List of Supplementary Material

Supplementary Methods

Further details of biopsies

Endoscopic GI tissue biopsies were performed on patients with symptoms consistent with classical or late-onset acute GI-GvHD. Of the 50 patients who had GI biopsies included in the study, 31 patients had GI biopsies before D+180 whereas 19 patients had GI biopsies after D+180; of these only 5 had previously had biopsies at earlier time points. Biopsies performed after initiation of steroids or with evidence of CMV colitis were excluded. All individual biopsies were considered for assessment of cellular content and histological GvHD. Median values for RARα^{hi} cells for all available biopsies from each patient (at the earliest time point if biopsies were available at multiple time points) were used for association with clinical severity of GI-GVHD and GI-GvHD-related mortality.

Technical details of antibody staining for conventional IHC

For single staining, slides were incubated with primary antibodies for 40 minutes at room temperature following antigen retrieval, **Supplementary Table 1**. Sections were then stained using SuperSensitive™ Polymer-Horse Radish Peroxidase (HRP) IHC Kit (Bio-Genex, San Ramon, CA), diaminobenzidine (DAB) chromogen, and hematoxylin. For double-staining, CD4 antibody staining was followed by Polymer-HRP and SG chromogen, then CD8 staining, Polymer-HRP and AEC chromogen.

Patient blood samples, flow cytometry and cluster analysis

PBMC were isolated from peripheral blood from AHST patients at the time of presentation with symptoms of GI-GvHD or skin GvHD but prior to initiation of steroid therapy, from AHST

patients without GvHD at similar time-points, and from healthy controls. Flow cytometry was performed using directly-conjugated monoclonal antibodies, **Supplementary Table 5.** For combined surface and intracellular staining cells were labelled with surface markers then fixed and permeabilised (Fix perm kit, Biolegend). Dead cells were excluded using DAPI (Sigma) or for intracellular staining Zombie Yellow fixable viability stain (Biolegend). For enumeration of rare cell populations (<1%) a minimum of 50 positive events were acquired. Analysis of flow cytometric data was performed using FlowJo software version 10 (Treestar). tSNE plots were generated and cluster analysis was performed using Phenograph (https://bioconductor.org/packages/cytofkit/).[1] In some experiments CD4 T-cells were identified as $CD3^+$ $CD8^-$ cells.

References

1. Levine JH, Simonds EF, Bendall SC, et al. Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell.* 2015;162(1):184-197.

Supplementary Table 1 Clinico-pathogical details for AHST patients with GI-biopsies without histological evidence of GvHD

UPN	Biopsy site	Histology	Clinical diagnosis
42	Lower GI	Normal	Antibody-related diarrhea
43	Upper GI	Gastritis/duodenitis	Gastritis/duodenitis
44	Upper GI	Coeliac disease	Coeliac disease
45	Lower GI	Normal	Antibody-related diarrhea
46	Lower GI	Normal	Non-specific diarrhea
47	Lower GI	Spirochaetosis	Spirochaetosis
48	Upper GI	Normal	Hiatus hernia
49	Lower GI	Normal	Non-specific diarrhea
50	Upper GI	Gastritis/duodenitis	Gastritis/duodenitis

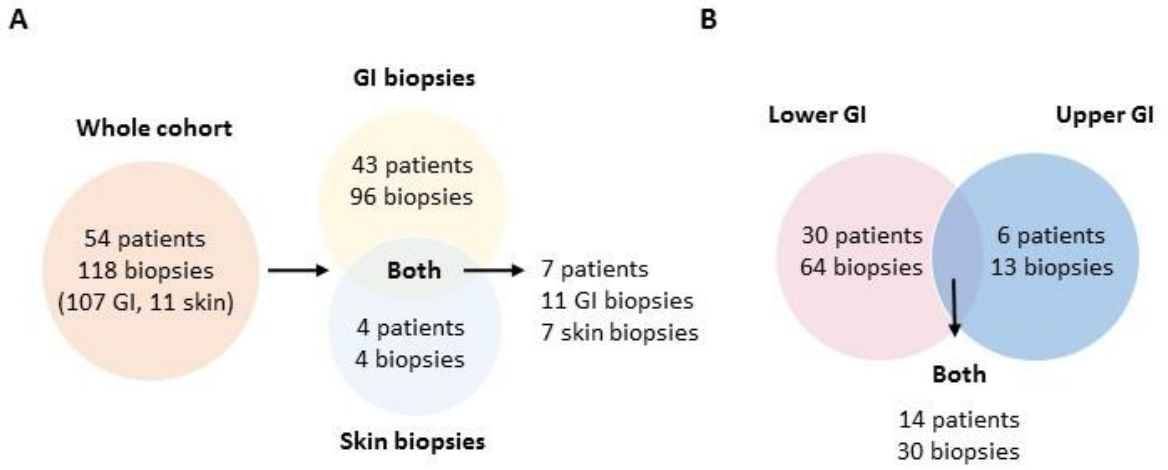
Supplementary Table 2 Antibodies for IHC

Antigen	Details	Supplier	Clone	Working dilution
CRABP1/II	Mouse monoclonal	Santa Cruz	F-9	1:3000
RARα	Rabbit polyclonal	Santa Cruz	N/A	1:4000
CD8	Mouse monoclonal	Dako	C8/144B	1:400
CD4	Mouse monoclonal	Nova Castra	4B12	1:500
T-bet	Mouse monoclonal	Santa Cruz	SC21003	1:500
ROR-γt	Mouse monoclonal	Merck Millipore	6F3.1	1:1000
IL-23p19	Rabbit polyclonal	Atlas Antibodies	N/A	1:500
IL-23R	Rabbit Polyclonal	Novus Biologicals	N/A	1:500
IL-33	Mouse monoclonal	Enzo	Nessy-1	1:500
IL-33R (ST2)	Rabbit Polyclonal	Atlas	N/A	1:500
FOXP3	Mouse monoclonal	Abcam	263A/E7	1:500

