Cell Reports Methods, Volume *3*

Supplemental information

A topic modeling approach reveals the dynamic T cell composition of peripheral blood during cancer immunotherapy

Xiyu Peng, Jasme Lee, Matthew Adamow, Colleen Maher, Michael A. Postow, Margaret K. Callahan, Katherine S. Panageas, and Ronglai Shen

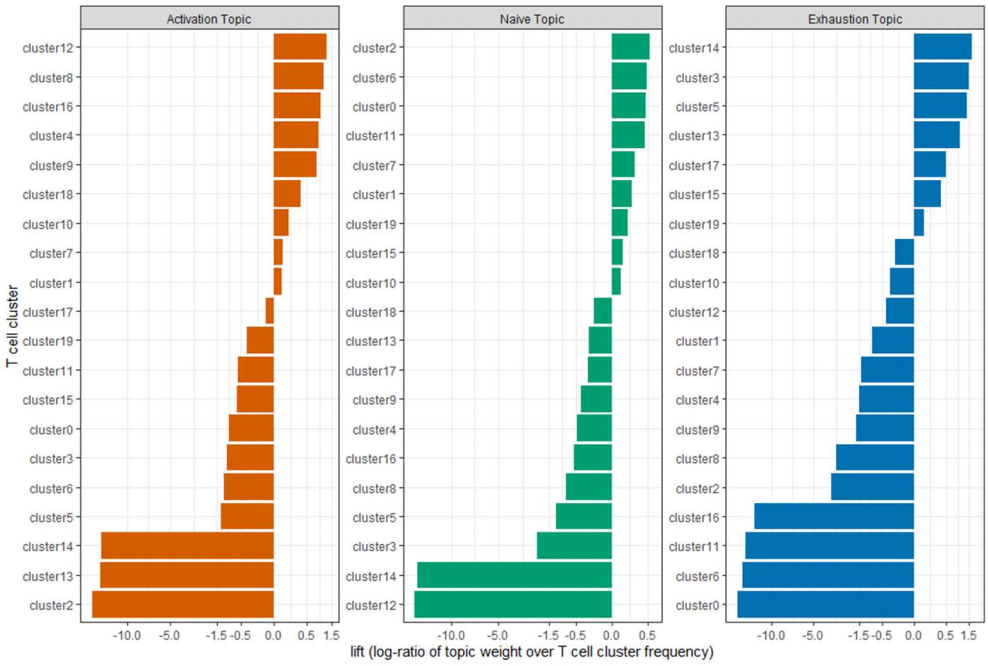
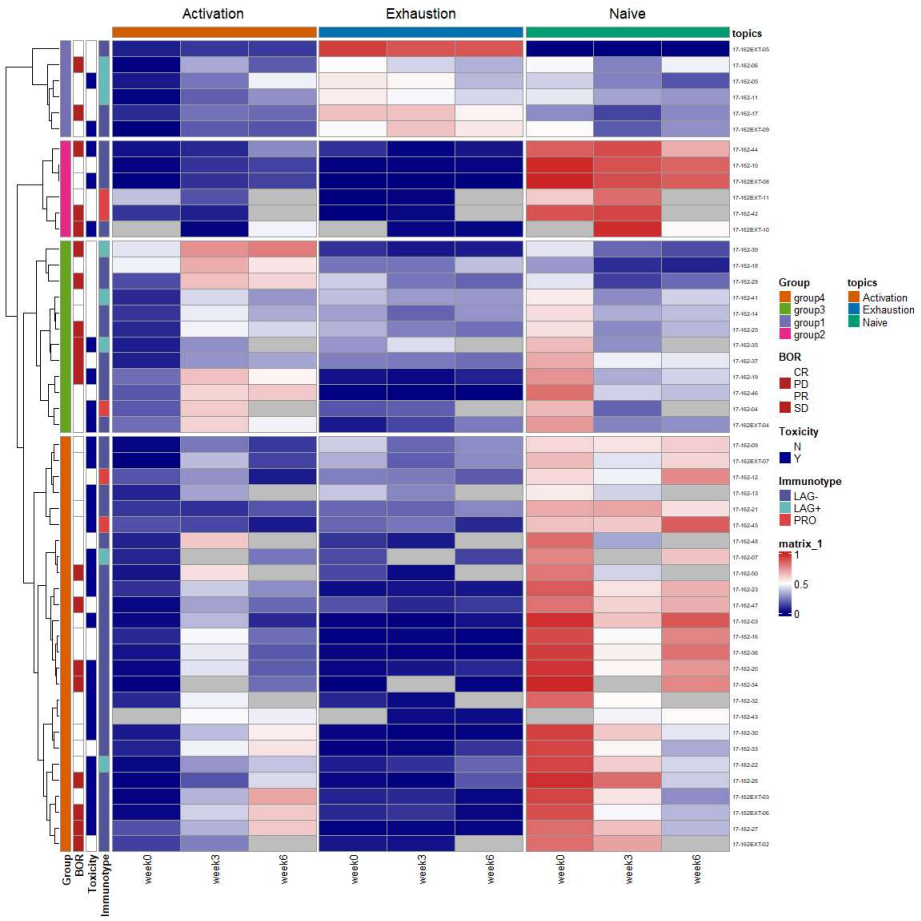


