Supplementary Figure legends

Supplementary Figure 1. Prognostic and GEP analysis for clinicopathological factors. (**A-B**) Forest plots showing the effects of clinicopathological factors on overall survival (OS, panel A) and progression-free survival (PFS, panel B) of the PT-DLBCL cohort. (**C**) Heatmaps for significantly differentially expressed genes (FDR <0.05, 198 genes for age>60 and 318 genes for tumor size≥7cm)

Supplementary Figure 2. Prognostic analysis for genetic alterations in PT-DLBCL. (**A**) Genetic subtypes did not show significant impact on progression-free survival (PFS) of patients with PT-DLBCL. Patients with a <5% variant allele frequency (VAF) of *TP53* mutation by DNA sequencing showed a trend of poorer overall survival (OS) than those with wild-type *TP53* DNA sequence. (**B**) MSI-high by DNA sequencing analysis was associated with significantly better PFS in PT-DLBCL. CIBERSORT analysis for corresponding RNA-seq data showed that MSI-high was associated with a lower mean proportion of M0 (unpolarized) macrophages in the PT-DLBCL tissue samples. (**C**) High numbers of chromosomes with copy number variations (CNVs) was associate with a significantly poorer PFS in the PT-DLBCL cohort. There were no associations between CNV numbers and age>60 or tumor size≥5cm. (**D**) By FISH analysis, patients with *BCL2* gene amplification, >4 polyploid, or *BCL2* gene rearrangement had significantly poorer PFS than other patients with PT-DLBCL. *BCL6* and *MYC* gene alterations by FISH analysis did not show significant prognostic effects.

Supplementary Figure 3. Unsupervised clustering and prognostic analysis for RNA-seq data in DLBCL. (**A**) Visualization of significantly differentially expressed 533 genes between PT-

DLBCL and systemic DLBCL cohorts by unsupervised clustering. (**B**) Unsupervised clustering in the PT-DLBCL cohort by genes upregulated in systemic DLBCL compared with PT-DLBCL. Significant prognostic differences were observed between the two clusters of PT-DLBCL. (**C**) The TLT cluster had a significantly poorer progression-free survival (PFS) than the ME cluster in the overall PT-DLBCL cohort and subcohort of patients less than 61 years old. (**D**) Top: Unsupervised clustering of the PT-DLBCL cohort by 150 genes upregulated in PT-DLBCL compared with systemic DLBCL (see more details on TLT and ME clusters in the main text). Bottom: Significantly differentially expressed 870 genes between the TLT and ME clusters. Survival curves below showed that the prognostic effect of TLT vs. ME was independent of MCD-like genetic subtype and CNV ≥3.

Supplementary Figure 4. *MYD88* mutation and prognostic analysis in PT-DLBCL. (**A**) Scatter plots show *BTK* mRNA levels by RNA-seq in PT-DLBCL and systemic DLBCL with wild-type or mutated *MYD88* according to RNA-seq variant analysis or DNA sequencing results. PT-DLBCL with wild-type *MYD88* showed a non-significant trend of better survival compared with systemic DLBCL with wild-type *MYD88*. Enrichment dot plot shows enriched GO terms and KEGG pathways in 63 genes upregulated in PT-DLBCL with wild-type *MYD88* compared with PT-DLBCL with mutated *MYD88*. (**B**) Significantly differentially expressed 469 genes between patients with and without *MYD88* mutations by DNA sequencing analysis. (**C**) No significant differences in patient survival were observed between PT-DLBCL and systemic DLBCL, between patients with wild-type *MYD88*, and between patients with *MYD88* mutation by DNA sequencing analysis. (**D**) Significantly differentially expressed 163 genes between PT-DLBCL patients with *MYD88* mutation and systemic DLBCL patients with *MYD88* mutation by DNA sequencing analysis.

