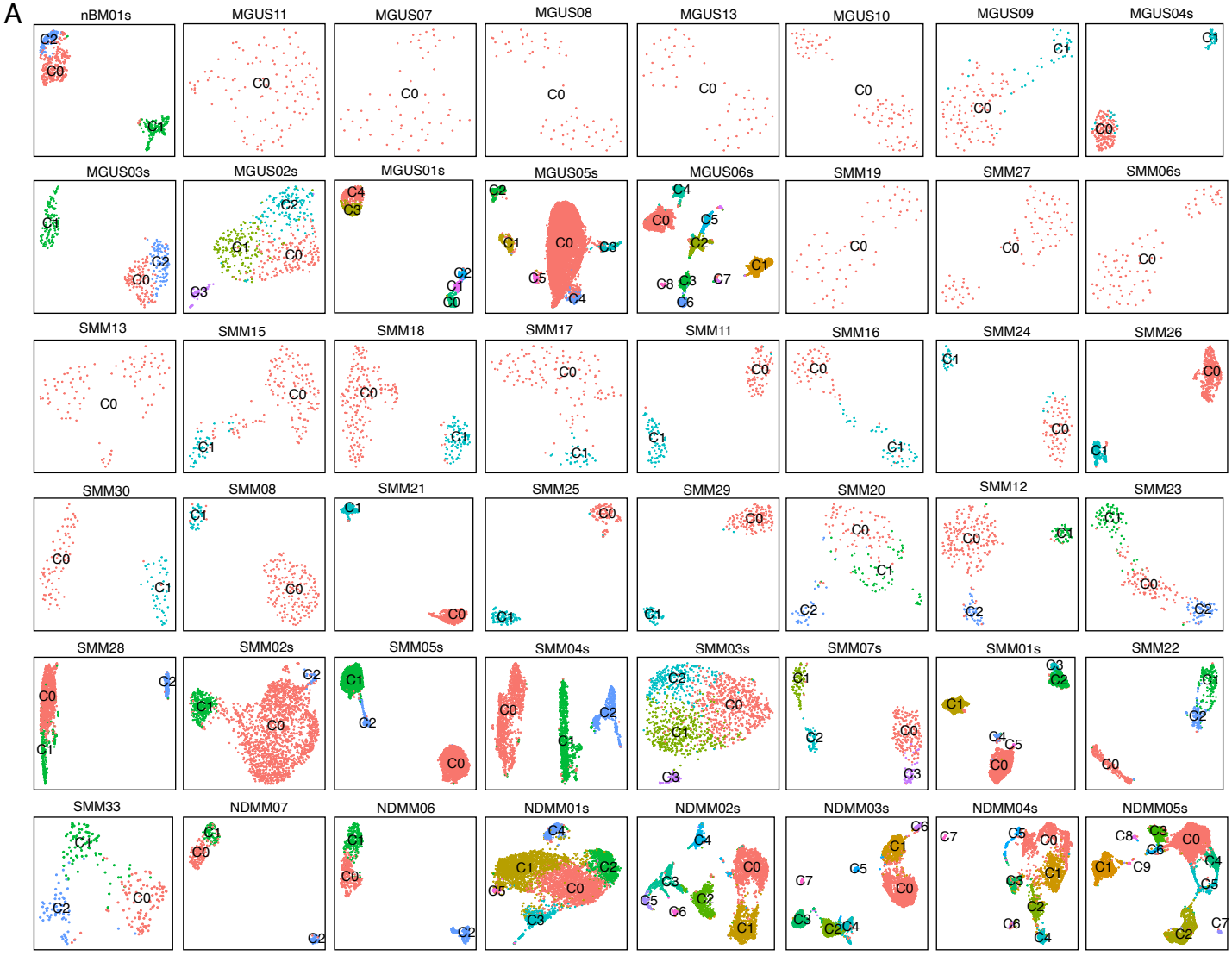
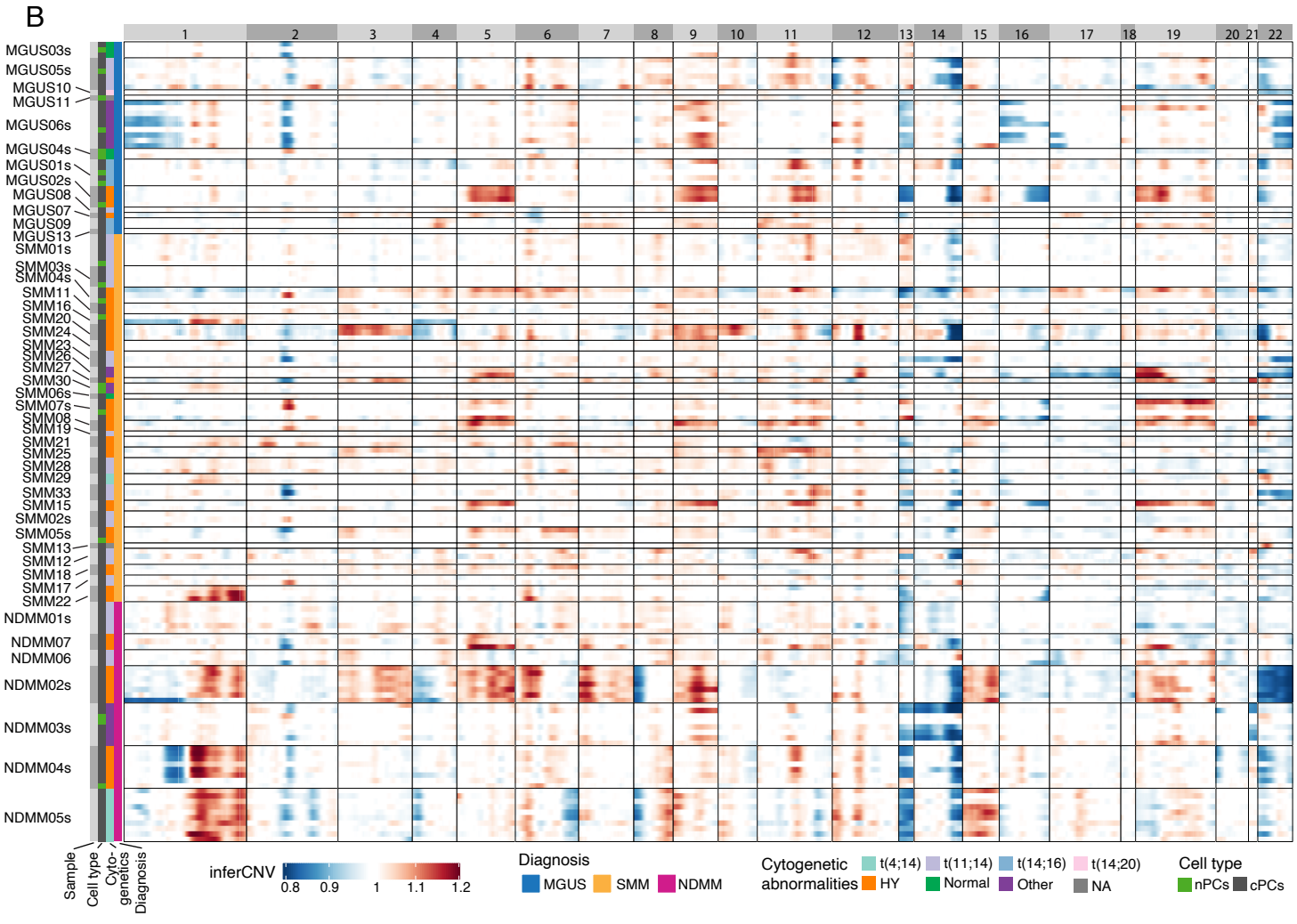
Supplementary Figures and Legends

