Integrated Phosphoproteomic and Metabolomic Profiling Reveals NPM-ALK-Mediated Phosphorylation of PKM2 and Metabolic Reprogramming in Anaplastic Large Cell Lymphoma

SUPPLEMENTAL METHODS

Phosphoproteomics: Sixty million cells were treated for 6 hours with CEP or DMSO. Cells were lysed with 9M urea/20mM HEPES pH8.0/0.1%SDS and a cocktail of phosphatase inhibitors (sodium orthovanadate, sodium pyrophosphate and β-glycerophosphate). For each sample, 6mg of protein were reduced with 4.5mM DTT for 30min at 60°C, alkylated with 10mM iodoacetamide for 30min at room temperature in the dark and then digested with trypsin overnight at 37°C. Samples were desalted on a C18 cartridge (Sep-Pak plus C18 cartridge, Waters), then dried before any further processing. Each sample was prepared in triplicate. Metal oxide affinity chromatography (MOAC) was performed using titanium dioxide (TiO_2) microparticles from GL Sciences Inc (Titansphere® Phos-TiO, GL Sciences Inc., Torrance, CA) in a microparticles-to-protein ratio of 6/1 (w/w). Hydrophilic phosphopeptides were eluted with 5% ammonium hydroxide solution and hydrophobic phosphopeptides were eluted with 5% pyrrolidine solution. After elution, peptides were dried. The equivalent of 5mg of protein was further enriched for phosphorylated tyrosine peptides by overnight immunoprecipitation using a cocktail of anti-phosphotyrosine antibodies (4G10, Millipore /PT-66, Sigma/p-Tyr-100, Cell Signaling Technology). After elution, phosphotyrosine peptides were dried.

Samples from MOAC were reconstituted in 25µl loading buffer (0.1% TFA/2% acetonitrile) while samples from MOAC followed by pY-IP were reconstituted in 35 µl loading buffer. LTQ Orbitrap XL(ThermoFisher) in-line with Paradigm MS2 (Michrom bioresources) was employed for acquiring high-resolution MS and MS/MS data. Ten µl of the peptides was loaded onto a sample trap (Captrap, Bruker-Michrom) that was in-line with a nano-capillary column (Picofrit, 75µm i.d.x 15µm tip, New Objective) packed in-house with 10 cm of MAGIC AQ C18 reverse phase material (Michrom Bioresource). Two

different gradient programs, one each for MOAC and phosphotyrosine immunoprecipitation samples, were used for peptide elution. For MOAC samples, a gradient of 5-40% buffer B (95% acetonitrile/1% acetic acid) in 135 min and 5 min wash with 100% buffer B followed by 30 min of re-equilibration with buffer A (2% acetonitrile/1% acetic acid) was used. For phosphotyrosine immunoprecipitation samples, which were much less complex mixture of peptides, 5-40% gradient with buffer B was achieved in 75 min followed by 5 min wash with buffer B and 30 min re-equilibration. Flow rate was ~0.3µl/min. Peptides were directly introduced into the mass spectrometer using a nano-spray source. Orbitrap was set to collect 1 MS scan between 400-2000 m/z (resolution of 30,000 @ 400 m/z) followed by data dependent CID spectra on top 9 ions in LTQ (normalized collision energy ~35%). Technical duplicate data for each of the MOAC and triplicate data for phosphotyrosine immunoprecipitation samples were acquired.

RAW files were converted to mzXML using msconvert. The mass spectrometry data was then searched with the Swissprot Human taxonomic protein database (2012_04 release) appended with common proteomics contaminants and reverse sequences as decoys. Searches were performed with X!Tandem (version 2010.10.01.1) using the k-score plugin. The search results were then post-processed through the Trans-Proteomic Pipeline (TPP) using PeptideProphet and ProteinProphet. Spectral counts for the Trans-Proteomic Pipeline (TPP) results were obtained for each cell line using the spectral counting software ABACUS.

Metabolomics: Samples were removed from -80°C storage and placed on ice for the duration of the experiment. Triple-distilled H2O (300 µL) was added to each sample tube containing a frozen pellet, and two Process Blanks (PB) were created by placing same volume of H20 into empty tubes. Samples were re-suspended, transferred to the wells of a 1mL deep-well plate containing 2-3 3mm glass beads and homogenized for 5 minutes at room temperature (RT) by a Hard Tissue homogenizer (VWR).

