INDEX OF SUPPLEMENTARY INFORMATION

Supplementary Methods

Cell lines and cell culture conditions

For cell lines, NK-S1 was established from a previously described NKTL xenograft.[1] NKYS, SNK1, SNK6 and SNT8 were provided by Dr. Norio Shimizu (Tokyo Medical and Dental University). KAI-3 and KHYG1 were purchased from JCRB Cell Bank and HUT78 and NK92 were purchased from ATCC. NK-S1 was cultured in DMEM medium (Life Technologies) supplemented with 10% heat inactivated FBS (Thermo Fischer Scientific) and 10% equine serum (Thermo Fischer Scientific). HANK1 and SNT8 were maintained in Artemis medium-2 (Nihon Techno Service Co. Ltd.) supplemented with 2% human plasma serum (Clontech) and 100 IU/ml or 700 IU/ml IL-2 (MACS), respectively. HUT78 was cultured in IMDM medium (Thermo Fischer Scientific) supplemented with 20% FBS. KAI-3, KHYG1 and NKYS were maintained in RPMI medium supplemented with 10% FBS and 100 IU/ml IL-2. NK92, SNK1 and SNK6 were cultured in RPMI medium supplemented with 10% FBS, 10% equine serum and 100 IU/ml IL-2. All cell lines were routinely checked and negative for mycoplasma contamination as assessed by the MycoAlert™ Mycoplasma Detection Kit (Lonza).

Deep-targeted capture sequencing

The 188-gene capture panel used in this study was designed with SureDesign of Agilent Technologies. Sample preparation was done using SureSelectXT Target Enrichment System for Illumina Paired-End Sequencing Library following manufacturer's instructions and was sequenced on the HiSeq2000 platform (Illumina). Paired-end sequencing reads were aligned to the human reference genome NCBI GRCh37-lite using BWA-MEM and PCR duplicates were removed using SAMtools. Genome Analyzer Toolkit (GATK) was used to recalibrate base quality scores and to realign reads for the callings of single nucleotide substitutions (SNVs) and insertion/deletions (indels). Only reads of mapping quality ≥ 30 and number of mismatches within a 40-bp window ≤ 3 were considered by the Genotyper. SNVs that passed additional quality filters (QD ≥ 3, variant depth ≥ 10) were also retained. Candidate variants were annotated using wAnnovar. For tumor samples without matching normals, the calls underwent stringent filtering to remove probable germline variants. We first removed off-target variants and those that fell in simple repeat regions. We retained variants with a total depth of at least 20 and with a

VAF of at least 3%. We also removed variants which were present in our local bespoke germline database and had a MAF greater than 0.5% in 1000 Genomes database, ExAC16* (Exome Aggregation Consortium), ESP17* (Exome Sequencing Project) database, CG46* database and NCI-60* database. Final list of variants were then curated by manual Integrative Genomics Viewer (IGV) inspection.

Immunohistochemistry (IHC)

pSTAT3 and PD-L1 expression were scored based on the percentage of positive tumor cells and staining intensity. A positive score means that there was expression in at least 1% of tumor cells. The intensity was graded as 0, negative staining; 1+, mild expression; 2+, moderate expression; 3+, strong expression. A Histo-score based on the percentage of stained cells multiplied by the staining intensity was calculated, as previously described.[2]

Whole-transcriptome sequencing (WTS)

RNA from snap frozen tumor tissues and cell lines were extracted and RNA yield and quality were assessed as previously described.[3] Sequencing libraries were prepared using the TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero (Illumina) and WTS was performed on HiSeq 2500 with 2x101 bp read length. RNA reads were aligned using STAR[4] to a combined reference of hs37d5 and EBV-1 in a 2-pass mode. The gene counts were normalized by DESeq2[5] and the significance in differential expression was calculated using two-sided independent t-test

STAT3 Modeling

The crystal structure of mouse STAT3 determined in complex with DNA at 2.25Å resolution (PDB ID: 1BG1)[6] was used as the template, sharing 99% sequence identity for 76% human STAT3 sequence. A total of 500 homology models were generated for human STAT3 sequence covering residues 136-715. These 500 models were evaluated by the DOPE (discrete optimized protein energy) score[7] and the best ranking model was selected. Structural models of human STAT3 homodimer was obtained on the basis of the best ranking monomer model through docking simulation using PatchDock/FireDock. The first 100 solutions from PatchDock[8] were further refined by FireDock,[9] and ranked by a binding score that includes contact energy,

softened van der Waals interactions, partial electrostatics, and additional estimations of the binding free energy. The best ranking dimer model was selected.

Supplementary Figure Legends

Supplementary Figure S1. **Frequently mutated genes in the JAK/STAT signaling pathway identified by targeted capture sequencing within each PTCL and NKTL subtype.** Distribution of mutations in the JAK/STAT signaling pathway in (A) ALCL, (B) CTCL, (C) NKTL and (D) PTCL-NOS.

Supplementary Figure S2. Novel STAT3 p.E616K mutation is sensitive to pharmacologic inhibition in NK-S1 cells. NK-S1 cells were transduced with empty vector, $STAT3^{WT}$ and p.E616K expression vectors. Cell viability assays with DMSO vehicle and Stattic (0.025 μM, 0.05 μM, 0.1 μM, 0.25 μM, 0.5 μM, 1.0 μM, 2.5 μM, 5.0 μM and 10.0 μM) were performed in NK-S1 cells cultured in media without IL-2 for 72 hours. Figure is a representative of 3 independent experiments.

