Genome-guided design of a defined mouse microbiota that confers colonization resistance against Salmonella enterica serovar Typhimurium.

Genome-guided design of a defined mouse microbiota that confers colonization resistance against *Salmonella enterica* serovar Typhimurium

Sandrine Brugiroux[1], Markus Beutler[1§], Carina Pfann[3§], Debora Garzetti[1,2§], Hans-Joachim Ruscheweyh[4], Diana Ring[1], Manuel Diehl[1], Simone Herp[1], Yvonne Lötscher[5], Saib Hussain[1], Boyke Bunk[6], Rüdiger Pukall[6], Daniel H. Huson[4], Philipp C. Münch[1,7], Alice C. McHardy[7,8], Kathy D. McCoy[9], Andrew J. Macpherson[9], Alexander Loy[3], Thomas Clavel[10], David Berry[3] and Bärbel Stecher[1,2*]

[1] Max von Pettenkofer Institute of Hygiene and Medical Microbiology, Ludwig-Maximilians-University of Munich, Munich, Germany

[2] German Center for Infection Research (DZIF); Partner Site Munich

[3] Division of Microbial Ecology, Department of Microbiology and Ecosystem Science, Research Network Chemistry meets Microbiology, University of Vienna, A-1090 Vienna, Austria

[4] Center for Bioinformatics, University of Tübingen, Tübingen, Germany.

[5] Institute of Microbiology, ETH Zürich, Zürich, Switzerland.

[6] DSMZ - German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany

[7] Computational Biology of Infection Research, Helmholtz Centre for Infection Research, Braunschweig, Germany.

[8] Department of Algorithmic Bioinformatics, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[9] Maurice Müller Laboratories, Department of Clinical Research (DKF), UVCM, University Hospital, Bern, Switzerland

[10] ZIEL Institute for Food and Health, Technische Universität München, Freising, Germany

[§] These authors contributed equally

[*] For correspondence: Stecher@mvp.uni-muenchen.de

Abstract

Protection against enteric infections, also termed colonization resistance (CR), results from mutualistic interactions of the host and its indigenous microbes. The gut microbiota of humans and mice is highly diverse and it is therefore challenging to assign specific properties to its individual members. Here, we used a collection of murine bacterial strains and a modular design approach to create a minimal bacterial community that, once established in germ-free mice, provided CR against the human enteric pathogen *Salmonella enterica* serovar Typhimurium (*S.* Tm). Initially, a community of twelve strains, termed Oligo-Mouse Microbiota (Oligo-MM^{12}) representing members of the major bacterial phyla in the murine gut was selected. This community was stable over consecutive mouse generations and provided CR against *S.* Tm infection, albeit not to the degree of a conventional complex microbiota. Comparative (meta)genome analyses identified functions represented in a conventional microbiome but absent from the Oligo-MM^{12}. By genome-informed design, we created an improved version of the Oligo-MM community harboring three facultative anaerobic bacteria from the Mouse Intestinal Bacterial Collection (miBC) that provided conventional-like CR. In conclusion, we established a highly versatile experimental system that showed efficacy in an enteric infection model. Thus, in combination with exhaustive bacterial strain collections and systems-based approaches, genome-guided design can be used to generate insights into microbe-microbe and microbe-host interactions for the investigation of ecological and disease-relevant mechanisms in the intestine.

Introduction

The mammalian gastrointestinal tract represents a complex ecosystem that offers niches for hundreds of microbial species. This endogenous microbial community – the gut microbiota - provides essential functions to its host, including colonization resistance (CR), which is the ability of the microbiota to preclude infection by enteric pathogens such as *Salmonella* spp. [1]. Disturbance of the microbiota due to antibiotic usage can result in a transiently increased susceptibility to infections by a broad variety of pathogens [2]. Due to the complexity of interactions between the microbiota, the environment, and the host, the underlying basis of CR is still insufficiently understood.

Molecular approaches allow detailed investigation of the composition and metabolic diversity of the gut microbiota and its association with human diseases [3]. However, in order to dissect redundant metabolic functions or parallel signaling pathways within the gut microbiota, it is key to experimentally manipulate the community and its individual members, e.g. by using gnotobiotic animal models. Pioneering studies in gnotobiotic mice range from highly reductionist models involving one- or two-member communities up to consortia of intermediate and high complexity [4].

Humanized mouse models, i.e. mice colonized with complex communities or a defined set of human-derived microorganisms have been used as model systems to gain fundamental insights into the causal role of the microbiota in human diseases and confirm associations identified by human metagenome analysis [5]. Yet, it must be highlighted that mutualistic microbiota-host association has been established through long-term co-evolution of both partners. The gut microbiota of wild and laboratory mice harbors over 10 different bacterial phyla [6]. Although human and mouse microbiota are similar at the phylum level, substantial differences exist at lower taxonomic levels and between metagenomes [7]. Microbiota-host interactions in humanized mice do not necessarily recapitulate the situation in humans [8]. Therefore, mice colonized with mouse-derived bacteria represent a more natural tool to study physiological and host-specific microbiota-host interactions. While great efforts are made to isolate and sequence representative strains of the human microbiota [9], the number of mouse-derived bacteria in public strain collections is limited, which largely constrains the design of defined mouse microbial communities.

Previously, we showed that colonizing germ-free mice with members of the Altered Schaedler Flora (ASF), a bacterial consortium consisting of eight mouse-derived strains [10], is insufficient to provide CR against *Salmonella enterica* serovar Typhimurium (*S.* Tm) [11]. In the present study, we established a bacterial consortium, termed Oligo-Mouse-Microbiota comprising twelve bacterial isolates from the mouse intestine (Oligo-MM^{12}). Enforced with three facultative anaerobes, the Oligo-MM^{12} provides complete CR against *S.* Tm, which indistinguishable from

conventional mice. Thus, the Oligo-MM^{12} serves as a fundamental model microbiota that can be modularly reduced or expanded by additional strains for detailed investigation of microbe-microbe and microbe-host interactions in the context of enteric infections and other (patho)physiological conditions.

