Identification of an N-acetylneuraminic acid-presenting bacteria isolated from a healthy human microbiome

Zhen Han¹, Peter S. Thuy-Boun¹, Wayne Pfeiffer², Vincent F. Vartabedian³, Ali Torkamani^{4,5}, John R. Tejjaro³, Dennis W. Wolan^{1,5}

¹Department of Molecular Medicine, The Scripps Research Institute, La Jolla, CA 92037

²San Diego Supercomputer Center, University of California, San Diego, La Jolla CA 92093

³Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA 92037

⁴Scripps Research Translational Institute, The Scripps Research Institute, La Jolla, CA 92037

⁵Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037

*Correspondence: wolan@scripps.edu

Abstract

N-acetylneuraminic acid is the most abundant sialic acid in humans and is generally expressed as the terminal sugar on intestinal mucus glycans. Several pathogenic bacterial species harvest sialic acid from the mucus, diet, as well as other intestinal sources and display this sugar on their own surface to evade sialic acid-binding immunoglobulin-type lectin (Siglec)-mediated host immune surveillance. While previous studies have identified bacterial enzymes associated with sialic acid catabolism, no reported methods permit the selective labeling, tracking, and quantitation of sialic acid-presenting microbes within complex multi-microbial systems. Here, we apply an interdisciplinary approach combining metabolic labeling, click chemistry, metagenomic, and whole-genome sequencing to selectively track and identify sialic acid-presenting microbes from a cultured healthy human fecal microbiome. We isolated and identified a new strain of *Escherichia coli* that incorporates sialic acid onto its own surface. Analysis of the sequence data reveals that this *E. coli* strain encodes for the NanT, NeuA, and NeuS genes necessary for harvesting environmental sialic acid and generating the capsular polysialic acid. We envision that this method is applicable to the detection and quantitation of sialic acid-presenting bacteria from human, animal, and environmental microbiomes, as well as investigating the importance of other carbohydrates to commensal and pathogenic bacteria.

Introduction

Several methodologies and model systems have been critical to understanding gut microbiome taxonomic composition and the influence these compositions exert on host physiology and intestinal disease pathology (1-3). Metagenomics and metabolomics as well as germ-free and gnotobiotic mouse models have helped establish how specific microbes and small-molecule metabolites modulate host responses to a variety of diseases (4-7). These methods, in combination with the genetic manipulation of specific bacterial strains, has revealed key catabolizing enzymes and pathogen-associated

molecular patterns (PAMPs) associated with microbiome-related metabolism, host immune activation, and bacterial infection (8-10). Despite these methods and models, the ability to selectively label and track bacteria with specific catabolic functionalities in a highly complex and metabolically active microbiome remains limited.

N-acetylneuraminic acid (Neu5Ac, termed sialic acid/SA here) is one of over thirty known sialic acid analogs. It is the predominant form of sialic acid in humans and is presented as the terminal residue on surface-exposed glycans, glycoproteins, and glycolipids (11). SA is recognized by immunoinhibitory sialic acid-binding immunoglobulin-like lectins (Siglecs) to prevent autoimmunity (12). Interestingly, select human pathogenic microbes have evolved to express this human SA epitope on their own cellular surface to evade host immune surveillance and clearance (8,13-15). For example, group B streptococcus (GBS), a common cause of sepsis in human newborns, presents terminal α 2,3-linked SA on its capsular polysaccharide (CPS) to bind Siglecs expressed by neutrophils, macrophages, and platelets and block immune activation (8,16,17). Sequencing analysis revealed that GBS uses a tripartite transporter to translocate environmental SA onto its CPS to augment this immune evasion (8,18). We posited that certain gut commensal organisms likely apply similar SA-mediated protective mechanisms to gain survival advantages. As such, the selective labeling and identification of these microbes from complex microbiome samples may reveal new SA-regulated host-microbe crosstalk mechanisms that could potentially be exploited for therapeutic development aimed at gut microbiome-related diseases.

Here, we report the application of an SA-based azide-containing probe N-acetyl-9-azido-9-deoxy-neuraminic acid (Neu5Ac9N₃/Sia9N₃), combined with flow cytometry and metagenomic and whole-genome sequencing, to selectively label, isolate, and identify SA-presenting bacteria from a complex cultured human fecal microbiome sample. Broadly, applications of metabolic probe labeling followed by click chemistry-tagging for visualization have been used to image and track pre-labeled bacterial components, including peptidoglycans and lipopolysaccharides (19-21). However, there is limited use of SA-based probes to label bacterial glycans, especially in a highly complex microbial sample. With respect to eukaryotic labeling, “clickable” SA-based probes and its metabolic precursor (*i.e.*, sialic acid-alkyne, sialic acid-azide, and mannosamine-alkyne) have been widely used for the study of sialoglycans in both *in vitro* and *in vivo* systems (22-26). We sought to expand the application of SA-based metabolic probes to selectively label SA-presenting bacteria from a distal gut microbiome. Fecal microbiome samples collected from a healthy human volunteer were cultivated in the presence of the Sia9N₃ metabolic probe (27) and bacterial incorporation of Sia9N₃ was examined by flow cytometry analysis following copper-free azide coupling with azadibenzocyclooctyne-conjugated biotin (ADIBO-BTN/DBCO-BTN) and streptavidin-Alexa 647 staining (28). We identified a healthy human fecal microbiome sample containing bacteria that can readily incorporate Sia9N₃. With fluorescence activated cell sorting (FACS) and 16S rDNA sequencing analyses, we found that the Sia9N₃-incorporating bacteria belonged to the *Escherichia* genus. Isolation of the SA-incorporating bacteria via colony screening with whole-genome sequencing analysis

