Title: Utilizing Genomics to Identify Novel Immunotherapeutic Targets in Multiple Myeloma High-Risk Subgroups

Authors: Enze Liu[1], Oumaima Jaouadi[1], Riya Sharma[1], Nathan Becker[1], Travis S. Johnson[2], Parvathi Sudha[1], Vivek S. Chopra[3], Faiza Zafar[3], Habib Hamidi[3], Charlotte Pawlyn[4], Attaya Suvannasankha[1], Rafat Abonour[1], Brian A. Walker[1]

Affiliations:

1-Melvin and Bren Simon Comprehensive Cancer Center, Division of Hematology and Oncology, School of Medicine, Indiana University, Indianapolis, IN, USA

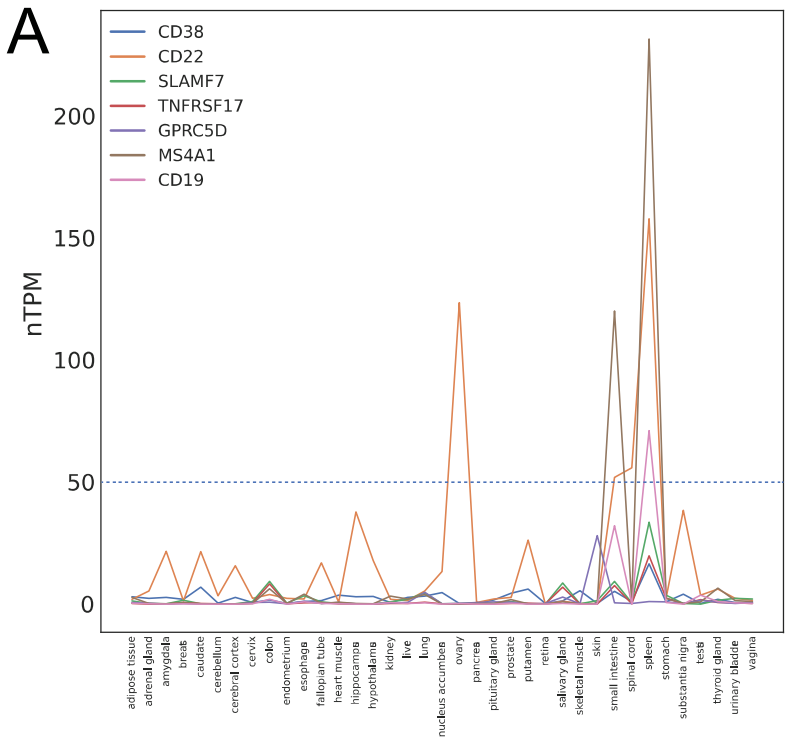
2-Department of Biostatistics & Health Data Sciences, School of Medicine, Indiana University, Indianapolis, IN, USA

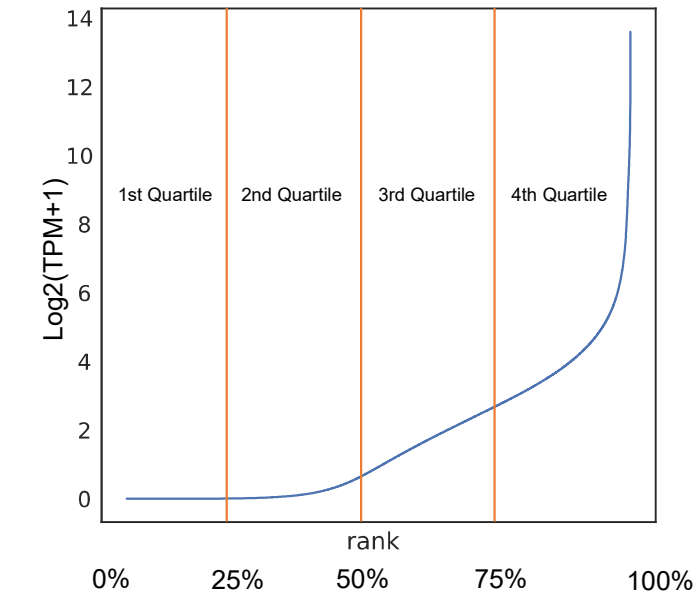
3-Genentech Inc., South San Francisco, CA, USA

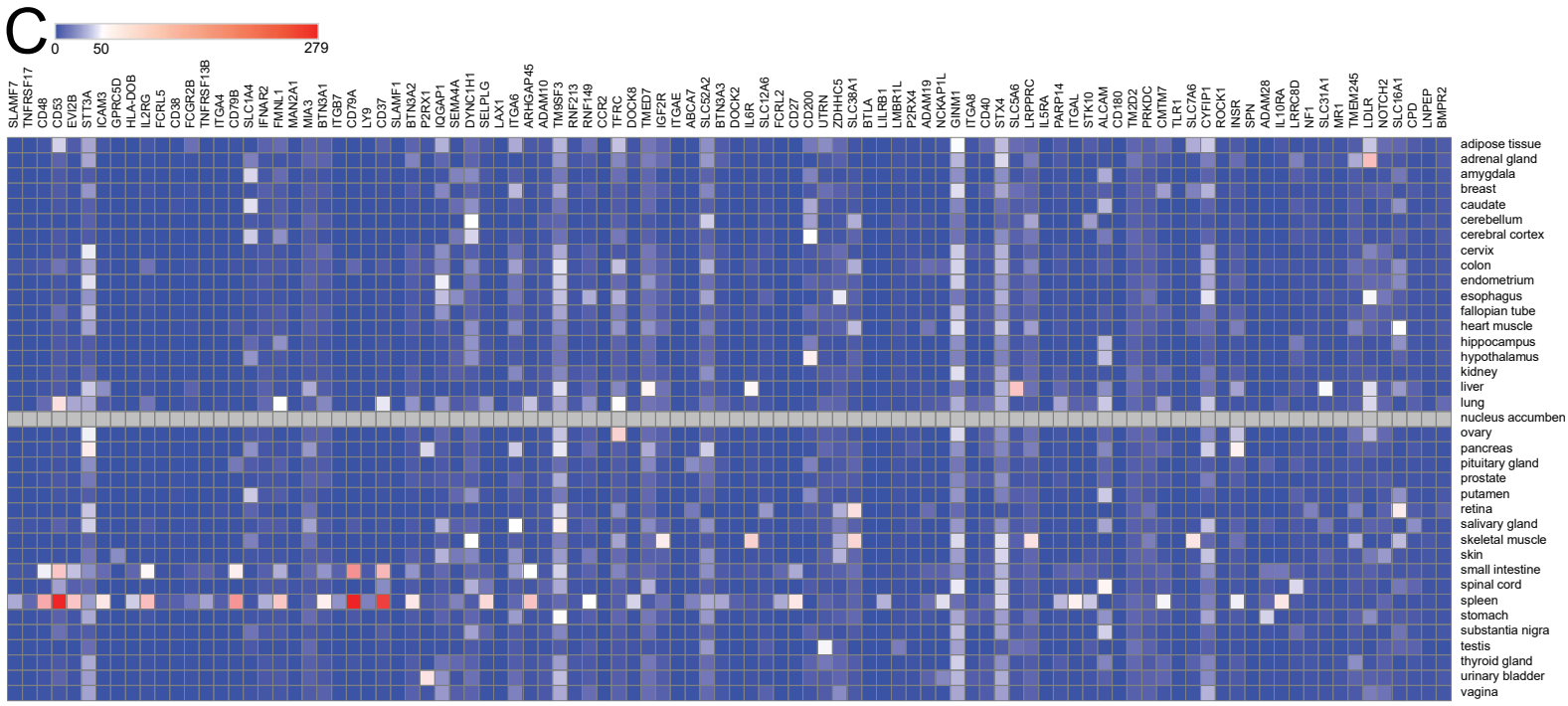
4-Insititute of Cancer Research, 15 Cotswold Road, London, UK.

Correspondence: Brian A. Walker, 1550 NW 10th Ave, Suite PAP-511, University of Miami, Miami, FL, USA, 33136. brianwalker@miami.edu

Supplementary Figures


A
CD38
CD22
SLAMF7
TNFRSF17
GPRC5D
MS4A1
CD19
nTPM
200
150
100
50
0
adipose tissue
adrenal gland
amygdala
breast
caudate
cerebellum
cerebral cortex
cervix
colon
endometrium
esophagus
fallopian tube
heart muscle
hippocampus
hypothalamus
kidney
liver
lung
nucleus accumbens
ovary
pancreas
pituitary gland
prostate
putamen
retina
salivary gland
skeletal muscle
skin
small intestine
spinal cord
spleen
stomach
substantia nigra
testis
thyroid gland
urinary bladder
vagina


B
Log2(TPM+1)
14
12
10
8
6
4
2
0
1st Quartile
2nd Quartile
3rd Quartile
4th Quartile
rank
0%
25%
50%
75%
100%


C
0
50
279
SLAMF7
TNFRSF17
CD48
CD53
EVI2B
STT3A
ICAM3
GPRC5D
HLA-DOB
IL2RG
FCRL5
CD38
FCGR2B
TNFRSF13B
ITGA4
CD79B
SLC1A4
IFNAR2
FMNL1
MAN2A1
MIA3
BTN3A1
ITGB7
CD79A
LY9
CD37
SLAMF1
BTN3A2
P2RX1
IQGAP1
SEMA4A
DYNC1H1
SELPLG
LAX1
ITGA6
ARHGAP45
ADAM10
TM9SF3
RNF213
RNF149
CCR2
TFRC
DOCK8
TMED7
IGF2R
ITGAE
ABCA7
SLC52A2
BTN3A3
DOCK2
IL6R
SLC12A6
FCRL2
CD27
CD200
UTRN
ZDHHC5
SLC38A1
BTLA
LILRB1
LMBR1L
P2RX4
ADAM19
NCKAP1L
GINM1
ITGA8
CD40
STX4
SLC5A6
LRPPRC
IL5RA
PARP14
ITGAL
STK10
ALCAM
CD180
TM2D2
PRKDC
CMTM7
TLR1
SLC7A6
CYFIP1
ROCK1
INSR
SPN
ADAM28
IL10RA
LRRC8D
NF1
SLC31A1
MR1
TMEM245
LDLR
NOTCH2
SLC16A1
CPD
LNPEP
BMPR2
adipose tissue
adrenal gland
amygdala
breast
caudate
cerebellum
cerebral cortex
cervix
colon
endometrium
esophagus
fallopian tube
heart muscle
hippocampus
hypothalamus
kidney
liver
lung
nucleus accumben
ovary
pancreas
pituitary gland
prostate
putamen
retina
salivary gland
skeletal muscle
skin
small intestine
spinal cord
spleen
stomach
substantia nigra
testis
thyroid gland
urinary bladder
vagina

