MICROBIOLOGY

Lectin-Seq: A method to profile lectin-microbe interactions in native communities

Robert L. McPherson[1,2†], Christine R. Isabella[1,2†], Rebecca L. Walker[3], Dallis Sergio[3], Sunhee Bae[1], Tony Gaca[3], Smrithi Raman[1], Le Thanh Tu Nguyen[2,4], Darryl A. Wesener[5,6], Melanie Halim[1,2], Michael G. Wuo[1,2‡], Amanda Dugan[1,2], Robert Kerby[7], Soumi Ghosh[8], Federico E. Rey[7], Catherine Dhennezel[3,9], Gleb Pishchany[3], Valerie Lensch[1], Hera Vlamakis[2,3], Eric J. Alm[2,3,6], Ramnik J. Xavier[2,3,9], Laura L. Kiessling[1,2*]

Soluble human lectins are critical components of innate immunity. Genetic models suggest that lectins influence host-resident microbiota, but their specificity for commensal and mutualist species is understudied. Elucidating lectins' roles in regulating microbiota requires an understanding of which microbial species they bind within native communities. To profile human lectin recognition, we developed Lectin-Seq. We apply Lectin-Seq to human fecal microbiota using the soluble mannose-binding lectin (MBL) and intelectin-1 (hItln1). Although each lectin binds a substantial percentage of the samples (10 to 20%), the microbial interactomes of MBL and hItln1 differ markedly in composition and diversity. MBL binding is highly selective for a small subset of species commonly associated with humans. In contrast, hItln1's interaction profile encompasses a broad range of lower-abundance species. Our data uncover stark differences in the commensal recognition properties of human lectins.

INTRODUCTION

Humans host trillions of microbes (*1*), but the mechanisms that we use to recognize and distinguish between different microbial species are not fully defined. Microbes are decorated with cell surface glycans that act as a primary point of contact between host and microbe (*2*). Microbial glycan composition is unique between species and could serve as a molecular readout of cellular identity for host factors (*3*). Soluble human carbohydrate-binding proteins termed lectins distinguish cell surface glycans and can achieve species- and strain-level binding specificity (*4*–*6*). Lectins are implicated in immune defense against pathogens, and lectin-microbe interactions can be key determinants of cell fate for microbial pathogens (*7*, *8*). Genetic models suggest that lectins play notable roles in regulating broad groups of microbial species in the gut microbiota (*9*–*12*). However, the microbial specificity of human lectins in host-resident communities has not been elucidated. Critically, it is unclear whether lectins in such communities bind a wide range of species or restrict their interactions to select microbes. Such information is vital to developing a mechanistic understanding of how lectins regulate the microbiota.

Lectins are commonly classified by their monosaccharide specificity, but their cell binding specificity can be influenced by glycan accessibility, density, and secondary interactions (*13*). In some cases, monosaccharide specificity can be used to predict lectin-microbe interactions (*14*), but in many cases, the saccharides displayed on a microbe's surface are neither known nor predictable. Previous studies have primarily characterized lectin-microbe interactions using monoculture binding assays (*11*, *14*). Single-species binding assays have drawbacks: Bacteria can modulate their glycan displays under different growth conditions, potentially leading to differences in the glycan composition of native and cultured microbes (*15*, *16*). Moreover, relevant lectin binding occurs in mixed communities. We postulate that competition for lectin binding might occur in such physiologically relevant settings.

RESULTS

To assess lectin selectivity within mixed communities, we first evaluated the binding of the human lectin intelectin-1 (hItln1) to bacterial isolates from the human gut microbiota. Previously, we found that hItln1 exhibits serotype-specific binding to isolates of *Streptococcus pneumoniae* (*14*). Others have noted binding to different bacterial species considered pathogenic, suggesting that hItln1 targets pathogens (*17*, *18*). Because *hITLN1* is expressed in the small intestine (*19*), we assessed the binding of recombinant hItln1 to a collection of human gut microbial strains by flow cytometry. The lectin binds to strains from diverse phyla (Fig. 1, A and B, and table S1). The ability of hItln1 to bind to gut-resident species that can contribute to human health emphasizes the need to further characterize human lectin binding to commensal and mutualist microbes.

[1]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [2]Center for Microbiome Informatics and Therapeutics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [3]The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [4]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [5]Edison Family Center for Genome Sciences & Systems Biology, Washington University School of Medicine, Saint Louis, MO 63110, USA. [6]Center for Gut Microbiome and Nutrition Research, Washington University School of Medicine, Saint Louis, MO 63110, USA. [7]Department of Bacteriology, University of Wisconsin-Madison, Madison, WI 53706, USA. [8]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [9]Center for Computational and Integrative Biology, Department of Molecular Biology, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA.
*Corresponding author. Email: kiesslin@mit.edu
†These authors contributed equally to this work.
‡Present address: Discovery Chemistry Department, Merck Research Laboratories, Boston, MA 02115, USA.

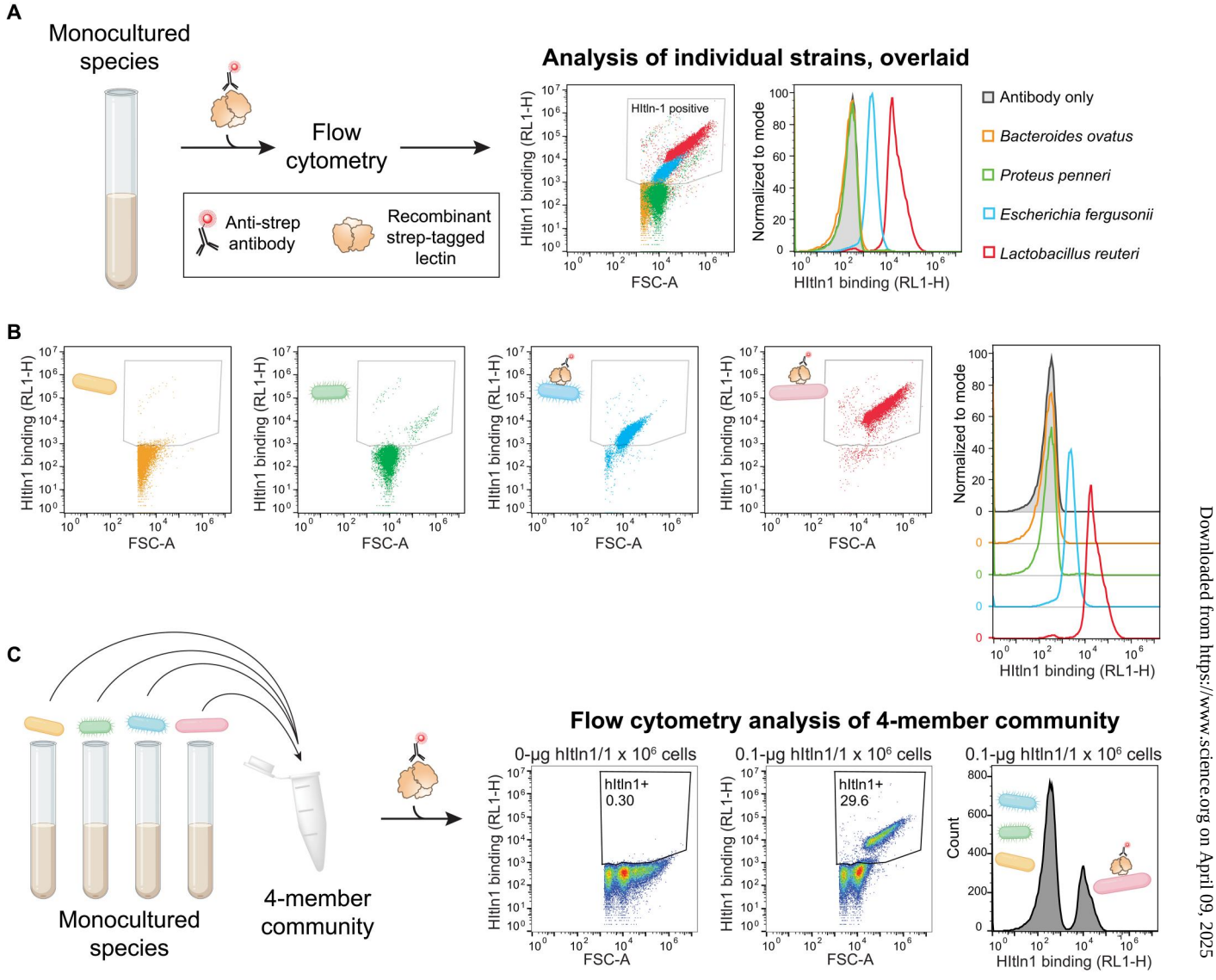

Fig. 1. HItln1 binding in a synthetic mixed microbial community. (**A**) Schematic of flow cytometry–based assay to measure soluble lectin binding to microbial isolates. Overlay of data from 0.1-µg hItln1/1 × 10^6 cells condition shown in (B) with the hItln1-positive binding region indicated. (**B**) Flow cytometry analysis of microbial isolates plotted as FSC against hItln1 binding (anti-Strep Oyster 645). Panels show 0.1-µg hItln1/1 × 10^6 cells condition. Staggered histograms of flow cytometry data plot cell count as a percent of the maximum against lectin binding (anti-Strep Oyster 645). Dot plots and histograms are color-coded according to the legend shown in (A). (**C**) Schematic of flow cytometry–based assay to measure soluble lectin binding to a synthetic mixture of microbial isolates. Flow cytometry analysis of synthetic mixed microbial community plotted as FSC against hItln1 binding (anti-Strep Oyster 645). Panels show 0-µg hItln1/1 × 10^6 cells and 0.1-µg hItln1/1 × 10^6 cells conditions. Histogram of flow cytometry data plots cell count as a percent of the maximum against lectin binding (anti-Strep Oyster 645) for the 0.1-µg hItln1/1 × 10^6 cells condition. Data are representative of two independent experiments.

We postulated that individual microbial species within a community might compete for lectin binding. Such competition would be especially relevant for lectins like hItln1 that would encounter bacterial populations at mucosal barriers. We therefore generated a synthetic community from two binding species (*Escherichia fergusonii* and *Lactobacillus reuteri*) and two nonbinding species (*Bacteroides ovatus* and *Proteus penneri*) that were each distinguishable by their scattering properties (Fig. 1, A and B). In this mixed community of 50% hItln1 binders (Fig. 1C), the lectin bound to only 29.6% of cells. Even at higher hItln1 concentrations, 50% binding was never achieved (fig. S1A). The forward scatter (FSC) and autofluorescence (FL1-H; fluorescein isothiocyanate) profile of the hItln1-bound population matches that of *L. reuteri* (Fig. 1, A to C, and fig. S1, B to D), indicating that *L. reuteri* outcompetes *E. fergusonii* for hItln1 binding in this simple mixed community. These data show that that species within microbial mixtures with high avidities for a given lectin can outcompete those with weaker binding. Thus, understanding human lectin function requires assessing recognition within diverse communities. These findings provide impetus to profile lectin binding to bacterial species within native communities.

