nature medicine

LETTERS

https://doi.org/10.1038/s41591-018-0324-z

Healthy infants harbor intestinal bacteria that protect against food allergy

Taylor Feehley[1,9], Catherine H. Plunkett[1,9], Riyue Bao[2,3,9], Sung Min Choi Hong[1], Elliot Culleen[1], Pedro Belda-Ferre[1], Evelyn Campbell[1], Rosita Aitoro[4], Rita Nocerino[4], Lorella Paparo[4], Jorge Andrade[2,3], Dionysios A. Antonopoulos[5,6], Roberto Berni Canani[4,7,8] and Cathryn R. Nagler[1*]

There has been a striking generational increase in life-threatening food allergies in Westernized societies[1,2]. One hypothesis to explain this rising prevalence is that twenty-first century lifestyle practices, including misuse of antibiotics, dietary changes, and higher rates of Caesarean birth and formula feeding have altered intestinal bacterial communities; early-life alterations may be particularly detrimental[3,4]. To better understand how commensal bacteria regulate food allergy in humans, we colonized germ-free mice with feces from healthy or cow's milk allergic (CMA) infants[5]. We found that germ-free mice colonized with bacteria from healthy, but not CMA, infants were protected against anaphylactic responses to a cow's milk allergen. Differences in bacterial composition separated the healthy and CMA populations in both the human donors and the colonized mice. Healthy and CMA colonized mice also exhibited unique transciptome signatures in the ileal epithelium. Correlation of ileal bacteria with genes upregulated in the ileum of healthy or CMA colonized mice identified a clostridial species, *Anaerostipes caccae*, that protected against an allergic response to food. Our findings demonstrate that intestinal bacteria are critical for regulating allergic responses to dietary antigens and suggest that interventions that modulate bacterial communities may be therapeutically relevant for food allergy.

Work from our laboratory and others has demonstrated that the fecal microbial communities of infants with CMA are markedly different from those of their healthy counterparts[5,6]. Based on these results, as well as evidence that members of the microbiota can be allergy protective[7], we used a gnotobiotic mouse model to investigate whether commensal bacteria have a causal role in protection against an allergic response to the cow's milk allergen β-lactoglobulin (BLG). Germ-free mice were colonized with human feces from four healthy and four immunoglobulin E (IgE)-mediated CMA infant donors who were matched for age, gender, and mode of birth[8,9] (Supplementary Table 1). It has previously been reported that diet is important for the stable colonization of germ-free mice with human feces[10]. To support the growth of human bacteria in the murine hosts, mice received feces from formula-fed healthy or CMA infants and were fed the same formulas consumed by their human infant donors in addition to plant-based mouse chow. The CMA infant donors received an extensively hydrolyzed casein formula to manage ongoing allergic symptoms, whereas the healthy donors received a standard cow's milk-based formula[5]. Initial transfer recipients were used as living repositories for subsequent experiments (see Online Methods).

Groups of germ-free mice and mice colonized with either the healthy or CMA infant microbiota were sensitized with BLG and the mucosal adjuvant cholera toxin. Consistent with previous reports[7,11], germ-free mice, devoid of any bacterial colonization, were highly susceptible to anaphylactic responses to food, as evidenced by a drop in core body temperature (Fig. 1a) and production of BLG-specific IgE and IgG1 (Fig. 1b,c). We also measured a substantial reduction in core body temperature in mice colonized with fecal samples from each of the four CMA donors in response to BLG challenge (Fig. 1a). Sensitized CMA-colonized mice produced significantly higher serum concentrations of BLG-specific IgE (Fig. 1b), IgG1 (Fig. 1c) and mouse mast cell protease-1 (mMCPT-1) (Fig. 1d) compared with healthy-colonized mice. Notably, all of the mice that received the four healthy infant microbiotas were protected from an anaphylactic response to BLG challenge; their core body temperature post-challenge was significantly different from that measured in germ-free or CMA-colonized mice (Fig. 1a). Histological analysis did not reveal any evidence of pathology or inflammation in ileal or colonic tissue samples taken post-challenge (Extended Data Fig. 1) or after long-term colonization (Extended Data Fig. 2). Microbial analysis revealed that community diversity and evenness were similar between healthy- and CMA-colonized mouse groups (Extended Data Fig. 3). To examine whether the cow's-milk-containing formula contributed to microbiota-independent protection against anaphylaxis in the healthy-colonized mice, we performed additional fecal transfers from breast-fed healthy and CMA donors (Supplementary Table 2). Recipient mice received only plant-based mouse chow. Mice colonized with feces from a breast-fed healthy donor were protected from an anaphylactic response to BLG sensitization and challenge. However, mice colonized with feces from a breast-fed CMA donor exhibited a significantly greater drop in core body temperature compared with healthy-colonized mice (Extended Data Fig. 4a) and higher levels of BLG-specific IgE (Extended Data Fig. 4b). We also compared sensitization to BLG in germ-free mice fed water or Enfamil. Both groups of mice responded robustly to sensitization with BLG (Extended Data Fig. 5). There was no significant difference in their drop in core body temperature post-challenge

[1]Department of Pathology and Committee on Immunology, The University of Chicago, Chicago, IL, USA. [2]Center for Research Informatics, The University of Chicago, Chicago, IL, USA. [3]Department of Pediatrics, The University of Chicago, Chicago, IL, USA. [4]Department of Translational Medical Science, Section of Pediatrics, University of Naples Federico II, Naples, Italy. [5]Department of Medicine, The University of Chicago, Chicago, IL, USA. [6]Biosciences Division, Argonne National Laboratory, Argonne, IL, USA. [7]European Laboratory for the Investigation of Food-Induced Diseases, University of Naples Federico II, Naples, Italy. [8]CEINGE Advanced Biotechnologies, University of Naples Federico II, Naples, Italy. [9]These authors contributed equally: Taylor Feehley, Catherine H. Plunkett, Riyue Bao. *e-mail: cnagler@bsd.uchicago.edu

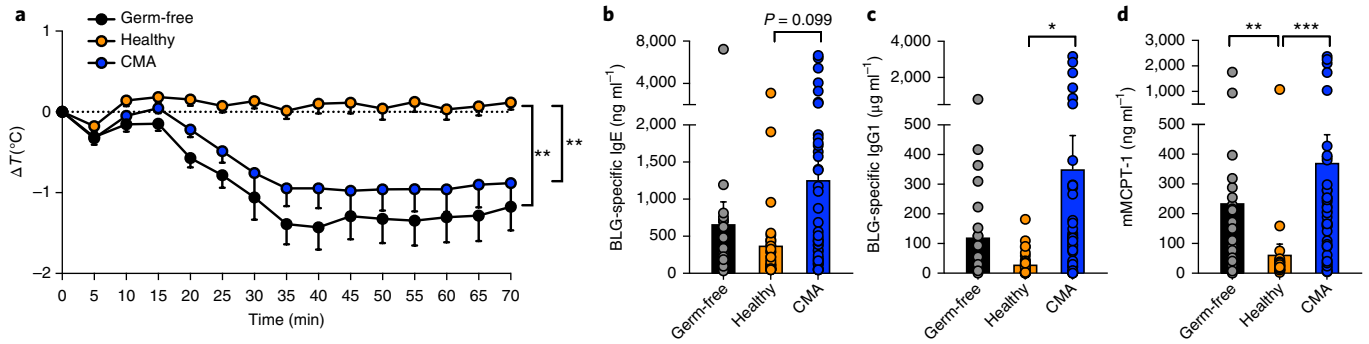


Fig. 1 | Transfer of healthy, but not CMA, infants' microbiota protects against an allergic response to food. a, Change in core body temperature at indicated time points following first challenge with BLG in germ-free mice and in mice colonized with feces from each of eight donors (four healthy, four CMA; see Supplementary Table 1) that had been sensitized with BLG plus cholera toxin; $n = 42$ CMA, 31 healthy and 24 germ-free mice, with 4–12 mice for each of the eight donors, collected from two independent experiments. **b**–**d**, Serum BLG-specific IgE (**b**), BLG-specific IgG1 (**c**), and mMCPT-1 (**d**) from mice in **a**. For **a**, circles represent mean, and error bars represent s.e.m. For **b**–**d**, circles represent individual mice, and bars represent mean + s.e.m. Linear mixed-effect models were used to compare groups in **a**–**d** with the BH-FDR method for multiple testing correction. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

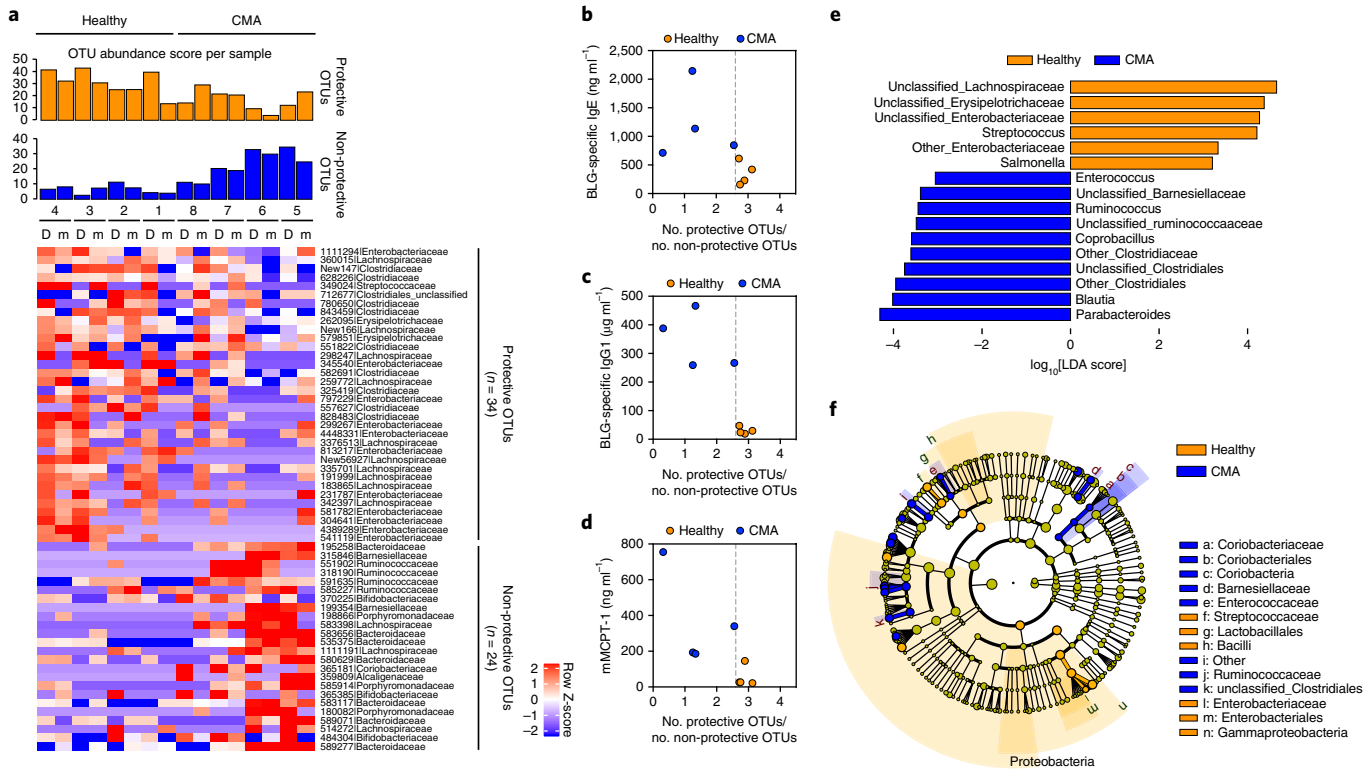


Fig. 2 | Analysis of fecal samples from eight human infant donors reveals taxonomic signatures that correlate with allergic phenotype. a, Heat map of OTUs differentially abundant between CMA and healthy donors. Rows show 58 OTUs identified as different at FDR controlled at 0.10 and present in at least four human fecal samples and at least two groups of colonized mice (see Supplementary Table 3). Columns depict each donor (D) or colonized mouse group (m). $n = 2$–3 technical replicates per donor and $n = 1$–4 mice per colonized mouse group, with feces taken at 2 and 3 weeks post-colonization (see Online Methods). The bar graphs above the heat map represent the abundance score of potentially protective (orange) or non-protective (blue) OTUs calculated for each donor or mouse group. **b**–**d**, The ratio of protective over non-protective OTUs (see Extended Data Fig. 6b) derived from colonized mice in **a** plotted against levels of BLG-specific IgE (**b**), BLG-specific IgG1 (**c**) and mMCPT-1 (**d**) from all mice in Fig. 1. Each circle represents average results from all mice colonized with each of the four healthy (orange) or CMA (blue) donor's feces. **e**, LEfSe analysis of taxa that were significantly enriched in healthy-colonized mice (orange) or CMA-colonized mice (blue) from samples in **a** ($n = 8$ mice in healthy group and $n = 9$ mice in CMA group, with fecal samples collected at 2 and 3 weeks post-colonization). **f**, Cladogram showing the community composition of colonized mouse samples from **a**, with the taxa detected as differentially abundant by LEfSe analysis colored by group (orange, healthy; blue, CMA). The discrete false discovery rate (DS-FDR) method was used to compare groups in **a** and the Kruskal–Wallis test in **e** (see Online Methods).

or in serum concentrations of BLG-specific IgE or IgG1; serum mMCPT-1 was, however, suppressed in mice fed the formula containing cow's milk.