suggested that the isolated bacteria is a new strain of *E. coli* that likely employs the NanT-NeuA-NeuS system to integrate environmental sialic acid to form polysialic acid on its capsular polysaccharide structure (29,30).

Results

FACS-based screening helps identify microbiome constituents capable of Sia9N₃ incorporation.

We surveyed fecal samples collected from one healthy human volunteer over the course of three months. Fecal microbes were cultured in a Gifu media (31) supplemented with Sia9N₃ at a physiologically relevant concentration (200 μM) (32) under anaerobic atmosphere for 20 h. Microbes were centrifuged, extensively washed, subjected to DBCO-BTN conjugation with streptavidin-Alexa 647, and examined for fluorophore incorporation with flow cytometry analysis (**Fig. 1**). Among three microbiota samples tested (designated H1, H2, and H3), more than 60% of the bacterial cells from one of the cultured human fecal samples (e.g., H2) readily displayed fluorescence in response to Sia9N₃ treatment (**Fig. 1**). Furthermore, a dose-dependent assay revealed that bacterial fluorescence labeling is dependent on concentrations of both Sia9N₃ and DBCO-BTN. Use of 200 μM Sia9N₃ in culture media with 50 μM DBCO-BTN for azide conjugation resulted in the maximum percentage of labeled bacterial cells (**Figs. 2A and B**). Further increases in the concentration of either labeling reagent only amplified the fluorescence intensity on the labeled bacterial cells without increasing background fluorescence or numbers of labeled cells, which suggested that Sia9N₃ selectively labeled a subset of SA-catabolizing bacteria (**Fig. S1**).

Fig. 1. Metabolic labeling of SA-presenting microbes from a fecal microbiome sample. An interdisciplinary approach combining metabolic labeling, click chemistry, and flow cytometry revealed a human fecal microbiome sample that contains SA-presenting bacteria.

We next assessed if Neu5Ac would compete with Sia9N₃ for surface expression by bacteria when supplemented into the 20 h cultures. Addition of excess Neu5Ac resulted in reduced labeling of bacteria by Sia9N₃ (**Figs. 2C and S2A**), which suggested that bacteria may use the same or partially overlapped transporters and enzymes for SA and Sia9N₃ transportation and catabolism. Additionally,

these results indicate that Sia9N₃ is biologically relevant as a metabolic probe since the substitution of hydroxyl to azide on C9 did not significantly affect its interaction with the SA-binding proteins. Labeled bacteria were also subjected to fluorescence microscopic analysis to visualize where the bacteria are labeled (*i.e.*, intracellular, periplasmic, cell wall, etc.). These experiments provided visual evidence that Sia9N₃ or its azido-metabolites are prominently delivered to the bacterial cell surface (Fig. 2D). This finding is consistent with the observation that certain bacteria can present environmental SA onto its surface capsular polysaccharides and/or lipopolysaccharides (33).

Fig. 2. Characterization of Sia9N₃ induced fluorescence staining with flow cytometry and confocal fluorescence microscopy. Labeling of SA-presenting microbes is Sia9N₃ (A) and DBCO-BTN (B) dose-dependent. C. Addition of excess natural SA partially competed for fluorescence labeling. D. Confocal fluorescence microscopic imaging showed that Sia9N₃ induced fluorescence labeling on the bacterial cell surface.

Bacterial surface Sia9N₃ is removed by microbial sialidases. Many gut bacteria can express sialidases to release SA from sialoglycans, providing themselves or other bacteria with a source of carbon, nitrogen, and building blocks for cell wall biosynthesis (34,35). To determine if the Sia9N₃ probe is presented on bacteria to form sialoglycans, we treated the metabolically labeled H2 cultured

sample with a purified recombinant sialidase (BT0455, NCBI gene ID: 1071627) (36) from *Bacteroides thetaiotaomicron*, followed by sequential DBCO-BTN conjugation, streptavidin-Alexa 647 staining, and flow cytometry. Sialidase treatment dramatically reduced the number of fluorescently labeled bacterial cells and fluorescence intensity (Figs. 3A and S2B). This result suggested that the majority of bacterial fluorescent labeling is due to Sia9N₃ transported into certain bacterial cells and presented onto the cell surface to form sialoglycans. The inability to completely eliminate fluorescence labeling is likely due to substrate specificity of BT-0455 that includes α 2,3-, α 2,6-, and α 2,8-linked sialic acid substrates (36). Sia9N₃ may be installed on glycans unrecognizable by BT-0455 and/or Sia9N₃ is catabolized and surface expressed as a modified SA or other sugar.

