Immunity, Volume *44*

Supplemental Information

Mass Cytometry of the Human Mucosal Immune System Identifies Tissue- and Disease-Associated Immune Subsets

Vincent van Unen, Na Li, Ilse Molendijk, Mine Temurhan, Thomas Höllt, Andrea E. van der Meulen-de Jong, Hein W. Verspaget, M. Luisa Mearin, Chris J. Mulder, Jeroen van Bergen, Boudewijn P.F. Lelieveldt, and Frits Koning

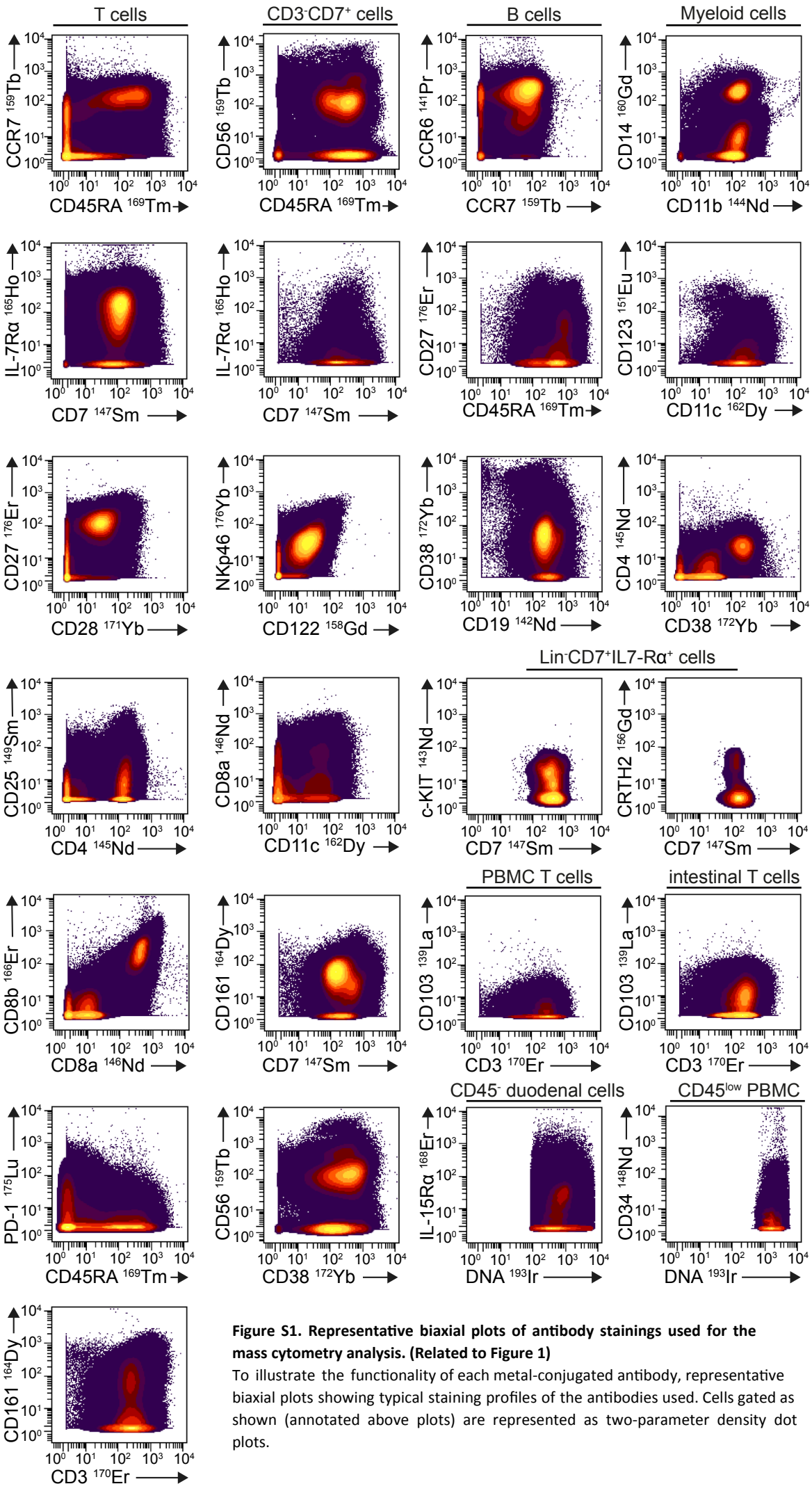

Figure S1. Representative biaxial plots of antibody stainings used for the mass cytometry analysis. (Related to Figure 1)

To illustrate the functionality of each metal-conjugated antibody, representative biaxial plots showing typical staining profiles of the antibodies used. Cells gated as shown (annotated above plots) are represented as two-parameter density dot plots.