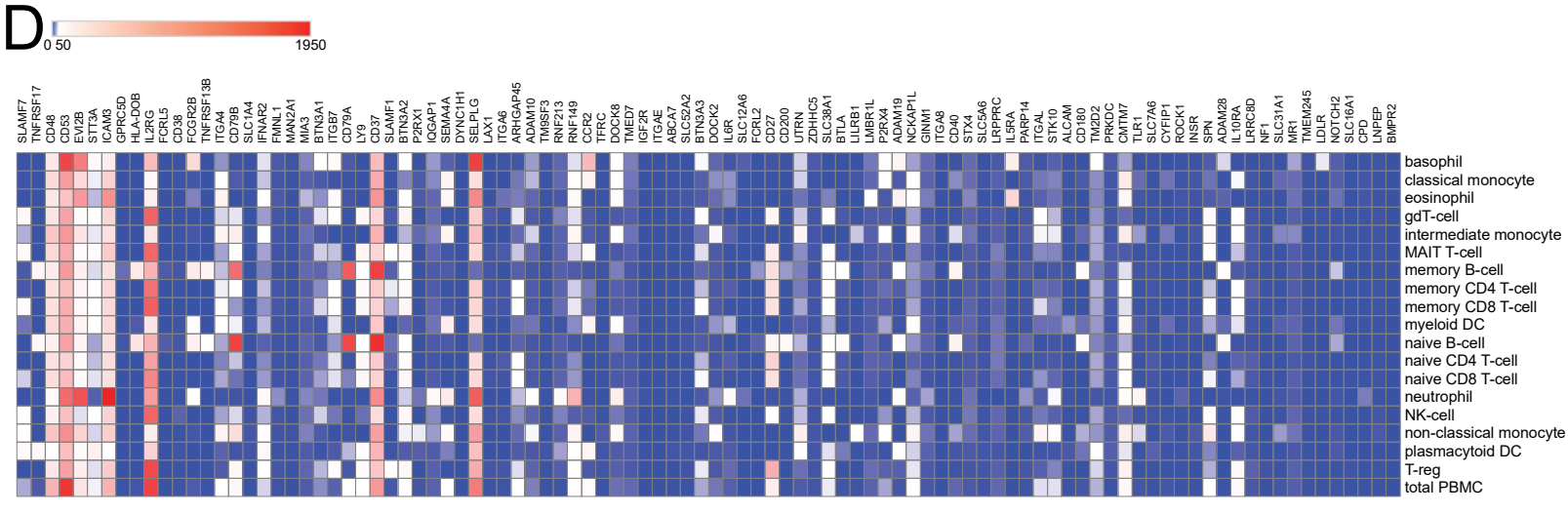
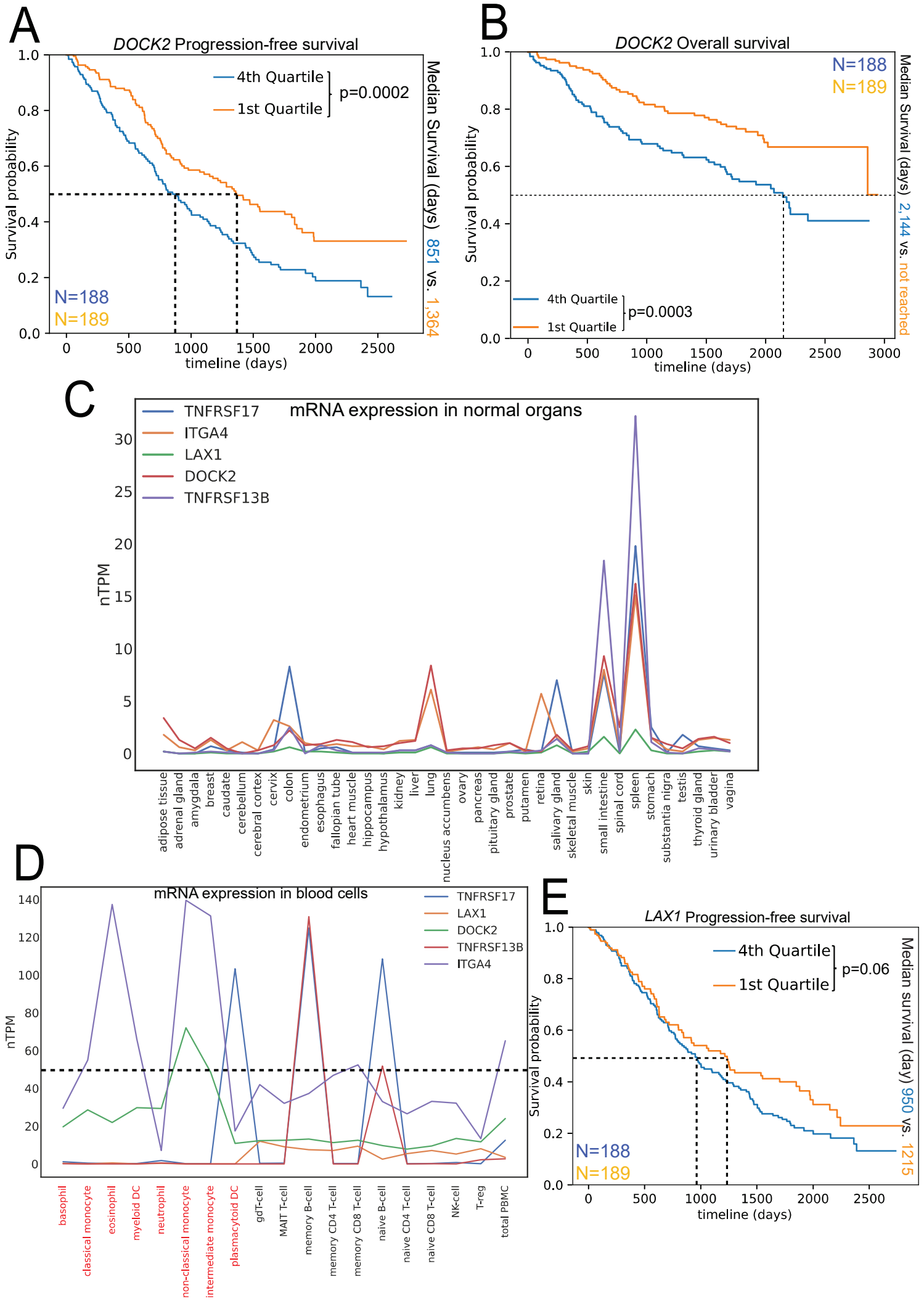
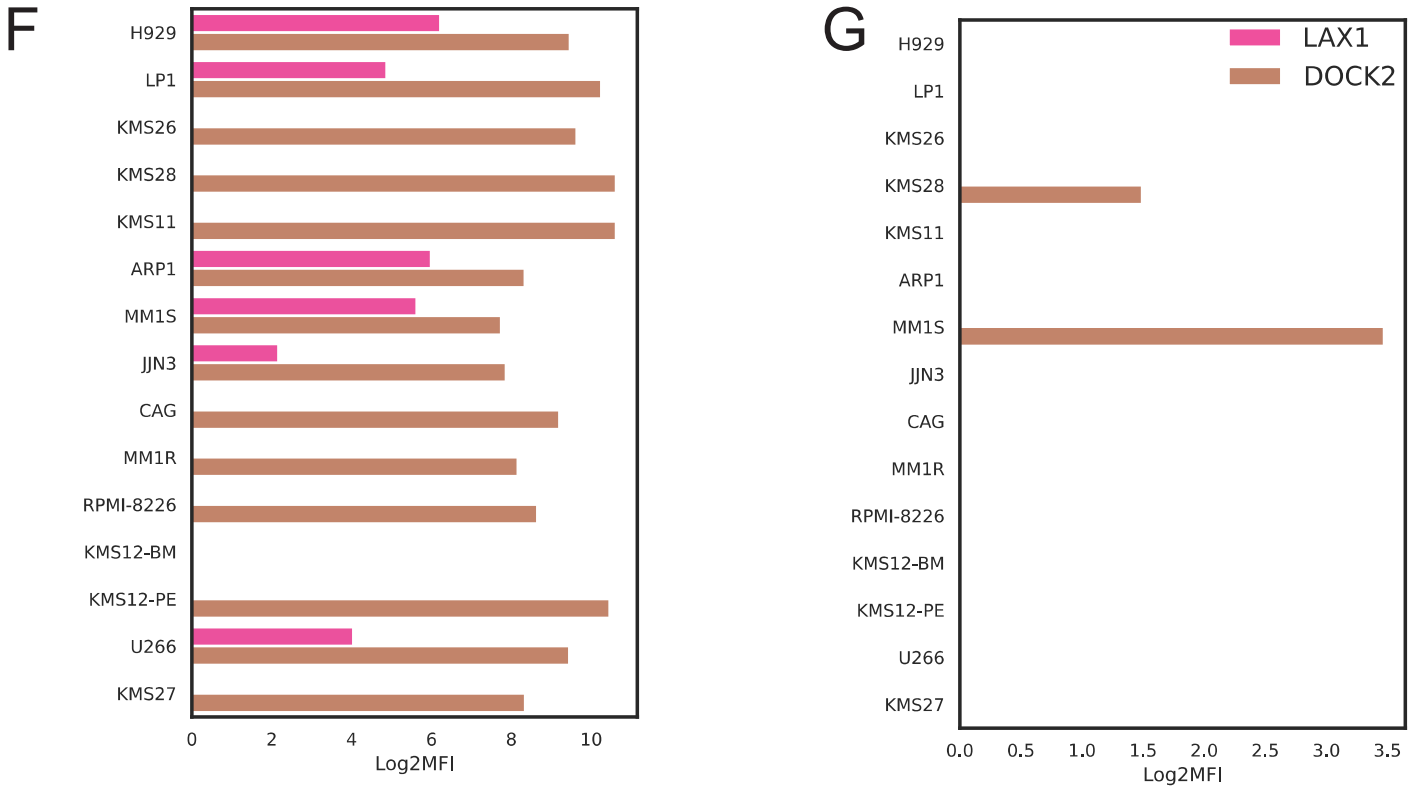
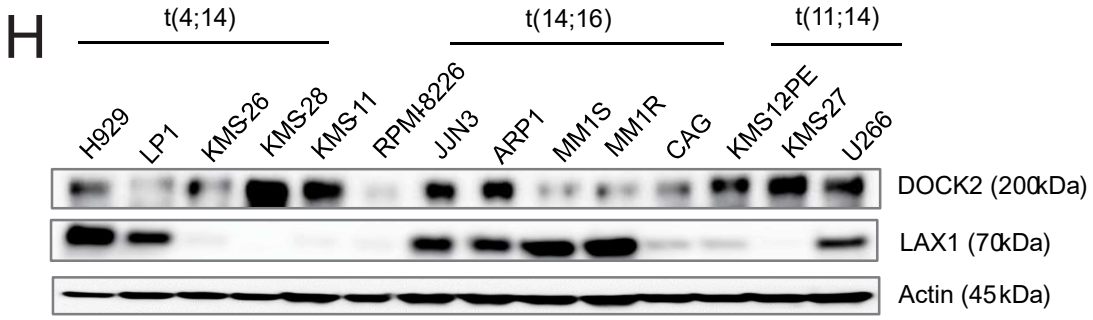