Fig. 3. Bacterially incorporated surface Sia9N₃ can be removed by microbial sialidases over time. A. Treatment of labeled microbes before DBCO-BTN conjugation with a recombinant sialidase significantly reduces fluorescence labeling. The incorporated Sia9N₃ is removed by microbial sialidases over time as fluorescence labeling was gradually removed during microbiota cultivation (B) and microbial sialidase activity was constantly detected in the culture media, 500 pM recombinant BT0455 was used as a positive control (C). D. 16S rDNA sequencing revealed the change of taxonomic composition during H2 microbiome culture.

Sialic acid can be released from sialoglycans by microbial sialidases as demonstrated (Fig. 3A), which can promote the growth of certain SA-utilizing pathogens in the gut and lead to inflammatory diseases (32). To explore if Sia9N₃ can also be released from the bacterially presented sialoglycans by other bacteria within the cultured sample, we tracked Sia9N₃ incorporation over time during the plateau phase of bacterial growth. Briefly, the H2 fecal microbiome sample was grown with Sia9N₃ for 14 to 66 h and the bacteria were quantitated for Sia9N₃ incorporation by flow cytometry as described. Interestingly, detection of Sia9N₃ labeling was maximally present at 14 h and gradually decreased until almost all labeling was eliminated by 40 h of incubation (Fig. 3B). Cultured media from each sample

for every timepoint was assessed for sialidase activity, as measured by catalytic turnover of 4-methylumbelliferyl-N-acetylneuraminic acid (4-MUANA). Measurable sialidase activity was observed across all samples and suggests that bacterially secreted sialidases could remove Sia9N₃ from bacterial sialoglycans over time (Fig. 3C). In parallel to the sialidase analysis, 16S rDNA sequencing showed that the compositional diversity of the cultured microbiota increased over time (Fig. 3D). At 14 h, the microbiome collection was dominated by *Escherichia/Shigella* (58% relative abundance), *Clostridium* (21%), and *Enterococcus* (15%). Common commensal organisms, including *Citrobacter*, *Peptoniphilus*, *Bacteroides*, and *Parabacteroides*, began to appear after 24 h of culture and their relative abundances peaked at 16%, 13%, 10%, and 5%, respectively, at 66 h of culture. The corresponding relative abundance of *Escherichia/Shigella*, *Clostridium*, and *Enterococcus* dropped to 33%, 2%, and 0%, respectively. Our results suggest that there is no correlation between Sia9N₃ removal and abundance of secreted sialidase activity within the cultured media. While the bloom of other commensal bacteria at later time points may suppress the detection of Sia9N₃-presenting microbes, we posit that the complete elimination of Sia9N₃ detection is likely due to sialidase activity. The sialidases responsible for the removal of Sia9N₃ may not be present at early time points and are introduced by late-blooming commensal organisms, such as *Bacteroides*, that appear at later time points, as determined by 16S rDNA analysis. Importantly, this shift in taxonomic composition may reflect protective mechanisms by which commensal bacteria target virulence factors expressed on pathogenic/pathobiont bacteria by exposing surface antigens for host recognition. Alternatively, commensal bacteria could be modifying environmental conditions to reduce the growth of pathogenic/pathobiont bacteria. All together, these data demonstrated that Sia9N₃ incorporation by bacteria in culture is dynamically regulated by other bacteria, nutrients, and bacterial sialidases.

Fig. 4. 16S rDNA metagenomic sequencing analyses. A. H2 microbiome sample was cultured with or without 200 μ M Sia9N₃, which does not significantly alter taxonomic composition of 20-h H2 microbiome culture. Sia9N₃ presenting bacteria from H2 microbiome culture were labeled with fluorescence and isolated with FACS. **A and B.** 16S rDNA sequencing showed that the *Escherichia/Shigella* genus was significantly enriched after cell sorting.

The isolated Sia9N₃-presenting bacteria is a new strain of *Escherichia coli*. We next separated the Sia9N₃ incorporating bacteria from a 20-h culture of the H2 microbiome sample with FACS. The isolated fraction that yielded a high fluorescence signal due to Sia9N₃ incorporation was subjected to 16S rDNA sequencing and revealed that the *Escherichia* genus was significantly enriched post sorting (Figs. 4A and B). We next subjected the cultured and primary H1-H3 microbiome samples to 16S

rDNA sequencing to determine if differences in the bacterial taxonomic composition accounted for the selective labeling of the H2 microbiome sample only. Of note, 20 h of culture significantly altered the composition of all samples, and therefore, made the Sia9N₃-presenting bacteria detectable in H2 as the *Escherichia* genus was significantly enriched compared to the starting fecal sample (**Fig. S3**). In addition, although only H2 contains Sia9N₃-presenting bacteria, *Escherichia* dominated all three cultured microbiome samples as the most abundant genus. These results suggest that *Escherichia* presented in the H1-3 cultured microbiome samples may differ at the strain level as only the *Escherichia* in H2 has the catabolic genes to incorporate Sia9N₃. An additional possibility is that the isolated *Escherichia* organism from H2 is present across all samples; however, the SA-presenting genes are not activated in the cultured H1 and H3 samples.