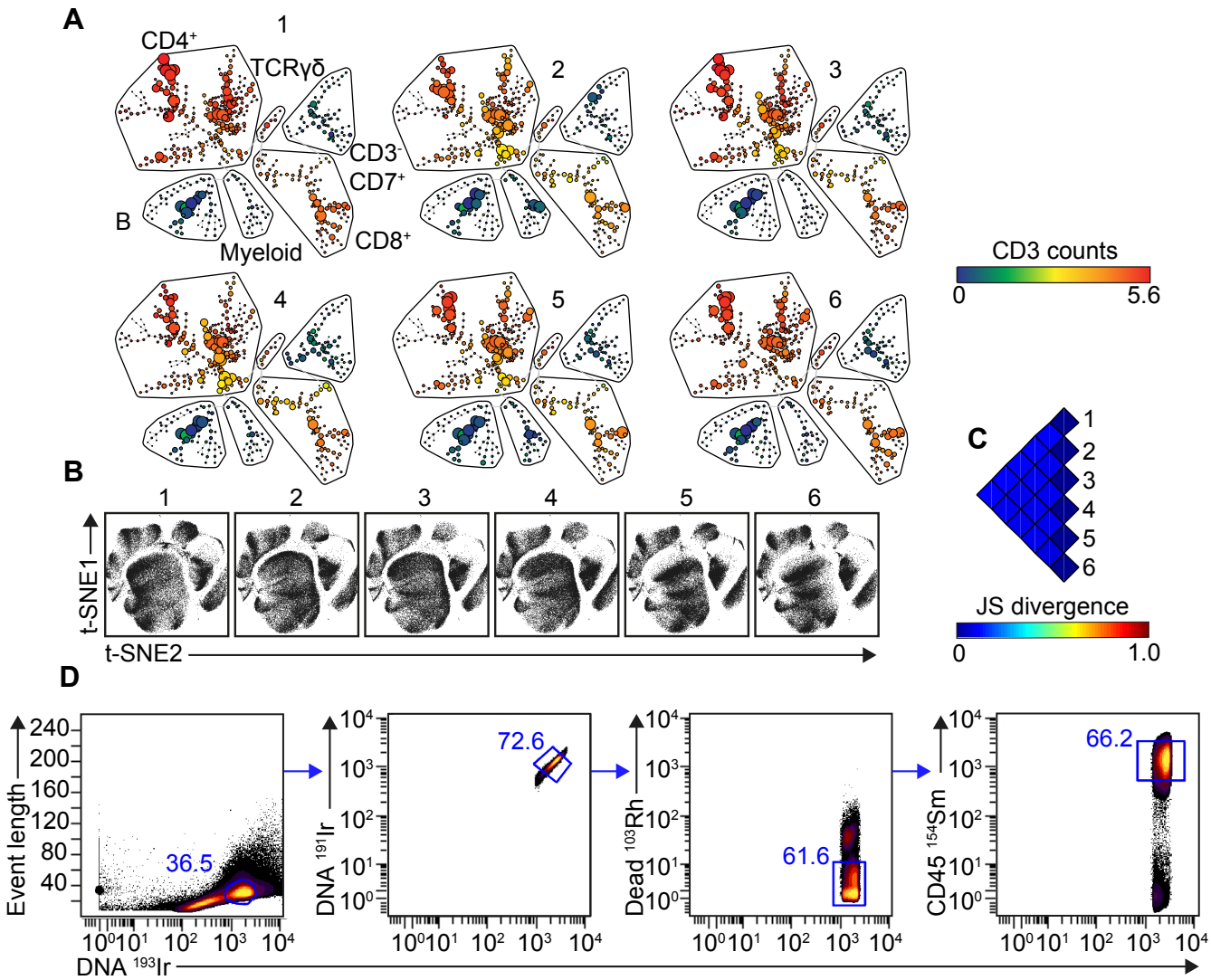

Figure S2. SPADE and t-SNE analysis of the internal peripheral blood control. (Related to Figure 1)