Figure S1: Selection of representative clusters for each topic, related to Figure 3. Lift of clusters for each topic, plotted on a signed square root scale. The metric lift gives high weights to clusters that appear less frequently in other topics. Those clusters that have high lift statistics are identified as representatives of single topics.

a



b

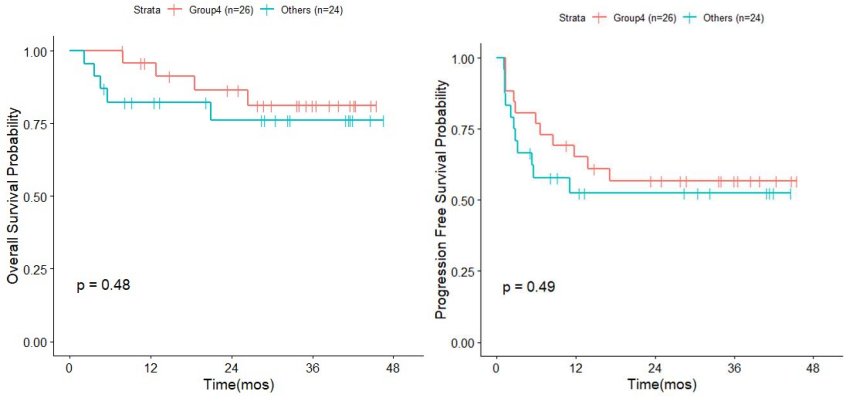


Figure S2: Identification of patient subgroups with LDA, related to Figure 4. a) Patient subgroups revealed by hierarchical clustering on sample topic proportions. b) Kaplan-Meier analysis of OS and PFS stratified by patient subgroup.