Supplementary Figure S3 Homology model of STAT3 homodimer. Amino acid residues p.D427 and p.E616 are located at the dimerization interface of the SH2 domain of the STAT3 protein.

Supplementary Figure S4. Expression of PD-L1 is induced by STAT3 in Ba/F3 cells. Ba/F3 cells were transduced with empty vector, $STAT3^{WT}$ and p.E616K expression vectors and maintained in media without IL-3 for 6 hours. These cells were harvested for (A) RT-qPCR to detect *PD-L1* mRNA, and (B) flow cytometry to detect membranous PD-L1 expression. PD-L1 mRNA was represented as fold change relative to empty vector and normalized against housekeeping gene NONO. All results are expressed as means ± SD of 3 independent experiments. *$P<0.05$ compared with empty vector, #$P<0.05$ compared with $STAT3^{WT}$.

Supplementary Figure S5. Effect of STAT3 inhibition on *PD-L1* expression in Ba/F3 cells. Ba/F3 cells transduced with p.E616K expression vectors were incubated with DMSO vehicle or Stattic (0.5 and 1μM) for 24 hours. These cells were harvested for (A) Western blotting with antibodies against pSTAT3, total STAT3 and PD-L1, (B) RT-qPCR to detect *PD-L1* mRNA. PD-L1 mRNA was represented as fold change relative to DMSO and normalized against housekeeping gene NONO All results are expressed as means ± SD of 3 independent experiments *$P<0.05$ compared with DMSO vehicle group.

Supplementary Figure S6. Binding of STAT3 to the *PD-L1* gene promoter in Ba/F3 cells. Protein cell lysates harvested from Ba/F3 cells transduced with empty vector, $STAT3^{WT}$ and p.E616K expression vectors were analyzed in the ChIP assay using an anti-STAT3 rabbit polycloncal antibody and primers pairs within the *PDL1* promoter region and a negative control region (NR) outside the promoter region. Enrichment data were analyzed by calculating the immunoprecipitated DNA as a percentage of input DNA for each sample

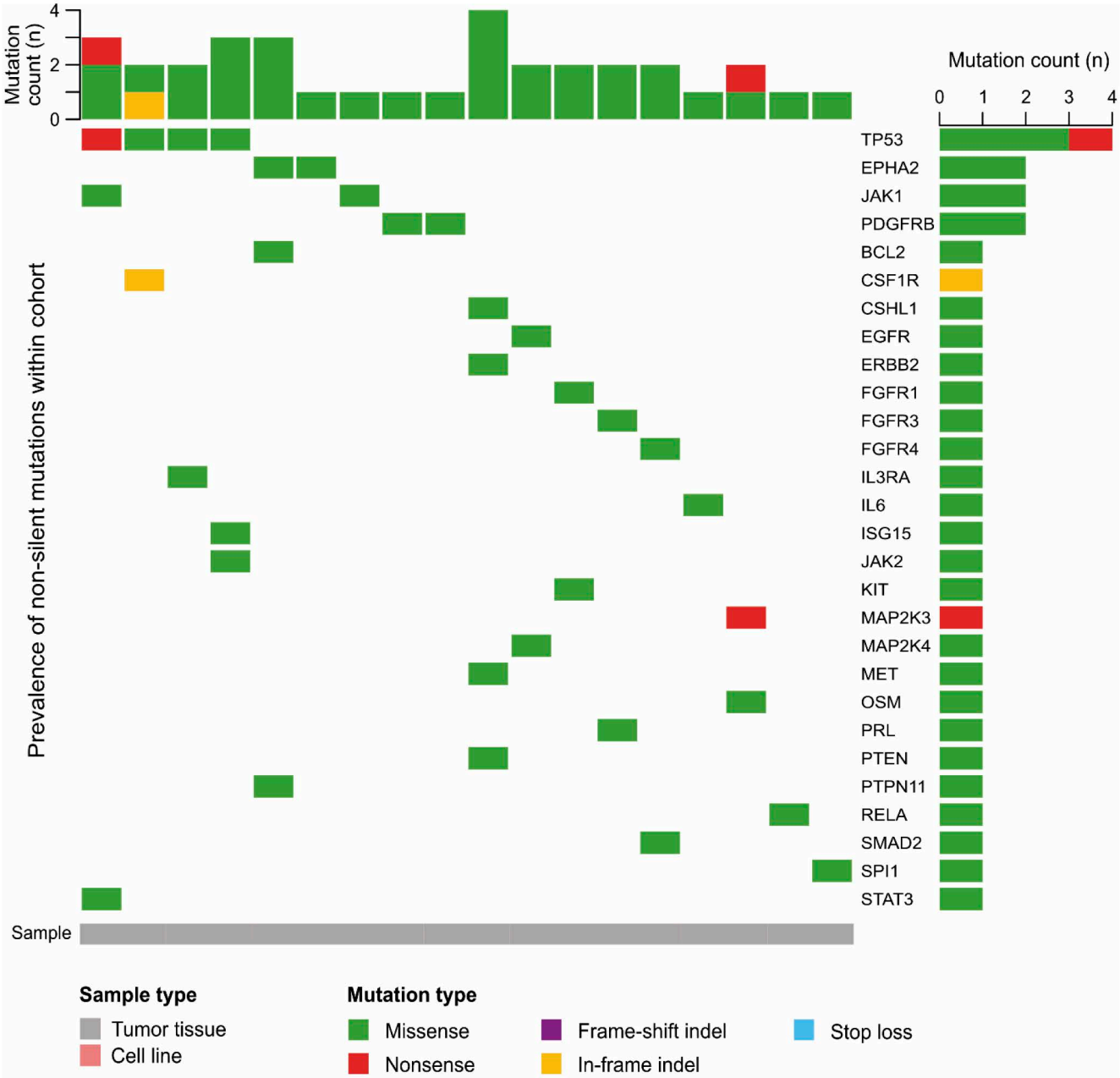
Supplementary Figure S7. Associations between immunohistochemical staining score for PD-L1 with STAT3 phosphorylation, STAT3 mutation status and clinical parameters in NKTL tumor. Data are presented as scatter plots.

