Article

Bile salt hydrolase acyltransferase activity expands bile acid diversity

https://doi.org/10.1038/s41586-024-07017-8

Received: 9 September 2022

Accepted: 2 January 2024

Published online: 07 February 2024

Check for updates

Douglas V. Guzior[1,2], Maxwell Okros[1], Madison Shivel[1,3], Bruin Armwald[1,3], Christopher Bridges[1,2], Yousi Fu[1], Christian Martin[1], Anthony L. Schilmiller[4], Wendy M. Miller[5,6], Kathryn M. Ziegler[5,6], Matthew D. Sims[5,6], Michael E. Maddens[5,6], Stewart F. Graham[5,6,7], Robert P. Hausinger[1,2] & Robert A. Quinn[1]✉

Bile acids (BAs) are steroid detergents in bile that contribute to the absorption of fats and fat-soluble vitamins while shaping the gut microbiome because of their antimicrobial properties[1–4]. Here we identify the enzyme responsible for a mechanism of BA metabolism by the gut microbiota involving amino acid conjugation to the acyl-site of BAs, thus producing a diverse suite of microbially conjugated bile acids (MCBAs). We show that this transformation is mediated by acyltransferase activity of bile salt hydrolase (bile salt hydrolase/transferase, BSH/T). *Clostridium perfringens* BSH/T rapidly performed acyl transfer when provided various amino acids and taurocholate, glycocholate or cholate, with an optimum at pH 5.3. Amino acid conjugation by *C. perfringens* BSH/T was diverse, including all proteinaceous amino acids except proline and aspartate. MCBA production was widespread among gut bacteria, with strain-specific amino acid use. Species with similar BSH/T amino acid sequences had similar conjugation profiles and several *bsh/t* alleles correlated with increased conjugation diversity. Tertiary structure mapping of BSH/T followed by mutagenesis experiments showed that active site structure affects amino acid selectivity. These MCBA products had antimicrobial properties, where greater amino acid hydrophobicity showed greater antimicrobial activity. Inhibitory concentrations of MCBAs reached those measured natively in the mammalian gut. MCBAs fed to mice entered enterohepatic circulation, in which liver and gallbladder concentrations varied depending on the conjugated amino acid. Quantifying MCBAs in human faecal samples showed that they reach concentrations equal to or greater than secondary and primary BAs and were reduced after bariatric surgery, thus supporting MCBAs as a significant component of the BA pool that can be altered by changes in gastrointestinal physiology. In conclusion, the inherent acyltransferase activity of BSH/T greatly diversifies BA chemistry, creating a set of previously underappreciated metabolites with the potential to affect the microbiome and human health.

In the liver, primary bile acids (BAs), cholic acid (CA) and chenodeoxycholic acid are synthesized from cholesterol and subsequently ligated at the C24 carboxyl group to either glycine or taurine to form glycocholic acid (GCA), glycochenodeoxycholic acid, taurocholic acid (TCA) and taurochenodeoxycholic acid, which collectively constitute our primary BA pool[5,6]. Conjugated primary BAs are stored in the gallbladder and then excreted into the gut on consumption of a meal, where they aid lipid digestion and function as key signalling molecules, detergents and antimicrobials[7,8]. After passing through the small intestine, BAs are reabsorbed at the terminal ileum and transported to the liver for reconjugation with glycine or taurine, if necessary, in a mechanism known as enterohepatic circulation (EHC)[9,10]. Approximately 5% escape recirculation and enter the large intestine where our gut microbiota can modify their structures extensively[1,3,11,12].

On entering the colon, BAs undergo several transformations by our gut microbiota, including deconjugation of glycine or taurine, dehydroxylation, oxidation and epimerization of the steroid core[1–3,13]. Recently, another microbial transformation of BAs was described: amino acid conjugation[11]. This process initially included the description of three new acyl-conjugated BAs: phenylalanocholic acid (PheCA), leucocholic acid (LeuCA) and tyrosocholic acid (TyrCA). The first bacterium identified to produce these conjugates was *Enterocloster bolteae*, a member of the Lachnospiraceae[11]. Recent work has shown that various amino acids can be conjugated to many base BAs by a wide variety of

[1]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI, USA. [2]Department of Microbiology, Genetics & Immunology, Michigan State University, East Lansing, MI, USA. [3]College of Osteopathic Medicine, Michigan State University, East Lansing, MI, USA. [4]Mass Spectrometry and Metabolomics Core, Michigan State University, East Lansing, MI, USA. [5]Corewell Health, William Beaumont University Hospital, Royal Oak, MI, USA. [6]Oakland University, William Beaumont School of Medicine, Rochester, MI, USA. [7]Beaumont Research Institute, Royal Oak, MI, USA. ✉e-mail: quinnrob@msu.edu

gut bacteria, creating the microbially conjugated bile acids (MCBAs) that greatly diversify the mammalian BA pool[1,12]. However, the enzyme and biochemical mechanism responsible for MCBA production has not yet been determined.

Here we show that microbial BA conjugation with amino acids is mediated by previously unknown acyltransferase activity of the N-terminal nucleophile (Ntn) enzyme bile salt hydrolase, henceforth referred to as bile salt hydrolase/transferase (BSH/T). This activity is present in BSH/T from numerous bacterial taxa, resulting in unique MCBA profiles depending on BSH/T amino acid sequence. This acyl conjugation greatly diversifies the BA pool creating some MCBAs with antimicrobial properties that can enter EHC.

BSH/T acyl transfer characterization

The first BSH/T to be purified and have its hydrolase activity characterized was from *Clostridium perfringens* (CpBSH/T)[14]. Owing to its established interaction with conjugated BAs, also known as bile salts, we investigated the capacity of CpBSH/T to exchange the conjugated amino acid. Enzyme-catalysed hydrolysis of bile salts occurs through a covalently bound cysteinyl intermediate[15] (Fig. 1a, steps 1 and 2) and CpBSH/T was active for hydrolysis over a broad pH range (pH 3–7; Extended Data Fig. 1b). When incubated with TCA and an equimolar mix of 20 essential amino acids, CpBSH/T rapidly hydrolysed TCA to CA (Extended Data Fig. 1b), as expected, in addition to catalysing acyl conjugation of CA with a variety of amino acids (Fig. 1b,c). Indeed, 16 of 20 amino acids became linked to CA, except aspartocholic acid (AspCA), methionocholic acid (MetCA), prolocholic acid (ProCA) and valocholic acid (ValCA) were not produced under these conditions (Extended Data Fig. 1d). CpBSH/T may catalyse acyl transfer through the reaction of amino acids with a covalently bound intermediate, where an amino acid acts as a nucleophile in lieu of water (Fig. 1a, steps 1 and 3). We observed acyltransferase activity across a broad pH range (Fig. 1b), with an optimum at pH 5.3 (Extended Data Fig. 1a,c) based on the summed abundance of MCBAs following 120 min of incubation at 37 °C. This value is slightly higher than the previously reported pH 4.5–4.9 optimum for TCA hydrolysis[16,17]. At peak activity, acyl transfer activity reaches 7.0% of hydrolysis activity. That is, one amino acid was incorporated for every 15 TCA molecules hydrolysed to CA, showing that acyl transfer by BSH/T is substantial.

We then sought to determine if a similar panel of amino acids would be conjugated when provided GCA instead of TCA. CpBSH/T incubated with GCA at pH 5.0 transferred 11 of 19 supplied amino acids, excluding glycine because of its availability after hydrolysis (Fig. 1d and Extended Data Fig. 1d). Surprisingly, CpBSH/T produced MetCA and ValCA, otherwise absent when provided TCA. The reduced number of amino acids transferred (11 of 19 compared to 16 of 20 when provided TCA) may be a consequence of competition by high glycine concentrations following hydrolysis (Extended Data Fig. 1d).

We also demonstrated that CpBSH/T can ligate amino acids directly to CA (Fig. 1e), probably occurring through a covalent intermediate (Fig. 1a, steps 4 and 3). CpBSH/T successfully ligated 12 of 20 amino acids (Extended Data Fig. 1d), including valine and methionine which were not observed with TCA transfer. However, consistent with TCA transfer, ProCA and AspCA were not observed. The absence of proline conjugation may be a result of its unique secondary amine preventing proper nucleophilic attack; previous reports have not observed proline conjugation[12].

To investigate the acyl transfer kinetics of CpBSH/T, we first determined a saturation concentration for TCA hydrolysis, as performed previously[14] and found 8 mM TCA more than sufficient to saturate the hydrolysis reaction (Fig. 1f). Investigating constants for phenylalanine transfer resulted in linear kinetics for PheCA production with increasing phenylalanine concentration (up to 5 mM; Fig. 1g). These linear kinetics observed for PheCA formation are consistent with the formation of the enzyme-CA adduct, followed by a rate-determining nucleophilic attack to achieve hydrolysis or amino acid acyl transfer[18]. Kinetics of phenylalanine ligation to CA were also linear and showed rates that were nearly 10% of those for acyl transfer to TCA (Fig. 1h), again supporting the rate-determining reaction after formation of an enzyme-CA intermediate.

BSH/T sequence shapes conjugation profiles

After confirming MCBA production by BSH/T in vitro, we analysed genomes of 29 strains subsequently screened for MCBA production to determine if phylogenetic relatedness correlated with conjugation capability (Fig. 2a and Supplementary Table 1). These included Actinomycetia, Verrucomicrobiae, Gammaproteobacteria, Bacilli and Clostridia, with a focus on the Lachnospiraceae family in Clostridia. Of these strains, 19 produced at least one MCBA (Fig. 2b) and production was particularly prevalent among the Lachnospiraceae, with only one species unable to produce MCBAs. Verrucomicrobiae (*Akkermansia muciniphila*) was the only class without observable production. The most robust MCBA producers, *Lactiplantibacillus plantarum*, *Ruminococcus gnavus*, *Enterococcus faecalis* and *Bifidobacterium bifidum* subsp. *infantis*, were phylogenetically disparate, indicating little association between evolutionary relatedness and MCBA production (Fig. 2a). Hierarchical clustering was used to group the bacterial conjugation profiles into five clusters on the basis of Bray–Curtis dissimilarity (Fig. 2b and Extended Data Fig. 1e). Cluster 1 strains showed robust conjugation of a wide variety of amino acids. Strains in cluster 2 favoured glycine and alanine conjugation, whereas cluster 3 preferentially conjugated small, hydrophilic amino acids. Cluster 4 showed extensive lysine conjugation and cluster 5 profiles were dominated by aspartate conjugation (Fig. 2b). The most robust MCBA producer, *L. plantarum*, lies in cluster 1 and produced 16 of 18 observed MCBAs (Fig. 2b). Although clustering showed little phylogenetic correlation, clusters 3 and 5 were primarily associated with members of the *Lachnospiraceae* (Fig. 2a).

We then mined the genomes of all 29 species for *bsh/t* presence to investigate relationships between translated protein sequences and MCBA profiles (Fig. 2c and Supplementary Table 1). Two species, *Clostridium sporogenes* and *Lacrimispora aerotolerans*, possess annotated *bsh/t* but did not produce MCBAs. Further analysis showed valine in place of a traditional start codon in *C. sporogenes* BSH/T, which may explain why MCBA production was absent in this bacterium. By contrast, *Clostridium scindens* ATCC 35704 produced MCBAs although lacking an annotated *bsh/t*. Analysis of 35 publicly available *C. scindens* genomes for *bsh/t* presence showed only strain Q4 contained a predicted *bsh/t* (Extended Data Fig. 2a and Supplementary Table 2). *C. scindens* Q4 BSH/T had high amino acid sequence similarity to BSH/T from *R. gnavus* and other sequences with BSH/T cluster 1 (Extended Data Fig. 2b). The absence of *bsh/t* in MCBA-producing *C. scindens* ATCC 35704 suggests that other enzymes capable of BA conjugation remain to be discovered.

The remaining 18 MCBA producers had at least one annotated or predicted *bsh/t* in their genome with some, such as *E. bolteae*, containing at least three. BSH/T phylogenetic tree topology (Fig. 2c) showed limited correlation to the five MCBA profile clusters (Fig. 2b). However, there were three main BSH/T lineages: group I, containing a set of diverse and robust MCBA producers; group II, primarily associating with MCBA clusters 3 and 5; and group III, showing significant sequence divergence from the other groups and little association with MCBA profiles (Fig. 2b,c). The last group may represent sequences with a high degree of similarity to other enzymes in the Ntn-hydrolase superfamily, indicating that these BSH/T homologues may have other functions. *E. bolteae* and *Enterocloster clostridioformis* contain BSH/T sequences from all three groups, yet *E. bolteae* produced a diverse MCBA profile whereas GCA dominated MCBAs produced by *E. clostridioformis*.