Protein concentration was determined by mixing a 2µL aliquot of homogenized sample with 198µL of Bradford reagent (Sigma B6916). Assay was performed at RT in a clear, flat bottom, shallow 96-well plate. The plate was vortexed and absorption data collected on BioTek uQuant with KC Junior software (v1.41.8) within 15 minutes after mixing sample and reagent. A 14 point calibration curve was created using bovine albumin (Sigma B4287). All measurements were performed in duplicate on the same plate.

Sample volumes for metabolomic analyses were determined by normalization to a protein content of 8 mg/mL using the formula: 8/x (mg/mL) *100= µL sample volume aliquoted. Pooled human plasma controls (HB, pooled human plasma from in-house reserve of American Red Cross supply stored at -80°C) were processed in parallel with other samples following this step.

For de-proteination the calculated volume of each sample was transferred to a 2mL deep-well plate and diluted with 8 volumes of 100% ethanol containing 13C algal amino acid mixture (Sigma-Aldrich 426199) as a recovery standard. The plate was then vortexed, centrifuged at 4750 RPM, 4°C for 10 min, and supernatant transferred to a new 2 mL deep-well plate. A pre-washed filter plate (Nunc 278010) was attached and the assembly centrifuged at 4750 RPM, 4°C for 10 min to further clarify the samples.

Two identical aliquots of each sample were transferred into glass inserts placed in custom-made 96 well aluminum plates and dried down under the stream of nitrogen for 1.5 hr. One set of the aliquots was than prepared for GC-MS analysis and another for LC-MS.

LC-MS Method: Dried LC-MS samples were re-suspended in H2O (27µL) containing injection standards (tBoc-L-Alanine, tBoc-L-Asparagine, tBoc-L-Phenylalanine). Two solvent blanks (SB, 100% H2O with t-Boc standards) and 2 water blanks (BB, 100% H2O) were added at the end of the sample prep. All samples were then capped, vortexed at 2000 RPM for 2 minutes, and transferred to the instrument for analysis.

Samples were analyzed on an Agilent 1200 LC / 6530 qTOF LC-MS system using a Waters Acquity HSS T3 reversed-phase column (1.8 μ, 50 mm x 2.1 mm ID). For positive ion detection, mobile phase A was 100% water with 0.1% formic acid and mobile phase B was 100% methanol with 0.1% formic acid. For negative ion detection, formic acid was replaced with 0.1% ammonium bicarbonate. The gradient was as follows: 0-0.5 min 1% B, 0.5-2 min 1-99% B, 2-6 min 99% B, 6-9 min 1% B. The flow rate was 0.35 mL/min and column temperature kept at 40°C. The injection volumes for positive and negative mode were 7 μL and 12 μL respectively. Jetstream ESI source gas temperature was set to 350°C, drying gas flow rate - 10 L/min, nebulizer pressure - 30psig, sheath gas temperature - 350°C and flow rate - 11 L/min. Capillary current was set to 10.0 μA and VCap voltage - to 3500V. Inline mass calibration was performed by delivering reference masses through a dedicated inlet on mass spectrometer. Reference masses were 99.09 and 922.009 m/z (Positive mode) and 62.0009 and 981.9956 m/z (Negative mode). Data analysis was performed using Find by Formula algorithm of the Agilent MassHunter Qualitative analysis software. Library for the analysis was constructed partially from the authentic standards and partially by creating "Known Unknowns" from mass spectral features consistently present in a large proportion of the samples. Molecular features were identified by Molecular Feature Extractor (Agilent MassHunter component) and processed using in-house software.

GC-MS Method: The duplicate plate destined for GC analysis was transferred to a dry box (nitrogen atmosphere) and 60μL of derivitization reagent (BSTFA in anhydrous pyridine containing butylated hydroxytoluene as a derivatization standard) was added to each dried sample. Process blanks, solvent blanks (SB, 100% derivative reagent) and blank blanks were processed identically. Samples were sealed using crimp-caps, and the plate was removed from the dry box, vortexed (2000 rpm, 2 min), and incubated (60°C, 30 min) on a heat block. The GC plate was allowed to cool to RT, then transferred to the GC autosampler. GC analyses were done on an Agilent 7890A-5975C inert XL MSD GCMS instrument. Samples were injected (1 uL, split 10:1, 250C) onto a HP-5MS 5% phenyl-methyl Silox (30 m x 250 uM x

0.25 uM) column and eluted (1 mL/min H2, 60C(0.5 min hold), 10 C/min to 80C, 50 C/min to 325C, 4.5

min at 325C for 11.9 min total run time), transferred to the MSD unit (280C), ionized (EI, 70V), and

scanned from 800-50 m/z after a solvent delay of 3 min (source 230C, quad at 150C). Data was analyzed

using Agilent CHEMSTATION software.

