Diversity and function of microbial lipases within the mammalian gut

Thomas C. A. Hitch[1*], Johannes M. Masson[1], Theresa Streidl[1], Thomas Fischöder, Lothar Elling and Thomas Clavel[1*]

[1]Functional Microbiome Research Group, Institute of Medical Microbiology, University Hospital of RWTH Aachen, Aachen, Germany

[2]Institute for Biotechnology and Helmhotlz-Institute for Biomedical Engineering, RWTH Aachen University, Aachen, Germany

*To whom correspondence should be addressed:
thitch@ukaachen.de (bioinformatics); tclavel@ukaachen.de (all other aspects of the work)

Abstract

Background: Current estimates suggest the majority of microbial proteins within the mammalian gut lack meaningful annotation. One such functional group are microbial lipases (EC:3.1.1.3), which can alter host access and utilisation of dietary fat. In this paper, we describe the diversity of lipolytic bacteria, including *in vitro* characterisation of a new lipase.

Results: Metagenomic sequence-based network analysis identified that the majority of microbial lipases in the gut of three host species (human, mouse, pig) belong to two unique clusters. These clusters were characterized by the presence of two novel motifs, AHSKGG and TTxxTPH, which may play a key functional role due to co-localisation in the active site, as identified by structural modelling. Analysis of metagenomic assembled genomes (MAGs) indicated that the majority of lipase-positive species belong to the phylum Firmicutes, although all dominant phyla within the human gut were represented by positive species. Metabolic analysis of these genomes identified a high prevalence of glycerol rather than fatty acid metabolism. The occurrence of microbial lipases determined across ~800 metagenomic gut samples depended on dietary fat consumption, with lipase expression increased in lard fed compared to palm oil fed mice. A representative lipase encoded within the genome of the species *Clostridium symbiosum* was cloned and its characterization confirmed the *in silico* prediction and provided detailed annotation to 373 proteins.

Conclusions: Microbial lipases within the gut represent a conserved group characterized by unique amino acid sequence motifs. While an increase in microbial lipase occurrence was positively associated with dietary fat intake, lipase-producing species seemed unable to metabolise the released fatty acids. In this paper, we provide a global analysis of the functional importance and diversity of microbial lipases within the intestine of mammals, which will improve the resolution of future sequence-based studies and open avenues for mechanistic experiments based on isolates.

Keywords: host metabolism, high-fat diet, triglycerides, gut microbiome, microbial lipases

Introduction

The gut microbiota is a complex ecosystem that consists of hundreds of microbial species collectively producing millions of proteins [1–4]. Mechanistic understanding of their role in health and disease is made difficult due to this complexity, leading to reductionist approaches such as the use of minimal microbial consortia [5,6]. Even when the complexity of the microbiota is reduced, mechanistic understanding is hampered by the lack of meaningful annotations for a majority of microbial functions within the gut. The most recent estimate for the human gut microbiome suggested that 40% of proteins lack any annotation and an additional 20% are assigned to 'Unknown Function' [7]. One approach for improving functional annotations is the targeted analysis of specific functions, as previously conducted for β-glucuronidases within both the human and murine gut [8,9] and also for bile acid metabolism [10].

The prevailing knowledge about fat absorption is that is occurs within the small intestine, specifically in the proximal region [11,12], according to studies conducted in the 50's [11,13,14], 60's [15] and 70's [12,16–18]. However, in light of findings which suggest fat absorption differs greatly depending on fat amount [19] and type provided [20], as well as the potential role of the microbiome [21], these assumptions should be re-evaluated and confirmed to provide comprehensive overview of mammalian fat absorption. To date, the role of microbial lipases on fat degradation and host metabolism has only been investigated a handful of times. Oral administration of a transgenic lipase-producing bacterium was shown to significantly increase fat absorption in pigs and improved steatorrhea [22]. Carp, *Cyprinus carpio*, fed with LipG1, a microbial lipase isolated from the gut of carp, significantly increased lipolytic activity and weight gain when fed palm oil, a cheap and industrially advantageous feed [23]. Despite these studies highlighting both the health and economic benefits to better understand microbial lipases and fat degradation, microbial lipases within the gut are currently understudied compared to host-derived lipases [24–26]. Triglyceride lipases (EC:3.1.1.3) catalyse the release of free fatty acids from triglyceride molecules. While functionally diverse, lipases and esterases share structural and sequence similarity as they are members of the protein family of alpha/beta hydrolases [27]. Microbial lipases and esterases have previously been grouped into 8 families based on sequence similarity, of which only family 1 represents 'true lipases' [28]. The 'true' lipases were termed so due to their maximal activity towards water-insoluble long-chain triglycerides and grouped into 11 sub-families [29]. The high level of sequence similarity between lipases is partially due to multiple conserved structural features including the alpha-beta fold and motifs such as the GxSxG motif at the active site [30].

Within this paper, we comprehensively studied the lipases present within the human, mouse and pig gut using a combination of metagenomics, microbial cultivation, and biochemical methods. This includes functional annotation of both known and novel lipases and studying their prevalence across feeding trials and the taxonomic diversity of lipase-producing species. We also characterised a member of a previously unknown family of gut bacterial lipases.

Methods

Databases

EGGNOG: EGGNOG consists of non-supervised orthologous gene clusters formed via Smith-Waterman reciprocal best matches [31] and includes the original COGs [32], which have since been expanded. COG1075 is assigned the function 'triacylglycerol lipase' and contains 1,260 proteins from 633 species. The proteins originating from bacteria (67.5%) were used to form the EGGNOG database.

EC: Triglyceride lipases are assigned the enzyme commission (EC) notation 3.1.1.3, specifying them as first hydrolases, secondly acting on ester bonds, thirdly carboxylic ester hydrolases, and finally as triglyceride lipases. The UNIPROT database was searched for genes assigned this EC notation, which returned 162 proteins that formed the EC database.

Lipase engineering database (LED): This database is a custom database of *alpha/beta*-hydrolase proteins encompassing six super families made up of 22 homologous families created from 5,278 proteins. Since its creation in 2000, the lipase engineering database has been used multiple times to study the sequence and structural similarity of hydrolases [30].

Lipase annotation

Proteins from the pig [3], mouse [4] and human [2] metagenomic gene catalogues were annotated against the EC, EGGNOG and LED databases using DIAMOND BLASTP (v0.9.12.114) [33]. As previous research suggests protein families are best captured using 30-40% identity and 35-70% coverage [34], annotations were required to cover ≥40% of both the query and subject sequence at a minimum sequence identity of 40%. Moreover, proteins needed to have a match within two of the three databases and to contain the GxGxG motif identified to be common in lipases [30].

Metagenomic and meta-transcriptomic sequence alignment

Fecal metagenomic data from 144 and 46 mice fed standard chow or a high-fat diet, respectively, [35], 57 healthy human individuals [36], and 157 pigs [3] were studied to provide average relative abundances of lipases within each host. In addition, metagenomes from 565 human individuals with varying body mass indices (BMIs) were studied to identify if individuals with varying degrees of obesity have altered lipase occurrence. Individuals were stratified into four groups based on their BMI: lean to overweight, < 30; obese class 1, 30 – 34.9; obese class 2, 35 – 39.9; and obese class 3, > 40 [37]. To identify the effect of dietary fat on the occurrence of lipase genes, faecal metagenomes from 10 individuals following a ketogenic diet (high fat, high protein, low carbohydrate) for 14 days were studied, including samples at day 0, 3, 7 and 14 [38]. Metatranscriptomic data from mice fed a high-fat diet enriched with primary bile acids and based either on lard (n = 5) or palm (n = 7) were also analysed to assess the effect of fat type on the expression of microbial lipases [39]. PALADIN (v1.4.1) [40] was used to align each meta-genomic/transcriptomic dataset obtained from the literature to the gene catalogues using the default settings and outputting to samtools view (v0.1.19) for conversion into BAM format [41]. For each dataset, the relevant host gene catalogue formed the reference for alignment.

Network analysis and multiple sequence alignment

Identity between the gut lipases and the type proteins for each known lipolytic protein family [29] was identified using DIAMOND with the same cutoffs as for lipase annotation. The protein matches were entered into Cytoscape [42] along with their shared identity for visualisation. The motifs specific for member sequences of both 'Cluster 1' and 'Cluster 2' (see results section) were identified via cluster-specific alignment with ClustalW [43] within the MEGAX software package [44] and visualised using weblogo [45].

