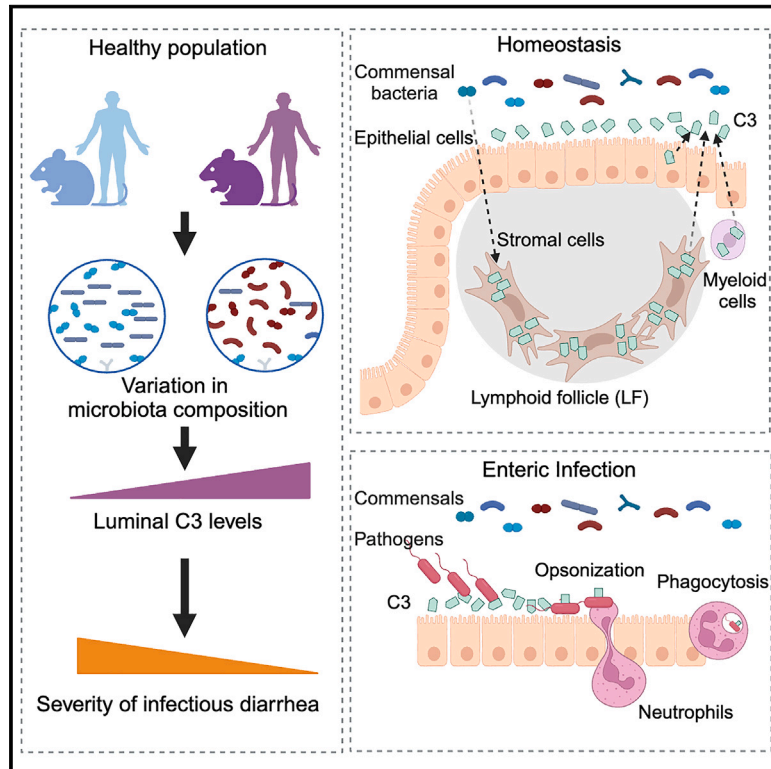


# Gut complement induced by the microbiota combats pathogens and spares commensals

## Graphical abstract



## Authors

Meng Wu, Wen Zheng, Xinyang Song, ..., John J. Mekalanos, Christophe Benoist, Dennis L. Kasper

## Correspondence

dennis\_kasper@hms.harvard.edu

## In brief

Long thought to be a defense system unique to the bloodstream, complement is shown here to be synthesized locally in the mammalian intestine, where it helps tune the tolerance to the gut microbiota while combating pathogenic bacteria.

## Highlights

- The gut has an independent complement system synthesized by intestinal cells
- In mice and humans, C3 is induced by the microbiota and varies between individuals
- Basal luminal C3 levels are inversely correlated with the severity of infectious diarrhea
- Stromal cells located in intestinal lymphoid follicles are a main source of luminal C3



## Article

# Gut complement induced by the microbiota combats pathogens and spares commensals

Meng Wu,<sup>1,7,8</sup> Wen Zheng,<sup>1,8</sup> Xinyang Song,<sup>1</sup> Bin Bao,<sup>2</sup> Yuanyou Wang,<sup>3,4</sup> Deepshika Ramanan,<sup>1</sup> Daping Yang,<sup>1</sup> Rui Liu,<sup>5</sup> John C. Macbeth,<sup>5</sup> Elyza A. Do,<sup>5</sup> Warrison A. Andrade,<sup>6</sup> Tiandi Yang,<sup>1</sup> Hyoung-Soo Cho,<sup>1</sup> Francesca S. Gazzaniga,<sup>1</sup> Marit Ilves,<sup>1</sup> Daniela Coronado,<sup>1</sup> Charlotte Thompson,<sup>1</sup> Saiyu Hang,<sup>5</sup> Isaac M. Chiu,<sup>1</sup> Jeffrey R. Moffitt,<sup>3,4</sup> Ansel Hsiao,<sup>5</sup> John J. Mekalanos,<sup>4</sup> Christophe Benoist,<sup>1</sup> and Dennis L. Kasper<sup>1,9,\*</sup>

<sup>1</sup>Department of Immunology, Harvard Medical School, Boston, MA 02115, USA

<sup>2</sup>Division of Gastroenterology, Boston Children's Hospital, and Harvard Medical School, Boston, MA 02115, USA

<sup>3</sup>Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA 02115, USA

<sup>4</sup>Department of Microbiology, Harvard Medical School, Boston, MA 02115, USA

<sup>5</sup>Department of Microbiology & Plant Pathology, University of California, Riverside, CA 92521, USA

<sup>6</sup>Genentech LLC, South San Francisco, CA 94080, USA

<sup>7</sup>Present address: Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO 63110, USA

<sup>8</sup>These authors contributed equally

<sup>9</sup>Lead contact

\*Correspondence: [dennis\\_kasper@hms.harvard.edu](mailto:dennis_kasper@hms.harvard.edu)

<https://doi.org/10.1016/j.cell.2023.12.036>

## SUMMARY

Canonically, the complement system is known for its rapid response to remove microbes in the bloodstream. However, relatively little is known about a functioning complement system on intestinal mucosal surfaces. Herein, we report the local synthesis of complement component 3 (C3) in the gut, primarily by stromal cells. C3 is expressed upon commensal colonization and is regulated by the composition of the microbiota in healthy humans and mice, leading to an individual host's specific luminal C3 levels. The absence of membrane attack complex (MAC) components in the gut ensures that C3 deposition does not result in the lysis of commensals. Pathogen infection triggers the immune system to recruit neutrophils to the infection site for pathogen clearance. Basal C3 levels directly correlate with protection against enteric infection. Our study reveals the gut complement system as an innate immune mechanism acting as a vigilant sentinel that combats pathogens and spares commensals.

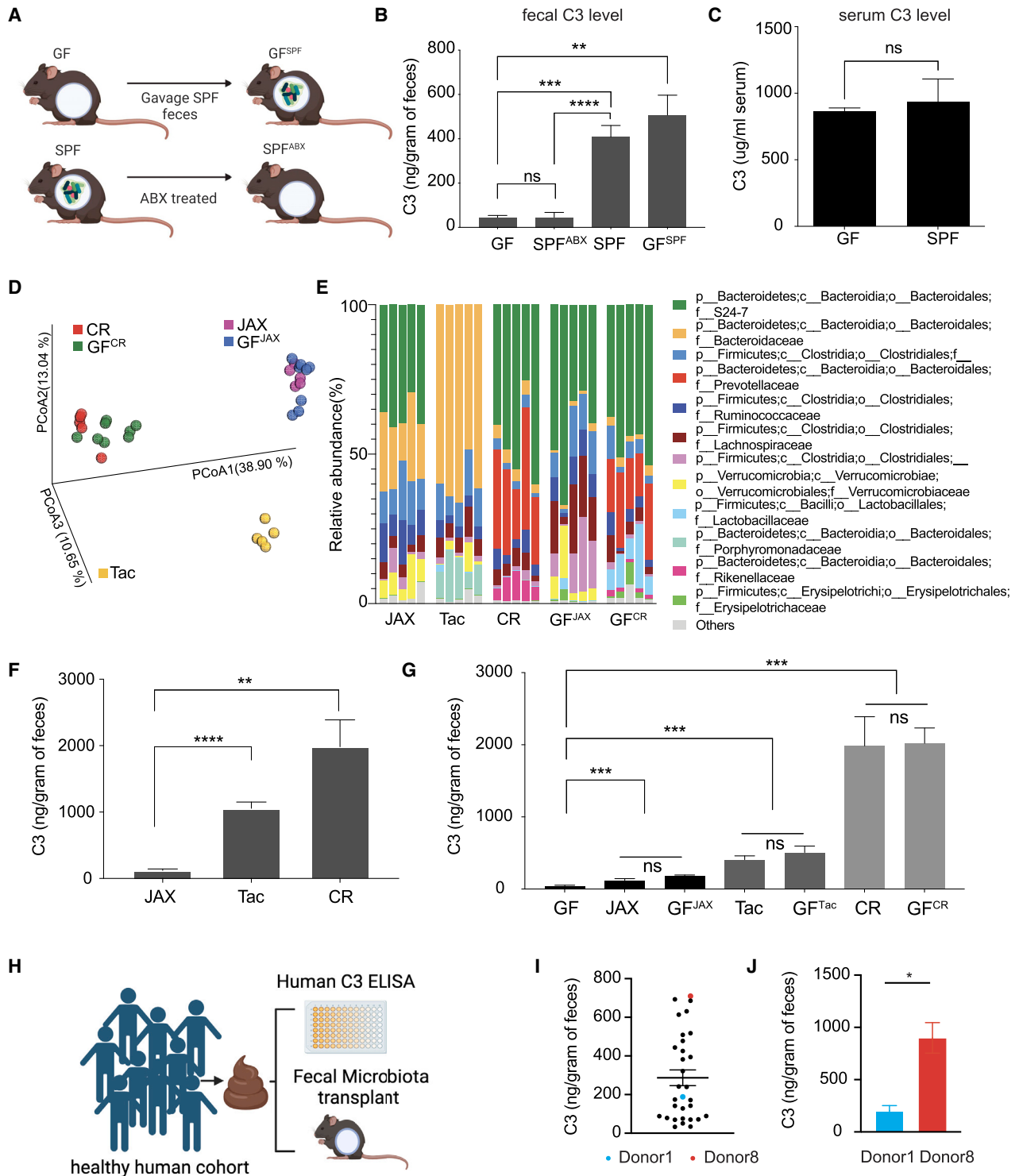
## INTRODUCTION

The complement system comprises more than fifty secreted proteins and membrane receptors and regulators that function in a highly coordinated fashion to form one of the first lines of innate immune defense against the invasion of host tissues by microbes. In blood, complement can be activated through the classical, lectin, and alternative pathways, which converge at complement component 3 (C3), the central element of the complement cascade. Deposition of a fragment of C3 (C3b) on the surface of bacterial cells is essential for promoting phagocytosis and killing by neutrophils and macrophages, as well as direct lysis of many gram-negative bacteria by activation of the membrane attack complex (MAC). Complement is classically considered a hepatocyte-derived and plasma-resident system, acting in both the blood and interstitial fluids.<sup>1</sup> Studies have demonstrated that C3 is also produced outside the liver by immune and nonimmune cells.<sup>2–8</sup> During the COVID-19 pandemic, studies found that the complement system is one of the most highly induced pathways in lung epithelial cells following

SARS-CoV-2 infection.<sup>9</sup> Complement cascade genes are also enriched among differentially upregulated genes in rectal biopsies from patients with inflammatory bowel disease (IBD) relative to healthy controls.<sup>10</sup> However, the scale and function of the complement system within the gastrointestinal tract remain unclear. The gut is densely populated with microbiota, including the bacteria, archaea, fungi, protists, and viruses. If a complement system were present in the gut, how would a functional complement system establish a harmonious co-existence with resident commensal organisms on the one hand and contribute to elimination of pathogens during infections on the other?

In this study, we show that complement protein C3, the central component of complement system, is synthesized by intestinal cells and secreted into the lumen. Luminal C3 levels increase during colonization with commensals and are regulated by the composition of the host's gut microbiota leading to individual host-specific luminal C3 levels. We report that the intestinal complement system operates as an independent and autonomous entity, distinguished from the circulatory complement system through its selective expression of specific complement





**Figure 1. Gut complement factor C3 is regulated by the host microbiota**

(A) Mouse models used to evaluate fecal C3 levels in (B). GF, C57BL/6 germ-free mice; SPF, specific-pathogen-free C57BL/6 mice harbor a conventional murine microbiota; SPF<sup>ABX</sup>, SPF mice treated with antibiotics for 2 weeks; GF<sup>SPF</sup>, germ-free (GF) mice colonized with SPF microbiota.

(B) Fecal C3 levels in GF, SPF<sup>ABX</sup>, SPF, and GF<sup>SPF</sup>.

(C) C3 levels in serum of GF and SPF mice.

(legend continued on next page)

components, thus allowing for precise modulation of responses toward commensals and pathogens. Microbiota-dependent complement C3 production in the gut underpins a relationship between commensal gut microbes while also playing an important role in the pathogenesis of infectious diarrhea. Understanding the complement system on mucosal surfaces could pave the way for microbiota-targeted precision medicine approaches to prevent and treat enteric diseases.

## RESULTS

### C3 is produced locally within the intestine in a microbiota-dependent manner

To investigate the role of the commensal gut microbiota in host complement production, we measured C3 protein levels in fecal and serum samples of age- and gender-matched wild type (WT) germ-free (GF), and specific-pathogen-free (SPF) C57BL/6 mice by ELISA (Figure 1A). We found that SPF mice have significantly higher levels of intestinal C3 than GF animals while the C3 levels in the serum of GF mice were comparable with those of SPF mice (Figures 1B and 1C). This result suggested that there are microbiota-dependent C3 differences in the gut lumen, and these differences are not due to systemic changes in serum complement levels.

To further test whether the production of C3 in the lumen could be triggered by exposure to microbes postweaning or is programmed during early development, we colonized 4-week-old GF mice with SPF microbiota (Figure 1A). 2 weeks after introducing an SPF mouse fecal slurry into GF mice (GF<sup>SPF</sup>), there were comparable levels of C3 in fecal specimens from the GF<sup>SPF</sup> mice as in fecal samples of SPF animals. We examined whether removing members of the gut microbiota reduced fecal C3 levels. After 2 weeks of broad-spectrum antibiotic treatment, antibiotic-treated SPF mice (SPF<sup>ABX</sup>) had no detectable live bacteria in their feces by culturing under both aerobic and anaerobic conditions, and fecal C3 levels were like those of GF mice (Figure 1B).

We investigated whether luminal C3 levels were affected by the composition of the microbiota. We measured C3 levels in 8- to 10-week-old C57BL/6 mice from three commercial breeders: Jackson Laboratory (JAX), Taconic (Tac), and Charles River (CR). We found that inbred animals of the same strain from each vendor harbored a distinct microbiota (Figures 1D, 1E, and S1A–S1C) and had significantly different concentrations of fecal C3 (Figure 1F). We then transplanted the gut microbiota from each of these three mouse vendors into GF mice and observed that colonization and C3 levels recapitulated donor signatures

(Figure 1G). Oral gavage of FITC-dextran in CR mice showed similar barrier permeability compared with JAX mice (Figure S1D). Blind histology check scored both JAX and CR mice at zero with no histologic sign of disease (Figure S1E). These data suggest that the composition of the microbiota regulates levels of gut C3 production and secretion.

We subsequently investigated whether our findings were relevant to humans. We measured fecal C3 levels from a cohort of healthy human donors without chronic gut inflammation<sup>11</sup> and found substantial individual variation in C3 levels in these fecal samples (range of 33–710 ng/g of feces) (Figures 1H and 1I). We transplanted the fecal microbial communities from either a high-C3-producing (donor 8) or a low-C3-producing (donor 1) human donor into GF mice and measured C3 levels in mouse fecal specimens 2 weeks after colonization (Figure 1H). Importantly, gnotobiotic mice colonized with donor 1 fecal slurry had low fecal C3 levels while mice colonized with donor 8 microbiota had high fecal C3 levels (Figure 1J). These findings suggest that microbiota from different humans can induce different levels of fecal C3 in the mouse gut.

### *Prevotella* spp. are one of the drivers of high-C3 production in the gut

To identify which specific bacterial species contribute to the induction of high C3 levels in SPF animals, we examined the mouse microbiota of JAX and CR as representative of low-C3 and high-C3 communities, respectively. We co-housed 4-week-old JAX and CR mice for 2 weeks and found that the co-housed JAX mice had increased fecal C3 levels to that of CR mice, while the co-housed CR mice retained their original baseline C3 production (Figure 2A). 16S rRNA gene sequencing of fecal samples showed that co-housed JAX mice acquired a microbiota very similar to the CR microbiota (Figures 2B and S2A–S2D), which suggested the presence of CR-specific strains that could induce high C3 production. To identify possible candidate strains, we employed a method known as “triangulation of microbe-phenotype relationships.”<sup>12</sup> In brief, we co-housed different groups of 4-week-old JAX and CR mice for 1 or 3 days to allow for microbiota transfer. We found that co-housing JAX mice with CR mice for either 1 or 3 days resulted in C3 levels similar to those seen in mice gavaged with CR feces (Figure 2C). 16S rRNA gene sequencing revealed co-housing JAX and CR mice for 1 day resulted in CR-like signatures in recipient JAX mice (JAX<sup>coh\_1 day</sup>) both in terms of the presence of bacterial taxa ( $\alpha$ -diversity) and the relative abundance of taxa ( $\beta$ -diversity) (Figures 2B and S2A–S2D). Differential abundance analysis of the microbiota composition of JAX and JAX<sup>coh\_1 day</sup> mice

(D) Principal coordinate analysis (PCoA) of unweighted UniFrac distance measurements based on the 16S rRNA gene sequencing data of the composition of bacterial communities in fecal samples collected from C57BL/6 mice from three commercial vendors: JAX, Tac, and CR; or GF mice colonized with JAX (GF<sup>JAX</sup>) or CR microbiota (GF<sup>CR</sup>) ( $q = 0.001$ , PERMANOVA).

(E) Family level abundance analysis of fecal microbiota from the mice in (D).

(F) Fecal C3 levels in C57BL/6 mice from JAX, Tac, and CR.

(G) Fecal C3 levels of gnotobiotic mice colonized for 2 weeks with fecal samples from the indicated commercial sources.

(H) Experimental scheme to evaluate fecal C3 levels in a healthy human cohort and their relationship to the gut microbiota.

(I) Fecal C3 levels in 31 healthy adult human donors.

(J) Fecal C3 levels in gnotobiotic mice monocolonized with fecal microbiota from either human donor 1 or donor 8.

ns, not significant, \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ , unpaired Student's  $t$  test. Error bars represent mean  $\pm$  SEM.

See also Figure S1.



identified *Prevotella* spp., *Rikenellaceae* spp., and *Lactobacillus salivarius* as significantly more abundant in the microbiota of JAX<sup>coh\_1 day</sup> compared with JAX mice (Figures 2D, 2E, S2E, and S2F; Table S1). This suggested that these bacteria could be associated with high fecal C3 levels and be potential high-C3 inducers.

To identify commensal strains that could also correlate with C3 production in humans, we compared C3 levels of individual human fecal samples with commensal taxa. Due to the extensive individual variation of bacterial species, we performed a genus-level analysis. We found that the genus *Prevotella* also positively correlated with fecal C3 levels across the cohort (Spearman  $r = 0.5487$ ,  $p = 0.029$ ), while the genus *Actinomyces* was negatively correlated with C3 levels (Spearman  $r = -0.6991$ ,  $p = 0.0025$ ) (Table S2).

Given the positive correlation between the abundance of genus *Prevotella* and high fecal C3 levels in both humans and mice, we sought to further examine its role in C3 production. We first isolated a *Prevotella* spp. (strain CR1) from CR mouse feces using selective growth conditions<sup>13</sup> and colonized 4-week-old SPF JAX mice with CR1 by oral gavage. We used *Proteus mirabilis* (strain CR2) isolated from CR mice as a control because it was only present in the CR microbiota; the presence of CR2 showed no significant positive correlation with C3 induction in our analysis. We found that in contrast to mice colonized with *P. mirabilis* CR2, JAX mice colonized with *Prevotella* strain CR1 showed a significant increase of fecal C3 production compared with untreated JAX control mice. Nevertheless, colonization with *Prevotella* CR1 did not entirely recapitulate the effect of a CR-microbiota gavage, suggesting that there are additional bacterial strains in CR mice that could induce high C3 production (Figure 2F). These data are consistent with our finding that multiple bacterial strains were identified as significantly positively associated with high fecal C3 levels in JAX<sup>coh\_1 day</sup> mice (Figure 2D). To further understand if this heightened C3 induction relied on interactions between *Prevotella* spp. CR1 and other microbes, we monocolonized GF mice with *Prevotella* spp. CR1 or *P. mirabilis* CR2. We found that *Prevotella* spp. CR1 and *P. mirabilis* CR2 induced comparable levels of C3 production following a 2-week period of monocolonization (Figure 2G).

To comprehensively assess the ability of bacterial strains to stimulate intestinal complement C3 production, we chose fifty distinct bacterial strains derived from humans, representing the broad diversity of human gut microbiota. Additionally, we included the mouse-derived strain, segmented filamentous bacteria (SFB), known for its potent Th17 response. We then introduced each of these strains individually to 4-week-old GF mice for 2 weeks to assess each microbe's individual capacity to induce complement C3. Fecal C3 levels from these mice revealed that in a monocolonized setting, each of these bacterial strains possess some capacity to trigger the production of intestinal complement C3 (Figure 2H).

