Unanticipated interacting features of human gut-associated lymphoid tissues link microbiota, intestinal immunity and autoimmunity

Lucia Montorsi[1†], Michael J. Pitcher[1†], Yuan Zhao[1], Alicia Demonti[1,2], Thomas J. Tull[3], Chiara Dionisi[1], Pawan Dhami[4], Richard J. Ellis[5], Cynthia Bishop[5], Jeremy Sanderson[1,6], Sahil Jain[7], David D'Cruz[1,7], Deena Gibbons[1], Thomas H Winkler[8], Mats Bemark[9], Francesca D. Ciccarelli[10], Jo Spencer[1*]

[†]These authors contributed equally to this work

[1]School of Immunology and Microbial Sciences, King's College London; London, UK.

[2] École Normale Supérieure de Lyon, Claude Bernard Lyon 1 University; Lyon, France

[3]St. John's Institute of Dermatology, King's College London; London, UK

[4]Genomics Research Platform and Single Cell Laboratory at Guy's and St Thomas' NHS Foundation Trust; London, UK

[5]Flow Cytometry and Immune Monitoring Platform at Guy's and St Thomas' NHS Foundation Trust; London, UK

[6]Department of Gastroenterology, Guy's and St Thomas' Foundation Trust, London, UK

[7]Louise Coote Lupus Unit, Guy's and St Thomas' NHS Foundation Trust; London, UK

[8]Division of Genetics, Department of Biology, Friedrich-Alexander-University Erlangen-Nürnberg (FAU); Erlangen, Germany.

[9]Department of Microbiology and Immunology, University of Gothenburg; Gothenburg, Sweden

[10]Cancer Systems Biology Laboratory, The Francis Crick Institute; London, UK

*Corresponding author:

jo.spencer@kcl.ac.uk

https://orcid.org/0000-0002-7202-2431
Department of Immunobiology
Guy's Hospital, Borough Wing
St Thomas' St
London SE1 9RT, UK

Abstract

Gut-associated lymphoid tissue (GALT) is organised lymphoid tissue that is chronically activated by the intestinal microbiota. It generates the IgA response that is critical for intestinal homeostasis. By iterative application of multiplexed technologies, we identify enrichment of double-negative 2 (DN2:CD27$^-$IgD$^-$CD21^{lo}CD11c^{hi}) B cells in GALT, where they comprise the majority of intraepithelial and subepithelial B cells. We show that DN2 B cells in GALT interact with DC in the sub-epithelial dome that express *DNASE1L3* and microbicides. Unlike in mice, *DNASE1L3* in humans does not associate with apoptotic debris, but is located between sampled bacteria and host tissue where it is co-expressed with *C1Q*, consistent with management of bacterial debris. Thus we demonstrate that DN2 B cells that are otherwise associated with lupus nephritis, and DNASE1L3 and C1q that are lupus autoantigens, are microbiota-associated, interacting components of normal intestinal immunity.

Introduction

Gut-associated lymphoid tissue (GALT) is the organised lymphoid tissue on the boundary between the contents of the gut lumen and the host [1]. It is separated from the lumen by the follicle associated epithelium (FAE) that supports the sampling of particulate gut contents by microfold (M) cells and dendritic cells (DC) [1] [2]. As a consequence, GALT is chronically activated throughout life [3]. The region of lymphoid tissue below the FAE is the subepithelial dome (SED) where initial encounter between antigens and the immune system occurs [4] [5] [6].

The B cell response to sampled gut antigen, including class switching to IgA in GALT, can occur via T cell (CD40) dependent and independent mechanisms; the latter being supported by receptor ligand systems including APRIL that binds to its receptor TACI [7] [8]. The IgA generated has specificity for sampled bacterial antigens, including glycans, that may be shared between bacterial species [9].

Lymphocytes activated in GALT become imprinted; that is, they acquire expression of homing receptors such as a4b7 integrin and chemokine receptors CCR9 and CCR10 that are induced by locally generated retinoic acid. Having acquired the capacity to home, they enter the blood and disseminate widely [10]. The homing receptors allow extravasation into and retention within the intestinal lamina propria, thus seeding large areas of gut with effector cells [11]. The dimeric IgA secreted throughout the gut is transported into the lumen where it interacts with the microbiota maintaining diversity and homeostatic balance [12] [9]. In addition to generating IgA plasma cells, the chronically activated state of GALT supports the diversification and propagation of marginal zone B cells that also circulate in blood and become concentrated as a population in the splenic marginal zone [13] [14].

Despite the importance of human GALT for maintenance of mucosal and systemic health, very little is known of the tissue microarchitecture and cellular interactions that support its activity. Analysis of B cells subsets in GALT have previously identified naïve B cells (CD27$^-$CD45RB$^-$IgM$^+$IgD$^+$) surrounding germinal centres (GC; CD38++CD10++). These are in turn are surrounded by populations of marginal zone B cells (MZB; CD27$^+$CD45RB$^+$IgM$^+$IgD$^+$) alongside their precursors (MZP; CD27$^-$CD45RB$^+$IgM$^+$IgD$^+$) and memory cells (CD27$^+$CD45RB$^+$IgD$^-$) [15]. It has been assumed from early work using cellular morphology and enzyme immunohistochemistry that the B cells in the SED, including those that infiltrate the FAE, are MZB though analysis by imaging mass cytometry (IMC) with a small panel of antibodies failed to confirm this [16] [15]. The subepithelial region is known to be rich in DC

expressing lysozyme (LysoDC) and A Proliferation Inducing Ligand (APRIL) [17,2,18]. APRIL supports T cell independent B cell function including class switching to IgA independent of TGFb or interaction between CD40 and CD40L[19,7,8]. However, there is little additional information on the cellular composition of the SED in humans.

Here, we applied multiplexed technologies iteratively to characterise the microanatomy of human GALT. We initially analysed isolated cells from GALT compared to matched blood using mass cytometry and observed that the most enriched B cell subset in GALT is the double negative 2 (DN2; CD27$^-$IgD$^-$CD21^{lo}CD11c^{hi}) population. This subset, that is associated with interferon driven autoimmunity in lupus, Sjögren's disease and infection [20,21], was found to be located in the SED and FAE using imaging mass cytometry. By spatial transcriptomics resolved to single cells we identify that the SED is enriched in expression of the lupus associated autoantigen *DNASE1L3* [22]. Whereas in mice DNASE1L3 is associated with clearance of apoptotic debris [23,22], we show here that its distribution in humans is inconsistent with this, but rather indicates a role for DNASE1L3 alongside another lupus autoantigen, C1q [24], in handling bacterial debris on the mucosal front line, adjacent to DN2 B cells. We suggest that these normal interacting features of GALT could be functionally caricatured or targeted aberrantly in lupus nephritis.

Results

Double negative (CD27⁻IgD⁻) cells are enriched in human GALT

Fresh paired cell suspensions of blood and GALT from healthy individuals,[15] were analysed using mass cytometry after labelling with metal-tagged antibodies (Supplementary Table 1). CD19+ cells were selected for viSNE dimentionality reduction that was run based on the expression of B cell lineage markers CD10, CD20, CD21, CD24, CD27, CD38, CD45RB, CD138, IgD, IgM, IgG and IgA (Fig. 1a). Cell similarities were visualised using 'spanning-tree progression analysis of density-normalized events' (SPADE) based on the viSNE coordinates. Nodes were manually grouped into SPADE 'bubbles' according to phenotypes of major B cell subsets (Fig. 1b, c) that were then visualised on the viSNE, and the relative frequencies of cells were represented in pie charts (Fig. 1d, e). In addition to GC cells and plasmablasts (PB)/plasma cells (PC) that were expected to be enriched in GALT compared to blood, we also saw enrichment of double negative B cells, in particular DN2/3 cells, which were differentiated from DN1 by their relatively lower expression of CD21(Fig.1 b, c, f) [25].

We analysed the expression of markers that were not used in the dimension reduction or clustering algorithms in a heatmap of mean marker expression (Fig. 1g). Notably DN2/3 cells, together with memory B cells, MZB and MZP as already described, were enriched in expression of FcRL4 in GALT suggesting proximity to epithelium [26] [27] [15]. They also expressed lower CD40 and HLA-DR than other B cell subsets suggesting reduced capacity for T cell interaction.

Thus DN2/3 cells were enriched in GALT where their cell surface marker expression suggested epithelial proximity and T cell independence.

Distribution of B cell subsets in human GALT

To investigate the microanatomic locations of B cell subsets, we analysed paraffin sections of human GALT from ileum, colon and appendix by imaging mass cytometry, using a panel of metal tagged antibodies (IMC Panel 1 in Supplementary Table 2). Following segmentation, quality control, and identification of major cell subtypes in GALT by machine learning (Supplementary Fig. 1a-d), B cells were identified by their expression of CD20, including those that infiltrate between epithelial cells in the FAE (Fig. 2a and Supplementary Fig. 1d). B cell boundaries were visualised on the images (Fig. 2b). Dimensionality reduction and clustering was performed on the B cells using the B lineage associated markers in the panel and clusters were merged and classified according to marker expression (Fig. 2c, d). The B cell subsets were then visualised on the image (Fig. 2e). As expected, GC B cells

5

were surrounded by naïve B cells, MZP, MZB and memory B cells in concentric circles [15]. Two clusters of DN cells were observed; one in the epithelium that carried the E-Cadherin marker from adjacent epithelial cells termed intraepithelial DN (IEDN) cells and DN2 cells in the SED that were high in CD11c (Fig. 2d - f). Members of all B cell clusters were also observed in the GC (Fig. 2e).

Because of the close proximity of DN cells to the CD11c expressing dendritic cells (DC) in the SED, we confirmed that B cells in the SED expressed CD11c using an RNAscope probe against *ITGAX* (gene encoding CD11c), visualised by confocal microscopy (Fig. 2g). By this method we were also able to confirm that a subset of IEDN cells were also DN2.

The relative distributions of B cell subsets described above including the dominance of DN2 cells in the SED and FAE was observed non only in GALT from colon, but also in ileum and appendix (Fig. 2e and Supplementary Fig. 1e, f).

Thus DN cells, in particular DN2, are found on the mucosal front line in the SED and FAE of GALT. DN2 cells are known to be enriched in blood in lupus and in blood and tissues in Sjogren's disease [20].

Interrogation of the sub-epithelial dome by spatial transcriptomics
Having identified enrichment of DN2 cells in GALT and their localization in the SED, we were interested to uncover features of that microenvironment that may influence their function. To that end we used 10X Genomics Visium spatial transcriptomics to identify features of the SED in GALT from human colon (n=5 GALT sites in 2 colon samples). In these, we first confirmed that the distribution of genes *MS4A1* (CD20), *IGHD* and *IGHM*, known to be associated with lymphoid tissue, matched their expected patterns of expression (Fig. 3a).

