Figure S1

A

IMmotion150 (n = 263)

Identification of gene signatures based on association with clinical outcome
T-effector: CD8A, IFNG, PRF1, EOMES, CD274
Angiogenesis: VEGFA, KDR, ESM1, PECAM1, CD34, ANGPTL4

↓

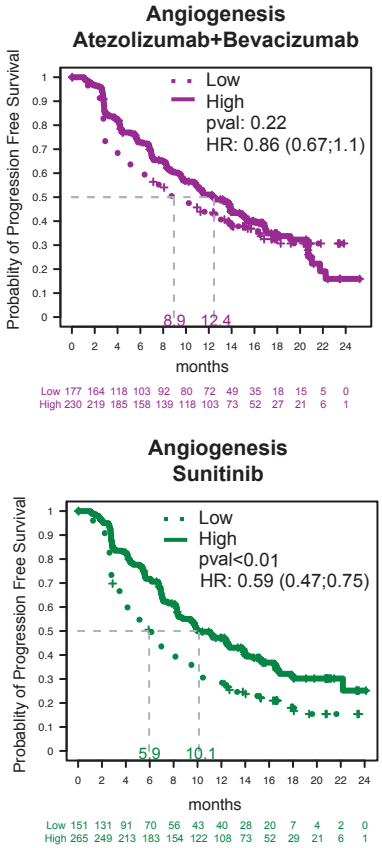
Absolute cutoff selection based on PFS HR
T-effector: 2.93 (40% prevalence)
Angiogenesis: 5.82 (50% prevalence)

↓

IMmotion151 (n = 823)

Pre-specified analysis of association with PFS
Unstratified HR and log-rank tests were used for PFS analyses
in biomarker-evaluable patients

B



C

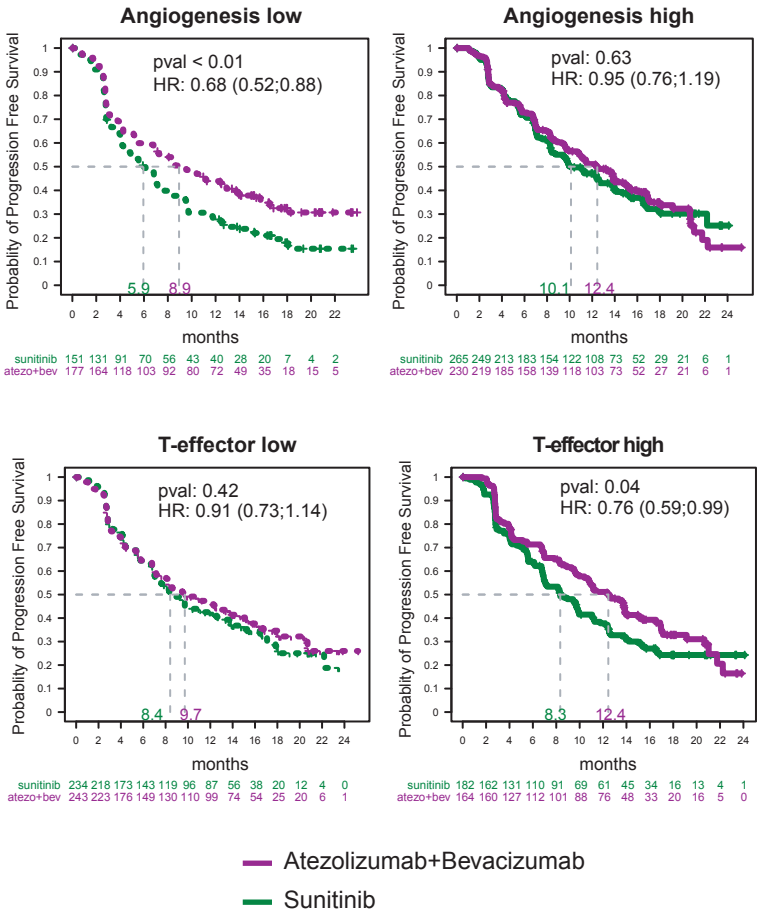


Figure S1: Validation of angiogenesis and T-effector transcriptional signature from IMmotion150 in IMmotion151, Related to STAR Methods. A. Workflow depicting the validation strategy for Angiogenesis and T-effector signatures established in IMmotion150. **B**. Kaplan-Meier curves of progression free survival (PFS) by treatment arm (*purple*, atezolizumab+bevacizumab; *green*, sunitinib) in patients with angiogenesis low (interrupted line) or high (continuous line) tumors. **C**. Kaplan-Meier curves of PFS by treatment arm (*purple*, atezolizumab+bevacizumab; green, sunitinib) in patients with angiogenesis low or high and patients with T-effector low or high tumors.