Together, these data demonstrate that diverse bacterial strains, when introduced individually to GF mice, have the capacity to independently stimulate intestinal C3 production. Notably, certain strains, such as the *Prevotella* strain CR1 from the CR microbiota can induce elevated luminal C3 levels when introduced into a diverse microbiota environment, and such heightened induction requires interactions with other microorganisms.

### Stromal cells are the major intestinal cell population transcribing complement protein C3 during homeostasis

Although complement proteins in blood are primarily produced and secreted by the liver,<sup>14–16</sup> recent studies have also shown extrahepatic C3 production by various immune and nonimmune cell types.<sup>2–9</sup> While intestinal epithelial cells (IECs) have been suggested to produce C3 under experimental and inflammatory conditions,<sup>17–19</sup> the cellular sources for intestinal C3 at homeostasis remain unclear.

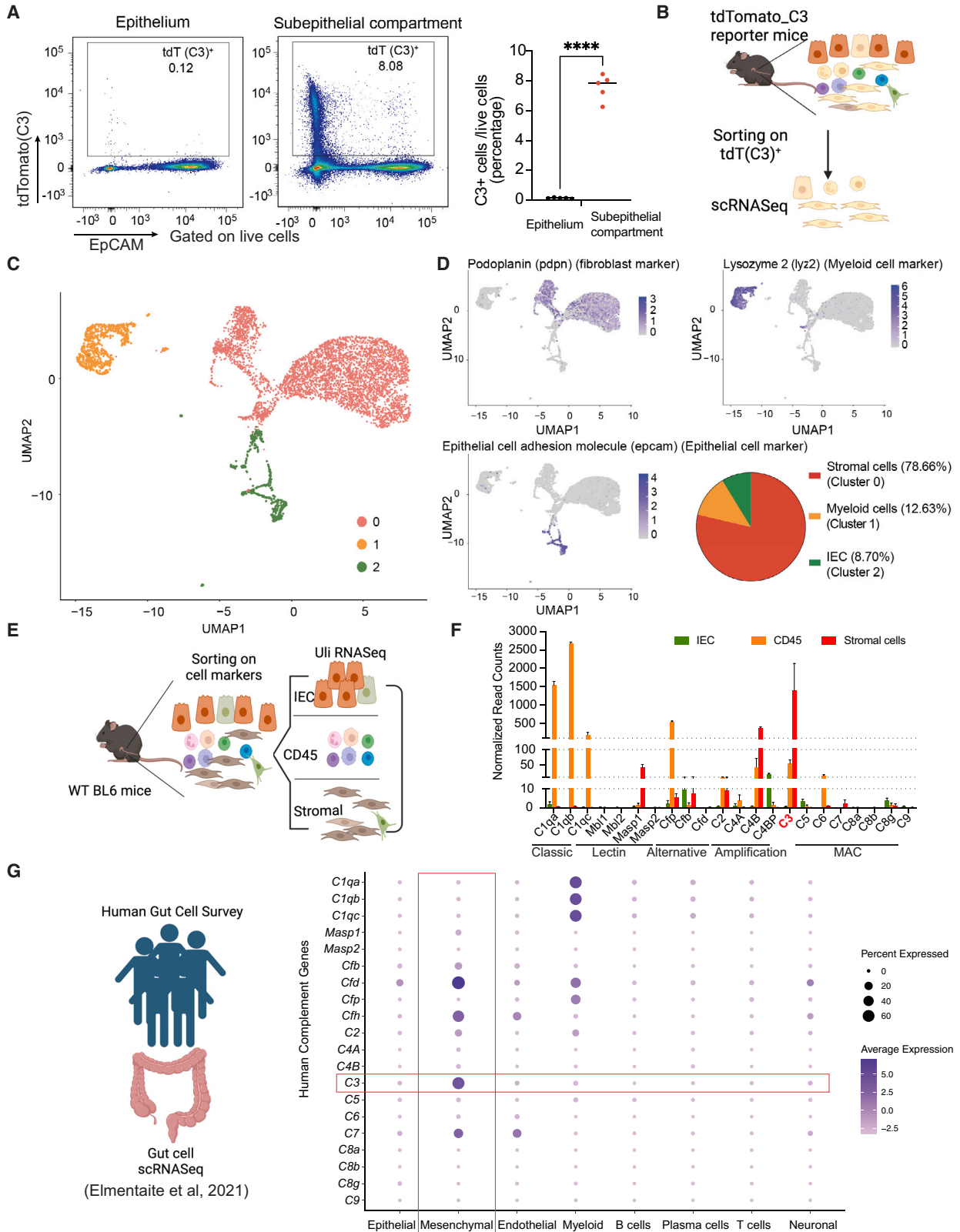
To identify which intestinal cell types express C3 under homeostasis, we used single-cell RNA sequencing (scRNA-seq) to perform a transcriptional survey of C3 expression across intestinal populations. Flow cytometry analysis on colonic cells from the C3<sup>IRES-tdTomato</sup> C57BL/6<sup>7</sup> found the subepithelial compartment, which includes both the lamina propria and the muscularis of the colon, as the dominant source of C3-expressing cells, compared with the epithelium (Figures 3A and S3A). Therefore, we focused on the subepithelial compartment for our scRNA-seq analysis. We isolated C3-expressing cells (tdTomato-positive) from the subepithelial

#### Figure 2. *Prevotella* spp. induce production of high luminal C3 levels in a complex host microbiota

- (A) Fecal C3 levels of JAX or CR mice by themselves or co-housed for 2 weeks with mice from the other vendor (JAX<sup>coh\_CR</sup>, JAX mice co-housed with CR mice; CR<sup>coh\_JAX</sup>, CR mice co-housed with JAX mice). Statistical analyses were performed comparing treated mice with their untreated controls.
- (B) Family level abundance analysis of fecal microbiota from the mice in (A) and (C).
- (C) Fecal C3 levels in JAX or CR mice co-housed for either 1 or 3 days, also in JAX mice gavaged with CR microbiota (JAX<sup>gav</sup>). Statistical analyses were performed comparing treated mice to their untreated control.
- (D) Microbiome alterations at the species level in JAX mice co-housed with CR mice for 1 day (JAX<sup>coh\_1 day</sup>) compared with JAX controls. Colored dots indicate the bacterial abundance differences; lines indicate the 95% confidence interval (CI) calculated using the bootstrap method. Bacterial species were identified by 16S rRNA gene sequencing and analyzed using the Wilcoxon rank-sum test. Red: increase; blue: decrease. Significance was determined by W value calculated using analysis of composition of microbiomes (ANCOM); asterisk indicates the species with significant alterations. Species with a W value above 20 are shown.
- (E) Relative abundance of *Prevotella* spp. in JAX or CR mice, as well as in JAX<sup>coh\_1 day</sup>, or in JAX<sup>gav</sup>.
- (F) Fecal C3 levels in JAX mice compared with JAX mice colonized with either *Proteus mirabilis*, *Prevotella* spp., or CR microbiota.
- (G) Fecal C3 levels in gnotobiotic mice monocolonized with either *Proteus mirabilis* CR2 or *Prevotella* spp. CR1.
- (H) Fecal C3 levels in gnotobiotic mice monocolonized with 50 distinct human-derived bacterial strains and the mouse-derived bacterial strain segmented filamentous bacteria (SFB).

ns, not significant, \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ , unpaired Student's t test. Error bars represent mean  $\pm$  SEM.

See also Figure S2.



(legend on next page)

compartment of both male and female C3<sup>IRES-tdTomato</sup> C57BL/6 mice and performed scRNA-seq on these cells (Figure 3B). Uniform manifold approximation and projection (UMAP) analysis revealed three major cell populations clustered by gene expression profile (Figure 3C). Using established genetic markers, we identified three major C3-expressing populations in the colon: stromal cells (cluster 0, marked by podoplanin [*pdpn*], representing ~79% of C3-expressing cells); myeloid cells (cluster 1, marked by lysozyme 2 [*lyz2*], representing ~12% of C3-expressing cells); and epithelial cells (cluster 2, marked by EpCAM [*epcam*], representing 9% of C3-expressing cells) (Figure 3D).

We then confirmed our observations by flow cytometric analysis (Figure S3B). We confirmed that stromal cells (EpCAM<sup>-</sup>CD45<sup>-</sup>CD31<sup>-</sup>PDPN<sup>+</sup> cells) formed the major C3-transcribing population in the subepithelial compartment under homeostatic conditions, representing 80% of the C3-expressing population. As expected from the scRNA-seq data, CD45<sup>+</sup> cells (~15%) and IECs (~5%) were the populations with the second and third highest C3 expression, respectively (Figure S3C).

To validate our reporter mouse results, we examined the endogenous C3 transcriptional levels of these three colonic cell populations using WT C57BL/6 mice. We used fluorescence-activated cell sorting (FACS) to isolate stromal cells, CD45<sup>+</sup>, and IEC populations from the subepithelial compartment of the colon of WT C57BL/6 mice (Figure 3E). Using bulk RNA-seq analysis, we found C3 transcriptional levels in stromal cells were significantly higher than in IECs and CD45<sup>+</sup> cells (Figure 3F). These results were consistent with our scRNA-seq and flow cytometry findings using C3<sup>IRES-tdTomato</sup> reporter mice. Taken together, these data suggested that under homeostasis, intestinal C3 is largely expressed by stromal cells located in the subepithelial compartment of the intestine.

To identify human intestinal cellular sources for C3, we analyzed scRNA-seq data from the Human Gut Cell Survey,<sup>20</sup> which included 79,929 intestinal cells from five anatomical sites in healthy adult human donors. B, T, endothelial, epithelial, mesenchymal, myeloid, neuronal, and plasma cells were identified based on their unique gene expression profiles. After retrieving the C3 expression profile from these human intestinal cells, we identified mesenchymal cells as the highest C3-transcribing population. Dot plot of C3 expression across cell types showed that mesenchymal cells have the largest percentage of cells expressing C3, as well as the highest average expression levels (Figure 3G). Our observations in mice are consistent with the findings using human datasets.

### Complement system gene expression profile and function in the intestine during homeostasis

Since complement proteins in blood usually function in groups and in sequence, we analyzed the intestinal transcriptional expression of complement proteins other than C3 in both mice and humans. We evaluated several major complement components at the transcriptional level using RNA-seq data from the three sorted WT populations described above. Based on normalized read counts for respective genes within each population, we defined transcript expression levels as low (<10/million reads), medium (10–100/million reads), or high (>100/million reads) (Figure 3F). The analysis showed that the classical-pathway component C1q (including *C1qa*, *C1qb*, and *C1qc*) is highly expressed in CD45<sup>+</sup> cells, consistent with a recent report that macrophages form the predominant source of C1q in the murine intestine.<sup>21</sup> Interestingly, mannose-binding lectin (MBL) pathway proteins showed very low expression levels in all three populations with the exception of *Masp1*, which was expressed at a medium level in stromal cells. The alternative pathway component *Cfp* was highly transcriptionally expressed in CD45<sup>+</sup> cells and expressed at low levels in epithelial and stromal cells. *C2* and *C4A* exhibited very low transcriptional expression in all three cell types, while *C4B* was highly expressed in stromal cells and moderately expressed in CD45<sup>+</sup> cells. The genes for proteins downstream of C3, including complement MAC components C5, C6, C7, C8, and C9 showed very low or undetectable levels of transcriptional expression (Figure 3F). Overall, this analysis suggested that in the murine colon, C1q, *Cfp*, and *C4B* are the only other major complement components (other than C3) with high transcriptional expression in a homeostatic environment.

Similarly to murine intestinal cells, human intestinal samples analyzed by scRNA-seq also possessed high expression of classical-pathway components *C1qa*, *C1qb*, and *C1qc* in the CD45<sup>+</sup> myeloid compartment. Human MBL pathway expression also remained equivalent to that of mice. Interestingly, alternative pathway components *Cfd* and *Cfh* were highly expressed in mesenchymal cells, and *Cfp* and *Cfd* were highly expressed in myeloid cells. While MAC member C7 was detected in mesenchymal and endothelial cells, none of the other MAC components (e.g., C5, C6, C8, or C9) were observed at detectable levels (Figure 3G). Based on these results, we conclude that mice and humans share similar expression of C3 and other main components of the gut complement system.

Next, to determine if the intestinal complement system can eliminate commensal organisms, thereby influencing the

### Figure 3. Stromal cells are the predominant cell population expressing C3 during homeostasis

- (A) Representative plots of the live colonic epithelial and subepithelial compartment of C3<sup>IRES-tdTomato</sup> reporter mice showing the frequency of C3-tdTomato<sup>+</sup> cells and the statistics.
- (B) Overall experimental scheme using C3<sup>IRES-tdTomato</sup> reporter mice to identify cellular sources for C3 in the colon.
- (C) UMAP shows scRNA-seq profiles of tdTomato-positive mouse colonic cells (dots) from C3<sup>IRES-tdTomato</sup> reporter mice, clustered into three cell types as indicated by the colors.
- (D) UMAP with cells organized by cell type and colored by *pdpn*, *lyz2*, and *epcam* expression levels.
- (E) Overall experimental scheme using WT C57BL/6 mice to verify cellular sources for C3 in the colon.
- (F) Complement system gene expression levels in WT C57BL/6 mice of the three indicated populations by ultra low input (ULI)-RNA-seq.
- (G) Dot plot showing average expression levels (color) of genes in the human complement system (rows) in several colonic cell populations (columns) from human adult colon samples<sup>20</sup>; dot size represents the percent of cells within each parent population (column) that express these genes (row).

\*\*\*\*p < 0.0001, unpaired Student's t test. Error bars represent mean ± SEM.

See also Figure S3.

composition of the gut microbiota, we compared the gut microbiota composition in C3-deficient mice and their WT littermate controls. Our analysis revealed no significant differences in the microbiota composition (Figure S3D), both in terms of  $\alpha$ -diversity and  $\beta$ -diversity (Figures S3E and S3F). Given the absence of the components (C5–C9) required for the formation of the MAC in the gut, coupled with the limited presence of phagocytes in the lumen at homeostasis, our results suggested that the basal levels of C3 during homeostasis do not lead to active clearance of commensal organisms.

### C3-expressing stromal cells are primarily located in colonic lymphoid follicles and produce C3 upon bacterial stimulation

To locate C3-expressing cells in the colon, we performed RNA scope analysis using C3-specific antisense probes on formalin-fixed, paraffin-embedded (FFPE) sections from the colon of WT C57BL/6 mice and found that the majority of C3-expressing cells were located in the isolated lymphoid follicles (ILFs) (Figure 4A). ILFs are tertiary lymphoid organs which consist of collections of B, T, myeloid, and stromal cells, and their development has been suggested to be linked to microbial exposure.<sup>22</sup> Morphologically, the majority of C3-expressing cells in ILFs are stromal cells. We confirmed this using immunofluorescent antibodies against tdTomato (Figures S4A–S4C) and PDPN on frozen colon cryo-sections from C3<sup>IRES-tdTomato</sup> BL6 mice. Our staining showed co-localization of C3 transcript and PDPN (Figure 4B) and confirmed that stromal cells in ILFs are the major cell population expressing C3.

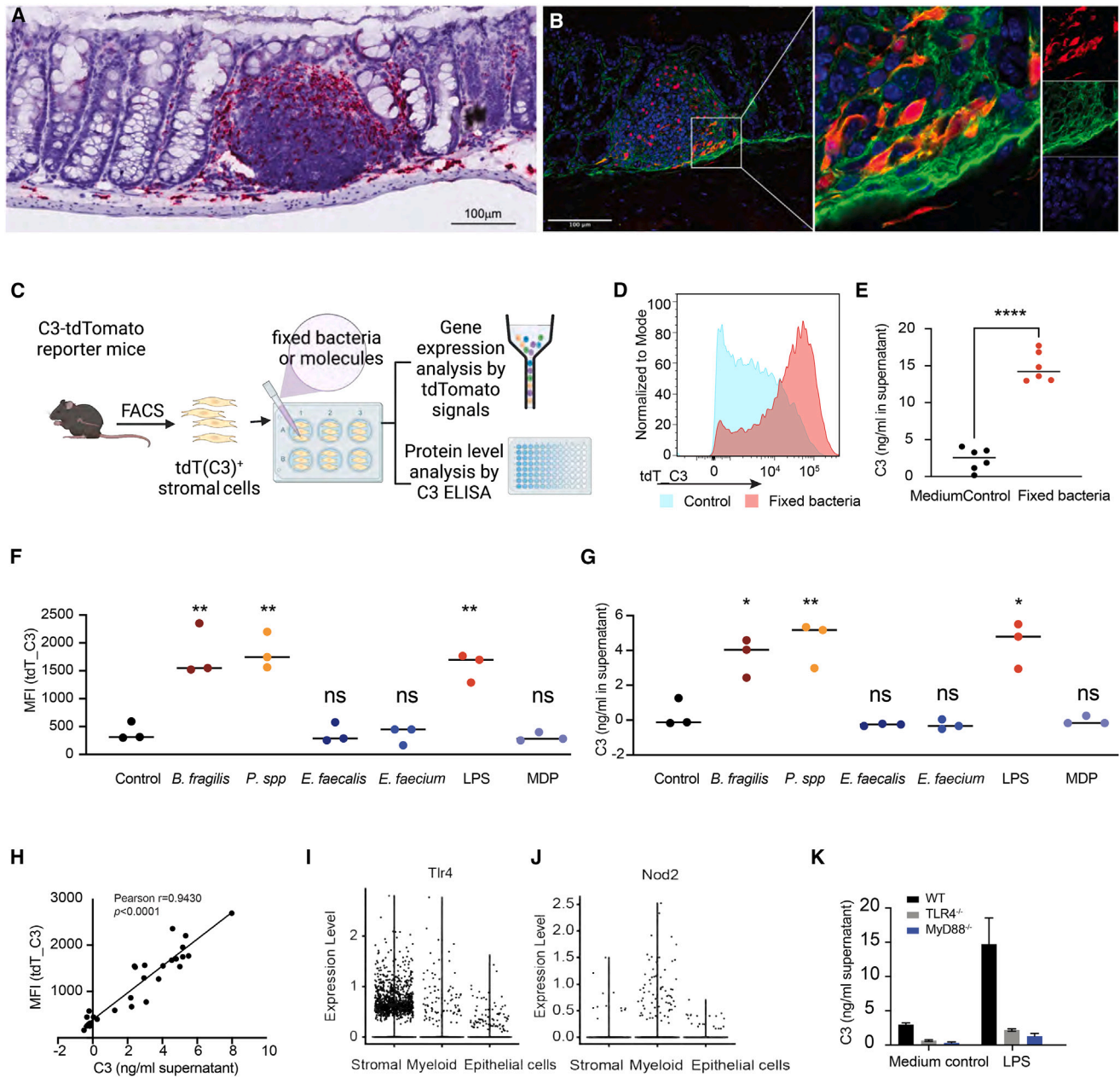
To further test if stromal cells can produce and secrete C3 protein, we isolated colonic stromal cells from C3<sup>IRES-tdTomato</sup> C57BL/6 mice and measured C3 gene expression and protein levels before and after bacterial stimulation (Figure 4C). For this, we formalin-fixed bacterial cells obtained from the feces of CR mice and used these to stimulate cultured primary stromal cells. We found increased tdTomato signal (indicating C3 transcript expression) upon bacterial exposure compared with unstimulated controls (Figure 4D). We also observed higher C3 protein levels in the extracellular medium of stromal cells 24 h after bacterial stimulation (Figure 4E). The expression levels of tdTomato significantly correlated with protein levels of C3 in the medium (Figure 4H), confirming that this reporter system was a robust measure of C3 protein production. To further dissect the ability of bacterial strains in stimulating C3 production, we picked four representative bacterial strains: two gram-negative species, *Bacteroides fragilis* 9343 and *Prevotella* isolate CR1, and two gram-positive species, *Enterococcus faecalis* TX0104 and *Enterococcus faecium* TX1330. We cultured the strains individually in their designated culture media, followed by formalin fixation. We found both formalin-fixed gram-negative bacterial strains upregulated C3 expression and C3 protein production (Figures 4F and 4G). In contrast, the two formalin-fixed gram-positive bacterial strains did not exhibit this effect (Figures 4F and 4G). Intriguingly, heat-killed *Enterococcus faecalis* TX0104 was able to induce C3 transcript expression (Figure S4D). Together, these results suggest that bacterial molecules from both gram-negative and gram-positive bacteria have the capacity to stimulate stromal cell C3 production.