We used three analysis methods to characterise genes selectively expressed in the SED to better understand the microenvironment in which DN2 cells reside. First, we manually selected the spots that corresponded to the lymphoid tissue by reference to the haematoxylin and eosin (H&E) staining, and within this, separated those belonging to the SED and the follicle (Fig. 3a, b and Supplementary Fig. 2a, b). Due to the resolution of this technique, spots overlaying the FAE were included in the SED group for the analysis. We then used the Seurat package to find markers that differed between the SED and the rest of the lymphoid tissue and expressed it as a heathmap (Fig. 3c) [28]. Among the most differentially expressed genes were *CCL20* and *CCL23* that are associated with the FAE in mouse models, and *ITGAX* that is expressed by DC in the SED, which confirmed that our

approach was successful [29]. Providing further validation, other lymphoid markers included *BCL6* and *TCL1A* identifying the GC and naïve B cells were instead preferentially expressed in the follicle area [30] [31]. Interestingly, the most differentially expressed gene between the SED and the rest of the follicle was *DNASE1L3*, with the complement *C1Q, C3* and *C1R* also enriched in the SED (Fig. 3c, d).

We then undertook an undirected approach to analysis. Spots across the whole section were clustered based on their transcriptomic profiles and location, and for each image the cluster that coincided with lymphoid tissue by reference to the H&E staining was selected (Fig. 3e,f; and Supplementary Fig. 2d, e). The lymphoid tissue clusters from different images were then subclustered to generate 5 regional clusters within the lymphoid tissue (Fig. 3g). These were classified according to expression of lineage related genes (Fig. 3g, h; and Supplementary Fig. 2g). We again observed enrichment of *DNASE1L3* expression in the subcluster that equated to the SED/epithelium, and to a lesser extent within a subcluster corresponding to the T cell zone (Fig. 3h). Complement components *C1Q, C3* and *C1R* had a similar distribution to *DNASE1L3*, with enrichment within SED and the T cell zone (Fig. 3h).

DNASE1L3 is a recently described lupus autoantigen[22] [23]. In addition, loss of function mutations in *DNASE1L3* result in the generation of anti-DNA antibodies. Like *DNASE1L3*, loss of function mutations in *C1Q* can predispose to the production of anti-DNA antibodies and C1q can also be a target of autoantibodies in lupus [32] [33] [34]. In contrast, C3 is associated differently with lupus: plasma C3 levels inversely correlate with disease state in SLE and antibodies are occationally seen against one of its breakdown products [35]. These features were of particular interest to us having observed the lupus associated DN2 B cell subset in this microenvironment. We therefore asked whether expression of selected SLE associated genes *C1QB, C1QA, C1R, C1S, C3, DNASE1L3, ACP2, AGER, CTBB, DNASE1, FAS* and *TMEM173* [36] [37] [38] [39] [40] generally spatially overlapped across the tissue irrespective of zonation (Fig. 3I and Supplementary Fig 2h, i). Expression of TLR7 and TLR9 was too low to be included. Whilst there was strong colocalization of *DNASE1L3* and *C1Q*, and to some extent *C3* and *C1R*, there was no colocalization between *DNASE1L3* and the other lupus related genes, including genes associated with apoptosis (Fig. 3I).

Spatial transcriptomics therefore identifies enrichment of genes associated in different ways with lupus nephritis in the SED of GALT. Loss of function mutations in *DNASE1L3* [32] or *C1Q* [34] result in the production of antibodies to DNA. They are also lupus associated autoantigens [22] [24]. In contrast C3 and C1r are consumed in active lupus, resulting in lower serum levels [35].

DNASE1L3 expression by DC in the subepithelial region of GALT

To characterise the cells producing DNASE1L3 in SED, we validated the expression of *DNASE1L3* in single cells using RNAscope combined with imaging mass cytometry (IMC) in regions of interest (ROI) from colon (n=7), ileum (n=4) and appendix (n=3) using IMC Panel 2 (Supplementary Table 2). Some antigenic targets, including CD11c, were damaged by the RNAscope tissue processing, thus the distribution of CD11c expression was visualised using the corresponding RNAscope probe (*ITGAX*). Consistent with the spatial transcriptomic analysis, we observed abundant *DNASE1L3* expression in the SED (Fig. 4a-c; Suppelementary Fig. 3a-c). Furthermore, the technique allowed us to demonstrate that *DNASE1L3* was expressed by *ITGAX*+CD68+ DC cells in the SED (Fig. 4a-c). Expression by DC with proximity to epithelium was confirmed computationally across all samples (Fig. 4d).

Surprisingly, *DNASE1L3* expression was absent in the GC from GALT (Fig. 4a-c) [41] [23]. This was confirmed computationally across all ROI with GC (n=10) (Fig. 4e and Supplementary Fig. 3a-c). Lack of expression of *DNASE1L3* in GC was also confirmed in 2 ROIs from each of 2 mesenteric lymph node (n=4) and 1 ROI from each of 2 tonsils (n=2) analysed (Supplementary Fig. 4a, b). Since the GC is a known site of high apoptotic activity, these observations are in contrast with the accepted function of DNASE1L3, which is to degrade DNA from apoptotic cells [42].

Occasionally, *DNASE1L3* expression was observed in the T cell zone in sparse *ITGAX*+CD68+ cells consistent with other work [43], though at a lower level of expression than in the SED (Fig. 4f).

Due to the location of *DNASE1L3*+ cells in the SED, which is know to be a site of antigen sampling, and their absence from GC we hypothesized that *DNASE1L3* might be associated with anti-microbial activity [17]. Consistent with this hypothesis, *DNASE1L3* expressing cells in this region also expressed anti-microbial enzyme lysozyme and the superoxide generating enzyme NOX2 that is also associated with microbial killing (Fig. 4g, h). *DNASE1L3*+ cells in the SED had a greater tendency to express lysozyme or NOX2 than the less frequent *DNASE1L3*+ cells located elsewhere in the tissues (Fig. 4g-i).

Having observed DN2 B cells, that are otherwise associated with lupus nephritis, in the SED and DC expressing lupus autoantigen in *DNASE1L3* the same microenviroment, we asked whether these cell types might interact. By RNAscope coupled with confocal microscopy we observed interactions between *ITGAX*+CD11c+CD20+ B cells, likely DN2, and CD11c+*ITGAX*+ dendritic cells in the SED (Fig. 4j).

Thus DC that express *DNASE1L3* interact with DN2 B cells in the SED, while expression of the enzyme could not be detected in the GC.

DNASE1L3 expression is not associated with apoptotic debris in humans, though it is in mice

The enrichment of DNASE1L3 in the SED and its absence from the GC was surprising, because the most commonly described role for DNASE1L3 is the removal of apoptotic debris, including those generated by B cell selection in the GC response [42] [41] [44] [23]. To ask whether the SED is an area enriched in apoptotic debris in humans, we incorporated the TUNEL assay into IMC Panel 1 (Supplementary Table 2). To overcome the incompatibility of the RNAscope protocol and the TUNEL protocol we applied IMC Panel 1 to sections serial to those that identified *DNASE1L3* distribution by RNAscope with IMC Panel 2 (Supplementary Table 2). Data from ROI containing GALT from colon (n=6), ileum (n=5) and appendix (n=5) were acquired. We observed large cells in the GC with abundant cytoplasmic accumulation of TUNEL[+] material that equated to GC tingible body macrophages, which are known to uptake apoptotic debris (Fig. 5, a-c and Supplementary Fig. 5a, b). TUNEL[+] macrophages were CD68[+] and either CD11c[+] (predominantly in the light zone) or CD163[+] (predominantly in the dark zone). The epithelium of the gut is known to be constantly renewed, but this happens by extrusion of epithelial cells into the gut lumen [45]. Consistent with this, TUNEL signal and efferocytosis were not features of the SED of GALT, where DNASE1L3 is expressed (Fig. 5, a-c; Supplementary Fig. 5a,b).

Much of the evidence for the role of DNASE1L3 in digestion of apoptotic debris *in vivo* includes murine models [23] [22] [44]. We therefore analysed the distribution of *DNASE1L3* in mouse GALT represented by Peyer's patches (n=4) using RNAscope and confocal microscopy. We observed a marked difference in the expression of *DNASE1L3* in mouse and human GALT. Whilst *DNASE1L3* transcripts were not detected in human GALT GC (Fig. 4), it was expressed in both the SED and in the GC in mice where it colocalized with lysozyme (Fig. 5d) (n=4 Peyer's Patches). This suggests that *DNASE1L3* expressing cells in the SED in mice were predominantly LysoDC [17]. The GC cells that expressed *DNASE1L3* and lysozyme in mice also contained pycnotic nuclear fragments (Fig. 5d). Through the use of the TUNEL assay we show that the lysozyme[+] cells in the GC that contained pycnotic nuclear fragments also displayed strong TUNEL signal in the cytoplasm (Fig. 5e), confirming that these were tingible body macrophages performing efferocytosis. Although the majority of efferocytic cells were concentrated in the GC of mice, LysoDC with cytoplasmic TUNEL

9

staining were occasionally observed in the SED (Fig. 5e) which contrasted with what we observed in humans (Fig. 5a-c).

Thus we identified major differences in the distributions of DNASE1L3 in humans and mice. In humans the distribution of DNASE1L3 is inconsistent with a major function in digestion of apoptotic debris.

Expression of DNASE1L3 by lamina propria subepithelial macrophages

Because of the observed expression of *DNASE1L3* in GALT proximal to the epithelial boundary, and its known inducibility [46], we asked whether lamina propria macrophages might also express *DNASE1L3* (n=4). We observed abundant lamina propria macrophages that were predominantly *ITGAX*CD163$^+$CD68$^+$ (Fig. 6a,b). Macrophages with this phenotype expressed *DNASE1L3* when in subepithelial locations closer to the intestinal lumen but only rarely when adjacent to the bases of the crypts (Fig. 6a,b)

The difference in phenotype between cells expressing *DNASE1L3* in lamina propria and SED suggests that *DNASE1L3* expression is associated with subepithelial location rather than myeloid cell phenotype or lineage. This is consistent with its previously observed inducibility[46].

Relationship between expression of DNASE1L3 and complement components

Expression of *C1Q*, like *DNASE1L3*, was found to be enriched in the SED by spatial transcriptomics (Fig. 3c,d) and they appeared spatially associated (Fig. 3h, i). We therefore asked whether *DNASE1L3* and *C1Q* were co-expressed by the same cells in GALT. Using RNAscope coupled with confocal microscopy, we observed co-expression of these two transcripts in the SED in appendix (n=2) and colon (n=2) (Fig. 7a and Supplementary Fig. 5c). Some GC cells were weakly positive for *C1Q* but all were negative for *DNASE1L3* (Fig. 7a).