Results

Isolation and taxonomic characterization of Oligo-Mouse-Microbiota (Oligo-MM^{12}) strains

To establish a defined minimal microbiota that confers CR against *S.* Tm, we started by isolating strains which were representative of the bacterial phylum-level diversity of a conventional microbiota. Intestinal content of specified pathogen-free (SPF) mice was cultured on rich, non-selective anaerobic culture media. Twelve strains representing the five most prevalent and abundant phyla of the intestinal microbiota of laboratory mice were selected for a bacterial consortium of medium-complexity, referred to as "Oligo-Mouse-Microbiota" (Oligo-MM^{12}; **Supplementary Table 1**; **Supplementary Figure. 1 and 2**). 16S rRNA gene sequence alignment against the highly curated EzTaxon database [12] indicated that four strains of yet uncultured taxa were among the isolates (proposed names in quotation marks). Strain YL27 ('Muribaculum intestinale') is a member of a family ('Muribaculaceae') within the order Bacteroidales, previously classified as S24-7. YL27 forms a separate lineage related to the genera *Coprobacter* and *Barnesiella*. Furthermore, strain KB18 is the first representative of a genus within the Lachnospiraceae ('Acutalibacter muris'). Strain YL45 is the first representative of a genus within the *Sutterellaceae* ('Turicimonas muris'), and strain I48 is proposed to be a species ('Bacteroides caecimuris'). Detailed description of these taxa is provided in [13]. In summary, six members of Oligo-MM^{12} were assigned to the phylum Firmicutes, two strains to the Bacteroidetes, one strain to the Actinobacteria one strain to the Betaproteobacteria and one strain to the Verrucomicrobia (**Supplementary Table 1**).

Next, we investigated whether the consortium can stably colonize the mouse intestine over several generations, which is an important requirement for a model community in order to generate isobiotic mouse lines. Germ-free mice (two breeding pairs; parental generation) were inoculated with the frozen mixtures of the 12 strains and bred in germ-free isolators to the F6 generation. We developed a strain-specific hydrolysis-probe based qPCR assay, which allowed specific and sensitive absolute quantification of 16S rRNA gene copy numbers (**Fig. 1; Supplementary Table 3**). All strains but *Bifidobacterium longum* subsp. *animalis* YL2 were detected in fecal samples of individual mice of the consecutive generations, indicating stable colonization and vertical transmission of 11 strains. Based on these data, we conclude that *B. longum* subsp. *animalis* YL2 either does not colonize or is below the detection limit.

The Oligo-MM^{12} consortium confers partial protection against oral *Salmonella* serovar Typhimurium infection

Enteric *Salmonella enterica* serovar Typhimurium (*S*. Tm) infection in mice is critically influenced by the microbiota and is therefore widely used as a model for studying microbiota–pathogen interactions in the gut [14]. Mice colonized with members of the ASF are highly susceptible to intestinal *S*. Tm colonization [11] (**Supplementary Figure 3**). Transplantation of a complex microbiota from conventional mice restores CR against *Salmonella* to the level of conventional mice [11]. To assess the potential of the Oligo-MM^{12} to confer CR against *S*. Tm infection, we transplanted mice already colonized with five members of the Altered Schaedler Flora (ASF360, ASF361, ASF457, SB2 [ASF502] and ASF519; ASF^{5}) with the Oligo-MM^{12}. ASF^{5} mice with or without additional transplantation of cecum content from conventional mice (CON) were used as controls (**Fig. 2A)**.

Successful microbiota transplantation was confirmed by 16S rRNA gene amplicon sequencing of feces at day 40 post-transplantation (p.t.). As expected, the number of observed species in ASF^{5}+CON was much increased compared to the other groups (**Fig. 2B**). Compositional analysis indicated successful transfer of Oligo-MM^{12} taxa (**Fig. 2C ; Supplementary Figure 4A**; **Supplementary Table 5 and 6**). Diversity analysis comparing microbiota community composition among the three different transplant groups revealed that the transplant type alone could explain 84% of the overall variance ($p<0.001$) and the constrained ordination showed a clear clustering of the samples (**Fig. 2D**). Furthermore, we observed a clear separation of ASF^{5}+CON, ASF^{5}+control and ASF^{5}+Oligo-MM^{12} by single-linkage hierarchical clustering (**Fig. 2E**). Compared to the ASF^{5}+Oligo-MM^{12}, members of additional bacterial families were detected in the ASF^{5}+CON (**Fig. 2C; Supplementary Table 5**).

At day 40 p.t., all mice were orally infected with *S.* Tm (*S.* Tm^{avir}; 5x10^{7} cfu). Gut inflammation induced by *S.* Tm wildtype is known to cause dysbiosis and compromise CR [15]. Therefore, we used a *S.* Tm *invG sseD* mutant strain, which is able to efficiently colonize the gut but is defective in tissue invasion and induction of inflammation. Fecal pathogen loads were determined one day post-infection (p.i.). Mice were sacrificed at day 2 p.i. and pathogen loads were determined in the cecum content and mesenteric lymph nodes (mLN) (**Figure 2F-H**). Strikingly, Oligo-MM^{12} transplanted mice displayed increased CR at day 1 p.i. as ASF^{5}+CON transplanted mice (**Figure 2F**). At day 2 p.i., *S.* Tm loads in the cecum content increased but were still significantly lower than in ASF^{5} control mice (**Figure 2G**). Interestingly, Oligo-MM^{12}-transplanted mice exhibited a lower relative cecal weight than ASF^{5} mice (**Figure 2I**), which may indicate a microbiota-induced “normalization” of intestinal physiology (**Supplementary Figure 4BC**) [16].