Fig. 5. Isolation and passaging of the Sia9N₃ incorporating bacterial strain from H2 microbiota. **A.** Colony screening of H2 microbiota isolated one Sia9N₃ incorporating bacterial strain. **B.** The Sia9N₃ incorporating feature is readily passaged to daughter generations.

To identify the bacterial genome of the Sia9N₃-presenting H2 *Escherichia*, we randomly selected 10 colonies from an H2 culture plate followed by Sia9N₃ metabolic labeling assay. Out of 10 colonies screened, liquid culture in Gifu media of colony 2 passage 1 yielded nearly 100% incorporation of Sia9N₃ (C2-P1, **Fig. 5A**). Furthermore, 3 randomly picked colonies from a C2 culture plate (C2 passage-2, C2-P2-1, -2, and -3) also showed 100% Sia9N₃ incorporation, suggesting that the Sia9N₃ catabolic genes were readily passaged to, and active in, the daughter generations (**Fig. 5B**). The ZH-C2 colony was subjected to shotgun whole-genome sequencing in an attempt to identify the *Escherichia* bacterial strain.

The genome of our newly isolated bacteria (ZH-C2) cannot be fully mapped to any previously sequenced *Escherichia* bacteria in the NCBI RefSeq database, but is most similar to pathogenic *E. coli* strains UMEA 3174-1 (accession PRJNA186306) and VR50 (accession PRJEA61445) with average nucleotide identities of 99.85% and 99.76%, respectively (**Fig. 6A** and **Supplemental Table, tab 1**). Our *E. coli* ZH-C2 strain has a 5.23 kb genome that is similar in size to UMEA 3174-1 and VR50 and is much larger than the 4.64 kb genome of the prototypical nonpathogenic strain K-12 MG1655 (MG1655, accession PRJNA647163) (**Supplemental Table, tab 1**). ZH-C2 shares over 90% identity in

sequence to most MG1655 genes; however, ZH-C2 consists of 1,111 missing or low-homology genes compared to MG1655, while MG1655 has 347 missing or low-homology genes relative to ZH-C2 (**Fig. 6A** and **Supplemental Table, tab 2**). The majority of gene products have no identified function; however, GO term analyses demonstrate ZH-C2 encodes for an increased diversity and number of additional proteins relative to MG1655 involved in DNA repair and replication, polysaccharide transport, virion assembly, as well as type II secretion (**Fig. 6B** and **Supplemental Table, tab 3**). With respect to molecular function, a diverse set of additional genes are encoded in ZH-C2, notably predicted cobalamin binding, ATPase, transferase, and porin activities (**Fig. 6C** and **Supplemental Table, tab 3**) with a significantly larger collection of ZH-C2 proteins likely localized to the outer membrane compared to MG1655 (**Fig. 6D** and **Supplemental Table, tab 3**). Conversely, the MG1655 genome consists of a number of additional genes associated with DNA-mediated recombination and integration, as well as cell adhesion, DNA binding and DNA recombinase activities (**Fig. 6B-D** and **Supplemental Table, tab 3**).

Fig. 6. Comparison of encoded *E. coli* ZH-C2 genes to reference non-pathogenic and pathogenic *E. coli*.

Venn diagram of CD-HIT analysis depicting conservation and differences of gene products between *E. coli* strains ZH-C2, K-12 MG1655, UMEA 3174-1, and VR50 (A). Protein sequences were clustered at CD-HIT sequence similarity cut-off of 90%. Biological process (B), molecular function (C), and cellular component (D) GO Terms of unique genes (less than 70% coverage and/or less than 90% identity) found in ZH-2 compared to MG1655, UMEA 3174-1, VR50, and the RefSeq reference genome collection. The number of genes used for each GO term analysis is in parentheses. GO Terms listed as “other” include <2 genes and are expanded in the **Supplementary Table**.

MG1655 encodes for the sialic acid catabolizing genes, including Nan A, C, E, K, M, the regulatory gene NanR, and sialic acid transporter gene NanT. However, as for *E. coli* strains UMEA 3174-1 and VR50, ZH-C2 encodes for: 1) NeuA, the cytidine monophosphate (CMP)-sialic acid synthetase which produces the sialyltransferase donor CMP-sialic acid; and 2) NeuS, the polysialyltransferase which synthesizes capsular polysialic acids from CMP-sialic acid (29). The lack of the NeuA and NeuS genes highlights the inability of MG1655 to activate and present Sia9N₃ on its cell surface (**Fig. S4**). Moreover, the sugar-phosphate isomerase KpsF/GutQ (accession WP_001296394, ZH-C2_01729), known to be associated with polysialic acid deposition onto the capsule is also identified in the ZH-C2 genome (with 100% identity to UMEA 3174-1 and VR50) but not in MG1655 (37). Altogether, the genomic analyses suggest that the cell-surface fluorescence labeling we observed is caused by deposition of polySia9N₃ on the bacterial capsule.