Complement components *C3* and *C1R* were also identified by spatial transcriptomics to be expressed in the lymphoid tissues in the SED and the periphery of the follicle. RNAscope analysis revealed that *C3* and *C1R* were co-expressed but in different cells to those expressing *C1Q* and consequently *DNASE1L3* (Fig. 7b, Supplementary Fig. 5d) in appendix (n=2) and colon (n=2). *C3* and *C1R* were also found to be co-expressed with vimentin that is expressed by stromal cells (Fig. 7c) in appendix (n=2) and colon (n=2).

DNASE1L3 expression relative to the distribution of translocated bacteria

10

The FAE of GALT and associated DC are known to sample particulate antigens, including bacteria, from the gut lumen, and DNASE1L3 enrichment could be related to sampling or digestion of microbial DNA [17] [2] [6]. Due to the high immunogenicity of bacterial DNA and shared epitopes with mammalian DNA, rapid degradation of bacterial DNA would be important in this location to avoid pathogenic anti-double stranded DNA antibody formation, potentially involving the adjacent DN2 B cells.

To assess the presence of bacteria in the SED, we used an RNAscope probe against a region of the *16S* transcript which is conserved across all Eubacteria in ileum (n=2), appendix (n=2) and colon (n=2). Bacterial RNA was detected in the FAE and the SED most proximal to the epithelium (Fig. 8a, b), in the proximity to *DNASE1L3*[+] cells, while it was not detected in the rest of the follicle below.

To investigate this further we incorporated the same RNAscope probe into IMC panel 3 that also included major lineage markers (Supplementary Table 2). Bacterial RNA in proximity with the FAE appeared to mostly be localized in the extracellular space. However, they were sometimes colocalized with E-cadherin, which is suggestive of bacterial transport into the SED (Fig. 8c, d), an activity typically ascribed to M cells[6]. Bacterial RNA were also often colocalized with CD68 and *DNASE1L3,* while they were not observed to colocalise with other cell types (Fig. 8f-g). This confirms that bacteria can cross human FAE and can be taken-up by *DNASE1L3*[+] DC in the SED or remain extracellular, though likely temporarily.

Thus we observe translocation of bacteria into GALT, either within the FAE, the subepithelial extracellular space or inside *DNASE1L3* expressing DC.

Discussion

Here we identify that DN2 B cells are abundant at the boundary between the microbiota and the host in the SED and FAE regions of GALT. These sites are enriched in APRIL that supports the T cell independent B-cell response and class switching towards IgA [8] [18]. Consistent with T cell independence, GALT DN cells have relatively low expression of CD40 and class II MHC that are associated with cognate T cell help [8] [7]. DN2 are enriched in the blood in lupus nephritis where they can be precursors of autoreactive plasma cells driven by IFNg [47] [20]. This pathway could be active in GALT that is known to be rich in IFNg producing cells [48] and to generate IgA via T cell independent pathways [8] [8] [49].

We show that DN2 B cells are intermixed with DC that co-express *DNASE1L3* and *C1Q.* Because of their enrichment in the SED in a region between sampled microbiota and the host we propose that DNASE1L3 and C1q are involved in management of bacteria sampled from the gut lumen. In particular, potential for high concentration of secreted DNASE1L3 in the SED is consistent with it having a role in digestion of bacterial DNA as observed previously in other contexts [50]. Bacterial DNA could be potentially pathogenic and could drive aberrant activity of DN2 B cells that are situated together with the *DNASE1L3-*expressing DC [20].

Whilst a potential role of DNASE1L3 in digesting bacterial DNA in the gut is relatively straight forward to envisage, the potential role of C1q is less clear [24] [51]. C1q can mediate bacterial interaction with host cells and is involved in the initiation of the classical pathway of complement activation that involves complement fixing antibodies [24]. Although IgG tends not to be secreted on mucosal surfaces in health, the other major complement fixing antibody class, IgM, is expressed [52]. Clones of IgM only memory cells in GALT and IgM secreting plasma cells in adjacent lamina propria are well described and an accepted feature of the intestinal immune system. Mucosal IgM specifically could be a mediator of bacterial killing in subepithelial regions following bacterial sampling by M cells.

Expression of *C1Q* has been identified as a marker of an embryonically derived subpopulation of macrophages in human lymphoid tissue that also express low levels of CD36 and *DNASE1L3* [43]. Our observation that *DNASE1L3* can be expressed by different subsets of macrophages and DC with contrasting phenotypes suggests that it may not contribute to lineage definition; rather this distribution supports inducibility [46]. It has been shown previously that DNASE1L3 can be induced by IL-4 and retinoic acid, and it is also already known that retinoic acid can be produced in the SED where it is involved in

12

imprinting the expression of gut-homing receptors such as integrin a4b7 [46] [10]. It is thus possible that DNASE1L3 is also induced as part of a similar mucosal imprinting process.

DNASE1L3 and C1q (or more specifically their absence or inactivation) are notably associated with SLE and anti-DNA antibody formation [32] [53] [54] [22]. The most cited antigenic driver of anti-DNA responses is DNA derived from apoptotic cells that is flipped to the external surface of dying cells [42]. In humans this is supported by the distribution of DNA on dying cells *in vitro* and there is also evidence that apoptotic debris generated by GC responses can be aberrantly distributed in the splenic GC of lupus patients [42] [55]. However, here we show that DNASE1L3 is not expressed in the GC in human GALT, tonsil or lymph node but is located in the SED. DNASE1L3 is found both intra- and extracellularly, can digest bacterial DNA and even contribute to the disruption of biofilm [50]. Mouse models have contributed in many ways to the current view of the relationship between apoptosis and lupus pathogenesis, yet much data may not be translationally relevant because of the marked species difference in the distribution of DNASE1L3 that we identify here [56]. Thus, we propose that the role of apoptotic debris as drivers of anti-DNA responses in SLE may not be the full story, and that DNASE1L3 in the SED may be critical to prevent anti-DNA antibody production by DN2 cells after recognition of bacterial DNA.

TLR7 and TLR9 are both receptors for pathogen genetic material that are likely to be expressed in gut-associated lymphoid tissues, though expression levels were too low to detect their transcripts reliably in our data. They are also both associated with aberrant reponses to their ligands in lupus, though in different ways [57] [58] [59]. Whereas over expression of TLR7 results in more severe disease, TLR9 can be protective [59]. TLR9 recognition of CpG has shown to be involved in class switching to IgA [49]. This system that is driven by microbial ligands could be impacted by DNASE1L3 deficiency in lupus.

Coeliac disease and lupus nephritis are both autoimmune diseases where autoantigens (gliadin and DNA respectively) are substrates for enzymes (tissue transglutaminase 2 ([TG2]) and DNASE1L3 respectively), and where both the substrate and the enzyme are targets for antibodies [60]. Here we show that like coeliac disease, the enzyme and substrate in lupus nephritis are also associated with the gut epithelial boundary in GALT. In coeliac disease TG2 deamidates gliadin peptides derived from dietary gluten [60]. The resulting modified peptide complements the binding groove of HLA-DQ8 or HLA-DQ2 haplotypes from where it can be presented to gliadin specific T cells [61]. B cells specific for gliadin can acquire help from T cells activated in this way resulting in the production of anti-gliadin antibodies [62].

TG2 can also be taken up in complexes with gliadin and can itself then become target of an autoimmune response so that antibodies are made to both enzyme and substrate [62] [63]. It is not so clear how tolerance is broken in the relationship between DNA and DNASE1L3. It was recently suggested that the main pathway for exposing the gut immune system to both gluten and TG2 would be through uptake from the gut lumen[64]. Future comparisons between lupus and coeliac disease could be helpful for the identification of potential parallel mechanisms, and potential T cell involvement in the breakage of tolerance.

In conclusion, we describe a high prevalence of DN2 B cells that are known to be capable of T cell independent responses at the boundary between the microbiota and the host in GALT [20]. We also observe by spatial transcriptomics resolved to single cells that the most enriched transcript in FAE and SED encodes the lupus associated autoantigen DNASE1L3, which is expressed together with C1q. DNASE1L3 has a known role in degrading pathogen derived DNA [50]. In line with the uptake of bacterial antigens into the SED, we propose that DNASE1L3 and C1q have key functions managing bacterial antigens on the epithelial boundary. Their activity could prevent exposure of DN2 B cells to excessive microbial DNA which could otherwise lead to anti-DNA responses.

14

Methods

Human tissues

All tissues used in this study were collected with informed consent, and ethical approval from UK Research Ethics Committees administered through the Integrated Research Application System. Cells and tissues were analysed with Research Ethics Committee (REC) approval number 11/LO/1274.

Gut biopsies and paired blood samples used for Mass Cytometry were from 5 healthy patients undergoing investigative colonoscopy who had no evidence of inflammation or other pathology. Mononuclear cells from gut biopsies and blood, were isolated as described previously and were stained on day of collection. These were 3 males aged 56, 56, 72 and 2 females aged 30 and 61.

For IMC, Visium and RNAscope analysis, FFPE tissues were obtained from anonymous patients undergoing appendicectomy, tonsillectomy or surgery for removal of colorectal cancer. FFPE blocks from normal colon (n=6), appendix (n=5), ileum (n=5), mesenteric lymph node (n=2), tonsil (n=2) were used. In each case the tissues were histologically normal.

Mouse tissue

Transverse sections of small intestine Peyer's Patches were cut from FFPE small intestine from four C57BL/6 mice obtained from Charles River, Sulzfeld.

Mass cytometry

Mass cytometry antibodies were purchased pre-conjugated from Fluidigm or protein-free from suppliers indicated in Supplementary Table 1 and conjugated in-house using MaxPar labelling kit (Fluidigm) as per manufacturer's instructions. Viability staining was done as follows: 2 x 106 cells washed in cell-staining medium (C-SM, PBS with 0.5% BSA and 2 mM EDTA), incubated for 20 minutes in 1 mL of pre-diluted rhodium-intercalator (Fluidigm, 1:500 in PBS). After washing cells twice in C-SM, the cell pellet was resuspended in 10 μL Kiovig Fc block (Baxter, 5 mg/ mL) for 10 minutes at room temperature followed by the metal-conjugated titrated antibody cocktail in a total volume of 100 μL on ice for 30 minutes. After washing cells twice more in C-SM, cells were fixed overnight in 2% PFA in PBS at 4°C. Cells were then centrifuged and stained with 500 μl of 1μM intercalating Ir (Fluidigm) in perm buffer (0.3% saponin [Fluidigm] in C-SM) and 100 μL intercalatin Ir (Fluidigm) for 30 minutes at room temperature. Finally, cells were washed twice with PBS and twice with Milli-Q water

before resuspended at 0.5 x 106 cells/mL in Milli-Q water with EQ beads (DVS Sciences). Data was then immediately acquired using a Fluidigm Helios mass cytometer, and exported in .fcs format. The mild collagenase digestion of gut biopsies reduced the expression of TACI and CCR6 by mass cytometry; we excluded it from all further analysis.