Functional genomic analysis of the Oligo-MM^{12} consortium

In order to gain insights into the potential functional capabilities of the individual Oligo-MM members, the genomes of the twelve strains were sequenced and analyzed (**Supplementary Table 2**). In addition, genomes of the eight ASF strains were included in the comparative analysis, as this consortium is still widely used to generate gnotobiotic mice with a stable and defined mouse-derived microbiota [17, 18], although strains are not publically available. Identified open reading frames were annotated and the predicted protein sequences were separately matched against the KEGG database [19]. Presence and completeness of KEGG modules was determined for individual genomes and used for hierarchical clustering (**Fig. 3**). It is noteworthy that KEGG module analysis is biased towards gene sets, pathways and functional groups of well-characterized bacteria (e.g. *Escherichia coli*, *Bacillus subtilis*, *Bacteroides thetaiotaomicron*). Depending on the strain, we only found between 18-57% of positive KEGG BLAST hits for the predicted protein sequences for each genome (**Supplementary Table 2**). Accordingly, the unmatched predicted protein sequences, which were not included in our analysis, may involve functions and pathways present in these strains.

Hierarchical clustering of the KEGG modules in general led to grouping of the strains according to their phylogenetic membership. One cluster contained modules highly conserved in the majority of strains (**Fig. 3, cluster 6; Supplementary Table 7**, including modules such as the Sec secretion system (M00335), DNA polymerase III complex (M00260), RNA polymerase (M00183), ribosome (M00187/M00179), glycolysis core module (M00001/M00002) and various amino-acid and nucleoside biosynthesis pathways. Besides, we identified a cluster of modules including a high number of carbohydrate uptake systems (phosphotransferase systems (PTS), peptide- and amino-acid transporters), enriched in Gram-positive strains (**Fig. 3, cluster 4; Supplementary Table 7**). Among others, modules more prominent in Gram-negative strains comprised lipopolysaccharide (LPS) biosynthesis (M00320; M00060,) and biotin biosynthesis modules (M00577, M00573, M00123) (**Fig. 3, cluster 5; Supplementary Table 7**). One cluster harbored modules enriched in the class Bacilli (**Fig. 3, cluster 3; Supplementary Table 7**).

Next, we determined the fraction of KEGG modules of a conventional mouse microbiota (CON) represented in the collective genomes of Oligo-MM^{12}. For comparison, we also included a combined metagenome of all eight ASF strains (ASF^{8}), as this consortium has been widely used in the past. We generated artificial metagenomes *in silico* by merging contigs of individual strains and compared them to eight metagenomic datasets from conventional mice. In total, 229 of the 471 modules identified in all metagenomic datasets were conserved among all three consortia, whereas 157 were only found in the conventional metagenome (**Fig. 4; cluster 1; Supplementary Table 8**), including functions like methanogenesis and coenzyme M biosynthesis (M00356, M00357, M00358, M00563), various archaeal functions (M00184,

M00327, M00343) and photosynthesis (M00161, M00163). We speculate that the latter may originate from plant-derived diet, as photosynthetic bacteria are not autochthonous inhabitants of the mammalian gut. Interestingly, cytochrome modules (M00151, M00152, M00155, M00156, M00162, M00416, M00417) were also specific to the CON metagenome. Cytochrome oxidases form a superfamily of proteins that act as the terminal enzymes of respiratory chains. Therefore, besides in plants and eukaryotic mitochondria, they are typically found in aerobic, microaerophilic and facultative anaerobic bacteria. Overall, the defined consortia ASF[8] and Oligo-MM[12] together covered 66,6% of the KEGG modules of the CON microbiome. The Oligo-MM[12] metagenome contributed 61 modules that were not represented among the ASF[8] consortium (**Fig. 4; cluster 3**).

Rational design of a consortium conferring increased CR based on predictions from functional genomic analysis

CR conferred by a conventional microbiota is higher compared to the Oligo-MM[12] (**Fig. 2FG**). Based on the comparative metagenomic analysis, we hypothesized that microbiota functions (i.e. KEGG modules) of the conventional metagenome may, to some extent, complement functions absent in Oligo-MM[12] (**Fig. 4; cluster 1,2**). We sought to obtain evidence for our hypothesis that CR in gnotobiotic mice can be promoted by increasing the coverage of KEGG modules found in a conventional microbiota. Facultative anaerobic bacteria are underrepresented in the Oligo-MM[12] (e.g. Lactobacilli, Enterobacteriaceae). On the other hand, the conventional metagenome harbors modules involved in bacterial respiration, such as cytochromes. Therefore, facultative anaerobic bacteria may be important to promote CR against *S.* Tm, for example by competition for oxygen or anaerobic electron acceptors such as nitrate.

To verify this idea, we used three sequenced facultative anaerobic bacteria (FA[3]), *E. coli* Mt1B1, *Streptococcus danieliae* ERD01G and *Staphylococcus xylosus* 33-ERD13C from a recently established strain collection of mouse intestinal bacteria (miBC;[13]) (**Supplementary Figure 5; Supplementary Table 9**). An artificial FA[3] metagenome was generated, KEGG modules were identified and included to the comparative hierarchical cluster analysis (**Supplementary Figure 6; Supplementary Table 10**). In total, 487 KEGG modules were represented by all three artificial and the CON metagenome. The FA[3] consortium covered with 332 modules a high fraction of the CON metagenome (68,2%). These KEGG modules comprised 63 modules that were absent in the defined microbial communities (ASF[8] and Oligo-MM[12]; **Supplementary Figure 6; cluster 3**). Among others, KEGG modules involved in respiration (e.g. cytochrome aa3-600 menaquinol oxidase M00416, cytochrome o ubiquinol oxidase M00417, cytochrome c oxidase M00154) were found in the FA[3] metagenome (**Supplementary Figure 6; Supplementary Table 10**). Of note, *E. coli* Mt1B1 contributed the majority of new modules, which likely reflects the fact that *E. coli* is by far the most studied bacterial species to date. *E.*

coli Mt1B1 showed striking similarities to *S.* Tm SL1344 on the metagenomic level, which reflects close phylogenetic relationship of these two species (**Supplementary Figure 5; Supplementary Table 9**).