(A) Six PBMC samples from the same individual taken at different time points during the entire study period are shown after collective SPADE analysis of live single CD45$^+$ cells. Size of the nodes is proportional to the respective number of clustered cells. Color bars represent ArcSinh5-transformed values for CD3 marker expression. The identities of major immune lineages are annotated on the basis of lineage marker expression (CD4$^+$ = CD4$^+$ T cells, CD8$^+$ = CD8$^+$ T cells and B = B cells) (for details SPADE analysis see Figure 1) (B) Collective t-SNE dimensionality reduced single-cell data containing live single CD45$^+$ cells from 6 PBMC samples as described in panel a analyzed are plotted. Each sample was down-sampled to 50,000 events prior to the t-SNE analysis. (for details t-SNE analysis see Figures 2 and 4) (C) Pairwise Jensen-Shannon (JS) divergence plot of the collective t-SNE maps from the 6 PBMC samples analyzed. A lower JS divergence value indicates more similarity between a pair of t-SNE maps as shown in panel b. (D) Live, single CD45+ cells gating strategy. Representative mass cytometry plots of a duodenum biopsy showing sequential gates with percentages. Event length is a mass cytometric measurement for the amount of scans it took to acquire a given ion cloud.

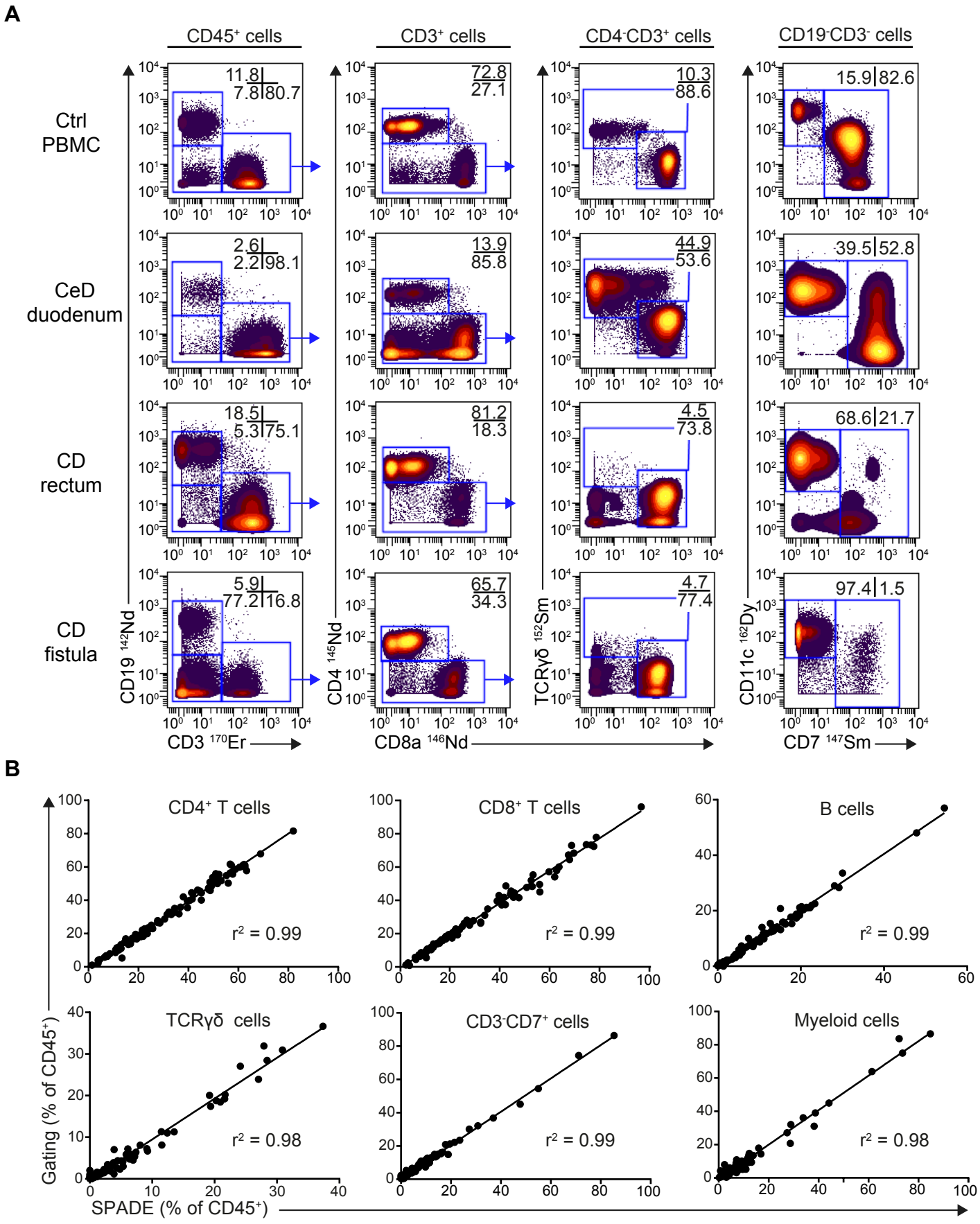

Figure S3.Traditional gating yields similar frequencies of major immune lineage cells as with SPADE analysis. (Related to Figure 1)
(A) Representative samples are shown to demonstrate the gating strategy applied to identify the major immune lineages: $CD4^+$ T cells, $CD8^+$ T cells, B cells, TCRγδ cells, $CD3^-CD7^+$ cells and myeloid cells. Cells gated as shown (annotated above plots) are represented as two-parameter density dot plots. (B) Correlation of cell frequencies obtained through traditional gating and with SPADE analysis are shown for the 102 samples as analyzed with linear regression. A dot represents a single sample.