Analysis of mass cytometry data

FCS files loaded into Cytobank (https://mrc.cytobank.org/) underwent quality control and were filtered to live CD19+ cells only as described previously [15] [13]. Equal numbers of events were extracted from each sample (n=9214) and viSNE dimensionality reduction was run on the combined events using markers CD45RB, IgD, CD20, IgA, CD138, CD21, CD38, CD10, CD27, CD24, IgG and IgM. SPADE was used for clustering based on viSNE co-ordinates, and the resulting nodes were manually placed into B cell subset bubbles based on their expression of key markers (Fig. 1b). Marker expression, subset classification and viSNE coordinates were then exported and loaded into R for visualisation and quantification.

Imaging mass cytometry

Sections were stained as described previously [65]. Briefly, slides were deparaffinized, rehydrated, and subjected to antigen retrieval using a pressure cooker. Tissues were then blocked and incubated with the mix of metal-conjugated antibodies specified in the panel overnight at 4°C and subsequently incubated with the DNA intercalator Iridium Cell-ID Intercalator-Ir (Standard BioTools) before being air-dried. Slides were then inserted into the Hyperion Imaging System and photographed to aid region selection. Regions of about 1 mm2 were selected and laser-ablated at 200-Hz frequency at 1 μm/pixel resolution.

Analysis of imaging mass cytometry

Cell segmentation was performed following IMCSegmentationPipeline [66]. Briefly, raw images for each ROI were extracted and a Random Forest Classifier was applied to each ROI based on manual labels of nucleus, cytoplasm and background on a set of training data [67]. Mean pixel intensities of each channel were calculated for each cell, and loaded into R, with analysis performed as per [68]. Intensities were arcsinh transformed for analysis, spillover correction of intensities was performed and quality control removed cells with size smaller than 11 pixels or with low counts of the DNA channels . Cells were phenotyped by a gating technique per image in the cytomapper R package [69]. Cells were gated as follows, with negative gates used to ensure mutually exclusive classifications: B cells (CD20+), T cells (CD3+), B-T neighbours (CD20+CD3+), Macrophage/DC (CD11c+ or CD11b+ or CD68+), Endothelial (CD31+) and Epithelial (E-Cadherin+). Any cells in the B-T neighbours classification, or cells with more than 1 classification, were not given an assignment. A

random forest classifier was then trained on the classified cells (split 70%/30% to training and testing data respectively) and applied to the unclassified cells in order to give each cell a phenotype [70]. Data were filtered only cells classified as B cells. Batch correction of marker intensity values for B cells was performed using the Harmony package on a principal component analysis consisting of main B cell markers [71]. Due to the presence of B cells within the epithelium, E-cadherin was included within the markers. Cells were then clustered, and a uniform manifold approximation and projection (UMAP) was performed. Clusters were manually classified based on their marker expression and visualised back on the images.

Spatial transcriptomics

Two FFPE blocks (STA and STB) containing normal colon were used to perform Visium spatial transcriptomics according manufacturer's instructions as per demonstrated protocol CG000408 Revision A and CG000409 Revision A. Briefly, $6 \times 6 mm^2$ sections from FFPE blocks were cut at 5mm and placed on the Visium slide inside fiducial frames. Slides were then incubated for 2h at $42°C$ and H&E stained. After imaging using the S60 Nanozoomer scanner, the coverslip was removed, and slides were de-crosslinked. Human whole transcriptome probes (v1 chemistry), consisting of a pair of specific probes for ~18,000 genes, were added to each section for hybridizing to their target genes. Hybridized and ligated probes were released from the tissue sections and captured on the spatially barcoded oligonucleotides on the capture areas. Sequencing libraries were generated from the probes and sequenced on the NextSeq 2000. Sequencing data was processed using Space Ranger pipelines (10x Genomics) for downstream analysis.

Analysis of spatial transcriptomics

Transcriptomic counts of spots in each image were loaded into Seurat [28]. Using Loupe Browser 6.1.0 (10X Genomics), spots were manually annotated as belonging to the subepithelial dome (classification 'SED'), rest of lymphoid follicle (classification 'Follicle'), or outside of the follicle, and these manual cell annotations were attached to the Seurat objects. New Seurat objects using only spots in the SED and Follicle classifications were created and expression data for these were normalised using the SCTransform function [72]. Marker genes that were differentially expressed between the SED and Follicle classifications were determined using the FindMarkers function, with a Wilcoxon rank sum test and thresholds of 0.25 log-fold change and minimum expression in 10% of spots in either classification. Markers were visualised back on the images using custom R scripts.

The full set of spatial transcriptomic data was then analysed using the BayesSpace package in R, with raw counts log-normalised via the spatialPreprocess function, using 15 principal components and 2000 highly variable genes [73]. For spatial clustering, the number of clusters was determined by inspection of tuning plots created using the qTune function, with 7 clusters for sample STA and 8 clusters for sample STB. The spatialCluster function was then used to spatially cluster the data for each sample individually, with spots being assigned to clusters based on similarity of their transcriptomes along with a Markov Chain Monte Carlo algorithm using a spatial prior to encourage neighbouring spots to cluster together. For each sample, the cluster deemed to best represent lymphoid tissue was selected, and the Seurat objects were filtered to spots in these clusters. Data from both samples were merged, and a principal component analysis was run. Harmony was used to correct any batch effects between the samples [71]. Further sub-clustering of spots was done using this harmonised reduction, resulting in a set of sub-clusters present in both images. Marker genes for each sub-cluster were found using FindMarkers, and sub-clusters were relabelled based on their expression of cell lineage markers.

A list of SLE associated genes related genes was collated, and correlations between the expression of each of these genes with the lymphoid spots were calculated using the RcmdrMisc package [74] using Spearman correlation.

RNAScope

RNAscope experiments were performed with Multiplex Fluorescent Reagent Kit V2 Assay (ACDBio #323110) following manufacturer's instructions on 5µm thick FFPE sections. Briefly, slides were dewaxed blocked and subjected to antigen retrieval using a steamer for 22min. Slides were then incubated at 40°C for 30min with the Protease Plus and for 2h with the mix of probes. The appropriate HRP-tagged oligos were then incubated for 15min followed by incubation with TSA-FITC (1:75, #NEL741001KT) or TSA Vivid Fluorophore 570 (1:1500, BioTechne #7526) or TSA Vivid Fluorophore 650 (1:1500, BioTechne #7527) diluted in TSA buffer. The enzymatic reaction was then blocked using HRP Blocker for 15min. When more than one probe was assessed simultaneously, a different set of oligos was incubated followed by another TSA reagent. The following ACDbio probes were used in different combinations: human-C1qB (#565931), human-C1R (#508951-C2), human-C3 (#430701-C3), human-DNASE1L3 (#576121-C2 and 576121-C1), mouse-DNASE1L3 (#819171), Eubacteria 16S (#464461-C3). Slides were finally mounted with Vectashield (VectorLab, H-1200) supplemented with DAPI diluted 1:250 (BioLegend, 422801) and coverslipped. Negative and positive control probes provided were also run on serial sections.

18

For the co-detection of RNA and protein, after the incubation with HRP Blocker, slides were washed, blocked and incubated in primary overnight at 4°C. Primary antibodies used were: Vimentin 1:500 (Standard BioTools, #3143027D), CD3 1:250 (Dako, #A0452, Lysozyme 1:400 (Biolegend, #860001), CD11c 1:300 (Abcam, #ab216655) CD20 (Biolegend, #382802). Slides were washed twice in PBS-T for 5min and incubated for 1h with anti-rabbit AlexaFluor555 (Invitrogen, #A32794) or anti-rabbit AlexaFluor647 (Invitrogen, #A32795TR) diluted 1:500. Slides were then washed and coverslipped as described above. As blocking reagent and antibody diluent a solution containing 5% BSA in PBS-T was used. Slides were imaged using the Eclipse Ti-A Inverted confocal microscope using 20x, 40x or 60X objectives.

For the quantification, images from the different transcripts analysed were separated in single RGB channel and imported in CellProfiler [75]. Images were then made binary using the "Threshold" module and masked. The proportion of pixels masked was then divided by total pixels of the transcript of interest.

IMC-RNAscope combined protocol

Slides were processed as above for RNAscope protocol using DNASE1L3 (ACDbio, #576121-C1) and ITGAX (ACDbio, #419151-C2) and Eubacteria 16S (#464461-C3) probes. C1 oligos were developed with TSA-Digoxygenin (Akoya Biosciences, NEL748001KT), C2 oligos with TSA-FITC, (Akoya Biosciences, #NEL741001KT) and C3 oligos with TSA-Biotin (Akoya Biosciences, NEL749A001KT) diluted 1:100 in TSA buffer. After the RNAscope protocol, slides were washed, blocked, and stained using Panel 2 and 3 (Supplementary Table 2) as described in the IMC paragraph above. Panels 2 and 3 incorporated anti-Digoxygenin (RD System, #611621), anti-FITC (Standard BioTools, #3174006B) and anti-Biotin (Biolegend, #409002) metal-tagged antibodies to detect the RNAscope staining output.

For quantification of co-expression of DNASE1L3 and Lysozyme or DNASE1L3 and NOX2 expressing cells, DNASE1L3+ cells were segmented using the FindPrimaryObjects module in CellProfiler using the DNASE1L3 channel as input [75]. The expression of all markers was calculated for each cell and arcsinh transformed, and cells were manually labelled as belonging to the SED using a custom R script.

TUNEL assay and Lysozyme co-staining

TUNEL assay was performed using the In Situ Cell Death Detection Kit, Fluorescein (Roche, #11684795910) according to the manufacturers' instructions including an antigen retrieval in

a pressure cooker with Basic Buffer (RD System, #CTS013). Slides were then blocked and incubated overnight at 4°C with anti-Lysozyme antibody diluted 1:400 (Biolegend, #860001). Slides were washed and incubated with anti-rabbit Alexa Fluor 555 diluted 1:500 (Invitrogen, #A32794), coverslipped, and imaged as described above.

IMC-TUNEL combined protocol

TUNEL assay was performed using the In Situ Cell Death Detection Kit, Fluorescein (Roche, #11684795910) as described above. Slides were then blocked and stained with Panel 1 (Supplementary Table 2) as described in the IMC paragraph above. An anti-FITC (Standard BioTools, #3174006B) metal-tagged antibody was incorporated into the panel to detect the output of the TUNEL assay.

