Human Secretory IgM Emerges from Plasma Cells Clonally Related to Gut Memory B Cells and Targets Highly Diverse Commensals Coated by Secretory IgA

Giuliana Magri[1,11,*], Laura Comerma[1,11], Marc Pybus[1], Jordi Sintes[1], David Lligé[1], Daniel Segura-Garzón[1], Sabrina Bascones[1], Ada Yeste[1], Emilie K. Grasset[2,3], Cindy Gutzeit[2], Mathieu Uzzan[2], Meera Ramanujam[4], Menno C. van Zelm[5], Raquel Albero-González[6], Ivonne Vazquez[6], Mar Iglesias[6,7], Sergi Serrano[6,7], Lucía Márquez[8], Elena Mercade[9], Saurabh Mehandru[2], Andrea Cerutti[1,2,10,12,*]

[1]Program for Inflammatory and Cardiovascular Disorders, Institut Hospital del Mar d'Investigacions Mèdiques (IMIM), Barcelona, 08003, Spain

[2]Department of Medicine, Immunology Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

[3]Department of Medicine, Center for Molecular Medicine at Karolinska University Hospital, Karolinska Institutet, Stockholm, 171 76, Sweden

[4]Immunology and Respiratory Disease Research, Boehringer Ingelheim Pharmaceuticals, Ridgefield, CT 06877, USA

[5]Department of Immunology and Pathology, Central Clinical School, Monash University, Melbourne, Victoria, 3004, Australia

[6]Pathology Department, Hospital del Mar, Barcelona, 08003, Spain

[7]Universitat Autònoma de Barcelona, Barcelona, 08003, Spain.

[8]Department of Gastroenterology, Hospital del Mar Research Institute, Barcelona, 08003, Spain

[9]Microbiology Department of Biology Health and Environment, University of Barcelona, Barcelona, Spain

[10]Catalan Institute for Research and Advanced Studies (ICREA), Barcelona, 08003, Spain

[11]These authors contributed equally

[12]Lead Contact

*Correspondence: gmagri@imim.es (G.M), acerutti@imim.es (A.C)

Summary

Secretory immunoglobulin A (SIgA) enhances host-microbiota symbiosis, whereas SIgM remains poorly understood. We found that gut IgM^+ plasma cells (PCs) were more abundant in humans than mice and clonally related to a large repertoire of memory IgM^+ B cells disseminated throughout the intestine but rare in systemic lymphoid organs. Besides sharing a gut-specific signature with memory IgA^+ B cells, memory IgM^+ B cells interrelated with some IgA^+ clonotypes and switched to IgA in response to T cell-independent or T cell-dependent signals. These signals induced abundant IgM, which recognized mucus-embedded commensals as SIgM from clonally affiliated PCs did. Unlike its murine counterpart, human SIgM recognized bacteria dually coated by SIgA that were characterized by increased richness and diversity compared to IgA-only-coated or uncoated bacteria. Thus, SIgM may emerge from pre-existing memory rather than newly activated naïve IgM^+ B cells and could help SIgA to anchor highly diverse commensal communities to mucus.

Introduction

Complex commensal communities generally referred to as microbiota colonize the gut mucosa soon after birth and have a broad impact on host metabolism, immune system development and gut homeostasis (Kamada et al., 2013). A central element of gut homeostasis is SIgA, an antibody that promotes symbiotic host-microbiota interactions by binding commensals inhabiting the small intestine and, to a lesser extent, the large intestine (Bunker et al., 2015; Kawamoto et al., 2012).

SIgA responses to commensals mostly initiate at gut follicular sites of antigen-entry, including Peyer's patches. At these sites, sampling of commensals by microfold cells, macrophages and dendritic cells promotes a sustained germinal center (GC) reaction involving cognate interaction of B cells with T cells, followed by B cell induction of IgM-to-IgA class switching and affinity maturation through class switch recombination (CSR) and somatic hypermutation (SHM), respectively (Kawamoto et al., 2014). High-affinity and IgA-expressing B cells emerging from GCs up-regulate gut-homing receptors and progressively differentiate into IgA-secreting plasma cells (PC-A), which home to the gut lamina propria (LP) (Macpherson et al., 2008). These PC-A release polymeric IgA, which translocates across epithelial cells to generate intraluminal SIgA that coats mucus-embedded commensals (Kubinak and Round, 2016).

The T cell-dependent (TD) pathway is complemented by a T cell-independent (TI) pathway entailing activation of follicular and possibly extrafollicular B cells by various cells of the innate immune system (Tsuji et al., 2008). Complementary TD and TI responses generate circulating IgA class-switched memory (ME-A) B cells and cooperatively shape the architecture of the microbiota during the development of an individual (Planer et al., 2016). However, the TD pathway may become predominant over time due to continuous accumulation of ME-A B cells (Lindner et al., 2015; Lindner et al., 2012).

ME-A B cells emerge from gut inductive sites along with PC-A and continuously diversify their B cell receptor repertoire via a microbiota-stimulated adaptation process involving induction of SHM and PC-A differentiation in pre-existing GCs (Lindner et al., 2012).

In addition to PC-A, the gut mucosa contains IgM-secreting PCs (PC-M) that release SIgM into the lumen through the polymeric Ig receptor (pIgR) (Macpherson et al., 2008). In mice, SIgM production increases upon induction of colonic damage to prevent lethal dissemination of commensals (Kirkland et al., 2012). This information aside, little is known about SIgM. A better understanding of this mucosal antibody may further elucidate host-microbiota interactions in health and disease states. Indeed, primary antibody deficiency patients selectively lacking SIgA rarely develop inflammatory bowel disease, which is instead more common and very severe in patients lacking both SIgM and SIgA (Agarwal and Mayer, 2009).

We found that human PC-M coexisted with a large but previously unrecognized repertoire of gut-specific memory IgM^+ (ME-M) B cells that were clonally related to PC-M as well as some ME-A B cells and PC-A. Besides inducing IgM-to-IgA CSR, ME-M cells exposed to TD or TI signals secreted copious IgM that targeted mucus-embedded commensals as SIgM from PC-M did. Bacteria recognized by SIgM were dually coated by SIgA and showed increased diversity compared to SIgA-only-coated or uncoated bacteria. Thus, SIgM may emerge from pre-existing ME-M B cells rather than newly activated naïve B cells and could help SIgA to anchor highly diverse microbial communities to gut mucus.

Results

Gut PC-M Are More Abundant in Humans Than Mice

Gut PC-A have been extensively studied, but little is known about gut PC-M, which account for about 10%-20% of all gut PCs in humans (Macpherson et al., 2008). Flow cytometry (FCM) identified PC-M in addition to PC-A in histologically normal resected tissues from terminal ileum and proximal colon samples of individuals undergoing right hemicolectomy due to malignancy, polyps or angiodysplasia (**Figure 1A** and **S1A**). PC-M were consistently more abundant in the ileum compared to the colon (**Figure 1A,** and **1B**) and expressed a phenotype similar to that of class-switched PCs (PC-SW), which included PC-A in addition to some PC-G/E (**Figure 1C** and **S1B**). Compared to human gut naïve B cells, PC-M showed increased expression of CD138, CCR10 and/or CCR9. Up-regulation of these PC-associated molecules was coupled with down-regulation of HLA-DR, CD20, and the follicle-associated chemokine receptors CCR7 and CXCR4. Notwithstanding their resemblance to PC-M, PC-SW expressed more CD138, which could reflect a more advanced maturation stage (Nutt et al., 2015) (**Figure 1C** and **S1B**).

Quantitative real-time PCRs (qRT-PCRs) further determined that, compared to human intestinal naïve B cells, PC-M contained more transcripts for BLIMP-1 and IRF4, two transcription factors required for PC differentiation (Nutt et al., 2015). In addition, PC-M contained more transcripts for BCL-2, an anti-apoptotic protein that may sustain human PC survival in the intestine (Nair et al., 2016) (**Figure S1C**). Of note, immunofluorescence analysis (IFA) followed by tissue-based cell counting demonstrated that, compared to the human ileum LP, the small intestinal LP from wild type C57BL/6 mice featured fewer PC-M accumulating intracellular IgM and a lower PC-M/PC-A ratio (**Figure 1D** and **1E**). FCM confirmed that the frequency of PC-M was negligible compared to that of PC-A in the mouse small and large intestines (**Figure S1D** and **S1E**).

Then, we further verified whether circulating PC-M expressed gut-homing properties. Compared to PC-A, a larger fraction of circulating PC-M co-expressed β7 and CCR9, whereas PC-G/E showed

little or no β7 and CCR9 (**Figure 1F** and **1G**). Thus, PC-M predominantly inhabit the human small intestine, are rare in the mouse small intestine, resemble gut PC-A, and include a circulating fraction expressing gut-homing receptors.

Gut PC-M Coexist with a Large Repertoire of ME-M B Cells

Considering that gut PC-A emerge from ME-A B cells diversifying at gut inductive sites (Lindner et al., 2015), we hypothesized that PC-M coexisted with ME-M B cells in the human gut and followed published gating strategies and comparable isolation procedures to segregate $IgM^+IgD^-CD27^+CD38^-$ ME-M B cells from $IgM^+IgD^{high}CD27^-CD38^-$ naïve, $IgM^+IgD^{low}CD27^+CD38^-$ marginal zone (MZ) and $IgM^-IgD^-CD27^+CD38^-$ class-switched memory (ME-SW) B cells (Berkowska et al., 2011; Descatoire et al., 2014; Klein et al., 1998). FCM showed that ME-M B cells were abundant in ileum and colon but rare in blood, spleen and tonsil (**Figure 1H** and **1I**). Consistent with the prevailing localization of PC-M in the ileum, ME-M B cells were enriched in the ileum compared to the colon and showed a phenotype similar to that of ileal ME-SW B cells and splenic MZ B cells (**Figure 1H-J** and **S1F**).

Next, we dissected the intestinal geography of ME-M B cells by tissue IFA. Tissue-based cell counting revealed that ME-M B cells were rare in the gut LP and comparably abundant in ILFs with or without GC (**Figure 2A** and **2B**). Consistently, gut ME-M B cells expressed follicle-retaining CCR7 and CXCR4 receptors as much as gut naïve and ME-SW B cells did (**Figure 2C**). However, kappa-deleting recombination excision circle assays (van Zelm et al., 2007) detected molecular fingerprints of GC proliferation in gut ME-M and ME-SW but not naïve B cells (**Figure 2D**).

Given that ME-A B cells up-regulate α4β7 and CCR9 gut-homing receptors as they recirculate from gut follicular inductive sites to the gut LP (Nair et al., 2016), we further determined β7 and CCR9 expression by circulating ME-M B cells. Compared to circulating ME-A B cells and MZ B cells, a larger fraction of circulating ME-M B cells co-expressed β7 and CCR9, whereas naïve B cells and ME-G/E B cells showed little or no β7 and CCR9 co-expression (**Figure 2E** and **2F**).

We then ascertained whether the gut mucosa generated ME-M B cells at an early age. IFA identified human intestinal ME-M B cells together with PC-M and PC-A as early as 1.5 months of age (**Figure 2G**). Tissue-based cell counting and FCM indicated that ME-M B cells remained stable over time, whereas PC-M increased through the first ten years of life (**Figure S1G** and **S1H**). Thus, human gut PC-M co-exist with a large and stable repertoire of ME-M B cells that emerge early in life, predominantly inhabit gut follicles and include a circulating fraction expressing gut-homing receptors.

Gut ME-M B Cells Express a Gene Signature Reflecting Antigen Experience

Human intestinal ME-M B cells were further characterized through transcriptomics. These studies were preceded by a morphological analysis that confirmed the specificity of our sorting procedures (**Figure 3A**).

The global gene expression profile of ME-M B cells was elucidated through cross-comparative strategies involving memory B cell subsets from gut or spleen, including splenic MZ B cells. In this approach, each subset was compared to tissue-specific naïve B cells. Supervised hierarchical clustering and Robust Multi-Array Average expression analysis indicated that ME-M B cells expressed a transcriptome distinct from that of naïve B cells, but similar to that of ME-SW B cells, irrespective of the tissue of origin (**Figure 3B**, and **S2A**). Venn diagrams further determined that gut ME-M B cells expressed a common memory signature entailing 215 transcripts, which were also differentially expressed by gut ME-SW B cells as well as splenic ME-SW and MZ B cells (**Figure S2B**). This common memory gene signature included increased expression of TACI (*TNFRSF13B)* along with decreased expression of IgD (*IGHD*), CD72 and the transcriptional suppressors of PC differentiation FOXO1, BTLA and BACH2 (**Figure S2C** and **S2D**) (Kurosaki et al., 2015; Nutt et al., 2015). Thus, human gut ME-M B cells express a common memory gene signature reflecting increased propensity to undergo activation, proliferation and PC differentiation.

Gut ME-M B Cells Express a Tissue-Specific Gene Signature

As shown by unsupervised hierarchical gene clustering and principal component analysis, gut ME-M and ME-SW B cells also expressed a tissue-specific signature. Indeed, these B cells clustered together but away from splenic ME-SW and MZ B cells, whereas naïve B cells grouped together independently of their tissue of origin (**Figure 3B**, **3C, S2E** and **S2F**). Accordingly, Volcano Plot, Venn diagram and heat map diagrams coupled with qRT-PCR showed 305 transcripts differentially expressed by gut ME-M and ME-SW B cells but not splenic ME-SW and MZ B cells compared to naïve B cells (**Figure 3D-F** and **S3A**).

This tissue-specific memory signature included increased expression of transcripts for 1) activation-induced receptors such as FcRL4, CD11c (*ITGAX*) and *SIGLEC6*; 2) IgA response-related transcription factors such as RUNX2 and *RORA*; 3) PC-inducing molecules such as IL-10, *IL10RA*, CD70 and the transcription factor *ZBTB32*; and 4) epithelium-targeting chemokine receptors such as CCR1, CCR2 and CCR9 (**Figure 3D-F** and **S3A**) (Ehrhardt et al., 2005; Jash et al., 2016; Nair et al., 2016; Rubtsov et al., 2015; Wang et al., 2012; Watanabe et al., 2010). Finally, Ingenuity's Upstream Regulator Analysis indicated that, compared to splenic ME-SW and MZ B cells, gut ME-M and ME-SW B cells expressed functional gene sets predicting enhanced signaling from IL-2, IL-5, IL-6, IL-10, IL-15 and GM-CSF pathways (**Figure 3G**), which are linked to B cell activation and PC differentiation (Nutt et al., 2015).

Consistent with these data, FCM showed that gut ME-M and ME-SW B cells expressed activation traits encompassing up-regulation of FcRL4, CD43 and CD11c combined with down-regulation of CD62 ligand expression (**Figure 3H**). However, compared to gut ME-SW B cells, gut ME-M B cells also showed enhanced MZ-like traits (**Figure 3I**), including increased expression of the complement receptors CD21, CD35, CD1c, and the adhesion molecule CD66a (Descatoire et al., 2014; Seifert et al., 2015). Notwithstanding these similarities, supervised gene expression analysis,

FCM and Gene Set Enrichment Analysis (GSEA) indicated that gut ME-M B cells were distinct from splenic MZ B cells (**Figure S3B-E**).

Indeed, gut ME-M B cells expressed less CD84 (**Figure S3B**), a member of the SLAM family highly expressed by innate-like lymphocytes (Sintes et al., 2010). Furthermore, gut ME-M B cells expressed more gene products implicated in IL-2 and IL-6 receptor signaling via STAT proteins, but less gene products implicated in NOTCH signaling (**Figure S3C-E**). Thus, human gut ME-M B cells express a tissue-specific memory gene signature reflecting increased immune activation and PC but not MZ differentiation.

Gut ME-M B Cells Are Clonally Related to PC-M and Some PC-A

We next characterized the Ig heavy chain variable (IGHV) and joining (IGHJ) gene repertoires of gut ME-M B cells from paired ileum and colon samples to broadly determine their degree of similarity and diversity with naïve B cells and PC-M as well as class-switched ME-A B cells and PC-A. Pearson's correlation coefficient analysis of IGHV gene usage indicated that gut ME B cell and PC subsets differed from gut naïve B cells but hierarchically clustered with each other based on the expressed isotype, tissue of origin and cell type (**Figure 4A** and **S4A**). The antigen-driven IGHV gene reconfiguration of gut ME B cell and PC subsets was further inferred from their negative selection of IGHV1-18, IGHV1-69, IGHV4-34 and IGHJ6 genes (**Figure 4B** and **S4A**) and positive selection of IGHV3-7, IGHV3-23 and IGHJ4 genes (**Figure S4A** and **S4B**).

Compared to gut naïve B cells, gut ME B cell and PC subsets expressed mutated IGHV genes encoding antigen-binding variable regions with shorter H-CDR3 average length (**Figure 4C**, **S4C** and **S4D**), two additional hallmarks of antigen-driven selection (Lindner et al., 2015; Lindner et al., 2012; Tipton et al., 2015). In addition to these commonalities, gut ME B cell and PC subsets showed significant differences. In particular, ME-M B cells and PC-M featured fewer IGHV gene mutations and shorter H-CDR3 compared to ME-A B cells and PC-A from ileum but not colon (**Figure 4C**, **S4C** and **S4D**). Moreover, ME-M B cells expressed more Igλ compared to ME-A B

cells from ileum but not colon (**Figure 4D**). These differences could reflect the involvement of ME-M B cells in antigen-driven selection and differentiation programs distinct from those regulating ME-A B cells.

We then investigated the dynamics of the inferred gut antibody repertoire from each donor by collapsing all clones into clonal families and comparing the resulting "core repertoire" to the original "expanded repertoire". Clonal families expressing both IgM and IgA or inhabiting both ileum and colon were expanded compared to the core repertoire, whereas clonal families expressing only IgM or IgA or inhabiting either ileum or colon were not. These observations indicate that clonal expansion is coupled with increased IgM-to-IgA class switching and intestinal dissemination (**Figure S4E**). We next visualized the ontogenetic affiliation of gut ME-M B cells with other ME B cell and PC subsets through circos plots. While most of these clonally related ME-M B cells were linked to PC-M inhabiting identical or distinct gut segments, fewer but large ME-M clonotypes were linked to ME-A B cells and/or PC-A (**Figure 4E**, **4F**, **S5A** and **S5B**).

Clonotypic affiliations were further corroborated through the calculation of the Morisita-Horn overlap index, which ascribes 0 and 1 values to unrelated and identical sequences, respectively (Lindner et al., 2015). Ileal ME-M clonotypes showed very robust relatedness with ileal PC-M clonotypes and less robust relatedness with all other subset clonotypes (**Figure 4G** and **S5C**). Further dissection of clonal families through lineage tree reconstruction analysis suggested that some ileal ME-M B cells generate colonic PC-M or ileal ME-A B cells upon re-entering GC pathways that induced SHM with or without CSR (**Figure 4H** and **4I**). Additional ileal ME-M B cells from ileum may generate colonic PC-A by entering extra-GC pathways inducing CSR but not SHM (**Figure 4J**). Accordingly, tissue IFA detected B cells expressing the CSR/SHM-inducing enzyme AID and the proliferation molecule Ki-67 in both GC and extra-GC areas from ileum, including ILFs (**Figure 4K**).

The presence of IgM-to-IgA CSR in some gut ME-M B cells was confirmed by RT-PCR and DNA sequencing studies, which detected germline Iα1-Cα1 and Iα2-Cα2 as well as switch circle Iα1/2-Cμ transcripts in ME-M but not control naïve B cells from human ileum (**Figure 4L** and **S5D**). Ileal GC B cells predictably contained Iα1-Cα1 and Iα1/2-Cμ, whereas PC-M unexpectedly contained Iα1/2-Cμ and Iα2-Cα2 but lacked Iα1-Cα1 (**Figure 4L**). These PC-M may emerge from recently activated ME-M clones concomitantly receiving IgM-to-IgA CSR-inducing signals. Thus, human ME-M B cells disseminate to both ileum and colon and may differentiate to PC-M and class-switched PC-A by re-entering GCs or progressing through extra-GCs pathways.

Gut ME-M B Cells Secrete IgM and Switch to IgA Upon TD or TI Stimulation

The differentiation potential of human gut ME-M B cells was further explored by evaluating their proliferation, CSR and PC differentiation upon exposure to TD (CD40L and IL21) or TI (CpG DNA combined or not with BAFF and APRIL) signals. IL-10 was supplemented to maximize CSR and PC differentiation (Macpherson et al., 2008).

As shown by FCM-based CFSE dilution assays, TD signals induced comparable proliferation of gut ME-M and naïve B cells, whereas TI signals induced proliferation of gut ME-M but not naïve B cells, particularly in the presence of BAFF and APRIL (**Figure 5A** and **5B**). Additional FCM assays showed that ME-M B cells differentiated into proliferating $CD38^{high}CFSE^{low}$ plasmablasts in response to either TD or TI signals, whereas naïve B cells did so only in response to TD signals (**Figure 5C**). Of note, a sizable fraction of plasmablasts emerging from activated ME-M but not naïve B cells expressed IgA but lacked IgM (**Figure 5D**), a hallmark of IgM-to-IgA CSR.

Consistent with these data, ME-M B cells secreted copious IgM and less abundant IgA in response to TD or TI signals, whereas naïve B cells showed weaker IgM and IgA responses to TD but not TI signals (**Figure 5E**). We then wondered whether CSR targeted gut ME-M B cells expressing FcRL4, which defines activated tissue-based memory B cells (Ehrhardt et al., 2005). Indeed, TD signals generated plasmablasts from both $FcRL4^{-}$ and $FcRL4^{+}$ ME-M B cells, but induced IgM-to-

IgA CSR only in FcRL4$^+$ ME-M B cells (**Figure 5F**). Accordingly, gut FcRL4$^+$ but not FcRL4$^-$ ME-M B cells contained transcripts for the CSR (and SHM)-inducing enzyme AID (**Figure 5G**). Thus, human gut ME-M B cells proliferate and generate PC-M in response to TD or TI signals, which further induce IgM-to-IgA CSR in the FcRL4$^+$ fraction of ME-M B cells.

Gut ME-M B Cells and PC-M Release IgM to Mucus-Embedded Commensals

We next developed an EBV-based protocol to determine whether human gut ME-M B cells produced IgM to autologous mucus-embedded bacteria (**Figure 6A**). Initial ELISAs showed that IgM from ileal ME-M B cells recognized commensal antigens such as phosphorylcholine, β-glucan, laminarin, galactose-α-1,3-galactose and capsular polysaccharides, whereas IgM from ileal naïve B cells did not (**Figure 6B** and **S6A**). Further FCM assays showed that IgM from ileal ME-M B cells recognized mucus-embedded commensals more efficiently than IgM from ileal naïve B cells did (**Figure 6C**).

