DNA mismatch repair defect and intratumor heterogeneous deficiency differently impact immune responses in diffuse large B-cell lymphoma

Supplemental Table 1. Median IHC scores (percentage of cells) of MMR protein expression by IHC staining in control samples (reactive tonsil, colon and prostate cancer) and DLBCL tissues.

	Reactive tonsil controls, germinal center B cells	Reactive tonsil controls, non-germinal center B cells	Colon/breast/prostate cancer tissues	DLBCL tissues
MSH6	30%	10%	30%	70%
MSH2	100%	10%	80%	100%
MLH1	80%	10%	30%	90%
PMS2	90%	10%	80%	90%

Supplemental Table 2. Clinical and molecular features of patients with high MMR protein expression or with MMR gene mutations (genetic dMMR)

	MSH6^{high}	MSH2^{high}	MLH1^{high}	PMS2^{high}	dMMR
Characteristic					
Gender					
Male	101	115	125	116	10
Female	63	73	80	69	8
Age, years					
≤60	64	76	85	74	11
>60	100	112	120	111	7
Stage of disease					
I-II	65	69	82	75	12
III-IV	93	114*	117	105	6
Serum LDH level					
Normal	61	66	75	67	5
Elevated	84	99	107	96	10
ECOG performance status					
0-1	113	131	149	129	16
≥2	27	33	29	33	1
No. of extranodal sites involved					
0-1	114	129	143	135	14
≥2	42	49	52	41	3
IPI risk group					
0-2	79	96	110	98	14
3-5	76**	83	86	79	3
B-symptoms					
Absence	106	120	137	123	11
Presence	52	62	61	55	7
Tumor size					
<5 cm	81	91	106	91	9
≥5 cm	56	62	68	68	7
Cell-of-origin subtype					
GCB	68	76	81	77	11**
ABC	69	84	96*	86	2
Unclassifiable	10	11	11	8	2
p53^{+}					
No	73	92	113	97	10
Yes	73***	79***	74	76**	4
MYC^{+}					
No	37	37	47	42	8
Yes	125***	148***	155***	142***	9
Ki-67^{high}					
No	44	55	60	50	5
Yes	120***	131**	142*	135***	11
BCL2^{high}					
No	76	88	97	82	12
Yes	86*	97*	104	101***	4

CRM1/XPO1^{+}					
No	95	116	131	114	8
Yes	57***	70**	74*	58	10
BCL6^{high}					
No	41	48	61	55	6
Yes	122***	138***	142*	129*	11
FOXP1^{high}					
No	55	59	75	62	11*
Yes	109***	127***	127***	121***	6
MUM1^{+}					
No	59	73	77	70	12
Yes	105***	113**	125***	115***	5
NFKB1 p50^{+}					
No	107***	122***	121	123***	9
Yes	38	48	62	48	5
NFKB2 p52^{+}					
No	124***	138***	144***	140***	9
Yes	25	37	38	31	5

Abbreviations: ECOG, eastern cooperative oncology group; LDH, lactate dehydrogenase; IPI: International Prognostic Index; GCB, germinal center B-cell–like; ABC, activated B-cell–like.

Notes:

Cutoffs for p53^{+}, MYC^{+}, Ki-67^{high}, BCL2^{high}, XPO1^{+}, BCL6^{high}, FOXP1^{high}, MUM1^{+}, p50^{+}, and p52^{+} were >15%, ≥40%, ≥70%, >30%, >50%, ≥80%, ≥40%, ≥20%, and >10%, respectively.

The parameters were compared between MMR protein-high and -low (data not shown) patients and between dMMR and non-dMMR (data not shown) patients using Chi-square test. Significant differences (higher frequency) are marked by asterisks. *, $P \leq 0.05$; **, $P \leq 0.01$, ***, $P \leq 0.001$.

Supplemental Table 3. Enriched Gene Ontology (GO) biological process, molecular function, and cellular components in the gene signatures identified by high MMR protein expression

	MSH6^{high}	MSH2^{high}	MLH1^{high}	PMS2^{high}
Upregulated				
GO Biological Process	Mitotic cell cycle; DNA replication and chromosome cycle; cell cycle; DNA metabolism; S phase of mitotic cell cycle; nucleobase\, nucleoside\, nucleotide and nucleic acid metabolism; DNA replication; cell proliferation; mitosis; M phase of mitotic cell cycle; M phase; nuclear division	Mitotic cell cycle	(DNA metabolism, EASE score 0.0002)	
GO Molecular Function	Nucleic acid binding		Nucleic acid binding	(DNA binding, EASE score 0.0007)
GO Cellular Component	Intracellular; nucleus	Intracellular		
Downregulated				
GO Biological Process	Immune response; defense response; response to biotic stimulus; response to external stimulus; antigen processing; antigen presentation; T-cell activation; antigen presentation\, endogenous antigen; response to pest/pathogen/parasite; antigen processing\, endogenous antigen via MHC class I; lymphocyte activation	Immune response; defense response; response to biotic stimulus; response to external stimulus; response to pest/pathogen/parasite; antigen presentation; antigen presentation\, endogenous antigen; antigen processing\, endogenous antigen via MHC class I; antigen processing; cellular defense response; response to stress; T-cell activation; cell activation; immune cell activation; response to wounding; lymphocyte activation	Defense response; immune response; response to biotic stimulus; response to external stimulus; response to pest/pathogen/parasite; T-cell activation; cell communication	(Defense response, EASE score 0.0005); immune response, EASE score 0.0005)
GO Molecular Function	Signal transducer activity; receptor signaling protein activity; MHC class I receptor activity; receptor activity	MHC class I receptor activity; signal transducer activity; receptor signaling protein activity; receptor activity		
GO Cellular Component	Plasma membrane; integral to plasma membrane; integral to membrane; membrane	Integral to membrane; integral to plasma membrane; plasma membrane; membrane; lytic vacuole; lysosome	Plasma membrane; integral to plasma membrane; integral to membrane	

The enrichment scores for the above categories by Expression Analysis Systematic Explorer (EASE) software were <0.0001.

Supplemental Table 4. Enriched (EASE score <0.0001) GO biological process, GO molecular function, and GO cellular components in the gene signature of dual MutSα^{high} (MSH6^{high}/MSH2^{high}) and in that of MutLα^{high} (MHL1^{high}/PMS2^{high}) expression

	GO Biological Process	GO Molecular Function	GO Cellular Component
MutSα^{high} *vs.* MutSα^{low}			
Downregulated in MutSα^{high}	Defense response; Immune response; Response to biotic stimulus; Response to external stimulus; Response to pest/pathogen/parasite; DNA replication and chromosome cycle; Response to stress; Cellular defense response; T-cell activation; Antigen processing\, endogenous antigen via MHC class I; Humoral immune response; Cell activation; Immune cell activation; Cellular process; Antigen presentation\, endogenous antigen; Humoral defense mechanism (sensu Vertebrata); Lymphocyte activation; Antigen processing; Antigen presentation; Inflammatory response.	MHC class I receptor activity; Signal transducer activity; Defense/immunity protein activity; Coreceptor activity; Receptor activity.	T-cell receptor complex; Immunological synapse.
Upregulated in MutSα^{high}	Mitotic cell cycle; S phase of mitotic cell cycle; Response to wounding; DNA replication; Cell proliferation; Cell cycle; DNA dependent DNA replication.		
MutLα^{high} *vs.* MutLα^{low}			
Downregulated in MutLα^{high}	Immune response; Defense response; Response to biotic stimulus; Response to external stimulus; Response to pest/pathogen/parasite; Response to stress; Response to wounding; Cell communication; Inflammatory response; Signal transduction; Innate immune response; Cellular process; Humoral immune response; Humoral defense mechanism (sensu Vertebrata); Antigen presentation\, endogenous antigen; Cellular defense response; Antigen processing\, endogenous antigen via MHC class I; T-cell activation; Cell	Signal transducer activity; Receptor activity; Defense/immunity protein activity; Cytokine binding; Transmembrane receptor activity; Interleukin binding; Hematopoietin/interferon-class (D200-domain) cytokine receptor activity; Interleukin receptor activity; Growth factor binding.	Integral to membrane; Plasma membrane; Integral to plasma membrane; Membrane; Immunological synapse; T-cell receptor complex.

	adhesion; Antigen presentation.		
Upregulated in MutLα^{high}	Mitotic cell cycle; DNA replication and chromosome cycle; Cell cycle; S phase of mitotic cell cycle; Cell proliferation; Response to DNA damage stimulus; Response to endogenous stimulus; DNA repair; Nucleobase\, nucleoside\, nucleotide and nucleic acid metabolism; DNA replication.	Nucleic acid binding.	Nucleus; Chromosome.

Abbreviations: EASE: Expression Analysis Systematic Explorer (software); GO: Gene Ontology.

Supplemental Figure Legends

Supplemental Figure 1. Lollipop diagrams for visualization of the DNA mismatch repair (MMR) gene mutations detected in the DLBCL cohort and scatter plots for NGS mutation numbers.

(Left) Lollipop plots for MMR gene mutations.