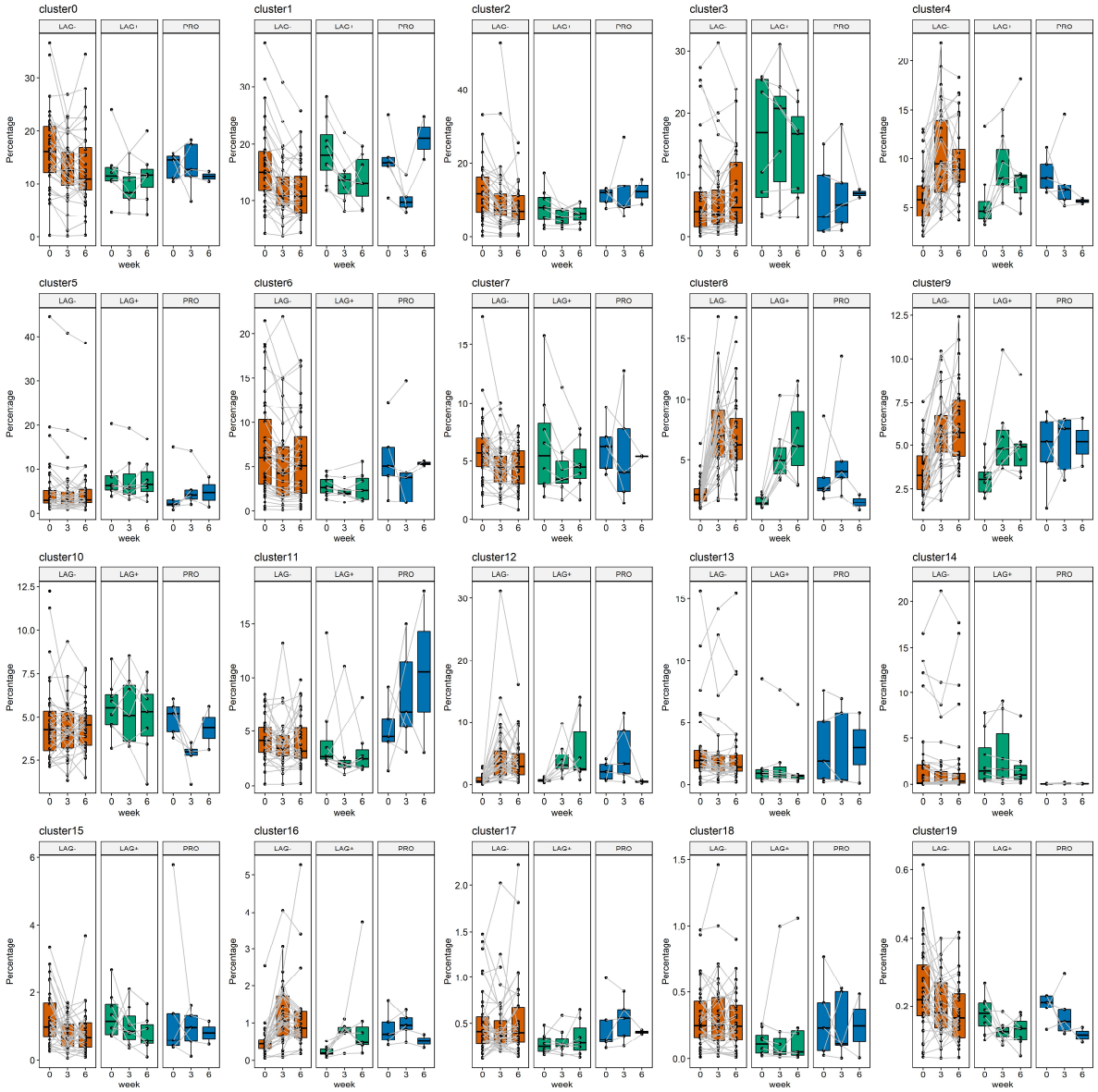


Figure S3: Pharmacodynamics of single clusters across different immunotypes, related to Figure 4.

