Supplementary Material

Clinical and biological subtypes of B-cell lymphoma revealed by microenvironmental signatures.

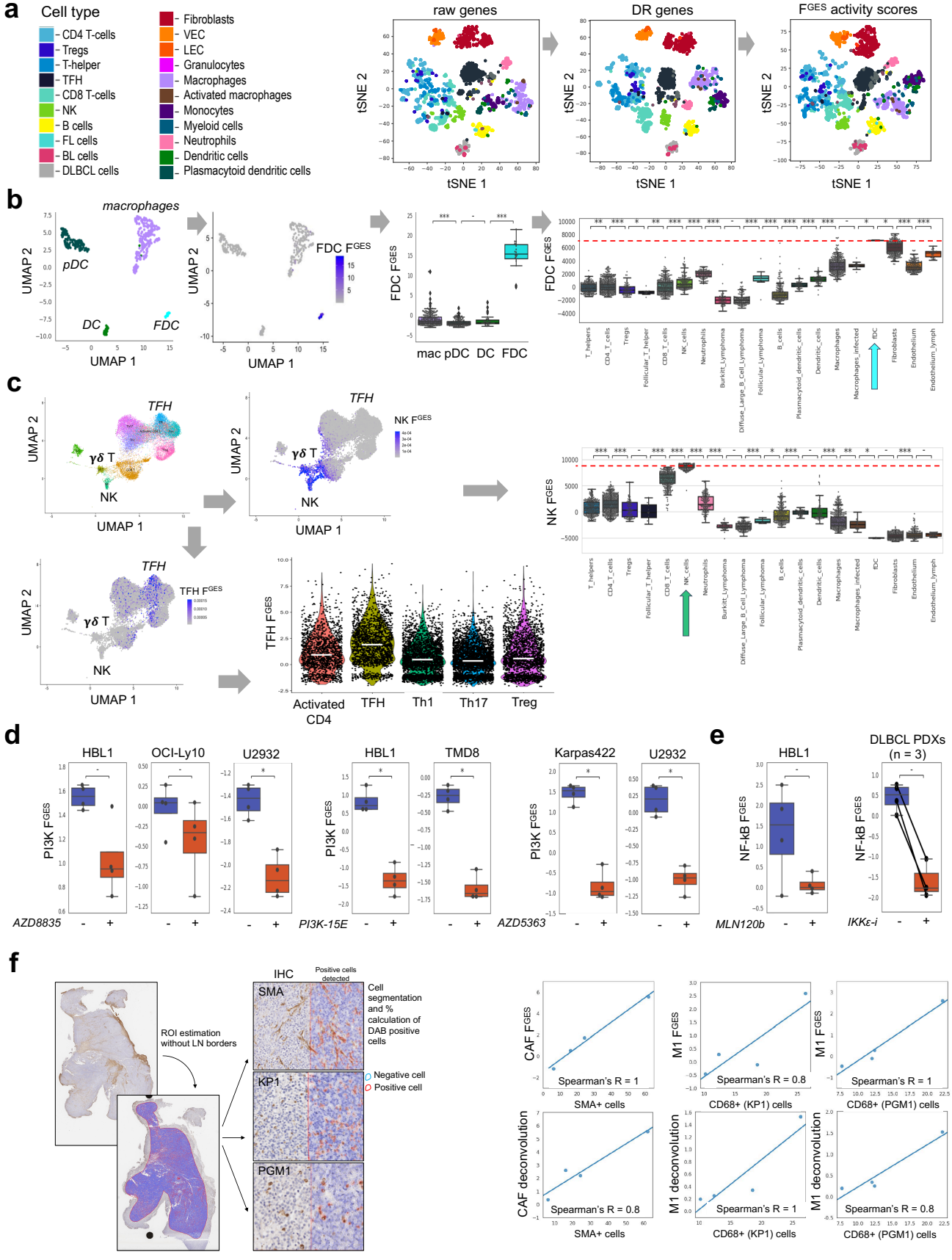
a
Cell type
CD4 T-cells
Tregs
T-helper
TFH
CD8 T-cells
NK
B cells
FL cells
BL cells
DLBCL cells
Fibroblasts
VEC
LEC
Granulocytes
Macrophages
Activated macrophages
Monocytes
Myeloid cells
Neutrophils
Dendritic cells
Plasmacytoid dendritic cells
raw genes
DR genes
FGES activity scores
tSNE 1
tSNE 2
b
macrophages
pDC
DC
FDC
UMAP 1
UMAP 2
FDC FGES
mac pDC DC FDC
T_helpers
CD4_T_cells
Tregs
Follicular_T_helper
CD8_T_cells
NK_cells
Neutrophils
Burkitt_Lymphoma
Diffuse_Large_B_Cell_Lymphoma
Follicular_Lymphoma
B_cells
Plasmacytoid_dendritic_cells
Dendritic_cells
Macrophages
Macrophages_infected
fDC
Fibroblasts
Endothelium
Endothelium_lymph
c
TFH
γδ T
NK
NK FGES
TFH FGES
Activated CD4
TFH
Th1
Th17
Treg
d
HBL1
OCI-Ly10
U2932
TMD8
Karpas422
PI3K FGES
AZD8835
PI3K-15E
AZD5363
e
NF-kB FGES
MLN120b
DLBCL PDXs (n = 3)
IKKε-i
f
IHC
Positive cells detected
SMA
KP1
PGM1
ROI estimation without LN borders
Cell segmentation and % calculation of DAB positive cells
Negative cell
Positive cell
CAF FGES
M1 FGES
Spearman's R = 1
Spearman's R = 0.8
SMA+ cells
CD68+ (KP1) cells
CD68+ (PGM1) cells
CAF deconvolution
M1 deconvolution

Figure S1. a: t-SNE projections of 5,069 RNA-sequencing profiles of different cell types representing progressive steps in data reduction dimensionality analysis from raw genes to F^{GES} activity scores. Different colors represent unique cell types**. b:** Follicular dendritic cells (FDC) F^{GES} validation in single cell RNA-seq profiles of macrophages, plasmacytic dendritic cells (pDC), dendritic cells (DC) and FDCs from (https://www.biorxiv.org/content/10.1101/2020.04.28.054775v2). FDC F^{GES} expression in RNA-seq from isolated lymphoma and LME cell subtypes. Arrow shows FDCs. **c:** Natural killer (NK) and follicular helper T cells (FHT) F^{GES} validation in single cell RNA-seq profiles of FHT, NK, gamma-delta T cells, activated CD4 cells, Th1, Th17 and Tregs from(1). NK F^{GES} expression in RNA-seq from isolated lymphoma and LME cell subtypes. Arrow shows NKs. (bottom) TFH F^{GES} expression in a panel of single cell RNA-seq T cells. **d-e:** PI3K and NF-kB F^{GES} validation in RNA-seq from ABC-DLBCL cells lines treated with PI3K and AKT inhibitors(2,3) (d) and HBL1 treated with IKKb inhibitor MLN120b(2) and 3 matched DLBCL PDX treated with IKKε/TBK1 inhibitor(4) (e). Data from datasets GSE92619, GSE94610 and GSE121159. **f:** example of immunohistochemistry (IHC) quantification (cell segmentation and calculation of positive cells) from anti-SMA and two anti-CD68 antibodies in DLBCL tissue. Spearman's correlation of F^{GES} for CAF and M1 macrophages (top) and deconvolution for CAF and M1 macrophages (bottom) with their respective cell quantification from IHC in 5 DLBCL cases. **p<0.05, **p<0.01, ***p<0.001.*