(Right) Numbers of non-silently mutated (MUT) genes on the NGS panel in patients with *MSH6*, *MSH2*, *MLH1*, or *PMS2* mutation (deficient DNA mismatch repair, dMMR) were significantly higher than those in patients without dMMR by 2-tailed unpaired *t* test. Each dot represents one patient. Bars represent the mean numbers.

Supplemental Figure 2. Mutation and expression analysis for four essential DNA mismatch repair (MMR) proteins.

(A) Left scatter plots: DLBCL-NOS cases with any MMR gene mutation (deficient MMR, dMMR) compared with those cases without had increased numbers of non-silently mutated (MUT) genes on the NGS panel.

Stacked bar graph: combined dMMR cases compared with wild-type cases had a significantly higher proportion of patients with ≥6 MUT genes.

Right scatter plots: dMMR was associated with lower PD-1^+ CD4 T cell counts by fluorescent multiplex IHC (and percentage expression: $P = 0.020$ by *t* test and $P = 0.019$ by Mann-Whitney test, figure not shown). Cases with *MSH6* or *MLH1* mutation(s) had a lower mean level of PD-1 expression in CD4 T cells (and in overall T cells) in the overall cohort (and GCB-DLBCL, $P = 0.022$, figure not shown). *MSH6* and *MLH1* mutations (but not *MSH2* and *PMS2* mutations) were associated with a trend of lower PD-1^+ T cell densities and a trend of higher total T cell densities with marginal significance.

(B) Left two stacked bar graphs: Majority of dMMR cases were GCB subtype and assigned with B cell state S1 subtype by Lymphoma Ecotyper software (https://ecotyper.stanford.edu/).

Right three scatter plots: dMMR cases compared with DLBCLs with four wild-type MMR genes had higher mean fractions of plasma cell state S3 (in plasma cells) and Treg cell state S2 (in Tregs) and lower fraction of dendritic cell state S3 (in dendritic cells), which were associated with favorable prognosis in publicly available DLBCL cohorts in a previous study by Steen et al (Ref. 37). Differences in median fractions were even larger (figures not shown). Greenish colors of cell state labels indicate favorable prognostic associations, and red color indicates unfavorable prognostic association shown by Ref. 37. Each dot represents one patient. Bars represent the mean fractions. Significance by 2-tailed unpaired *t* test: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

(C) Only *MSH2* mutations in the GCB subtype of DLBCL-NOS were associated with lower MSH2 protein expression; no significant associations between gene mutations and protein expression were observed for other MMR proteins.

In ABC-DLBCL, cases with 5-70% of cells being MSH6^+ by IHC had a lower mean number of MUT genes on the NGS panel than case with >70% of cells being MSH6^+. MSH2 high expression (cutoff, 100%) and MutSα+ (i.e., dual high MSH6 and MSH2 expression) were associated with higher MUT numbers in ABC and lower MUT numbers in GCB. A similar pattern was observed in MYC$^-$ (<40%) and MYC$^+$ DLBCL-NOS cases, but the associations were not significant.

Each dot represents one patient. Bars represent the mean mutated gene numbers.

Supplemental Figure 3. Immunohistochemistry (IHC) results for four essential MMR proteins in DLBCL.

(A) Histograms showing the distribution of IHC scores of the four MMR proteins in the overall DLBCL cohort. MSH6 had a higher number of patients with low expression compared with the other three MMR proteins.

(B) Pearson correlations between MMR protein expressions. Each dot represents one patient.

Supplemental Figure 4. MMR protein expression was associated with MYC, p53, and Ki-67 expression in DLBCL.

(A) Pearson correlations between MSH6 or PMS2 protein expression and MYC or p53 expression. Each dot represents one patient.

(B) Left three scatter plots: p53+, MYC+, and Ki-67 high IHC scores (cutoffs, >15%, ≥40%, and ≥70%, respectively) were associated with high IHC scores of MSH6 expression by Mann-Whitney test. Bars represent the median levels (differences in mean levels were also highly significant: $P < 0.0001$ by *t* test). Each dot represents one patient.

Right two scatter plots: cases with high MSH6 expression (>70%; low and high expression were denoted with negative and positive signs, respectively) had higher mean IHC scores of MYC and p53 expression by *t* test (results by Mann-Whitney test were also significant). Bars represent the mean levels. Each dot represents one patient.

Supplemental Figure 5. Comparisons of MMR protein expression levels among DLBCL subtypes.

(A) Top line: DLBCL cases were subtyped based on the most abundant CD4 T cell state in each DLBCL tissue; MSH6, PMS2, and MLH1 expression in DLBCL-NOS cases showed differences between Ecotyper-assigned subtypes (all were higher in S1, MSH6 was lower in S2).

Bottom line: DLBCL cases were subtyped based on the most abundant CD8 T cell state in each DLBCL tissue; MSH6, PMS2, and MSH2 expression in DLBCL-NOS cases showed differences between Ecotyper-assigned subtypes (all were higher in S1; also, MSH6 was lower in S2 and S4, and slightly higher in S3).

(B) Comparisons of four MMR protein expression levels among LymphGen genetic subtypes of DLBCL.

Bars represent the mean levels. Each dot represents one patient.

Supplemental Figure 6. Scattered boxplots showing significant differences in absolute counts of 13 immunophenotypic markers evaluated by fluorescent multiplex IHC between patients with and without high expression of MSH6, MSH2, MLH1, or PMS2.

All the four MMR proteins showed significantly negative association with absolute counts of $CD3^+$ T cells, $CD4^+$ T cells, $CD45RO^+$ cells, and $FOXP3^+$ cells. In addition, three MMR proteins (except for MLH1) showed significant association with decreased $CD8^+$ T cells and $PD\text{-}1^+$ cells (also trend of decrease in $MLH1^{high}$ compared with $MLH1^{low}$). Significant differences are marked by asterisks. *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$ by a two-sided Mann-Whitney *U* test.

Supplemental Figure 7. Correlative analyses for high MMR protein expression.

(A) Frequency of T cells was calculated by digitally quantified T cell count divided by the sum of $PAX5^+$, $CD3^+$, $CD68^+$, and $CD56^+$ counts by fluorescent multiplex IHC in each DLBCL sample. High expressionsof MMR proteins were associated with significantly lower frequencies of T ($CD3^+$) cells, CD4 T cells, and CD8 T cells.

(B) In both GCB and ABC subtypes, lower MSH6 expression was associated with significantly higher frequencies of T cells, higher absolute counts per square millimeter of T cells and PD-1^+ T cells, but not $CD68^+$ cells. MutSα^+ (dual high expression of MSH6 and MSH2) was associated with higher percentage of PD-1 expression in CD4 T cells in the GCB subtype with borderline significance.

(C) MutLα^+ (dual high expression of MLH1 and PMS2) was associated with significantly lower T cell abundance and lower PD-1^+ T cell density in both GCB and ABC subtypes (both *t* and *U* test except for PD-1^+ T cell density in ABC being significant only by *U* test), lower PD-1 expression levels in CD4 T cells in the GCB-DLBCL subtype, and a lower mean Treg percentage in CD4 T cells in the ABC-DLBCL subtype. In contrast, in GCB-DLBCL, MutLα^+ was associated with a higher mean Treg percentage in CD4 T cells. MutSα^+ was also associated with a lower mean Treg percentage in CD4 T cells in the ABC-DLBCL subtype with borderline significance.

Note: Each dot in scatter plots represents one patient. Bars represent the mean levels except for the PD-1^+CD3^+ cell density in the first plot of panel C (median densities). *P* values are by unpaired *t* test unless marked as *U* test. In plots for four groups, low and high expression were denoted with negative and positive signs, respectively.

Supplemental Figures 8-10. Comparisons of Ecotyper-defined cell states (fractions per cell type) and ecotypes (fractions of in DLBCL tissues) between cases with and without high MMR protein expression.

Significance by 2-tailed unpaired *t* test: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$.

Supplemental Figure 11. Correlative and prognostic analyses in MYC^- and MYC^+ subsets of the DLBCL-NOS cohort.

(A) Top line: In both MYC^- and MYC^+ subsets of the DLBCL cohort, $MSH6^{high}$ expression was associated with significantly lower proportion of T cells among total counts of $PAX5^+$, $CD3^+$, $CD68^+$, and $CD56^+$, and higher MYC and p53 IHC scores by unpaired *t* test (*U* test with medians shown in plots are in Figure 6). In the MYC^- DLBCL subset, $MSH6^{high}$ expression was associated with significantly poorer overall survival (OS) independent of GCB/ABC subtypes. In MYC^+ patients (in which $MSH6^{high}$ did not show prognostic effect), low $CD3^+$ T cell densities were associated with poorer OS with marginal significance; however, for patients without low $CD3^+$ T cell densities, T cell densities did not show further significant prognostic effects in MYC^+ patients.

Bottom line: the adverse prognostic effect of $MSH6^{high}$ expression in MYC^- DLBCL did not appear to be caused by decreased $CD3^+$ T cells, because high T cell densities and PD-1^+ T cell densities actually showed significant adverse prognostic effects in overall MYC^- patients, $MSH6^{low}$ MYC^- patients (Figure 6B), and $MSH6^{high}$ MYC^- patients. High $CD68^+$ cell densities were also associated with poorer OS in MYC^- DLBCL with borderline significance.