D
0 50
1950
basophil
classical monocyte
eosinophil
gdT-cell
intermediate monocyte
MAIT T-cell
memory B-cell
memory CD4 T-cell
memory CD8 T-cell
myeloid DC
naive B-cell
naive CD4 T-cell
naive CD8 T-cell
neutrophil
NK-cell
non-classical monocyte
plasmacytoid DC
T-reg
total PBMC

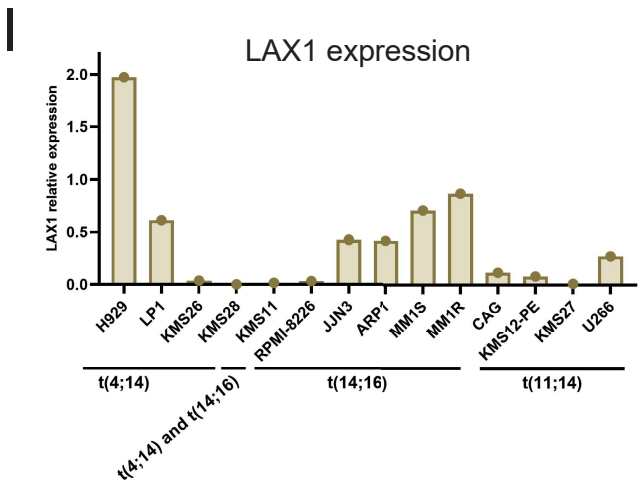
Fig. S1. Characteristics of candidate targets identified from MM populations.

A. RNA expression level of current CAR-T cell targets of blood cancers in healthy organs documented in the human protein atlas (THPA). **B**. Distribution of mRNA expression of 19,892 protein coding genes in MMRF and IU cohort. **C**. Expression level of 98 candidates in healthy organs documented in the human protein atlas (THPA). **D**. Expression level of 98 candidates in myeloid blood cells documented in 'The Human Protein Atlas' database.


A
DOCK2 Progression-free survival
4th Quartile
1st Quartile
p=0.0002
N=188
N=189
Survival probability
timeline (days)
Median Survival (days) 851 vs. 1,364
B
DOCK2 Overall survival
N=188
N=189
4th Quartile
1st Quartile
p=0.0003
Survival probability
timeline (days)
Median Survival (days) 2,144 vs. not reached
C
mRNA expression in normal organs
TNFRSF17
ITGA4
LAX1
DOCK2
TNFRSF13B
nTPM
adipose tissue
adrenal gland
amygdala
breast
caudate
cerebellum
cerebral cortex
cervix
colon
endometrium
esophagus
fallopian tube
heart muscle
hippocampus
hypothalamus
kidney
liver
lung
nucleus accumbens
ovary
pancreas
pituitary gland
prostate
putamen
retina
salivary gland
skeletal muscle
skin
small intestine
spinal cord
spleen
stomach
substantia nigra
testis
thyroid gland
urinary bladder
vagina
D
mRNA expression in blood cells
TNFRSF17
LAX1
DOCK2
TNFRSF13B
ITGA4
nTPM
basophil
classical monocyte
eosinophil
myeloid DC
neutrophil
non-classical monocyte
intermediate monocyte
plasmacytoid DC
gdT-cell
MAIT T-cell
memory B-cell
memory CD4 T-cell
memory CD8 T-cell
naive B-cell
naive CD4 T-cell
naive CD8 T-cell
NK-cell
T-reg
total PBMC
E
LAX1 Progression-free survival
4th Quartile
1st Quartile
p=0.06
N=188
N=189
Survival probability
timeline (days)
Median survival (day) 950 vs. 1215


F
H929
LP1
KMS26
KMS28
KMS11
ARP1
MM1S
JJN3
CAG
MM1R
RPMI-8226
KMS12-BM
KMS12-PE
U266
KMS27
0
2
4
6
8
10
Log2MFI
G
LAX1
DOCK2
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Log2MFI


H
t(4;14)
t(14;16)
t(11;14)
H929
LP1
KMS-26
KMS-28
KMS-11
RPMI-8226
JJN3
ARP1
MM1S
MM1R
CAG
KMS12PE
KMS27
U266
DOCK2 (200kDa)
LAX1 (70kDa)
Actin (45kDa)


I
LAX1 expression
LAX1 relative expression
2.0
1.5
1.0
0.5
0.0
H929
LP1
KMS26
KMS28
KMS11
RPMI-8226
JJN3
ARP1
MM1S
MM1R
CAG
KMS12-PE
KMS27
U266
t(4;14)
t(4;14) and t(14;16)
t(14;16)
t(11;14)

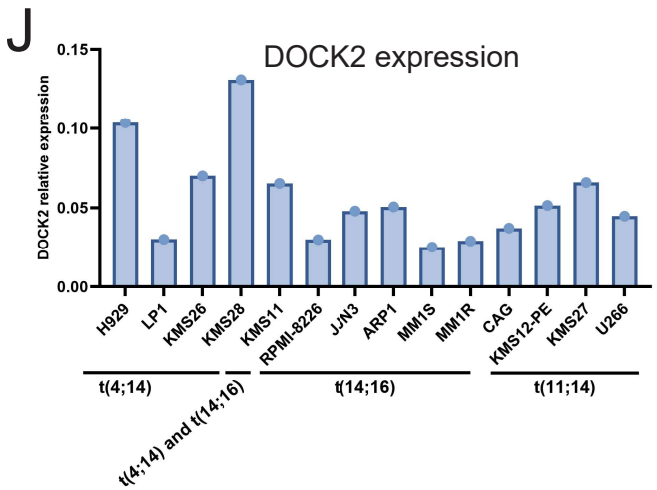

J
DOCK2 expression
DOCK2 relative expression
0.15
0.10
0.05
0.00
H929
LP1
KMS26
KMS28
KMS11
RPMI-8226
JJN3
ARP1
MM1S
MM1R
CAG
KMS12-PE
KMS27
U266
t(4;14)
t(4;14) and t(14;16)
t(14;16)
t(11;14)

Fig. S2. Toxicity and association with patient survival of identified population-based candidates.

A. Progression-free survival of patients with different *DOCK2* expression levels. **B**. Overall survival of patients with different *DOCK2* expression levels. **C**. Expression level of existing targets and selected candidates in healthy organs. **D**. Expression level of existing targets and selected candidates in blood cells. **E**. Progression-free survival of patients with different *LAX1* expression. Flow cytometry results of LAX1 and DOCK2 were identified in **F** and **G**. **H**. Expression level of LAX1 and DOCK2 detected by immunoblotting. Their expression level was quantified in barplots **I** and **J**. Statistical test in Kaplan-Meier curves: Log-rank test. Highlighted labels in **D**: myeloid blood cells.