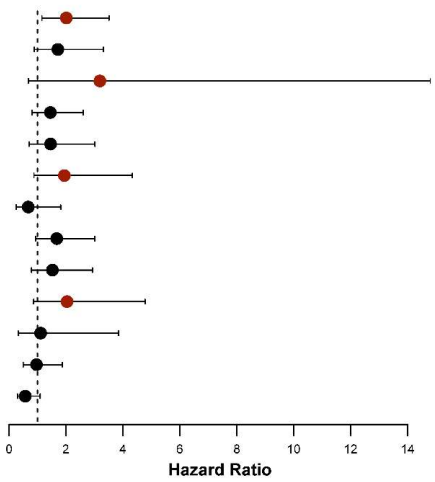
Supplementary Figure 5. (**A**) Boxplot for the abundances of 22 immune cell subpopulations by CIBERSORT analysis in G1 to G5 clusters. (**B**) Volcano plot and heatmap by unsupervised clustering for significantly differentially expressed microRNAs between PT-DLBCL and systemic DLBCL. (**C**) Unsupervised clustering of the PT-DLBCL microRNA profiling cohort using 40 microRNAs, including top 20 up and top 20 down microRNAs in PT-DLBCL compared with systemic DLBCL. The cluster with high expression of 16 upregulated microRNAs had significantly poorer OS and PFS in the PT-DLBCL cohort. (**D**) Heatmap showing differential expression of genes with $q<0.06$ between two PT-DLBCL patient groups in panel **C** clustered by microRNAs.

Supplementary Figure 6. Expression of top 20 up (**A**) and top 20 down (**B**) microRNAs in our PT-DLBCL compared with systemic DLBCL cohorts in normal and diseased tissues of testis, brain, blood, and lymph nodes in the miTED database.

Suppl. Figure 1

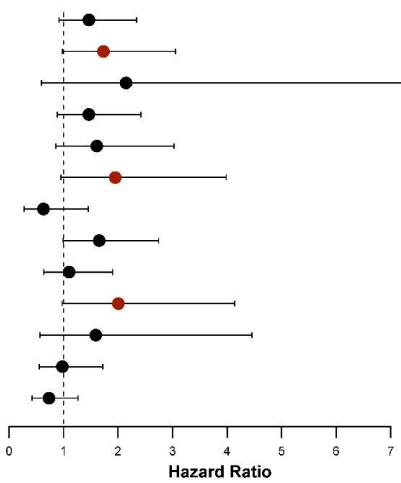
A. OS

	P value	Hazard Ratio
Age>60	0.020	2.013(1.155- 3.510)
Stage3- 4	0.067	1.716(0.888- 3.318)
ECOG>1	0.009	3.191(0.688- 14.810)
LDH elevated	0.197	1.453(0.810- 2.606)
Extranodal involvement	0.248	1.461(0.708- 3.017)
IPI>2	0.042	1.945(0.874- 4.330)
B- symptom	0.506	0.675(0.251- 1.813)
Tumor size...5cm	0.092	1.677(0.934- 3.009)
Left vs. right or bilateral	0.196	1.523(0.791- 2.932)
Regional nodal involvement	0.037	2.037(0.869- 4.775)
Faraway nodal involvement	0.862	1.111(0.321- 3.843)
ABC vs. GCB	0.915	0.965(0.498- 1.870)
Rituximab	0.056	0.573(0.299- 1.098)

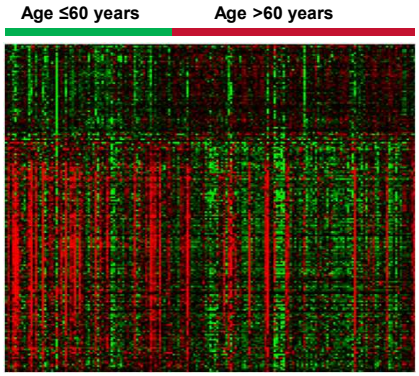


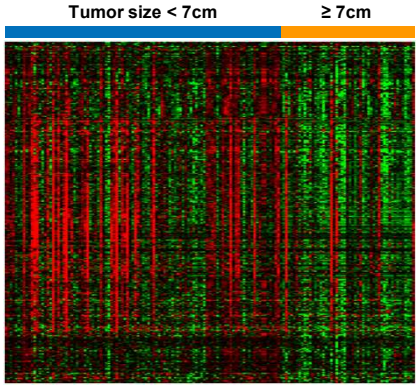
B. PFS

	P value	Hazard Ratio
Age>60	0.120	1.466(0.919- 2.338)
Stage3- 4	0.028	1.732(0.983- 3.053)
ECOG>1	0.090	2.147(0.596- 7.736)
LDH elevated	0.122	1.465(0.886- 2.420)
Extranodal involvement	0.084	1.611(0.858- 3.025)
IPI>2	0.019	1.949(0.954- 3.985)
B- symptom	0.367	0.632(0.275- 1.453)
Tumor size...5cm	0.056	1.654(0.997- 2.743)
Left vs. right or bilateral	0.723	1.102(0.639- 1.898)
Regional nodal involvement	0.017	2.006(0.971- 4.142)
Faraway nodal involvement	0.281	1.591(0.568- 4.454)
ABC vs. GCB	0.940	0.979(0.556- 1.724)
Rituximab	0.222	0.734(0.427- 1.264)