To identify specific bacterial molecules participating in the induction of C3 in stromal cells, we examined expression levels of several pattern-recognition receptors (PRRs) in colonic C3-expressing cells using our scRNA-seq data. We found that intestinal C3-expressing stromal cells express *Tlr4*, *Tlr2*, *Tlr3*, *Nod1*, but almost no detectable levels of *Nod2*, *Tlr1*, *Tlr8*, or *Tlr9* (Figures 4I, 4J, and S4E). Therefore, we chose two microbe-associated molecular patterns (MAMPs) to test their ability to induce C3 expression in stromal cells: lipopolysaccharide (LPS), a known TLR4 ligand, and muramyl dipeptide (MDP), a known NOD2 ligand. LPS upregulated the transcription and secretion of C3 in primary colonic stromal cells while MDP did not (Figures 4F and 4G). To identify innate cellular signaling pathways used by colonic stromal cells to detect LPS and trigger C3 production, we isolated stromal cells from TLR4<sup>-/-</sup>, MyD88<sup>-/-</sup>, and WT BL6 mice. Compared with WT control cells, LPS stimulation induced less C3 in both TLR4- and MyD88-knockout stromal cells, indicating that the TLR4-MyD88 pathway is critical for LPS sensing and triggering of C3 production in intestinal stromal cells (Figure 4K). In addition, we selected some TLR2 ligands, including peptidoglycan, lipoteichoic acid, and Pam3CSK4, to assess their capacity to induce C3. Our results showed that these bacterial or bacteria-like components are also capable of stimulating C3 production (Figure S4D).

### Luminal complement C3 is critical for protection against *Citrobacter rodentium* infection

To investigate the function of mucosal complement C3 on host health, we assessed the role of C3 during infection against the murine pathogen *Citrobacter rodentium*. This pathogen is the mouse counterpart to model human pathogen enteropathogenic *E. coli* (EPEC) and enterohemorrhagic *E. coli* (EHEC).<sup>23</sup> We orally infected weaning-age C3-deficient mice and age-matched WT control mice. C3-deficient mice were more sensitive to *C. rodentium* infection, with only a 50% survival rate, while all WT mice survived the pathogen challenge ( $p < 0.05$ ) (Figure 5A). Adult C3-deficient mice also showed significantly more symptoms such as weight loss than WT mice (Figure 5B). Symptom onset occurred 1 week after infection and peaked 5 days later. We found significantly higher levels of *C. rodentium* shed in the feces of C3-deficient mice compared with WT mice at 11 days post-infection (dpi) (Figure 5C). Since *C. rodentium* is known for the formation of “attaching and effacing” (A/E) lesions in the colonic epithelium during infection, we examined the location of *C. rodentium* using specific rRNA FISH probes at 7 dpi. There were significantly higher numbers of *C. rodentium* invasion in the epithelium of C3-deficient mice compared with WT mice (Figures 5D–5F) ( $p < 0.005$ ).

To determine whether mucosal C3 was deposited on bacterial cells during infection, we infected mice with *C. rodentium* LB1, a strain which constitutively expresses GFP.<sup>24</sup> Flow cytometric analysis on GFP<sup>+</sup> *C. rodentium* cells in fecal samples with anti-C3 antibody showed that the microbes were coated with C3 at 7 dpi (Figure 5G). To rule out whether the classical or lectin pathways are required for this phenotype, we infected C1qa- and MBL-deficient mice with GFP<sup>+</sup> *C. rodentium*, we found similar levels of fecal C3 and C3-coated GFP<sup>+</sup> *C. rodentium* cells (Figures S5A–S5C), suggesting that the deposition of C3 does



**Figure 4. Location and function of C3-expressing intestinal stromal cells**

(A) RNAscope *in situ* hybridization of C3-specific antisense mRNA probe (red) in mouse colon samples from WT C57BL/6 mice.

(B) Immunofluorescence microscopy of colonic sections from C3<sup>IRES-tdTomato</sup> reporter mice stained against tdTomato (red), PDPN (green), or DNA (blue).

(C) Overall experimental scheme for the isolation of colonic C3<sup>+</sup> stromal cells from C3<sup>IRES-tdTomato</sup> reporter mice, which are subsequently stimulated by fixed bacteria or molecules and tested for C3 transcript expression and protein production.

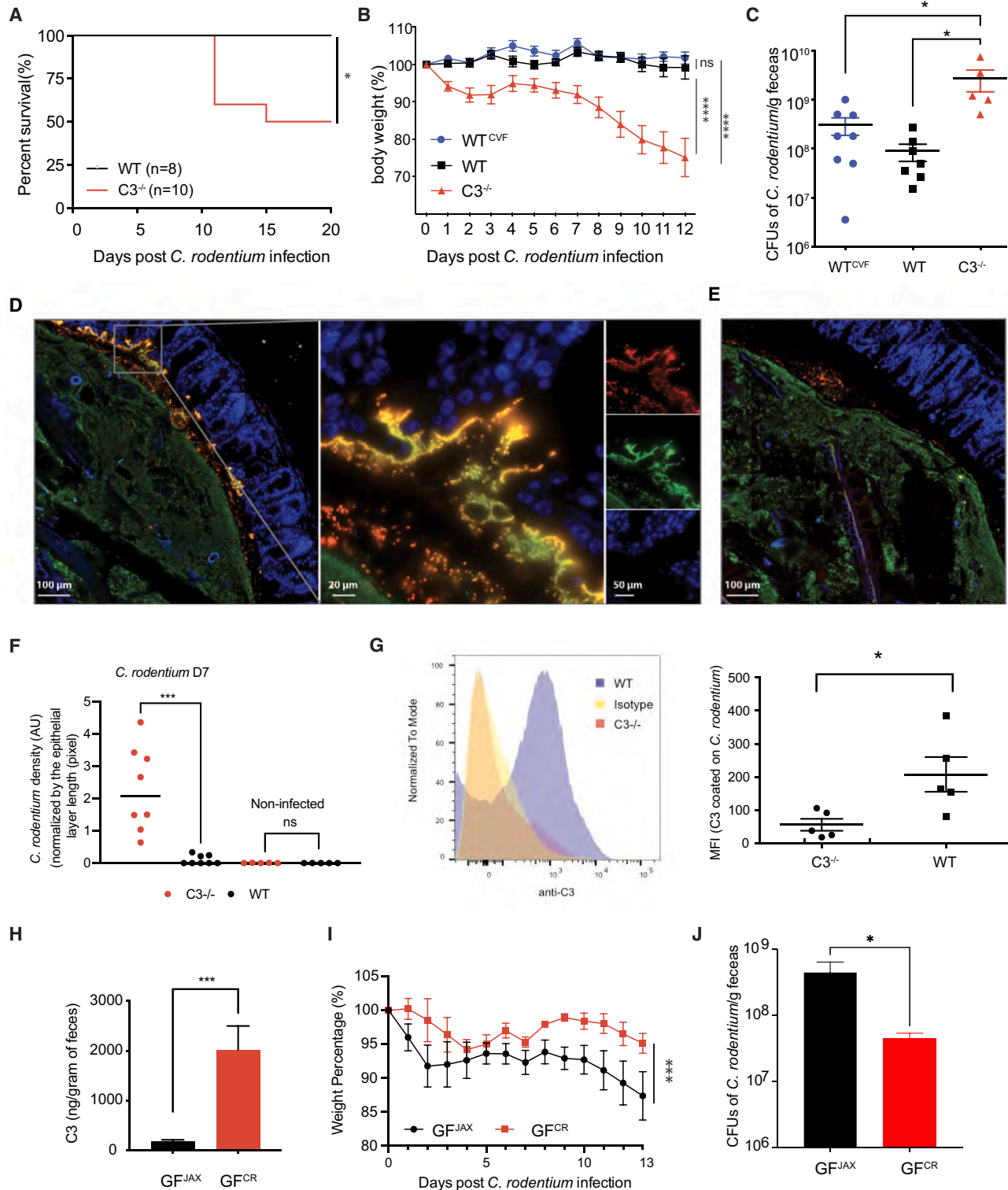
(D–G) Transcriptional expression levels of C3 (D and F) and C3 protein concentration in the culture medium (E and G) of colonic stromal cells stimulated *in vitro* by fixed bacterial strains individually or in combination.

(H) Correlation of C3 transcriptional expression levels and C3 protein levels in the culture media from (D)–(G).

(I and J) Transcriptional expression levels of *tlr4* (I) and *nod2* (J) in the three indicated cell types from scRNA-seq of C3-expressing colonic cells.

(K) C3 concentration in the culture medium of colonic stromal cells from WT, TLR4<sup>-/-</sup>, or MyD88<sup>-/-</sup> mice stimulated by LPS (1 μg/mL) or medium only, control. \*p < 0.05; \*\*p < 0.01; \*\*\*\*p < 0.0001, unpaired Student's t test. Error bars represent mean ± SEM.

See also Figure S4.



**Figure 5. Complement factor C3 is critical for protection against *Citrobacter rodentium* infection**

(A) Survival of ~4-week-old C57BL/6 WT and C3-deficient mice following challenge with *C. rodentium* ( $5 \times 10^8$  CFU/mouse).

(B) Weight loss following challenge with *C. rodentium* ( $5 \times 10^8$  CFU/mouse) in three groups of ~8-week-old mice: WT, C3-deficient ( $C3^{-/-}$ ), and WT depleted of serum complement with cobra venom factor ( $WT^{CVF}$ ).

(C) Fecal CFU counts of *C. rodentium* at 11 dpi in untreated WT; WT were treated with cobra venom factor (CVF) ( $WT^{CVF}$ ), or C3-deficient mice ( $C3^{-/-}$ ).

(legend continued on next page)

not rely on the classical or lectin pathway. This was consistent with the report that C1q is not essential for protection against *C. rodentium* infection.<sup>21</sup>

To confirm the role of luminal C3, as opposed to plasma-circulating C3, in *C. rodentium* infection, we treated SPF mice with cobra venom factor (CVF) to specifically deplete plasma-circulating C3 without depleting gut luminal C3 (Figures S5D and S5E). Subsequently, mice were challenged with *C. rodentium* and monitored for symptoms. We found CVF treatment did not compromise the resistance to *C. rodentium* infection displayed by WT mice. There was no significant difference in weight loss or bacterial burden in CVF-treated compared with non-CVF-treated WT controls (Figures 5B and 5C). These results indicate that mucosal complement C3 is sufficient to confer resistance to this enteric pathogen.

We next tested if modulation of luminal C3 by different gut microbiota could affect host susceptibility to *C. rodentium* challenge. We colonized GF mice with JAX or CR microbiota resulting in GF<sup>JAX</sup> with low luminal C3 level or GF<sup>CR</sup> with high luminal C3 level, respectively (Figure 5H). When mice were challenged with a high dose of *C. rodentium* ( $2 \times 10^{10}$  colony-forming unit [CFU]/mouse), we found that GF<sup>JAX</sup> mice lost significantly more weight during the infection compared with GF<sup>CR</sup> mice (Figure 5I). GF<sup>JAX</sup> mice also shed significantly higher numbers of *C. rodentium* in their feces compared with GF<sup>CR</sup> mice at 11 dpi (Figure 5J). These results demonstrated that the level of luminal C3 driven by the unique compositions of the host gut microbiota modulated the degree of susceptibility to enteric infection by *C. rodentium*.

Our results with the murine enteric pathogen *C. rodentium* suggest that mucosal C3 can opsonize select enteric pathogens during enteric infections and that this can occur independently of the classical or lectin pathways. Variations in luminal C3 levels due to the host commensal microbiota can impact the host's susceptibility to enteric pathogen infection.

### Luminal C3 production increases significantly during enteric infection

We reasoned that, if luminal C3 levels play a major role in controlling enteric infections, *C. rodentium* infection may lead to upregulation of mucosal C3 at the transcriptional and protein levels. To test this hypothesis, C3<sup>IRES-tdTomato</sup> reporter mice were infected with *C. rodentium*; transcriptional levels of C3 were measured by flow cytometry; and fecal C3 protein levels were measured by ELISA (Figure 6A). We found that fecal C3 protein levels increased significantly during infection (Figure 6B), which correlated with C3 gene transcriptional expression as measured by the total tdTomato signal in cells from the colonic subepithelial compartment (Figure 6C). Gating on tdTomato-positive cells, we

found that stromal cells, CD45<sup>+</sup> cells, and IECs are also major sources of C3 production during infection (Figures S6A and S6B).

Compared with uninfected controls, mice infected with *C. rodentium* had an increased proportion of stromal cells expressing C3 by 5 dpi (30% versus 12%), which coincides with the early phase of infection as *C. rodentium* begins to attach to enterocytes.<sup>25</sup> The percentage of stromal cells transcribing C3 remained at ~30% throughout the infection. CD45<sup>+</sup> cells transcribing C3 increased from 5% to ~30% by day 5 and reached a level of ~40% by day 8. This level of expression was maintained in CD45<sup>+</sup> cells through day 14. IECs transcribing C3 increased from ~0.5% to 5% of total IECs by day 5 and reached a level of ~10% by day 11, which continued to rise to ~15% through day 14 (Figure 6D). In terms of cell numbers—due to the large number of total IECs in the colon, IEC hyperplasia, and the infiltration of CD45<sup>+</sup> cells into the gut during infection—the composition of the C3-transcribing population changed from predominantly stromal cells to predominantly CD45<sup>+</sup> and IECs as the infection progressed (Figures 6E and S6B).

We next performed bulk RNA-seq on sorted IECs, CD45<sup>+</sup>, and stromal cells from uninfected and infected C57BL/6 mice (Figure 6F). We found significant increases in C3 gene expression in all three populations at 7 and 11 dpi (Figure 6G), suggesting that all three populations contributed to C3 transcriptional expression. Nonetheless, stromal cells still showed the highest C3 mRNA expression levels during infection when compared with other host cell populations (Figure 6G), consistent with the results in reporter mice.

We also looked for increased transcript expression of other complement system proteins and compared infected and noninfected controls of the same three cell populations in WT C57BL/6 mice (Figure 6H). We found that in addition to C3, both *Cfb* and *C4B* are significantly upregulated in stromal and CD45<sup>+</sup> cells. However, the majority of genes encoding MAC components (C5–C9) were not expressed at detectable levels by RNA-seq in any of the three sorted populations, which is consistent with low C5 protein levels in WT fecal specimens from 11 dpi (Figure S6C).

### C3-dependent neutrophil-mediated phagocytosis is critical for protection against *C. rodentium* infection

We reported above that C3-deficient mice had significantly greater pathogen loads during infection than C3-sufficient mice (Figure 5C). Thus, we wanted to determine the major mechanisms used by the gut complement system to limit pathogen load. Since components of the MAC (C5–C9) are not upregulated or secreted by cells in the colon during infection, we examined the role of opsonophagocytosis in controlling *C. rodentium* infection. Flow cytometric analysis of intestinal cells isolated from C3<sup>IRES-tdTomato</sup>

(D and E) Images of cross-section of colons containing fecal matter from *C. rodentium*-infected mice that are C3 deficient (D) or WT (E). Samples were stained with *C. rodentium* specific rRNA FISH probes (red), Eub338 rRNA FISH probes (green), and DAPI (blue).

(F) *C. rodentium* density under the epithelial layer of C3-deficient or WT mice, normalized by colon length (n = 8).

(G) Flow cytometry histogram and quantification of C3 protein on GFP<sup>+</sup> *C. rodentium* at 7 dpi.

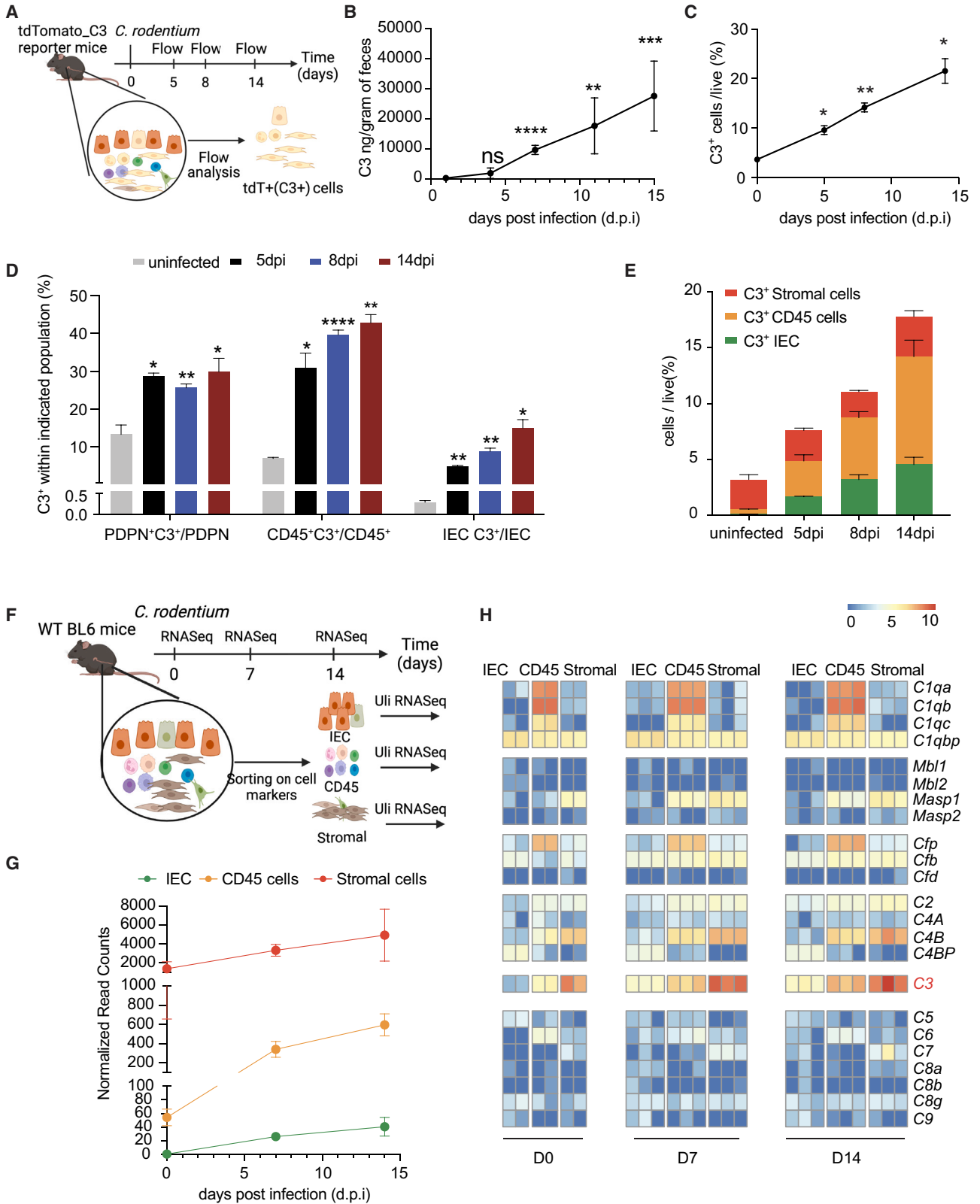
(H) Fecal C3 levels of gnotobiotic mice 2 weeks after colonization with JAX or CR microbiota.

(I) Weight loss of gnotobiotic mice colonized with JAX or CR microbiota following challenge with *C. rodentium* ( $2 \times 10^{10}$  CFUs/mouse).

(J) Fecal CFU enumeration of *C. rodentium* on 11 dpi in gnotobiotic mice colonized prior to challenge with JAX or CR microbiota.

ns = not significant, \*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.001, Student's t test. Error bars represent mean  $\pm$  SEM.

See also Figure S5.



(legend on next page)

reporter mice showed a population within CD45<sup>+</sup> cells that incrementally upregulated C3 expression throughout the course of infection (Figure 7A). Whereas in a homeostatic environment only ~4% of C3-expressing CD45<sup>+</sup> cells express high C3 levels, this population rose to ~50% at 14 dpi (Figure 7B). Further analysis of this high-C3-expressing population revealed they were myeloid cells (CD11b<sup>+</sup>), with the majority being neutrophils, followed by macrophages and monocytes (Figures 7C, S7A, and S7B).

We then analyzed C3 expression dynamics in these three myeloid populations during infection. While only a subset of macrophages and monocytes expressed C3 (Figures S7C–S7E), neutrophils uniformly expressed high levels of C3 (Figures 7D and 7E). Among the myeloid cell population, neutrophils had the highest tdTomato mean fluorescence intensity (MFI) and maintained this level throughout the infection (Figures 7D and S7C–S7E). We found that neutrophils were robustly recruited to the colonic lamina propria and epithelial layer during infection (Figures 7F and 7G) where they became the major source of high-C3-expressing CD45<sup>+</sup> cells (Figures 7C and S7B). The dominant presence and activity of neutrophils is consistent with the concept that phagocytosis of microbes is imperative for C3-dependent bacterial clearance.