Supplementary Table 3 Antibodies for flow cytometry

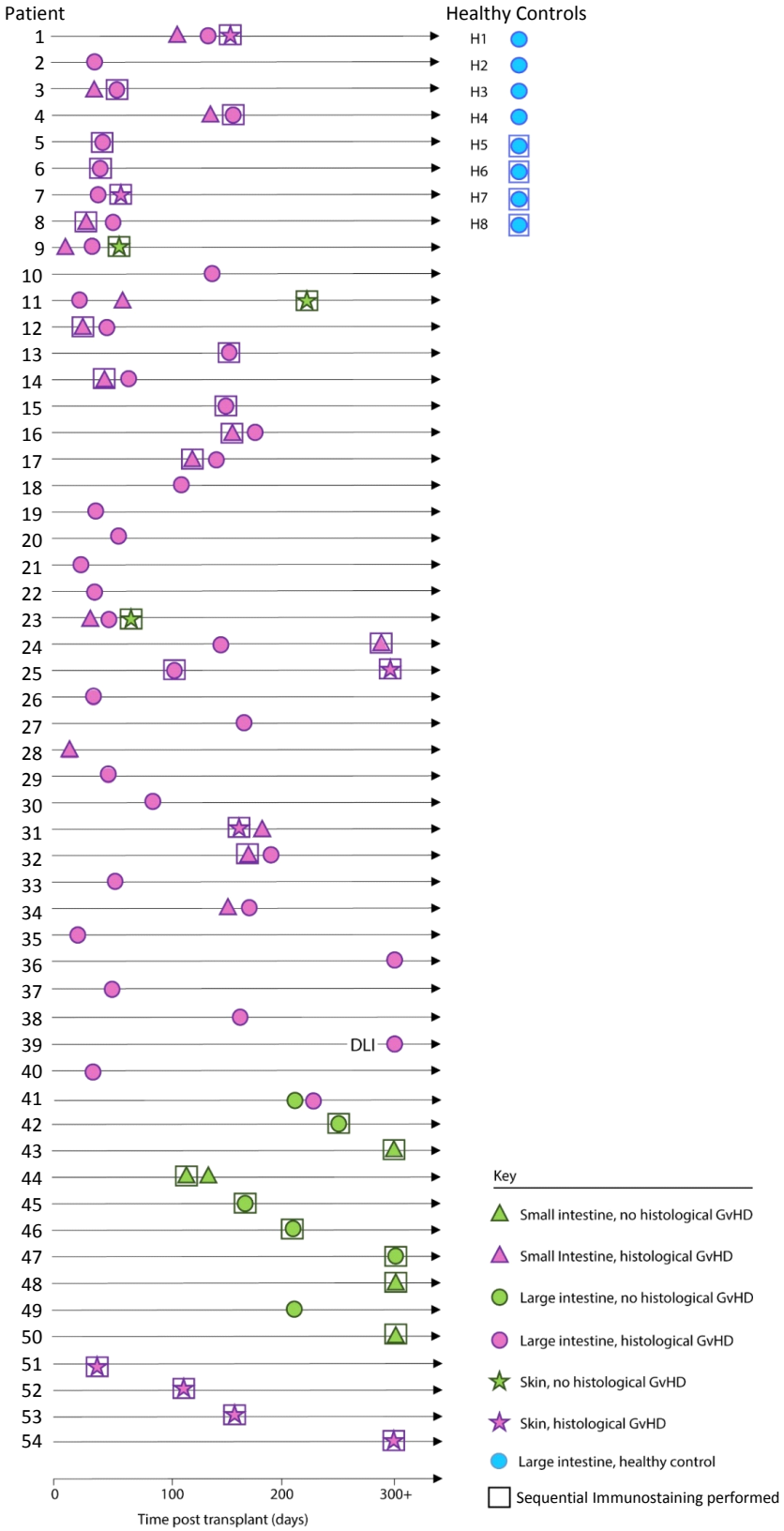
Antigen	Fluorochrome	Supplier	Clone
CD3	PerCP PE-Cy5	Biolegend Biolegend	OKT3 HIT3a
CD4	Alexa700	BD pharmingen	RPA-T4
CD8	BV605 PE-Cy7 Alexa700	Biolegend Biolegend Biolegend	RPA-T8 HIT8a SKI
RARα	FITC PE	USBiological Life Sciences	Rabbit polyclonal
α_4	APC	R&D systems	7.2R
β_7	PE APC	Biolegend	FIB504
CCR9	PE-Cy7 APC	Biolegend	L053E8
CD25	BV421	Biolegend	BC96
T-bet	PerCP 5.5	ebioscience	4B10
IL-23R	APC PE	R&D systems	218213
FOXP3	PE	Biolegend	206D

Supplementary Figure 1



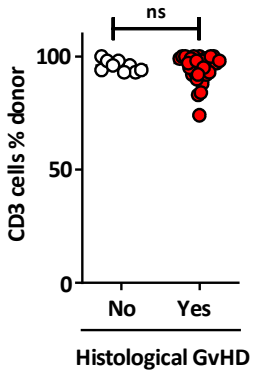
Supplementary Figure 1 Numbers of AHST patients and tissues biopsies included in the study, **A**, and numbers of AHST patients and tissue biopsies from lower and upper GI tract included in the study, **B**.

Supplementary Figure 2



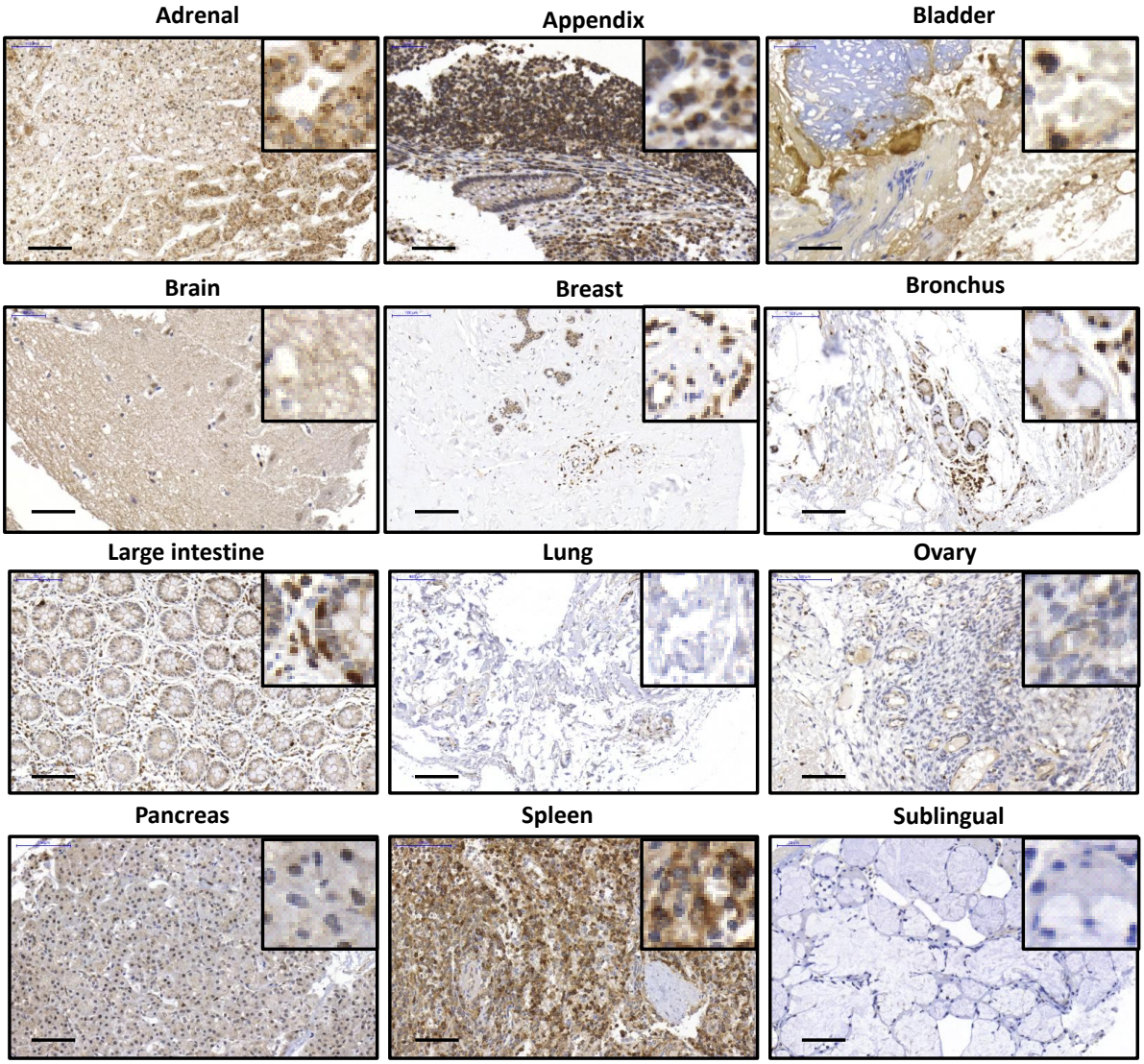
Supplementary Figure 2 Diagram of sites and times of biopsies included in the analysis. In some cases multiple biopsies from different sites within the upper and/or lower GI tract were studied at the same time point