Statistics

All statistical tests were done in Prism v9 unless stated otherwise. Two-tailed paired t tests were performed using parametric tests for comparisons of GALT and PBMC in mass cytometry data, as well as for paired comparisons of IMC data. For differential gene expression in spatial transcriptomic data, Wilcoxon tests with Bonferroni correction for multiple comparisons were used in the Seurat package in R version 4.2.2 [28]. Statistically significant gene correlations in spatial transcriptomic data were identified using Spearman correlation in R along with pairwise P values corrected for multiple inference using Holm's method.

Figure legends

Fig. 1 Comparison of B cells in GALT and PBMC using mass cytometry

a viSNE plot of B cells (gated CD19+) from GALT and PBMC combined, showing expression of markers used for dimensionality reduction. **b** Overview of B cell subset classifications and phenotypes used in SPADE analysis. **c** Representative SPADE plots of matched GALT (top row) and PBMC (bottom row) from the same donor. Node size indicates number of cells and colour indicates median expression level for each marker. **d** viSNE plot showing distribution of B cell subsets for all donors, split by GALT (left) and PBMC (right). **e** Pie charts showing relative proportions of each B cell subtype for all donors for GALT (left) and PBMC (right). Colouring matches that shown in **d**. **f** Matched relative proportions of DN1 (left) and DN2/3 (right) for each sample. ** = P<0.01. **g** Heatmap showing the mean expression (scaled by row) of key markers for each subset split by GALT (left) and PBMC (right).

Fig. 2 Analysis of B cells in GALT using imaging mass cytometry

a IMC acquired images of GALT in human colon showing key lineage marker expression, with inset of epithelium and sub-epithelial dome region. Cyan = CD20, magenta = CD3, yellow = CD68, orange = E-cadherin. **B** Segmentation of cells (grey) from same ROI as shown in **a**, with cells computationally classified as B cells highlighted (red). **C** UMAP of B cells with assigned subtypes. **d** Heatmap of mean expression of markers used for dimensionality reduction for B cell subtypes. **e** B cell subtypes overlaid on ROI from **a** and **b**. **f** IMC of ROI from **a**, showing DN-related markers CD27 (green), IgD (red) and CD11c (magenta). Computationally derived borders of DN2 cells (white) are overlaid and show that they are found in a region that is CD27-IgD-CD11c+ near the epithelium. **G** Confocal microscopy image showing the distribution of DAPI (blue), CD20 (red) and *ITGAX* (green) in a representative appendix follicle. CD20$^+$*ITGAX*$^+$ cells are highlighted with arrows. Scalebars = 100μm.

Fig. 3 Spatial transcriptomics analysis of GALT

a H&E staining of GALT in colon, along with the expression of B cell markers in spatial transcriptomics spots showing follicle location. **B** Manual classification of spots as either sub-epithelial dome (SED, black) or rest of follicle (white). **C** Heatmap of genes differentially expressed between SED and Follicle spots. **D** Expression of DNASE1L3 and complement components within the follicle, with spots coloured by expression and outlined as per

21

classifications in **b**. **e** Unsupervised clustering of spots using BayesSpace, with spots clustered based on gene expression and similarity to neighbouring spots. **F** Spots classified as lymphoid tissues (Cluster 7). **G** Sub-clustering of lymphoid tissue spots from **f**. **h** Dot plot of gene expression within each subcluster, showing expression of cell lineage markers, DNASE1L3 and complement components. **I** Correlation matrix of DNASE1L3, complement components and genes associated with apoptosis in lupus, showing correlation of expression within the lymphoid tissue spots in **f**. *=P < 0.05; **=P < 0.01; ***=P < 0.001, ****=P<0.0001. Scalebars = 500μm for **a**, **b** and **d**, and 2mm for **e**, **f** and **g**.

Fig. 4 DNASE1L3 expression in the GALT subepithelial region

A IMC-derived images illustrating spatial structure of an appendix follicle, including sub-epithelial dome (SED, green) and germinal centre (GC, grey), shown through expression of CD3 (T cells, brown), Ki67 (proliferation, cyan) and E-Cadherin (epithelium, orange). **B** Same field as in **a** showing staining of CD68 (red) and E-Cadherin (orange), as well as gene expression of *DNASE1L3* (green) and *ITGAX* (encoding for CD11c, blue) through RNAscope. **C** Enlarged profiles of SED and GC regions from **b** showing individual and merged marker expression. **D** Quantification of distance to epithelium for cells segmented on CD68 and *ITGAX* classified as either DNASE1L3$^+$ or DNASE1L3$^-$. **e** Quantification of *DNASE1L3* expression on cells segmented on CD68 and *ITGAX* classified as either GC or non-GC. **f** IMC-derived images of ileum follicle illustrating CD3 (brown) and E-Cadherin (orange) protein localization alongside *DNASE1L3* (green) and *ITGAX* (blue) gene expression. **G** Visualisation of *DNASE1L3* (green), NOX2 (blue) and lysozyme (red) distribution within the same image as in **a**. **h** Enlargement of SED region depicted in **g**. **I** Quantification of the average expression of NOX2 and lysozyme in *DNASE1L3*$^+$ cells and classified as residing in either the SED or in non-SED regions. **j** Confocal microscopy image showing DAPI (blue), CD20 (cyan), *ITGAX* (green) and CD11c (red) expression in an appendix follicle. *ITGAX*$^+$CD20$^+$ cells assessed using RNAscope are highlighted with cyan arrows. CD11c$^+$*ITGAX*$^+$CD20$^-$ cells are highlighted with red arrows. Statistics between groups were assessed by two-tailed T test. *=P < 0.05, **=P < 0.01 and ****=P<0.0001. Scalebars =100mm

Fig. 5 Distribution of apoptotic debris and DNASE1L3 in human and distribution of DNASE1L3 in mouse

a Visualisation of apoptotic debris (TUNEL, red), cell nuclei (Iridium, blue), CD68 (green) and E-Cadherin (orange) obtained using an IMC-TUNEL combined protocol in a section serial to

that shown in Fig 2A. **b** Distribution of CD11c (yellow) and CD163 (cyan) on same image as in (A). **c** Magnification of regions in **a** and **b**, showing individual and merged images for each marker shown in **a** and **b**. **d** RNAscope analysis showing the distribution of cell nuclei (DAPI, blue), *DNASE1L3* (cyan) and lysozyme (magenta) in mouse GALT (Peyer's Patch) assessed through confocal microscopy, including demarcation of subepithelial dome (SED; green) and germinal centre (GC; grey) **e** Visualisation of apoptotic debris with TUNEL assay (green) and lysozyme protein expression (magenta) through confocal microscopy in mouse GALT. Scalebars =100μm

Fig. 6 Expression of ITGAX and DNASE1L3 in the colonic lamina propria

a IMC image of colonic lamina propria showing the distribution of E-Cadherin (orange), CD68 (red), *ITGAX* (blue) alongside those of CD163 (magenta) and *DNASE1L3* (green) **b**. Scalebars =100μm

Fig. 7 Co-expression of C1Q with DNASE1L3 and C1R with C3

a RNAscope coupled with confocal microscopy showing cell nuclei (DAPI, blue), *DNASE1L3* (green) and *C1Qb* (red) transcripts in the SED and GC of an appendix follicle. **b** RNAscope coupled with confocal microscopy showing cell nuclei (DAPI, blue), *C1r* (cyan), *C3* (magenta) and *C1Qb* (red) transcripts in an appendix follicle. **c** RNAscope coupled with confocal microscopy showing cell nuclei (DAPI, blue), Vimentin protein (yellow), and *C3* (magenta) and *C1R* (cyan) transcripts in a human appendix follicle. Scalebars =100μm

Fig. 8 Localization of internalized bacteria

a Confocal microscopy images showing distribution of *DNASE1L3* (red) and *Bacterial 16S* (green) transcripts in a colonic follicle. Bacterial RNA was assessed using a probe against a region of the 16S transcript conserved across all Eubacteria by RNAscope. DAPI was used as nuclear counterstain. **b** Magnification of the image in **a**. **c** RNAscope coupled with IMC analysis showing the distribution of *Bacterial 16S,* E-Cadherin and nuclear counterstain (Iridium). **d-g** magnification of image shown in **c** alongside other markers for different cell lineages. Scalebars = 100μm

Fig. S1 Computational analysis of imaging mass cytometry data

a Quality control of segmented cells on images, removing cells with small size or lack of DNA staining (grey = cells passing quality control, red = cells failing quality control). **b** Pie charts showing the proportions of cell types in the gated cells (left), and in the final classifications from gating and machine learning (right). **c** Relative importance of each marker in the IMC panel to the machine learning algorithm. **d** Overlay of cell types on colon sample shown in **Fig. 2a**, **e** Overlay of B cell subtypes on human appendix sample. **f** Overlay of B cell subtypes on human ileum sample.

Fig. S2 Computational analysis of spatial transcriptomic data

a H&E staining of spatial transcriptomic samples STA and STB. **b** Manual classification of spots as being either in the sub-epithelial dome ('SED', black) or rest of follicle ('Follicle', white). **c**. Negative log-likelihood plot for varying spatial cluster numbers (q) used to determine ideal spatial cluster number per image. **d** Spatial clusters for sample STB. **e** Lymphoid tissue spot assignment for sample STB. **f** Harmonised principal component analysis showing overlap of spots from STA (purple) and STB (green). **g** Sub-cluster regions for sample STB. **h** Correlation R values for gene correlations as shown in **Fig. 1i** Correlation P values for gene correlations as shown in **Fig. 1i**.

Fig. S3 DNASE1L3 distribution in Appendix, Ileum and Colon.

Images from IMC coupled with RNAscope analysis showing the distribution of CD68 (red) and CD45RB (white) localization alongside *DNASE1L3* (green) gene expression in **a** Appendix, **b** Ileum, **c** colon.

Fig. S4 *DNASE1L3* distribution in mLN and Tonsil

a, b Images from IMC coupled with RNAscope analysis showing the distribution of CD68 (red) and CD45RB (white) localization alongside *DNASE1L3* (green) gene expression in **a** mesenteric lymph node (mLN) and **b** Tonsil. GC (grey) and T cell zone (magenta) are highlighted.

Fig. S5 Distribution of TUNEL positive apoptotic debris in appendix and colon and quantification of *DNASE1L3* and complement component co-expression.

a, b IMC-derived image showing distribution of CD68 (green), cell apoptosis (TUNEL, red), CD45RB (white) and E-cadherin (orange) in colon and appendix follicles. Regions of

24

subepithelial dome (SED, green) and germinal centre (GC, grey) are highlighted and enlarged. Scalebar=100mm. **c,d** Quantification of RNAscope analysis shown in Figure 7. **c** Percentage of $C1Qb^+DNASE1L3^+$ double positive pixels over total $DNASE1L3^+$ pixels. **d** Percentage of double positive pixels over total pixels for $C1Qb^+C1R^+$/total $C1R^+$, $C1Qb^+C3^+$/total $C3^+$, $C3^+C1R^+$/total $C1R^+$ and $C3^+C1R^+$/total $C3^+$.