Protein modelling

The structure of protein sequences was modelled using the I-TASSER suite [46]. COFACTOR [47] was utilised with TM-ALIGN [48] to compare the predicted structure of the proteins to those within the Protein Data Bank (PDB) [49]. Models were viewed in Pymol (v1.8).

Biochemical characterisation

The lipase sequence (WP_003509004.1) from *Clostridium symbiosum* DSM 29356 [50] was modified *in silico* by removal of the transmembrane region (first 150 N-terminal amino acids) and addition of a polyhistidine-tag (His_6) at the *C*-terminus. The new construct (LipΔ150), integrated into a pET-15b vector was purchased from BioCat (Heidelberg, Germany) and transformed into *Escherichia coli* BL21 (DE3). The transformed *E. coli* was cultured in the presence of ampicillin and a low IPTG concentration (0.1 mM) at low temperatures (25 °C) to support a functional folding of the expressed LipΔ150. Cells were harvested after 22 h by centrifugation and frozen at -20°C.

An aliquot of the frozen cells was solved in lysis buffer (50 mM Tris/HCl; 300 mM NaCl; 10 mM imidazole; pH 8) to a concentration of 150 mg/ml. The cells were disrupted by sonication cooled on ice (amplitude 25%; impulse length 3 x 30 s). Cell debris were removed by centrifuging the lysate (16,000 g, 30 min, 4°C). The supernatant was filtered (0.45 µm) and the target protein purified via immobilized metal-ion affinity chromatography (IMAC) using a 5 mL HisTrap™ HP column (GE Healthcare, UK). The lysis buffer (50 mM Tris/HCl; 300 mM NaCl; 10 mM imidazole; pH 8) was used for equilibration of the column and as washing buffer. Proteins were recovered using the elution buffer (50 mM Tris/HCl; 300 mM NaCl; 300 mM imidazole; pH 8) and collected in 0.5 ml fractions. The protein content during the washing

and elution steps was monitored using the Äkta purifier 100 system (GE Healthcare, UK) and double-checked with a Bradford assay. Protein containing fractions from the elution step were pooled and used for further purification.

Imidazole was removed using a PD-10 Desalting Column (GE Healthcare, UK) with an imidazole-free equilibration buffer (50 mM Tris/HCl; 300 mM NaCl; pH 8). Samples of approximately 200 µl were collected, a Bradford assay along with the *p*NP assay (C8) (see below) were performed to identify the lipase fractions, which were then pooled. The final protein concentration was determined by a BSA-Bradford assay (R^2 0.99). The purity of the isolated protein was verified with a native polyacrylamide gel electrophoresis (PAGE).

Lipolytic activity, referred to only when activity was confirmed, was tested utilizing the degradation of para-nitrophenyl (*p*NP) bound fatty acid substrates [51–53]. The test was carried out by mixing 83.3 µl of a testing buffer with 16.7 µl protein solution and the absorption measured at 410 nm over a period of 45 min at 37°C. The standard testing buffer (final concentrations in the 100 µl reaction mix: 50 mM Tris/HCl; 20 mM NaCl; 4.46 mM Tween20; 0.3 mM *p*NP-octanoate (C8); pH 8) was used during purification and for comparison in every test. A negative control was prepared for every tested condition by adding 16.7 µl of lipase free equilibration buffer (50 mM Tris/HCl; 300 mM NaCl; pH 8) to 83.3 µl of the corresponding testing buffer (*i.e.* containing *p*NP substrates of varying chain length). Protein precipitation was observed when testing the effect of both PMSF (10mM) and Orlistat (10mM), preventing overtime measurement. Instead, after 45 min incubation at 37°C, samples (including all controls) were centrifuged (4,000 g, 10 min) and the supernatant measured. For quantification, a standard curve was created using the *p*NP assay with defined concentrations of para-nitrophenol (R^2 = 0.99). Results were normalized by subtraction of the negative control values and the standard curve was used to determine the concentration of released fatty acids. Statistical significance was determined using a t-test with Bonferroni multiple test correction adjusted p-values reported.

Bacterial cultivation and lipolytic activity assay

Hungate tubes with 9 ml of BHI media was used to cultivate *Escherichia coli* Mt1B1 (negative control) and *Clostridium symbiosum* DSM 29356. After 24 hours, 50 µl of each culture was placed onto Rhodamine B agar plates [54]. As positive control, 50 µl of porcine pancreas lipase (Sigma-Aldrich, USA) with a concentration of 20 mg/ml was added and the plates were incubated for 2 h at 37°C. Fluorescence, caused by lipolytic activity was observed under UV light using the GelStudio SA (Analytic Jena, Germany). To identify if the lipolytic activity was cell-bound, each isolate was cultured in Hungate tubes with 9 ml BHI media supplemented with 0.5% cysteine (0.05 g/ml), 0.5% DTT (0.02 g/ml), and 200 µl olive oil. After 7 days of cultivation, 6 ml of dense culture was centrifuged (10,000 g, 10 min) to collect biomass. The pellet and 50 µl of the supernatant were put onto Rhodamine B agar plates [54]. As positive control, 50 µl of porcine pancreas lipase (Sigma-Aldrich, USA) with a concentration of 20 mg/ml was added and the plates were incubated for 2 h at 37°C. Fluorescence, caused by lipolytic activity was observed under UV light using the GelStudio SA (Analytic Jena, Germany).

Genome analysis of catabolism pathways

Metagenomically assembled genomes (MAGs) were obtained from Pasolli *et al.* (2019) [55], along with metadata for their originating sample, assignment to species genome bins (SGBs), prevalence and relative abundance [55]. Taxonomy was assigned to each MAG using GTDB-TK, release 89 [56,57]. MAGs were annotated using PROKKA (v1.13.3) [58] and KEGG orthologs (KOs) [59] identified using the PROKKA2KEGG tool (https://github.com/SilentGene/Bio-py/tree/master/prokka2kegg). Each MAGs KOs were searched against previously published lipolytic catabolism modules [60].

Results

Specific microbial lipase repertoires within the gut

The millions of proteins from each gene catalogue (human, mouse, pig) were annotated against three reference databases, each representing a different method for the identification of lipases. False positive identification was prevented by secondary filtering based on proteins being required to have matches within two of the three databases. Tertiary filtering based on the requirement of proteins to contain the common serine hydrolase motif GxSxG further reduced the number of total matches by >45% (Table **1**).

In our recent work [29], we grouped all so far known lipases into 35 families based on both their sequence similarity and biochemical characteristics. The identified gut lipases were placed into this system based on sequence identity using an all-against-all annotation approach (Figure **1a**). While some gut lipases shared similarity with Family 1.1, 1.2, 1.3,1.4 and one protein shared distant similarity with Family 7, the majority grouped together to form two loosely connected clusters, labelled as Cluster 1 and 2.

Further sequence analysis of the proteins belonging to either cluster 1 or 2 allowed to identify specific amino acid motifs. Within cluster 1, the motifs AHSKGG and TTxxTPH both occurred in all 98 member proteins (Figure **1b**). Both motifs were also highly conserved within cluster 2, occurring in ≥90% of the 194 member proteins (Figure **1c**). Although the GxSxG motif was used during identification of these lipases, the conserved version was identified to contain isoleucine and histidine as the variable second and forth positions, respectively. The type proteins of each lipase family/subfamily were scanned for the presence of both of the novel motifs to identify if they are unique markers of gut lipases. AHSKGG was identified as unique to cluster 1 and 2 while the TTxxTPH motif was identified within the type proteins for family 1.2 and 1.5.

To better understand the variability within this collection of gut lipases, the length of each protein was studied in relationship to their host of origin (Figure **1d**). The mean length was identified to be 410 amino acids, although sizes ranged from 158 to 706. Interestingly, 40.5% of pig lipases were less than 300 amino acids long, however, out of the ten longest lipases, four originated from the pig gut, and the other six from humans.