Analysis of fecal samples from the eight formula-fed human infant donors (Supplementary Table 1) identified 58 operational taxonomic units (OTUs) that were differentially abundant between

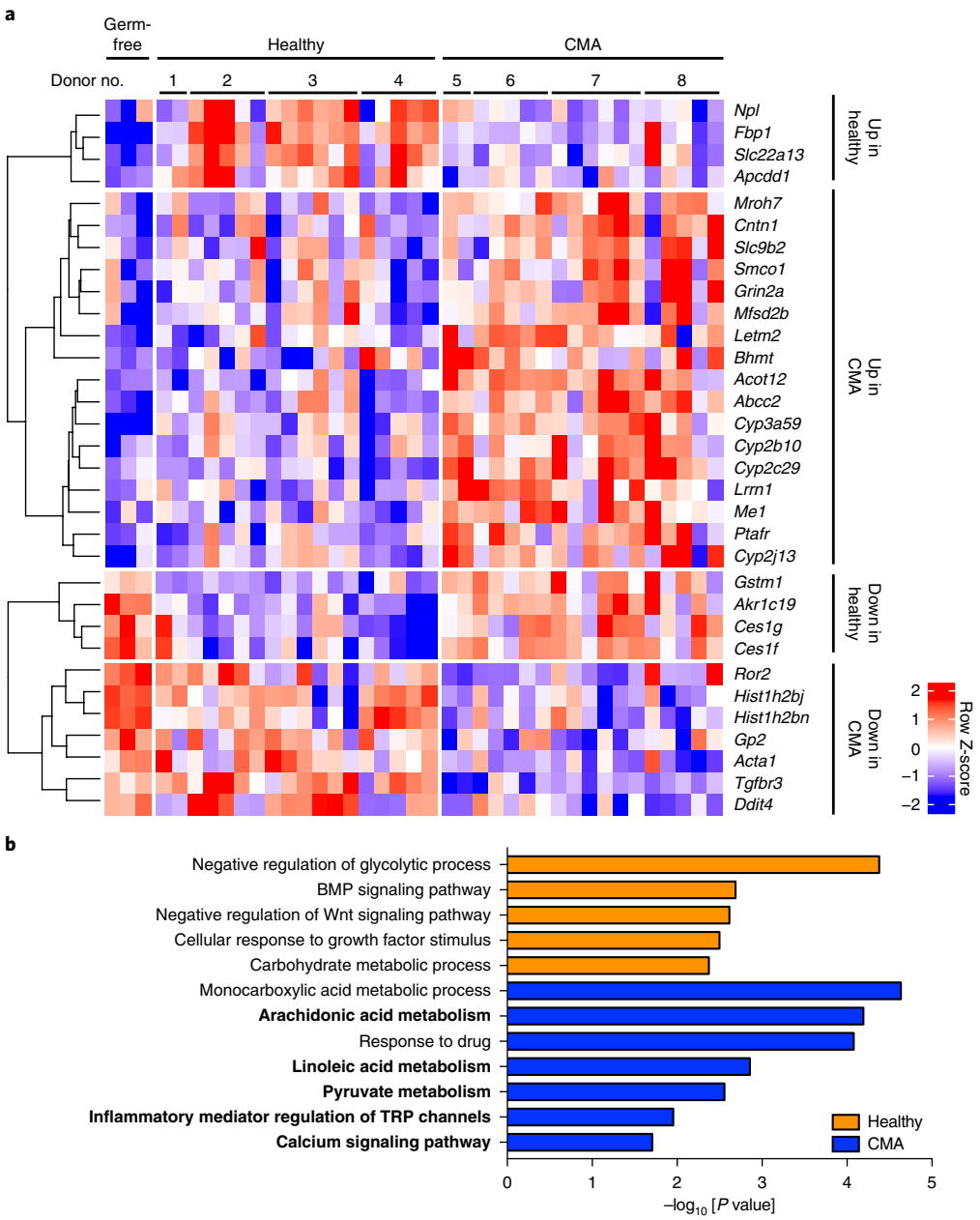


Fig. 3 | Unique ileal transcriptome signatures distinguish healthy- and CMA-colonized mice. a, Heat map of 32 DEGs in ileal IECs isolated from germ-free ($n=3$), healthy-colonized ($n=18$) or CMA-colonized ($n=18$) mice collected from at least two independent experiments at 7 d post-colonization (see Supplementary Table 4). Each column depicts an individual mouse colonized with donor feces as indicated. Four types of gene expression changes are shown: (1) up in healthy: genes that are upregulated in healthy mice relative to both CMA and germ-free; (2) up in CMA: genes that are upregulated in CMA mice relative to both healthy and germ-free; (3) down in healthy: genes that are downregulated in healthy mice relative to both CMA and germ-free; and (4) down in CMA: genes that are downregulated in CMA mice relative to healthy and germ-free. **b**, GO terms (light) and KEGG pathways (bold) significantly enriched in DEGs from **a** that are associated with healthy-colonized (orange) or CMA-colonized (blue) mice. Hypergeometric testing was used in **b** with the BH-FDR method for multiple testing correction (see Online Methods).

healthy and CMA infants (Fig. 2a and Supplementary Table 3). Given that variation exists between each donor and murine transfer recipient at the single-OTU level, we examined whether donor-derived microbiome composition differences were able to distinguish the colonized mouse groups. As an aggregated measure to present the data, we calculated the number of potentially 'protective' (more abundant in healthy donors; $n=34$) and potentially 'non-protective' (more abundant in CMA donors; $n=24$) OTUs to produce a presence/absence ratio for each donor (Extended Data Fig. 6a; see Online Methods). In addition, we calculated a score weighted towards each OTU based on its relative abundance in the sample (hereafter called the abundance score) (Fig. 2a; see Online Methods). When the OTU abundance score was plotted against the presence/absence ratio, donors segregated by ratio into the healthy and CMA groups (Extended Data Fig. 6b). This threshold also separated the CMA- and healthy-colonized mice by their biological phenotype (Extended Data Fig. 6b), demonstrating that this donor-derived aggregated microbiota signature is validated in the murine transfer recipients. The significantly higher protective/non-protective OTU ratio in healthy infants relative to those with CMA was

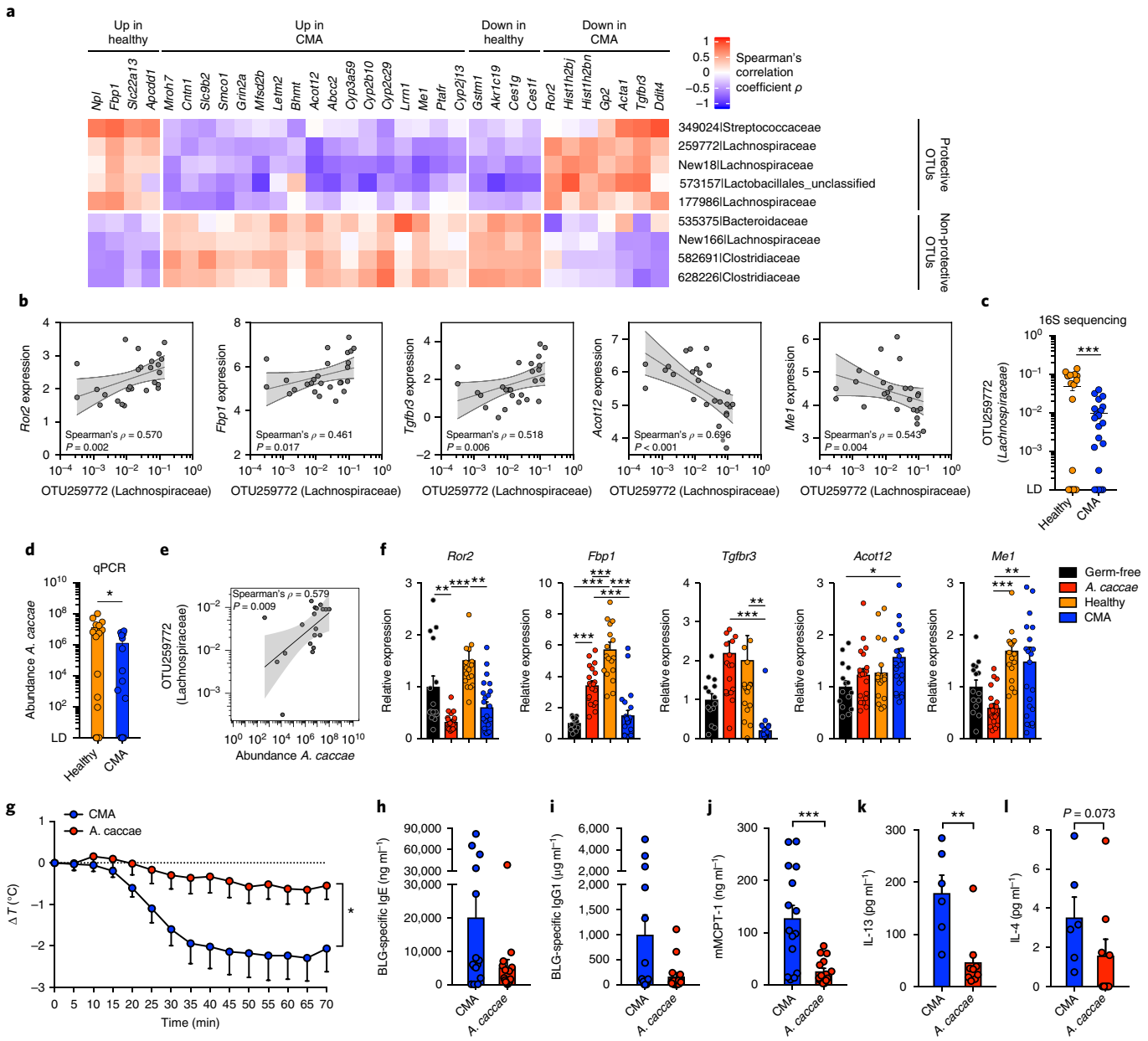


Fig. 4 | Correlation of ileal OTUs with DEGs in the ileum of healthy-colonized mice identifies a clostridial species, *A. caccae*, that protects against an allergic response to food. a, Heat map showing Spearman's rank correlation coefficient between relative abundance of ileal OTUs (row) and expression of DEGs (column) from CMA versus healthy mouse ileal IEC samples (see Fig. 3a and Online Methods). **b**, Spearman's correlation between abundance of OTU259772 (Lachnospiraceae) from the ileal 16S dataset (see Supplementary Table 5) and RNA-seq gene expression in ileal IECs of *Ror2*, *Fbp1*, *Tgfbr3*, *Acot12*, and *Me1*. Circles represent individual mice, and shaded bands indicate 95% confidence interval fitted by linear regression. **c**,**d**, Abundance of OTU259772 (Lachnospiraceae) by 16S sequencing (**c**) and abundance of *A. caccae* by qPCR (**d**) in ileal samples from healthy- and CMA-colonized mice. LD indicates samples that were below the limit of detection for the assay. **e**, Spearman's correlation between abundance of OTU259772 (Lachnospiraceae; 16S sequencing) and abundance of *A. caccae* (qPCR) in ileal samples from healthy- and CMA-colonized mice. Circles represent individual mice, and shaded bands indicate 95% confidence interval fitted by linear regression. Ileal samples that were above LD in both 16S and qPCR experiments are shown ($n = 19$). **f**, Gene expression of *Ror2*, *Fbp1*, *Tgfbr3*, *Acot12*, and *Me1* in ileal IECs isolated from germ-free mice and from healthy- and CMA-colonized mice or mice monocolonized with *A. caccae* by qPCR. Data is normalized to *Hprt* as the housekeeping gene and shown as the fold change in expression from germ free, set as 1. **g**, Change in core body temperature at indicated time points following first challenge with BLG in BLG plus cholera toxin-sensitized CMA and *A. caccae*-monocolonized mice. **h**–**j**, Serum BLG-specific IgE (**h**), BLG-specific IgG1 (**i**) and mMCPT-1 (**j**) from mice in **g**. **k**,**l**, IL-13 (**k**) and IL-4 (**l**) in culture supernatants of splenocytes from CMA or *A. caccae*-colonized mice killed 24 h post-challenge and stimulated for 72 h with BLG. For **c**, **d**, **f**, and **h**–**l**, circles represent individual mice, and bars represent mean + s.e.m. For **g**, circles represent mean, and bars represent s.e.m. For **a** and **b**, $n = 18$ healthy-colonized or 18 CMA-colonized mice per group. For **c** and **d**, $n = 19$ healthy-colonized or 21 CMA-colonized mice per group. For **f**, $n = 14$ germ-free, 20 *A. caccae*-colonized, 18 healthy-colonized or 23 CMA-colonized mice per group. For **g**–**j**, $n = 16$ CMA-colonized and 16 *A. caccae*-colonized mice collected from three independent experiments with two different CMA donors (5 and 6), and bars represent mean + s.e.m. For **k** and **l**, $n = 6$ CMA-colonized and 9 *A. caccae*-colonized mice from 1 experiment, circles represent individual mice, and bars represent mean + s.e.m. The DS-FDR method was used to compare groups in **c**, two-sided Student's *t* test in **d**, one-way analysis of variance with Bonferroni multiple testing correction in **f** or linear mixed-effect models in **g**, and two-sided Student's *t* test in **h**–**l** after log transformation. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

independently corroborated in an unrelated set of samples from the same Neapolitan cohort by re-analysis of 16S fecal sample data collected in a previously published study[5] (Extended Data Fig. 7). The donor-derived OTU ratio also separated healthy- and CMA-colonized mice when plotted against biomarkers of allergic disease, including BLG-specific IgE (Fig. 2b), BLG-specific IgG1 (Fig. 2c) and mMCPT-1 (Fig. 2d). Notably, linear discriminant effect size (LEfSe) analysis (Fig. 2e,f) showed that members of Lachnospiraceae, a family in the Clostridia class that was previously implicated in protection against allergic sensitization to food, were enriched in the healthy colonized mice[7].