We then established an FCM-based strategy to quantify binding of PC-derived SIgM to mucus-embedded microbiota from paired human ileum and colon samples (**Figure 6D**, **S6B** and **S6C**). Initial ELISAs detected free SIgM in mucus, though in lesser amounts than free SIgA (**Figure 6E**). FCM determined that most mucus samples included significant SIgA$^+$SIgM$^+$, SIgA$^+$SIgM$^-$ and SIgA$^-$SIgM$^-$ but negligible SIgA$^-$SIgM$^+$ microbiota fractions (**Figure 6F**). The frequency of these fractions was variable among individuals and between ileum and colon from the same individual. Consistent with the virtual lack of PC-M in the murine gut and published results (Bunker et al., 2015), the microbiota from the small and large intestines of wild type mice included IgA$^-$SIgM$^-$ and SIgA$^+$SIgM$^-$ but not SIgA$^+$SIgM$^+$ bacterial fractions (**Figure 6G**). Enhancing gut microbiota complexity by housing mice outside the specific pathogen-free (SPF) barrier increased neither SIgA$^+$SIgM$^+$ bacteria nor PC-M nor PC-A, but did increase SIgA$^+$SIgM$^-$ bacteria (**Figure S6D** and **S6E**). Thus, human gut ME-M B cells recognize commensals as clonally related PC-M do. These

latter generate homeostatic SIgM responses that target SIgA-coated bacteria in humans but not wild type mice.

SIgM from Gut PC-M Binds Highly Diverse Commensals Dually Coated by SIgA

We next devised a strategy to profile mucus-embedded bacteria from human ileum or colon by 16S ribosomal RNA (rRNA) gene sequencing and found inter-individual variability at phylum and genus levels (**Figures S7A-C**). Differences in bacterial composition between ileum and colon from the same donor were less pronounced than between donors, as shown by unsupervised hierarchical clustering (**Figure S7C**). Rarefaction plots calculating Faith's phylogenetic diversity and Shannon Index (H') confirmed that phylogenetic richness and species diversity varied among donors but not between tissues (**Figure S7D** and **S7E**).

To comparatively profile $SIgA^{+}SIgM^{+}$, $SIgA^{+}SIgM^{-}$ and $SIgA^{-}SIgM^{-}$ fractions of mucus-embedded bacteria in a comprehensive and unbiased manner, we combined SIgM/A-based sorting with 16S rRNA gene sequencing (**Figure 7A**). The composition of each fraction varied at both phylum and genus levels, though all fractions showed more Bacteroidetes and Firmicutes than Proteobacteria and Actinobacteria (**Figure 7B** and **7C**). Also phylogenetic richness and microbial species diversity varied among fractions, with an overall decrease from $SIgA^{+}SIgM^{+}$ to $SIgA^{+}SIgM^{-}$ and $SIgA^{-}SIgM^{-}$ fractions (**Figure 7D**). This finding correlated with differences in phylum composition, including proportionally fewer Bacteroidetes but more Firmicutes in $SIgA^{+}SIgM^{+}$ compared to $SIgA^{+}SIgM^{-}$ and $SIgA^{-}SIgM^{-}$ fractions (**Figure 7E**).

To identify microbial species accounting for the above phylum differences at the OTU level, we used a log-based enrichment index. A hierarchical clustering algorithm applied to a conservative selection of OTU-based enrichment indexes showed that $SIgA^{+}SIgM^{+}$ and $SIgA^{+}SIgM^{-}$ fractions grouped together, separately from the $SIgA^{-}SIgM^{-}$ fraction (**Figure 7F**). Of twelve OTUs showing a significantly different enrichment index, seven *Lachnospiraceae* and *Ruminococcaceae* were enriched in $SIgA^{+}SIgM^{+}$ compared to $SIgA^{-}SIgM^{-}$ bacteria (**Figure 7F** and **S7F**). Accordingly,

FCM-based coating assays executed by incubating a selected set of culturable commensals with graded amounts of IgM from four independent EBV-transformed ME-M B cell lines determined that gut IgM strongly bound Firmicutes such as *B. cereus*, *R. intestinalis* (belonging to *Lachnospiraceae*) and *R. lactatiformans* (belonging to *Ruminococcaceae*). In addition, IgM showed elevated binding to Bacteroidetes such as *B. vulgatus*, but little or no binding to other Bacteroidetes or Proteobacteria such as *B. fragilis*, *B. thetaiotamicron* and *E. coli* (**Figure S7G**). Thus, human SIgM may cooperate with SIgA to implement mucus retention of diverse microbial communities, including Firmicutes with putative beneficial functions.

Discussion

We have shown that human gut PC-M clonally related to a large and previously unrecognized repertoire of ME-M B cells that predominantly inhabited gut-associated follicles. Besides undergoing IgM-to-IgA CSR in response to TD or TI signals, gut ME-M B cells secreted abundant IgM, which recognized mucus-embedded commensals as SIgM from PC-M did. Of note, SIgM-coated bacteria were dually targeted by SIgA and showed increased diversity and distinct composition compared to uncoated or SIgA-only-coated bacteria. Thus, SIgM may help SIgA to anchor non-redundant microbial communities to mucus.

The key role of SIgA in gut homeostasis can be inferred from the emergence of dysbiosis in mice lacking B cells, IgA, AID or pIgR (Kubinak and Round, 2016). In addition to dysbiosis, patients with antibody deficiency can develop gut inflammation, including inflammatory bowel disease (Agarwal and Mayer, 2009). This complication is more frequent in common variable immunodeficiency cases with combined SIgM and SIgA depletion (Agarwal and Mayer, 2009), suggesting that human gut homeostasis requires microbiota targeting by both SIgM and SIgA. Accordingly, we found that PC-M accumulated in the human but not mouse gut mucosa and further demonstrated that SIgM coated human but not mouse gut bacteria in combination with SIgA.

Remarkably, human gut PC-M established extensive clonal relationships with a large repertoire of gut ME-M B cells that were rare in systemic or mucosal extra-intestinal lymphoid organs, including spleen and tonsils. The prominent gut tropism of ME-M B cells was further indicated by studies showing robust $\alpha4\beta7$ and CCR9 co-expression on a large fraction of circulating ME-M B cells and PC-M, which exceeded circulating ME-A B cells and PC-A expressing identical gut-homing receptors. Of note, $\alpha4\beta7$ and CCR9 induction mostly occurs in lymphoid structures from the small intestine and promotes migration of gut ME-A B cells and immature PC-A to the small intestinal LP (Macpherson et al., 2008). Accordingly, gut ME-M B cells predominantly inhabited Peyer's patches and ILFs from the small intestinal mucosa, whereas PC-M mostly accumulated in the small

intestinal LP. Similar to PC-A, gut ME-M B cells and PC-M became detectable as early as 1.5 months after birth. While PC-M further accumulated over the first ten years of life, ME-M B cells remained numerically stable over time. These results suggest that SIgM may shape the microbiota of a developing individual in cooperation with SIgA (Planer et al., 2016).

Our identification of clonally related ME-M B cells and PC-M in the human gut extends evidence from mouse systemic immunization models indicating that humoral memory is not merely comprised of ME-G and ME-A B cells, but further extends to ME-M B cells (Dogan et al., 2009; Kurosaki et al., 2015; Pape et al., 2011). Besides expressing canonical memory molecules such as CD24, CD27 and CD148, human gut ME-M B cells featured post-GC expression of mutated IGHV genes and negative selection of IGHV1-69, IGHV4-34 and IGHJ6 genes, which encode antibodies enriched in self-reactivity (Tipton et al., 2015). Furthermore, some ME-M B cells showed clonal properties consistent with re-entry into GC pathways promoting SHM in addition to PC-M differentiation.

Diversification of human gut PC-M from pre-existing memory specificities echoes works showing homeostatic or immunization-induced diversification of gut ME-A B cells in GC from Peyer's patches (Bemark et al., 2016; Lindner et al., 2015). In addition to PC-M, human gut ME-M B cells generated some ME-A B cells and PC-A by entering either GC pathways coupled with SHM and CSR or GC-independent pathways promoting CSR but not SHM. This conclusion was supported by lineage tree reconstruction analysis of high-throughput IGHV gene sequencing data, detection of AID in activated FcRL4^{+} gut ME-M B cells responsive to IgM-to-IgA CSR-inducing signals, identification of IgM-to-IgA CSR in unfractioned ME-M B cells, and detection of AID in B cells from both GC and extra-GC areas.

In mice, systemic ME-G and ME-M B cells were thought to rapidly induce PC-G or a secondary GC reaction upon re-exposure to antigen, respectively (Dogan et al., 2009; Pape et al., 2011). This view has been recently modified by mouse studies indicating that systemic ME-M B cells can

rapidly differentiate into IgG class-switched plasmablasts in response to TI or TD signals (Krishnamurty et al., 2016; Zuccarino-Catania et al., 2014). Accordingly, human intestinal ME-M B cells progressed along intertwined GC-dependent and GC-independent pathways that promoted IgM-to-IgA CSR in addition to PC differentiation. Compared to gut naïve B cells, which showed weaker and TD-biased antibody responses, gut ME-M B cells comparably induced plasmablasts secreting IgM or IgA in response to either TD or TI signals. Of note, similar signals activate rotavirus-specific gut ME-M B cells (Narvaez et al., 2012). Local TD and TI cues could further imprint ME-M B cells with a tissue-specific memory signature similar to that described in tonsillar ME-G/A B cells (Ehrhardt et al., 2005). This signature included CD11c, Siglec-6, CCR9, IL-10, IL-10Rα and RORα up-regulation, which reflects non-inflammatory activation and mucosal homing.

Gut ME-M B cells also expressed a common memory signature shared with gut and splenic ME-SW B cells. This signature included FcRL4 and TACI up-regulation, which suggests increased activation by TI signals from TLR ligands and BAFF/APRIL (He et al., 2010; Sohn et al., 2011), as well as CD72, FOXO1, BACH2 and BTLA4 down-regulation, which reflects increased propensity to undergo proliferation and PC differentiation (Kurosaki et al., 2015). Similar to splenic MZ B cells, gut ME-M B cells expressed more CD1c, CD21, CD35 and CD66a compared to gut ME-SW B cells, a finding reminiscent of recent studies suggesting that splenic MZ B cells originate from gut precursors (Vossenkamper et al., 2013). However, the global transcriptome of gut ME-M B cells clustered separately from that of splenic MZ B cells. Compared to these latter, gut ME-M B cell expressed more CD84 and contained more gene products linked to PC-inducing IL-2 and IL-6 signaling, but fewer gene products linked to NOTCH signaling, which is required for MZ B cell differentiation (Descatoire et al., 2014).

Besides establishing affiliations with IgM^+ and IgA^+ clonotypes, ME-M B cells formed a core repertoire of clonally organized families that emerged early in life and did not show age-dependent accumulation. Thus, ME-M B cells may form a stable but functionally plastic pool of “immune

sentinels" within sites of antigen entry, such as ILFs. These sites contained two parallel repertoires of unswitched ME-M and naïve B cells expressing follicle-targeting CCR7 and CXCR4 receptors. As indicated by their unique phenotypic, molecular, transcriptional and functional traits, these B cell subsets may offer unique solutions to different problems. While ME-M B cells may initiate quickly developing SIgM and SIgA responses to rapidly match transient changes of the microbiota, naïve B cells may induce *de novo* SIgM and SIgA responses to counter more durable microbial perturbations. Consistent with their involvement in the homeostatic control of commensals, ME-M B cells produced IgM to mucus-embedded bacteria as SIgM from clonally related PC-M did.

The presence of homeostatic SIgM responses in humans but not mice may reflect the lower complexity of the mouse gut microbiota (Kamada et al., 2013). However, neither PC-M nor SIgM-coated bacteria increased in non-SPF mice harboring a more complex microbiota, suggesting that B cell-intrinsic differences also play a role. Accordingly, ME-M B cells from orally immunized mice have been shown to colonize the spleen and bone marrow but not gut follicles (Bemark et al., 2016). Remarkably, SIgM coated a fraction of the human microbiota that was also targeted by SIgA. Compared to uncoated or SIgA-only-coated bacteria, bacteria dually coated by SIgM and SIgA showed increased diversity, a parameter linked to gut homeostasis (Kamada et al., 2013). By activating complement (Kirkland et al., 2012), SIgM may constrain the growth of common commensals, thereby helping SIgA to optimize microbiota diversity (Kubinak and Round, 2016). Accordingly, SIgA deficiency causes dysbiosis and inflammatory bowel disease when combined with SIgM deficiency, as in patients with common variable immunodeficiency (Agarwal and Mayer, 2009).

Compared to uncoated bacteria, bacteria dually coated by SIgM and SIgA as well as bacteria coated by SIgA alone were enriched in Firmicutes, including *Lachnospiraceae* and *Ruminococcaceae*. By degrading dietary polysaccharides into short-chain fatty acids with immunoregulatory and SIgA-inducing functions (Arpaia et al., 2013; Kim et al., 2016), these commensals may enhance

protection against obesity and infection (Biddle et al., 2013; Cho et al., 2012; Petrof et al., 2013). Should this be the case, PC-M and clonally related PC-A may maximize gut retention of non-redundant microbial consortia through a mechanism involving mucus interaction with the pIgR-derived secretory fragment of SIgM and SIgA (Macpherson et al., 2008). This dual anchoring strategy may have evolved to preserve microbiota homeostasis under common pathological conditions selectively depleting SIgA. Similar to a large fraction of gut IgA (Benckert et al., 2011), gut IgM showed evidence of polyreactivity, because it targeted common microbial products and some Bacteroidetes such as *B. vulgatus* in addition to Firmicutes. In summary, SIgM may emerge from pre-existing memory rather than newly activated naïve IgM^+ B cells and could help SIgA to select, control and retain highly diverse and putatively beneficial commensal communities within gut mucus.

Author Contributions

G.M. and L.C. designed and performed experiments, analyzed results, discussed data and wrote the manuscript; M.P. designed experiments, analyzed and discussed sequencing data and wrote the manuscript; J.S. performed research, analyzed results and discussed data; D.L., D.S.G., S.B., A.Y. and M.U. performed experiments; C.G., E.K.G. and M.R. discussed the data; M.V.Z. and E.M provided reagents and discussed data; R.A., I.V., M.I., S.S. and L.M. provided tissue and microbiota samples; S.M designed experiments and discussed data; A.C. designed experiments, discussed data and wrote the manuscript.

Acknowledgments

We thank the Microarray Analysis Service from IMIM for global gene transcriptome analysis, the Genomic Core Facility from UPF and Christopher M. Tipton from Emory University for help with NGS, and the FACS Unit from UPF and CRG for help with cell sorting. Supported by European Advanced Grant (ERC-2011-ADG-20110310), MINECO (SAF2014-52483-R), AGAUR (2014 SGR 832), US NIH grants P01 AI61093, R01 AI57653 and U01 AI95613, Boeringher grant 0249-2537 (to A.C.), NIH grant R01 DK 112296-01 (to A.C. and S.M.), and Fondo de Investigación Sanitaria ISCIII (to G.M and L.C.).

References

Agarwal, S., and Mayer, L. (2009). Pathogenesis and treatment of gastrointestinal disease in antibody deficiency syndromes. J Allergy Clin Immunol *124*, 658-664.

Arpaia, N., Campbell, C., Fan, X., Dikiy, S., van der Veeken, J., deRoos, P., Liu, H., Cross, J.R., Pfeffer, K., Coffer, P.J., and Rudensky, A.Y. (2013). Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. Nature *504*, 451-455.

Barak, M., Zuckerman, N.S., Edelman, H., Unger, R., and Mehr, R. (2008). IgTree: creating Immunoglobulin variable region gene lineage trees. J Immunol Methods *338*, 67-74.

Bemark, M., Hazanov, H., Stromberg, A., Komban, R., Holmqvist, J., Koster, S., Mattsson, J., Sikora, P., Mehr, R., and Lycke, N.Y. (2016). Limited clonal relatedness between gut IgA plasma cells and memory B cells after oral immunization. Nat Commun *7*, 12698.

Benckert, J., Schmolka, N., Kreschel, C., Zoller, M.J., Sturm, A., Wiedenmann, B., and Wardemann, H. (2011). The majority of intestinal IgA+ and IgG+ plasmablasts in the human gut are antigen-specific. J Clin Invest *121*, 1946-1955.

Berkowska, M.A., Driessen, G.J., Bikos, V., Grosserichter-Wagener, C., Stamatopoulos, K., Cerutti, A., He, B., Biermann, K., Lange, J.F., van der Burg, M., *et al.* (2011). Human memory B cells originate from three distinct germinal center-dependent and -independent maturation pathways. Blood *118*, 2150-2158.

Biddle, A., Stewart, L., Blanchard, J., and Leschine, S. (2013). Untangling the genetic basis of fibrolytic specialization by Lachnospiraceae and Ruminococcaceae in diverse gut communities. Diversity *5*, 627-640.

Bunker, J.J., Flynn, T.M., Koval, J.C., Shaw, D.G., Meisel, M., McDonald, B.D., Ishizuka, I.E., Dent, A.L., Wilson, P.C., Jabri, B., *et al.* (2015). Innate and Adaptive Humoral Responses Coat Distinct Commensal Bacteria with Immunoglobulin A. Immunity *43*, 541-553.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I., *et al.* (2010). QIIME allows analysis of high-throughput community sequencing data. Nat Methods *7*, 335-336.

Cho, I., Yamanishi, S., Cox, L., Methe, B.A., Zavadil, J., Li, K., Gao, Z., Mahana, D., Raju, K., Teitler, I., *et al.* (2012). Antibiotics in early life alter the murine colonic microbiome and adiposity. Nature *488*, 621-626.

Descatoire, M., Weller, S., Irtan, S., Sarnacki, S., Feuillard, J., Storck, S., Guiochon-Mantel, A., Bouligand, J., Morali, A., Cohen, J., *et al.* (2014). Identification of a human splenic marginal zone B cell precursor with NOTCH2-dependent differentiation properties. J Exp Med *211*, 987-1000.

Dogan, I., Bertocci, B., Vilmont, V., Delbos, F., Megret, J., Storck, S., Reynaud, C.A., and Weill, J.C. (2009). Multiple layers of B cell memory with different effector functions. Nat Immunol *10*, 1292-1299.

Ehrhardt, G.R., Hsu, J.T., Gartland, L., Leu, C.M., Zhang, S., Davis, R.S., and Cooper, M.D. (2005). Expression of the immunoregulatory molecule FcRH4 defines a distinctive tissue-based population of memory B cells. J Exp Med *202*, 783-791.

Gupta, N.T., Vander Heiden, J.A., Uduman, M., Gadala-Maria, D., Yaari, G., and Kleinstein, S.H. (2015). Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. Bioinformatics *31*, 3356-3358.

He, B., Santamaria, R., Xu, W., Cols, M., Chen, K., Puga, I., Shan, M., Xiong, H., Bussel, J.B., Chiu, A., *et al.* (2010). The transmembrane activator TACI triggers immunoglobulin class switching by activating B cells through the adaptor MyD88. Nat Immunol *11*, 836-845.

Jash, A., Wang, Y., Weisel, F.J., Scharer, C.D., Boss, J.M., Shlomchik, M.J., and Bhattacharya, D. (2016). ZBTB32 Restricts the Duration of Memory B Cell Recall Responses. J Immunol *197*, 1159-1168.

Kamada, N., Seo, S.U., Chen, G.Y., and Nunez, G. (2013). Role of the gut microbiota in immunity and inflammatory disease. Nat Rev Immunol *13*, 321-335.

Kawamoto, S., Maruya, M., Kato, L.M., Suda, W., Atarashi, K., Doi, Y., Tsutsui, Y., Qin, H., Honda, K., Okada, T., *et al.* (2014). Foxp3(+) T cells regulate immunoglobulin a selection and facilitate diversification of bacterial species responsible for immune homeostasis. Immunity *41*, 152-165.

Kawamoto, S., Tran, T.H., Maruya, M., Suzuki, K., Doi, Y., Tsutsui, Y., Kato, L.M., and Fagarasan, S. (2012). The inhibitory receptor PD-1 regulates IgA selection and bacterial composition in the gut. Science *336*, 485-489.

Kim, M., Qie, Y., Park, J., and Kim, C.H. (2016). Gut Microbial Metabolites Fuel Host Antibody Responses. Cell Host Microbe *20*, 202-214.

Kirkland, D., Benson, A., Mirpuri, J., Pifer, R., Hou, B., DeFranco, A.L., and Yarovinsky, F. (2012). B cell-intrinsic MyD88 signaling prevents the lethal dissemination of commensal bacteria during colonic damage. Immunity *36*, 228-238.

Klein, U., Rajewsky, K., and Kuppers, R. (1998). Human immunoglobulin (Ig)M+IgD+ peripheral blood B cells expressing the CD27 cell surface antigen carry somatically mutated variable region genes: CD27 as a general marker for somatically mutated (memory) B cells. J Exp Med *188*, 1679-1689.

Krishnamurty, A.T., Thouvenel, C.D., Portugal, S., Keitany, G.J., Kim, K.S., Holder, A., Crompton, P.D., Rawlings, D.J., and Pepper, M. (2016). Somatically Hypermutated Plasmodium-Specific IgM(+) Memory B Cells Are Rapid, Plastic, Early Responders upon Malaria Rechallenge. Immunity *45*, 402-414.

Kubinak, J.L., and Round, J.L. (2016). Do antibodies select a healthy microbiota? Nat Rev Immunol *16*, 767-774.

Kurosaki, T., Kometani, K., and Ise, W. (2015). Memory B cells. Nat Rev Immunol *15*, 149-159.

Lindner, C., Thomsen, I., Wahl, B., Ugur, M., Sethi, M.K., Friedrichsen, M., Smoczek, A., Ott, S., Baumann, U., Suerbaum, S., *et al.* (2015). Diversification of memory B cells drives the continuous adaptation of secretory antibodies to gut microbiota. Nat Immunol *16*, 880-888.

Lindner, C., Wahl, B., Fohse, L., Suerbaum, S., Macpherson, A.J., Prinz, I., and Pabst, O. (2012). Age, microbiota, and T cells shape diverse individual IgA repertoires in the intestine. J Exp Med *209*, 365-377.

Macpherson, A.J., McCoy, K.D., Johansen, F.E., and Brandtzaeg, P. (2008). The immune geography of IgA induction and function. Mucosal Immunol *1*, 11-22.