We next characterized the microbe binding of two lectins with nonoverlapping glycan specificities by assessing their interactions with members of native microbial communities. The human gut harbors the most complex and diverse microbial community associated with the human body, and snapshots of the community are readily accessible through stool (*20*). We queried binding by hItln1, as its expression in the gut (*19*), its biochemical affinity for glycans exclusively found on microbes (*14*, *21*), and our initial binding data (Fig. 1) all suggest that it may preferentially interact with gut-resident microbes. In addition, *hITLN1* is associated with diseases linked to microbiome dysbiosis, suggesting that it may regulate gut microbiota (*22*, *23*). However, interactions between hItln1 and intestinal microbial communities have not been assessed, and its complete microbial binding repertoire is not known.

To contrast with our studies of hItln1, we evaluated the specificity of human mannose-binding lectin (MBL) for human gut microbial communities. Decades of research show that MBL binding promotes pathogen clearance (*24*–*26*), supporting a role in immune defense. A deficiency of MBL in murine models causes changes in the gut microbiota (*12*), but whether these changes are due to direct MBL-microbe interactions is unclear. *MBL* is primarily expressed in the liver, but low levels of expression have been detected in the intestine (*27*). While an understanding of the direct interactions between MBL and gut-resident microbes is lacking, studies applying the MBL carbohydrate recognition domain suggest that MBL binds an extremely broad range of Gram-negative and Gram-positive pathogenic bacteria isolates (97 of 112 isolates tested) (*28*). We therefore hypothesized that MBL would serve as a positive control as a lectin that was expected to show binding to a broad range of gut microbes.

We used flow cytometry to analyze the binding of recombinant MBL and hItln1 to stool homogenates collected from healthy human donors (Fig. 2A). Calcium was added, as both lectins require it for glycan binding (*29*, *30*). We assessed MBL and hItln1 at a range of concentrations (fig. S2, A and B) and found each bound a substantial population of intact microbial cells [Fig. 2, B (middle column) and C]. MBL binding was saturated at ~20 μg/ml, where it bound 10 to 20% of cells in stool homogenate [Fig. 2, B (middle column) and C, and fig. S2B]. HItln1 interacted with a comparable proportion of cells at 10 and 20 μg/ml [Fig. 2, B (middle column) and C, and fig. S2, B and C]; therefore, all subsequent experiments used this lectin concentration. The addition of the calcium chelator ethylenediaminetetraacetic acid (EDTA) reduced binding [Fig. 2B (right column)]. Lectin-specific inhibitors, mannan for MBL (*5*) or glycerol for hItln1 (*14*), also were inhibitory (fig. S3, A and B). These findings indicate that the detected microbial binding interactions are glycan dependent. The addition of recombinant human lectins is a critical aspect of this analysis. We did not detect notable levels of endogenous MBL and hItln1 on the tested human samples (fig. S4, A and B). These oligomeric lectins bind through multivalent interactions, which are kinetically labile. Thus, the lectins should be washed away during the stool homogenization process. A recent study demonstrated that interactions between murine Itln1 and stool microbiota are tenuous, and special handling is required to preserve these interactions during the harvesting and processing of stool (*31*). Consistent with these data, we found that application of the stool homogenization conditions to lectin-treated samples reduced the binding of recombinant MBL or hItln1 (fig. S4C). Together, the data indicate that MBL and hItln1 bind cells through glycan-dependent interactions and engage a similar proportion of cells within the gut microbiome.

Our initial data indicate that MBL is not wildly promiscuous but rather exhibits a similar level of specificity for microbes as hItln1. MBL and hItln1 bind orthogonal sets of glycans on a microbial glycan array (Fig. 2D and table S2). HItln1 enriched for glycans with 3-deoxy-D-*manno*-oct-2-ulosonic acid (KDO), D-*glycero*-D-*talo*-oct-2-ulosonic acid (KO), D-glycerol 1-phosphate, and galactofuranose residues (table S2) (*14*, *21*), epitopes that are absent from mammalian glycomes but widely distributed in bacteria (*32*). In contrast, MBL preferentially bound glycans containing mannose and *N*-acetyl glucosamine residues, and a recent study suggests that 10 to 30% of murine stool microbiota display mannose or *N*-acetylglucosamine (*33*). Despite the similar levels of bacterial engagement (Fig. 2C), each lectin′s orthogonal glycan recognition properties should translate to specificity for different microbes. To test this possibility, we coincubated fluorescently labeled MBL and hItln1 with healthy human stool homogenate and analyzed binding by flow cytometry. Most lectin-bound cells were bound by either MBL or hItln1 with less than 5% binding both (Fig. 2, E and F). Thus, MBL and hItln1 bind distinct bacterial populations.

Our next objective was to identify the microbial taxa recognized by each lectin. To maximize microbial diversity in our analyses, we included stool homogenate from five healthy and nine clinically diagnosed inflammatory bowel disease (IBD) patient donors. Although the composition of healthy and IBD microbiomes differ (*34*), similar levels of lectin binding occurred between healthy and diseased samples (fig. S5A). Stool homogenates from these 14 donors were stained with MBL or hItln1, and the lectin-bound and unbound cells were isolated by fluorescence-activated cell sorting (FACS). The sorted fractions were analyzed by metagenomic sequencing and assigned species-level taxonomy using MetaPhlAn2.0 (*35*) to determine the relative abundance of microbial species in each fraction. Last, we calculated a probability ratio (P_R) (*36*) for each species to determine whether it was enriched or depleted in the lectin-bound fraction. We refer to this workflow as Lectin-Seq (Fig. 3).

The microbe-binding properties of MBL were remarkably specific. Lectin-Seq revealed that MBL-bound fractions were predominantly composed of *Faecalibacterium prausnitzii* (Fig. 4A, fig. S6, and table S4). *F. prausnitzii* was present in the MBL-bound fractions of 12 of 14 stool samples and, on average, accounted for >40% relative abundance across all MBL-bound fractions (Fig. 4, A and B, and fig. S6). In the two remaining samples, we did not detect *F. prausnitzii* in the input fraction (Fig. 4B). In stool samples that contain *F. prausnitzii*, the microbe was enriched in the MBL-bound fraction of 8 of 12 samples, as indicated by a positive P_R (Fig. 4C). In the remaining four samples, *F. prausnitzii* had a negative P_R close to zero, indicating that this microbe was present at similar levels in the MBL-bound and unbound fraction (Fig. 4C).

F. prausnitzii is an abundant member of the gut microbiota; in the samples analyzed, the average percent relative abundance was ~12% (table S4). To determine whether MBL recognition correlates with species abundance, we asked if the MBL-bound fraction was enriched for other abundant species. Of the top 10 most abundant species (average abundance in input fraction; fig. S7A), only *F. prausnitzii* had an average $P_R > 0$ for the MBL-bound fraction, indicating that MBL binding does not track with species abundance (Fig. 4D). MBL enriched for an additional 35 lower relative

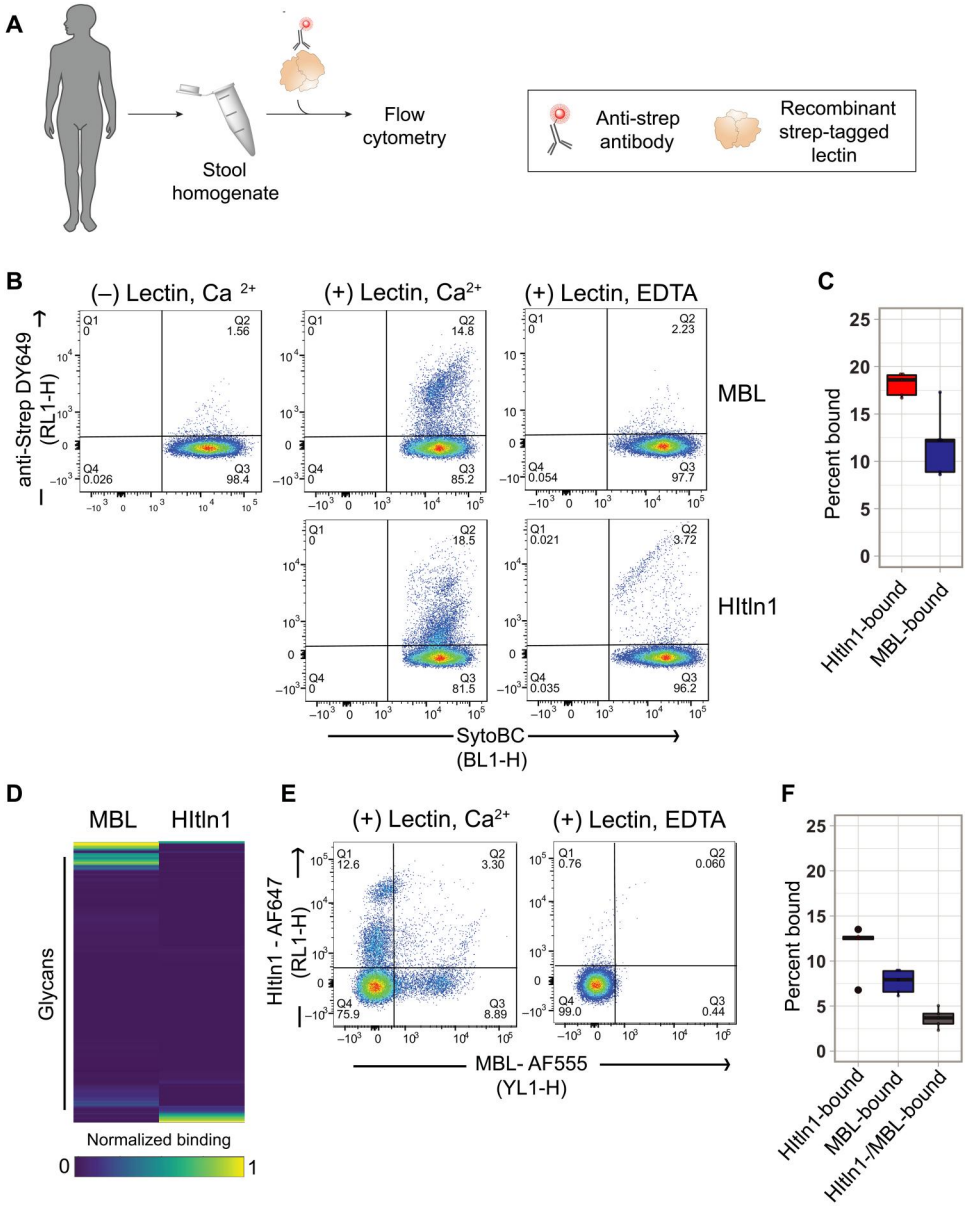

Fig. 2. MBL and hItln1 binding to microbial populations in human stool samples. (**A**) Schematic of flow cytometry–based assay to measure soluble lectin binding to stool microbiome. (**B**) Flow cytometry analysis of healthy human stool plotted as SYTO BC against lectin binding (anti-Strep DY649). Panels show no lectin control (column 1), and strep-MBL (top row), or strep-hItln1 (bottom row) binding in the presence of Ca^{2+} (column 2) or EDTA (column 3). (**C**) Quantification of data shown in (A). Percent bound is calculated by determining the percentage of cells that fall in the strep-pos/SYTO BC-pos fraction (Q2). Data shown (box plots represent the means and SD; reported values are means and SEM) are from the (+) strep-hItln1/Ca^{2+} (red; n = 5; 18.1 ± 0.5%) and (+) strep-MBL/Ca^{2+} (blue; n = 5; 11.9 ± 1.6%) samples. (**D**) Heatmap depicting normalized binding of MBL and hItln1 to microbial glycan array. Each row represents binding to a single glycan (mean of three spots). Rows were grouped by hierarchical clustering (Euclidean). (**E**) Flow cytometry analysis of the stool microbiome to determine differential hItln1 and MBL binding. Flow cytometry data are plotted as hItln1-binding (AF-647) against MBL-binding (AF-555). Panels show lectin binding in the presence of (+) lectin/Ca^{2+} (panel 1) or (+) lectin/EDTA (panel 2). (**F**) Quantification of data (box plots represent the means and SD; reported values are means and SEM) shown in (E), panel 2. Percent bound is calculated by determining the percentage of cells that fall in the hItln1-pos/MBL-neg (red; n = 5; 11.6 ± 1.2%), hItln1-neg/MBL-pos (blue; n = 5; 7.7 ± 0.6%), and hItln1-pos/MBL-pos fractions (dark gray; n = 5; 3.6 ± 0.5%).