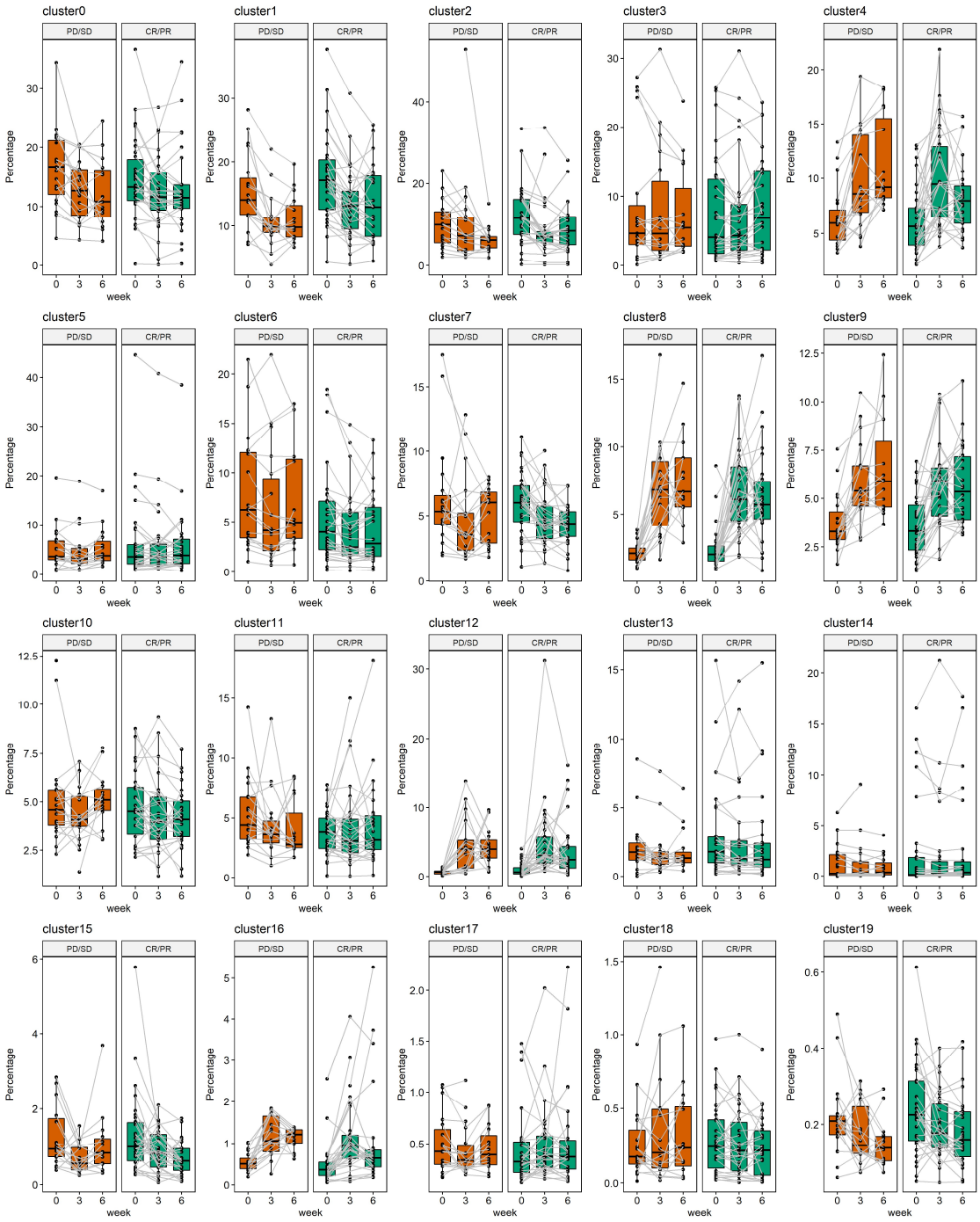


Figure S4: Pharmacodynamics of single clusters across different responses, related to Figure 4.

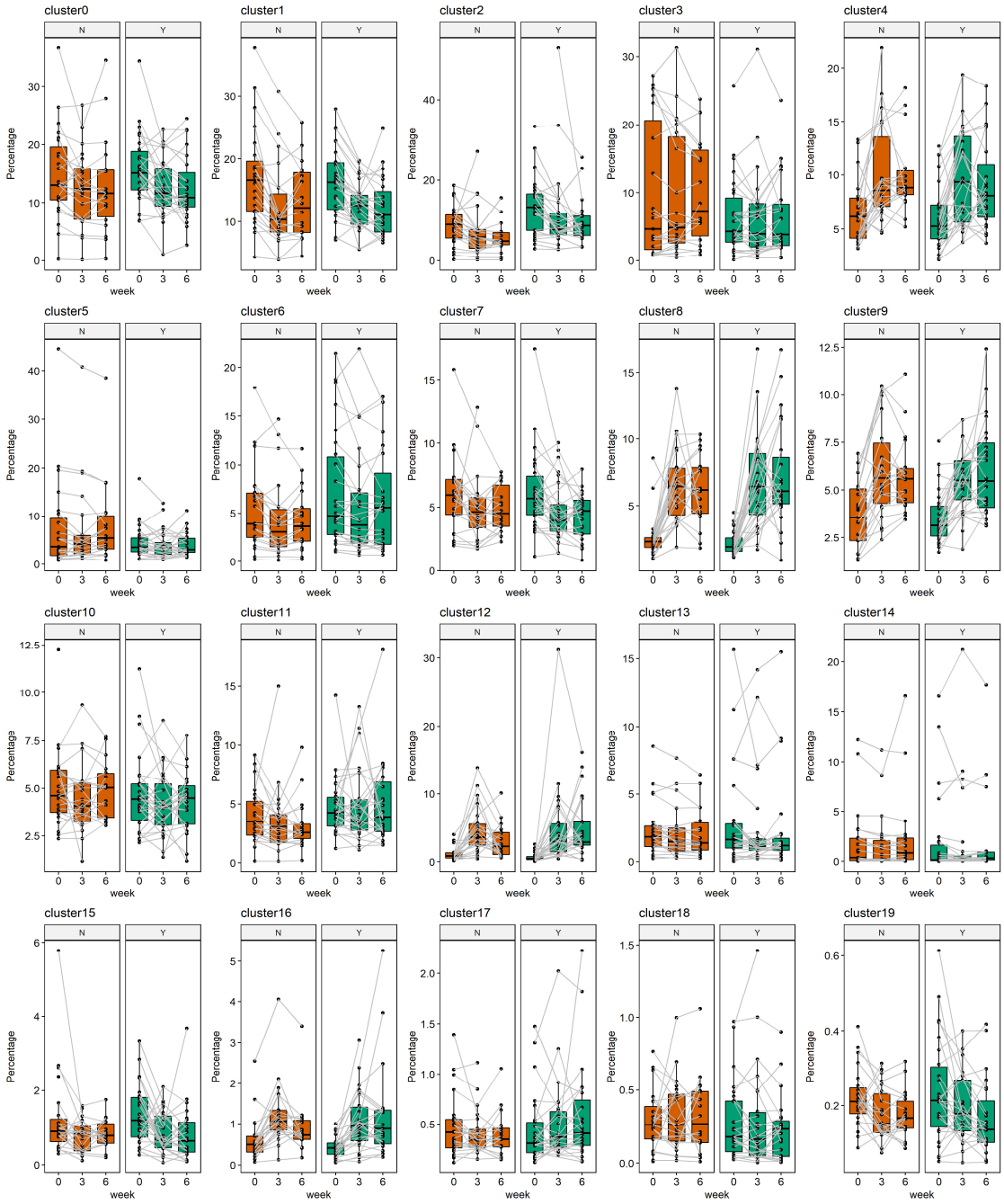
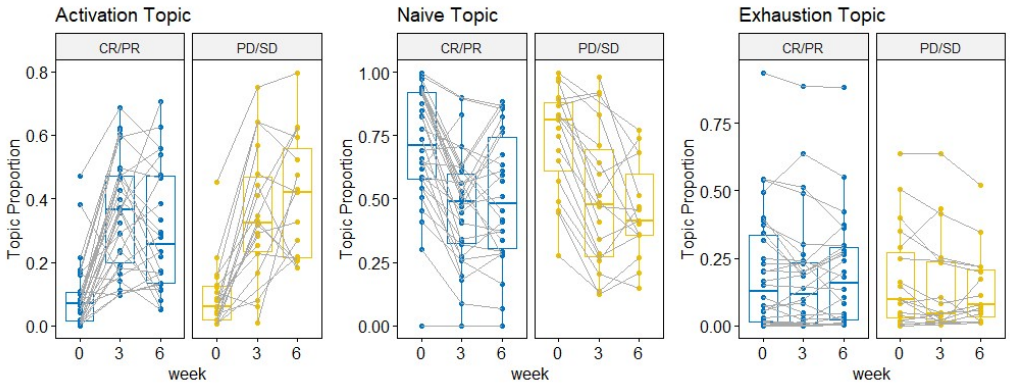