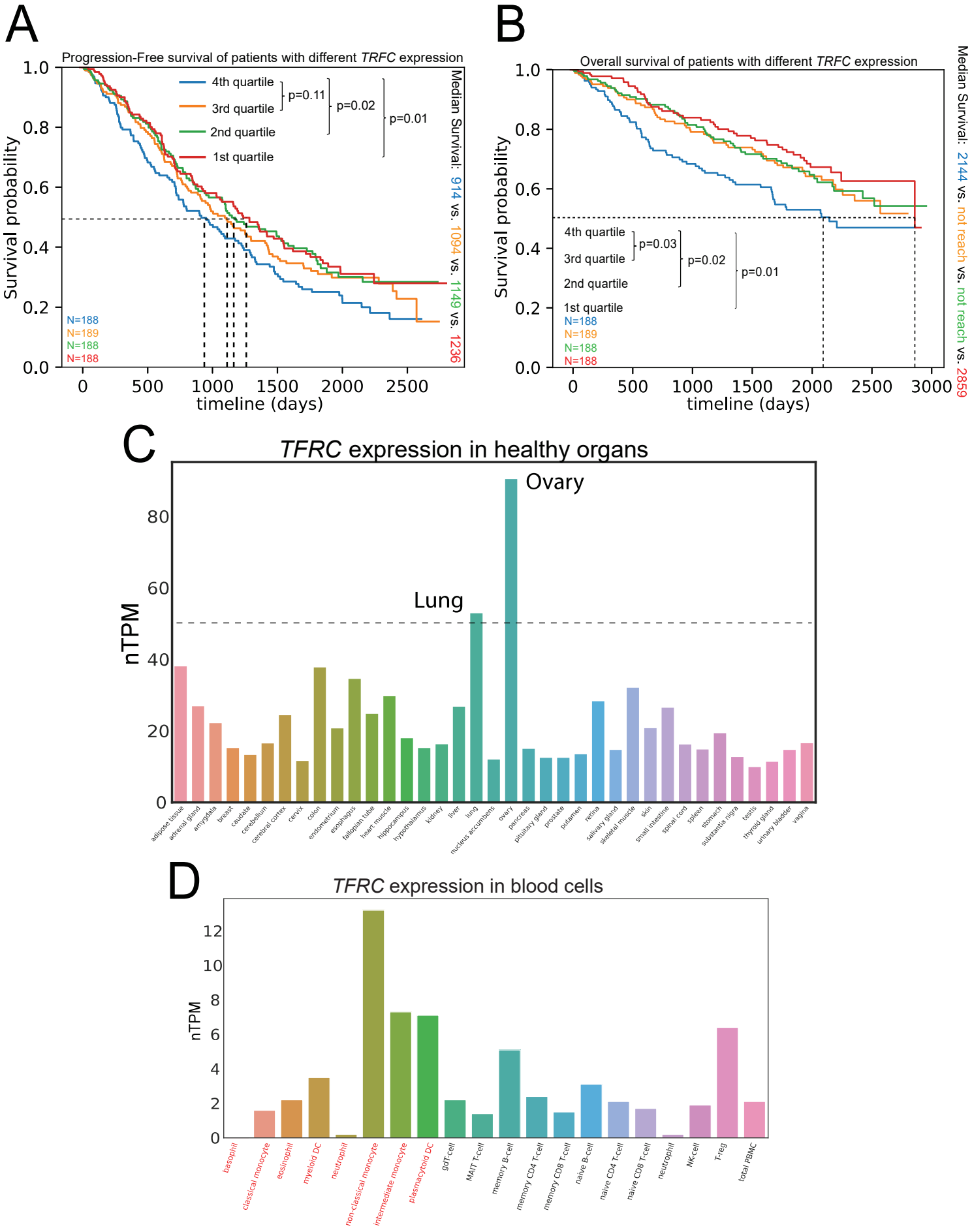

A
Progression-Free survival of patients with different TRFC expression
Survival probability
timeline (days)
4th quartile
3rd quartile
2nd quartile
1st quartile
p=0.11
p=0.02
p=0.01
N=188
N=189
N=188
N=188
Median Survival: 914 vs. 1094 vs. 1149 vs. 1236
B
Overall survival of patients with different TRFC expression
Survival probability
timeline (days)
4th quartile
3rd quartile
2nd quartile
1st quartile
p=0.03
p=0.02
p=0.01
N=188
N=189
N=188
N=188
Median Survival: 2144 vs. not reach vs. not reach vs. 2859
C
TFRC expression in healthy organs
nTPM
Ovary
Lung
adipose tissue
adrenal gland
amygdala
breast
caudate
cerebellum
cerebral cortex
cervix
colon
endometrium
esophagus
fallopian tube
heart muscle
hippocampus
hypothalamus
kidney
liver
lung
nucleus accumbens
ovary
pancreas
pituitary gland
prostate
putamen
retina
salivary gland
skeletal muscle
skin
small intestine
spinal cord
spleen
stomach
substantia nigra
testis
thyroid gland
urinary bladder
vagina
D
TFRC expression in blood cells
nTPM
basophil
classical monocyte
eosinophil
myeloid DC
neutrophil
non-classical monocyte
intermediate monocyte
plasmacytoid DC
gdT-cell
MAIT T-cell
memory B-cell
memory CD4 T-cell
memory CD8 T-cell
naive B-cell
naive CD4 T-cell
naive CD8 T-cell
neutrophil
NK-cell
T-reg
total PBMC

Fig. S3. Characteristics of a highly essential candidate *TRFC*

A. Progression-free survival of patients with different *TFRC* expression levels. **B**. Overall survival of patients with different *TFRC* expression levels. **C**. *TFRC* expression level in healthy organs. **D**. *TFRC* expression level in blood cells. Statistical test in Kaplan-Meier curves: Logrank test. Highlighted labels in **D**: myeloid blood cells.

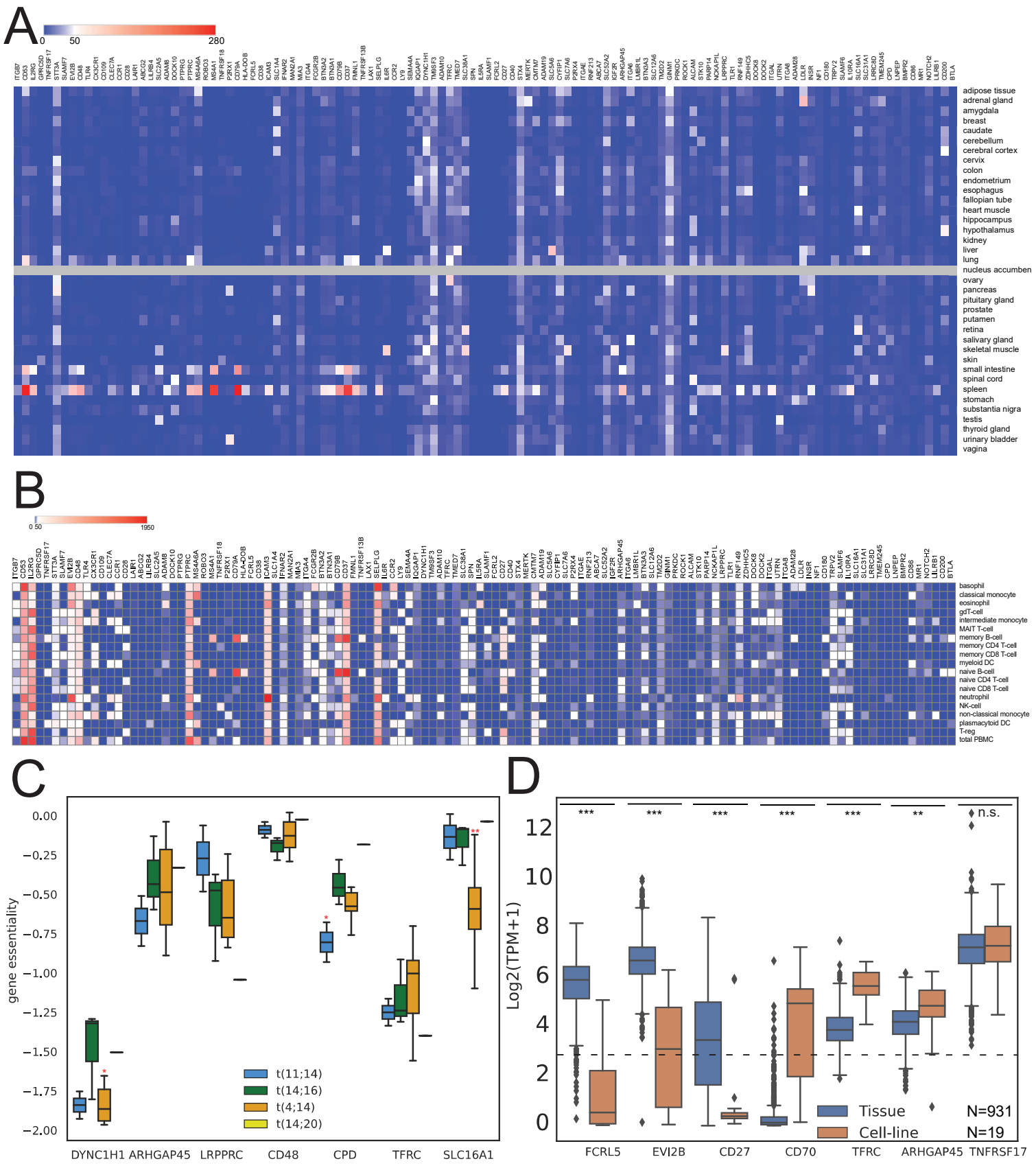
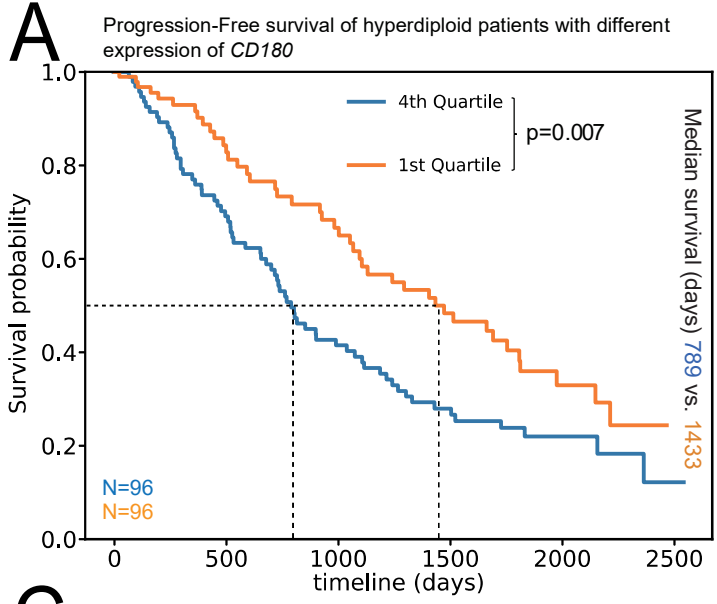
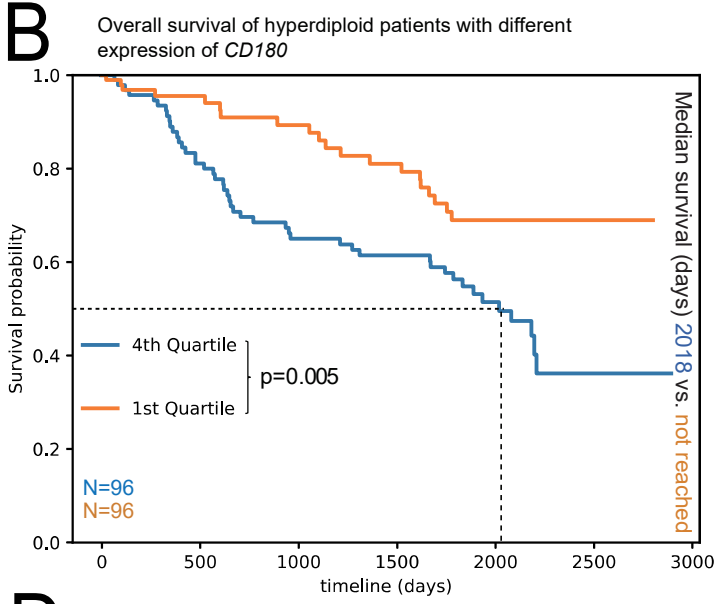
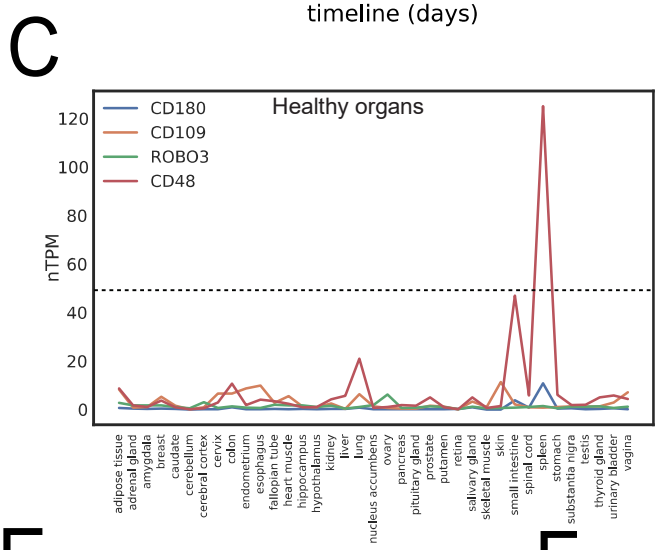
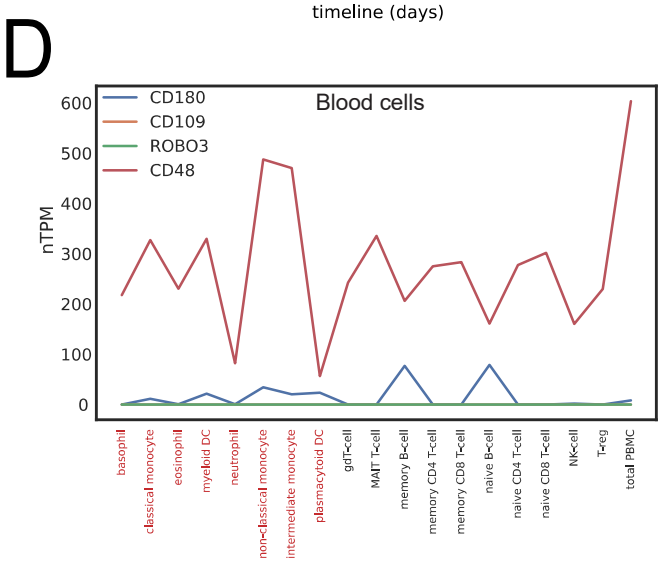