To test the correlation of FA^3-specific KEGG modules and CR, Oligo-MM^{12} mice were transplanted with a mixture of the 3 strains and CR against *S.* Tm was tested as described before (**Fig.5AF**). Oligo-MM^{12} mice transplanted with FA^3 exhibited a pronounced increase in CR at day 1 and day 2 p.i. (**Fig.5BC**). Strikingly, *S.* Tm colonization loads were reduced to a median of $1.5x10^4$ cfu/g, which is the level of CR in ASF^5 mice transplanted with a complex conventional microbiota (**Fig. 2FG; Fig. 6**). *S.* Tm colonization of the mLN was not significantly reduced by FA^3 transplantation (**Fig.5D**), and also the relative cecal weight was unaltered (**Fig.5E**). Most importantly, transplantation of FA^3 into ASF mice did not lead to an increase in CR, suggesting that FA^3 can only confer protection in combination with a specific set of strains, such as the Oligo-MM^{12} (**Supplementary Figure 7**). Finally, we set out to verify that the newly identified consortium of 15 strains also protects against disease induced by Wild type *S.* Tm (*S.* Tm^{Wt}). For comparison, we included groups of ASF^4 and conventional C57Bl/6 mice (CON). All mice were orally infected with *S.* Tm^{Wt} and daily monitored for signs of disease (**Supplementary Figure 8A**). ASF^4 were highly susceptible to colonization, gut inflammation and systemic infection by *S.* Tm^{Wt} and had to be taken out of the experiment after 2-3 days as they showed terminal signs of disease. All other mice were sacrificed 4 days p.i.. *S.* Tm^{Wt} efficiently colonized the gut of Oligo-MM^{12} mice by day 3-4 p.i. and animals concomitantly developed signs of gut inflammation. In contrast, Oligo-MM^{12} mice transplanted with FA^3 were equally well protected as CON mice against *S.* Tm^{Wt} colonization and disease (**Supplementary Figure 8**).

Discussion

Early culture-based studies demonstrated that CR-mediating bacteria can be isolated, grown *in vitro* and confer CR upon reintroduction in germ-free or antibiotic-treated mice [20]. Yet, none of these isolates have been deposited in public culture collections. In the present study, we progressively assembled a defined consortium of mouse intestinal bacteria by “genome-guided design” that provides full CR against *S.* Tm. Enteric *Salmonella* infection can be divided into two main stages: an initial expansion phase where *Salmonella* can grow to high luminal loads (10^9 cfu/g), followed by triggering of inflammation and pathogen growth in the inflamed gut in Enterobacteriaceae ‘blooms’. In this study, we focused on the first stage of *S.* Tm infection, which is largely influenced by the gut microbiota. Loss of microbial diversity in the course of antibiotic-therapy can open up ecological niches that are exploited by infecting pathogens [21]. In fact, similar mechanisms may underlie the increased susceptibility to *S.* Tm infection of mice colonized with a low-complex gut microbiota, such as the ASF. Indeed, microbial community

complexity overall correlated with CR in our experiments (**Fig. 6**), which is consistent with Rolf Freter´s nutrient niche theory [22]. This theory states that populations of indigenous bacteria are controlled by substrate competition in a complex community. A strain (e.g. pathogen) can only invade if it is able to use a specific limiting nutrient more efficiently than the rest of the community. Accordingly, CR against pathogens is mediated by efficient restriction of all available nutrient niches by a complex microbial community. Competition for carbohydrates [21] but also hydrogen [23], iron [24] and anaerobic electron acceptors (e.g. NO_3) [25] have been shown to substantially influence *S.* Tm gut ecosystem invasion and colonization. Modules for iron uptake systems, carbohydrate utilization, PTS systems, B12 transport and Trimethylamine-N-oxide (TMAO) respiration are represented in FA^3 and Oligo-MM^{12} metagenomes. These pathways may as well play a role in CR against *S.* Tm in this model as *Salmonella* presumably depends on utilizing these nutrients in the mammalian gut [26].

S. Tm consumes epithelial-derived oxygen by aerobic respiration when CR is impaired [27]. In fact, cytochrome modules (M00416; M00417; M00153; M00154; M00155) and dissimilatory nitrate reduction (M00530; M00471; M00472; M00483) were identified in the FA^3 or Oligo-MM^{12} metagenomes (**Supplementary Table 10**). This suggests, that, in particular, competition for oxygen, i.e. by facultative anaerobic commensals, might play a role in CR. Intriguingly, the effect of streptomycin treatment on CR has previously been ascribed to its selective elimination of these facultative bacteria [28]. This is consistent with our finding that transplantation of *E. coli*, *Streptococcus danieliae*, and *Staphylococcus xylosus* restored CR of Oligo-MM^{12} to the level of conventional mice. Hierarchical KEGG module clustering revealed a striking similarity of *E. coli* Mt1B1 and *S.* Tm SL1344, suggesting that they might occupy a similar niche in the gut. Of note, the FA^3 consortium only increased CR in the background of the more complex Oligo-MM^{12} consortium, but not in ASF mice (**Supplementary Figure 7**). This strongly indicates that protection by FA^3 in Oligo-MM^{12} mice occurs via occupation of niche space (e.g. consumption of nutrient resources) that can no longer be exploited by *S.* Tm. In order to characterize the physical niches occupied by *S.* Tm and the closely related *E. coli* Mt1B1 strain, we performed FISH of cecal tissue of control mice or mice transplanted with FA^3 and infected with *S.* Tm for 2 days (**Fig. 5 GHI**) and characterized localization of *S.* Tm^{avir} and *E. coli* Mt1B1 in the cecum with regard to luminal or epithelial enrichment. Interestingly, *E. coli* was found to be significantly enriched at the epithelium compared to the lumen (mean increase: 1.85-fold). However, for *S.* Tm in the absence of *E. coli*, no such enrichment was observed and *S.* Tm evenly colonized the gut lumen and epithelial regions in the control mice with overall ~4% abundance (**Fig. 5 GH**). This implies that, in Oligo-MM^{12} mice, *S.* Tm is not enriched at the epithelium (e.g. higher levels of oxygen or mucins) but rather is evenly distributed in the gut lumen, which is in contrast to streptomycin-treated mice [29]. However, since *E. coli* is only slightly enriched at the epithelium and otherwise shows a similar distribution/abundance as *S.* Tm, it is conceivable that *E. coli*

297 could actually fill up the niche space which is preferred by *S.* Tm in Oligo-MM[12] mice and thereby
298 prevents infection.