To explore the hypothesis that mucosal C3 was acting as an opsonin to enhance bacterial phagocytosis, we measured the magnitude of bacterial uptake by neutrophils in the mucosal environment. WT mice were infected with GFP<sup>+</sup> *C. rodentium*, and neutrophils from the mucosal epithelial and lamina propria layers were collected and analyzed by flow cytometry at 7 dpi. We found a significant fraction of GFP<sup>+</sup> neutrophils in both layers (Figures 7H and 7I), suggesting that neutrophils had phagocytosed GFP<sup>+</sup> *C. rodentium* within the host mucosal environment.

In addition, we wanted to determine whether gut complement was important to the clearance of other enteric pathogens. Therefore, we tested the human pathogen EHEC in a murine environment. We infected 6-week-old C3-deficient or WT C57BL/6 mice with GFP<sup>+</sup> EHEC strain EDL931 via the oral gastric route. We found that adult C3-deficient mice were highly susceptible to EHEC infection, showing only a 16% survival rate. In contrast, WT mice were substantially more resistant to EHEC infection with only 1 out of 11 mice succumbing to the infection (Figure S7G). We collected fecal samples from WT and C3-deficient mice at 4 dpi and performed flow cytometric analysis using an anti-C3 antibody to assess C3 opsonization of EHEC. C3 was detected on fecal GFP<sup>+</sup> EHEC EDL931, demonstrating that EHEC can also be significantly opsonized by C3 in the lumens of WT mice (Figure S7H).

## DISCUSSION

The complement cascade is an evolutionarily conserved part of our immune system. Initially identified within the circulatory system, its scope has recently broadened to encompass various locations including the lungs<sup>5,9,26</sup> as well as intracellular functions.<sup>3,5,7</sup> This study has revealed the presence of a gut complement system, separate and distinct from the systemic complement system. The components of the gut complement system were selectively expressed by local colonic cells and regulated by the composition of the gut microbiota during homeostasis. Moreover, its functionality lies in the ability to establish harmony with commensal organisms while effectively eliminating pathogens.

C3 protein exists in the lumen of healthy mice and humans and its level is regulated by the composition of gut microbiota resulting in host-specific gut C3 levels. There was a local increase in C3 transcription levels in colonic cell populations and an increase in C3 protein levels in the lumen during enteropathogenic infection. However, the proteins forming the MAC are not expressed at a detectable level in the gut, neither in homeostasis nor infection. This suggests that bactericidal lysis by the MAC is not a function of the gut complement system, and this may be beneficial to spare the commensals microorganisms in the gut. Instead, the gut complement system appears to use complement-mediated phagocytosis as the major downstream pathway to clear pathogens. Phagocytes are not generally present in the gut lumen and neutrophils are only recruited to the gut during infection. This fine-tuning of the gut complement system reveals an innate immunity mechanism that differentiates gut commensals from pathogens.

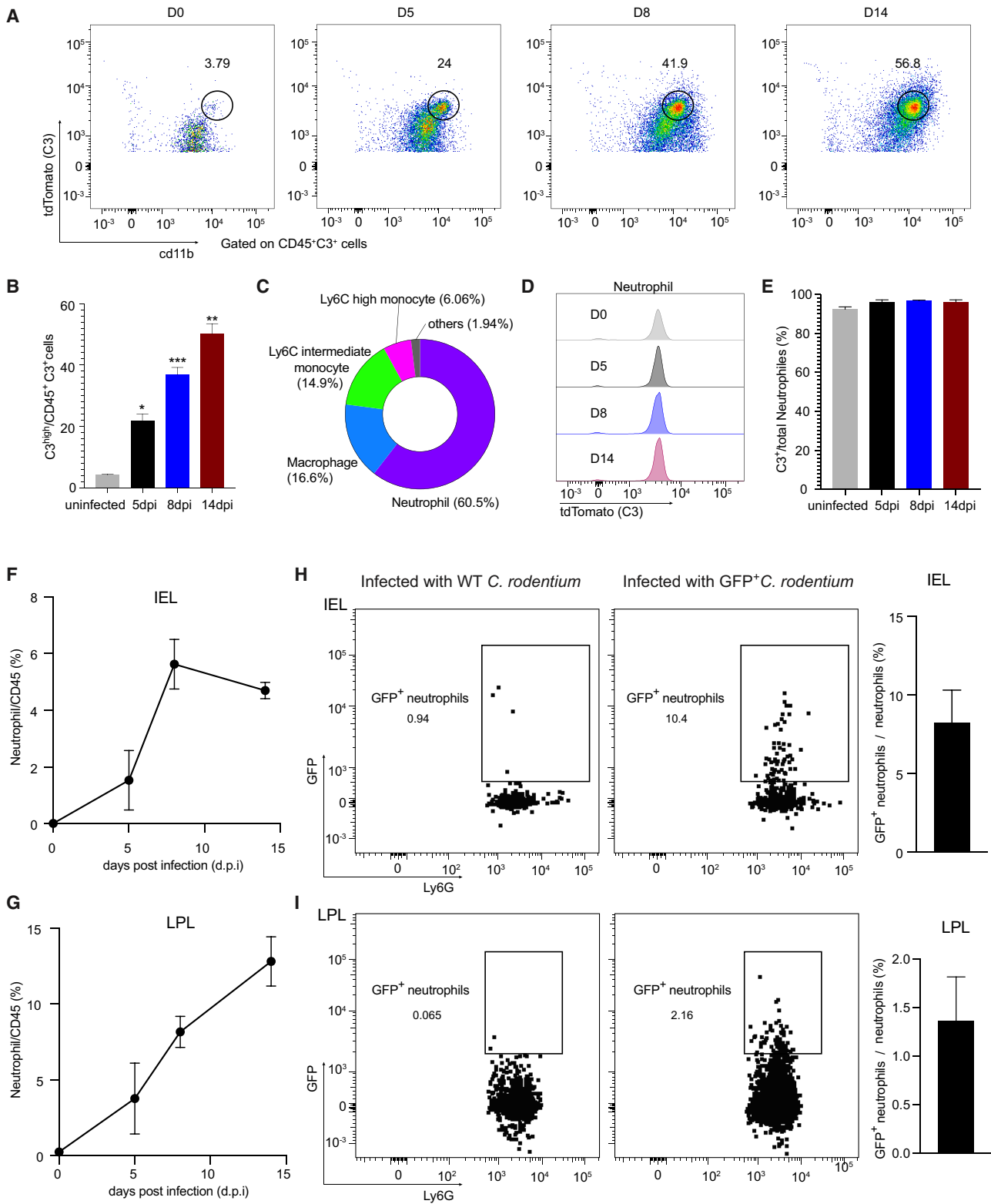
Our transcriptional analysis revealed that C1q is also highly expressed in gut myeloid cells. However, using C1q-deficient mice, we found that C1q is dispensable for defending the host from enteropathogens in the lumen, a finding that is consistent with a recent report.<sup>21</sup> Our study and others' show complement system components may play different roles locally than they do systemically.<sup>3,5,7,21,26</sup> Fecal C3 was previously connected to noninfectious diseases. Local gut complement C3 deposition was observed on the surface of the colonic epithelium in biopsies or resection specimens from IBD patients.<sup>27</sup> It will be interesting to explore additional functions of complement C3 in the gut.

Our data showed that the expression levels and cellular sources of various complement components in mice and humans are strongly conserved. This conservation emphasizes the importance of local complement. The finding that stromal cells are the major source of C3 in both mice and humans during

### Figure 6. Fecal levels of complement C3 increase significantly during infection

- (A) Overall experimental timeline of flow cytometric analysis of C3-expressing cells in C3<sup>IRE5-tdTomato</sup> reporter mice after *C. rodentium* infection.  
(B) Fecal C3 protein levels at 1, 4, 7, 11, and 15 dpi.  
(C) Percentage of live cells isolated from the colonic lamina propria of C3<sup>IRE5-tdTomato</sup> reporter mice transcriptionally expressing C3 at 0, 5, 8, and 14 dpi.  
(D) Percentage of cells transcribing C3 among stromal cells, CD45<sup>+</sup> cells, and IECs.  
(E) Percentage of stromal cells, CD45<sup>+</sup> cells, or IECs transcribing C3 among the total live cell population isolated from colonic lamina propria at indicated days post-infection.  
(F) Overall experimental timeline of RNA-seq analysis of sorted cell populations in WT C57BL/6 mice after *C. rodentium* infection.  
(G) RNA-seq analysis of C3 gene expression levels in the three indicated cell populations in WT C57BL/6 mice.  
(H) Heatmap of complement system gene expression levels in the three indicated cell populations of WT C57BL/6 mice.  
ns, not significant, \*\*p < 0.01, \*\*\*p < 0.001, \*\*\*\*p < 0.0001; Student's t test. Error bars represent mean ± SEM.

See also Figure S6.



**Figure 7. C3-mediated neutrophil phagocytosis is crucial for clearing *Citrobacter rodentium* during enteric infection**

(A and B) Representative gating (A) and frequencies (B) of the high-C3-expressing subset in colonic subepithelial CD45<sup>+</sup>C3<sup>+</sup> cells during homeostasis (day 0) and at 5, 8, and 14 days post *C. rodentium* infection.

(legend continued on next page)

homeostasis was unexpected. Traditionally, stromal cells are nonhematopoietic, nonepithelial, and non-endothelial structural cells that provide essential connective tissue (stroma).<sup>28</sup> They define the architecture of the tissue microenvironment by depositing and remodeling extracellular matrix (ECM) components. Recent studies have uncovered important functions for these cells in initiating, supporting, or suppressing innate and adaptive immune responses in lymphoid and nonlymphoid tissues.<sup>29–34</sup> Here, we established an additional stromal cell function, which defends against pathogens through the expression of complement C3. Although we've discovered that bacterial components can stimulate C3 production via TLR4 and TLR2 signaling pathways, the underlying regulatory mechanisms are still largely unknown and warrant additional study.

Studies have revealed that the composition of the gut microbiota can affect susceptibility to infectious diarrhea.<sup>11</sup> Our findings that basal luminal C3 levels are determined by the host gut commensal microbiota and are correlated to the host's susceptibility to infectious disease are intriguing. It expands our understanding of the role of the microbiota in shaping variability in host infection susceptibility and severity and reinforces the important concept that the microbiota can indirectly influence a host's response to infection by priming immune defense mechanisms before a pathogen is encountered.

Diarrheal diseases are the third leading cause of child mortality globally, and death rates are especially high in low- and middle-income countries. In these regions, malnutrition is a major public health concern.<sup>35</sup> Diet is a major factor that shapes the composition of the gut microbiota. Malnourished children exhibit delayed maturation of their gut microbiome and have greater susceptibility to enteric pathogens.<sup>36,37</sup> Malnutrition and enteric disease susceptibility may be both a cause and a consequence of each other. Pathogenic *E. coli* (including EHEC and EPEC) are thought to be the most common enteric pathogens found to be associated with diarrheal disease in low- and middle-income countries. In our mouse models, we illustrated the link between microbiota-dependent C3 levels and the host's susceptibility to infectious diseases. This suggests that disease vulnerability and severity might be tied to microbiota changes due to malnutrition. Such findings underscore the need for deeper investigations in human cohorts. This could be pivotal in leveraging the therapeutic capabilities of the microbiota against infectious diseases via modulation of the host's immune system.

### Limitations of the study

While our study highlights that the induction of C3 production is modulated by the composition of the gut microbiota, and specific bacterial strains have the capacity to induce heightened

C3 production within a complex community environment, the precise molecular mechanisms underlying this process need further investigation. When introducing the high-C3-inducing bacterial strain to GF mice in a monocolonized state, we did not observe significantly greater C3 levels than other bacterial strains. However, this same strain was able to trigger significantly elevated C3 levels within a complex microbiota environment. This implies that interactions between this strain and other microbes in the environment are crucial for elevated C3 induction phenotypes. Similar dependencies on microbial interactions for immunomodulatory functions have been elucidated in other studies.<sup>38–42</sup> For instance, *Helicobacter hepaticus* alone is insufficient to induce colitis in interleukin (IL)-10-deficient mice,<sup>39–41</sup> and the presence of other bacteria is necessary for Enterobacteriaceae to provoke inflammation.<sup>42</sup> Microbial interactions have the potential to alter the metabolites produced, influence surface protein or polysaccharide synthesis, and even guide microbes to distinct niches, thereby establishing unique interactions with the host. These host-microbiota interactions are intricate and further studies are needed.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Mice
  - Human samples
- METHOD DETAILS
  - Antibiotic treatment and fecal transfer
  - 16S rRNA gene sequencing of fecal samples and data analysis
  - Isolation of *Prevotella spp.* strains from Charles River (CR) mice gut microbiota
  - C3 and C5 ELISA
  - Gut permeability assay
  - Preparation of intestinal cells and flow cytometry
  - Stromal cell isolation and culture
  - scRNASeq profiling of C3-expressing cells in colon
  - Gene expression profiling by population RNAseq
  - Analysis of published single-cell data

(C) Percentage of neutrophils, macrophages, Ly6C-intermediate, and Ly6C-high monocytes in the high-C3-expressing CD45<sup>+</sup> cells at 14 days post *C. rodentium* infection.

(D) Representative histograms of C3 gene expression in neutrophils at 0, 5, 8, and 14 dpi in C3<sup>IRE5-tdTomato</sup> reporter mice.

(E) Percentage of the C3-expressing neutrophils among total neutrophils from colonic lamina propria at indicated days post-infection.

(F and G) Percentage of neutrophils in CD45<sup>+</sup> cells in epithelium (F) or colonic lamina propria (G) at 0, 5, 8, and 14 dpi.

(H and I) Representative gating and frequency of GFP<sup>+</sup> *C. rodentium* uptake in neutrophils in the colonic intraepithelial (H) and lamina propria (I) layers at 7 dpi. Left panels show the mice infected with WT *C. rodentium* as control for GFP<sup>+</sup> gating. Right panels are the mice infected with GFP<sup>+</sup> isogenic *C. rodentium*. \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001, Student's t test. Error bars represent mean ± SEM.

See also Figure S7.

- *Citrobacter rodentium* infection and C3-deposition analysis
- Cobra venom factor (CVF) treatment of wildtype mice
- Enterohemorrhagic Escherichia Coli (EHEC) infection
- Mouse C3 RNAscope
- Immunofluorescent staining of C3-expressing cells in C3<sup>IRES-tdTomato</sup> C57BL/6 mice
- Bacterial probe staining
- 16S-rRNA-specific probes for *C. rodentium*
- Bacterial imaging and data analysis
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cell.2023.12.036>.

### ACKNOWLEDGMENTS

We thank the staff in our animal facility for their support in animal husbandry, and all Kasper laboratory members for their comments and support. We thank Ian Magill, Niket Patel, Gang Wang, Yutao Wang, Teshika Jayewickreme, and Juliana Lee for their help with flow cytometry analysis, scRNA-seq, and ultra low input (Uli) RNA-seq. We thank Dr. John Atkinson (Washington University in St. Louis) and Dr. Philip Ahern (Cleveland Clinic) for advice and discussions. We thank Dr. Claudia Kemper (National Institute of Health) for C3<sup>IRES-tdTomato</sup> reporter mice. We thank Dr. Lynn Bry (Brigham and Women's Hospital) for *C. rodentium* DBS100 and GFP<sup>+</sup> *C. rodentium* LB1. We thank the Harvard Gnotobiotic Core, the Harvard Immunology Flow Cytometry Core, the Microscopy Resources on the North Quad (MicRoN) Core, the Harvard Rodent Histopathology Core, and the Histology Core at BIDMC for their help and support. This work was supported in part by NIH, NIAID 5R01AI148273 to D.L.K., and in part by NIH R01AI157106 and R35GM124724 to A.H. BioRender was used to create the graphical abstract and schematic illustration.

### AUTHOR CONTRIBUTIONS

Planning and conceptualization, M.W., W.Z., and D.L.K.; flow cytometry, sample collection, and ELISAs, M.W., W.Z., X.S., B.B., T.Y., H.-S.C., F.S.G., M.I., D.C., and C.T.; mouse microbiome work and analysis, M.W.; human microbiome work and analysis, M.W., R.L., J.C.M., E.A.D., and A.H.; imaging, M.W., W.Z., Y.W., D.R., D.Y., W.A.A., and S.H.; scRNA-seq, UliRNA-seq, and computational analysis, M.W.; provision of key resources, S.H., I.M.C., J.R.M., A.H., J.J.M., C.B., and D.L.K.; writing and editing manuscript, M.W., W.Z., A.H., J.J.M., and D.L.K. with help from all authors.

### DECLARATION OF INTERESTS

J.R.M. is a co-founder, stakeholder, and advisor for Vizgen, Inc.

Received: January 24, 2023

Revised: September 25, 2023

Accepted: December 30, 2023

Published: January 26, 2024

### SUPPORTING CITATIONS

The following references appear in the supplemental information:<sup>43</sup>

### REFERENCES

1. Kunz, N., and Kemper, C. (2021). Complement Has Brains—Do Intracellular Complement and Immunometabolism Cooperate in Tissue Homeostasis and Behavior? *Front. Immunol.* **12**, 629986.
2. Pratt, J.R., Abe, K., Miyazaki, M., Zhou, W., and Sacks, S.H. (2000). In situ localization of C3 synthesis in experimental acute renal allograft rejection. *Am. J. Pathol.* **157**, 825–831.
3. Liszewski, M.K., Kolev, M., Le Friec, G., Leung, M., Bertram, P.G., Fara, A.F., Subias, M., Pickering, M.C., Drouet, C., Meri, S., et al. (2013). Intracellular complement activation sustains T cell homeostasis and mediates effector differentiation. *Immunity* **39**, 1143–1157.
4. Lubbers, R., van Essen, M.F., van Kooten, C., and Trouw, L.A. (2017). Production of complement components by cells of the immune system. *Clin. Exp. Immunol.* **188**, 183–194.
5. Kulkarni, H.S., Elvington, M.L., Perng, Y.-C., Liszewski, M.K., Byers, D.E., Farkouh, C., Yusef, R.D., Lenschow, D.J., Brody, S.L., and Atkinson, J.P. (2019). Intracellular C3 protects human airway epithelial cells from stress-associated cell death. *Am. J. Respir. Cell Mol. Biol.* **60**, 144–157.
6. King, B.C., Kulak, K., Krus, U., Rosberg, R., Golec, E., Wozniak, K., Gomez, M.F., Zhang, E., O'Connell, D.J., Renström, E., et al. (2019). Complement component C3 is highly expressed in human pancreatic islets and prevents  $\beta$  cell death via ATG16L1 interaction and autophagy regulation. *Cell Metab.* **29**, 202–210.e6.
7. Kolev, M., West, E.E., Kunz, N., Chauss, D., Moseman, E.A., Rahman, J., Freiwald, T., Balmer, M.L., Lötscher, J., Dimeloe, S., et al. (2020). Diapedesis-induced integrin signaling via LFA-1 facilitates tissue immunity by inducing intrinsic complement C3 expression in immune cells. *Immunity* **52**, 513–527.e8.
8. Friščić, J., Böttcher, M., Reinwald, C., Bruns, H., Wirth, B., Popp, S.-J., Walker, K.I., Ackermann, J.A., Chen, X., Turner, J., et al. (2021). The complement system drives local inflammatory tissue priming by metabolic reprogramming of synovial fibroblasts. *Immunity* **54**, 1002–1021.e10.
9. Yan, B., Freiwald, T., Chauss, D., Wang, L., West, E., Mirabelli, C., Zhang, C.J., Nichols, E.-M., Malik, N., Gregory, R., et al. (2021). SARS-CoV-2 drives JAK1/2-dependent local complement hyperactivation. *Sci. Immunol.* **6**, eabg0833.
10. Lloyd-Price, J., Arze, C., Ananthakrishnan, A.N., Schirmer, M., Avila-Pacheco, J., Poon, T.W., Andrews, E., Ajami, N.J., Bonham, K.S., Brislawn, C.J., et al. (2019). Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* **569**, 655–662.
11. Alavi, S., Mitchell, J.D., Cho, J.Y., Liu, R., Macbeth, J.C., and Hsiao, A. (2020). Interpersonal gut microbiome variation drives susceptibility and resistance to cholera infection. *Cell* **181**, 1533–1546.e13.
12. Surana, N.K., and Kasper, D.L. (2017). Moving beyond microbiome-wide associations to causal microbe identification. *Nature* **552**, 244–247.
13. Gálvez, E.J.C., Iljazovic, A., Amend, L., Lesker, T.R., Renault, T., Thiemann, S., Hao, L., Roy, U., Gronow, A., Charpentier, E., et al. (2020). Distinct polysaccharide utilization determines interspecies competition between intestinal *Prevotella* spp. *Cell Host Microbe* **28**, 838–852.e6.
14. Alper, C.A., Johnson, A.M., Birtch, A.G., and Moore, F.D. (1969). Human C'3: evidence for the liver as the primary site of synthesis. *Science* **163**, 286–288.
15. Morris, K.M., Aden, D.P., Knowles, B.B., and Colten, H.R. (1982). Complement biosynthesis by the human hepatoma-derived cell line HepG2. *J. Clin. Invest.* **70**, 906–913.
16. Walport, M.J. (2001). Complement. First of two parts. *N. Engl. J. Med.* **344**, 1058–1066.
17. Moon, R., Parikh, A.A., Szabo, C., Fischer, J.E., Salzman, A.L., and Hasegawa, P.O. (1997). Complement C3 production in human intestinal epithelial cells is regulated by interleukin 1 $\beta$  and tumor necrosis factor  $\alpha$ . *Arch. Surg.* **132**, 1289–1293.
18. Andoh, A., Fujiyama, Y., Bamba, T., and Hosoda, S. (1993). Differential cytokine regulation of complement C3, C4, and factor B synthesis in human intestinal epithelial cell line, Caco-2. *J. Immunol.* **151**, 4239–4247.
19. Sünderhauf, A., Skibbe, K., Preisker, S., Ebbert, K., Verschoor, A., Karsenti, C.M., Kemper, C., Huber-Lang, M., Basic, M., Bleich, A., et al. (2017). Regulation of epithelial cell expressed C3 in the intestine – Relevance for the pathophysiology of inflammatory bowel disease? *Mol. Immunol.* **90**, 227–238.

