Gut bacteria responding to dietary change encode sialidases that exhibit preference for red meat-associated carbohydrates

Livia S. Zaramela1,10, Cameron Martino1,2,10, Frederico Alisson-Silva2,3,4,9,10, Steven D. Rees3,5, Sandra L. Diaz3,4, Léa Chuzel6, Mehul B. Ganatra6, Christopher H. Taron6, Patrick Secrest3,4, Cristal Zuñiga1,10, Jianbo Huang5, Dionicio Siegel5, Geoffrey Chang5, Ajit Varki3,4 and Karsten Zengler1,3,7,8*

Dietary habits have been associated with alterations of the human gut resident microorganisms contributing to obesity, diabetes and cancer. In Western diets, red meat is a frequently eaten food, but long-term consumption has been associated with increased risk of disease. Red meat is enriched in N-glycolyneuraminic acid (Neu5Gc) that cannot be synthesized by humans. However, consumption can cause Neu5Gc incorporation into cell surface glycans, especially in carcinomas. As a consequence, an inflammatory response is triggered when Neu5Gc-containing glycans encounter circulating anti-Neu5Gc antibodies. Although bacteria can use free sialic acids as a nutrient source, it is currently unknown if gut microorganisms contribute to releasing Neu5Gc from food. We found that a Neu5Gc-rich diet induces changes in the gut microbiota, with Bacteroidales and Clostridiales responding the most. Genome assembly of mouse and human shotgun metagenomic sequencing identified bacterial sialidases with previously unobserved substrate preference for Neu5Gc-containing glycans. X-ray crystallography revealed key amino acids potentially contributing to substrate preference. Additionally, we verified that mouse and human sialidases were able to release Neu5Gc from red meat. The release of Neu5Gc from red meat using bacterial sialidases could reduce the risk of inflammatory diseases associated with red meat consumption, including colorectal cancer and atherosclerosis.

Red meat is enriched in Neu5Gc that cannot be synthesized by humans due to an evolutionary loss of a functional CMP-Neu5Ac hydroxylase (CMAH). Metabolic incorporation of Neu5Gc into tissues requires glycosidically-bound Neu5Gc for reasons currently unknown. In contrast, free Neu5Gc is utilized by gut microbes or cleared rapidly by the kidneys through the urine. Intestinal bacteria can release host-derived sialic acids from mucusal mucins and glycolipids by expressing sialidases. It is also known that gut commensal and pathogenic bacteria can use sialic acids as a carbon source. However, to our knowledge, every bacterial sialidase tested prefers N-acetylenuraminic acid (Neu5Ac) to Neu5Gc. It is currently unknown how bound Neu5Gc is metabolized in the gut. Once free, sialic acids can be taken up through membrane-associated transporters and used as carbon, nitrogen or energy sources, or used to sialylate bacterial cell surface glycans. In addition, changes in the intestinal concentration of sialic acids, for example induced by inflammation, can alter the expression of bacterial genes involved in sialic acid catabolism promoting intestinal dysbiosis. Due to the importance of sialic acids in microbe–host interaction within the gut, we investigated if a Neu5Gc-rich diet could provoke changes in bacterial metabolism and alter the gut microbiome. We compared the microbiota composition in faecal samples of Cmah−/− mice that were fed a sialic acid-free (soy) diet, a Neu5Gc-rich porcine submaxillary mucin (PSM) diet or a Neu5Ac-rich edible bird’s nest (EBN) diet (Supplementary Table 1).

Faecal contents were scraped from the mouse colons, and the microbiome of Cmah−/− mice and wild-type (WT) mice fed soy, PSM or EBN diets was determined through 16S rRNA gene amplification sequencing. Bray–Curtis dissimilarity showed a significant difference in the bacterial genotypes present in Cmah−/− and WT mice, indicating that a mouse’s inability to synthesize endogenous Neu5Gc significantly impacted microbiome composition (Fig. 1a) and substantiated the need for a human-like Cmah−/− mouse model in our experimentation. We found that changes in the microbial composition in Cmah−/− mice were diet-dependent, with Clostridiales and Bacteroidales contributing significantly to the variations observed amongst the diets (Fig. 1b and Supplementary Fig. 2). Independent of genotype, the microbiome of PSM-fed mice was significantly less diverse compared to the microbiome of those fed soy and EBN diets (P < 0.05) (Supplementary Fig. 3a). Human-like Cmah−/− mice revealed similar taxonomic profiles at the family level amongst the three diet groups (Supplementary Fig. 3b). However, at the genus level, Helicobacter, Intestinimonas and Candidatus Saccharibacteria genera incertae sedis were significantly enriched in the EBN group compared to PSM (Fig. 1c). Moreover, Bacteroides, Barnesiella, Clostridium group III, Parabacteroides, Roseburia and Turicibacter were significantly enriched in the PSM group compared to EBN

1Department of Pediatrics, University of California, San Diego, CA, USA. 2Bioinformatics and Systems Biology Program, University of California, San Diego, CA, USA. 3Department of Medicine and Cellular and Molecular Medicine, University of California, San Diego, CA, USA. 4Glycobiology Research and Training Center, San Diego, CA, USA. 5Krieger School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, CA, USA. 6New England Biolabs, Ipswich, MA, USA. 7Department of Bioengineering, University of California, San Diego, CA, USA. 8Center for Microbiome Innovation, University of California, San Diego, CA, USA. 9Present address: Paulo de Goes Institute of Microbiology, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil. 10These authors contributed equally: Livia S. Zaramela, Cameron Martino, Frederico Alisson-Silva. *e-mail: kzenzler@ucsd.edu