Tolerance to dietary antigens begins with their absorption in the small intestine[4,12]. Most commensal bacteria reside in the colon; in the small intestine, bacteria are most numerous in the ileum[13]. The interaction of these bacteria with intestinal epithelial cells (IECs) is central to regulation of immunity at the host-microbe interface[13,14]. We therefore isolated ileal IECs from groups of mice colonized by each of the eight infant donors and quantified gene expression by RNA sequencing (RNA-seq; Fig. 3a). We found that healthy-colonized mice upregulated a unique set of ileal genes compared with CMA-colonized mice (Fig. 3a and Supplementary Table 4). For example, *Fbp1*, which encodes a key gluconeogenic enzyme abundantly expressed in epithelial cells of the small intestine[15], was significantly upregulated across all healthy-colonized mice (Fig. 3a). Reduced expression of *Fbp1* has been associated with a metabolic switch from oxidative phosphorylation to aerobic glycolysis[16,17], which alters oxygenation of the epithelium and contributes to dysbiosis[18]. *Tgfbr3* and *Ror2* were downregulated in the ileum of CMA-colonized mice relative to healthy-colonized mice (Fig. 3a). *Tgfbr3* encodes a receptor for the growth factor TGF-β and is abundantly expressed in the small intestine of suckling rats[19]. Soluble TGFβ RIII and TGF-β2 are present at high concentrations in breast milk; activation of TGF-β signaling by Wnt5a is mediated through *Ror2* and is important for epithelial repair[20]. By contrast, *Acot12* and *Me1*, genes involved in pyruvate metabolism, were upregulated in the ileum of CMA-colonized mice relative to healthy-colonized mice. These metabolic and molecular processes are reflected in the Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways that were significantly altered in CMA- and healthy-colonized mice (Fig. 3b).

To determine whether the fecal OTU signatures identified in Fig. 2 are also reflective of ileal bacterial populations, we examined the correlation between ileal OTUs and the fecal signature in healthy- and CMA-colonized mice (Extended Data Fig. 8a). We found that the majority of the taxa change in the same direction (increase or decrease in abundance) in healthy relative to CMA between mouse fecal and ileal samples (Extended Data Fig. 8b,c). The identification of differential gene expression in ileal IECs from healthy- and CMA-colonized mice (Fig. 3a) suggested that ileal bacteria regulate host immunity to contribute to allergic sensitization. Integrative analysis of ileal bacteria and ileal differentially expressed genes (DEGs) revealed nine OTUs significantly and consistently correlated with genes upregulated in the ileum of healthy- or CMA-colonized mice (Fig. 4a). Notably, three out of five of the protective OTUs associated with DEGs upregulated in the ileum of healthy-colonized mice are members of the family Lachnospiraceae; 70% of the OTUs in our previously identified allergy protective murine Clostridia consortium belong to this family[7]. A BLAST search of assembled 16S sequences against the National Center for Biotechnology Information database (16S ribosomal RNA, Bacteria and Archaea) revealed that all three protective Lachnospiraceae OTUs upregulated in the healthy-colonized mice (259772, New18, and 177986) have *Anaerostipes caccae* as the closest matching species. In particular, OTU259772 was annotated with *A. caccae* in a previous study of human infant feces and diet[21]. *A. caccae* is non-spore-forming, utilizes lactate and acetate and produces butyrate[22,23]. Spearman's correlations between Lachnospiraceae OTU259772 and several highly correlated ileal DEGs of interest (*Ror2*, *Fbp1*, *Tgfbr3*, *Acot12*, and *Me1*) from Fig. 3a are depicted in Fig. 4b. Analysis of ileal and fecal samples using quantitative PCR (qPCR) with previously validated species-specific primers[24] provided independent confirmation of the enrichment of *A. caccae* in healthy-colonized mice (Fig. 4c–e and Extended Data Fig. 9a–c). Abundance of *A. caccae* in ileal samples also correlated with DEGs from ileal IECs (Extended Data Fig. 10). Of note, two of the highly correlated DEGs (*Acot12* and *Me1*) are involved in pyruvate metabolism. Butyrate is an important energy source for colonic epithelial cells[25]. Butyrate drives oxygen consumption by colonocytes through β-oxidation, thereby maintaining a locally hypoxic niche for butyrate-producing obligate anaerobes[26]. Under conditions of dysbiosis, colonocytes generate energy via glycolysis, a process that includes production of pyruvate as a key intermediate[27]. It is tempting to speculate that the negative correlation between the abundance of butyrate-producing *A. caccae* and pyruvate metabolism-related genes in IECs from CMA-colonized mice is reflective of metabolic shifts in ileal epithelial function under conditions of dysbiosis.

We next examined whether *A. caccae* can mimic the changes in gene expression and protection against anaphylaxis associated with the healthy microbiota by monocolonizing germ-free mice (see Online Methods). Some of the genes significantly upregulated in healthy-colonized mice (*Fbp1*, *Tgfbr3*) were also significantly upregulated in *A. caccae*-monocolonized mice (Fig. 4f) compared with germ-free or CMA-colonized mice. *Acot12* expression was significantly upregulated in CMA-colonized mice, but not in healthy-colonized or *A. caccae*-monocolonized mice (Fig. 4f). BLG plus cholera toxin-sensitized *A. caccae*-monocolonized mice were protected against an anaphylactic response to BLG challenge. As in Fig. 1, CMA-colonized mice exhibited a marked drop in core body temperature indicative of anaphylaxis (Fig. 4g). Both the changes in core body temperature and serum concentrations of mMCPT-1 were significantly reduced in *A. caccae*-monocolonized mice compared with CMA-colonized mice (Fig. 4g,j). Antigen-specific, Th2-dependent antibody (serum BLG-specific IgE and IgG1; Fig. 4h,i) and cytokine (IL-13 and IL-4; Fig. 4k,l) responses were all reduced in *A. caccae*-monocolonized mice.

Anaerobic, mucosa-associated bacteria in the Clostridia class have attracted considerable interest because of their reported roles in the maintenance of intestinal homeostasis through induction of regulatory T cells[28,29], production of immunomodulatory metabolites[30,31] and regulation of colonization resistance[32]. Previous studies have focused mostly on the colon. We now also place these immunomodulatory bacteria in the ileum, at the site of food absorption, and our findings demonstrate their causal role in protection against an anaphylactic response to food. Mechanistic analysis of the Clostridia-associated changes in ileal gene expression described herein is likely to reveal additional pathways critical to the maintenance of tolerance to dietary antigens. The model described in this report does not address whether the allergic state drives dysbiosis[33] or whether dysbiosis precedes allergy. Indeed, many factors are likely to contribute to the development of food allergies. Our data demonstrate that the commensal bacteria have an important role in preventing allergic responses to food and provide proof of concept for the development of microbiome-modulating strategies to prevent or treat this disease.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41591-018-0324-z.

Received: 27 March 2018; Accepted: 5 December 2018;
Published online: 14 January 2019

References

1. Sicherer, S. et al. Critical issues in food allergy: a National Academies consensus report. *Pediatrics* **140**, e20170194 (2017).
2. Iweala, O. I. & Burks, A. W. Food allergy: our evolving understanding of its pathogenesis, prevention, and treatment. *Curr. Allergy Asthma Rep.* **16**, 37 (2016).
3. Wesemann, D. R. & Nagler, C. R. The microbiome, timing, and barrier function in the context of allergic disease. *Immunity* **44**, 728–738 (2016).
4. Plunkett, C. H. & Nagler, C. R. The influence of the microbiome on allergic sensitization to food. *J. Immunol.* **198**, 581–589 (2017).
5. Berni Canani, R. et al. *Lactobacillus rhamnosus* GG-supplemented formula expands butyrate-producing bacterial strains in food allergic infants. *ISME J.* **10**, 742–750 (2016).
6. Bunyavanich, S. et al. Early-life gut microbiome composition and milk allergy resolution. *J. Allergy Clin. Immunol.* **138**, 1122–1130 (2016).
7. Stefka, A. T. et al. Commensal bacteria protect against food allergen sensitization. *Proc. Natl Acad. Sci. USA* **111**, 13145–13150 (2014).
8. Dominguez-Bello, M. G. et al. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc. Natl Acad. Sci. USA* **107**, 11971–11975 (2010).
9. Mueller, N. T., Bakacs, E., Combellick, J., Grigoryan, Z. & Dominguez-Bello, M. G. The infant microbiome development: mom matters. *Trends Mol. Med.* **21**, 109–117 (2015).
10. Blanton, L. V. et al. Gut bacteria that prevent growth impairments transmitted by microbiota from malnourished children. *Science* **351**, aad3311 (2016).
11. Cahenzli, J., Koller, Y., Wyss, M., Geuking, M. B. & McCoy, K. D. Intestinal microbial diversity during early-life colonization shapes long-term IgE levels. *Cell Host Microbe* **14**, 559–570 (2013).
12. Pabst, O. & Mowat, A. M. Oral tolerance to food protein. *Mucosal Immunol.* **5**, 232–239 (2012).
13. Honda, K. & Littman, D. R. The microbiota in adaptive immune homeostasis and disease. *Nature* **535**, 75–84 (2016).
14. Thaiss, C. A., Zmora, N., Levy, M. & Elinav, E. The microbiome and innate immunity. *Nature* **535**, 65–74 (2016).
15. Yanez, A. J. et al. Broad expression of fructose-1,6-bisphosphatase and phosphoenolpyruvate carboxykinase provide evidence for gluconeogenesis in human tissues other than liver and kidney. *J. Cell. Physiol.* **197**, 189–197 (2003).
16. Ostroukhova, M. et al. The role of low-level lactate production in airway inflammation in asthma. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **302**, L300–L307 (2012).
17. Zhu, Y. et al. NPM1 activates metabolic changes by inhibiting FBP1 while promoting the tumorigenicity of pancreatic cancer cells. *Oncotarget* **6**, 21443–21451 (2015).
18. Berger, C. N. et al. *Citrobacter rodentium* subverts ATP flux and cholesterol homeostasis in intestinal epithelial cells *in vivo*. *Cell Metab.* **26**, 738–752 (2017).
19. Zhang, M., Zola, H., Read, L. & Penttila, I. Identification of soluble transforming growth factor-β receptor III (sTβIII) in rat milk. *Immunol. Cell Biol.* **79**, 291–297 (2001).
20. Miyoshi, H., Ajima, R., Luo, C. T., Yamaguchi, T. P. & Stappenbeck, T. S. Wnt5a potentiates TGF-β signaling to promote colonic crypt regeneration after tissue injury. *Science* **338**, 108–113 (2012).
21. Planer, J. D. et al. Development of the gut microbiota and mucosal IgA responses in twins and gnotobiotic mice. *Nature* **534**, 263–266 (2016).
22. Schwiertz, A. et al. *Anaerostipes caccae* gen. nov., sp. nov., a new saccharolytic, acetate-utilising, butyrate-producing bacterium from human faeces. *Syst. Appl. Microbiol.* **25**, 46–51 (2002).
23. Duncan, S. H., Louis, P. & Flint, H. J. Lactate-utilizing bacteria, isolated from human feces, that produce butyrate as a major fermentation product. *Appl. Environ. Microbiol.* **70**, 5810–5817 (2004).
24. Kurakawa, T. et al. Diversity of intestinal *Clostridium coccoides* group in the Japanese population, as demonstrated by reverse transcription-quantitative PCR. *PLoS ONE* **10**, e0126226 (2015).
25. Donohoe, D. R. et al. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab.* **13**, 517–526 (2011).
26. Byndloss, M. X. et al. Microbiota-activated PPAR-γ signaling inhibits dysbiotic Enterobacteriaceae expansion. *Science* **357**, 570–575 (2017).
27. Donohoe, D. R., Wali, A., Brylawski, B. P. & Bultman, S. J. Microbial regulation of glucose metabolism and cell-cycle progression in mammalian colonocytes. *PLoS ONE* **7**, e46589 (2012).
28. Atarashi, K. et al. Induction of colonic regulatory T cells by indigenous *Clostridium* species. *Science* **331**, 337–341 (2011).
29. Atarashi, K. et al. T_{reg} induction by a rationally selected mixture of Clostridia strains from the human microbiota. *Nature* **500**, 232–236 (2013).
30. Furusawa, Y. et al. Commensal microbe-derived butyrate induces differentiation of colonic regulatory T cells. *Nature* **504**, 446–450 (2013).
31. Yano, J. M. et al. Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis. *Cell* **161**, 264–276 (2015).
32. Kim, Y. G. et al. Neonatal acquisition of Clostridia species protects against colonization by bacterial pathogens. *Science* **356**, 315–319 (2017).
33. Noval Rivas, M. et al. A microbiota signature associated with experimental food allergy promotes allergic sensitization and anaphylaxis. *J. Allergy Clin. Immunol.* **131**, 201–212 (2013).