Figure S2

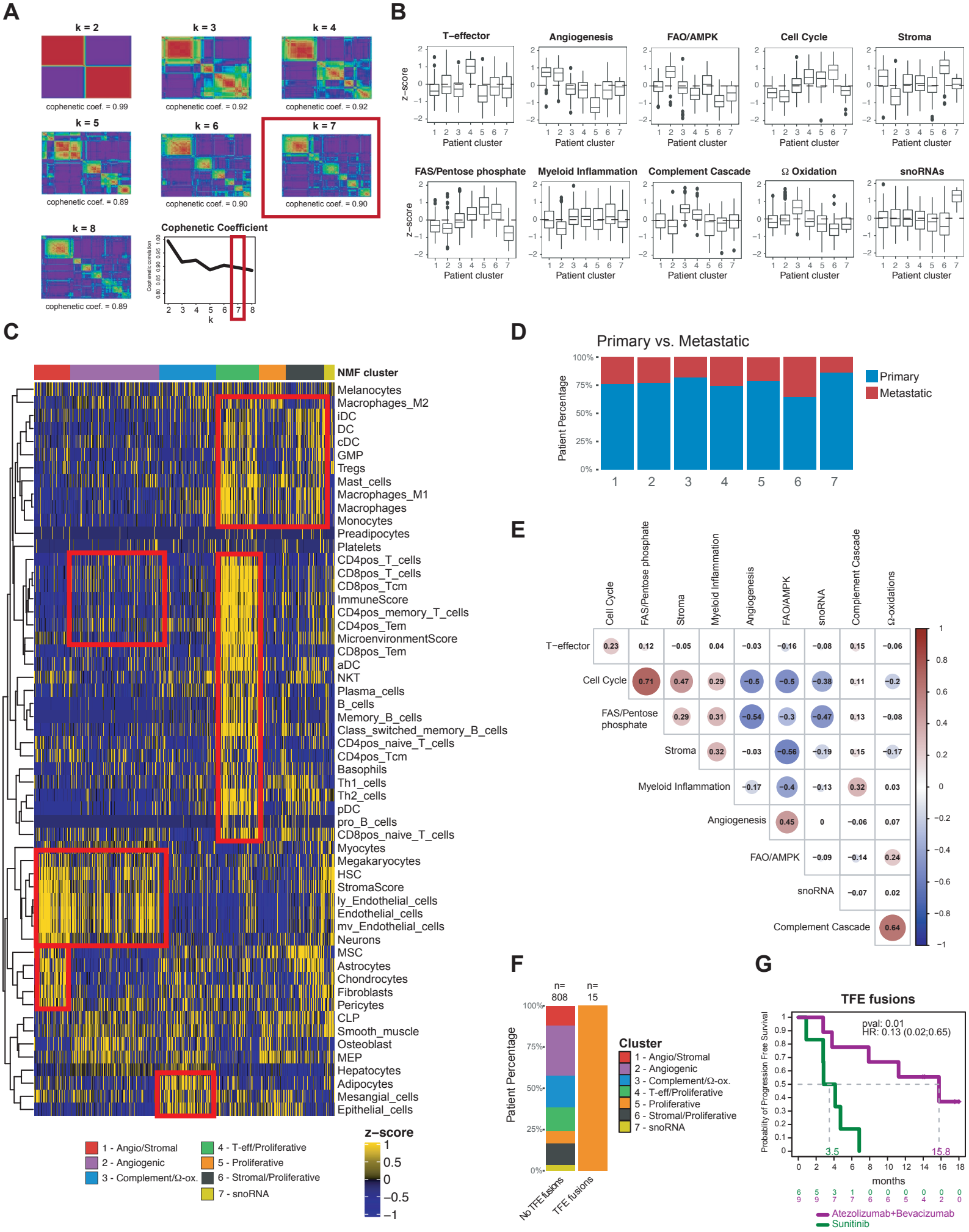

A
k = 2
cophenetic coef. = 0.99
k = 3
cophenetic coef. = 0.92
k = 4
cophenetic coef. = 0.92
k = 5
cophenetic coef. = 0.89
k = 6
cophenetic coef. = 0.90
k = 7
cophenetic coef. = 0.90
k = 8
cophenetic coef. = 0.89
Cophenetic Coefficient
Cophenetic correlation
k
B
T−effector
Angiogenesis
FAO/AMPK
Cell Cycle
Stroma
FAS/Pentose phosphate
Myeloid Inflammation
Complement Cascade
Ω Oxidation
snoRNAs
z−score
Patient cluster
C
NMF cluster
Melanocytes
Macrophages_M2
iDC
DC
cDC
GMP
Tregs
Mast_cells
Macrophages_M1
Macrophages
Monocytes
Preadipocytes
Platelets
CD4pos_T_cells
CD8pos_T_cells
CD8pos_Tcm
ImmuneScore
CD4pos_memory_T_cells
CD4pos_Tem
MicroenvironmentScore
CD8pos_Tem
aDC
NKT
Plasma_cells
B_cells
Memory_B_cells
Class_switched_memory_B_cells
CD4pos_naive_T_cells
CD4pos_Tcm
Basophils
Th1_cells
Th2_cells
pDC
pro_B_cells
CD8pos_naive_T_cells
Myocytes
Megakaryocytes
HSC
StromaScore
ly_Endothelial_cells
Endothelial_cells
mv_Endothelial_cells
Neurons
MSC
Astrocytes
Chondrocytes
Fibroblasts
Pericytes
CLP
Smooth_muscle
Osteoblast
MEP
Hepatocytes
Adipocytes
Mesangial_cells
Epithelial_cells
1 - Angio/Stromal
2 - Angiogenic
3 - Complement/Ω-ox.
4 - T-eff/Proliferative
5 - Proliferative
6 - Stromal/Proliferative
7 - snoRNA
z−score
D
Primary vs. Metastatic
Patient Percentage
Primary
Metastatic
E
Cell Cycle
FAS/Pentose phosphate
Stroma
Myeloid Inflammation
Angiogenesis
FAO/AMPK
snoRNA
Complement Cascade
Ω-oxidations
T−effector
F
n= 808
n= 15
Patient Percentage
No TFE fusions
TFE fusions
Cluster
1 - Angio/Stromal
2 - Angiogenic
3 - Complement/Ω-ox.
4 - T-eff/Proliferative
5 - Proliferative
6 - Stromal/Proliferative
7 - snoRNA
G
TFE fusions
pval: 0.01
HR: 0.13 (0.02;0.65)
Probability of Progression Free Survival
months
3.5
15.8
Atezolizumab+Bevacizumab
Sunitinib