Taxonomic and functional insights into lipase-positive species

With the release of large datasets of metagenomic assembled genomes (MAGs) from the human microbiome, many of which represent uncultured bacteria, we are now able to gain functional insights into previously unstudied species [55,61,62]. Proteins from 154,723 MAGs previously reconstructed using metagenomes from different human body sites [55] were annotated using the lipase annotation method utilized for the gene catalogues. In total, 13,677 lipases were identified belonging to 11,011 genomes (7.1%). The majority of these proteins were assigned to Cluster 2 (n = 7,604). However, many were also identified as sharing highest similarity with the lipolytic enzyme family 1 sub-families (n = 1,487). Of those assigned to Cluster 2, 98.3% contained the three motifs (GxSxG, TTxxTPH and AHSKGG) along with 98.5% of the 653 proteins assigned to Cluster 1.

To study the taxonomy of lipase-positive species further, the MAGs were grouped based on ANI values, reducing the 154,723 MAGs into 4,930 species genome bins (SGB). For each SGB, a single representative MAG was selected, of which 235 (4.8%) were identified to contain a lipase (Figure **2a**). Taxonomically, these lipase-positive species belong to 37 bacterial families across the phyla Firmicutes, Actinobacteria, Proteobacteria, Fusobacteria, Bacteroidetes, Spirochaetes and Cyanobacteria. Of the most prevalent SGBs (based on number of MAGs assigned to the bin), only 8 of the top 23 could be assigned to a named species, whilst the rest represented metagenomically inferred species of novel lineages (Figure **2b**). The median relative abundance of lipase-positive species was nearly twice that of the lipase-negative species (1.06% Vs 0.68%), suggesting that they are dominant members of the human microbiome (Figure **2c**). Although dominant, almost 50% of the SGBs were identified to represent unknown species or genera according to GTDB-TK (Figure **2d**), showing that the majority of lipase-positive species lack a cultured representative. The majority of the SGB representatives originated from gut samples (n = 176, 74.9%) and represented the major phyla from the gut (Firmicutes, Bacteroides, Actinobacteria, Proteobacteria and Fusobacteria). The most commonly reconstructed lipase-positive species was *Ruminococcus_D bicirculans*, a dominant commensal of the human gut [63] which contains a lipase belonging to Cluster 2 (Figure **2b**).

The second largest body site from which lipase-positive species were identified was the skin (45 SGB, 19.1%), of which the majority (29 SGB, 64.4%) belonged to the phylum Actinobacteria (Figure **2a**). Additionally, nine Firmicutes species belonging to the genus *Staphylococcus* were identified. *Staphylococcus* species have previously been shown to include lipase producers from the skin microbiota [64]. Both *Staphylococcus epidermidis* and *Cutibacterium acnes*, common skin colonisers and some of the most prevalent reconstructed species within the utilised dataset (Figure **2b**), were identified to be lipase-positive, although their lipases were both classified as belonging to Family 1.

The presence of lipases within a high proportion of gut colonizing-bacteria suggests an evolutionary and metabolic benefit to triglyceride degradation, such as for energy generation or biomass production [65]. Hence, we further examined metabolic potential of the 235 lipase-positive SGBs (Figure **2e**). Lipolytic activity releases two products: free fatty acids that can undergo beta-oxidation, and glycerol which can be metabolized to enter glycolysis. Three models of lipid catabolism were studied [60]. Whilst aerobic and anaerobic models of beta-oxidation were examined [66], no isolates contained the anaerobic beta-oxidation pathway and 5.5% of lipase-positive SGBs contained the complete aerobic beta-oxidation pathway. Conversely, 41.7% of SGBs contained the glycerol degradation II pathway. Additionally, when genomes containing incomplete pathways were included, a further 28.1% of SGBs were identified to utilize glycerol.

Taxonomic and functional shifts of the gut microbiota have been linked to diets rich in fat [35,39]. The increased availability of fat in the luminal environment may select for species expressing lipases. Hence, metagenomic data from 144 mice fed a chow diet and 40 fed a high-fat diet was studied [35]. A significant increase in lipase relative abundance was observed in the high-fat mice (Figure **2f**). Such a diet-driven increase in mice was confirmed to occur within humans by studying the longitudinal effect of a ketogenic diet on 10 individuals using published metagenomic data [38]. After 14 days of high fat intake, significantly higher, individual-specific counts of lipase genes were detected in faeces (Figure **2g**). Previous research has shown that germfree mice are resistant to diet-induced obesity when the fat used is lard, but not if palm oil is used [20]. Therefore, we investigated the effect of both lard and palm oil on lipase expression within the murine gut of SPF mice and found significantly higher lipase expression in the gut microbiome of lard-fed mice (Figure **2h**).

Characterization of a novel microbial lipase

The lack of experimental validation for microbial functions has been highlighted as one of the major limitations and bottlenecks within meta-omic studies [7]. Hence, we utilised cultured isolates and biochemical testing to provide *in vitro* evidence for their functional assignment. Genomes from our isolates within the mouse intestinal bacterial collection (www.dsmz.de/miBC) [50] were searched against the predicted lipases to identify cultured host species. A complete sequence match (100% identity and coverage) was identified between a member of cluster 2 and a protein within *Clostridium symbiosum* DSM 29356. The isolate was cultured in a lipid-rich medium for 24h and tested for lipolysis on Rhodamine B agar. The fluorescence detected from the *Clostridium symbiosum* cell solution, compared to the *Escherichia coli* negative control and porcine lipase-positive control, confirm its lipolytic activity (Figure **3a**). To determine if this enzymatic activity was cell bound or excreted, the isolate was cultured for 1 week, after which the cell pellet and supernatant were tested separately, with activity being identified only in the cell pellet fraction (Figure **3b**). Sequence analysis confirmed that the protein contains the GxSxG, AHSKGG, and TTxxTPH motifs, as well as five transmembrane regions at its N terminus, suggesting it is a membrane-bound lipase (Figure **3c**). The predicted model of the lipase confirmed a low normalised B-factor value for the transmembrane region, indicating that this region is unstable and may represent a transmembrane region (Figure **3d**).

Structural modelling predicted that removal of the transmembrane region at the 150 amino acid would leave a functional and stable enzymatic region (Figure **3e-g**). Comparison of the predicted structure of the truncated lipase, LipΔ150, against the PDB database using COFACTOR confirmed that the five top matches, based on TM-score, were lipases. Functional annotation of the structure assigned the EC as 3.1.1.3 and the gene ontology term 'triglyceride lipase activity (GO:0004806)'. Ligand binding was also predicted, based on a TM-SCORE of 0.81, of 5LipA bound to octyl-phosphinic acid 1,2-bis-octylcarbamoyloxy-ethyl ester (OCP), a structural homolog for triglycerides which binds irreversibly to the catalytic serine, facilitating crystallography (Figure **3e**). The AHSKGG, TTxxTPH and GxSxG motifs were all identified to co-localise with OCP within the lipase structure (Figure **3f**). Co-localisation of the motifs and OCP was studied further by space-filled modelling the motifs and OCP molecule (Figure **3g**). The conserved nature of these motifs, along with their co-localisation with OCP suggests that they are not only hallmarks of gut lipase clusters but are of functional importance.

The LipΔ150 sequence was cloned with a poly-histidine tag into a pET-15b expression vector in *Escherichia coli*. Once expressed and purified, the enzyme was biochemically characterized to test its predicted function. LipΔ150 was confirmed to be a lipase, as activity was detected against all lengths of fatty acids tested, ranging from C2 to C18 (Figure **3h**). However, a preference was shown for medium chain fatty acids under the conditions tested, as the highest concentration of released fatty acids was observed with C8 (133 ± 3 μM), and the least was with C2 (57 ± 5 μM).

Due to the use of the GxSxG motif in the prediction of these enzymes, PMSF was used to identify if LipΔ150 is a serine-hydrolase. Additionally, Orlistat, a lipase inhibitor commonly used for weight loss, was tested to verify if it has the potential to also inhibit microbial lipases within the GI tract. Both PMSF (Figure **3i**) and Orlistat (Figure **3j**) were observed to inhibit 100% of the lipase activity and led to precipitation of the protein. Lipolytic activity was also observed to be temperature-dependant, with maximum activity occurring at 45°C (Figure **3k**). As LipΔ150 represents a novel lipase based on its low sequence similarity to existing families, we propose it as the first member of the new lipolytic enzyme family XXXVI [29].