Nair, N., Newell, E.W., Vollmers, C., Quake, S.R., Morton, J.M., Davis, M.M., He, X.S., and Greenberg, H.B. (2016). High-dimensional immune profiling of total and rotavirus VP6-specific intestinal and circulating B cells by mass cytometry. Mucosal Immunol *9*, 68-82.

Narvaez, C.F., Feng, N., Vasquez, C., Sen, A., Angel, J., Greenberg, H.B., and Franco, M.A. (2012). Human rotavirus-specific IgM Memory B cells have differential cloning efficiencies and switch capacities and play a role in antiviral immunity in vivo. J Virol *86*, 10829-10840.

Nutt, S.L., Hodgkin, P.D., Tarlinton, D.M., and Corcoran, L.M. (2015). The generation of antibody-secreting plasma cells. Nat Rev Immunol *15*, 160-171.

Pape, K.A., Taylor, J.J., Maul, R.W., Gearhart, P.J., and Jenkins, M.K. (2011). Different B cell populations mediate early and late memory during an endogenous immune response. Science *331*, 1203-1207.

Petrof, E.O., Gloor, G.B., Vanner, S.J., Weese, S.J., Carter, D., Daigneault, M.C., Brown, E.M., Schroeter, K., and Allen-Vercoe, E. (2013). Stool substitute transplant therapy for the eradication of Clostridium difficile infection: 'RePOOPulating' the gut. Microbiome *1*, 3.

Planer, J.D., Peng, Y., Kau, A.L., Blanton, L.V., Ndao, I.M., Tarr, P.I., Warner, B.B., and Gordon, J.I. (2016). Development of the gut microbiota and mucosal IgA responses in twins and gnotobiotic mice. Nature *534*, 263-266.

Rubtsov, A.V., Rubtsova, K., Kappler, J.W., Jacobelli, J., Friedman, R.S., and Marrack, P. (2015). CD11c-Expressing B Cells Are Located at the T Cell/B Cell Border in Spleen and Are Potent APCs. J Immunol *195*, 71-79.

Seifert, M., Przekopowitz, M., Taudien, S., Lollies, A., Ronge, V., Drees, B., Lindemann, M., Hillen, U., Engler, H., Singer, B.B., and Kuppers, R. (2015). Functional capacities of human IgM memory B cells in early inflammatory responses and secondary germinal center reactions. Proc Natl Acad Sci U S A *112*, E546-555.

Sintes, J., Romero, X., de Salort, J., Terhorst, C., and Engel, P. (2010). Mouse CD84 is a pan-leukocyte cell-surface molecule that modulates LPS-induced cytokine secretion by macrophages. J Leukoc Biol *88*, 687-697.

Sohn, H.W., Krueger, P.D., Davis, R.S., and Pierce, S.K. (2011). FcRL4 acts as an adaptive to innate molecular switch dampening BCR signaling and enhancing TLR signaling. Blood *118*, 6332-6341.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Tipton, C.M., Fucile, C.F., Darce, J., Chida, A., Ichikawa, T., Gregoretti, I., Schieferl, S., Hom, J., Jenks, S., Feldman, R.J., *et al.* (2015). Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. Nat Immunol *16*, 755-765.

Tsuji, M., Suzuki, K., Kitamura, H., Maruya, M., Kinoshita, K., Ivanov, II, Itoh, K., Littman, D.R., and Fagarasan, S. (2008). Requirement for lymphoid tissue-inducer cells in isolated follicle formation and T cell-independent immunoglobulin A generation in the gut. Immunity *29*, 261-271.

van Zelm, M.C., Szczepanski, T., van der Burg, M., and van Dongen, J.J. (2007). Replication history of B lymphocytes reveals homeostatic proliferation and extensive antigen-induced B cell expansion. J Exp Med *204*, 645-655.

Vander Heiden, J.A., Yaari, G., Uduman, M., Stern, J.N., O'Connor, K.C., Hafler, D.A., Vigneault, F., and Kleinstein, S.H. (2014). pRESTO: a toolkit for processing high-throughput sequencing raw reads of lymphocyte receptor repertoires. Bioinformatics *30*, 1930-1932.

Vossenkamper, A., Blair, P.A., Safinia, N., Fraser, L.D., Das, L., Sanders, T.J., Stagg, A.J., Sanderson, J.D., Taylor, K., Chang, F., *et al.* (2013). A role for gut-associated lymphoid tissue in shaping the human B cell repertoire. J Exp Med *210*, 1665-1674.

Wang, N.S., McHeyzer-Williams, L.J., Okitsu, S.L., Burris, T.P., Reiner, S.L., and McHeyzer-Williams, M.G. (2012). Divergent transcriptional programming of class-specific B cell memory by T-bet and RORalpha. Nat Immunol *13*, 604-611.

Watanabe, K., Sugai, M., Nambu, Y., Osato, M., Hayashi, T., Kawaguchi, M., Komori, T., Ito, Y., and Shimizu, A. (2010). Requirement for Runx proteins in IgA class switching acting downstream of TGF-beta 1 and retinoic acid signaling. J Immunol *184*, 2785-2792.

Wettenhall, J.M., and Smyth, G.K. (2004). limmaGUI: a graphical user interface for linear modeling of microarray data. Bioinformatics *20*, 3705-3706.

Ye, J., Ma, N., Madden, T.L., and Ostell, J.M. (2013). IgBLAST: an immunoglobulin variable domain sequence analysis tool. Nucleic Acids Res *41*, W34-40.

Zuccarino-Catania, G.V., Sadanand, S., Weisel, F.J., Tomayko, M.M., Meng, H., Kleinstein, S.H., Good-Jacobson, K.L., and Shlomchik, M.J. (2014). CD80 and PD-L2 define functionally distinct memory B cell subsets that are independent of antibody isotype. Nat Immunol *15*, 631-637.

Figure Legends

Figure 1. Human PC-M Accumulate in the Gut Together with ME-M B Cells and Include a Circulating Counterpart Expressing Gut-Homing Receptors

(A) FCM of IgM and IgA on $CD19^{+}CD38^{high}CD10^{-}$ PCs from human ileum and colon. (B) Frequency of PC-M cells among total PCs. (C) FCM of selected surface molecules on naïve (N) B cells, PC-M and PC-SW from human ileum. Numbers indicate mean fluorescence intensity (MFI). (D) IFA of IgM, IgA and DNA in human ileum and mouse small intestine (SI) lamina propria (LP). Original magnification, 20×. Scale bars, 50 μm. (E) Number of PC-M (top), PC-A (center) per mm^2 of LP and PC-M/PC-A ratio (bottom) from human or mouse SI assessed following tissue IFA. Data summarize 6 different tissue samples where at least 4 high-power microscopic fields were analyzed. (F, G) Representative FCM (F) and frequency (G) of $\beta7^{+}CCR9^{+}$ cells in human circulating PC-M, PC-A and PC-G/E. (H) Representative FCM showing IgM vs IgD staining on $CD19^{+}CD38^{-}CD10^{-}$ B cells from different human tissues. (I) Frequency of ME-M B cells from tissues shown in (H). (J) FCM of IgD, CD24, CD27 and CD148 on naïve, ME-M and ME-SW B cells from ileum. Data show one representative result (A,F,H) of 12 (B), 8 (G) or 52 (I) experiments or are from one experiment of at least 3 with similar results (C,D,J). Small horizontal lines indicate the mean (± s.e.m.); Two-tailed unpaired Student's t-test (B and E) and one-way ANOVA with Tukey's *post-hoc* test (I). *p <0.05, **p <0.01, ***p <0.001. See also **Figure S1**.

Figure 2. Human Gut ME-M B Cells Inhabit Mucosal Follicles, Show Post-GC Traits, Include a Circulating Counterpart Expressing Gut-Homing Receptors, and Emerge Early in Life

(A) IFA of IgM, IgD and DNA in human ileum tissue sections. Boxes correspond to enlarged right images. Original magnification, 20× (left) or 60× (right). Scale bars, 50 μm. (B) Number of $IgM^{+}IgD^{-}$ ME-M B cells from human intestine assessed by counting cells/mm^2 following tissue IFA. Data summarize results from 5 different tissue samples where at least 4 microscopic fields were analyzed (C) FCM of CCR7 and CXCR4 on naïve, ME-M and ME-SW B cells from human

ileum and colon. (D) Replication history analyzed by KREC assay. Dashed line corresponds to past cell divisions in control GC B cells from human tonsils. (E, F) FCM (E) and frequency (F) of human circulating $\beta7^{+}CCR9^{+}$ B cells. (G) IFA of IgM, IgD, IgA and DNA in intestinal tissues from children. Boxes correspond to enlarged right images. Original magnification, 10× (top-left), 20× (top-right), 4× (mid-left), 40× (mid-right), 10× (bottom-left), and 40× (bottom-right). Scale bars, 50 μm. Data are from one of at least 3 experiments with similar results (A, C, G), summarize 3 experiments (D) or show one representative result (E) of 24 experiments (F). Results are presented as the mean (± s.e.m.). *p <0.05, **p <0.01, ***p <0.001 (two-tailed unpaired Student's t-test). See also **Figure S1**.

Figure 3. Human Gut ME-M B Cells Are Transcriptionally Distinct from Naïve B Cells and Express a Tissue-Specific Memory Gene Signature That Includes Multiple Activation Traits

(A) May-Grünwald-Giemsa staining on sorted cells. Original magnifications, 100×. (B) Dendrogram of unsupervised agglomerative hierarchical cluster analysis and gene expression heat map diagram displaying genes differentially expressed by naïve, ME-M and ME-SW B cells sorted from human ileum or naïve, MZ and ME-SW from spleen. (C) Scatter plot depicting Robust Multi-array Average normalized expression. (D) Venn diagram showing transcripts exclusively differentially expressed by ME-M and ME-SW B cells versus naïve B cells from human ileum. (E) Volcano plot representation of genes differentially expressed by ME-M B cells versus naïve B cells from human ileum. Selected genes are highlighted. (F) qRT-PCR of mRNAs encoding selected genes in B cell populations as in (B). Results are presented as relative expression (RE) compared to naïve B cells. (G) Ingenuity's upstream regulator comparison analysis showing selected upstream regulators among cytokine and growth factors differentially expressed (p value > $|\log_{10}5|$). (H-I) FCM of selected surface molecules on naïve, ME-M and ME-SW B cells from human ileum. Data are from one of three experiments with similar results (A,H,I), summarize results from 4 biological replicates (B,C,D,E,G) or summarize three independent experiments from at least three different

donors (F). Error bars, s.e.m.; *p <0.05; **p <0.01; ***p <0.001 (two-tailed unpaired Student's t-test). See also **Figure S2** and **S3**.

Figure 4. Human Gut ME-M B Cells Express a Post-GC Mutational Profile and Clonally Relate to Gut PC-M and Some PC-A

(A) Pearson's correlation coefficient matrix of IGHV gene usage by naïve B cells and paired ME-M B cells, ME-A B cells, PC-M and PC-A from human ileum or colon grouped by hierarchical clustering algorithm according to common gene set usage. Circle size and color saturation indicate correlation strength. (B,C) Relative mean frequency of IGHJ6 gene usage and mean number of IGHV gene mutations per 100 bp in gut B cell and PC subsets as in (A). (D) Frequency of Igλ-expressing cells calculated by FCM. (E, F) Circos plots depicting clonal relationships and IGHV gene mutations (G) Morisita-Horn index showing clonal repertoire overlap between ME-M B cells from human ileum and all other B cell subsets. (H, I, J) Lineage tree reconstruction of inferred clonal families (colored circles) and their inferred germline (GL) and intermediate precursors (grey circles). Edges and numbers indicate mutations accumulated along the lineage tree. (K) IFA of ILFs from human ileum stained for AID, Ki-67 and DNA. Insets correspond to boxed areas in main images. Original magnification, 20× (top) or 40× (bottom). Scale bars, 50 μm. (L) RT-PCR of switch circle Iα-Cμ transcripts (SCTs) as well as germline Iμ-Cμ, Iα-Cα1 and Iα-Cα2 transcripts (GTs) in naïve, ME-M and IgM^+ GC B cells or PC-M from human ileum. Data are from one representative donor (E,F,H,I,J) or summarize results from 4 different donors (A,B,C,D,G). Error bars, s.d. *p <0.05, **p <0.01, ***p <0.001 (Welch's t-test and one-way ANOVA with Tukey's post-hoc test). See also **Figure S4 and S5**.

Figure 5. Human Gut ME-M B Cells Undergo Proliferation, PC Differentiation, IgM Secretion and IgA Class Switching in Response to TD or TI Signals

(A, B) FCM of CFSE dilution profiles (A) and MFI of CFSE staining (B) in naïve and ME-M B cells from human ileum cultured for 5 days with control (ctrl) medium alone (black), CD40L, IL-21 and IL-10 (red), CpG and IL-10 (grey) or CpG, IL-10, BAFF and APRIL (blue). (C) FCM of CFSE and CD38 on naïve (top) and ME-M (bottom) B cells from human ileum cultured as in (B). (D) FCM of IgM and IgA on $CD38^{hi}CFSE^{low}$ plasmablasts emerging upon stimulation as in (B). (E) ELISA of IgM and IgA secreted by naïve and ME-M cells from human ileum cultured as in (B). (F) FCM of CD38, CD27, IgM and IgA on sorted $FcRL4^{-}$ (top) or $FcRL4^{+}$ (bottom) ME-M B cells from human ileum cultured for 5 days with medium alone (ctrl) or with CD40L, IL-21 and IL-10. (G) qRT-PCR analysis of mRNA encoding AID (*AICDA*) in naïve, $FcRL4^{-}$ ME-M and $FcRL4^{+}$ ME-M B cells from human ileum. Results are presented as relative expression (RE) compared to that of naïve B cells, which was set as 1. Error bars represent s.e.m; $*p < 0.05$, $**p < 0{,}01$, $***p < 0{,}001$ (two-tailed unpaired Student's *t*-test). Data represent one representative experiment of two with similar results (A-D,F) or summarize at least three different experiments (E,G).

Figure 6. Human IgM from Gut ME-M B Cells and SIgM from Gut PC-M Target Mucus-Embedded Commensals, Including Bacteria Dually Coated by SIgA

(A) Strategy used to test reactivity of IgM secreted by human ME-M B cells and control naïve B cells from human ileum. LCL, lymphoblastoid cell line. (B) ELISA measuring reactivity for phosphorylcholine (PCh) and β-glucan of IgM secreted by EBV-transformed naïve or ME-M B cells from human ileum. (C) Reactivity of IgM from naïve or ME-M B cells to SYTO BC^{+} mucus embed microbiota. (D) Experimental strategy used to measure free and microbiota-bound SIgM and SIgA from human gut mucus. (E) ELISA of free SIgM (blue) and SIgA (red) in mucus samples. (F, G) FCM of SIgA and SIgM bound to viable SYTO BC^{+} microbiota from human ileum or colon mucus (F) or from small intestine (SI) and large intestine (LI) lumen of wild type C57BL/6 mice housed under SPF conditions (G). Data summarize at least 3 (B, C, bottom graph), 12 (E) or 20 (F, G, right graphs) experiments in addition to showing results from one representative experiment (C,

cytogram and profiles; F, left cytograms; G, left cytograms). Two-tailed paired Student's t-test (C,E) and one-way ANOVA with Tukey's *post-hoc* test (F,G) were used for statistical analysis. Error bars represent s.e.m; *p <0.05, **p <0.01, ***p <0.001. See also **Figure S6.**

Figure 7. **Human Mucus-Embedded Commensals Dually Coated by SIgM and SIgA Show Increased Richness and Diversity Compared to Uncoated or SIgA-Only-Coated Commensals.**

(A) Strategy used to profile mucus-embedded microbiota from human ileum or colon. (B,C) Relative abundance of phyla (B) and genera (C). Top 15 most abundant genera are depicted. (D) Rarefaction plots depicting phylogenetic richness (left) and species diversity (right) of microbiota fractionated by IgM/A-sorting as in (A). OTU, operational taxonomic unit. (E) Box plot distribution of mean relative abundance of phyla from bacteria fractioned as in (A). (F) Heatmap diagram of mean Enrichment Index (EI) among donors at the OTU level. Hierarchical clustering was used to group fractions and OTUs. OTUs with significantly different EI among fractions are shown in bold. Data are from one sequencing data set from two different gut segments of one representative donor (B,C) or summarize 10 sequencing data sets from two different gut segments of 5 donors (D-F). Error bars represent s.d; *p <0.05, **p <0.01, ***p <0.001 (Welch's *t*-test). See also **Figure S7.**

STAR METHODS

Contact for Reagent and Resource Sharing

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Andrea Cerutti (acerutti@imim.es).

Experimental Model and Subject Details

Human Tissue and blood Specimens

Histologically normal tissue samples from terminal ileum and ascending colon were obtained from 50 patients undergoing right hemicolectomy due to colonic tumors, unresectable polyps or angiodysplasia. The age of these patients ranged from 28 to 89 years (mean 68.8 years) and the male/female ratio was 1:1.2. Peripheral mononuclear cells were isolated from buffy coats and splenocytes from histologically normal spleens from deceased organ donors or individuals undergoing post-traumatic splenectomy. Tonsils were obtained from adult patients with follicular hyperplasia. The use of blood and tissue samples was approved by the Ethical Committee for Clinical Investigation of the Institut Hospital del Mar d' Investigacions Mèdiques (CEIC-IMIM 2011/4494/I). Fresh tissue and mucus samples and formalin-fixed and paraffin-embedded tissue sections were collected from the Mar Biobanc tissue repository with patient-signed informed consent. All tissue samples were assigned coded identifiers and relevant clinical information remained concealed.

Mice

Female C57BL/6 mice (Charles River Laboratories) were bred in-house in the animal facility of the Barcelona Biomedical Research Park under specific pathogen free (SPF) or conventional housing conditions. All mice were used at 8-12 weeks of age. Mice were euthanized by cervical dislocation and animal procedures were approved by the Ethic Committee of the Barcelona Biomedical Research Park and performed according to Spanish and European legislations.

Method Details

Sample processing

For the isolation of mononuclear cells from human intestinal samples, mucosa and submucosa were dissected from muscularis externa and cut into 2-to-3-mm pieces. These pieces were first washed in calcium and magnesium-free Hanks' balanced salt solution (HBSS) before incubation at 37°C for 20 min in HBSS containing 5mM Dithiothreitol (DTT) and 1 mM Ethylenediaminetetraacetic acid (EDTA). The tissue pieces were transferred into a falcon tube with 30ml of HBSS and shaken vigorously for 10 seconds twice. The supernatant, containing the intra-epithelial lymphocytes fraction, was discarded. The remaining fraction was digested by incubation for 40 min at 37°C with stirring in a solution of HBSS containing 1 mg/ml collagenase IV (Thermo), 50ng/ml DNAse (New England Biolabs) and 0.5% human serum (Sigma). Lamina propria (LP) suspensions were passed through a 70-µm filter, washed, and resuspended in RPMI 1640 medium (Thermo) with 10% fetal bovine serum (FBS). To isolate murine LP lymphocytes, upon excision of PPs, SI and LI segments were opened longitudinally and cut into 5-mm pieces. These pieces were subsequently processed following the same procedure used for human intestinal samples. Human splenocytes and tonsillar mononuclear cells were obtained from fresh samples by enzymatic digestion of the tissue for 40 min at 37 °C in a solution of HBSS containing 1 mg/ml collagenase IV (Thermo), 50ng/ml DNAse (New England Biolabs) and 0.5% human serum (Sigma), followed by separation on a Ficoll-Hypaque gradient (GE Healthcare). Peripheral blood mononuclear cells (PBMCs) were obtained from heparinized blood samples by separation on Ficoll-Hypaque gradient.

Flow cytometry

Cells were incubated at 4°C with Fc-blocking reagent (Miltenyi Biotec) before the addition of the appropriate 'cocktails' of fluorochrome-labeled monoclonal antibodies (mAbs). Dead cells were excluded through the use of 4'-6'-diamidine-2'-phenylindole (DAPI) (Sigma). Cells were acquired

with LSR Fortessa (BD Biosciences) and data were further analyzed with FlowJo V10 software (TreeStar).

Cell Sorting

For cell sorting, cell suspension were incubated at 4ºC with Fc-blocking reagent (Miltenyi Biotec) and stained for 30min with the following monoclonal antibodies: anti-CD45 AF700 (clone: HI30), anti-CD19 PE-cy7 (clone: HIB19), anti-CD38APC-Cy7 (clone: HIT2), anti-CD10 PE (clone: HI10a), anti-IgM BV605 (clone: MHM-88) (all from Biolegend) anti-CD27 PerCpCy5.5 (clone: M-T271) (BD Biosciences), and anti-IgD FITC (Southern). $CD45^{+}CD19^{+}CD38^{dull}CD10^{-}IgD^{++}$ $IgM^{+}CD27^{-}$ naïve B cells, $CD45^{+}CD19^{+}CD38^{dull}CD10^{-}IgD^{+}IgM^{++}CD27^{+}$ MZ B cells, $CD45^{+}CD19^{+}CD38^{int}CD10^{+}$IgD-$IgM^{+}CD27^{+}$ GC-M, $CD45^{+}CD19^{+}CD38^{dull}CD10^{-}IgD^{-}IgM^{+}$ $CD27^{+}$ ME-M B cells, $CD45^{+}CD19^{+}$ $CD38^{dull}CD10^{-}$ $IgD^{-}IgM^{-}CD27^{+}$ ME-SW B cells, $CD45^{+}CD19^{+}CD38^{int}CD10^{+}IgD^{-}CD27^{+}$ GC B cells, $CD45^{+}CD19^{+}CD38^{++}CD10^{-}IgD^{-}gM^{+}CD27^{+}$ PC-M and $CD45^{+}CD19^{+}CD38^{++}CD10^{-}$ $IgD^{-}IgM^{+}CD27^{+}$ PC-SW were sorted with a FACSAria II (BD Biosciences) after exclusion of dead cells through DAPI staining. For sorting of $FCRL4^{+}$ and $FCRL4^{-}$ ME-M, anti-FCRL4 APC (clone: 413D12) was added to the 'cocktails'. The purity of cells sorted this way was consistently >95%.

Cell Cultures

Human sorted intestinal naïve and ME-M B cells (1×10^{5}/well) were seeded in 96-well U-bottomed plates (Thermo) and cultured for 6-7 days in complete RPMI 1640 medium (Thermo) supplemented with 10% FBS, penicillin and streptomycin (10 U/ml) with or without 200 ng/ml megaCD40L (Enzo Life Science), 50 ng/ml IL-10 (Peprotech), 500 ng/ml IL-21 (Peprotech), 1μg/ml CpG ODN-2006 (Invivogen), 500 ng/ml BAFF (Alexis) and 100 ng/ml Mega APRIL (Alexis).