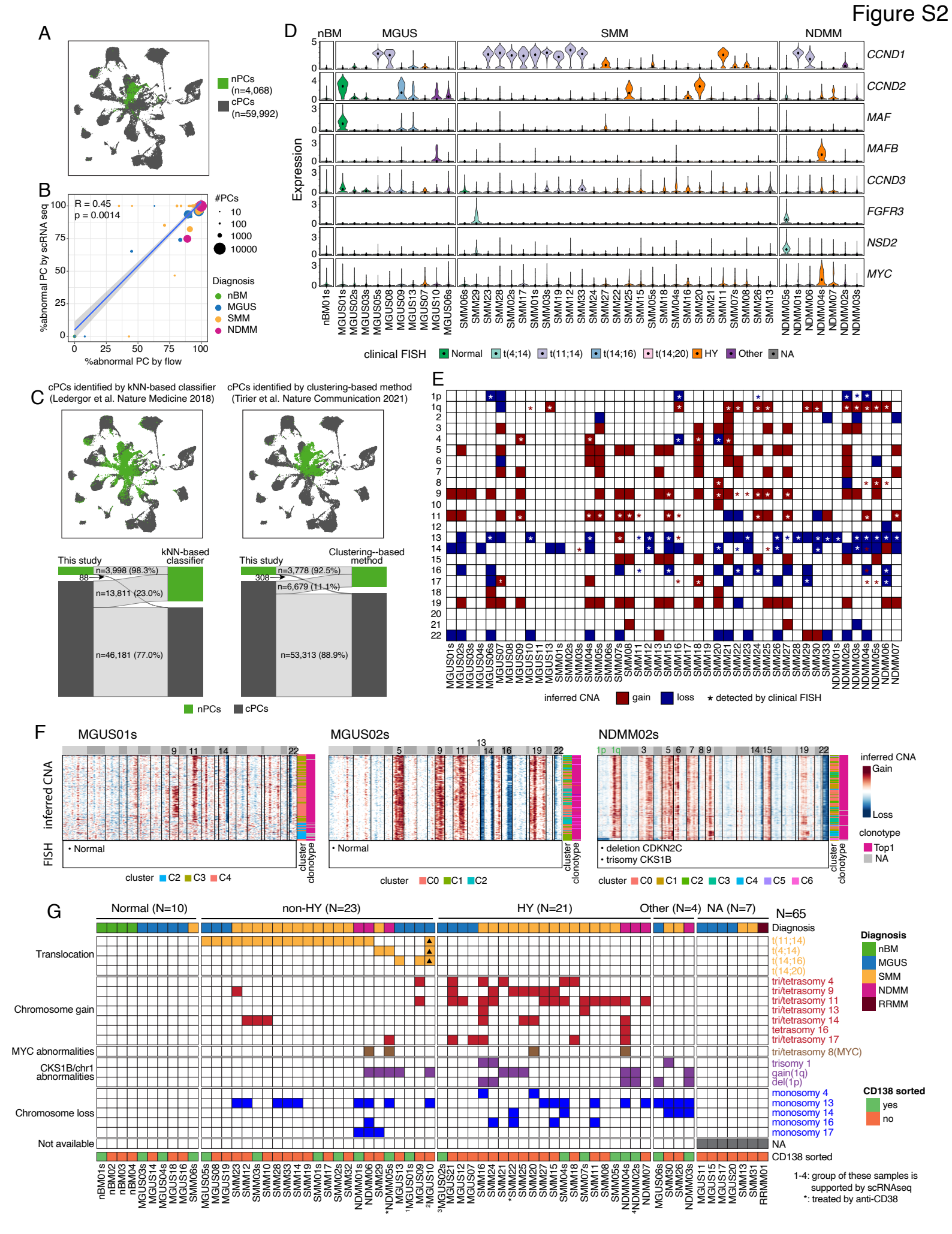
Supplementary Figure S1. Intra-tumoral heterogeneity as shown by unsupervised clustering and intra-sample subcluster level CNA profiles, related to Figure 1. (A) UMAP visualization of unsupervised clustering analysis of PCs at patient level in 48 patients with more than 50 PCs. Patients were grouped by their diagnosis. **(B)** Heatmap showing genome wide inferred CNA profiles for 150 PC clusters from 47 patients except the normal one. Clusters were firstly grouped by their patient origins and then by diagnosis. Columns represent genomic regions, ordered by genome position across all chromosomes. Rows represent cluster averaged CNAs, with one row per cluster.


A
nBM01s
MGUS11
MGUS07
MGUS08
MGUS13
MGUS10
MGUS09
MGUS04s
MGUS03s
MGUS02s
MGUS01s
MGUS05s
MGUS06s
SMM19
SMM27
SMM06s
SMM13
SMM15
SMM18
SMM17
SMM11
SMM16
SMM24
SMM26
SMM30
SMM08
SMM21
SMM25
SMM29
SMM20
SMM12
SMM23
SMM28
SMM02s
SMM05s
SMM04s
SMM03s
SMM07s
SMM01s
SMM22
SMM33
NDMM07
NDMM06
NDMM01s
NDMM02s
NDMM03s
NDMM04s
NDMM05s
48 samples (1 nBM, 12 MGUS, 28 SMM, 7 NDMM), 153 intra-sample subclusters


B
MGUS03s
MGUS05s
MGUS10
MGUS11
MGUS06s
MGUS04s
MGUS01s
MGUS02s
MGUS08
MGUS07
MGUS09
MGUS13
SMM01s
SMM03s
SMM04s
SMM11
SMM16
SMM20
SMM24
SMM23
SMM26
SMM27
SMM30
SMM06s
SMM07s
SMM08
SMM19
SMM21
SMM25
SMM28
SMM29
SMM33
SMM15
SMM02s
SMM05s
SMM13
SMM12
SMM18
SMM17
SMM22
NDMM01s
NDMM07
NDMM06
NDMM02s
NDMM03s
NDMM04s
NDMM05s
Sample
Cell type
Cyto-genetics
Diagnosis
inferCNV
0.8 0.9 1 1.1 1.2
Diagnosis
MGUS SMM NDMM
Cytogenetic abnormalities
t(4;14) t(11;14) t(14;16) t(14;20)
HY Normal Other NA
Cell type
nPCs cPCs