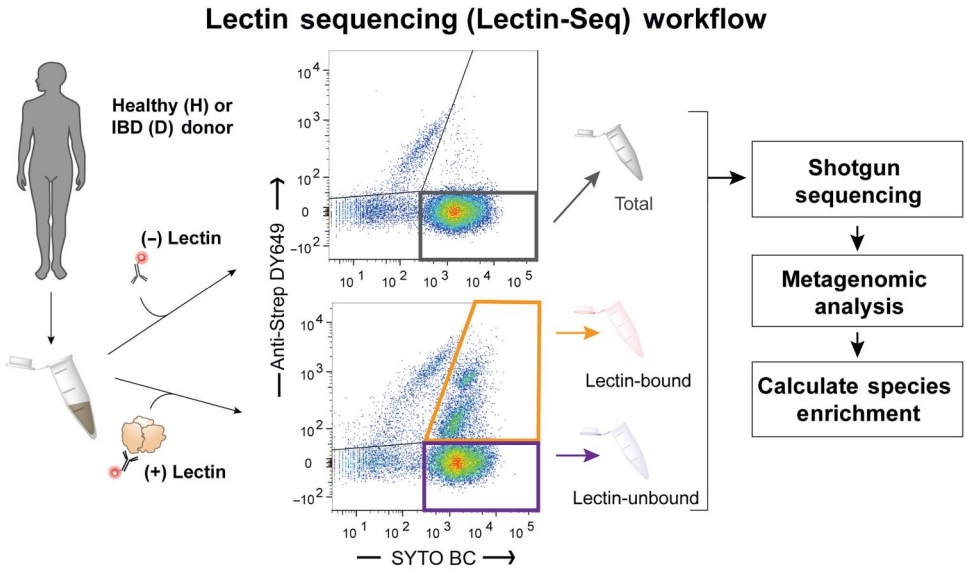

Fig. 3. Lectin-Seq schematic. Lectin-Seq workflow for identifying lectin-bound microbes from human stool samples.

abundance species (average input abundance, <0.01); however, more than half of these species occur in only one stool sample (table S4 and fig. S8A). These data suggest that MBL exhibits an unexpected level of specificity for distinct species.

Only two species apart from *F. prausnitzii* were enriched by MBL in least five samples: *Akkermansia muciniphila* and *Ruminococcus lactaris* (Fig. 4A, table S4, and fig. S8A). Notably, MBL binding to *A. muciniphila* was detected only in IBD-diagnosed donor samples including the two samples that lacked *F. prausnitzii* (Fig. 4A and fig. S7, B and C). *F. prausnitzii* levels are reduced in patients with IBD (*37*). Thus, when MBL′s preferred binding partner is largely absent, it engages other species. In addition, *A. muciniphila* associated with IBD may have an altered glycan display that leads to differential MBL binding. Microbes can modulate their glycan displays under different growth conditions (*15*, *16*), and *A. muciniphila* genomes are enriched for genes, including putative GT2 and GT4 glycosyltransferases (*38*), that are capable of dynamically controlling the display of mannose-containing glycans. We also note that *A. muciniphila* isolates from different individuals are genetically and phenotypically diverse (*38*, *39*).

MBL and hItln1 bound a similar absolute proportion of cells in samples, but the taxa bound by hItln1 and MBL had little overlap. MBL preferred the single species *F. prausnitzii*, but hItln1-bound fractions consisted of more diverse populations (figs. S6 and S8B and table S4). No single species accounted for >15% of the average percent relative abundance in hItln1-bound fractions (Fig. 4E, figs. S6 and S8B, and table S4), and the targeted species were not found in most of the enriched fractions. Nine species were found in five or more hItln1-bound fractions; however, the P_R values for these species varied across samples, indicating that enrichment by hItln1 is donor dependent (Fig. 4F and fig. S8B). Four species—*A. muciniphila*, *Bifidobacterium longum*, *Ruthenibacterium lactatiformans*, and *Coprococcus comes*—had an average positive P_R across all samples (Fig. 4F). These four species each made up less than 2% of the input across all samples and were not detected in the input fractions of most donor samples (Fig. 4G). Thus, while both lectins engage a large percentage of the commensal population, hItln1 binds to low-abundance species within the human microbiome, while MBL prefers *F. prausnitzii*, a common and prevalent commensal.

To compare Lectin-Seq to in vitro strain binding, we tested MBL and hItln1 binding to select isolates of highly enriched species. MBL bound all tested strains of *A. muciniphila* and *F. prausnitzii*, supporting the robust enrichment of the species detected in human samples (Fig. 5A). We also verified that MBL and hItln1 did not bind isolates of species that were depleted from the respective lectin-bound fractions in stool homogenate (fig. S9). However, we observed some strain-specific interactions. For example, MBL did not bind any of the tested isolates of *R. lactaris* (Fig. 5A) despite its ability to enrich for this species from human samples. This discrepancy may be due to differences between the *R. lactaris* glycan displays in culture and in vivo. Alternately, interactions between *R. lactaris* and MBL may require unknown factors present in stool homogenate that are not present in culture such as bacteria-associated dietary polysaccharides or host immune factors. Lectin strain specificity was even more pronounced for hItln1, which bound two of three isolates of *B. longum*, one of two isolates of *A. muciniphila*, and no isolates of *C. comes* (Fig. 5B). Because strain identity varies across individual microbiota (*40*), the donor-dependent enrichment of these species by hItln1 is likely due to strain-specific lectin interactions. These findings highlight the importance of evaluating lectin binding to endogenous bacterial populations. Overall, the binding data support our findings from human samples and provide a mechanistic explanation for the donor-dependent interaction profile of hItln1.

To identify the molecular targets of MBL on the surface of microbial isolates, we profiled the glycan displays of isolates of *F. prausnitzii* and *A. muciniphila* isolates using plant-derived lectins with known glycan specificities. Most plant lectins had little ability to bind to the microbes enriched by the human lectins. Still, *F. prausnitzii* was robustly bound by concanavalin A (ConA) and wheat germ agglutinin (WGA; fig. S10A and table S7), while *A.*

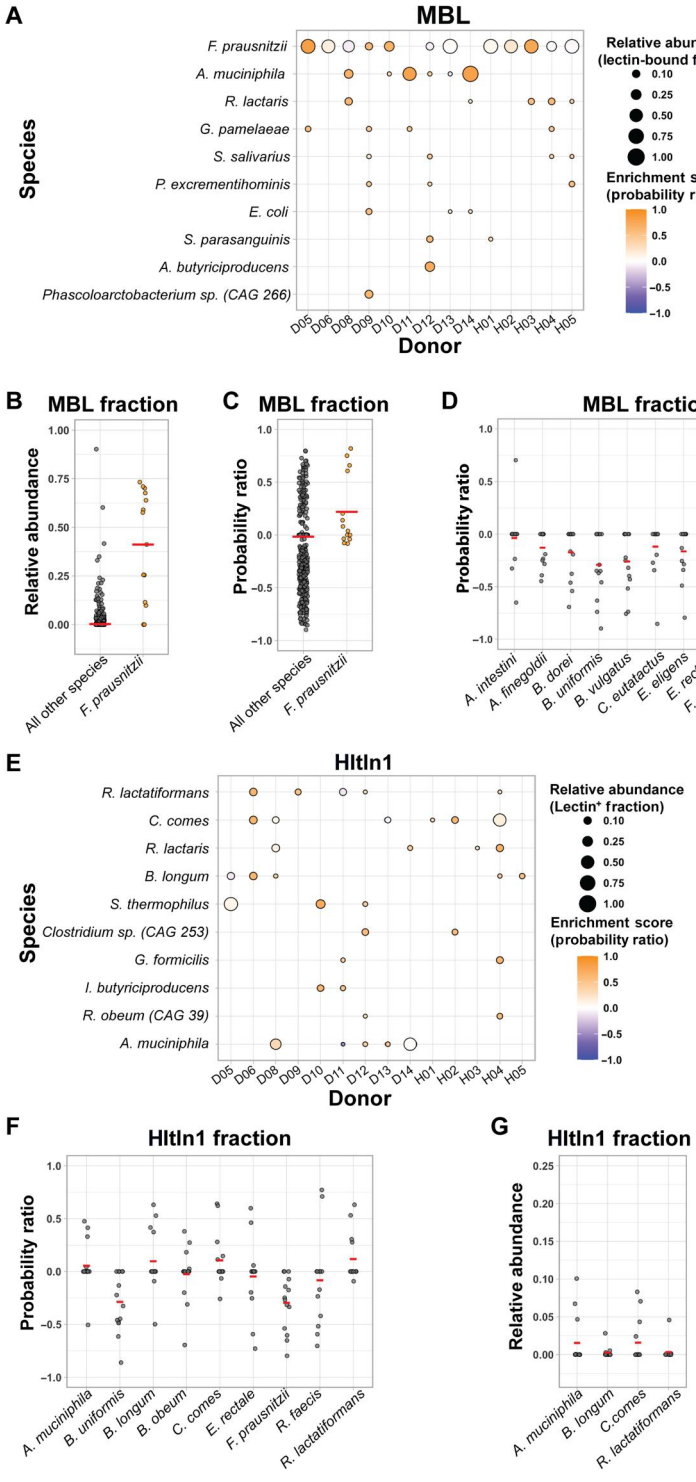

Fig. 4. Identification and quantification of MBL and hltln1 bound bacterial species from human stool samples. (**A** and **E**) "Bubble plot" depicting the enrichment of top enriched bacterial species in the lectin-bound fraction (defined as the probability ratio) donor stool samples bound by either MBL (A) or hltln1 (E). Panels show top enriched bacterial species as calculated by average probability ratio. Probability ratio of each species is represented by the color of each dot, and fractional abundance is represented by the size of the dot. Data shown are from 14 donor stool homogenate samples (9 IBD-diagnosed labeled D05, D06, and D08-13) and 5 healthy donors (labeled sH01-05). (**B** and **C**) Dot plot of the (B) relative abundance or (C) probability ratio of *F. prausnitzii* or all other species in the lectin$^+$ fraction (MBL$^+$ fraction; $n = 14$). Red line denotes the mean. (**D**) Dot plot of the probability ratio of the top 10 most abundant species (input fraction; $n = 14$) in the MBL$^+$ fraction. Red line denotes the mean. (**F**) Dot plot of the probability ratio of species bound in more than five hltln1$^+$ fractions across all 14 donor stool homogenates. Red line denotes the mean. (**G**) Dot plot of the relative abundance (input fraction; $n = 14$) of species with an average $P_R > 0$ (hltln1$^+$ fraction; $n = 14$). Red line denotes the mean.