Proliferation Assay

Cell proliferation was assessed with carboxyfluorescein succinimidyl ester (CFSE) using CellTrace CFSE Cell Proliferation Kit (Thermo). Briefly, sorted lymphocytes were resuspended at $1x10^{6}$

cell/ml in PBS supplemented with 5% FBS and incubated with 1.25μM CFSE solution for 5 minutes at room temperature (RT). Stained cells were extensively washed and cultured for 6-7 days in complete RPMI 1640 medium supplemented or not with specific stimuli. Cell division was assessed by measuring the decrease in CFSE fluorescence via flow cytometry.

Generation of EBV-Transformed B Cells

For the generation of EBV-transformed B cell lines, sorted B cells were seeded at 5 x 10^4 cells/well in 96 U-bottom plates (Thermo) in complete RPMI 1640 medium (Thermo) medium containing 2.5 μg/ml CpG ODN-2006 (invivogen) and 30% supernatant from the EBV-producing marmoset B cell line B95-8 (ECACC). Proliferating cells were maintained in culture for 2-3 weeks and then frozen. Culture supernatants containing polyclonal immunoglobulins were stored at -80 ºC.

ELISA

Total and antigen-specific IgM and IgA from culture supernatants were detected by home-made ELISA. Briefly, 96 well ELISA plates (Thermo) were coated over night with goat anti-human Ig-UNLB (Southern Biotech) at 1ug/ml. To measure Ab-reactivity to specific antigens, ELISA plates were coated with either β-D-glucan (50 μg/ml; Sigma) or capsular polysaccharides (1 μg/ml; ATCC), Gal-α1,3-gal-HSA (3 atom spacer) (10 μg/ml; Dextra Laboratories), laminarin (50 μg/ml, Sigma) or L-α-Phosphatidylcholine (0.5 μg/ml; Bioresearch Technologies). For total Ig, serial dilutions of cell culture supernatants were added for 2 h. for antigen-specific Ig, supernatants were used at 20 μg/ml Ig concentration and three 1:10 dilutions in PBS. All ELISAs were developed using HRP-labeled goat anti-human IgM (0.2 μg/ml; cappel) or IgA Fc Ab (0.25 μg/ml; Southern Biotech) and TMB substrate reagent set (BD Bioscience). OD450 was measured and Ab-reactivity was calculated after subtraction of background (OD450 of culture supernatants on PBS coated plates).

Immunofluorescence analysis

Formalin-fixed paraffin-embedded human tissue sections 3-μm in thickness were treated in xylene, a decreasing alcohol gradient and distilled water to achieve de-waxing and rehydration of the tissue. Heat induced epitope retrieval was performed for 15 min in citrate buffer (pH 6) or Tris-EDTA buffer (pH 9). After epitope retrieval, tissue sections were permeabilized with 0.2% Triton X-100 in PBS, blocked with 5% bovine serum albumin and 5% Fc receptor blocking (Miltenyi Biotec) and stained with various combinations of antibodies to specific antigens. Biotinylated antibodies were detected with streptavidin–Alexa Fluor conjugates. Nuclear DNA was visualized with DAPI and coverslips applied with FluorSave reagent (Merck Millipore). Images were acquired either with a Leica TCS SP5 Upright confocal microscope (Leica) or a Nikon Eclipse Ni-E microscope (Nikon) and were further analyzed with ImageJ software.

Giemsa staining

Cytospins were performed from sorted intestinal population at 800 rpm for 5 min using a Cytospin 4 apparatus (Thermo). Approximately 5,000 cells per subset were dried overnight on albumin-coated slides and stained with Giemsa Stain Kit (Jenner-Wright) (Agilent).

RNA extraction and Reverse Transcription

Total cellular RNA was isolated with the RNeasy Micro kit (Qiagen) by following the manufacturer's protocol. Approximately 2 ng of RNA with were reversed transcribed into cDNA using TaqMan® Reverse Transcription Reagents and Random examers (Thermo).

Quantitative Real Time and standard PCR

Quantitative real time PCR (qRT-PCR) were performed in 384-well plates containing Power Sybr Green PCR Master Mix (Thermo) and specific primer pairs (**Table S1**) and analyzed on QuantStudio 12K Flex Real-Time PCR System (Thermo). Gene expression was normalized to that of the gene encoding β-actin for each sample. For the analysis of germline Iμ-Cμ, Iα-Cα1 and Iα-$C\alpha_2$ transcripts, RT-PCRs were carried out using specific primers (**Table S2**) in a 50μl PCR volume with AmpliTaq Gold PCR Mastermix (Thermo). Nested RT-PCR analysis of Iα-Cμ circle transcript

was carried out using two sets of specific primer pairs (**Table S2**) and the following cycling conditions. In the first RT-PCR round, external primers were used in an initial denaturing step at 95°C for 9 min followed by 30 cycles comprised of 94°C for 30 sec, 60°C for 1 min and 72ºC for 10 min. In the second RT-PCR round, internal primers were used in an initial denaturing step at 95°C for 9 min followed by 25 cycles comprised of 94°C for 30 sec, 60°C for 1 min and 72ºC for 10 min. PCR products were subjected to Sanger sequencing for confirmation.

KREC assay

Genomic DNA was isolated from sorted B cell subsets with QIAamp DNA Mini Kit (Qiagen). The replication history of B cell subsets was determined using the κ-deleting recombination excision circle (KREC) assay as described previously (van Zelm et al., 2007). This assay is based on a quantification of coding joints and signal joints of an Igκ-deleting rearrangement (intron RSS-Kde) by qRT-PCR. The ΔCT between the signal joint and the coding joint exactly represents the number of cell divisions a B cell has undergone. The previously established control cell line U698 DB01 (van Zelm et al., 2007) contains 1 coding and 1 signal joint per genome and was used to correct for minor differences in efficiency of both real-time quantitative-PCR assays.

Global Transcriptome Analysis

Total cellular RNA was isolated with the RNeasy Micro kit (Qiagen) from sorted B cell subsets by following the manufacturer's protocol. RNA integrity was assessed using Agilent 2100 Bioanalyzer (Agilent). Only samples with high integrity (RNA integrity number ≥7) were used for transcriptome analysis. Amplification, labeling and hybridizations were performed according to protocol Ovation® Pico WTA System V2 and Encore™ Biotin Module (NuGEN) and then hybridized to GeneChip Human Gene 2.0 ST Array (Affymetrix) in a GeneChip Hybridization Oven 640. Washing and scanning were performed using the Expression Wash, Stain and Scan Kit (Affymetrix) and the Affymetrix GeneChip System including GeneChip Fluidics Station 450 and GeneChip Scanner 3000 7G. After quality control, raw data were background corrected, quantile-normalized

and summarized to a gene-level using the robust multi-chip average (RMA) system. For the detection of differentially expressed genes, a linear model was fitted to the data and empirical Bayes moderated statistics were calculated using the *limma package* from Bioconductor (Wettenhall and Smyth, 2004). Adjustment of *p* values was performed by the determination of false discovery rates (FDR) using the Benjamini-Hochberg procedure. Genes with an adjusted *p* value less than 0.05 and with an absolute fold change value above 1.5 were selected as significant. Spearman's rank correlation was used to study correlation between normalized gene expressions of the different comparisons. Analyses were performed with R and standard packages. Comparison Upstream Regulator Analysis with Ingenuity Pathway Analysis (Ingenuity Systems, www.ingenuity.com) and GSEA from the Molecular Signature Database (MSigDB) (Subramanian et al., 2005) were used to identify similarities and differences among samples.

Next Generation Sequencing of Ig Gene Repertoires

Aliquots of cDNA products from sorted B cells (**Table S3**) were mixed with high performance liquid chromatography-purified primers specific for VH1-VH6 framework region 1 (50 nM) and primers specific for Cα or Cμ (250 nM) containing corresponding Illumina Nextera sequencing tags (**Table S4**) in a PCR volume of 25 μl (4 μl template cDNA) with High Fidelity Platinum PCR Supermix (Thermo). Amplification was performed using the following conditions. An initial step of 95 °C for 5 min was followed by 35 cycles including 95 °C for 30 sec, 58 °C for 30 sec, and 72 °C for 30 sec, supplemented with a final extension step of 72 °C for 5 min. Products were purified with miniElute PCR purification Kit (Qiagen) and Nextera indices were added via PCR with the following conditions: 72 °C for 3 min, 98 °C for 30 sec, 5 cycles of 98 °C for 10 sec, 63 °C for 30 sec, and 72 °C for 3 min. Ampure XP beads (Beckman Coulter Genomics) were used for purification of the PCR products, which were subsequently pooled and denatured. Single-strand products were paired-end sequenced on a MiSeq (Illumina) with the 500 Cycle v2 Kit (2 x 250 bp). In total, 4,898,226 IGHV gene sequences from four donors were obtained through next generation

sequencing (**Table S3**). Paired-end raw sequencing reads were processed into donor-specific Ig gene sequences and grouped in clonal families using a bioinformatics pipeline based on pRESTO, IgBLAST and Change-O (Gupta et al., 2015; Vander Heiden et al., 2014; Ye et al., 2013). First, raw reads with a Phred score lower than 20 were filtered out, and primers for CH and VH genes were masked. Next, corresponding paired-end reads were aligned and merged together (minimum overlap of 6 nucleotides) and annotated with donor, cell type and isotype origin. Finally, donor-specific Ig subsets were combined (prior to VDJ annotation) to estimate donor-specific clonal families. VDJ calling and assignment were performed using the default parameters for IgBLAST and querying the latest downloaded human IGH IMGT database from February 2016. Donor-specific clonal groups were inferred using a distance-based clustering method implemented in Change-O. Sequences with the same VH and JH annotation, identical H-CDR3 region length and Hamming distance higher than 85% were considered to belong to the same clonal group or family. Finally, germline sequences for each clone were reconstructed using the annotated VDJ information as implemented in Change-O. Donor-specific clonal groups were represented through RCircos plots to visualize the relationships between B cell subsets at clonal level. Morisita-Horn Index estimates were calculated using the *divo* R package. In both analyses, a rarefied sample of 10,000 clones per subset was used to account for possible sample size biases between donors, tissues and cell types. MHI bootstrap estimates were calculated by resampling 1,000 times a subsample of 10,000 clones for each tissue, cell type and donor. VH and JH gene usage was estimated for each donor and B cell subset, with bars representing average values among donors. SHM levels were estimated by averaging the number of mutations compared to the inferred germline sequence of each clone belonging to a given B cell subset and donor. Donor-specific clonally related B cell lineage trees were reconstructed and plotted using IgTree Software (Barak et al., 2008).

Mucus Collection and Processing

Mucus was obtained by scrapping off the epithelial surface of macroscopically unaffected fresh tissue samples from terminal ileum and ascending colon of patients undergoing right hemicolectomy. Intestinal contents from murine SI and LI segments were removed by running forceps along a given intestinal segment and placed in a 1.5-ml Eppendorf tube on ice. Aliquots of microbial samples from mucus or intestinal contents were weighed and resuspended at 0.1 mg/μL in PBS with protease inhibitors from SIGMAFAST Protease Inhibitor Tablets (Sigma). Samples were homogenized by vigorously vortexing for 5 min and then centrifuged at 400 g for 5 min to pellet large debris. The supernatant was filtered through a sterile 70μm cell strainer and centrifuged at 8000 g for 5 minutes to pellet microbes. At this stage, supernatants were saved and frozen at −80ºC for the analysis of free SIgM and Sig and microbial pellets were used for bacterial flow cytometry and FACSorting.

Bacterial Flow Cytometry and FACSorting

To measure endogenous SIgM and SIgA bound to intestinal bacteria, microbial pellets were resuspended in PBS 5% FBS and incubated for 30 min on ice with the following combination of antibodies: anti-human IgM APC (clone: SA-DA4) (Southern Biotech) and anti-human IgA PE (Miltenyi Biotec). Finally, bacterial samples were washed and resuspended in PBS with SYTO BC (Thermo) for 15 min on ice to perform FCM analysis. To measure reactivity of IgM from EBV-transformed B cells against intestinal microbiota, endogenous bacteria-bound SIgs were stripped following incubation for 3 min in an acidic sodium citrate buffer (40 mM sodium citrate and 140 mM NaCl, pH 3.0). Then, microbes were quickly spun down to remove the buffer and PBS was added for pH neutralization. These samples were incubated with supernatants from EBV-transformed B cells (at 50 μg/ml total IgM) for 30 min on ice. After washing, microbial pellets were resuspended in MACS 5% FBS and incubated for 30 min on ice with anti-human IgM APC (clone: SA-DA4) (Southern Biotech). Finally, samples were washed and resuspended in PBS with SYTO BC (Thermo) for 15 minutes on ice for FCM analysis. To measure reactivity of IgM from EBV-

transformed B cells against specific bacterial species (*Escherichia coli, Bacillus cereus, Bacteroides vulgatus, Bacteroides fragilis Bacteroides thetaiotaomicron*- all from ATCC and *Ruthenibacterium lactatiformans* and *Roseburia intestinalis* from DSMZ), 10^5 to 10^6 heat inactivated bacteria (65ºC for 20 minutes) were incubated for 15 minutes at RT with graded amounts of IgM from EBV-transformed ME-M B cell lines (from 0.18 to 15 μg/ml total IgM). After washing, microbial pellets were resuspended in PBS 5% FBS and incubated for 15min with anti-human IgM APC (clone: SA-DA4) (Southern Biotech) in the presence of SYTO BC (Thermo). Contamination was minimized by passing all buffers and media through sterile 0.22-μm filters before use. In all settings, bacterial FCM was performed using a FORTESSA Cytometer (BD Biosceince) with low forward scatter (FSC) and side scatter (SSC) thresholds to allow bacterial detection. FSC and SSC were set to a Log scale and samples were gated FSC^+SSC^+SYTO BC^+ and then assessed for IgA and IgM staining. Microbial samples were sorted using a FACSAria II (BD Biosciences) instrument. Threshold settings were set to the minimal allowable voltage for SSC and 50,000 events were collected from $SIgA^-SIgM^-$, $SIgA^+SIgM^-$ or $SIgA^+SIgM^+$ fractions gates as shown in Figure 7A. Each fraction (typically 50 μL) was stored at −20°C before performing PCR and sequencing of bacterial 16S rRNA genes. Multiple precautions were taken to minimize potential contamination of FACSorted fractions, including collecting samples from the flow cytometer droplet stream (sheath fluid) immediately before each sorting to allow assessment of any potential contaminants in fluid lines.

Bacterial 16S rRNA Gene Analysis

DNA was extracted from unsorted mucus samples using PureLink™ Microbiome DNA Purification Kit (Thermo) following the manufacturer's instructions and amplicons of V3-V4 regions of 16S rRNA genes were generated as described below. 16S rRNA amplicons from FACSorted bacteria were generated by adding 2.5 μl of each bacterial fraction directly to Platinum® PCR SuperMix High Fidelity (Thermo) containing PCR primers that target 16S V3 and V4 region (**Table S4**) in triplicate 20 μl reactions. 16S rRNA analysis was performed on samples collected from the flow

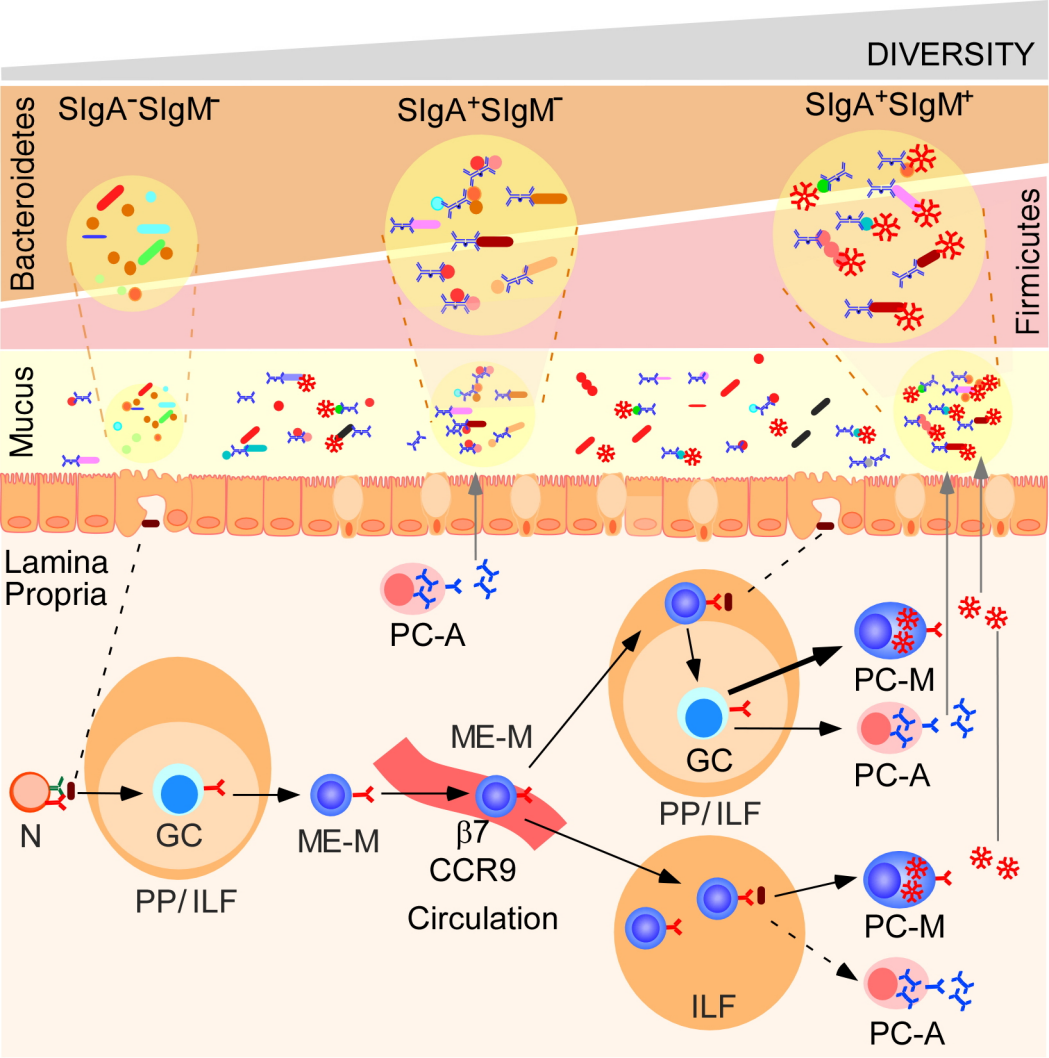
cytometer droplet stream before every sort (sheath fluid), which permitted the identification of sequences that did not originate from the sorted sample. The following PCR conditions were used: an initial denaturation step at 95 °C for 10 min was followed by 35 cycles that included 95 °C for 30 sec, 55 °C for 30 sec, and 72 °C for 30 sec, with an ending step of 72 °C for 5 min. Triplicate reactions were pooled and subjected to 1% agarose gel electrophoresis to verify the presence of a PCR product (these gels also contained negative control reactions). Pooled amplicons were purified with AMPure XP magnetic beads (Agencourt) and subjected to multiplexed sequencing (paired-end 250 nucleotide reads) on a MiSeq instrument (Illumina) with the 500 Cycle V2 Kit (2 x 250 bp). Paired-end reads were filtered (Phred > 19) and merged using the *fastq-join* algorithm. De-multiplexed reads were clustered into OTUs with a 97% identity sequence using the gg_13_5 release from Greengenes database as well as the default open reference QUIIME pipeline for Ilumina reads (Caporaso et al., 2010). Similar to FACSorted fractions, sheath fluid samples were sequenced and processed to identify putative contaminant OTUs. A comparison against "high-biomass" samples (mucus) was performed to pick the most frequent contaminant OTU found in the sheath fluid but not in "high-biomass" samples. This reference contaminant OTU was then used to proportionally remove all the other contaminant OTUs found in either sheath fluid or FACSorted samples. Finally, an 'abundance-filtered data set' was generated by selecting OTUs that were detected at >0.1% relative abundance in each sample. This OTU table was then rarefied to the minimum sample's depth (27529 reads). Rarefied alpha diversity plots for Shannon Index and Phylogenetic Diversity (PD_whole_tree) were generated using default QIIME scripts. A log-transformed Enrichment Index (EI) was calculated for each OTU and sample according to the formula shown in **Figure 7F**. Only OTUs present in input samples at least 3 times and at a frequency higher than 0.1% were used to build the final EI distributions. The heatmap plot was generated using mean values from EI distributions. OTUs and fractions were then grouped through a hierarchical clustering algorithm.