20. Elmentaite, R., Kumasaka, N., Roberts, K., Fleming, A., Dann, E., King, H.W., Kleshchevnikov, V., Dabrowska, M., Pritchard, S., Bolt, L., et al. (2021). Cells of the human intestinal tract mapped across space and time. *Nature* *597*, 250–255.
21. Pendse, M., De Selle, H., Vo, N., Quinn, G., Dende, C., Li, Y., Salinas, C.N., Srinivasan, T., Prohete, D.C., Crofts, A.A., et al. (2023). Macrophages regulate gastrointestinal motility through complement component 1q. *eLife* *12*, e78558.
22. Bouskra, D., Brézillon, C., Bérard, M., Werts, C., Varona, R., Boneca, I.G., and Eberl, G. (2008). Lymphoid tissue genesis induced by commensals through NOD1 regulates intestinal homeostasis. *Nature* *456*, 507–510.
23. Mullineaux-Sanders, C., Sanchez-Garrido, J., Hopkins, E.G.D., Shenoy, A.R., Barry, R., and Frankel, G. (2019). *Citrobacter rodentium*–host–microbiota interactions: immunity, bioenergetics and metabolism. *Nat. Rev. Microbiol.* *17*, 701–715.
24. Belzer, C., Liu, Q., Carroll, M.C., and Bry, L. (2011). The role of specific IgG and complement in combating a primary mucosal infection of the gut epithelium. *Eur. J. Microbiol. Immunol. (Bp)* *1*, 311–318.
25. Silberger, D.J., Zindl, C.L., and Weaver, C.T. (2017). *Citrobacter rodentium*: a model enteropathogen for understanding the interplay of innate and adaptive components of type 3 immunity. *Mucosal Immunol.* *10*, 1108–1117.
26. Sahu, S.K., Ozantürk, A.N., Kulkarni, D.H., Ma, L., Barve, R.A., Dannull, L., Lu, A., Starick, M., McPhatter, J., Garnica, L., et al. (2023). Lung epithelial cell-derived C3 protects against pneumonia-induced lung injury. *Sci. Immunol.* *8*, eabp9547.
27. Halstensen, T.S., and Brandtzaeg, P. (1991). Local complement activation in inflammatory bowel disease. *Immunol. Res.* *10*, 485–492.
28. Krausgruber, T., Fortelny, N., Fife-Gernedl, V., Senekowitsch, M., Schuster, L.C., Lercher, A., Nemic, A., Schmidl, C., Rendeiro, A.F., Bergthaler, A., et al. (2020). Structural cells are key regulators of organ-specific immune responses. *Nature* *583*, 296–302.
29. Davidson, S., Coles, M., Thomas, T., Kollias, G., Ludewig, B., Turley, S., Brenner, M., and Buckley, C.D. (2021). Fibroblasts as immune regulators in infection, inflammation and cancer. *Nat. Rev. Immunol.* *21*, 704–717.
30. Costa, A., Kieffer, Y., Scholer-Dahirel, A., Pelon, F., Bourachot, B., Cardon, M., Sirven, P., Magagna, I., Fuhrmann, L., Bernard, C., et al. (2018). Fibroblast heterogeneity and immunosuppressive environment in human breast cancer. *Cancer Cell* *33*, 463–479.e10.
31. Elyada, E., Bolisetty, M., Laise, P., Flynn, W.F., Courtois, E.T., Burkhart, R.A., Teinor, J.A., Belleau, P., Biffi, G., Lucito, M.S., et al. (2019). Cross-species single-cell analysis of pancreatic ductal adenocarcinoma reveals antigen-presenting cancer-associated fibroblasts. *Cancer Discov.* *9*, 1102–1123.
32. Puram, S.V., Tirosh, I., Parkh, A.S., Patel, A.P., Yizhak, K., Gillespie, S., Rodman, C., Luo, C.L., Mroz, E.A., Emerick, K.S., et al. (2017). Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell* *171*, 1611–1624.e24.
33. Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* *352*, 189–196.
34. Calon, A., Lonardo, E., Berenguer-Llargo, A., Espinet, E., Hernando-Mombona, X., Iglesias, M., Sevillano, M., Palomo-Ponce, S., Tauriello, D.V.F., Byrom, D., et al. (2015). Stromal gene expression defines poor-prognosis subtypes in colorectal cancer. *Nat. Genet.* *47*, 320–329.
35. Caulfield, L.E., de Onis, M., Blössner, M., and Black, R.E. (2004). Undernutrition as an underlying cause of child deaths associated with diarrhea, pneumonia, malaria, and measles. *Am. J. Clin. Nutr.* *80*, 193–198.
36. Chen, R.Y., Mostafa, I., Hibberd, M.C., Das, S., Mahfuz, M., Naila, N.N., Islam, M.M., Huq, S., Alam, M.A., Zaman, M.U., et al. (2021). A microbiota-directed food intervention for undernourished children. *N. Engl. J. Med.* *384*, 1517–1528.
37. Charbonneau, M.R., O'Donnell, D., Blanton, L.V., Totten, S.M., Davis, J.C.C., Barratt, M.J., Cheng, J., Guruge, J., Talcott, M., Bain, J.R., et al. (2016). Sialylated milk oligosaccharides promote microbiota-dependent growth in models of infant undernutrition. *Cell* *164*, 859–871.
38. Cadwell, K., Patel, K.K., Maloney, N.S., Liu, T.-C., Ng, A.C.Y., Storer, C.E., Head, R.D., Xavier, R., Stappenbeck, T.S., and Virgin, H.W. (2010). Virus-plus-susceptibility gene interaction determines Crohn's disease gene Atg16L1 phenotypes in intestine. *Cell* *141*, 1135–1145.
39. Dieleman, L.A., Arends, A., Tonkonogy, S.L., Goerres, M.S., Craft, D.W., Grenther, W., Sellon, R.K., Balish, E., and Sartor, R.B. (2000). *Helicobacter hepaticus* does not induce or potentiate colitis in interleukin-10-deficient mice. *Infect. Immun.* *68*, 5107–5113.
40. Nagalingam, N.A., Robinson, C.J., Bergin, I.L., Eaton, K.A., Huffnagle, G.B., and Young, V.B. (2013). The effects of intestinal microbial community structure on disease manifestation in IL-10<sup>-/-</sup> mice infected with *Helicobacter hepaticus*. *Microbiome* *1*, 15.
41. Whary, M.T., Taylor, N.S., Feng, Y., Ge, Z., Muthupalani, S., Versalovic, J., and Fox, J.G. (2011). *Lactobacillus reuteri* promotes *Helicobacter hepaticus*-associated typhlocolitis in gnotobiotic B6.129P2-IL-10tm1Cgn (IL-10<sup>-/-</sup>) mice. *Immunology* *133*, 165–178.
42. Garrett, W.S., Gallini, C.A., Yatsunenkov, T., Michaud, M., DuBois, A., Delaney, M.L., Punit, S., Karlsson, M., Bry, L., Glickman, J.N., et al. (2010). Enterobacteriaceae act in concert with the gut microbiota to induce spontaneous and maternally transmitted colitis. *Cell Host Microbe* *8*, 292–300.
43. McBride, S.M., Fischetti, V.A., Leblanc, D.J., Moellering, R.C., and Gilmore, M.S. (2007). Genetic diversity among *Enterococcus faecalis*. *PLoS One* *2*, e582.
44. Moffitt, J.R., Bambah-Mukku, D., Eichhorn, S.W., Vaughn, E., Shekhar, K., Perez, J.D., Rubinstein, N.D., Hao, J., Regev, A., Dulac, C., et al. (2018). Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science* *362*, eaau5324.
45. Choi, K.-H., and Schweizer, H.P. (2006). mini-Tn7 insertion in bacteria with single attTn7 sites: example *Pseudomonas aeruginosa*. *Nat. Protoc.* *1*, 153–161.
46. Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghathli, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* *37*, 852–857.
47. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. *Cell* *184*, 3573–3587.e29.
48. Geva-Zatorsky, N., Sefik, E., Kua, L., Pisman, L., Tan, T.G., Ortiz-Lopez, A., Yanortsang, T.B., Yang, L., Jupp, R., Mathis, D., et al. (2017). Mining the human gut microbiota for immunomodulatory organisms. *Cell* *168*, 928–943.e11.
49. Alavi, S., and Hsiao, A. (2020). Protocol for microbiome transplantation in suckling mice during *Vibrio cholerae* infection to study commensal-pathogen interactions. *Star Protoc.* *1*, 100200.
50. Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., and Kirschner, M.W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* *161*, 1187–1201.
51. McInnes, L., Healy, J., and Melville, J. (2020). UMAP: uniform manifold approximation and projection for dimension reduction. <https://doi.org/10.48550/arXiv.1802.03426>.
52. Thevaranjan, N., Puchta, A., Schulz, C., Naidoo, A., Szamosi, J.C., Verschoor, C.P., Loukov, D., Schenck, L.P., Jury, J., Foley, K.P., et al. (2017). Age-associated Microbial Dysbiosis Promotes Intestinal Permeability, Systemic Inflammation, and Macrophage Dysfunction. *Cell Host & Microbe* *21*, 455–466.e4. <https://doi.org/10.1016/j.chom.2017.03.002>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Syrian Hamster anti-mouse Podoplanin antibody	BioLegend	Cat#: 127402 RRID: AB_1089187
Goat anti-tdTomato Polyclonal Antibody	OriGene Technologies	Cat#: AB8181-200 RRID: AB_2722750
Donkey anti-goat Cy3	Jackson ImmunoResearch	Cat #: 705-166-147 RRID: AB_2340413
Rabbit anti- Syrian Hamster Alexa 488	Jackson ImmunoResearch	Cat #: 307-546-003 RRID: AB_2339598
Mouse anti-CD45 Pacific Blue	BioLegend	Cat#: 103126 RRID: AB_493535
Mouse anti-Ly-6C Brilliant Violet 510	BioLegend	Cat # 128033 RRID: AB_2562351
Mouse anti-CD31 Brilliant Violet 605	BioLegend	Cat # 102427 RRID: AB_2563982
Mouse anti-CD326 (Ep-CAM) FITC	BioLegend	Cat #: 118207 RRID: AB_1134106
Mouse anti-Podoplanin PE/Cyanine7	BioLegend	Cat #: 127412 RRID:AB_10613648
Mouse anti-I-A/I-E (MHCII) Pacific Blue	BioLegend	Cat # 107619 RRID: AB_493528
Mouse anti-CD11b Brilliant Violet 605	BioLegend	Cat # 101257 RRID: AB_2565431
Mouse anti-Ly-6G FITC	BioLegend	Cat #: 127606 RRID: AB_1236494
Mouse anti-CD45 PerCP/ Cyanine5.5	BioLegend	Cat #: 103132 RRID: AB_893340
Mouse anti-CD11c PE/Cyanine7	BioLegend	Cat #: 117318 RRID: AB_493568
Mouse anti-F4/80 APC	BioLegend	Cat #: 123116 RRID: AB_893481
Mouse anti-Ly-6G PE	BioLegend	Cat #: 127607 RRID: AB_1186104
Recombinant rabbit monoclonal anti-mouse C3	Abcam	Cat #: ab259700 RRID: AB_3076186
Rabbit anti-mouse C3 antibody	Santa Cruz Biotechnology	Cat #: sc-20137 RRID: AB_634667
Goat anti-rabbit IgG-Alexa 647 antibody	Invitrogen	Cat # A-21244 RRID: AB_2535812
<b>Bacterial and virus strains</b>		
<i>Prevotella</i> spp. strain CR1	This study	strain CR1
<i>Proteus mirabilis</i> strain CR2	This study	strain CR2
<i>Citrobacter rodentium</i> DBS100	Dr. Lynn Bry <sup>24</sup>	DBS100
GFP <sup>+</sup> <i>C. rodentium</i> LB1	Dr. Lynn Bry <sup>24</sup>	LB1
Enterohemorrhagic <i>Escherichia coli</i> EDL931	ATCC	EDL931
GFP <sup>+</sup> EHEC EDL931 DLK1	This study	DLK1
Commensal bacteria strains used in <a href="#">Figure 2H</a>	This study	<a href="#">Table S3</a>

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Biological samples</b>		
Healthy human volunteer donor fecal samples	Alavi et al. <sup>11</sup>	N/A
<b>Chemicals, peptides, and recombinant proteins</b>		
Collagenase, Type II	Gibco	Cat# 17101015
Dispase II	Gibco	Cat# 17105041
Dithiothreitol	Millipore-Sigma	Cat# D0632-10G
UltraPure 0.5M EDTA, pH 8.0	Invitrogen	Cat# 15575020
UltraPure™ Phenol:Chloroform:Isoamyl Alcohol (25:24:1, v/v)	Invitrogen	Cat# 15593049
LPS-EK (LPS from <i>E. coli</i> K12) Ultrapure	InvivoGen	Cat# tlr1-pek1ps
Pam3CSK4	InvivoGen	Cat# tlr1-prms
PGN-SA (PGN from <i>Staphylococcus aureus</i> )	InvivoGen	Cat# tlr1-pgns2
LTA-SA (Purified LTA from <i>S. aureus</i> )	InvivoGen	Cat# tlr1-pslta
Trypticase soy agar (TSA) with 5% Sheep Blood	Thermo Scientific	Cat# R01200
Brucella agar with 5% Sheep Blood, Hemin and Vitamin K	Thermo Scientific	Cat# R01254
Fluid Thioglycollate Medium	BD, Becton, Dickinson and Company	Cat# 220888
Laked Brucella Blood Agar with Kanamycin and Vancomycin (LKV) Agar	Hardy Diagnostics	Cat# J87
Bacto Yeast Extract (for PYG broth)	BD, Becton, Dickinson and Company	Cat# 212750
Bacto Proteose Peptone No.3 (for PYG broth)	BD, Becton, Dickinson and Company	Cat# 211693
EDTA-free Protease Inhibitor Cocktail	Roche	Cat# 04693159001
Cobra Venom Factor (CVF)	CompTech	Cat# A150
<b>Critical commercial assays</b>		
QIAquick PCR Purification kit	QIAGEN	Cat# 28106
Qubit dsDNA HS Assay	Thermo Fisher	Cat# Q32854
Mouse C3 ELISA kit	Abcam	Cat# ab157711
Mouse C5 ELISA kit	Abcam	Cat# ab264609
Human C3 ELISA kit	Abcam	Cat# ab108823
Chromium Next GEM Single Cell 3' Reagent Kit v3.1	10X Genomics	Cat# CG000315
RNAScope Probe-Mm-C3	Advanced Cell Diagnostics	Cat# 417841
<b>Deposited data</b>		
Raw 16S sequencing	This paper	PRJNA928752
UliRNASeq data of sorted gut cells in WT mice	This paper	GSE248940
scRNASeq data of C3-expressing intestinal cells	This paper	GSE 250268
<b>Experimental models: Organisms/strains</b>		
Mouse: C57BL/6	Jackson Laboratory	Stock #: 000664
Mouse: C57BL/6	Taconic Biosciences	Stock#: C57BL/6NTac
Mouse: C57BL/6	Charles River Laboratories	Stock #: 027
Mouse: B6.129S4-C3 <sup>tm1Crr/J</sup>	Jackson Laboratory	Stock #:029661
Mouse: B6-C1qa <sup>tm1d(EUCOMM)Wtsi/Tenn/J</sup>	Jackson Laboratory	Stock #:031675
Mouse:B6.129S4-Mbl1 <sup>tm1Kata</sup> Mbl2 <sup>tm1Kata/J</sup>	Jackson Laboratory	Stock #:006122
Mouse: B6(Cg)-Tlr4 <sup>tm1.2Karp/J</sup>	Jackson Laboratory	Stock #:029015
Mouse: B6.129P2(SJL)-Myd88 <sup>tm1.1Defr/J</sup>	Jackson Laboratory	Stock #:009088
Mouse:C3 <sup>IRES-tdTomato</sup> C57BL/6	Provided by Dr. Claudia Kemper	Kolev et al. <sup>7</sup>
Mouse: Germ Free C57BL/6 mice	Harvard Medical School Gnotobiotic Core Facility	N/A

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Oligonucleotides</b>		
gPrevo-F primer: CACRGTAACGATGGATGCC	Gálvez et al. <sup>13</sup>	N/A
gPrevo-R primer: GGTCGGGTTGCAGACC	Gálvez et al. <sup>13</sup>	N/A
C. rodentium-16S-rRNA-specific probe: TTACCTACCAACAAGCTAATCCCATCTGGAA GAGTGAGTAGTAGTGGAGT	This study	N/A
Eubacterial 338 FISH probe1: TCCCTACTGCT GCCTCCCGTAGGAGTCTGGATGTGATGGAA GTTAGAGGGT	This study	N/A
Eubacterial 338 FISH probe2: /5Acryd/TCCCTACTGCTGCCTCCCGTAGGAG TCTGGATTGAGTGGATGGAGTGTA	This study	N/A
Readout Probe1: Alexa750-ACTCCACTACTACTACTCT	Moffitt et al. <sup>44</sup>	N/A
Readout Probe2: Cy5-ACCCTCTAACTTCCATCACA	Moffitt et al. <sup>44</sup>	N/A
<b>Recombinant DNA</b>		
plasmid pUC18T-mini-Tn7T-Tp-gfpmut3	Choi and Schweizer <sup>45</sup>	N/A
<b>Software and algorithms</b>		
FlowJo V10	Treestar	<a href="http://www.flowjo.com">www.flowjo.com</a>
FIJI Is Just Image J version 2	ImageJ	<a href="https://fiji.sc">https://fiji.sc</a>
QIIME2	Bolyen et al. <sup>46</sup>	<a href="https://qiime2.org/">https://qiime2.org/</a>
Seurat R toolkit	Hao et al. <sup>47</sup>	<a href="https://satijalab.org/seurat/">https://satijalab.org/seurat/</a>
<b>Other</b>		
FACSAria	BD	N/A
FACSymphony	BD	N/A
LSR-II	BD	N/A
MACSQuant cytometer	Miltenyi Biotec	N/A

**RESOURCE AVAILABILITY****Lead contact**

Please direct requests for resources and reagents to lead contact, Dennis Kasper ([dennis\\_kasper@hms.harvard.edu](mailto:dennis_kasper@hms.harvard.edu)).

**Materials availability**

This study did not generate new unique reagents.

**Data and code availability**

- Raw 16S rRNA gene sequencing data, single-cell RNA-seq data, and Ultra-low-input RNASeq have been deposited at NCBI and are publicly available as of the date of publication. Accession numbers are listed in the [key resources table](#). Microscopy data reported in this paper will be shared by the [lead contact](#) upon request.
- There is no original code from this study.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

**EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS****Mice**

C57BL/6 mice were purchased from Jackson Laboratories, Taconic Biosciences, or Charles River Laboratories. B6.129S4-C3<sup>tm1Crr</sup>/J, B6-C1qa<sup>tm1d(EUCOMM)Wtsi/Tenn</sup>/J, B6.129S4-Mbl1<sup>tm1Kata</sup> Mbl2<sup>tm1Kata</sup>/J, B6(Cg)-Tlr4<sup>tm1.2Karp</sup>/J, and B6.129P2(SJL)-Myd88<sup>tm1.1Defr</sup>/J mice were purchased from Jackson Laboratories. C3-tdTomato reporter (C3<sup>ires-tdTomato</sup> C57BL/6) mice were provided by Dr. Claudia Kemper (National Institute of Health). Mice were bred and housed in an SPF animal facility at Harvard Medical School (HMS). GF

C57BL/6 mice were maintained at the Harvard Medical School Gnotobiotic Core Facility. Age-matched 6- to 12-week-old littermate male and female mice were used for experiments. All mice were used in accordance with animal care guidelines from Harvard Medical School Standing Committee on Animals and the National Institutes of Health.