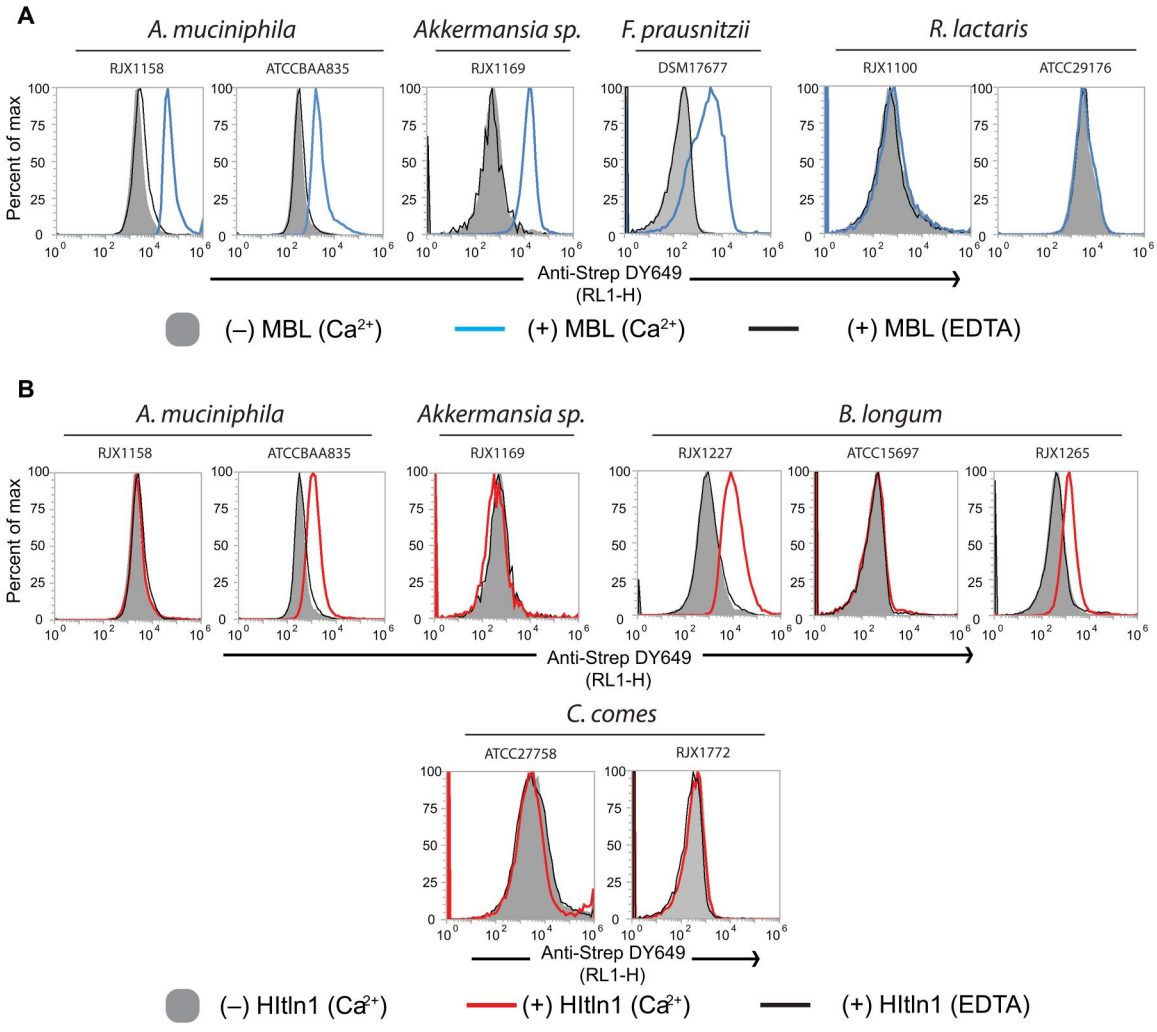

Fig. 5. MBL- and hItIn1-binding to bacterial isolates. (**A** and **B**) Summarized flow cytometry analyses of MBL (A) or hItIn1 (B) binding to bacterial isolates. Isolates are grouped by species. Individual histograms plot cell counts as a percent of the maximum signal against lectin binding (anti-Strep DY649). Each plot shows data from (–) lectin (Ca^{2+}; solid gray), (+) MBL/Ca^{2+} (blue trace) or (+) hItIn1/Ca^{2+} (red trace), and (+) lectin/EDTA (black trace) binding conditions. Data are representative of two independent experiments.

muciniphila was bound by ConA (fig. S10B and table S7). The known glycan specificity of ConA is for mannose and glucose, while WGA has specificity for *N*-acetylglucosamine (table S7). Because MBL binds glycans having mannose and *N*-acetylglucosamine residues, we tested the ability of ConA and WGA to inhibit MBL binding to *F. prausnitzii* (fig. S10, C and D). The plant lectins were effective competitors, suggesting that the human lectin binds the same *F. prausnitzii* glycans. These results are further supported by a recent genome analysis of 84 *F. prausnitzii* strains, which found that all genomes analyzed contain genes from the GT94 glycosyltransferase family (*41*). GT94 is composed of a mannosyltransferase that catalyzes the addition of alpha-linked mannose residues to xanthan chains in exopolysaccharide (*42*). Together, these findings provide a molecular basis for MBL binding specificity in native microbial communities.

DISCUSSION

Historically, the physiological roles of soluble lectins are ascribed to pathogen recognition and clearance (*6*), but emerging data suggest a role for murine lectins in regulating gut microbiota (*9*–*12*). Whether the glycan-dependent interactomes of human lectins encompass fungal species in the native mycobiome is not clear. We did not detect substantial levels of fungal DNA in our samples (table S4); however, fungi and bacteria may compete for lectin interactions in vivo. We did, however, find that human lectins bind selectively to distinct bacterial species in native communities.

Lectin-Seq showed that MBL robustly binds *F. prausnitzii*, an abundant member of the gut microbiome. We observed some, albeit less, binding to *A. muciniphila* and *R. lactaris*; the former was present especially in disease samples lacking *F. prausnitzii*. The specificity of MBL binding to mutualist and commensal species is unprecedented and contrasts with previous data indicating that MBL exhibits relatively indiscriminate binding to a broad range of pathogenic microbes, including *Escherichia coli* and

Enterococcus faecalis, two abundant gut microbes (*43*). This observation aligns with our synthetic community results (Fig. 1) that lectins show greater specificity in a community setting, such as the highly complex community found in the gut microbiome, compared to a single isolate binding experiment. Such specificity suggests that MBL could be co-opted as a tool to rapidly quantify the abundance of these species in uncharacterized stool samples.

The selectivity of MBL for host-resident bacteria suggests that its biological functions go beyond pathogen surveillance. *MBL* is expressed at low levels in the intestine (*27*). *Akkermansia* levels are decreased in MBL null mice (*12*), but whether this decrease is due to direct interactions with murine MBL or if this finding translates to humans was unclear. Our finding that MBL binds to *F. prausnitzii* suggests that it provides support for direct bacterial interactions as a part of its influence on the microbiome. In the blood, MBL binding to pathogens results in their clearance (*24*); therefore, in healthy individuals, MBL is poised to eliminate specific gut-resident microbes that breach epithelial barriers. In patients diagnosed with IBD or colorectal cancer (CRC), a reduction in barrier integrity can lead to blood, and therefore high levels of MBL, entering different regions of the gut. *F. prausnitzii* levels are broadly reduced in patients with IBD (*37*) and, more specifically, in patients with CRC who report blood in their stool (*44*). The binding of MBL to *F. prausnitzii* is consistent with these observations.

Lectin-Seq revealed a notable difference between MBL and hItln1. The latter does not bind a high-abundance species but rather exhibits donor-dependent binding to a broad group of low-abundance microbes. HItln1-bound microbes from the stool microbiome are highly diverse: The population includes Gram-positive and Gram-negative species capable of critical metabolic activities such as short-chain fatty acid production and mucin degradation. Unlike some intestinal lectins (*11*), hItln1 has no known microbial toxicity and, unlike MBL, lacks canonical domains for complement recruitment (*14*). The limited literature on hItln1 function suggests a range of innate immune roles, some of which are conflicting. Notably, none of these considers the native environment nor the bacterial recognition profile of the lectin. For example, the interaction of hItln1 with select pathogens activates macrophage- and neutrophil-mediated killing (*17*, *18*, *45*); however, hItln1-mediated neutrophil activation requires serum components, suggesting that the bactericidal functions of hItln1 would be limited to the blood where it is present at low concentrations. Alternatively, hItln1 is reported to suppress lipopolysaccharide-induced macrophage activation (*46*, *47*). Moreover, if we consider our dataset of hItln1 binders, *A. muciniphila* is an important mucin degrader linked to human health. Murine Itln1 has also been shown to bind to *A. muciniphila* in mouse stool samples (*31*). This microbe colonizes the human intestine in the first year of life (*48*), which is when *hITLN1* expression in the intestine is at its highest levels. Also identified as a top-binder in our hItln1 dataset, and abundant in the healthy infant microbiome is *B. longum* (*49*). Thus, our data are difficult to reconcile with a model in which hItln1 functions by eliminating these microbes.

Our data provide insight into the interplay of lectin recognition and microbial glycans. A major finding from our study is that soluble human lectins have notably disparate microbial interaction profiles and that binding is highly specific within a community context. MBL is highly specific for a prevalent human gut commensal, while hItln1 binds a diverse group of low-abundance species. Furthermore, the orthogonal glycan specificities of these lectins reflect the importance of microbial glycans for species-specific host-microbe interactions. Still unclear is whether MBL and hItln1 have evolved to recognize ligands differentially displayed across the microbiota or if specific members of the microbiota have evolved mechanisms, allowing them to adjust their glycan displays to engage or evade these lectins. Future efforts to identify the microbial genetic determinants of lectin binding will inform the mechanistic basis and evolutionary dynamics of these host-microbe interactions. Our data highlight the utility of Lectin-Seq for uncovering lectin specificity and the need to broadly characterize the interactomes of soluble lectins in native microbial communities.