Supplementary Figure S8. LMP1 is not associated with STAT3 activation in NKTL tumor and cell lines. (A) Expression of LMP1 mRNA examined by whole-transcriptome sequencing in 8 NKTL cell lines and 12 NKTL tumor tissues. (B) Protein expression of LMP1 in NKTL cell lines at basal state examined by Western blot.

Supplementary References

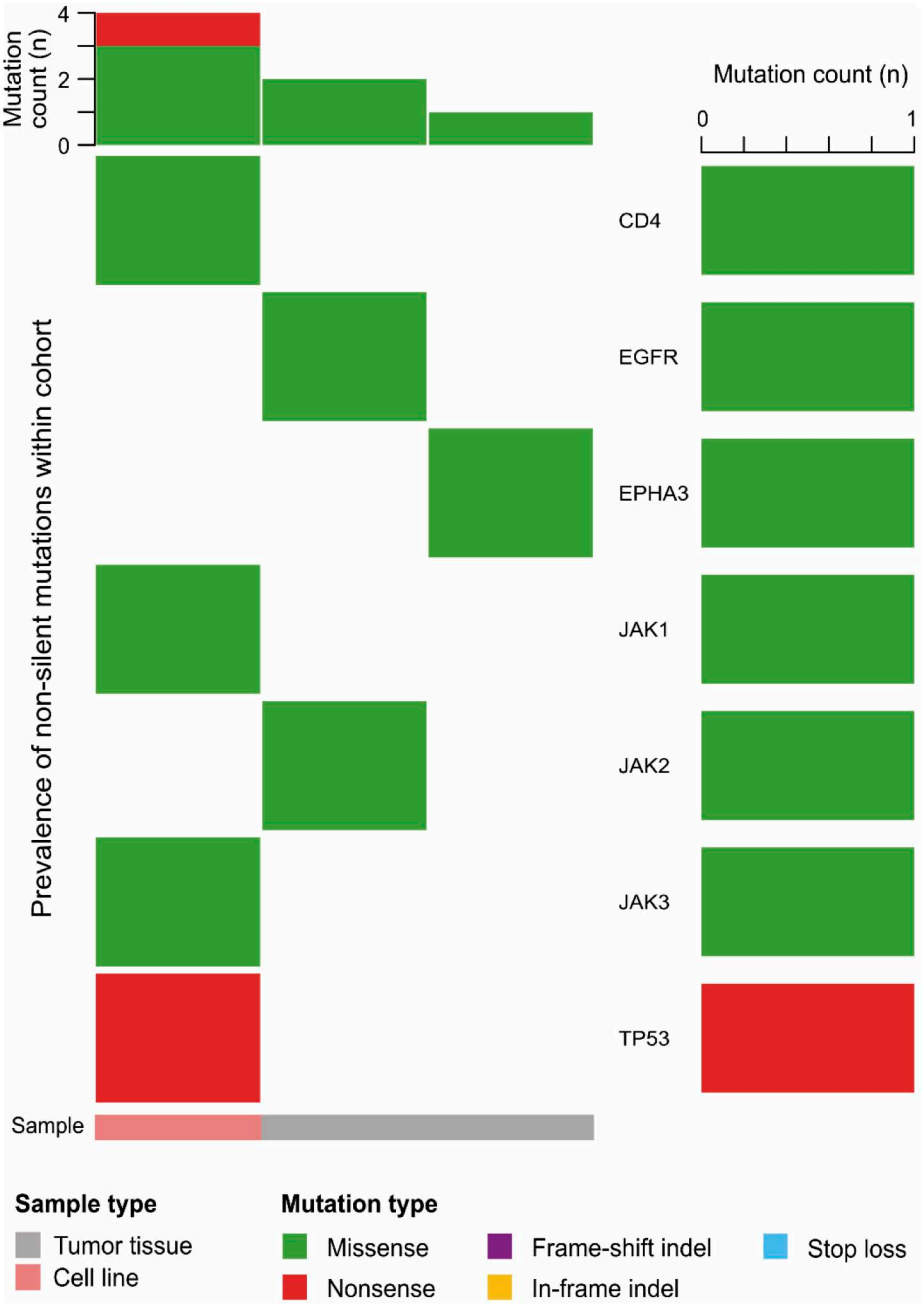
1. Loong SL, Hwang JS, Lim ST, Yap SP, Tao M, Chong TW, et al. An Epstein-Barr virus positive natural killer lymphoma xenograft derived for drug testing. Leuk Lymphoma. 2008;49(6):1161-7.
2. Igarashi T, Teramoto K, Ishida M, Hanaoka J, Daigo Y. Scoring of PD-L1 expression intensity on pulmonary adenocarcinomas and the correlations with clinicopathological factors. ESMO Open. 2016;1(4):e000083.
3. Nairismagi ML, Tan J, Lim JQ, Nagarajan S, Ng CC, Rajasegaran V, et al. JAK-STAT and G-protein-coupled receptor signaling pathways are frequently altered in epitheliotropic intestinal T-cell lymphoma. Leukemia. 2016;30(6):1311-9.
4. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.
5. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology. 2014;15(12):550.
6. Becker S, Groner B, Muller CW. Three-dimensional structure of the Stat3beta homodimer bound to DNA. Nature. 1998;394(6689):145-51.
7. Shen MY, Sali A. Statistical potential for assessment and prediction of protein structures. Protein science : a publication of the Protein Society. 2006;15(11):2507-24.
8. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic acids research. 2005;33(Web Server issue):W363-7.
9. Mashiach E, Schneidman-Duhovny D, Andrusier N, Nussinov R, Wolfson HJ. FireDock: a web server for fast interaction refinement in molecular docking. Nucleic acids research. 2008;36(Web Server issue):W229-32.