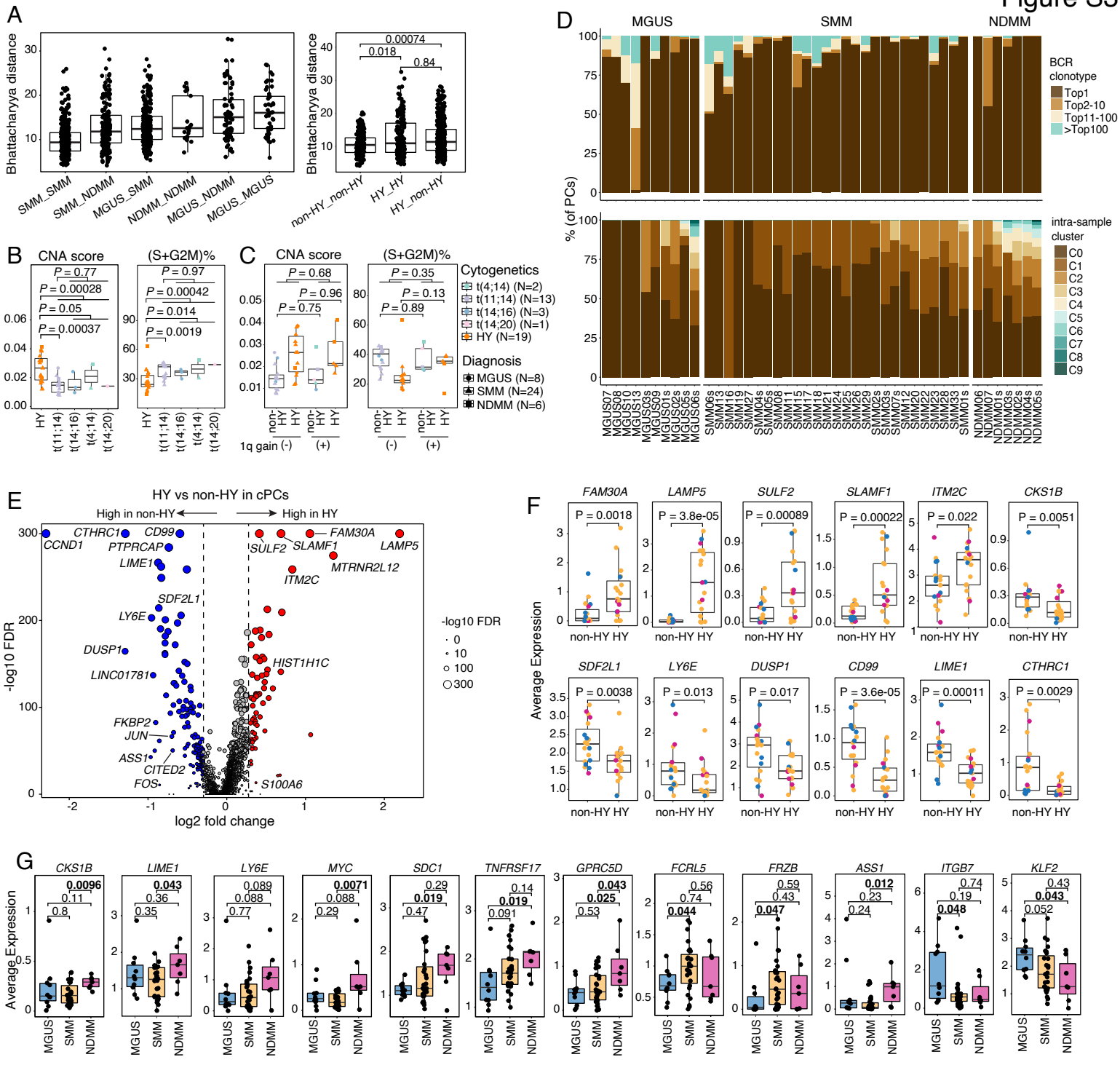
Supplementary Figure S2. Distribution of cells across patients, comparison of FISH and scRNAseq for abnormalities detection, and determination of genomic drivers of transformation for each patient, related to Figure 1. (A) UMAP visualization of unsupervised clustering analysis of all PCs (n = 64,078) from 48 patients with more than 50 PCs. Cells were color coded by normal or clonal. **(B)** Comparison of proportion of abnormal plasma cells in all plasma cells measured by scRNAseq and flow cytometry. **(C)** UMAP showing the distribution of cPCs identified by KNN-based classifier (top left) or clustering-based method (top right). Alluvial plot at the bottom showed true positive, false negative, true negative, false negative of KNN-based (bottom left) or clustering-based (bottom right) method for identifying cPCs when using the results in this study as truth. **(D)** Violin plots showing expression of translocation related genes in cPCs in 44 patients or in nPCs in 1 healthy donor as shown in Fig. 1E but grouped by their diagnosis and color coded by their cytogenetic abnormalities determined by clinical FISH. **(E)** Comparisons of chromosome abnormalities detected by clinical FISH and scRNAseq. Chromosome and arm level gain and loss from scRNAseq were visually judged from inferred CNA profiles. **(F)** Comparisons of FISH results and inferred copy number alternation profiles from sing cell RNA sequencing data for MGUS01s, MGUS02s and NDMM02s. Chromosome regions with obvious CNAs were labelled on the top. **(G)** Fluorescence in situ hybridization (FISH) results of 65 samples in this study as shown in Fig. 1B but grouped by their cytogenetic abnormalities.

Figure S2


A
nPCs (n=4,068)
cPCs (n=59,992)
B
R = 0.45
p = 0.0014
%abnormal PC by scRNA seq
%abnormal PC by flow
#PCs
10
100
1000
10000
Diagnosis
nBM
MGUS
SMM
NDMM
C
cPCs identified by kNN-based classifier (Ledergor et al. Nature Medicine 2018)
cPCs identified by clustering-based method (Tirier et al. Nature Communication 2021)
This study
kNN-based classifier
n=3,998 (98.3%)
88
n=13,811 (23.0%)
n=46,181 (77.0%)
This study
Clustering--based method
n=3,778 (92.5%)
308
n=6,679 (11.1%)
n=53,313 (88.9%)
nPCs
cPCs
D
nBM
MGUS
SMM
NDMM
Expression
CCND1
CCND2
MAF
MAFB
CCND3
FGFR3
NSD2
MYC
clinical FISH
Normal
t(4;14)
t(11;14)
t(14;16)
t(14;20)
HY
Other
NA
E
inferred CNA
gain
loss
detected by clinical FISH
F
MGUS01s
MGUS02s
NDMM02s
inferred CNA
FISH
Normal
deletion CDKN2C
trisomy CKS1B
cluster
clonotype
Gain
Loss
Top1
NA
G
Normal (N=10)
non-HY (N=23)
HY (N=21)
Other (N=4)
NA (N=7)
N=65
Diagnosis
Translocation
Chromosome gain
MYC abnormalities
CKS1B/chr1 abnormalities
Chromosome loss
Not available
CD138 sorted
t(11;14)
t(4;14)
t(14;16)
t(14;20)
tri/tetrasomy 4
tri/tetrasomy 9
tri/tetrasomy 11
tri/tetrasomy 13
tri/tetrasomy 14
tetrasomy 16
tri/tetrasomy 17
tri/tetrasomy 8(MYC)
trisomy 1
gain(1q)
del(1p)
monosomy 4
monosomy 13
monosomy 14
monosomy 16
monosomy 17
NA
yes
no
1-4: group of these samples is supported by scRNAseq
*: treated by anti-CD38