Supplementary Figure 3



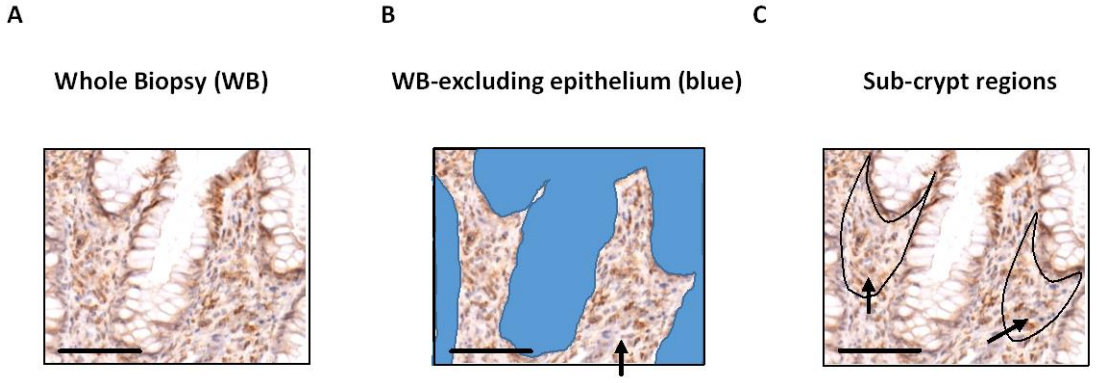
Supplementary Figure 3 Peripheral blood chimerism on purified CD3 T-cells in patients at the time of first GI-biopsy. Horizontal lines are medians. P value is for Mann Whitney test. ns, not significant, (P>0.10)

Supplementary Figure 4



Supplementary Figure 4 RARα staining patterns in normal tissues used as positive and negative controls to optimize RARα antigen retrieval and staining protocols and quality control for patient sample. Colour change reagent is DAB and positive staining is brown. Tissues are counterstained with hematoxylin. Scale bar is 100µm.

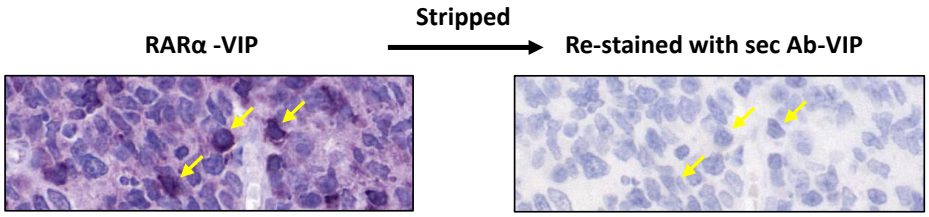
Supplementary Figure 5



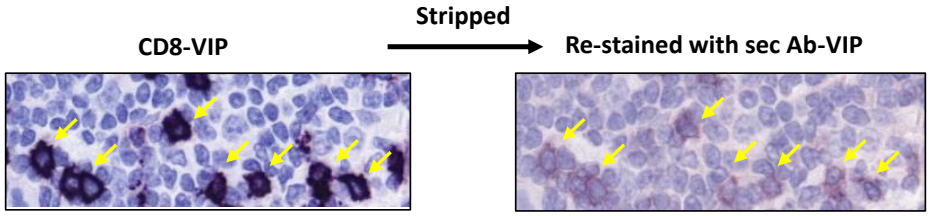
Supplementary Figure 5 Approach to regional analysis of GI biopsies for IHC enumeration of RARα^{hi} mononuclear cells. A, whole biopsy stained for RARα and DAB chromogen and counterstained with hematoxylin. Diffuse staining for RARα (brown) in epithelial cells is seen. **B**, Whole biopsy with epithelial cell areas (blue) excluded. Area of biopsy analyzed is arrowed. **C** Regional analysis restricted to sub-crypt regions (demarcated in black, arrowed). A minimum of 3 sub-crypt regions and total of 200 mononuclear cells were analyzed for each individual biopsy. Scale bar is 50μm.