Supplementary Figure S1A

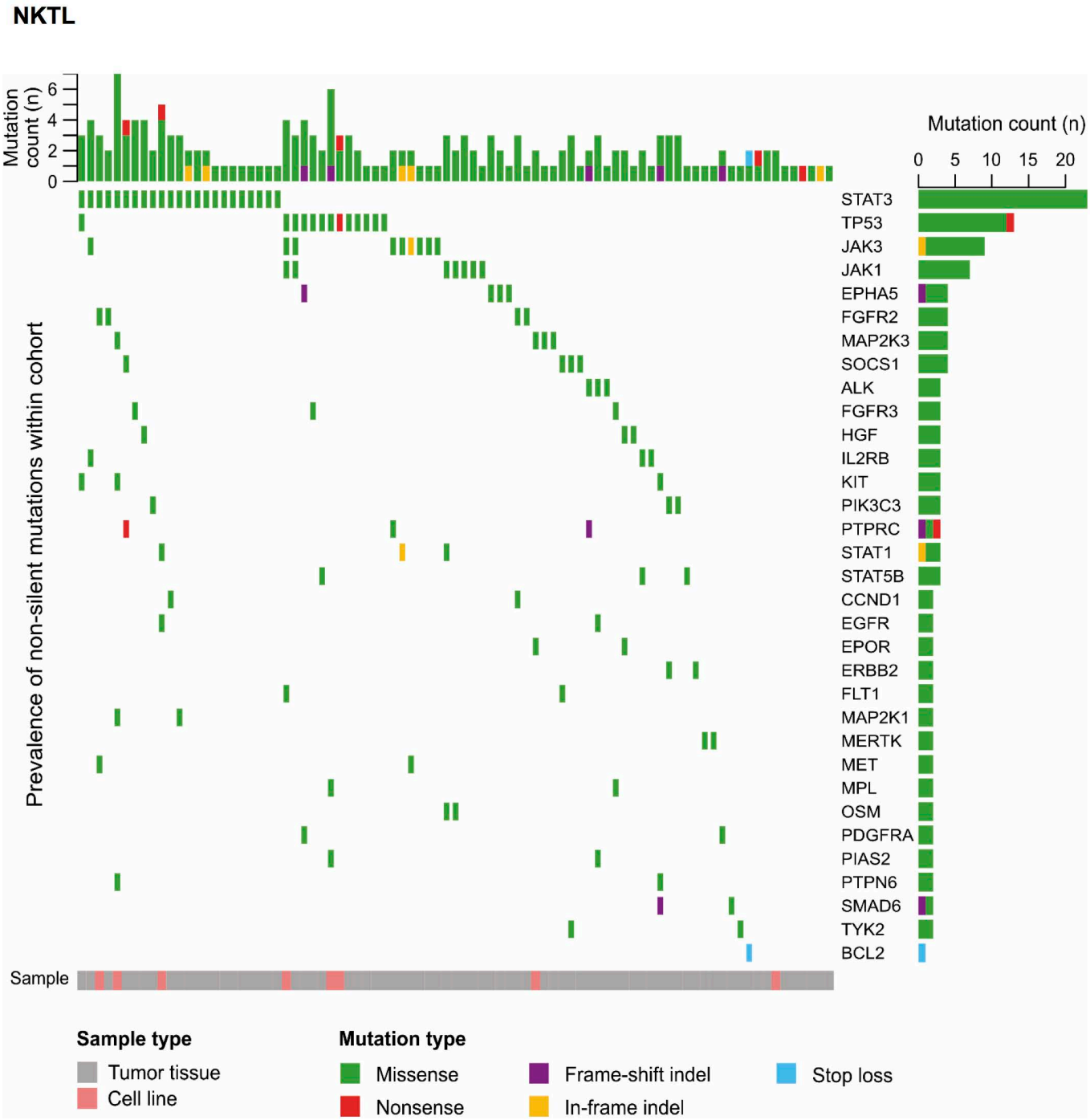


Supplementary Figure S1B

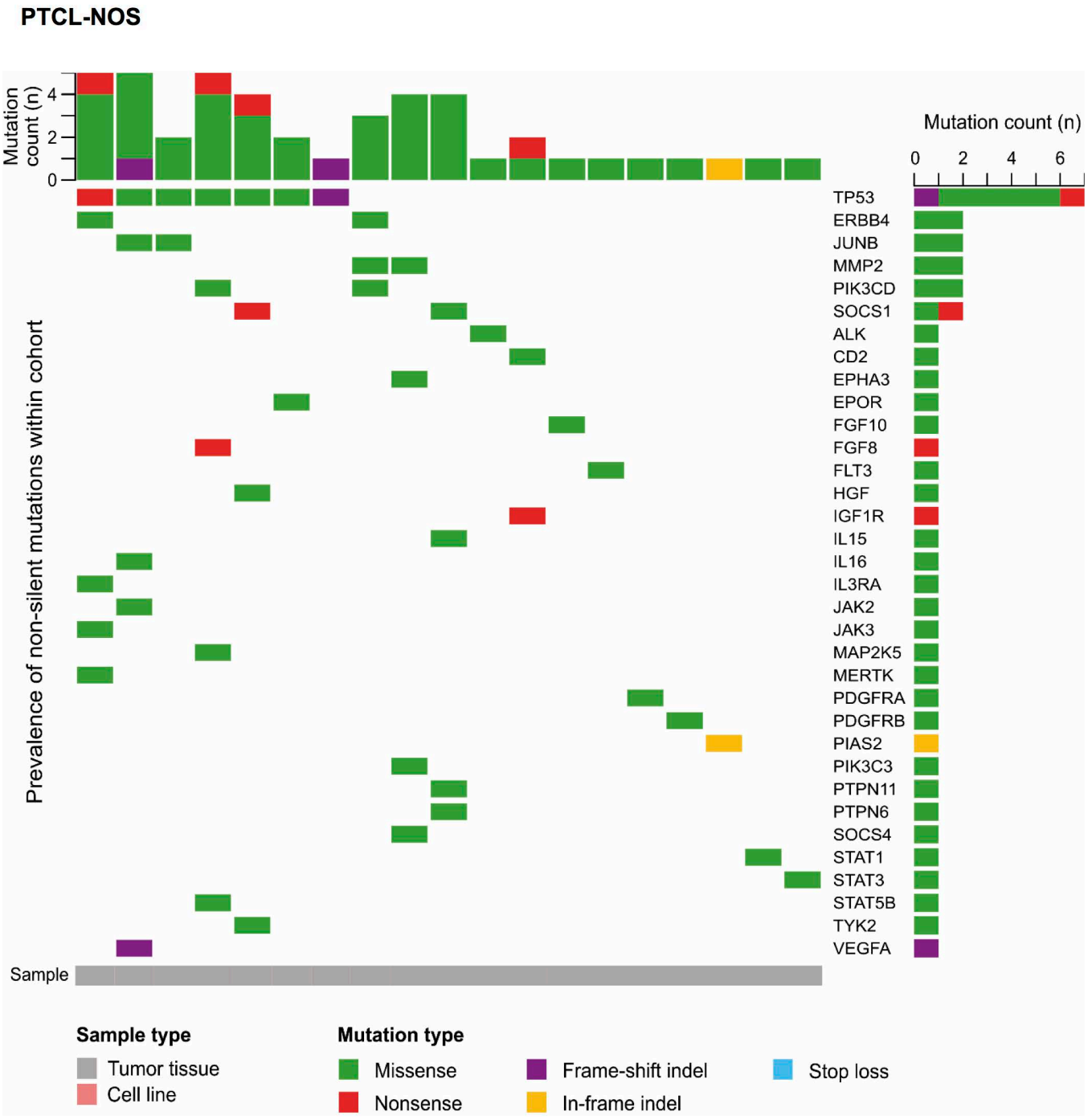
CTCL



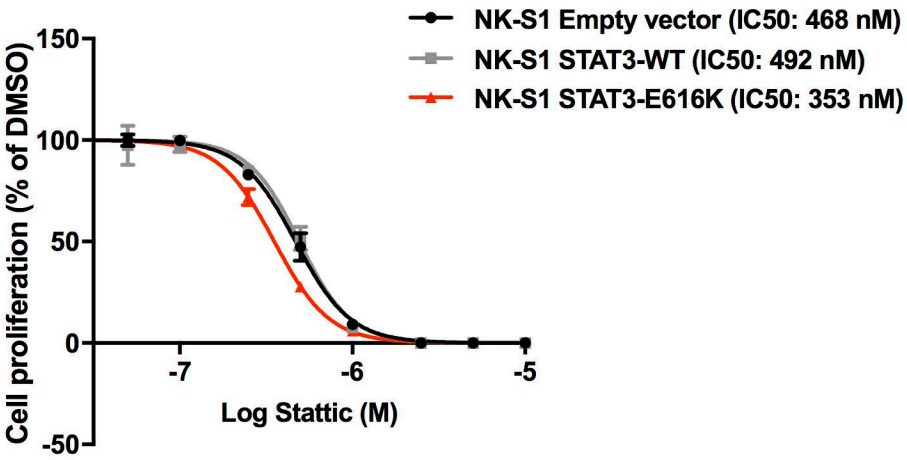