2082 NATURE MICROBIology | VOL 4 | DECEMBER 2019 | 2082-2089 | www.nature.com/naturemicrobiology
Fig. 1 | Composition of gut microbial community of mice fed on soy, PSM or EBN diet. a, Beta-diversity analysis of WT versus Cmah−/− mouse. Pairwise Bray–Curtis dissimilarities were plotted against the first and second principal coordinates (PC1 and PC2) (ANOSIM $R = 0.979$, $P = 0.001$). WT samples include $n = 5$ biologically independent animals and $n = 15$ independent experiments per diet. Cmah−/− samples include $n = 3$ biologically independent animals and $n = 9$ independent experiments. b, Beta-diversity analysis of Cmah−/− mouse fed on soy, PSM or EBN diet. Pairwise Bray–Curtis dissimilarities were plotted against PC1 and PC2 (ANOSIM $R = 0.831$, $P = 0.001$). Grey arrows represent the significant vector fitting with each of the PCoA (principal coordinates analysis) ordinations. The most representative taxa are indicated on the plot. c, Differentially abundant bacterial genera on PSM and EBN diet. The increased abundance of bin13 in PSM suggests its prominent role in Neu5Gc metabolism in the gut. As such, bin13 sialidases might possess biochemical properties that make them better suited for Neu5Gc release. To test this notion, each of the five sialidases genes in bin13 were amplified by PCR (Supplementary Fig. 9). Additionally, the examination of individual sialidase genes revealed several diet-dependent sialidases (Fig. 1d). To evaluate the sialidases genes present in the microbiome, the combined metagenomes were co-assembled and 51 genome bins containing 21 sialidase genes were identified (Supplementary Fig. 7, Supplementary Table 3). Amongst the bins with annotated sialidase genes, bin13, whose closest relative was B. thetaiotaomicron (Supplementary Table 3), was the most abundant in PSM compared to EBN diets (Fig. 1e).
and heterologously expressed in vitro, purified and assayed for its substrate preference. Sialidase activity measurements were performed using different enzyme concentrations (0.5–10 µg) at three different pHs (6.5, 7 and 8) (Fig. 2a and Supplementary Fig. 10). Four out of five sialidases showed preferential Neu5Gc activity in at least one of the pHs tested (Fig. 2a). To the best of our knowledge, no previously characterized exo-sialidases have been shown to prefer Neu5Gc over Neu5Ac28–30 (Supplementary discussion). Additionally, we tested in vivo sialidase activity in fresh faecal samples. Clarified faecal pellets from mice fed with PSM tend to preferentially release Neu5Gc compared to clarified faecal pellets from mice fed with soy (Fig. 2d).

The most compelling enzyme from the substrate specificity study in all tested conditions was sialidase26. Sialidase26 had protein sequence motifs characteristic of the GH33 family of sialidases (Supplementary Figs. 11 and 12). Despite this substrate preference, sequence residues that are predicted to interact with terminal sialic acids in the catalytic site are highly conserved with structurally studied sialidases exhibiting unknown Neu5Gc preference over Neu5Ac (Supplementary Fig. 13 and Supplementary discussion). To explain the structural underpinnings of Neu5Gc preference, we used X-ray crystallography to determine the structure of sialidase26 both alone (PDB 6MRX, 2.0 Å resolution) and in complex with the inhibitors DANA-Ac (N-acetyl-2,3-dehydro-2-deoxyneuraminic acid) and DANA-Gc (N-glycolyl-2,3-dehydro-2-deoxyneuraminic acid) (Protein Data Bank (PDB) 6MRV, 1.8 Å resolution and PDB 6MYV, 2.2 Å resolution, respectively) (Supplementary Table 4). Sialidase26 structure is common to GH33 sialidases (Fig. 2b), including Y509 nucleophilic engagement of C5 following E398 charge activation, D228 acid–base catalysis of the glycosidic bond at C5, and C2 stabilization by an Arg triad (R203, R414, R478). However, typical sialic acid stabilizing interactions are lost, including Glu engagement of W507 R478 Y509 R414 E398 Q364 W299 F343 D228 D271 D271.
the glycerol moiety C7-C9 (T397 in sialidase26), C10 stability by an Arg residue (now a Leu), and inward movement of W507 into the binding pocket (although this is restored in the DANA-Gc co-crystal structure). Co-crystallization of sialidase26 with DANA-Gc indicated an overall fold and ligand placement similar to that of DANA-Ac (Fig. 2c). The hydroxyl group at the end of the C5 acetamido group of DANA-Gc points towards the binding pocket residues, forming H contacts (hydrogen bonds) with D271 and probably increasing its stability. Amino acid substitutions at this position to Leu (D271L) or Asn (D271N) significantly lowered sialidase protein activity and eliminated the Neu5Gc preferential cleavage (Fig. 2a). This, in combination with the changes described above, probably explains sialidase26 preference for Neu5Gc.

To extend our findings to humans, we reanalysed faecal shotgun metagenomes from the Hadza (Supplementary Table 5), a genetically distinct indigenous ethnic group that resides in remote Tanzania31. The Hadza are hunter–gatherers and change their diet periodically throughout the year according to food availability. In dry seasons, their diet is enriched with meat and tubers, whereas in wet seasons it consists largely of honey and berries31,32. Raw reads from Hadza metagenomes were mapped to bin13. We observed that bin13 was significantly more abundant in microbiome samples taken during the dry season compared to the wet season (Fig. 3a). To evaluate the sialidases genes present in the Hadza microbiome, the combined metagenomes were co-assembled and 24 genome bins containing 51 sialidase genes were identified (Supplementary Table 5).
letters

