	58/M	ALCL, ALK negative	Other
15	81/M	ALCL, ALK negative	Other
16	18/M	ALCL, ALK positive	ALK
17	77/M	ALCL, ALK negative	Other
18	13/F	ALCL, cutaneous	Other
19	54/F	ALCL, ALK negative	Other
20	50/M	ALCL, ALK positive	ALK
21	61/M	ALCL, cutaneous	Other
22	6/F	ALCL, ALK positive	ALK
23	66/F	ALCL, ALK negative	Other
24	60/M	ALCL, ALK negative	Other
25	14/M	ALCL, ALK positive	ALK
26	29/F	ALCL, ALK positive	ALK
27	16/M	ALCL, ALK positive	ALK
28	48/M	ALCL, ALK negative	Other
29	77/M	ALCL, ALK negative	DUSP22
30	75/M	ALCL, cutaneous	Other
31	68/F	ALCL, ALK negative	DUSP22

Supplemental Table 2. Summary of ALCLs used in FFPE validation experiments*

Experiment	Genetic Subtype	WHO Diagnosis			
		ALCL, ALK positive	ALCL, ALK negative	ALCL, cutaneous	Total
pSTAT3 IHC† (Fig. 2B): n=334, mean age=54 yr, M:F=1.4					
	ALK	98	0	0	98
	DUSP22	0	45	19	64
	Other	0	113	59	172
	Total	98	158	78	334
RNAseq (Fig. 3C-D): n=53, mean age=56 yr, M:F=1.3					
	ALK	15	0	0	15
	DUSP22	0	10	3	13
	Other	0	15	10	25
	Total	15	25	13	53
Methylation arrays (Fig. 4D-F; suppl. Fig. 2D-F): n=63, mean age=54 yr, M:F=1.9					
	ALK	15	0	0	15
	DUSP22	0	12	4	16
	Other	0	20	12	32
	Total	15	32	16	63
PD-L1 IHC (Fig. 6D): n=152, mean age=54 yr, M:F=1.4					
	ALK	44	0	0	44
	DUSP22	0	22	5	27
	Other	0	49	32	81
	Total	44	71	37	152
HLA-DR IHC (Fig. 7D): n=148, mean age=54 yr, M:F=1.5					
	ALK	47	0	0	47
	DUSP22	0	21	6	27
	Other	0	47	27	74
	Total	47	68	33	148

IHC, immunohistochemistry.
*All cases studied in validation experiments were non-overlapping with the 31 cases in the frozen discovery set (supplemental Table 1).
†All IHC studies were performed on whole tissue sections except 32 cases tested for pSTAT3 from a prospective Danish cohort, for which tissue microarrays were analyzed.[15,16]

Supplemental Table 3. Top gene sets negatively associated with ALCLs in Cluster 1

Name	MSigDB Type	NES	*P*	FDR
HALLMARK_IL6_JAK_STAT3_SIGNALING	HALLMARK	-2.463551	0	0
HALLMARK_TNFA_SIGNALING_VIA_NFKB	HALLMARK	-2.4395595	0	0
HALLMARK_EPITHELIAL_MESENCHYMAL_ TRANSITION	HALLMARK	-2.3364592	0	0
HALLMARK_COAGULATION	HALLMARK	-2.3157318	0	0
HALLMARK_INTERFERON_GAMMA_RESPONSE	HALLMARK	-2.287144	0	0
HALLMARK_INFLAMMATORY_RESPONSE	HALLMARK	-2.271289	0	0
HALLMARK_COMPLEMENT	HALLMARK	-2.252647	0	0
REACTOME_PLATELET_ACTIVATION_ SIGNALING_AND_AGGREGATION	REACTOME	-2.2501082	0	0
KEGG_ECM_RECEPTOR_INTERACTION	KEGG	-2.2365189	0	0
REACTOME_INTEGRIN_CELL_SURFACE_ INTERACTIONS	REACTOME	-2.2300467	0	0

NES, normalized enrichment score; *P*, nominal *P*-value; FDR, false discovery rate q-value.

Supplemental Table 4. Top gene sets positively associated with ALCLs in Cluster 1

Name	MSigDB Type	NES	*P*	FDR
HALLMARK_E2F_TARGETS	HALLMARK	3.0004435	0	0
HALLMARK_G2M_CHECKPOINT	HALLMARK	2.611709	0	0
HALLMARK_MYC_TARGETS_V1	HALLMARK	2.5923479	0	0
REACTOME_DNA_REPLICATION	REACTOME	2.3953528	0	0
REACTOME_CELL_CYCLE	REACTOME	2.3702123	0	0
REACTOME_MITOTIC_M_M_G1_PHASES	REACTOME	2.3683693	0	0
REACTOME_PROCESSING_OF_CAPPED_INTRON_CONTAINING_PRE_MRNA	REACTOME	2.3569436	0	0
REACTOME_CELL_CYCLE_MITOTIC	REACTOME	2.3207977	0	1.04E-04
REACTOME_MRNA_PROCESSING	REACTOME	2.3102338	0	9.24E-05
REACTOME_CHROMOSOME_MAINTENANCE	REACTOME	2.3100226	0	8.32E-05

NES, normalized enrichment score; *P*, nominal *P*-value; FDR, false discovery rate q-value.