Within the gastrointestinal tract, bile acids are known to act as emulsifier to increase host lipid absorption within the small intestine. However, *in vitro* and *in vivo* experimentations have shown that bile acids can decrease lipase activity due to reducing the pH [69,70]. Therefore, the effect of varying concentrations of emulsification both above and below the critical micelle concentration (CMC) of Tween20 (0.035 mM at 37°C) was examined (Figure **3l**). Lipolytic activity was significantly increased when the concentration of emulsifier was at or below the CMC. This is likely due to being optimized to work at the physiological concentration of total bile acids in murine GI tract, which was reported to be 0.011 ± 0.001 mM [71] whilst CMC for individual bile acids are >1 mM [72,73]. The possibility of pH altering lipolytic activity was also investigated, with pH of 6 or below preventing activity and maximum activity being observed at pH 9 (Figure **3m**).

Discussion

The lack of detailed functional assignment is commonly cited as a limitation of metagenomic and genomic studies, limiting our insight into microbial processes. Therefore, we have generated the first collection of microbial lipases present within the mammalian GI tract which are freely available at github.com/thh32/GutLipases. The conserved, yet unique sequences of these lipases, as compared to those isolated from other environments, was highlighted by the identification of two motifs which are both nearly ubiquitous across all identified microbial lipases within the gene catalogues and MAG collection. We hypothesis that these motifs, along with the conserved GxSxG motif, are key to the activity of the microbial lipases within the gut due to their co-localisation both together and with bound molecules.

The taxonomic range of lipase-positive species was investigated via *in-silico* analysis of MAGs, which identified that lipases are present within species belonging to each of the major phyla and families identified within mammalian body-associated microbiomes. These results are contra to previous findings suggesting that few species, mostly from the phylum Proteobacteria, are lipolytic [60]. This discrepancy may be explained by the targeted annotation applied within this manuscript compared to the general annotations used to identify KEGG orthologs in previous studies [74].

Many of the most common lipase-positive SGBs were assigned to the genera *Ruminococcus* and *Coprococcus*, both of which are dominant members of the mammalian gut microbiome due to their diverse functional roles [75,76]. Interestingly, the mean relative abundance of lipase-positive SGB was twice that of the lipase negative species, suggesting that these species are dominant and potentially key members of the ecosystem. Due to the lack of cultured representatives for many of the lipase-positive species, in vivo experimentation is not possible, however with the renaissance of cultivation [50,77–79], additional species may soon be isolated facilitating further analysis of the metabolism and ecological importance of lipase producers within the mammalian gut.

Due to the lack of cultured representatives, the metabolism of the SGBs themselves was investigated based on three major lipolytic modules. The lack of aerobic or anaerobic beta-oxidation from the majority of lipase-positive SGB suggests that the metabolism of fatty acids is not the purpose of their lipolytic activity. In contrast, the majority of lipase-positive species were predicted to utilise glycerol, indicating the same species degrading triglycerides may be able to further utilise the glycerol unit [80]. Contra to this, almost 30% of lipase-positive species had no identifiable pathway for the utilisation of the lipolytic products. Instead of metabolising either the fatty acids or glycerol, it is possible that these species utilise diacylglycerol molecules in the generation of membrane phospholipids during replication and cell remodelling [81].

Direct measurement of lipolytic activity along the GI tract can be difficult due to the diverse nature of intestinal content and sample variation. However, biochemical characterisation of the new gut microbial lipase has provided some hints into their localisation and niche. Activity was observed to decrease with increasing detergent concentration, particularly above the CMC concentration. In addition to this, lipolytic activity was completely remove below pH 7, suggesting that microbial lipases within the gut have evolved to work in an environment with sub-optimal detergent concentrations and at a pH at or above 7. Along the GI tract, the bile concentration decreases whilst the pH and microbial load increase, suggesting that the microbial lipases have evolved to be active in the distal gut. For validation of this theory as well as to enhance our understanding of microbial lipase localisation along the GI tract, *in vivo* experiments will be required.

Conclusions

While many microbial functions remain unknown, this study provides detailed insights into microbial encoded lipases within host body habitats by combining metagenomics and cultivation. The observed diet-dependent occurrence of lipolytic gut bacteria, the potential role of lipases in bacterial metabolism, and the characterization of a new microbial lipase from a cultured bacterium open new avenues to investigate the functional input of microbiomes in triglyceride degradation and fat metabolism.

Ethics approval and consent to participate

Not applicable

Consent for publication

All co-authors have approved this manuscript for publication.

Competing interests

No competing interests

Funding

TCAH was funded by an internal University Hospital RWTH Aachen START grant titled ‘KETOGUT’ as well as a RWTH Start-Up grant titled ‘ProtoBiome’. Access to the computational resources of the RWTH cluster were provided under grant; rwth2608. TC received funding from the German Research Foundation (DFG) via SPP1656 and Project-ID 403224013 – SFB 1382, Gut-liver axis.

Authors contributions

TCAH and TC conceived the study. TCAH conducted all bioinformatic analyses and supervised daily laboratory work. JM, TS, and TF conducted wet lab experiments. TCAH, JM, TS, and TC interpreted data. LE and TC provided guidance and access to essential materials. TCAH and TC secured funding and wrote the paper. All authors reviewed the manuscript and agreed with its final content.

Figure and Table legends

Table 1: Annotation of mammalian gene catalogues for the identification of lipases. The table indicate the total number of proteins within each hosts gene catalogue, in million, followed by the number identified as being lipases throughout the different steps of the workflow.

Figure 1: Protein sequence analysis of microbial lipases. a. Network analysis of sequence identities shared between the identified microbial lipases. Each node represents a single lipase coloured based on its host of origin. Edges length and width is proportional to the sequence identity shared between the two nodes. Cluster 2 is highlighted with a red box. The conserved motifs identified within both cluster 1 (**b**) and cluster 2 (**c**) along with their respective positions in the alignment and coloured based on amino acid features. **d**. Protein length variation across the lipases, coloured based on their host of origin.

Figure 2: Metagenomic occurrence of lipases, taxonomic diversity of their host species, and capacity for lipolytic catabolism. a. Phylogenomic tree of lipase-positive SGBs generated using phylophlan [67] and visualised with iTOL [68]. External rings depict important features of SGBs as follows (inner to outer rings): those SGBs which contained a lipase belonging to either family I, Cluster 1 or Cluster 2; metadata associated to the sample the SGB representative was reconstructed from: body site, age of the host, and if from a Western country or not. **b**. The taxonomy and size (number of MAGs) of all lipase-positive SGBs with a size > 100. **c**. Mean relative abundance of both lipase-positive and lipase-negative SGBs in the human gut. **d**. Novelty of lipase-positive SGBs according to their GTDB taxonomic assignment. **e.** The occurrence of each module for lipolytic catabolism across the SGBs is stated in blue and next to this in orange is the number with the incomplete pathway (>50% of the total KOs in the pathway). **f-h**. Relative abundances of the identified lipases across metagenomic data from mice fed a chow vs. high-fat diet (HFD) (**f**), human subjects on a ketogenic diet (**g**) and metatranscriptomic data from the gut of mice fed a HFD based on either palm oil or lard (**h**).

Figure 3: Structural modelling and biochemical characterisation of LipΔ150. a. Rhodamine B agar assay for *Clostridium symbiosum* DSM 29356, *Escherichia coli* Mt1B1 (negative control) and porcine lipase viewed under both UV (60 ms) and white light (400 ms). Lipase activity is indicated by fluorescence, seen as white in the figures. **b**. Rhodamine B agar assay for the cell pellet and supernatant fractions of *C. symbiosum*, cultured for 7 days, along with the porcine lipase-positive control, viewed under both UV and white light. **c**. Prediction of trans-membrane regions via TMHMM. **d**. Structural model of the complete lipase protein coloured according to the B-normalised value from green (low) to blue (high). **e-g**. Structural modelling of LipΔ150 bound to the substrate OCP (green) on the surface (**e**), using a ribbon model (**f**) and highlighting only the conserved motifs (**g**). In both **f** and **g**, the conserved motifs are coloured red for GISHG, orange for AHSKGG, and yellow for TTINTPH. **h-m**. Biochemical characterisation of LipΔ150 was conducted using a range of fatty acid lengths (**h**), inhibition via PMSF (**i**) and orlistat (**j**), temperature-dependant activity (**k**), emulsifier concentration (**l**), and pH (**m**). Statistical significance (t-test with Bonferroni multiple test correction) is represented as follows: * < 0.05, ** < 0.01, *** < 0.005.