Fig. 14, Supplementary Table 6). The binHz19, whose closest relative was Alconstipes sp., was equally abundant in microbiomes from both seasons (Fig. 3b and Supplementary Table 6). BinHz19 also contained a sialidase (sialidaseHz136) with the greatest sequence similarity to sialidase26 identified in our mouse study, suggesting widespread distribution of this sialidase homologue amongst mammals (Supplementary Fig. 15). SialidaseHz136 showed preferential activity for Neu5Gc over Neu5Ac in all tested conditions (Fig. 3c). X-ray crystallography of sialidaseHz136 (PDB 6MN), 2.2 Å resolution) exhibited similar predicted engagement of conserved residues with sialic acid substrate when compared to sialidase26 (Fig. 3d), despite a shift in the acetamido-interacting Asp residue. This difference supports the notion that sialidaseHz136 is capable of metabolizing structurally diverse glycans, a concept that would be of benefit to an individual with seasonal variations in diet. To determine if sialidase26 and sialidaseHz136 can release Neu5Gc from food sources directly we tested their activity on beef, pork and PSM chow as substrate. Both enzymes showed pronounced sialidase activity by releasing Neu5Gc from all food tested (Fig. 3e).

Our findings indicate that sialidases exhibiting Neu5Gc preference can be widespread in the mammalian intestine. To determine if Neu5Gc-specific sialidases are restricted to enteric bacteria, we used functional metagenomic screening to identify exo-sialidases with a Neu5Gc specificity in a terrestrial environment. A fosmid library containing ~40 kilobase (kb) inserts of total environmental DNA isolated from soil from an organic composting facility was constructed in Escherichia coli. Lysates from 5,576 clones were screened for hydrolysis of 4MU-α-Neu5Gc (Supplementary Fig. 16). Clones showing activity were tested subsequently and re-tested for their ability to cleave 4MU-α-Neu5Ac. Two clones (C19 and C22) showed significant hydrolysis of 4MU-α-Neu5Gc but only minor activity on 4MU-α-Neu5Ac (Fig. 3f). The DNA sequence of cloned inserts from C19 and C22 encoded a single bacterial sialidase gene. Both enzymes showed exo-sialidase activity and had a marked preference for Neu5Gc hydrolysis in vitro compared to several other known bacterial exo-sialidases (Fig. 3f). Analysis of the deduced amino acid sequences of C19 and C22 showed that they are highly similar to each other (56% identity), both belong to the G43 family of bacterial exo-sialidases (Supplementary Fig. 17) and are most similar to proteins from terrestrial bacteria of the genera Rhodopirellula and Verrucomicrobia. Considered together, these data support the conclusion that exo-sialidases with Neu5Gc preference also exist in terrestrial environments where degrading biologic material from animal sources might require such preference.

In summary, we identified several sialidases with previously undescribed preference for Neu5Gc, which are enriched in the gut microbiota of mice and humans on consumption of a Neu5Gc-rich diet. Previous studies from our group showed that dietary free Neu5Gc is not incorporated in Cmah−/− (ref. 11) mice. Thus, the cleavage of Neu5Gc from foods entering the gut can potentially prevent incorporation of this non-human sugar into the colonial tissue. We also hypothesize that the gut microbiome with an under-representation of bacteria with Neu5Gc-prefering sialidases could result in increased xenosialitis and be a potential contributing factor to inflammation-mediated promotion of diseases. Although further in vivo characterization needs to be done, our results lay the foundation to define a strategy for translation of pre- or probiotics to prevent incorporation or to eliminate Neu5Gc from tissues of red meat eaters, thereby reducing the risk of xenosialitis and other diseases associated with red meat consumption.

Methods

In vivo sampling, feeding and animal diet. WT lineage C57BL/6 was purchased from Harlan Laboratories and human-like Cmah−/− mice generated as previously described11. In brief, the generation of the human-like Cmah−/− transgenic mice was performed by target deletion of exon 6 of the Cmah gene (similar deletion that evolved in humans) using loxp sites andCre recombinase expression in embryonic stem cells. Transgenic mice were generated by the University of California, San Diego (UCSD), Transgenic Mouse Core. All animal experiments were approved by the UCSD Institutional Animal Care and Use Committee under the protocol number S01227. All animals were maintained in the UCSD vivarium according to guidelines, with 12-h diurnal lighting and access to food and water ad libitum. Sample processing and analysis were not blinded at any step. Samples size was chosen based on the cost, mice, diet and vivarium availability, and the minimum number of samples to perform statistical analysis. WT and Cmah−/− mice were raised in the same vivarium room and fed with the same water source. The cages were kept side by side in the same cage rack to minimize external influence on the gut microbiome. Age and sex-matched female mice used in the study were maintained on sialic acid free-soy based diet (Dyets; 110951) from weaning until 8–10 weeks of age to prevent previous exposure to sialic acid. To evaluate the effect of the dietary Neu5Gc in the gut microbiome, sex-matched Cmah−/− mice at 10 weeks of age were caged in three groups of five mice each and fed during 4 weeks with the same soy-based diet that was either enriched with Neu5Gc (PSM), Neu5Ac (EBN) or kept in sialic acid-free soy as control. The animals were euthanized in a CO2 chamber and the colon tissues were cut open with blunted scissors for fresh collection of faecal samples by scraping it straight from the tissue. Colonic faecal scrappings from five mice of each diet type were used for 16S rRNA gene amplification and three mice of each diet type were used for metagenomic shotgun sequencing analysis. The feeding protocol was chosen based on previous evidence from our group showing that feeding Cmah−/− mice with PSM over a period of weeks can cause mouse tissue incorporation of Neu5Gc at levels historically similar to the levels seen in adult humans who have eaten red meat for many years12. The PSM diet contained 250 µg of Neu5Gc per gram of chow to mimic the amount of Neu5Gc present in beef, the most consumed form of red meat in the Western diet.