Supplemental Table 5. Clinical outcomes in patients with systemic ALCL in Cluster 1

Case*	*DUSP22* Rearrangement	Time to Follow-up (months)	Status at Follow-up
1	Yes	92	Dead
2	No	1	Dead
9	No	38	Dead
10	Yes	188	Alive
13	No	1	Dead
29	Yes	95	Alive
31	Yes	73	Alive

*Phenotypic features of these cases have been reported previously.[6] See also supplemental Table 1. All cases were ALK-negative (see main manuscript, Figure 1A).

Supplemental Table 6. Characteristics of 105 systemic ALK-negative ALCLs evaluated for overall survival*

Characteristic	Value
Age (yr)	
mean	59
range	17-89
Sex	
male	70
female	35
IPI	
0-2	42
3-5	19
missing	44
DUSP22 rearrangement	
present	30
absent	75
$pSTAT3^{Y705}$ immunohistochemistry	
positive†	46
negative	59

IPI, international prognostic index.
*****Includes patients from both discovery set (n=15) and validation set (n=90). Cases from the discovery set correspond to cases 1, 2, 5, 10, 11, 13, 14, 15, 17, 19, 23, 24, 28, 29, and 31 in Supplemental Table 1.
†Based on published cutoff of ≥20% nuclear staining.[17]

Supplemental Table 7. Prognostic factors for overall survival in 105 systemic ALK-negative ALCLs

	Univariate analysis			Multivariate analysis		
Variable	***P***	**HR**	**CI95%**	***P***	**HR**	**CI95%**
Age (>60)	*0.0021*	2.70	[1.43, 5.26]	*0.0335*	2.06	[1.06, 4.13]
Sex (female)	0.3297	1.38	[0.72, 2.58]	0.8258	0.93	[0.46, 1.81]
IPI (0-2)	*0.0475*	0.50	[0.26, 0.99]	*0.0076*	0.33	[0.14, 0.74]
DUSP22-R (present)	*<0.0001*	0.18	[0.05, 0.45]	*0.0001*	0.14	[0.04, 0.41]
pSTAT3 (positive*)	0.3968	1.32	[0.69, 2.49]	0.9928	1.00	[0.47, 2.19]

CI95%, 95% confidence interval; *DUSP22*-R, *DUSP22* rearrangement; HR, hazard ratio.
*Based on published cutoff of ≥20% nuclear staining.[17]

Supplemental Table 8. Genes in ALK Signature

Affy Probe ID	Gene Symbol	Fold Change*	FDR†
208212_s_at	*ALK*	669.39	1.25213E-23
209369_at	*ANXA3*	115.395	4.75273E-07
210305_at	*PDE4DIP*	105.582	1.04126E-06
226145_s_at	*FRAS1*	95.9181	2.45539E-08
211372_s_at	*IL1R2*	83.1767	1.68122E-05
230496_at	*AMER2*	70.977	2.78768E-09
1552767_a_at	*HS6ST2*	64.9973	0.000118984
221111_at	*IL26*	51.4451	0.000197398
235465_at	*AMER2*	45.3223	2.46816E-09
219295_s_at	*PCOLCE2*	37.515	2.30121E-06
204105_s_at	*NRCAM*	31.4487	8.35519E-07
205872_x_at	*PDE4DIP*	29.3072	2.66545E-07
213338_at	*TMEM158*	29.1076	3.92559E-05
228580_at	*HTRA3*	28.3271	1.16482E-08
204811_s_at	*CACNA2D2*	27.1019	1.69182E-08
229435_at	*GLIS3*	27.0546	1.82841E-06
209700_x_at	*PDE4DIP*	26.3127	8.02019E-08
1553681_a_at	*PRF1*	25.5974	2.09494E-05
1557143_at	*CSMD2*	25.2701	2.7608E-09
205227_at	*IL1RAP*	24.5748	6.61719E-08
211751_at	*PDE4DIP*	24.2715	8.02007E-07
236984_at	*C4orf26*	24.1849	0.000244588
230258_at	*GLIS3*	23.749	2.80979E-05
202833_s_at	*SERPINA1*	22.6065	1.46031E-05
220603_s_at	*MCTP2*	20.7203	6.24267E-07
228285_at	*TDRD9*	20.6104	0.000147332
227055_at	*METTL7B*	19.5324	8.74266E-07
243541_at	*IL31RA*	18.7172	7.13968E-06
205458_at	*MC1R*	18.4522	1.39741E-06
209765_at	*ADAM19*	15.9154	1.14453E-05
205578_at	*ROR2*	15.7427	3.99299E-06
229951_x_at	*LOC101060353*	14.998	2.0775E-08
211429_s_at	*SERPINA1*	14.6205	8.27393E-06
229538_s_at	*IQGAP3*	14.3839	0.000067806
242931_at	*LONRF3*	13.4778	5.78943E-07
223991_s_at	*GALNT2*	12.957	4.11138E-07
1569095_at	*LOC731424*	12.2529	8.91369E-05
208211_s_at	*ALK*	12.0543	5.97251E-07
206341_at	*IL2RA*	11.7023	2.39215E-05
217787_s_at	*GALNT2*	10.8619	2.40142E-06
231514_at	*C1orf94*	10.7105	9.47922E-06
224507_s_at	*MGC12916*	10.6575	4.46437E-06
202856_s_at	*SLC16A3*	10.2489	7.08763E-05