(B) Analyses for PMS2 expression in MYC^- and MYC^+ subsets. Only in the MYC^- subset, was $PMS2^{high}$ associated with significantly poorer OS. In both MYC^- and MYC^+ subsets, $PMS2^{high}$ compared with $PMS2^{low}$ patients had significantly higher levels of T cells, PD-1^+ T cells (*P* value was significant in the MYC^+ subset only), and MYC expression.

Note: in the two scatter plots for four groups on the right, low and high expression were denoted with negative and positive signs, respectively.

Supplemental Figures 12-13. Comparisons of Ecotyper-defined cell states and ecotypes in DLBCL tissues between MSH6^{low} and MSH6^{high} patients separately in MYC^{-} and MYC^{+} subsets of the DLBCL-NOS cohort.

Only in MYC^{-} patients, low MSH6 expression (≤70%) was associated with decreased B cell state S1 (GC-like) and a non-significant trend of increased CD4 T cell state S3 (naïve) fraction.

Only in MYC^{+} patients, low MSH6 expression was associated with decreased abundance of LE9 ecotype, increased dendritic cell state S1, decreased dendritic cell state S4, decreased NK cell state S1, and increased endothelial cell state S2.

Significance: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$.

Supplemental Figure 14. Correlative analysis for *HDAC* gene expression.

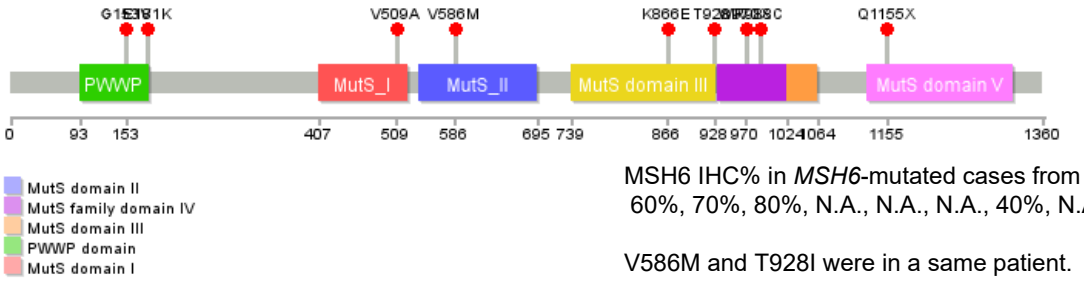
(A-C) Pearson correlations between *HDAC1*/*HDAC3* gene expression and expression of four MMR genes (positive correlations, panels A and B) and between *HDAC10* gene expression and expression of four MMR genes **(**negative correlations, panel C) in GSE31312.

(D) Scatter plots to show that GCB-DLBCL cases with ≥6 non-silently mutated (MUT) genes by targeted NGS had significantly lower *HDAC1* and *HDAC3* expression in GSE31312 than GCB-DLBCL cases with 0-5 MUT genes. One dot represents one patient. Bars represent mean values.

(E) Heatmap for four MMR gene transcription levels in triplet DLBCL cell samples treated with tucidinostat/venetoclax alone or in combination. NC: normal control; CHI, chidamide (tucidinostat); VEN: venetoclax; COM, combination (chidamide plus venetoclax).

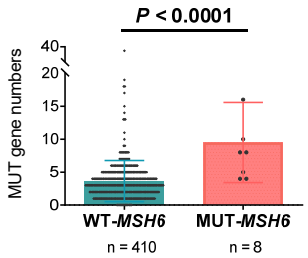
Suppl. Figure 1

MSH6 mutations

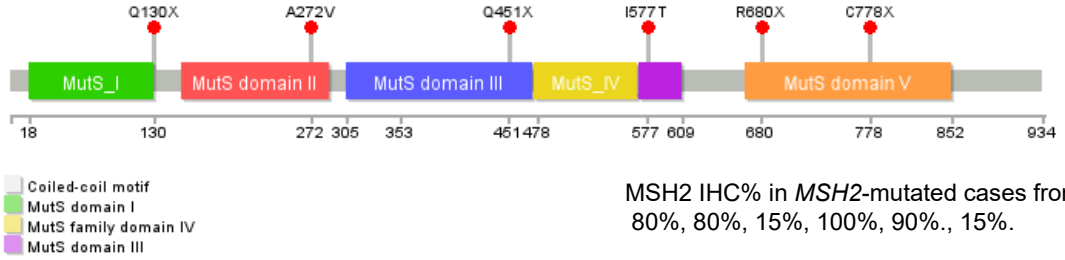


MSH6 IHC% in *MSH6*-mutated cases from left to right:
60%, 70%, 80%, N.A., N.A., N.A., 40%, N.A., 100%.

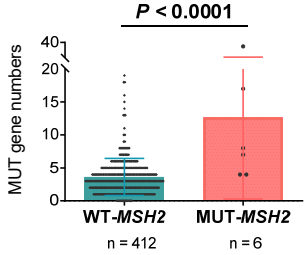
V586M and T928I were in a same patient.



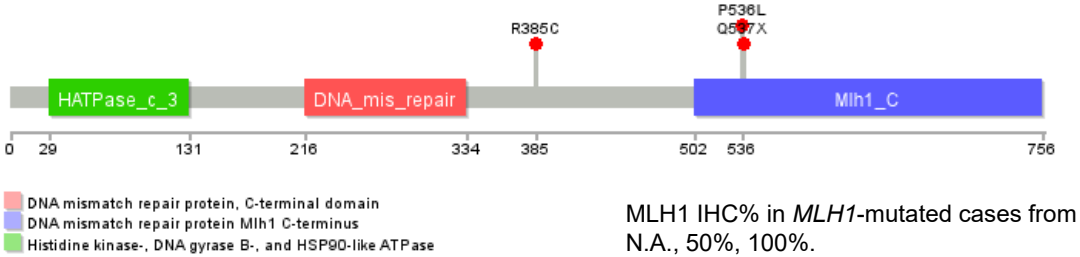
MSH2 mutations



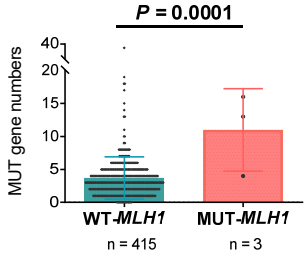
MSH2 IHC% in *MSH2*-mutated cases from left to right:
80%, 80%, 15%, 100%, 90%., 15%.



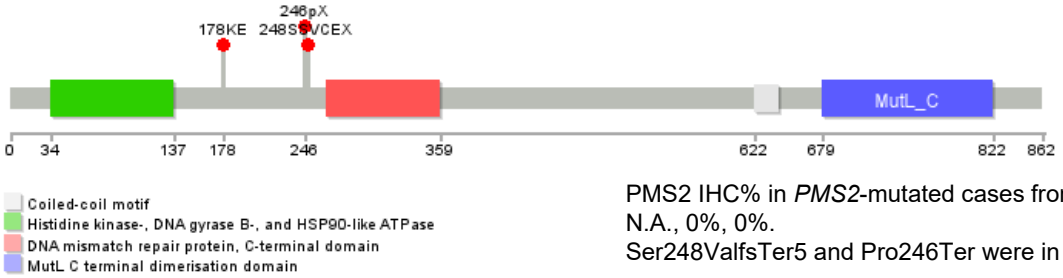
MLH1 mutations



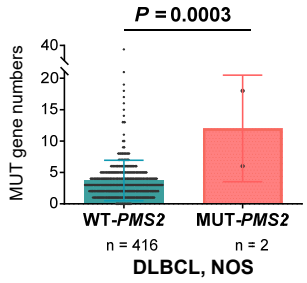
MLH1 IHC% in *MLH1*-mutated cases from left to right:
N.A., 50%, 100%.