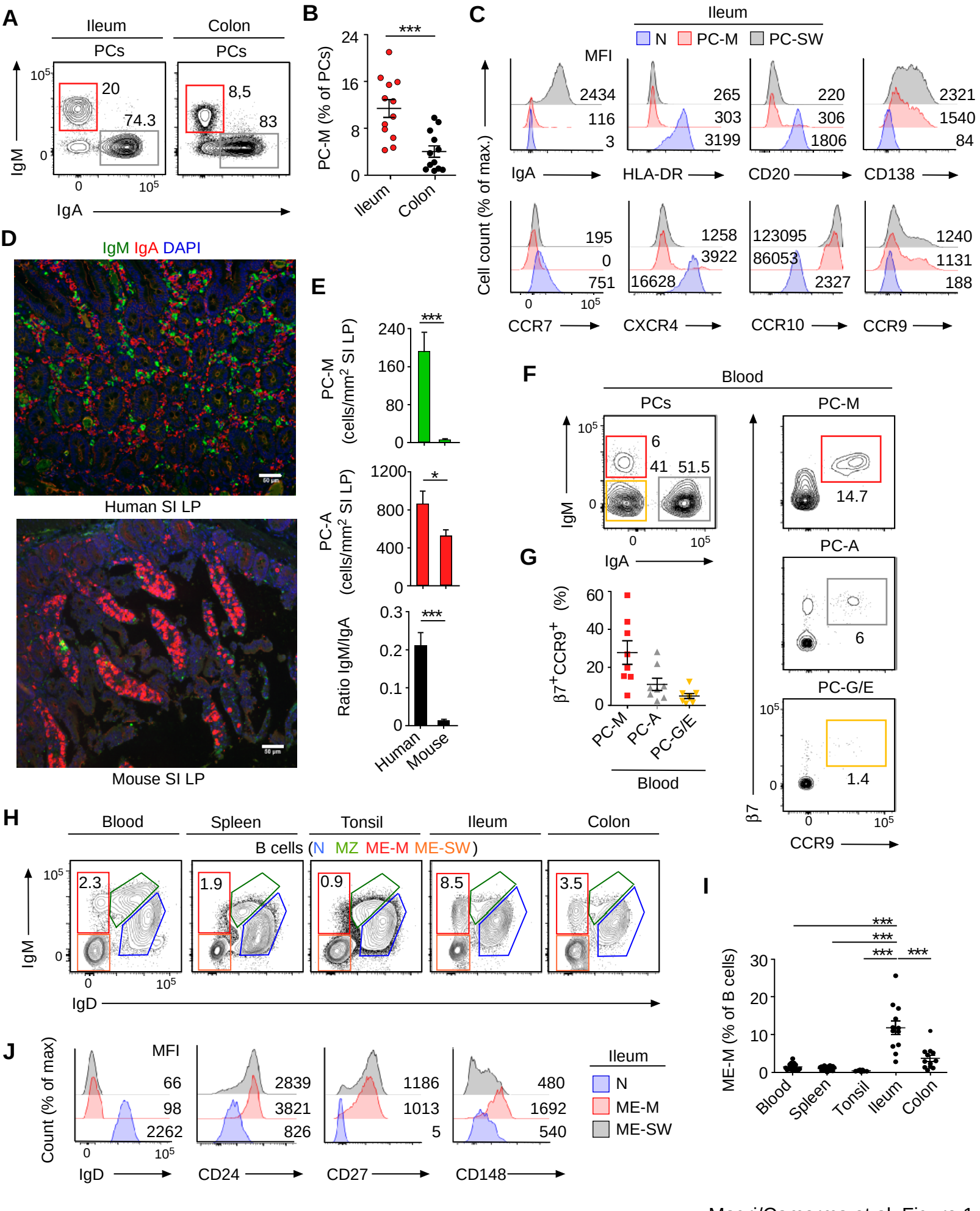
Quantification and Statistical Analysis

Differences between means from independent groups were assessed using Prism 5.03 software (GraphPad) and R studio. For comparison of two groups, *P* values were determined by unpaired two-tailed Student's *t*-test, unless otherwise indicated. For comparison of more than two groups, significant values were calculates via one-way ANOVA with Tukey's *post-hoc* test. *p* values < 0.05 were considered significant. P-values are indicated on plots and in figure legends. ($* \ p < 0.05$, $** \ p < 0.01$, $*** \ p < 0.001$)

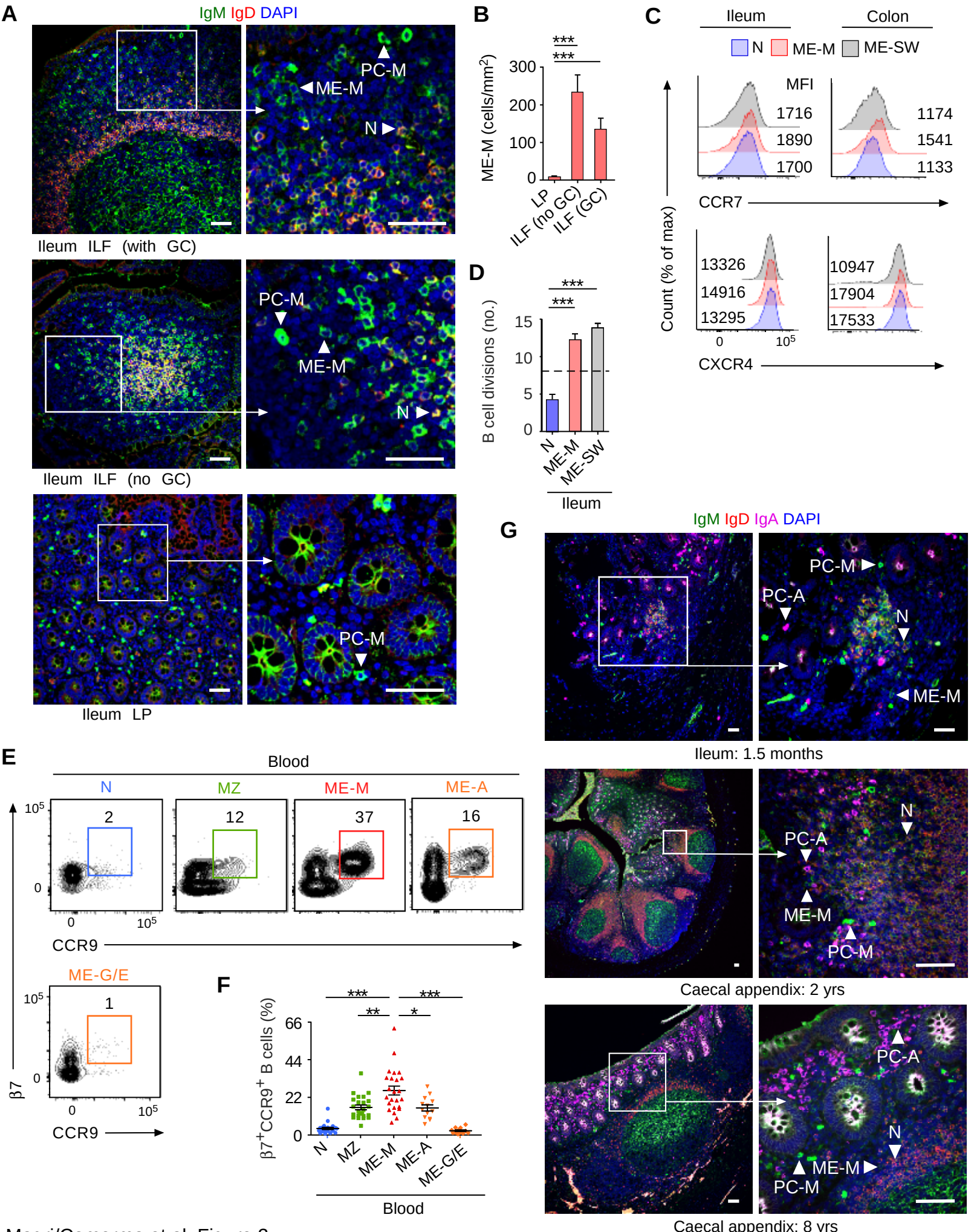
Data and Software Availability

The Gene Expression Omnibus accession number for the global gene transcriptional analysis reported in this paper is GSE89282. Sequencing data are publicly available under BioProject accession number PRJNA355402.

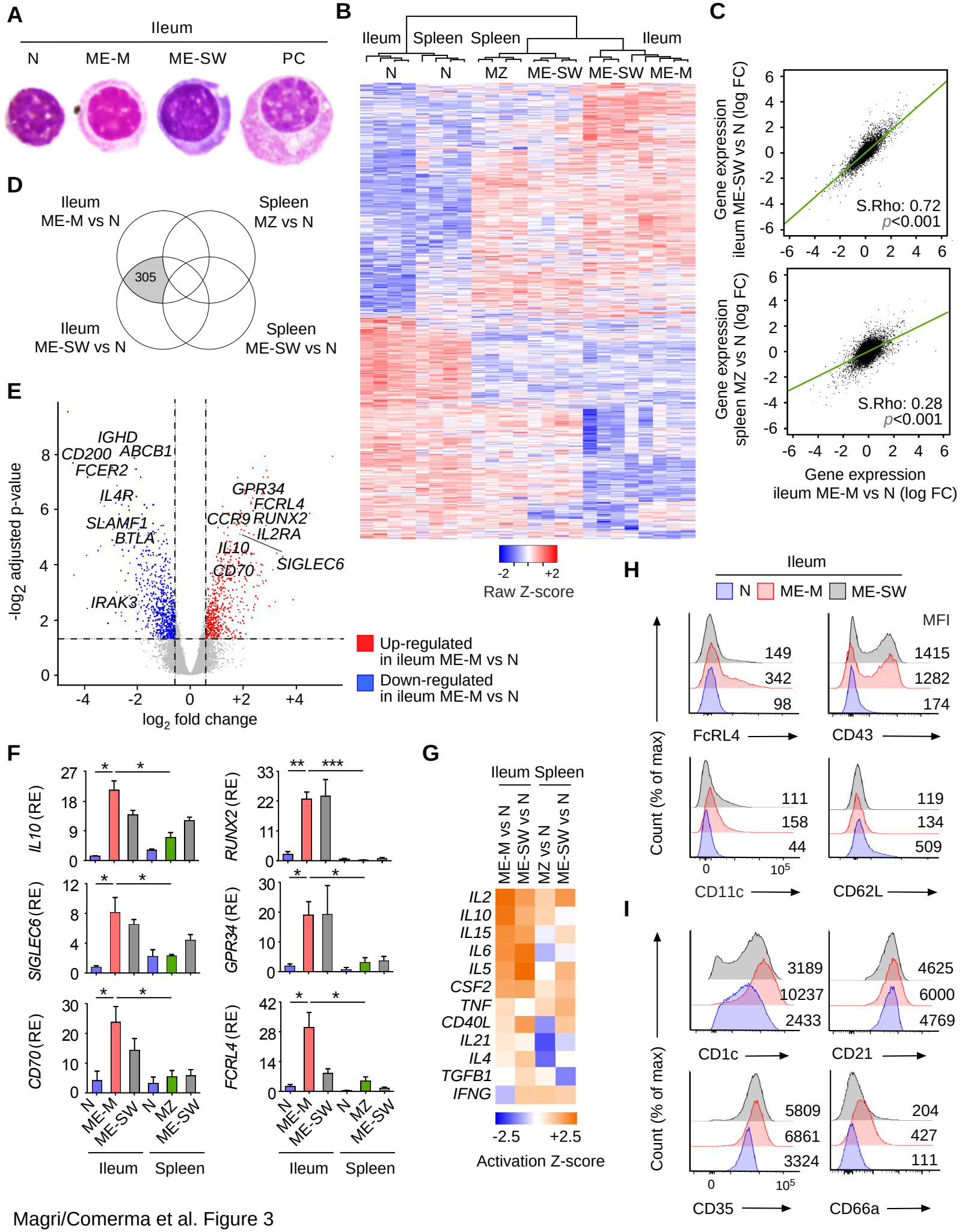

DIVERSITY
SIgA⁻SIgM⁻
SIgA⁺SIgM⁻
SIgA⁺SIgM⁺
Bacteroidetes
Firmicutes
Mucus
Lamina Propria
PC-A
PC-M
PC-A
GC
PP/ILF
ME-M
N
GC
PP/ILF
ME-M
β7
CCR9
Circulation
PC-M
PC-A
ILF