237461_at	*NLRP7*	10.0334	4.90624E-05
231118_at	*ANKRD35*	9.77106	5.43627E-05
218693_at	*TSPAN15*	9.20985	2.88604E-05
211026_s_at	*MGLL*	8.80714	4.88516E-05
222692_s_at	*FNDC3B*	8.76132	2.71516E-06
216620_s_at	*ARHGEF10*	8.65901	1.63233E-06
202464_s_at	*PFKFB3*	8.63605	2.14932E-05
222693_at	*FNDC3B*	8.27008	7.07582E-06
218618_s_at	*FNDC3B*	8.16665	5.8774E-07
217788_s_at	*GALNT2*	8.14349	2.27516E-06
228946_at	*INTU*	8.05283	7.52163E-07
1557523_at	*ATP6AP1L*	7.96348	0.000120045
224508_at	*MGC12916*	7.8976	5.37727E-06
239930_at	*GALNT2*	7.52359	5.4807E-08
219985_at	*HS3ST3A1*	7.21212	5.06385E-05
207357_s_at	*GALNT10*	7.21161	5.09135E-09
226944_at	*HTRA3*	6.55548	1.75569E-05
200770_s_at	*LAMC1*	6.55249	7.96629E-05
233016_at	*LOC100506546*	5.91107	0.00210425
218788_s_at	*SMYD3*	5.0459	7.58014E-06

*ALK-positive ALCL vs. ALK-negative ALCL (both systemic and primary cutaneous).
†False discovery rate step up/down Q-value.

Supplemental Table 9. Genes in DUSP22 Signature

Affy Probe ID	Gene Symbol	Fold Change*	FDR†
213245_at	*ADCY1*	133.137	2.98121E-07
220565_at	*CCR10*	96.1915	8.57885E-07
1556096_s_at	*UNC13C*	86.7452	0.000101333
208059_at	*CCR8*	61.621	2.23946E-10
220138_at	*HAND1*	57.3219	9.54306E-09
218796_at	*FERMT1*	53.7862	1.30469E-06
219496_at	*SOWAHC*	38.9448	8.54478E-06
227034_at	*SOWAHC*	36.6142	2.42157E-06
209016_s_at	*KRT7*	35.3159	1.83446E-13
236222_at	*MAATS1*	33.8439	2.8991E-06
1556095_at	*UNC13C*	33.3021	0.000669274
235049_at	*ADCY1*	31.4285	1.32403E-06
228367_at	*ALPK2*	31.3028	5.44611E-05
60474_at	*FERMT1*	30.4465	1.5204E-06
230964_at	*FREM2*	30.1503	7.82223E-10
1553645_at	*CCDC141*	27.8437	1.63281E-05
210394_x_at	*SSX4*	25.2963	0.000708766
205893_at	*NLGN1*	24.574	2.94986E-12
236565_s_at	*LARP6*	23.126	5.38451E-05
208195_at	*TTN*	23.0428	0.000049739
219932_at	*SLC27A6*	22.4375	0.000340336
225996_at	*LONRF2*	20.8478	0.000590502
218651_s_at	*LARP6*	19.8317	8.23296E-08
1554528_at	*MAATS1*	19.7634	0.000041741
219400_at	*CNTNAP1*	19.7144	7.25695E-06
207176_s_at	*CD80*	19.5636	2.3089E-11
221606_s_at	*HMGN5*	18.4732	0.00147441
1556488_s_at	*MAATS1*	18.4504	1.90656E-05
1569969_a_at	*UNC13C*	17.8085	0.000838471
231963_at	*ANKRD33B*	17.6795	6.43555E-05
230782_at	*SORD*	17.2015	2.2089E-06
1554147_s_at	*MAATS1*	16.5533	3.04175E-05
1563933_a_at	*PLD5*	16.2737	0.000321549
227812_at	*TNFRSF19*	16.1246	0.000212158
231517_at	*ZYG11A*	16.0256	0.000028375
213342_at	*YAP1*	15.9626	1.67212E-08
1555689_at	*CD80*	15.5778	2.68972E-08
203661_s_at	*TMOD1*	15.1419	0.00180256
224895_at	*YAP1*	14.3545	1.9865E-06
203662_s_at	*TMOD1*	13.76	0.000235545
205978_at	*KL*	13.0683	0.000027388
230864_at	*NIM1*	13.0417	2.07252E-06
214720_x_at	*SEPT10*	13.008	1.17841E-06