Competing interests:

Authors declare that they have no competing interests.

References

1   Brandtzaeg, P. *et al.* The B-cell system of human mucosae and exocrine glands. *Immunol Rev* **171**, 45-87 (1999).

2   Lelouard, H., Fallet, M., de Bovis, B., Meresse, S. & Gorvel, J. P. Peyer's patch dendritic cells sample antigens by extending dendrites through M cell-specific transcellular pores. *Gastroenterology* **142**, 592-601 e593, doi:10.1053/j.gastro.2011.11.039 (2012).

3   Gustafson, C. E. *et al.* Limited expression of APRIL and its receptors prior to intestinal IgA plasma cell development during human infancy. *Mucosal immunology* **7**, 467-477, doi:10.1038/mi.2013.64 (2014).

4   Komban, R. J. *et al.* Activated Peyer's patch B cells sample antigen directly from M cells in the subepithelial dome. *Nat Commun* **10**, 2423, doi:10.1038/s41467-019-10144-w (2019).

5   Yamanaka, T. *et al.* M cell pockets of human Peyer's patches are specialized extensions of germinal centers. *European journal of immunology* **31**, 107-117, doi:10.1002/1521-4141(200101)31:1<107::aid-immu107>3.0.co;2-4 (2001).

6   Torow, N. *et al.* M cell maturation and cDC activation determine the onset of adaptive immune priming in the neonatal Peyer's patch. *Immunity*, doi:10.1016/j.immuni.2023.04.002 (2023).

7   Litinskiy, M. B. *et al.* DCs induce CD40-independent immunoglobulin class switching through BLyS and APRIL. *Nature immunology* **3**, 822-829, doi:10.1038/ni829 (2002).

8   Grasset, E. K. *et al.* Gut T cell-independent IgA responses to commensal bacteria require engagement of the TACI receptor on B cells. *Sci Immunol* **5**, doi:10.1126/sciimmunol.aat7117 (2020).

9   Kabbert, J. *et al.* High microbiota reactivity of adult human intestinal IgA requires somatic mutations. *The Journal of experimental medicine* **217**, doi:10.1084/jem.20200275 (2020).

10   Mora, J. R. Homing imprinting and immunomodulation in the gut: role of dendritic cells and retinoids. *Inflamm Bowel Dis* **14**, 275-289, doi:10.1002/ibd.20280 (2008).

11   Spencer, J. & Sollid, L. M. The human intestinal B-cell response. *Mucosal immunology* **9**, 1113-1124, doi:10.1038/mi.2016.59 (2016).

12   Fagarasan, S. *et al.* Critical roles of activation-induced cytidine deaminase in the homeostasis of gut flora. *Science* **298**, 1424-1427, doi:10.1126/science.1077336 (2002).

13   Tull, T. J. *et al.* Human marginal zone B cell development from early T2 progenitors. *The Journal of experimental medicine* **218**, doi:10.1084/jem.20202001 (2021).

14   Weller, S. *et al.* T-independent responses to polysaccharides in humans mobilize marginal zone B cells prediversified against gut bacterial antigens. *Sci Immunol* **8**, eade1413, doi:10.1126/sciimmunol.ade1413 (2023).

15   Zhao, Y. *et al.* Spatiotemporal segregation of human marginal zone and memory B cell populations in lymphoid tissue. *Nat Commun* **9**, 3857, doi:10.1038/s41467-018-06089-1 (2018).
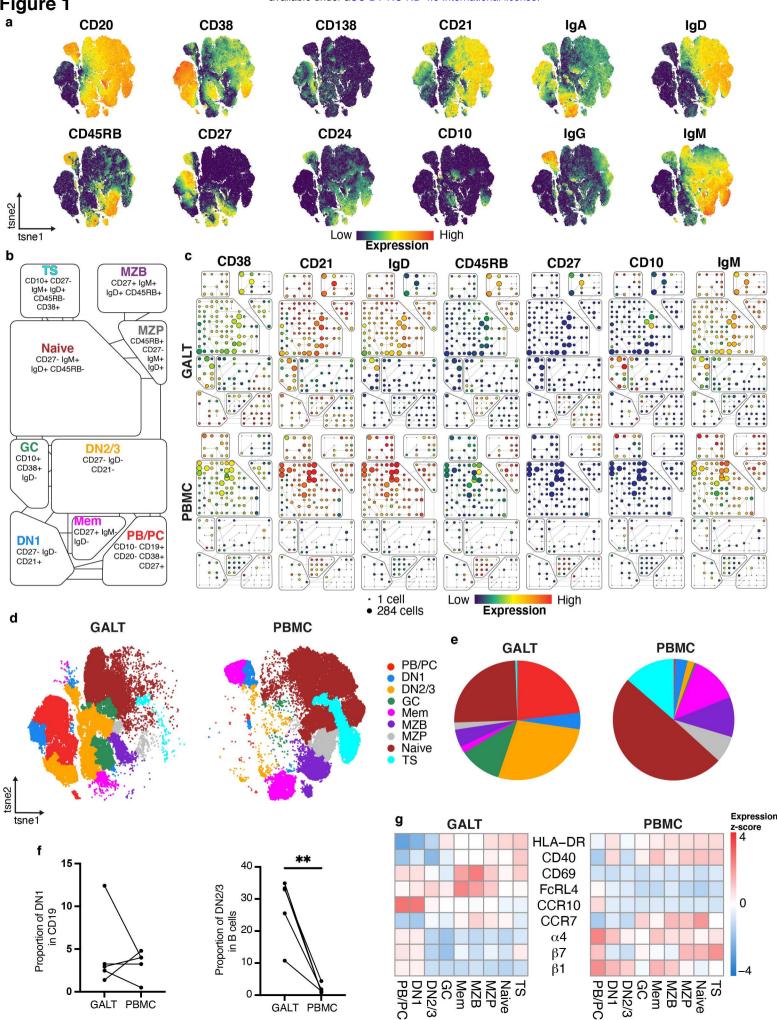
16   Spencer, J., Finn, T., Pulford, K. A., Mason, D. Y. & Isaacson, P. G. The human gut contains a novel population of B lymphocytes which resemble marginal zone cells. *Clin Exp Immunol* **62**, 607-612 (1985).

17   Lelouard, H. *et al.* Pathogenic bacteria and dead cells are internalized by a unique subset of Peyer's patch dendritic cells that express lysozyme. *Gastroenterology* **138**, 173-184 e171-173, doi:10.1053/j.gastro.2009.09.051 (2010).

18   Barone, F., Patel, P., Sanderson, J. D. & Spencer, J. Gut-associated lymphoid tissue contains the molecular machinery to support T-cell-dependent and T-cell-independent class switch recombination. *Mucosal immunology* **2**, 495-503, doi:10.1038/mi.2009.106 (2009).

19   von Bulow, G. U., van Deursen, J. M. & Bram, R. J. Regulation of the T-independent humoral response by TACI. *Immunity* **14**, 573-582 (2001).

20   Jenks, S. A. *et al.* Distinct Effector B Cells Induced by Unregulated Toll-like Receptor 7 Contribute to Pathogenic Responses in Systemic Lupus Erythematosus. *Immunity* **49**, 725-739 e726, doi:10.1016/j.immuni.2018.08.015 (2018).

21   Matthew Woodruff, R. R., Kevin Cashman, Doan Nguyen, Ankur Saini, Natalie Haddad, Ariel Ley, Shuya Kyu, J. Christina Howell, Tugba Ozturk, Saeyun Lee, Weirong Chen, Jacob Estrada, Andrea Morrison-Porter, Andrew Derrico, Fabliha Anam, Monika Sharma, Henry Wu, Sang Le, Scott Jenks, Christopher M Tipton, Wiliam Hu, F. Eun-Hyung Lee, Ignacio Sanz. Dominant extrafollicular B cell responses in severe COVID-19 disease correlate with robust viral-specific antibody production but poor clinical outcomes. *medRxiv* **https://www.medrxiv.org/content/10.1101/2020.04.29.20083717v2** (2020).

22   Hartl, J. *et al.* Autoantibody-mediated impairment of DNASE1L3 activity in sporadic systemic lupus erythematosus. *The Journal of experimental medicine* **218**, doi:10.1084/jem.20201138 (2021).

23   Sisirak, V. *et al.* Digestion of Chromatin in Apoptotic Cell Microparticles Prevents Autoimmunity. *Cell* **166**, 88-101, doi:10.1016/j.cell.2016.05.034 (2016).

24   Pickering, M. C. & Botto, M. Are anti-C1q antibodies different from other SLE autoantibodies? *Nat Rev Rheumatol* **6**, 490-493, doi:10.1038/nrrheum.2010.56 (2010).

25   Stamper, H. B., Jr. & Woodruff, J. J. Lymphocyte homing into lymph nodes: in vitro demonstration of the selective affinity of recirculating lymphocytes for high-endothelial venules. *The Journal of experimental medicine* **144**, 828-833, doi:10.1084/jem.144.3.828 (1976).

26   Amara, K. *et al.* B cells expressing the IgA receptor FcRL4 participate in the autoimmune response in patients with rheumatoid arthritis. *Journal of autoimmunity*, doi:10.1016/j.jaut.2017.03.004 (2017).

27   Ehrhardt, G. R. *et al.* Discriminating gene expression profiles of memory B cell subpopulations. *The Journal of experimental medicine* **205**, 1807-1817, doi:10.1084/jem.20072682 (2008).

28   Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587 e3529, doi:10.1016/j.cell.2021.04.048 (2021).