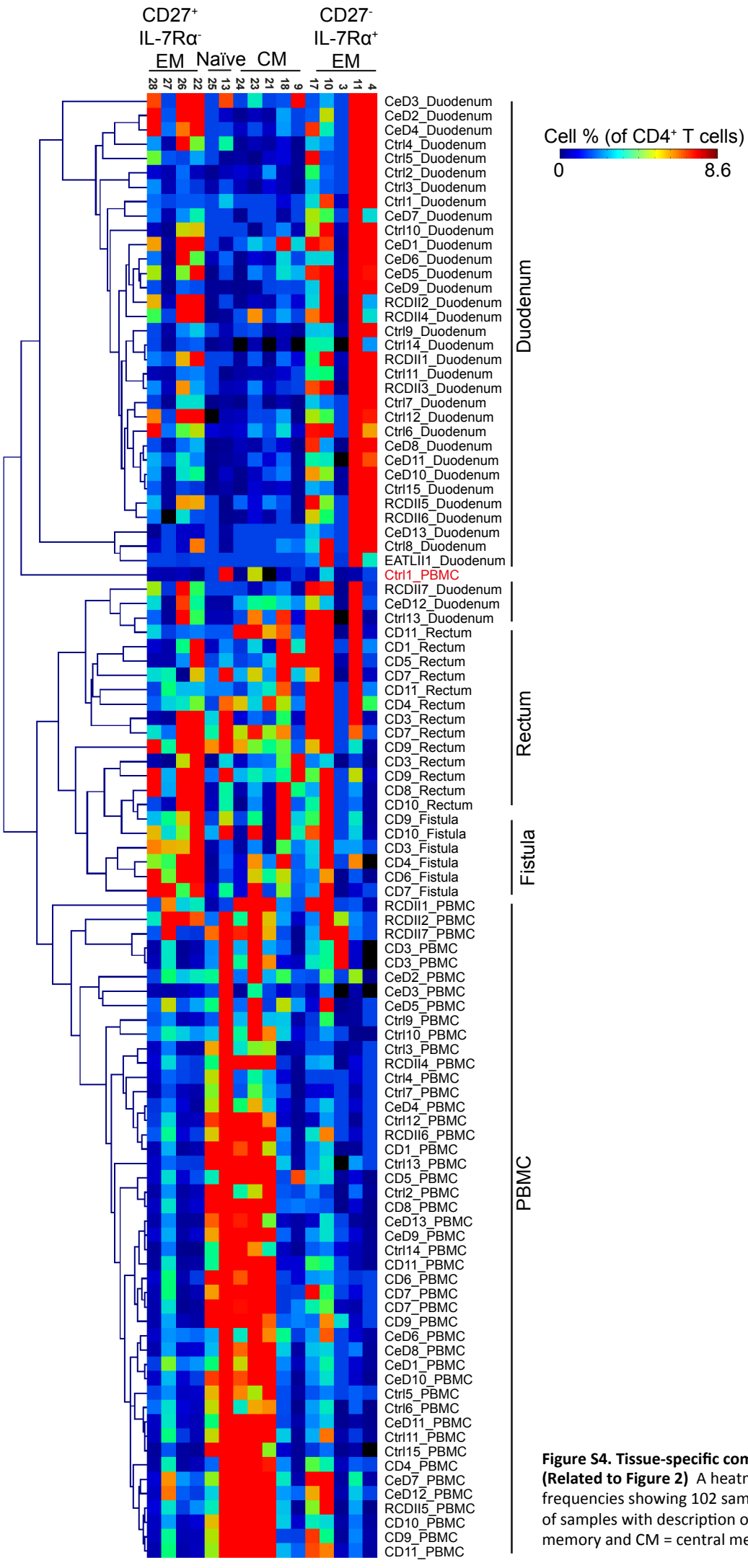

Figure S4. Tissue-specific composition of CD4$^+$ T cell subsets. (Related to Figure 2) A heatmap summary of subset frequencies showing 102 samples and hierarchical clustering of samples with description of tissue type. EM = effector memory and CM = central memory.