Despite being isolated from a healthy human fecal sample, our ZH-C2 strain may be a pathogenic/pathobiont *E. coli*. The two most similar genomes, UMEA 3174-1 and VR50, were both isolated from urinary tract infections (**Supplemental Table, tab 1**). Established pathogenic genes, such as the type III secretion system are conserved across ZH-C2, UMEA 3174-1, and VR50 (38). Notwithstanding the high conservation, notable differences were also observed. ZH-C2 has 628 and 680 low homology/unique genes relative to UMEA 3174-1 and VR50, respectively. Conversely, UMEA and VR50 have 507 and 393 low homology/unique genes, respectively, to ZH-C2 as determined by sequences with <90% identity and/or covering <70% of the sequence length (**Supplemental Table, tabs 4,5**). Interestingly, there is an enrichment of predicted GO terms associated with DNA-mediated transposition, oxidation-reduction processes, pilus assembly, and carbohydrate metabolic activities in ZH-C2 relative to both UMEA 3174-1 and VR50, as well as a localization of proteins to the membrane (**Fig. 6B-D** and **Supplemental Table, tab 3**). Those missing functions in ZH-C2 found in UMEA 3174-1 and VR50, include transcriptional regulation, viral release, nucleic acid binding, and recombinase activity (**Fig. 6B-D** and **Supplemental Table, tab 3**).

With respect to all available non-redundant RefSeq reference genomes, ZH-C2 has 207 low homology or unique proteins with approximately 65 sequences consisting of <100 amino acids (**Fig. 6B-D** and **Supplemental Table, tab 6**). Almost all of these sequences have been identified in one or more of the other >20K *E. coli* strains that have been sequenced to date and are in the NCBI RefSeq/GenBank database. However, ZH-C2 encodes for ZH-C2_04397 and ZH-C2_04599 that are 1,174 and 475 amino acids long, respectively, and have not been observed previously in any *E. coli* strain. ZH-C2_04397 is a multi-subunit protein consisting of an N-terminal autoinducer-2 kinase, an autotransporter barrel domain-containing lipoprotein, and a C-terminal HipA-like Ser/Thr kinase (**Fig. S5**) (39,40). While most *E. coli* encode for a linked autotransporter lipoprotein and C-terminal HipA-like

Ser/Thr kinase, only one other *E. coli* has an N-terminal autoinducer-2 kinase connected to the autotransporter lipoprotein (LsrK in *E. coli* KTE98, accession EOV99460.1). However, this *E. coli* KTE98 gene lacks the C-terminal HipA-like Ser/Thr kinase. Most *E. coli* do encode for an autoinducer-2 kinase as a single gene. ZH-C2_04599 also is a multidomain gene consisting of an N-terminal transposase and a C-terminal EamA transporter (**Fig. S6**) (41). Again, both domains are well-represented across *E. coli* as single genes. The biological importance of these multi-functional connected domains will be subject to future investigations.

Discussion

Azide-modified glucosamine, mannosamine, fucose, and 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) have been used to metabolically label bacterial cells *in vitro* and *in vivo* (19,20,42). Here, we demonstrate the application of azide-modified N-acetylneuraminic acid to selectively label and track Neu5Ac-presenting bacteria from a complex human fecal microbiome sample for the first time. Using flow cytometry and metagenomic sequencing, we identified a new strain of *E. coli* that can present environmental Neu5Ac on its surface via the NanT-NeuA-NeuS pathway. Our results also suggest a dynamic process by which select distal gut bacteria can exploit environmental sialic acid for self-decoration that, in turn, is likely subject to removal by sialidase activity provided by other microbiome constituents (e.g., *B. thetaiotaomicron*).

The Sia9N₃-presenting activity of our new *E. coli* ZH-C2 strain is only detected in sample H2. Additional studies will investigate if this new *E. coli* is missing from the H1 and H3 cultures or present with a silenced NanT-NeuA-NeuS pathway during the *ex vivo* culture. Previous studies suggest both possibilities, as the gut microbiome is regulated by diet and crosstalk amongst bacteria as well as the host and suggests a dynamic composition of different bacterial strains over time (43,44). Conversely, certain bacteria, including *Bacteroides fragilis*, are able to switch capsular polysaccharide structures enabling the colonization specific intestinal niches (45). The gradual decrease in *E. coli* Sia9N₃ surface expression over time will also be further interrogated. We anticipate that the loss of Sia9N₃ is caused by other commensal organisms that gain prominence over time in the *ex vivo* culture and remove the *E. coli* poly-SA as a biologically relevant mechanism for host recognition and clearance. Additionally, surface-expressed poly-SA may be harvested by the parental *E. coli* for growth, as environmental conditions and nutrient availability shift over time.