Monosaccharide and amino acid composition of the diets. Monosaccharide composition of the PSM and EBN diets was analysed by High-Performance Anion-Exchange Chromatography Coupled with Pulsed Electrochemical Detection (HPAEC-PAD) using Dionex ICS-3000 system (Thermo Fischer Scientific) equipped with CarboPac PA1 column 4 × 250 mm, 4 µm, with a 4 × 50 mm Guard. Briefly, 500 µg of each diet was dissolved in 200 µl of Milli-Q water. The samples were hydrolysed by adding an equal volume of 4N Trifluoroacetic acid (final concentration of 2N TFA) at 100°C for 4h. The hydrolysed samples were centrifuged at 400g for 2 min and evaporated under a flow of dry nitrogen. Once dried, samples were resuspended in 200 µl of Milli-Q water and 50% of each sample was injected. The separation of monosaccharide peaks was achieved using the solvents water (A), 100 mM NaOH with 5 mM NaOAc (B) and 100 mM NaOH with 250 mM NaOAc (C) with pulsed amperometric detector and standard quad wave form, and the gradient conditions in Table 1. The amino acid composition of both diets was performed by gas chromatography mass spectrometry (GC-MS) 1BDMS (dimethyl-tart-butyl-silyl) derivatives quantitation as previously described13. HPAEC-PAD and GC-MS analysis were performed by the Glycotechnology Core at the Glycobiology Research and Training Center, UCSD (https://medschool.ucsd.edu/research/GRTC/services/glycometrics/Pages/default.aspx).

Table 1 | Solvents and gradient conditions used for high-performance anion exchange chromatography

| Solvents          | Gradient setting |
|-------------------|------------------|
| A: Water          | Time %A %B %C |
| B: 100 mM NaOH with 5 mM NaOAc | 0 84 16 0 |
| C: 100 mM NaOH with 250 mM NaOAc | 20 84 16 0 |
|                  | 21 0 100 0      |
| Pulsed amperometric detector | 31 0 100 0      |
| Waveform: standard quad | 32 84 16 0      |
|                   | 50 84 16 0      |

16S rRNA sequencing and analysis. Total genomic DNA was extracted using MoBio PowerFecal DNA isolation kit (MoBio) following the manufacturer’s instructions. Purified DNA was amplified and processed according to the Illumina 16S protocol (https://support.illumina.com/documents/documentation/chemistry_dry-prep-guide-16/16s-metagenomic-library-prep-guide-15044223-b.pdf). 16S rRNA libraries were generated from three or five biologic replicates and three independent experiments per diet group. Libraries were quality assessed using quantitative PCR (qPCR) and Bioanalyzer (Agilent Technologies), and subsequently sequenced using two MiSeq 600 cycle kits (Illumina). Adapters were trimmed from the Illumina data using Trimmomatic v.0.36 (ref. 14). 16S analysis
was performed with Usearch denoising and the RDPRibosomal Database Project (RDP) 0.89 rRNA database v.16. Sequences were analysed using the Usearch v.10 (ref.) following the MiSeq 2x250 16S V4 pipeline (https://www.drive5.com/usearch/manual/examples.html). In brief, paired-end reads were merged using fastq-mergears (-fastq_maxdiffs 10; -fastq_pctid 10). Sequences with a distance-based similarity of 97% or greater were grouped into operational taxonomic units (OTU) using cluster_otsu (−minsize 2). OTU tables were rarefied to 10,000 observations per sample. OTU-based microbial diversity and dissimilarity metrics were estimated using R package vega2 v.2.5.2. Bray-Curtis distance was used for beta-
diversity analysis. Vector fitting with each of the PCoA ordinations was performed using the function envfit from R package vegan v.2.5-2. Significant vectors were selected following the criteria (R ≥ 0.7 and P ≤ 0.01). Bacteria genera with relatively abundant below 1% were not considered for differential abundance analysis. Statistical differences in abundance between diets were calculated using the non-
parametric Mann-Whitney U rank sum test with Holm correction for multiple hypotheses when appropriate.

Shotgun metagenome sequencing and analysis. As described above, total genomic DNA was extracted using MoBio PowerFecal DNA isolation kit (MoBio) following the manufacturer's instructions. Purified DNA from three biological replicates per diet group was prepared for shotgun metagenome sequencing using the Nextera XT library preparation method with the average fragment size of 450 base pairs (bp) (Illumina). Libraries were quality assessed using qPCR and a Bioanalyzer (Agilent Technologies) and subsequently sequenced using MiSeq 2x 250 bp. An average of 2.5 million non-mouse reads were generated per library. Adaptors were trimmed from the Illumina data using Trimomatic (v.0.36). Sequences were filtered of possible human and mouse contamination by aligning the trimmed reads against reference databases using Bowtie2 v.2.2.3 (ref.) with the following parameters (−D 20 −R 3 −N 1 −L 20
−very-sensitive-local). Overlapped reads were merged using FLASH v.1.2.11 (ref.). Merged and unmapped reads were assembled using SPAdes with the following parameters (−k 21,33,55,77,99,127 −meta −merge). Differential binning was performed using MetaBAT2 v.2.12.1 with minimum contig length of 1,500 bp (ref.). Bin quality (completeness and contamination) was evaluated using CheckM v.1.0.7 (ref.). Taxonomic classification (closest phylogenetic neighbour) was assessed using the RAST (rapid annotation using subsystem technology) interface P3-scripts.

Analysis of previously published shotgun metagenomic data. Previously published focal shotgun metagenomic data from Hadza hunter-gatherer individuals were obtained from Sequence Read Archive repository under the project IDs PRJNA393210 and PRJNA393218. Forty shotgun metagenomic data from individuals were analysed, 20 samples of which were collected during the wet season and 20 during the dry season (Supplementary Table 5). All published shotgun metagenomic data were processed as described in the previous section.