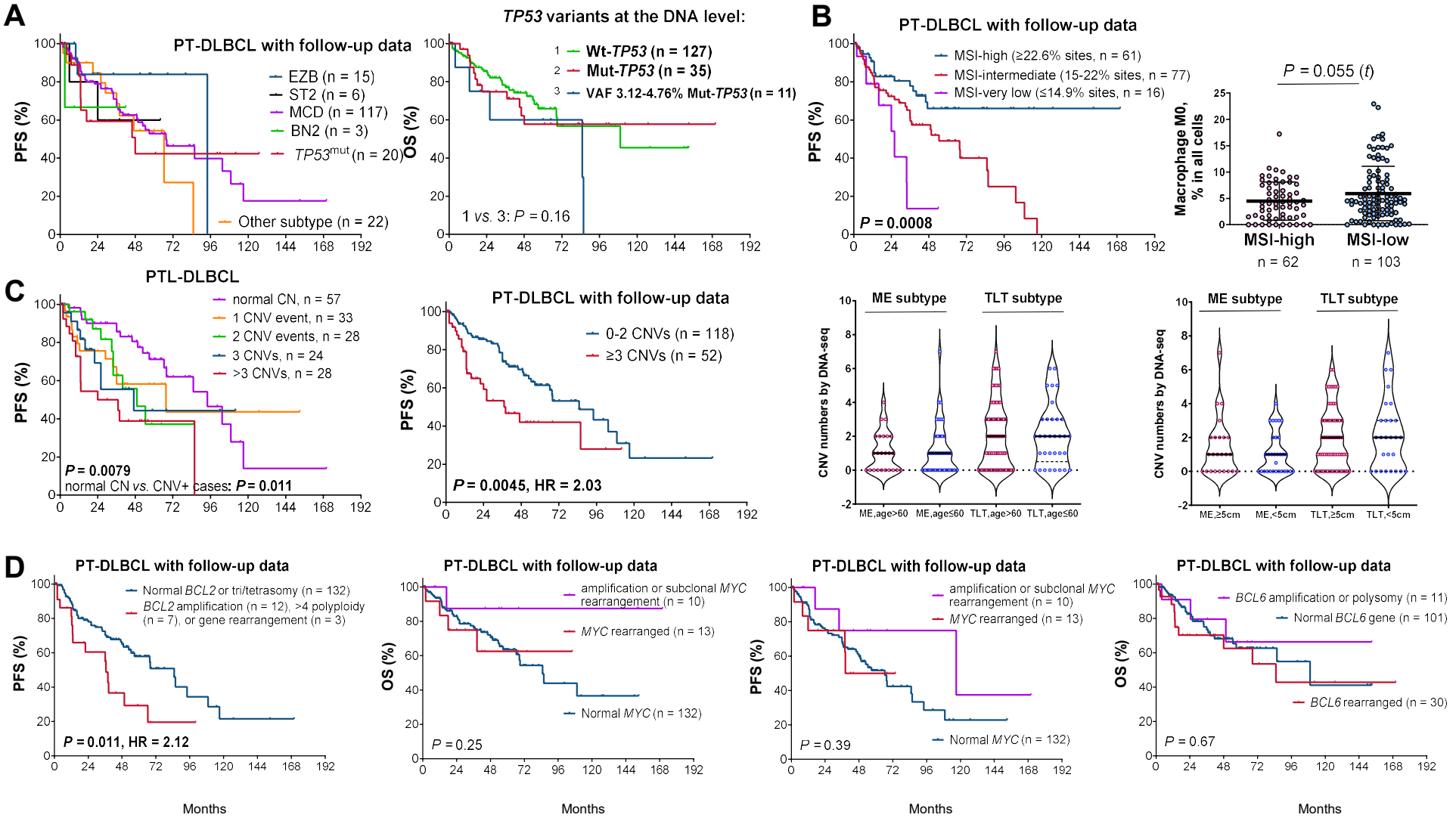


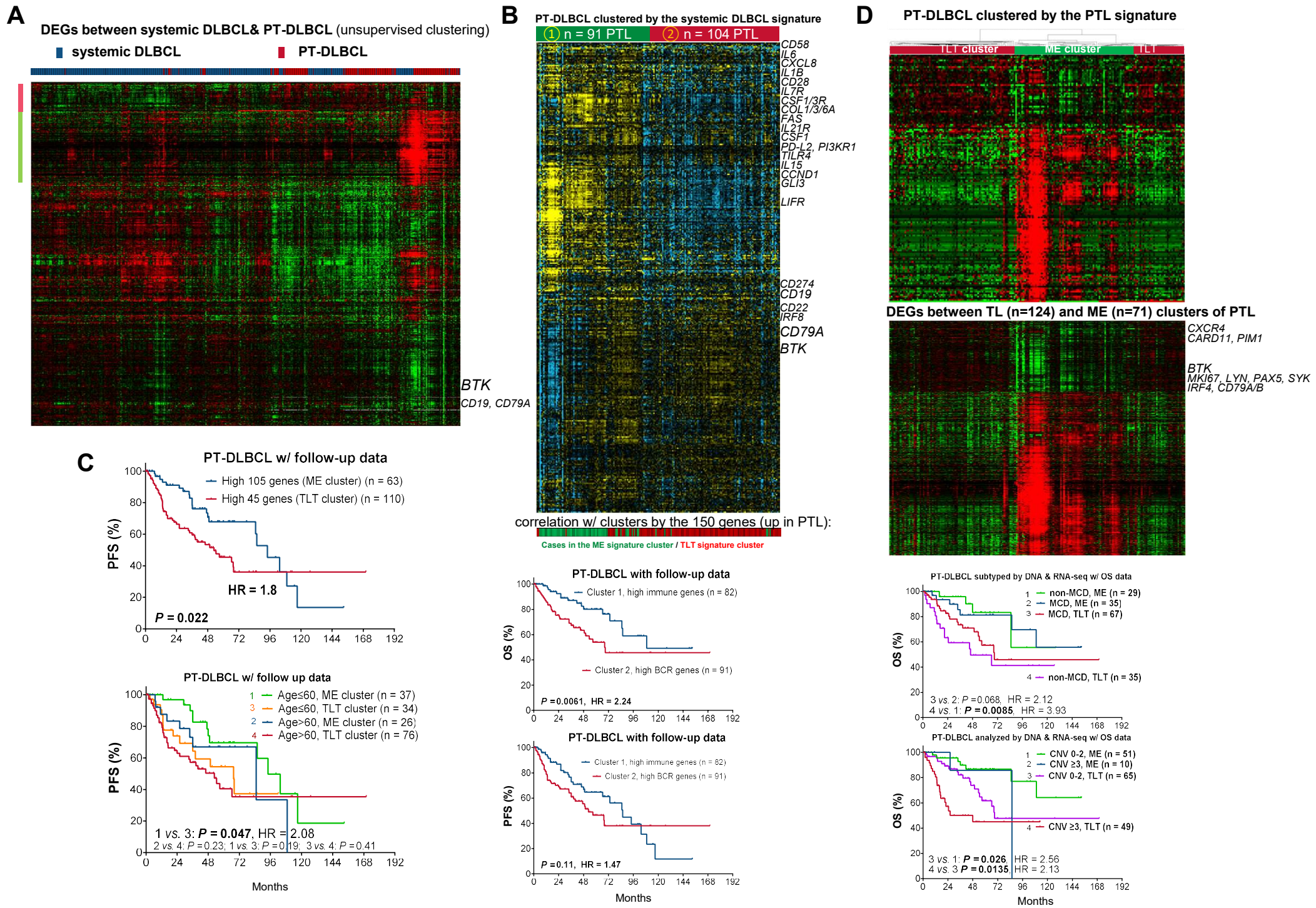
C





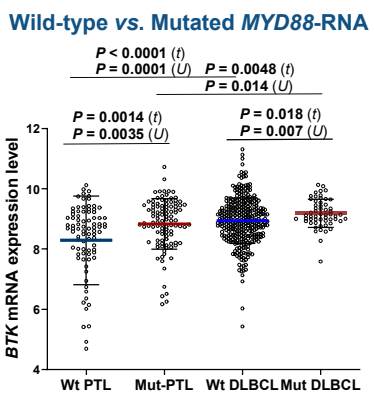
Suppl. Figure 2

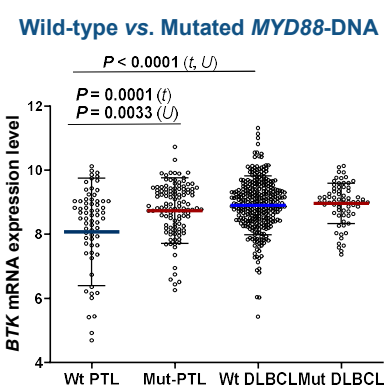


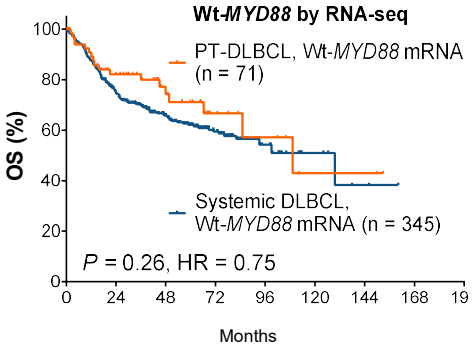

A
DEGs between systemic DLBCL& PT-DLBCL (unsupervised clustering)
systemic DLBCL
PT-DLBCL
BTK
CD19, CD79A
B
PT-DLBCL clustered by the systemic DLBCL signature
① n = 91 PTL
② n = 104 PTL
CD58
IL6
CXCL8
IL1B
CD28
IL7R
CSF1/3R
COL1/3/6A
FAS
IL21R
CSF1
PD-L2, PI3KR1
TILR4
IL15
CCND1
GLI3
LIFR
CD274
CD19
CD22
IRF8
CD79A
BTK
correlation w/ clusters by the 150 genes (up in PTL):