Data analysis and normalization: Data from the three MS platforms were compiled for analysis. Data

were normalized by both column (sample) and row (metabolite) prior to analysis.

Metabolomic Pathway Analysis: MetaboAnalyst 2.0 was used to generate significant pathways in

metabolomics dataset. The raw data was inserted into the software and the phenotype label was set to

"discrete" (DMSO vs. CEP). Identified compounds were matched based on KEGG ID. Missing values

were replaced by a small value. The data was normalized by creating a pooled average sample from

DMSO group. No data transformation or data filtering was applied, while data was scaled by autoscaling

(mean-centered and divided by the standard deviation of each variable). Global test was used for

pathway enrichment analysis and out-degree centrality was used for pathway topology analysis. The p-

value is assigned based on the coverage and change observed in a given metabolic pathway. A pathway

impact score was assigned based on the location of the identified metabolites in a given pathway. A

significant threshold was assigned at y=1/x for subsequent analysis.

Metabolic Flux: Cells were treated with DMSO or CEP (300nM) for 6 hours in serum free media. Media

was then changed to one containing ^{13}C-glucose for 30 minutes. A mixture of methanol chloroform and

water were used to extract metabolites from pelleted cells. Following centrifugation at 10000rpm for 5

min, supernatant containing metabolites were transferred to autosampler vials for LC-MS analysis.

Agilent 1200 chromatography (Santa Clara, CA), Luna (Phenomenex, Torrance, CA) NH2 HILIC

(hydrophilic interaction chromatography) column was used for chromatographic separation. Agilent

6200 series Time-of-flight mass spectrometer was used for detection and quantification. Data were processed by MassHunter workstation software, version B.04 (Agilent, Santa Clara, CA).

Table S1: Phosphoproteomic results for metabolic proteins.

		Replicate 1		Replicate 2		Replicate 3		Merged		Merged Difference	Avg Fold Change
		DMSO	CEP	DMSO	CEP	DMSO	CEP	DMSO	CEP		
Glycolysis	**MOAC**										
	ENO1	6	0	15	7	0	0	21	7	**-14**	0.55
	LDHA	9	7	6	6	0	0	15	13	**-2**	0.93
	LDHB	6	7	11	8	0	0	17	15	**-2**	0.96
	PDHA1	9	10	21	8	6	10	36	28	**-8**	1.03
	PFKP	0	1	0	0	0	0	0	1	**1**	1.33
	PGK1	12	7	3	4	7	12	22	23	**1**	1.16
	PGM1	2	6	2	13	0	1	4	20	**16**	3.00
	PKM2	7	4	4	5	0	0	11	9	**-2**	0.94
	ACSS2	2	6	3	0	0	0	5	6	**1**	1.19
	TPI1	6	11	6	9	6	5	18	25	**7**	1.33
	pY-IP										
	ENO1	12	2	60	41	15	14	87	57	**-30**	0.62
	ALDOA	0	0	0	0	1	0	1	0	**-1**	0.83
	GAPDH	0	0	2	0	5	8	7	8	**1**	0.94
	LDHA	18	31	15	11	15	14	48	56	**8**	1.12
	LDHB	6	2	12	4	19	17	37	23	**-14**	0.57
	PFKL	0	0	6	0	0	0	6	0	**-6**	0.71
	PGK1	0	0	3	0	0	1	3	1	**-2**	1.08
	PKM2	12	3	49	44	30	16	91	63	**-28**	0.59
	TPI1	0	0	0	0	0	1	0	1	**1**	1.33
Pentose Phosphate Pathway	**MOAC**										
	PRPS1L1	0	0	0	0	1	0	1	0	**-1**	0.83
	PFKP	0	1	0	0	0	0	0	1	**1**	1.33
	PGM1	2	6	2	13	0	1	4	20	**16**	3.00
	PGM2	7	1	6	6	0	4	13	11	**-2**	2.08
	pY-IP										
	ALDOA	0	0	0	0	1	0	1	0	**-1**	0.83
	PFKL	0	0	6	0	0	0	6	0	**-6**	0.71
	TKT	0	0	6	2	7	4	13	6	**-7**	0.68
TCA Cycle	**MOAC**										
	ACLY	12	4	11	9	4	8	27	21	**-6**	1.01
	PDHA	9	10	21	8	6	10	36	28	**-8**	1.03
	pY-IP										
	SUCLA2	0	0	0	0	4	4	4	4	**0**	1.00
	ACLY	20	51	108	90	52	46	180	187	**7**	1.40