References

[1] Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., et al., 2010. A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464(7285): 59–65, Doi: 10.1038/nature08821.

[2] Li, J., Jia, H., Cai, X., Zhong, H., Feng, Q., Sunagawa, S., et al., 2014. An integrated catalog of reference genes in the human gut microbiome. Nature Biotechnology 32(8): 834–41, Doi: 10.1038/nbt.2942.

[3] Xiao, L., Estellé, J., Kiilerich, P., Ramayo-Caldas, Y., Xia, Z., Feng, Q., et al., 2016. A reference gene catalogue of the pig gut microbiome. Nature Microbiology 1(October 2017), Doi: 10.1038/nmicrobiol.2016.161.

[4] Lesker, T.R., Durairaj, A.C., Gálvez, E.J.C., Lagkouvardos, I., Baines, J.F., Clavel, T., et al., 2020. An Integrated Metagenome Catalog Reveals New Insights into the Murine Gut Microbiome. Cell Reports 30(9): 2909-2922.e6, Doi: 10.1016/j.celrep.2020.02.036.

[5] Brand, M.W., Wannemuehler, M.J., Phillips, G.J., Proctor, A., Overstreet, A.M., Jergens, A.E., et al., 2015. The altered schaedler flora: Continued applications of a defined murine microbial community. ILAR Journal 56(2): 169–78, Doi: 10.1093/ilar/ilv012.

[6] Brugiroux, S., Beutler, M., Pfann, C., Garzetti, D., Ruscheweyh, H.J., Ring, D., et al., 2016. Genome-guided design of a defined mouse microbiota that confers colonization resistance against Salmonella enterica serovar Typhimurium. Nature Microbiology 2(November): 1–12, Doi: 10.1038/nmicrobiol.2016.215.

[7] Thomas, A.M., Segata, N., 2019. Multiple levels of the unknown in microbiome research. BMC Biology: 17–20.

[8] Pollet, R.M., D'Agostino, E.H., Walton, W.G., Xu, Y., Little, M.S., Biernat, K.A., et al., 2017. An Atlas of β-Glucuronidases in the

Human Intestinal Microbiome. Structure 25(7): 967-977.e5, Doi: 10.1016/j.str.2017.05.003.

[9] Creekmore, B.C., Gray, J.H., Walton, W.G., Biernat, K.A., Little, M.S., Xu, Y., et al., 2019. Mouse Gut Microbiome-Encoded β-Glucuronidases Identified Using Metagenome Analysis Guided by Protein Structure. MSystems 4(4): 1–13, Doi: 10.1128/mSystems.00452-19.

[10] Funabashi, M., Grove, T.L., Wang, M., Varma, Y., McFadden, M.E., Brown, L.C., et al., 2020. A metabolic pathway for bile acid dehydroxylation by the gut microbiome. Nature 582(7813): 566–70, Doi: 10.1038/s41586-020-2396-4.

[11] BORGSTROM, B., DAHLQVIST, A., LUNDH, G., SJOVALL, J., 1957. Studies of intestinal digestion and absorption in the human. The Journal of Clinical Investigation 36(10): 1521–36, Doi: 10.1172/JCI103549.

[12] Clark, S.B., Lawergren, B., Martin, J. V., 1973. Regional intestinal absorptive capacities for triolein: an alternative to markers. American Journal of Physiology 225(3): 574–85, Doi: 10.1152/ajplegacy.1973.225.3.574.

[13] BORGSTRÖM, B., 1952. On the Mechanism of the Intestinal Fat Absorption. Acta Physiologica Scandinavica, Doi: 10.1111/j.1748-1716.1954.tb01158.x.

[14] Borgström, B., 1957. Determination of pancreatic lipase in human small intestinal content. Scandinavian Journal of Clinical and Laboratory Investigation 9(3): 226–8, Doi: 10.3109/00365515709079961.

[15] Aberdeen, V., Shepherd, P.A., Simmonds, W.J., 1960. CONCURRENT MEASUREMENT, IN UNANAISTHETIZED RATS, OF INTESTINAL TRANSPORT AND FAT ABSORPTION FROM THE LUMEN. Q J Exp Physiol Cogn Med Sci.

[16] Ammon, H. V., Phillips, S.F., 1973. Inhibition of colonic water and electrolyte absorption by fatty acids in man. Gastroenterology 65(5): 744–9, Doi: 10.1016/s0016-5085(19)33009-4.

[17] Snipes, R.L., 1977. Morphological appearance of fat in the epithelial cells of different portions of the intestines in mice. Experientia 33(10): 1335–6, Doi: 10.1007/BF01920166.

[18] Wu, A.L., Clark, S.B., Holt, P.R., 1975. Transmucosal triglyceride transport rates in proximal and distal rat intestine in vivo. Journal of Lipid Research 16(4): 251–7.

[19] Booth, C.C., Alldis, D., Read, A.E., 1961. Studies on the site of fat absorption. Gut 2(1).

[20] Kübeck, R., Bonet-Ripoll, C., Hoffmann, C., Walker, A., Müller, V.M., Schüppel, V.L., et al., 2016. Dietary fat and gut microbiota interactions determine diet-induced obesity in mice. Molecular Metabolism 5(12): 1162–74, Doi: 10.1016/j.molmet.2016.10.001.

[21] Rabot, S., Membrez, M., Bruneau, A., Gérard, P., Harach, T., Moser, M., et al., 2010. Germ-free C57BL/6J mice are resistant to high-fat-diet-induced insulin resistance and have altered cholesterol metabolism. The FASEB Journal 24(12): 4948–59, Doi: 10.1096/fj.10-164921.

[22] Drouault, S., Juste, C., Marteau, P., Renault, P., Corthier, G., 2002. Oral treatment with Lactococcus lactis expressing Staphylococcus hyicus lipase enhances lipid digestion in pigs with induced pancreatic insufficiency. Applied and Environmental Microbiology 68(6): 3166–8, Doi: 10.1128/AEM.68.6.3166-3168.2002.

[23] Ran, C., He, S., Yang, Y., Huang, L., Zhou, Z., 2015. A novel lipase as aquafeed additive for warm-water aquaculture. PLoS ONE 10(7): 1–15, Doi: 10.1371/journal.pone.0132049.

[24] Huang, Y.S., Lin, X., Redden, P.R., Horrobin, D.F., 1995. In vitro hydrolysis of natural and synthetic γ-linolenic acid-containing triacylglycerols by pancreatic lipase. Journal of the American Oil Chemists' Society 72(6): 625–31, Doi: 10.1007/BF02635646.

[25] Lowe, M.E., 1992. The catalytic site residues and interfacial binding of human pancreatic lipase. The Journal of Biological Chemistry 267(Bio 101): 17069–73.

[26] Huggins, K.W., Camarota, L.M., Howles, P.N., Hui, D.Y., 2003. Pancreatic Triglyceride Lipase Deficiency Minimally Affects Dietary Fat Absorption but Dramatically Decreases Dietary Cholesterol Absorption in Mice. Journal of Biological Chemistry 278(44): 42899–905, Doi: 10.1074/jbc.M303422200.

[27] Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S., et al., 1992. The α/β hydrolase fold. Protein Engineering, Design and Selection 5(3): 197–211, Doi: 10.1093/protein/5.3.197.

[28] Arpigny, J.L., Jaeger, K.-E., 1999. Bacterial lipolytic enzymes: classification and properties. Biochemical Journal 343(1): 177, Doi: 10.1042/0264-6021:3430177.

[29] Hitch, T.C.A., Clavel, T., 2019. A proposed update for the classification and description of bacterial lipolytic enzymes. PeerJ 7(e7249), Doi: 10.7717/peerj.7249.

[30] Pleiss, J., Fischer, M., Peiker, M., Thiele, C., Schmid, R.D., 2000. Lipase engineering database: Understanding and exploiting sequence-structure-function relationships. Journal of Molecular Catalysis - B Enzymatic 10(5): 491–508, Doi: 10.1016/S1381-

1177(00)00092-8.