### Human samples

All human samples were part of a study approved by the UCR Institutional Review Board and followed NIH guidelines.<sup>11</sup> We collected intact fecal samples from a cohort of healthy adult volunteers at the University of California, Riverside. Inclusion criteria were: 1) age between 18 and 40 years, 2) must be able to provide signed and dated informed consent, 3) must be willing and able to provide stool specimen. Exclusion criteria were: 1) systemic antibiotic usage (oral, intramuscular, or intravenous) in the 2 weeks prior to sampling; 2) acute disease at time of enrollment (presence of moderate or severe illness with or without fever); 3) diarrhea (liquid or very loose stools not associated with a change in diet) in the 2 weeks prior to sampling; 4) active uncontrolled GI disorders or diseases including Inflammatory bowel disease (IBD), ulcerative colitis, Crohn's disease, or indeterminate colitis, persistent, infectious gastroenteritis, colitis, or gastritis, and chronic constipation; 5) Major surgery of the GI tract, excluding cholecystectomy and appendectomy, but including major bowel resection at any time. Age and sex information for study participants is not provided in this manuscript due to privacy agreements. The focus of this study remains on the variability in fecal C3 levels among healthy adult individuals. Fecal samples were collected aseptically from each person at UCR and immediately preserved at  $-80^{\circ}\text{C}$  and shipped on dry ice to Harvard Medical School for subsequent experiments.

## METHOD DETAILS

### Antibiotic treatment and fecal transfer

Antibiotics were given via drinking water. For broad-spectrum antibiotic treatment (ABX), 0.5g/L of Vancomycin, 1g/L of Neomycin, 1g/L of Ampicillin and 1g/L of Metronidazole were added to drinking water. Antibiotic water was refreshed every three days. To confirm antibiotic activity, fecal matter was resuspended in PBS (1 ml/pellet) and cultured on Trypticase soy agar (TSA) and Brucella blood agar plates in aerobic and anaerobic conditions, respectively. For mouse microbiota fecal transfer, fecal samples collected from C57BL/6 WT mice from different vendors and resuspended in PBS at 100 mg/ml in the anaerobic chamber. Aliquots were kept at  $-80^{\circ}\text{C}$ . GF mice were orally gavaged with 200  $\mu\text{l}$  of stock. For the co-housing experiment, various sets of 4-week-old JAX and CR mice were paired in a 1:1 ratio for specified durations, including two weeks, one day, or three days, after which they were separated. Fecal samples were subsequently collected two weeks after the initial co-housing day and examined for C3 levels using ELISA or determining microbiota composition using 16S rRNA gene sequencing. For human microbiota fecal transfer, human fecal samples were collected and prepared as 25% glycerol stock as described.<sup>1</sup> Fecal samples were normalized to be representative of 16S rRNA gene levels in 300ul of  $\text{OD}_{600}=0.4$  culture of defined community. GF mice were orally gavaged with 300ul normalized human fecal slurries. For monocolonized mice, GF mice were orally inoculated by gavage with a liquid culture of single bacteria strain at 4 weeks of age and kept in an IsoCage (Isolator at cage level: airtight cages with high negative pressure, TECNIPLAST) as described.<sup>48</sup> Each group of mice was housed in gnotobiotic IsoCages under sterile conditions for 2 weeks. Fecal material was collected 2 weeks after bacterial inoculation to verify monocolonization.

### 16S rRNA gene sequencing of fecal samples and data analysis

Bacterial genomic DNA from frozen stool samples was extracted using phenol:chloroform:isoamyl alcohol and then purified further with the QIAquick PCR Purification kit (Qiagen). Purified DNA was quantified by Qubit dsDNA HS Assay (Thermo Fisher Cat# Q32854) and normalized. Amplicons were purified and quantified by Qubit dsDNA HS Assay and combined with equal mass to make a pooled library. The pooled library was multiplexed sequenced (Illumina MiSeq, 251 nt x 2 pair-end reads with 12 nt index reads) through Harvard's Biopolymers Facility. Raw sequencing data was processed with QIIME2.<sup>46</sup> In brief, raw sequencing data was imported to QIIME2 and demultiplexed, then DADA2 was used for sequence quality control and feature table construction. The feature table was used for beta diversity analysis, taxonomic analysis, and differential abundance testing using QIIME2. Beta group significance was determined by permutational analysis of variance (PERMANOVA). Identification of taxa associated with different groups was determined using Analysis of Composition of Microbiomes (ANCOM). Human fecal 16S rRNA gene library sequencing were described previously.<sup>49</sup> Spearman correlation between human gut microbiota at genus level and C3 levels was performed in R.

### Isolation of *Prevotella spp.* strains from Charles River (CR) mice gut microbiota

Fecal samples were collected from Charles River mice and homogenized in BBL thioglycollate media in an anaerobic chamber. Serial dilutions from the homogenized fecal sample were plated on Laked Brucella Blood Agar with Kanamycin and Vancomycin (LKV) Agar (Hardy Diagnostics) and cultured at  $37^{\circ}\text{C}$  in the anaerobic chamber for three days. Individual pigmented colonies were selected from these plates and cultured in peptone-yeast-glucose (PYG) broth for two days in the anaerobic chamber. Stocks were made from the liquid cultures and stored in  $-80^{\circ}\text{C}$ . Bacterial gDNA were extracted from the liquid cultures using phenol:chloroform:isoamyl alcohol and then purified further with the QIAquick PCR Purification kit (Qiagen). *Prevotella*-specific 16S rRNA gene primers (gPrevo-F and R-)<sup>13</sup> were used to confirm the stocks as *Prevotella spp.* and named as *Prevotella CR1*.

### C3 and C5 ELISA

C3 or C5 levels in feces or serum of mice were measured with a Mouse C3 ELISA kit (Abcam, ab157711) or a Mouse C5 ELISA kit (Abcam, ab264609) according to the manufacturer's instructions. C3 levels in human feces were measured with a Human C3 ELISA kit (Abcam, ab108823) according to the manufacturer's instructions. Briefly, fecal samples were collected and resuspended in dilution buffer at 100 mg/ml, centrifuged at 14,000g for 10 minutes at 4°C, and the supernatant was then used for C3/C5 levels measurement. Blood was collected from mice and centrifuged at 3000 rpm for 10 minutes after coagulation. Serum was collected for C3 concentration measurements.

### Gut permeability assay

A gut permeability assay was performed as previously described.<sup>52</sup> Briefly, mice were fasted for four hours before oral gavage of 600 mg/kg FITC-3-5-kDa dextran (Sigma, FD4). Serum was collected six hours after gavage as well as control mice without gavage and analyzed using a Synergy HT plate reader (Bio-Tek) with 485 nm excitation and 530 nm emission.

### Preparation of intestinal cells and flow cytometry

Intestinal tissues were prepared as previously described.<sup>48</sup> Briefly, distal colon were collected, treated with 30 mL of RPMI containing 1 mM dithiothreitol, 20 mM EDTA, and 2% FBS at 37°C for 15 min to isolate epithelial cells and intraepithelial lymphocytes. The intestinal tissues were then minced and dissociated in RPMI containing collagenase II (1.5 mg/mL; GIBCO), dispase (0.5 mg/mL), and 1% FBS, with constant stirring at 37°C for 45 min. Single-cell suspensions were then filtered and washed with 10% FBS RPMI solution. Cells isolated above were stained with surface antibodies in a dilution of 1:200 on ice for 20 minutes in staining buffer (2% FBS in PBS) and washed twice with the same buffer. Cells were acquired with a BD FACSymphony (Beckman Coulter) or a BD LSR-II and analyzed in Flowjo.

### Stromal cell isolation and culture

Colonic C3-expressing stromal cells were isolated from C3<sup>IRES-tdTomato</sup> C57BL/6 mice. Briefly, single cell suspensions were prepared as described above from C3<sup>IRES-tdTomato</sup> C57BL/6 mice. Live C3-expressing stromal cells were sorted as DAPI<sup>-</sup>Epcam<sup>-</sup>CD45<sup>-</sup>CD31<sup>+</sup>PDPN<sup>+</sup>tdTomato<sup>+</sup> using a FACSAria (BD) into RPMI medium containing 10% FBS, then seeded on 96-well flat cell culture dishes (10,000 cells/well) in DMEM supplemented with 10% FBS and 1% penicillin/streptomycin at 37°C in the presence of 5% CO<sub>2</sub>. Colonic stromal cells were isolated from TLR4<sup>-/-</sup>, MyD88<sup>-/-</sup>, and WT BL6 mice using similar protocol as sorting DAPI<sup>-</sup>Epcam<sup>-</sup>CD45<sup>-</sup>CD31<sup>+</sup>PDPN<sup>+</sup> from single colonic cell suspensions. After 3 days of culture, live cells were collected, counted, and seeded as 10,000 cells/well in DMEM supplemented with 10% FBS and 1% penicillin/streptomycin at 37°C overnight. Bacterial stimulants were fixed with 1% formalin for 24 hours and checked for no viability or killed at 120°C for 30 mins. Bacterial stimulants were normalized to OD<sub>600</sub> 1 in RPMI medium and added 200ul to stromal cells, RPMI medium are used as control. LPS or MDP were dissolved in RPMI medium at the indicated concentrations and 200 ul were given to stromal cells.

### scRNASeq profiling of C3-expressing cells in colon

scRNASeq of colonic C3-expressing cells isolated from C3<sup>IRES-tdTomato</sup> C57BL/6 mice was performed using 10x Genomics Chromium Next GEM Single Cell 3' Reagent Kit (V3.1 chemistry). Briefly, colons from one male and one female C3<sup>IRES-tdTomato</sup> C57BL/6 mice were collected, and single cell suspensions were prepared as described above. Live C3-expressing cells were sorted as DAPI<sup>-</sup>tdTomato<sup>+</sup> using FACSAria (BD) into DMEM containing 10% FBS. A total of ~10,000 C3-expressing cells from each mouse were collected and pooled together. Samples were loaded into Chip G per the user guide from 10x Genomics. The Chip G was then run on a 10x Chromium Controller. 100ul of GEM emulsion was taken from the chip, and inspected visually for inconsistencies in the recovered volume, the uniformity of the emulsion, or the relative amount of partitioning oil, and incubated at RT in a thermocycler (45 min at 53°C, 5 min at 85°C, then held at 4°C until the next step). The emulsion was then broken by adding 125ul of Recovery Agent. The cDNA was then purified using Dynabeads MyOne SILANE. Per the 10x 3'v3.1 user guide, the 35ul of the purified cDNA was mixed with 15ul 10x cDNA primers (PN 200089) and 50ul of 10x Amp Mix (PN 200047), then amplified on a thermocycler (98°C for 3 minutes, [98°C for 15 seconds, 63°C for 20 seconds, 72°C for 60 seconds], 12 repeats of the bracketed steps, 72°C for 60 seconds, then held at 4°C). The amplified cDNA was then purified and separated based on size using a 0.6x Beckman Coulter SPRIselect Reagent cleanup. The larger bound fragments in the bead pellet containing transcript-derived cDNA were eluted in 35ul of Qiagen Buffer EB for use in Gene Expression library construction. Library size was measured by an Agilent Bioanalyzer 2100 High Sensitivity DNA assay and quantified using a Qubit dsDNA HS Assay kit on a Qubit 4.0 Fluorometer.

A portion of the transcript-derived cDNA was fragmented, ligated to an Illumina Read 2 Sequencing Adaptor, and indexed with a unique 10x T Set A index, yielding the GEX library. 50ng of the transcript-derived cDNA was mixed with nuclease-free water up to a volume of 20ul, then mixed with 15ul of nuclease-free water, 5ul of 10x Fragmentation Buffer (PN 2000091), and 10ul of 10x Fragmentation Enzyme (PN 2000090), loaded onto a thermocycler, and incubated at 32°C for 2 minutes, then 65°C for 35 minutes, and then held at 4°C. The fragmented product was purified with a 0.6x SPRI, with 75ul of the supernatant saved. The supernatant was purified with a 0.8x SPRI, with the pellet-bound DNA saved and eluted in 50ul. All 50ul of eluted sample was mixed with 20ul 10x Ligation Buffer (PN 2000092), 10ul DNA ligase (PN 220110), and 20ul of Adaptor Oligos (2000094), then incubated on a thermocycler at 20°C for 15min. This product was purified using a 0.8x SPRI cleanup, with the pellet-bound DNA saved and eluted in 30ul. These

30ul were then mixed with 20ul of a unique 10X T Set A index and 50ul of 10x Amp Mix, then incubated on a thermocycler at 98°C for 45 seconds, [98°C for 20 seconds, 54°C for 30 seconds, 72°C for 20 seconds], 13 repeats of the bracketed steps, 72°C for 60 seconds, then held at 4°C. The sample was then purified with a 0.6x SPRI, saving and transferring 150ul of the supernatant. The transferred supernatant was purified with a 0.8x SPRI and the pellet-bound DNA was eluted in 35ul. The finished library was measured on an Agilent BioAnalyzer and quantified on a Qubit 4.0. This library was then sequenced on an Illumina Novaseq SP 100, yielding approximately ~366M reads in the GEX.

Raw sequencing data in the form of binary files were converted into FASTQ files using the `mkfastq` command from CellRanger v6.1.0. Then, the FASTQ files were demultiplexed into gene expression libraries using CellRanger's `count` command, and only reads mapping unambiguously and with less than 2 mismatches were kept. For each single-cell library, reads were mapped to the mouse mm10 transcriptome using CellRanger. Duplicate reads, those mapping to multiple regions or having a low alignment score (MAPQ < 10) were filtered out. A

final gene expression matrix with genes in rows and cells in columns was then constructed. Data were then analyzed using the Seurat R toolkit<sup>47</sup> (<https://satijalab.org/seurat/>). Briefly, cells were first filtered based on the UMI count and mitochondrial gene percentage. Cells with < 500 UMIs or > 10 % mitochondrial genes were excluded from further analysis. Gene expression values for each cell were then normalized by the total expression, multiplied by an arbitrary scale factor of 10,000, and log-transformed. A principal component analysis (PCA) was performed on the top 200 most variable genes that were expressed in > 1% of the cells. The number of statistically significant PCs was determined by comparison with PCA over a randomized matrix as described previously.<sup>50</sup> The data were then visualized using the Uniform Manifold Approximation and Projection (UMAP) dimensionality reduction algorithm<sup>51</sup> based on the significant PCs. Clustering was done on selected PCs using the `FindNeighbors` followed by the `FindClusters` functions, and the identity of the main clusters was determined based on the expression of common identifiers: stromal (*pdpn*), epithelial (*epcam*), myeloid (*lyz2*).

### Gene expression profiling by population RNAseq

A thousand live epithelial (EpCAM<sup>+</sup>), CD45<sup>+</sup>, and stromal cells (EpCAM<sup>-</sup>CD45<sup>-</sup>CD31<sup>+</sup>PDPN<sup>+</sup>) were doubly sorted from colons of WT SPF mice by BD FACS Aria II. Sorted cells were directly lysed with 5mL of TCL buffer (Qiagen, 1031576) containing 1% 2-mercaptoethanol (Sigma-Aldrich, M6250), then flash frozen using dry-ice and stored at -80°C. Smart-Seq2 libraries for low-input RNA-seq were prepared by the Broad Technology Labs and were subsequently sequenced through the Broad Genomics Platform according to the Immgen protocol ([https://www.immgen.org/img/Protocols/ImmGenULI\\_RNAseq\\_methods.pdf](https://www.immgen.org/img/Protocols/ImmGenULI_RNAseq_methods.pdf)). The GCT and CLS files generated and used for downstream analysis. After normalization and filtered low count reads by DESeq2, fragments/counts per million were calculated for all the genes in the complement systems. Histogram and heatmaps with main components in complement system were depicted in Figures 3F, 6G, and 6H.

### Analysis of published single-cell data

For the dotplot gene expression profiles of complement system genes in human colonic cells, a full scRNASeq dataset of 428,000 intestinal cells<sup>20</sup> was downloaded from <https://www.gutcellatlas.org/>, and then filtered as to keep cells with more than 500 genes, less than 25% mitochondrial reads, and genes with expression in more than three cells. A Scrublet (v.0.2.1) score cut-off of 0.25 was applied to exclude doublet. A global-scaling normalization method "LogNormalize" was applied to normalize the feature expression measurements for each cell by the total expression, multiplies this by a scale factor (10,000 by default), and log-transformed. Complement system gene expression table was extracted from the normalized table and plotted the expression levels using Dotplot in R.

### Citrobacter rodentium infection and C3-deposition analysis

*C. rodentium* DBS100 and GFP<sup>+</sup> *C. rodentium* LB1 were provided by Dr. Lynn Bry (Brigham and Women's Hospital).<sup>24</sup> Survival studies were performed by orally infected 5X10<sup>8</sup> CFU of *C. rodentium* DBS100 with twenty-one-day-old WT and C3<sup>-/-</sup> C57BL/6 mice and followed for 28 days. Adult (8–12 weeks) WT and C3<sup>-/-</sup> C57BL/6 mice were used to evaluate the roles of complement C3 in infection using *C. rodentium* doses between 5X10<sup>8</sup>–1X10<sup>10</sup> CFU as indicated. Animal survival, weight, and disease symptoms were monitored daily for 3 weeks. Fecal burden represented by CFU/g of feces were monitored from day 7 to day 14. For C3 deposition analysis, age matched WT and C3<sup>-/-</sup> mice were oral infected with 5X10<sup>8</sup> CFU of GFP<sup>+</sup> *C. rodentium* LB1. At day 7, fecal pellets were collected and resuspended in PBS containing a Roche proteinase inhibitor cocktail (Roche Cat. No. 04693159001). Cell suspensions were centrifuged at 50X g for 10 min at 4°C in order to pellet large particulate material. The supernatant fluids were filtered through a 40µm filter and the flow-through was then centrifuged to collect the bacterial cell pellet. Fecal bacteria were then washed by resuspending the pellet in PBS, centrifuging, and resuspending with a proteinase inhibitor twice. The washed bacteria cells were resuspended in PBS containing 2% BSA and incubated for 30 minutes at 4°C before adding 1/10 volume of rabbit anti-mouse C3 antibody (sc-20137). After incubating the bacterial cells at 4°C for 1 hour, cells were washed twice and resuspended in PBS containing 1:10,000 diluted goat anti-rabbit IgG-647 antibody (Invitrogen). After 30mins at 4°C, bacterial cells were washed again and resuspended in PBS before they were analyzed with a MACSQuant cytometer (Miltenyi Biotec). C3 deposition analysis in age-matched WT, MBL KO, and C1qa KO mice were performed in similar procedure with DyLight405 (ab201798) conjugated recombinant rabbit monoclonal anti-mouse C3 antibody (ab259700).

### Cobra venom factor (CVF) treatment of wildtype mice

12.5ug CVF per 25 grams of body weight were given to WT mice via intraperitoneal injection one day before *C. rodentium* infection and on days 4, 8, and 12 after infection. Control mice received saline injections. To test CVF depletion of C3, serum and fecal samples from CVF- or saline-treated mice were collected at day 0, day 1 and day 4 post CVF treatment to measure C3 levels using ELISA.

### Enterohemorrhagic Escherichia Coli (EHEC) infection

EHEC strain EDL931 were purchased from the American Type Culture Collection. The plasmid pUC18T-mini-Tn7T-Tp-gfpmut3<sup>45</sup> was electroporated into EHEC EDL931 competent cells. The correct transformant was selected and confirmed to be positive for GFP by PCR as well as by flow cytometry, and was designated as GFP<sup>+</sup> EHEC. 6-7-week-old WT and C3<sup>-/-</sup> C57BL/6 mice were orally gavaged with Streptomycin (1mg/gram of body weight), then 1X10<sup>9</sup> EHEC were given 24h after streptomycin treatment. Food was withdrawn 4 hours before giving EHEC and return soon after infection. For C3 deposition analysis, age-matched WT and C3<sup>-/-</sup> mice were orally infected with 1X10<sup>9</sup> CFU with GFP<sup>+</sup> EHEC. At day 4 post-infection, fecal pellets were collected, and fecal bacteria were isolated as described earlier and stained with Dylight405 conjugated recombinant rabbit monoclonal anti-mouse C3 antibody (ab259700). GFP<sup>+</sup> EHEC cells were identified as FITC<sup>+</sup> by flow cytometry on a MACSQuant cytometer.