Acknowledgements

We thank the children and families for their participation in this study. We are grateful to D. Wesemann, E. Forbes-Blom, G. Nunez, M. Rothenberg, M. Alegre and J. Colson for discussion. We thank S. Wang, M. Bauer and A. Kemter for assistance with some experiments, M. Jarsulic for technical assistance with computing infrastructure, K. Hernandez for discussion of the statistical results and C. Weber for histopathological evaluation of all intestinal sections. Statistical consultation was also provided by M. Giurcanu of the University of Chicago Biostatistics Laboratory. We are grateful to B. Theriault and her staff at the University of Chicago Gnotobiotic Research Animal Facility for superb animal care and experimental support. This work was supported by the Sunshine Charitable Foundation (C.R.N.), a pilot award from the University of Chicago Institute for Translational Medicine (CTSA ULI TR000430, C.R.N.), National Institutes of Health (NIH) grants AI134923 (C.R.N.), DK42086 (D.A.A.) and an Italian Ministry of Health grant PE-2011-02348447 (R.B.C.). The Center for Research Informatics is funded by the Biological Sciences Division at the University of Chicago with additional support provided by the Institute for Translational Medicine/Clinical and Translational Award (NIH 5UL1TR002389-02), and the University of Chicago Comprehensive Cancer Center Support Grant (NIH P30CA014599). Bioinformatics analysis was performed on Gardner High-Performance Computing clusters at the Center for Research Informatics at the University of Chicago. A provisional US patent application (62/755,945) was filed on 5 November 2018.

Author contributions

T.F., C.H.P., R.B., R.B.C., and C.R.N. designed the study. C.H.P. and T.F. performed mouse experiments with help from P.B.-F., R.A., E.Culleen, E.Campbell, and S.M.C.H. R.B., P.B.-F., and J.A. performed bioinformatics analysis. T.F., C.H.P., R.B., P.B.F., and C.R.N. analyzed results. R.B.C., R.N., and L.A. cared for patients and provided donor fecal samples. D.A.A. provided protocols and assisted with the colonization of germ-free mice with human feces and *A. caccae*. T.F., C.H.P., R.B., R.B.C., and C.R.N. wrote the manuscript. All authors read and commented on the manuscript.

Competing interests

C.R.N. is president and co-founder of ClostraBio, Inc. The other authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41591-018-0324-z.

Supplementary information is available for this paper at https://doi.org/10.1038/s41591-018-0324-z.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to C.R.N.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019

Methods

Gnotobiotic mouse husbandry. All mice were bred and housed in the Gnotobiotic Research Animal Facility (GRAF) at the University of Chicago. GRAF is an operational facility of the University of Chicago Animal Resource Center. Mice were maintained in Trexler-style flexible film isolator housing units (Class Biologically Clean) with Ancare polycarbonate mouse cages (catalog number N10HT) and Teklad Pine Shavings (7088; sterilized by autoclave) on a 12 h light/dark cycle at a room temperature of 20–24 °C. Mice were provided with autoclaved sterile water, USP grade, at pH 5.2 *ad libitum*. Bedding was changed weekly; cages of formula-fed mice required near-daily bedding changes due to leakage of formula from the bottles. All mice were fed Purina Lab Diet 5K67, stored in a temperature-controlled environment in accordance with *The Guide for the Care and Use of Laboratory Animals*, 8th edition. The diet was sterilized by autoclaving at 121 °C × 30 min. The sterility of the isolators was checked weekly by both cultivation and 16S ribosomal RNA analysis of fecal samples by qPCR. Cultivation was in BHI, nutrient and Sabouraud broth at 37 °C aerobic and anaerobic and 42 °C aerobic for 96 h. All mice are initially screened upon rederivation or receipt for all internal and external parasites, full serology profile and/or PCR, bacteriology, and gross and histologic analysis of major organs through either IDEXX Radil or Charles River Lab using an Axenic Profile Screen. Germ-free C3H/HeN mice were transferred within the facility from T. Golovkina (University of Chicago).

Preparation of human fecal samples. Healthy (non-allergic) fecal samples were obtained from participants in a vaccination program. These subjects were not at risk for atopic disorders, and their clinical history was negative for any allergic condition. Infants with CMA were diagnosed at a tertiary pediatric allergy center (Pediatric Allergy Program at the Department of Translational Medical Science of the University of Naples Federico II); for complete patient information see Supplementary Tables 1 and 2. All aspects of this study were conducted in accordance with the Declaration of Helsinki and approved by the ethics committee of the University of Naples Federico II. Written informed consent was obtained from the parents/guardians of all children involved in the research. Fresh fecal samples were collected in the clinic in sterile tubes, weighed, mixed with 2 ml LB broth plus 30% glycerol per 100–500 mg, aliquoted into sterile cryovials and immediately stored at −80 °C. Samples were shipped on dry ice to the University of Chicago, where they were stored at −80 °C until homogenization. To colonize mice, frozen fecal samples were introduced into an anaerobic chamber and thawed. Thawed feces were mixed with 3 mm borosilicate glass beads in a sterile 50 ml tube with 2.5 ml pre-reduced PBS plus 0.05% cysteine and vortexed gently to dissociate. The resulting homogenate was filtered through a 100 µm filter. This homogenization and filtration process was repeated three more times, and the final filtrate was mixed with an equal volume of 30% glycerol plus 0.05% cysteine. This solution was aliquoted into Balch tubes with rubber stoppers for transport and introduction into the gnotobiotic isolator. The remaining fecal solution was frozen in aliquots at −80 °C.

Colonization of germ-free mice. All mice were weaned at 3 weeks of age onto a plant-based mouse chow (Purina Lab Diet 5K67) and colonized with human feces at weaning. Germ-free mice received autoclaved sterile water. Both male and female mice were used for all experiments. Each experiment was littermate controlled. All mice were identified by unique five-digit ear tags. All work was performed in accordance with the University of Chicago Institutional Biosafety and Animal Care and Use Committees. Each human infant donor transfer was maintained in its own flexible film isolator to avoid cross-contamination. In all experiments, repository mice were created from human fecal donors by intragastric gavage of germ-free mice with 500 µl of freshly prepared infant fecal homogenate. These repositories were then used to colonize subsequent experimental mice via mouse-to-mouse transfer by intragastric gavage of mouse feces. Fecal samples from both repository and experimental mice were examined regularly by 16S rRNA analysis, which demonstrated that mouse-to-mouse transfer from repository to experimental mice by gavage was highly reproducible and stable over time. For colonization of experimental mice, a freshly voided fecal pellet from a repository mouse was homogenized in 1 ml of sterile PBS, and 250 µl of this homogenate was used to gavage one recipient mouse. For mice fed infant formula, the drinking water was replaced by formula 4 h prior to colonization. Mice colonized with healthy infant feces were given Enfamil Infant (Mead Johnson Nutrition), and CMA-colonized mice were given extensively hydrolyzed casein formula, Nutramigen I (Mead Johnson Nutrition) *ad libitum*. Both dry and liquid forms of the formulas were utilized. Dry formula was mixed with autoclaved sterile water, USP grade, according to the manufacturer's instructions. All formulas were refreshed daily.

For *A. caccae*-monocolonized mice, *A. caccae* (DSM-14662; Deutsche Sammlung von Mikroorganismen und Zellkulturen) was cultured in an anaerobic chamber (Coy, Model B) in reduced Schaedler's broth (Remel) overnight at 37 °C to an optical density (OD600) of 1.08, and 250 µl (approximately 2.5×10^8 CFU) was gavaged to germ-free mice. These mice were monitored for colonization by qPCR and were maintained as living repositories. For colonization of experimental mice, Enfamil Infant formula (liquid) was added to the drinking water 4 h prior to colonization. A freshly voided fecal pellet from a repository mouse was then homogenized in 1 ml of sterile PBS, and 250 µl of this homogenate was used to gavage one recipient mouse. Monocolonization with *A. caccae* was confirmed by qPCR with species-specific primers (Supplementary Table 6). In addition, monocolonization with *A. caccae* was confirmed by 16S rRNA-targeted sequencing of fecal samples collected at sacrifice for all experimental mice.

16S rRNA-targeted sequencing. Bacterial DNA was extracted using the Power Soil DNA Isolation Kit (MoBio). At the Environmental Sample Preparation and Sequencing Facility at Argonne National Laboratory, 16S rRNA gene amplicon sequencing was performed on an Illumina MiSeq instrument. Procedures described in ref. [34] were used to generate 151–base pair (bp) paired-end reads from the fecal samples with 12-bp barcodes. The V4 region of the 16S rRNA gene was PCR amplified with region-specific primers (515F–806R) that include sequencer adapter sequences used in the Illumina flowcell. The microbiota signature cohort consisting of infant donor fecal samples and gnotobiotic mouse fecal and ileal samples ($n = 99$) was analyzed by Quantitative Insights into Microbial Ecology (QIIME) v.1.9[35]. Raw reads were trimmed to remove low-quality bases; paired-end 3′ overlapping sequences were merged using SeqPrep (https://github.com/jstjohn/SeqPrep). The open reference OTU picking protocol was used at 97% sequence identity against the Greengenes database (August 2013 release)[36]. Sequences were aligned with PyNAST[37]. Taxonomic assignments were made with the uclust consensus taxonomy assigner[38]; predicted chimeric sequences were removed using ChimeraSlayer (v.20110519; http://microbiomeutil.sourceforge.net). Data were rarefied to an even depth of 3,160 reads for the donor and colonized mouse cohort ($n = 99$, consisting of donor fecal samples, mouse fecal samples at 2 and 3 weeks post-colonization, and mouse ileal samples) and 10,050 reads for the mouse cohort shown in Extended Data Fig. 8c ($n = 70$, consisting of paired fecal and ileal samples from the 35 mice at 1 week post-colonization). Alpha (Shannon index) and beta diversity metrics were compared between CMA and healthy groups using the Mann–Whitney–Wilcoxon test (non-parametric) and PERMANOVA with weighted UniFrac distance in R package vegan (v.2.4.5)[39], respectively. Pielou's evenness index J' was computed by $J' = \frac{H'}{lnS}$, where H' is the Shannon index, ln is natural logarithm, and S is the maximum number of OTUs. Discrete false discovery rate (DS-FDR)[40] was used to identify differentially abundant bacterial taxa between fecal communities of the CMA and healthy groups with parameters 'transform_type = normdata, method = meandiff, alpha = 0.10, numperm = 1000, fdr_method = dsfdr' (accessed 02/26/2018; https://github.com/biocore/dsFDR). Compared with the Benjamini–Hochberg false discovery rate (BH-FDR) method, the DS-FDR method has increased power with limited sample size and is robust to sparse data structure (low proportion of non-zero values in the microbe abundance table) and is therefore uniquely suited for data analysis of microbe communities[40]. The DS-FDR algorithm does not compute adjusted P values; instead, it estimates the FDR from a permutation test (default 1,000 permutations), which controls the FDR at the desired level (0.10). As such, it computes the raw P values, test statistics and rejected hypotheses in the output (Supplementary Tables 3 and 5). In each comparison, OTUs present in fewer than four samples were removed prior to applying the DS-FDR test. LEfSe analysis was used to identify taxa significantly enriched in the CMA or healthy group compared with the other, using the per-sample normalization value of 1,000,000 and default values for other parameters[41]. In LEfSe analysis, the linear discriminant analysis (LDA) score was computed for taxa differentially abundant between the two groups. A taxon at $P < 0.05$ (Kruskal–Wallis test) and $\log_{10}[\text{LDA}] \geq 2.0$ (or ≤ -2.0) was considered significant. For Fig. 2a, after differential abundance testing in donor CMA versus healthy comparison using DS-FDR, we further filtered the significant OTUs by requiring presence in at least two mouse groups, leaving a total of 58 OTUs for further analysis. An OTU ratio was calculated by dividing the total number of potentially protective OTUs (more abundant in healthy) by the total number of potentially non-protective OTUs (more abundant in CMA) per sample. In addition, an OTU abundance score was computed taking into consideration the abundance of 58 OTUs identified in CMA relative to healthy donor fecal samples shown in Fig. 2a. First, data transformation was applied on the relative abundance to bring the signal close to Gaussian distribution. The relative abundance of each OTU was multiplied by a constant (1×10^6) to bring all values to larger than 1, $\log_{10}$ transformed and scaled by dividing the value by their root mean square across samples. The abundance of potentially non-protective OTUs was multiplied by −1. Next, the sum of the transformed abundance of the 58 OTUs was calculated to generate the aggregate score. To validate the OTU ratio differences in the independent cohort, we re-analyzed the 16S sequencing data of fecal samples collected from the healthy and CMA infants ($n = 38$) in ref. [5] using the same analysis protocol described above, with data rarefied to an even depth of 6,424 reads. Among the 58 OTUs shown in Fig. 2a, 55 OTUs were assigned with known reference IDs and 3 with new reference IDs (Supplementary Table 3). The new reference OTU IDs are not comparable between different analysis cohorts, so we focused on the OTUs with known reference IDs. Out of 55 known OTUs, 52 were matched in the re-analyzed independent cohort and used for the calculation of protective/non-protective OTU ratio depicted in Extended Data Fig. 7.

Food allergen sensitization and challenge. Protocols were adapted from ref. [7]. All mice were weaned onto a plant-based mouse chow (Purina Lab Diet 5K67)

at 3 weeks of age. Germ-free mice received autoclaved sterile water. For mice colonized with feces from infant donors, or monocolonized with *A. caccae*, the drinking water was replaced by formula 4 h prior to colonization. Mice colonized with healthy feces or *A. caccae* received Enfamil; CMA-colonized mice received Nutramigen (both from Mead Johnson). On day 0, one week post-weaning (germ-free) or colonization (healthy/*A. caccae*/CMA), all mice were fasted for 4 h and then given a gavage of 200 mM sodium bicarbonate. After 30 min, mice were given 20 mg BLG (Sigma-Aldrich) plus 10 μg cholera toxin (List Biologicals). This protocol was repeated weekly for 5 weeks. For formula-fed mice, formula was replaced by sterile water for the week after the last sensitization. Prior to challenge on day 42, mice were fasted for 4 h and given sodium bicarbonate by gavage. Two doses of 100 mg BLG each were then administered via intragastric gavage 30 min apart. Core body temperature was measured in a blinded fashion prior to allergen challenge and every 5 min after the first challenge until at least 30 min after the second challenge using a rectal probe (PhysiTemp). Serum was collected 1 h after the second challenge to measure mMCPT-1 levels. Serum was collected 24 h after challenge for antibody measurements.