Supplementary Figure S1C



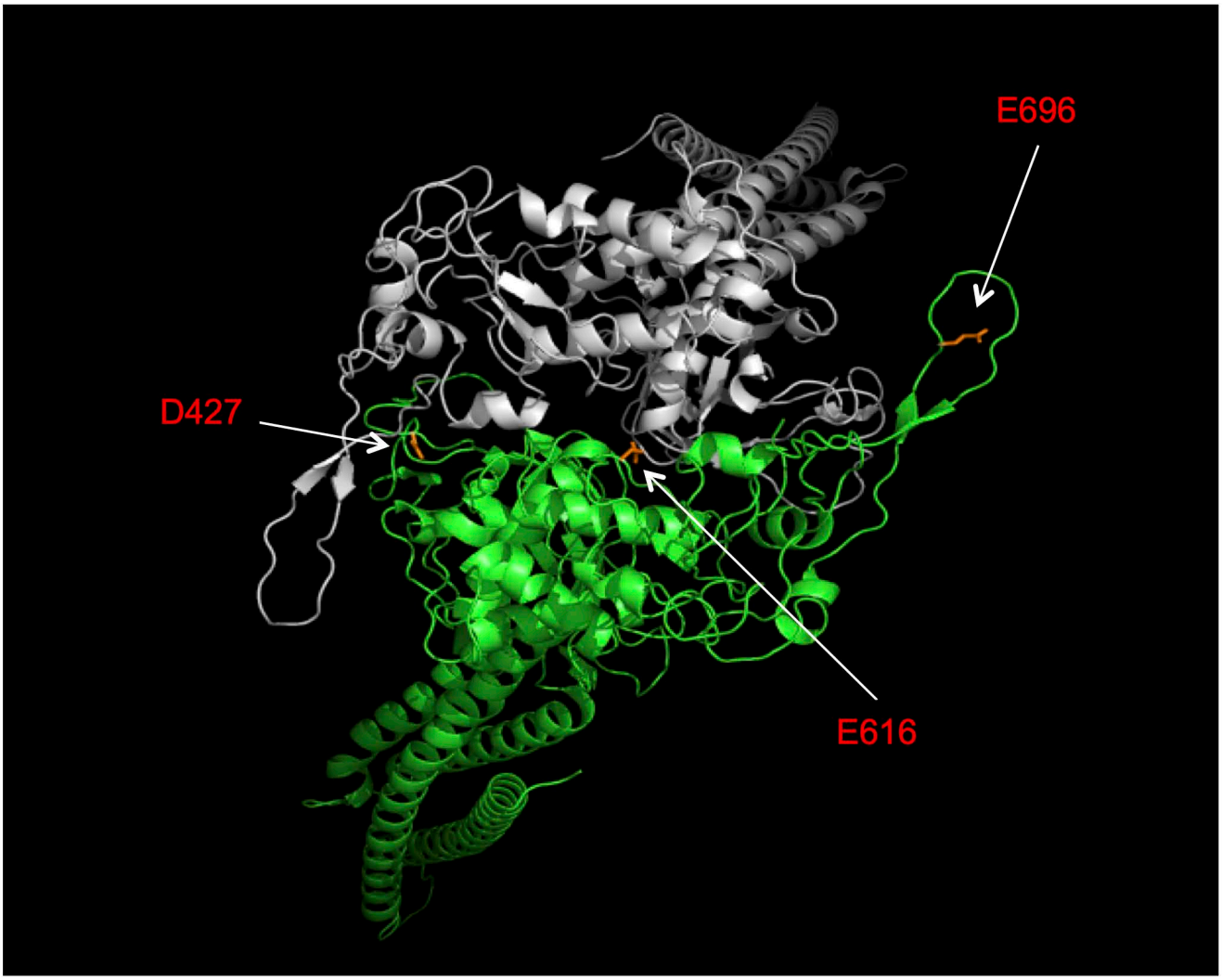
Supplementary Figure S1D



Supplementary Figure S2

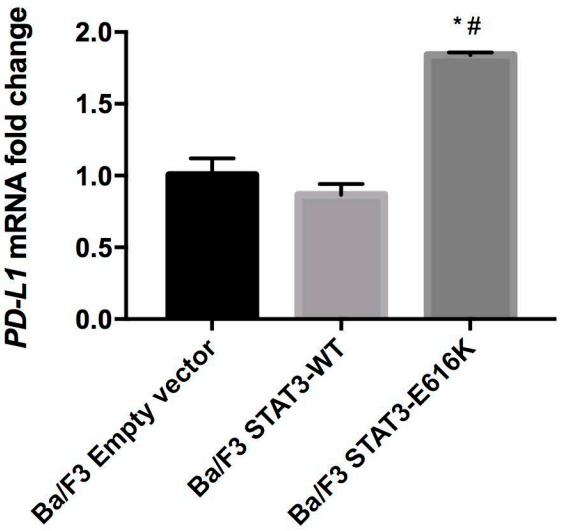


Supplementary Figure S3