Protein expression and assay. Target sialidase sequences from shotgun metagenomic data were PCR amplified from genomic DNA isolated as described above or ordered from Integrated DNA Technologies, subcloned into a pET19b expression vector with a C-terminal 10XHis tag and N-terminal truncation to remove any signal peptide sequence (predicted by SignalP4.1, CBS) and transformed into the E. coli (MilliporeSigma) using electroporation methods. Cells were grown to optical density 0.6–0.8 (OD 600) in multiple 1-L cultures at 37 °C and induced overnight at 25 °C with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG). Harvested cells were resuspended in lysis buffer (50 mM HEPES pH 8.0 buffer, 50 mM NaCl and 1 mM TCEP (tris(2-carboxyethyl) phosphine hydrochloride)) with DNasel and hen egg white lysozyme, lysed with a T5-Series cell disruptor (Constant Systems) at 15 KPSI (kilo-pound per square inch), and spun for 45 min at 186,000g with a Ti45 ultracentrifugation rotor (Beckman Coulter) to remove cell debris. Purification was performed as described below based on purification of a putative Bacteroides neuraminidase that was provided by the Protein Structure Initiative (BACCAC_01090, Joint Center for Structural Genomics, to be published), with modifications to imidazole stringency based on the purified sialidase. Supernatant was loaded on a 5-ml HisTrap Ni affinity column (nicking-charged columns for high resolution histidine-tagged protein purification) on an Akta Explorer purification system (GE Healthcare Life Sciences) with 20–40 mM imidazole added, washed with Running Buffer (50 mM HEPES pH 8.0 buffer, 300 mM NaCl, 40–60 mM imidazole, 1% glycerol and 1 mM TCEP) and eluted with Eulution Buffer (20 mM HEPES pH 8.0 buffer, 300 mM imidazole, 1% glycerol and 1 mM TCEP). Samples were concentrated using 10–30 kDa Amicon centrifugal filters (MilliporeSigma) at 1,500g for 1 ml and dialyzed in 5 ml desalting buffer using the Akta system into Running Buffer (20 mM HEPES pH 8.0 buffer, 200 mM NaCl). The resulting protein sample was diluted as required for functional studies.

Assay for sialidase activity. In vitro activity. Sialidases purified as described above were quantified using SDS–PAGE image analysis with BSA (bovine serum albumin) references (Bio-Rad) and absorbance using the Nanophotometer P320 (Implen), with extinction coefficients calculated with ExPASy Translate (https://web.expasy.org/translate). Sialidase activity assays were performed in a dilution series: 0.5 µg, 2.5 µg, 5 µg and 10 ng of each enzyme were incubated with equal amounts of 2ACNA, 2ACNB and mouse serum non-reducing terminal sialyl oligosaccharides (0.64 µmol l−1, Neu5Gc: 0 pmol l−1 and WT mouse serum (Neu5Ac: 196.29 µmol l−1, Neu5Gc: 1,337.45 pmol l−1) for 1 h at 37 °C. An additional 10 µg of each enzyme was inactivated by heat for 5 min at 95 °C. The samples were kept at −20 °C until derivatization and analysed by HPLC as described below. The DMB reagent was made with the following recipe: 14 mM DMB (1,2-diamino-
4,5-methylenedioxybenzene, Sigma D4787), 1.8 mM sodium hydrosulfite (Sigma 157953), 1.0 M 2-mercaptoethanol (Sigma M3148) and 40 mM trifluoroacetic acid (Sigma T6508), and incubated at 50 °C for 2.5 h. The DMB-derived samples were analysed on a Dionex Ultra3000 HPLC System using a Phenomenex Gemini 5 μC18 250 μX 4.6-mm HPLC column at room temperature, eluted in isocratic mode with 85% water, 7% methanol and 8% acetic acid. The same protocol was used to evaluate the sialidases activity in three different pH ranges.

In vivo activity. In vivo activity using clarified faecal samples: Faecal pellets were collected from 10-week-old C57Bl/6J mice caged in two groups of four mice and three mice, which were fed over 2 weeks with PSM or sialic acid-free soy, respectively. All the animals’ information and maintenance are the same as described in In vivo sampling, feeding and animal diet. Fresh faecal pellets were homogenized (10% w/v) in Tris–HCl pH 7 buffer with protease inhibitor (Roche 1183617001) using Pellet Pestle Cordless Motor (Kimble 794590). Samples were spun at 4,000g for 3 min and the supernatant was transferred to SpinX filters (Costar 3600). Samples were centrifuged at 14,000g for 10 min and the flow through was transferred to a clean tube. Total protein concentration was determined using Pierce BCA (bicinchoninic acid) Protein Assay (Thermo Scientific 23225). To measure sialidase activity in clarified faecal samples, 2 µg of total protein per sample was incubated with equal amount of human-like CAZymes (CAZy database) using the following recipe for 2 h at 37 °C: 10 ng of each enzyme was incubated with 2 µg of total protein per sample was inactivated by heat for 5 min at 95 °C. Samples were cleaned using 10 kDa Amicon centrifugal filters (MilliporeSigma), 14,000g for 25 min at 4 °C. Samples were lyophilized and resuspended in 50 µl of sterile water. Derivatization and HPLC analysis was performed as described in In vivo activity.