Cases in the ME signature cluster / TLT signature cluster
C
PT-DLBCL w/ follow-up data
High 105 genes (ME cluster) (n = 63)
High 45 genes (TLT cluster) (n = 110)
PFS (%)
HR = 1.8
P = 0.022
PT-DLBCL w/ follow up data
1 Age≤60, ME cluster (n = 37)
3 Age≤60, TLT cluster (n = 34)
2 Age>60, ME cluster (n = 26)
4 Age>60, TLT cluster (n = 76)
PFS (%)
1 vs. 3: P = 0.047, HR = 2.08
2 vs. 4: P = 0.23; 1 vs. 3: P = 0.19; 3 vs. 4: P = 0.41
Months
PT-DLBCL with follow-up data
Cluster 1, high immune genes (n = 82)
Cluster 2, high BCR genes (n = 91)
OS (%)
P = 0.0061, HR = 2.24
PT-DLBCL with follow-up data
Cluster 1, high immune genes (n = 82)
Cluster 2, high BCR genes (n = 91)
PFS (%)
P = 0.11, HR = 1.47
Months
D
PT-DLBCL clustered by the PTL signature
TLT cluster
ME cluster
TLT
DEGs between TL (n=124) and ME (n=71) clusters of PTL
CXCR4
CARD11, PIM1
BTK
MKI67, LYN, PAX5, SYK
IRF4, CD79A/B
PT-DLBCL subtyped by DNA & RNA-seq w/ OS data
1 non-MCD, ME (n = 29)
2 MCD, ME (n = 35)
3 MCD, TLT (n = 67)
4 non-MCD, TLT (n = 35)
OS (%)
3 vs. 2: P = 0.068, HR = 2.12
4 vs. 1: P = 0.0085, HR = 3.93
PT-DLBCL analyzed by DNA & RNA-seq w/ OS data
1 CNV 0-2, ME (n = 51)
2 CNV ≥3, ME (n = 10)
3 CNV 0-2, TLT (n = 65)
4 CNV ≥3, TLT (n = 49)
OS (%)
3 vs. 1: P = 0.026, HR = 2.56
4 vs. 3 P = 0.0135, HR = 2.13
Months