Distal gut bacteria can degrade and ferment complex carbohydrates to provide an energy source and/or immunomodulatory molecules for themselves, the host, and/or other gut bacteria (1,32,46,47). In return, host cells can produce glycans which can provide nutritional advantages for certain microbes (48). Such mutualistic relationships make the tracking and study of carbohydrate catabolism by individual bacteria in complex microbiomes complicated and challenging. Consistent with previous studies, we found that cultured microbiome samples do not fully represent the taxonomic diversity of the original source due to growth advantages of certain bacteria under anaerobic systems with limited nutritional sources from the selected media (31,49). Notwithstanding, our approach enriched for a new

SA-presenting strain of *Escherichia coli* isolated from a human gut microbiome. We posit that SA-presenting bacteria can be identified from different microbial sources, including respiratory, urinary, and vaginal microbiomes in both health and disease (*i.e.*, detection of pathogens in blood and stool clinical samples). Additionally, the scavenging of environmental sialic acid is associated with pathogenic bacterial virulence (50) and proteins involved in the process may represent new targets for the discovery of antibacterial small molecules.

Conclusions

Our integrated assay system detects sialic acid-presenting bacteria in real-time from complex microbiome samples. We envision that this labeling and tracking strategy can assist in identifying new SA-presenting bacteria and developing diagnostic tools and drug discovery models. Furthermore, this workflow may be used to investigate the installation of other environmental glycans by microbiota.

Supplemental Information

Detailed methods and reagents used in this study are included in the supplemental information.

Acknowledgements

We thank Jenny Cornell, Jessica Ledesma, and Steven Robert Head at Next Generation Sequencing Core at The Scripps Research Institute for 16S rDNA and whole-genome sequencing experiments. We thank Padmaja Natarajan and Alain Domissy at Center for Computational Biology at The Scripps Research Institute for analyzing the 16S rDNA sequencing data. We thank Qiyun Zhu at the University of California, San Diego for assistance with the whole-genome sequencing data analysis. This work was supported by NIH award R21 AI139744 to D.W.W. and Scripps Research Translational Institute, an NIH-NCATS Clinical and Translational Science Award (CTSA; UL1 TR002550) to W.P. and A.T.

Author contributions

Z.H. and D.W.W. designed the research; Z.H., P.T.B., V.F.V., and J.R.T. performed the experiments; Z.H., W.P., P.T.B., A.T., and D.W.W. analyzed the data; Z.H. and D.W.W. composed the paper and all authors edited and approved its contents.

Declaration of interests

The authors declare no conflict of interest.

References

1. Rooks, M. G., and Garrett, W. S. (2016) Gut microbiota, metabolites and host immunity. *Nat. Rev. Immunol.* **16**, 341-352
2. Valdes, A. M., Walter, J., Segal, E., and Spector, T. D. (2018) Role of the gut microbiota in nutrition and health. *BMJ* **361**, k2179
3. Boulange, C. L., Neves, A. L., Chilloux, J., Nicholson, J. K., and Dumas, M. E. (2016) Impact of the gut microbiota on inflammation, obesity, and metabolic disease. *Genome Med.* **8**, 42
4. Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., *et al.* (2010) A human

- gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59-65
5. Korem, T., Zeevi, D., Suez, J., Weinberger, A., Avnit-Sagi, T., Pompan-Lotan, M., *et al.* (2015) Growth dynamics of gut microbiota in health and disease inferred from single metagenomic samples. *Science* **349**, 1101-1106
6. Shepherd, E. S., DeLoache, W. C., Pruss, K. M., Whitaker, W. R., and Sonnenburg, J. L. (2018) An exclusive metabolic niche enables strain engraftment in the gut microbiota. *Nature* **557**, 434-438
7. Tanoue, T., Morita, S., Plichta, D. R., Skelly, A. N., Suda, W., Sugiura, Y., *et al.* (2019) A defined commensal consortium elicits CD8 T cells and anti-cancer immunity. *Nature* **565**, 600-605
8. Uchiyama, S., Sun, J., Fukahori, K., Ando, N., Wu, M., Schwarz, F., Siddiqui, S. S., Varki, A., Marth, J. D., and Nizet, V. (2019) Dual actions of group B streptococcus capsular sialic acid provide resistance to platelet-mediated antimicrobial killing. *Proc. Natl. Acad. Sci. USA* **116**, 7465-7470
9. Allen, S., Zaleski, A., Johnston, J. W., Gibson, B. W., and Apicella, M. A. (2005) Novel sialic acid transporter of *Haemophilus influenzae*. *Infect. Immun.* **73**, 5291-5300
10. Zhang, Z. J., Pedicord, V. A., Peng, T., and Hang, H. C. (2020) Site-specific acylation of a bacterial virulence regulator attenuates infection. *Nat. Chem. Biol.* **16**, 95-103
11. Crocker, P. R., and Varki, A. (2001) Siglecs, sialic acids and innate immunity. *Trends Immunol.* **22**, 337-342
12. Macauley, M. S., Crocker, P. R., and Paulson, J. C. (2014) Siglec-mediated regulation of immune cell function in disease. *Nat. Rev. Immunol.* **14**, 653-666
13. Vimr, E., Lichtensteiger, C., and Steenbergen, S. (2000) Sialic acid metabolism's dual function in *Haemophilus influenzae*. *Mol. Microbiol.* **36**, 1113-1123
14. Xiang, S. L., Zhong, M., Cai, F. C., Deng, B., and Zhang, X. P. (2006) The sialic acid residue is a crucial component of *C. jejuni* lipooligosaccharide ganglioside mimicry in the induction Guillain-Barre syndrome. *J. Neuroimmunol.* **174**, 126-132
15. Khatua, B., Ghoshal, A., Bhattacharya, K., Mandal, C., Saha, B., Crocker, P. R., and Mandal, C. (2010) Sialic acids acquired by *Pseudomonas aeruginosa* are involved in reduced complement deposition and Siglec mediated host-cell recognition. *FEBS Lett.* **584**, 555-561
16. Marques, M. B., Kasper, D. L., Pangburn, M. K., and Wessels, M. R. (1992) Prevention of C3 deposition by capsular polysaccharide is a virulence mechanism of type III group B streptococci. *Infect. Immun.* **60**, 3986-3993
17. Carlin, A. F., Lewis, A. L., Varki, A., and Nizet, V. (2007) Group B streptococcal capsular sialic acids interact with siglecs (immunoglobulin-like lectins) on human leukocytes. *J. Bacteriol.* **189**, 1231-1237
18. Pezzicoli, A., Ruggiero, P., Amerighi, F., Telford, J. L., and Soriani, M. (2012) Exogenous sialic acid transport contributes to group B streptococcus infection of mucosal surfaces. *J. Infect. Dis.* **206**, 924-931
19. Geva-Zatorsky, N., Alvarez, D., Hudak, J. E., Reading, N. C., Erturk-Hasdemir, D., Dasgupta, S., von Andrian, U. H., and Kasper, D. L. (2015) *In vivo* imaging and tracking of host-microbiota interactions via metabolic labeling of gut anaerobic bacteria. *Nat. Med.* **21**, 1091-1100
20. Hudak, J. E., Alvarez, D., Skelly, A., von Andrian, U. H., and Kasper, D. L. (2017) Illuminating vital surface molecules of symbionts in health and disease. *Nat. Microbiol.* **2**, 17099
21. Wang, W., Lin, L., Du, Y., Song, Y., Peng, X., Chen, X., and Yang, C. J. (2019) Assessing the viability of transplanted gut microbiota by sequential tagging with D-amino acid-based metabolic probes. *Nat. Commun.* **10**, 1317
22. Feng, L., Hong, S., Rong, J., You, Q., Dai, P., Huang, R., Tan, Y., Hong, W., Xie, C., Zhao, J., and Chen, X. (2013) Bifunctional unnatural sialic acids for dual metabolic labeling of cell-surface sialylated glycans. *J. Am. Chem. Soc.* **135**, 9244-9247
23. Prescher, J. A., Dube, D. H., and Bertozzi, C. R. (2004) Chemical remodelling of cell surfaces in living animals. *Nature* **430**, 873-877
24. Bull, C., Heise, T., Adema, G. J., and Boltje, T. J. (2016) Sialic acid mimetics to target the sialic acid-Siglec axis. *Trends Biochem. Sci.* **41**, 519-531
25. Moons, S. J., Adema, G. J., Derks, M. T., Boltje, T. J., and Bull, C. (2019) Sialic acid glycoengineering using N-acetylmannosamine and sialic acid analogs. *Glycobiology* **29**, 433-445
26. Gilormini, P. A., Lion, C., Vicogne, D., Levade, T., Potelle, S., Mariller, C., Guerardel, Y., Biot, C., and Foulquier, F. (2016) A sequential bioorthogonal dual strategy: ManNAI and SiaNAI as distinct tools to unravel sialic acid metabolic pathways. *Chem. Commun.* **52**, 2318-2321
27. Mbua, N. E., Li, X., Flanagan-Steet, H. R., Meng, L., Aoki, K., Moremen, K. W., Wolfert, M. A.,