299 **B**esides competition for nutrients, the microbiota can produce an arsenal of antibacterial and
300 anti-virulence molecules, such as short-chain fatty acids (SCFAs), secondary bile acids, and
301 bacteriocins that could antagonize pathogen growth, viability and virulence. *E. coli* Mt1B1 did not
302 produce a colicin with activity against *S.* Tm. However, it encodes a type VI secretion systems
303 (T6SS), which may be involved in *S.* Tm inhibition [30].

304 **I**n conclusion, we showed that genome-guided assembly of defined microbial consortia from a
305 collection of diverse mouse-derived bacterial strains is an efficient experimental approach for
306 generating a minimal consortium with a defined function for their host. As a proof-of-concept, we
307 have assembled a minimal consortium of fifteen bacterial strains that provided full CR against *S.*
308 Tm. We show that prominent facultative anaerobic bacteria play a major role in full protection
309 against *S.* Tm infection. Of note, this is the first defined consortium of cultured bacteria that
310 provides CR against *S.* Tm and therefore represents a unique experimental model to further
311 elucidate the underlying molecular and ecological CR mechanisms. In recent years, initiatives
312 have been launched to generate a comprehensive genomic catalogue of type strains
313 represented in the human microbiota. Expansion of the catalogue of publically-available strains
314 by mouse-derived microorganisms, as recently implemented [13], will spur the creation of
315 functionally defined and simplified microbial consortia for application in gnotobiotic animals, as
316 exemplified in the present study. Combined with genetic manipulation of the host, such strain
317 collections can yield unprecedented insights into host-microbiota interactions.

Conflict of interest

The authors declare no conflict of interest.

Methods

Bacterial strains. ASF356, ASF360, ASF361, ASF457, ASF500 and ASF519 were provided by Charles River, 251 Ballardvale Street, Wilmington, MA 01887. SB2, a re-isolate of ASF502, was obtained by cultivation on Schaedler blood agar using feces from a C57Bl/6 mouse colonized with ASF strains (ETH, Zurich). The Oligo-MM^{12} strains were isolated as described previously [13]. Strains *Clostridium innocuum* I46 (DSM 26113), 'Bacteroides caecimuris' I48 (DSM 26085) and *Lactobacillus reuteri* I49 (DSM 32035) were isolated from fecal pellets of C57Bl/6 mice (Janvier, Le Genest St. Isle, Rodent Center HCI, ETH Zurich) on Wilkins-Chalgren agar (Oxoid) supplemented with 4 g/l glucose, 10 µg/l hemin, 0.4 g/l L-cystine 5%, 15g/l agar and 5% defibrinated sheep blood (WSB) after growth under anoxic atmosphere (7% H_2, 10% CO_2, rest N_2) at 37 °C for 3-5 days. *Bifidobacterium longum* subsp. *animalis* YL2 (DSM26074), 'Muribaculum intestinale' YL27 (DSM 28989), *Flavonifractor plautii* YL31 (DSM 26117), *Clostridium clostridioforme* YL32 (DSM 26114), *Akkermansia muciniphila* YL44 (DSM 26109), 'Turicimonas muris' YL45 (DSM 26109) and *Blautia coccoides* YL58 (DSM 26115) were isolated from cecum content of C57Bl/6 mice (Janvier, Le Genest St. Isle, Rodent Center HCI, ETH Zurich). Cecum content was resuspended in anaerobic diluting fluid as described [31] and either directly plated on AII agar (18.5 g/l Brain-Heart-Infusion, 5g/l yeast extract, 15g/l trypticase soy broth, 2.5g/l K_2HPO_4, 10 µg/l hemin, 0.5 g/l glucose, 0.33 g/l palladium chloride, 42 mg/l, 50mg/l cysteine, Na_2CO_3, 5 µg/l menadione, 3 % fetal calf serum (complement-inactivated) and 15 g/l agar as modified from [31]) or plated on WSB agar after heat- (10min, 80°C) or $CHCl_3$-treatment. Plates were incubated under anoxic atmosphere (7 % H_2, 10 % CO_2, rest N_2) at 37°C for 7 days. 'Acutalibacter muris' KB18 (DSM 26090) and *Enterococcus faecalis* KB1 (DSM 32036) were isolated from feces of a SPF TCR$^{MOG92\text{-}106}$/I-A^s transgenic (RR) SJL/J mouse [32], which was heat-inactivated at 70°C for 15 min and frozen at -20°C. *E. faecalis* KB1 was isolated from brain heart infusion (BHI) agar. 'Acutalibacter muris' KB18 was isolated by limited dilution in liquid BHI. *Escherichia coli* Mt1B1 (DSM 28618), *Streptococcus danieliae* ERD01G (DSM 22233) [33] and *Staphyloccocus xylosus* 33-ERD13C (DSM 28566) were obtained from a recently established collection of mouse intestinal bacteria (MIBC) [13]. The Oligo-MM^{12} strains are part of this collection and can be obtained via the German Collection of Microorganisms and Cell Cultures (www.dsmz.de/miBC).