Figure S2: Non-negative matrix factorization (NMF) cluster analysis, Related to Figure 1. **A.** Selection of cluster number based on consensus matrices for k=2 to k=8, and measure of cophenetic coefficient stability at various values of k. k=7, with a cophenetic coefficient of 0.90, was chosen. **B.** Transcriptional z-scores for the 10 signatures presented in the heatmap in Figure 1D were represented as boxplots by patient cluster. **C.** Hierarchical clustering of deconvolution z-scores obtained from xCell. Samples are ordered by NMF cluster. **D**. Distribution of primary and metastatic tumors in NMF clusters. **E.** Correlations between transcriptional signatures across IMmotion151 data set. Signature z-scores were computed for each of the 823 samples from IMmotion151 and Pearson correlations between signatures were calculated in a pairwise fashion. Positive and negative correlations are shown as *red* and *blue* circles respectively. The diameter of the circles is proportional to the absolute Pearson R value, which is also numerically displayed in the circles. **F**. Barchart representing the distribution of NMF clusters in tumors with or without *TFE* fusions. Fusions in *TFE3* and *TFEB* were grouped together. Tumors from 12 patients had *TFE3* fusions and 3 patients had *TFEB* fusions. **G**. Kaplan-Meier curve of progression free survival (PFS) by treatment arm (*purple*, atezolizumab+bevacizumab; *green*, sunitinib) in patients with *TFE*-fusions