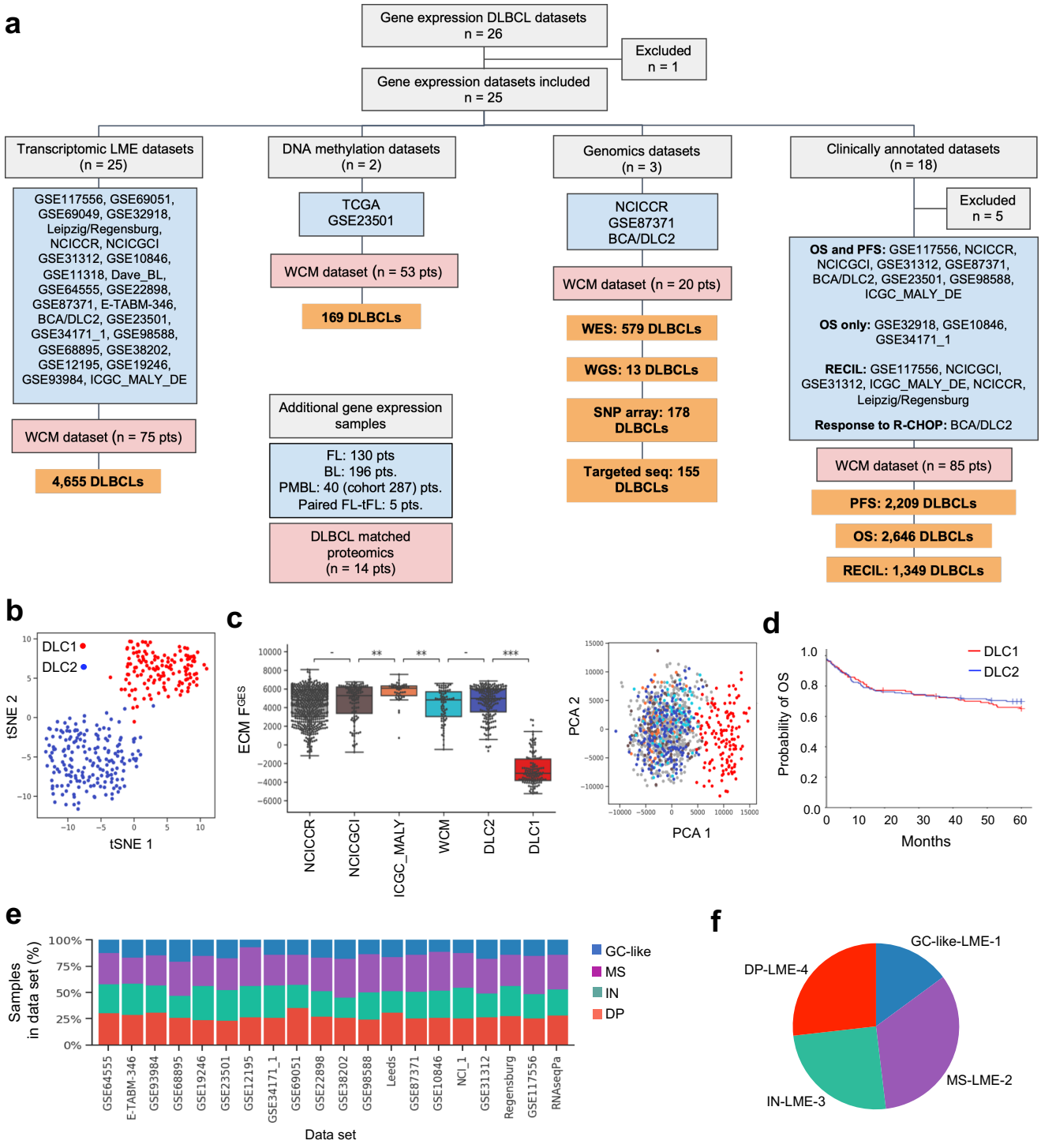

Figure S2. a. Flowchart of datasets included in the study. Part of one dataset (DCL1) was excluded because contained isolated lymphoma cells, another dataset with mutations was excluded due to low number of cases analyses and 5 clinically annotated datasets were excluded since they OS/PFS curves were not representative of an unselected DLBCL population. **b:** tSNE projection in the F^{GES} space of the

sample excluded (DCL1) vs. samples containing microenvironment cells (DCL2) from the same dataset. **c:** DLC1 samples had significantly lower F^{GES} ECM signature and were outliers on the PCA projection in F^{GES} space. **d.** Kaplan-Meier curves for OS of DCL1 (excluded) vs. DCL2 (included) datasets. **e:** LME distribution per compiled datasets showing individual dataset contribution to LME characterization. **f:** LME distribution in all the samples analyzed (n = 4,655)