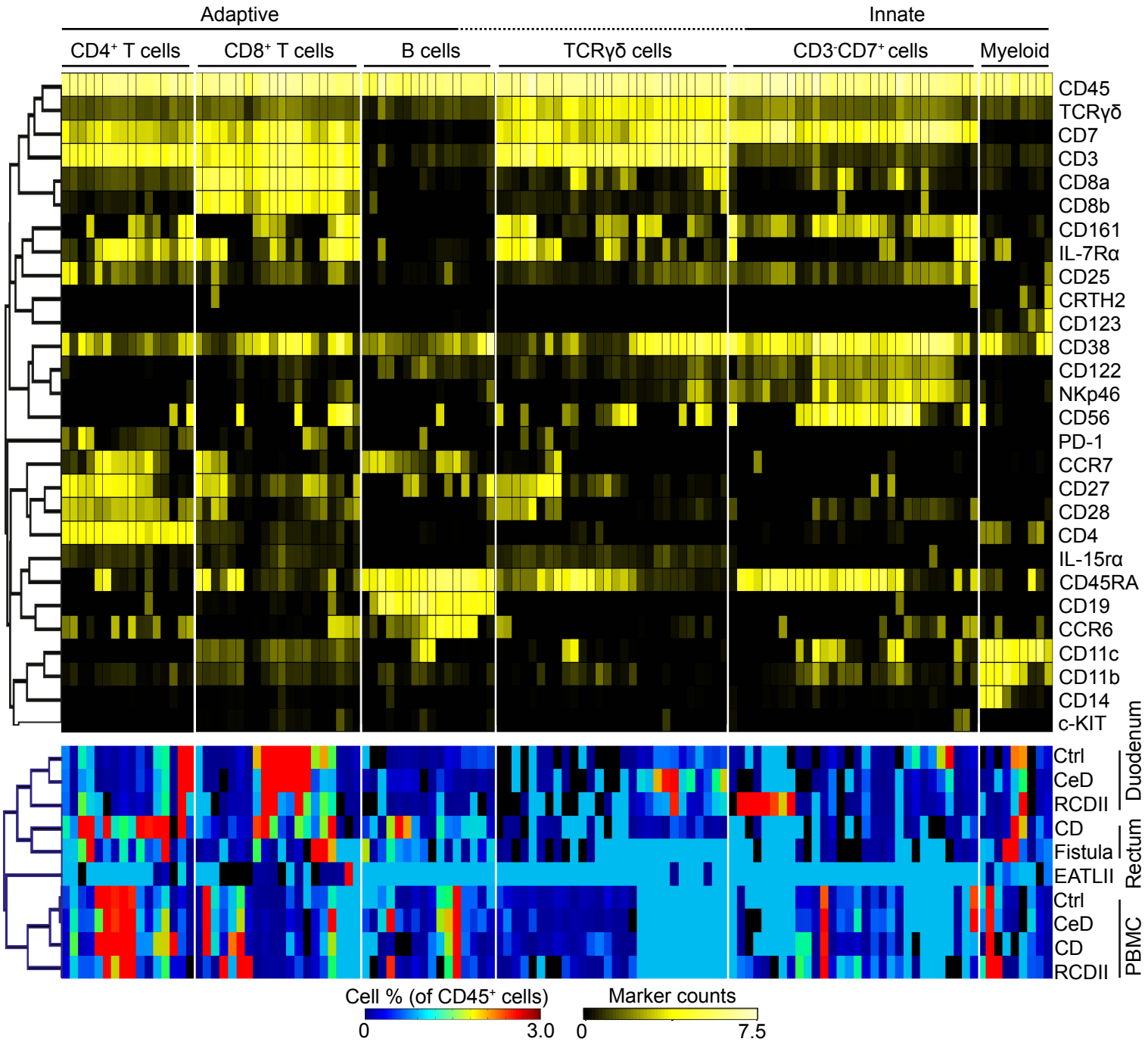

Figure S5. Marker expression profiles and composition of immune subsets in peripheral blood and intestine. (Related to Figure 3 and 5)
Heatmaps showing (A) characterization of cell populations (median ArcSinh5-transformed values of marker expression; black-to-yellow scale), (B) composition (average cell percentages; rainbow scale) and hierarchical clustering of markers and samples.