Food sources. Beef (New York steak) and pork (breakfast sausage) were purchased at Whole Foods Market and cut into small pieces using sterile blades. PSM chow was powdered using a crucible and pestle. Beef, pork and PSM were homogenized (10% w/v) in Tris–HCl pH 7 buffer with protease inhibitor (Roche 1183617001) using Pellet Pestle Cordless Motor (Kimble 794590)). Samples were spun at 4,000g for 3 min and the supernatant was transferred to SpinX filters (Costar 3600). Samples were centrifuged at 14,000g for 10 min and the flow through was transferred to a clean tube. Total protein concentration was determined using Pierce BCA (bicinchoninic acid) Protein Assay (Thermo Scientific 23225). To measure sialidase activity in clarified faecal samples, 2 µg of total protein per sample was incubated with equal amount of human-like CAZymes (CAZy database) using the following recipe for 2 h at 37 °C: 10 ng of each enzyme was incubated with 2 µg of total protein per sample was inactivated by heat for 5 min at 95 °C. Samples were cleaned using 10 kDa Amicon centrifugal filters (MilliporeSigma), 14,000g for 25 min at 4 °C. Samples were lyophilized and resuspended in 50 µl of sterile water. Derivatization and HPLC analysis was performed as described in In vivo activity.

Neu5Gc2en synthesis. Neu5Gc2en (DANA-Gc) was synthesized as previously published with minor modifications. Briefly, Neu5Gc (Sigma–Aldrich) was treated with Dowex 50W-X8 (H+)-form resin in MeOH for 2 h at 20 °C to form the methyl ester. This ester was then treated with acetic anhydride and pyridine for 4.25 h at 20 °C to generate the milliporeperacetyl methyl ester, which was purified by column chromatography (50:1 CHCl3–MeOH). This sample was treated with TMSOTf under dry nitrogen at 0 °C in MeCN for 6 h to induce elimination. The unsaturated compound was obtained by chromatography (toluene/acetone, 3:1–2:1). The acetyl groups were cleaved by treatment with NaOH in MeOH over 12 h, followed by neutralization with H2 resin.
Crystallization and structure determination of sialidase26 and sialidaseH136e. Purified sialidase26 was concentrated to 8 g l−1 and set on sitting drop trays in 1:1 volume ratios with the mother liquor (20% PEG (polyethylene glycol) 6000, 0.1 M Tris–HCl pH 8.0) at 4°C. For ligand co-crystals, concentrated sialidase26 was incubated for 1 h at room temperature with 5 mM N-acetyl-2,3-dehydro-2-deoxyxoyuronic acid (Neu5Ac2en or DANA, Sigma-Aldrich) or DANA-Gc (synthesized). Crystals appeared after 3–4 d and grew to full size in 8–10 d.

2-deoxyneuraminic acid (Neu5Ac2en or DANA, Sigma-Aldrich) was incubated for 1 h at room temperature with 5 mM λ-excitation Plus 384 Microplate Reader (Molecular Devices) at 6 h, 24 h and 48 h. Positive PHENIX.REFINE in the PHENIX software suite v.1.13 (ref. 50) against a maximum density maps in Coot v.0.8.9.1 (ref. 51). Geometry restraints for DANA were performed with PHENIX.MULTI_CRYSTAL_AVERAGE for five cycles following experimental constraints for 20 sugars, 24 fibre-related metabolites, 12 fatty acids, 20 amino acids and 88 minerals, vitamins and other metabolites. The statistical significance of sialidase activity was computed using Student’s two-tailed t-test. The significance levels are *P<0.05, **P<0.01 and ***P<0.001. All analysis was performed in biologically independent animals or independent experiments as indicated in the text. Samples were not randomized. All data were used in the analysis.

Reconstructed metabolic models analysis. Genome-scale network reconstructions combine detailed biochemical and physiological information, providing insights into the metabolism for subsequent manipulation strategies or to control metabolism53. The scope of these models encompasses the characterization of the metabolic behaviour of target microorganisms. We evaluated growth phenotypes of microorganisms with sialidase associated with sialidase metabolism (Enzyme Commission number 3.2.1.18). Seventeen microorganisms of the gut microbiota containing sialidases were identified by screening the repository of the gut microbiome metabolic models5. Growth rates were simulated using flux balance analysis. All metabolic models were constrained using a Western diet (45% fat, 35% carbohydrate, 20% protein), containing experimental constraints for 20 sugars, 24 fibre-related metabolites, 12 fatty acids, 20 amino acids and 88 minerals, vitamins and other metabolites. Experimental constraints are reported in detail in the gut microbiota repository5. All metabolic models were simulated using the Gurobi Optimizer v.5.6.3 (Gurobi Optimization) solver in MATLAB (The MathWorks) with the COBRA (constraint-based reconstruction and analysis) Toolbox5. Additionally, the contribution to growth of the metabolites associated with sialidase activity (for example, N-acetylneuraminic acid) was determined using shadow prices simulations.

Statistics and reproducibility. Statistical analysis was performed using R programming language. The statistical significance of differential relative abundance (16S rRNA amplicon and shotgun metagenomics) was computed using no-parametric two-sided Wilcoxon rank sum test with Holm correction for multiple hypotheses. The statistical significance of sialidase activity was computed using Student’s two-tailed t-test. The significance levels are *P<0.05, **P<0.01 and ***P<0.001. All analysis was performed in biologically independent animals or independent experiments as indicated in the text. Samples were not randomized. All data were used in the analysis.

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

Data availability
Sequencing data supporting the findings of this study are available under accession number PRJNA505660. X-ray crystallographic data that support the findings of this study have been deposited in the RCSB (Research Collaboratory for Structural Bioinformatics) Protein Data Bank (accession codes: 6MRX, 6MRY, 6MYV and 6MN1).

Code availability
The code used to generate the figures and for statistical analysis can be accessed from the corresponding author upon request.