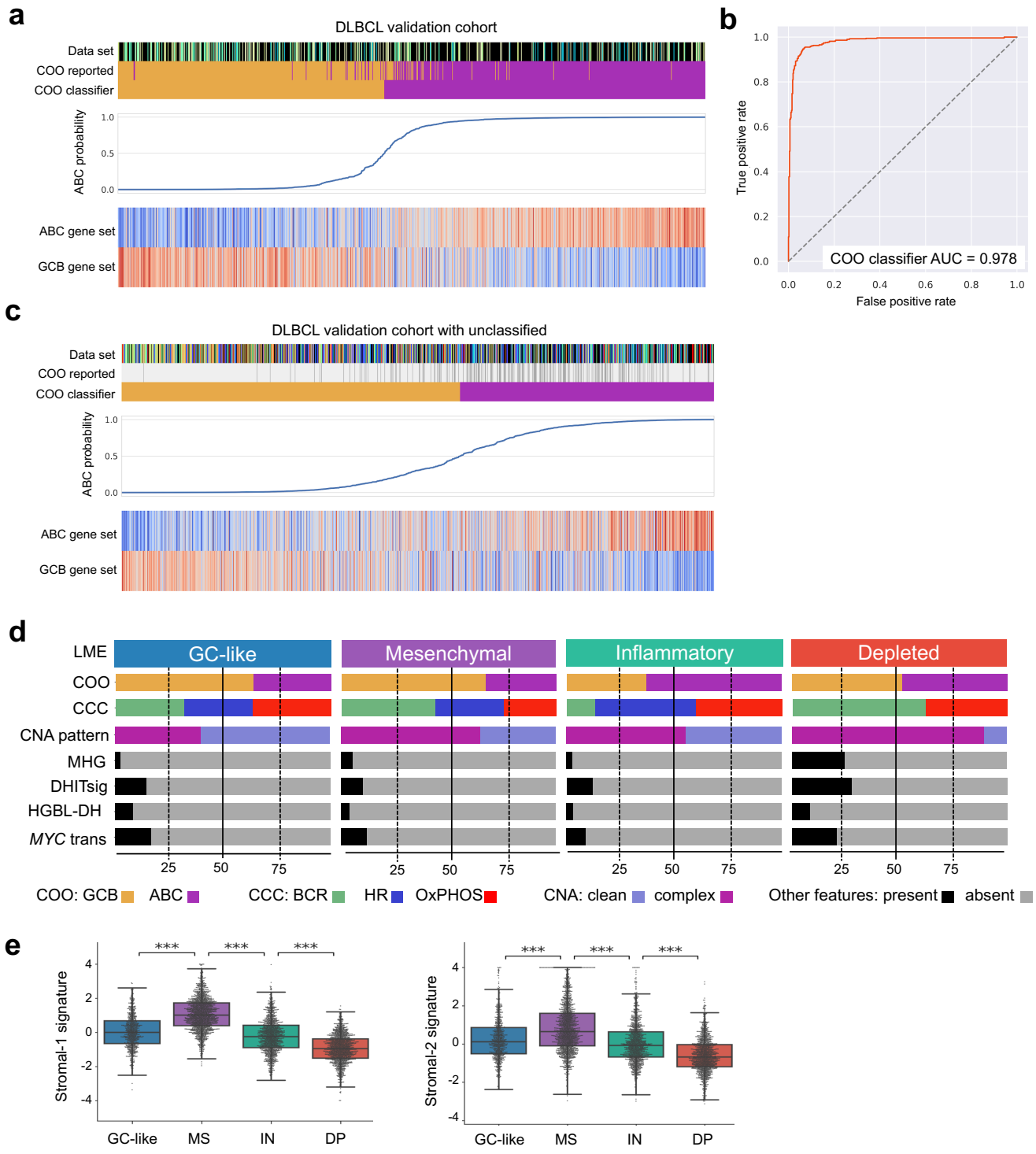

Figure S3. a: *Top:* COO classifier performance on the validation dataset comparing reported vs. predicted COO class. *Middle:* ABC-DLBCL probability. *Bottom:* ABC- and GCB-DLBCL signatures as reported. **b:** ROC curves comparing reported vs. classified. **c:** COO classifier performance in the full cohort using samples with previously reported (gray lines) and unknown (white lines) COO status. ABC-

DLBCL probability and ABC- and GCB-DLBCL signatures. **d:** Distribution of COO, CCC, CNA pattern, MHG, DHITsig, HGBL-DH and MYC translocation categories across LME subtypes. **e:** Distribution of stromal-1 and stromal-2 gene expression signatures according to LME subtype.

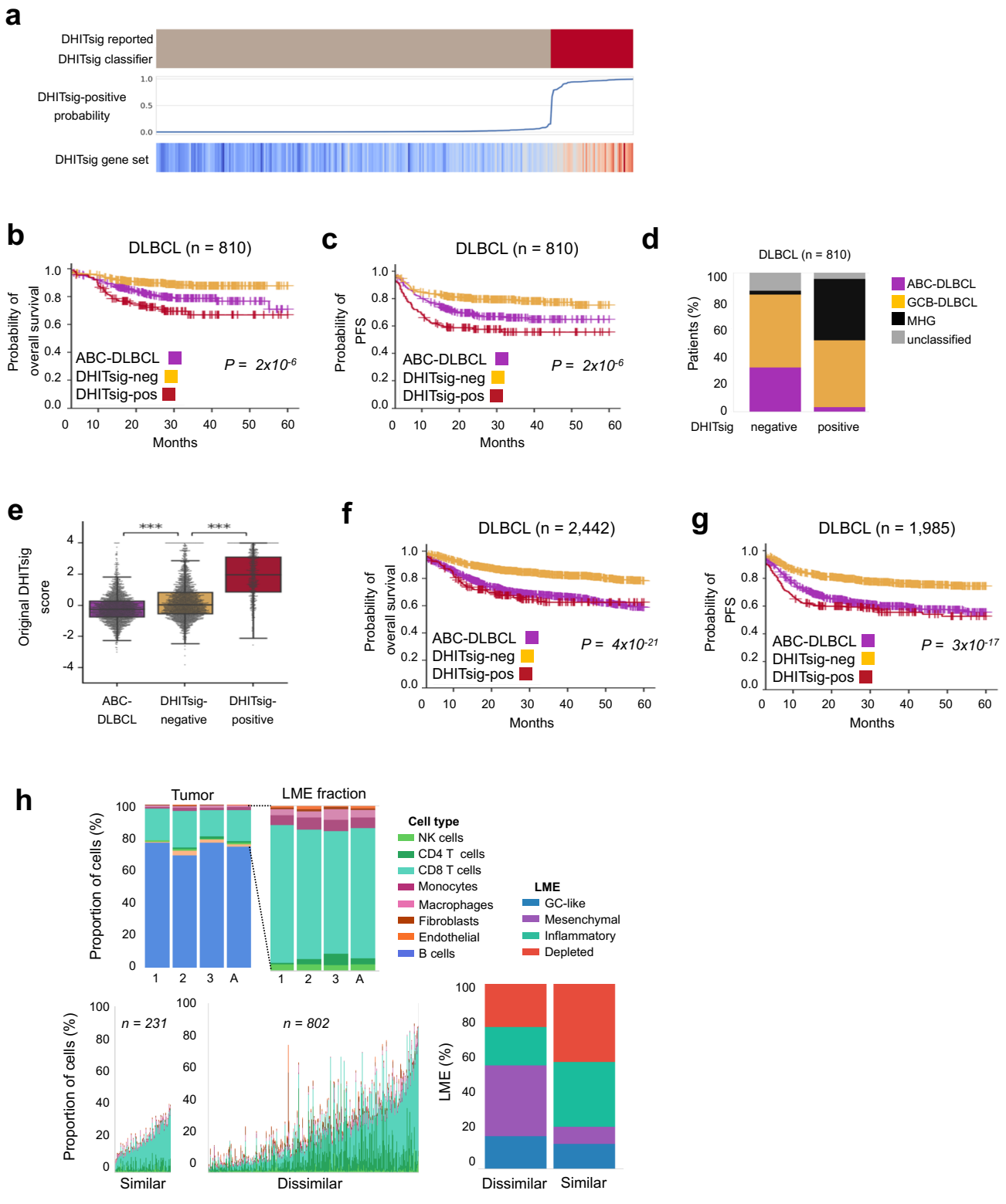

Figure S4. a: *Top:* DHITsig classifier performance on the validation dataset comparing reported vs. predicted DHITsig (red present, gray absent) in GCB-DLBCLs. *Middle:* DHIT-sig-positive probability.

Bottom: DHITsig signature as reported. **b-c:** Five-year OS and PFS curves for the GSE117556 DLBCL cohort (n = 810). **d:** overlap between several classifications showing the distribution of COO subtypes (ABC and GCB) and MHG DLBCLs into DHITsig negative vs. positive cases. UNC: unclassified. e: DHITsig score in the complete DLBCL cohort. **f-g:** Five-year OS and PFS curves for the full cohort (n = 2,442 and 1,985 cases, respectively). **h:** Characterization of the LME in the *myc* + *bcl2* expressing genetically engineered mice B-cell lymphoma model by the LME similarity score. LME human DLBCLs with the highest 20% of similarity score (n = 231) to *myc* + *bcl2* expressing LME showed enrichment of DP-LME (41%).

a

LME-2 / LME-1	N
LME-4 / LME-3	N
LME-3 / LME-4	N
LME-3 / LME-3	Y
LME-2 / LME-2	Y
LME-4 / LME-4	Y
LME-1 / LME-4	N
LME-4 / LME-4	Y
LME-2 / LME-4	N
LME-3 / LME-3	Y
LME-3 / LME-3	Y
LME-4 / LME-4	Y
LME-2 / LME-2	Y
LME-2 / LME-2	Y
LME-4 / LME-4	Y
LME-2 / LME-2	Y
LME-1 / LME-1	Y
LME-4 / LME-4	Y
LME-2 / LME-2	Y
LME-2 / LME-2	Y

b

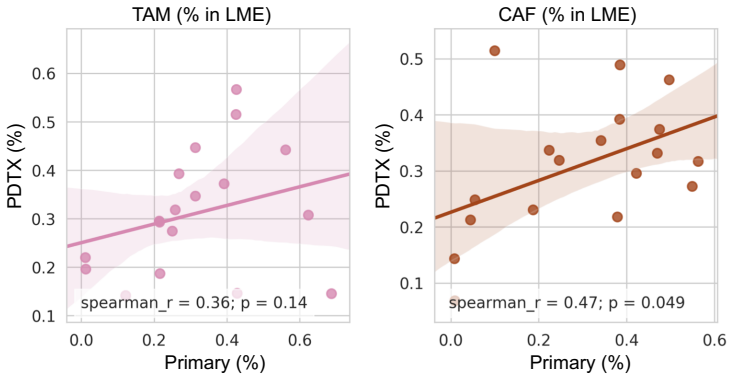

Figure S5. a: LME categories (by cell deconvolution) between primary DLBCL and their matched PDTXs. Conserved LME categories are indicated as "Y". **b:** Spearman correlation between proportion of TAM and CAF in the primary sample (X-axis) and the matched PDTX sample (Y-axis).