Figure S3

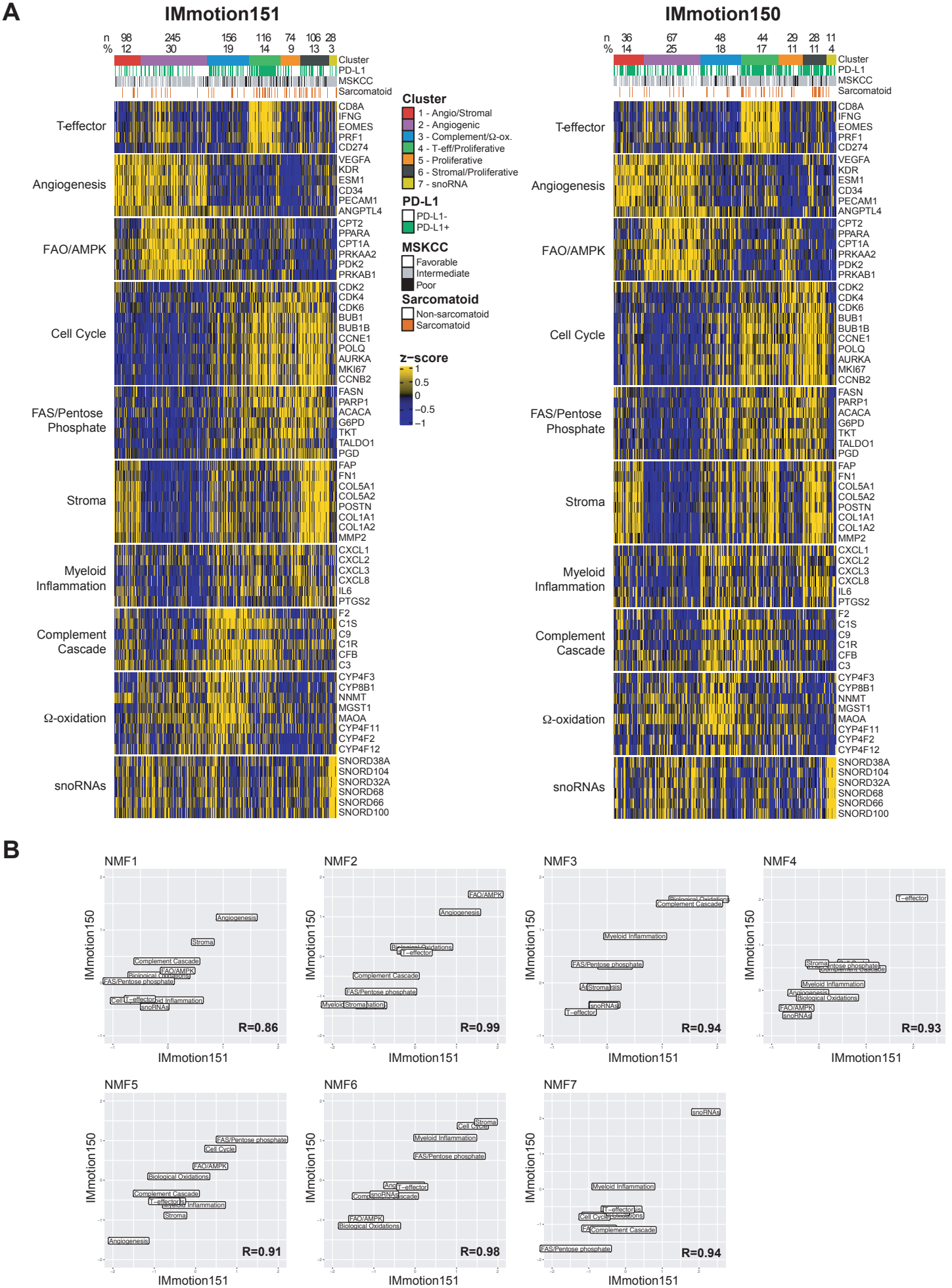

A
IMmotion151
IMmotion150
n 98 245 156 116 74 106 28
% 12 30 19 14 9 13 3
n 36 67 48 44 29 28 11
% 14 25 18 17 11 11 4
Cluster
PD-L1
MSKCC
Sarcomatoid
T-effector
CD8A
IFNG
EOMES
PRF1
CD274
Angiogenesis
VEGFA
KDR
ESM1
CD34
PECAM1
ANGPTL4
FAO/AMPK
CPT2
PPARA
CPT1A
PRKAA2
PDK2
PRKAB1
Cell Cycle
CDK2
CDK4
CDK6
BUB1
BUB1B
CCNE1
POLQ
AURKA
MKI67
CCNB2
FAS/Pentose Phosphate
FASN
PARP1
ACACA
G6PD
TKT
TALDO1
PGD
Stroma
FAP
FN1
COL5A1
COL5A2
POSTN
COL1A1
COL1A2
MMP2
Myeloid Inflammation
CXCL1
CXCL2
CXCL3
CXCL8
IL6
PTGS2
Complement Cascade
F2
C1S
C9
C1R
CFB
C3
Ω-oxidation
CYP4F3
CYP8B1
NNMT
MGST1
MAOA
CYP4F11
CYP4F2
CYP4F12
snoRNAs
SNORD38A
SNORD104
SNORD32A
SNORD68
SNORD66
SNORD100
Cluster
1 - Angio/Stromal
2 - Angiogenic
3 - Complement/Ω-ox.
4 - T-eff/Proliferative
5 - Proliferative
6 - Stromal/Proliferative
7 - snoRNA
PD-L1
PD-L1-
PD-L1+
MSKCC
Favorable
Intermediate
Poor
Sarcomatoid
Non-sarcomatoid
Sarcomatoid
z−score
1
0.5
0
−0.5
−1
B
NMF1
R=0.86
NMF2
R=0.99
NMF3
R=0.94
NMF4
R=0.93
NMF5
R=0.91
NMF6
R=0.98
NMF7
R=0.94
IMmotion150
IMmotion151

Figure S3: Validation of non-negative matrix factorization (NMF) clusters in IMmotion150, Related to Figure 1. A machine learning algorithm using random forest was developed to train a model in IMmotion151 to classify patients into NMF groups based on the 3,072 (top 10%) most variable genes in this data set (STAR Methods). This model was applied to assign patients from IMmotion150 into each cluster. **A.** The heatmap generated in Figure 1D in IMmotion151 (*left panel*) was then derived in IMmotion150 (*right panel*). Signature patterns across patient clusters were highly conserved between IMmotion151 and IMmotion150 datasets. **B.** X-Y charts representing the mean aggregate z-score for our ten transcriptional signatures in IMmotion151 (x-axis) and IMmotion150 (y-axis) for each NMF group. The Pearson R value is represented on each plot.