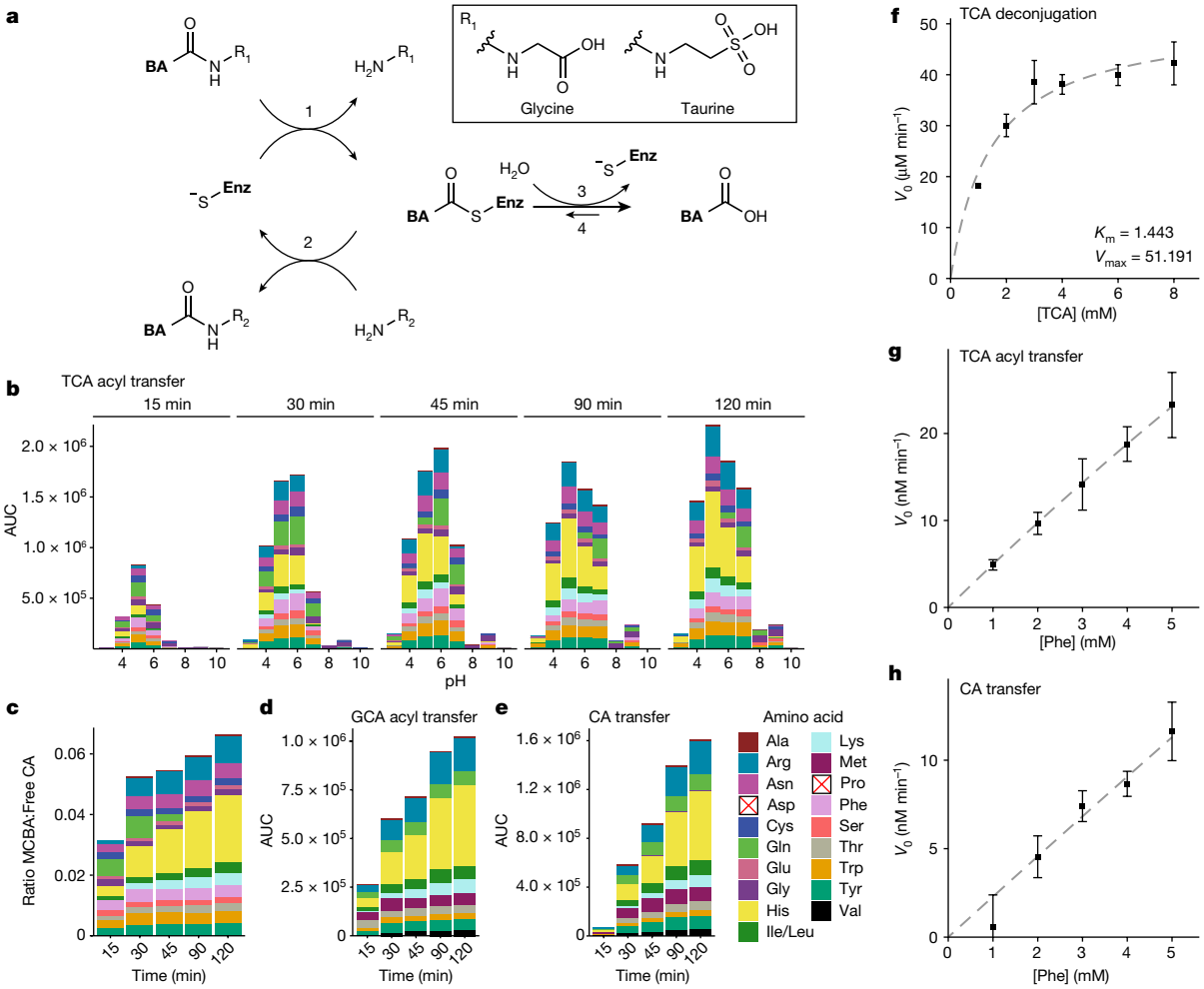


Fig. 1 | *C. perfringens* BSH/T produces a broad range of MCBAs at acidic pH. **a**, Chemical reaction steps catalysed by BSH/T. The enzyme is capable of (1) reacting with conjugated primary BAs through nucleophile attack using Cys2 to form a covalently bound enzyme-CA intermediate followed by (2) hydrolytic release of the BA or (3) reaction with other amino acids by an acyl transfer reaction, resulting in formation of MCBAs. In addition, MCBAs can be generated by (4) direct formation of the enzyme-CA intermediate from CA with subsequent acyl transfer. Enz, enzyme. **b**, Stacked area-under-the-curve (AUC) profiles of MCBA products following CpBSH/T incubation with TCA and an equimolar amino acid mixture over a broad pH range (3.0–10.0), across time. **c**, Ratio of mean summed MCBA abundance to CA abundance, derived from acyltransferase activity and hydrolase activity of CpBSH/T incubated with TCA and an equimolar amino acid mixture at pH 5.0. **d**,**e**, MCBA profiles at pH 5.0 following CpBSH/T incubation with an equimolar amino acid mixture and 2.5 mM GCA (**d**) or 2.5 mM CA (**e**). Prolocholic acid (ProCA) and aspartocholic acid (AspCA) were not present in any samples. $n = 3$ independent reactions per pH, per BA. **f**, Deconjugation kinetics for commercial CpBSH/T when incubated with TCA. K_m, Michaelis constant; V_{max}, maximum rate of reaction. **g**,**h**, Reaction kinetics for the formation of PheCA when incubating 8 mM TCA (**g**) or 8 mM CA (**h**) with CpBSH/T and 1–5 mM phenylalanine. Data are presented as mean ± s.e.m.; $n = 3$ independent reactions per concentration.

Analysis of BSH/T amino acid sequence alignment showed an amino acid substitution that was potentially responsible for divergence in the conjugation profiles observed. Asn82 (Fig. 2d,e; *C. perfringens* BSH/T as reference[17,19]) is reported as being highly conserved in BSH/T sequences in previous studies[20]. However, we show that this position is instead a Tyr in BSH/T from several Lachnospiraceae species, most residing in BSH/T group II. This residue lies in the active site of the BSH/T crystal structure from *C. perfringens* (PDB ID 2BJG), adjacent to the carboxylate of cocrystalized deoxycholic acid (DCA; Extended Data Fig. 1f)[19] and directly at the location of the amide bond of TCA in *Lactobacillus salivarius* BSH/T cocrystallized with TCA (PDB ID 8BLT; Fig. 2e)[21]. We therefore propose that BSH/T sequence variation at the active site determines its capacity for BA conjugation and substituted both Asn82 and Cys2 of CpBSH/T to test this hypothesis.

Alteration of Asn82 to Tyr82 (N82Y) in CpBSH/T shaped the amino acid conjugation pool in a similar fashion observed for organisms encoding either of these variants. *Escherichia coli* expressing N82Y variants demonstrated significant deficits in BA conjugation, with decreased abundance of glutamatocholic acid (GluCA), lysocholic acid (LysCA) and LeuCA when grown in media containing 1 mM GCA or 1 mM TCA (Fig. 2f). These trends were also seen when grown in medium containing 1 mM CA. However, AlaCA was significantly enriched in the N82Y variant compared to wild type (WT) when provided TCA. GCA was also enriched in the N82Y variant, although this was not statistically significant. By contrast, Cys2 substitution (C2A) resulted in complete ablation of BA conjugation, regardless of substrate. C2A variants were also unable to hydrolyse TCA, whereas WT and N82Y variants completely hydrolysed TCA (Extended Data Fig. 3a) and most GCA (Extended Data Fig. 3b) to CA.

Antimicrobial efficacy of MCBAs

Free BAs are known to exert antimicrobial activity by damaging cell membranes and chromosomal DNA[4], a mechanism not limited to

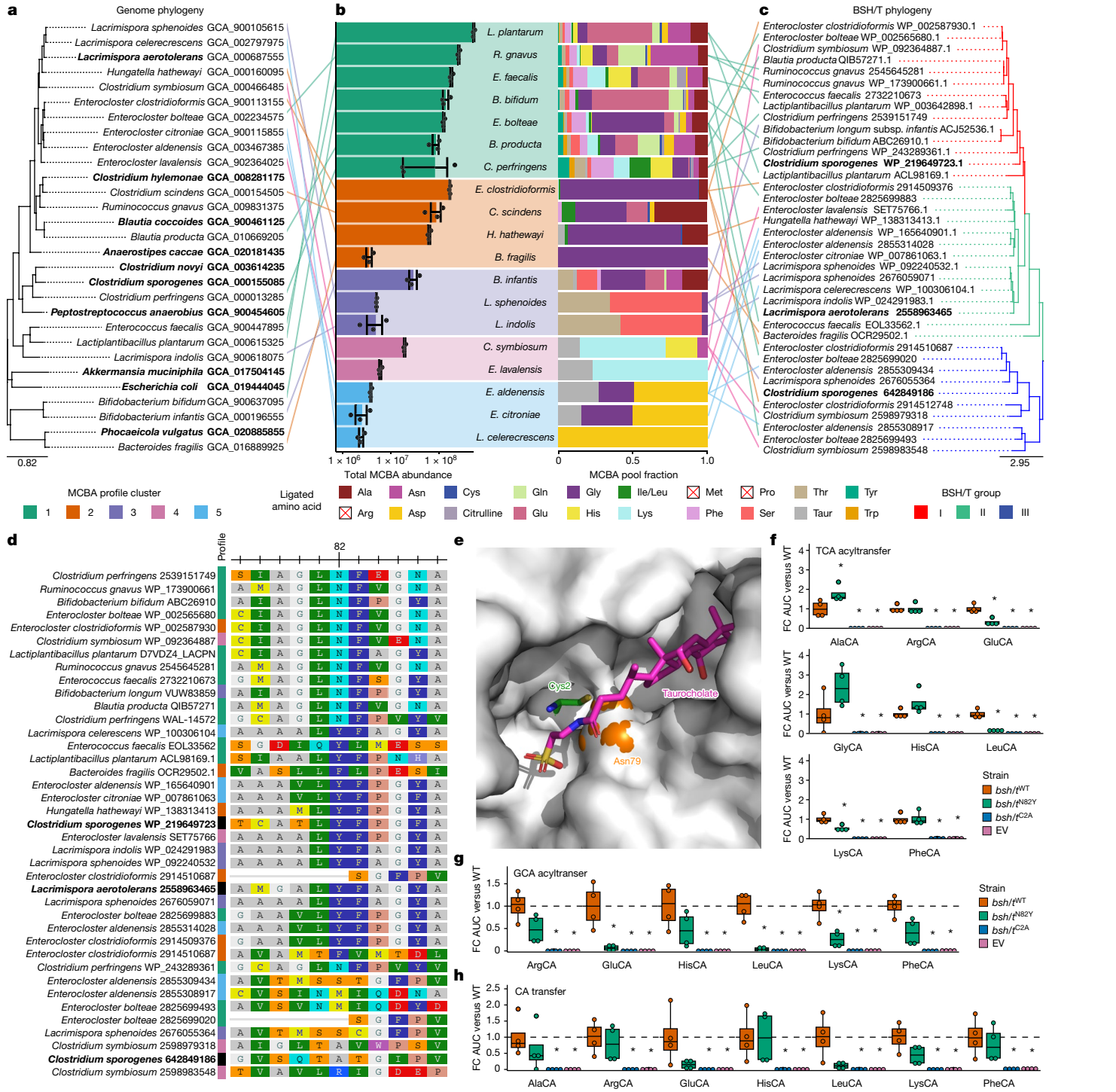


Fig. 2 | MCBA product identities correlate with BSH/T amino acid sequences. **a**, Genome phylogenetic tree for all strains screened, with strains in bold denoting those that did not produce MCBAs. **b**, Summed MCBA abundances and profiles from gut bacteria grown in the presence of 1 mM CA coloured by MCBA profile cluster. Data presented as mean ± s.e.m.; $n = 3$ independent cultures. **c**, Phylogenetic relatedness of BSH/T amino acid sequences showing three clusters of related sequences. Lines connect genome and BSH/T sequences to the product profiles for each strain. Line colour corresponds to MCBA profile cluster. **d**, BSH/T amino acid sequence alignment highlighting conserved Asn82 or Tyr82 for Clostridia-like and Lachnospiraceae-like BSH/T sequences, respectively, with MCBA profile cluster identified next to the strain and BSH/T accession number. **e**, Structure of *L. salivarius* BSH/T (PDB ID 8BLT)[21,40] in complex with TCA (molecular surface representation) showing the proximity of Asn79 (Asn82 in CpBSH/T) and catalytic Cys2 with the amide bond of TCA. **f**–**h**, Fold-change (FC) in abundance of MCBAs produced by *C. perfringens* BSH/T with substitutions in Asn82 (bsh/t^{N82Y}) or Cys2 (bsh/t^{C2A}) compared to WT when expressed by *E. coli* DH5α incubated with 1 mM TCA (**f**), 1 mM GCA (**g**) or 1 mM CA (**h**), using endogenous amino acids for BA conjugation. EV denotes pBAD18-Cm without insert; $n = 4$ independent cultures. Data in **f**–**h** are presented as boxplots where the middle lines are the median, lower and upper hinges represent the first and third quartiles, upper whiskers extend to maxima and lower whiskers extend to minima. Statistical significance in **f**–**h** was determined by Wilcoxon rank-sum test against WT with Benjamini–Hochberg *P* value correction, $^*P < 0.05$.

bacterial cells. This antimicrobial activity is a well-known property of secondary BAs[22–25], whereas conjugated primary BAs are less antimicrobial[4,26]. We therefore suggested that microbial BA conjugation may be a means for bacteria to modulate BA toxicity. To test this, we first determined the effects of medium supplementation with 1 mM CA or individual MCBAs on *E. bolteae*, the first organism identified to produce MCBAs. *E. bolteae* showed increased growth in the presence of any MCBA but growth with CA showed a slight detriment (Fig. 3a). We chose to use 1 mM CA as it is known to be inhibitory against most BA-susceptible bacteria and represents the higher range of native BA concentrations in the human gastrointestinal tract[27,28]. Impacts of MCBA administration on further species showed variable antimicrobial efficacy. The most marked reductions in growth were observed for *Clostridium hylemonae*, *Blautia coccoides*, *Peptostreptococcus anaerobius*, *L. aerotolerans* and *Lacrimispora indolis*, of which only *L. indolis* produced MCBAs (Fig. 3a). *P. anaerobius* and *L. aerotolerans* showed the most marked deficit as growth was significantly reduced if not completely inhibited (Fig. 3b). Antimicrobial efficacy depended on the amino acid conjugated where hydrophobic conjugates showed the strongest effects, particularly PheCA and LeuCA (Fig. 3b). Importantly, these effects were not observed for host conjugates GCA or TCA, indicating that microbial conjugation with these non-canonical amino acids can increase BA toxicity. *L. aerotolerans* showed growth defects when grown in LeuCA, with an effective dose nearly half that of CA (median effective dose ED_{50} = 236 μM versus 425 μM; Extended Data Fig. 4a,b) but showed slightly increased resistance to PheCA (ED_{50} = 460 μM; Extended Data Fig. 4c) compared with CA. However, PheCA effectively inhibited *P. anaerobius* growth at nearly two-thirds the concentration of CA (ED_{50} = 287 μM versus 388 μM; Extended Data Fig. 4d,e) whereas LeuCA and TyrCA were half as effective (ED_{50} = 557 μM and 521 μM, respectively; Extended Data Fig. 4f,g).

Given the in vitro effects of MCBAs in bacterial monoculture, we sought to determine if these effects translate in vivo. Administration of 100 mg kg^{-1} of PheCA, SerCA, TCA or mock control to C57BL/6J mice through oral gavage for 13 days resulted in significant caecal and faecal microbiome community shifts, in which several ASVs showed the same treatment-specific trends in communities in both sample types (Supplementary Information, Extended Data Fig. 5 and Supplementary Tables 3 and 4). This indicates that the antimicrobial activity of MCBAs may translate to community shifts in vivo at this high concentration. To further explore these microbiome effects, we developed a more natural manner of administering MCBAs using a modified protocol from ref. 29 which involves feeding mice peanut butter pellets with known concentrations of BAs. Mice fed MCBAs with the peanut butter feeding method (PBFM) at 10 mg kg^{-1} resulted in gastrointestinal concentrations similar to those measured in the human gut (Supplementary Table 5; refs. 30,31). However, at this physiologically relevant concentration and feeding method, microbiome changes were more subtle. Significant differences in faecal microbiome structure were observed overall (Extended Data Fig. 6, timepoint-nested permutational multivariate analysis of variance (PERMANOVA); F = 1.4901, P = 0.003) but differences in individual timepoints were not. These experiments demonstrate that MCBAs have altered microbial activities compared to native BAs but these effects may only occur at the highest concentrations encountered in vivo.

MCBAs enter enterohepatic circulation

An important question about MCBAs is whether or not they can enter EHC, a tightly regulated process for recycling BAs starting at the terminal ileum that returns BAs to the liver through the hepatic portal vein (HPV)[9]. To investigate the propensity for MCBAs to enter EHC, C57BL/6J mice were fed 100 mg kg^{-1} of SerCA through the PBFM[29]. SerCA was detected in all gastrointestinal tissues analysed, blood from the HPV, and appeared in faecal pellets after 24 h (Extended Data Fig. 7).