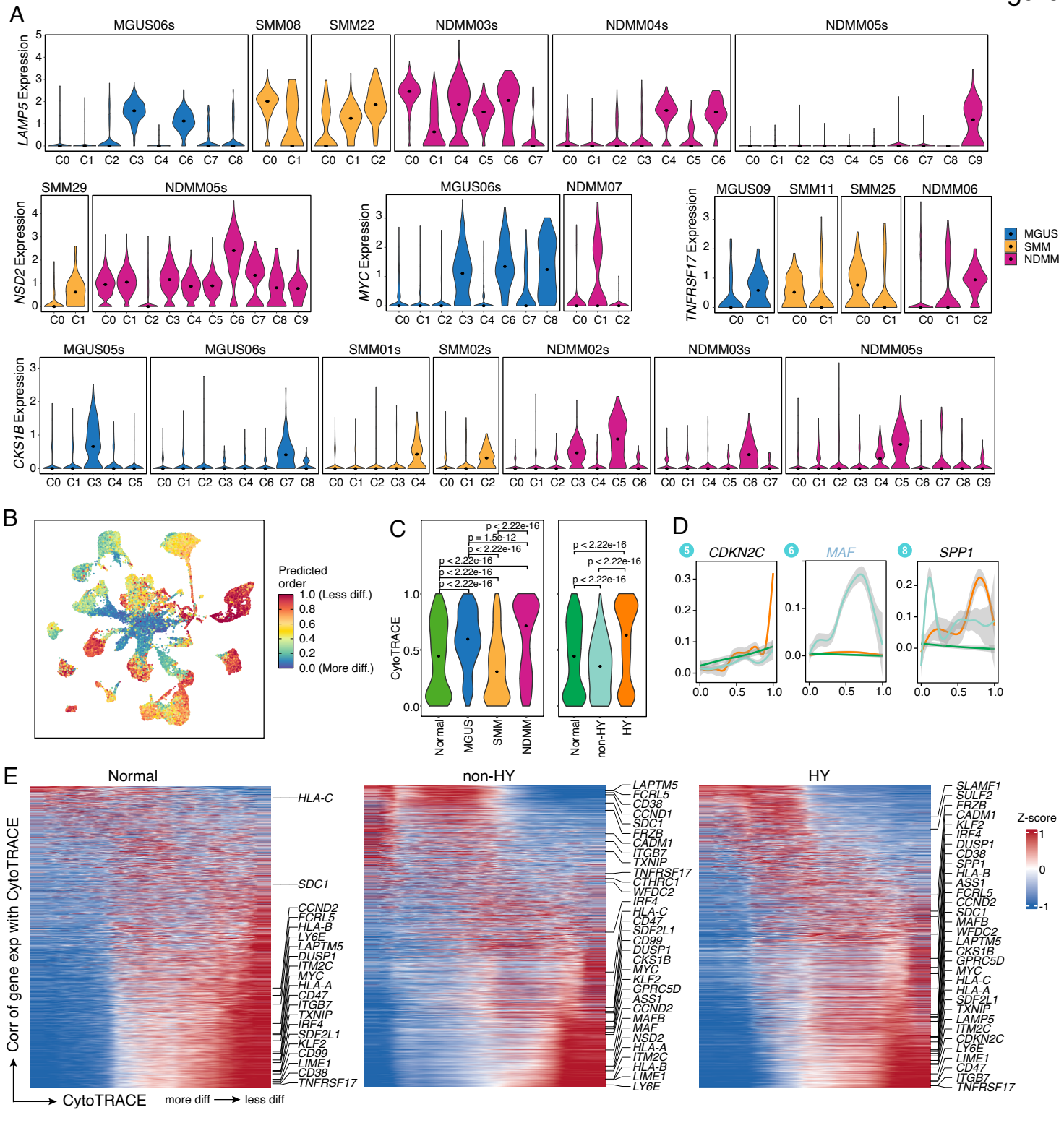
Supplementary Figure S3. Characterization of cPCs, related to Figure 2. (A) Boxplot showing the comparisons of Bhattacharyya distance between different diagnosis or cytogenetics groups. Groups were ordered by their median values. **(B)** Boxplots showing the comparisons of sample-averaged CNA score and proportion of cells at S and G2M phase between HY and different translocation groups. Dots were color coded by their cytogenetic abnormalities and shape coded by diagnosis. **(C)** Boxplots showing the comparisons of sample-averaged CNA score and proportion of cells at S and G2M phase between samples with (N=10) and without (N=28) 1q gain. Dots were color coded by their cytogenetics abnormalities and shape coded by disease stages. **(D)** Stacked bar charts showing BCR clonotype (upper) and intra-sample subcluster (lower) distributions in each sample grouped by disease stage. Clonotypes were grouped and colored by clonotype frequency categories. **(E)** Volcano plots of differentially expressed genes between cPCs from non-HY samples and HY samples. FDR, two-sided Wilcoxon rank sum test with Bonferonni correction. Dashed line, log2FC > 0.3 or < -0.3. Labels, top ranked DEGs based on log2FC. **(F)** Boxplots showing the comparisons of averaged expression of top DEGs in cPCs between non-HY (N=19) and HY (N=19) at sample level. Dots were color coded by diagnosis. **(G)** Boxplots showing the comparisons of averaged expression of MM related key genes in cPCs between MGUS (N=10), SMM (N=27) and NDMM (N=7).