231361_at	*NLGN1*	12.7358	2.21712E-07
229603_at	*BBS12*	12.6859	0.000268725
226864_at	*PKIA*	12.5919	6.14051E-06
215189_at	*KRT86*	12.5202	3.5401E-06
203088_at	*FBLN5*	12.4872	4.58398E-06
205619_s_at	*MEOX1*	12.0361	6.85434E-06
227177_at	*CORO2A*	11.8642	0.000678382
226908_at	*LRIG3*	11.7454	2.24482E-06
238755_at	*RASSF10*	11.0122	7.47932E-05
230876_at	*ZNF883*	10.7791	4.18992E-05
229774_at	*CXXC4*	10.5063	0.000197028
1555719_a_at	*MAATS1*	10.4253	0.000245574
213280_at	*RAP1GAP2*	10.1531	7.4983E-08
1554519_at	*CD80*	10.0657	1.73774E-06
221035_s_at	*TEX14*	9.81904	2.05795E-05
228266_s_at	*HDGFRP3*	9.80324	0.000818756
212698_s_at	*SEPT10*	9.64867	8.88833E-08
229437_at	*MIR155*	9.33478	4.86986E-05
226536_at	*NSMCE2*	9.09913	8.82073E-10
1553663_a_at	*NPB*	8.98901	6.73369E-07
211674_x_at	*CTAG1A*	8.78657	0.000394788
1555370_a_at	*CAMTA1*	8.78596	8.34753E-06
205599_at	*TRAF1*	8.77786	5.32546E-08
230698_at	*CALN1*	8.67295	0.00274279
232010_at	*FSTL5*	8.64239	0.000335892
215733_x_at	*CTAG2*	8.63384	0.00105744
213268_at	*CAMTA1*	8.62828	3.24934E-05
214642_x_at	*MAGEA10-MAGEA5*	8.55268	0.00139503
215543_s_at	*LARGE*	8.47441	3.10083E-07
239178_at	*FGF9*	8.43953	5.43345E-06
202936_s_at	*SOX9*	8.3515	0.00138447
222061_at	*CD58*	8.15296	2.06624E-05
220277_at	*CXXC4*	8.1411	0.000120937
219740_at	*VASH2*	8.13114	0.00215882
219670_at	*BEND5*	8.03367	0.000670229
244764_at	*HIVEP3*	8.02761	2.52698E-06
234980_at	*TMEM56*	8.01249	0.00292626
235333_at	*B4GALT6*	7.90906	1.25696E-05
228080_at	*LAYN*	7.68927	0.000211574
1555168_a_at	*CALN1*	7.64897	6.97583E-06
229778_at	*C12orf39*	7.60338	2.34701E-06
211470_s_at	*SULT1C2*	7.58552	0.000106616
204612_at	*PKIA*	7.56383	0.000203188
229545_at	*FERMT1*	7.40012	1.72114E-06
238870_at	*KCNK9*	7.23953	0.00234992

235911_at	*MFI2*	7.2393	0.00026593
239282_at	*CCDC41*	7.09754	7.31864E-05
228796_at	*CPNE4*	6.98846	0.000487359
227506_at	*SLC16A9*	6.95895	0.000112276
217127_at	*CTH*	6.89798	0.00120004
206508_at	*CD70*	6.71923	0.000172111
228061_at	*CCDC126*	6.63146	7.25408E-06
203358_s_at	*EZH2*	6.55301	2.40315E-05
228547_at	*NRXN1*	6.30992	0.000434215
218625_at	*NRN1*	6.27098	0.00118673
228653_at	*SAMD5*	6.16281	1.40148E-06
209525_at	*HDGFRP3*	6.01183	3.17432E-05
228414_at	*KCNMA1*	5.85987	4.45097E-06
203771_s_at	*BLVRA*	5.82788	0.000436772
205538_at	*CORO2A*	5.72834	0.0005575
239975_at	*HLA-DPB2*	5.70986	0.000861296
210018_x_at	*MALT1*	5.48523	1.287E-07
208309_s_at	*MALT1*	5.38762	3.13025E-07
216945_x_at	*PASK*	5.35804	2.25552E-05
206376_at	*SLC6A15*	5.33454	7.68194E-06
235977_at	*LONRF2*	5.17776	3.98329E-05
225532_at	*CABLES1*	5.10685	5.12974E-05
231188_at	*ZSCAN2*	4.86644	9.46276E-06
232487_at	*SFT2D1*	4.82013	0.000627543
227166_at	*DNAJC18*	4.81288	1.43446E-06
206085_s_at	*CTH*	4.58973	0.00167605
216323_x_at	*TUBA3C*	4.52398	3.8899E-06
200824_at	*GSTP1*	4.49254	0.000232085

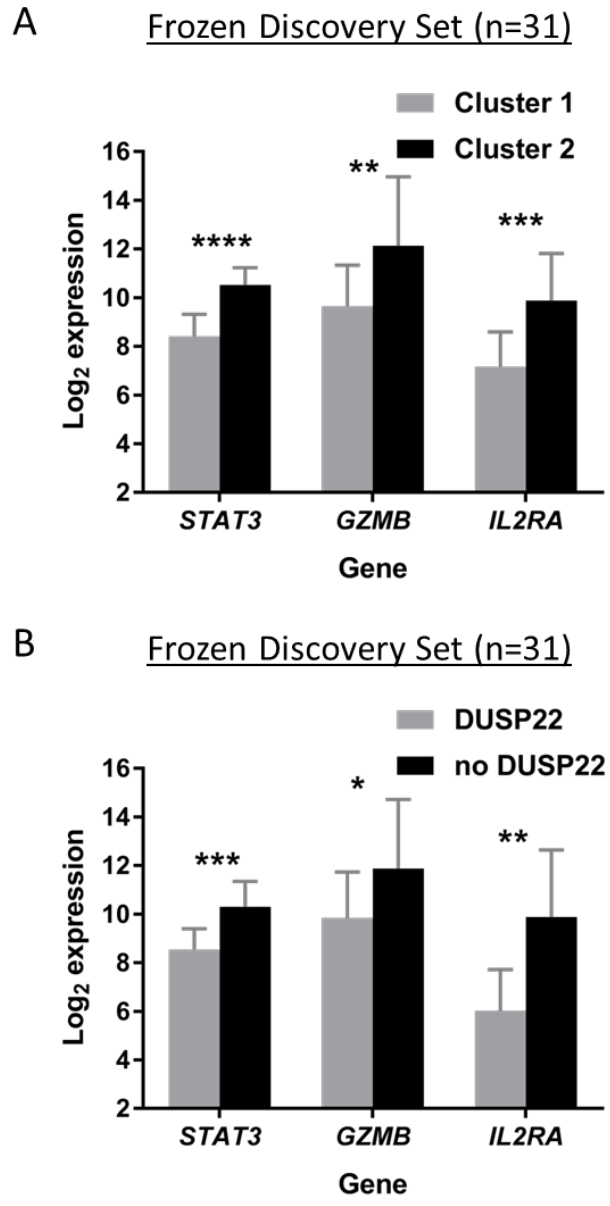
*ALCLs with *DUSP22* rearrangement vs. ALCLs without *DUSP22* rearrangement.
†False discovery rate step up/down Q-value.