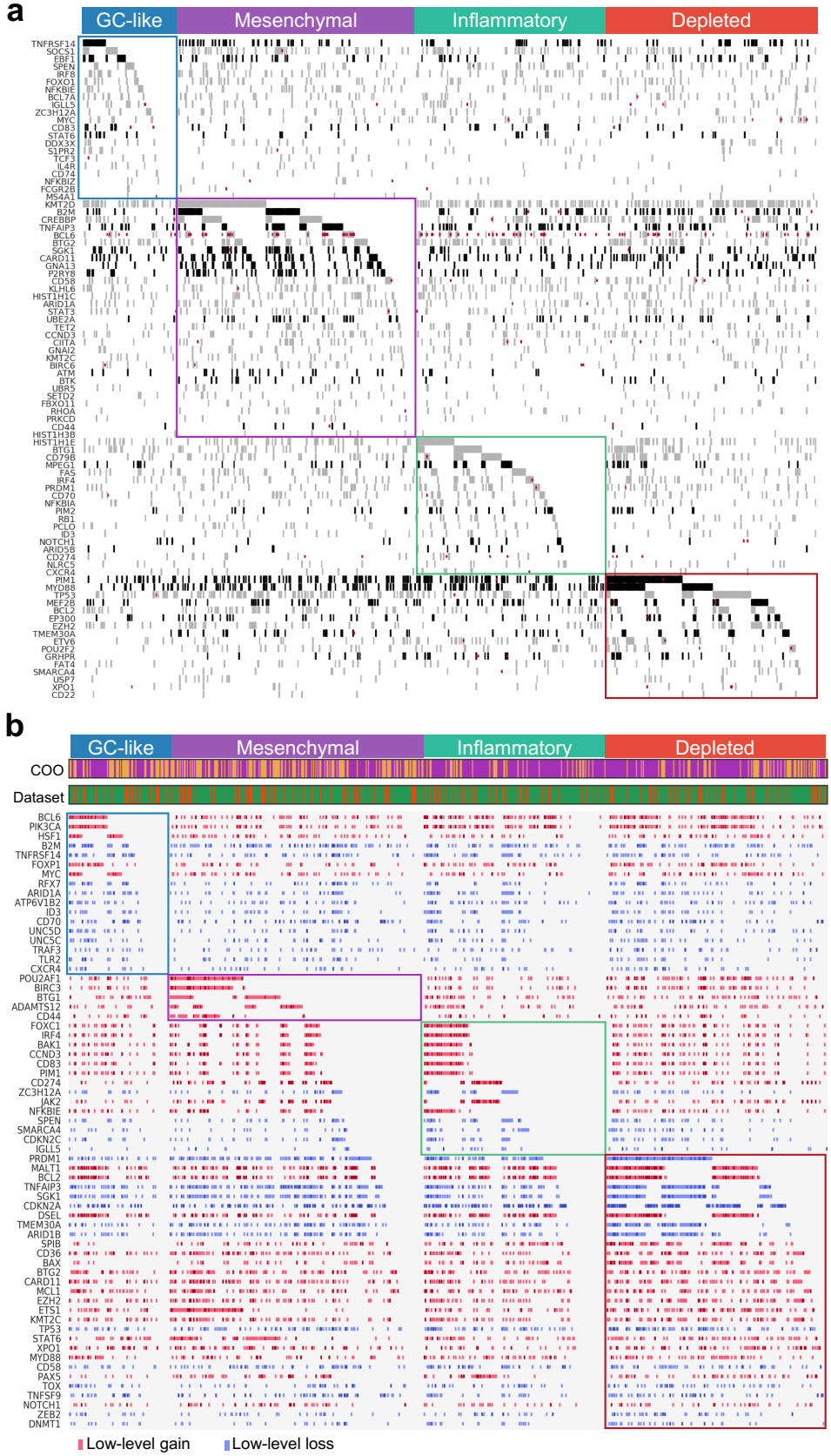
a
GC-like
Mesenchymal
Inflammatory
Depleted
TNFRSF14
SOCS1
EBF1
SPEN
IRF8
FOXO1
NFKBIE
BCL7A
IGLL5
ZC3H12A
MYC
CD83
STAT6
DDX3X
S1PR2
TCF3
IL4R
CD74
NFKBIZ
FCGR2B
MS4A1
KMT2D
B2M
CREBBP
TNFAIP3
BCL6
BTG2
SGK1
CARD11
GNA13
P2RY8
CD58
KLHL6
HIST1H1C
ARID1A
STAT3
UBE2A
TET2
CCND3
CIITA
GNAI2
KMT2C
BIRC6
ATM
BTK
UBR5
SETD2
FBXO11
RHOA
PRKCD
CD44
HIST1H3B
HIST1H1E
BTG1
CD79B
MPEG1
FAS
IRF4
PRDM1
CD70
NFKBIA
PIM2
RB1
PCLO
ID3
NOTCH1
ARID5B
CD274
NLRC5
CXCR4
PIM1
MYD88
TP53
MEF2B
BCL2
EP300
EZH2
TMEM30A
ETV6
POU2F2
GRHPR
FAT4
SMARCA4
USP7
XPO1
CD22
b
GC-like
Mesenchymal
Inflammatory
Depleted
COO
Dataset
BCL6
PIK3CA
HSF1
B2M
TNFRSF14
FOXP1
MYC
RFX7
ARID1A
ATP6V1B2
ID3
CD70
UNC5D
UNC5C
TRAF3
TLR2
CXCR4
POU2AF1
BIRC3
BTG1
ADAMTS12
CD44
FOXC1
IRF4
BAK1
CCND3
CD83
PIM1
CD274
ZC3H12A
JAK2
NFKBIE
SPEN
SMARCA4
CDKN2C
IGLL5
PRDM1
MALT1
BCL2
TNFAIP3
SGK1
CDKN2A
DSEL
TMEM30A
ARID1B
SPIB
CD36
BAX
BTG2
CARD11
MCL1
EZH2
ETS1
KMT2C
TP53
STAT6
XPO1
MYD88
CD58
PAX5
TOX
TNFSF9
NOTCH1
ZEB2
DNMT1
Low-level gain
Low-level loss
High-level gain
High-level loss

Figure S6. a: Mutation plot for the most common mutated genes in lymphoma according to the LME subtypes. Mutations are grouped into LME subtypes according to their prevalence. Mutations with significant enrichment (q-value < 0.1) in a particular LME subtype are in black. Gene translocations are in red. **b:** CNA plot for the 65 most frequently altered genes in lymphoma according to the LME subtypes. CNAs are grouped into LME subtypes according to their prevalence.