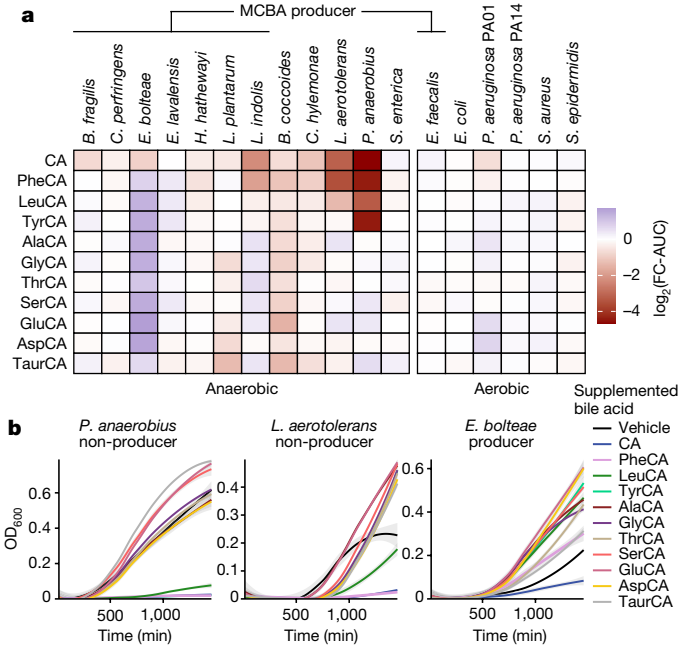


Fig. 3 | MCBAs show varied antimicrobial properties. a, Average log_2-fold-change in area under the growth curve (FC-AUC) between BA-treated cultures and control. **b**, Representative growth curves for *P. anaerobius* and *L. aerotolerans*, species showing growth detriments in the presence of 1 mM CA and CA conjugated with hydrophobic amino acids, in addition to *E. bolteae*, demonstrating slight increases in FC-AUC for all MCBAs administered. Growth curve data presented as smoothed mean optical density at 600 nm (OD_{600}) with 95% confidence interval shaded behind the line. n = 3 independent cultures for anaerobic growth and 4 independent cultures for aerobic growth.

We then sought to determine the ability of MCBAs with various conjugated amino acids to enter EHC. Mice were fed a mixture of eight MCBAs for 5 days (10 mg kg^{-1} of AlaCA, AspCA, GluCA, LeuCA, PheCA, SerCA, ThrCA and TyrCA) through PBFM. These MCBAs were observed throughout the gastrointestinal tract and in the liver, kidney, serum and gallbladder with particularly high abundances of PheCA and SerCA across all samples (Extended Data Fig. 8 and Supplementary Table 6). However, these were also detected at low concentrations in mock-fed controls, probably due to basal concentrations of production in vivo. Therefore, we further verified the ability of both SerCA and PheCA to enter EHC in a follow-up experiment including equimolar amino acid + BA controls (matching 10 mg kg^{-1} of individual BA). Interestingly, SerCA concentrations were eight-fold higher in the gallbladder of SerCA-treated animals than PheCA in PheCA-treated animals (359 ± 80.4 μM versus 41.6 ± 12.4 μM) and SerCA was detected consistently in the liver (0.95 ± 0.20 μM) when PheCA was not (Fig. 4a,b and Supplementary Table 6). Gallbladder and liver samples from amino acid + BA controls contained low concentrations of these compounds, supporting de novo conjugation and subsequent circulation, but these concentrations were significantly less than MCBA-fed animals. It is possible that the limited EHC observed with PheCA was due to specific hydrolysis by pancreatic carboxypeptidases, as reported for TyrCA and other conjugated BAs[32,33]. However, neither carboxypeptidase showed activity when incubated with PheCA or SerCA while each catalysed near-complete hydrolysis of positive controls (Extended Data Fig. 9). It is also possible that preferential microbial hydrolysis of PheCA in the gut, which has been recently described, may have occurred[34]. These experiments support that MCBAs can enter EHC intact when fed to mice at physiologically relevant concentrations, potentially affecting liver and BA metabolism and that the degree of EHC depends on the amino acid conjugated.

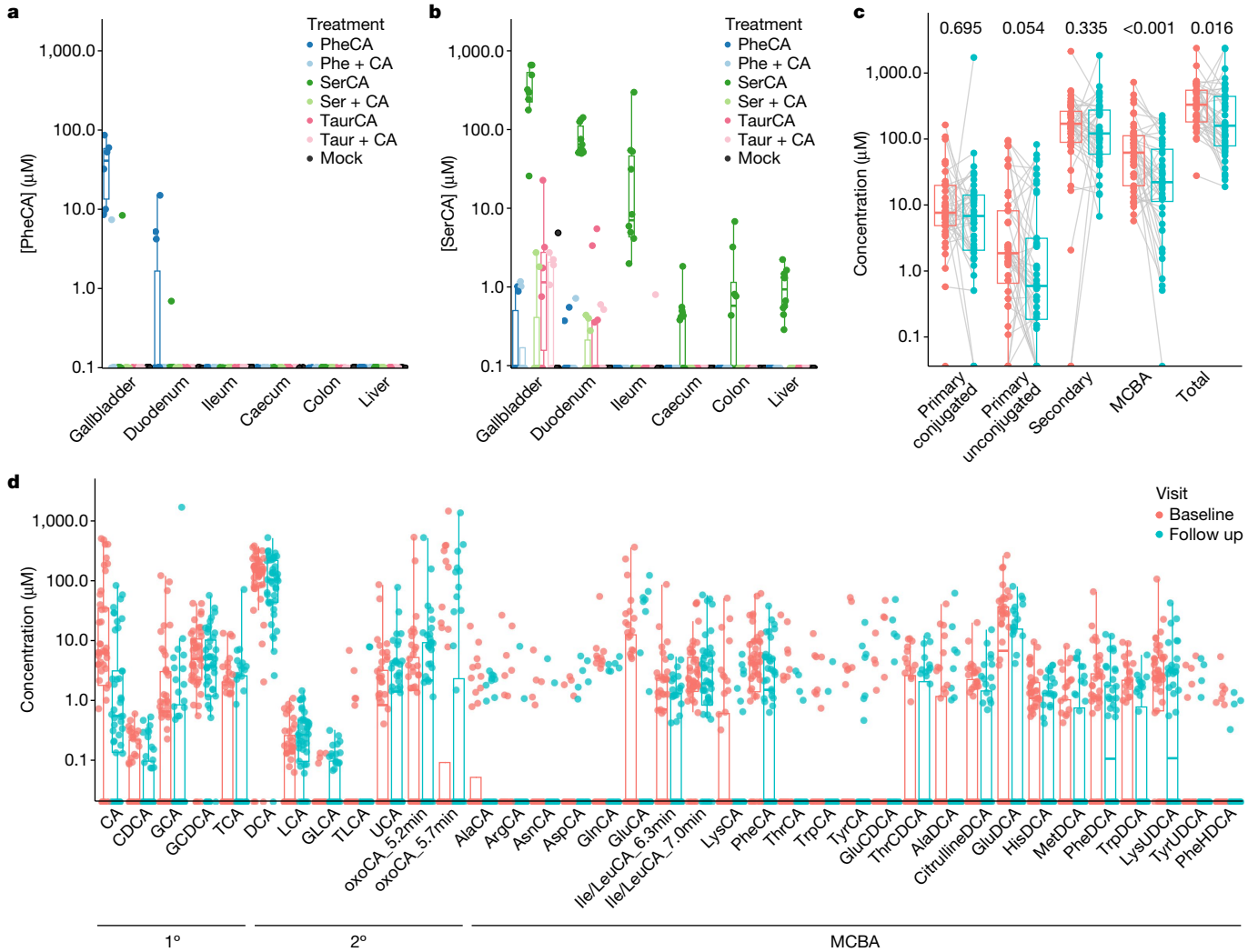


Fig. 4 | BA concentrations in mouse tissue samples following MCBA feeding and in human faeces of patients undergoing sleeve gastrectomy. **a**,**b**, Concentrations of PheCA (**a**) and SerCA (**b**) in mouse tissue samples following 10-day dosing with the indicated treatments through PBFM. PheCA is highly abundant in the gallbladder and present in the duodenum of PheCA-fed mice, while SerCA is enriched in all tissues sampled, including the colon and liver. $n = 5$ male and 5 female mice per treatment. **c**,**d**, BA class shifts in a patient population undergoing sleeve gastrectomy before (baseline) and 3 months after (follow up) surgery, with significant decreases in total MCBA concentration and total BA (**c**) and changes in individual BA concentrations in that cohort (**d**). $n = 44$ patients, with paired samples at each timepoint. Data in **a**–**d** are presented as boxplots where the middle lines are the median, lower and upper hinges represent the first and third quartiles, upper whiskers extend to maxima and lower whiskers extend to minima. Significance between timepoints determined by Wilcoxon signed-rank tests with Benjamini–Hochberg P value correction.

Human bariatric surgery affects faecal MCBAs

To investigate whether MCBA concentrations change in the context of human gastrointestinal health, we analysed faecal samples from patients who underwent sleeve gastrectomy as a treatment modality for obesity. The concentrations of MCBAs and other BAs were quantified from samples collected before surgery and 3 months postoperation. A diverse complement of MCBAs were detected, including at least 25 unique compounds (Fig. 4d and Supplementary Table 7) with an average total MCBA concentration of 77.7 ± 11.9 μM. In comparison, primary conjugated BAs were measured at 34.2 ± 19.6 μM, free primary BAs at 10.8 ± 2.4 μM and secondary BAs at 223 ± 32.7 μM. This is evidence that faecal MCBAs can reach concentrations at or above primary BAs in faeces and approximately one-third the concentration of secondary BAs (Supplementary Table 8). Furthermore, collective analysis of BA chemistry showed significant reductions in faecal concentrations of MCBAs ($P = 8.9 \times 10^{-4}$) and total BAs ($P = 0.016$) after sleeve gastrectomy but conjugated primary BAs, free BAs and secondary BAs were not (Fig. 4c). This supports that MCBAs are a substantial component of the human BA pool and are affected by surgical treatment for obesity.

Conclusions

We identified BSH/T as an enzyme responsible for MCBA production. This acyltransferase activity of BSH/T changes our understanding of one of the best-studied and ubiquitous enzymes in the human gut microbiome[35]. Transferase activity is a substantial feature of the enzyme, favoured at pH close to that of lower human gastrointestinal tract. The pH of the small intestine remains fairly neutral, around pH 7.0, whereas caecal pH can drop to 5.42–6.50 but then increase to around 7.0 in the colon[36,37]. A recent study showed that MCBAs have higher concentrations in the upper gastrointestinal tract than stool which may be explained by pH variability along the gut altering the conjugation or hydrolysis activity of BSH/T[30].

We also show that sequence variation in BSH/T, long known to exist among gut bacteria, is associated with unique amino acid conjugation

profiles[38,39]. Sequence comparisons suggested that differences at residue 82 (Asn versus Tyr) reflected differences in both the total MCBA production and the diversity of amino acids used in conjugation. This was then confirmed by mutation of *C. perfringens bsh/t*. MCBA abundances were significantly reduced in *E. coli* expressing *bsh/t*^{N82Y} compared to WT. Interestingly, there was an enrichment of MCBAs containing small amino acids (alanine and glycine) in cultures expressing the N82Y variant. There may be selective pressure on certain BSH/T enzymes to conjugate particular amino acids akin to preferential hydrolysis of TCA or GCA[35]. We show that this selectivity may be driven by modulating inherent BA toxicity, depending on the amino acid conjugated but further work to determine their affects in vivo are required. Hydrophobic conjugates (PheCA and LeuCA) were more antimicrobial than hydrophilic conjugates but antimicrobial effects were only observed at the highest concentrations of MCBAs detected in humans. Although not tested in this study, other base BAs with stronger antimicrobial properties, such as DCA or LCA, may become more potent when conjugated. Further study of conjugates with the strongest antimicrobial effects will help elucidate their potential as agents of microbial warfare in the human gut or, in contrast, as a means of detoxification.

The fate of MCBAs in vivo remains a significant question. Recent work has shown that BSH/T can hydrolyse these compounds in vitro[34], as can pancreatic carboxypeptidases A and B[33]. Our experiments in mice using PheCA and SerCA showed they can enter EHC and were not deconjugated by pancreatic carboxypeptidases in vitro. Liver and gallbladder concentrations of SerCA were higher than PheCA, indicating some selection for entry into these organs. Whether this was due to specific selectivity by BA transporters in the ileum, favoured microbial hydrolysis or other forms of metabolism remains unknown. Follow-up studies will be needed to better understand the fate of MCBAs and their effects on the host and its microbiome. As physiological and microbiological effects of MCBAs become clearer, so too does the potential for their use as treatments for gastrointestinal-related diseases.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-024-07017-8.