Supplementary Figure 6

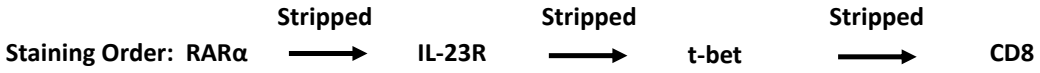
A



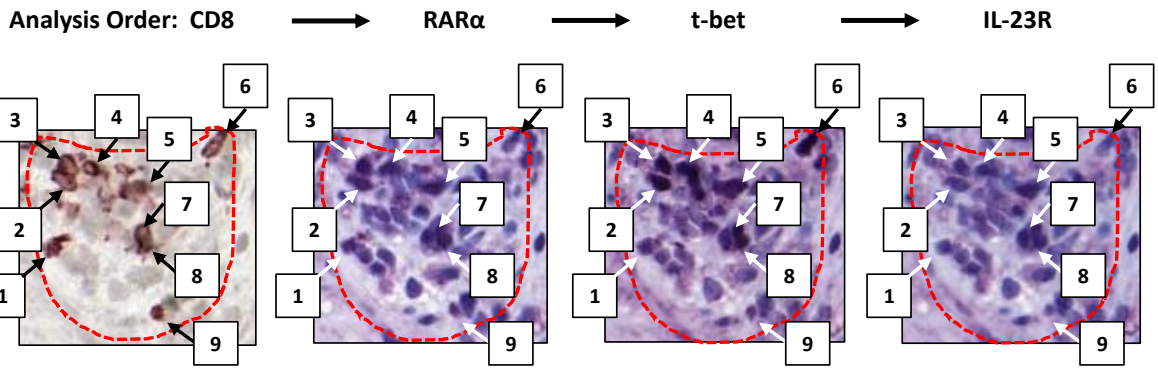
B



C



D



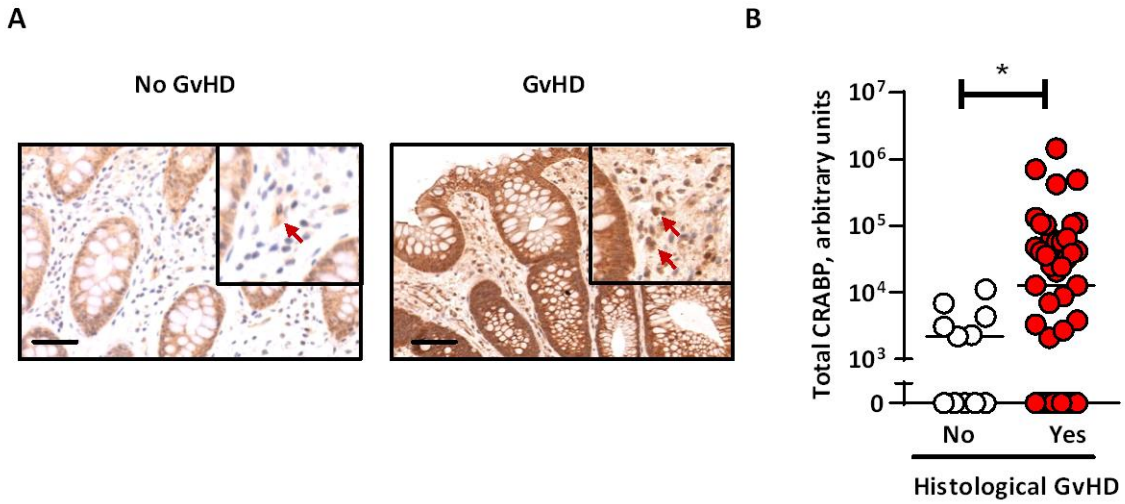
E

Cell identity	CD8	RARα	T-bet	IL-23R
1	+	hi	-	-
2	+	hi	+	-
3	+	hi	+	-
4	+	hi	+	-
5	+	hi	+	+
6	+	lo	+	-
7	+	hi	+	+
8	+	hi	+	+
9	+	lo	-	-

Supplementary Figure 6 Approach to confirm adequate stripping of primary antibodies and determination of co-expression patterns on mononuclear cells using sequential immunostaining

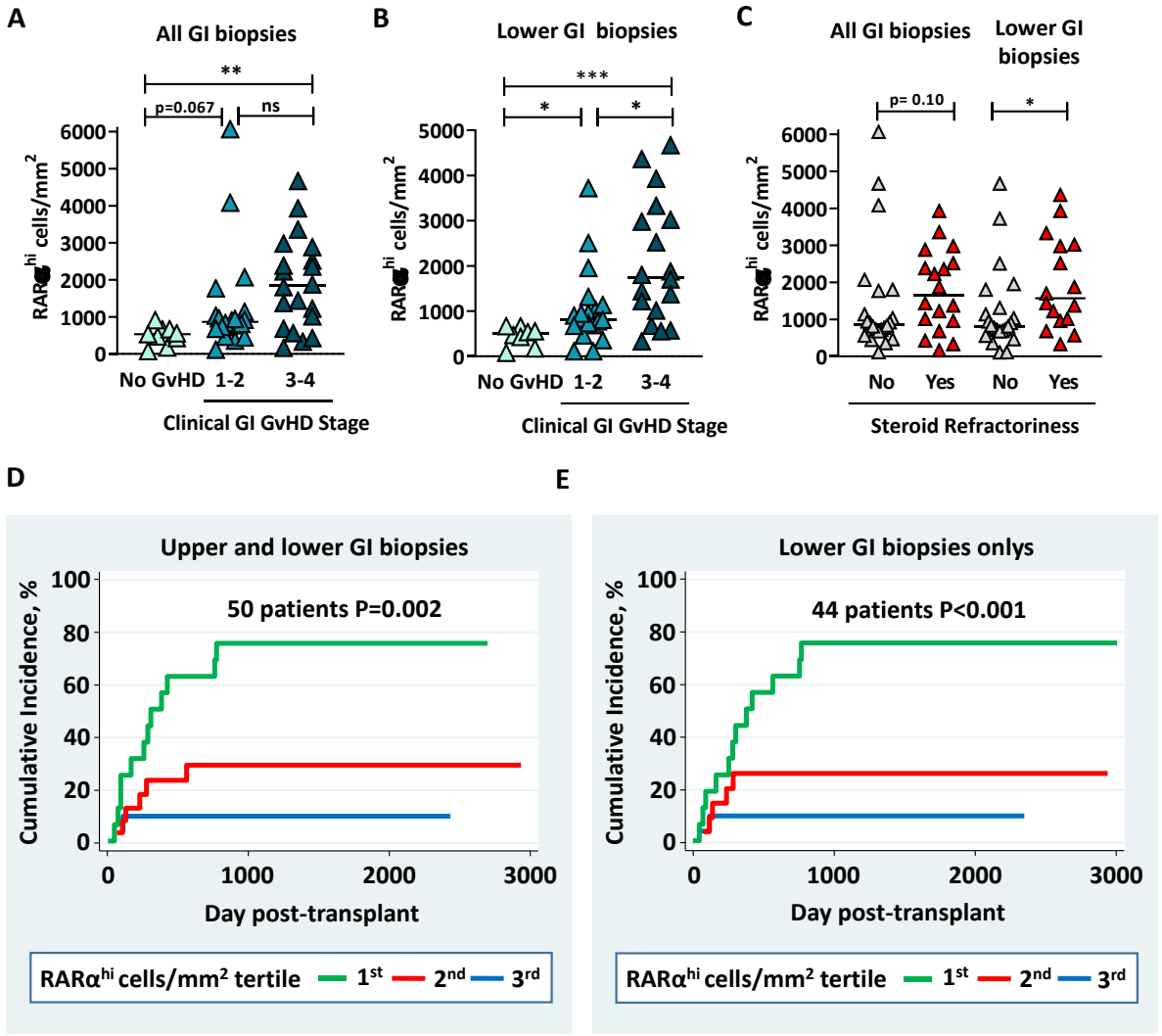
A RARα-VIP staining could be completely stripped. RARα-VIP staining (purple) of healthy lymph node tissue. Cells with high RARα staining are denoted with yellow arrows, left panel. Stripping of antibody was complete demonstrated by the absence of residual signal after re-staining with secondary antibody and VIP chromogen, right panel. Similar results were obtained with IL23-R-VIP and T-bet-VIP. **B** In contrast CD8 staining could not be completely stipped. CD8-VIP staining (purple) of healthy lymph node tissue. CD8 T-cells are denoted with yellow arrows, left panel. Stripping of antibodies was incomplete with DAB or VIP (shown) demonstrated by the presence of residual signal after re-staining with secondary antibody and VIP chromogen, right panel. In view of this CD8 staining was performed last in the sequential immunostaining process. **C** Staining order of sequential immunostaining pipeline. **D** Co-expression patterns of mononuclear cells were identified by electronic marking. An example is shown for $CD8^+$ T-cell analysis. All individual $CD8^+$ T-cells (identified with DAB chromogen, brown) in sub-crypt regions, demarcated with dashed red lines, were electronically marked in Pannoramic Viewer (3DHistotech, Budapest, Hungary) for a minimum of 3 sub-crypt regions/biopsy. Each individually marked cell was then mapped to images of the same sections sequentially immunostained for other markers 9with VIP chromogen, purple). Scale bar is 25µm. **E** $CD8^+$ T-cells were then scored for expression other markers (hi/lo for RARα, pos/neg for T-bet and IL-23R) to identify the expression pattern of each individual $CD8^+$ T-cell. The absolute number/unit area and proportion of total cells (minimum 200/biopsy) with each distinct expression pattern was collated.