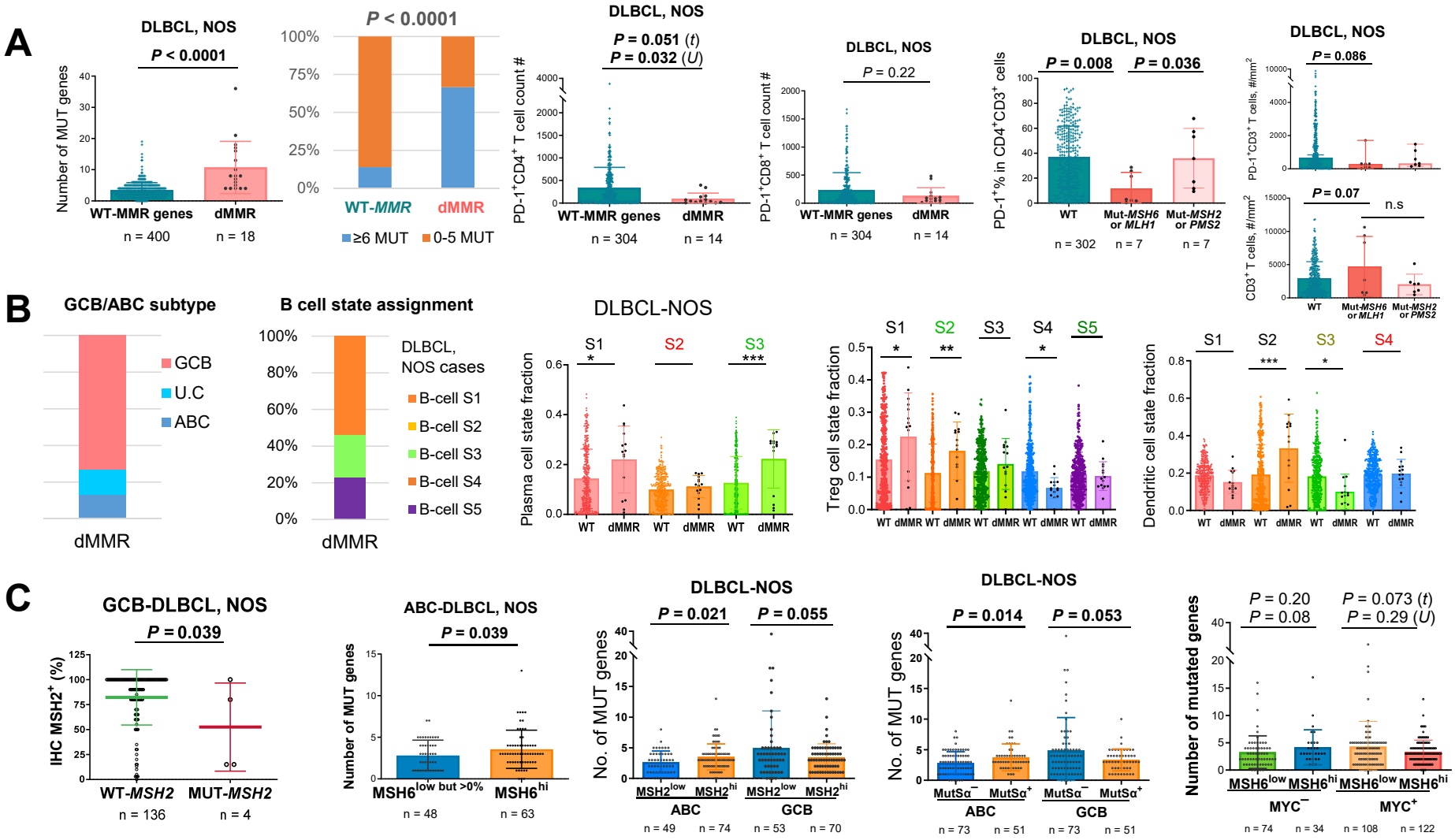
PMS2 mutations



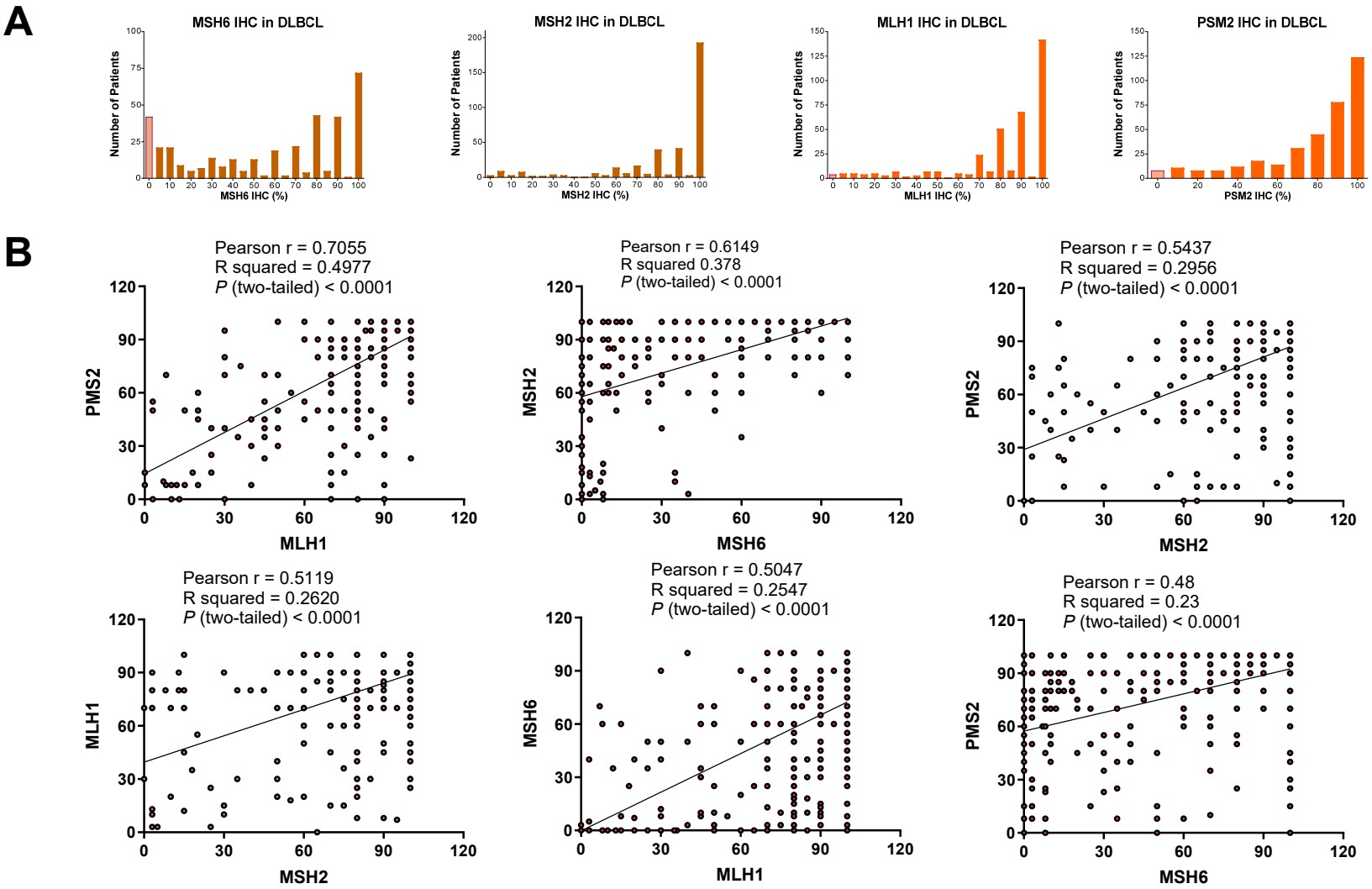
PMS2 IHC% in *PMS2*-mutated cases from left to right:
N.A., 0%, 0%.
Ser248ValfsTer5 and Pro246Ter were in a same patient.