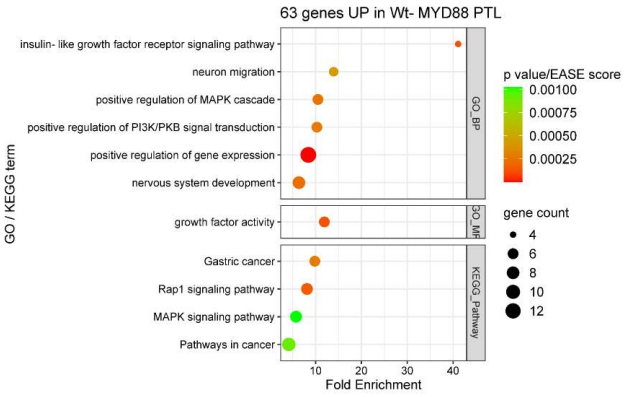
Suppl. Figure 4

A



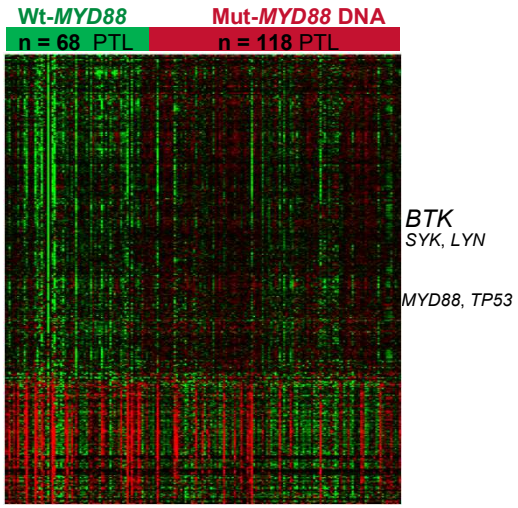




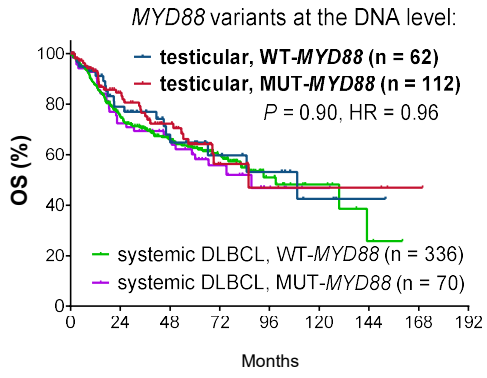