Supplementary Figure 7



Supplementary Figure 7 A Representative GI-biopsies stained for CRABP1/!! using conventional IHC and DAB chromogen from AHST patients with and without histological GvHD. Scale bar is 50um, **B** Total CRABPI/II (calculated as the product of the number of CRABPI/II$^+$ mononuclear cells and the mean intensity of staining of cells for CRABPI/II) in sub-crypt regions of GI-biopsies from AHST patients with and without histological GvHD. Horizontal lines on graphs depict medians P value is for Mann Whitney test. * = $p<0.05$

Supplementary Figure 8

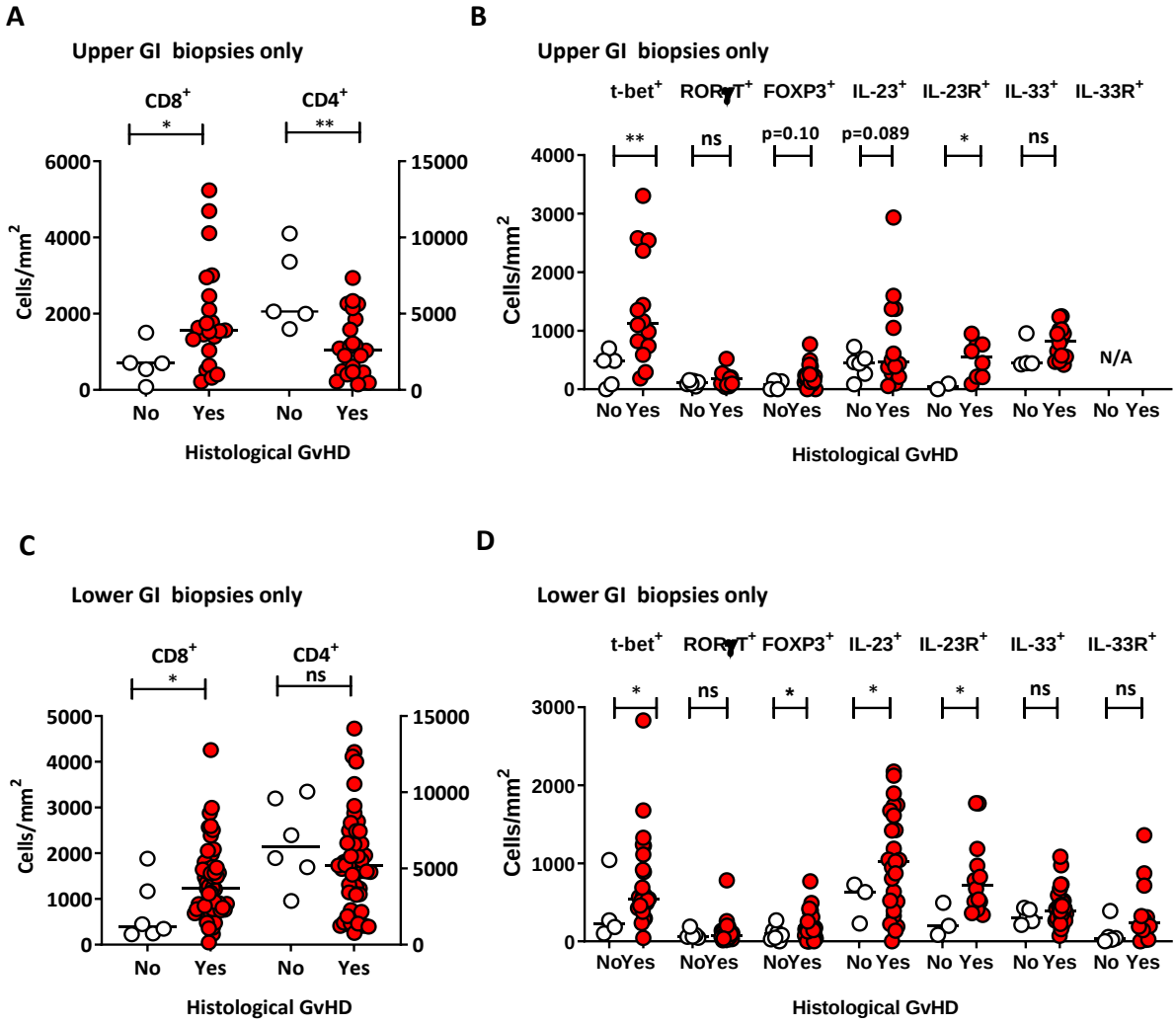


Supplementary Figure 8 RARα^{hi} mononuclear cell numbers are associated with GI-GvHD severity, steroid refractoriness and GvHD-related mortality

A RARα^{hi} mononuclear cell numbers in sub-crypt regions of upper and lower GI-biopsies in AHST patients grouped according to clinical GI-GvHD stage. **B** RARα^{hi} mononuclear cell numbers in sub-crypt regions of lower GI-biopsies in AHST patients grouped according to clinical GI-GvHD stage. **C** RARα^{hi} mononuclear cell numbers in sub-crypt regions of GI-biopsies in GI-GvHD patients with GI-GvHD grouped according to response to steroid therapy. **D** Cumulative incidence of death from GI-GvHD in AHST patients grouped into tertiles by RARα^{hi} mononuclear cell numbers in sub-crypt regions of upper and lower GI-biopsies. **E** Cumulative incidence of death from GI-GvHD in AHST patients grouped into tertiles by RARα^{hi}