Supplemental Table 10. Summary of differential methylation in ALCLs with *DUSP22* rearrangements

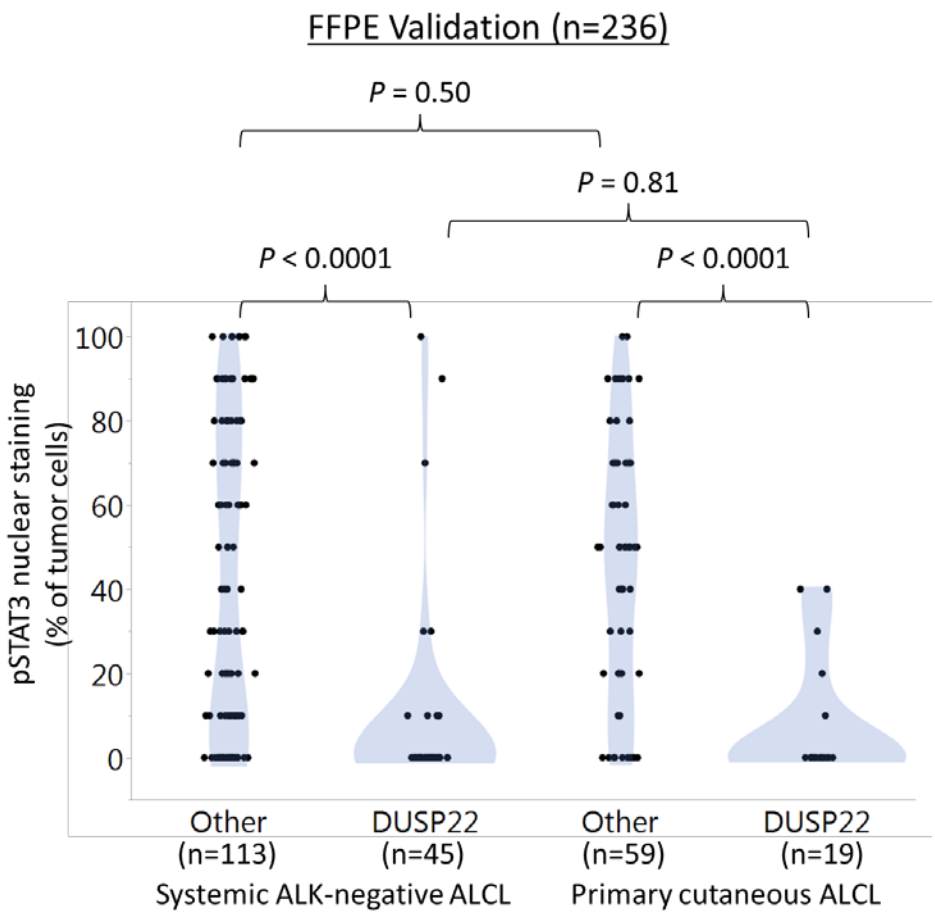
	Number of DMRs (Frozen RRBS Data)				Number of DMCs (FFPE Array Data)			
	DUSP22 vs. no DUSP22		DUSP22 vs. ALK		DUSP22 vs. no DUSP22		DUSP22 vs. ALK	
Region type	**Hyperm.**	**Hypom.**	**Hyperm.**	**Hypom.**	**Hyperm.**	**Hypom.**	**Hyperm.**	**Hypom.**
Promoter	313	4729	1066	1863	2750	13182	4801	7098
3'UTR	93	567	256	612	353	2543	611	1518
5'UTR	20	170	278	431	2201	9533	3713	5314
Intron	1739	12259	4822	18194	n/a	n/a	n/a	n/a
Exon	389	3750	1397	4963	n/a	n/a	n/a	n/a
Body	n/a	n/a	n/a	n/a	7410	46549	12693	26705
CpGi	268	7868	843	1567	1217	5718	3866	2832
CpGs	733	6404	1216	3505	3607	22598	6205	13685
LTR	220	1962	209	2203	591	6415	1049	3331
LINE	200	1572	285	1903	917	8870	1470	4578
SINE	1279	7463	982	6169	719	5993	1200	3418

CpGi, CpG island; CpGs, CpG shore; DMC, differentially methylated CpG probe; DMR, differentially methylated region; hyperm., hypermethylated; hypom., hypomethylated; LINE, long interspersed nuclear element; LTR, long terminal repeat; RRBS, reduced representation bisulfite sequencing; SINE, short interspersed nuclear element; UTR, untranslated region.