Top KEGG pathway: Rap1signaling pathway

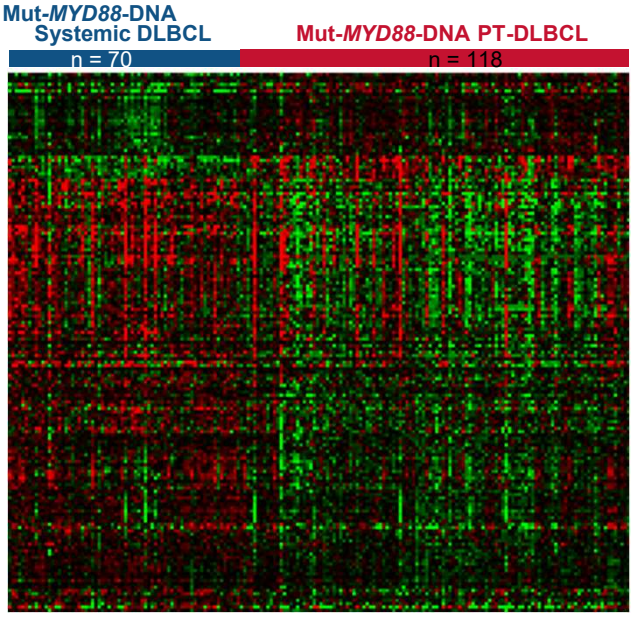
B



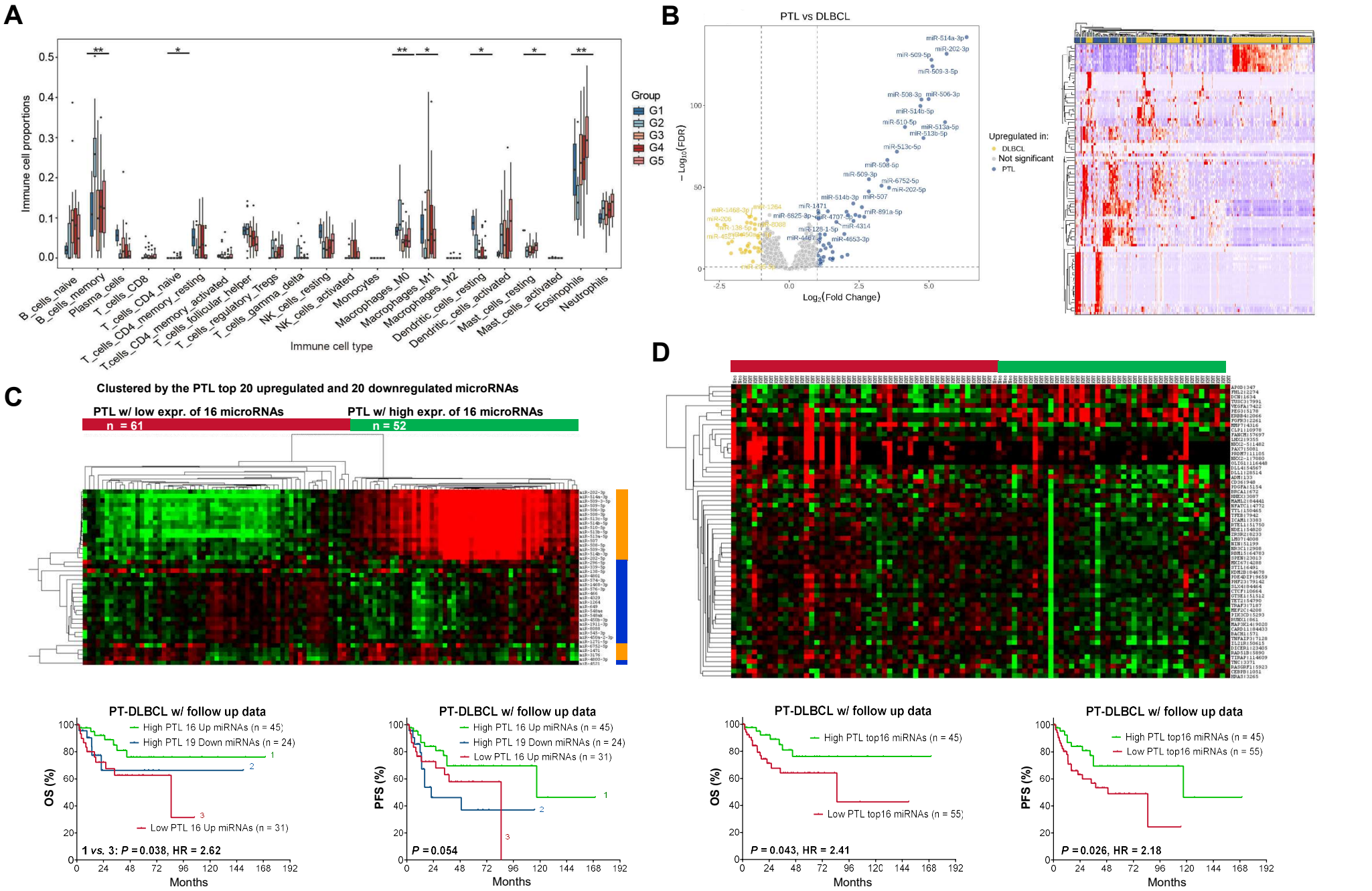
C



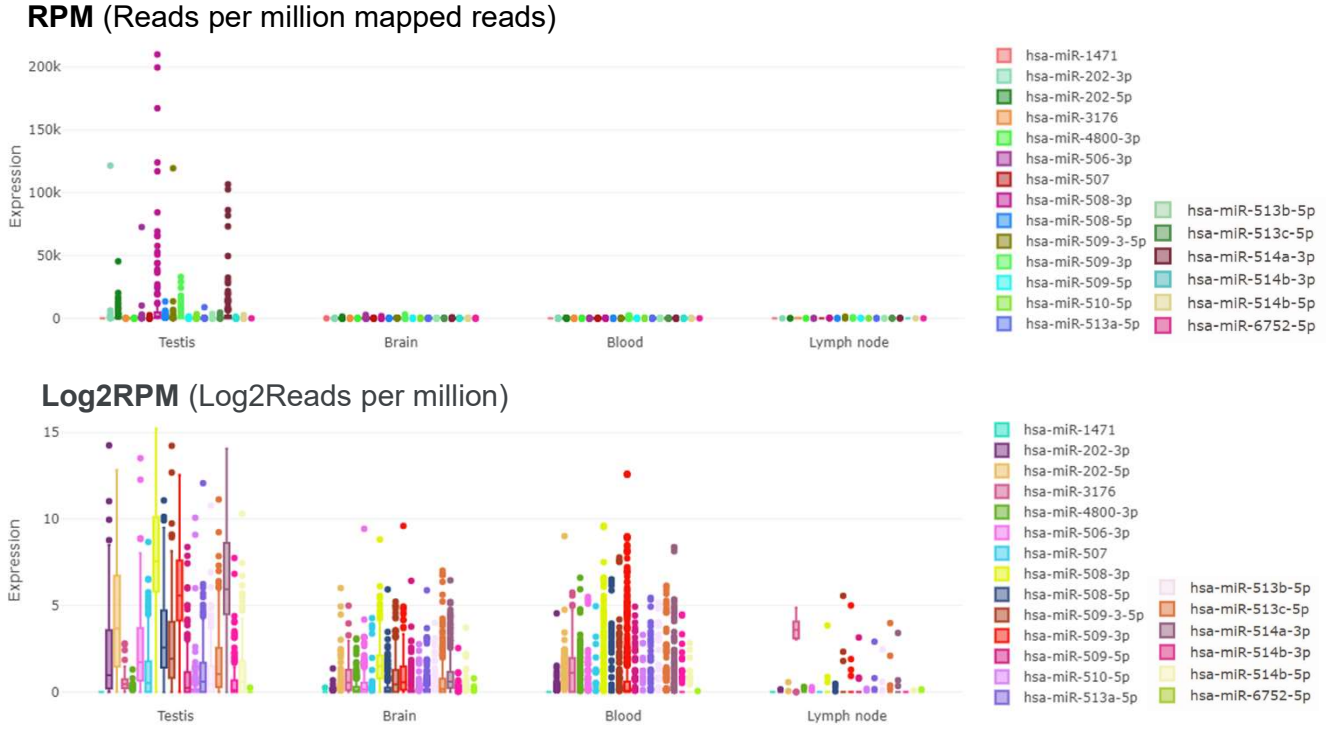