ELISAs. mMCPT-1 was quantified in serum collected 1 h after the second challenge according to the manufacturer's protocol (eBioscience). BLG-specific ELISAs were performed using protocols modified from ref. [7]. Briefly, plates were coated overnight at 4 °C with 100 μg ml^{-1} BLG in 100 mM carbonate-bicarbonate buffer (pH 9.6). Plates were blocked for 2 h at room temperature with 3% BSA. Samples were added in 1% BSA and incubated overnight at 4 °C. Assays were standardized with BLG-specific antibodies (IgE or IgG1) purified on a BLG-conjugated-CNBr-Sepharose affinity column using sera from mice immunized with BLG plus alum[42]. BLG-specific antibodies were detected with goat anti–mouse IgE-UNLB (Southern Biotech) and rabbit anti–goat IgG-AP (Thermo Fisher Scientific), then developed with *p*-NPP (KPL Labs) or IgG1-HRP (Southern Biotech) and developed with TMB (Sigma-Aldrich). See Supplementary Table 7 for a list of all antibodies used in this report.

For cytokine analysis, spleens were harvested 24 h post-challenge from *A. caccae*- or CMA-colonized mice sensitized with BLG plus cholera toxin for 5 weeks. Splenocytes were stimulated at a concentration of 2×10^6 cells ml^{-1} at 37 °C, 10% CO_2 with 10 mg ml^{-1} BLG (Sigma-Aldrich) in cDMEM with 4% FCS (HyClone), 10 mM HEPES (Gibco), 100 U ml^{-1} penicillin/streptomycin (Gibco) and 55 μM 2-mercaptoethanol (Gibco). Cytokine concentrations in 72 h culture supernatants were determined by ELISA for IL-13 and IL-4 (both from Invitrogen).

Epithelial cell isolation. As in sensitization experiments, mice were weaned at 3 weeks of age and placed on infant formula prior to colonization. Seven days after colonization, mice were euthanized and ileum was removed. For IEC isolation, tissues were cleaned and inverted as described in ref. [43]. IECs were collected by inflating inverted tissue in Cell Recovery Solution (Corning) every 5 min for 30 min. IEC samples were lysed in TRIZol (Thermo Fisher Scientific), and RNA was extracted with a PureLink RNA Mini Kit (Ambion) plus on-column DNase treatment (PureLink DNase Set, Ambion).

RNA-seq. RNA libraries were prepared using a TruSeq Stranded Total Library Preparation Kit with Ribo-Zero human/mouse/rat (Illumina). Samples were sequenced at the University of Chicago Functional Genomics Core, using 50-bp single-end reads chemistry on a HiSeq2500 instrument, with sequencing replicates in two lanes. The quality of raw reads was assessed by FastQC (v.0.11.5)[44]. The QC30 score across 39 RNA-seq samples was 96.81% ± 0.06% (mean ± s.e.m), which represents the percentage of bases with quality score ≥Q30. Alignment to the mouse reference transcriptome was performed with Gencode gene annotation (v.M16, GRCm38) by Kallisto (v.0.43.1) with the strand-specific mode[45]. This mode implements a k-mer-based pseudo-alignment algorithm to accurately quantify transcripts from RNA-seq data while robustly detecting errors in the reads. The average mapping rate was 62.77% ± 1.10% (mean ± s.e.m) based on the Kallisto pseudo-alignments to the reference transcriptome. On average, 35 million raw sequencing reads were generated per sample, and 22 million were mapped to the transcriptome using Kallisto. Transcript-level abundance was quantified specifying strand-specific protocol, summarized into gene level using tximport (v.1.4.0)[46], normalized by the trimmed mean of M values method, and $\log_2$-transformed. Genes expressed in at least six samples (counts per million reads > 3) were kept for further analysis. Genes differentially expressed between groups were identified using the limma voom algorithm with precision weights (v.3.34.5)[47]. The duplicateCorrelation function was used to estimate the correlation among mouse samples with 'Donor' (1 to 8) as the blocking factor. The lmFit function was used to fit all mouse samples ($n = 39$; 18 CMA-colonized, 18 healthy-colonized and 3 germ-free) into one linear model incorporating the correlation structure computed from above. Contrasts were set as CMA versus healthy, CMA versus germ-free and healthy versus germ-free to identify DEGs in each comparison. Genes that are significantly differentially expressed between CMA and healthy, and also different from germ-free mice, were identified using a two-step procedure: first, genes were detected as different in CMA versus healthy comparison with fold change ≥ 1.5 or ≤ −1.5 at FDR corrected *P* value smaller than 0.10; second, genes from the first step were further filtered by fold change ≥ 1.5 or ≤ −1.5 in CMA versus germ-free or healthy versus germ-free comparison at FDR 0.05. A more stringent FDR threshold (0.05) was applied in the second step to prioritize potentially true positives when compared with the negative control (germ-free). Multiple testing correction was performed using the BH-FDR method[48]. A total of 32 DEGs passed these thresholds, which represent 4 types of gene expression changes in colonized mice: (1) 'up in healthy': genes that are upregulated in healthy mice relative to both CMA and germ-free; (2) 'up in CMA': genes that are upregulated in CMA mice relative to both healthy and germ-free; (3) 'down in healthy': genes that are downregulated in healthy mice relative to both CMA and germ-free; and (4) 'down in CMA': genes that are downregulated in CMA mice relative to healthy and germ-free. The four groups of DEGs are shown in Figs. 3a and 4a. GO and KEGG pathways significantly enriched in the 32 DEGs of interest were identified using clusterprofiler (v.3.6.0)[49] at FDR corrected *P* value smaller than 0.10 (BH-FDR method, hypergeometric test). The DEGs were split into two groups for this analysis: 'healthy' included all genes that were 'up in healthy' and 'down in CMA', and 'CMA' included all genes that were 'up in CMA' and 'down in healthy'. Correlation between DEGs and ileal OTUs significantly differentially abundant between CMA and healthy samples were computed using Spearman's rank correlation method, followed by applying filters, first to keep the OTUs that show significant correlation with at least 1 DEG from the designated group at $P < 0.05$. For potentially protective OTUs (more abundant in healthy), they are correlated with genes from the 'up in healthy' or 'down in healthy' group; for potentially non-protective OTUs (more abundant in CMA), they are correlated with genes from the 'up in CMA' or 'down in CMA' group. Second, filters were applied to keep the OTUs that show a relatively consistent trend of positive correlation (Spearman's $\rho > 0.20$) across at least 60% of the DEGs from the designated group. For potentially protective OTUs, they are correlated with genes from the groups 'up in healthy' and 'down in CMA' joined; for potentially non-protective OTUs, they are correlated with genes from groups 'up in CMA' and 'down in healthy' joined. Finally, filters were applied to keep OTUs present in at least three CMA and three healthy mice. Nine ileal OTUs passed these correlation filters, and these are shown in Fig. 4a. Correlation between the relative abundance of OTUs and gene expression was calculated using Spearman's correlation method with samples that are above the limit of detection for the assay.

qPCR. Gene expression was measured by qPCR as described in ref. [7]. In brief, complementary DNA was prepared from RNA using the iScript cDNA Synthesis kit (BioRad). Gene expression was measured with PowerUp SYBR green master mix (Applied Biosystems) according to the manufacturer's instructions. Primers are derived from refs. [20] and[50–54] and are listed in Supplementary Table 8. Expression of genes of interest was normalized to *Hprt*. Relative expression was measured using ΔΔCt centered around the geometric mean; germ-free mice were used as a reference.

The presence of *A. caccae* in fecal and ileal samples was confirmed using qPCR as described in ref. [24]. Bacterial DNA was extracted using the Power Soil DNA Isolation Kit (MoBio), and qPCR was performed using PowerUp SYBR green master mix (Applied Biosystems) using 4 μl of each primer at 10 μM working dilution and 2 μl of bacterial DNA. Primers are listed in Supplementary Table 6. The cycling conditions for the reaction consisted of an activation cycle of 50 °C for 2 min followed by one cycle of 95 °C for 10 min and 40 cycles at 94 °C for 20 s, 55 °C for 20 s and 72 °C for 50 s. The fluorescent probe was detected in the last step of this cycle. A melt curve was performed at the end of the PCR to confirm the specificity of the PCR product. Relative abundance is expressed as 2^{-Ct} normalized to total 16S rRNA copies per gram of fecal material and multiplied by a constant (1×10^{25}) to bring all values above 1.

Histopathologic analysis. For histological analysis, 3 mm pieces of mid-colon and mid-ileum tissue were fixed in either 10% formalin for H&E staining or Carnoy's fixative for periodic acid–Schiff (PAS) staining. Sectioning and staining were performed by the Human Tissue Resource Center at the University of Chicago. All sections were reviewed by a gastrointestinal pathologist in a blinded fashion.

Statistical analysis. Prism 7.0 (GraphPad) was used to perform one-way (Fig. 4f) ANOVA with Bonferroni correction for multiple comparisons and Student's *t* test (Fig. 4d and Extended Data Fig. 9b), as indicated in the figure legends. The DS-FDR method was used to identify significant OTUs (Figs. 2a and 4c and Extended Data Fig. 9a) comparing CMA with the healthy group. The BH-FDR method was used for multiple testing correction in RNA-seq analysis (Fig. 3a) and Gene Ontology enrichment analysis (Fig. 3b). Shannon diversity and Pielou's evenness were compared using a non-parametric Mann–Whitney–Wilcoxon test (Extended Data Fig. 3a,b). Analysis of protective/non-protective OTU ratio in the larger, independent cohort of infants was performed using two-sided Wilcoxon rank sum test (Extended Data Fig. 7). The biological responses of different donor colonized mice to sensitization with BLG (Figs. 1a–d and 4g and Extended Data Figs. 4a and 5a) were explored with linear mixed-effect models[55] based on restricted maximum likelihood in R (lmerTest v.3.0.1)[56]. Group (germ-free, healthy and CMA in Fig. 1a; *A. caccae* and CMA in Fig. 4g; healthy BFD and CMA BFD in Extended Data Fig. 4a; H_2O and Enfamil in Extended Data Fig. 5a) temperature changes across time (both linear and quadratic) were modeled as

temperature = group + time × group + time × time × group with random intercepts and slopes estimated for individual mice. Contrasts of group temperature trends were performed using *t*-tests with the BH-FDR adjustment for multiple comparisons. To control for cases where groups contained multiple donors (Fig. 1a), we updated the previous model to include mice nested within each donor as a random effect and repeated the contrasts. Since the results of the two models were concordant, we chose to report the results from the first model for consistency of methods. For Fig. 1b–d, antibody concentrations were log transformed and modeled as log[concentration] = group with donors as a random effect. Contrasts for group differences were performed using the previously mentioned methods. For Fig. 4h–l and Extended Data Figs. 4b–d and 5b–d, antibody and cytokine concentrations were log transformed and compared using *t*-tests. The data analysis commands (including the data files and R markdown files for reproducibility) are available from the authors upon request.

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Code availability

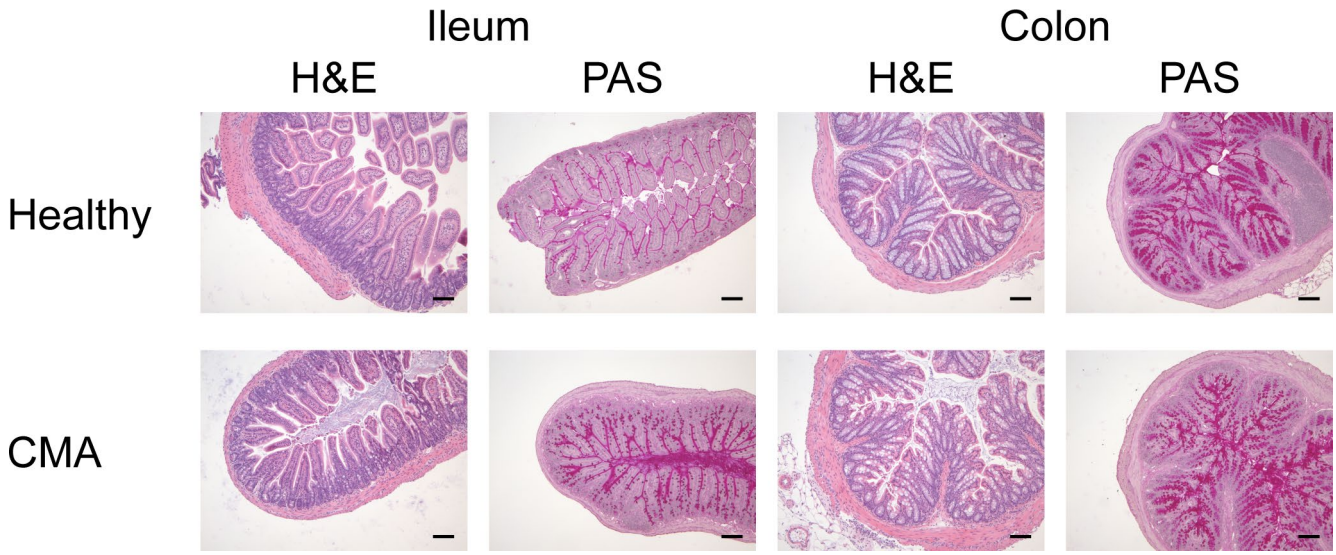
The open-source analysis software used in this study is publicly available and referenced as appropriate. Custom codes are available from the corresponding author upon request.