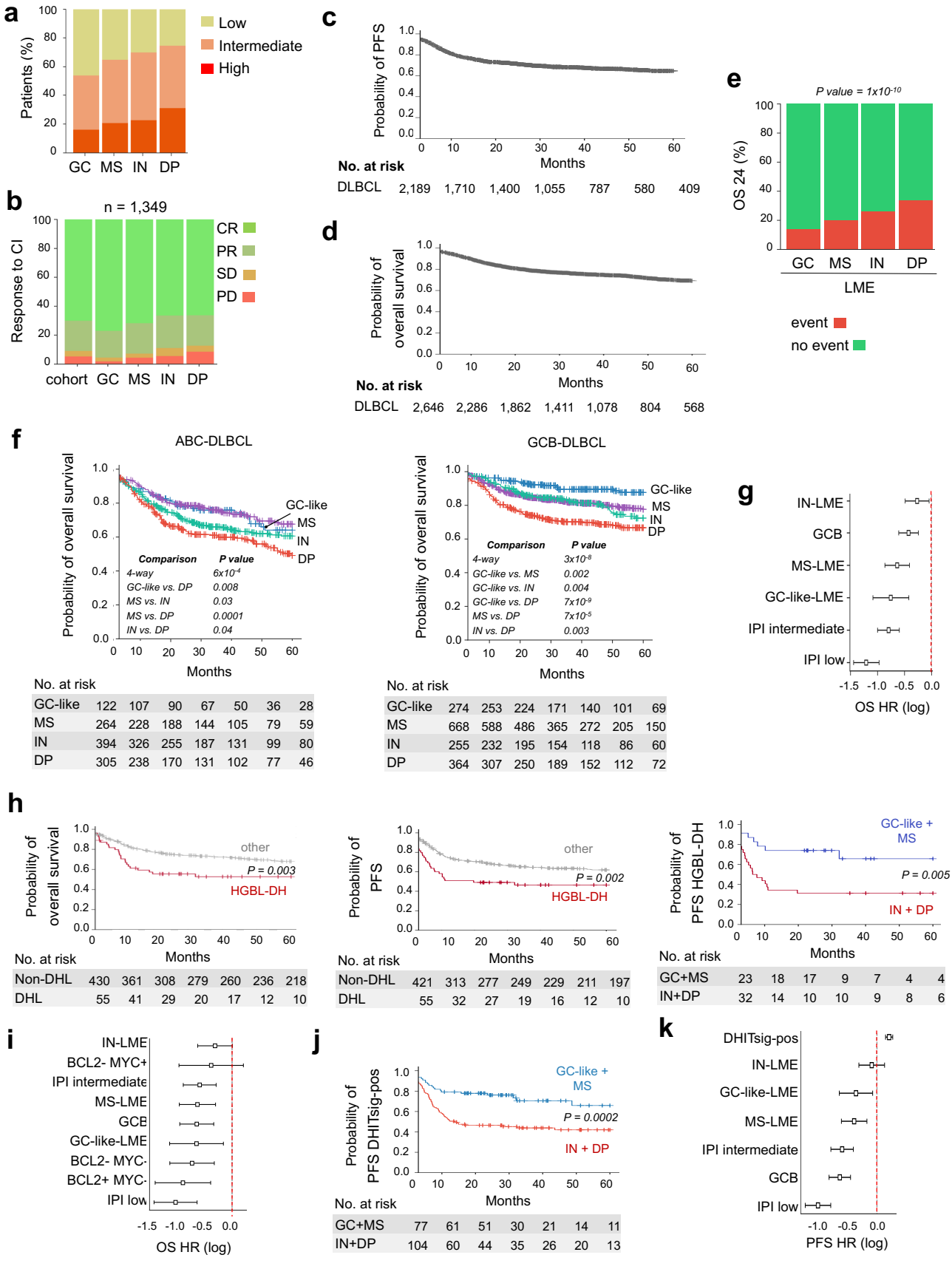
a
Patients (%)
GC MS IN DP
Low
Intermediate
High
b
n = 1,349
Response to CI
cohort GC MS IN DP
CR
PR
SD
PD
c
Probability of PFS
Months
No. at risk
DLBCL 2,189 1,710 1,400 1,055 787 580 409
d
Probability of overall survival
Months
No. at risk
DLBCL 2,646 2,286 1,862 1,411 1,078 804 568
e
P value = 1x10-10
OS 24 (%)
GC MS IN DP
LME
event
no event
f
ABC-DLBCL
Probability of overall survival
GC-like
MS
IN
DP
Comparison P value
4-way 6x10-4
GC-like vs. DP 0.008
MS vs. IN 0.03
MS vs. DP 0.0001
IN vs. DP 0.04
Months
No. at risk
GC-like 122 107 90 67 50 36 28
MS 264 228 188 144 105 79 59
IN 394 326 255 187 131 99 80
DP 305 238 170 131 102 77 46
GCB-DLBCL
Probability of overall survival
GC-like
MS
IN
DP
Comparison P value
4-way 3x10-8
GC-like vs. MS 0.002
GC-like vs. IN 0.004
GC-like vs. DP 7x10-9
MS vs. DP 7x10-5
IN vs. DP 0.003
Months
No. at risk
GC-like 274 253 224 171 140 101 69
MS 668 588 486 365 272 205 150
IN 255 232 195 154 118 86 60
DP 364 307 250 189 152 112 72
g
IN-LME
GCB
MS-LME
GC-like-LME
IPI intermediate
IPI low
OS HR (log)
h
Probability of overall survival
other
P = 0.003
HGBL-DH
Months
No. at risk
Non-DHL 430 361 308 279 260 236 218
DHL 55 41 29 20 17 12 10
Probability of PFS
other
P = 0.002
HGBL-DH
Months
No. at risk
Non-DHL 421 313 277 249 229 211 197
DHL 55 32 27 19 16 12 10
Probability of PFS HGBL-DH
GC-like + MS
P = 0.005
IN + DP
Months
No. at risk
GC+MS 23 18 17 9 7 4 4
IN+DP 32 14 10 10 9 8 6
i
IN-LME
BCL2- MYC+
IPI intermediate
MS-LME
GCB
GC-like-LME
BCL2- MYC-
BCL2+ MYC-
IPI low
OS HR (log)
j
Probability of PFS DHITsig-pos
GC-like + MS
P = 0.0002
IN + DP
Months
No. at risk
GC+MS 77 61 51 30 21 14 11
IN+DP 104 60 44 35 26 20 13
k
DHITsig-pos
IN-LME
GC-like-LME
MS-LME
IPI intermediate
GCB
IPI low
PFS HR (log)

Figure S7. a: Distribution of IPI status (low: 0-1, intermediate: 2-3 and high: 4-5) according to the four LME subtypes. **b:** Outcome (RECIL criteria) to rituximab-containing chemoimmunotherapy (CI) regimens in an unselected cohort of 1,349 DLBCL patients according to their LME category. **c:** Five-year progression-free survival (PFS) curve for the DLBCL cohort of 2,189 patients used in our analyses. **d:** Five-year overall survival (OS) curve for the DLBCL cohort of 2,646 patients used in our analyses. **e:** OS at 24 months (OS24) in DLBCL patients according to the LME category. Censored patients are not shown. **f:** Kaplan-Meier models of OS according to LME subtype in ABC- and GCB-DLBCLs. Only statistically significant comparisons shown. **g.** OS hazard ratio (HR) plots (log with 95% confidence interval) for LME category, COO subgroup and IPI. **h:** Kaplan-Meier models of OS and PFS for HGBL-DH vs. non-HGBL-DH and PFS for HGBL-DH segregated into favorable prognosis LMEs (GC-like and MS) vs. unfavorable prognosis LMEs (IN and DP). **i:** OS hazard ratio (HR) plots (log with 95% confidence interval, n = 486) for LME category, COO subgroup, IPI and BCL2 and MYC translocation status. **j:** Five-year PFS curves for DHITsig positive patients segregated into favorable prognosis LMEs (GC-like and MS) vs. unfavorable prognosis LMEs (IN and DP). **k.** PFS hazard ratio (HR) plots (log with 95% confidence interval, n = 2,024) for LME category, COO subgroup, IPI and DHITsig status.