D



Suppl. Figure 5

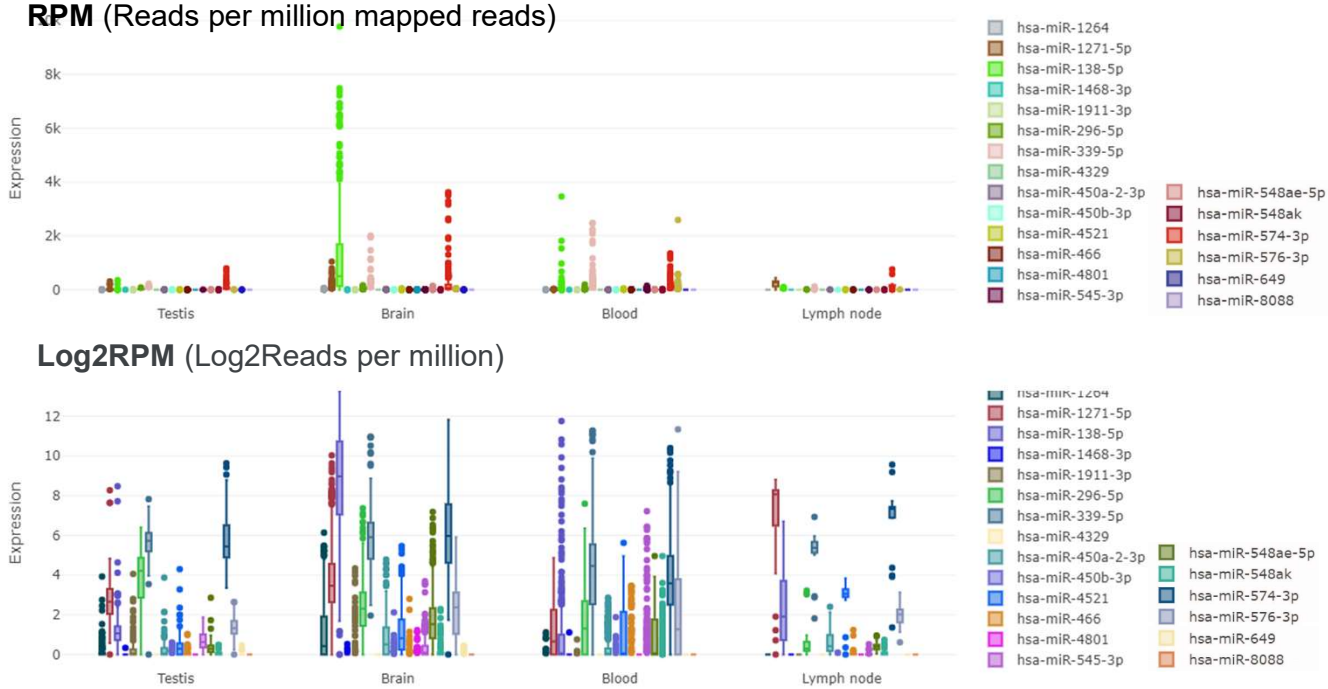


(A)



(B)

20 down-regulated microRNAs in PT-DLBCL



Tissue-Disease Sankey plot