Culture conditions. For cryopreservation, 1 vol. bacterial mixture or individual strains was mixed to 1 vol. 20% glycerol solution supplemented with palladium black crystals [Sigma] and 1ml aliquots were prepared in 1.5ml glass vials, sealed with butyl-rubber stoppers (Wheaton)

and frozen at -80°C within 1h. Culture purity and identity for individual strains was confirmed by 16S rRNA gene sequencing for each cryostock. Anaerobic media, solutions and glass bottles were pre-reduced at least 2 days before use under anoxic conditions (3% H_2, rest N_2) in an anaerobic chamber. For cultivation, 1ml cryostocks containing palladium black and 10% glycerol were thawn at 37 °C in a water bath. A single vial was inoculated into a 100ml Wheaton serum bottle (Sigma) sealed with a butyl rubber stopper containing 10ml of pre-reduced growth medium. Liquid cultures were gassed (7% H_2, 10% CO_2, rest N_2) and incubated until bacterial growth was observed. Anaerobic bacterial strains were grown either in BHI (37g/L BHI [Oxoid] 0.25g/L Cysteine-HCl.H_2O, 0.25g/L $Na_2S.9H_2O$) or in anaerobic Akkermansia medium [34] (AAM; 18g/L BHI [Oxoid], 15g/L trypticase soy broth, 5g/L yeast extract, 0.25g/L, 2.5g/L K_2HPO_4, 1mg/L hemin, 0.5g/L D-Glucose, 0.5mg/L menadione, 3% heat-inactivated fetal calf serum, 0.25g/L hog gastric mucin [Sigma; only for YL44], cysteine-HCl.H_2O, 0.25g/L $Na_2S.9H_2O$). *S.* Tm^{Wt} [35] and *S.* Tm^{avir} (*invG*, *sseD* ::*aphT*) [36] were grown on LB agar supplemented with streptomycin (50µg/ml) and kanamycin (30µg/ml). For mouse infection, a single colony was inoculated in 3 ml LB medium containing 0.3M NaCl ($LB^{0.3}$) and grown at 37°C on a wheel rotor for 12h. The culture was diluted 1:20 in fresh $LB^{0.3}$ medium and grown at 37 °C for another 4h. The subculture was washed in ice-cold PBS and the bacterial pellet was re-suspended in PBS. For infection, mice were gavaged with $5x10^7$ cfu *S.* Tm strains.

Plasmids. Plasmids harboring 16S rRNA genes used for the qPCR assay are listed in **Supplementary Table 4**. To generate plasmids pM1411-1, pM1412-4, pM1413-1, pM1414-1, pM1417-1, pM1456-1, pM1457-1, pM1459-1 and pM1460-1, the 16S rRNA gene was amplified using primer fD1-EcoRV-XbaI, fD2-EcoRV-XbaI and rP1-EcoRV-BamHI (**Supplementary Table 3**). For plasmids pSAB3 to pSAB13 pDK4, pDK5, and pDK6, primer fD1/2 and rP1 were used (**Supplementary Table 3**). pSAB3 to pSAB13 as well as pDK4 to pDK6 were generated using the CloneJET™ PCR Cloning kit (Thermo Fisher Scientific) following manufacturer's instructions. Inserts were sequenced using primers pJet1-FP and pJet1-RP. For pM1456-1, pM1457-1, pM1459-1 and pM1460-1 purified PCR products were inserted in the linear pCR®2.1-TOPO® vector using the TOPO® Cloning kit (Invitrogen) following manufacturer's instructions. To generate plasmids pM1411-1, pM1412-4, pM1413-1, pM1414-1 and pM1417-1, PCR products were cloned into linearized pSB-Bluescript SK II (Stratagene). To generate pM1411-1, pM1412-4 and pM1413-1, pSB-Bluescript SK II was linearized using EcoRV; for pM1414-1 using NotI/HindIII and for pM1417-1 using NotI/BamHI.

Taxonomic assignment of the Oligo-MM^{12} strains. PCR was performed as previously described using primers targeting the entire 16S rRNA gene (**Supplementary Table 4**). Full length 16S rRNA gene sequences were generated by PCR using primers fD1-EcoRV-XbaI/fD2-EcoRV-XbaI and rP1-EcoRV-BamHI (**Supplementary Table 3**), subcloned and sequenced from

both ends. Sequences were trimmed and assembled using the software CLC DNA Workbench (version 6.0.2). 16S rRNA gene sequences were blasted against the Ribosomal Database Project (RDP) [37], NCBI blast [38], Greengenes [39] databases or aligned against SILVA [40] using SINA. The latter was preferred for taxonomic assignment. The isolates were also identified using EzTaxon-e [12]. Strains assumed to represent novel taxa (i.e. <97 % 16S rRNA sequence identity to a bacterium with a validated name) were identified by phylogenomic, enzymatic and chemotaxonomic analysis [13].

Genome sequencing. Bacterial cultures were prepared in filtered (0.2 µm) media. Genomic DNA (gDNA) extraction was performed by phenol/chloroform/isoamylalcohol extraction. Briefly, bacterial pellets of Gram-positive strains were resuspended in TE buffer supplemented with 0.5% SDS and 20mg/ml lysozyme and bacterial suspensions were incubated at 37°C for 90min. Suspensions were supplemented with 0.1µg/ml proteinase K and incubated at 55°C for 60min. The bacterial pellets of Gram-negative strains were directly resuspended in TE buffer supplemented with 0.5% SDS and 0.1µg/ml proteinase K, then incubated at 55°C for 1h. 0.64 M NaCl and 0.1 volume (vol.) CTAB/NaCl buffer (10% CTAB, 0.7M NaCl) were added and incubated at 65 °C for 10min. Afterwards, gDNA was extracted using phenol/ chloroform/ isoamylalcohol (25 :24 :1). The genome sequence of 'Acutalibacter muris' KB18 was obtained on the PacBio sequencing platform (DSMZ, Braunschweig, Germany). For the remaining 11 Oligo-MM strains, whole-genome shotgun sequencing was performed on the Illumina Miseq sequencing platform (Eurofins Genomics GmbH, Ebersberg, Germany). Libraries of 500-bp insert size were prepared from the isolated genomic DNAs and sequenced as 300-bp paired-end runs on an Illumina MiSeq v3 instrument. Raw Illumina reads were de novo assembled using SPAdes, version 3.5.0 [41], careful mode, with a minimum read coverage cutoff of 20 and minimum contig-length of 500 bp. The quality of the Illumina draft genome assemblies was assessed with QUAST [42]. Illumina and PacBio raw sequences have been deposited in the Sequence Read Archive (SRA) NCBI database. Accession numbers are given in **Supplementary Table 2**.