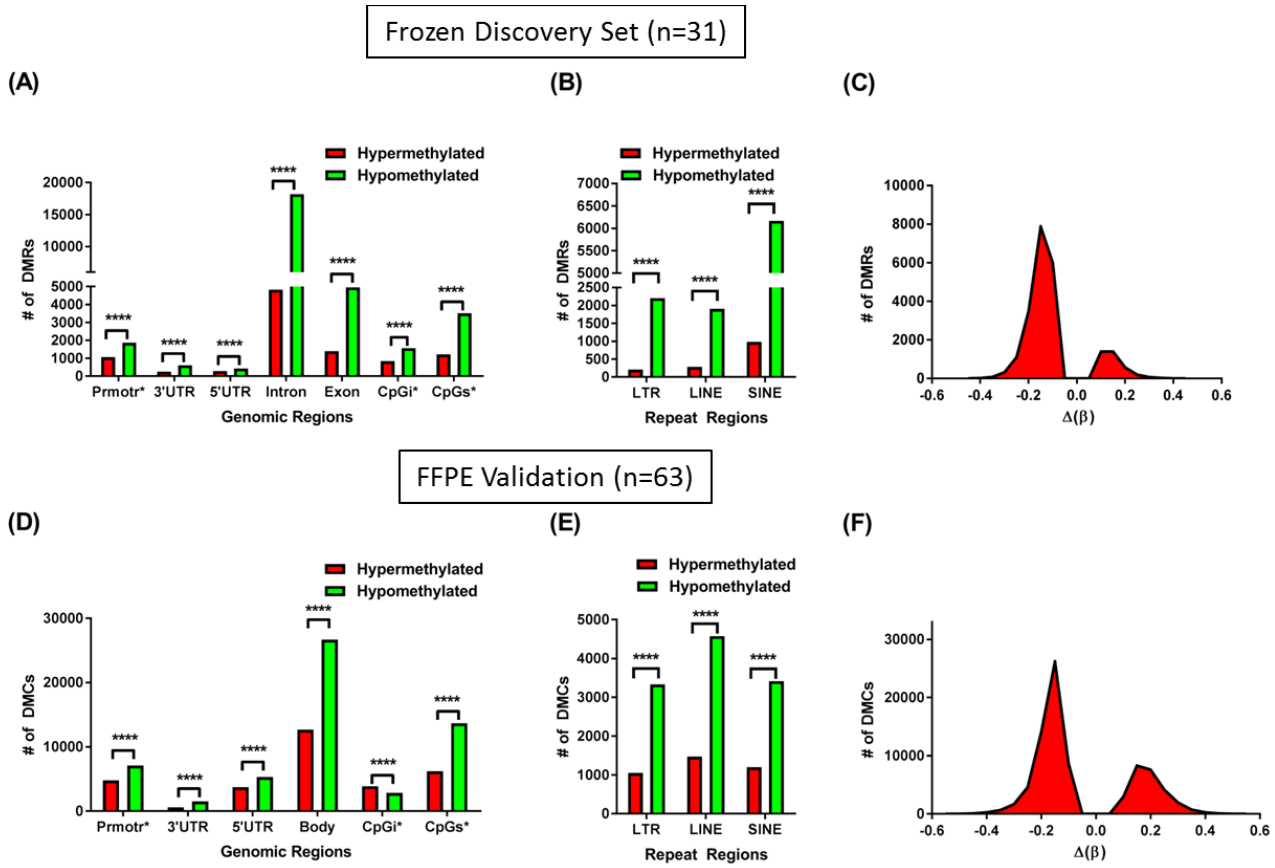
SUPPLEMENTAL FIGURES



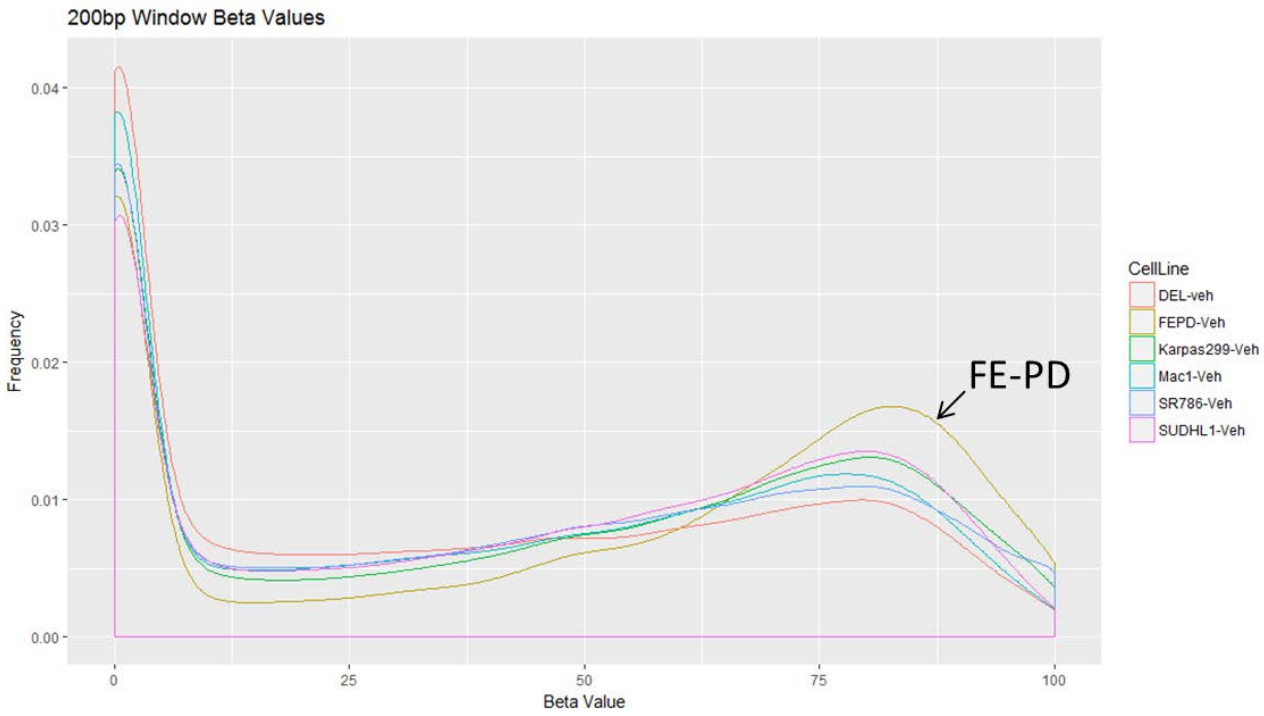
Supplemental Figure 1. Expression of *STAT3* and representative target genes. (A) Cases in Cluster 1 containing all ALCLs with *DUSP22* rearrangements show decreased expression of *STAT3* as well as the known STAT3 targets *GRZB* (encoding granzyme B) and *IL2RA* (encoding CD25). (B) ALCLs with *DUSP22* rearrangements similarly show lower expression of all 3 genes than ALCLs without *DUSP22* rearrangements. Data are shown as means ± standard deviations. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$ (Wilcoxon test).