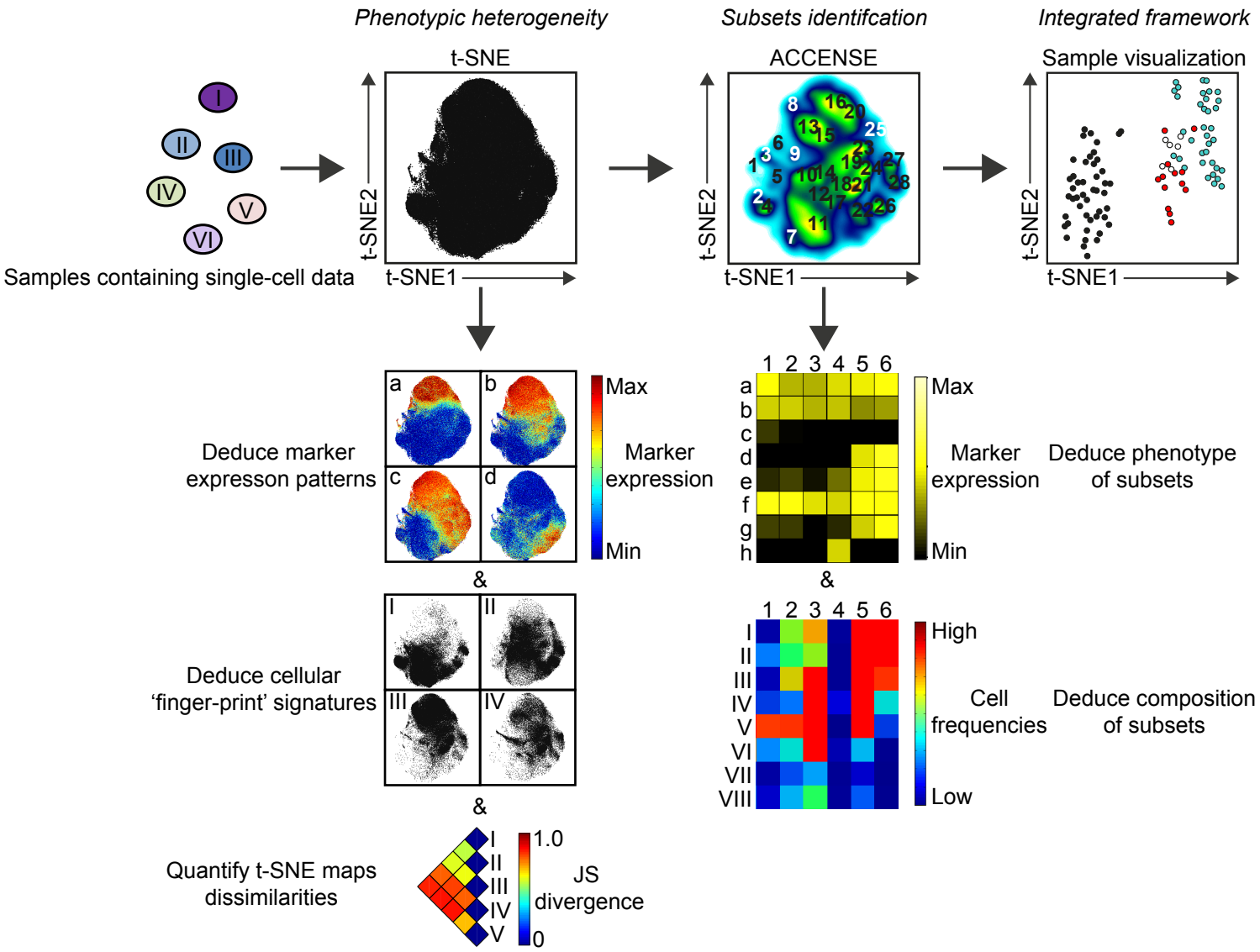

Figure S6. High-dimensional mass cytometry analysis pipeline. (Related to experimental procedures; data analysis)
Samples containing single-cell data of cell populations of interest (e.g. based on SPADE analysis or gated subsets) are tagged with integer sample identification values, marker expression values are transformed using hyperbolic arcsin with a cofactor of 5, and cells events are downsampled to match sample cell number. Collective t-SNE dimensionality reduction was performed on the single-cell data derived from all the samples. A dot represents a single cell and color overlays illustrate marker expression values to deduce marker expression patterns. t-SNE maps are stratified for tissues and disease states to illustrate cellular 'finger-print' signatures. Pairwise Jensen-Shannon (JS) divergence analysis quantifies dissimilarity between a pair of t-SNE maps. On the basis of the map created by t-SNE, a machine-learning cell cluster detection approach (ACCENSE) delineates cell subsets. Heatmaps illustrate the phenotypic signature of the identified immune subsets and illustrate the composition of subsets within the samples. The cell frequency data of all the identified immune subsets are used to visualize samples in an integrated framework using the t-SNE algorithm, where a dot represents a single sample and the color overlay illustrates clinical information. Markers are indicated in letters, samples in roman letters and cell subsets in numbers.