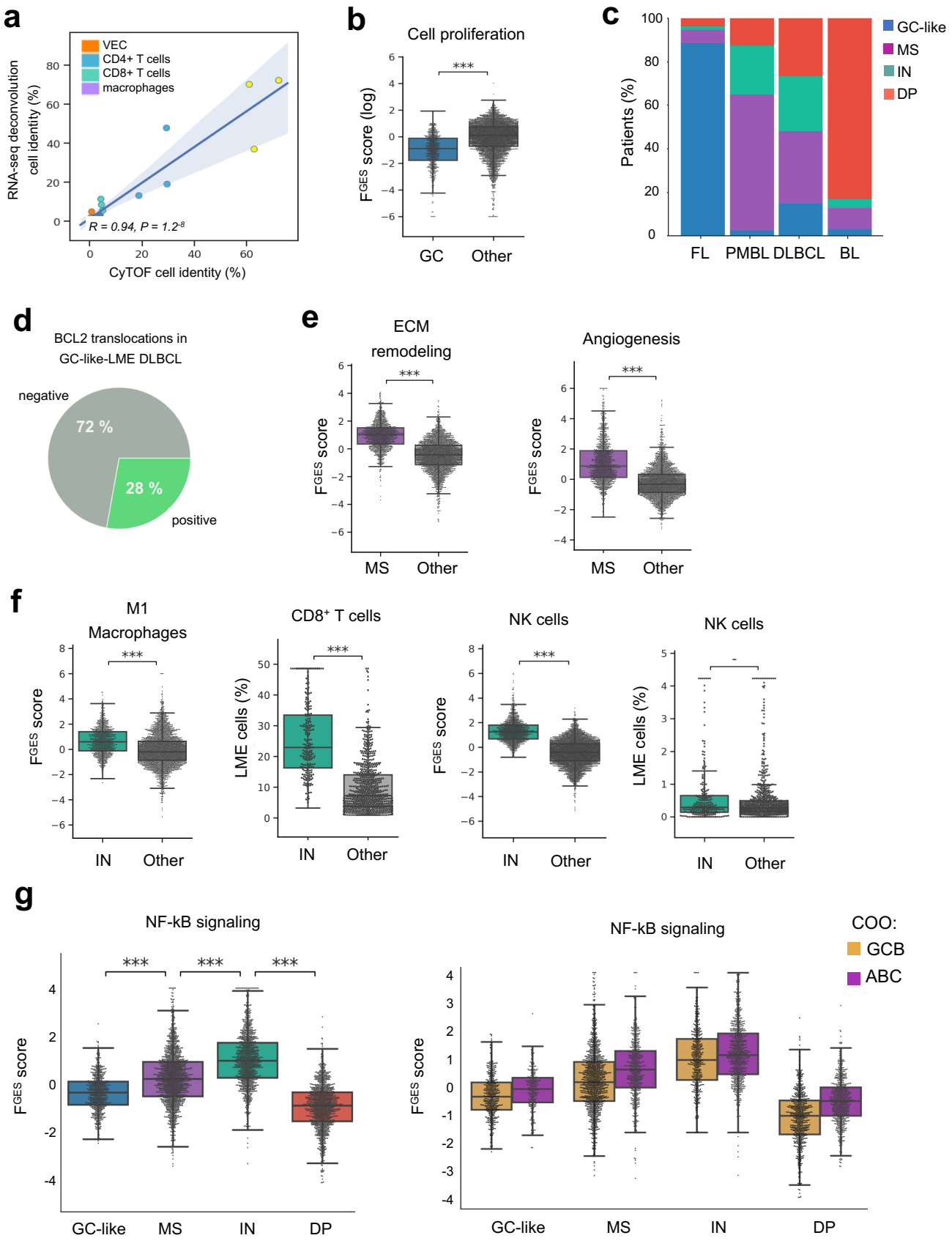
a
VEC
CD4+ T cells
CD8+ T cells
macrophages
RNA-seq deconvolution cell identity (%)
CyTOF cell identity (%)
R = 0.94, P = 1.2-8
b
Cell proliferation
FGES score (log)
***
GC
Other
c
Patients (%)
GC-like
MS
IN
DP
FL
PMBL
DLBCL
BL
d
BCL2 translocations in GC-like-LME DLBCL
negative
72 %
28 %
positive
e
ECM remodeling
Angiogenesis
FGES score
***
MS
Other
f
M1 Macrophages
CD8+ T cells
NK cells
NK cells
FGES score
LME cells (%)
***
IN
Other
g
NF-kB signaling
NF-kB signaling
COO:
GCB
ABC
FGES score
***
GC-like
MS
IN
DP

Figure S8. a: Comparison of cell identities and proportions between CyTOF and deconvolution algorithms in tonsil samples. **b:** F^{GES} for cell proliferation in GC-like-LME DLBCLs vs. other LMEs. **c:** DLBCL LME distribution according to B-cell lymphoma subtypes follicular lymphoma (FL), primary mediastinal B-cell lymphoma (PMBL) and Burkitt lymphoma (BL). **d:** BCL2 translocations in GC-like-LME DLBCL patients. **e:** F^{GES} for ECM remodeling and angiogenesis in MS-LME DLBCLs vs. other LMEs. **f:** Proportion of LME cells significantly enriched in the IS-LME subtype obtained by cell deconvolution algorithms or F^{GES}. **g:** NF-kB F^{GES} in DLBCL samples according to LME and to LME by COO. ****p<0.001.*