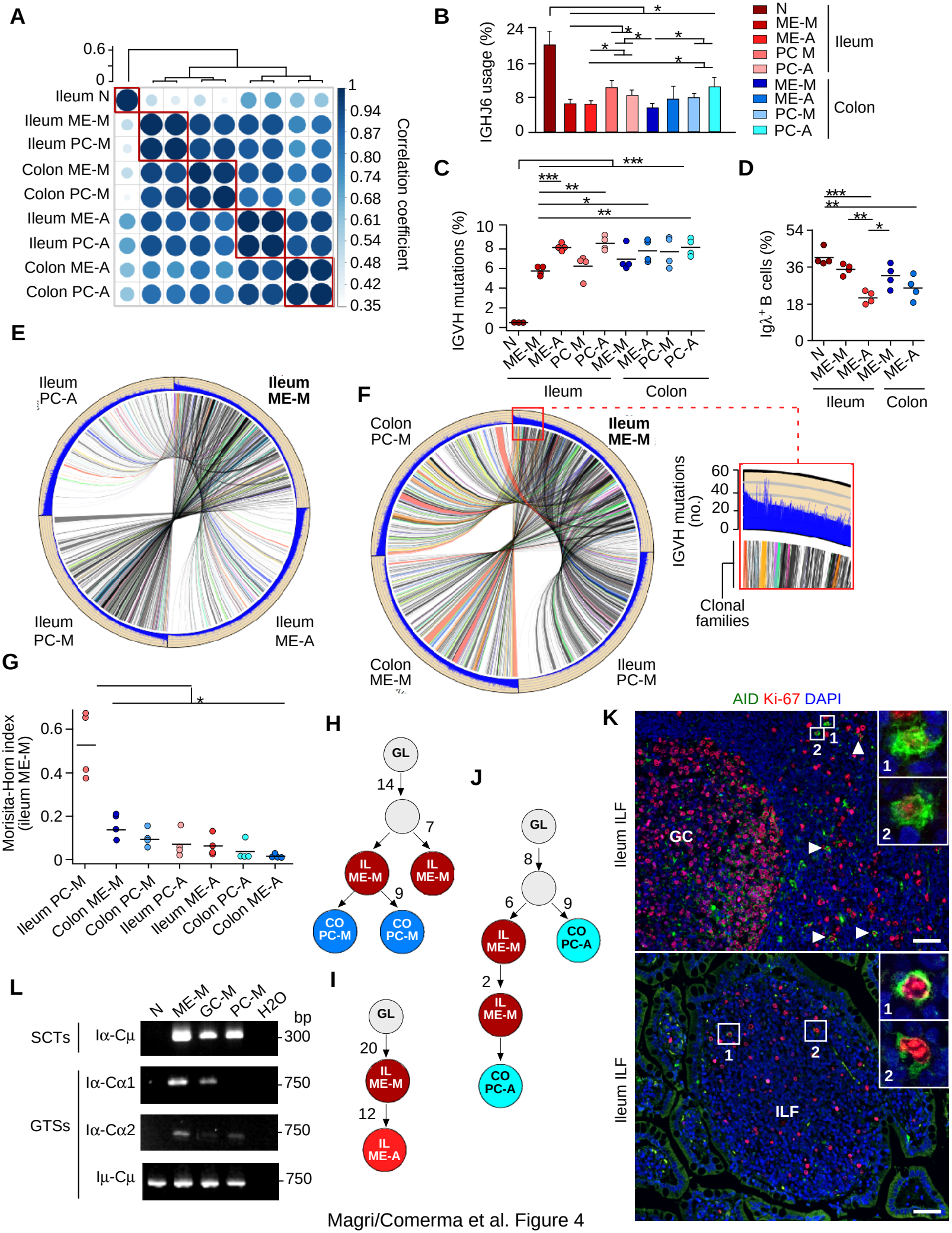
Magri/Comerma et al. Figure 1



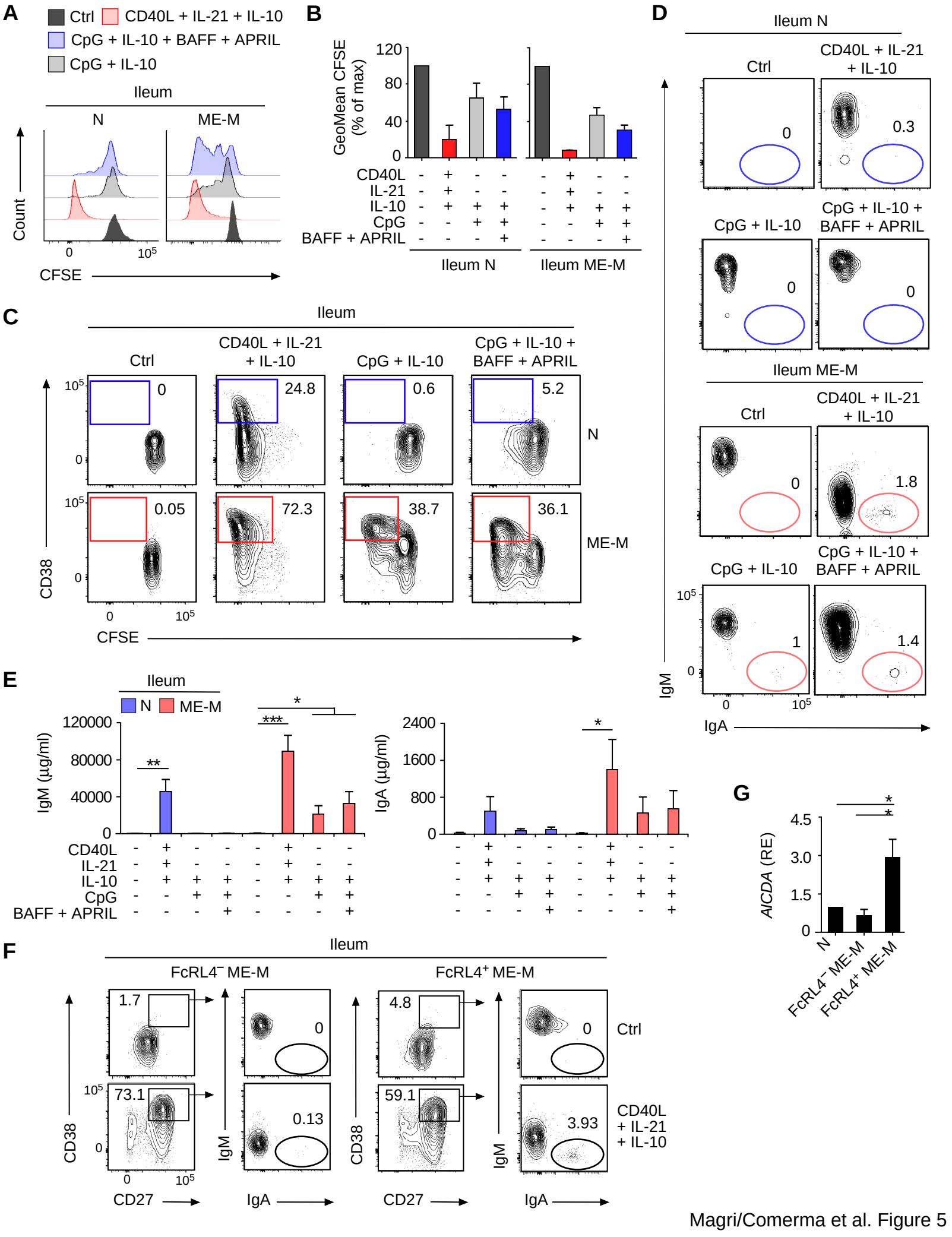
Magri/Comerma et al. Figure 2



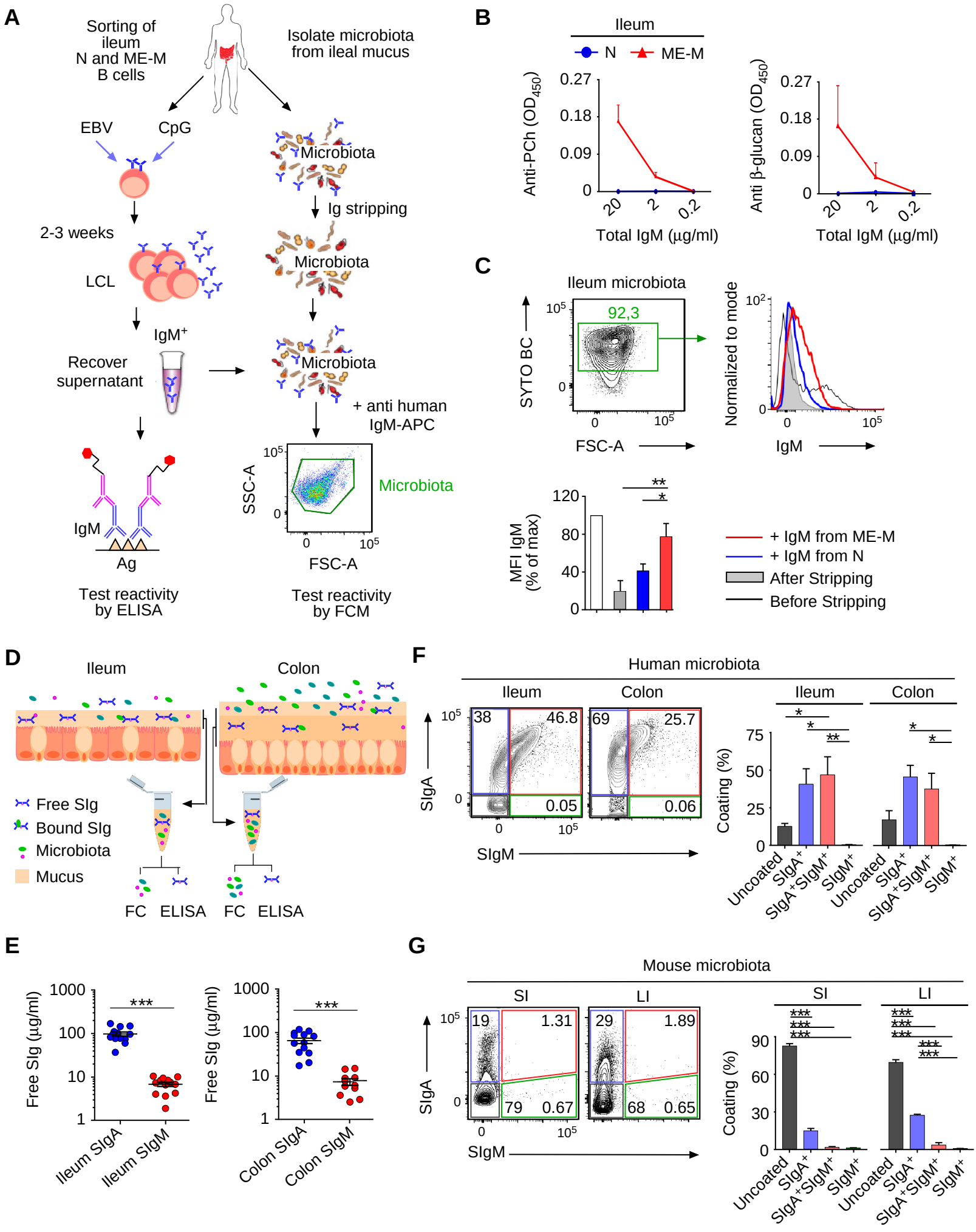
Magri/Comerma et al. Figure 3



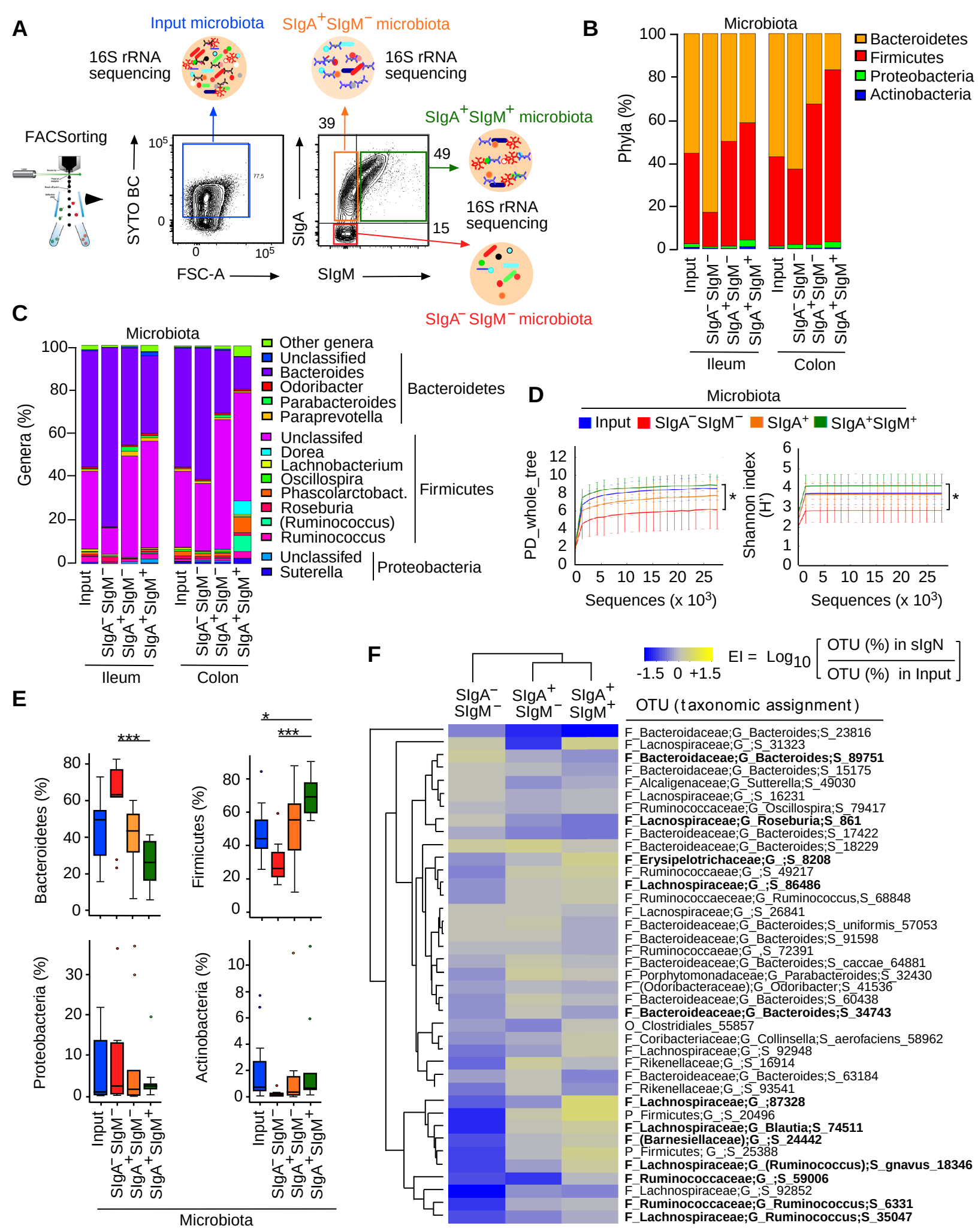
Magri/Comerma et al. Figure 4



Magri/Comerma et al. Figure 5



Magri/Comerma et al. Figure 6



Magri/Comerma et al. Figure 7

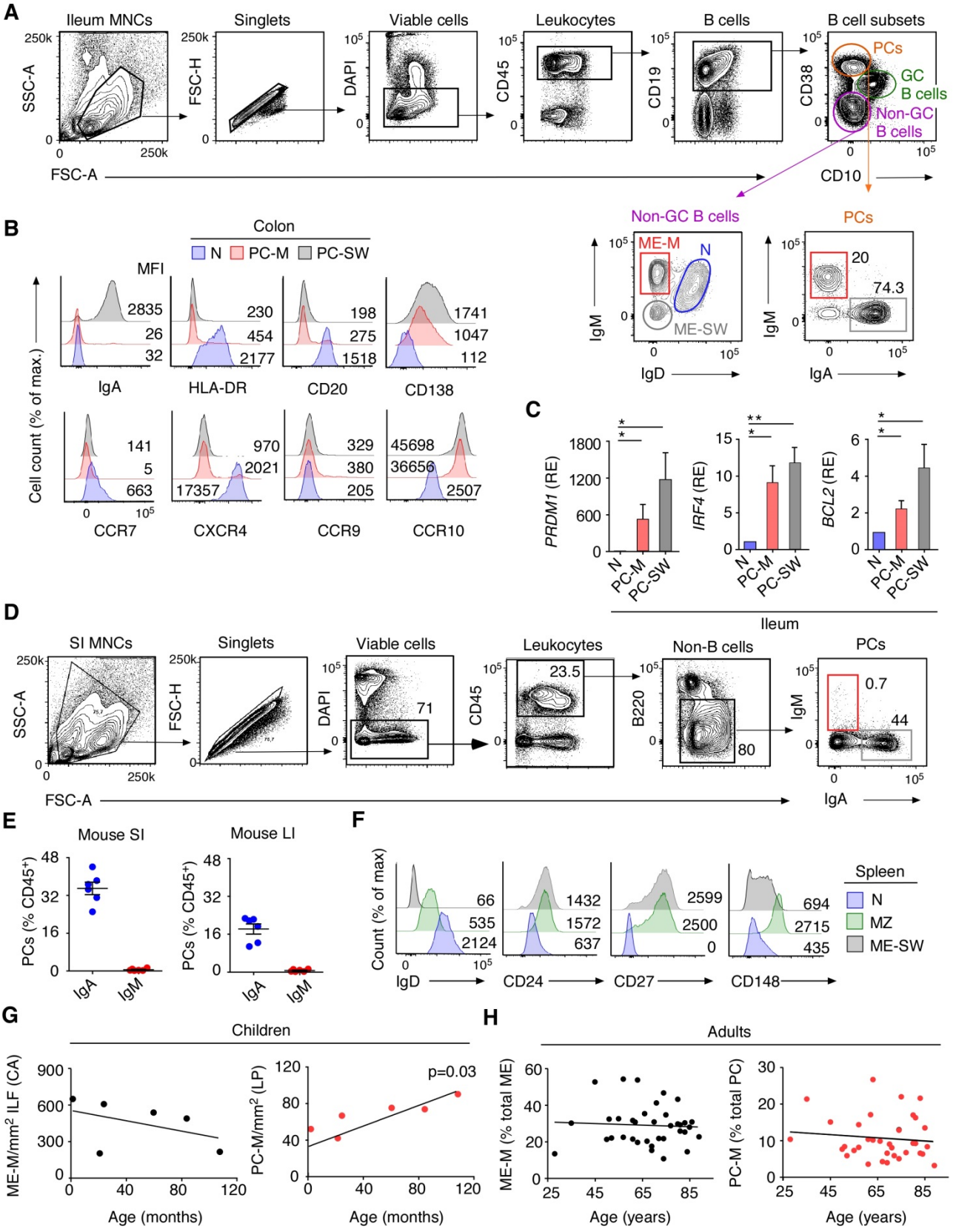

A
Ileum MNCs
Singlets
Viable cells
Leukocytes
B cells
B cell subsets
SSC-A
FSC-H
DAPI
CD45
CD19
CD38
FSC-A
CD10
PCs
GC B cells
Non-GC B cells
B
Colon
N
PC-M
PC-SW
MFI
Cell count (% of max.)
IgA
HLA-DR
CD20
CD138
CCR7
CXCR4
CCR9
CCR10
Non-GC B cells
PCs
ME-M
ME-SW
IgM
IgD
IgA
C
PRDM1 (RE)
IRF4 (RE)
BCL2 (RE)
Ileum
D
SI MNCs
Non-B cells
B220
E
Mouse SI
Mouse LI
PCs (% CD45+)
F
Spleen
MZ
Count (% of max)
CD24
CD27
CD148
G
Children
ME-M/mm2 ILF (CA)
PC-M/mm2 (LP)
p=0.03
Age (months)
H
Adults
ME-M (% total ME)
PC-M (% total PC)
Age (years)

Figure S1 (Related to Figure 1 and Figure 2). Human PC-M Accumulate in the Gut Together with a Large Repertoire of ME-M B Cells

(A) FCM of CD45, CD19, CD38, CD10, IgM and IgD and IgA on viable $DAPI^-$ mononuclear cells from human ileum. IgM and IgA expression by gated PCs is also shown in Figure 1A. (B) FCM of IgA, HLA-DR, CD20, CD138, CCR7, CXCR4, CCR9 and CCR10 on naïve B cells, PC-M and PC-SW from human colon. N, naïve B cells. Numbers indicate MFI. (C) qRT-PCR analysis of mRNAs encoding BLIMP-1 (*PRDM1*), IRF4 and BCL2 in naive, PC-M and PC-SW from human ileum. Results are normalized to those of mRNA encoding β-actin (*ACTB*) and presented as relative expression (RE) compared to that on naive B cells, which was set as 1. (D) FCM of CD45, B220, IgM and IgA on viable $DAPI^-$ mononuclear cells from mouse SI. (E) Frequency of PC-A and PC-M in mouse small and large intestines assessed by FCM. (F) FCM of IgD, CD24, CD27 and CD148 on N, MZ and ME-SW B cells from human spleen. (G, H) Correlation between number of ME-M (left) or PC-M (right) and ages in children (G) and adults (H) assessed by tissue-based cell counting following IFA (G) or FCM (H). Data are from one of at least 3 experiments with similar results (B,F), summarize 3 (C), 14 (E), 7 (G) or more than 30 (H) experiments or show one representative experiment from 12 (A) or 6 (D) with similar results. Error bars, s.e.m. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$ (two-tailed unpaired Student's t-test).

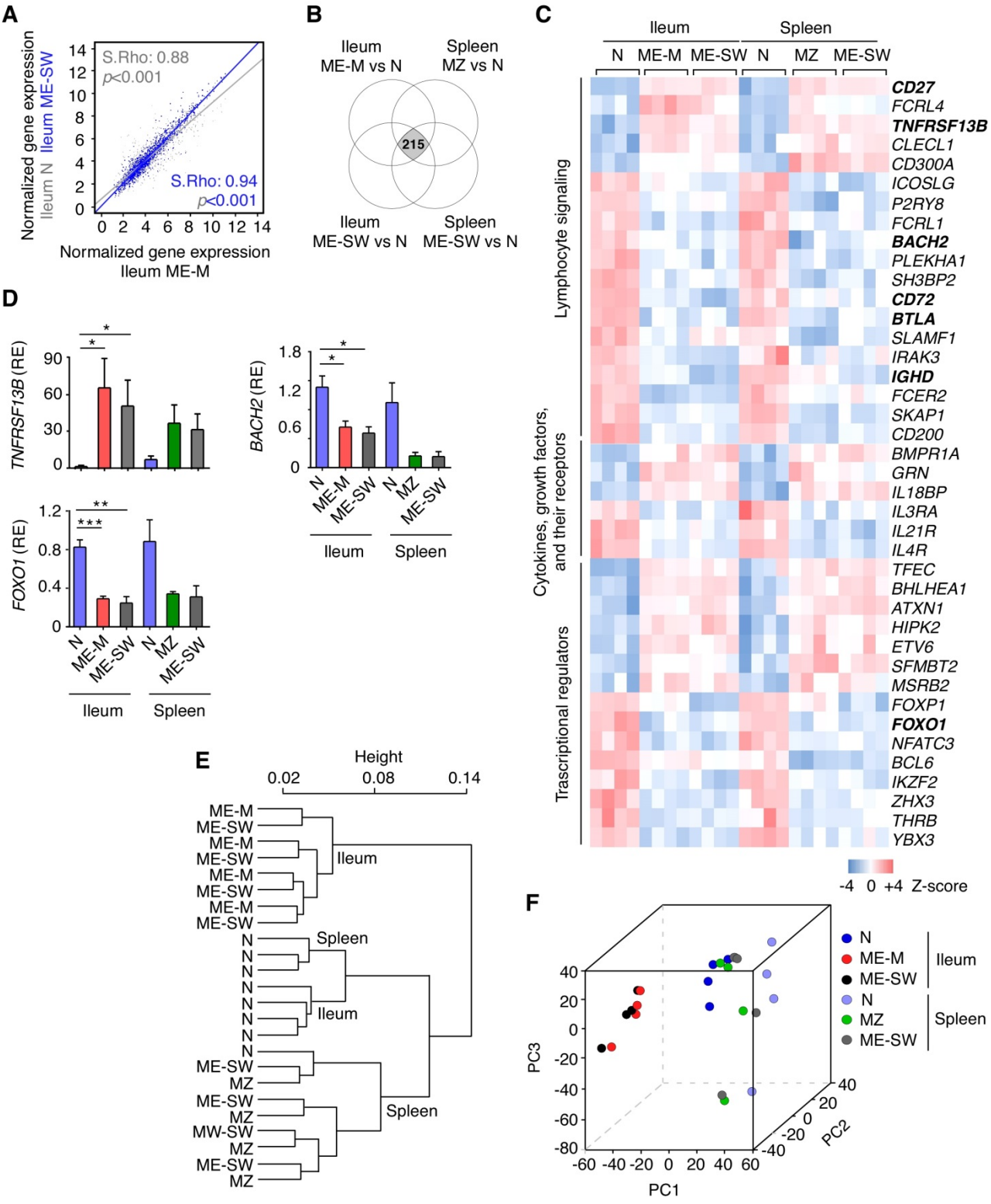

A
Normalized gene expression
Ileum N Ileum ME-SW
S.Rho: 0.88
p<0.001
S.Rho: 0.94
p<0.001
Normalized gene expression
Ileum ME-M
B
Ileum
ME-M vs N
Spleen
MZ vs N
215
Ileum
ME-SW vs N
Spleen
ME-SW vs N
C
Ileum
Spleen
N
ME-M
ME-SW
N
MZ
ME-SW
Lymphocyte signaling
Cytokines, growth factors, and their receptors
Trascriptional regulators
CD27
FCRL4
TNFRSF13B
CLECL1
CD300A
ICOSLG
P2RY8
FCRL1
BACH2
PLEKHA1
SH3BP2
CD72
BTLA
SLAMF1
IRAK3
IGHD
FCER2
SKAP1
CD200
BMPR1A
GRN
IL18BP
IL3RA
IL21R
IL4R
TFEC
BHLHEA1
ATXN1
HIPK2
ETV6
SFMBT2
MSRB2
FOXP1
FOXO1
NFATC3
BCL6
IKZF2
ZHX3
THRB
YBX3
-4 0 +4 Z-score
D
TNFRSF13B (RE)
BACH2 (RE)
FOXO1 (RE)
N
ME-M
ME-SW
N
MZ
ME-SW
Ileum
Spleen
E
Height
0.02
0.08
0.14
ME-M
ME-SW
Ileum
N
Spleen
Ileum
ME-SW
MZ
Spleen
MW-SW
F
PC1
PC2
PC3
N
ME-M
ME-SW
Ileum
N
MZ
ME-SW
Spleen

Figure S2 (Related to Figure 3). Human Gut ME-M B Cells Express a Common Memory Gene Signature Shared with Gut or Splenic ME-SW and MZ B Cells.

(A) Scatter plot depicting gene expression distributions normalized through a Robust Multi-Array Average approach that compared human ileum naïve B cells versus ileum ME-M B cells (grey) as well as human ileum ME-SW B cells versus ME-M B cells (blue). Regression lines are shown and Spearman's Rho (S.Rho) correlation is computed for each scatter point cloud. N, naïve B cells. (B) Venn diagram representing genes commonly differentially expressed (adjusted *P* value <0.05; |log2FC| >0.58) by ME-M and ME-SW B cells from human ileum as well as MZ and ME-SW B cells from human spleen. For each B cell subset, differential expression was established from comparisons with tissue-specific naive B cells. Shared differentially expressed genes correspond to number in grey area. (C) Heat map diagrams showing a selection of immune response-related genes belonging to the common human B cell memory signature shown in (B) and grouped by curated functional categories. Each row represents an individual gene and each column a biological replicate. The color bar depicts normalized intensity values (adjusted P value <0.05; |log2FC| >0.58). Highlighted genes are discussed in the text. (D) qRT-PCR of transcripts encoding TACI (*TNFRSF13B*), BACH2 or FOXO1 in naïve, ME-M and ME-SW B cells from human ileum as well as naïve, MZ and ME-SW B cells from human spleen. Results are normalized to those of mRNA encoding β-actin (*ACTB*) and presented as relative expression (RE) compared with that of naive B cells, which was set as 1. (E) Unsupervised hierarchical clustering analysis of global gene expression profiles from human ileal and splenic B cells listed in (D) using correlation distance as well as the Ward's linkage method. Dendrogram height parameter indicates dissimilarity between clusters. (F) Principal component (PC) analysis of global gene expression profiles from human and ileal B cells listed in (D). Data show results from one experiment including 4 biological replicates from each tissue and from each B cell subset (A,B,C.F) or summarize at least three different experiments (D). Error bars, s.e.m. **p* <0.05, ***p* <0.01, ****p* <0.001 (Welch's *t*-test).

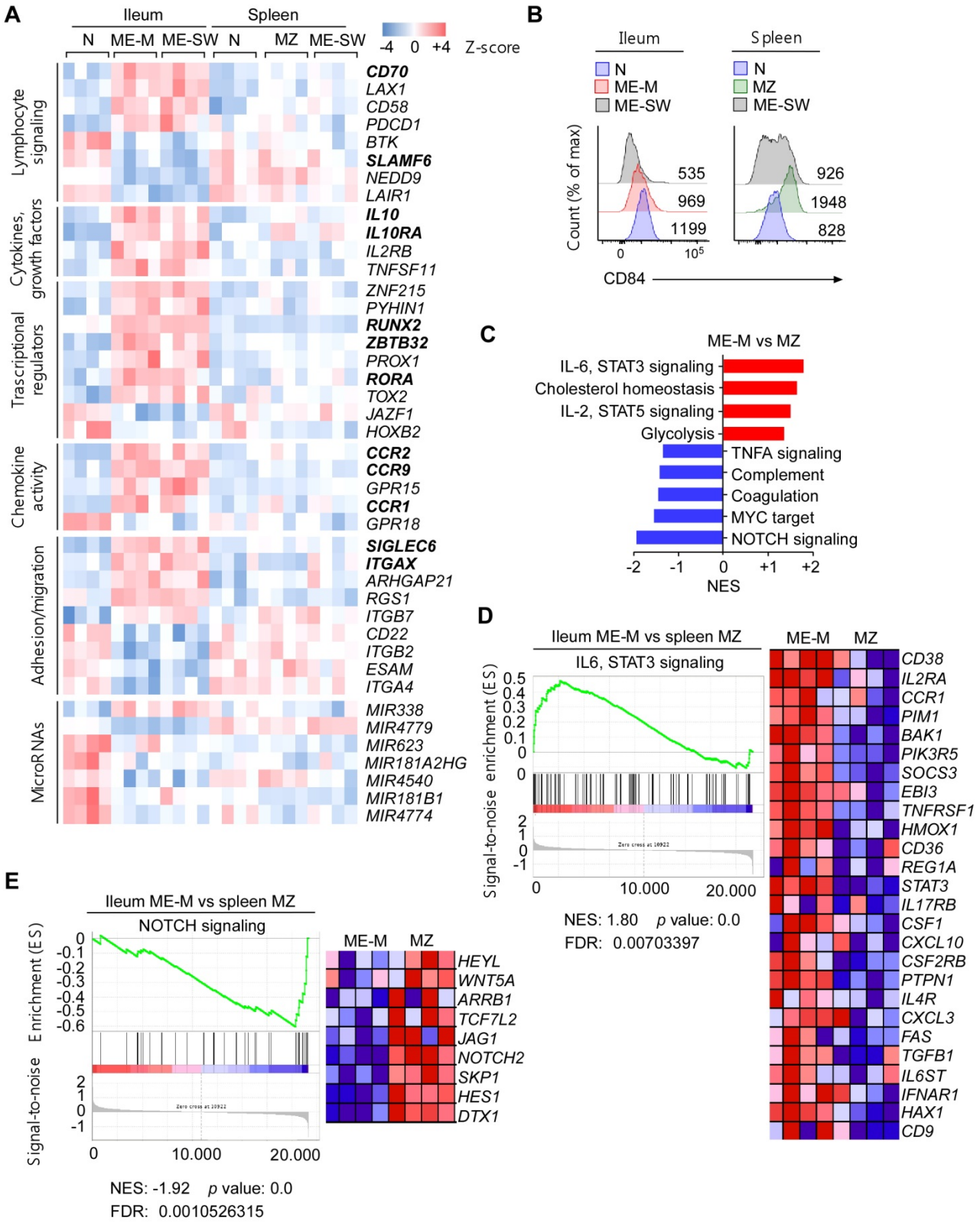

A
Ileum
Spleen
N
ME-M
ME-SW
N
MZ
ME-SW
-4 0 +4
Z-score
Lymphocyte signaling
CD70
LAX1
CD58
PDCD1
BTK
SLAMF6
NEDD9
LAIR1
Cytokines, growth factors
IL10
IL10RA
IL2RB
TNFSF11
Trascriptional regulators
ZNF215
PYHIN1
RUNX2
ZBTB32
PROX1
RORA
TOX2
JAZF1
HOXB2
Chemokine activity
CCR2
CCR9
GPR15
CCR1
GPR18
Adhesion/migration
SIGLEC6
ITGAX
ARHGAP21
RGS1
ITGB7
CD22
ITGB2
ESAM
ITGA4
MicroRNAs
MIR338
MIR4779
MIR623
MIR181A2HG
MIR4540
MIR181B1
MIR4774
B
Ileum
Spleen
N
ME-M
ME-SW
N
MZ
ME-SW
Count (% of max)
535
969
1199
926
1948
828
0
10^5
CD84
C
ME-M vs MZ
IL-6, STAT3 signaling
Cholesterol homeostasis
IL-2, STAT5 signaling
Glycolysis
TNFA signaling
Complement
Coagulation
MYC target
NOTCH signaling
-2 -1 0 +1 +2
NES
D
Ileum ME-M vs spleen MZ
IL6, STAT3 signaling
Signal-to-noise enrichment (ES)
0.5 0.4 0.3 0.2 0.1 0
0
2 1 0 -1
0
10.000
20.000
NES: 1.80 p value: 0.0
FDR: 0.00703397
ME-M MZ
CD38
IL2RA
CCR1
PIM1
BAK1
PIK3R5
SOCS3
EBI3
TNFRSF1
HMOX1
CD36
REG1A
STAT3
IL17RB
CSF1
CXCL10
CSF2RB
PTPN1
IL4R
CXCL3
FAS
TGFB1
IL6ST
IFNAR1
HAX1
CD9
E
Ileum ME-M vs spleen MZ
NOTCH signaling
Enrichment (ES)
0 -0.1 -0.2 -0.3 -0.4 -0.5 -0.6
Signal-to-noise
2 1 0 -1
0
10.000
20.000
NES: -1.92 p value: 0.0
FDR: 0.0010526315
ME-M MZ
HEYL
WNT5A
ARRB1
TCF7L2
JAG1
NOTCH2
SKP1
HES1
DTX1

Figure S3 (Related to Figure 3). Human Gut ME-M B Cells Express a Tissue-Specific Memory Gene Signature Shared with Gut ME-SW But Not Splenic ME-SW and MZ B Cells.