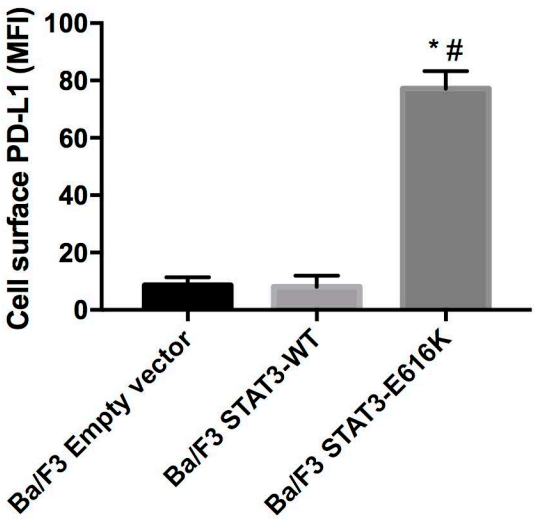


Supplementary Figure S4

A

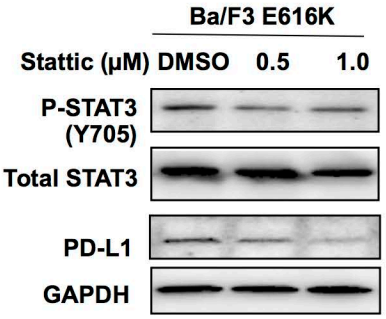


B

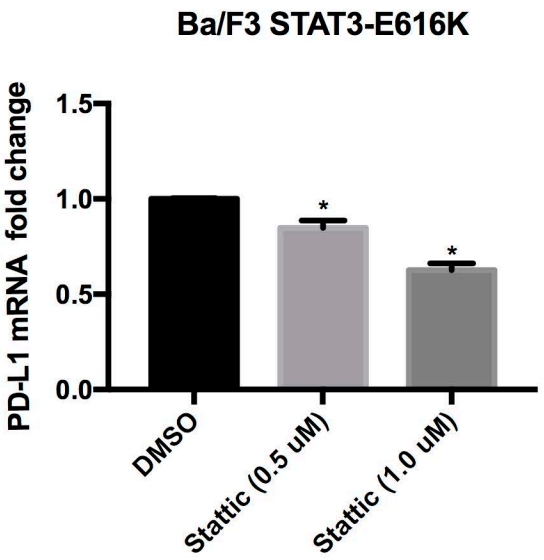


Supplementary Figure S5

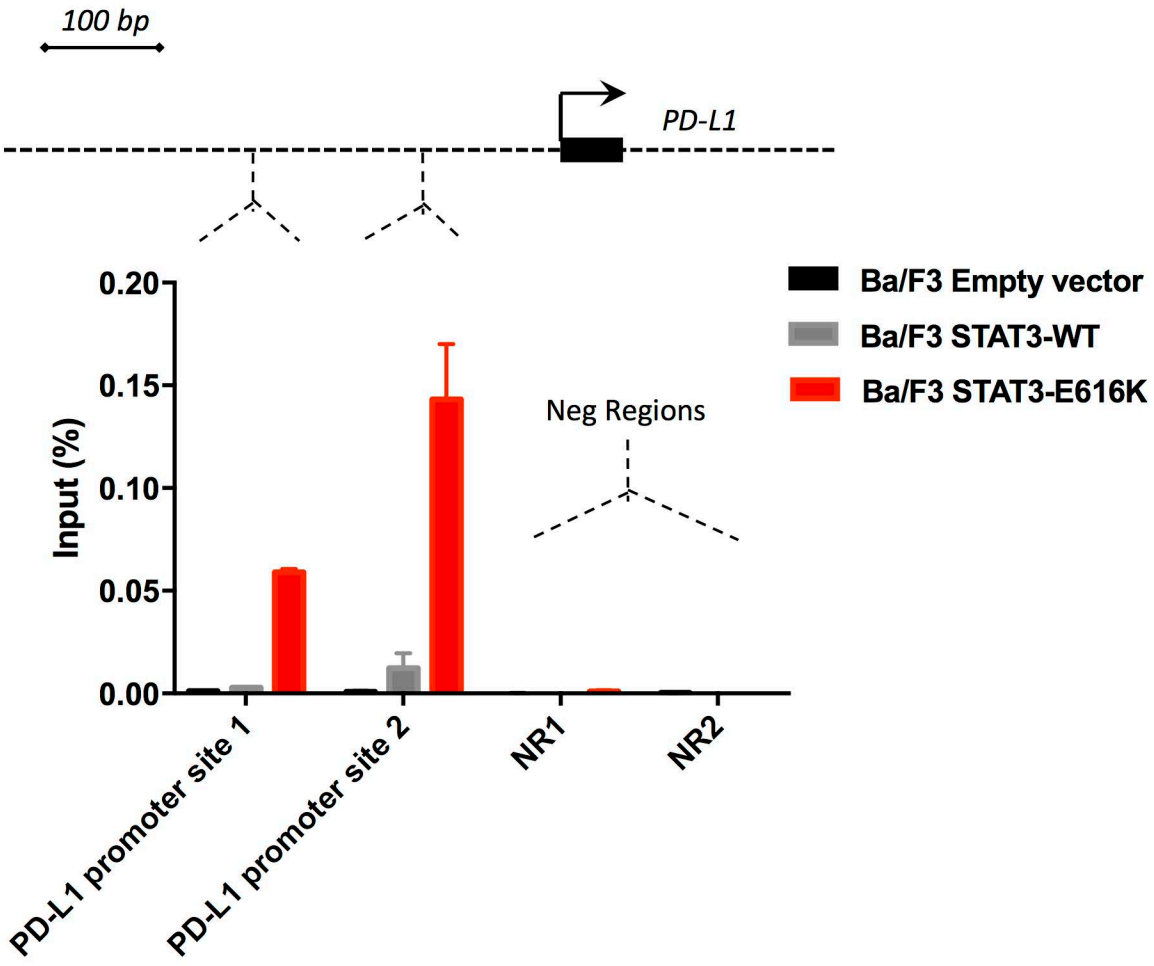