Figure S3


A
Bhattacharyya distance
SMM_SMM
SMM_NDMM
MGUS_SMM
NDMM_NDMM
MGUS_NDMM
MGUS_MGUS
0.00074
0.018
0.84
non-HY_non-HY
HY_HY
HY_non-HY
B
CNA score
(S+G2M)%
P = 0.77
P = 0.00028
P = 0.05
P = 0.00037
P = 0.97
P = 0.00042
P = 0.014
P = 0.0019
HY
t(11;14)
t(14;16)
t(4;14)
t(14;20)
C
CNA score
(S+G2M)%
P = 0.68
P = 0.96
P = 0.75
P = 0.35
P = 0.13
P = 0.89
1q gain (-)
(+)
Cytogenetics
t(4;14) (N=2)
t(11;14) (N=13)
t(14;16) (N=3)
t(14;20) (N=1)
HY (N=19)
Diagnosis
MGUS (N=8)
SMM (N=24)
NDMM (N=6)
D
MGUS
SMM
NDMM
% (of PCs)
BCR clonotype
Top1
Top2-10
Top11-100
>Top100
intra-sample cluster
C0
C1
C2
C3
C4
C5
C6
C7
C8
C9
E
HY vs non-HY in cPCs
High in non-HY
High in HY
-log10 FDR
log2 fold change
CTHRC1
CD99
CCND1
PTPRCAP
LIME1
SDF2L1
LY6E
DUSP1
LINC01781
FKBP2
JUN
ASS1
CITED2
FOS
FAM30A
SULF2
SLAMF1
LAMP5
MTRNR2L12
ITM2C
HIST1H1C
S100A6
-log10 FDR
0
10
100
300
F
FAM30A
LAMP5
SULF2
SLAMF1
ITM2C
CKS1B
SDF2L1
LY6E
DUSP1
CD99
LIME1
CTHRC1
Average Expression
non-HY HY
G
CKS1B
LIME1
LY6E
MYC
SDC1
TNFRSF17
GPRC5D
FCRL5
FRZB
ASS1
ITGB7
KLF2
Average Expression
MGUS
SMM
NDMM

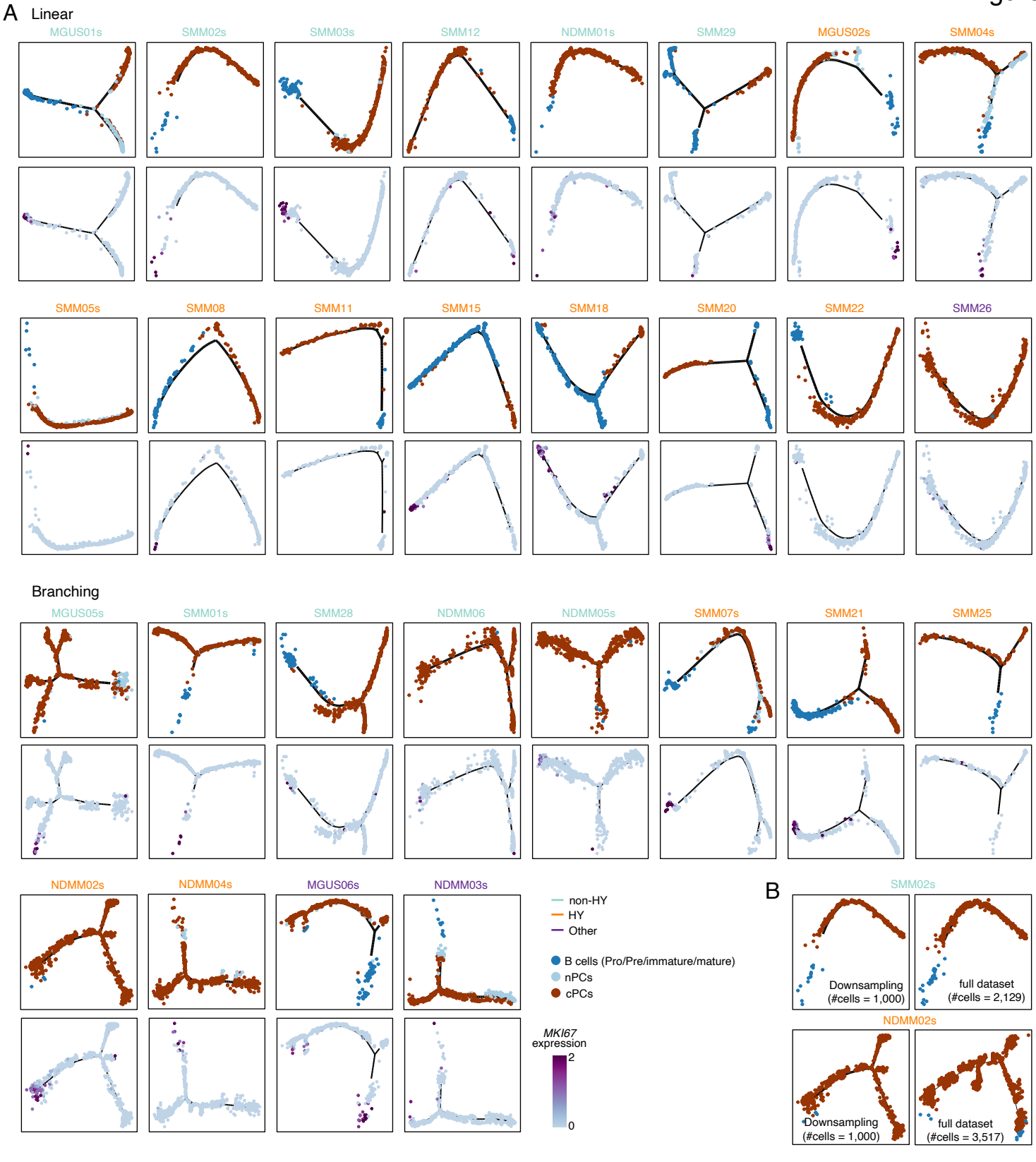
Supplementary Figure S4. Dynamics of oncogene expression, related to Figure 3. (A) Violin plots showing expression of *LAMP5*, *NSD2*, *MYC*, *TNFRSF17* and *CKS1B* in subclusters of related samples. **(B)** UMAP visualization of all PCs color coded by CytoTRACE score. **(C)** Violin plots showing distribution of CytoTRACE scores in cells with different disease stages and different cytogenetic abnormalities. **(D)** Two-dimensional plots showing the dynamic changes in expression levels of *CDKN2C*, *MAF* and *SPP1* (which show very low expression level) along the CytoTRACE score. Shaded band, 95th confidence interval. **(E)** Heatmap showing the expression of the top 2,000 highly variable genes along the CytoTRACE score in normal, non-HY and HY cells separately. Genes were ordered by their correlation with CytoTRACE score. Genes in (A) were labelled on the right.