(A) Heat map diagrams showing a selection of immune response-related genes belonging to the tissue-specific signature detected in human ileal ME-M and ME-SW B cells. Genes are grouped by curated functional categories. Each row represents an individual gene and each column a biological replicate. The color bar depicts normalized intensity values (adjusted *p* value <0.05; |log2FC| >0.58). Highlighted genes are discussed in the text. N, naïve B cells. (B) FCM of CD84 (encoded by *SLAMF6*) on naive, ME-M and ME-SW B cells from human ileum as well as splenic naive, MZ and ME-SW B cells from human spleen. Numbers indicate MFI. (C) GSEA showing coordinated gene sets from Molecular Signature Database (MSigDB) significantly enriched (nominal *p* value <0.05) in human ileum ME-M B cells (red) compared to human splenic MZ B cells (blue) and ranked for the normalized enriched score (NES). (D) Left graph: GSEA of "IL-6, STAT3 signaling" category genes differentially expressed by human ileal ME-M compared to human splenic MZ B cells. Right heat map: gene expression levels of the leading edge subset. (E) Left graph: GSEA of "NOTCH signaling pathway" category genes differentially expressed by human ileal ME-M compared to human splenic MZ B cells. Right heat map: gene expression levels of the leading edge subset. Data show results from one experiment including 4 biological replicates from each tissue and from each B cell subset (A-E).

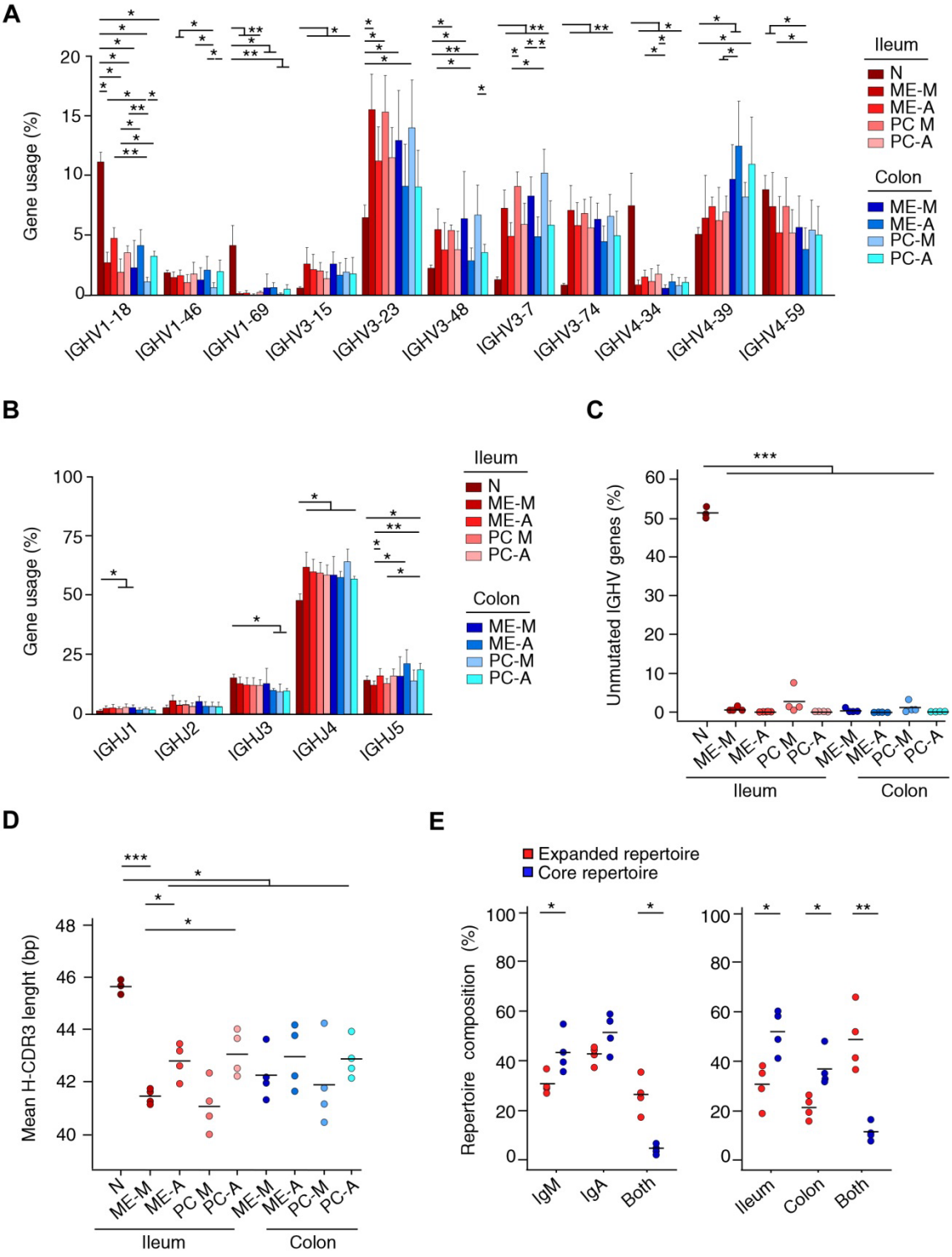
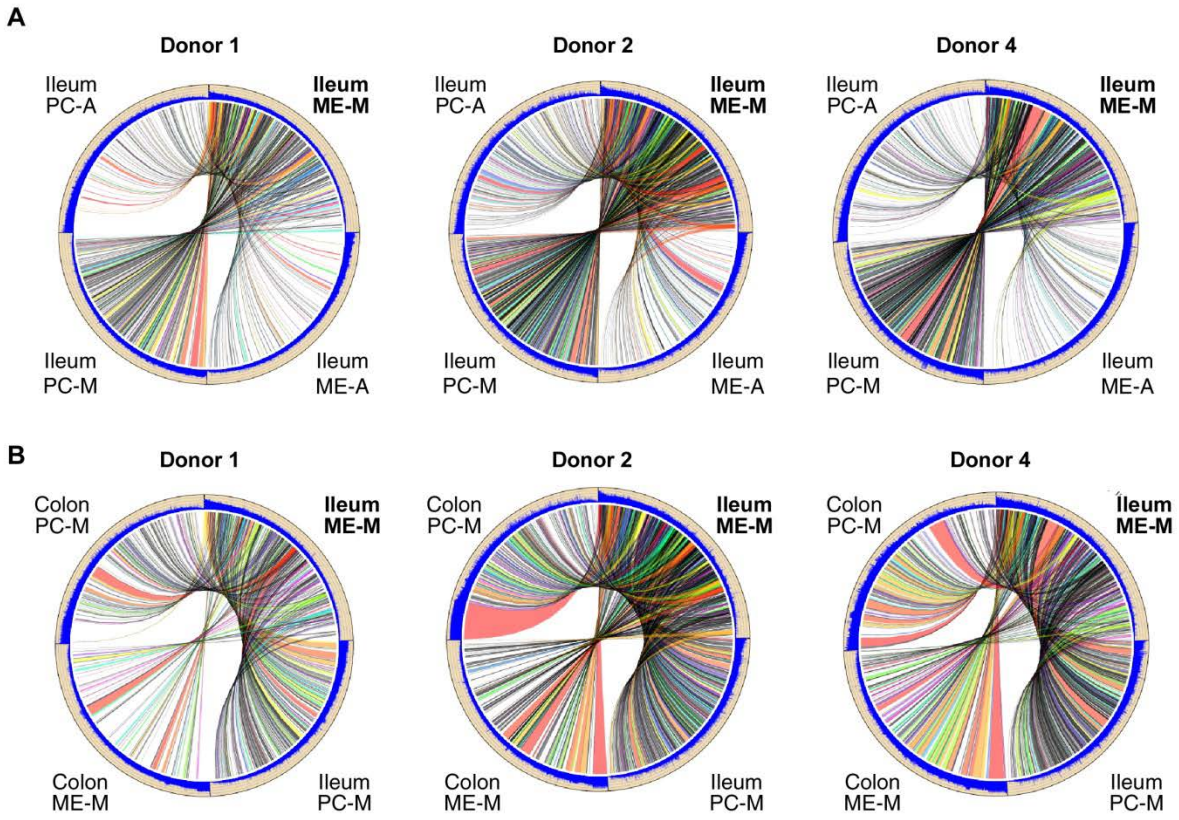
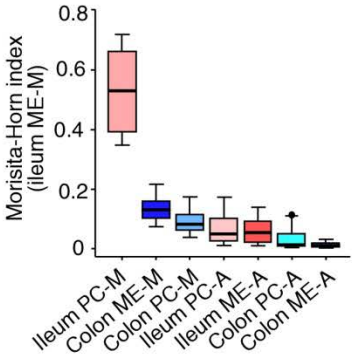

A
Gene usage (%)
IGHV1-18
IGHV1-46
IGHV1-69
IGHV3-15
IGHV3-23
IGHV3-48
IGHV3-7
IGHV3-74
IGHV4-34
IGHV4-39
IGHV4-59
Ileum
N
ME-M
ME-A
PC M
PC-A
Colon
ME-M
ME-A
PC-M
PC-A
B
Gene usage (%)
IGHJ1
IGHJ2
IGHJ3
IGHJ4
IGHJ5
C
Unmutated IGHV genes (%)
Ileum
Colon
D
Mean H-CDR3 lenght (bp)
E
Expanded repertoire
Core repertoire
Repertoire composition (%)
IgM
IgA
Both
Ileum
Colon
Both

Figure S4 (Related to Figure 4). Human Gut ME-M B Cells Express Molecular Fingerprints Denoting Transition Through GCs and Ongoing IgM-to-IgA CSR.

(A) IGHV gene usage (mean relative frequency) by naïve B cells, ME-M B cells, ME-A B cells, PC-M and PC-A from human ileum (shaded red) or colon (shaded blue) measured by next generation sequencing. Only IGHV genes significantly differentially used are shown. (B) IGHJ1, IGHJ2, IGHJ3, IGHJ4 and IGHJ5 gene usage (mean relative frequency) by human B cell and PC subsets as in (A) measured by next generation sequencing. (C) Proportion of unmutated IGHV gene sequences (no mutations compared to their putative germline counterpart) in human B cell and PC subsets as in (A) measured by next generation sequencing. (D) Mean value of H-CDR3 length (measured as number of bp) in human B cell and PC subsets as in (A) measured by next generation sequencing. (E) Clone and clonal family composition from each donor according to antibody isotype expression and tissue localization. The core repertoire was built by collapsing clones into clonal families, whereas the expanded repertoire corresponds to the entirety of sequenced clones. All samples were rarefied to 10,000 reads. Data summarize experiments from 4 donors. Error bars, s.d. *p <0.05, **p <0.01, ***p <0.001 (two-tailed Student's t-test).


A
Donor 1
Donor 2
Donor 4
Ileum PC-A
Ileum ME-M
Ileum PC-M
Ileum ME-A
B
Donor 1
Donor 2
Donor 4
Colon PC-M
Ileum ME-M
Colon ME-M
Ileum PC-M


C
Morisita-Horn index (ileum ME-M)
0.8
0.6
0.4
0.2
0
Ileum PC-M
Colon ME-M
Colon PC-M
Ileum PC-A
Ileum ME-A
Colon PC-A
Colon ME-A

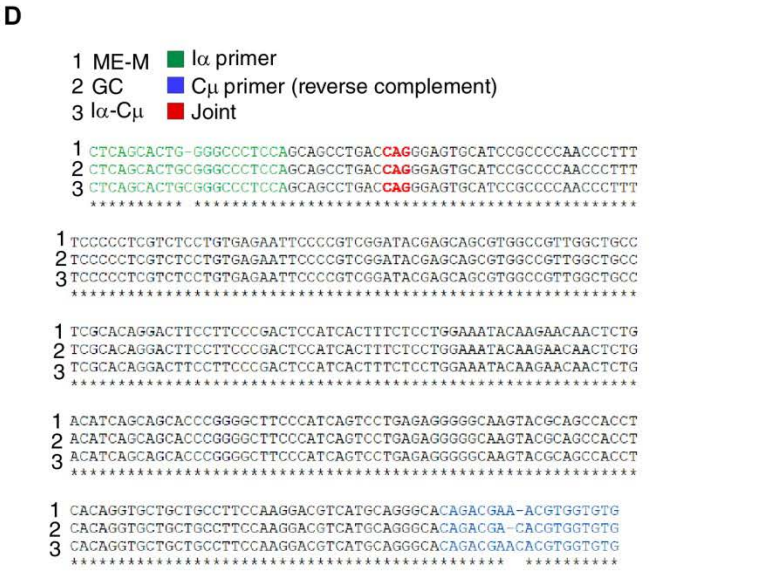

D
1 ME-M
2 GC
3 Iα-Cμ
Iα primer
Cμ primer (reverse complement)
Joint
1 CTCAGCACTG-GGGCCCTCCAGCAGCCTGACCAGGGAGTGCATCCGCCCCAACCCTTT
2 CTCAGCACTGCGGGCCCTCCAGCAGCCTGACCAGGGAGTGCATCCGCCCCAACCCTTT
3 CTCAGCACTGCGGGCCCTCCAGCAGCCTGACCAGGGAGTGCATCCGCCCCAACCCTTT
1 TCCCCCTCGTCTCCTGTGAGAATTCCCCGTCGGATACGAGCAGCGTGGCCGTTGGCTGCC
2 TCCCCCTCGTCTCCTGTGAGAATTCCCCGTCGGATACGAGCAGCGTGGCCGTTGGCTGCC
3 TCCCCCTCGTCTCCTGTGAGAATTCCCCGTCGGATACGAGCAGCGTGGCCGTTGGCTGCC
1 TCGCACAGGACTTCCTTCCCGACTCCATCACTTTCTCCTGGAAATACAAGAACAACTCTG
2 TCGCACAGGACTTCCTTCCCGACTCCATCACTTTCTCCTGGAAATACAAGAACAACTCTG
3 TCGCACAGGACTTCCTTCCCGACTCCATCACTTTCTCCTGGAAATACAAGAACAACTCTG
1 ACATCAGCAGCACCCGGGGCTTCCCATCAGTCCTGAGAGGGGGCAAGTACGCAGCCACCT
2 ACATCAGCAGCACCCGGGGCTTCCCATCAGTCCTGAGAGGGGGCAAGTACGCAGCCACCT
3 ACATCAGCAGCACCCGGGGCTTCCCATCAGTCCTGAGAGGGGGCAAGTACGCAGCCACCT
1 CACAGGTGCTGCTGCCTTCCAAGGACGTCATGCAGGGCACAGACGAA-ACGTGGTGTG
2 CACAGGTGCTGCTGCCTTCCAAGGACGTCATGCAGGGCACAGACGA-CACGTGGTGTG
3 CACAGGTGCTGCTGCCTTCCAAGGACGTCATGCAGGGCACAGACGAACACGTGGTGTG

Figure S5 (Related to Figure 4). Human Gut ME-M B Cells Clonally Relate to Gut PC-M and Some ME-A and PC-A and Show Ongoing IgM-to-IgA CSR.

(A,B) Circos plots from three donors additional to donor 3 shown in main Figure 4E and 4F. Clonal relationships and IGHV gene mutations in ME-M B cells, ME-A B cells, PC-M and PC-A from human ileum (A) and in ME-M B cells and PC-M from human ileum and colon (B) are depicted. Each sample was rarefied to 10,000 clones to establish proper comparisons. Outer track: number of mutations per sequence; inner space: clonal relationships of ME-M B cells with other B cell and PC subsets. (C) Morisita-Horn bootstrapping distributions showing clonal repertoire overlap (0 = no overlap; 1 = full overlap) in ME-M B cells from human ileum relatively to ME-M B cells, ME-A B cells, PC-M and PC-A from human ileum or colon. (D) Multiple sequence alignment of RT-PCR-amplified switch circle Iα-Cμ transcripts from ME-M and GC B cells from human ileum and their reference sequence.

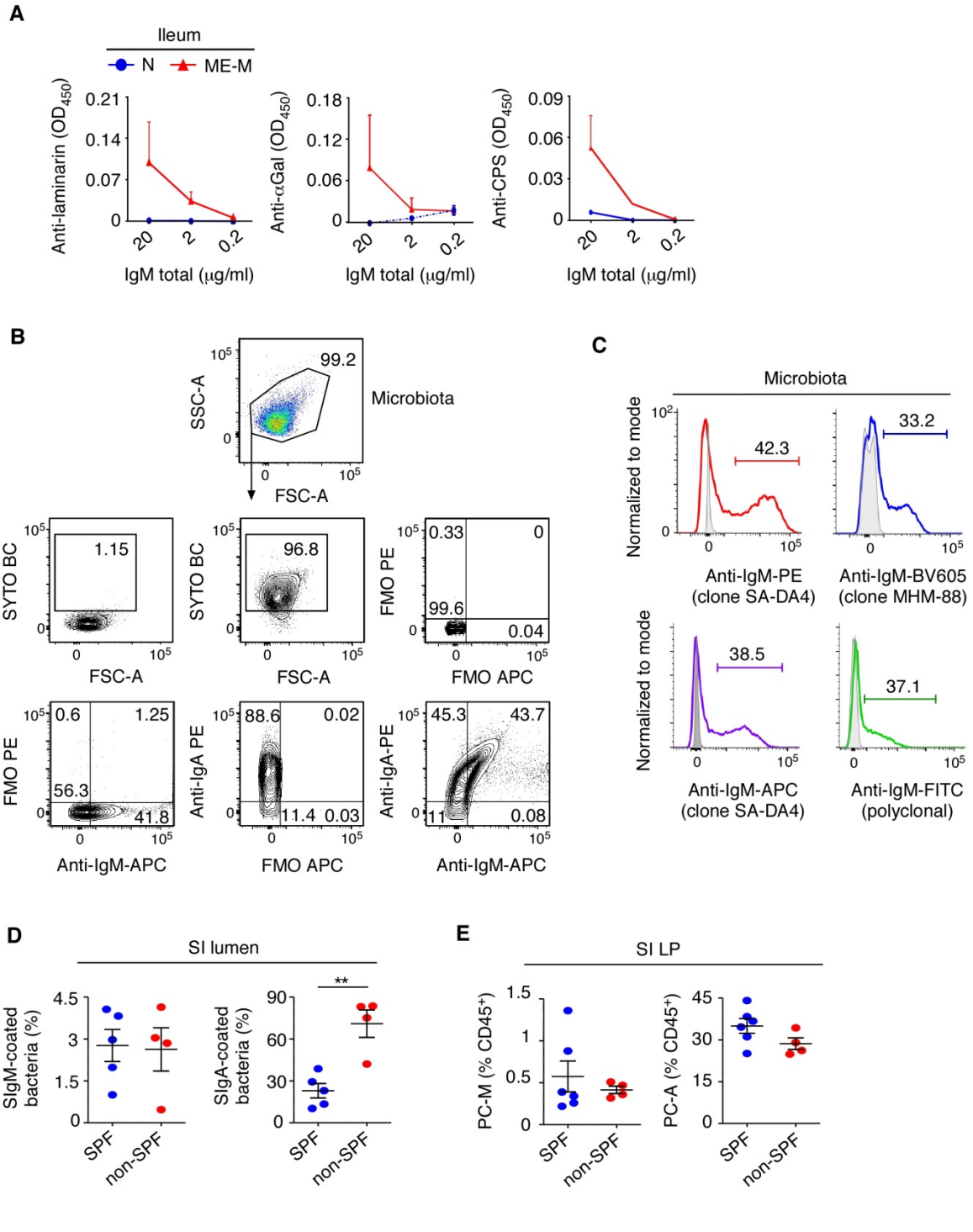

A
Ileum
N
ME-M
Anti-laminarin (OD450)
Anti-αGal (OD450)
Anti-CPS (OD450)
IgM total (μg/ml)
B
Microbiota
SSC-A
FSC-A
SYTO BC
FMO PE
FMO APC
Anti-IgA PE
Anti-IgA-PE
Anti-IgM-APC
C
Microbiota
Normalized to mode
Anti-IgM-PE (clone SA-DA4)
Anti-IgM-BV605 (clone MHM-88)
Anti-IgM-APC (clone SA-DA4)
Anti-IgM-FITC (polyclonal)
D
SI lumen
SIgM-coated bacteria (%)
SIgA-coated bacteria (%)
SPF
non-SPF
E
SI LP
PC-M (% CD45+)
PC-A (% CD45+)

Figure S6 (Related to Figure 6). Human IgM from Gut ME-M B Cells and SIgM from Gut PC-M Bind Autologous Mucus-Embedded Microbiota.

(A) ELISA measuring binding of laminarin, galactose-□-1,3-galactose (αGal) and capsular polysaccharide (CPS) antigens to IgM antibodies secreted by EBV-transformed ME-M or control naïve B cells from human ileum. (B) FCM of human gut microbiota stained with the DNA-labeling compound SytoBC and further reacted with differently fluorochrome-labeled isotype-matched control antibodies or antibodies to IgM or IgA. Numbers indicate % of positive cells. (C) FCM analysis of SIgM-bound gut microbiota using various monoclonal antibodies or a polyclonal antibody to human IgM. (D) FCM of SIgM and SIgA bound to viable SYTO BC^+ bacteria from the small intestine of C57BL/6 mice housed under SPF and non-SPF conditions. (E) FCM of PC-M and PC-A from the small intestine of C57BL/6 mice housed under SPF and non-SPF conditions. Data summarize 3 experiments (A), show one representative experiment of at least 3 with similar results (B,C), or show 2 experiments with at least 3 mice/group (D). Error bars, s.e.m. $^{**}p < 0.01$ (two-tailed unpaired Student's t-test).