Table S1. CyTOF antibody panel. (Related to experimental procedures; antibodies)

Antigen	Tag	Clone	Com.	Cat#	Conc.
CD56	^{176}Yb	NCAM16.2	DVS	3176008B	1:100
CD4	^{145}Nd	RPA-T4	DVS	3145001B	1:100
CD8a	^{146}Nd	RPA-T8	DVS	3146001B	1:100
CD3	^{170}Er	UCHT1	DVS	3170001B	1:100
CD14	^{160}Gd	M5E2	DVS	3160001B	1:100
CD19	^{142}Nd	HIB19	DVS	3142001B	1:100
TCRgd	^{152}Sm	11F2	DVS	3152008B	1:100
CD45	^{154}Sm	HI30	DVS	3154001B	1:100
CD45RA	^{169}Tm	HI100	DVS	3169008B	1:100
CD27	^{167}Er	O323	DVS	3167002B	1:100
CD38	^{172}Yb	HIT2	DVS	3172007B	1:100
CD127	^{165}Ho	AO19D5	DVS	3165008B	1:100
CD11b	^{144}Nd	ICRF44	DVS	3144001B	1:100
CD7	^{147}Sm	CD7-6B7	DVS	3147006B	1:100
CD34	^{148}Nd	581	DVS	3148001B	1:100
C-Kit	^{143}Nd	104D2	DVS	3143001B	1:100
CD161	^{164}Dy	HP-3G10	DVS	3164009B	1:100
CD123	^{151}Eu	6H6	DVS	3151001B	1:100
CCR6	^{141}Pr	G034E3	DVS	3141003A	1:100
CD25	^{149}Sm	2A3	DVS	3149010B	1:100
CCR7	^{159}Tb	G043H7	DVS	3159003A	1:100
PD-1	^{175}Lu	EH 12.2H7	DVS	3175008B	1:100
CD11c	^{162}Dy	Bu15	DVS	3162005B	1:100
CD8b	^{166}Er	SIDI8BEE	eBio	14-5273	1:50
CRTH2	^{156}Gd	BM16	BioL	350102	1:50
TCRab	^{150}Nd	1P26	BioL	306702	1:50
IL-21R	^{153}Eu	2G1-K12	BioL	347802	1:50
IL-15Ra	^{168}Er	eBioJM7A4	eBio	14-7159-82	1:50
CD103	^{139}La	Ber-ACT8	BioL	350202	1:100
CD28	^{171}Yb	CD28.2	BioL	302902	1:100
NKp46	^{174}Yb	9E2	BioL	331902	1:100
CD122	^{158}Gd	TU27	BioL	339002	1:200