Figure S5: Pharmacodynamics of single clusters across different levels of toxicity, related to Figure 4.

a



b

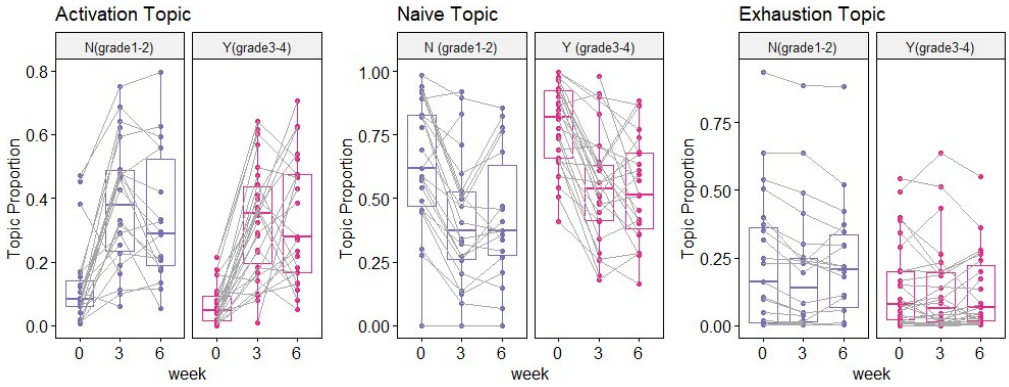
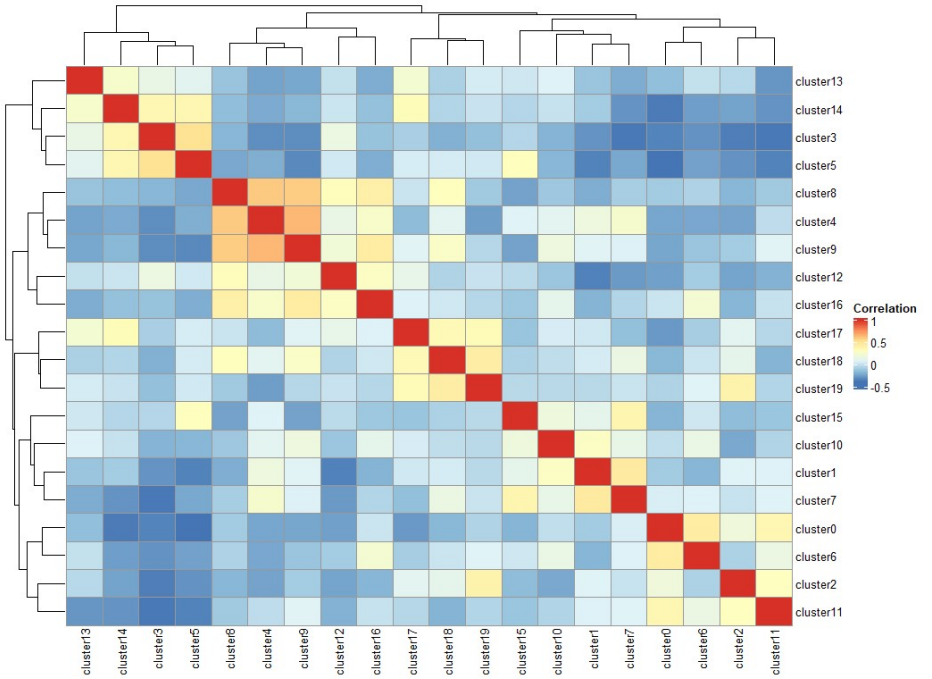