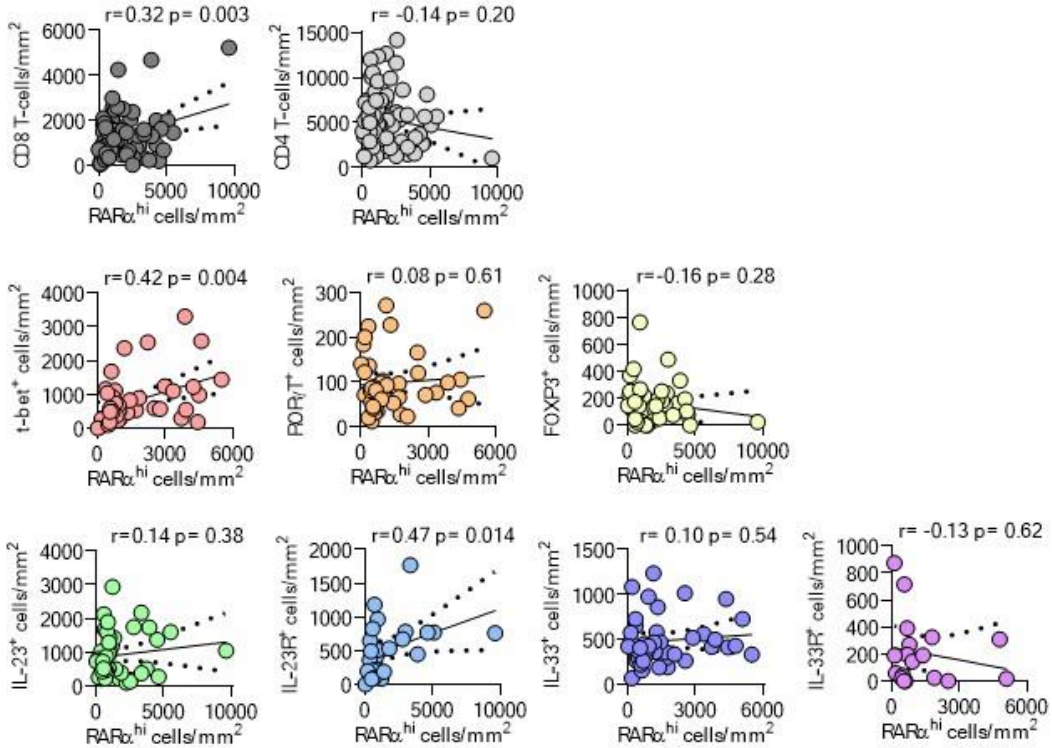
mononuclear cell numbers in sub-crypt regions of lower GI-biopsies. Horizontal lines are medians. P values are for ANOVA tests with correction for multiple comparisons (**A** and **B)**, Mann Whitney (**C**), and Grays test (**D**-**E**), **** $p<0.0001$ *** $p<0.001$ ** $p<0.01$, * $p<0.05$. ns, not significant, ($p>0.10$)

Supplementary Figure 9



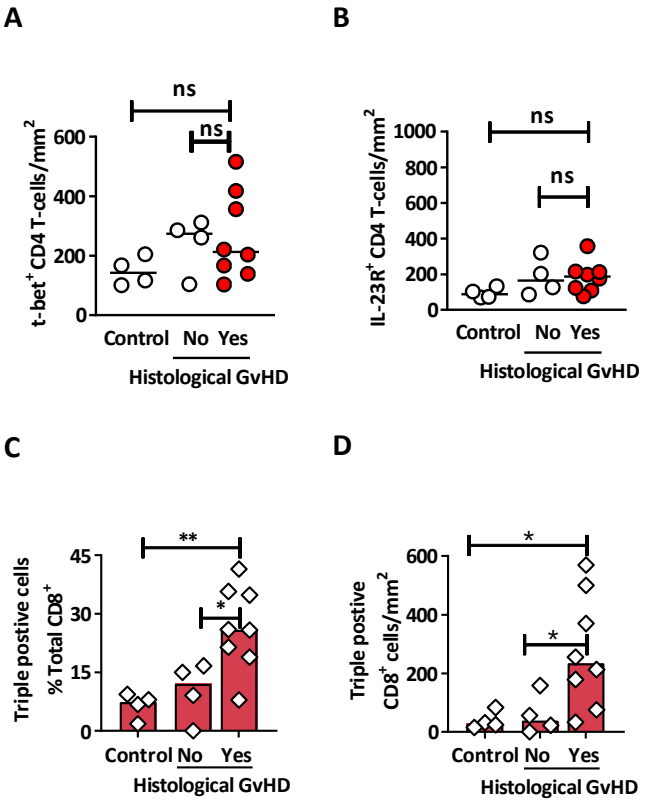
Supplementary Figure 9 CD8 and CD4 T-cell numbers in sub-crypt regions of upper (**A**) and lower (**C**) GI-biopsies from AHST patients with and without histological GvHD and mononuclear cell number expressing T-cell transcription factors, cytokines and cytokine receptors in sub-crypt regions of upper (**B**) and lower (**D**) GI-biopsies from AHST patients with and without histological GvHD. Horizontal lines are medians. P values are for Mann Whitney (**A-D**) tests ** p<0.01, * p<0.05., ns, not significant (p>0.10)

Supplementary Figure 10



Supplementary Figure 10 Correlations of RARα^{hi} mononuclear cell numbers and cells expressing T-cell lineage markers, transcription factors, cytokines and cytokine receptors in sub-crypt regions of upper and lower GI-biopsies from AHST patients. Solid and dotted lines denote linear regression and 95% confidence intervals respectively. r and p values are for Pearson correlations.