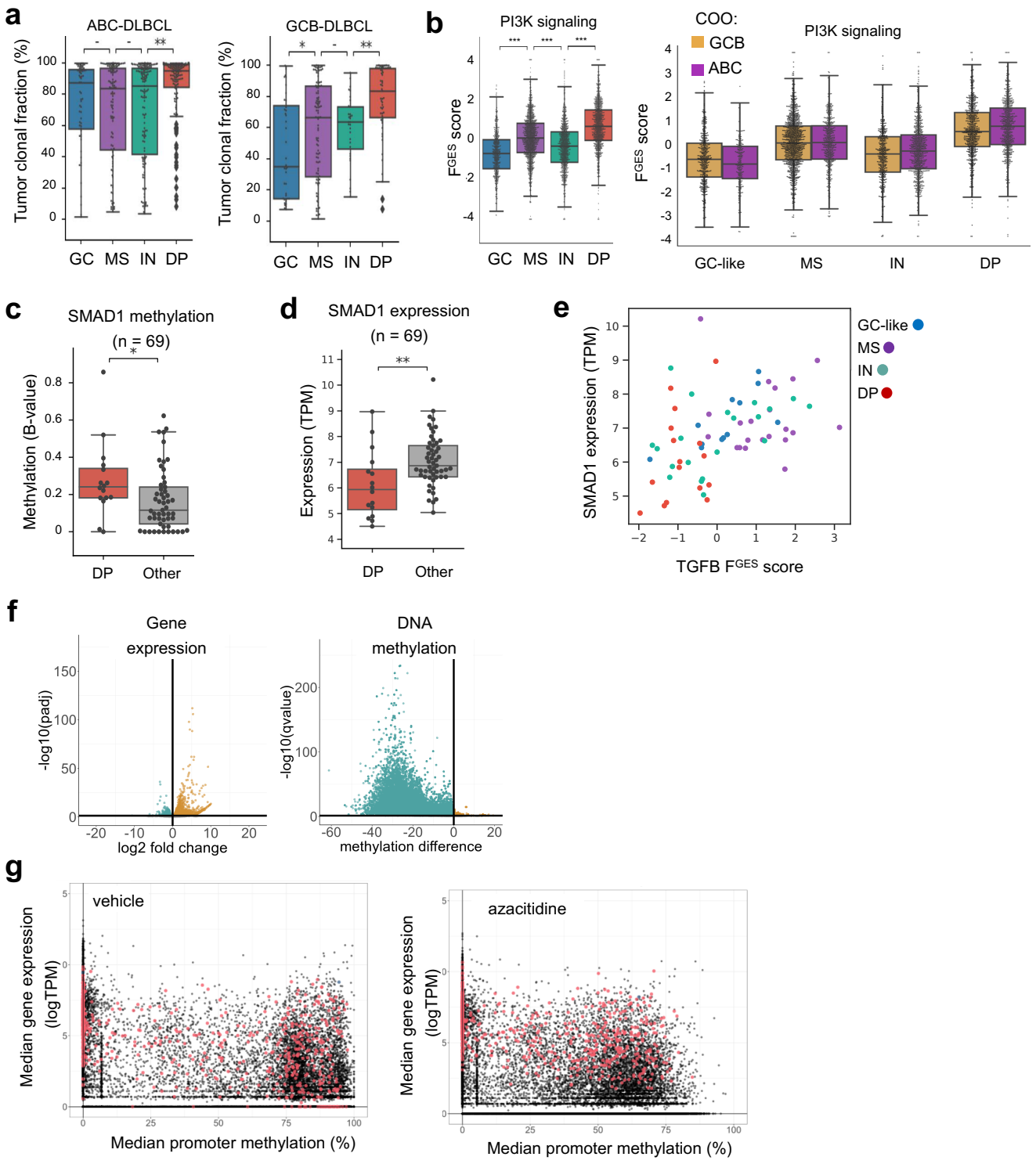

Figure S9. a: Tumor clonal fraction determined by IGH clonality in ABC- and GCB-DLBCL distributed by LME. **b:** PI3K F^{GES} in DLBCL samples according to LME and to LME by COO. **c-d:** SMAD1 promoter methylation (B-value) and SMAD1 expression in 69 DLBCL samples. **e:** scatter plot between SMAD1 expression and TGFB pathway activity in 69 DLBCL samples according to their LME subtype. **f:** *Left:*

volcano plot of differential gene expression. Upregulated and downregulated genes with values of p-adjusted < 0.05 upon azacitidine treatment are shown in orange and green, respectively. Gray dots represent non-significant genes according to the selected p-adjusted value threshold. *Right:* volcano plot of differential promoter region methylation. Increased and decreased promoter methylation with p-adjusted < 0.05 upon azacitidine treatment are shown in orange and green, respectively. Gray dots represent non-significant regions according to the selected threshold. **g:** Scatter plot representing the median promoter methylation and median gene expression per gene for vehicle-treated samples (left) and azacitidine treated samples (right). Genes that show both significant gene expression changes (padj< 0.05) and significant differences in promoter region methylation (q-value <0.05, absolute methylation difference > 10) are depicted in red. **p<0.05, **p<0.01, ***p<0.001*

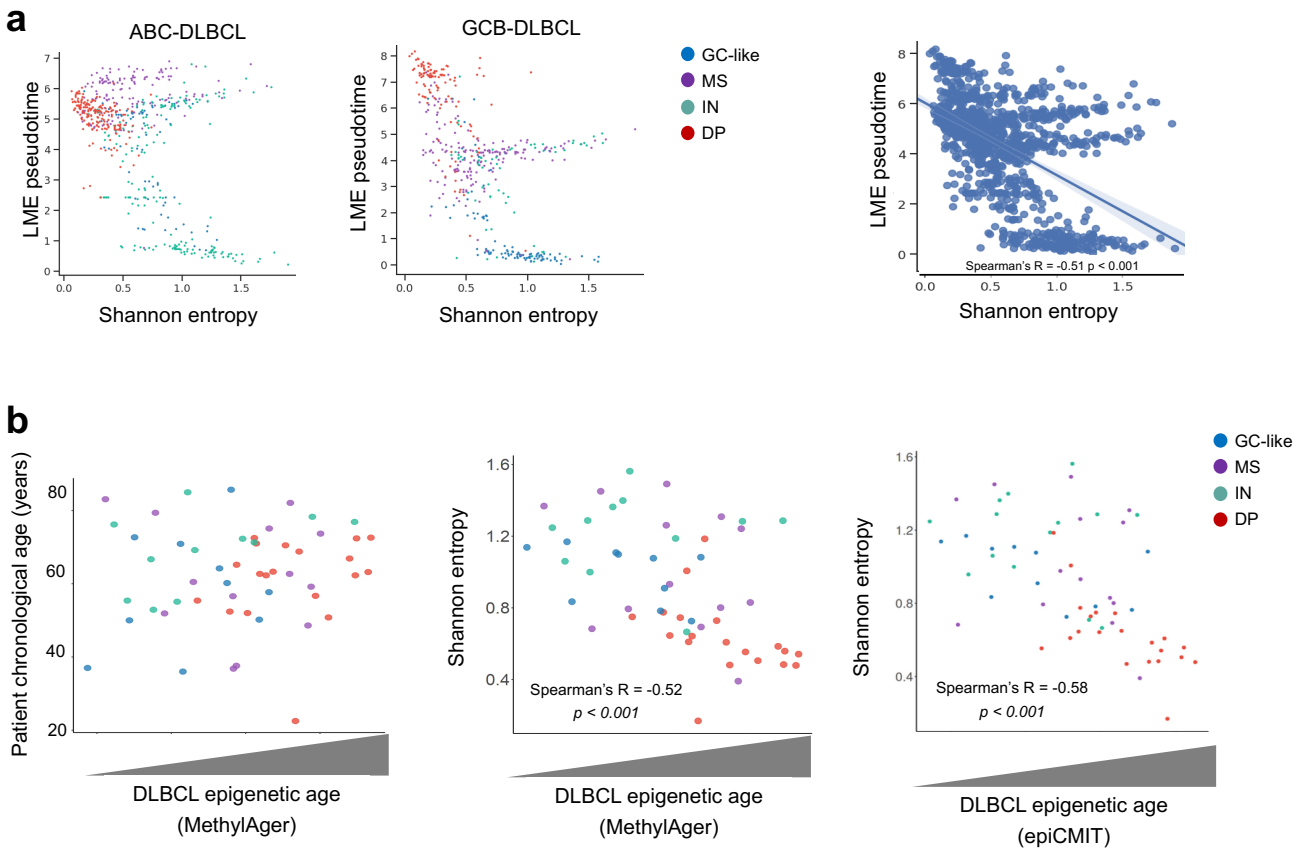

Figure S10. a: Correlation between LME pseudotime (Y-axis) and LME Shannon entropy (X-axis) in ABC- and GCB-DLBCLs. LME cluster indicated by colors. Correlation line and statistics for all the samples. **b:** Correlation between the chronological age of the patient (Y-axis) and the epigenetic age of the lymphoma (X-axis) measured using MethylAger scoring. Correlations between LME Shannon entropy (Y-axis) and the epigenetic age of the lymphoma (X-axis) measured using MethylAger and epiCMIT methods.