A



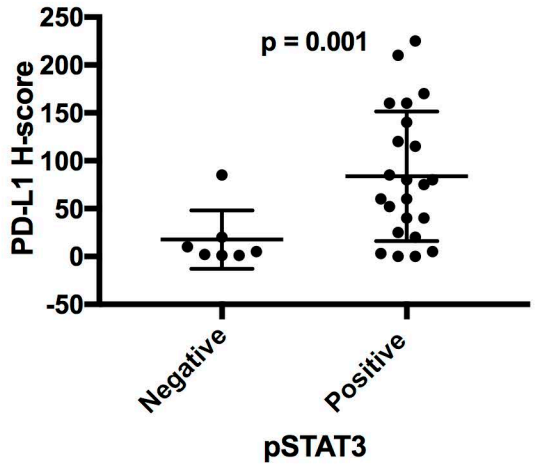
B

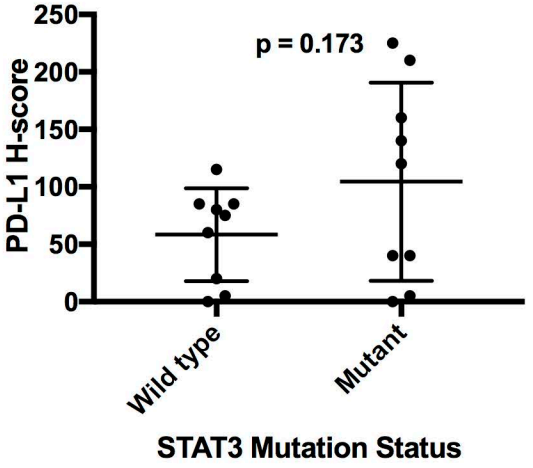


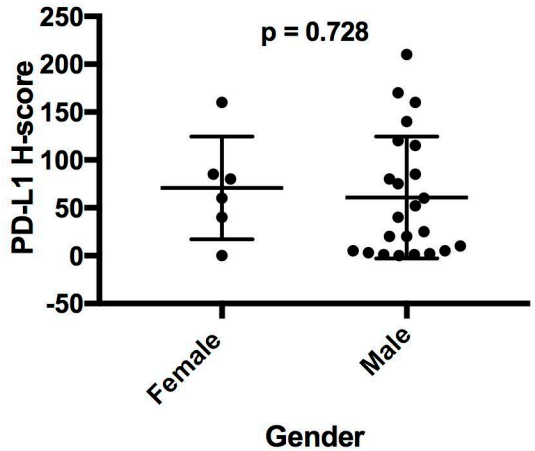
Supplementary Figure S6

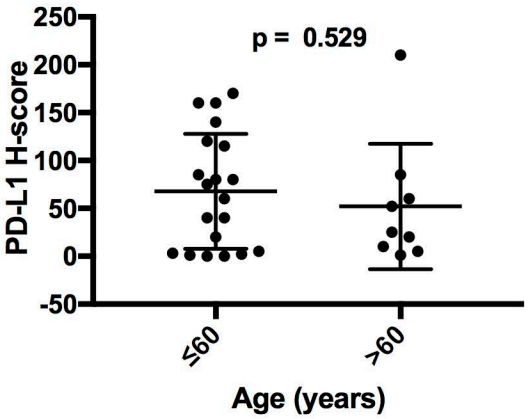