A
0
50
280
adipose tissue
adrenal gland
amygdala
breast
caudate
cerebellum
cerebral cortex
cervix
colon
endometrium
esophagus
fallopian tube
heart muscle
hippocampus
hypothalamus
kidney
liver
lung
nucleus accumben
ovary
pancreas
pituitary gland
prostate
putamen
retina
salivary gland
skeletal muscle
skin
small intestine
spinal cord
spleen
stomach
substantia nigra
testis
thyroid gland
urinary bladder
vagina
B
0 50
1950
basophil
classical monocyte
eosinophil
gdT-cell
intermediate monocyte
MAIT T-cell
memory B-cell
memory CD4 T-cell
memory CD8 T-cell
myeloid DC
naive B-cell
naive CD4 T-cell
naive CD8 T-cell
neutrophil
NK-cell
non-classical monocyte
plasmacytoid DC
T-reg
total PBMC
C
gene essentiality
0.00
−0.25
−0.50
−0.75
−1.00
−1.25
−1.50
−1.75
−2.00
t(11;14)
t(14;16)
t(4;14)
t(14;20)
DYNC1H1
ARHGAP45
LRPPRC
CD48
CPD
TFRC
SLC16A1
D
Log2(TPM+1)
12
10
8
6
4
2
0
***
***
***
***
***
**
n.s.
Tissue
Cell-line
N=931
N=19
FCRL5
EVI2B
CD27
CD70
TFRC
ARHGAP45
TNFRSF17

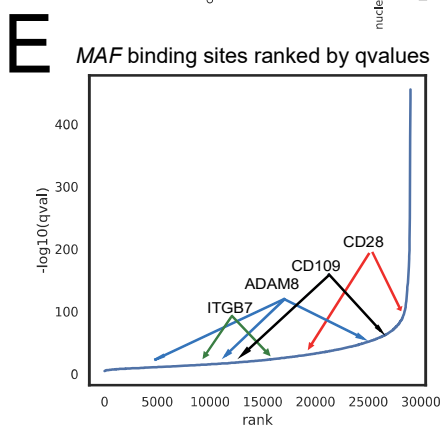
Fig. S4. Characteristics of candidates identified from primary subtypes. A. Expression level of 120 candidates in healthy organs documented in the human protein atlas (THPA). **B**. Expression level of 120 candidates in blood cells documented in 'The Human Protein Atlas' database. **C**. Gene essentiality differences of selected candidates in 18 MM cell-lines from DepMap. **D**. Expression heterogeneity of candidates between tissue and cell-lines. Statistical tests in **C** and **D**: Mann-Whitney U test. Highlighted stars in **C**: significance level of candidates with higher essentiality in their subgroups than average in MM cell-lines. Highlighted stars in **D**: significance level of candidates with differential expression. Significance level: * $p<0.05$; ** $p<0.01$, *** $p<0.001$.

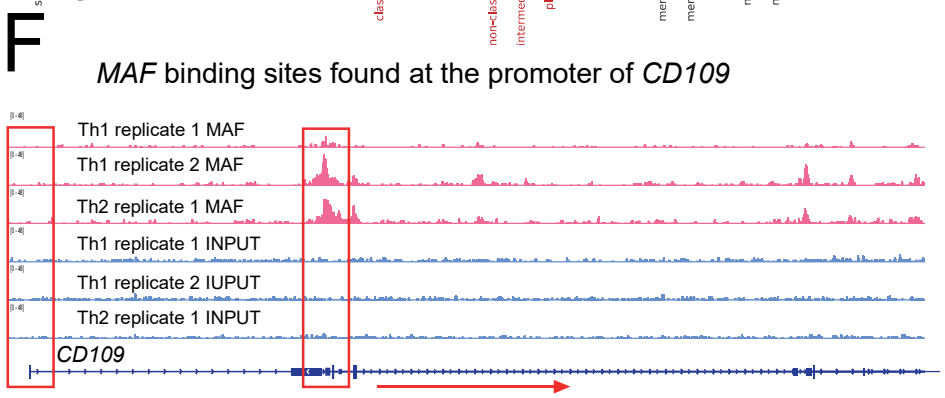

A
Progression-Free survival of hyperdiploid patients with different expression of CD180
Survival probability
timeline (days)
4th Quartile
1st Quartile
p=0.007
N=96
N=96
Median survival (days) 789 vs. 1433

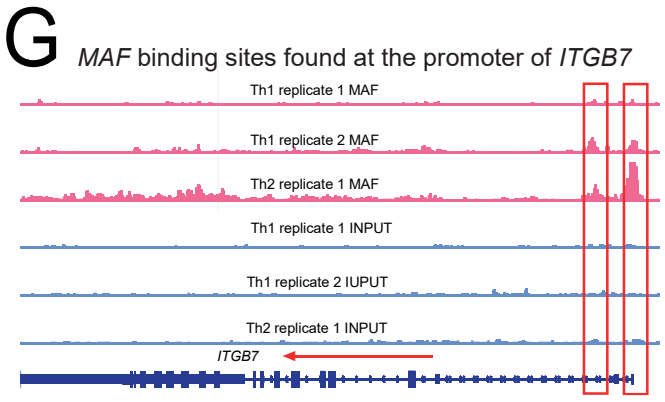

B
Overall survival of hyperdiploid patients with different expression of CD180
Survival probability
timeline (days)
4th Quartile
1st Quartile
p=0.005
N=96
N=96
Median survival (days) 2018 vs. not reached


C
Healthy organs
CD180
CD109
ROBO3
CD48
nTPM


D
Blood cells
CD180
CD109
ROBO3
CD48
nTPM


E
MAF binding sites ranked by qvalues
-log10(qval)
rank
CD28
CD109
ADAM8
ITGB7


F
MAF binding sites found at the promoter of CD109
Th1 replicate 1 MAF
Th1 replicate 2 MAF
Th2 replicate 1 MAF
Th1 replicate 1 INPUT
Th1 replicate 2 IUPUT
Th2 replicate 1 INPUT
CD109