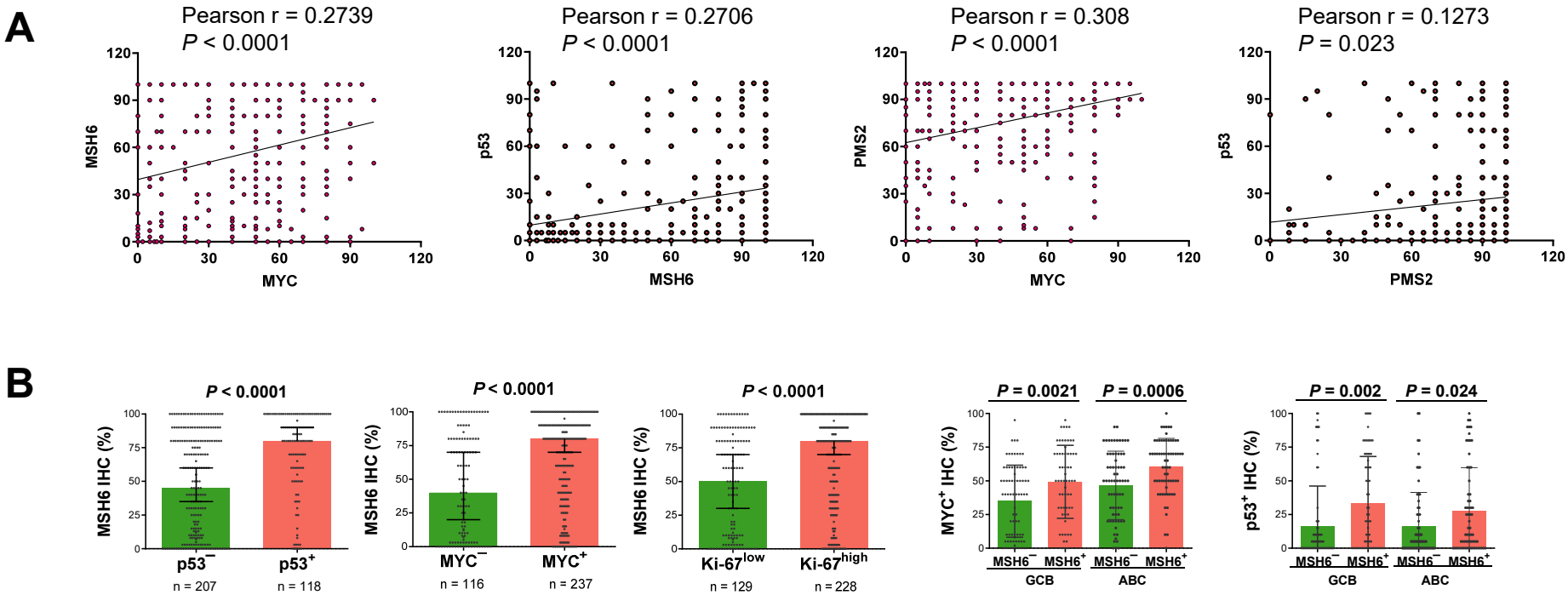
Suppl. Figure 2



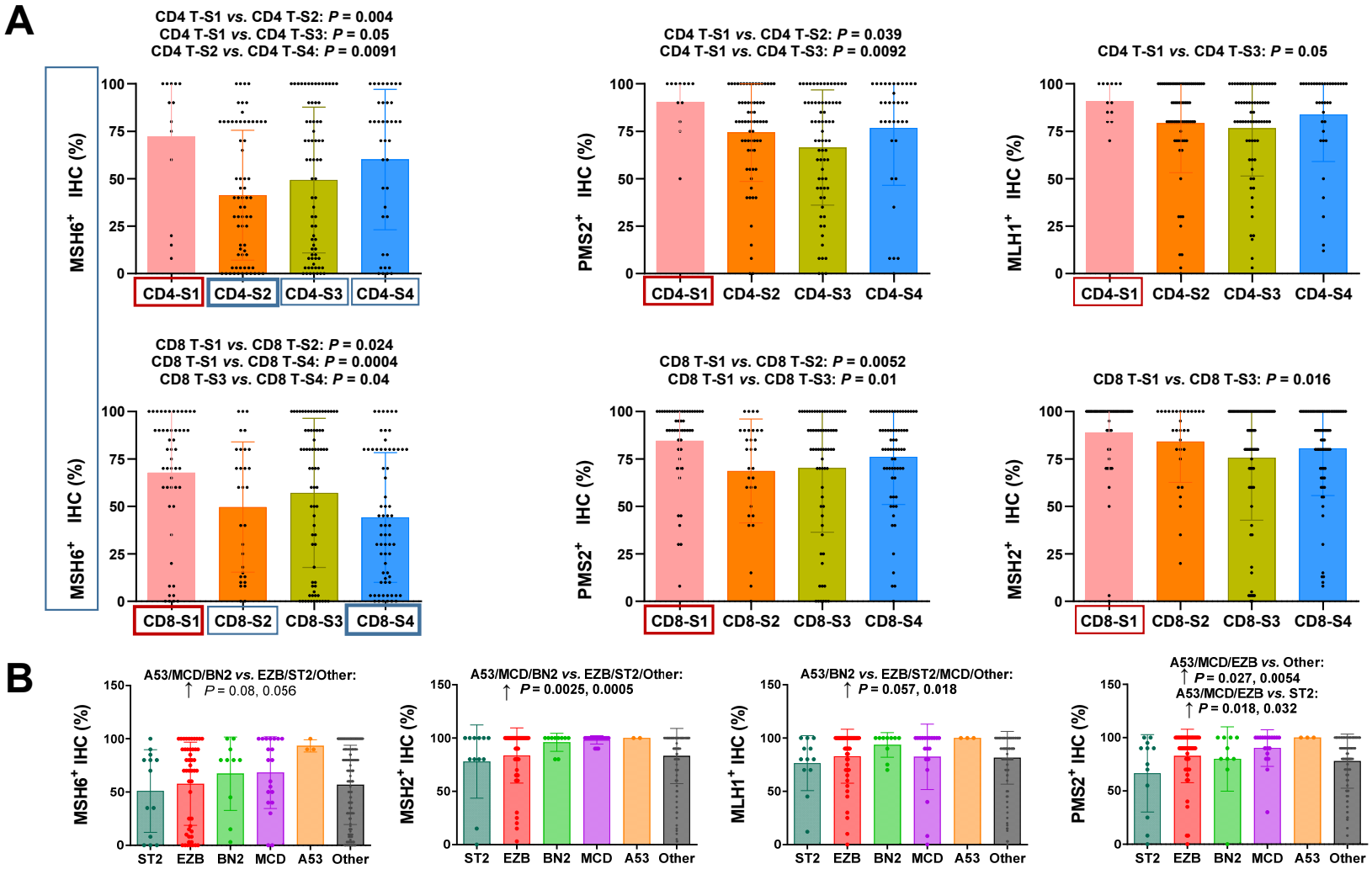
Suppl. Figure 3



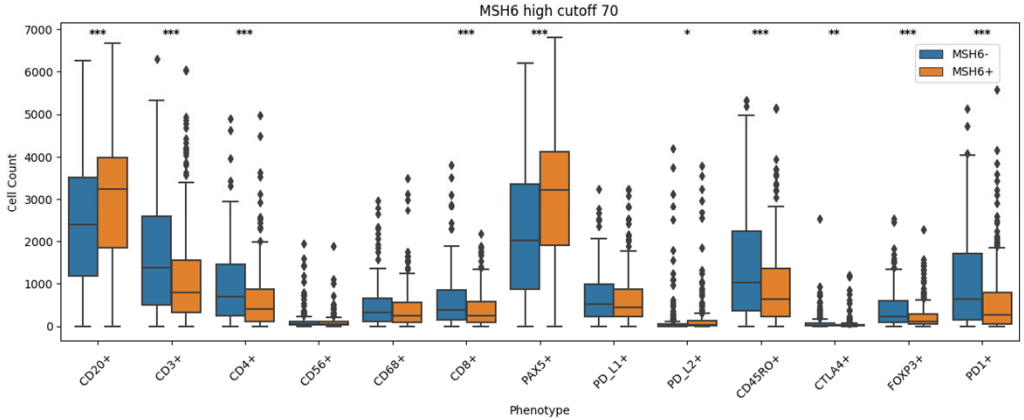
Suppl. Figure 4

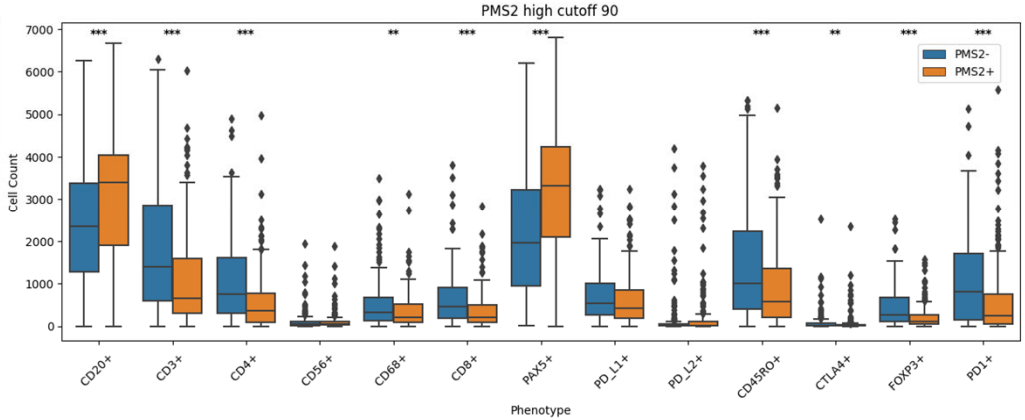