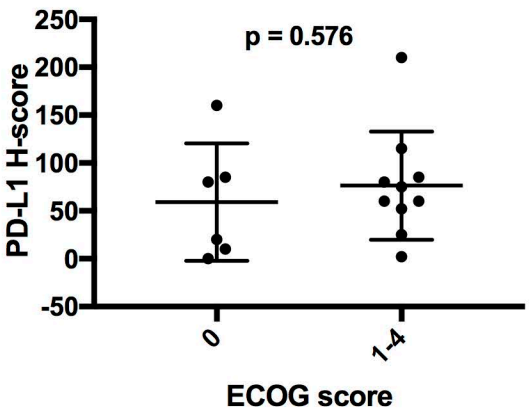
Supplementary Figure S7

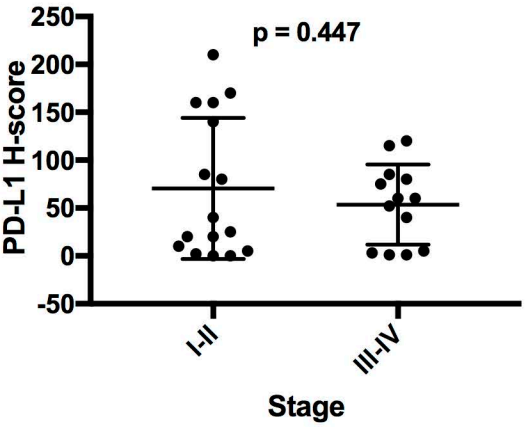


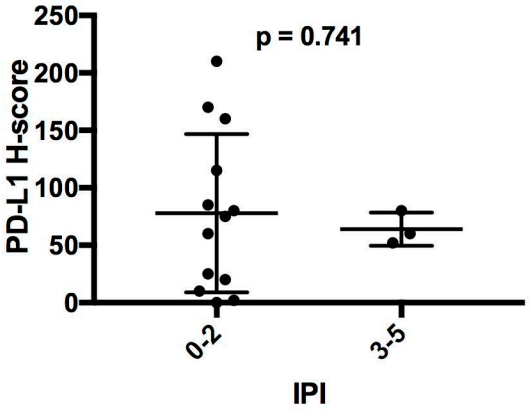












Supplementary Figure S8