1. Guzior, D. V. & Quinn, R. A. Review: microbial transformations of human bile acids. *Microbiome* **9**, 140 (2021).
2. Ridlon, J. M., Harris, S. C., Bhowmik, S., Kang, D. J. & Hylemon, P. B. Consequences of bile salt biotransformations by intestinal bacteria. *Gut Microbes* **7**, 22–39 (2016).
3. Ridlon, J. M., Kang, D. J. & Hylemon, P. B. Bile salt biotransformations by human intestinal bacteria. *J. Lipid Res.* **47**, 241–259 (2006).
4. Sannasiddappa, T. H., Lund, P. A. & Clarke, S. R. *In vitro* antibacterial activity of unconjugated and conjugated bile salts on *Staphylococcus aureus*. *Front. Microbiol.* **8**, 1581 (2017).
5. Hofmann, A. F. The continuing importance of bile acids in liver and intestinal disease. *Arch. Intern. Med.* **159**, 2647–2658 (1999).
6. Russell, D. W. The enzymes, regulation and genetics of bile acid synthesis. *Annu. Rev. Biochem.* **72**, 137–174 (2003).
7. de Aguiar Vallim, T. Q., Tarling, E. J. & Edwards, P. A. Pleiotropic roles of bile acids in metabolism. *Cell Metab.* **17**, 657–669 (2013).
8. Shin, D. J. & Wang, L. in *Bile Acids and Their Receptors* Vol. 256 (eds Fiorucci, S. & Distrutti, E.) 51–72 (Springer, 2019).
9. Hofmann, A. F. The enterohepatic circulation of bile acids in mammals: form and functions. *Front. Biosci.* **14**, 2584–2598 (2009).
10. Dawson, P. A. & Karpen, S. J. Intestinal transport and metabolism of bile acids. *J. Lipid Res.* **56**, 1085–1099 (2015).
11. Quinn, R. A. et al. Global chemical effects of the microbiome include new bile–acid conjugations. *Nature* **579**, 123–129 (2020).
12. Lucas, L. N. et al. Dominant bacterial phyla from the human gut show widespread ability to transform and conjugate bile acids. *mSystems* **6**, e00805–e00821 (2021).
13. Lee, J. W. et al. Formation of secondary allo-bile acids by novel enzymes from gut Firmicutes. *Gut Microbes* **14**, 2132903 (2022).
14. Gopal-Srivastava, R. & Hylemon, P. B. Purification and characterization of bile salt hydrolase from *Clostridium perfringens*. *J. Lipid Res.* **29**, 1079–1085 (1988).
15. Lodola, A. et al. A catalytic mechanism for cysteine N-terminal nucleophile hydrolases, as revealed by free energy simulations. *PLoS ONE* **7**, e32397 (2012).
16. Coleman, J. P. & Hudson, L. L. Cloning and characterization of a conjugated bile acid hydrolase gene from *Clostridium perfringens*. *Appl. Environ. Microbiol.* **61**, 2514–2520 (1995).
17. Rossocha, M., Schultz-Heienbrok, R., Von Moeller, H., Coleman, J. P. & Saenger, W. Conjugated bile acid hydrolase is a tetrameric N-terminal thiol hydrolase with specific recognition of its cholyl but not of its tauryl product. *Biochemistry* **44**, 5739–5748 (2005).
18. Hinberg, I. & Laidler, K. J. The kinetics of reactions catalyzed by alkaline phosphatase: the effects of added nucleophiles. *Can. J. Biochem.* **50**, 1360–1368 (1972).
19. Rossocha, M., Schultz-Heienbrok, R., Von Moeller, H., Coleman, J. P. & Saenger, W. Crystal structure of conjugated bile acid hydrolase from *Clostridium perfringens* in complex with reaction products taurine and deoxycholate. *Biochemistry* **44**, 5739–5748 (2005).
20. Foley, M. H., O'Flaherty, S., Barrangou, R. & Theriot, C. M. Bile salt hydrolases: gatekeepers of bile acid metabolism and host–microbiome crosstalk in the gastrointestinal tract. *PLoS Pathog.* **15**, e1007581 (2019).
21. Karlov, D. S. et al. Characterization of the mechanism of bile salt hydrolase substrate specificity by experimental and computational analyses. *Structure* **31**, 629–638 (2023).
22. Bernstein, H., Bernstein, C., Payne, C. M., Dvorakova, K. & Garewal, H. Bile acids as carcinogens in human gastrointestinal cancers. *Mutat. Res.* **589**, 47–65 (2005).
23. Bernstein, C. et al. Carcinogenicity of deoxycholate, a secondary bile acid. *Arch. Toxicol.* **85**, 863–871 (2011).
24. Jia, W., Xie, G. & Jia, W. Bile acid–microbiota crosstalk in gastrointestinal inflammation and carcinogenesis. *Nat. Rev. Gastroenterol. Hepatol.* **15**, 111–128 (2018).
25. Cao, H. et al. Secondary bile acid-induced dysbiosis promotes intestinal carcinogenesis. *Int. J. Cancer* **140**, 2545–2556 (2017).
26. Kurdi, P., Kawanishi, K., Mizutani, K. & Yokota, A. Mechanism of growth inhibition by free bile acids in Lactobacilli and Bifidobacteria. *J. Bacteriol.* **188**, 1979–1986 (2006).
27. Hamilton, J. P. et al. Human cecal bile acids: concentration and spectrum. *Am. J. Physiol. Gastrointest. Liver Physiol.* **293**, G256–G263 (2007).
28. Northfield, T. C. & McColl, I. Postprandial concentrations of free and conjugated bile acids down the length of the normal human small intestine. *Gut* **14**, 513–518 (1973).
29. Zapata, R. C., Zhang, D., Chaudry, B. & Osborn, O. Self-Administration of drugs in mouse models of feeding and obesity. *J. Vis. Exp.* https://doi.org/10.3791/62775 (2021).
30. Shalon, D. et al. Profiling the human intestinal environment under physiological conditions. *Nature* **617**, 581–591 (2023).
31. Gentry, E. C. et al. Reverse metabolomics for the discovery of chemical structures from humans. *Nature* https://doi.org/10.1038/s41586-023-06906-8 (2023).
32. Huijghebaert, S. M. & Hofmann, A. F. Influence of the amino acid moiety on deconjugation of bile acid amidates by cholylglycine hydrolase or human faecal cultures. *J. Lipid Res.* **27**, 742–752 (1988).
33. Huijghebaert, S. M. & Hofmann, A. F. Pancreatic carboxypeptidase hydrolysis of bile acid–amino acid conjugates: selective resistance of glycine and taurine amidates. *Gastroenterology* **90**, 306–315 (1986).
34. Foley, M. H. et al. Bile salt hydrolases shape the bile acid landscape and restrict *Clostridioides difficile* growth in the murine gut. *Nat. Microbiol.* **8**, 611–628 (2023).
35. Dong, Z. & Lee, B. H. Bile salt hydrolases: structure and function, substrate preference and inhibitor development. *Protein Sci.* **27**, 1742–1754 (2018).
36. Maurer, J. M. et al. Gastrointestinal pH and transit time profiling in healthy volunteers using the IntelliCap system confirms ileo-colonic release of ColoPulse Tablets. *PLoS ONE* **10**, e0129076 (2015).
37. Evans, D. F. et al. Measurement of gastrointestinal pH profiles in normal ambulant human subjects. *Gut* **29**, 1035–1041 (1988).
38. Jia, B., Park, D., Hahn, Y. & Jeon, C. O. Metagenomic analysis of the human microbiome reveals the association between the abundance of gut bile salt hydrolases and host health. *Gut Microbes* **11**, 1300–1313 (2020).
39. Jones, B. V., Begley, M., Hill, C., Gahan, C. G. M. & Marchesi, J. R. Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. *Proc. Natl Acad. Sci. USA* **105**, 13580–13585 (2008).
40. Karlov, D. S. et al. *Structure of Lactobacillus salivarius (Ls) bile salt hydrolase(BSH) in complex with taurocholate (TCA)* (Worldwide Protein Data Bank, 2023); https://doi.org/10.2210/pdb8blt/pdb.

Methods

Reaction conditions for enzyme characterization and determining acyl transfer kinetics

Lyophilized *C. perfringens* BSH/T (BSH/T, Creative Enzymes) was resuspended in 0.1 M phosphate buffer at pH 7.0 to a concentration of 2 units μl^{-1}. Stocks of 100 mM GCA and TCA were prepared in water and CA stocks were produced in dimethylsulfoxide. Each enzyme reaction was run in triplicate in bicarbonate, Tris or citrate-phosphate buffer at the indicated pH using 0.1 units μl^{-1} of enzyme, 2.5 mM BA (TCA, GCA or CA) and 125 μM complete amino acid mixture (Promega) for 6.25 μM individual amino acid concentration. Reactions were incubated at 37 °C and 30 μl aliquots of the reaction were quenched by the addition of 45 μl of ice-cold methanol at each timepoint for a final concentration of 60% methanol (v/v). Extracts were stored at −80 °C before mass spectrometry analysis.

We examined the kinetics of the acyl transfer reaction using phenylalanine, CA and TCA. Phenylalanine stocks were prepared in sterile water followed by 0.2 μm filtering. Reactions were prepared in 0.1 M citrate buffer at pH 5.3 with a final concentration of 0.05 units μl^{-1} of enzyme and 8% dimethylsulfoxide. Reactions were sampled at 1, 5, 10, 15 and 20 min and quenched with cold methanol. Reactions were brought to a final concentration of 50% methanol (v/v). TCA deconjugation kinetics were determined by the addition of 1–8 mM, as previously described[13]. Reaction velocities for phenylalanine transfer were determined by the addition of 1–5 mM phenylalanine and 8 mM CA or TCA. Extracts were stored at −80 °C before mass spectrometry analysis.

Bacterial strains, media and growth conditions

All media used in anaerobic experiments were prereduced for at least 24 h in the anaerobic chamber before inoculation and cultures were grown in an atmosphere containing 98% nitrogen and 2% hydrogen. Bacterial cultures were grown from glycerol freezer stocks in reinforced clostridial medium (RCM, Merck), brain heart infusion medium (BHI, Merck) and BHI supplemented with 5 $\mu g\ ml^{-1}$ of hemin, 1 $\mu g\ ml^{-1}$ of vitamin K, 10 $g\ l^{-1}$ of yeast extract and 0.5 $g\ l^{-1}$ of L-cysteine (BHIS). A full list of strains used in this work can be found in Supplementary Table 1.

C. perfringens bsh/t was amplified and inserted into pBAD18-Cm through Gibson assembly (New England Biolabs) followed by cloning into *E. coli* DH5α. Plasmid purification was performed through Mini-Prep (Qiagen) and inserts were verified through PCR. Site-directed mutagenesis was performed for codons of residues Asn82 and Cys2 using a Q5 site-directed mutagenesis kit (New England BioLabs) with mutations confirmed through Sanger sequencing. Primers can be found in Supplementary Table 9. To compare MCBA production profiles, overnight cultures of each strain were diluted to optical density measured at 600 nm OD_{600} = 0.01 in LB with a final concentration of 1 mM CA/TCA/GCA or 1% dimethylsulfoxide, 1 $mg\ ml^{-1}$ of arabinose, 100 μM taurine and 20 $\mu g\ ml^{-1}$ of chloramphenicol. Cultures were incubated aerobically for 24 h at 37 °C with 220 rpm shaking. OD_{600} was measured and metabolite extraction was performed by diluting whole cultures 2:3 (v/v) in 100% ice-cold methanol in 1.7 ml microcentrifuge tubes (Axygen) followed by overnight incubation at 4 °C. Extracts were then centrifuged at 10,000*g* for 5 min to pellet cell debris followed by storage at −80 °C before liquid chromatography–tandem mass spectrometry analysis (LC–MS/MS).

In vitro screen for MCBA production

Overnight cultures were grown from freezer stocks, anaerobically at 37 °C. Once OD_{600} reached at least 0.10, cultures were then diluted to a final OD_{600} of 0.01 in media with or without 1 mM CA and 100 μM taurine in 96-deep-well plates (Thermo Fisher Scientific). Taurine is absent in all standard media formulations and was therefore supplemented to screen TCA production. Plates were sealed with a rubber mat (Thermo Fisher Scientific) and incubated for 24 h at 37 °C under anaerobic conditions. The OD_{600} was again measured following incubation and metabolites were extracted as previously described.

Untargeted metabolomics for BA analysis

Bacterial culture extracts were diluted 1:1 (v/v) in 50% methanol containing 2.5 $\mu g\ ml^{-1}$ of phenol red internal standard before LC–MS/MS. Extracts for measuring enzyme activity were not diluted before LC–MS/MS analysis. LC was performed using a Thermo Scientific Vanquish Autosampler and an Acquity UPLC BEH C-18 column, 2.1 × 100 mm^2 (Waters). MS was performed using a Thermo Scientific Q Exactive hybrid quadrupole-orbitrap mass spectrometer running in positive ion mode. All analyses used a 10 μl injection volume, 0.4 ml min^{-1} flow rate and 60 °C column temperature. Samples were eluted using a linear solvent gradient of water (A) and acetonitrile (B), each containing 0.1% formic acid, across a 12 min chromatographic run as follows: 0–1 min, 2% B; 1–8 min, 2–100% B; 8–12 min, 100% B; 10–12 min, 2% B. Data were collected using electrospray ionization in positive mode. MS^1 data were collected using a 35,000 resolution, AGC target of 1×10^6, maximum injection time of 100 ms and a scan range set from 100 to 1,500 *m/z* (during min 1–10). Data-dependent MS^2 spectra were collected for the top five most abundant peaks identified in MS^1 survey scans. Files were converted to mzXML format through GNPS Vendor Conversion and submitted the Global Natural Products Social Molecular Networking Database (GNPS, University of California at San Diego) for molecular networking and spectral identification[41].

For *bsh/t* expression and mutagenic studies, converted files were submitted to GNPS for classic molecular networking to identify MCBAs present in each sample. Peak area-under-curve (AUC) abundances were calculated using Thermo Scientific XCaliber software based on *m/z* and retention time of each MCBA annotated by GNPS (Supplementary Table 10).

Targeted metabolomics for BA quantification

For mouse and human samples, BAs were quantified by running a standard mix of known concentrations of various BAs (Supplementary Table 11) as an eight-point standard curve using the same LC–MS/MS as described above. The standard curve samples were added to the end of the run for all samples for which quantification was applied. Data were processed with MZmine3 (v.3.3.0)[42] to obtain AUC abundance for BAs present in both standards and analysed samples. BA concentrations were determined on the basis of the equation of the curve that fit either a linear or power function, depending on the ionization behaviour of individual BAs. Because many MCBAs were detected for which there were no available standards, we used the standard curve from AlaCA to calculate pseudoconcentrations of these compounds.

LC–MS for BSH/T conjugation kinetics

Enzyme extracts were run without dilution. LC was performed as described above and MS^1 data were collected as described above, with the exception that 70,000 resolution was used. PheCA concentrations were calculated using Thermo Scientific XCaliber software and an eight-point standard curve containing labelled standards.

Assaying MCBA impacts on bacterial growth

Pure cultures were grown in reduced RCM and incubated overnight under anaerobic conditions at 37 °C. The OD_{600} was measured followed by dilution to a final OD_{600} of 0.01 in clear-bottom 96-well plates containing RCM supplemented with 1 mM BA or vehicle (dimethylsulfoxide). Growth curves were generated through a BioTek Synergy HTX plate reader equipped with Gen5 imaging software (v.3.10, Agilent). Plates were incubated at 37 °C under aerobic or anaerobic conditions, as described above, with 205 cycles per min orbital shaking. OD_{600} was measured every 15 min for 24 h for anaerobic growth, every 20 min for 20 h for aerobic growth. Measurements were blank-corrected and subsequent growth curve analyses were performed in R (v.4.2.2)[43] and

R studio (v.2023.06.2 + 561, Posit). Growth curves were analysed using the 'growthcurver' R package (v.0.3.1)[44] and comparisons were drawn between fold-change differences in the logarithmic AUC. ED_{50} values in Extended Data Fig. 4 were determined using terminal OD_{600} after growing *L. aerotolerans* or *P. anaerobius* in RCM supplemented with 0–1,000 µM PheCA, TyrCA, LeuCA or CA for 24 h.

Phylogenetic analysis and BSH/T visualization

Genomic sequences were acquired from GenBank and BSH/T amino acid sequences were obtained from the Joint Genome Institute and the National Center for Biotechnology Information (NCBI) protein databases (Supplementary Table 1). Phylogenetic trees were constructed using FastTree and visualized in R using the 'ggtree' package (v.3.6.2)[45]. BSH/T sequences were aligned using the NCBI constraint-based aligner tool (COBALT)[46]. Parameters of the alignment were set to defaults, including an E-value of 0.003, word size of 4 and maximum cluster distance of 0.8. The three-dimensional structure of CpBSH/T (PDB ID 2BJG; refs. 17,19) was visualized using the Protein Data Bank online structure viewer, Mol*Viewer[47]. The three-dimensional structure of *L. salivarius* BSH/T (PDB ID 8BLT; refs. 21,40) was visualized using PyMOL (v.2.5.4, Schrödinger).