MATERIALS AND METHODS

Protein expression, purification, and Western blotting

Cloning of full-length N-terminal strepII-tagged hItln1 into the pcDNA4/myc-HisA vector backbone (Life Technologies) was described previously (*14*). Human MBL (*MBL2*, NM_000242) gBlocks DNA Fragment (Integrated DNA Technologies) was Gibson ligated into a linearized pcDNA4 plasmid [primers reverse (5′-GGTGAAGCTTAAGTTTAAACG-3′) and forward (5′-TGAGGATCCACTAGTCCAGTG-3′)]. StrepII tag (5′-TGGAGCCATCCGCAGTTTGAAAAG-3′) was inserted C-terminal of amino acid 20 into pcDNA4-hMBL2 using inverse polymerase chain reaction (PCR) mutagenesis using forward primer (5′-TGGAGCCATCCGCAGTTTGAAAAGGAAACTGTGACCTGTGAGG-3′) and reverse primer (5′-TTCTGAGTAAGACGCTGCC-3′). Correct insertion was verified by DNA sequencing (Quintara Biosciences).

HItln1 and MBL were each expressed by transient transfection of suspension-adapted human embryonic kidney (HEK) 293T cells. Cells were transfected at 1.8×10^6 cells/ml in growth medium [Dulbecco's modified Eagle's medium (Thermo Fisher Scientific, catalog no. 11995) supplemented with 10% heat-inactivated fetal bovine serum (FBS), penicillin-streptomycin (50 U/ml), 4 mM L-glutamine, and 1× nonessential amino acids] using Lipofectamine 2000 (Thermo Fisher Scientific) following the manufacturer's protocol. Six hours after the transfection, culture medium was exchanged to FreeStyle F17 expression medium (Thermo Fisher Scientific) supplemented with penicillin-streptomycin (50 U/ml), 4 mM L-glutamine, 1× nonessential amino acids, 0.1% heat-inactivated FBS, and 0.1% Pluronic F-68 (Thermo Fisher Scientific). Transiently transfected cells were cultured up to 4 days or until viability was below 60%. The conditioned expression medium was then harvested by centrifugation and sterile filtration.

For purification, $CaCl_2$ was added to the harvested conditioned expression medium to a final concentration of 10 mM, and avidin (7 mg/ml) was added at 12 μl per ml of expression media (IBA, catalog no. 2-0204-015; per the IBA protocol). Protein was captured onto 2 ml of Strep-Tactin Superflow High-Capacity resin (IBA Lifesciences, catalog no. 2-1208-002) equilibrated with Hepes-Ca buffer [20 mM Hepes (pH 7.4), 150 mM NaCl, and 10 mM $CaCl_2$]. The resin was then washed with Hepes-EDTA buffer [20 mM Hepes (pH 7.4), 150 mM NaCl, and 1 mM EDTA] and eluted with 5 mM *d*-desthiobiotin (Sigma-Aldrich) in Hepes-EDTA buffer. StrepII-tagged hItln1 was concentrated with a 30,000–molecular weight cutoff (MWCO) Amicon Ultra centrifugal filter. StrepII-

tagged MBL was concentrated with a 10,000-MWCO Vivaspin 6 centrifugal filter (GE). All proteins were buffer exchanged to Hepes-EDTA buffer for storage. Protein concentrations were determined by absorbance at 280 nm. Extinction coefficients and molecular weights were calculated for the mature trimeric product of each protein (without the signal peptide) using the ProtParam tool (web.expasy.org/protparam). StrepII-tagged hItln1 had a calculated ε = 239,775 cm^{-1} M^{-1} and an estimated molecular mass of 102,024 Da. StrepII-tagged MBL had a calculated ε = 71,595 cm^{-1} M^{-1} and an estimated molecular mass of 75,570 Da. Recombinant hItln1 and MBL were characterized by Western blotting using monoclonal mouse anti-MBL2 antibody (3E7; 1:50; Thermo Fisher Scientific, MA1-40145) and polyclonal rabbit anti-hItln1 antibody (1:2000; Proteintech, 11770-1-AP).

Culturing of microbial isolates

Bacterial isolates along with their respective media used in this study are detailed in tables S1 and S6. All isolates were struck out from frozen stock onto solid medium in an anaerobic chamber and incubated anaerobically at 37°C for 24 to 72 hours (isolate dependent). Liquid Mega Medium (isolates in table S1) was prepared and reduced in a Coy Laboratory anaerobic chamber (5% H_2, 20% CO_2, and 75% N_2) as previously described (*50*). Bacteria in table S6 were grown in CHG medium [brain heart infusion (37 g $liter^{-1}$, Sigma-Aldrich) supplemented with 5% sterile-filtered FBS (Sigma-Aldrich), 1% vitamin K1–hemin solution (BD Biosciences), 1% trace mineral supplement (American Type Culture Collection), 1% trace vitamin supplement (American Type Culture Collection), 1 g $liter^{-1}$ D-(+)-cellobiose (Sigma-Aldrich), 1 g $liter^{-1}$ D-(+)-maltose (Sigma-Aldrich), 1 g $liter^{-1}$ D-(+)-fructose (Sigma-Aldrich), and 0.5 g $liter^{-1}$ L-cysteine (Sigma-Aldrich), or YCFAC medium (Yeast Casitone Fatty Acids Agar with Carbohydrates, Anaerobe Systems)] (*51*). Liquid cultures were established for each isolate through the inoculation of a single colony into 5 to 10 ml of media. Upon the observation of turbid growth (16 to 72 hours), liquid cultures were combined with a 1× phosphate-buffered saline (PBS) + 40% glycerol + 0.001% L-cysteine solution at a 1:1 ratio, aliquoted into a series of cryovials, and stored at −80°C before flow cytometry. A portion of each culture was subjected to 16*S* ribosomal DNA V1-V9 PCR amplification, spin column purification, and Sanger sequencing, with taxonomic verification via standard nucleotide BLAST via BLASTN suite software (National Institutes of Health, National Library of Medicine; https://blast.ncbi.nlm.nih.gov/Blast.cgi).

Flow cytometry of microbial isolates

All buffers were sterile-filtered before use, and all centrifugation steps were performed at 5000 relative centrifugal force (RCF) at 4°C. Microbial isolates were screened for viability by propidium iodide staining by flow cytometry before lectin binding. We observed differences in binding to viable and unviable cells; therefore, only viable isolates were used for validation. Microbial isolates washed twice with 2 ml of PBS-BSA-T [PBS (pH 7.4; Gibco), 0.1% (w/v) bovine serum albumin (US Biological, A1311), and 0.05% (v/v) Tween 20 (Sigma-Aldrich)], pelleted, and resuspended in PBS-BSA-T. Staining was performed at optical density at 600 nm (OD_{600}) of 0.2 for all samples. Staining solutions were Hepes-Ca-BSA-T with SYTO BC (1:1000; Thermo Fisher Scientific) and StrepMAB classic DY649-conjugate (1:150; IBA Lifesciences) for the unstained samples. All lectin-stained Ca^{2+} samples were stained as follows: Hepes-Ca-BSA-T, SYTO BC (1:1000; Fig. 5 and fig. S5), StrepMAB classic Oyster 645 (1:250; Fig. 1 and table S1) or StrepMAB classic DY649-conjugate (1:250; Fig. 5 and fig. S5), and recombinant StrepII-tagged lectin (20 μg/ml). For the lectin-stained EDTA control samples, the staining conditions were as follows: Hepes-EDTA-BSA-T [20 mM Hepes (pH 7.4), 150 mM NaCl, 1 mM EDTA, 0.1% BSA, 0.05% Tween 20], SYTO BC (1:1000), StrepMAB classic DY649-conjugate (1:250), and recombinant StrepII-tagged lectin (20 μg/ml). Staining was performed at 4°C for 1 hour before being diluted four times without washing and analyzed on a BD Accuri C6 flow cytometer (Fig. 1 and table S1) or Thermo Fisher Scientific Attune flow cytometer (Fig. 5 and fig. S6).

For plant lectin staining, microbial isolates were prepared as described above before staining. Biotinylated plant lectins were purchased from Vector Laboratories. All plant lectin-stained samples were stained as follows: Hepes-Ca-BSA-T, SYTO BC (1:1000), Strep-Tactin XT DY649 (1:250), and biotinylated plant lectin (20 μg/ml). Staining was performed at 4°C for 1 hour before being diluted four times without washing and analyzed on a Thermo Fisher Scientific Attune flow cytometer.

For competition assays, microbial isolates were prepared as described above before staining. Unlabeled plant lectins were purchased from Vector Laboratories. All competition assay samples were preincubated with indicated concentrations of unlabeled plant lectins at 4°C for 1 hour. Samples were then stained as follows: Hepes-Ca-BSA-T, SYTO BC (1:1000), StrepMAB classic DY649-conjugate (1:250), and StrepII-tagged MBL (20 μg/ml). Staining was performed at 4°C for 1 hour before being diluted four times without washing and analyzed on a Thermo Fisher Scientific Attune flow cytometer.

HItln1 binding to synthetic mixed community

Strains of bacteria were prepared as described above. Cell density was quantified on a BD Accuri C6 flow cytometer, and equal numbers of cells from each species were combined into a single sample at a final cell concentration of 5×10^7 to 1×10^8 cells/ml. Indicated amounts of hItln1 were added per 1×10^6 cells in the final sample (final concentrations ranged from 5 to 50 μg/ml). Cells were stained and analyzed as described above.

Direct labeling of proteins with fluorophores

Recombinant StrepII-tagged hItln1 and MBL were directly conjugated with fluorophores using NHS-ester conjugation chemistry. HItln1 was labeled with Alexa Fluor 647 NHS Ester (Succinimidyl Ester) (Thermo Fisher Scientific), and MBL was labeled with Alexa Fluor 555 NHS Ester (Succinimidyl Ester) (Thermo Fisher Scientific) following the manufacturer's protocol as follows: Proteins were labeled in storage buffer [Hepes-EDTA (pH 7.4) as described above] by the addition of dye dissolved in dimethyl sulfoxide at a ratio of 1:4 (w:w, dye:protein) and incubation for 1 hour at room temperature with mixing. Unreacted dye was removed by desalting using a Zeba 2 ml 7 K MWCO Spin Desalting Column (Thermo Fisher Scientific) following the manufacturer's protocol. Molar ratio of fluorophore: Protein was determined from ultraviolet-visible absorbance spectrum using the extinction coefficients: ε (Alexa Fluor 555) = 155,000 cm^{-1} M^{-1} and ε (Alexa Fluor 647) = 270,000 cm^{-1} M^{-1}.