[31] Jensen, L.J., Julien, P., Kuhn, M., von Mering, C., Muller, J., Doerks, T., et al., 2008. eggNOG: Automated construction and annotation of orthologous groups of genes. Nucleic Acids Research 36(SUPPL. 1): D250-4, Doi: 10.1093/nar/gkm796.

[32] Tatusov, R.L., 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Research 28(1): 33–6, Doi: 10.1093/nar/28.1.33.

[33] Buchfink, B., Xie, C., Huson, D.H., 2014. Fast and sensitive protein alignment using DIAMOND. Nature Methods 12(1): 59–60, Doi: 10.1038/nmeth.3176.

[34] Veeramachaneni, V., Makałowski, W., 2004. Visualizing sequence similarity of protein families. Genome Research 14(6): 1160–9, Doi: 10.1101/gr.2079204.

[35] Xiao, L., Feng, Q., Liang, S., Sonne, S.B., Xia, Z., Qiu, X., et al., 2015. A catalog of the mouse gut metagenome. Nature Biotechnology 33(10): 1103–8, Doi: 10.1038/nbt.3353.

[36] Feng, Q., Liang, S., Jia, H., Stadlmayr, A., Tang, L., Lan, Z., et al., 2015. Gut microbiome development along the colorectal adenoma-carcinoma sequence. Nature Communications 6, Doi: 10.1038/ncomms7528.

[37] De Lorenzo, A., Soldati, L., Sarlo, F., Calvani, M., Di Lorenzo, N., Di Renzo, L., 2016. New obesity classification criteria as a tool for bariatric surgery indication. World Journal of Gastroenterology 22(2): 681–703, Doi: 10.3748/wjg.v22.i2.681.

[38] Mardinoglu, A., Wu, H., Bjornson, E., Zhang, C., Hakkarainen, A., Räsänen, S.M., et al., 2018. An Integrated Understanding of the Rapid Metabolic Benefits of a Carbohydrate-Restricted Diet on Hepatic Steatosis in Humans. Cell Metabolism 27(3): 559-571.e5, Doi: 10.1016/j.cmet.2018.01.005.

[39] Just, S., Mondot, S., Ecker, J., Wegner, K., Rath, E., Gau, L., et al., 2018. The gut microbiota drives the impact of bile acids and fat source in diet on mouse metabolism. Microbiome 6(1): 1–18, Doi: 10.1186/s40168-018-0510-8.

[40] Westbrook, A., Ramsdell, J., Schuelke, T., Normington, L., Bergeron, R.D., Thomas, W.K., et al., 2017. PALADIN: Protein alignment for functional profiling whole metagenome shotgun data. Bioinformatics 33(10): 1473–8, Doi: 10.1093/bioinformatics/btx021.

[41] Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25(16): 2078–9, Doi: 10.1093/bioinformatics/btp352.

[42] Shannon, P., Markiel, A., Owen Ozier, 2., Baliga, N.S., Wang, J.T., Ramage, D., et al., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Research 13(13): 2498–504, Doi: 10.1101/gr.1239303.metabolite.

[43] Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22(22): 4673–80, Doi: 10.1093/nar/22.22.4673.

[44] Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution 35(6): 1547–9, Doi: 10.1093/molbev/msy096.

[45] Crooks, G., Hon, G., Chandonia, J., Brenner, S., 2004. WebLogo: a sequence logo generator. Genome Research 14: 1188–90, Doi: 10.1101/gr.849004.1.

[46] Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang, Y., 2014. The I-TASSER suite: Protein structure and function prediction. Nature Methods 12(1): 7–8, Doi: 10.1038/nmeth.3213.

[47] Roy, A., Yang, J., Zhang, Y., 2012. COFACTOR: An accurate comparative algorithm for structure-based protein function annotation. Nucleic Acids Research 40(W1): 471–7, Doi: 10.1093/nar/gks372.

[48] Zhang, Y., Skolnick, J., 2005. TM-align: A protein structure alignment algorithm based on the TM-score. Nucleic Acids Research 33(7): 2302–9, Doi: 10.1093/nar/gki524.

[49] Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., et al., 2000. The Protein Data Bank. Nucleic Acids Research 28(1): 235–42, Doi: 10.1093/nar/28.1.235.

[50] Lagkouvardos, I., Pukall, R., Abt, B., Foesel, B.U., Meier-Kolthoff, J.P., Kumar, N., et al., 2016. The Mouse Intestinal Bacterial Collection (miBC) provides host-specific insight into cultured diversity and functional potential of the gut microbiota. Nature Microbiology 1(10): 16131, Doi: 10.1038/nmicrobiol.2016.131.

[51] Peng, Y., Fu, S., Liu, H., Lucia, L.A., 2016. Accurately Determining Esterase Activity via the Isosbestic Point of p-Nitrophenol. BioResources 11(4), Doi: 10.15376/biores.11.4.10099-10111.

[52] Zhang, Y., Sun, W., Elfeky, N.M., Wang, Y., Zhao, D., Zhou, H., et al., 2020. Self-assembly of lipase hybrid nanoflowers with bifunctional Ca2+ for improved activity and stability. Enzyme and Microbial Technology 132(2): 109408, Doi: 10.1016/j.enzmictec.2019.109408.

[53] de Oliveira, B.H., Coradi, G.V., de Oliva-Neto, P., do Nascimento, V.M.G., 2020. Biocatalytic benefits of immobilized Fusarium sp. (GFC) lipase from solid state fermentation on free lipase from submerged fermentation. Industrial Crops and Products 147(November 2019): 112235, Doi: 10.1016/j.indcrop.2020.112235.

[54] Kouker, G., Jaeger, K.E., 1987. Specific and sensitive plate assay for bacterial lipases. Applied and Environmental Microbiology 53(1): 211–3, Doi: 10.1128/aem.53.1.211-213.1987.

[55] Pasolli, E., Asnicar, F., Manara, S., Zolfo, M., Karcher, N., Armanini, F., et al., 2019. Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle. Cell 176(3): 649-662.e20, Doi: 10.1016/j.cell.2019.01.001.

[56] Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P., et al., 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. BioRxiv (August): 256800, Doi: 10.1101/256800.

[57] Chaumeil, P.-A., Mussig, A.J., Hugenholtz, P., Parks, D.H., 2019. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 36(November 2019): 1925–7, Doi: 10.1093/bioinformatics/btz848.

[58] Seemann, T., 2014. Prokka: Rapid prokaryotic genome annotation. Bioinformatics 30(14): 2068–9, Doi: 10.1093/bioinformatics/btu153.

[59] Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., Kanehisa, M., 1999. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Research 27(1): 29–34, Doi: 10.1093/nar/27.1.29.

[60] Vieira-Silva, S., Falony, G., Darzi, Y., Lima-Mendez, G., Garcia Yunta, R., Okuda, S., et al., 2016. Species-function relationships shape ecological properties of the human gut microbiome. Nature Microbiology 1(8), Doi: 10.1038/nmicrobiol.2016.88.

[61] Parks, D.H., Rinke, C., Chuvochina, M., Chaumeil, P.-A., Woodcroft, B.J., Evans, P.N., et al., 2017. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. Nature Microbiology 903: 1–10, Doi: 10.1038/s41564-017-0012-7.

[62] Ormerod, K.L., Wood, D.L.A., Lachner, N., Gellatly, S.L., Daly, J.N., Parsons, J.D., et al., 2016. Genomic characterization of the uncultured Bacteroidales family S24-7 inhabiting the guts of homeothermic animals. Microbiome 4(1): 36, Doi: 10.1186/s40168-016-0181-2.

[63] Wegmann, U., Louis, P., Goesmann, A., Henrissat, B., Duncan, S.H., Flint, H.J., 2014. Complete genome of a new Firmicutes species belonging to the dominant human colonic microbiota ('Ruminococcus bicirculans') reveals two chromosomes and a selective capacity to utilize plant glucans. Environmental Microbiology 16(9): 2879–90, Doi: 10.1111/1462-2920.12217.

[64] Kwaszewska, A., Sobiś-Glinkowska, M., Szewczyk, E.M., 2014. Cohabitation—relationships of corynebacteria and staphylococci on human skin. Folia Microbiologica 59(6): 495–502, Doi: 10.1007/s12223-014-0326-2.