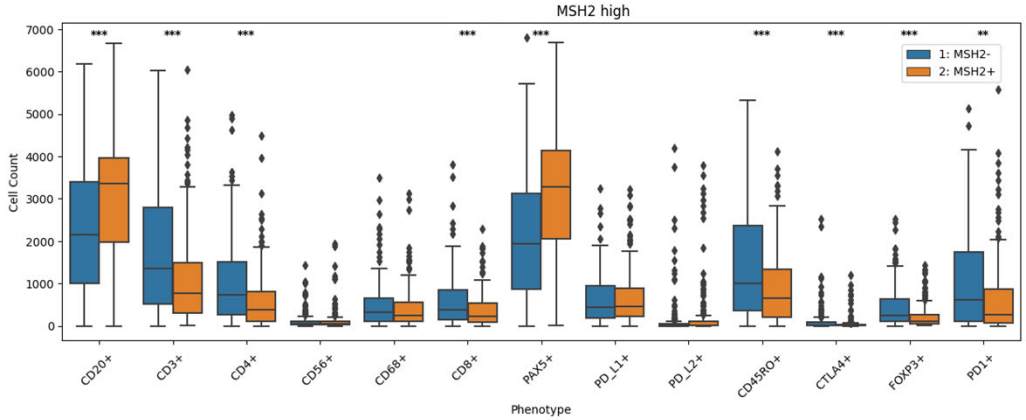
Suppl. Figure 5

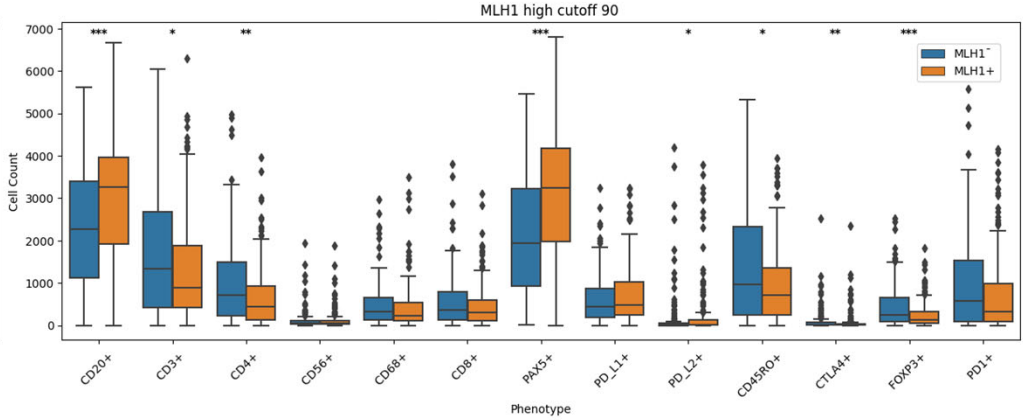


Suppl. Figure 6

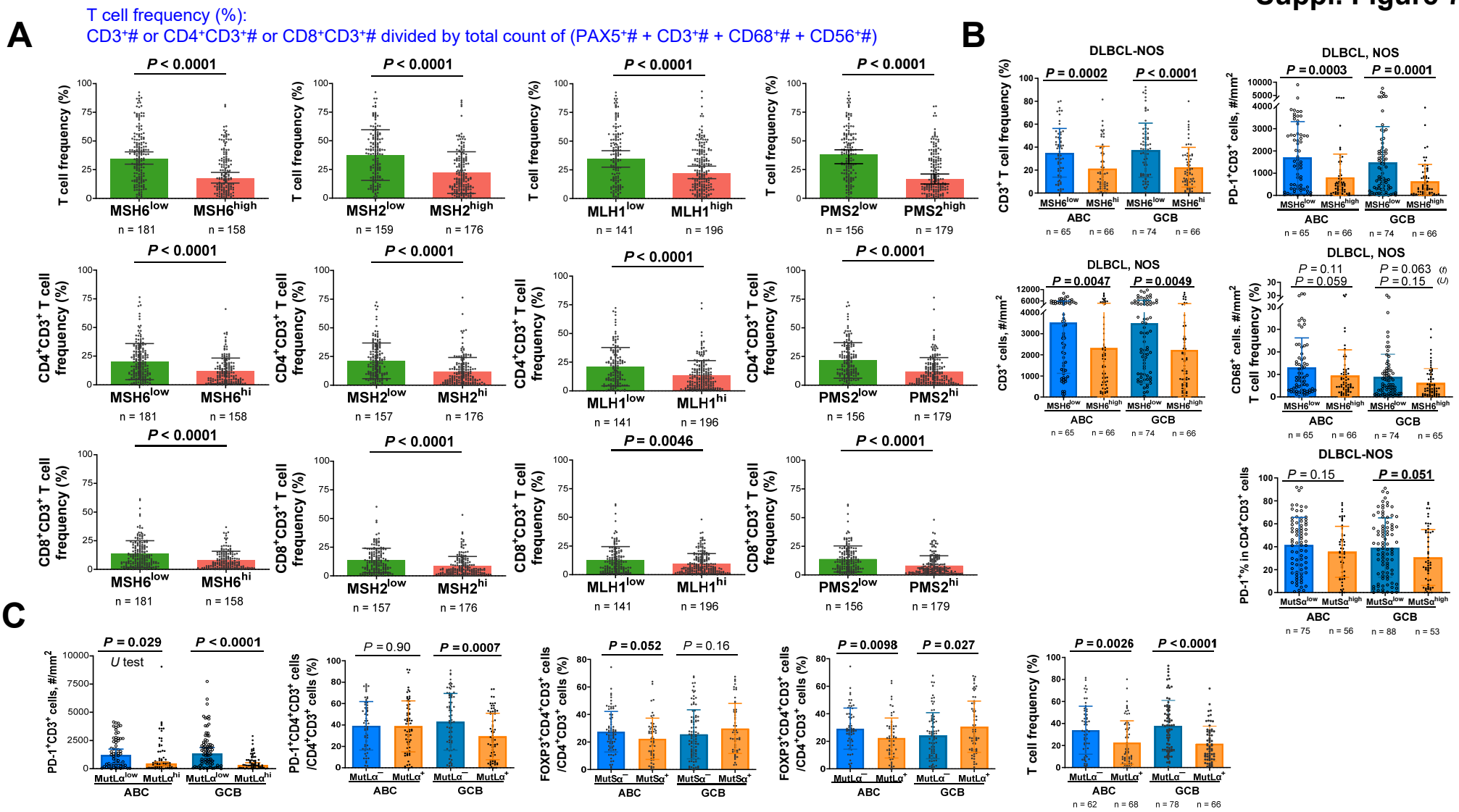




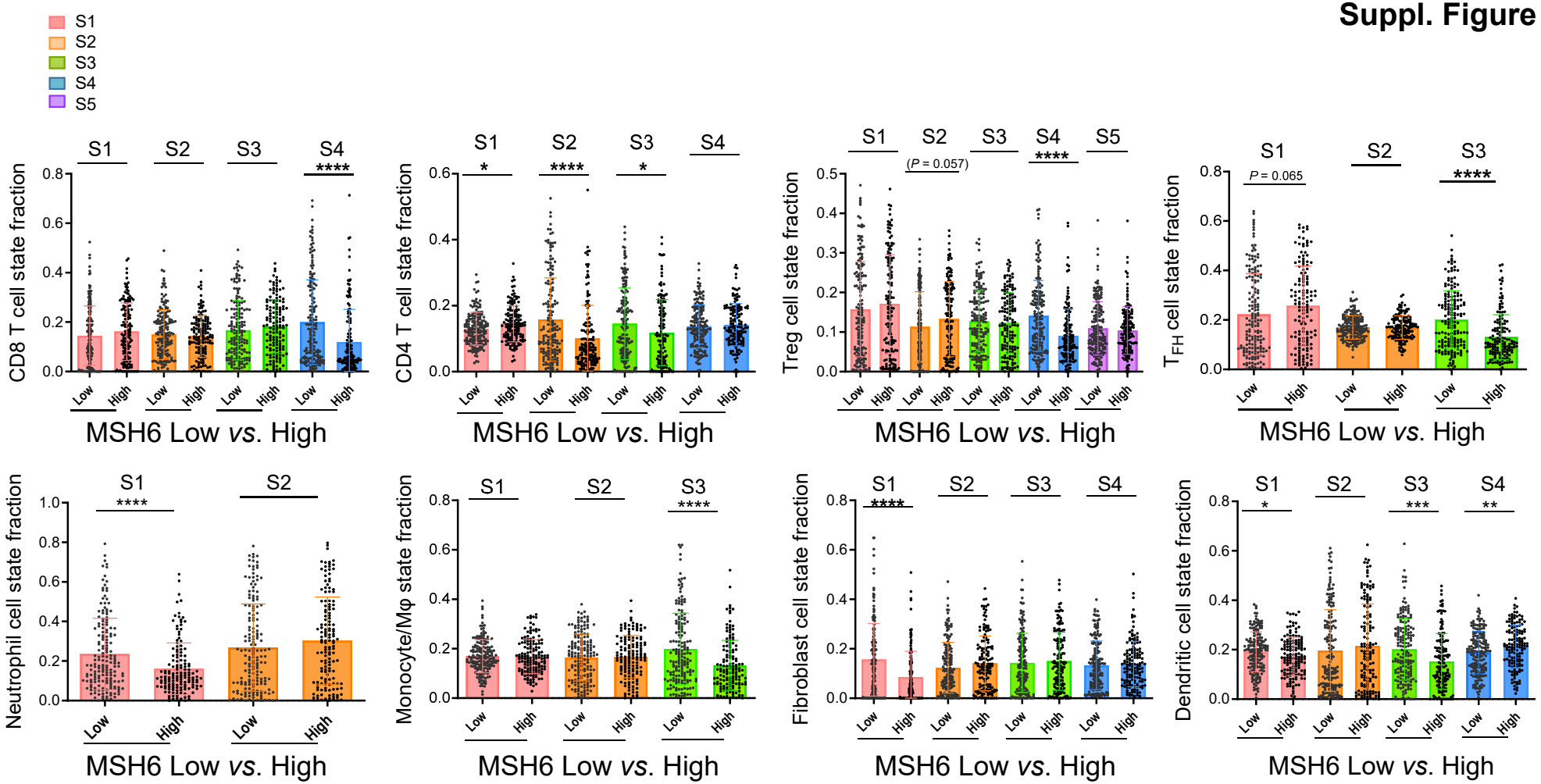