### Mouse C3 RNAScope

Mouse C3 RNAScope *in situ* hybridization was performed using reagents and protocols from Advanced Cell Diagnostics. Briefly, colons from WT SPF mice were collected, fixed in 4% PFA, embedded in paraffin, and 5 μm sections were allowed to dry overnight. Sections were rehydrated twice in xylene for 5 min each, followed by two washes in 100% ethanol and one wash in 95% ethanol, 1 min each. After rehydration, the samples were incubated for 10 min in hydrogen peroxide, boiled in an antigen retrieval buffer for 15 min, followed by digestion with proteinase for 15 minutes at 40°C. Slides were washed twice with *in situ* hybridization (ISH) wash buffer, and hybridized with C3 probes (ACD#417841) for 2 hours at 40°C. After hybridization, amplification steps and staining were performed according to the protocol (RNAScope Fluorescent Reagent kit Assay).

### Immunofluorescent staining of C3-expressing cells in C3<sup>IRE5-tdTomato</sup> C57BL/6 mice

Distal colons were collected for staining. The intestines were opened and fixed in Silgard dish with 4% PFA at 4°C overnight. Subsequently, the tissues were dehydrated using 30% sucrose at 4°C overnight. The tissues were then embedded in OCT (Optimal cutting temperature compound) and frozen. 10μm sections were blocked with 10% donkey and goat serum in PBST (0.5% Triton-X100/PBS), then incubated with primary antibody (goat anti-tdTomato, 1:500; syrian hamster anti-pdpr, 1:500) overnight, followed by secondary antibody staining in PBST for 1 hour, and then stained with Hoechst for 5 mins, mounted with Prolong Antifade mounting medium, and sealed prior to imaging. Sections were imaged using a confocal microscopy at 63X and analyzed in Fiji. The C3-expressing cells isolated from C3<sup>IRE5-tdTomato</sup> C57BL/6 and WT mice were cultured and stained with goat anti-tdTomato in a 96-well plate with glass-like polymer bottom (Cellvis P96-1.5P) and imaged using a widefield microscopy to test anti-tdTomato antibody (Figures S4A–S4C).

### Bacterial probe staining

Colon tissues with fecal contents were collected and fixed by immersion in methacarn solution at 4°C for 48 hours. Tissues were then washed twice with 100% methanol (Sigma, MX0480-6), followed by twice with 100% ethanol (Fisher Scientific, BP2818500). Samples were then embedded in paraffin and cut into 4-μm-thick sections. In preparation for staining, the tissue paraffin slices were deparaffinized by heating at 60°C followed by immersion in four changes of room-temperature xylene (Sigma, 534056). The sample was then washed twice at room-temperature in 100% ethanol and rehydrated through a series of 95% then 70% ethanol incubations. The slices were then washed once with FISH-probe wash buffer<sup>44</sup> for 5 minutes, and incubated with a probe mixture comprised of the *C. rodentium* rRNA FISH probe, the Eub338 probes, and FISH hybridization buffer.<sup>44</sup> The samples were then incubated at 37 °C for 48 hours. Background fluorescence was removed by embedding the samples in a thin polyacrylamide gel and then cleared and treated with protease and detergent as described previously.<sup>44</sup> Samples were then washed with 2× SSC five times for imaging immediately or stored at 4 °C.

### 16S-rRNA-specific probes for *C. rodentium*

A 30-nt *C. rodentium*-16S-rRNA-specific probe was designed using a previous pipeline,<sup>44</sup> and this sequence was concatenated with an additional 20-nt 'readout' sequence complementary to a previously designed fluorescently labeled FISH probe.<sup>44</sup> Two, 30-nt eubacterial FISH probes were designed by extending the standard Eub388 probe sequence to a total length of 30-nt, and concatenating different 20-nt 'readout' sequences. One Eub388 FISH probe carries a 5' acrydite moiety, which served to anchor all eubacterial 16S rRNA during the tissue clearing process. As the sequence of the eubacterial probes do not overlap that of the *C. rodentium*-specific probe, we anticipated *C. rodentium* would be labeled by both.

### Bacterial imaging and data analysis

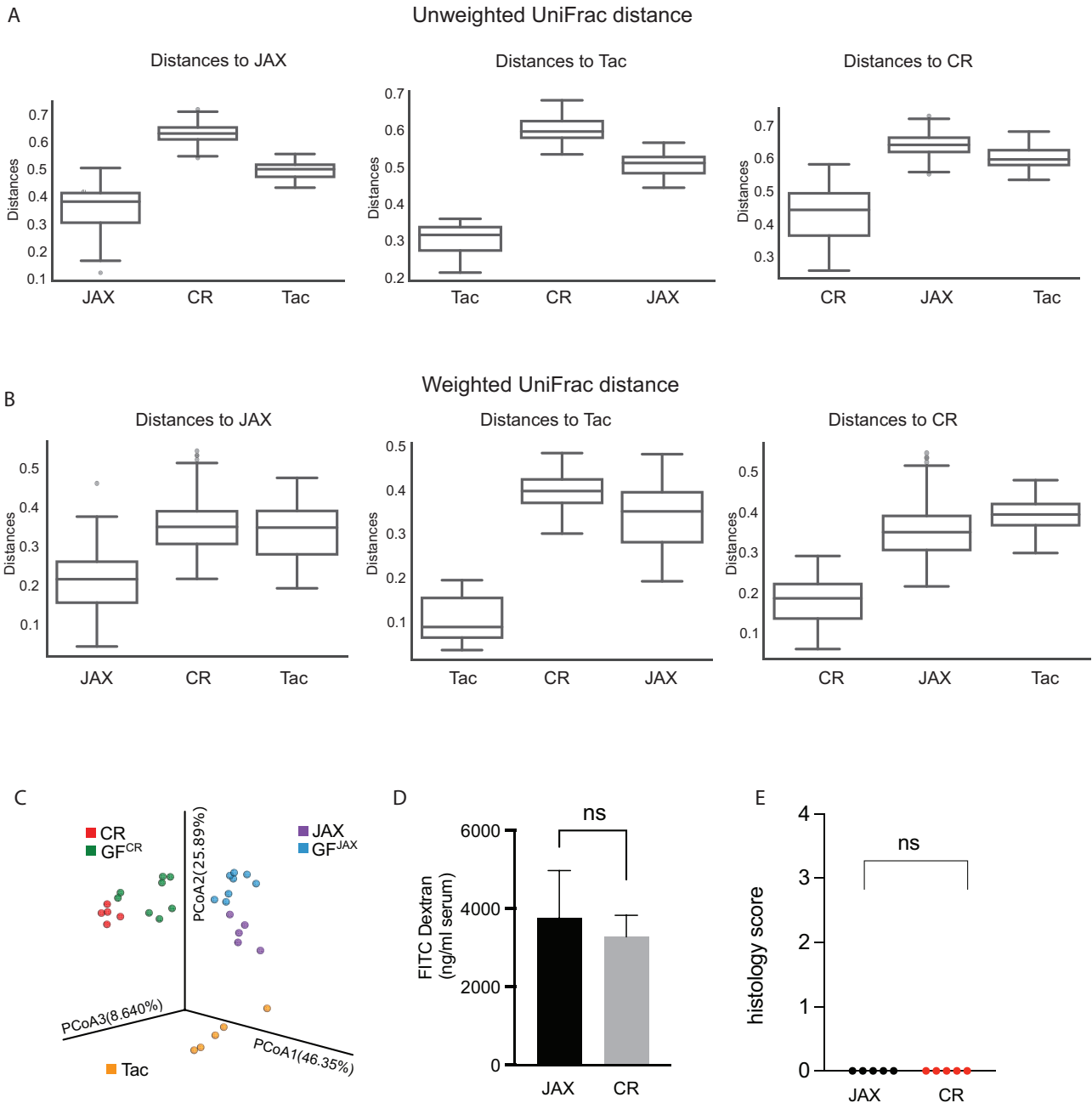
To prepare the sample for imaging, the sample was stained with two fluorescently-labeled readout probes complementary to the readout sequences on each of the *C. rodentium*-specific and Eub338 probes, as well as DAPI. The samples were imaged using

buffers and protocols described previously.<sup>44</sup> Briefly, the samples were imaged on a home-built epi-fluorescence microscope using 95 mW of 750-nm illumination (Alexa750), 170mW of 635-nm illumination (Cy5), 26 mW of 545-nm illumination (orange fiducial beads), 77 mW of 473-nm illumination (Alexa488), and 72 mW of 408-nm illumination (DAPI) provided by a Celesta light engine (Lumencor). Illumination was reflected to the sample via a penta-band dichroic (Semrock FF421/491/567/659/776-Di01-25x36) and emitted light filtered with a penta-notch filter (Semrock FF01-391/477/549/639/741-25; FF01-441/511/593/684/817-25). Samples were imaged with a  $\times 60$  CFI PlanApo oil objective (Nikon) and two high-performance CMOS cameras (Hamamatsu, Flash 4.0). The mosaic for the cross-section colon slides was generated by stitching single field-of-views together. The epithelial boundaries were identified based on the 405-channel using Fiji. The *C. rodentium* signals were measured inside the epithelial boundaries by thresholding the *C. rodentium*-specific channel on intensity and summing the number of pixels above the threshold in Fiji. As different cross-sections had different sizes, the measured *C. rodentium* signals were normalized by the measured length of the epithelial-lumen boundary in each slice.

### QUANTIFICATION AND STATISTICAL ANALYSIS

Data from independent experiments are representative of at least two independent replicates or as pooled data. Each experiments contains three-five animals/group. Statistical parameters including numbers, averages, deviation, and statistical tests are reported in the figures and corresponding legends. For all microscopy analysis, images were blinded prior to scoring. Data are represented as mean  $\pm$  standard error (SEM) throughout the figures. Statistical analyses were performed in GraphPad Prism or R.

# Supplemental figures



**Figure S1. Microbiome analysis of fecal samples and gut barrier function assays in C57BL/6 mice from three commercial vendors, related to Figure 1**

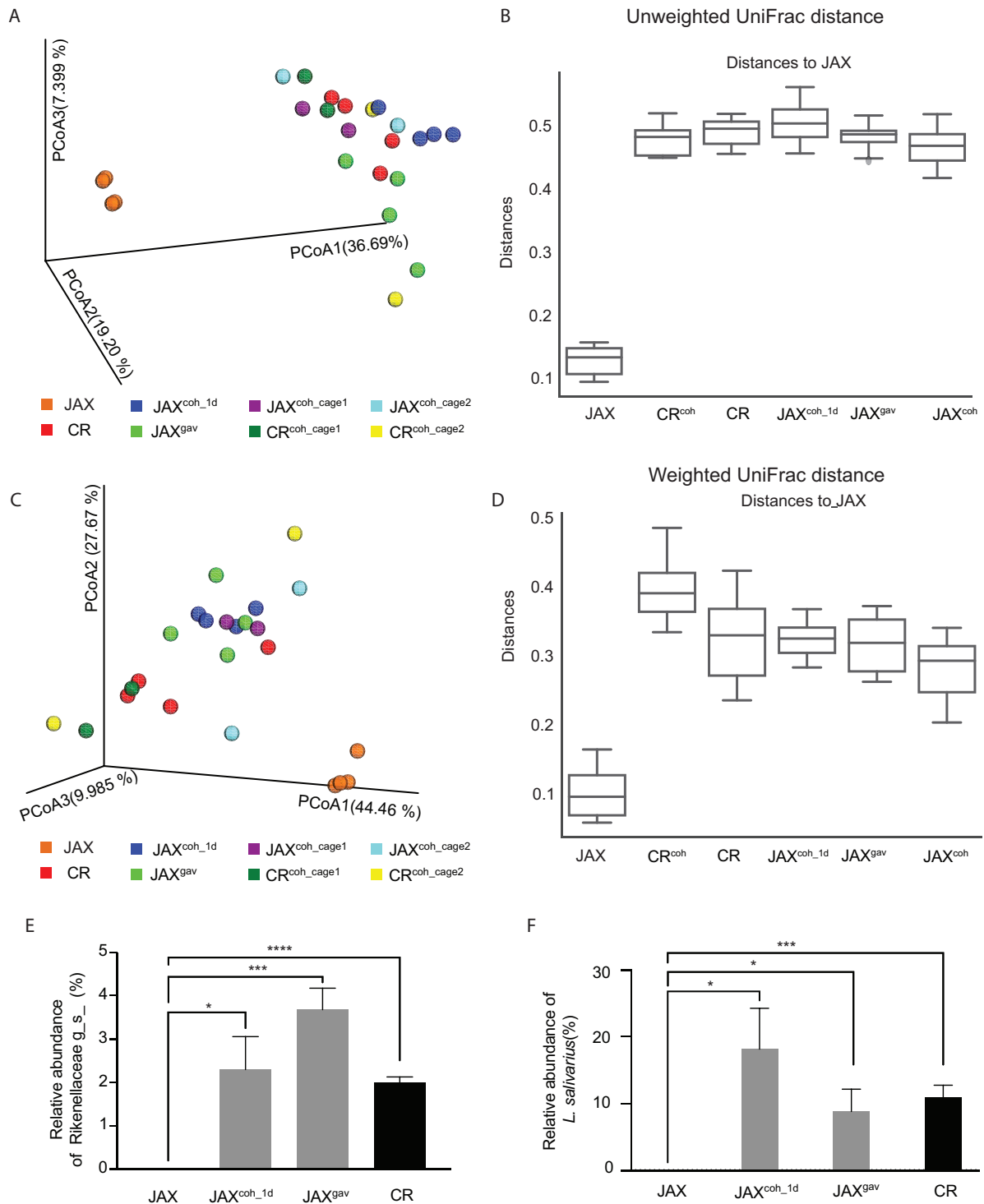
(A and B) Pairwise unweighted UniFrac distances (A) and weighted UniFrac distances (B) of fecal bacterial community composition within and across three commercial vendors: JAX, Jackson; Tac, Taconic; and CR, Charles River.

(C) Principal coordinate analysis (PCoA) of weighted UniFrac distance measurements based on the 16S rRNA gene sequencing data of the fecal bacterial composition in samples collected from C57BL/6 JAX, TAC, CR, or gnotobiotic mice colonized with JAX (GF<sup>JAX</sup>) or CR microbiota (GF<sup>CR</sup>) ( $q = 0.001$ , PERMANOVA).

(D) Fluorescence measurement of serum collected from JAX and CR mice orally gavaged with FITC-dextran.

(E) Histology scoring of colons collected from JAX and CR mice.

ns, not significant, unpaired Student's *t* test. Error bars represent mean ± SEM.

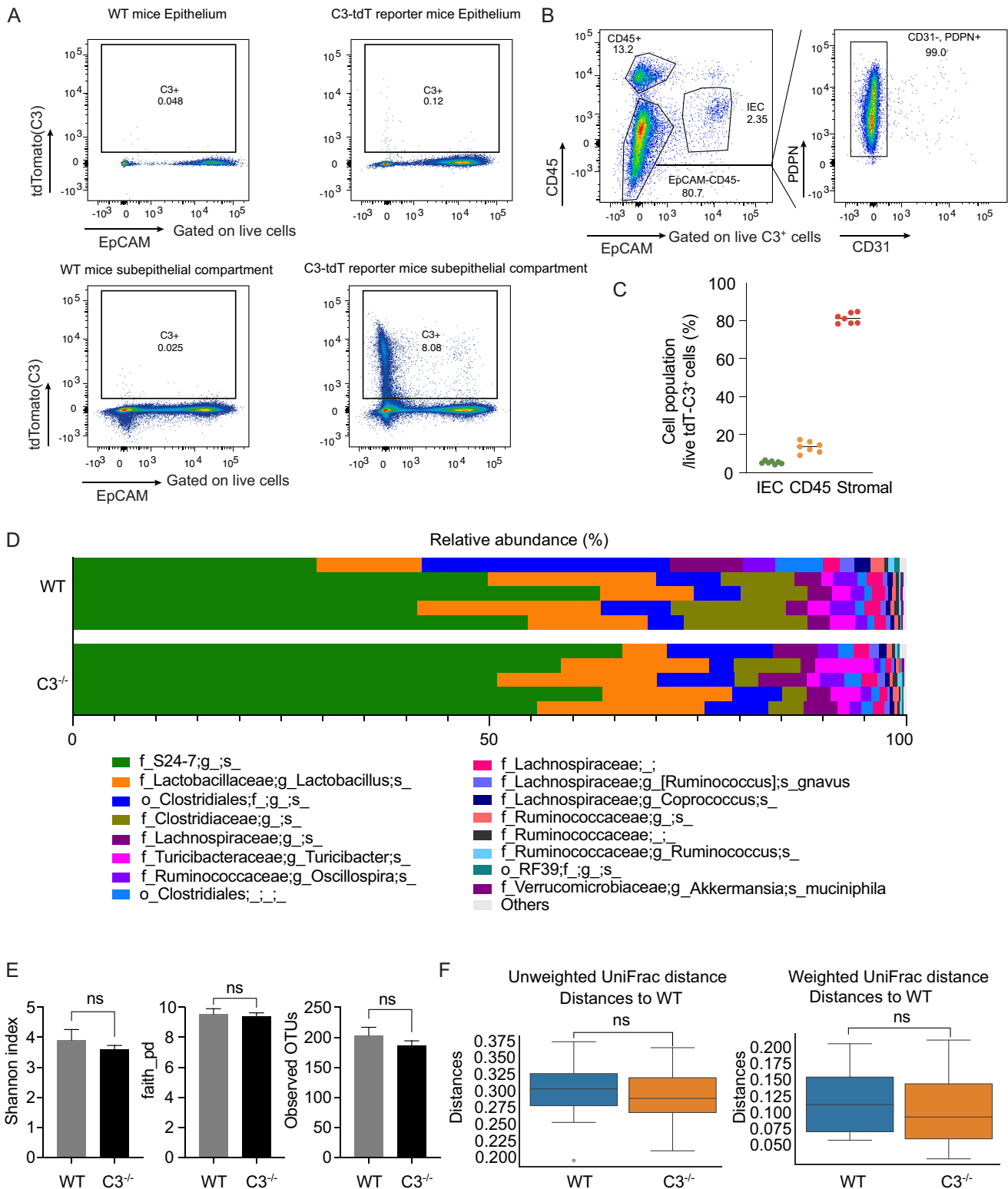


**Figure S2. Microbiome analysis of fecal samples in co-housing experiments, related to Figure 2**

(A and C) Principal coordinate analysis (PCoA) of unweighted UniFrac distance (A) and weighted UniFrac distance (C) measurements based on the 16S rRNA gene sequencing data of fecal bacterial composition in samples collected from mice as indicated (q = 0.001 between JAX to all other groups, PERMANOVA).

(B and D) Pairwise unweighted UniFrac distances (B) and weighted UniFrac distances (D) of the composition of bacterial communities in fecal samples within JAX and across groups compared with JAX.

(E and F) Relative abundance of *Rikenellaceae* spp. and *L. salivarius* in JAX or CR mice, as well as in JAX mice co-housed with CR mice for 1 day (JAX<sup>coh\_1d</sup>), or in JAX mice gavaged with CR microbiota (JAX<sup>gav</sup>). ns, not significant, \*p < 0.05, \*\*\*p < 0.001, \*\*\*\*p < 0.0001, unpaired Student's t test. Error bars represent mean ± SEM.



**Figure S3. Flow cytometric analysis of C3-expressing colonic cells in  $C3^{IRES-tdTomato}$  reporter mice, related to Figure 3**

(A) Flow cytometry gating of live colonic epithelial and subepithelial compartment of  $C3^{IRES-tdTomato}$  reporter mice showing the frequency of C3-tdTomato<sup>+</sup> cells with WT mice as control.

(legend continued on next page)

---

(B) Flow cytometry gating of live C3<sup>+</sup> colonic cells in subepithelial compartment of C3<sup>RES-tdTomato</sup> reporter mice showing the frequencies of CD45<sup>+</sup>, IEC, CD45<sup>-</sup>EpCAM<sup>-</sup>, and CD31<sup>-</sup>PDPN<sup>+</sup> cells.