- Steet, R., and Boons, G. J. (2013) Selective exo-enzymatic labeling of N-glycans on the surface of living cells by recombinant ST6Gal I. *Angew. Chem. Int. Ed.* **52**, 13012-13015
28. Arumugam, S., Chin, J., Schirrmacher, R., Popik, V. V., and Kostikov, A. P. (2011) [18F]azadibenzocyclooctyne ([18F]ADIBO): a biocompatible radioactive labeling synthon for peptides using catalyst free [3+2] cycloaddition. *Bioorg. Med. Chem. Lett.* **21**, 6987-6991
29. Vimr, E. R., Kalivoda, K. A., Deszo, E. L., and Steenbergen, S. M. (2004) Diversity of microbial sialic acid metabolism. *Microbiol. Mol. Biol. Rev.* **68**, 132-153
30. Lin, B. X., Qiao, Y., Shi, B., and Tao, Y. (2016) Polysialic acid biosynthesis and production in *Escherichia coli*: current state and perspectives. *Appl. Microbiol. Biotechnol.* **100**, 1-8
31. Rettedal, E. A., Gumpert, H., and Sommer, M. O. (2014) Cultivation-based multiplex phenotyping of human gut microbiota allows targeted recovery of previously uncultured bacteria. *Nat. Commun.* **5**, 4714
32. Huang, Y. L., Chassard, C., Hausmann, M., von Itzstein, M., and Hennet, T. (2015) Sialic acid catabolism drives intestinal inflammation and microbial dysbiosis in mice. *Nat. Commun.* **6**, 8141
33. Severi, E., Hood, D. W., and Thomas, G. H. (2007) Sialic acid utilization by bacterial pathogens. *Microbiology* **153**, 2817-2822
34. Juge, N., Tailford, L., and Owen, C. D. (2016) Sialidases from gut bacteria: a mini-review. *Biochem. Soc. Trans.* **44**, 166-175
35. Vimr, E., and Lichtensteiger, C. (2002) To sialylate, or not to sialylate: that is the question. *Trends Microbiol.* **10**, 254-257
36. Park, K. H., Kim, M. G., Ahn, H. J., Lee, D. H., Kim, J. H., Kim, Y. W., and Woo, E. J. (2013) Structural and biochemical characterization of the broad substrate specificity of *Bacteroides thetaiotaomicron* commensal sialidase. *Biochim. Biophys. Acta* **1834**, 1510-1519
37. Cieslewicz, M., and Vimr, E. (1997) Reduced polysialic acid capsule expression in *Escherichia coli* K1 mutants with chromosomal defects in kpsF. *Mol. Microbiol.* **26**, 237-249
38. Ogura, Y., Ooka, T., Iguchi, A., Toh, H., Asadulghani, M., Oshima, K., Kodama, T., Abe, H., Nakayama, K., Kurokawa, K., Tobe, T., Hattori, M., and Hayashi, T. (2009) Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **106**, 17939-17944
39. Maeda, Y., Lin, C. Y., Ishida, Y., Inouye, M., Yamaguchi, Y., and Phadtare, S. (2017) Characterization of YjjJ toxin of *Escherichia coli*. *FEMS Microbiol. Lett.* **364**
40. Gradowski, M., Baranowski, B., and Pawlowski, K. (2020) The expanding world of protein kinase-like families in bacteria: forty families and counting. *Biochem. Soc. Trans.*, BST20190712
41. Arthur, M., Molinas, C., Depardieu, F., and Courvalin, P. (1993) Characterization of Tn1546, a Tn3-related transposon conferring glycopeptide resistance by synthesis of depsipeptide peptidoglycan precursors in *Enterococcus faecium* BM4147. *J. Bacteriol.* **175**, 117-127
42. Wang, W., Zhu, Y., and Chen, X. (2017) Selective imaging of Gram-negative and Gram-positive microbiotas in the mouse gut. *Biochemistry* **56**, 3889-3893
43. Bibbo, S., Ianiro, G., Giorgio, V., Scaldaferrì, F., Masucci, L., Gasbarrini, A., and Cammarota, G. (2016) The role of diet on gut microbiota composition. *Eur. Rev. Med. Pharmacol. Sci.* **20**, 4742-4749
44. Priya, S., and Blekhnman, R. (2019) Population dynamics of the human gut microbiome: change is the only constant. *Genome Biol.* **20**, 150
45. Krinos, C. M., Coyne, M. J., Weinacht, K. G., Tzianabos, A. O., Kasper, D. L., and Comstock, L. E. (2001) Extensive surface diversity of a commensal microorganism by multiple DNA inversions. *Nature* **414**, 555-558
46. Chassard, C., and Lacroix, C. (2013) Carbohydrates and the human gut microbiota. *Curr. Opin. Clin. Nutr. Metab. Care* **16**, 453-460
47. Pacheco, A. R., Curtis, M. M., Ritchie, J. M., Munera, D., Waldor, M. K., Moreira, C. G., and Sperandio, V. (2012) Fucose sensing regulates bacterial intestinal colonization. *Nature* **492**, 113-117
48. Hooper, L. V., Xu, J., Falk, P. G., Midtvedt, T., and Gordon, J. I. (1999) A molecular sensor that allows a gut commensal to control its nutrient foundation in a competitive ecosystem. *Proc. Natl. Acad. Sci. USA* **96**, 9833-9838
49. Lagier, J. C., Khelaifia, S., Alou, M. T., Ndongo, S., Dione, N., Hugon, P., et al. (2016) Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat. Microbiol.* **1**, 16203
50. Chang, Y. C., and Nizet, V. (2014) The interplay between Siglecs and sialylated pathogens.

Glycobiology **24**, 818-825