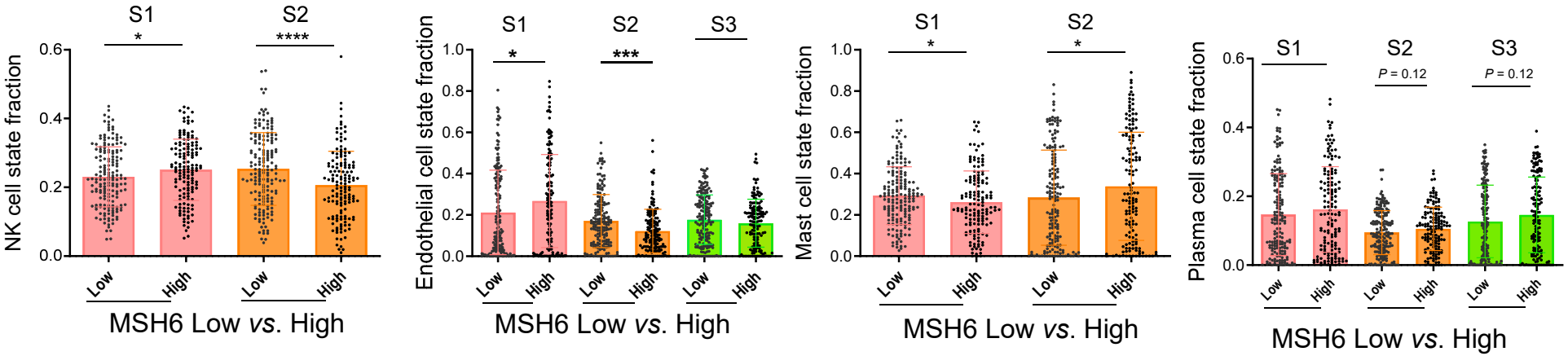
Suppl. Figure 7

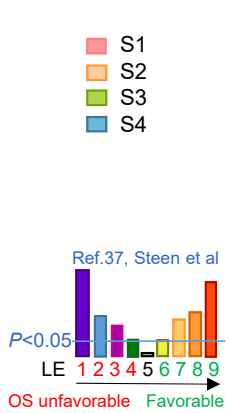


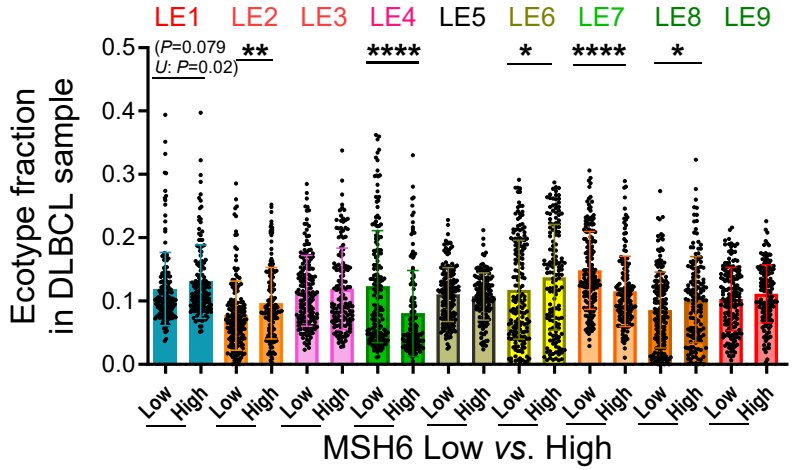
Suppl. Figure 8



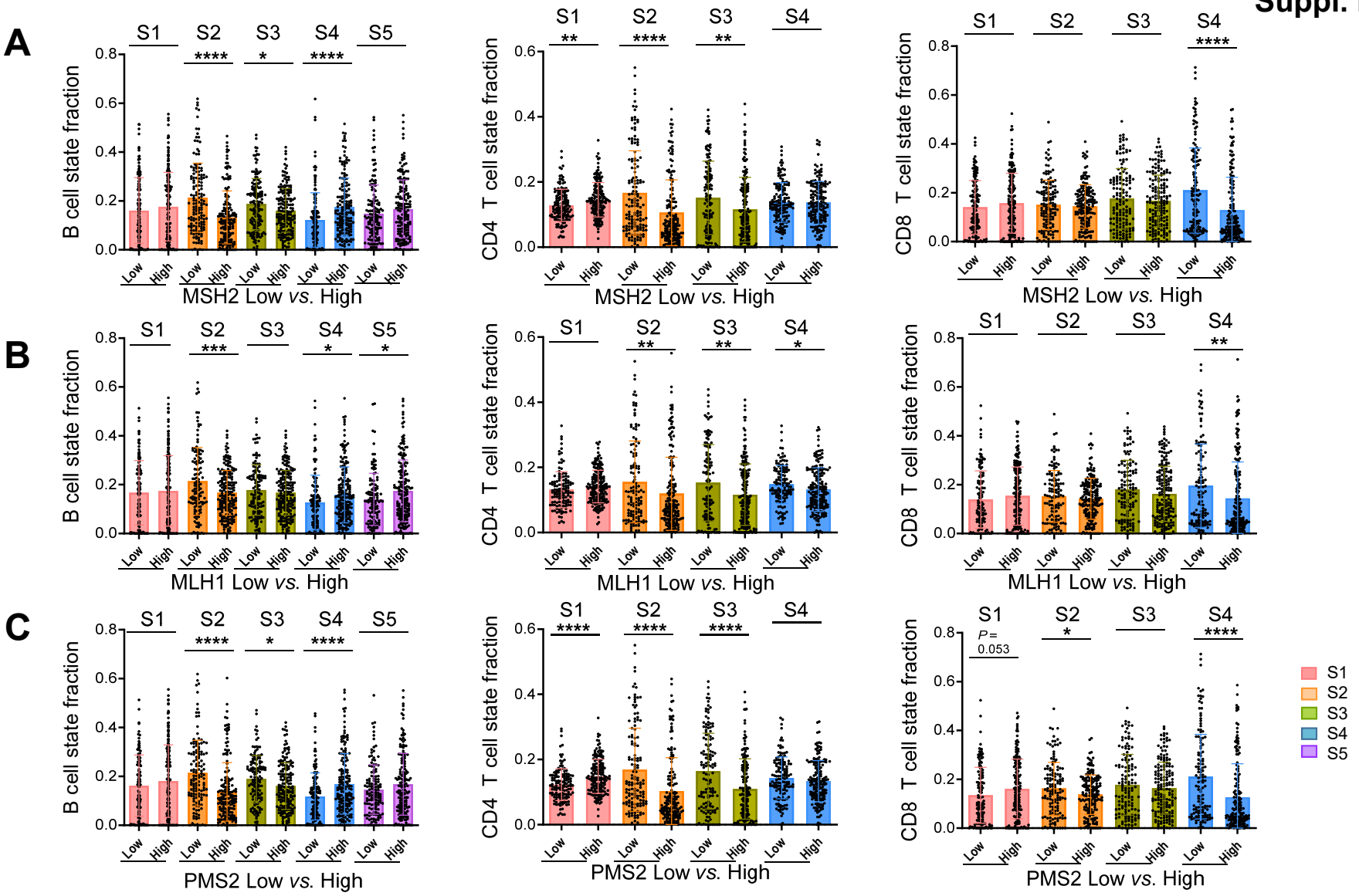
Suppl. Figure 9