Data availability

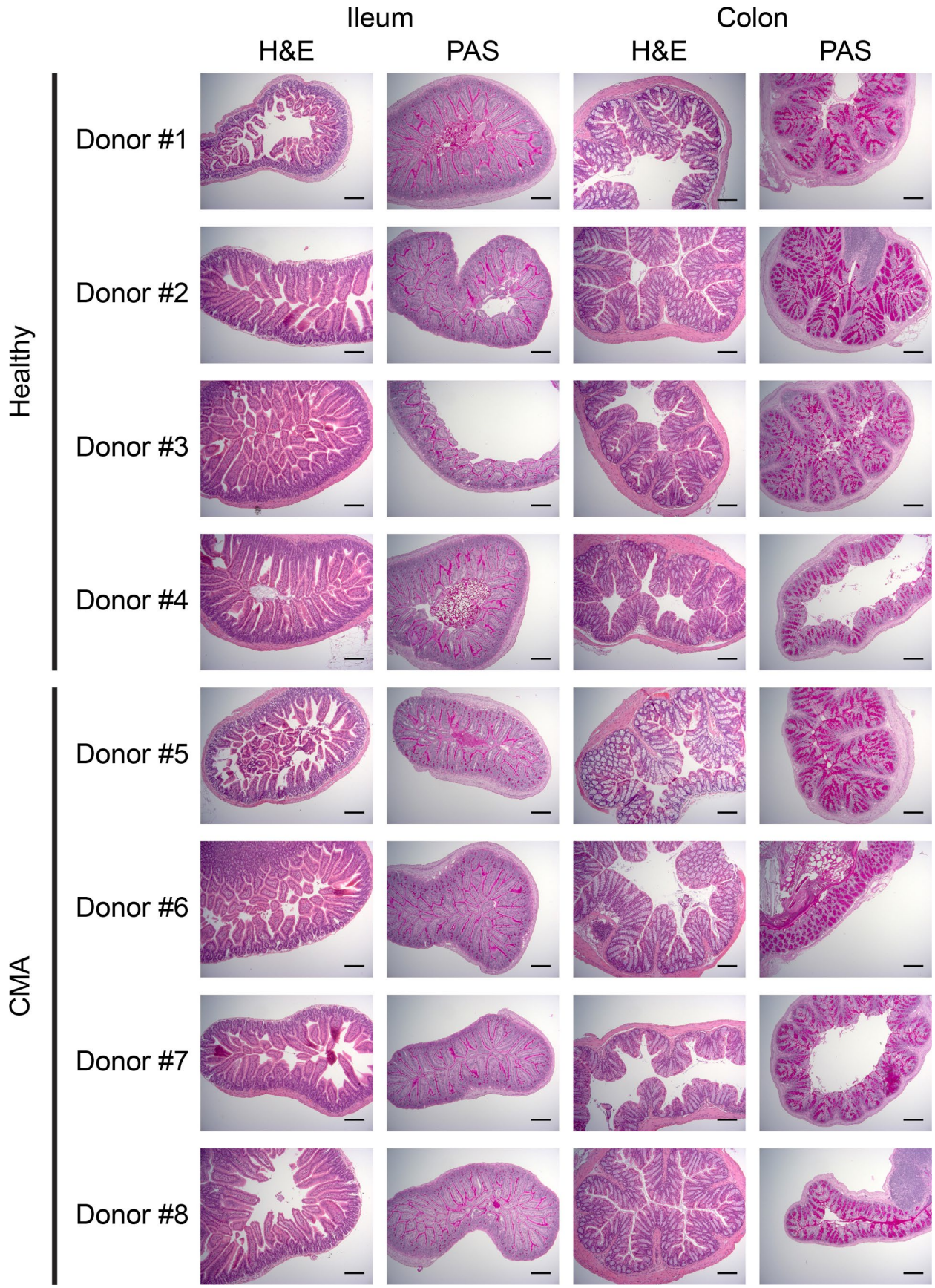
The data that support the findings of this study are available from the corresponding author upon request. The 16S rRNA and RNA-seq raw FastQ sequencing files were deposited into the National Center for Biotechnology Information Sequence Read Archive and are available under the accession numbers SRP130620 and SRP130644, respectively. Additional processed data reported in this study are available upon request.

References

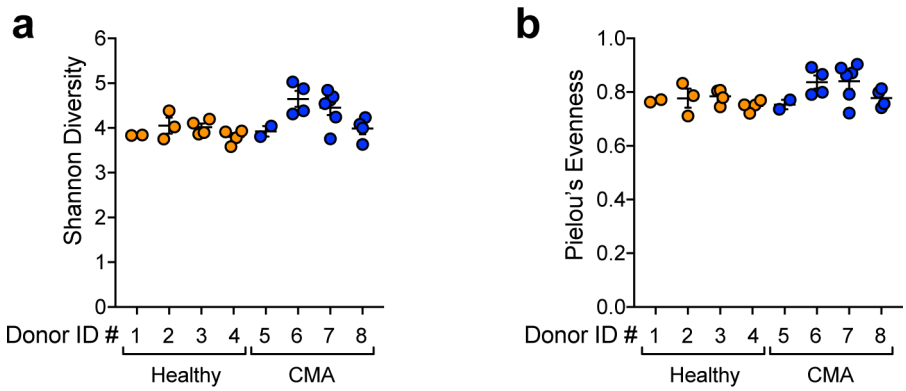
34. Caporaso, J. G. et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **6**, 1621–1624 (2012).
35. Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
36. DeSantis, T. Z. et al. Greengenes, a chimera-checked 16 S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72**, 5069–5072 (2006).
37. Caporaso, J. G. et al. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* **26**, 266–267 (2010).
38. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
39. Oksanen, J., et al. Package 'vegan': Community Ecology Package. R package v.2.4.5 (2017).
40. Jiang, L. et al. Discrete false-discovery rate improves identification of differentially abundant microbes. *mSystems* **2**, e00092-17 (2017).
41. Segata, N. et al. Metagenomic biomarker discovery and explanation. *Genome Biol.* **12**, R60 (2011).
42. Bashir, M. E., Louie, S., Shi, H. N. & Nagler-Anderson, C. Toll-like receptor 4 signaling by intestinal microbes influences susceptibility to food allergy. *J. Immunol.* **172**, 6978–6987 (2004).
43. Nik, A. M. & Carlsson, P. Separation of intact intestinal epithelium from mesenchyme. *Biotechniques* **55**, 42–44 (2013).
44. Andrew, S. FastQC: a quality control application for high throughput sequence data. *Babraham Institute* http://www.bioinformatics.babraham.ac.uk/projects/fastqc (2016).
45. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
46. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research* **4**, 1521 (2015).
47. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
48. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).
49. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
50. Upadhyay, V. et al. Lymphotoxin regulates commensal responses to enable diet-induced obesity. *Nat. Immunol.* **13**, 947–953 (2012).
51. Liu, X. et al. Warburg effect revisited: an epigenetic link between glycolysis and gastric carcinogenesis. *Oncogene* **29**, 442–450 (2010).
52. Roelen, B. A., Lin, H. Y., Knezevic, V., Freund, E. & Mummery, C. L. Expression of TGF-βs and their receptors during implantation and organogenesis of the mouse embryo. *Dev. Biol.* **166**, 716–728 (1994).
53. Ellis, J. M., Bowman, C. E. & Wolfgang, M. J. Metabolic and tissue-specific regulation of acyl-CoA metabolism. *PLoS ONE* **10**, e0116587 (2015).
54. Al-Dwairi, A., Pabona, J. M., Simmen, R. C. & Simmen, F. A. Cytosolic malic enzyme 1 (ME1) mediates high fat diet-induced adiposity, endocrine profile, and gastrointestinal tract proliferation-associated biomarkers in male mice. *PLoS ONE* **7**, e46716 (2012).
55. Pinheiro, J. C. & Bates, D. M. *Mixed Effects in Models S and S Plus* (Springer, New York, 2000).
56. Kuznetsova, A., Brockhoff, P. B., Rune, H. & Christensen, B. lmertest package: tests in linear mixed effects models. *J. Stat. Softw.* **82**, 1–26 (2017).



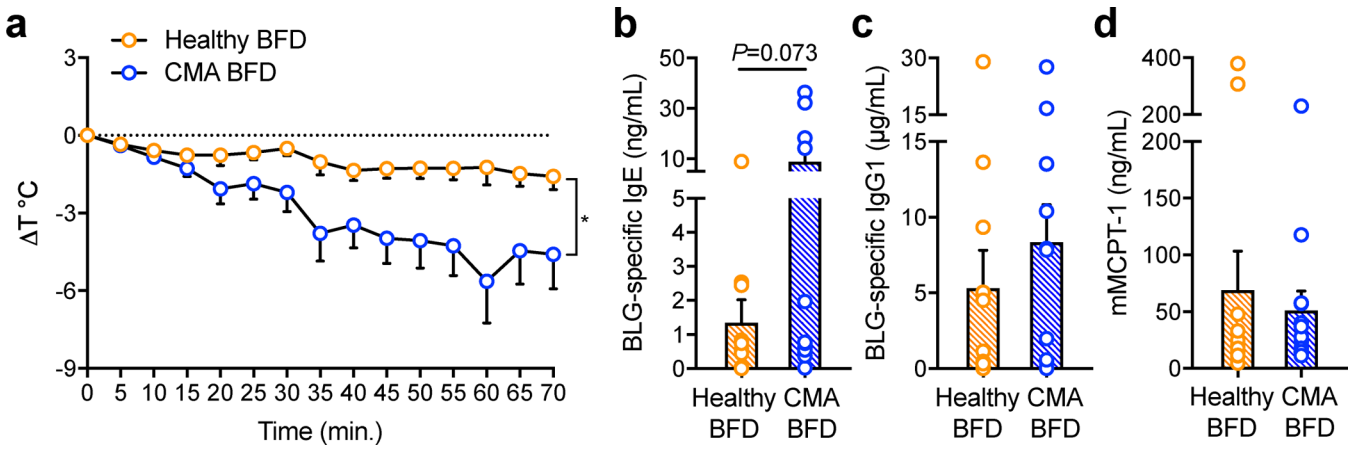
Extended Data Fig. 1 | Sensitization of healthy- or CMA-colonized mice with BLG plus cholera toxin does not result in intestinal pathology. Representative images of histological samples from BLG plus cholera toxin-sensitized healthy- or CMA-colonized mice 24 h post-challenge for donors 1 (healthy) and 5 (CMA; see Supplementary Table 1). All sections stained with H&E or PAS, as indicated. Scale bars, 100 μm.



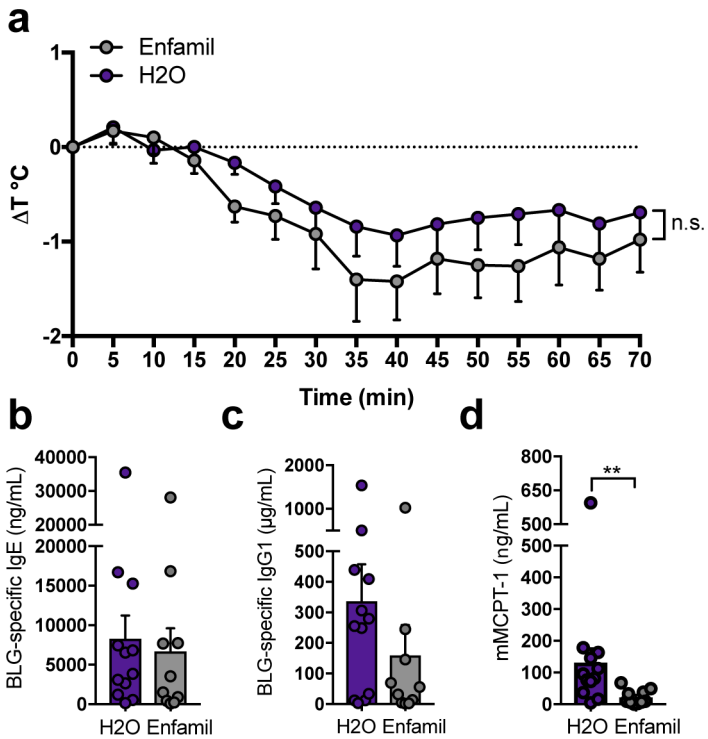
Extended Data Fig. 2 | Long-term colonization of germ-free mice with feces from healthy or CMA infants does not lead to intestinal pathology. Representative images of histological samples from unsensitized healthy- or CMA-colonized mice collected 5 to 6 months post-colonization for donors described in Supplementary Table 1. All sections stained with H&E or PAS, as indicated. Scale bars, 100 μm.