Figure S6: Pharmacodynamics of topic proportions across different responses and levels of toxicity, related to Figure 4.

a



b

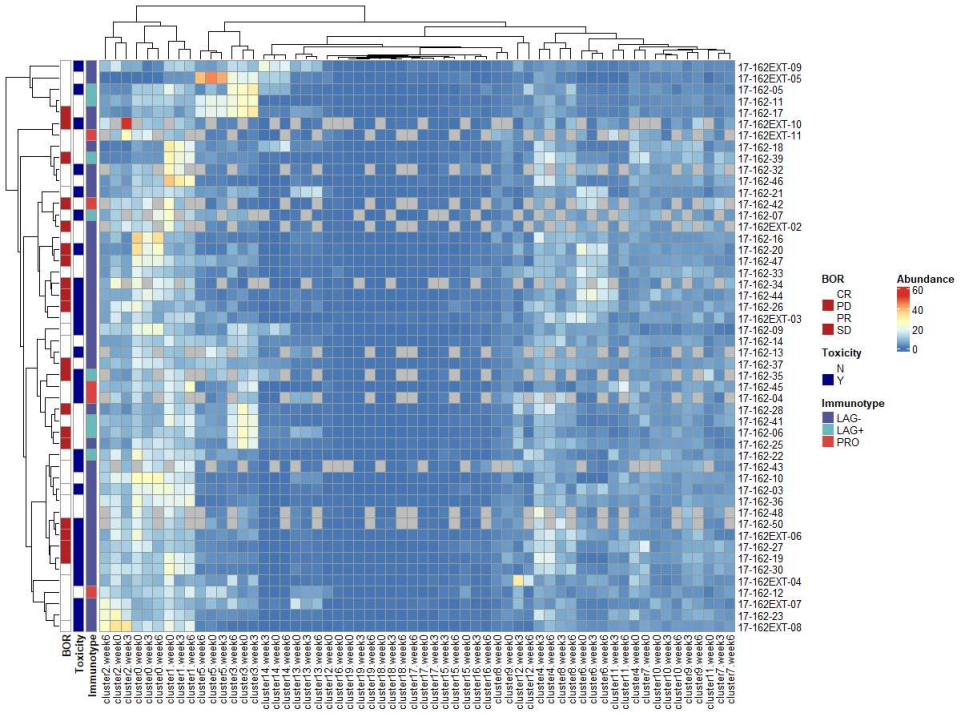


Figure S7: Hierarchical clustering based on a) Pearson correlation of clusters and b) cluster abundances, related to Figure 4.