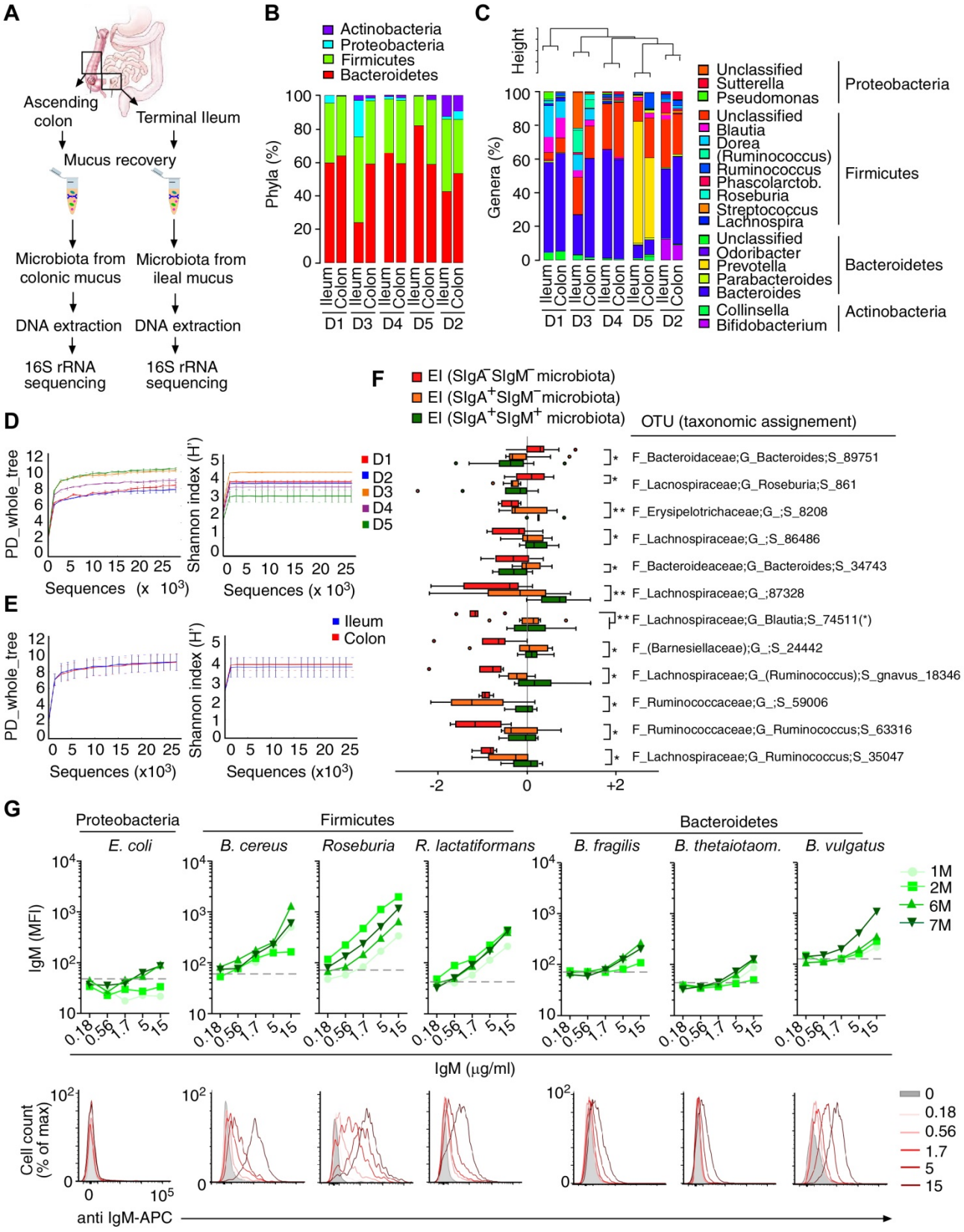

A
Ascending colon
Terminal Ileum
Mucus recovery
Microbiota from colonic mucus
Microbiota from ileal mucus
DNA extraction
DNA extraction
16S rRNA sequencing
16S rRNA sequencing
B
Actinobacteria
Proteobacteria
Firmicutes
Bacteroidetes
Phyla (%)
Ileum Colon
D1 D3 D4 D5 D2
C
Height
Genera (%)
Unclassified
Sutterella
Pseudomonas
Proteobacteria
Unclassified
Blautia
Dorea
(Ruminococcus)
Ruminococcus
Phascolarctob.
Roseburia
Streptococcus
Lachnospira
Firmicutes
Unclassified
Odoribacter
Prevotella
Parabacteroides
Bacteroides
Bacteroidetes
Collinsella
Bifidobacterium
Actinobacteria
D
PD_whole_tree
Sequences (x 10^3)
Shannon index (H')
Sequences (x 10^3)
D1
D2
D3
D4
D5
E
PD_whole_tree
Sequences (x10^3)
Shannon index (H')
Sequences (x10^3)
Ileum
Colon
F
EI (SIgA-SIgM- microbiota)
EI (SIgA+SIgM- microbiota)
EI (SIgA+SIgM+ microbiota)
OTU (taxonomic assignement)
F_Bacteroidaceae;G_Bacteroides;S_89751
F_Lacnospiraceae;G_Roseburia;S_861
F_Erysipelotrichaceae;G_;S_8208
F_Lachnospiraceae;G_;S_86486
F_Bacteroideaceae;G_Bacteroides;S_34743
F_Lachnospiraceae;G_;87328
F_Lachnospiraceae;G_Blautia;S_74511(*)
F_(Barnesiellaceae);G_;S_24442
F_Lachnospiraceae;G_(Ruminococcus);S_gnavus_18346
F_Ruminococcaceae;G_;S_59006
F_Ruminococcaceae;G_Ruminococcus;S_63316
F_Lachnospiraceae;G_Ruminococcus;S_35047
G
Proteobacteria
Firmicutes
Bacteroidetes
E. coli
B. cereus
Roseburia
R. lactatiformans
B. fragilis
B. thetaiotaom.
B. vulgatus
IgM (MFI)
1M
2M
6M
7M
IgM (µg/ml)
Cell count (% of max)
anti IgM-APC
0
0.18
0.56
1.7
5
15

Figure S7 (Related to Figure 7). Human Mucus from Ileal and Colonic Samples Shows Inter-Individual Variability, But Predominantly Contains Bacteroidetes and Firmicutes.

(A) Experimental strategy followed to perform 16S rRNA gene analysis in mucus samples scrapped from the epithelial surface of human ileum or colon specimens. (B and C) Relative composition of phyla (B) and genera (C) in ileal and colonic mucus samples. Each bar corresponds to one sample and the 16 most abundant genera are shown. Unclassified refers to the Greengenes reference database. Hierarchical clustering was applied to group samples according to relative genera composition. D, donor. (D) Phylogenetic diversity (left) and species diversity (right) of mucus-embedded microbiota samples calculated by PD_whole_tree and Shannon metrics, respectively. (E) Phylogenetic diversity (left) and species diversity (right) of mucus-embedded microbiota samples from human ileum and colon calculated as in (D). (F) Boxplot distribution of Enrichment Index (EI) for selected OTUs with statistically significant differences among $SIgA^{-}SIgM^{-}$, $SIgA^{+}SIgM^{-}$ and $SIgA^{+}SIgM^{+}$ microbiota fractions. (G) FCM of IgM from EBV-transformed human ileum ME-M B cells bound to SYTO BC^{+} bacterial isolates. A summary of the data (top) and representative histograms (bottom) are shown. Dashed line indicates background fluorescence. Data are from two distinct gut segments of 5 (B-F) donors or from ileal ME-M B cells of 4 donors (G). Error bars, s.d. *p <0.05, **p <0.01, ***p <0.001 (Welch's t-test).

Table S1 (Related to STAR METHODS). Primers for RT-PCR

Target gene		Primer sequence
ACTB	S	GGATGCAGAAGGAGATCACT
	AS	CGATCCACACGGAGTACTTG
AICDA	S	AGAGGCGTGACAGTGCTACA
	AS	TGTAGCGGAGGAAGAGCAAT
CD70	S	TACGTATCCATCGTGATG
	AS	GTTGGTGCAGAGTGTGTC
FCRL4	S	TCAGCTGGGAGAAGAAGAGGAA
	AS	GAGTTATCTGGGTGTTGTGTCTTTACC
GPR34	S	GAAAGGTTGCGACTATTACCAAA
	AS	GTGAAAGTGCTAAATGACATATTCCTC
IL10	S	ACCTGCCTAACATGCTTCGAG
	AS	TGTCCAGCTGATCCTTCATTTG
IRF4	S	GCCAAGATTCCAGGTGACTC
	AS	ATCGTAGCCCCTCAGGAAAT
PRDM1	S	GTGGTATTGTCGGGACTTTGCAG
	AS	TCGGTTGCTTTAGACTGCTCTGTG
RUNX2	S	CCAACCCACGAATGCACTATC
	AS	TAGTGAGTGGTGGCGGACATAC
SIGLEC6	S	AAGGGGCTGATGTTCCAGTG
	AS	ATGCAGCATTGTCCCTCCTC
TNFRSF13B	S	CAGACAACTCGGGAAGGTACC
	AS	GCCACCTGATCTGCACTCAGCTTC
BACH2	S	CCAGCAATGACTCAGGCATC
	AS	TCATGAGTCTTGTCGCTGGT
FOXO1	S	TTATGACCGAACAGGATGATCTTG
	AS	TGTTGGTGATGAGAGAAGGTTGAG
BCL2	S	CCTCCCTGGCCTGAAGAAGA
	AS	TGTCCTTCGGCGTGGAAATC

Table S2 (Related to STAR METHODS). Primers for CSR

Target gene		Primer sequence
Iμ	S	GTGATTAAGGAGAAACACTTTGAT
Cμ 243	AS	CACACCACGTGTTCGTCTG
Cμ 268	AS	GTTGCCGTTGGGGTGCTGGAC
Iα 1/2	S	CAGCAGCCCTCTTGGCAGGCAGCCAG
Iα internal	S	CTCAGCACTGCG GGCCCTCCA
Cα1	AS	GGGTGGCGGTTAGCGGGGTCTTGG
Cα2	AS	TGTTGGCGGTTAGTGGGGTCTTCA

Table S3 (Related to STAR METHODS). Number of Sorted Cells and Filtered Reads for Ig Repertoire Analysis

Donor #	Tissue	Sample	# Sorted cells	# Filtered reads
1	Ileum	ME-M	232000	205300
1	Ileum	ME-A	221000	113424
1	Ileum	PC-M	12127	189639
1	Ileum	PC-A	64880	186256
1	Colon	ME-M	150000	90391
1	Colon	ME-A	129000	209843
1	Colon	PC-M	31000	208868
1	Colon	PC-A	268000	103512
2	Ileum	ME-M	220000	201540
2	Ileum	ME-A	280000	251415
2	Ileum	PC-M	52000	105529
2	Ileum	PC-A	200000	98980
2	Colon	ME-M	42000	90107
2	Colon	ME-A	44000	218355
2	Colon	PC-M	36800	250019
2	Colon	PC-A	200000	290760
3	Ileum	ME-M	500000	378724
3	Ileum	ME-A	500000	320316
3	Ileum	PC-M	300000	236172
3	Ileum	PC-A	300000	123548
3	Colon	ME-M	267000	197134
3	Colon	ME-A	500000	134617
3	Colon	PC-M	40000	366219
3	Colon	PC-A	500000	443847
4	Ileum	ME-M	180000	204923
4	Ileum	ME-A	500000	140012
4	Ileum	PC-M	45000	56453
4	Ileum	PC-A	170000	88109
4	Colon	ME-M	20000	139985
4	Colon	ME-A	100000	134624
4	Colon	PC-M	33000	3194
4	Colon	PC-A	500000	160899

Table S4 (Related to STAR METHODS). Primers for Next Generation Sequencing

Target gene		Primer sequence
VH1a	S	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAGGTKCAGCTGGTGCAG
VH1b	S	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGSAGGTCCAGCTGGTACAG
VH1c	S	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCARATGCAGCTGGTGCAG
VH2	S	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAGGTCACCTTGARGGAG
VH3	S	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGTCCCTGAGACTCTCCTGT
VH4	S	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACCCTGTCCCTCACCTGC
VH5	S	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGCAGCTGGTGCAGTCTGGAG
VH6	S	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAGGACTGGTGAAGCCCTCG
VH7	S	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAGGTGCAGCTGGTGCAA
Cμ	AS	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCAGGAGACGAGGGGGAAAAGG
Cα	AS	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGAAGACCTTGGGGCTGGTCG
16S	S	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG
16S	AS	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rat anti human AID (clone: EK2 5G9)	Cell Signaling	4959S RRID:AB_10692771
Anti-human CD1c PE (clone: L1S1)	Biolegend	331505 RRID:AB_1089000
Anti-human CD10 PE (clone: HI10a)	Biolegend	312203 RRID:AB_312203
Anti-human CD10 APC/Cy7 (clone: HI10a)	Biolegend	312212 RRID:AB_2146550
Anti-human CD11c PE (clone: B-ly6)	BD Biosciences	555392 RRID:AB_395793
Anti-human CD19 PE/Cy7 (clone: HIB19)	Biolegend	302215 RRID:AB_314245
Anti-human CD20 FITC (clone: 2H7)	BD Biosciences	555622 RRID:AB_395988
Anti-human CD21 PE(clone: B-ly4)	BD Biosciences	555422 RRID:AB_395816
Anti-human CD24 PE (clone: ML5)	BD Biosciences	555428 RRID:AB_395822
Anti-human CD27 PE (clone: O323)	eBioscience	12-0279-42 RRID:AB_10718394
Anti-human CD27 PerCP/Cy5.5 (clone: M-T271)	BD Biosciences	560612 RRID:AB_1727457
Anti-human CD35 PE (clone: E11)	Biolegend	333405 RRID:AB_2085021
Anti-human CD38 APC (clone: HIT2)	BD Biosciences	555462 RRID:AB_398599
Anti-human CD38 APC/Cy7 (clone: HIT2)	Biolegend	303534 RRID:AB_2561605
Anti-human CD43 PE (clone: 1G10)	BD Biosciences	560199 RRID:AB_1645655
Anti-human CD45 Alexa Fluor 700 (clone: HI30)	Biolegend	304024 RRID:AB_493761
Anti-human CD62L PE (clone: DREG-56)	Biolegend	304805 RRID:AB_314465
Anti-human CD66 a/c/e PE (clone: ASL-32)	Biolegend	342303 RRID:AB_1626288
Anti-human CD84 PE (clone: CD84.1.21)	Biolegend	326007 RRID:AB_2074766
Anti-human CD138 APC (clone: DL-101)	Biolegend	352307 RRID:AB_10901175
Anti-human CD138 PerCp/Cy5.5 (clone: DL-101)	Biolegend	352310 RRID:AB_10900979
Anti-human DEP-1/CD148 PE-conjugated (clone: 143-41)	R&D Systems	FAB1934P RRID:AB_2174832
Anti-human CCR10 PE (clone: 6588-5)	Biolegend	341503 RRID:AB_1595542
Anti-human CCR7 PE (clone: G043H7)	Biolegend	353203 RRID:AB_10916391
Anti-human CCR9 PE (clone: L053E8)	Biolegend	358903 RRID:AB_2562384
Anti-human CXCR4 PE (clone: 12G5)	Biolegend	306505 RRID:AB_314611
Anti-human CXCR4 APC (clone: L291H4)	Biolegend	359407 RRID:AB_2562428

Anti-human FcRL4 APC (clone: 413D12)	Biolegend	340205 RRID:AB_10710013
Anti-human FcRL4 PE (clone: 413D12)	Biolegend	340203 RRID:AB_1575103
Anti-human HLA-DR Alexa Fluor 700 (clone: L243)	Biolegend	307626 RRID:AB_493771
Anti-human HLA-DR Brilliant Violet 605 (clone:L243)	Biolegend	307639 RRID:AB_11219187
Anti-human Ki-67 (clone: MIB-1)	DAKO	M7240 RRID:AB_2142367
Anti-human Ig light chain λ (clone: MHL-38)	Biolegend	316607 RRID:AB_493626
Anti-human IgA APC (clone: IS118E10)	Miltenyi Biotec	130-093-113 RRID:AB_1036152
Anti-human IgA-BIOT	Southern Biotech	2052-08
Anti-human IgA FITC (clone: IS118E10)	Invitrogen	H14101 RRID:AB_2536555
Anti-human IgA PE (clone: IS118E10)	Miltenyi Biotec	130-093-128 RRID:AB_1036158
FLEX Polyclonal Rabbit Anti-Human IgD Ready-to-Use	DAKO	IS517
Anti-human IgD APC (clone: IA6-2)	BD Biosciences	561303 RRID:AB_10642578
Anti-human IgD PerCP/Cy5.5 (clone: IA6-2)	Biolegend	348208 RRID:AB_10641706
Anti-human IgD FITC	Southern Biotech	2032-02
Anti-human IgM Fc Secondary	Invitrogen	H15000 RRID:AB_2536556
Anti-human IgM APC (clone: SA-DA4)	Southern Biotech	9020-11
Anti-human IgM Brilliant Violet 605 (clone: MHM-88)	Biolegend	314523 RRID:AB_2562373
Anti-human/mouse Integrin β7 FITC (clone: FIB504)	Biolegend	321213 RRID:AB_830857
Anti-mouse IgD FITC (clone: 11-26c.2a)	BD Biosciences	558597 RRID:AB_647211
Anti-mouse CD45 PE/Cy7 (clone: 30-F11)	Biolegend	103113 RRID:AB_312978
Anti-mouse/human CD45R PE/Cy7 (clone: RA3-6B2)	Biolegend	103229 RRID:AB_492875
Anti-mouse CD138 Biotin (clone: 281-2)	Biolegend	142511 RRID:AB_2561980
Goat F(ab')2 Anti-Human IgA-PE	Southern Biotech	2052-09
Anti-mouse IgD Brilliant Violet 605 (clone: 11-26c.2a)	Biolegend	405727 RRID:AB_2562887
Anti-mouse IgM APC (clone: II/41)	BD	550676 RRID:AB_398464
Anti-mouse IgM FITC (clone: RMM-1)	Biolegend	466505 RRID:AB_315056
Anti-mouse IgA Secondary	Novus	NB7506 RRID:AB_10125106
Goat Anti-Mouse IgM-UNLB	Southern Biotech	1021-01
Goat Anti-Human Ig-UNLB	Southern Biotech	2010-01
Goat Anti-Human IgA-HRP	Southern Biotech	2050-05
Goat Anti-human IgM-HRP	Cappel	55255 RRID:AB_2334397
Bacterial and Virus Strains		
Escherichia coli	ATCC	25992
Bacillus cereus	ATCC	11778
Roseburia intestinalis	DSMZ	14610

Ruthenibacterium lactatiformans	DSMZ	100348
Bacteroides vulgatus	ATCC	8482
Bacteroides fragilis	ATCC	25285
Bacteroides thetaiotaomicron	ATCC	29148
Biological Samples		
Histologically normal tissue samples from terminal ileum and ascending colon	Patients undergoing right hemicolectomy-Hospital del Mar	N/A
Histological normal spleens	Hospital Clinic	N/A
Tonsils	Patients undergoing tonsillectomy-Hospital del Mar	N/A
Blood samples	BST-Barcelona	N/A
Chemicals, Peptides, and Recombinant Proteins		
4',6-Diamidine-2-phenylindole dihydrochloride (DAPI)	Sigma-Aldrich	28718-90-3
Ficoll-Paque™ PLUS	GE Healthcare	17144003
Collagenase IV	ThermoFischer	17104019
DNAse	New England Biolab	M0303S
β-D-glucan from barley	Sigma	9041-22-9
Capsular polisaccharides (serotypes 9N, 14, 19F, 23F)	ATCC	N/A
Gal-α1,3-gal-HSA (3 atom spacer)	Dextra Laboratories	NGP2203
Laminarin from Laminaria digitata	Sigma-Aldrich	9008-22-4
L-α-Phosphatidylcholine	Sigma-Aldrich	8002-43-5
megaCD40L	Enzo Life science	ALX-522-110-C010
TMB Substrate reagent set	BD bioscience	555214
IL-10	Peprotech	200-10
IL-21	Peprotech	200-21
CpG ODN-2006	ThermoFischer	tlrl-2006
BAFF (soluble)(human),(recombinant)	Enzo	ALX-522-025-C010
MEGACD40L®(soluble)(human),(recombinant)	Enzo	ALX-522-110-C010
APRIL (human) (H98) (multimeric)	Adipogen	AG-40B-0088
FcR Blocking Reagent, human	Miltenyi	130-059-901
FluorSave reagent	Merck Millipore	345789
Syto BC	ThermoFischer	S34855
Critical Commercial Assays		
Giemsa Stain Kit (Jenner-Wright)	Agilent	AR308
CellTrace CFSE cell proliferation Kit	ThermoFisher	C34554
TaqMan® Reverse Transcription Reagents	ThermoFisher	N8080234
Nextera XT DNA Sample Preparation Kits	Illumina	1502354
Pure Link Microbiome DNA Purification Kit	ThermoFisher	A29790
Quiamp DNA Mini Kit	Qiagen	Cat# 51304
RNeasy Micro Kit	Qiagen	Cat# 74004
AmpliTaq Gold PCR mastermix	ThermoFisher	4398881
Power SYBR® Green PCR Master Mix	ThermoFisher	4367659
Deposited Data		
Microarray data	This Paper	GSE89282
NGS data	This Paper	PRJNA355402
Experimental Models: Cell Lines		

EBV producing marmoset B cell line (B95-8)	ECACC	85011419
EBV transformed cell lines	This Paper	N/A
Experimental Models: Organisms/Strains		
Oligonucleotides		
Primers for RT-PCR see Table S1	This paper	N/A
Primers for CSR see Table S2	This paper	N/A
Primers for NGS see table S4	This paper	N/A
Recombinant DNA		
Software and Algorithms		
FlowJo v10.0.7	FlowJO	https://www.flowjo.com/
Fuji (ImageJ)	ImageJ	https://fiji.sc/
Gene Set Enrichment Analysis	(Subramanian et al., 2005)	http://software.broadinstitute.org/gsea/index.jsp
limmaGUI	(Wettenhall and Smyth, 2004)	https://www.bioconductor.org/packages/release/bioc/html/limmaGUI.html
pRESTO	(Vander Heiden et al., 2014)	https://presto.readthedocs.io/en/version-0.5.3---license-change/
IgBLAST	(Ye et al., 2013)	https://www.ncbi.nlm.nih.gov/igblast/
Change-O	(Gupta et al., 2015)	https://changeo.readthedocs.io/en/version-0.3.4---license-change/
R package	N/A	http://www.R-project.org/
IgTree Software	(Barak et al., 2008)	http://immsilico2.lnx.biu.ac.il/Software.html
QUIIME	(Caporaso et al., 2010)	http://qiime.org/
Prism v.6.0	GraphPad	www.graphpad.com

Ingenuity Pathway Analysis	QIAGEN	www.qiagenbioinformatics.com
Other		