Figure S4

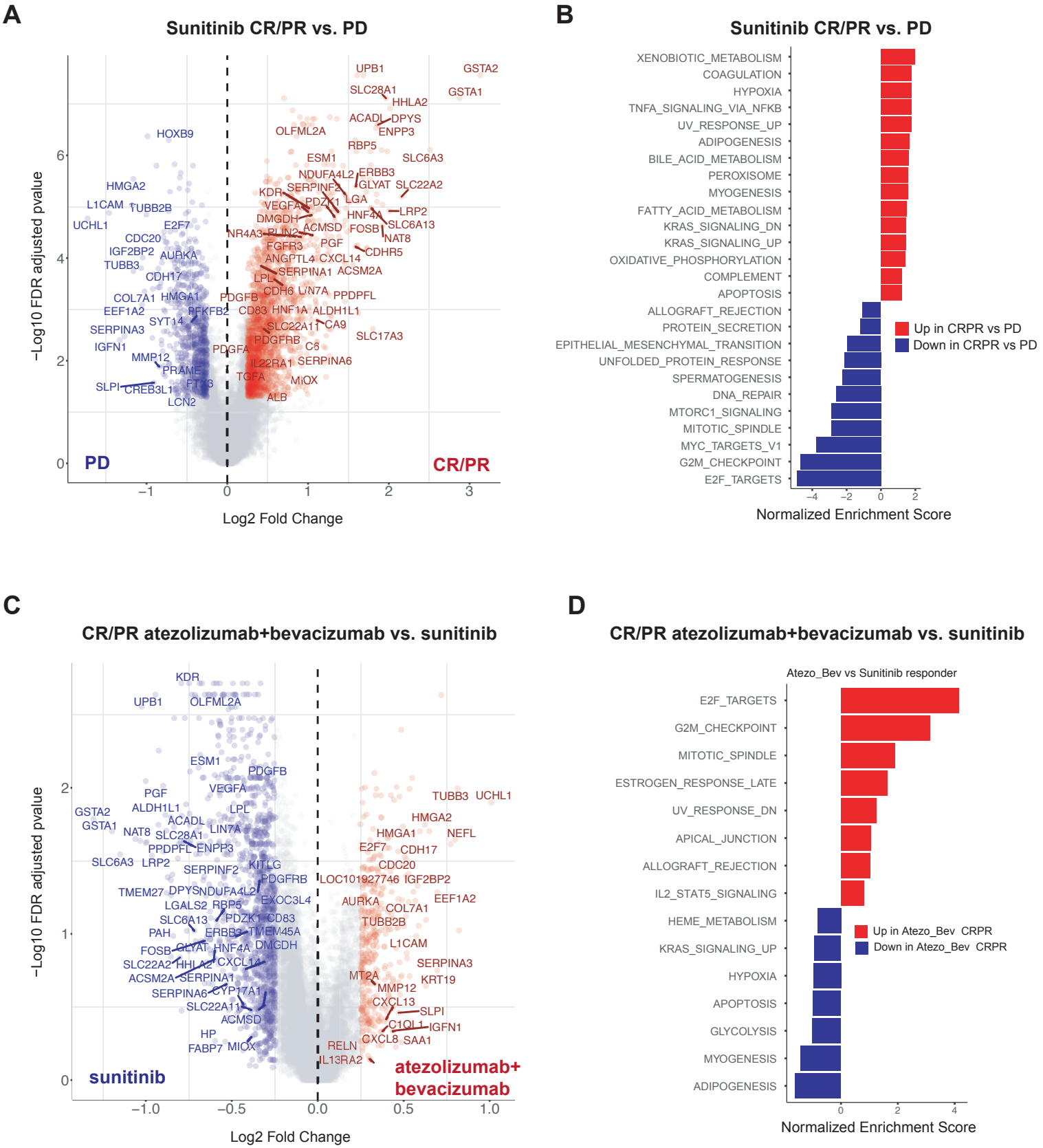
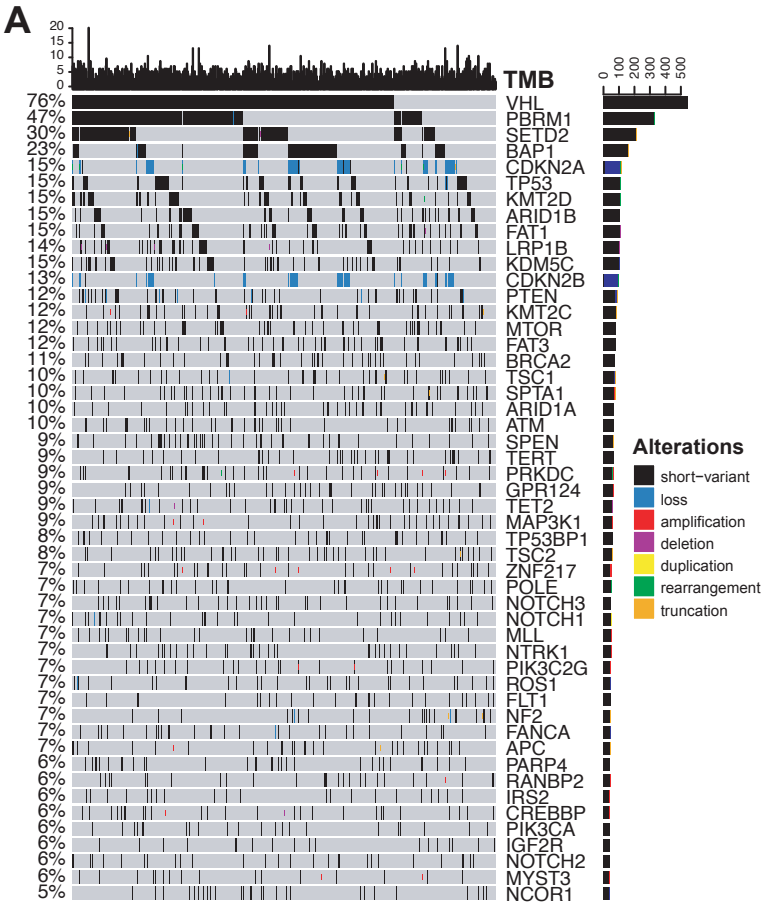
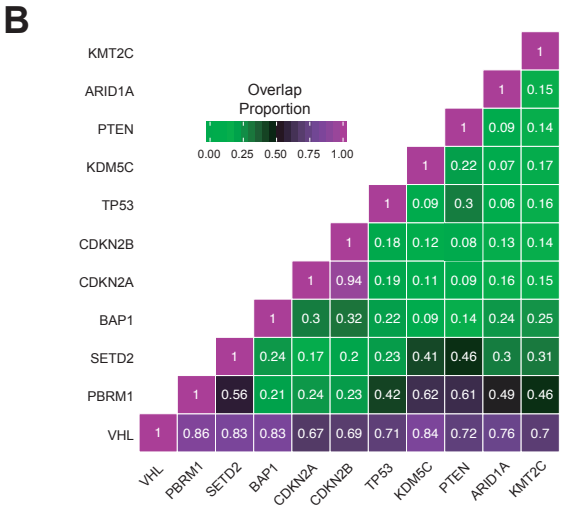
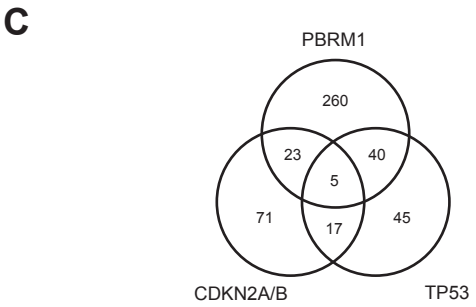

A
Sunitinib CR/PR vs. PD
−Log10 FDR adjusted pvalue
Log2 Fold Change
PD
CR/PR
B
Sunitinib CR/PR vs. PD
XENOBIOTIC_METABOLISM
COAGULATION
HYPOXIA
TNFA_SIGNALING_VIA_NFKB
UV_RESPONSE_UP
ADIPOGENESIS
BILE_ACID_METABOLISM
PEROXISOME
MYOGENESIS
FATTY_ACID_METABOLISM
KRAS_SIGNALING_DN
KRAS_SIGNALING_UP
OXIDATIVE_PHOSPHORYLATION
COMPLEMENT
APOPTOSIS
ALLOGRAFT_REJECTION
PROTEIN_SECRETION
EPITHELIAL_MESENCHYMAL_TRANSITION
UNFOLDED_PROTEIN_RESPONSE
SPERMATOGENESIS
DNA_REPAIR
MTORC1_SIGNALING
MITOTIC_SPINDLE
MYC_TARGETS_V1
G2M_CHECKPOINT
E2F_TARGETS
Up in CRPR vs PD
Down in CRPR vs PD
Normalized Enrichment Score
C
CR/PR atezolizumab+bevacizumab vs. sunitinib
−Log10 FDR adjusted pvalue
Log2 Fold Change
sunitinib
atezolizumab+
bevacizumab
D
CR/PR atezolizumab+bevacizumab vs. sunitinib
Atezo_Bev vs Sunitinib responder
E2F_TARGETS
G2M_CHECKPOINT
MITOTIC_SPINDLE
ESTROGEN_RESPONSE_LATE
UV_RESPONSE_DN
APICAL_JUNCTION
ALLOGRAFT_REJECTION
IL2_STAT5_SIGNALING
HEME_METABOLISM
KRAS_SIGNALING_UP
HYPOXIA
APOPTOSIS
GLYCOLYSIS
MYOGENESIS
ADIPOGENESIS
Up in Atezo_Bev CRPR
Down in Atezo_Bev CRPR
Normalized Enrichment Score