Received: 16 January 2019; Accepted: 16 August 2019;
Published online: 23 September 2019

References
1. Hall, A. B., Tolonen, A. C. & Xavier, R. J. Human genetic variation and the gut microbiome in disease. Nat. Rev. Genet. 16, 600–669 (2017).
2. Mann, N. Dietary lean red meat and human evolution. Evol. J. Nutr. 39, 71–79 (2000).
3. Etemadi, A. et al. Mortality from different causes associated with meat, heme iron, nitrates, and nitrates in the NIH-AARP Diet and Health Study: population based cohort study. BMJ 357, 1–11 (2017).
4. Alisson-Silva, F., Kawaishi, K. & Varki, A. Human risk of diseases associated with red meat intake: analysis of current theories and proposed role for metabolic incorporation of a non-human sialic acid. Mol. Aspects Med. 51, 16–30 (2016).
5. Sunraj, A. N. et al. A red meat-derived glycine promotes inflammation and cancer progression. Proc. Natl Acad. Sci. USA 112, 542–547 (2015).
6. Tongvarananukul, P. et al. Human uptake and incorporation of an immunogenic non-human dietary sialic acid. Proc. Natl Acad. Sci. USA 100, 12045–12050 (2003).
7. Sunraj, A. N., Laumb, H., Varki, N. & Varki, A. Involvement of a non-human sialic acid in human cancer. Front. Oncol. 4, 1–13 (2014).
8. Varki, A. Uniquely human evolution of sialic acid genetics and biology. Proc. Natl Acad. Sci. USA 107, 8399–8406 (2010).
9. Dhar, C., Sasmal, A. & Varki, A. From ‘serum sickness’ to ‘xenolitiasis’: past, present, and future significance of the non-human sialic acid Neu5c. Front. Immunol. 10, 807 (2019).
10. Almagro-Moreno, S. & Boyd, E. F. Sialic acid catabolism confers a competitive advantage to pathogenic Vibrio cholerae in the mouse intestine. Infect. Immun. 77, 3807–3816 (2009).
11. McDonald, N. D., Lubin, J.-B., Chowdhury, N. & Boyd, E. F. Host-derived sialic acids are an important nutrient source required for optimal bacterial fitness in vivo. MBio 7, e02237–15 (2016).
12. Lewis, A. L. & Lewis, W. G. Host sialoglycans and bacterial sialidases: a mucosal perspective. Cell. Microbiol. 14, 1174–1182 (2012).
13. Kawanishi, K. et al. Human species-specific loss of CMP-N-acetylneuraminic acid hydrolyase enhances atherosclerosis via intrinsic and extrinsic mechanisms. Proc. Natl Acad. Sci. USA 116, 16036–16045 (2019).
14. Bandi, K., Gregg, C. J., Chow, R., Varki, N. M. & Varki, A. Metabolism of vertebrate amino sugars with N-glycolyl groups: mechanisms underlying gastrointestinal incorporation of the non-human sialic acid xeno-autoantigen N-glycoly neuraminic acid. J. Biol. Chem. 287, 28852–28864 (2012).
15. Almogu-Moreno, S. & Boyd, E. F. Insights into the evolution of sialic acid catabolism among bacteria. BMC Evol. Biol. 9, 118 (2009).
16. Almogu-Moreno, S. & Boyd, E. F. Bacterial catabolism of non-ulosonic (sialic) acid and fitness in the gut. Gut Microbes 1, 45–50 (2010).
17. Li, J. & McClane, B. A. Nani sialidase can support the growth and survival of Clostridium perfringens strain F4996 in the presence of sialylated host macromolecules (mucin) or Caco-2 cells. Infect. Immun. 86, e00547–17 (2018).
18. Tailford, L. E. et al. Discovery of intracellular trans-sialidases in human gut microbiota suggests novel mechanisms of metabolic adaptation. Nat. Commun. 6, 7624 (2015).
19. Kim, S., Oh, D. B., Kang, H. A. & Kwon, O. Features and applications of bacterial sialidases. Appl. Microbiol. Biotechnol. 91, 1–15 (2011).
20. Juge, N., Tailford, L. & Owen, C. D. Sialidases from gut bacteria: a mini-review. Biochem. Soc. Trans. 44, 166–175 (2016).
21. Chokhawala, H. A., Yu, H. & Chen, X. High-throughput substrate specificity studies of sialidases using chemoselectively synthesized sialoside libraries. Chembiochem 8, 194–201 (2007).
22. Huang, Y. L., Chassard, C., Hausmann, M., Von Itzstein, M. & Hennet, T. Sialic acid catabolism drives intestinal inflammation and microbiotib dysbiosis in mice. Nat. Commun. 6, 8141 (2015).
23. Hedlund, M. et al. N-glycoly neuraminic acid deficiency in mice: implications for human biology and evolution. Mol. Cell. Biol. 27, 4340–4346 (2007).
24. David, L. A. et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature 505, 559–563 (2014).
25. Flint, H. J., Scott, K. P., Duncan, S. H., Louis, P. & Forano, E. Microbial degradation of complex carbohydrates in the gut. Gut Microbes 3, 289–306 (2012).
26. Magnusdóttir, S. et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. Nat. Biotechnol. 35, 81–89 (2017).
27. Huang, L. et al. DbcAN-seq: a database of carbohydrate-active enzyme (CAZYyme) sequence and annotation. Nucleic Acids Res. 46, D516–D521 (2018).
28. Owen, C. D. et al. Unravelling the specificity and mechanism of sialic acid recognition by the gut symbiont Ruminococcus gnavus (CAZyme) sequence and annotation. Appl. Environ. Microbiol. 78, 2114–2120 (2012).
29. Almogu-Moreno, S. & Boyd, E. F. Bacterial catabolism of non-ulosonic (sialic) acid and fitness in the gut. Gut Microbes 1, 45–50 (2010).
30. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D 66, 213–221 (2010).
31. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D 66, 486–501 (2010).
32. Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. Electronic ligand builder and optimization workbench (eLBOW): a tool for ligand coordinate and restraint generation. Acta Crystallogr. D 65, 1074–1080 (2009).
33. Tan, J., Zuniga, C. & Zengler, K. Unraveling interactions in microbial communities—from co-cultures to microbiomes. J. Microbiol. 53, 295–305 (2015).
34. Zengler, K. & Zaramela, L. S. The social network of microorganisms—how axotrophs shape complex communities. Nat. Rev. Microbiol. 16, 383–390 (2018).
35. Zuñiga, C., Zaramela, L. S. & Zengler, K. Elucidation of complexity and prediction of interactions in microbial communities. Microb. Biotechnol. 10, 1500–1522 (2017).
36. Schellenberger, J. et al. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. Nat. Protoc. 6, 1290–1307 (2011).