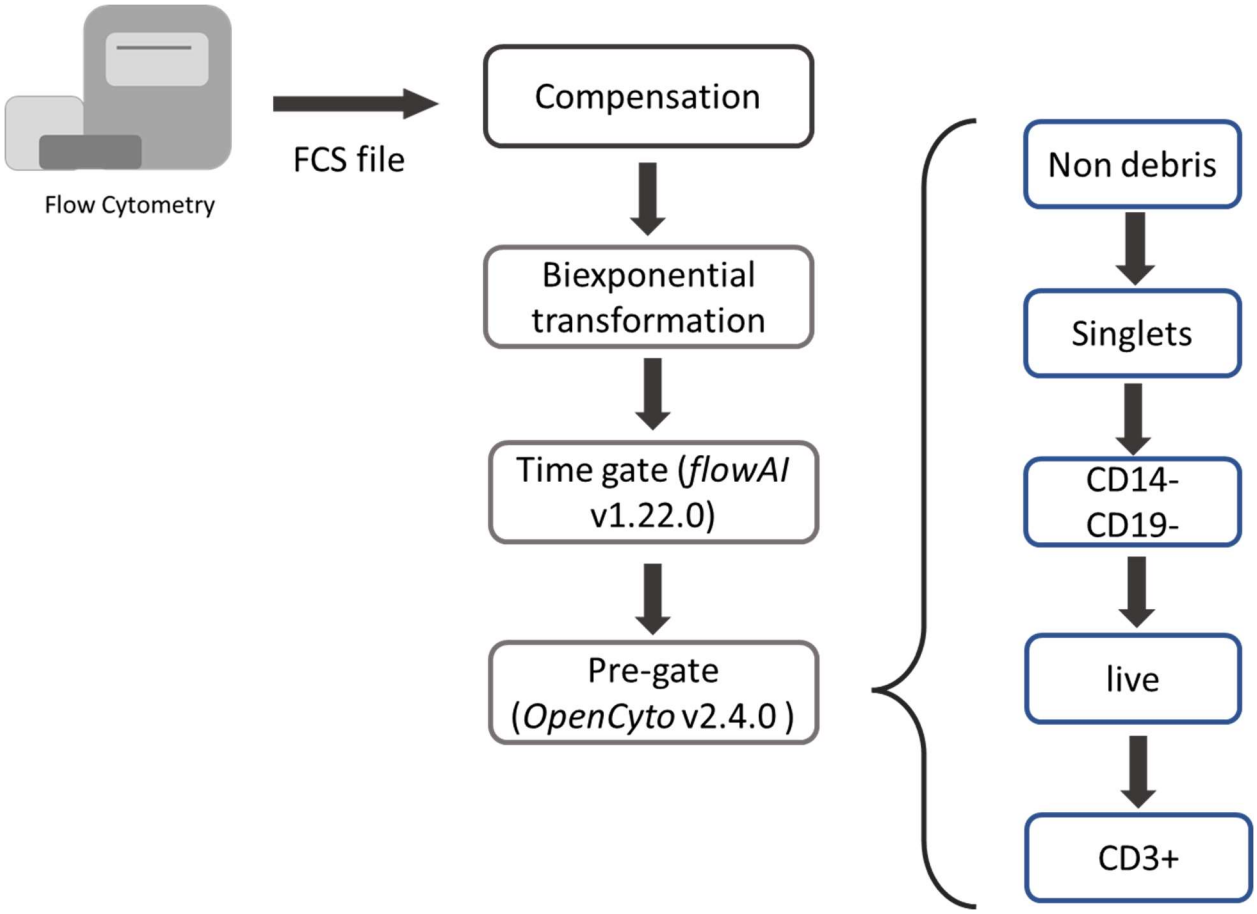


Figure S8: Pre-gating analysis on flow cytometry data, related to STAR Methods.