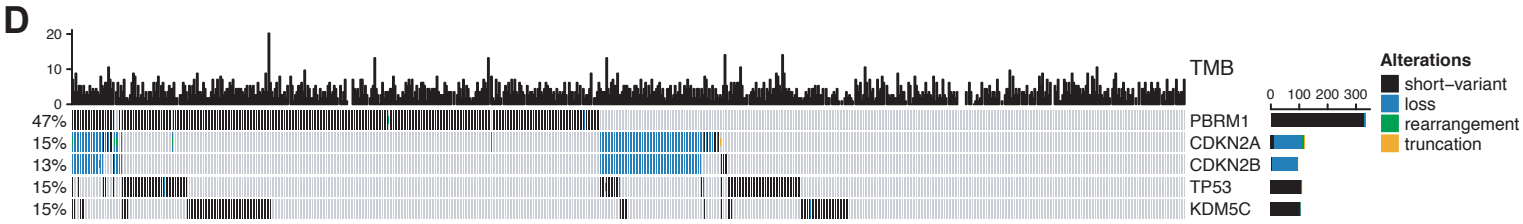
Figure S4: Analysis of differentially expressed genes and pathways in responder vs. non-responder patients, Related to Figure 2. A. Volcano plot depicting differentially expressed genes between responders (CR/PR) and non-responders (PD) in the sunitinib arm. Genes with FDR-corrected $p<0.05$ and absolute log-fold change >= 0.25 are represented in *red* or *blue*. **B.** Bar chart representing pathway enrichment scores for the top upregulated or downregulated MSigDb hallmark gene sets within the differentially expressed genes identified in *A*. **C.** Volcano plot depicting differentially expressed genes in responders (CR/PR) treated with atezolizumab+bevacizumab or sunitinib. Genes with FDR-corrected $p<0.05$ and absolute log-fold change >= 0.25 are represented in *red* or *blue*. **D.** Bar chart representing pathway enrichment scores for the top upregulated or downregulated MSigDb hallmark gene sets within the differentially expressed genes identified in *C*. CR, complete response; PR, partial response; PD, progressive disease.

Figure S5









E

Gene	Alteration Status	PFS HR	p-value
VHL	non-altered	0.77	0.167
VHL	altered	0.89	0.318
PBRM1	non-altered	0.75	0.032
PBRM1	altered	0.99	0.961
SETD2	non-altered	0.89	0.322
SETD2	altered	0.78	0.165
BAP1	non-altered	0.85	0.141
BAP1	altered	0.89	0.540
CDKN2A	non-altered	0.92	0.462
CDKN2A	altered	0.57	0.011
TP53	non-altered	0.91	0.394
TP53	altered	0.68	0.097
KMT2D	non-altered	0.87	0.177
KMT2D	altered	0.85	0.519
ARID1B	non-altered	0.84	0.108
ARID1B	altered	1.04	0.869
FAT1	non-altered	0.87	0.189
FAT1	altered	0.83	0.450
LRP1B	non-altered	0.86	0.147
LRP1B	altered	0.88	0.626
KDM5C	non-altered	0.87	0.183
KDM5C	altered	0.77	0.351
CDKN2B	non-altered	0.91	0.397
CDKN2B	altered	0.62	0.044
PTEN	non-altered	0.88	0.216
PTEN	altered	0.76	0.283
KMT2C	non-altered	0.93	0.484
KMT2C	altered	0.48	0.012
MTOR	non-altered	0.85	0.110
MTOR	altered	1	0.993
FAT3	non-altered	0.83	0.081
FAT3	altered	1.11	0.710
BRCA2	non-altered	0.85	0.127
BRCA2	altered	0.92	0.790
TSC1	non-altered	0.84	0.102
TSC1	altered	0.98	0.929
SPTA1	non-altered	0.91	0.337
SPTA1	altered	0.56	0.063
ARID1A	non-altered	0.92	0.433
ARID1A	altered	0.49	0.043
ATM	non-altered	0.91	0.358
ATM	altered	0.54	0.057
CDKN2A/B	non-altered	0.91	0.409
CDKN2A/B	altered	0.63	0.031

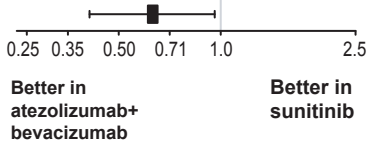


Figure S5: Extended analyses of somatic alterations in tumors from IMmotion 151, Related to Figures 3 and 4. A. Oncoprint depicting the top 50 most frequently altered genes in tumors from IMmotion151. **B.** Heatmap representing the overlap proportion between pairs of the most common somatic alterations in this dataset. Proportion is calculated as the ration of overlap between two groups over the size of the smaller group. The heatmap highlights minimal overlap between PBRM1 mutations and BAP1/CDKN2A/B alterations. **C.** Venn diagram representing the overlap between tumors somatically altered in PBRM1, CDKN2/B and TP53. **D.** Oncoprint depicting somatic alterations in PBRM1, CDKN2A/B, TP53 and KDM5C. **E.** Association between somatic alterations and PFS in genes altered in >10% of tumors. Forest plot depicting PFS hazard ratios comparing patients treated with atezolizumab+bevacizumab vs. sunitinib by somatic alteration status for each gene. Whiskers represent 95% confidence intervals.