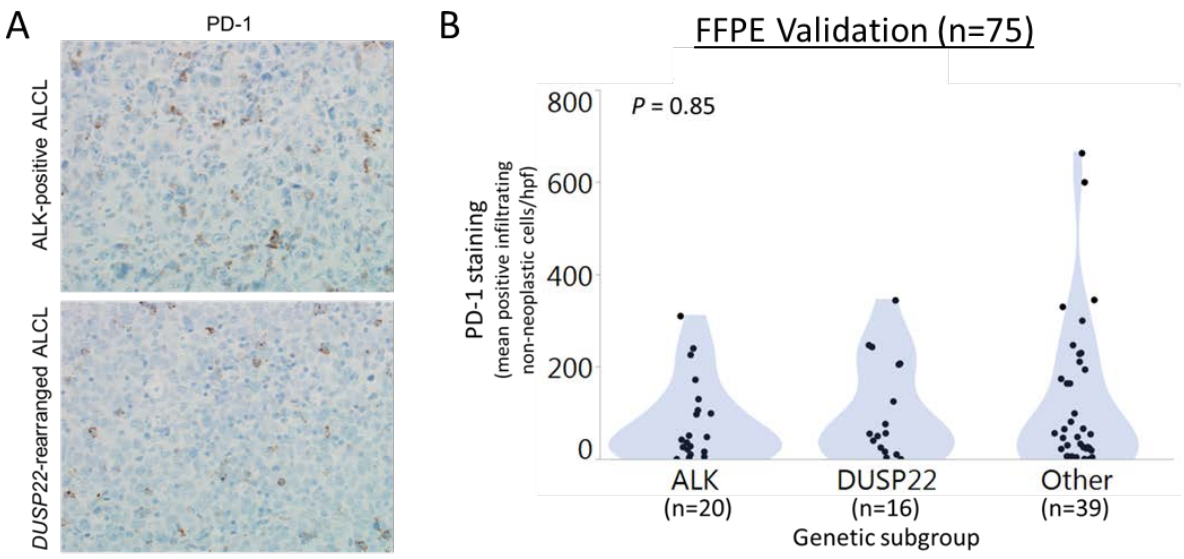
Supplemental Figure 2. Associations between *DUSP22* rearrangements and pSTAT3 expression in systemic and cutaneous ALK-negative ALCLs. The low expression of pSTAT3 in most cases of *DUSP22*-rearranged ALCL was similar for both systemic ALK-negative ALCL and primary cutaneous ALCL. Means ± standard deviations were: systemic ALK-negative ALCL without *DUSP22* rearrangement ("other"), 40.9 ± 36.1%; systemic ALK-negative ALCL with *DUSP22* rearrangement, 8.4 ± 22.5%; primary cutaneous ALCL without *DUSP22* rearrangement, 44.4 ± 31.0%; primary cutaneous ALCL with *DUSP22* rearrangements, 7.4 ± 14.1%. *P* values are shown (Wilcoxon test).



Supplemental Figure 3. Differential methylation between ALK-negative ALCLs with *DUSP22* rearrangements and ALK-positive ALCLs. (A) Reduced representation bisulfite sequencing (RRBS) of DNA extracted from frozen tissue in the discovery set. Differentially methylated regions (DMRs) reflect comparison of ALK-negative ALCLs with *DUSP22* rearrangements *versus* ALK-positive ALCLs and show marked hypomethylation in *DUSP22*-rearranged ALCLs across all types of genomic regions, including promoters, 3' and 5' untranslated regions (UTRs), introns, exons, CpG islands (CpGi), and CpG shores (CpGs). (B) Hypomethylation in *DUSP22*-rearranged ALCLs in the discovery set involves all types of non-coding regions, including long terminal repeats (LTRs), long interspersed nuclear elements (LINEs), and short interspersed nuclear elements (SINEs). (C) Histogram showing the distribution of methylation changes [Δ(β)] for DMRs across the genome in the discovery set. (D) MethylationEPIC BeadChip array analysis of DNA extracted from FFPE tissue in an independent validation set. Designations are similar to panel (A) except differentially methylated CpG probes (DMCs) are shown and intron and exon data are represented together as gene bodies. *DUSP22*-rearranged ALCLs are hypomethylated across all types of genomic regions. (E) Hypomethylation in *DUSP22*-rearranged ALCLs in the validation set involves all types of non-coding regions. (F) Histogram showing the distribution of methylation changes [Δ(β)] for DMRs across the genome in the validation set. ****$P \leq 0.0001$.