Figure S4


A
MGUS06s
SMM08
SMM22
NDMM03s
NDMM04s
NDMM05s
LAMP5 Expression
SMM29
NDMM05s
NSD2 Expression
MGUS06s
NDMM07
MYC Expression
MGUS09
SMM11
SMM25
NDMM06
TNFRSF17 Expression
MGUS
SMM
NDMM
MGUS05s
MGUS06s
SMM01s
SMM02s
NDMM02s
NDMM03s
NDMM05s
CKS1B Expression
B
Predicted order
1.0 (Less diff.)
0.8
0.6
0.4
0.2
0.0 (More diff.)
C
CytoTRACE
p < 2.22e-16
p = 1.5e-12
Normal
MGUS
SMM
NDMM
non-HY
HY
D
5
CDKN2C
6
MAF
8
SPP1
E
Normal
non-HY
HY
Corr of gene exp with CytoTRACE
CytoTRACE
more diff
less diff
Z-score
HLA-C
SDC1
CCND2
FCRL5
HLA-B
LY6E
LAPTM5
DUSP1
ITM2C
MYC
HLA-A
CD47
ITGB7
TXNIP
IRF4
SDF2L1
KLF2
CD99
LIME1
CD38
TNFRSF17
LAPTM5
FCRL5
CD38
CCND1
SDC1
FRZB
CADM1
ITGB7
TXNIP
TNFRSF17
CTHRC1
WFDC2
IRF4
HLA-C
CD47
SDF2L1
CD99
DUSP1
CKS1B
MYC
KLF2
GPRC5D
ASS1
CCND2
MAFB
MAF
NSD2
HLA-A
ITM2C
HLA-B
LIME1
LY6E
SLAMF1
SULF2
FRZB
CADM1
KLF2
IRF4
DUSP1
CD38
SPP1
HLA-B
ASS1
FCRL5
CCND2
SDC1
MAFB
WFDC2
LAPTM5
CKS1B
GPRC5D
MYC
HLA-C
HLA-A
SDF2L1
TXNIP
LAMP5
ITM2C
CDKN2C
LY6E
LIME1
CD47
ITGB7
TNFRSF17

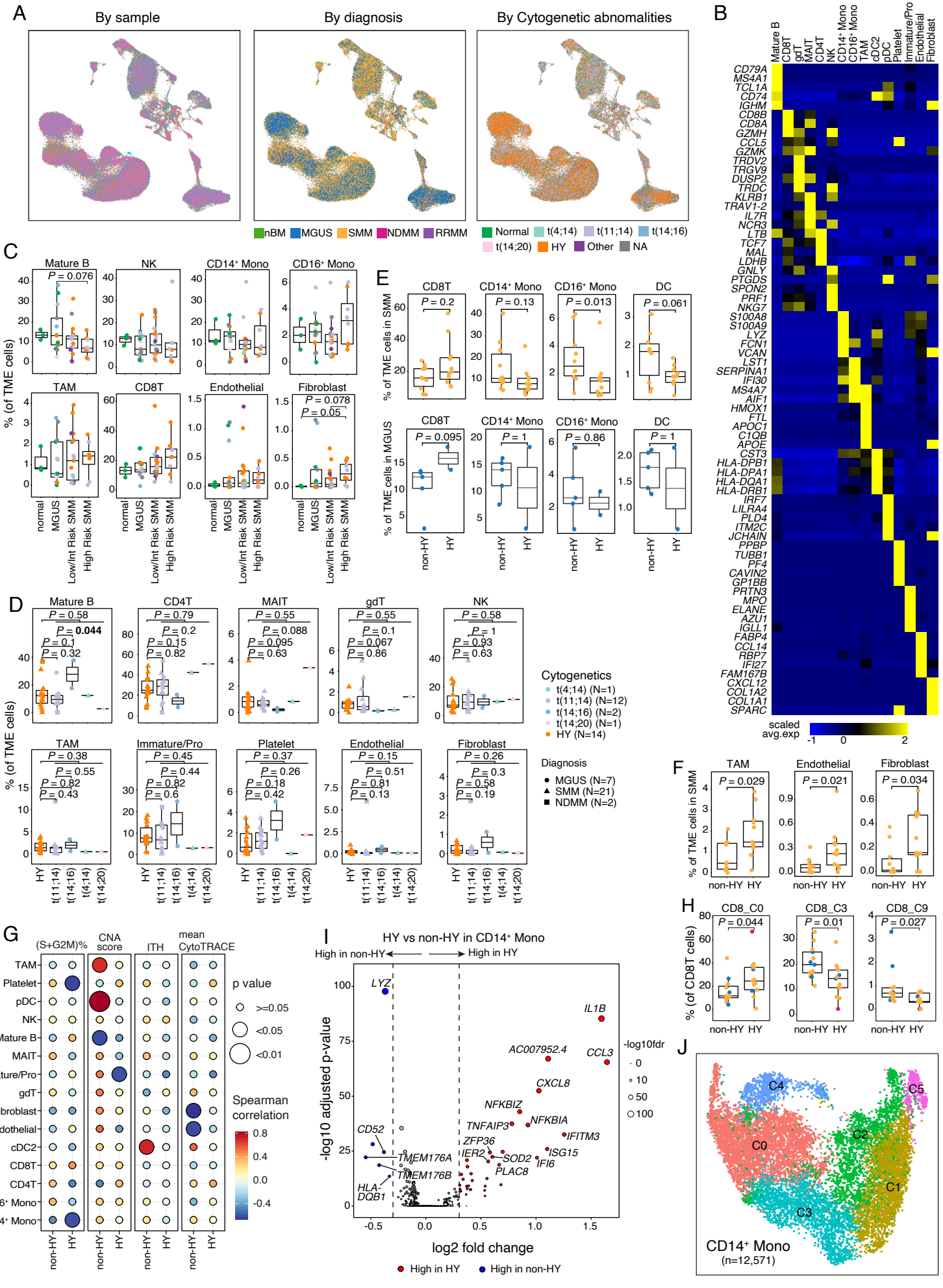
Supplementary Figure S5. Pseudotime trajectory analysis of B lineage cells, related to Figures 4 and 5. (A) Trajectories reconstructed using Monocle2 with cell type and expression of *MKI67* mapped on. Samples were grouped by visually inferred evolution patterns and sample names were color coded by their cytogenetic abnormalities. Only samples with more than 100 cPCs and 2 subclusters were shown. **(B)** Down-sampling pseudotime trajectory analysis of SMM02s and NDMM02s.