[65] Yao, J., Rock, C.O., 2015. How bacterial pathogens eat host lipids: Implications for the development of fatty acid synthesis therapeutics. Journal of Biological Chemistry 290(10): 5940–6, Doi: 10.1074/jbc.R114.636241.

[66] Campbell, J.W., Morgan-Kiss, R.M., Cronan, J.E., 2003. A new Escherichia coli metabolic competency: Growth on fatty acids by a novel anaerobic b-oxidation pathway. Molecular Microbiology 47(3): 793–805, Doi: 10.1046/j.1365-2958.2003.03341.x.

[67] Segata, N., Börnigen, D., Morgan, X.C., Huttenhower, C., 2013. PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. Nature Communications 4: 1–10, Doi: 10.1038/ncomms3304.

[68] Letunic, I., Bork, P., 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Research 47(W1): W256–9, Doi: 10.1093/nar/gkz239.

[69] Li, Y., Hu, M., McClements, D.J., 2011. Factors affecting lipase digestibility of emulsified lipids using an in vitro digestion model: Proposal for a standardised pH-stat method. Food Chemistry 126(2): 498–505, Doi: 10.1016/j.foodchem.2010.11.027.

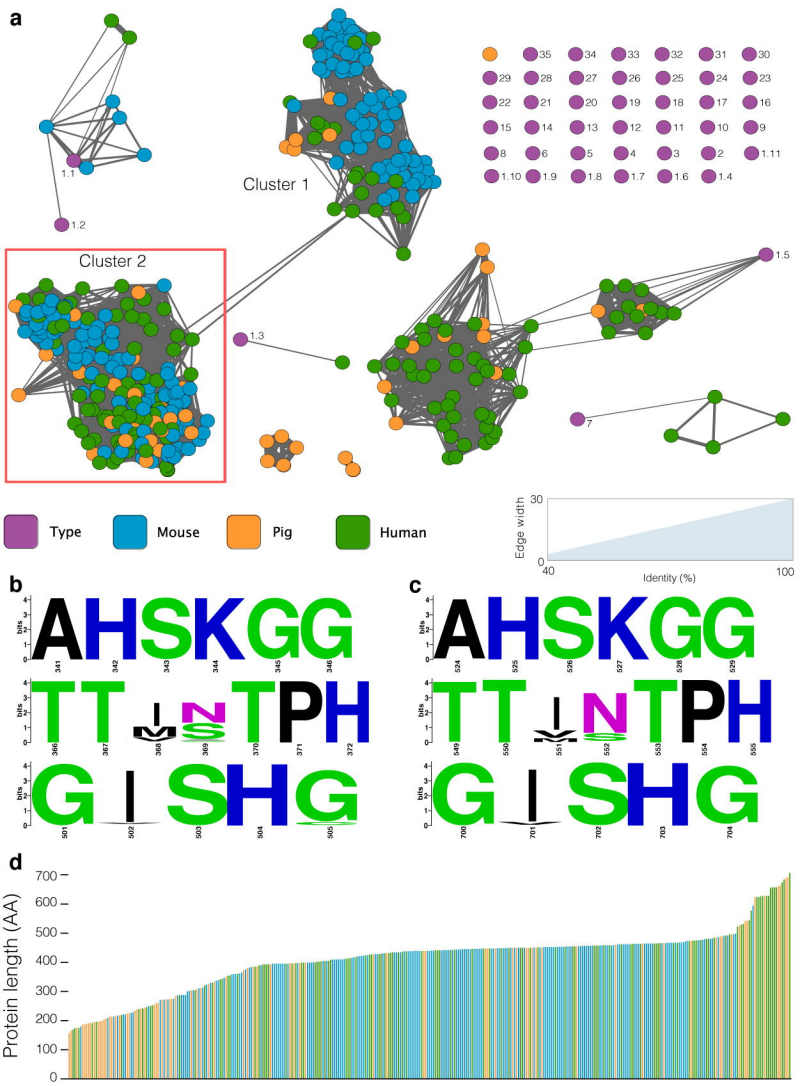
[70] Knarreborg, A., Jensen, S.K., Engberg, R.M., 2003. Pancreatic lipase activity as influenced by unconjugated bile acids and pH, measured in vitro and in vivo. Journal of Nutritional Biochemistry 14(5): 259–65, Doi: 10.1016/S0955-2863(03)00008-1.

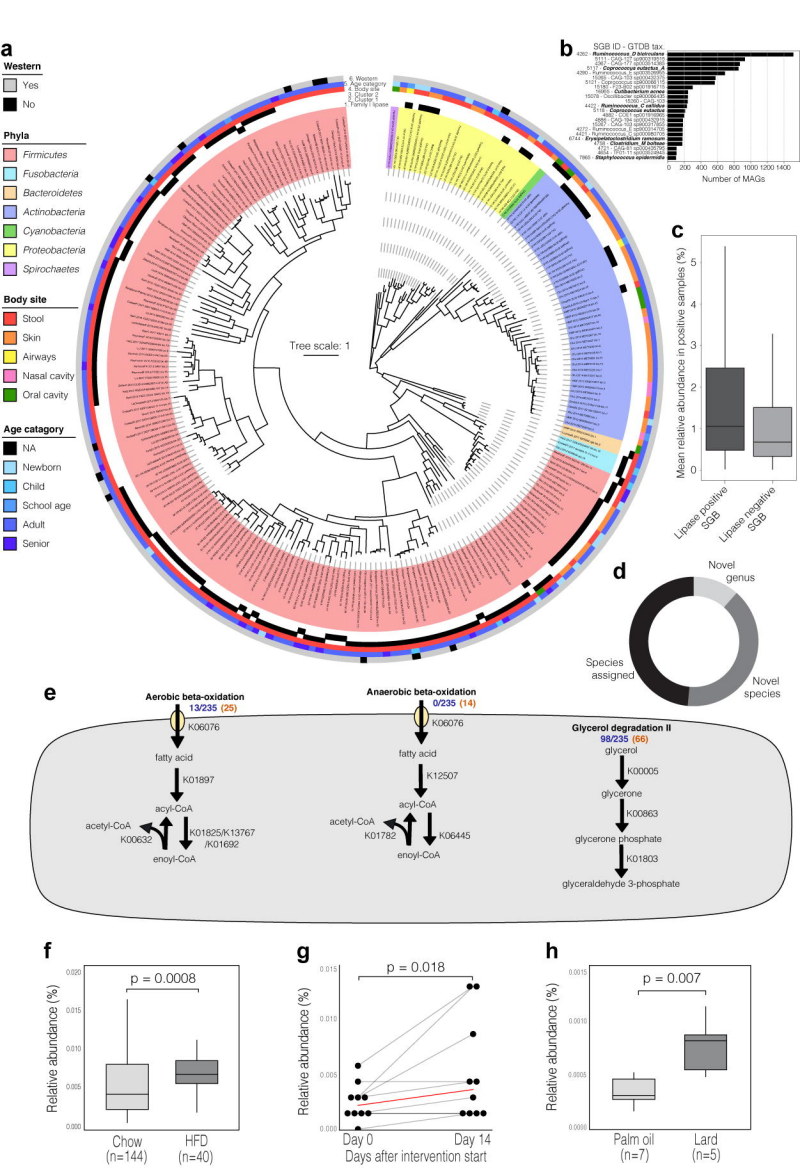
[71] Kuribayashi, H., Miyata, M., Yamakawa, H., Yoshinari, K., Yamazoe, Y., 2012. Enterobacteria-mediated deconjugation of taurocholic acid enhances ileal farnesoid X receptor signaling. European Journal of Pharmacology 697(1–3): 132–8, Doi: 10.1016/j.ejphar.2012.09.048.