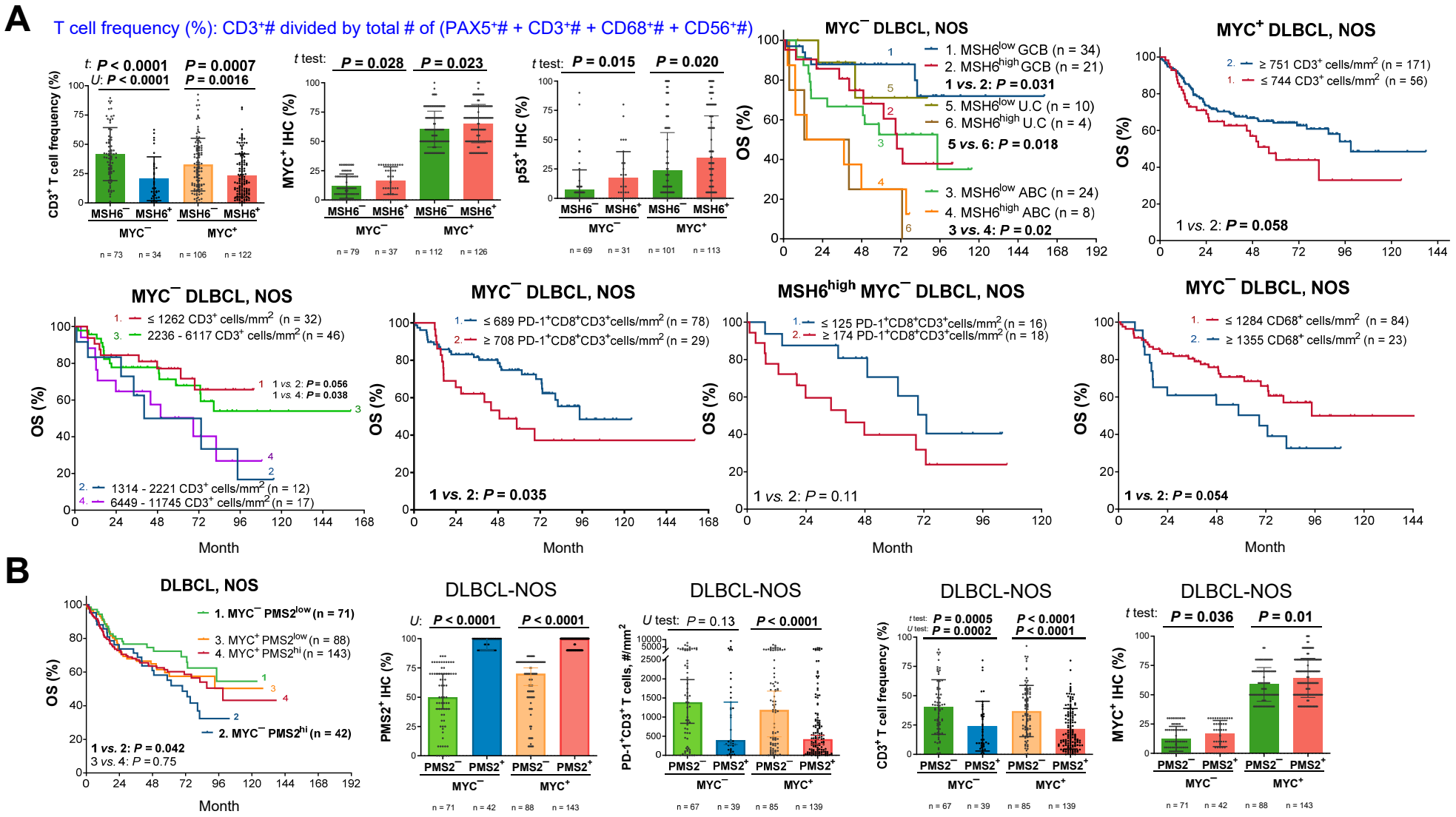




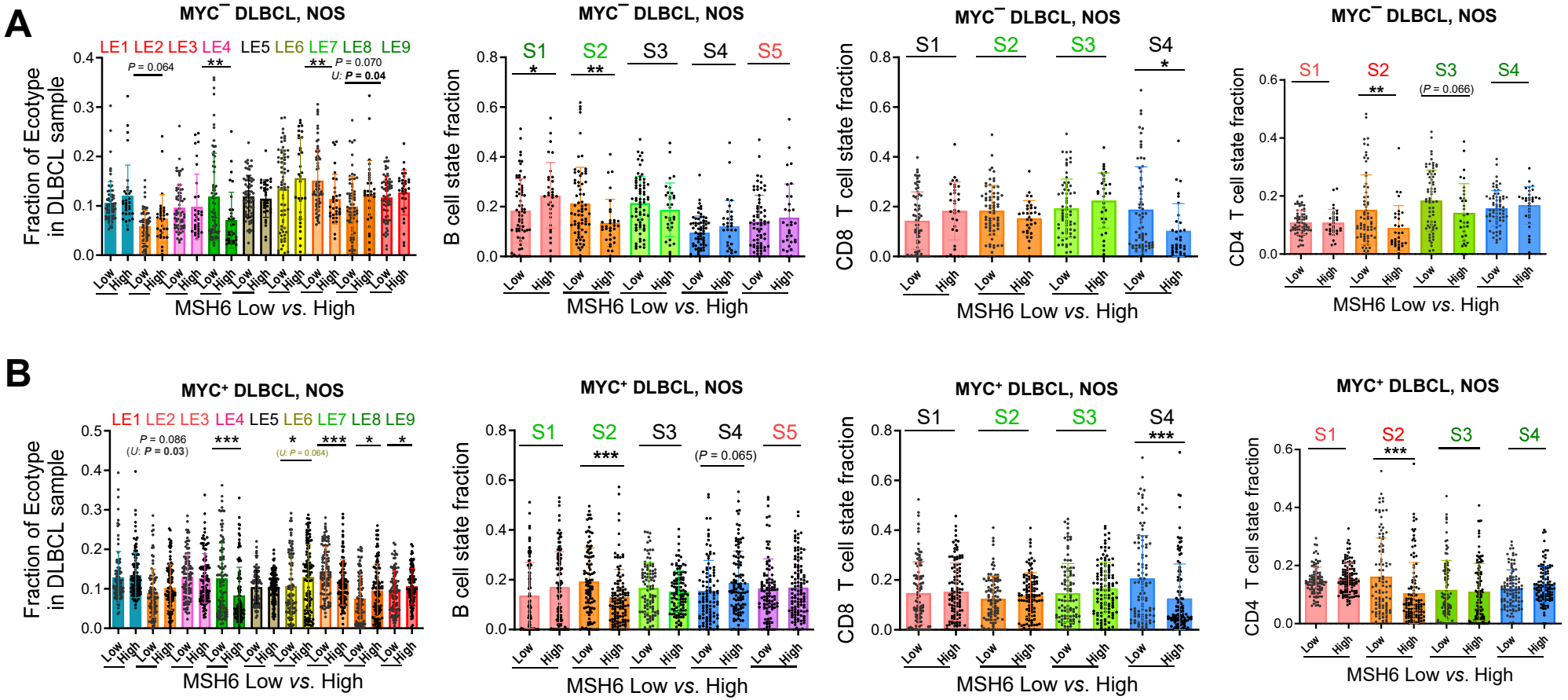
Suppl. Figure 10



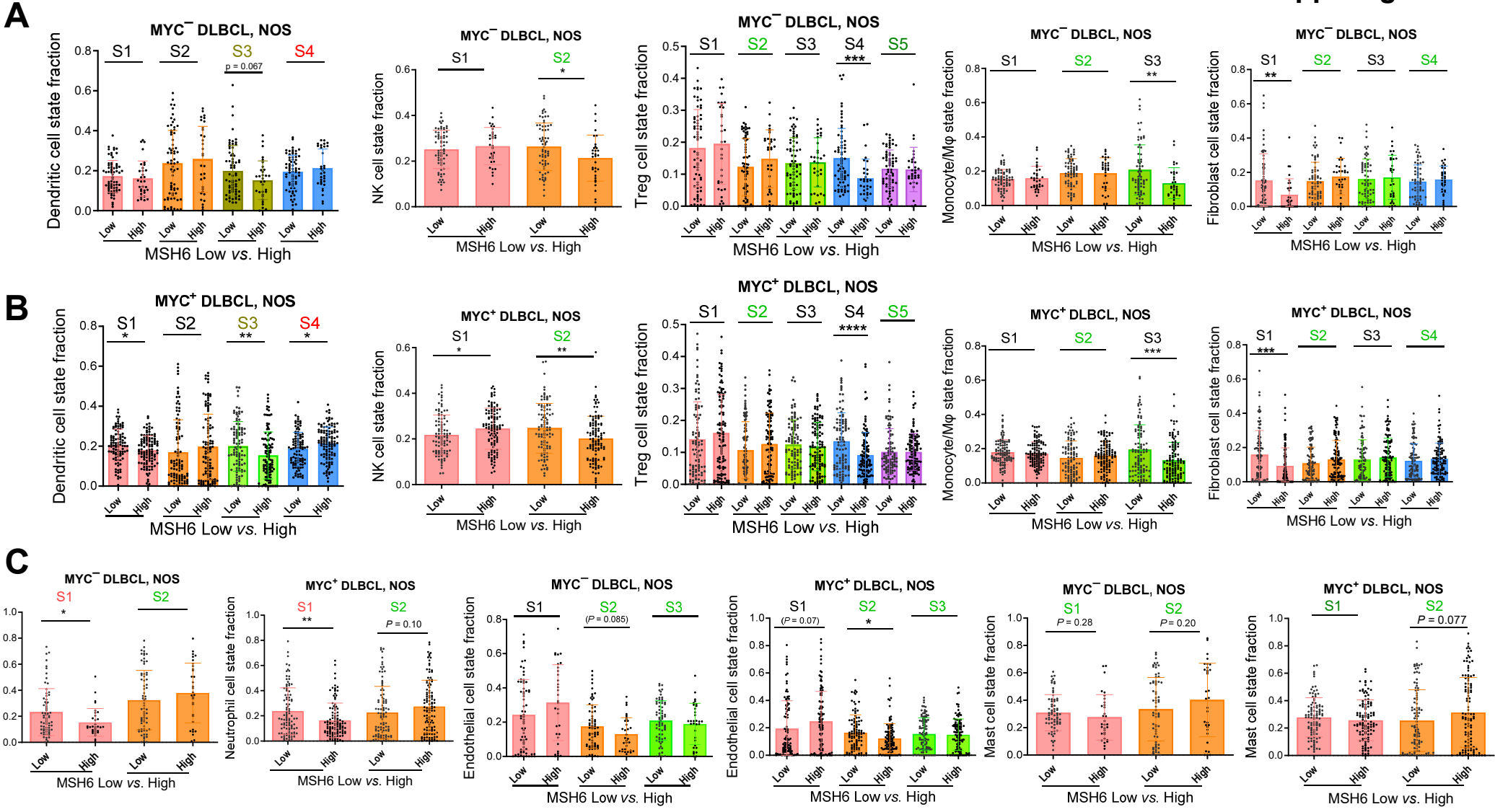
Suppl. Figure 11



Suppl. Figure 12



Suppl. Figure 13



Suppl. Figure 14