Figure S5


A
Linear
MGUS01s
SMM02s
SMM03s
SMM12
NDMM01s
SMM29
MGUS02s
SMM04s
SMM05s
SMM08
SMM11
SMM15
SMM18
SMM20
SMM22
SMM26
Branching
MGUS05s
SMM01s
SMM28
NDMM06
NDMM05s
SMM07s
SMM21
SMM25
NDMM02s
NDMM04s
MGUS06s
NDMM03s
non-HY
HY
Other
B cells (Pro/Pre/immature/mature)
nPCs
cPCs
MKI67 expression
2
0
B
SMM02s
Downsampling (#cells = 1,000)
full dataset (#cells = 2,129)
NDMM02s
Downsampling (#cells = 1,000)
full dataset (#cells = 3,517)

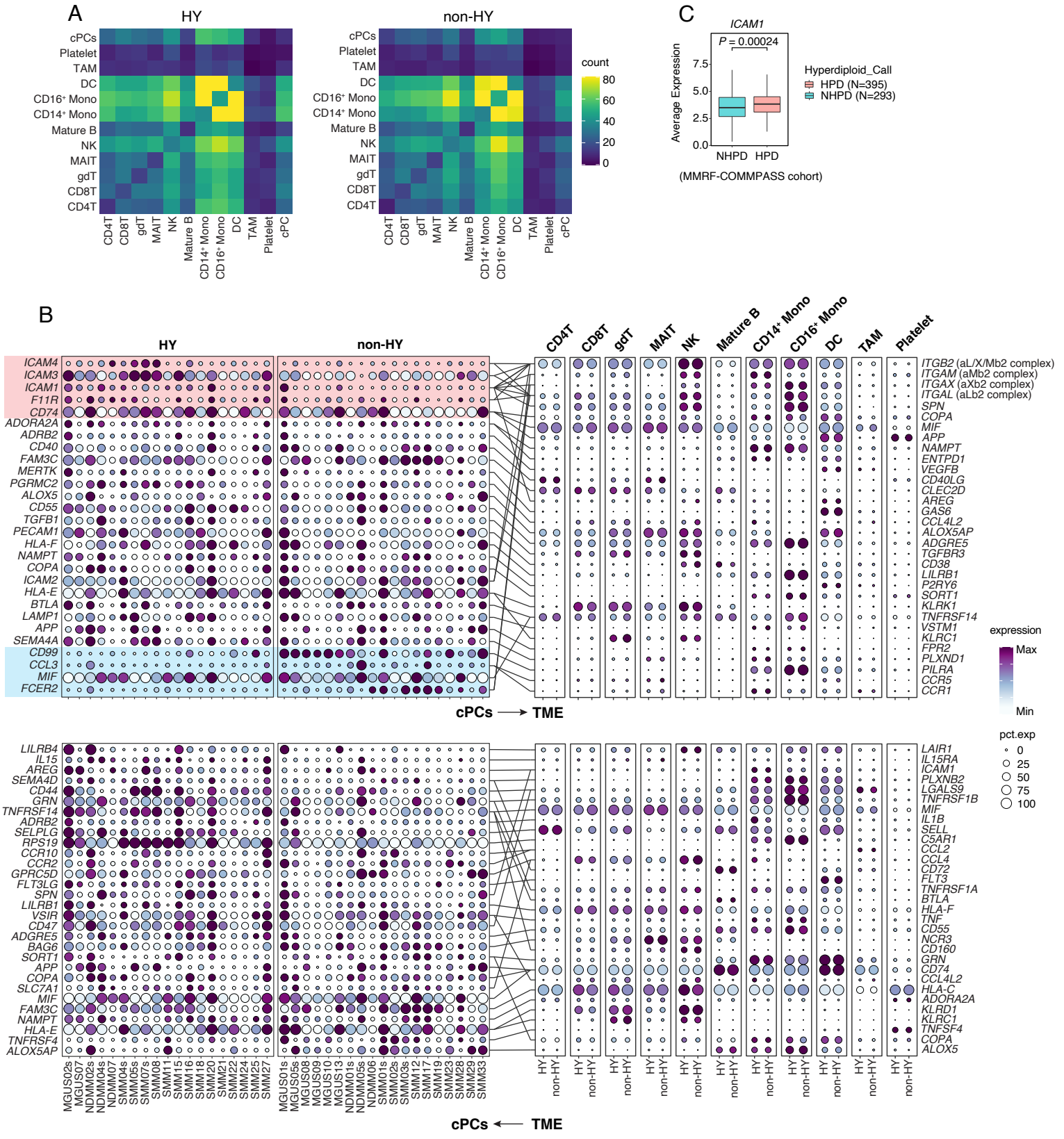
Supplementary Figure S6. Landscape of tumor microenvironment, related to Figure 6. (A) UMAP visualization of unsupervised clustering analysis of all TME cells (n = 120,677). Cells were color coded for their corresponding patient origins, diagnosis, and cytogenetic abnormalities. **(B)** Heatmap showing the scaled expression level of top 10 DEGs in the major cell types in TME. **(C)** Boxplots showing the comparisons of major cell type proportions among TME cells between nBM (N=3), MGUS (N=13), low/intermediate risk SMM (N=17) and high risk SMM (N=8). Dots are color coded by their cytogenetics abnormalities. Comparisons with P<0.1 are labeled. **(D)** Boxplots showing the comparisons of proportion of major cell types in TME between HY (N=14) and different translocation groups. Dots were color coded by their cytogenetic abnormalities and shape coded by diagnosis. **(E)** Boxplots showing the comparisons of proportion of $CD8^+$ T cells, $CD14^+$ monocytes, $CD16^+$ monocytes and DCs in TME between non-HY SMM (N=10) and HY SMM (N=11) samples (upper panel) and between non-HY MGUS (N=5) and HY MGUS (N=2) samples (lower panel). **(F)** Boxplots showing the comparisons of proportion of TAMs, Endothelial and Fibroblasts in TME between non-HY SMM (N=10) and HY SMM (N=11) samples. **(G)** Bubble plot showing the correlation of TME cellular compositions with ITH score, sample-averaged CNA score, proportion of cells at S and G2M phase, and sample-averaged CytoTRACE score of cPCs in non-HY and HY samples separately. **(H)** Boxplots showing the pairwise comparisons of proportion of CD8 T cell sub-clusters in CD8 T cells between non-HY (N=15) and HY (N=13) samples. Only sub-clusters reaching statistical significance were shown. **(I)** Volcano plots of differentially expressed genes in $CD14^+$ monocytes between non-HY and HY samples. FDR, two-sided Wilcoxon rank sum test with Bonferonni correction. Dashed line, log2FC > 0.3 or < -0.3. Labels, biologically important genes. **(J)** Uniform manifold approximation and

projection (UMAP) visualization of unsupervised sub-clustering analysis of $CD14^+$ Monocytes (n = 12,571).