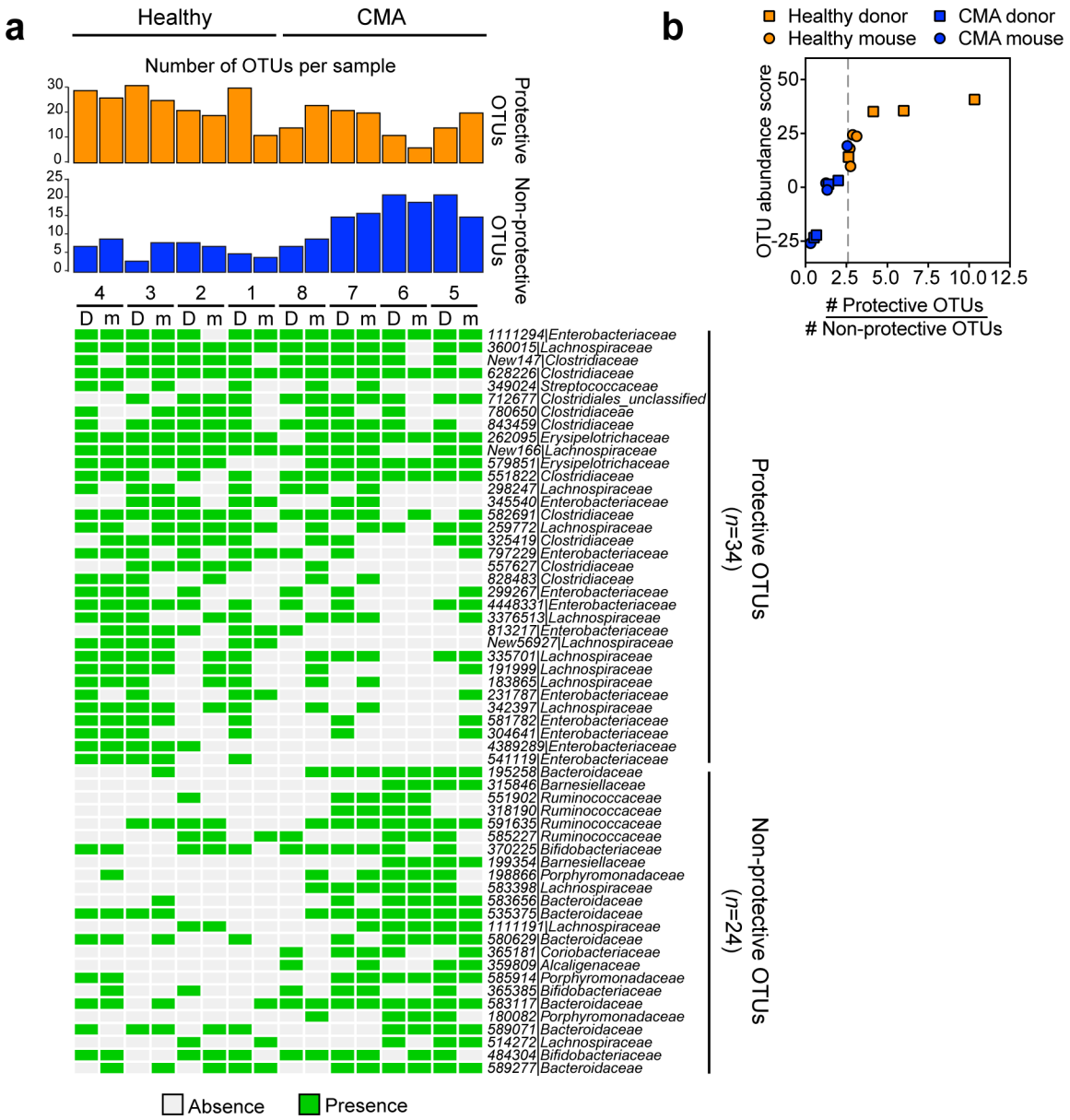
Extended Data Fig. 3 | Diversity analysis of fecal samples from healthy- or CMA-colonized mice. Shannon diversity index (**a**) and Pielou's evenness index (**b**) in feces from healthy-colonized (orange) and CMA-colonized (blue) mice from Fig. 2a. $n = 1$–4 mice per colonized mouse group with feces taken at 2 and 3 weeks post-colonization; see Online Methods). Each circle represents one fecal sample; bars represent mean +s.e.m. The eight human formula-fed fecal donors are described in Supplementary Table 1.



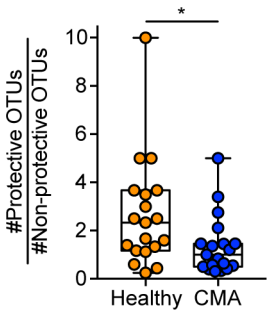
Extended Data Fig. 4 | Transfer of a healthy, exclusively breast-fed infant microbiota protects against an anaphylactic response to sensitization with BLG plus cholera toxin. a, Change in core body temperature at indicated time points following first challenge with BLG of mice colonized with feces from breast-fed healthy or CMA infant donors ($n = 13$ mice per group, collected from at least 2 independent experiments). **b**–**d**, Serum BLG-specific IgE (**b**), BLG-specific IgG1 (**c**) and mMCPT-1 (**d**) from mice in **a**. Four of the BLG plus cholera toxin-sensitized CMA-colonized mice died of anaphylaxis following challenge. For **a**, symbols represent mean, and bars represent s.e.m. For **b**–**d**, symbols represent individual mice, and bars represent mean + s.e.m. Linear mixed-effect models were used to compare groups in **a** and two-sided Student's *t*-test in **b** after log transformation. The two human breast-fed fecal donors are described in Supplementary Table 2. $^*P < 0.05$.



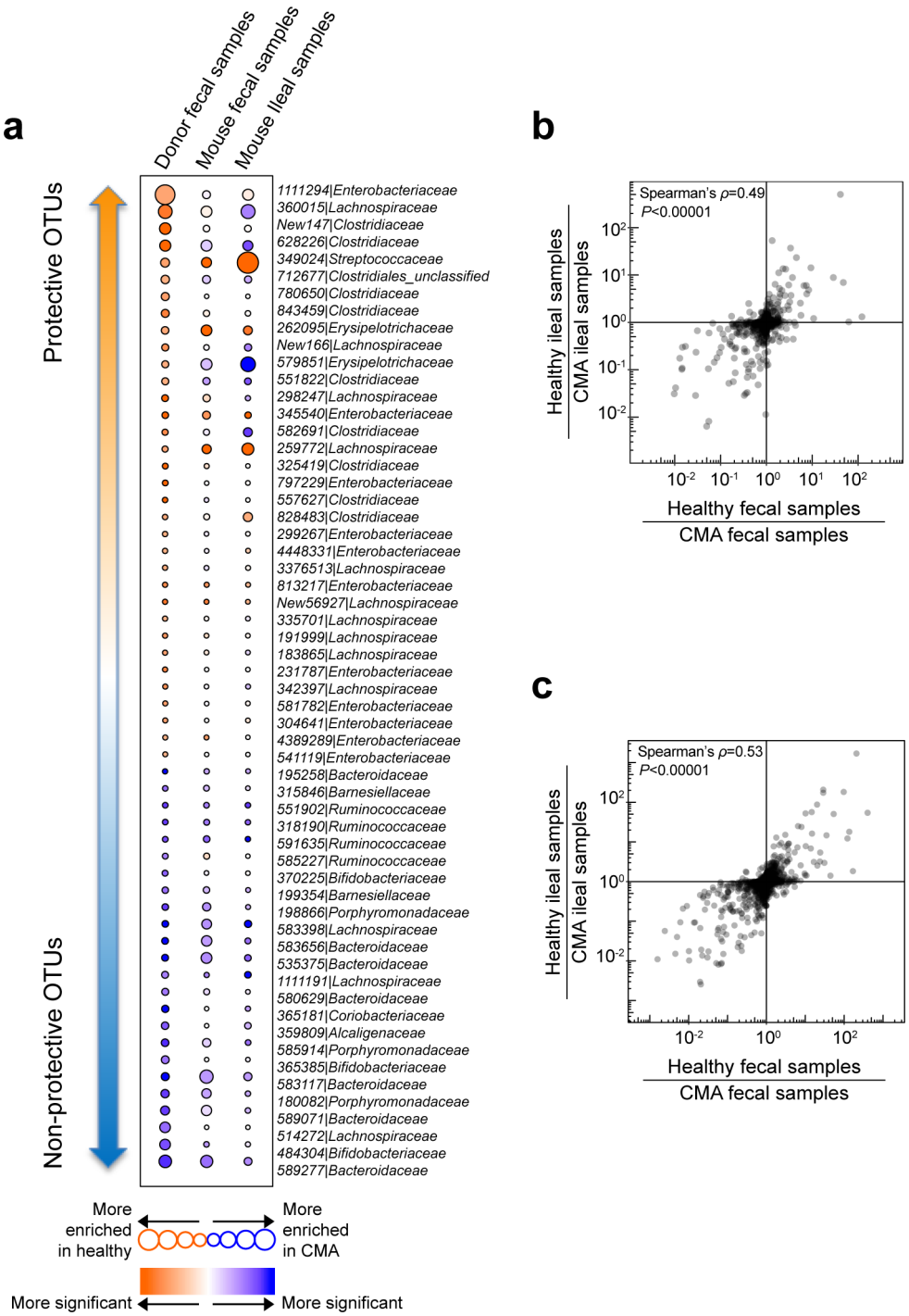
Extended Data Fig. 5 | Continuous exposure to cow's milk does not induce tolerance to BLG in germ-free mice fed with water or Enfamil and sensitized with BLG plus cholera toxin. **a**, Change in core body temperature at indicated time points following first challenge with BLG of mice fed with water ($n = 12$) or Enfamil ($n = 10$) collected from 3 independent experiments. **b**–**d**, serum BLG-specific IgE (**b**), BLG-specific IgG1 (**c**) and mMCPT-1 (**d**) from mice in **a**. For **a**, circles represent mean, and error bars represent s.e.m. For **b**–**d**, circles represent individual mice, and bars represent mean + s.e.m. Linear mixed-effect models were used to compare groups in **a** and two-sided Student's *t*-test in **b**–**d** after log transformation. $^{**}P < 0.01$. n.s. = not significant ($P = 0.36$).



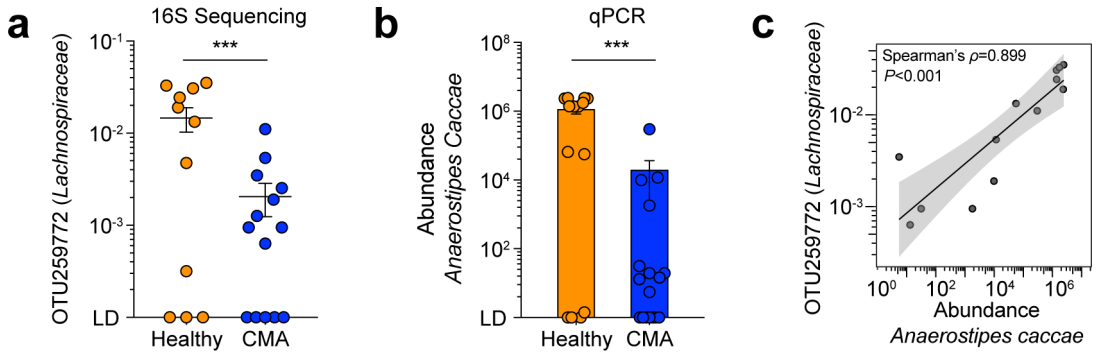
Extended Data Fig. 6 | Binary representation of protective and non-protective OTUs in CMA and healthy donors and colonized mouse groups. a, Binary map of the presence/absence ratio of protective/non-protective OTUs in CMA and healthy donors with the same layout as Fig. 2a. Columns depict each donor (D) or colonized mouse group (m). $n = 2$–3 technical replicates per donor and $n = 1$–4 mice per colonized mouse group, with feces taken at 2 and 3 weeks post-colonization; see Online Methods). Rows show 58 OTUs FDR controlled at 0.10 (see Online Methods) in human CMA versus healthy donor comparison, present in at least 4 human fecal samples and at least 2 groups of colonized mice (see Supplementary Table 3). The bar graphs above the grid map represent the total number of potentially protective (more abundant in healthy donors; orange) and potentially non-protective (more abundant in CMA donors; blue) OTUs in each individual donor or mouse group. The grid map represents presence (green) or absence (white) of protective and non-protective OTUs in each sample. **b**, A protective/non-protective OTU ratio was computed per individual donor or mouse group from **a**, taking into consideration the presence or absence of 58 OTUs. The donors and their murine transfer recipients are shown in squares and circles, respectively. The vertical dashed line represents a ratio of 2.6.



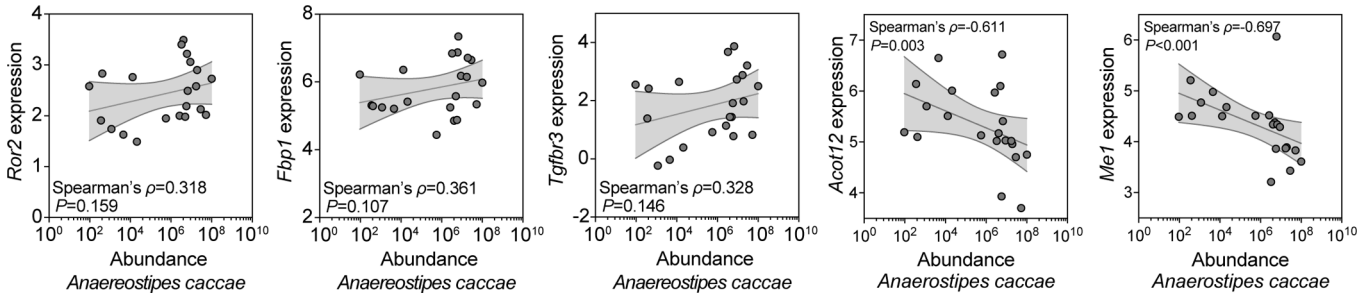
Extended Data Fig. 7 | Validation of protective/non-protective OTU ratio using a larger, independent cohort of healthy and CMA infant donors. Box plots showing the protective/non-protective OTU ratio (see Fig. 2 and Extended Data Fig. 6) in fecal samples from healthy ($n=19$) and CMA ($n=19$) infants from ref. [5]. The horizontal center line indicates the median, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the farthest data point within a maximum of 1.5 times the interquartile range (IQR). All individual points are shown, with each circle denoting a subject. Out of the 58 OTUs shown in Fig. 2a, 55 OTUs were assigned with known reference IDs and 3 with new reference IDs. The new reference OTU IDs are not comparable across the different analysis cohorts, so we focused on the OTUs with known reference IDs. Among the 55 known OTUs, 52 (29 protective OTUs and 23 non-protective OTUs) were detected in this cohort and were used for the ratio calculation (see Online Methods). The other 3 were not detected. Two-sided Wilcoxon rank sum test was used. $^*P<0.05$.