29   Anderle, P. *et al.* Novel markers of the human follicle-associated epithelium identified by genomic profiling and microdissection. *Gastroenterology* **129**, 321-327, doi:10.1053/j.gastro.2005.03.044 (2005).
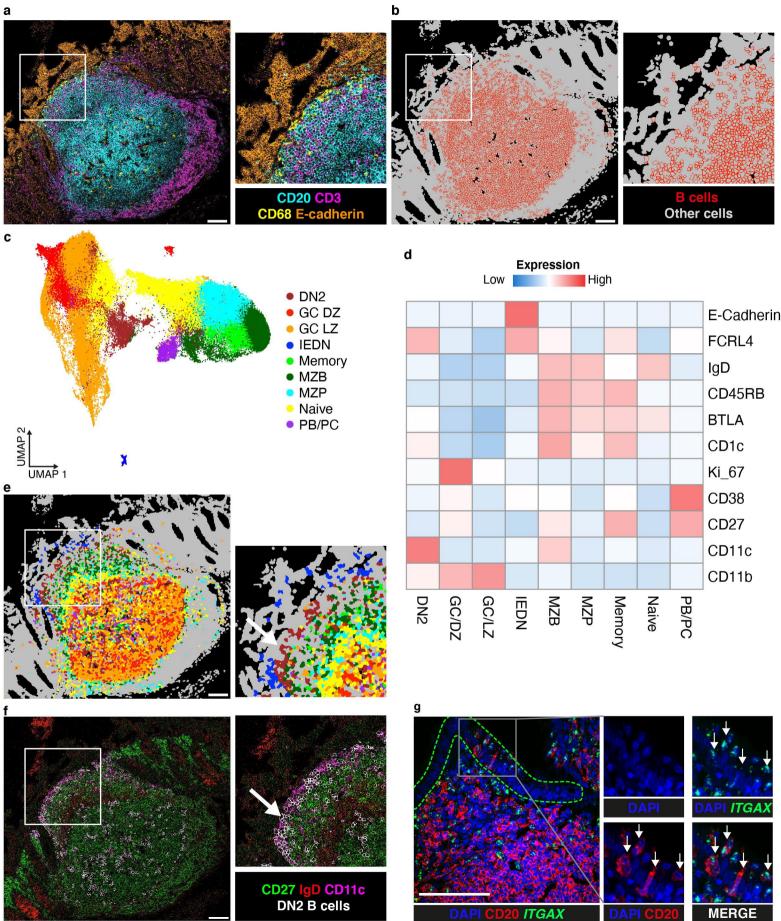
30   Cattoretti, G. *et al.* BCL-6 protein is expressed in germinal-center B cells. *Blood* **86**, 45-53 (1995).

31   King, H. W. *et al.* Single-cell analysis of human B cell maturation predicts how antibody class switching shapes selection dynamics. *Sci Immunol* **6**, doi:10.1126/sciimmunol.abe6291 (2021).

32   Al-Mayouf, S. M. *et al.* Loss-of-function variant in DNASE1L3 causes a familial form of systemic lupus erythematosus. *Nat Genet* **43**, 1186-1188, doi:10.1038/ng.975 (2011).

33   Gezgin Yildirim, D. & Bakkaloglu, S. A. Monogenic lupus caused by mutations in DNASE1L3: A rare cause of systemic lupus erythematosus in children. *Scand J Immunol* **95**, e13162, doi:10.1111/sji.13162 (2022).

34   Berkel, A. I. *et al.* Development of systemic lupus erythematosus in a patient with selective complete C1q deficiency. *Eur J Pediatr* **156**, 113-115, doi:10.1007/s004310050567 (1997).

35   Hsieh, S. C., Tsai, C. Y. & Yu, C. L. Potential serum and urine biomarkers in patients with lupus nephritis and the unsolved problems. *Open Access Rheumatol* **8**, 81-91, doi:10.2147/OARRR.S112829 (2016).

36   Hong, S. M., Chen, W., Feng, J., Dai, D. & Shen, N. Novel Mutations in ACP5 and SAMHD1 in a Patient With Pediatric Systemic Lupus Erythematosus. *Front Pediatr* **10**, 885006, doi:10.3389/fped.2022.885006 (2022).

37   Serveaux-Dancer, M. *et al.* Pathological Implications of Receptor for Advanced Glycation End-Product (AGER) Gene Polymorphism. *Dis Markers* **2019**, 2067353, doi:10.1155/2019/2067353 (2019).

38   Martinez-Valle, F. *et al.* DNase1 activity in systemic lupus erythematosus patients with and without nephropathy. *Rheumatology international* **30**, 1601-1604, doi:10.1007/s00296-009-1199-6 (2010).

39   Costa-Reis, P. & Sullivan, K. E. Monogenic lupus: it's all new! *Curr Opin Immunol* **49**, 87-95, doi:10.1016/j.coi.2017.10.008 (2017).

40   Lee, Y. H. & Song, G. G. Associations between the FAS -670 A/G, -1377 G/A, and FASL -844 T/C polymorphisms and susceptibility to systemic lupus erythematosus: a meta-analysis. *Clin Exp Rheumatol* **34**, 634-640 (2016).

41   Herrmann, M. *et al.* Impaired phagocytosis of apoptotic cell material by monocyte-derived macrophages from patients with systemic lupus erythematosus. *Arthritis Rheum* **41**, 1241-1250, doi:10.1002/1529-0131(199807)41:7<1241::AID-ART15>3.0.CO;2-H (1998).

42   Casciola-Rosen, L. A., Anhalt, G. & Rosen, A. Autoantigens targeted in systemic lupus erythematosus are clustered in two populations of surface structures on apoptotic keratinocytes. *The Journal of experimental medicine* **179**, 1317-1330, doi:10.1084/jem.179.4.1317 (1994).

43   Alaoui, L. *et al.* Functional specialization of short-lived and long-lived macrophage subsets in human tonsils. *The Journal of experimental medicine* **220**, doi:10.1084/jem.20230002 (2023).

44   Weisenburger, T. *et al.* Epistatic Interactions Between Mutations of Deoxyribonuclease 1-Like 3 and the Inhibitory Fc Gamma Receptor IIB Result in Very Early and Massive Autoantibodies Against Double-Stranded DNA. *Front Immunol* **9**, 1551, doi:10.3389/fimmu.2018.01551 (2018).

45   Blander, J. M. Death in the intestinal epithelium-basic biology and implications for inflammatory bowel disease. *FEBS J* **283**, 2720-2730, doi:10.1111/febs.13771 (2016).
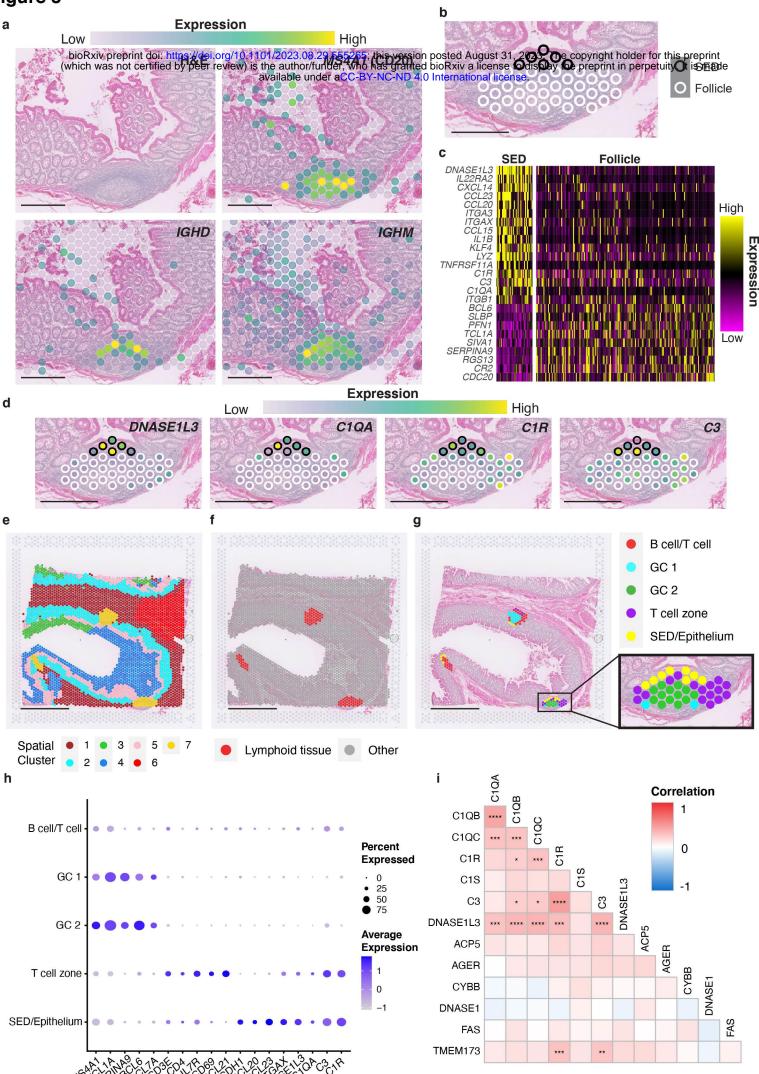
46   Inokuchi, S. *et al.* Homeostatic Milieu Induces Production of Deoxyribonuclease 1-like 3 from Myeloid Cells. *Journal of immunology (Baltimore, Md. : 1950)* **204**, 2088-2097, doi:10.4049/jimmunol.1901304 (2020).

47   Wangriatisak, K. *et al.* The expansion of activated naive DNA autoreactive B cells and its association with disease activity in systemic lupus erythematosus patients. *Arthritis research & therapy* **23**, 179, doi:10.1186/s13075-021-02557-0 (2021).

48   Hauer, A. C., Bajaj-Elliott, M., Williams, C. B., Walker-Smith, J. A. & MacDonald, T. T. An analysis of interferon gamma, IL-4, IL-5 and IL-10 production by ELISPOT and quantitative reverse transcriptase-PCR in human Peyer's patches. *Cytokine* **10**, 627-634, doi:10.1006/cyto.1997.0337 (1998).

49   Carsetti, R. *et al.* Lack of Gut Secretory Immunoglobulin A in Memory B-Cell Dysfunction-Associated Disorders: A Possible Gut-Spleen Axis. *Front Immunol* **10**, 2937, doi:10.3389/fimmu.2019.02937 (2019).

50   Lacey, K. A. *et al.* Secreted mammalian DNases protect against systemic bacterial infection by digesting biofilms. *The Journal of experimental medicine* **220**, doi:10.1084/jem.20221086 (2023).

51   Agarwal, V. & Blom, A. M. Roles of Complement C1q in Pneumococcus-Host Interactions. *Crit Rev Immunol* **35**, 173-184, doi:10.1615/critrevimmunol.2015012177 (2015).

52   Magri, G. *et al.* Human Secretory IgM Emerges from Plasma Cells Clonally Related to Gut Memory B Cells and Targets Highly Diverse Commensals. *Immunity* **47**, 118-134.e118, doi:10.1016/j.immuni.2017.06.013 (2017).

53   Mitchell, D. A. *et al.* C1q deficiency and autoimmunity: the effects of genetic background on disease expression. *Journal of immunology (Baltimore, Md. : 1950)* **168**, 2538-2543, doi:10.4049/jimmunol.168.5.2538 (2002).

54   Schaller, M., Bigler, C., Danner, D., Ditzel, H. J. & Trendelenburg, M. Autoantibodies against C1q in systemic lupus erythematosus are antigen-driven. *Journal of immunology (Baltimore, Md. : 1950)* **183**, 8225-8231, doi:10.4049/jimmunol.0902642 (2009).

55   Baumann, I. *et al.* Impaired uptake of apoptotic cells into tingible body macrophages in germinal centers of patients with systemic lupus erythematosus. *Arthritis Rheum* **46**, 191-201, doi:10.1002/1529-0131(200201)46:1<191::AID-ART10027>3.0.CO;2-K (2002).