| Nucleotide Metabolism										
MOAC										
POLD3	1	0	4	0	1	1	6	1	**-5**	0.57
TWISTNB	6	2	4	3	2	5	12	10	**-2**	1.08
POLA2	1	0	0	4	0	0	1	4	**3**	2.17
POLR2A	3	1	1	0	0	0	4	1	**-3**	0.67
TXNRD1	7	8	2	7	0	4	9	19	**10**	2.93
CAD	6	5	22	12	2	5	30	22	**-8**	1.14
PAICS	3	1	2	4	0	0	5	5	**0**	1.06
NUDT5	0	3	5	4	1	1	6	8	**2**	1.94
PRPS1L1	0	0	0	0	1	0	1	0	**-1**	0.83
IMPDH2	4	4	0	0	0	0	4	4	**0**	1.00
ATIC	2	5	2	6	0	0	4	11	**7**	1.78
PFAS	2	0	0	1	0	0	2	1	**-1**	1.11
PGM1	2	6	2	13	0	1	4	20	**16**	3.00
PKM2	7	4	4	5	0	0	11	9	**-2**	0.94
RRM2	27	28	17	16	9	12	53	56	**3**	1.09
pY-IP										
POLR3A	0	0	2	2	2	0	4	2	**-2**	0.78
POLR2A	0	0	8	5	12	4	20	9	**-11**	0.68
RRM1	0	0	7	5	4	1	11	6	**-5**	0.72
RRM2	2	0	0	0	0	8	2	8	**6**	3.44
TXNRD1	12	7	52	57	18	20	82	84	**2**	0.94
UMPS	0	3	6	6	6	6	12	15	**3**	2.00
CAD	0	0	0	0	6	6	6	6	**0**	1.00
POLR1B	0	0	1	0	0	0	1	0	**-1**	0.83
NUDT5	0	0	13	4	12	12	25	16	**-9**	0.79
IMPDH2	0	0	0	0	3	2	3	2	**-1**	0.92
ATIC	0	4	6	6	7	6	13	16	**3**	2.29
PKM2	12	3	49	44	30	16	91	63	**-28**	0.59

Supplemental Table 1: Phosphoproteomic results for metabolic proteins. The metabolic proteins that reside in glycolysis, the pentose phosphate pathway, the TCA cycle and nucleotide metabolism (based on KEGG annotation) are tabulated with their corresponding raw spectral counts in DMSO and CEP treated conditions.   Four biological replicates are shown.   Proteins are grouped based on their identification in the pS/T (MOAC) dataset or the pY (pY-IP) dataset. The merged column combines all 3 replicates.   The "Merged Difference" column displays the combined data from all three biological replicates as represented by the change in spectral counts following ALK inhibition.

Table S2: Metabolomic results for ALK regulated metabolites.