G
MAF binding sites found at the promoter of ITGB7
Th1 replicate 1 MAF
Th1 replicate 2 MAF
Th2 replicate 1 MAF
Th1 replicate 1 INPUT
Th1 replicate 2 IUPUT
Th2 replicate 1 INPUT
ITGB7

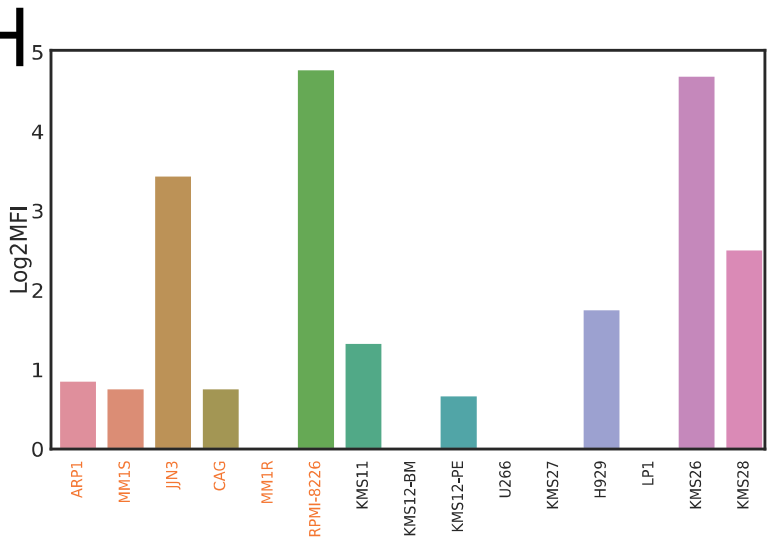
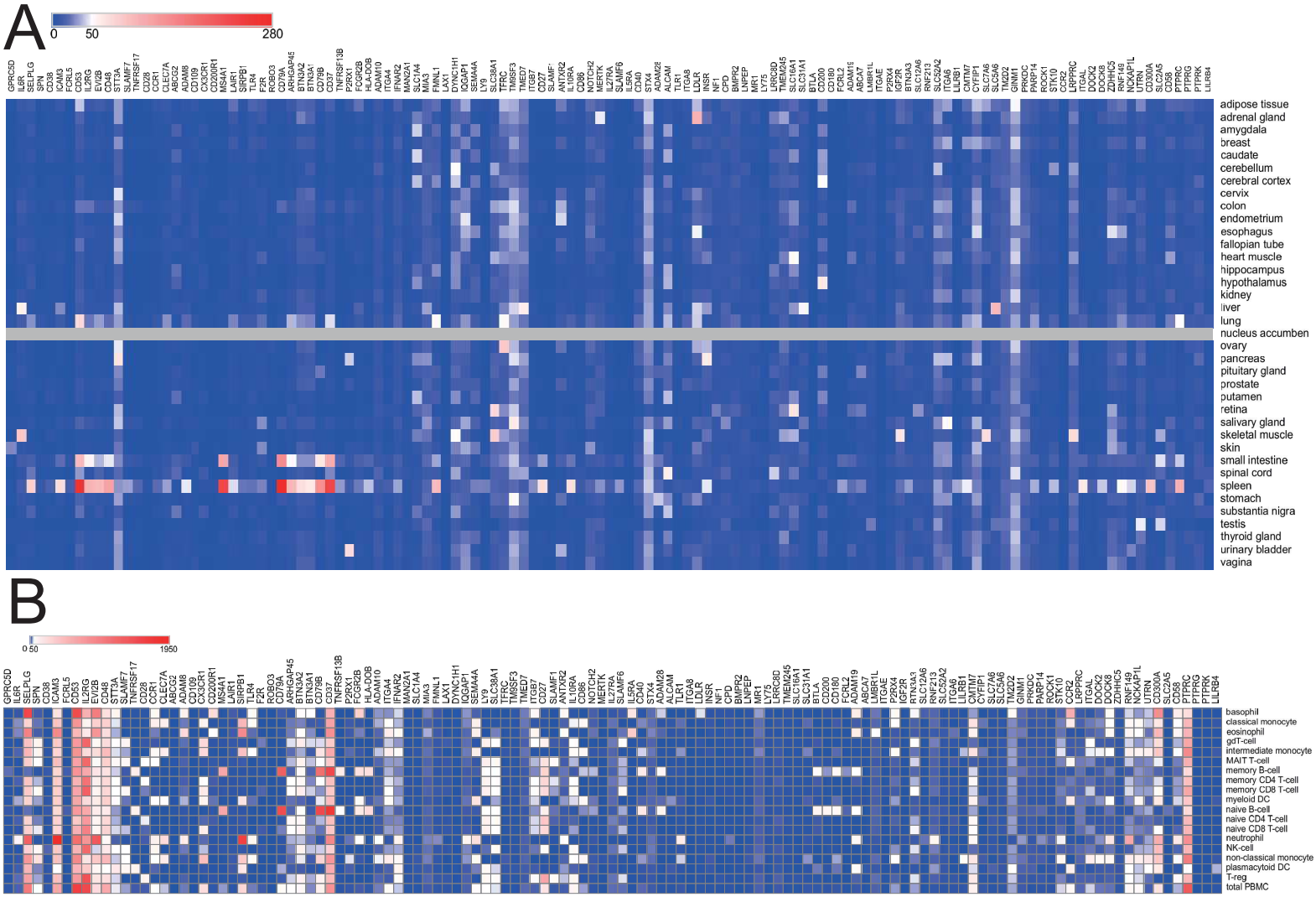
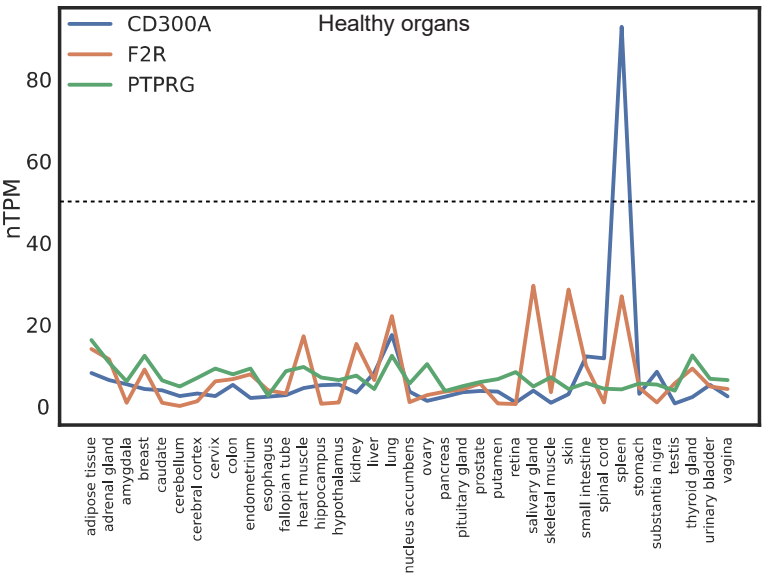

H
Log2MFI
ARP1
MM1S
JJN3
CAG
MM1R
RPMI-8226
KMS11
KMS12-BM
KMS12-PE
U266
KMS27
H929
LP1
KMS26
KMS28

Fig. S5. Toxicity and association with patient survival of identified subtype-specific candidates.

A. Progression-free survival of hyper-diploid patients with different *CD180* expression levels. **B**. Overall survival of hyper-diploid patients with different *CD180* expression levels. **C**. Selected candidates identified in subtypes and their expression level in healthy organs. **D**. Selected candidates identified in subtypes and their expression level in blood cells. **E**. Four candidates found with *MAF* binding sites from ChIP-seq result. **F**. *MAF* binding sites found at the promoter region of CD109. **G**. *MAF* binding sites found at the promoter region of *ITGB7*. **H**. Protein expression level (Log2MFI) of *CD109* detected by flow cytometry among 15 cell-lines. Statistical test in Kaplan-Meier curves: Logrank test. Highlighted labels in **D**: myeloid blood cells. Highlighted boxes in **F** and **G**: promoter regions. Red arrows: transcription direction. Highlighted cell-lines in **H**: t(14;16) cell-lines. KMS11 is a t(14;16) and t(4;14) cell-line.