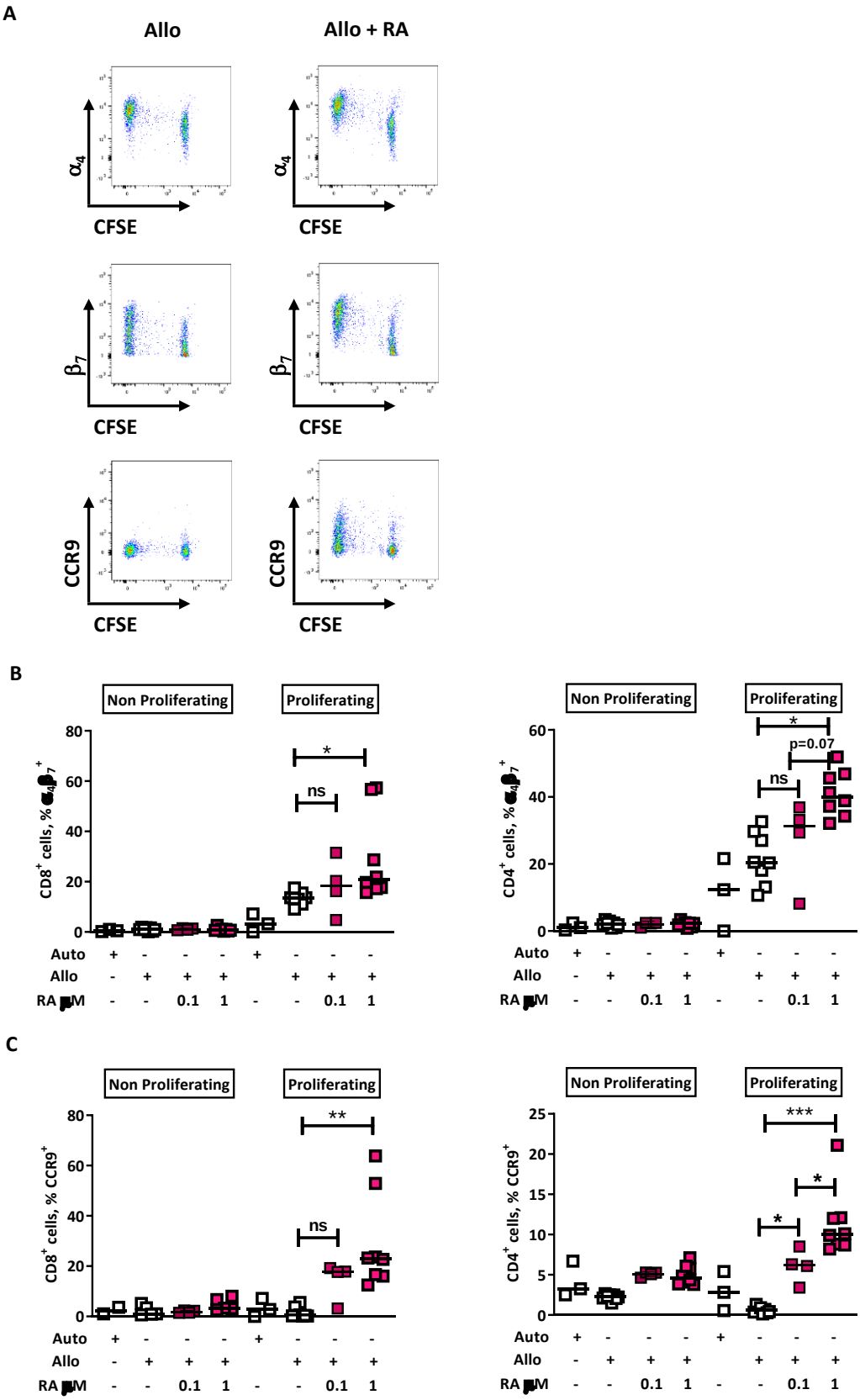
Supplementary Figure 11



Supplementary Figure 11 A

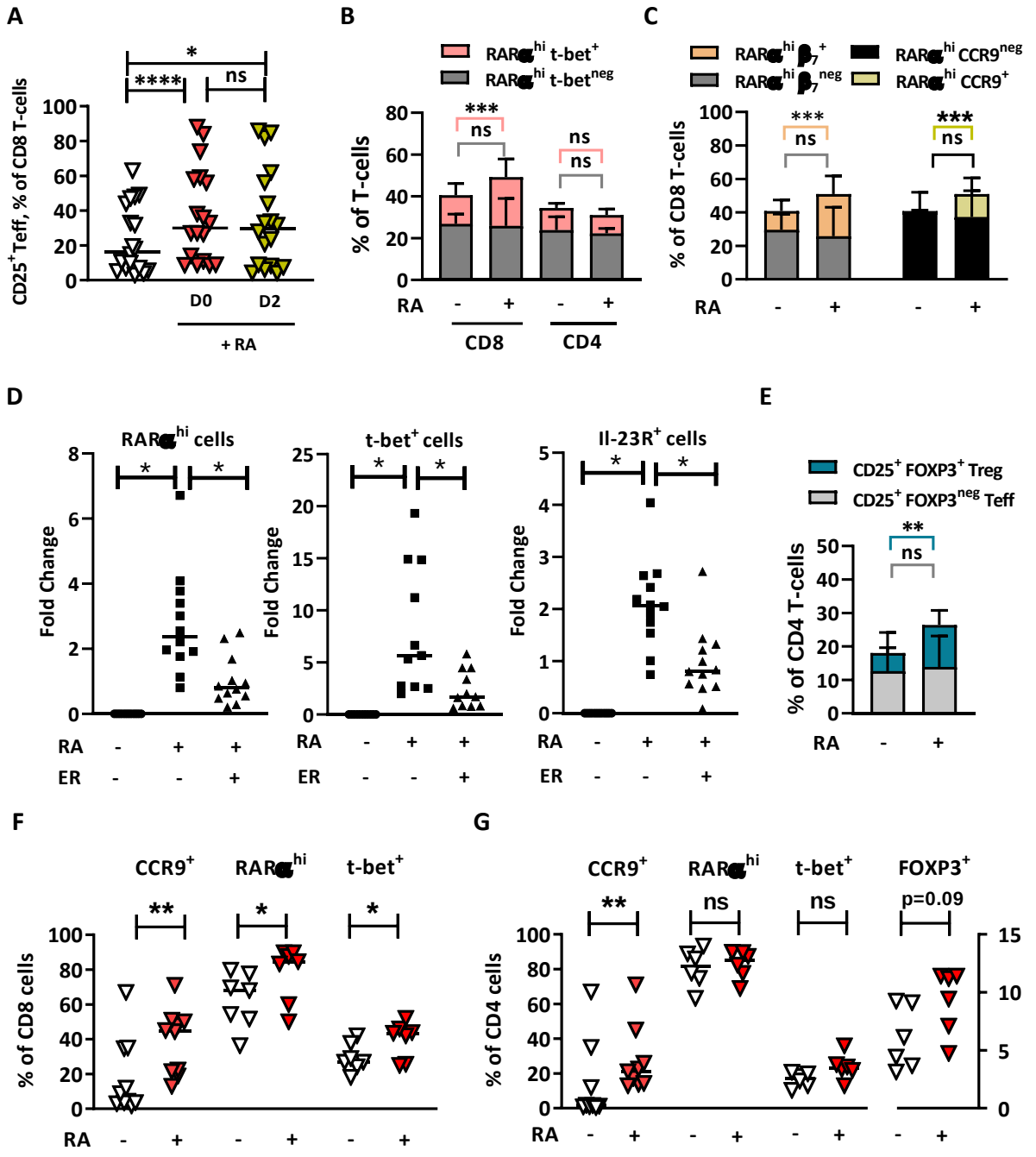
A-B CD4 T-cell numbers co-expressing T-bet (**A**) or IL-23R (**B**) in sub-crypt regions of upper and lower GI-biopsies from AHST patients with and without histological GvHD and lower GI-biopsies from healthy controls. **C** Sub-crypt triple positive cells (as percentage of CD8 T-cells) in lower GI-biopsies. **D** absolute numbers of sub-crypt triple positive CD8 T-cells in lower GI-biopsies. Bars are medians. P values are for ANOVA with post-test correction**.** ** $p<0.01$, * $p<0.05$, ns, not significant ($p>0.05$)