		Normalized Abundance								Fold change
		DMSO				CEP				
Pyrimidine Metabolism										
C00106	Uracil	21.4	16.6	9.9	19.6	6.2	7.7	9.9	8.7	**0.48**
C00119	5-Phospho-alpha-D-ribose 1-diphosphate	19.4	16.1	17.7	19.3	4.5	7.2	8	7.6	**0.38**
C00295	Orotate	29	23.3	18.5	18.1	4.3	2.7	2.8	1.4	**0.13**
C00299	Uridine	13.6	10.6	9.5	10.4	12.3	15.8	14	13.7	**1.26**
C00337	(S)-Dihydroorotate	31.7	27.1	23.5	17.8	0	0	0	0	**0**
C00364	dTMP	13.8	20.5	16.7	16.5	8.9	7.2	9	7.6	**0.48**
C00365	dUMP	13.6	13.7	14.3	13.2	9.7	13.1	11.3	11.1	**0.82**
C00380	Cytosine	24.3	23	22.8	29.4	0.2	0	0	0.3	**0**
C00429	5,6-Dihydrouracil	19.8	13.7	18.4	19.5	4.1	7.7	7.5	9.2	**0.4**
C00475	Cytidine	6.6	7.6	5.3	9.3	16.5	17.1	18.8	18.8	**2.47**
C00881	Deoxycytidine	3.6	4.7	9.2	4.5	23.1	4.8	25.2	24.9	**3.54**
C02354	2',3'-Cyclic CMP	47.8	0	24.3	27.9	0	0	0	0	**0**
Purine Metabolism										
C00020	AMP	18.6	22.2	13.5	18.2	3.4	5.5	14.2	4.3	**0.38**
C00117	D-Ribose 5-phosphate	21.2	22.1	18.7	20	3.8	5	4.7	4.4	**0.22**
C00119	5-Phospho-alpha-D-ribose 1-diphosphate	19.4	16.1	17.7	19.3	4.5	7.2	8	7.6	**0.38**
C00144	GMP	10.2	11.2	9.9	12.9	14.8	13.7	13.8	13.6	**1.27**
C00212	Adenosine	16.2	14.4	10.6	15.1	11	11	11.7	10.1	**0.78**
C00242	Guanine	10.4	10.3	8.1	11.1	13.5	15.1	16.6	14.8	**1.51**
C00294	Inosine	19.2	18.7	12.3	16.6	7.8	7.4	10	7.9	**0.5**
C00301	ADP-ribose	21.7	15.8	26.7	25.7	1.3	2.9	3.6	2.3	**0.11**
C00330	Deoxyguanosine	16.3	15	11.1	15.8	10.2	10.2	11.6	9.8	**0.72**
C00387	Guanosine	17.4	16.5	12.2	18.1	8.3	8.3	10.7	8.6	**0.56**
C00575	3',5'-Cyclic AMP	14.2	14.5	14.6	12.7	11.7	13.5	10.1	8.8	**0.79**
Glycine, Serine, Threonine, Metabolism										
C00300	Creatine	14.8	15.3	13.5	13.5	10.4	11.5	11	10	**0.75**
C00719	Betaine	14.7	14.9	15.3	14.1	6.4	10.4	10.6	13.5	**0.69**
Pentose Phosphate Pathway										
C00117	D-Ribose 5-phosphate	21.2	22.1	18.7	20	3.8	5	4.7	4.4	**0.22**
C00119	5-Phospho-alpha-D-ribose 1-diphosphate	19.4	16.1	17.7	19.3	4.5	7.2	8	7.6	**0.38**
C03752	2-Amino-2-deoxy-D-gluconate	0	0	0	0	16.6	19.6	48.4	15.4	**100**
Glycolysis										
C00111	Glycerone phosphate	19	20.6	17.1	16.1	5.7	7.3	6.1	8	**0.37**
C00186	(S)-Lactate	22.6	16.6	19.6	22.1	3.2	4.8	5.2	5.9	**0.24**

TCA Cycle/ Oxidative Phosphorylation										
C00042	NAD+	16.2	13.8	17.2	19.9	6.7	8.9	8.9	8.4	**0.49**
C00003	Succinate	17.6	19.8	22.3	18.1	6.4	4.7	6.6	4.4	**0.28**
Glycerophospholipid Metabolism										
C00111	Glycerone phosphate	19	20.6	17.1	16.1	5.7	7.3	6.1	8	**0.37**
C00307	CDP-choline	16.6	12.8	13.1	14.7	10.3	12.1	10.5	9.9	**0.75**
C00346	Ethanolamine phosphate	8.2	12.9	7.9	12.1	13.9	16.2	16.2	12.6	**1.43**
C00670	sn-glycero-3-Phosphocholine	4.8	4.5	4.1	8.4	14.4	26	12.7	25	**3.57**
Fatty Acid Metabolism/ Biosynthesis										
C00249	Hexadecanoic acid	16.7	16.3	17.2	16	7.1	11.2	11.9	3.7	**0.51**
C06424	Tetradecanoic acid	17.7	15.4	17.8	14.9	6.9	11.7	11.6	4	**0.52**
C00010	CoA	31.1	10	21.8	27.6	1.6	4.9	1.5	1.6	**0.11**
C00489	Glutarate	17.2	21.5	23.1	9.9	9.3	11	8	0	**0.39**