A
0
50
280
adipose tissue
adrenal gland
amygdala
breast
caudate
cerebellum
cerebral cortex
cervix
colon
endometrium
esophagus
fallopian tube
heart muscle
hippocampus
hypothalamus
kidney
liver
lung
nucleus accumbens
ovary
pancreas
pituitary gland
prostate
putamen
retina
salivary gland
skeletal muscle
skin
small intestine
spinal cord
spleen
stomach
substantia nigra
testis
thyroid gland
urinary bladder
vagina
B
0 50
1950
basophil
classical monocyte
eosinophil
gdT-cell
intermediate monocyte
MAIT T-cell
memory B-cell
memory CD4 T-cell
memory CD8 T-cell
myeloid DC
naive B-cell
naive CD4 T-cell
naive CD8 T-cell
neutrophil
NK-cell
non-classical monocyte
plasmacytoid DC
T-reg
total PBMC


C
CD300A
F2R
PTPRG
Healthy organs
nTPM
0
20
40
60
80
adipose tissue
adrenal gland
amygdala
breast
caudate
cerebellum
cerebral cortex
cervix
colon
endometrium
esophagus
fallopian tube
heart muscle
hippocampus
hypothalamus
kidney
liver
lung
nucleus accumbens
ovary
pancreas
pituitary gland
prostate
putamen
retina
salivary gland
skeletal muscle
skin
small intestine
spinal cord
spleen
stomach
substantia nigra
testis
thyroid gland
urinary bladder
vagina

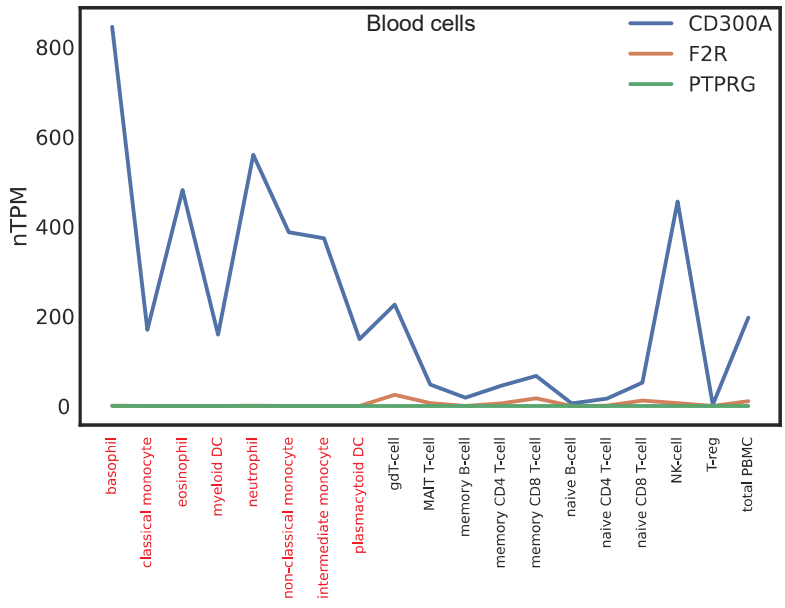
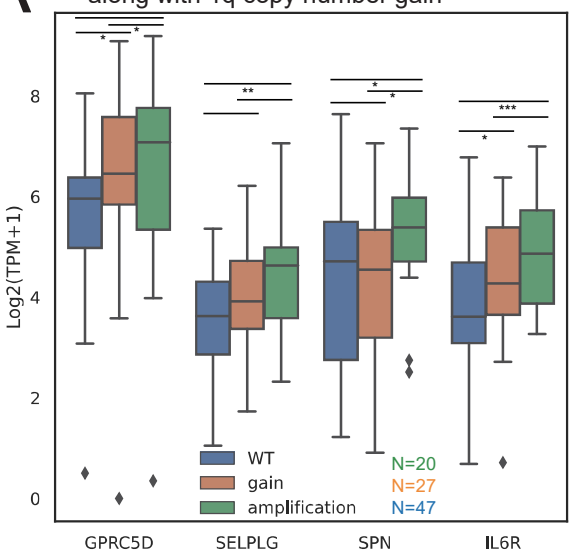
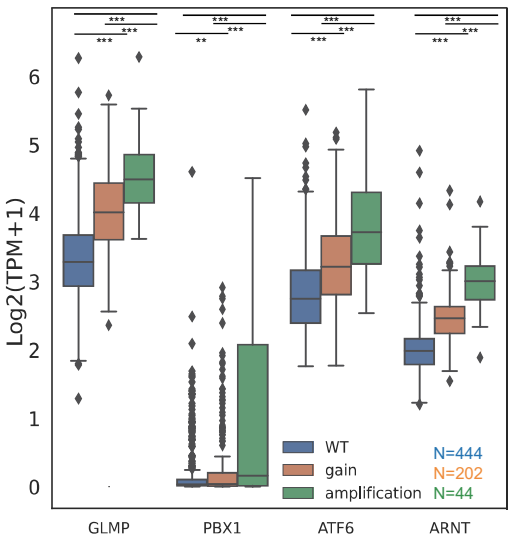

D
CD300A
F2R
PTPRG
Blood cells
nTPM
0
200
400
600
800
basophil
classical monocyte
eosinophil
myeloid DC
neutrophil
non-classical monocyte
intermediate monocyte
plasmacytoid DC
gdT-cell
MAIT T-cell
memory B-cell
memory CD4 T-cell
memory CD8 T-cell
naive B-cell
naive CD4 T-cell
naive CD8 T-cell
NK-cell
T-reg
total PBMC

Fig. S6. Characteristics of candidates identified from high-risk subtypes. A. Expression level of 125 candidates in healthy organs documented in the human protein atlas (THPA). **B**. Expression level of 125 candidates in blood cells documented in THPA. **C**. Three candidates identified in PR subgroup and their expression in healthy organs. **D**. Three candidates identified in PR subgroup and their expression in blood cells.

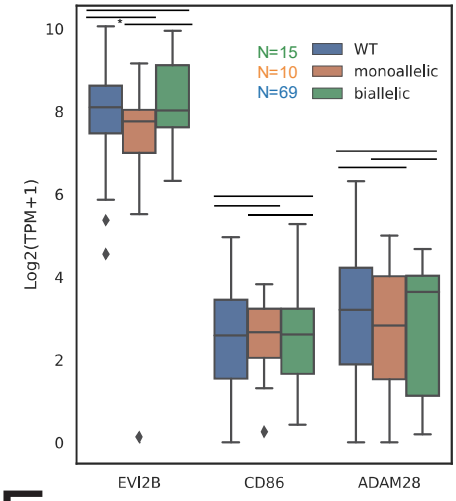
A


Elevated expression of several genes along with 1q copy number gain
Log2(TPM+1)
WT
gain
amplification
N=20
N=27
N=47
GPRC5D
SELPLG
SPN
IL6R

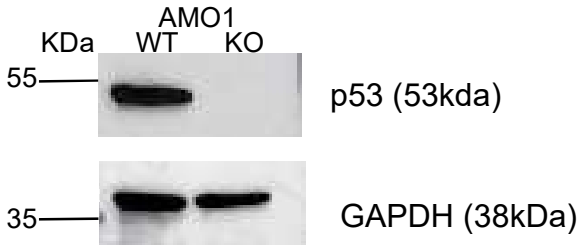
B


Expression of top ranked regulators of GPRC5D
Log2(TPM+1)
WT
gain
amplification
N=444
N=202
N=44
GLMP
PBX1
ATF6
ARNT

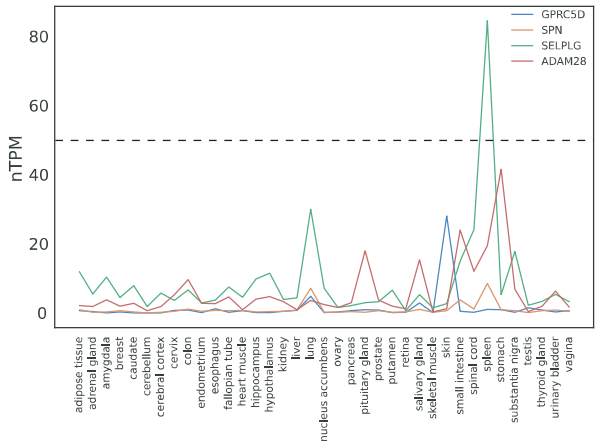
C


Expression of candidates in patients with different TP53 status
Log2(TPM+1)
N=15
N=10
N=69
WT
monoallelic
biallelic
EVI2B
CD86
ADAM28

D


AMO1
KDa
WT
KO
55
p53 (53kda)
35
GAPDH (38kDa)

E


nTPM
GPRC5D
SPN
SELPLG
ADAM28

F

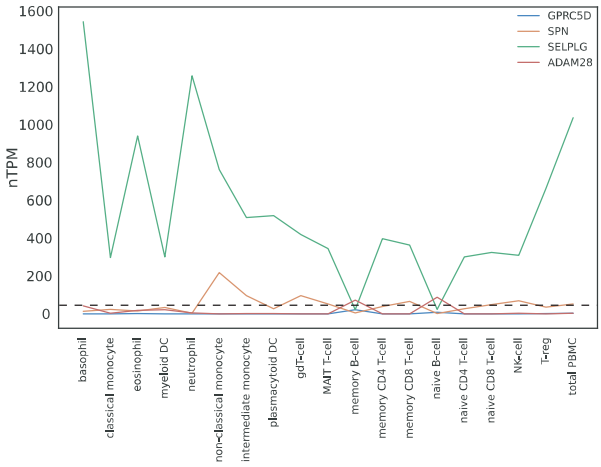

nTPM
GPRC5D
SPN
SELPLG
ADAM28

Fig. S7. Expression of selected candidates and their potential regulators

A. Elevated expression level of four candidates as 1q copy number gain in IU cohort. **B**. Elevated expression level of potential regulators of *GPRC5D* on chromosome 1q. **C**. Expression level of biallelic *TP53*-specific candidates in IU cohort. **D**. A western blot result indicating the p53 protein expression between a *TP53* knockout and a WT cell-line of AMO1. **E.** Expression of selected candidates in healthy organs. **F.** Expression of selected candidates in blood cells. Statistical test in **A**, **B** and **C**: two-sided Mann Whitney U test; significance level: * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