Supplementary Figure 12



Supplementary Figure 12 A Representative dot plots showing expression of gut-tropic integrins α_4 and β_7 and CCR9 on proliferative ($CFSE^{dim}$) and non-proliferative ($CFSE^{bright}$) T-cells after allostimulation with/without exogenous RA (1000µM). **B** Percentatge of CD8 and CD4 T-cells expressing integrins α_4 and β_7 after allostimulation in the absence or presence of RA (100-1000 µM). 4-8 independent experiments are shown. **C** Percentatge of CD8 and CD4 T-cells expressing iCCR9 after allostimulation in the absence or presence of RA (100-1000 µM). 4-8 independent experiments are shown. Auto, autologous control. Horizontal lines are medians. P values are for mixed effects models with post-test correction. *** $p<0.001$, ** $p<0.01$, * $p<0.05$. ns, not significant ($p>0.10$)

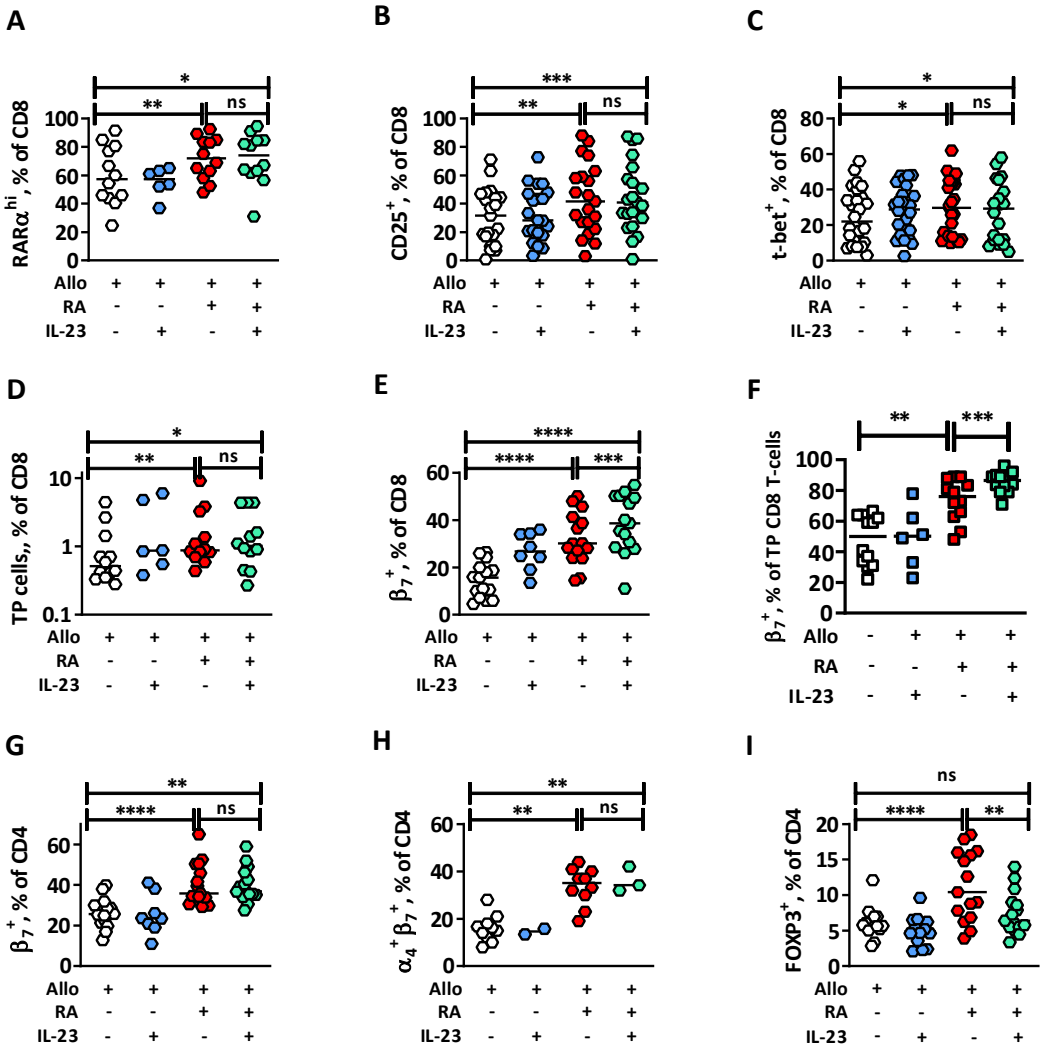
Supplementary Figure 13



Supplementary Figure 13 A Percentage of live alloproliferative CD8 T-cells expressing CD25 after allostimulation without or with exogenous RA (1μM) added either at D0 or D2. Results depict 18 independent experiments using healthy donor PBMC. **B** T-bet expression of live alloproliferative RARα^{hi} CD8 and CD4 T-cells after allostimulation in the absence or presence of exogenous RA (1μM). Results depict 15 independent experiments using healthy donor

PBMC. **C** β_7 and CCR9 expression of live alloproliferative RARα^{hi} CD8 T-cells after allostimulation in the absence or presence of exogenous RA (1µM). Results depict 15 independent experiments using healthy donor PBMC. **D** Fold-change in frequency of alloproliferative GI-tropic CD8 effector T-cells expressing high RARα, T-bet, or IL-23R after allostimulation with/without exogenous RA (1µM) and the RARα-specific inhibitor ER-50891 (ER). Results depict 12 independent experiments using healthy donor PBMC. **E** FOXP3 expression of live alloproliferative CD25^{+} CD4 T-cells after allostimulation in the absence or presence of exogenous RA (1µM). Results depict 24 independent experiments using healthy donor PBMC. **F-G** Proportion of CCR9^{+}, RARα^{hi}, T-bet^{+} and FOXP3^{+} alloproliferative CD8 (**F**) and CD4 (**G**) T-cells from peripheral blood of AHST patients after *ex vivo* allostimulation with/without RA (1µM). Results depict 7 independent experiments using 3 D+30 samples from AHST patients without GvHD. Horizontal lines and bars are medians. Error bars are interquartile ranges. P values are for ANOVA with post-test correction (**A and D**), Wilcoxon matched-pairs signed rank test (**B-C, E-G)**. **** p<0.0001 *** p<0.001 ** p<0.01, * p<0.05. ns , not significant, (p >0.10)

Supplementary Figure 14



Supplementary Figure 14 A-D proportion of alloproliferative CD8 T-cells expressing high RARα, CD25, T-bet and with a triple positive phenotype in healthy donor PBMC allostimulated with/without RA and/or IL-23**. E-F** proportion of alloproliferative CD8 T-cells (**E**) and TP CD8 T cells (**F**) expressing β_7 integrin, in healthy donor PBMC allostimulated with/without RA and/or IL-23**. G-H** Proportion of alloproliferative CD4 T-cells expressing β_7 integrin (**G**) or co-expressing α_4 and β_7 integrins (**H**) in healthy donor PBMC allostimulated with/without RA and/or IL-23**. I** Proportion of alloproliferative CD4 T-cells expressing FOXP3 in healthy donor PBMC allostimulated with/without RA and/or IL-23**.** Horizontal lines are medians. P values are for paired ANOVA or mixed effects models where groups contain uneven numbers of samples, with post-test correction. **** p<0.0001, *** p<0.001 **p<0.01 *p<0.05 ns p>0.10