Figure S1

A

SUDHL-1   SUPM2

CEP
300 nM for 6hr   −   +   −   +

p-ALK (Y1604)

ALK

Actin

B

% Viable
% Dead

% of Cell Population

Untreated 0hr   DMSO 6hr   CEP 6hr

C

PKM2 pY105
TATESFASDPIL pYRPVAVALDTK
Adjusted probability = 0.9948

D

PKM2 pY390
EAEAAIpYHLQLFEELRR
Adjusted Probability = 0.9945
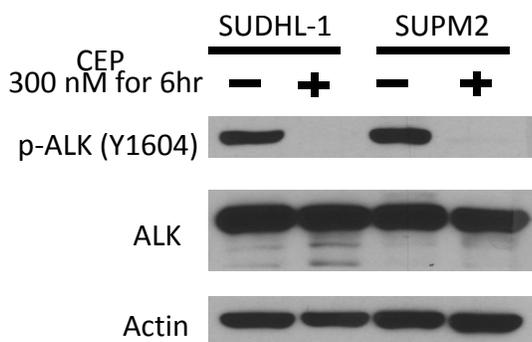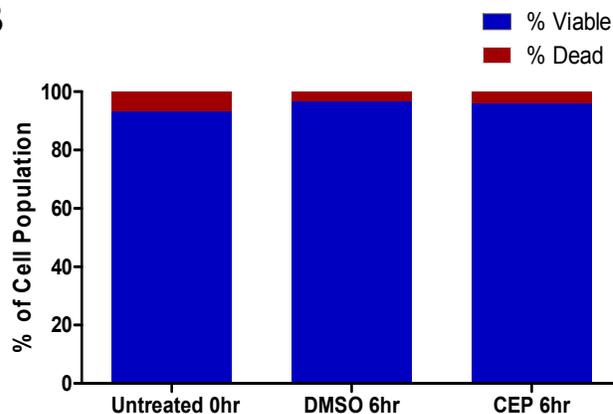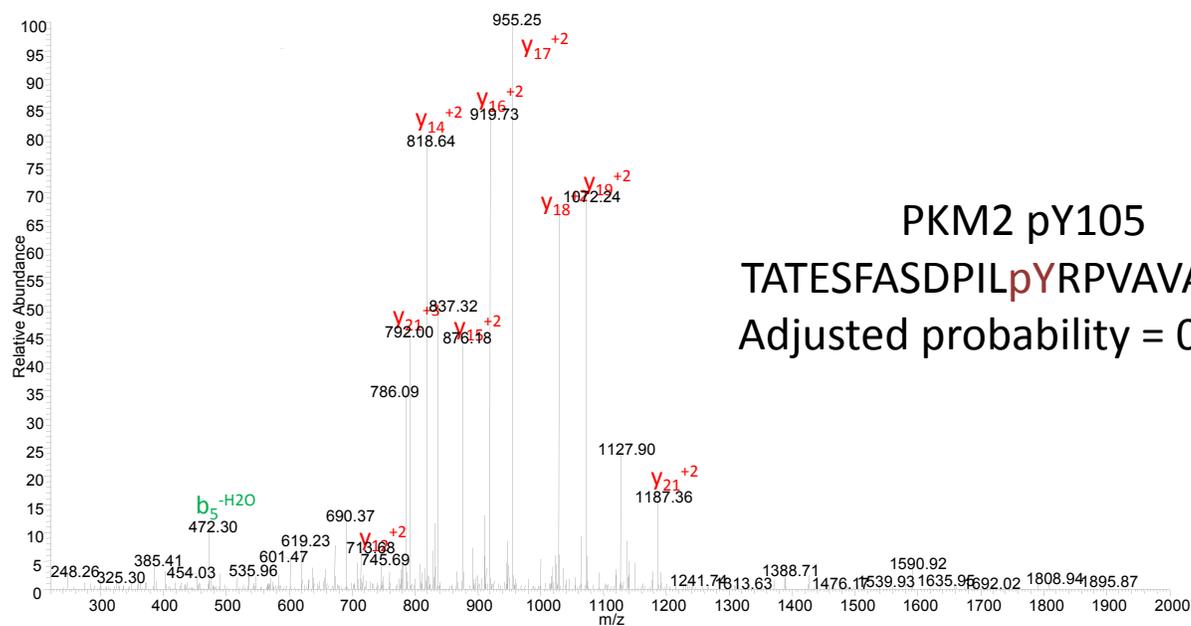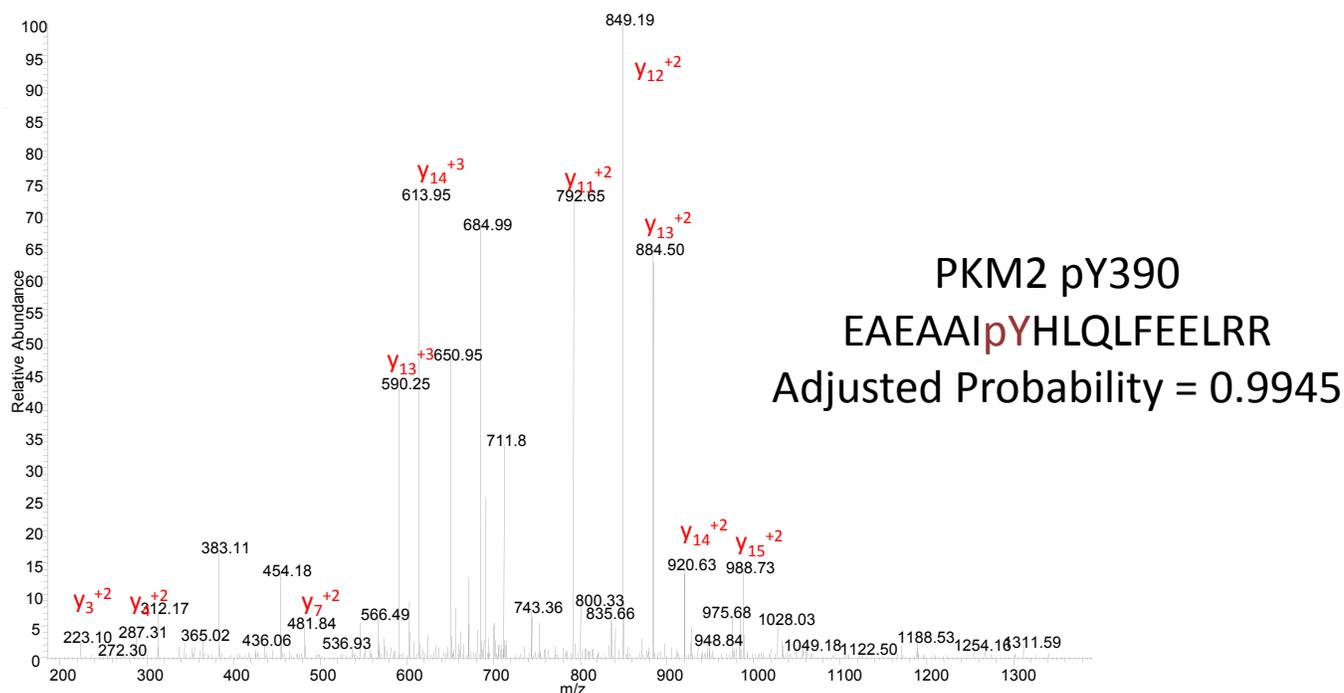
Supplemental Figure 1: Validation of conditions for phosphoproteomic studies and identification of PKM2.