Acknowledgements
We thank all Zengler-, Varki- and Chang lab members for helpful discussions. Research was supported in part by the National Institutes of Health under award no.R01GM13273 (to A.V.) and by the National Science Foundation under award no. 105-144435 (to G.C.). C.M. was supported by grants from the National Institutes of Health, USA (NIH grant no. T32GM080689) and by a Chancellor’s Research Excellence Scholarship (UCSD). F.A.-S. was partly supported by the Program Science Without Borders Bex 9254-13-7-Capes Brazil.

Author contributions
L.S.Z., C.M., F.A.-S., S.R., S.L.D., G.C., A.V. and K.Z. have filed a patent application with constraint-based models: the COBRA Toolbox v2.0. Nat. Protoc. 6, 1290–1307 (2011).

Competing interests
L.S.Z., C.M., F.A.-S., A.V. and K.Z. conceptualized the study. L.S.Z., F.A.-S. and K.Z. wrote the manuscript with input from all authors. L.S.Z., C.M. performed and performed the metabolic model analysis. L.C., M.B.G. and C.H.T. constructed the fosmid library and restraint generation. Acta Crystallogr. D 65, 1074–1080 (2009).

Additional information
Correspondence and requests for materials should be addressed to K.Z.
Reprints and permissions information is available at www.nature.com/reprints.
Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019
Reporting Summary

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

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

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

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

Metagenomic libraries were sequenced by La Jolla Institute for Immunology Sequencing Core and Illumina Basespace (online platform) was used to collect the raw data (fastq files). Sequencing data supporting the findings of this study are available under accession number PRJNA505660.

Data analysis

Trimmomatic version 0.36, Usearch v10, R package vegan v2.5-2, Bowtie2 v2-2.2.3, Flash version 1.2.11, Spades v3.12.0, MetaBat2 v2.12.1, CheckM v1.0.7 31, Bowtie2 v2-2.2.3 27, Diamond v0.8.24, ALS beamlines: 8.2.2, 8.2.1, Aimless v0.7.2, Phenix v1.13, Coot v0.8.9.1, PyMOL v1.8.3.2, XDS v2018-06-08, and CCP4 v7.0.073. R scripts to generate figure and for statistical analysis are available from the corresponding author upon request.

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

Data

Policy information about availability of data

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

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

Sequencing data supporting the findings of this study are available under accession number PRJNA505660. X-ray crystallographic data that support the findings of this study have been deposited in the RCSB Protein Data Bank (accession codes: 6MRX, 6MRV, 6MYV, and 6MNJ).
Field-specific reporting

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

☑ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

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

Sample size

Colonic fecal scraping from 3-5 mice of each diet type were used for 16S rRNA gene amplicon and three mice of each diet type were used for metagenomic shotgun sequencing analysis. Samples size was chosen based on the cost, mice, diet and vivarium availability, and the need for an N to offer sufficient statistical power.

Data exclusions

No data exclusions. For the experiments, the data are reported in full.

Replication

The number of mice studied, or number of biological observations studied for each experiment are provided in the methods. In brief, microbiome experiments were performed using 3-5 biologically independent animals, with 9-15 independent experiments. Enzyme activity analysis were performed using 3 independent experiments.

Randomization

Human-like Cmah-/- mice were bred onto a congenic C57BL/6 background and maintained in the University of California, San Diego vivarium according to Institutional Animal Care and Use Committee (IACUC) guidelines, with 12 hours diurnal lighting and access to food and water ad libitum. The cages were kept side by side in the same cage rack to minimize external influence on the gut microbiome. Age and sex matched female mice used in the study were maintained in sialic acid free soy based diet (Dyets, Inc.; 110951) from weaning until 8 to 10 weeks of age to prevent previous exposure to sialic acid.

Blinding

The experiments were not blinded once we conducted an supervised data analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

| n/a | Involved in the study |
|-----|------------------------|
| ☑   | Antibodies             |
| ☑   | Eukaryotic cell lines  |
| ☑   | Palaeontology          |
| ☑   | Animals and other organisms |
| ☑   | Human research participants |
| ☑   | Clinical data |

Methods

| n/a | Involved in the study |
|-----|------------------------|
| ☑   | ChIP-seq               |
| ☑   | Flow cytometry         |
| ☑   | MRI-based neuroimaging |

Animals and other organisms

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

Laboratory animals

Age and sex matched females. Wild type and cmah-/- mice from C57BL/6 lineage were used in this study. Samples were collected from mice with 10 weeks of age.

Wild animals

The study did not involve wild animals.

Field-collected samples

The study did not involve samples collected from the field.

Ethics oversight

All animal experiments were approved by the UC San Diego Institutional Animal Care and Use Committee (IACUC) under the protocol number S01227.

Note that full information on the approval of the study protocol must also be provided in the manuscript.