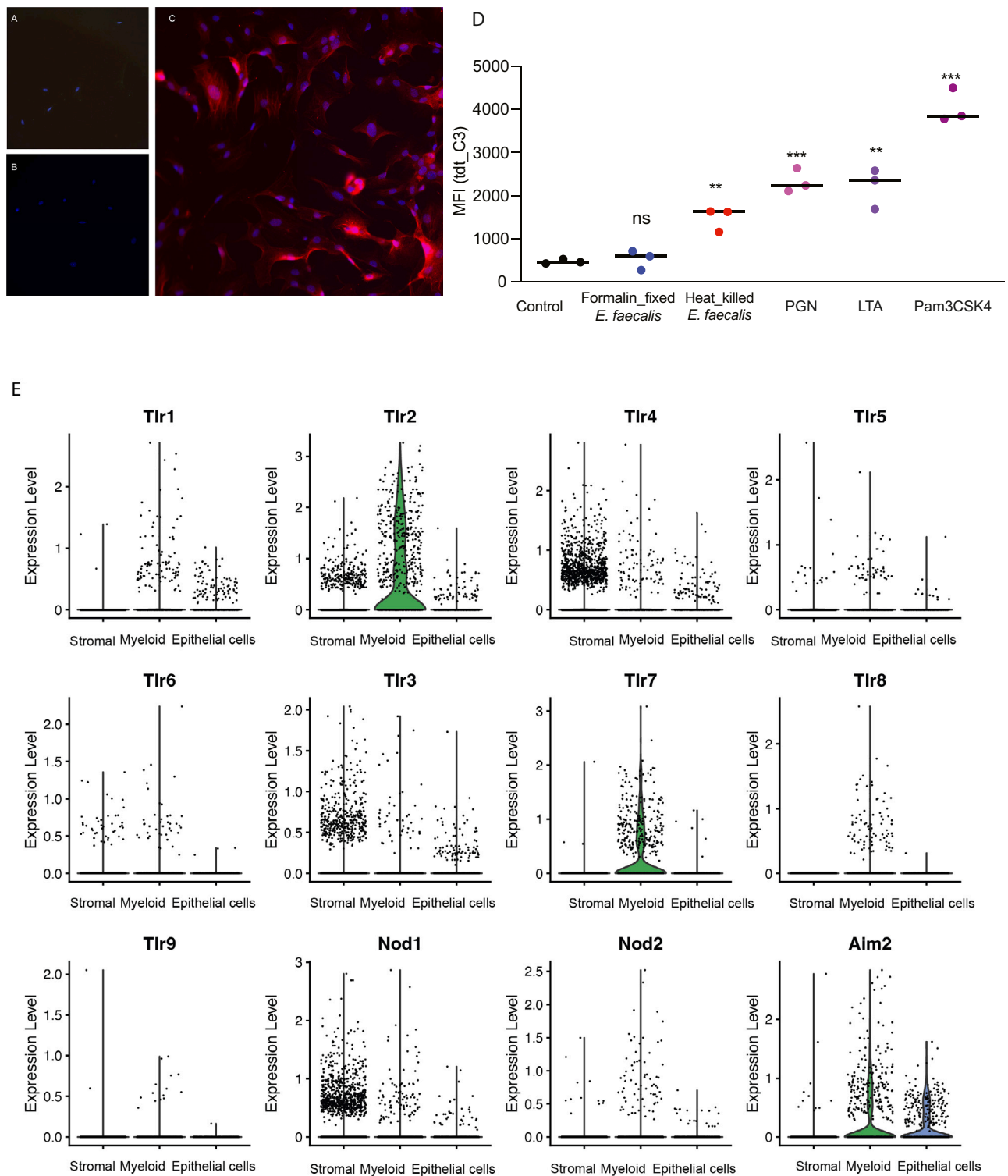
(C) Percentage of IECs, stromal, or CD45<sup>+</sup> cells in total colonic C3-expressing cells in the subepithelial compartment as assessed by flow cytometry in (B).

(D) Species-level abundance analysis of fecal microbiota from the WT and C3-deficient mice. ns, not significant, unpaired Student's t test.

(E) Shannon index, faith\_pd (phylogenetic diversity), and observed operational taxonomic units (OTUs) of bacterial communities in fecal samples within WT C57BL/6 mice and C3-deficient mice relative to WT C57BL/6 mice. ns, not significant, unpaired Student's t test.

(F) Pairwise unweighted UniFrac distances and weighted UniFrac distances of the composition of bacterial communities in fecal samples within WT C57BL/6 mice and C3-deficient mice relative to WT C57BL/6 mice. (q = 0.097 unweighted UniFrac distances between C3-deficient to WT mice, q = 0.103 weighted UniFrac distances between C3-deficient to WT mice, PERMANOVA).

Error bars represent mean ± SEM.



**Figure S4. Immunofluorescence microscopy of C3-expressing stromal cells and transcriptional expression levels of receptor genes in C3-expressing colonic cells, related to Figure 4**

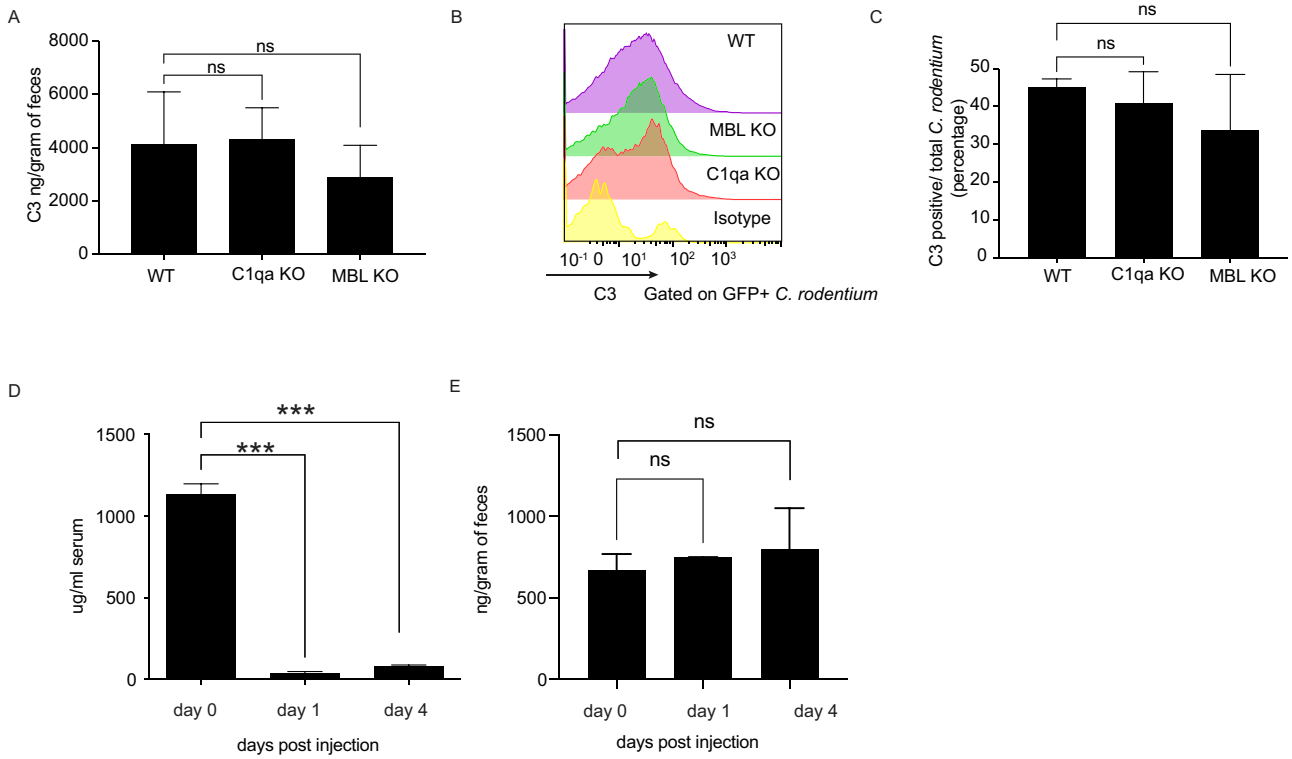
(A–C) Immunofluorescence microscopy of C3-expressing stromal cells isolated from the colon of C3<sup>IRES-tdTomato</sup> reporter mice, stained with a primary antibody against tdTomato (A), a secondary antibody against tdTomato (B), or both primary and secondary antibodies (red) against tdTomato (C), as well as Hoechst (blue).

(legend continued on next page)

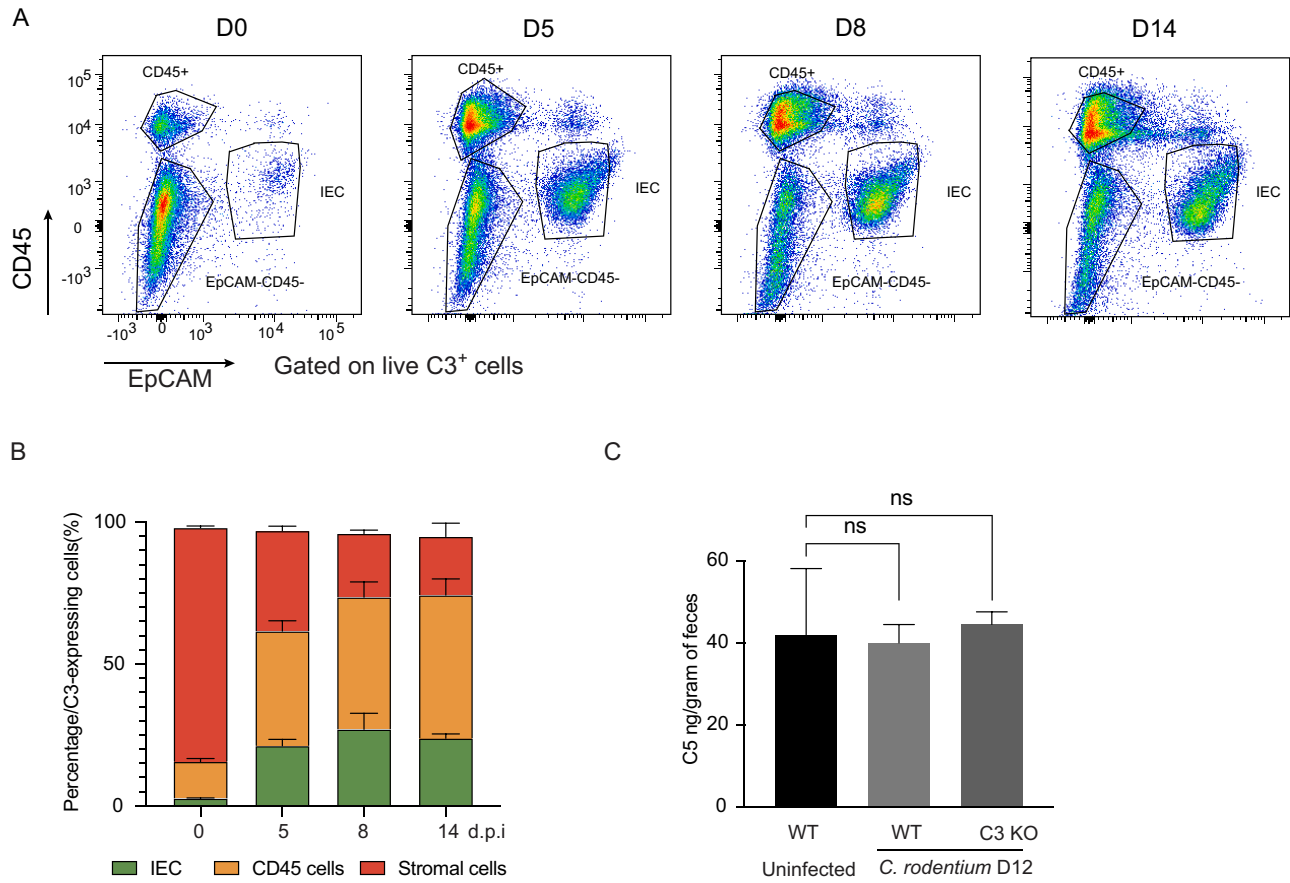
---

(D) Transcriptional expression levels of C3 in colonic stromal cells stimulated by formalin-fixed *E. faecalis* TX104, heat-killed *E. faecalis* TX104, peptidoglycan (100  $\mu\text{g}/\text{mL}$ ), lipoteichoic acid (100  $\mu\text{g}/\text{mL}$ ), Pam3CSK4 (1  $\mu\text{g}/\text{mL}$ ), or medium only as control. ns, not significant; \* $p < 0.05$ ; \*\* $p < 0.01$ , \*\*\*\* $p < 0.0001$ , unpaired Student's t test.

(E) Violin plot showing indicated receptor gene expression levels in C3-expressing stromal, myeloid, and epithelial cells identified in [Figure 3C](#) from scRNA-seq of tdTomato<sup>+</sup> mouse colonic cells using C3<sup>IRES-tdTomato</sup> reporter mice.



**Figure S5. C3 levels and activity in C1qa deficient, MBL-deficient mice, and WT mice treated with cobra venom factor, related to Figure 5**  
 (A) Fecal C3 levels at day 7 post *C. rodentium* infection in C1qa-deficient or MBL-deficient mice with WT mice as control.  
 (B and C) C3 opsonization of GFP+ *C. rodentium* at 7 dpi in WT, MBL-deficient, or C1qa-deficient mice by flow cytometric analysis using anti-C3 antibody or isotype control (B) and the frequency of C3+ *C. rodentium* (C).  
 (D and E) Serum C3 levels (D) and fecal C3 levels (E) in WT mice treated with cobra venom factor at day 0, 1, and 4 after treatment.  
 ns: not significant; \*\*\* p<0.001, unpaired student's t-test. Error bars represent mean  $\pm$  SEM.



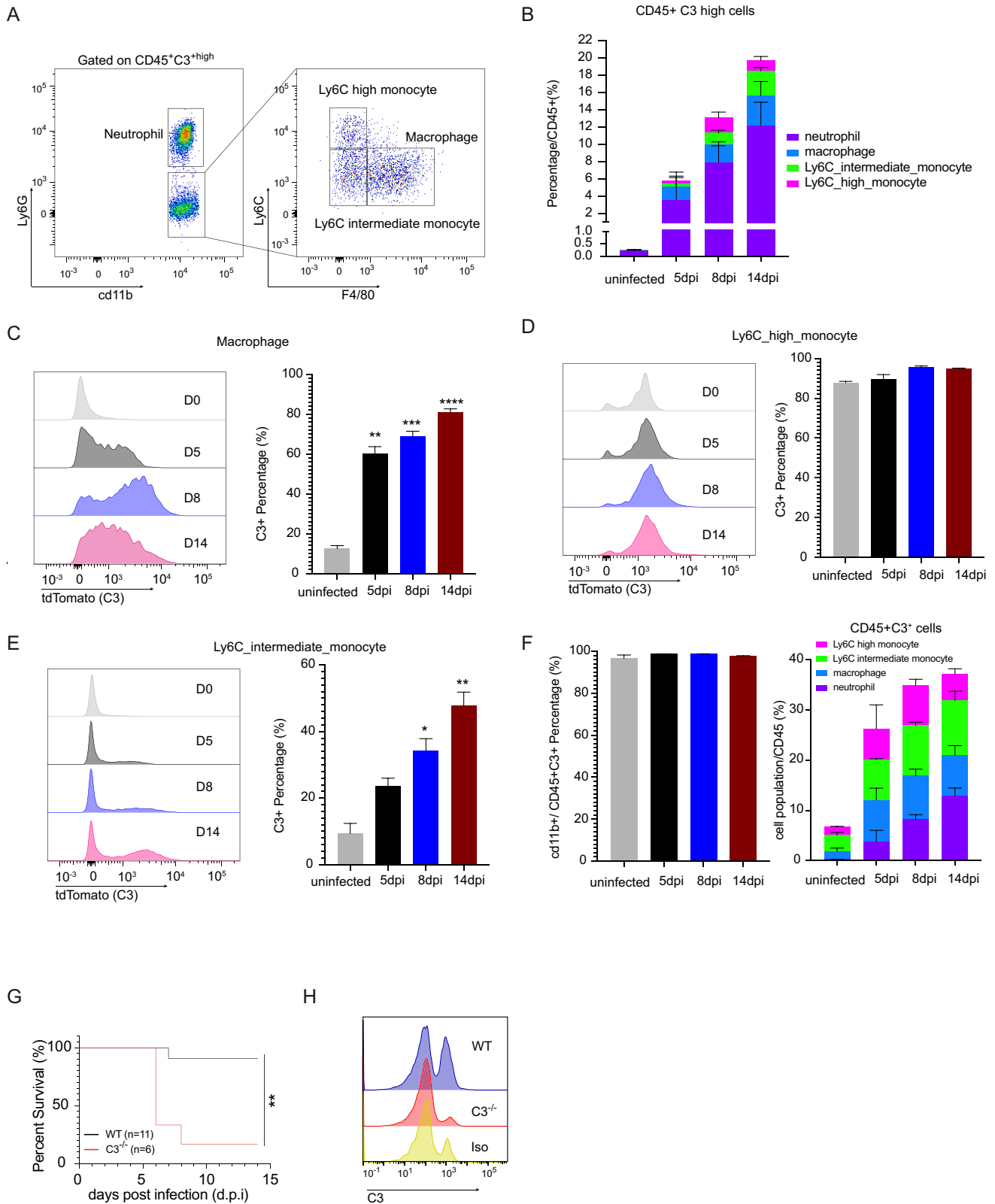
**Figure S6. Composition of total colonic C3-expressing cells and fecal C5 levels during *C. rodentium* infection, related to Figure 6**

(A) Flow cytometry gating of live C3<sup>+</sup> colonic cells in subepithelial compartment of C3<sup>IRES-tdTomato</sup> reporter mice showing the frequency of CD45<sup>+</sup>, IEC, or CD45<sup>-</sup>EpCAM<sup>-</sup> cells.

(B) Percentages of IEC, CD45<sup>+</sup>, or stromal cells in live C3<sup>+</sup> colonic cells in the subepithelial compartment of C3<sup>IRES-tdTomato</sup> reporter mice during *C. rodentium* infection.

(C) Fecal C5 levels in uninfected WT mice or infected WT and C3-deficient mice at day 12 post *C. rodentium* infection.

ns: not significant, unpaired student's t-test. Error bars represent mean ± SEM.



(legend continued on next page)

---

**Figure S7. Luminal C3 is critical for the protection against EHEC and immune profiling of colonic CD45<sup>+</sup>C3<sup>+</sup> cells during *C. rodentium* infection, related to Figure 7**

(A) Flow cytometry gating of high-C3-expressing CD45<sup>+</sup> colonic cells in the subepithelial compartment of C3<sup>IRES-tdTomato</sup> reporter mice at day 14 post *C. rodentium* infection.

(B) Percentage of high-C3-expressing neutrophils, macrophages, Ly6C intermediate-, and Ly6C high monocytes in CD45<sup>+</sup> cells at days 0, 4, 5 and 14 post *Citrobacter rodentium* infection.

(C) Representative histogram of C3 gene expression (MFI in C3-tdTomato reporter mice) in macrophages at 0, 5, 8, and 14 dpi in C3<sup>IRES-tdTomato</sup> reporter mice, and percentages of C3-expressing macrophages in the total macrophage population from the colonic lamina propria at indicated days post-infection.

(D) Representative histogram of C3 gene expression (MFI in C3-tdTomato reporter mice) in Ly6C-high monocytes at 0, 5, 8, and 14 dpi in C3<sup>IRES-tdTomato</sup> reporter mice, and percentages of the C3-expressing Ly6C-high monocytes in the total Ly6C-high monocytes from the colonic lamina propria at indicated days post-infection.

(E) Representative histogram of C3 gene expression (MFI in C3-tdTomato reporter mice) of Ly6C-intermediate monocytes at 0, 5, 8, and 14 dpi in C3<sup>IRES-tdTomato</sup> reporter mice, and percentages of the C3-expressing Ly6C-intermediate monocytes in the total Ly6C-intermediate monocytes from colonic lamina propria at indicated days post-infection.

(F) Percentages of CD11b<sup>+</sup> cells among the total colonic lamina propria CD45<sup>+</sup>C3<sup>+</sup> population at indicated days post-infection, and percentages of C3-expressing neutrophils, macrophages, Ly6C- intermediate and Ly6C-high monocytes in the CD45<sup>+</sup> population at days 0, 5, 8, and 14 post *Citrobacter rodentium* infection.

(G) Survival curve of C57BL/6 WT and C3-deficient mice (~6–7 weeks old) following oral gavage with enterohemorrhagic *E. coli* (EHEC) (10<sup>9</sup> CFUs/mice).

(H) Histogram of flow cytometric analysis of C3 opsonization of GFP<sup>+</sup> EHEC at 4 dpi in WT or C3-deficient mice.

\*p<0.05; \*\*p < 0.01; \*\*\*p < 0.001; \*\*\*\* p < 0.0001, unpaired student's t-test. Error bars represent mean ± SEM.