Extended Data Fig. 8 | The healthy versus CMA OTU abundance ratio is significantly correlated between mouse fecal and ileal samples. **a**, Bubble plots show a similar pattern in fecal ($n=8$ mice in healthy group, $n=9$ mice in CMA group, with fecal samples collected at 2 and 3 weeks post-colonization, same as in Fig. 2a) and ileal samples ($n=22$ mice in healthy group, $n=25$ mice in CMA group) from healthy- and CMA-colonized mice; 58 OTUs significantly differentially abundant between CMA and healthy donors are shown in the same order as in Fig. 2a. The size of the circle indicates the magnitude of relative abundance enrichment towards either CMA or healthy. Color intensity indicates the statistical significance computed using the DS-FDR permutation test (see Online Methods). **b**,**c**, The healthy versus CMA OTU abundance ratio is significantly correlated between mouse fecal and ileal samples. Each dot represents one individual OTU. For **b**, for each OTU, its average abundance was calculated at the group level among 8 healthy-colonized and 9 CMA-colonized mice for the fecal samples, and among 22 healthy-colonized and 25 CMA-colonized mice for the ileal samples. The ratios of OTU abundance in the feces are plotted on the *x* axis with the ratio of OTU abundance in the ileum on the *y* axis. For **c**, $n=35$ (15 healthy-colonized and 20 CMA-colonized) mice collected from at least 2 independent experiments were used for the calculation of both the fecal and ileal OTU abundance ratio, where fecal and ileal samples were collected from the same individual mice. For further details, see the Online Methods.



Extended Data Fig. 9 | Abundance of OTU259772 (Lachnospiraceae) and *A. caccae* are correlated in fecal samples from healthy- and CMA-colonized mice. **a**,**b**, Abundance of OTU259772 (Lachnospiraceae) from the 16S dataset (**a**) and abundance of *A. caccae* by qPCR (**b**) in fecal samples from healthy-colonized ($n=7$) and CMA-colonized ($n=8$) mice from Fig. 2. For each individual mouse, 1–2 fecal samples were collected at 2 and 3 weeks post-colonization. LD indicates samples that were below the limit of detection for the assay. **c**, Spearman's correlation between abundance of OTU259772 (Lachnospiraceae; 16S sequencing) and abundance of *A. caccae* (qPCR) in fecal samples from healthy- and CMA-colonized mice from Fig. 2. Fecal samples that were above LD in both 16S and qPCR experiments are shown ($n=13$). Each circle represents one fecal sample. For **a** and **b**, bars show mean + s.e.m. For **c**, shaded bands indicate 95% confidence interval fitted by linear regression. The DS-FDR method was used to compare groups in **a** and two-sided Student's *t*-test in **b**. ***$P<0.001$.



Extended Data Fig. 10 | Abundance of A. caccae in ileal samples correlates with gene expression in ileal IECs. Spearman's correlation between abundance of *A. caccae* by qPCR and RNA-seq gene expression of *Ror2*, *Fbp1*, *Tgfbr3*, *Acot12* and *Me1* in ileal IECs (see Fig. 3a). Circles show individual mice, and shaded bands indicate 95% confidence interval fitted by linear regression. $n = 36$ (18 healthy and 18 CMA-colonized) mice collected from at least 2 independent experiments. Samples with values above the limit of detection are shown (*A. caccae* abundance > 0).

nature research

Corresponding author(s): Dr Cathryn Nagler

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	No software was used
Data analysis	QIIME v1.9.1, SeqPrep v1.2, PyNAST v0.1, uclust v10.0.240, ChimeraSlayer v20110519, vegan v2.4.5, DS-FDR accessed 02262018, LEfSe v1.0, FastQC v0.11.5, Kallisto v0.43.1, tximport v1.4.0, limma v3.34.5, clusterprofiler v3.6.0, GraphPad Prism v7.0, lmerTest v3.0.1, R v3.4.3

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data reported in this study are tabulated in the main text and supplementary materials. The 16S sequencing and RNA-Seq data files were deposited into The NCBI

Sequence Read Archive (SRA) and are available under the accession no. SRP130620 and SRP130644, respectively. Custom code and additional processed data used in this study are available upon request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences

Study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Each experiment contained at least 2 littermate controlled mice per colonization group, and typically 5-6 mice per group. Each experiment was repeated independently at least twice. See figure legends for details on n for each display figure. Given the time constraint and limited resource of the mouse model, in certain experiments the sample size is dictated by the availability of mice during designated experimental time.
Data exclusions	5 healthy-colonized and 2 CMA colonized mice were excluded from final analysis in Figure 1a-d. One mouse was runted at weaning and sickly at the time of challenge and was excluded based on this criteria. An isolator housing 4 healthy-colonized mice was compromised with a contamination during the experiment and these mice were therefore excluded. 2 CMA colonized mice were excluded because they had litters during sensitization protocol.
Replication	Experimental findings were reproduced in at least two independent experiments.
Randomization	Both male and female mice were used in all groups for all experiments. Mice were randomly allocated into each treatment group by staff at the Gnotobiotic Research Facility at University of Chicago based on instructions provided by the researcher. Human samples were selected to have healthy and CMA participants that were age- and gender-matched, had a similar birth route and age at fecal collection (see Supplementary Table 1 for patient details).
Blinding	Temperature analysis for all challenges was performed in a blinded fashion by a second, independent lab member. All mice were identified by a unique 5 digit eartag which allowed for blinding during ELISA assays and other analyses as results were first analyzed and tabulated based on the 5 digit identifier before being matched to experimental groups.

Materials & experimental systems

Policy information about availability of materials

n/a	Involved in the study
☒	☐ Unique materials
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☐	☒ Research animals
☐	☒ Human research participants

Antibodies

Antibodies used	Goat anti-mouse IgE UNLB Southern Biotech Cat# 1110-01 (1:2,000) Rabbit anti-goat IgG AP ThermoFisher Cat# 31300 (1:5,000) Goat anti-mouse IgG1 HRP Southern Biotech Cat# 1070-05 (1:10,000) Purified Rat anti-mouse IgE (IgE coating Ab) BD Pharminagen Cat# 553413 (1:250) Purified mouse IgE k Isotype control (IgE standard) BD Pharmingen Cat# 557079 (1:250) Biotin Rat anti-mouse IgE (Biotin IgE secondary Ab) BD Pharmingen Cat# 553419 (1:500) Purified Goat anti-mouse IgG1 (IgG1 coating Ab) Southern Biotech Cat# 1070-01 (1:1,000) Purified mouse IgG1-UNLB (IgG1 standard) Southern Biotech Cat# 0102-01 (1:5,000) Goat anti-mouse IgG1-HRP (IgG1 secondary Ab) Southern Biotech Cat# 1070-05 (1:10,000) Goat anti-mouse IgA-UNLB (fecal IgA coat) Southern Biotech Cat# 1040-01 (1:500) Purified mouse IgA k isotype control (IgA standard) BD Pharminigen Cat# 553476 (1:2,500) Goat anti-mouse IgA-AP (AP IgA secondary Ab) Southern Biotech Cat# 1040-04 (1:2,500) Mouse IL-13 Uncoated ELISA kit: ThermoFisher Scientific, Cat# 88-7137-88, Lot# 187510001 Mouse IL-4 ELISA Ready-SET-Go!: Affymetrix eBioscience, Ref# 88-7044-88 Lot# E09342-1642 Mouse MCPT-1 Uncoated ELISA Kit: Thermofisher Scientific, Cat# 88-7503-88 Lot# 173830001

Validation	Validation of assays using antibodies is attached in Antibody Validation Form attached.

Research animals

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Animals/animal-derived materials	Male and female germ-free C3H/HeN mice were used for all mouse studies in this paper. See methods section for these details.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Stool samples were collected from 8 male infants, at age of 6 months, with less than 14 days of breastfeeding. See Supplementary Table 1 for all details on human participants for this study.

Method-specific reporting

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ Magnetic resonance imaging

nature research

Corresponding author(s): Dr Cathryn Nagler

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	No software was used
Data analysis	QIIME v1.9.1, SeqPrep v1.2, PyNAST v0.1, uclust v10.0.240, ChimeraSlayer v20110519, vegan v2.4.5, DS-FDR accessed 02262018, LEfSe v1.0, FastQC v0.11.5, Kallisto v0.43.1, tximport v1.4.0, limma v3.34.5, clusterprofiler v3.6.0, GraphPad Prism v7.0, lmerTest v3.0.1, R v3.4.3

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data reported in this study are tabulated in the main text and supplementary materials. The 16S sequencing and RNA-Seq data files were deposited into The NCBI

Sequence Read Archive (SRA) and are available under the accession no. SRP130620 and SRP130644, respectively. Custom code and additional processed data used in this study are available upon request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences

Study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Each experiment contained at least 2 littermate controlled mice per colonization group, and typically 5-6 mice per group. Each experiment was repeated independently at least twice. See figure legends for details on n for each display figure. Given the time constraint and limited resource of the mouse model, in certain experiments the sample size is dictated by the availability of mice during designated experimental time.
Data exclusions	5 healthy-colonized and 2 CMA colonized mice were excluded from final analysis in Figure 1a-d. One mouse was runted at weaning and sickly at the time of challenge and was excluded based on this criteria. An isolator housing 4 healthy-colonized mice was compromised with a contamination during the experiment and these mice were therefore excluded. 2 CMA colonized mice were excluded because they had litters during sensitization protocol.
Replication	Experimental findings were reproduced in at least two independent experiments.
Randomization	Both male and female mice were used in all groups for all experiments. Mice were randomly allocated into each treatment group by staff at the Gnotobiotic Research Facility at University of Chicago based on instructions provided by the researcher. Human samples were selected to have healthy and CMA participants that were age- and gender-matched, had a similar birth route and age at fecal collection (see Supplementary Table 1 for patient details).
Blinding	Temperature analysis for all challenges was performed in a blinded fashion by a second, independent lab member. All mice were identified by a unique 5 digit eartag which allowed for blinding during ELISA assays and other analyses as results were first analyzed and tabulated based on the 5 digit identifier before being matched to experimental groups.

Materials & experimental systems

Policy information about availability of materials

n/a	Involved in the study
☒	☐ Unique materials
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☐	☒ Research animals
☐	☒ Human research participants

Antibodies

Antibodies used	Goat anti-mouse IgE UNLB Southern Biotech Cat# 1110-01 (1:2,000) Rabbit anti-goat IgG AP ThermoFisher Cat# 31300 (1:5,000) Goat anti-mouse IgG1 HRP Southern Biotech Cat# 1070-05 (1:10,000) Purified Rat anti-mouse IgE (IgE coating Ab) BD Pharminagen Cat# 553413 (1:250) Purified mouse IgE k Isotype control (IgE standard) BD Pharmingen Cat# 557079 (1:250) Biotin Rat anti-mouse IgE (Biotin IgE secondary Ab) BD Pharmingen Cat# 553419 (1:500) Purified Goat anti-mouse IgG1 (IgG1 coating Ab) Southern Biotech Cat# 1070-01 (1:1,000) Purified mouse IgG1-UNLB (IgG1 standard) Southern Biotech Cat# 0102-01 (1:5,000) Goat anti-mouse IgG1-HRP (IgG1 secondary Ab) Southern Biotech Cat# 1070-05 (1:10,000) Goat anti-mouse IgA-UNLB (fecal IgA coat) Southern Biotech Cat# 1040-01 (1:500) Purified mouse IgA k isotype control (IgA standard) BD Pharminigen Cat# 553476 (1:2,500) Goat anti-mouse IgA-AP (AP IgA secondary Ab) Southern Biotech Cat# 1040-04 (1:2,500) Mouse IL-13 Uncoated ELISA kit: ThermoFisher Scientific, Cat# 88-7137-88, Lot# 187510001 Mouse IL-4 ELISA Ready-SET-Go!: Affymetrix eBioscience, Ref# 88-7044-88 Lot# E09342-1642 Mouse MCPT-1 Uncoated ELISA Kit: Thermofisher Scientific, Cat# 88-7503-88 Lot# 173830001

Validation	Validation of assays using antibodies is attached in Antibody Validation Form attached.

Research animals

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Animals/animal-derived materials	Male and female germ-free C3H/HeN mice were used for all mouse studies in this paper. See methods section for these details.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Stool samples were collected from 8 male infants, at age of 6 months, with less than 14 days of breastfeeding. See Supplementary Table 1 for all details on human participants for this study.

Method-specific reporting

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ Magnetic resonance imaging