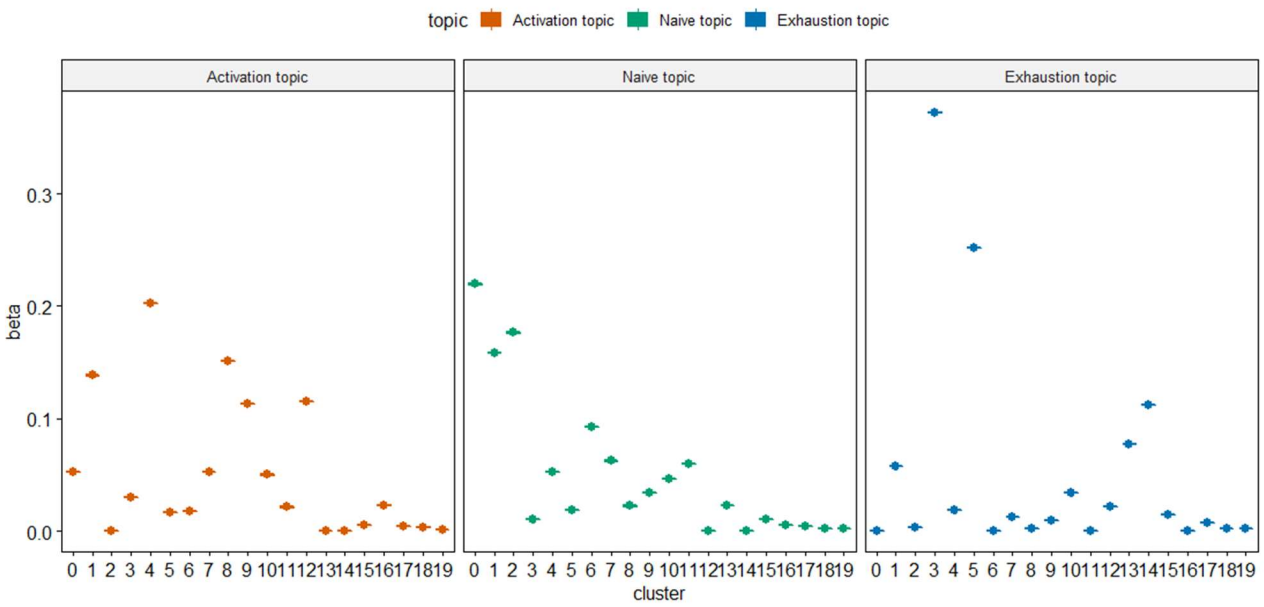


Figure S9: Estimation of the cell-type-by-topic matrix B by Gibbs Sampling under ten random starts, related to STAR Methods.

Table S1. Statistical analysis of single clusters associated with patient clinical outcomes and immunotypes, related to Figure 2.

Cluster	Immunotype		Response		Toxicity	
	Immunotype	Interaction with time	Response	Interaction with time	Toxicity	Interaction with time
0	ns	ns	ns	ns	ns	ns
1	ns	*	ns	ns	ns	ns
2	ns	ns	ns	ns	ns	ns
3	ns	ns	ns	ns	ns	ns
4	ns	***	ns	ns	ns	ns
5	ns	ns	ns	ns	ns	ns
6	*	*	ns	ns	ns	ns
7	ns	ns	ns	ns	ns	ns
8	ns	***	ns	ns	ns	ns
9	ns	ns	ns	ns	ns	ns
10	ns	ns	ns	ns	ns	ns
11	ns	*	ns	ns	ns	ns
12	ns	***	ns	ns	ns	ns
13	ns	ns	ns	ns	ns	ns
14	*	ns	ns	ns	ns	ns
15	ns	ns	ns	ns	ns	ns
16	ns	*	*	ns	ns	ns
17	ns	ns	ns	ns	ns	ns
18	ns	ns	ns	ns	ns	ns
19	ns	ns	ns	ns	ns	ns

*** P < 0.001; ** P < 0.01; * P < 0.05; ns, not significant. P-values for the main effect and the interaction effect with time were given by *nparLD* R package, based on patients with all three timepoints (n=37). P-values was adjusted by Benjamini-Hochberg method with a false discovery rate controlled at 5%.

Table S2. Cell markers in the X50 flow panel, related to STAR Methods.

Cell Marker	Role
CD3	Lineage_T-cell
CD4	Lineage_T-cell
CD8	Lineage_T-cell
CD14	Lineage_Monocyte
CD19	Lineage_B-cell
FoxP3	Lineage_Treg
CXCR5	Chemokine-Tfh cells
CCR4	Chemokine-Th2, Th17, Tregs cells
CD45RA	Differentiation_Memory
CCR7/ CD197	Differentiation_Memory
CD27	Differentiation_Costimulatory
CD28	Differentiation_Costimulatory
CD127	Differentiation
CD25	Differentiation
CD57	Differentiation
Eomes	Transcription Factor- T-cell expansion/proliferation
Tbet	Transcription Factor
Granzyme B	Functional: Cytotoxic T-Lymphocytes
Ki67	T cell proliferation marker/Activation
HLA-DR	Differentiation_Activation
CD38	Differentiation_Activation
ICOS/ CD278	Costimulatory marker
GITR	Costimulatory
CTLA-4/ CD152	Exhaustion
PD-1/ CD279	Activation/exhaustion
TIM-3	Exhaustion
LAG-3/ CD223	Co-inhibitory/Exhaustion
TIGIT	Activation/Exhaustion

Table S3. T cell Gating template used in *openCyto* R package, related to STAR Methods.

alias	pop	parent	dims	gating_method	gating_args	collapseDataForGating	groupBy	preprocessing_method	preprocessing_args
nonDebris	+	root	FSC-A	gate_mindensity					
singlets	+	nonDebris	FSC-A,FSC-H	singletGate					
cd14-cd19-	-	singlets	CD14 19	gate_mindensity					
live	-	cd14-cd19-	L_D	gate_mindensity					
cd3	+	live	CD3	gate_mindensity		TRUE	4		

In the pre-gating procedure, nonDebris, singlets, CD1419-, live, CD3+ cells were gated in the order described in the gating template, which was used as the input of *openCyto* R package (blank cells are the default to be used as the input). See details in the documentation of *openCyto* R package.