Glycan array analysis

Microbial glycan microarrays (MGM) were synthesized as previously described (*14*, *52*) and are commercially available from the National Center for Functional Glycomics. For each lectin, an MGM was thawed at room temperature for 30 min. The slide was placed in a slide holder tube and soaked in wash buffer [20 mM Hepes (pH 7.5) and 150 mM NaCl] for 10 min. The slide was placed onto a raised slide holder platform in a humidification chamber and a solution of StrepII-tagged lectin at a concentration of 10 μg/ml in binding buffer [20 mM Hepes (pH 7.5), 150 mM NaCl, 10 mM $CaCl_2$, 0.1% BSA, and 0.1% Tween 20]. The slide was incubated with gentle rocking at room temperature for 1 hour. The slide was then washed with binding buffer, wash buffer, and ddH_2O (four times each). The slide was placed back into the humidification chamber and incubated with DY549-conjugated anti-Strep monoclonal antibody at 4 μg/ml in binding buffer at room temperature for 1 hour. The slide was washed again with binding buffer, wash buffer, and ddH_2O (four times each). The slide was dried with a slide spinner and scanned using a GenePix 4000 (Molecular Devices) using the 532-nm laser. Raw data were averaged across technical replicates and plotted against sample ID numbers.

Approval for human sample research

Human stool samples were obtained from the Prospective Registry in IBD Study at MGH (PRISM) cohort, which was reviewed and approved by the Partners Human Research Committee (ref. 2004-P-001067). Stool samples from healthy subjects were obtained under protocol approved by the institutional review board at Massachusetts Institute of Technology [Institutional Review Board (IRB) protocol ID no. 1510271631]. Participants under both protocols provided informed consent, and all experiments adhered to the regulations of the review boards.

Preparation of human stool samples

Healthy human stool samples from IRB protocol ID no. 1510271631 were prepared as follows: Healthy subjects stool samples were brought into the anaerobic chamber within 2 hours of donation. Samples were homogenized using 1× PBS (pH 7.4), 0.1% L-cysteine at a ratio of 1 g of stool:2.5 ml of PBS. Glycerol (25% in 1X PBS and 0.1% L-cysteine) was added to the homogenate to a concentration of 12.5%, giving a final solution of 1 g of stool in 5 ml of 1× PBS, 0.1% L-cysteine, and 12.5% glycerol. The homogenate was removed from the anaerobic chamber and stored at −80°C before FACS analysis.

Fresh stool samples from adult donors with inflammatory bowel disease from the PRISM cohort were collected in anaerobic transport medium (Anaerobic Systems, AS-690), mailed at ambient temperature, and processed anaerobically within 24 to 72 hours of collection. In the anaerobic chamber, samples were decanted into sterile gentleMACS C tubes (Miltenyi Biotec, 130-093-237) and topped up to a total volume of 10 ml using 1× PBS + 40% glycerol + 0.001% L-cysteine solution. Stool was homogenized via gentleMACS Dissociator (Miltenyi Biotec, 130-093-235) for 3 cycles of 61 s on the "intestine" setting, aliquoted into a series of 2.0 ml of cryovials (Corning, 430659), and stored at −80°C before FACS analysis.

Flow cytometry and FACS of donor stool samples

All buffers were sterile-filtered before use, and all centrifugation steps were performed at 5000 RCF at 4°C. Stool homogenate was thawed on ice, and 200 μl of homogenate was washed twice with 2 ml of PBS-BSA-T [PBS (pH 7.4; Gibco), 0.1% (w/v) BSA (US Biological, A1311), and 0.05% (v/v) Tween 20 (Sigma-Aldrich)], pelleted, and resuspended in PBS-BSA-T. The material was then passed through a 35-μm cell strainer cap (Falcon), pelleted, and resuspended in 2 ml of Hepes-Ca-BSA-T [20 mM Hepes (pH 7.4), 150 mM NaCl, 10 mM $CaCl_2$, 0.1% BSA, and 0.05% Tween 20]. A 5-μl sample of the bacterial suspension was pelleted and frozen as the input sample for metagenomic sequencing. Staining was performed at OD_{600} of 0.2 for all samples. Staining solutions were Hepes-Ca-BSA-T with SYTO BC (1:1000; Thermo Fisher Scientific) and StrepMAB classic DY649-conjugate (1:250; IBA Lifesciences) for the unstained samples. All lectin-stained Ca^{2+} samples were stained as follows: Hepes-Ca-BSA-T, SYTO BC (1:1000), StrepMAB classic DY649-conjugate (1:250), and recombinant StrepII-tagged lectin (20 μg/ml) unless indicated otherwise. For the lectin-stained EDTA control samples, the staining conditions were as follows: Hepes-EDTA-BSA-T [20 mM Hepes (pH 7.4), 150 mM NaCl, 1 mM EDTA, 0.1% BSA, and 0.05% Tween 20], SYTO BC (1:1000), StrepMAB classic DY649-conjugate (1:250), and recombinant StrepII-tagged lectin (20 μg/ml) unless indicated otherwise. For the lectin-stained mannan and glycerol control samples, the staining conditions were as follows: Hepes-Ca-BSA-T, indicated concentrations of mannan (MilliporeSigma, M7504) or glycerol, SYTO BC (1:1000), StrepMAB classic DY649-conjugate (1:250), and recombinant StrepII-tagged lectin (20 μg/ml). For mock stool homogenization, samples were stained with lectin in Hepes-Ca-BSA-T as described above without SYTO BC and StrepMAB classic DY649-conjugate. Samples were then washed once with 1× PBS and 12.5% glycerol and thrice with 1× PBS. Washed samples were then stained with SYTO BC and StrepMAB classic DY649-conjugate as described above. Staining was performed at 4°C for 4 hours before being diluted 10× without washing and analyzed on a BD LSRII HTS flow cytometer with BD FACSDiva software or an Attune Flow Cytometer (Thermo Fisher Scientific) or diluted six times without washing and sorted by FACS on a FACSAria contained in a biosafety cabinet, with BD FACSDiva software. Stained samples were screened for viability by propidium iodide staining before lectin binding and sorting. Under our staining conditions, lectin binding did not alter cell viability (fig. S11). Cells were gated by forward scatter, side scatter, and SYTO BC staining and sorted into lectin-negative or lectin-positive fractions. For metagenomics, 1.5×10^6 cells were collected for the lectin-negative and lectin-positive samples. Sorted cells were centrifuged at 7000 RCF at 4°C for 5 min; supernatant was removed, and the cell pellet was stored at −20°C before nucleic acid extraction.

Stool samples were stained for endogenous hItln1 and MBL using monoclonal mouse anti-MBL2 antibody (3E7; 1:50; Thermo Fisher Scientific, MA1-40145) and polyclonal rabbit anti-hItln1 antibody (1:450; Proteintech, 11770-1-AP). Primary antibodies were detected using goat anti-Mouse AF555 (1:1000; Thermo Fisher Scientific, A-21422) and goat anti-Rabbit AF555 (1:1000; Thermo Fisher Scientific, A-31572). Stool samples were prepared as described above and stained in Hepes-Ca-BSA-T with SYTO BC (1:1000; Thermo Fisher Scientific) for 1 hour at 4°C and analyzed on an Attune Flow Cytometer (Thermo Fisher Scientific).

Flow cytometry on stool samples was performed at the The Swanson Biotechnology Center Flow Cytometry Facility housed in the Koch Institute (Massachusetts Institute of Technology, Cambridge, MA).

Nucleic acid extraction

Cell pellets were lysed by the HotSHOT lysis method (*53*, *54*). HotSHOT lysis offers advantages for workflows that deal with low volumes and cell counts; however, this method is inefficient at lysing certain classes of microbes. As a result, select species may be underrepresented across all samples. Notably, this should not affect comparisons between lectin$^+$ and lectin$^-$ fractions. Briefly, cell pellets were lysed by addition of 50 μl of HotSHOT lysis buffer [25 mM NaOH and 0.2 mM EDTA (pH 12)] and heating to 95°C for 10 min followed by addition of equal volume HotSHOT neutralization buffer (40 mM Tris-HCl, pH 5) and vortexed to combine. Lysed samples were centrifuged at 3000 RCF for 10 min to pellet debris and transferred to 96-well plates. The genomic DNA was then purified by AMPURE bead cleanup as follows: Ninety microliters of beads was added to 100 μl of DNA and incubated for 13 min at room temperature. Beads were separated on a magnet for 2 min, and supernatant was removed. Beads were washed twice with 200 μl of ethanol and air dried for 20 min while still on the magnet. The elution was performed by resuspending the beads thoroughly in 30 μl of H_2O and incubation for 7 min at room temperature. Beads were separated on a magnet, and 27 μl of eluted DNA was removed to a sterile plate.

Metagenomics library construction and sequencing

One microliters of purified DNA was directly used as input into a miniaturized version of the Nextera XT Library Preparation Kit (Illumina Inc.). All reactions were scaled to one-fourth their original volumes. Libraries were constructed according to the manufacturer's instructions with several modifications to accommodate low DNA concentrations. The amplicon tagmentation mix was diluted 1:10 in tagmentation DNA buffer to reduce the tagmentase:DNA ratio. Tagmentation time was also reduced from 5 min to 1 min. Both modifications were implemented to boost the insert size, thereby reducing read overlap and allowing for sampling of a larger proportion of the nucleotide sequence space. Last, the number of cycles in the library amplification PCR was increased from 12 to 20 to generate sufficient product for sequencing. The concentrations of sample libraries were normalized, and the libraries were pooled at equimolar concentrations and sequenced on a HiSeq 2500 at 200 cycles for 2 × 100 paired-end reads.

Metagenomic profiling and downstream analysis

Raw sequencing reads were first processed through Cutadapt (version 1.7.1) (*55*) and Trimmomatic (version 0.3.3) (*56*) to remove Nextera adapters, low-quality bases, and low-quality reads. Trimmed reads were then filtered for human contamination (Hg19) with KneadData (version 0.5.1). These filtered reads were then run through Metaphlan2 (version 2.6.0) for taxonomic profiling (*35*, *57*). Relative taxonomic abundances from Metaphlan2 were imported into the phyloseq package (version 1.30.0) for further refinement and visualization.

We categorized microbial species that were found in buffer controls as contaminants (most notably, the skin-resident microbe *Cutibacterium acnes*) and excluded these species from downstream analyses. Relative taxonomic abundances were used to calculate enrichment scores (probability ratios; P_R) in R using the IgAScores package (version 0.1.2) (*36*). To calculate probability ratio enrichments in the case of a species present in the lectin-bound fraction but undetected in the input, we added a pseudo-count of 0.00001 for the abundance of that species in the input sample. Data were further processed using TidyR (version 1.1.4) and visualized by ggplot2 (version 3.3.5).