A) Immunoblots of SU-DHL-1 and SUPM2 cell lysates treated with 300nM CEP for 6hrs in 1% FBS. B) The same treatment conditions from (a) were used to determine viability by trypan blue staining. C) A representative tandem mass spectrum of the PKM2 peptide containing the pY105 residue. D) A representative tandem mass spectrum of the PKM2 peptide containing the pY390 residue. Identified b and y ions are indicated.

Figure S2

A

|	ALK+ALCL				ALK-ALCL	SS	MF	NKTL	
	DEL	Karpas299	SUDHL-1	SUPM2	MAC1	HUT78	HH	NK92	YT
Peptides	8	9	6	11	0	0	0	3	4
Spec Counts	63	34	56	63	0	0	0	19	20

B

C

SUPM2 Cells

DEL Cells
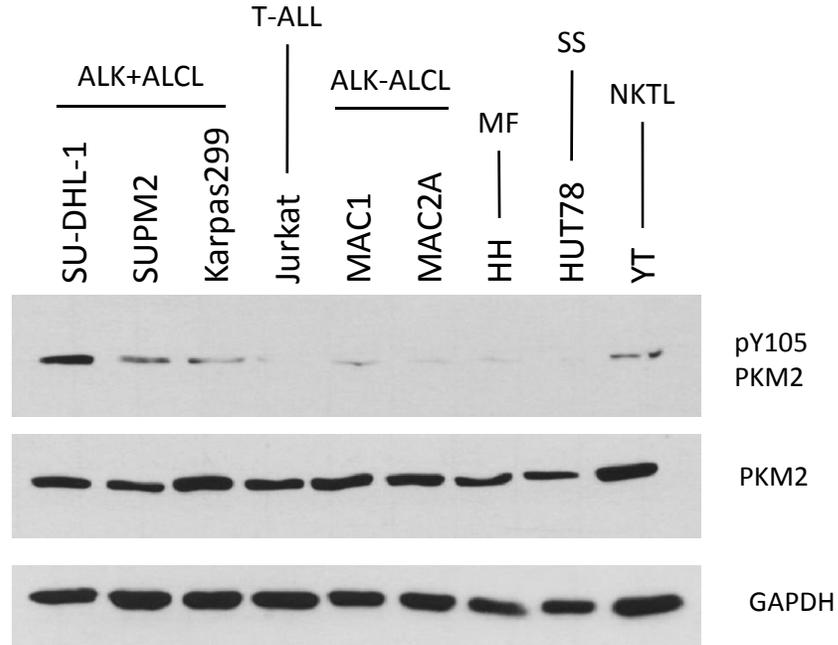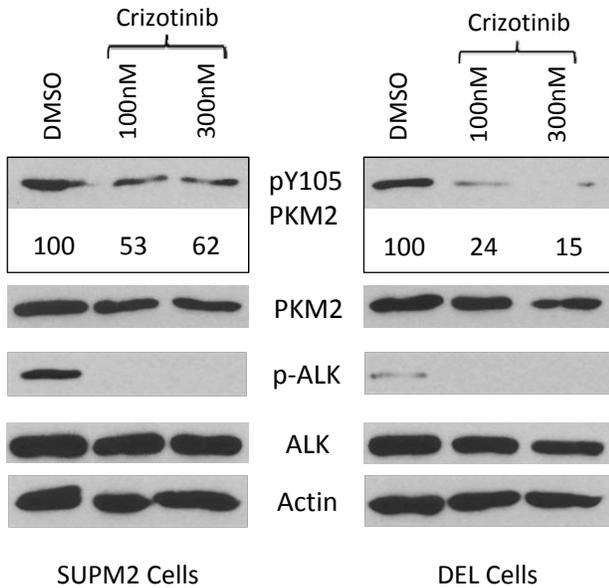
D
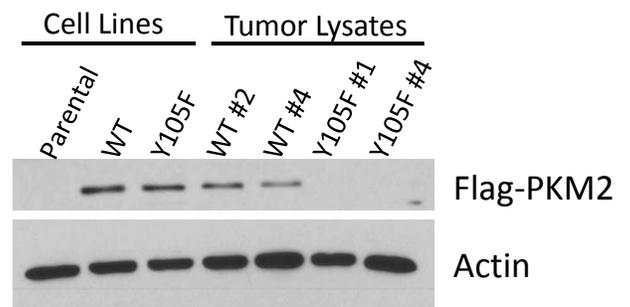
Supplemental Figure 2: pY105-PKM2 is regulated by NPM-ALK and shows differential expression across other T-cell derived lymphomas. A) Phosphoproteomic data from nine different T-cell lymphoma derived cell lines. Data shows the number of peptides and the spec counts for the identification of PKM2 across the different disease entities. B) Western blot analysis of different T-cell lymphoma derived cell lines for phosphorylated and total PKM2. Mycosis fungoides (MF), Sézary syndrome (SS), Natural killer/T-cell lymphoma (NKTL). C) Immunoblots of SUPM2 and DEL cell lysates after treatment with 100nM and 300nM Crizotinib for 16 hours. D) Western blot analysis of cell lines used for tumor injections and tumor lysates from several of the developed tumors from Figure 7C.