Metagenome analysis. Input files were assembled genomes of Oligo-MM, ASF [17] and FA strains [13]. Artificial metagenomes were created by merging contigs of each genome into a multi-fasta file. Sequences were merged and assembled with Ray-Meta version 2.3.1 with default parameters [43]. KEGG mapping was performed using a custom pipeline. Gene prediction was performed with Prodigal version 2.60 [44] and the predicted protein files of the four groups (Oligo-MM[12], ASF[8], FA[3] and conventional) were aligned separately against a reduced KEGG database provided by the lab of Dr. Curtis Huttenhower (http://huttenhower.sph.harvard.edu/) using RAPSearch version 2.23 [45]. For obtaining the reduced KEGG database, genes with no KO were removed from the full database (release 58) and then genes within each KO were clustered at 85% sequence identity. Custom Python scripts (available:

https://github.com/pseudonymcp/keggmapping, http://dx.doi.org/10.6084/m9.figshare.1404958; BLAST criteria: log10 e-value<=-5, bitscore>=60, Percent identity>=60) were used for alignment against the KEGG database to obtain information about the presence and completeness of each KEGG module, expressed as on a scale from 0 to 4 (0=complete,1=1 block missing, 2=2 blocks missing, 4=absent).

DNA extraction from intestinal contents. Small intestinal, cecal and fecal gDNA were extracted using the QIAamp DNA Stool Mini Kit (Qiagen) following the manufacturer's instructions with modifications. An initial bead-beating step using differentially sized beads (glass beads: 0.5-0.75 mm and zirkonia beads: <100 µm) was included and lysozyme (20 mg/ml) was added to the lysis buffer.

16S rRNA gene amplicon sequencing. PCR comprised two consecutive steps. In the first step, primers 338F-M13 and 1044R-rM13 targeting the 16S rRNA gene carrying 5'M13/rM13 adapters (**Supplementary Table 3**) were used to amplify the V3-V6 region of the bacterial 16S rRNA gene. One PCR reaction contained 500ηM of each primer, 2 x DreamTaq PCR Master Mix and 50ηg template DNA. PCR reaction was performed in duplicates using a peqSTAR 2X Gradient Thermocycler (Peqlab Biotechnology). PCR conditions were: 95°C for 10 min, followed by 20 cycles of 95°C for 30 s, 55°C for 30 s and 72°C for 45 s and a final elongation step at 72°C for 10 min. Duplicate reactions were pooled and purified using the NucleoSpin PCR Clean-up kit. Concentration of the purified PCR products was determined by Nanodrop (Peqlab Biotechnology). PCR products were used for a second PCR reaction to add a specific multiplex identifier (MID) sequence (**Supplementary Table 11**) and the 454-specific Lib-L tag using primers A-M13/B-rM13. PCR was performed using 400ηM of each primer (A-M13 and B-rM13). PCR conditions were: 95°C for 10 min, followed by 10 cycles of 95°C for 30 s, 60°C for 30 s and 72 °C for 60 s and a final elongation step at 72°C for 10 min. Amplicons were pooled, concentrated by ethanol precipitation, purified by gel electrophoresis and AMPure beads (Beckmann Coulter) and finally resuspended in ddH_2O. Amplicon-sequencing was performed at Eurofins, on a 454 GS FLX Titanium platform from one side (Lib-L-A) according to the recommended procedures for 454 Roche. The QIIME software package version 1.8 [46] was used for read pre-processing, OTU clustering, taxonomic assignment and alpha diversity analysis. Briefly, OTU clustering was performed at the 97% similarity level using an open-reference method, based on the SILVA database [40] and a custom database of the full length 16S rRNA gene sequences of the 12 Oligo-MM and 5 ASF strains (ASF360, ASF361, ASF457, ASF502, ASF519). The custom database was able to distinguish all strains except the Lactobacilli (ASF360, ASF361 and I49) at the 97 % sequence similarity level. Alpha diversity was determined using the metric of observed species as measure of within-sample diversity. In order to compare all samples at equal sequencing depth for diversity analyses, reads were normalized to the sample with lowest number of reads. To quantify the contribution of the transplant type to

differences in microbiota composition, we applied a permutational multivariate ANOVA based on distance matrices (PERMANOVA) and a canoncial analysis of principal coordinates (CAP) using the functions adonis, anova.cca and capscale with 5000 permutations of the vegan package in R. Only significant differences that were supported by both methods and all three dissimilarity measures (Bray-Curtis, weighted and unweighted UniFrac) were reported. In each case we controlled for false discovery rate (FDR) using the Benjamini and Hochberg procedure with alpha set to 0.05. The hclust function was used to perform the hierarchical cluster analysis of the dissimilarity matrix. R scripts are accessible at https://github.com/hzi-bifo/OligoMM.

Quantitative PCR of bacterial 16S rRNA genes. 16S rRNA specific primers and hydrolysis probes were designed using Primer Express 3 (Applied Biosystems, Life Technologies). For duplex qPCR assays, hydrolysis probes were 5'-labelled with either 6-carboxyfluorescein (FAM) or 6-carboxyhexafluorescein (HEX). Each probe was conjugated with the black hole quencher 1 (BHQ1) at the 3' end. Primers and probes were synthesized by Metabion (Planegg, Germany). qPCR conditions were established according to the MIQE guidelines (Minimum Information for Publication of Quantitative Real-Time PCR Experiments) [47]. All primers were designed for an optimal annealing temperature of 60°C. qPCR standard curves were determined using linearized plasmid as DNA template (**Supplementary Table 4**). Plasmid DNA was diluted in H_2O containing 100 ηg/µl yeast t-RNA (Roche). Standard curves were run on a Roche Lightcycler96 instrument in triplicates.