Supplementary Materials

This PDF file includes:
Figs. S1 to S11
Legends for tables S1 to S8
References

Other Supplementary Material for this manuscript includes the following:
Tables S1 to S8

REFERENCES AND NOTES

1. P. J. Turnbaugh, R. E. Ley, M. Hamady, C. M. Fraser-Liggett, R. Knight, J. I. Gordon, The human microbiome project. *Nature* **449**, 804–810 (2007).
2. A. Varki, *Essentials of Glycobiology*. (Cold Spring Harbor Laboratory Press, ed. 3, 2017), p. xxix.
3. H. L. Tytgat, S. Lebeer, The sweet tooth of bacteria: Common themes in bacterial glycoconjugates. *Microbiol. Mol. Biol. Rev.* **78**, 372–417 (2014).
4. H. Lis, N. Sharon, Lectins: Carbohydrate-specific proteins that mediate cellular recognition. *Chem. Rev.* **98**, 637–674 (1998).
5. W. I. Weis, K. Drickamer, W. A. Hendrickson, Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature* **360**, 127–134 (1992).
6. D. A. Wesener, A. Dugan, L. L. Kiessling, Recognition of microbial glycans by soluble human lectins. *Curr. Opin. Struct. Biol.* **44**, 168–178 (2017).
7. T. Fujita, Evolution of the lectin–complement pathway and its role in innate immunity. *Nat. Rev. Immunol.* **2**, 346–353 (2002).
8. S. R. Stowell, C. M. Arthur, M. Dias-Baruffi, L. C. Rodrigues, J.-P. Gourdine, J. Heimburg-Molinaro, T. Ju, R. J. Molinaro, C. Rivera-Marrero, B. Xia, D. F. Smith, R. D. Cummings, Innate immune lectins kill bacteria expressing blood group antigen. *Nat. Med.* **16**, 295–301 (2010).
9. J. H. Bergström, G. M. H. Birchenough, G. Katona, B. O. Schroeder, A. Schütte, A. Ermund, M. E. V. Johansson, G. C. Hansson, Gram-positive bacteria are held at a distance in the colon mucus by the lectin-like protein ZG16. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 13833–13838 (2016).
10. H. Sarashina-Kida, H. Negishi, J. Nishio, W. Suda, Y. Nakajima, M. Yasui-Kato, K. Iwaisako, S. Kang, N. Endo, H. Yanai, M. Asagiri, H. Kida, M. Hattori, A. Kumanogoh, T. Taniguchi, Gallbladder-derived surfactant protein D regulates gut commensal bacteria for maintaining intestinal homeostasis. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 10178–10183 (2017).
11. S. Vaishnava, M. Yamamoto, K. M. Severson, K. A. Ruhn, X. Yu, O. Koren, R. Ley, E. K. Wakeland, L. V. Hooper, The antibacterial lectin regiiiγ promotes the spatial segregation of microbiota and host in the intestine. *Science* **334**, 255–258 (2011).
12. M. Wu, F. Wang, J. Yang, P. Li, D. Yan, Y. Yang, W. Zhang, J. Ren, Z. Zhang, M. Wang, The responses of the gut microbiota to MBL deficiency. *Mol. Immunol.* **122**, 99–108 (2020).
13. W. I. Weis, K. Drickamer, Structural basis of lectin-carbohydrate recognition. *Annu. Rev. Biochem.* **65**, 441–473 (1996).
14. D. A. Wesener, K. Wangkanont, R. McBride, X. Song, M. B. Kraft, H. L. Hodges, L. C. Zarling, R. A. Splain, D. F. Smith, R. D. Cummings, J. C. Paulson, K. T. Forest, L. L. Kiessling, Recognition of microbial glycans by human intelectin-1. *Nat. Struct. Mol. Biol.* **22**, 603–610 (2015).
15. J. Gao, D. Liu, Z. Wang, Screening lectin-binding specificity of bacterium by lectin microarray with gold nanoparticle probes. *Anal. Chem.* **82**, 9240–9247 (2010).
16. K.-L. Hsu, K. T. Pilobello, L. K. Mahal, Analyzing the dynamic bacterial glycome with a lectin microarray approach. *Nat. Chem. Biol.* **2**, 153–157 (2006).
17. S. Andresen, K. Fantone, D. Chapla, B. Rada, K. W. Moremen, M. Pierce, C. M. Szymanski, Human intelectin-1 promotes cellular attachment and neutrophil killing of *Streptococcus pneumoniae* in a serotype-dependent manner. *Infect. Immun.* **90**, e00682–e0068221 (2022).

18. M. Sigal, M. . M. Reinés, S. Müllerke, C. Fischer, M. Kapalczynska, H. Berger, E. R. M. Bakker, H. J. Mollenkopf, M. E. Rothenberg, B. Wiedenmann, S. Sauer, T. F. Meyer, R-spondin-3 induces secretory, antimicrobial Lgr5^{+} cells in the stomach. *Nat. Cell Biol.* **21**, 812–823 (2019).
19. S. Tsuji, J. Uehori, M. Matsumoto, Y. Suzuki, A. Matsuhisa, K. Toyoshima, T. Seya, Human intelectin is a novel soluble lectin that recognizes galactofuranose in carbohydrate chains of bacterial cell wall. *J. Biol. Chem.* **276**, 23456–23463 (2001).
20. S. R. Gill, M. Pop, R. T. Deboy, P. B. Eckburg, P. J. Turnbaugh, B. S. Samuel, J. I. Gordon, D. A. Relman, C. M. Fraser-Liggett, K. E. Nelson, Metagenomic analysis of the human distal gut microbiome. *Science* **312**, 1355–1359 (2006).
21. C. M. McMahon, C. R. Isabella, I. W. Windsor, P. Kosma, R. T. Raines, L. L. Kiessling, Stereoelectronic effects impact glycan recognition. *J. Am. Chem. Soc.* **142**, 2386–2395 (2020).
22. J. C. Barrett, S. Hansoul, D. L. Nicolae, J. H. Cho, R. H. Duerr, J. D. Rioux, S. R. Brant, M. S. Silverberg, K. D. Taylor, M. M. Barmada, A. Bitton, T. Dassopoulos, L. W. Datta, T. Green, A. M. Griffiths, E. O. Kistner, M. T. Murtha, M. D. Regueiro, J. I. Rotter, L. P. Schumm, A. H. Steinhart, S. R. Targan, R. J. Xavier; NIDDK IBD Genetics Consortium, C. Libioulle, C. Sandor, M. Lathrop, J. Belaiche, O. Dewit, I. Gut, S. Heath, D. Laukens, M. Mni, P. Rutgeerts, A. Van Gossum, D. Zelenika, D. Franchimont, J.-P. Hugot, M. de Vos, S. Vermeire, E. Louis; Belgian-French IBD Consortium; Wellcome Trust Case Control Consortium, L. R. Cardon, C. A. Anderson, H. Drummond, E. Nimmo, T. Ahmad, N. J. Prescott, C. M. Onnie, S. A. Fisher, J. Marchini, J. Ghori, S. Bumpstead, R. Gwilliam, M. Tremelling, P. Deloukas, J. Mansfield, D. Jewell, J. Satsangi, C. G. Mathew, M. Parkes, M. Georges, M. J. Daly, Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat. Genet.* **40**, 955–962 (2008).
23. A. D. Pemberton, M. J. Rose-Zerilli, J. W. Holloway, R. D. Gray, S. T. Holgate, A single-nucleotide polymorphism in intelectin 1 is associated with increased asthma risk. *J. Allergy Clin. Immunol.* **122**, 1033–1034 (2008).
24. R. M. Dommett, N. Klein, M. W. Turner, Mannose-binding lectin in innate immunity: Past, present and future. *Tissue Antigens* **68**, 193–209 (2006).
25. O. Neth, D. L. Jack, A. W. Dodds, H. Holzel, N. J. Klein, M. W. Turner, Mannose-binding lectin binds to a range of clinically relevant microorganisms and promotes complement deposition. *Infect. Immun.* **68**, 688–693 (2000).
26. W. I. Weis, K. Drickamer, Trimeric structure of a C-type mannose-binding protein. *Structure* **2**, 1227–1240 (1994).
27. J. Seyfarth, P. Garred, H. O. Madsen, Extra-hepatic transcription of the human mannose-binding lectin gene (*mbl2*) and the MBL-associated serine protease 1–3 genes. *Mol. Immunol.* **43**, 962–971 (2006).
28. B. T. Seiler, M. Cartwright, A. L. M. Dinis, S. Duffy, P. Lombardo, D. Cartwright, E. H. Super, J. Lanzaro, K. Dugas, M. Super, D. E. Ingber, Broad-spectrum capture of clinical pathogens using engineered Fc-mannose-binding lectin enhanced by antibiotic treatment [version 1; peer review: 2 approved]. *F1000Research* **8**, 108 (2019).
29. K. Drickamer, M. E. Taylor, Recent insights into structures and functions of C-type lectins in the immune system. *Curr. Opin. Struct. Biol.* **34**, 26–34 (2015).
30. W. I. Weis, M. E. Taylor, K. Drickamer, The C-type lectin superfamily in the immune system. *Immunol. Rev.* **163**, 19–34 (1998).
31. J. D. Matute, J. Duan, M. B. Flak, P. Griebel, J. A. Tascon-Arcila, S. Doms, T. Hanley, A. Antanaviciute, J. Gundrum, J. L. M. Welch, B. Sit, S. Abtahi, G. M. Fuhler, J. Grootjans, F. Tran, S. T. Stengel, J. R. White, N. Krupka, D. Haller, S. Clare, T. D. Lawley, A. Kaser, A. Simmons, J. N. Glickman, L. Bry, P. Rosenstiel, G. Borisy, M. K. Waldor, J. F. Baines, J. R. Turner, R. S. Blumberg, Intelectin-1 binds and alters the localization of the mucus barrier-modifying bacterium Akkermansia muciniphila. *J. Exp. Med.* **220**, e20211938 (2023).
32. S. Herget, P. V. Toukach, R. Ranzinger, W. E. Hull, Y. A. Knirel, C. W. von der Lieth, Statistical analysis of the Bacterial Carbohydrate Structure Data Base (BCSDB): Characteristics and diversity of bacterial carbohydrates in comparison with mammalian glycans. *BMC Struct. Biol.* **8**, 35 (2008).
33. L. Oinam, H. Tateno, Glycan profiling by sequencing to uncover multicellular communication: Launching glycobiology in single cells and microbiomes. *Front. Cell Dev. Biol.* **10**, 919168 (2022).
34. J. D. Matute, J. Duan, M. B. Flak, P. Griebel, J. A. Tascon-Arcila, S. Doms, T. Hanley, A. Antanaviciute, J. Gundrum, J. L. M. Welch, B. Sit, S. Abtahi, G. M. Fuhler, J. Grootjans, F. Tran, S. T. Stengel, J. R. White, N. Krupka, D. Haller, S. Clare, T. D. Lawley, A. Kaser, A. Simmons, J. N. Glickman, L. Bry, P. Rosenstiel, G. Borisy, M. K. Waldor, J. F. Baines, J. R. Turner, R. S. Blumberg, Gut microbiota composition and functional changes in inflammatory bowel disease and irritable bowel syndrome. *Sci. Transl. Med.* **10**, eaap8914 (2018).
35. D. T. Truong, E. A. Franzosa, T. L. Tickle, M. Scholz, G. Weingart, E. Pasolli, A. Tett, C. Huttenhower, N. Segata, MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat. Methods* **12**, 902–903 (2015).
36. M. A. Jackson, C. Pearson, N. E. Ilott, K. E. Huus, A. N. Hegazy, J. Webber, B. B. Finlay, A. J. Macpherson, F. Powrie, L. H. Lam, Accurate identification and quantification of commensal microbiota bound by host immunoglobulins. *Microbiome* **9**, 33 (2021).
37. Y. Cao, J. Shen, Z. H. Ran, Association between Faecalibacterium prausnitzii reduction and inflammatory bowel disease: A meta-analysis and systematic review of the literature. *Gastroenterol. Res. Pract.* **2014**, 872725 (2014).
38. X. Guo, S. Li, J. Zhang, F. Wu, X. Li, D. Wu, M. Zhang, Z. Ou, Z. Jie, Q. Yan, P. Li, J. Yi, Y. Peng, Genome sequencing of 39 Akkermansia muciniphila isolates reveals its population structure, genomic and functional diveristy, and global distribution in mammalian gut microbiotas. *BMC Genomics* **18**, 800 (2017).
39. B. Becken, L. Davey, D. R. Middleton, K. D. Mueller, A. Sharma, Z. C. Holmes, E. Dallow, B. Remick, G. M. Barton, L. A. David, J. R. McCann, S. C. Armstrong, P. Malkus, R. H. Valdivia, Genotypic and phenotypic diversity among human isolates of *Akkermansia muciniphila*. *mBio* **12**, e00478–e00421 (2021).
40. Y. Yan, L. H. Nguyen, E. A. Franzosa, C. Huttenhower, Strain-level epidemiology of microbial communities and the human microbiome. *Genome Med.* **12**, 71 (2020).
41. Z. Bai, N. Zhang, Y. Jin, L. Chen, Y. Mao, L. Sun, F. Fang, Y. Liu, M. Han, G. Li, Comprehensive analysis of 84 *Faecalibacterium prausnitzii* strains uncovers their genetic diversity, functional characteristics, and potential risks. *Front. Cell. Infect. Microbiol.* **12**, 919701 (2022).
42. E. Drula, M. L. Garron, S. Dogan, V. Lombard, B. Henrissat, N. Terrapon, The carbohydrate-active enzyme database: Functions and literature. *Nucleic Acids Res.* **50**, D571–D577 (2022).
43. B. T. Seiler, M. Cartwright, A. L. M. Dinis, S. Duffy, P. Lombardo, D. Cartwright, E. H. Super, J. Lanzaro, K. Dugas, M. Super, D. E. Ingber, Broad-spectrum capture of clinical pathogens using engineered Fc-mannose-binding lectin enhanced by antibiotic treatment. *F1000Res* **8**, 108 (2019).
44. T. Chénard, M. Malick, J. Dubé, E. Massé, The influence of blood on the human gut microbiome. *BMC Microbiol.* **20**, 44 (2020).
45. S. Tsuji, M. Yamashita, D. R. Hoffman, A. Nishiyama, T. Shinohara, T. Ohtsu, Y. Shibata, Capture of heat-killed Mycobacterium bovis bacillus Calmette-Guérin by intelectin-1 deposited on cell surfaces. *Glycobiology* **19**, 518–526 (2009).
46. H. Kobayashi, K. Uchimura, T. Ishii, K. Takahashi, K. Mori, K. Tsuchiya, F. Furuya, Intelectin1 ameliorates macrophage activation *via* inhibiting the nuclear factor kappa B pathway. *Endocr. J.* **69**, 539–546 (2022).
47. J. Wang, Y. Gao, F. Lin, K. Han, X. Wang, Omentin-1 attenuates lipopolysaccharide (LPS)-induced U937 macrophages activation by inhibiting the TLR4/MyD88/NF-κB signaling. *Arch. Biochem. Biophys.* **679**, 108187 (2020).
48. M. C. Collado, M. Derrien, E. Isolauri, W. M. de Vos, S. Salminen, Intestinal integrity and Akkermansia muciniphila, a mucin-degrading member of the intestinal microbiota present in infants, adults, and the elderly. *Appl. Environ. Microbiol.* **73**, 7767–7770 (2007).
49. G. A. Lugli, L. Mancabelli, C. Milani, F. Fontana, C. Tarracchini, G. Alessandri, D. van Sinderen, F. Turroni, M. Ventura, Comprehensive insights from composition to functional microbe-based biodiversity of the infant human gut microbiota. *NPJ Biofilms Microbiomes* **9**, 25 (2023).
50. K. A. Romano, E. I. Vivas, D. Amador-Noguez, F. E. Rey, Intestinal microbiota composition modulates choline bioavailability from diet and accumulation of the proatherogenic metabolite trimethylamine-N-oxide. *mBio* **6**, e02481 (2015).
51. A. B. Hall, M. Yassour, J. Sauk, A. Garner, X. Jiang, T. Arthur, G. K. Lagoudas, T. Vatanen, N. Fornelos, R. Wilson, M. Bertha, M. Cohen, J. Garber, H. Khalili, D. Gevers, A. N. Ananthakrishnan, S. Kugathasan, E. S. Lander, P. Blainey, H. Vlamakis, R. J. Xavier, C. Huttenhower, A novel Ruminococcus gnavus clade enriched in inflammatory bowel disease patients. *Genome Med.* **9**, 103 (2017).
52. S. R. Stowell, C. M. Arthur, R. McBride, O. Berger, N. Razi, J. Heimburg-Molinaro, L. C. Rodrigues, J. P. Gourdine, A. J. Noll, S. von Gunten, D. F. Smith, Y. A. Knirel, J. C. Paulson, R. D. Cummings, Microbial glycan microarrays define key features of host-microbial interactions. *Nat. Chem. Biol.* **10**, 470–476 (2014).
53. G. E. Truett, P. Heeger, R. L. Mynatt, A. A. Truett, J. A. Walker, M. L. Warman, Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). *Biotechniques* **29**, 52–54 (2000).
54. J. D. Brewster, G. C. Paoli, DNA extraction protocol for rapid PCR detection of pathogenic bacteria. *Anal. Biochem.* **442**, 107–109 (2013).
55. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, (2011).
56. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
57. L. J. McIver, G. Abu-Ali, E. A. Franzosa, R. Schwager, X. C. Morgan, L. Waldron, N. Segata, C. Huttenhower, bioBakery: A meta'omic analysis environment. *Bioinformatics* **34**, 1235–1237 (2017).