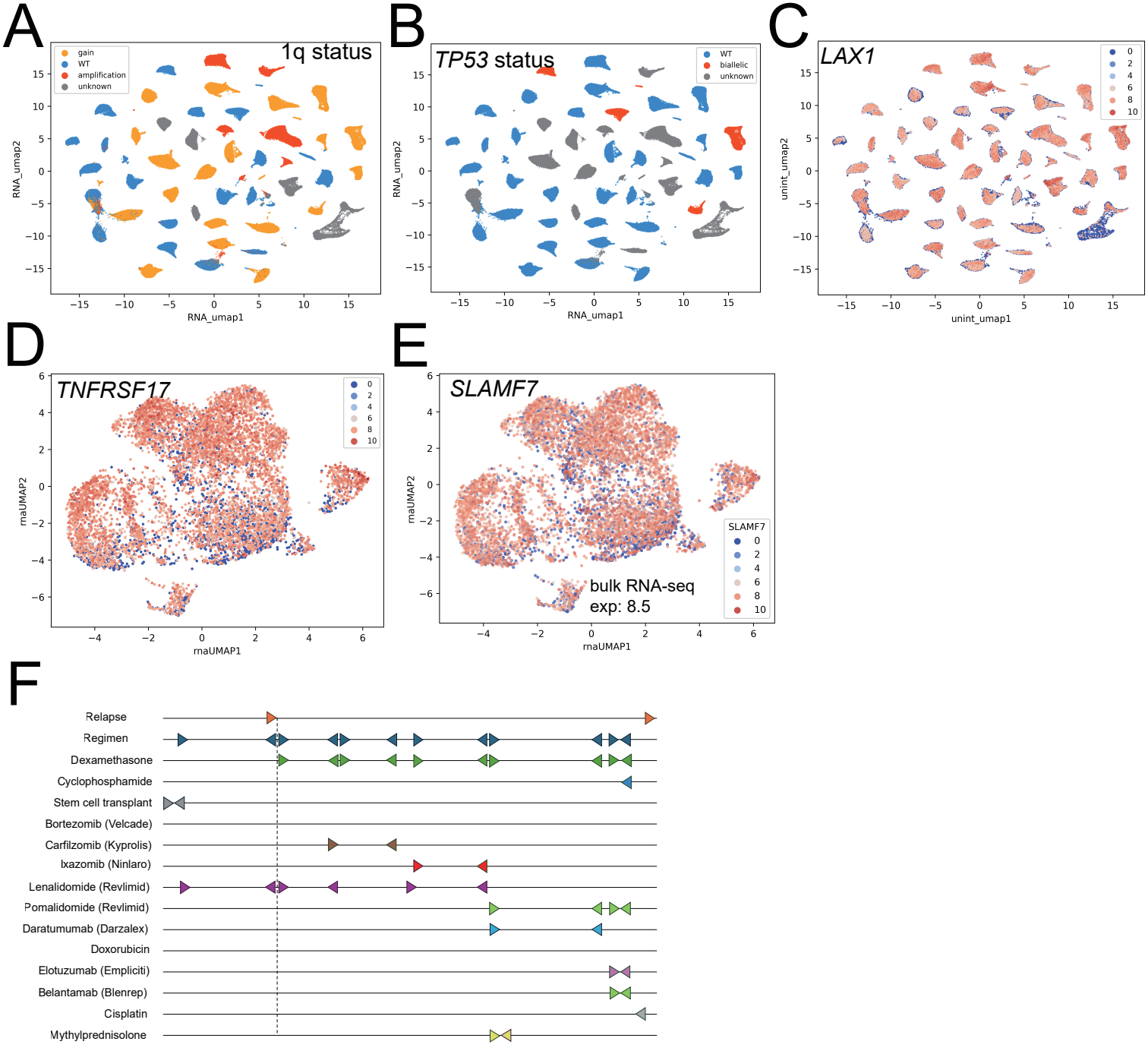
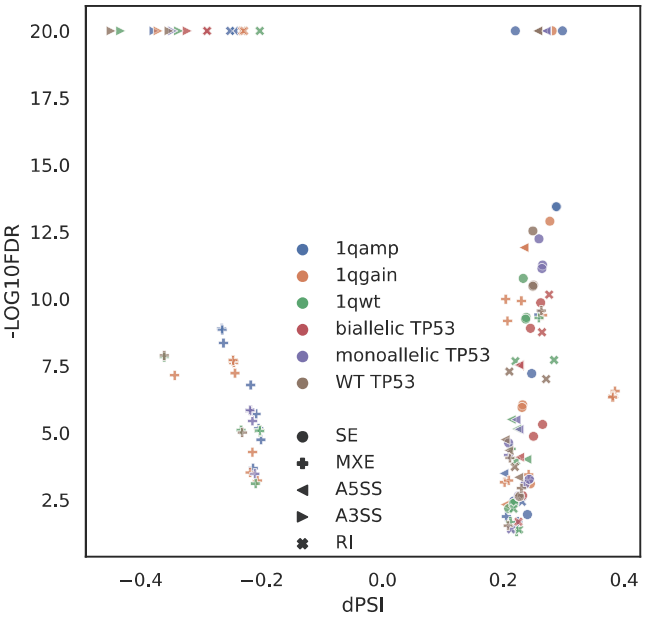

A
1q status
gain
WT
amplification
unknown
RNA_umap1
RNA_umap2
B
TP53 status
WT
biallelic
unknown
RNA_umap1
RNA_umap2
C
LAX1
0
2
4
6
8
10
unint_umap1
unint_umap2
D
TNFRSF17
rnaUMAP1
rnaUMAP2
E
SLAMF7
bulk RNA-seq
exp: 8.5
rnaUMAP1
rnaUMAP2
F
Relapse
Regimen
Dexamethasone
Cyclophosphamide
Stem cell transplant
Bortezomib (Velcade)
Carfilzomib (Kyprolis)
Ixazomib (Ninlaro)
Lenalidomide (Revlimid)
Pomalidomide (Revlimid)
Daratumumab (Darzalex)
Doxorubicin
Elotuzumab (Empliciti)
Belantamab (Blenrep)
Cisplatin
Mythylprednisolone

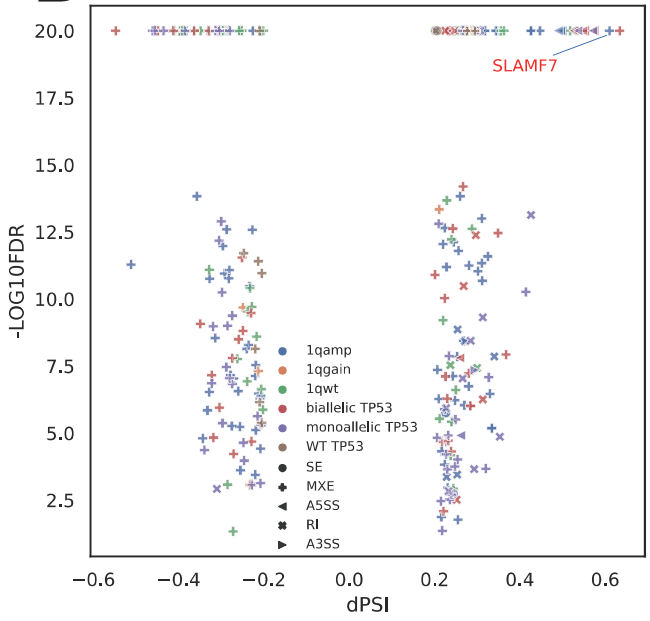
Fig. S8. Characteristics of candidates' expression in subclones.

A. 1q status of 49 patients with single-cell RNA-seq. **B**. *TP53* status of 49 patients with single-cell RNA-seq. **C**. single-cell expression of *LAX1* among subclones of a relapsed t(4;14) patients. **D**. single-cell expression of *TNFRSF17*/BCMA among subclones of a relapsed t(4;14) patients. **E**. expression level of *SLAMF7*/CD319 in a relapsed t(4;14) patient. **F**. Treatment timeline of the relapsed t(4;14) patient.

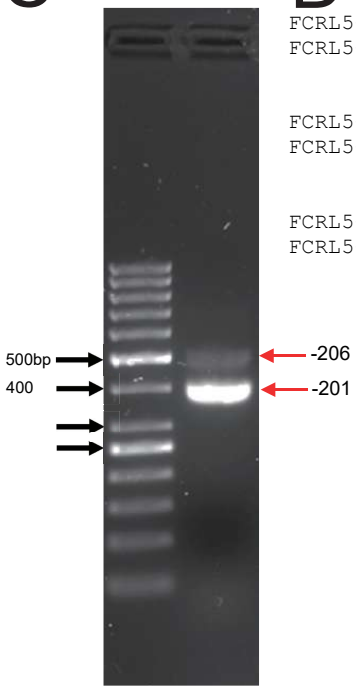
A



B



C



D

```
FCRL5-201       CGGCACCTCAGAACCACAGAGTCTCCTTCAAACACAGAAAGTGGAGCTTGCAGGATCAGCGAAG---------
FCRL5-206       CGGCACCTCAGAACCACAGAGTCTCCTTCAAACACAGAAAGTGGAGCTTGCAGGATCAGCGAAGCTGTAATCC

FCRL5-201       ------------------------------------------------------------------------
FCRL5-206       CAGCGCTTTGAGAGGCCTAGGTGAGCAGATCACTCGAGCCCAGGAGTTCAACATTAGCCTGGGCAGCATGAGG

FCRL5-201       -----------CTGAAGAAAAATCCAAGTGCACAGGGCTACTGAGAGGGGAGCCCTGGGCCTGGCATCTGTACTC
FCRL5-206       AAATCCCATCTCTGAAGAAAAATCCAAGTGCACAGGGCTACTGAGAGGGGAGCCCTGGGCCTGGCATCTGTACTC
```

Fig. S9. Alternative splicing landscapes of 125 high-risk identified candidates.

A. Alternative splicing landscapes among 125 high-risk candidates in MMRF cohort. **B**. Alternative splicing landscapes among 125 high-risk candidates in IU cohort. **C**. A RT-PCR experiment result indicating the expression of *FCRL5*-206. Columns in (**C**): 1. A PDX sample; 2. A SACHI cell-line 3. No reverse transcriptase control; 4. NTC (PCR negative control). Rows in (**C**): fragment size. Expected fragment size: cryptic exon: 495 bps; WT: 404 bps. **D**. cDNA sequences of exon 4, the cryptic exon (dashed basepairs) and exon3.

Supplementary Table legends

Table S1. Summary of candidates identified under all conditions.

Table S2. Association between expression of candidate found in population and patient survival.

Table S3. Association between expression of candidate found in primary subtypes and patient survival.

Table S4. Association between expression of candidate found in high-risk subtypes and patient survival.

Table S5. Primer designing in TP53 knockout experiments.