For *C. scindens* predictive *bsh/t* analysis, genomic sequences were obtained from NCBI (Supplementary Table 2; $n = 35$). Prodigal (v.2.6.3)[48] was used to predict protein-encoding genes present in each genome followed by searching for predicted BSH/T sequences using to DIAMOND (v.0.9.36.137)[49].

Microbial 16S rRNA gene sequencing

DNA from mouse faeces and tissue was extracted using the Quick-DNA Faecal/Soil Microbe Miniprep kit (Zymo) according to the manufacturer's instructions. To test extraction efficacy, full-length 16S ribosomal RNA genes were amplified using primers 27f and 1492r and analysed through gel electrophoresis. Subsequent microbiome sequencing was performed using Illumina compatible primers 515f and 806r to amplify the V4 hypervariable region of the 16S rRNA gene. Sequencing was performed at the Michigan State University RTSF Genomics Core following the protocol previously described by ref. 50. PCR products were batch normalized through SequalPrep DNA Normalization plate (Invitrogen) and product recovered from the plates was pooled. This pool was concentrated and cleaned up using a QIAquick Spin column (Qiagen) and AMPure XP magnetic beads (Beckman Coulter). Quality was checked and quantified using a combination of Qubit dsDNA HS (Thermo Fisher Scientific), 4200 TapeStation HS DNA1000 (Agilent) and Collibri Illumina Library Quantification qPCR assays (Invitrogen). This pool was loaded onto a MiSeq v.2 standard flow cell and sequencing was carried out in a 2 × 250 base pair paired end format using a MiSeq v.2 500 cycle reagent cartridge. Custom sequencing and index primers complementary to the 515f/806r oligomers were added to appropriate wells of the reagent cartridge. Base calling was done by Real Time Analysis v.1.18.54 (RTA, Illumina) and output of RTA was demultiplexed and converted to FastQ format with Bcl2fastq v.2.20.0 (Illumina).

Raw sequences were analysed using Qiita[51], a web-based QIIME 2 (ref. 52) analysis platform. Sequences were filtered on the basis of quality to generate amplicon sequence variants through the Deblur method[53]. Taxonomy was assigned using the q2-feature-classifier against the 99% SILVA 16S rRNA gene sequence database (release 138)[54,55]. Sample data were rarefied to 8,000 reads per sample and core diversity metrics, such as Bray–Curtis dissimilarity, were calculated. We performed statistical analysis in R and random forest classification was performed using the 'randomforest' package (v.4.7-1.1)[56,57].

Ethics statement

All mouse experiments were approved by the Institutional Animal Care and Use Committee at Michigan State University. Animal health was routinely assessed by laboratory technicians as well as the Michigan State University veterinary staff.

Animals, housing, BA dosing and sample collection

C57BL/6J mice were purchased from Jackson Laboratories and acclimated in the new facility for 1 week before BA administration through oral gavage. Cage changes were performed weekly in a laminar flow hood by core facility staff. Mice were housed under a 12 h cycle of light and darkness. Male and female 6-week-old C57BL/6J mice ($n = 5$ per sex, per group) were administered 100 mg kg^{-1} of TCA, PheCA or SerCA dissolved in corn oil through daily oral gavage for 14 d. A control group was administered corn oil alone (vehicle). Treatments were randomized upon receiving the mice and blinding was not used. Longitudinal faecal samples and weights were collected from individual mice daily throughout the duration of treatment. On day 13, mice were fasted for approximately 12 h to clear faecal material from the gut before necropsy and tissue collection. Animals used for necropsy and tissue collection were euthanized humanely through anaesthesia using isoflurane followed by cervical dislocation. Before analysis, phosphate buffered saline (Sigma) was added 3:1 (v/w) to faecal samples while 200 µl was added to caecum samples followed by homogenization through bead bashing at maximum speed for 10 min using a Bead Ruptor 96 (Omni International).

Male and female C57BL/6Crl mice were bred inhouse and pups were weaned at 3 weeks of age. At P38, mice were singly housed in cages lined with paper towels and weights were recorded. Mice were allowed to acclimate to singly housed environments for 3 days. After this acclimation period, paper towel lining was replaced and mice were fasted for 12 h before training (P41). The next day, mice were given a plain peanut butter pellet and were observed to determine whether they would consume the pellet. Normal chow was returned to the cages and all mice that consumed the pellet were advanced to the next stage of training in which they were given a plain peanut butter pellet at approximately the same time of day for three consecutive days. On the third day of non-fasted training (P45), the mice were again weighed. All mice that successfully completed training (consumed the plain peanut butter training pellet within 1 h) were included in subsequent studies. The pilot mixed MCBA dosing experiment lasted for 5 days (starting at P46, day 0). Each day, mice were weighed and treated with either mock peanut butter pellets or pellets containing 10 mg kg^{-1} of each of the following for a total dose of 80 mg kg^{-1}: AlaCA, AspCA, GluCA, LeuCA, PheCA, SerCA, ThrCA and TyrCA. Treatments were randomized for mice that successfully completed PBFM training and blinding was not used. Faecal samples were collected on days 0, 1, 3 and 5. Paper towel linings were replaced to refresh cages on days 0 and 3. On day 5 (P51), mice were euthanized through anaesthesia using isoflurane followed by cervical dislocation. Tissue samples were collected and flash frozen in liquid nitrogen immediately after collection.

The experimental phase of the individual MCBA study including equimolar amino acid-BA controls lasted for 10 days (starting at P46, day 0). Mice were administered MCBAs using the PBFM (first developed by ref. 29; Supplementary Information), a means of administering hydrophobic compounds to mice in a palatable and controlled manner. Each day, mice were weighed and fed peanut butter pellets containing 10 mg kg^{-1} of the following: TCA, equimolar taurine and CA (Taur + CA), SerCA, equimolar serine and CA (Ser + CA), PheCA, equimolar phenylalanine and CA (Phe + CA), or a vehicle containing just peanut butter (five mice per sex per group). Faecal samples were collected on days 0, 1, 4, 7 and 10. Paper towel linings were replaced to refresh cages on days 0, 3, 6 and 9. On day 10 (P56), mice were euthanized through anaesthesia using isoflurane followed by cervical dislocation. Tissue samples were collected and flash frozen in liquid nitrogen immediately after collection.

All mice were housed under a 12 h cycle of light and darkness and treatments, weights, faecal collections and paper towel changes were conducted in a laminar flow hood. Calculations for dosage for

treatments were done using weights at P45. Before analysis, phosphate buffered saline was added 3:1 (v/w) to faecal samples and 5:1 (v/w) to tissue samples followed by homogenization through bead bashing at 20 s^{-1} for 30 s with 1 min of rest three times using a Bead Ruptor 96 (Omni International). Statistical methods were not used to predetermine sample size and blinding was not used.

Sample collection from patients undergoing sleeve gastrectomy, processing and analysis

This prospective, single-arm study enroled 44 obese patients participating in our health system's (Corewell Health (formerly Beaumont Health), Royal Oak, MI) bariatric surgery programme and planning on sleeve gastrectomy (SG). This study was approved by the Beaumont Institutional Review Board (IRB no. 2017-201) and reviewed by the Michigan State University IRB (STUDY00003064*)*. All participants provided informed consent before participating in the study. Participants were recruited from the Royal Oak Weight Control Center, an affiliate of the Corewell Health William Beaumont University Hospital, during the pre-operative bariatric surgery process. This involves a medical work-up, surgical risk stratification and multidisciplinary education before moving on to surgery. Information about the study was presented at the free informational bariatric surgery seminar which prospective patients attend before starting the bariatric surgery programme. Fliers about the trial were posted and distributed at the Beaumont Weight Control Center and mailed to patients with the bariatric surgery approval letter. The approval letter is written by a Weight Control Center physician and indicates that a patient is approved from a medical and multidisciplinary team perspective to move forward with bariatric surgery. Inclusion criteria followed the National Institutes of Health criteria for bariatric surgery: body mass index (BMI) at or above 40 kg m^{-2} or a BMI of 35–40 kg m^{-2} with an obesity comorbidity such as type 2 diabetes, heart disease or obstructive sleep apnoea[58]. A further inclusion requirement was being between 18 and 70 years old. Patients were excluded if they had poorly controlled medical or psychiatric conditions which, in the opinion of the investigator, made the patient unlikely to be able to properly participate in the study. Biases include self-selection bias and that only patients planning the SG bariatric surgical procedure were included. It is unlikely that these biases had a significant impact on the results of this study.

Faecal specimens were provided by the William Beaumont Research Institute biorepository. Demographic data were collected, weight and height were measured and BMI was calculated on enrolment into the bariatric surgery programme. Weight was measured again on the morning of SG surgery and at 3 months post-SG. Faecal samples were collected pre-operatively and 3 months following SG.

Faecal samples were extracted 1:5 (w/v) in 70% LC–MS grade ice-cold methanol and BA concentrations were calculated on the basis of targeted analysis described above. One faecal sample was lost during extraction, therefore resulting in a cohort of 44 subjects with paired faecal samples before and after SG. The extracts were spun in a microcentrifuge at 12,000*g* to pellet protein and the methanol supernatant was diluted 1:1 (v/v) in 50% methanol before mass spectrometry analysis. LC–MS/MS protocols were the same as described above for all untargeted metabolomics analysis of microbial and mouse samples. Data were processed with MZmine and GNPS feature-based molecular networking as previously described.

Reporting summary

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

Data availability

Protein structures are available on the Protein Data Bank. *C. perfringens* BSH/T in complex with DCA and taurine, from refs. 17,19, is available under PDB ID 2BJG (https://doi.org/10.2210/pdb2BJG/pdb). *L. salivarius* BSH/T in complex with TCA, from refs. 21,40, is available under PDB ID 8BLT (https://doi.org/10.2210/pdb8BLT/pdb). Raw mass spectrometry data are publicly available in the MassIVE database (massive.ucsd.edu) for the in vitro screen for MCBA production under MSV000090234 (https://doi.org/10.25345/C5S756Q1B), for CpBSH/T variant analysis under MSV000092138 (https://doi.org/10.25345/C55D8NQ9V), for MCBA gavage samples under MSV000093173 (https://doi.org/10.25345/C57S7J35N), for mixed MCBA PBFM dosing under MSV000093171 (https://doi.org/10.25345/C5H98ZQ3R), for 100 mg^{-1} kg of SerCA PBFM dosing at MSV000093169 (https://doi.org/10.25345/C5RV0DB2C), 10 mg^{-1} kg of MCBA PBFM dosing under MSV000093172 (https://doi.org/10.25345/C5CJ87W9C) and for SG faecal samples under MSV000093167 (https://doi.org/10.25345/C51834C9N). GNPS molecular networks are available for the MCBA production screen at gnps.ucsd.edu/ProteoSAFe/status.jsp?task=565151309a874d5f97caa3f383c95382, for CpBSH/T incubation with 1 mM BA and equimolar amino acid mix at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=3dec8f7ab26d47098406a7e597825154 and https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=33da5da024ed44848770a4a02b119d9e, for the CpBSH/T mutagenesis experiment at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=30c88ca297a44f84be5fa32b376e5cb9 and for the SG faecal samples at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=f11eaab1cf1d43b1a5f754575d171e87. 16S rRNA gene amplicon data were deposited in the EMBL-EBI European Nucleotide Archive. Data from the 100 mg kg^{-1} gavage experiment can be found under project PRJEB68000, study accession ERP153011. Tissue data from the 10 mg kg^{-1} PBFM experiment can be found under project PRJEB68146, study accession ERP153132. Faecal data available from 10 mg kg^{-1} PBFM experiment are available under project PRJEB68149, study accession ERP153135. Analyses can be found on Qiita under analysis ID 53128 for the 100 mg kg^{-1} gavage and IDs 57407 and 57481 for tissue and faecal samples, respectively, from the 10 mg kg^{-1} PBFM experiment. Source data are provided with this paper.

41. Nothias, L.-F. et al. Feature-based molecular networking in the GNPS analysis environment. *Nat. Methods* **17**, 905–908 (2020).
42. Pluskal, T., Castillo, S., Villar-Briones, A. & Orešič, M. MZmine 2: modular framework for processing, visualizing and analyzing mass spectrometry-based molecular profile data. *BMC Bioinform.* **11**, 395 (2010).
43. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2022).
44. Sprouffske, K. & Wagner, A. Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinform.* **17**, 172 (2016).
45. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T. Y. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017).
46. Papadopoulos, J. S. & Agarwala, R. COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics* **23**, 1073–1079 (2007).
47. Sehnal, D. et al. Mol*Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Res.* **49**, W431–W437 (2021).
48. Hyatt, D. et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform.* **11**, 119 (2010).
49. Buchfink, B., Reuter, K. & Drost, H.-G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* **18**, 366–368 (2021).
50. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* **79**, 5112–5120 (2013).
51. Gonzalez, A. et al. Qiita: rapid, web-enabled microbiome meta-analysis. *Nat. Methods* **15**, 796–798 (2018).
52. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
53. Amnon, A. et al. Deblur rapidly resolves single-nucleotide community sequence patterns. *mSystems* **2**, e00191–16 (2017).
54. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
55. Yilmaz, P. et al. The SILVA and 'all-species Living Tree Project (LTP)' taxonomic frameworks. *Nucleic Acids Res.* **42**, D643–D648 (2014).
56. Breiman, L., Cutler, A., Liaw, A. & Wiener, M. randomForest: Breiman and Cutler's Random Forests for classification and regression. R package v. 4.7-1.1 (2022).
57. Liaw, A. & Wiener, M. Classification and regression by randomForest. *R News* **2**, 18–22 (2002).

58. *Potential Candidates for Bariatric Surgery* (National Institute of Diabetes and Digestive and Kidney Diseases, 2020); www.niddk.nih.gov/health-information/weight-management/bariatric-surgery/potential-candidates.

Acknowledgements We thank C. M. Waters, N. D. Hammer and K. Parent for generously providing some of the strains used in this work in addition to P. Lawson for his help and input acquiring strains from the Culture Collection, University of Gothenburg. We would also like to thank J. B. Gomez, E. N. Ottosen and K. C. Ford for their guidance. This work was funded by Michigan State University and the Global Grants for Gut Health, cosponsored by Yakult and Nature Research.

Author contributions D.V.G., S.F.G. and R.A.Q. designed the project. D.V.G. and R.A.Q. discovered acyl transfer activity by bile salt hydrolase and performed phylogenetic analysis. D.V.G. generated data for in vitro BSH/T characterization, species-specific MCBA production screens, MCBA antimicrobial activity and heterologous *bsh/t* expression experiments. D.V.G. and C.B. generated data for pancreatic carboxypeptidase activity. M.O. raised mice. M.O., M.S. and B.A. conducted work with animals. D.V.G., M.O., M.S., B.A., C.B. and Y.F. collected samples and generated data from animal experiments. W.M.M., K.M.Z., M.D.S. and M.E.M. coordinated human sample collection, treated patients, performed bariatric surgeries and completed all clinical follow-up. D.V.G., C.M. and R.A.Q. analysed data. M.O., A.L.S., S.F.G., R.P.H. and R.A.Q. guided experimental design and analysis. D.V.G. and R.A.Q. wrote the manuscript.