58. M. Derrien, E. E. Vaughan, C. M. Plugge, W. M. de Vos, *Akkermansia muciniphila* gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. *Int. J. Syst. Evol. Microbiol.* **54**, 1469–1476 (2004).
59. V. B. D. Skerman, V. McGOWAN, P. H. A. Sneath, Approved lists of bacterial names. *Int. J. Syst. Evol. Microbiol.* **30**, 225–420 (1980).
60. L. V. Holdeman, W. E. C. Moore, New genus, coprococcus, twelve new species, and emended descriptions of four previously described species of bacteria from human feces. *Int. J. Syst. Evol. Microbiol.* **24**, 260–277 (1974).
61. P. Mattarelli, C. Bonaparte, B. Pot, B. Biavati, Proposal to reclassify the three biotypes of *Bifidobacterium longum* as three subspecies: *Bifidobacterium longum* subsp. longum subsp. nov., *Bifidobacterium longum* subsp. infantis comb. nov. and *Bifidobacterium longum* subsp. suis comb. nov. *Int. J. Syst. Evol. Microbiol.* **58**, 767–772 (2008).
62. W. E. C. Moore, J. L. Johnson, L. V. Holdeman, Emendation of *Bacteroidaceae* and *Butyrivibrio* and descriptions of *Desulfomonas* gen. nov. and ten new species in the genera *Desulfomonas, Butyrivibrio, Eubacterium, Clostridium*, and *Ruminococcus*. *Int. J. Syst. Evol. Microbiol.* **26**, 238–252 (1976).
63. L. C. Reimer, A. Vetcininova, J. S. Carbasse, C. Söhngen, D. Gleim, C. Ebeling, J. Overmann, BacDive in 2019: Bacterial phenotypic data for high-throughput biodiversity analysis. *Nucleic Acids Res.* **47**, D631–D636 (2019).
64. G. Reuter, Designation of type strains for bifidobacterium species. *Int. J. Syst. Evol. Microbiol.* **21**, 273–275 (1971).
65. V. Scardovi, L. D. Trovatelli, G. Zani, F. Crociani, D. Matteuzzi, Deoxyribonucleic acid homology relationships among species of the genus Bifidobacterium. *Int. J. Syst. Evol. Microbiol.* **21**, 276–294 (1971).
66. R. F. Wang, W. W. Cao, C. E. Cerniglia, PCR detection and quantitation of predominant anaerobic bacteria in human and animal fecal samples. *Appl. Environ. Microbiol.* **62**, 1242–1247 (1996).

Acknowledgments: We would like to thank L. Besse for project management and the Broad Institute's Microbial 'Omics Core and Genomics Platform for sample processing and sequencing data generation. We also thank D. Plichta at the Broad Institute for helpful discussions. We are grateful to all study participants who provided samples and to H. Lau and K. Shannon for clinical sample coordination for the PRISM cohort. **Funding:** This work was supported by Center for Microbiome Informatics and Therapeutics (H.V., R.J.X., and L.L.K.), NIH grant R01HL148577 (F.E.R), National Institutes of Health grant R01AI055258 (L.L.K.), NIH grant U01CA231079 (L.L.K.), Helen Hay Whitney Foundation postdoctoral fellowship (R.L.M.), and NSF GRFP predoctoral fellowship (C.R.I.). **Author contributions:** Conceptualization: C.R.I., D.A.W., and L.L.K. Methodology: R.L.M., C.R.I., R.L.W., D.S., T.G., S.R., L.T.T.N., D.A.W., M.H., M.G.W., A.D., R.K., S.G., S.B., C.D., and G.P. Investigation: R.L.M., C.R.I., R.L.W., D.S., T.G., S.R., L.T.T.N., D.A.W., M.H., M.G.W., A.D., R.K., S.G., S.B., and C.D. Visualization: R.L.M., C.R.I., R.L.W., and M.G.W. Funding acquisition: R.L.M., C.R.I., F.E.R, H.V., R.J.X., and L.L.K. Project administration: H.V., E.J.A., R.J.X., and L.L.K. Supervision: H.V., E.J.A., R.J.X., and L.L.K. Writing (original draft): R.L.M., C.R.I., R.L.W., DS., M.H., H.V., and L.L.K. Writing (review and editing): R.L.M., C.R.I., S.B., and L.L.K. **Competing interests:** L.L.K., D.A.W., and F.E.R. are coinventors on a patent (Patent No. US 10,413,588 filed by the Wisconsin Alumni Research Foundation on 18 September 2019) pertaining to the results presented in this manuscript. The authors declare no other competing interests. **Data and materials availability:** Metagenomic sequences are available via Sequence Read Archive (www.ncbi.nlm.nih.gov/sra) with BioProject number PRJNA905698. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 10 July 2022
Accepted 27 June 2023
Published 28 July 2023
10.1126/sciadv.add8766