Supplemental Figure 4. Baseline DNA methylation in ALCL cell lines. Frequencies of beta values indicating baseline DNA methylation are shown for each of the cell lines studied. Although FE-PD bears a *DUSP22* rearrangement,[13] it demonstrated the highest degree of methylation among the ALCL cell lines. This is in contrast to tissue samples of ALCLs with *DUSP22* rearrangements, which show marked hypomethylation compared to ALCLs without *DUSP22* rearrangements (main text, Figure 4). This discrepancy likely represents an *in vitro* phenomenon, as many cancer cell lines demonstrate hypermethylation compared to the primary tumors from which they are derived.[18] Similar to other ALCL cell lines, FE-PD demonstrated up-regulation of the DUSP22 gene expression signature when hypomethylated pharmacologically using decitabine (main text, Figure 5A).



Supplemental Figure 5. PD-1 staining in ALCL by genetic subtype. (A) Immunohistochemistry for PD-1 in ALK-positive ALCL and *DUSP22*-rearranged ALCL shows scattered positive non-neoplastic cells in the microenvironment. Original magnification, 400×. (B) No significant difference was observed among genetic subtypes in the mean number of non-neoplastic PD-1-positive cells per high-powered field (hpf). N = 75; Mean ± S.D. for ALK, 83 ± 20; DUSP22, 106 ± 108; Other, 118 ± 158; $P = 0.85$ (Kruskal-Wallis test).

SUPPLEMENTAL REFERENCES

1. Akbani R, Ng PK, Werner HM, et al. A pan-cancer proteomic perspective on The Cancer Genome Atlas. *Nat Commun*. 2014;5:3887.

2. Tibes R, Qiu Y, Lu Y, et al. Reverse phase protein array: validation of a novel proteomic technology and utility for analysis of primary leukemia specimens and hematopoietic stem cells. *Mol Cancer Ther*. 2006;5:2512-21.

3. Sharma A, Oishi N, Boddicker RL, et al. Recurrent STAT3-JAK2 fusions in indolent T-cell lymphoproliferative disorder of the gastrointestinal tract. *Blood*. 2018;131:2262-6.

4. Kalari KR, Nair AA, Bhavsar JD, et al. MAP-RSeq: Mayo Analysis Pipeline for RNA sequencing. *BMC Bioinformatics*. 2014;15:224.

5. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15-21.

6. Hu G, Dasari S, Asmann YW, et al. Targetable fusions of the FRK tyrosine kinase in ALK-negative anaplastic large cell lymphoma. *Leukemia*. 2018;32:565-9.

7. Feldman AL, Law ME, Inwards DJ, Dogan A, McClure RF, Macon WR. PAX5-positive T-cell anaplastic large cell lymphomas associated with extra copies of the PAX5 gene locus. *Mod Pathol*. 2010;23:593-602.

8. Feldman AL, Dogan A, Smith DI, et al. Discovery of recurrent t(6;7)(p25.3;q32.3) translocations in ALK-negative anaplastic large cell lymphomas by massively-parallel genomic sequencing. *Blood*. 2011;117:915-9.

9. Walker DL, Bhagwate AV, Baheti S, et al. DNA methylation profiling: comparison of genome-wide sequencing methods and the Infinium Human Methylation 450 Bead Chip. *Epigenomics*. 2015;7:1287-302.

10. Akalin A, Kormaksson M, Li S, et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol*. 2012;13:R87.

11. Fortin JP, Triche TJ, Jr., Hansen KD. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. *Bioinformatics*. 2017;33:558-60.

12. Bose M, Wu C, Pankow JS, et al. Evaluation of microarray-based DNA methylation measurement using technical replicates: the Atherosclerosis Risk In Communities (ARIC) Study. *BMC Bioinformatics*. 2014;15:312.

13. Vasmatzis G, Johnson SH, Knudson RA, et al. Genome-wide analysis reveals recurrent structural abnormalities of TP63 and other p53-related genes in peripheral T-cell lymphomas. *Blood*. 2012;120:2280-9.

14. Wang X, Dasari S, Nowakowski GS, et al. Retinoic acid receptor alpha drives cell cycle progression and is associated with increased sensitivity to retinoids in T-cell lymphoma. *Oncotarget*. 2017;8:26245-55.

15. Pedersen MB, Danielsen AV, Hamilton-Dutoit SJ, et al. High intratumoral macrophage content is an adverse prognostic feature in anaplastic large cell lymphoma. *Histopathology*. 2014;65:490-500.

16. Pedersen MB, Hamilton-Dutoit SJ, Bendix K, et al. DUSP22 and TP63 rearrangements predict outcome of ALK-negative anaplastic large cell lymphoma: a Danish cohort study. *Blood*. 2017;130:554-7.

17. Khoury JD, Medeiros LJ, Rassidakis GZ, et al. Differential expression and clinical significance of tyrosine-phosphorylated STAT3 in ALK+ and ALK- anaplastic large cell lymphoma. *Clin Cancer Res*. 2003;9:3692-9.

18. Smiraglia DJ, Rush LJ, Fruhwald MC, et al. Excessive CpG island hypermethylation in cancer cell lines versus primary human malignancies. *Hum Mol Genet*. 2001;10:1413-9.