56   Shao, W. H. & Cohen, P. L. Disturbances of apoptotic cell clearance in systemic lupus erythematosus. *Arthritis research & therapy* **13**, 202, doi:10.1186/ar3206 (2011).

57   Leibler, C. *et al.* Genetic dissection of TLR9 reveals complex regulatory and cryptic proinflammatory roles in mouse lupus. *Nature immunology* **23**, 1457-1469, doi:10.1038/s41590-022-01310-2 (2022).

58   Wen, L. *et al.* Toll-like receptors 7 and 9 regulate the proliferation and differentiation of B cells in systemic lupus erythematosus. *Front Immunol* **14**, 1093208, doi:10.3389/fimmu.2023.1093208 (2023).

59   Marshak-Rothstein, A. Toll-like receptors in systemic autoimmune disease. *Nat Rev Immunol* **6**, 823-835, doi:10.1038/nri1957 (2006).

60   Stamnaes, J. & Sollid, L. M. Celiac disease: Autoimmunity in response to food antigen. *Semin Immunol* **27**, 343-352, doi:10.1016/j.smim.2015.11.001 (2015).

61   Jabri, B. & Sollid, L. M. T Cells in Celiac Disease. *Journal of immunology (Baltimore, Md. : 1950)* **198**, 3005-3014, doi:10.4049/jimmunol.1601693 (2017).

62   Iversen, R. *et al.* Efficient T cell-B cell collaboration guides autoantibody epitope bias and onset of celiac disease. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 15134-15139, doi:10.1073/pnas.1901561116 (2019).

63   Di Niro, R. *et al.* High abundance of plasma cells secreting transglutaminase 2-specific IgA autoantibodies with limited somatic hypermutation in celiac disease intestinal lesions. *Nat Med* **18**, 441-445, doi:10.1038/nm.2656 (2012).

64   Amundsen, S. F., Stamnaes, J., Lundin, K. E. A. & Sollid, L. M. Expression of transglutaminase 2 in human gut epithelial cells: Implications for coeliac disease. *PLoS One* **18**, e0287662, doi:10.1371/journal.pone.0287662 (2023).

65   Bortolomeazzi, M. *et al.* Immunogenomics of Colorectal Cancer Response to Checkpoint Blockade: Analysis of the KEYNOTE 177 Trial and Validation Cohorts. *Gastroenterology* **161**, 1179-1193, doi:10.1053/j.gastro.2021.06.064 (2021).

66   Zanotelli, V. R. *ImcSegmentationPipeline: A pixelclassification based multiplexed image segmentation pipeline*, <https://zenodo.org/record/3841961> (2022).

67   Berg, S. *et al.* ilastik: interactive machine learning for (bio)image analysis. *Nature methods* **16**, 1226-1232, doi:10.1038/s41592-019-0582-9 (2019).

68   Eling, N. *Analysis workflow for IMC data*, <https://bodenmillergroup.github.io/IMCDataAnalysis/> (2023).

69   Eling, N., Damond, N., Hoch, T. & Bodenmiller, B. cytomapper: an R/Bioconductor package for visualization of highly multiplexed imaging data. *Bioinformatics* **36**, 5706-5708, doi:10.1093/bioinformatics/btaa1061 (2021).

70   Kuhn, M. *caret: Classification and Regression Training, R package version 6.0-94*, <https://cran.r-project.org/web/packages/caret/> (2021).

71   Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nature methods* **16**, 1289-1296, doi:10.1038/s41592-019-0619-0 (2019).

72   Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* **20**, 296, doi:10.1186/s13059-019-1874-1 (2019).

73   Zhao, E. *et al.* Spatial transcriptomics at subspot resolution with BayesSpace. *Nat Biotechnol* **39**, 1375-1384, doi:10.1038/s41587-021-00935-2 (2021).

74   Fox, J. *RcmdrMisc: R Commander Miscellaneous Functions, R package version 2.7-2*, <https://cran.r-project.org/web/packages/RcmdrMisc/> (2022).

75   Stirling, D. R. *et al.* CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics* **22**, 433, doi:10.1186/s12859-021-04344-9 (2021).

Figure 1

Figure 2

a

CD20 CD3
CD68 E-cadherin

b

B cells
Other cells

c

- DN2
- GC DZ
- GC LZ
- IEDN
- Memory
- MZB
- MZP
- Naive
- PB/PC

UMAP 2
UMAP 1

d

Expression
Low — High

	DN2	GC/DZ	GC/LZ	IEDN	MZB	MZP	Memory	Naive	PB/PC
E-Cadherin									
FCRL4									
IgD									
CD45RB									
BTLA									
CD1c									
Ki_67									
CD38									
CD27									
CD11c									
CD11b									

e

f

CD27 IgD CD11c
DN2 B cells

g

DAPI | DAPI *ITGAX*
DAPI CD20 | MERGE

DAPI CD20 *ITGAX*

Figure 3

Figure 4

Figure 5

a Iridium TUNEL CD68 E-cadherin

b Iridium TUNEL CD11c CD163 E-cadherin

c Iridium TUNEL | Iridium CD68 | Iridium CD163 | Iridium CD11c | MERGE

d DAPI *DNASE1L3* | DAPI LYSOZYME | DAPI *DNASE1L3* LYSOZYME | DAPI | LYSOZYME | *DNASE1L3* | MERGE

e DAPI TUNEL LYSOZYME

Figure 6

Figure 7

a

SED

GC

DAPI *DNASE1L3* *C1Qb*

DAPI *DNASE1L3* | DAPI *C1Qb* | MERGE

b

DAPI *C1r* *C3* *C1Qb*

DAPI *C3* | DAPI *C1r*

DAPI *C1Qb* | MERGE

c

DAPI Vimentin *C3* *C1r*

DAPI *Vimentin* | DAPI *C1r*

DAPI *C3* | MERGE

Figure 8

a
DAPI *Bacterial 16S* *DNASE1L3*

SED

b

c
Iridium (Nuclei) *Bacterial 16S* E-cadherin

d
Bacterial 16S E-cadherin

Bacterial 16S Iridium (Nuclei)

e
Bacterial 16S CD68

Bacterial 16S *DNASE1L3*

f
Bacterial 16S CD3

Bacterial 16S CD31

g
Bacterial 16S CD38

Bacterial 16S FCRL4

Figure S1

a Pass QC / Fail QC — DNA vs size scatter plot

b Gated cells / Cells gated + machine learning pie charts

B Cell, Unlabelled, Endothelial Cell, Epithelial Cell, MacDC, T Cell

c CD20, E_Cadherin, CD3, CD31, CD68, CD45RB, CD40, NOX2, CD45RO, CD11c, CD74, CD11b, DNASE_1, GzB, Vimentin, PDL1, LAMP3, FCRL4, CD27, CD163, CD5, IgM, Lysozyme, CD8, CD38, CD83, CD1c, CD103, TUNEL, BTLA, KI_67, IgD — Importance

d B Cell, Endothelial Cell, Epithelial Cell, MacDC, T Cell

e Appendix — DN2, GC DZ, GC LZ, IEDN, Memory, MZB, MZP, Naive, PB/PC

f Ileum — DN2, GC DZ, GC LZ, IEDN, Memory, MZB, MZP, Naive, PB/PC

Figure S2

a

STA STB

b

● SED ○ Follicle

c

STA

STB

Negative log likelihood

Number of clusters (q)

d

Spatial Cluster

1
2
3
4
5
6
7
8

e

Lymphoid tissue

Other

f

STA
STB

harmony_2

harmony_1

g

B cell/T cell
T cell zone
GC 1
GC 2
SED/Epithelium

h Spatial correlation R values

	C1QA	C1QB	C1QC	C1R	C1S	C3	DNASE1L3	ACP5	AGER	CYBB	DNASE1	FAS
C1QB	0.43											
C1QC	0.32	0.30										
C1R	0.19	0.23	0.24									
C1S	0.10	0.24	0.10	0.21								
C3	0.11	0.22	0.25	0.40	0.18							
DNASE1L3	0.28	0.32	0.28	0.30	0.13	0.32						
ACP5	0.03	0.04	0.03	0.16	0.03	0.15	0.09					
AGER	−0.07	0.04	−0.01	0.05	0.07	0.07	0.04	0.04				
CYBB	−0.02	0.04	−0.10	0.04	0.16	0.05	0.07	−0.06	−0.01			
DNASE1	−0.08	−0.02	0.01	−0.06	0.02	−0.04	−0.09	0.07	−0.02	−0.09		
FAS	0.00	0.13	0.00	0.09	0.01	−0.02	0.12	0.05	0.07	0.11	−0.15	
TMEM173	0.07	0.13	0.09	0.28	0.05	0.22	0.12	0.10	0.05	−0.06	−0.08	0.13

i Spatial correlation P values

	C1QA	C1QB	C1QC	C1R	C1S	C3	DNASE1L3	ACP5	AGER	CYBB	DNASE1	FAS
C1QB	0.0001											
C1QC	0.0001	0.0003										
C1R	0.1869	0.0192	0.0111									
C1S	1.0000	0.0109	1.0000	0.0740								
C3	1.0000	0.0445	0.0066	0.0001	0.3460							
DNASE1L3	0.0007	0.0001	0.0009	0.0002	1.0000	0.0001						
ACP5	1.0000	1.0000	1.0000	0.9900	1.0000	1.0000	1.0000					
AGER	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000				
CYBB	1.0000	1.0000	1.0000	1.0000	0.9721	1.0000	1.0000	1.0000	1.0000			
DNASE1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000		
FAS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
TMEM173	1.0000	1.0000	1.0000	0.0009	1.0000	0.0528	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Figure S3

a

SED

GC

Appendix

E-cadherin CD68 *DNASE1L3* CD45RB

b

SED

GC

Ileum

E-cadherin CD68 *DNASE1L3* CD45RB

c

SED

GC

Colon

E-cadherin CD68 *DNASE1L3* CD45RB

Figure S4

a

mLN

GC

T cell zone

CD68 *DNASE1L3* CD45RB

b

Tonsil

T cell zone

CD68 *DNASE1L3* CD45RB

Figure S5

a Colon

b Appendix

SED
GC
E-cadherin TUNEL CD68 CD45RB

SED
GC
E-cadherin TUNEL CD68 CD45RB

c

C1Qb+ pixels / *DNASE1L3* total

% of overlapping pixels/ total

n=6 follicles
from 2 appendix, 2 colons

d

% of overlapping pixels/ total

n=9 follicles
from 2 appendix, 2 colons

- *C1Qb+* pixels / *C1R* total
- *C1Qb+* pixels / *C3* total
- *C3+* pixels / *C1R* total
- *C1R+* pixels / *C3* total