A
By sample
By diagnosis
By Cytogenetic abnormalities
nBM MGUS SMM NDMM RRMM
Normal t(4;14) t(11;14) t(14;16)
t(14;20) HY Other NA
B
Mature B
CD8T
gdT
MAIT
CD4T
NK
CD14+ Mono
CD16+ Mono
TAM
cDC2
pDC
Platelet
Immature/Pro
Endothelial
Fibroblast
CD79A
MS4A1
TCL1A
CD74
IGHM
CD8B
CD8A
GZMH
CCL5
GZMK
TRDV2
TRGV9
DUSP2
TRDC
KLRB1
TRAV1-2
IL7R
NCR3
LTB
TCF7
MAL
LDHB
GNLY
PTGDS
SPON2
PRF1
NKG7
S100A8
S100A9
LYZ
FCN1
VCAN
LST1
SERPINA1
IFI30
MS4A7
AIF1
HMOX1
FTL
APOC1
C1QB
APOE
CST3
HLA-DPB1
HLA-DPA1
HLA-DQA1
HLA-DRB1
IRF7
LILRA4
PLD4
ITM2C
JCHAIN
PPBP
TUBB1
PF4
CAVIN2
GP1BB
PRTN3
MPO
ELANE
AZU1
IGLL1
FABP4
CCL14
RBP7
IFI27
FAM167B
CXCL12
COL1A2
COL1A1
SPARC
scaled avg.exp
-1 0 2
C
% (of TME cells)
Mature B
NK
CD14+ Mono
CD16+ Mono
TAM
CD8T
Endothelial
Fibroblast
P = 0.076
P = 0.078
P = 0.05
normal
MGUS
Low/Int Risk SMM
High Risk SMM
E
% of TME cells in SMM
% of TME cells in MGUS
CD8T
CD14+ Mono
CD16+ Mono
DC
P = 0.2
P = 0.13
P = 0.013
P = 0.061
P = 0.095
P = 1
P = 0.86
P = 1
non-HY
HY
D
Mature B
CD4T
MAIT
gdT
NK
TAM
Immature/Pro
Platelet
Endothelial
Fibroblast
% (of TME cells)
P = 0.58
P = 0.044
P = 0.1
P = 0.32
P = 0.79
P = 0.2
P = 0.15
P = 0.82
P = 0.55
P = 0.088
P = 0.095
P = 0.63
P = 0.55
P = 0.1
P = 0.067
P = 0.86
P = 0.58
P = 1
P = 0.93
P = 0.63
P = 0.38
P = 0.55
P = 0.82
P = 0.43
P = 0.45
P = 0.44
P = 0.82
P = 0.6
P = 0.37
P = 0.26
P = 0.18
P = 0.42
P = 0.15
P = 0.51
P = 0.81
P = 0.13
P = 0.26
P = 0.3
P = 0.58
P = 0.19
HY
t(11;14)
t(14;16)
t(4;14)
t(14;20)
Cytogenetics
t(4;14) (N=1)
t(11;14) (N=12)
t(14;16) (N=2)
t(14;20) (N=1)
HY (N=14)
Diagnosis
MGUS (N=7)
SMM (N=21)
NDMM (N=2)
F
TAM
Endothelial
Fibroblast
% of TME cells in SMM
P = 0.029
P = 0.021
P = 0.034
non-HY HY
G
(S+G2M)%
CNA score
ITH
mean CytoTRACE
TAM
Platelet
pDC
NK
Mature B
MAIT
Immature/Pro
gdT
Fibroblast
Endothelial
cDC2
CD8T
CD4T
CD16+ Mono
CD14+ Mono
non-HY
HY
p value
>=0.05
<0.05
<0.01
Spearman correlation
0.8
0.4
0.0
-0.4
H
CD8_C0
CD8_C3
CD8_C9
% (of CD8T cells)
P = 0.044
P = 0.01
P = 0.027
non-HY HY
I
HY vs non-HY in CD14+ Mono
High in non-HY
High in HY
-log10 adjusted p-value
log2 fold change
LYZ
IL1B
AC007952.4
CCL3
CXCL8
NFKBIZ
NFKBIA
TNFAIP3
IFITM3
ZFP36
IER2
SOD2
ISG15
IFI6
PLAC8
CD52
TMEM176A
TMEM176B
HLA-DQB1
-log10fdr
0
10
50
100
High in HY
High in non-HY
J
C4
C5
C0
C2
C1
C3
CD14+ Mono
(n=12,571)

Supplementary Figure S7. Cell-cell interactions between cPCs and TME cells, related to Figure 7. (A) Heatmap showing the number of interactions between each major cell types (#cells > 1000) inferred by CellphoneDB in non-HY and HY samples separately. **(B)** Bubble plot showing the expression (color key) and frequency (size key) of each ligand and receptor in corresponding cell types in non-HY and HY side by side. Only pairs in Fig 6I are shown. **(C)** Boxplots showing the comparisons of expression of *ICAM1* between NHPD (non-HY) and HPD (HY) in MMRF-COMMPASS cohort.

Figure S7


A
HY
non-HY
cPCs
Platelet
TAM
DC
CD16⁺ Mono
CD14⁺ Mono
Mature B
NK
MAIT
gdT
CD8T
CD4T
count
80
60
40
20
0
C
ICAM1
P = 0.00024
Average Expression
7.5
5.0
2.5
0.0
NHPD
HPD
Hyperdiploid_Call
HPD (N=395)
NHPD (N=293)
(MMRF-COMMPASS cohort)
B
HY
non-HY
ICAM4
ICAM3
ICAM1
F11R
CD74
ADORA2A
ADRB2
CD40
FAM3C
MERTK
PGRMC2
ALOX5
CD55
TGFB1
PECAM1
HLA-F
NAMPT
COPA
ICAM2
HLA-E
BTLA
LAMP1
APP
SEMA4A
CD99
CCL3
MIF
FCER2
ITGB2 (aL/X/Mb2 complex)
ITGAM (aMb2 complex)
ITGAX (aXb2 complex)
ITGAL (aLb2 complex)
SPN
COPA
MIF
APP
NAMPT
ENTPD1
VEGFB
CD40LG
CLEC2D
AREG
GAS6
CCL4L2
ALOX5AP
ADGRE5
TGFBR3
CD38
LILRB1
P2RY6
SORT1
KLRK1
TNFRSF14
VSTM1
KLRC1
FPR2
PLXND1
PILRA
CCR5
CCR1
expression
Max
Min
cPCs → TME
LILRB4
IL15
AREG
SEMA4D
CD44
GRN
TNFRSF14
ADRB2
SELPLG
RPS19
CCR10
CCR2
GPRC5D
FLT3LG
SPN
LILRB1
VSIR
CD47
ADGRE5
BAG6
SORT1
APP
COPA
SLC7A1
MIF
FAM3C
NAMPT
HLA-E
TNFRSF4
ALOX5AP
LAIR1
IL15RA
ICAM1
PLXNB2
LGALS9
TNFRSF1B
MIF
IL1B
SELL
C5AR1
CCL2
CCL4
CD72
FLT3
TNFRSF1A
BTLA
HLA-F
TNF
CD55
NCR3
CD160
GRN
CD74
CCL4L2
HLA-C
ADORA2A
KLRD1
KLRC1
TNFSF4
COPA
ALOX5
pct.exp
0
25
50
75
100
MGUS02s
MGUS07
NDMM02s
NDMM04s
NDMM07
SMM04s
SMM05s
SMM07s
SMM08
SMM11
SMM15
SMM16
SMM18
SMM20
SMM21
SMM22
SMM24
SMM25
SMM27
MGUS01s
MGUS05s
MGUS08
MGUS09
MGUS10
MGUS13
NDMM01s
NDMM05s
NDMM06
SMM01s
SMM02s
SMM03s
SMM12
SMM17
SMM19
SMM23
SMM28
SMM29
SMM33
cPCs ← TME