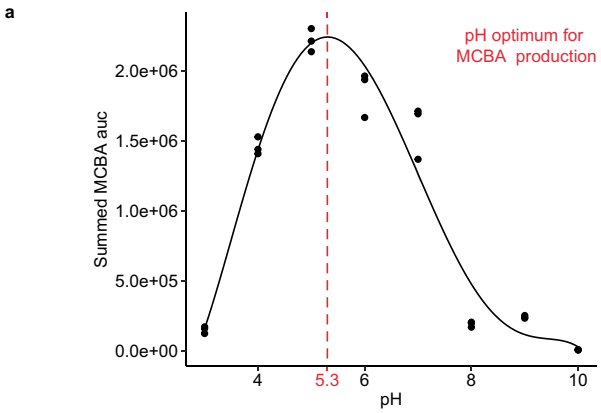
Competing interests The authors declare no competing interests.

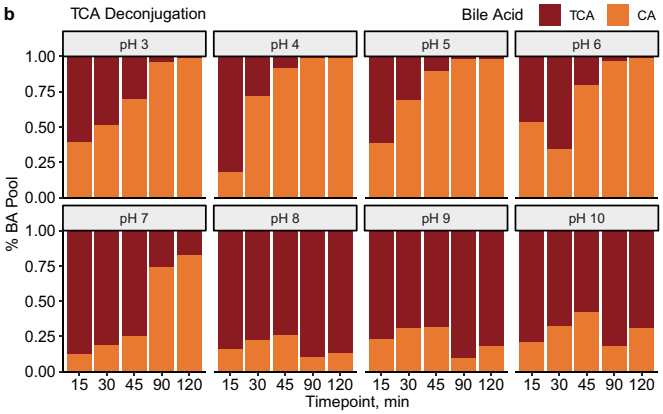
Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-024-07017-8.

Correspondence and requests for materials should be addressed to Robert A. Quinn.

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.



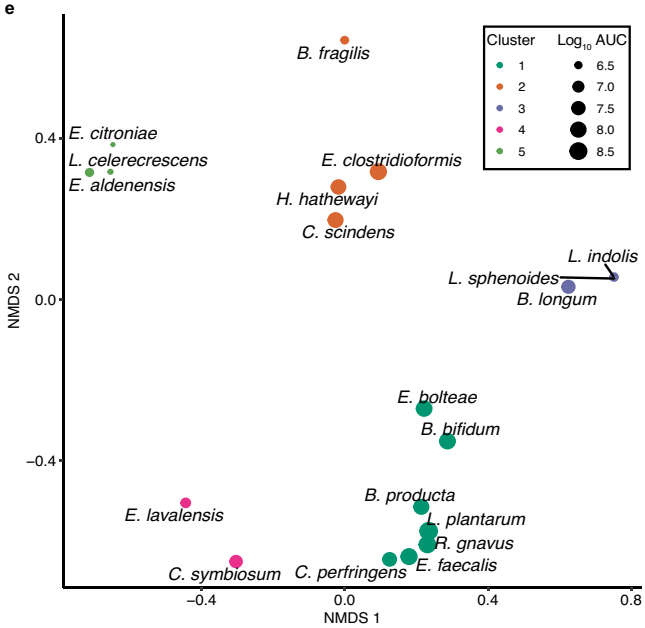


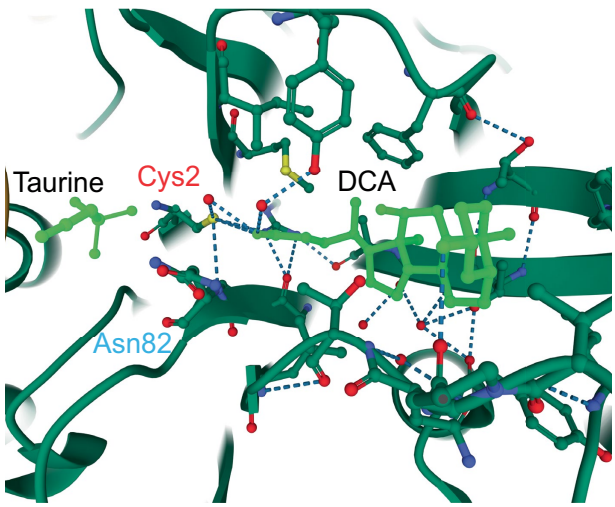
c

	Dependent variable:				
	y				
	-1	poly(2)	poly(3)	poly(4)	**poly(5)**
Constant	2.00e+06***	9.63e+05***	9.63e+05***	9.63e+05***	**9.63e+05*** **
	-4.90E+05	-1.06E+05	-4.75E+04	-4.79E+04	**-4.48E+04**
x	-1.60e+05**	-1.80e+06***	-1.80e+06***	-1.80e+06***	**-1.80e+06*** **
	-7.11E+04	-5.19E+05	-2.33E+05	-2.35E+05	**-2.20E+05**
x2		-2.89e+06***	-2.89e+06***	-2.89e+06***	**-2.89e+06*** **
		-5.19E+05	-2.33E+05	-2.35E+05	**-2.20E+05**
x3			2.14e+06***	2.14e+06***	**2.14e+06*** **
			-2.33E+05	-2.35E+05	**-2.20E+05**
x4				1.96E+05	**1.96E+05**
				-2.35E+05	**-2.20E+05**
x5					**-4.23e+05***
					-2.20E+05
Observations	24	24	24	24	**24**
R^2	0.187	0.672	0.937	0.939	**0.95**
Adjusted R^2	0.15	0.641	0.928	0.926	**0.936**
Residual Std. Error	7.98E+05 (df = 22)	5.19E+05 (df = 21)	2.33E+05 (df = 20)	2.35E+05 (df = 19)	**2.20E+05 (df = 18)**
F Statistic	5.061** (df = 1; 22)	21.526*** (df = 2; 21)	99.244*** (df = 3; 20)	73.479*** (df = 4; 19)	**67.932*** (df = 5; 18)**

d

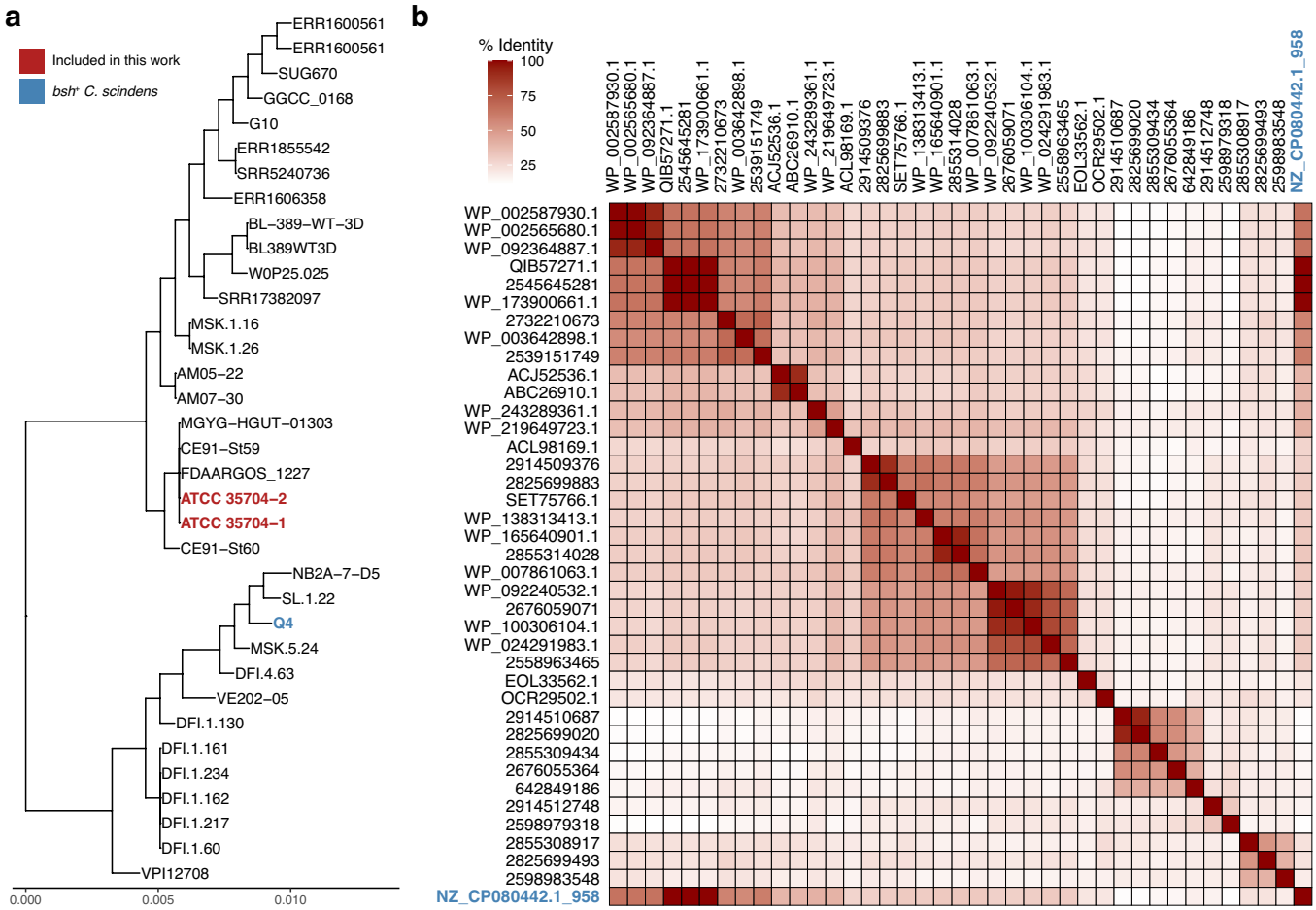
Amino Acid	TCA	CA	GCA
Ala	1.88E+04 (1.22E+02)	1.54E+04 (5.90E+02)	7.59E+03 (3.77E+02)
Arg	2.95E+05 (7.18E+03)	2.68E+05 (9.36E+03)	1.72E+05 (3.18E+03)
Asn	1.75E+05 (5.76E+03)	0.00E+00 (0.00E+00)	0.00E+00 (0.00E+00)
Cys	7.22E+04 (9.33E+01)	0.00E+00 (0.00E+00)	0.00E+00 (0.00E+00)
Gln	0.00E+00 (0.00E+00)	1.38E+05 (1.44E+03)	6.90E+04 (3.22E+03)
Glu	5.04E+04 (1.59E+02)	0.00E+00 (0.00E+00)	0.00E+00 (0.00E+00)
Gly	5.81E+04 (2.70E+03)	6.90E+03 (2.74E+03)	N/A
His	7.42E+05 (3.08E+04)	5.63E+05 (1.88E+04)	4.14E+05 (5.33E+03)
Ile/Leu	1.12E+05 (5.52E+03)	1.29E+05 (1.04E+03)	7.17E+04 (1.61E+03)
Lys	1.41E+05 (4.59E+03)	9.13E+04 (3.59E+03)	6.85E+04 (1.82E+03)
Met	0.00E+00 (0.00E+00)	1.16E+05 (2.76E+02)	6.40E+04 (2.59E+03)
Phe	1.29E+05 (4.83E+03)	0.00E+00 (0.00E+00)	0.00E+00 (0.00E+00)
Ser	6.60E+04 (1.08E+03)	0.00E+00 (0.00E+00)	0.00E+00 (0.00E+00)
Thr	8.54E+04 (2.66E+03)	7.32E+04 (3.22E+03)	3.95E+04 (1.42E+03)
Trp	1.36E+05 (4.85E+03)	4.84E+04 (6.61E+03)	3.20E+04 (1.27E+03)
Tyr	1.34E+05 (4.02E+03)	1.07E+05 (2.01E+03)	5.44E+04 (9.38E+02)
Val	0.00E+00 (0.00E+00)	5.48E+04 (1.41E+03)	3.33E+04 (3.73E+02)



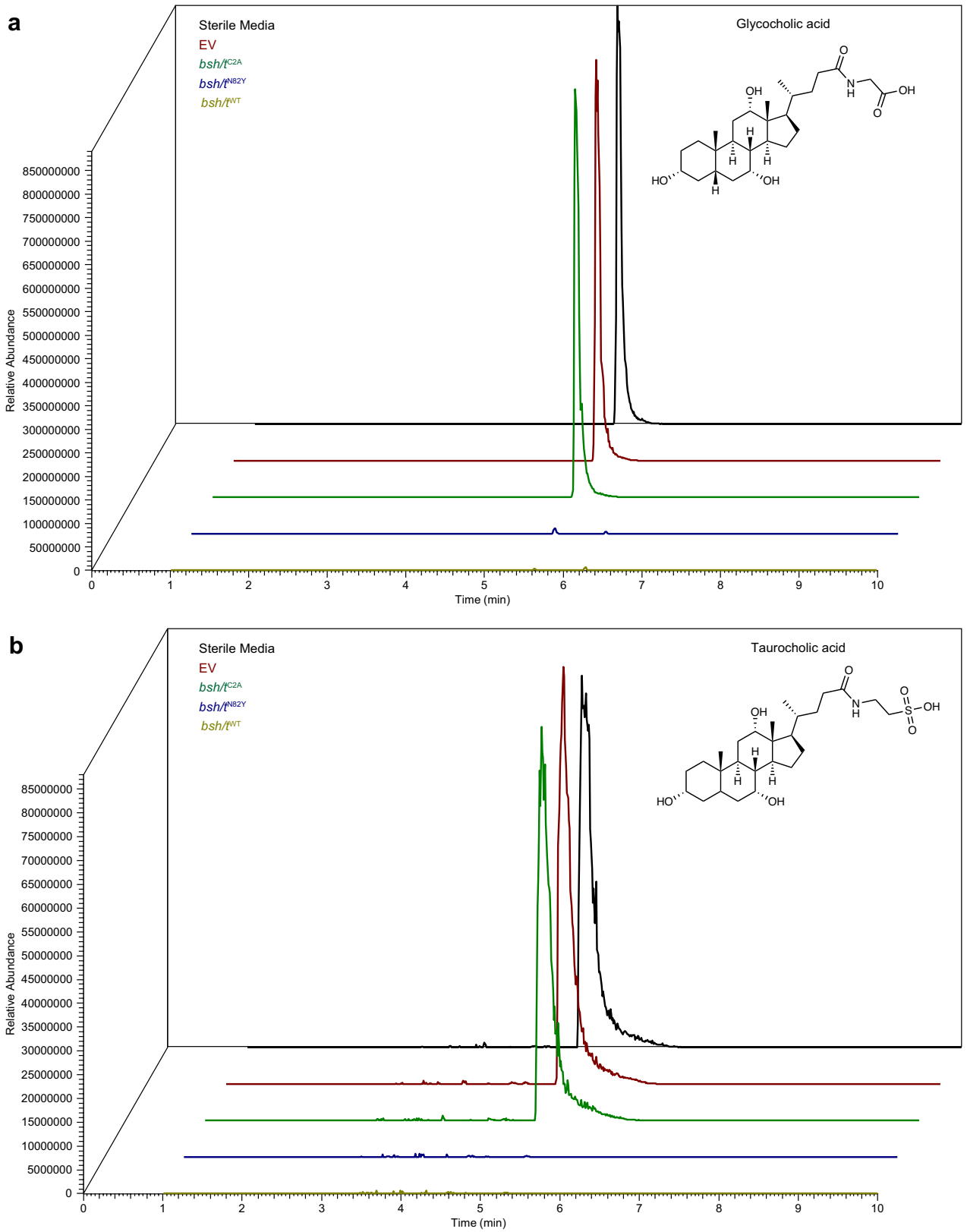


Extended Data Fig. 1 | See next page for caption.