DVS Sciences (DVS), eBioscience (eBio), and Biolegend (BioL).

The antibody panel was designed to obtain a global overview of the heterogeneity of the immune system. For this purpose, the panel included markers that distinguish the major immune cell lineages, i.e. $CD4^+$ T cells, $CD8^+$ T cells, TCRγδ cells, B cells, innate lymphocytes ($CD3^-CD7^+$) and myeloid cells. In addition, markers were included that allow the identification of phenotypically distinct subsets within those immune cell lineages, like naïve and memory cells, resting and activated cells. In addition, markers were included that provide information on additional differentiation stages of cells, homing properties of cells, and their potential responsiveness to humoral factors like cytokines and chemokines. Our choice in markers was biased towards phenotyping CD7-expressing lymphoid cells (T cells and $CD3^-CD7^+$ cells) rather than B cells and myeloid cells.

Table S2. Characteristics of control and patients with inflammatory intestinal diseases. (Related to experimental procedures; human samples)

Characteristics	Ctrl (N=15)	CeD (N=13)	RCDII[a] (N=5)	EATLII (N=1)	CD[b] (N=10)
AgeBiop, (mean, ±SEM)	33.1±5.6	34.5±6.5	74.8±1.6	66	37.3±2.9
Gender, N (%)					
Male	5 (33.3)	6 (46.2)	5 (100.0)	0 (0.0)	3 (30.0)
Female	10 (66.6)	7 (53.8)	0 (0.0)	1(100.0)	7 (70.0)
GFD, N (%)					NA
No	13 (86.7)	5 (38.5)	0 (0.0)	0 (0.0)	-
Yes	2 (13.3)	8 (61.5)	5 (100.0)	1 (100.0)	-
Type of Biopsy					
Duodenum	15	13	7	1	-
Rectum	-	-	-	-	13
Fistula	-	-	-	-	6
PBMC[d]	14	13	6	-	14
Inflamed gut biopsy[c], N (%)					
No	14 (93.3)	6 (46.2)	2 (28.6)	0 (0.0)	12 (66.7)
Yes	1 (6.7)	7 (53.8)	5 (71.4)	1 (100.0)	6 (33.3)
Marsh score biopsy N (%)				NA	NA
M0	14 (93.3)	6 (46.2)	2 (28.6)	-	-
M1	0 (0.0)	1 (7.7)	0 (0.0)	-	-
M2	0 (0.0)	0 (0.0)	0 (0.0)	-	-
M3a	1 (6.7)	1 (7.7)	2 (28.6)	-	-
M3b	0 (0.0)	4 (30.8)	1 (14.3)	-	-
M3c	0 (0.0)	1 (7.7)	2 (28.6)	-	-

[a]Two RCDII patients were biopsied twice at different time points.
[b]Four CD patients were rectally biopsied twice at different time points.
[c]All 13 rectum biopsies of CD patients were non-inflamed, whereas all 6 fistula samples were inflamed.
[d]PBMC of CD patients were analyzed cryopreserved, whereas all other PBMC samples were analyzed fresh.
Celiac disease (CeD), refractory celiac disease type II (RCDII), enteropathy-associated T cell lymphoma type II (EATLII), Crohn's Disease (CD), age at biopsy time point (AgeBiop), gluten-free diet (GFD), and peripheral blood mononuclear cells (PBMC).