Figure S6

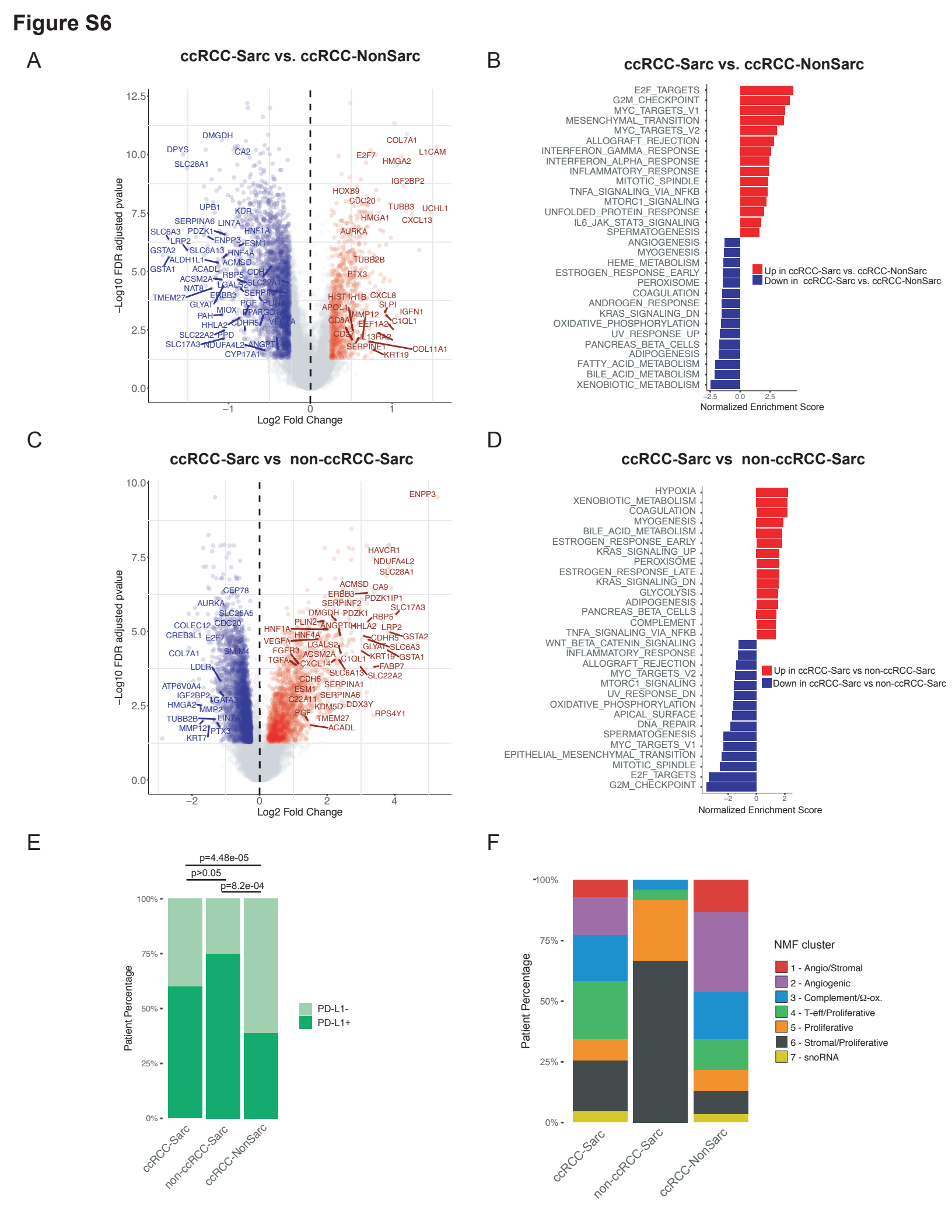

A
ccRCC-Sarc vs. ccRCC-NonSarc
−Log10 FDR adjusted pvalue
Log2 Fold Change
DMGDH
DPYS
CA2
SLC28A1
UPB1
KOR
SERPINA6
LIN7A
SLC6A3
PDZK1
HNF1A
LRP2
ENPP3
ESM1
GSTA2
SLC6A13
HNF4A
ALDH1L1
ACMSD
GSTA1
ACADL
CDH2
RBP5
ACSM2A
LGALS2
SLC22A1
NAT8
SERPINF2
TMEM27
ERBB3
GLYAT
PGF
PLIN2
MIOX
PPARGC1A
PAH
HHLA2
CDHR5
VEGFA
SLC22A2
PPD
SLC17A3
NDUFA4L2
ANGPTL4
CYP17A1
COL7A1
L1CAM
E2F7
HMGA2
IGF2BP2
HOXB9
CDC20
TUBB3
UCHL1
HMGA1
CXCL13
AURKA
TUBB2B
PTX3
HIST1H1B
CXCL8
APOL1
SLPI
MMP12
IGFN1
CD8A
EEF1A2
C1QL1
CD27
IL13RA2
SERPINE1
COL11A1
KRT19
B
ccRCC-Sarc vs. ccRCC-NonSarc
E2F_TARGETS
G2M_CHECKPOINT
MYC_TARGETS_V1
MESENCHYMAL_TRANSITION
MYC_TARGETS_V2
ALLOGRAFT_REJECTION
INTERFERON_GAMMA_RESPONSE
INTERFERON_ALPHA_RESPONSE
INFLAMMATORY_RESPONSE
MITOTIC_SPINDLE
TNFA_SIGNALING_VIA_NFKB
MTORC1_SIGNALING
UNFOLDED_PROTEIN_RESPONSE
IL6_JAK_STAT3_SIGNALING
SPERMATOGENESIS
ANGIOGENESIS
MYOGENESIS
HEME_METABOLISM
ESTROGEN_RESPONSE_EARLY
PEROXISOME
COAGULATION
ANDROGEN_RESPONSE
KRAS_SIGNALING_DN
OXIDATIVE_PHOSPHORYLATION
UV_RESPONSE_UP
PANCREAS_BETA_CELLS
ADIPOGENESIS
FATTY_ACID_METABOLISM
BILE_ACID_METABOLISM
XENOBIOTIC_METABOLISM
Up in ccRCC-Sarc vs. ccRCC-NonSarc
Down in ccRCC-Sarc vs. ccRCC-NonSarc
Normalized Enrichment Score
C
ccRCC-Sarc vs non-ccRCC-Sarc
−Log10 FDR adjusted pvalue
Log2 Fold Change
ENPP3
HAVCR1
NDUFA4L2
SLC28A1
ACMSD
CA9
CEP78
ERBB3
PDZK1IP1
AURKA
SERPINF2
SLC17A3
SLC25A5
DMGDH
PDZK1
RBP5
COLEC12
CDC20
PLIN2
ANGPTL4
HHLA2
LRP2
CREB3L1
E2F7
HNF1A
HNF4A
CDHR5
GSTA2
VEGFA
LGALS2
GLYAT
SLC6A3
COL7A1
SMIM4
FGFR3
ACSM2A
KRT19
GSTA1
TGFA
CXCL14
C1QL1
FABP7
LDLR
SLC22A2
SLC6A13
CDH6
ATP6V0A4
SERPINA1
ESM1
IGF2BP2
SERPINA6
C22A11
HMGA2
DDX3Y
MMP2
KDM5D
PGF
RPS4Y1
TUBB2B
LIN7A
TMEM27
MMP12
PTX3
ACADL
KRT7
Up in ccRCC-Sarc vs non-ccRCC-Sarc
Down in ccRCC-Sarc vs non-ccRCC-Sarc
Normalized Enrichment Score
HYPOXIA
XENOBIOTIC_METABOLISM
COAGULATION
MYOGENESIS
BILE_ACID_METABOLISM
ESTROGEN_RESPONSE_EARLY
KRAS_SIGNALING_UP
PEROXISOME
ESTROGEN_RESPONSE_LATE
KRAS_SIGNALING_DN
GLYCOLYSIS