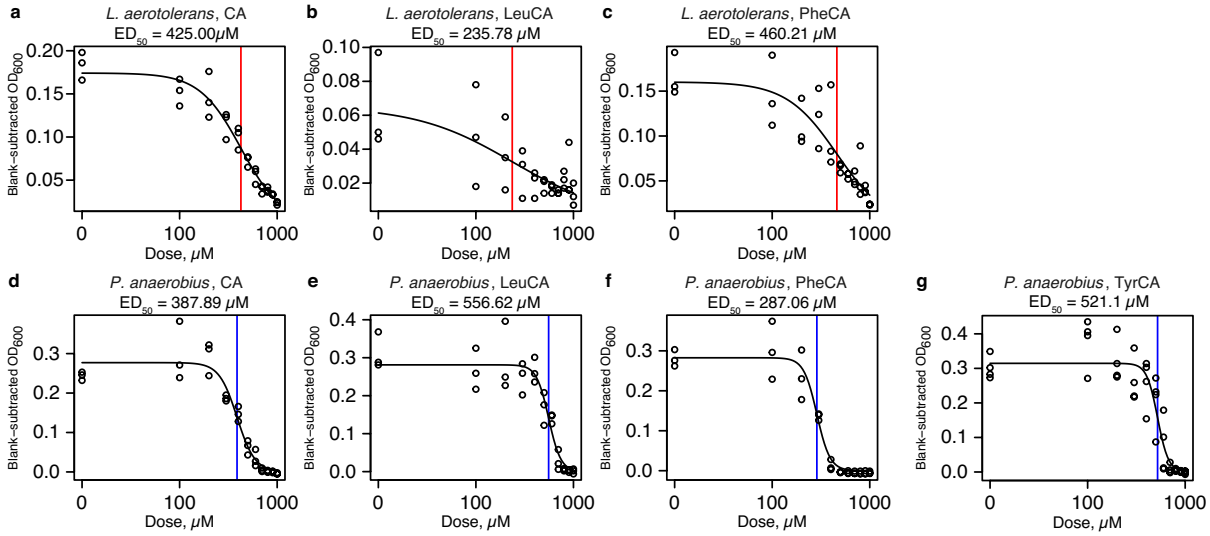
Extended Data Fig. 1 | MCBA abundance following purified CpBSH/T or microbial incubation with BAs. a, Summed MCBA auc after 120 min incubation of CpBSH/T, 2.5 mM TCA and 125 µM equimolar amino acid mix at various pH revealing pH-dependence of BA conjugation, $n = 3$ independent reactions. Red dashed line indicates the pH 5.3 optimum following derivation as determined by fitting a 5-factor polynomial equation. **b**, Proportion of TCA and CA in the bile acid pool when CpBSH/T was incubated with 8 mM TCA at different pH values across time. $n = 3$ replicates. **c**, Goodness of fit outputs for curve fitting to determine CpBSH/T pH optimum, values in parentheses as sem. Equation used to calculate pH optimum (5-factor, bolded) based on adjusted R^2. Coefficient significance determined by one-sided t test and model significance determined by one-way ANOVA without *P* value adjustment. $^*P < 0.1$, $^{**}P < 0.05$ $^{***}P < 0.01$. **d**, Relative abundances of MCBAs produced by purified CpBSH/T. Enzyme was individually provided 1 mM BA and an equimolar mix of amino acids, buffered at pH 5, and sampled after 120-min incubation at 37 °C. Data presented as mean auc abundance (sem), $n = 3$. **e**, Nonmetric data scaling using Bray-Curtis dissimilarity of amino acids used in BA conjugation, using average amino acid auc per strain. Colour represents cluster assigned based on cluster analysis and dot size represents the average total MCBA abundance. $n = 3$ independent cultures. **f**, *Cp*BSH/T (PBD ID: 2bjg)[17,19] cocrystalized with TDCA and residues important for BA deconjugation are highlighted in addition to Asn82, the residue playing a key role in BA reconjugation specificity.



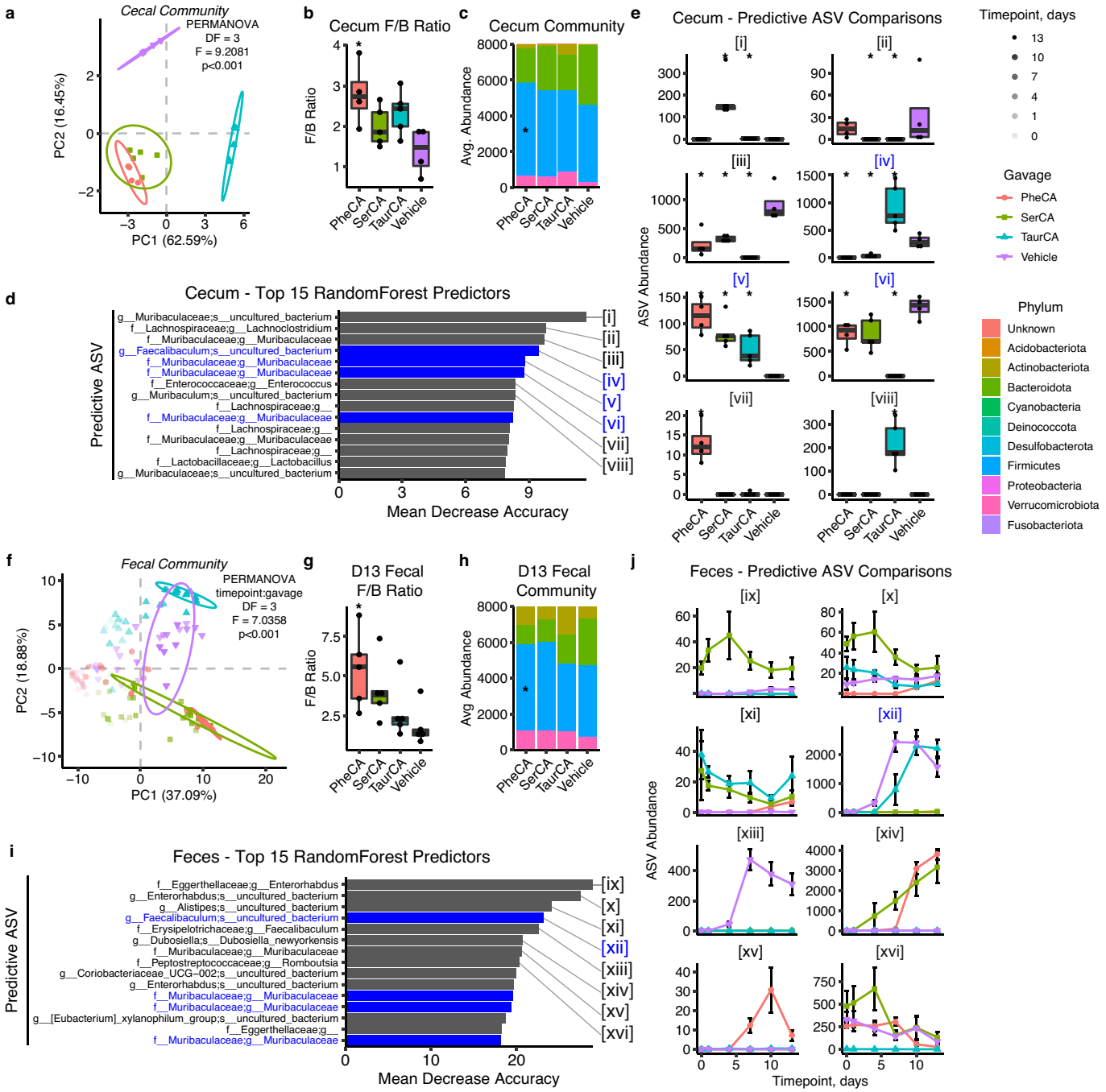
Extended Data Fig. 2 | *Clostridium scindens* genome analysis for putative *bsh/t* annotation. a, Phylogenetic analysis of 35 publicly available genomes for *C. scindens*. The ATCC type strain, used in this work, has two deposited genomes and is highlighted in red. The only strain with predicted *bsh/t* was *C. scindens* strain Q4, highlighted in blue. **b**, Pairwise BSH/T amino acid sequence similarity of all strains included in this work (matching Fig. 2c), now including the predicted BSH/T present in *C. scindens* strain Q4 (NZ_CP080442.1_958, based on Prokka analysis).



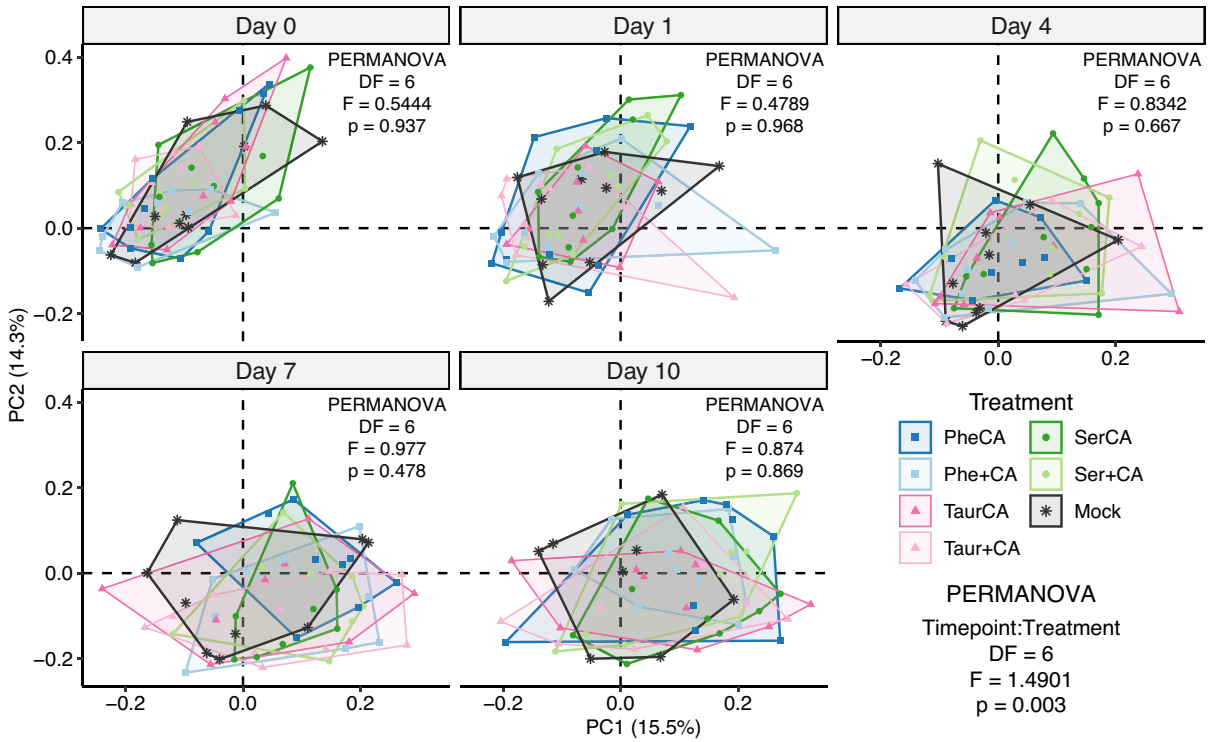
Extended Data Fig. 3 | GCA and TCA extracted ion chromatograms following 24 h induction of *C. perfringens* BSH/T variants in *E. coli*. Representative **a**, GCA and **b**, TCA extracted ion chromatograms showing significantly diminished in WT and N82Y variant strains with minimal change in the C2A variant and EV control.



Extended Data Fig. 4 | Amino acid-dependency of MCBA antimicrobial efficacy. Dose–response curves for *L. aerotolerans* when grown for 24 h in **a**, CA, **b**, LeuCA, or **c**, PheCA with calculated ED_{50} shown in red. Dose–response curves for *P. anaerobius* when grown for 24 h in **d**, CA, **e**, LeuCA, **f**, PheCA, or **g**, TyrCA with ED_{50} shown in blue. $n = 4$ independent cultures per strain.

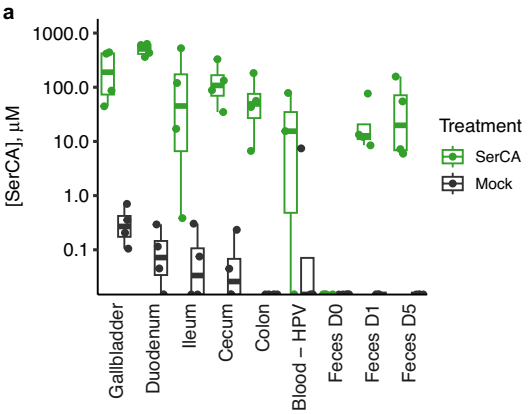


Extended Data Fig. 5 | Microbiome shifts in female mice following 100 mg kg^{-1} MCBA gavage. a Caecal and **f** faecal samples were subjected to principal coordinate analysis (PCoA) of microbiome community structure via Bray–Curtis dissimilarity after oral gavage of different MCBAs in C57BL/6 mice. Significance determined by PERMANOVA. The ratio of Firmicutes/Bacteroidota (F/B ratio) between gavage groups for **b** caecum and **g** faecal samples at day 13, with corresponding phylum-level community profiles for both **c** caecum and **h** faecal samples. Bar charts showing the top 15 bacterial groups impacting the mean decrease in accuracy of random forest classification based on MCBA gavage group for **d** caecal and **i** faecal samples from days 1–13, with higher values indicating greater contribution to the predictive model. ASVs highlighted in blue represent those that matched between caecal and faecal classifications. Comparisons between gavage groups for the top predictive ASVs in **e** caecal samples and **j** faecal samples over time, with blue graph titles indicating shared features between the top 15 predictors in both analyses. Caecum 16S analysis, n = 5 per group, 4 for PheCA; faecal 16S analysis, $n = 5$ per group, per timepoint. Ellipses were drawn at 95% confidence for caecum and day 13 faecal samples. Line plots show mean ± sem. In panels **b**, **c**, **g** and **h** statistical significance was determined by two-sided Wilcoxon rank-sums test against vehicle. *$P < 0.05$; **$P < 0.01$. Data in **b**, **e** and **j** are presented as boxplots where the middle lines are the median, lower and upper hinges represent the first and third quartiles, upper whiskers extend to maxima and lower whiskers extend to minima.