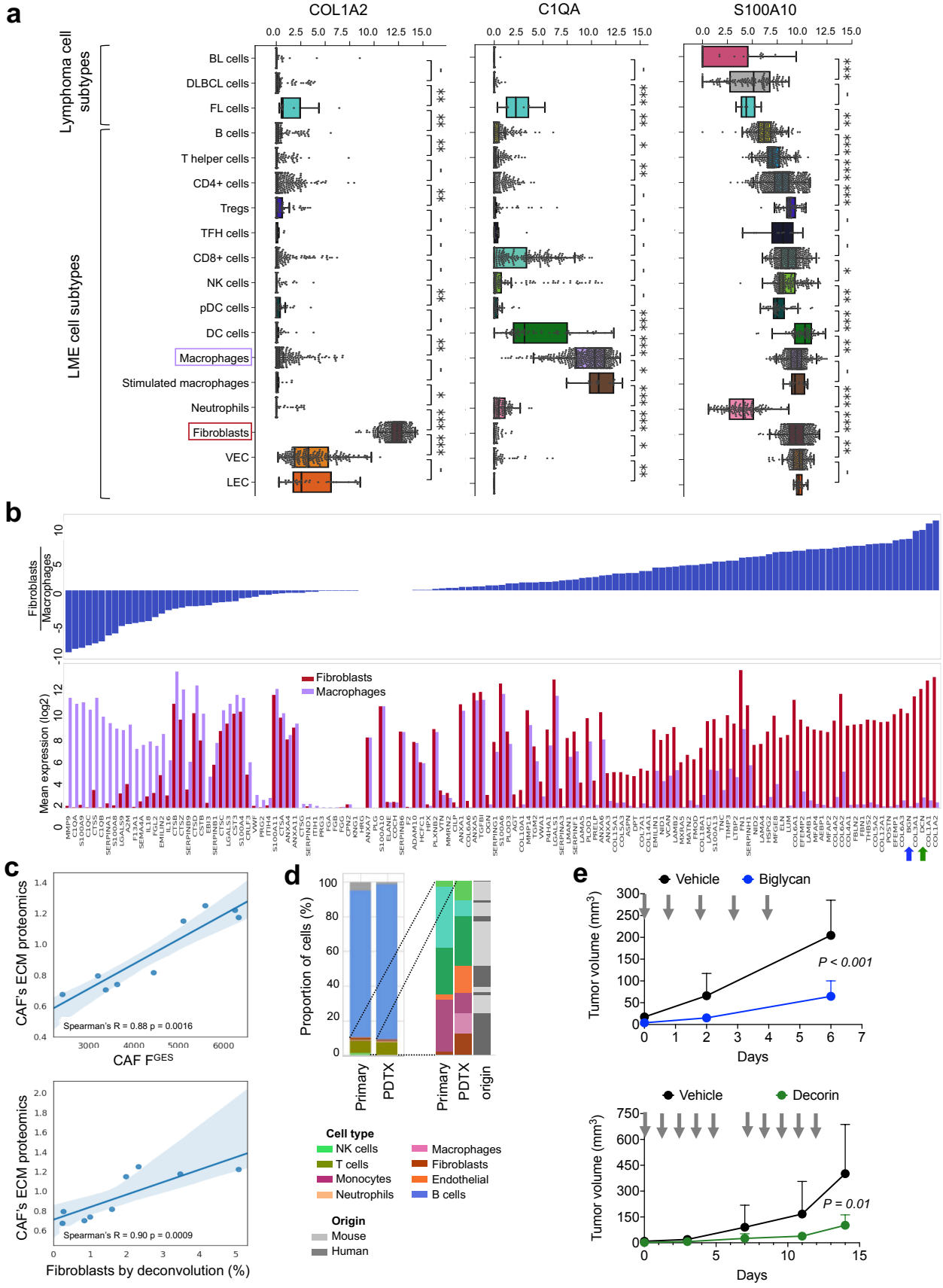
a
COL1A2
C1QA
S100A10
Lymphoma cell subtypes
LME cell subtypes
BL cells
DLBCL cells
FL cells
B cells
T helper cells
CD4+ cells
Tregs
TFH cells
CD8+ cells
NK cells
pDC cells
DC cells
Macrophages
Stimulated macrophages
Neutrophils
Fibroblasts
VEC
LEC
b
Fibroblasts
Macrophages
Mean expression (log2)
c
CAF's ECM proteomics
Spearman's R = 0.88 p = 0.0016
CAF F^GES
Spearman's R = 0.90 p = 0.0009
Fibroblasts by deconvolution (%)
d
Proportion of cells (%)
Primary
PDTX
origin
Cell type
NK cells
T cells
Monocytes
Neutrophils
Macrophages
Fibroblasts
Endothelial
B cells
Origin
Mouse
Human
e
Vehicle
Biglycan
Tumor volume (mm3)
P < 0.001
Days
Decorin
P = 0.01

Figure S11. a: Median gene expression of selected matrisome genes *COL1A2, C1QA* and *S100A10* in lymphoma and LME cellular subtypes. **b:** Ranking of transcripts from matrisome coding proteins according to the association with fibroblasts (red) or macrophages (purple). Ratio between fibroblasts and macrophages (top). Relative abundance of the transcript for each of these cell types (bottom). The arrows indicate the position of decorin (DCN) and biglycan (BGN). **c:** Spearman's correlation between the abundance of the 10 top proteins associated with CAFs (Y-axis) and the CAF FGES (top) and proportion of fibroblasts by signature deconvolution (bottom) in 9 DLBCL samples. **d:** Cellular composition by transcript deconvolution of primary DLBCL LME-DP and its PDTX2 model. Expansion of the LME cellular composition in the primary and PDTX2 model and indication whether these cells in PDTX2 are of human or mouse origin. **e.** DLBCL growth curve of DLBCL LME-DP PDTX2 treated with vehicle vs. BGN (top) or vehicle vs. DCN (bottom). Drugs were administered in the days indicated by arrows starting 10 days after tumor implantation (day 0). *$*p<0.05$, $**p<0.01$, $***p<0.001$.*

Supplementary Tables

Table S1. Gene sets of F^{GES}

Table S2. COO classifier features

Table S3. DHITsig classifier features

Table S4. Prevalence of mutations across LME categories

Table S5. Prevalence of CNAs across LME categories

Table S6. Upregulated and hypomethylated genes in murine A20 lymphomas upon azacitidine treatment and pathway analysis.

Table S7. DLBCL matrisome

References

1. James KR, Gomes T, Elmentaite R, Kumar N, Gulliver EL, King HW*, et al.* Distinct microbial and immune niches of the human colon. Nat Immunol **2020**;21(3):343-53 doi 10.1038/s41590-020-0602-z.
2. Erdmann T, Klener P, Lynch JT, Grau M, Vockova P, Molinsky J*, et al.* Sensitivity to PI3K and AKT inhibitors is mediated by divergent molecular mechanisms in subtypes of DLBCL. Blood **2017**;130(3):310-22 doi 10.1182/blood-2016-12-758599.
3. Kloo B, Nagel D, Pfeifer M, Grau M, Duwel M, Vincendeau M*, et al.* Critical role of PI3K signaling for NF-kappaB-dependent survival in a subset of activated B-cell-like diffuse large B-cell lymphoma cells. Proc Natl Acad Sci U S A **2011**;108(1):272-7 doi 10.1073/pnas.1008969108.
4. Carr M, Mamand S, Chapman KL, Perrior T, Wagner SD. IKKepsilon and TBK1 in diffuse large B-cell lymphoma: A possible mechanism of action of an IKKepsilon/TBK1 inhibitor to repress NF-kappaB and IL-10 signalling. J Cell Mol Med **2020**;24(19):11573-82 doi 10.1111/jcmm.15774.