ADIPOGENESIS
PANCREAS_BETA_CELLS
COMPLEMENT
TNFA_SIGNALING_VIA_NFKB
WNT_BETA_CATENIN_SIGNALING
INFLAMMATORY_RESPONSE
ALLOGRAFT_REJECTION
MYC_TARGETS_V2
MTORC1_SIGNALING
UV_RESPONSE_DN
OXIDATIVE_PHOSPHORYLATION
APICAL_SURFACE
DNA_REPAIR
SPERMATOGENESIS
MYC_TARGETS_V1
EPITHELIAL_MESENCHYMAL_TRANSITION
MITOTIC_SPINDLE
E2F_TARGETS
G2M_CHECKPOINT
E
p=4.48e-05
p>0.05
p=8.2e-04
Patient Percentage
PD-L1-
PD-L1+
ccRCC-Sarc
non-ccRCC-Sarc
ccRCC-NonSarc
F
Patient Percentage
NMF cluster
1 - Angio/Stromal
2 - Angiogenic
3 - Complement/Ω-ox.
4 - T-eff/Proliferative
5 - Proliferative
6 - Stromal/Proliferative
7 - snoRNA
ccRCC-Sarc
non-ccRCC-Sarc
ccRCC-NonSarc

Figure S6: Comparison of molecular features of clear cell RCC sarcomatoid (ccRCC-Sarc), clear cell RCC non-sarcomatoid (ccRCC-NonSarc) and non-clear cell sarcomatoid (non-ccRCC-Sarc tumors, Related to Figure 5. A. Volcano plot depicting differentially expressed genes between ccRCC-Sarc and ccRCC-NonSarc tumors. Genes with FDR-corrected $p<0.05$ and absolute log-fold change >= 0.25 are represented in *red* or *blue*. **B.** Bar chart representing pathway enrichment scores for the top upregulated or downregulated MSigDb hallmark gene sets within the differentially expressed genes identified in *A*. **C.** Volcano plot depicting differentially expressed genes between ccRCC-Sarc and non-ccRCC-Sarc tumors. Genes with FDR-corrected $p<0.05$ and absolute log-fold change >= 0.25 are represented in *red* or *blue***.** **D.** Bar chart representing pathway enrichment scores for the top upregulated or downregulated MSigDb hallmark gene sets within the differentially expressed genes identified in *C*. **E.** Barchart representing the distribution of PD-L1 expression by immunohistochemistry (IHC) in ccRCC-Sarc, non-ccRCC-Sarc and ccRCC-NonSarc tumors. P-value was obtained from Pearson's Chi-squared test conducted between each pair of conditions. **F.** Barchart representing distribution of non-negative matrix factorization (NMF) clusters in ccRCC-Sarc, non-ccRCC-Sarc and ccRCC-NonSarc tumors.