Extended Data Fig. 6 | Microbiome community shifts following 10 mg kg^{-1} MCBA dosing via PBFM. Timepoint-nested PERMANOVA reveals significant shifts by treatment, though significance is lost when tested within individual timepoints. $n = 5$ male, 5 female per treatment.

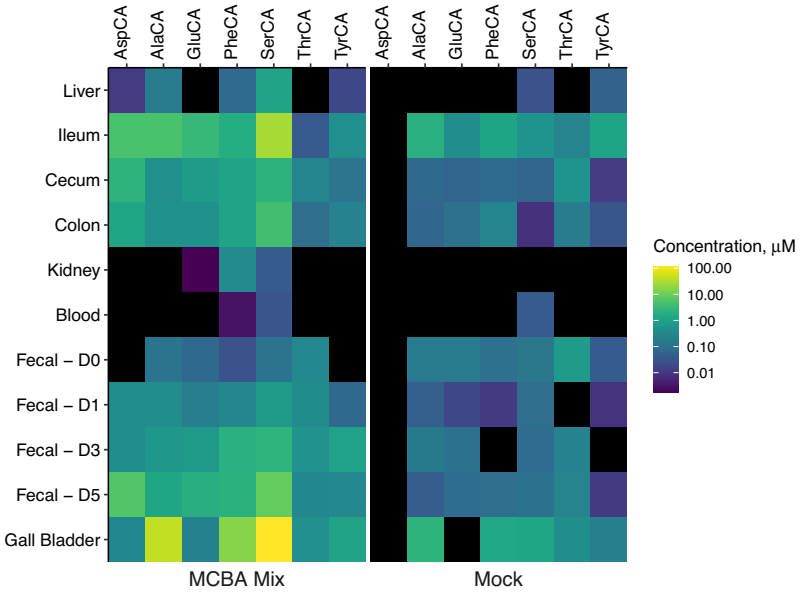
a



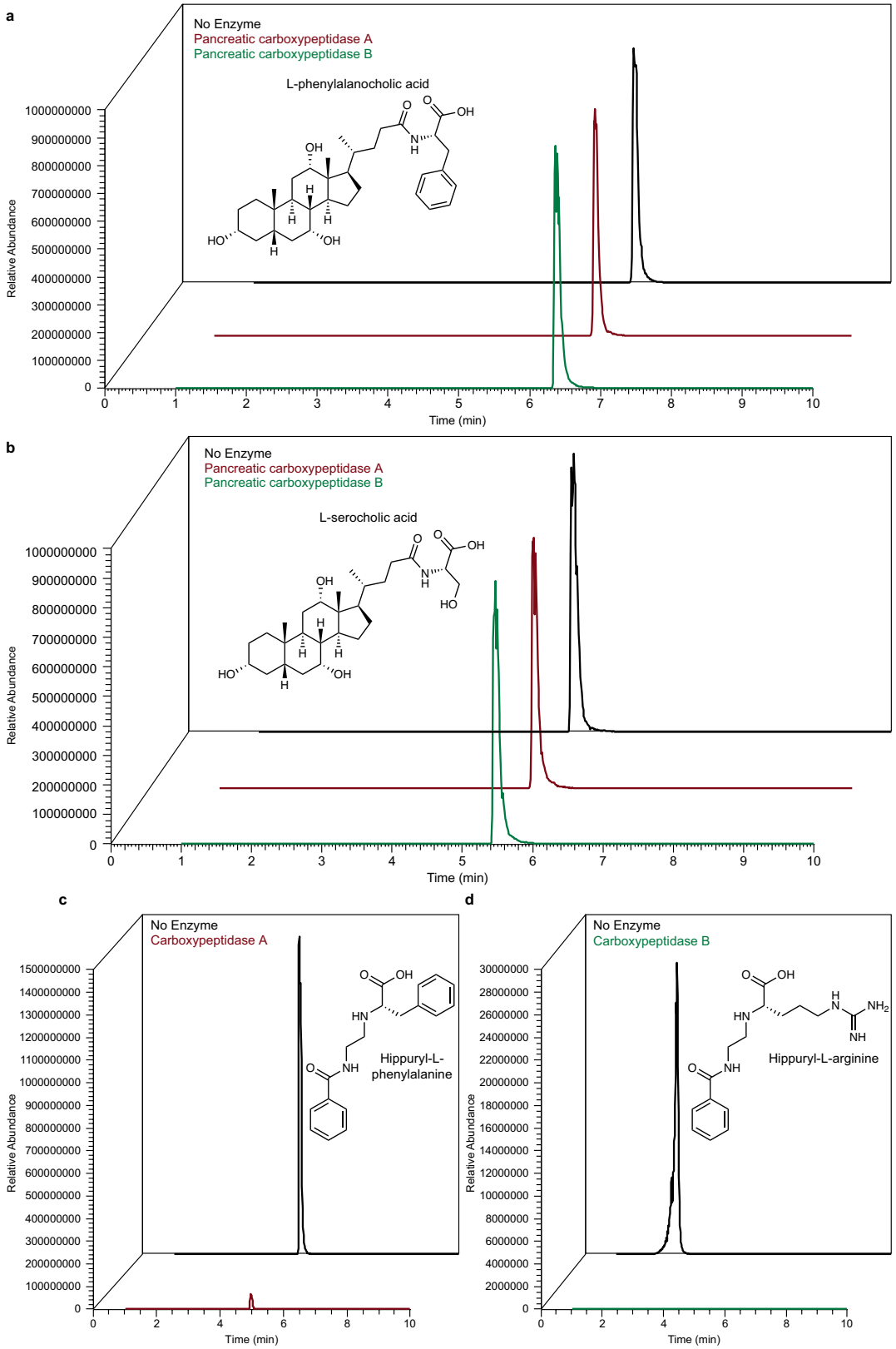
b

Sample	SerCA-treated	Mock
Blood - HPV	31.3 ± 24 µM	1.86 ± 1.86 µM
Gallbladder	247 ± 105 µM	0.328 ± 0.131 µM
Duodenum	506 ± 65.1 µM	0.103 ± 0.0628 µM
Ileum	166 ± 123 µM	0.0875 ± 0.069 µM
Colon	72 ± 38.2 µM	0 ± 0 µM
Cecum	146 ± 64.1 µM	0.0625 ± 0.053 µM
Feces - D0	0 ± 0 µM	0 ± 0 µM
Feces - D1	27.6 ± 16.3 µM	0 ± 0 µM
Feces - D5	56.5 ± 35.8 µM	0 ± 0 µM

Extended Data Fig. 7 | SerCA concentrations following 100 mg kg^{-1} feeding. **a**, SerCA concentrations in murine tissue and faecal samples following 100 mg kg^{-1} SerCA dosing via PBFM. Data are presented as boxplots where the middle lines are the median, lower and upper hinges represent the first and third quartiles, upper whiskers extend to maxima and lower whiskers extend to minima. **b**, Table showing SerCA concentration by sample type, presented as mean ± sem, $n = 4$ mice per group.



Extended Data Fig. 8 | MCBA concentrations in faecal and tissue samples following mixed MCBA dosing via PBFM. Data are presented as the average concentration of each MCBA included in the MCBA mix (80 mg kg^{-1} total, 10 mg kg^{-1} per individual MCBA). $n = 3$ treatment, 2 control.



Extended Data Fig. 9 | Extracted ion chromatograms of PheCA and SerCA exposed to pancreatic carboxypeptidases. When incubated with **a**, 1 mM PheCA or **b**, 1 mM SerCA, neither pancreatic carboxypeptidase A nor pancreatic carboxypeptidase B were able to deconjugate the supplemented MCBA while still showing near-complete elimination of native substrates **c**, hippuryl-L-phenylalanine and **d**, hippuryl-L-arginine for carboxypeptidase A and carboxypeptidase B, respectively. Reactions were performed in triplicate.

nature portfolio

Corresponding author(s): Robert A. Quinn

Last updated by author(s): Oct 27, 2023

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our Editorial Policies 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	Confirmed	
☐	☒	The exact sample size (*n*) 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	Thermo Scientific XCaliber (version 4.1.31.9), Axygen Gen5 (version 3.13).
Data analysis	R (version 4.2.2), RStudio (version 2023.03.0+386), and XCaliber (version 4.1.31.9) were used for data analysis. Notable R packages used were ggtree (version 3.6.2), growthcurver (version 0.3.1), and randomforest (version 4.7-1.1),

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Protein structures are available on the Protein Data Bank (at rscb.org). C. perfringens BSH/T in complex with DCA and taurine, from Rossocha et al., is available under PDB ID 2bjg (doi:10.2210/pdb2BJG/pdb). L. salivarius BSH/T in complex with TCA, from Karlov et al., is available under PDB ID 8blt (doi:10.2210/pdb8BLT/pdb).

Raw mass spectrometry data are publicly available within the MassIVE database (at massive.ucsd.edu) for the in vitro screen for MCBA production under MSV000090234 (doi:10.25345/C5S756Q1B), for CpBSH/T variant analysis under MSV000092138 (doi:10.25345/C55D8NQ9V), for MCBA gavage samples under MSV000093173 (doi:10.25345/C57S7J35N), for mixed MCBA PBFM dosing under MSV000093171 (doi:10.25345/C5H98ZQ3R), for 100 mg-1 kg SerCA PBFM dosing at MSV000093169 (doi:10.25345/C5RV0DB2C), 10 mg-1 kg MCBA PBFM dosing of under MSV000093172 (doi:10.25345/C5CJ87W9C), and for sleeve gastrectomy fecal samples under MSV000093167 (doi:10.25345/C51834C9N).

GNPS molecular networks are available for the MCBA production screen at gnps.ucsd.edu/ProteoSAFe/status.jsp?task=565151309a874d5f97caa3f383c95382, for CpBSH/T incubation with 1 mM BA and equimolar amino acid mix at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=3dec8f7ab26d47098406a7e597825154 and https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=33da5da024ed44848770a4a02b119d9e, for the CpBSH/T mutagenesis experiment at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=30c88ca297a44f84be5fa32b376e5cb9, and for the sleeve gastrectomy fecal samples at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=f11eaab1cf1d43b1a5f754575d171e87.

16S rRNA gene amplicon data were deposited in the Qiita database. Data from 100 mg kg-1 gavage can be found under study ID 14697 with analysis available under analysis ID 53128. Data from 10 mg kg-1 PBFM can be found under study IDs 15013 and 15021 with analysis available under analysis IDs 57407 and 57481.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Sex and gender were not considered in the study design as shift before and after sleeve gastrectomy were the main focus. Within this cohort, 4 patients were male and 40 patients were female.
Population characteristics	Average patient age was 48 and ranged from 18 to 66 years old. Average baseline BMI was 45, ranging from 36.4 to 68.4. Of the patients within this cohort, 27 were white, 16 were African American/black, 1 was Native/American Indian, and 1 was multiple races. All participants were non-Hispanic. 12 patients exhibited type 2 diabetes as co-morbidity, 24 exhibited obstructive sleep apnea as a co-morbidity, and 2 exhibited coronary artery disease as a co-morbidity, and 1 exhibited arterial fibrilation as a co-morbidity.
Recruitment	Participants were recruited from the Royal Oak Weight Control Center, an affiliate of the Corewell Health William Beaumont University Hospital, during the preoperative bariatric surgery process. This involves a medical work-up, surgical risk stratification, and multidisciplinary education prior to moving on to surgery. Information about the study was presented at the free informational bariatric surgery seminar that prospective patients attend before starting the bariatric surgery program. Fliers about the trial were posted and distributed at the Beaumont Weight Control Center and mailed to patients with the bariatric surgery approval letter. The approval letter is written by a Weight Control Center physician and indicates that a patient is approved from a medical and multidisciplinary team perspective to move forward with bariatric surgery. Inclusion criteria followed the National Institutes of Health criteria for bariatric surgery: body mass index (BMI) at or above 40 kg/m^2, or a BMI of 35 to 40 kg/m^2 with an obesity co-morbidity such as type 2 diabetes, heart disease, or obstructive sleep apnea. An additional inclusion requirement of being between 18 and 70 years old. Patients were excluded if they had poorly controlled medical or psychiatric conditions which, in the opinion of the investigator, made the patient unlikely to be able to properly participate in the study. Biases include self-selection bias and that only patients planning the sleeve gastrectomy bariatric surgical procedure were included. It is unlikely that these biases had a significant impact on the results of this study.
Ethics oversight	Beaumont Institutional Review Board, Michigan State University Institutional Review Board. All patients provided informed consent prior to participating in the study.

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

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	No statistical methods were used to predetermine sample sizes for mouse experiments.
Data exclusions	Fecal and cecal data from male mice involved in 100 mg/kg MCBA gavage were excluded from analysis due to less noticeable microbiome shifts being observed.
Replication	Multiple, individual experiments allowed for the confirmation that Clostridium perfringens BSH/T can produce MCBAs. Three biological replicates for bacterial culturing and screening for MCBA production allowed for verification of species-specific production of MCBAs. Three replicates for each commercial CpBSH/T experiment were sufficient in calculating deconjugation kinetic constants and observing linearity when investigating constants involved in phenylalanine conjugation to TCA or CA.

For the 100 mg/kg gavage study, 5 mice tested in each treatment group allowed for verification of the shifts observed in fecal and cecal microbiomes. For the 100 mg/kg SerCA PBFM study, 4 mice tested in each treatment group allowed for verification of the systemic movement of SerCA to various tissues throughout the body in addition to validating the novel bile acid dosing approach.

For the 10 mg/kg/MCBA mixed MCBA PBFM study, 3 mice in the treatment group and two in the control were sufficient for this pilot study validating SerCA presence in murine tissues in addition to supporting the selection of PheCA as the second MCBA to use for the 10 mg/kg individual MCBA study including more robust controls. 10 mice (5 female, 5 male) per treatment group were sufficient to reasonably quantify BA concentrations within various sample types, with this robust replication to confirm that findings were separate from amino acid+CA controls.

Randomization

Mice were randomized into treatment groups upon arrival from the vendor or reaching the appropriate age following on-site breading. Mice involved in peanut butter pellet experiments were randomized upon successfully completing PBFM training.

Patients within the bariatric surgery study were not randomized into treatment groups as temporal sampling was the primary measure to identify changes prior to surgery and three months following surgery.

Blinding

Blinding for mouse experiments was not relevant due to a single variable being tested (MCBA gavaged/fed).

Blinding for the bariatric surgery analysis was not relevant due to temporal changes before and after surgery were the primary measure. Patient samples were paired based on sampling time.

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 and archaeology
☐	☒ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	6-week-old C57BL/6J mice
Wild animals	This study did not involve wild animals.
Reporting on sex	Findings apply to both male and female mice, however female response to MCBA gavage appear more pronounced than male response in the gavage study. All other reported results include both sexes.
Field-collected samples	This study did not involve samples collected from the field.
Ethics oversight	The Institutional Animal Care and Use Committee (IACUC) at Michigan State University provided guidance and oversight.

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

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*

Outcomes

Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.