[72] Roda, A., Hofmann, A.F., Mysels, K.J., 1983. The influence of bile salt structure on self-association in aqueous solutions. Journal

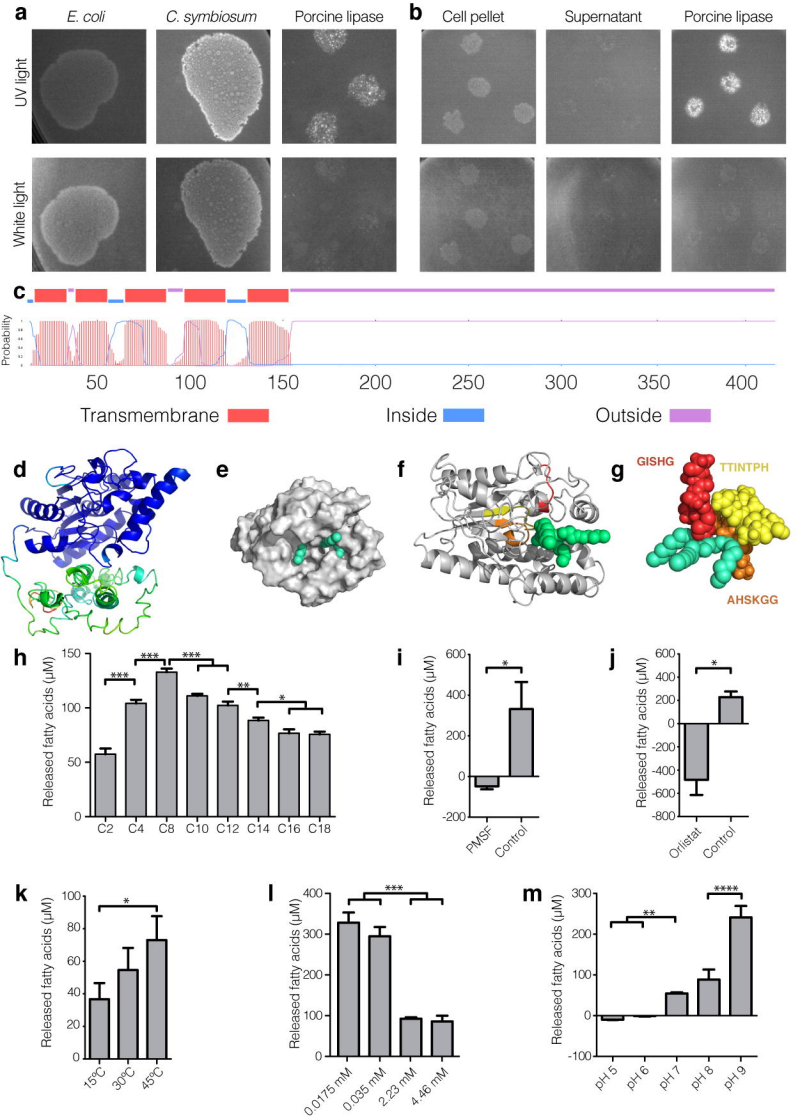
530 of Biological Chemistry 258(10): 6362–70.
531 [73] Simonović, B.R., Momirović, M., 1997. Determination of Critical Micelle Concentration of Bile Acid Salts by Micro-calorimetric
532 Titration. Mikrochimica Acta 127(1): 101–4, Doi: 10.1007/bf01243172.
533 [74] Markowitz, V.M., Szeto, E., Palaniappan, K., Grechkin, Y., Chu, K., Chen, I.M.A., et al., 2008. The integrated microbial genomes
534 (IMG) system in 2007: Data content and analysis tool extensions. Nucleic Acids Research 36(SUPPL. 1): 528–33, Doi:
535 10.1093/nar/gkm846.
536 [75] Berg Miller, M.E., Antonopoulos, D.A., Rincon, M.T., Band, M., Bari, A., Akraiko, T., et al., 2009. Diversity and strain specificity of
537 plant cell wall degrading enzymes revealed by the draft genome of Ruminococcus flavefaciens FD-1. PLoS ONE 4(8), Doi:
538 10.1371/journal.pone.0006650.
539 [76] Noriega, B.S., Sanchez-Gonzalez, M.A., Salyakina, D., Coffman, J., 2016. Understanding the Impact of Omega-3 Rich Diet on
540 the Gut Microbiota. Case Reports in Medicine 2016, Doi: 10.1155/2016/3089303.
541 [77] Poyet, M., Groussin, M., Gibbons, S.M., Avila-Pacheco, J., Jiang, X., Kearney, S.M., et al., 2019. A library of human gut bacterial
542 isolates paired with longitudinal multiomics data enables mechanistic microbiome research. Nature Medicine 25(9): 1442–52,
543 Doi: 10.1038/s41591-019-0559-3.
544 [78] Forster, S.C., Kumar, N., Anonye, B.O., Almeida, A., Viciani, E., Stares, M.D., et al., 2019. A human gut bacterial genome and
545 culture collection for improved metagenomic analyses. Nature Biotechnology 37(2): 186–92, Doi: 10.1038/s41587-018-0009-7.
546 [79] Seshadri, R., Leahy, S.C., Attwood, G.T., Teh, K.H., Lambie, S.C., Cookson, A.L., et al., 2018. Cultivation and sequencing of
547 rumen microbiome members from the Hungate1000 Collection. Nature Biotechnology 36(4): 359–67, Doi: 10.1038/nbt.4110.
548 [80] Xiao, L., Sonne, S.B., Feng, Q., Chen, N., Xia, Z., Li, X., et al., 2017. High-fat feeding rather than obesity drives taxonomical and
549 functional changes in the gut microbiota in mice. Microbiome 5(1): 1–12, Doi: 10.1186/S40168-017-0258-6.
550 [81] Sohlenkamp, C., Geiger, O., 2015. Bacterial membrane lipids: Diversity in structures and pathways. FEMS Microbiology Reviews
551 40(1): 133–59, Doi: 10.1093/femsre/fuv008.
552

Database	**Human (9.9 million)** [2]	**Mouse (4.5 million)** [4]	**Pig (7.7 million)** [3]
LED	10,988	5,658	6,490
EC	14	5	20
EGGNOG	239	274	192
Annotation in >2 databases	175	250	158
GxSxG motif	116	178	79


a
Cluster 1
Cluster 2
1.1
1.2
1.3
1.5
7
35
34
33
32
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
6
5
4
3
2
1.11
1.10
1.9
1.8
1.7
1.6
1.4
Type
Mouse
Pig
Human
Edge width
30
0
40
100
Identity (%)
b
c
d
Protein length (AA)
700
600
500
400
300
200
100
0



a
Western
Yes
No
Phyla
Firmicutes
Fusobacteria
Bacteroidetes
Actinobacteria
Cyanobacteria
Proteobacteria
Spirochaetes
Body site
Stool
Skin
Airways
Nasal cavity
Oral cavity
Age catagory
NA
Newborn
Child
School age
Adult
Senior
6. Western
5. Age catagory
4. Body site
3. Cluster 2
2. Cluster 1
1. Family lipase
Tree scale: 1
b
SGB ID - GTDB tax
Number of MAGs
c
Mean relative abundance in positive samples (%)
Lipase positive SGB
Lipase negative SGB
d
Novel genus
Species assigned
Novel species
e
Aerobic beta-oxidation
13/235 (25)
K06076
fatty acid
K01897
acyl-CoA
acetyl-CoA
K00632
K01825/K13767 /K01692
enoyl-CoA
Anaerobic beta-oxidation
0/235 (14)
K06076
fatty acid
K12507
acyl-CoA
acetyl-CoA
K01782
K06445
enoyl-CoA
Glycerol degradation II
98/235 (66)
glycerol
K00005
glycerone
K00863
glycerone phosphate
K01803
glyceraldehyde 3-phosphate
f
p = 0.0008
Relative abundance (%)
Chow (n=144)
HFD (n=40)
g
p = 0.018
Relative abundance (%)
Day 0
Day 14
Days after intervention start
h
p = 0.007
Relative abundance (%)
Palm oil (n=7)
Lard (n=5)


a
E. coli
C. symbiosum
Porcine lipase
UV light
White light
b
Cell pellet
Supernatant
Porcine lipase
c
Probability
50
100
150
200
250
300
350
400
Transmembrane
Inside
Outside
d
e
f
g
GISHG
TTINTPH
AHSKGG
h
Released fatty acids (µM)
C2
C4
C8
C10
C12
C14
C16
C18
i
Released fatty acids (µM)
PMSF
Control
j
Released fatty acids (µM)
Orlistat
Control
k
Released fatty acids (µM)
15°C
30°C
45°C
l
Released fatty acids (µM)
0.0175 mM
0.035 mM
2.23 mM
4.46 mM
m
Released fatty acids (µM)
pH 5
pH 6
pH 7
pH 8
pH 9