Efficiency of each qPCR reaction was calculated based on the slope of standard curves (qPCR efficiency: $(10^{(-1/\text{slope of standard curve})}-1) \times 100$) using 10-fold dilutions of template. Efficiencies for all qPCR reactions were within the range of 90-110 %. For all experiments the software LightCylcer96 version 1.1. reproduced standard curves based on single DNA template with known DNA quantity as well as the efficiency derived from the standard curve of each qPCR assay of the initial run of the standard curves. Specificity was confirmed by performing an assay for each primer/probe pair using a equimolar mixture of all linearized plasmids except for the one to be tested as template. One PCR reaction (total volume: 20µl) contained 300ηM of each primer, 250ηM of the corresponding hydrolysis probe (**Supplementary Table 3**), FastStart Essential DNA Probes Master (Roche) and 5ηg template gDNA. PCR reactions with DNA templates extracted from feces or cecal content were run in duplicates. PCR conditions were: 95°C for 10min, followed by 45 cycles of 95°C for 15s and 60°C for 1min. Fluorescence for each cycle was recorded after the step at 60°C. Quantification cycle (Cq) as well as the baseline were automatically determined by the software LightCycler96 version 1.1 (Roche).

The detection limit ranged between 2-338 16S rRNA gene copies/5ηg template gDNA (**Supplementary Table 3**). By correlating *S.* Tm cfu/g with respective values of 16S rRNA copies/5ηg template gDNA, the detection limit of the qPCR assay (25 copies/5ηg template

gDNA) corresponds to $1{,}4x10^5$ cfu *S.* Tm /g intestinal content. Due to the different individual detection limits of the qPCR assay for each primer/probe combination, the detection limit for Oligo-MM[12], ASF and FA[3] strains ranged between $1{,}1x10^4$ cfu/g to $1{,}9x10^6$ cfu/g (**Supplementary Table 3**).

Animal experiments. Germ-free C57Bl/6J mice were obtained from the Clean Mouse Facility (CMF, University of Bern, Switzerland) and directly inoculated with a mixture of ASF[5] (ASF360, ASF361, ASF457, SB2, ASF519) or ASF[4] strains (ASF356, ASF361, SB2, ASF519). C57Bl/6 mice stably colonized with the defined Oligo-MM[12] consortium were also designated as stable Defined Moderately Diverse Microbiota mouse (sDMDMm2) and were generated by Prof. McCoy and Prof. Macpherson (University of Bern, Switzerland) [48]. SPF C57Bl/6J mice harboring a conventional microbiota were purchased from Janvier (Le Genest-Saint-Isle). Gnotobiotic mice (all C57Bl/6J) were bred under germ-free conditions in flexible film isolators (Harlan Laboratories). A protocol was developed that allows straightforward inoculation of the consortium as a frozen mixture containing all individual strains. The advantage of this protocol is that the inoculum can be readily shipped and inoculations can be carried out in facilities lacking microbiological expertise or equipment. Gnotobiotic mice were orally and rectally inoculated with 100 µl bacterial mixture from frozen stocks (SPF cecum content, bacterial consortia). For *S.* Tm^{avir} infection, sex- and age-matched female or male animals (8-12 weeks) were orally gavaged with $5x10^7$ cfu *S.* Tm^{avir} under germ-free conditions and maintained in gnotocages (Han, Bioscape, Emmendingen). Mice were sacrificed by cervical dislocation and organs were removed aseptically. Live *S.* Tm loads in the cecal content and the mLN were determined by plating on MacConkey-agar (Roth). Histology of the cecum was done at necropsy. Cecum tissue was embedded in O.C.T. (Sakura, Torrance) and flash frozen. Cryosections (5µm) of the cecal tissue were H&E-stained and scored as described in detail in [15]. The parameters submucosal edema, infiltration, loss of goblet cells and epithelial damage were scored in a blinded manner according to the severity of inflammatory symptoms yielding a total score of 0–13 points. Animals were excluded from the experiment when they showed terminal signs of disease before the planned end of the experiment. Number of animals per group was calculated to detect biological relevant effects, i.e. defined by P ("cfu/g higher or lower than control group") = 0.9. alpha error = 0.05; beta-error = 0.2. Sample size ranged between 4-6 animals/group including reserve. All animal experiments were approved by the local authorities (Regierung von Oberbayern) and an ethics committee and performed according to the legal requirements.

Fluorescence in situ hybridization: Cecal tissue was fixed in 4% paraformaldehyde (overnight, 4°C), equilibrated in 20% sucrose (overnight, 4°C), embedded in O.C.T (Sakura) and flash frozen in liquid nitrogen. Sections (7µm) were mounted on glass slides and dried at RT overnight. Fluorescence *in situ* hybridization (FISH) was performed according to a standard protocol. Briefly, slides were dehydrated in 50%; 80% and 96% ethanol and air-dried. Double

3´and 5´-labelled 16S rRNA targeted probes (1:1 mix of Eub338I-2xCy5 (GCT GCC TCC CGT AGG AGT) and Eub338III-2xCy5 (GCT GCC ACC CGT AGG TGT) to target all bacteria and Ent186-2xCy3 (CCC CCW CTT TGG TCT TGC) were added to hybridization buffer (0.9M NaCl, 20mM Tris/HCl, 30% formamide, 0.01% SDS) to a final concentration of 5ng/µl and incubated for 3h at 46°C in a humid chamber. Slides were washed in buffer containing 0.1M NaCl, 20mM Tris/HCl, 5mM EDTA for 10min at 48°C, rinsed in ice-cold ddH2O and DNA was stained with 1µg/ml DAPI (Roth) and 0.1µg/ml SYTOX green (Invitrogen) for 30min at 4°C. After three times washing in ice-cold ddH_2O, slides were mounted with Vectashield (Vector Laboratories) and sealed with nail polish. Immunofluorescent staining was performed as described [15]. Briefly, slides were washed once in 1xPBS, blocked 1h in 10% normal goat serum in PBS at RT and afterwards incubated with a polyclonal rabbit anti-*Salmonella* B serum (Difco, 1:400, 10% normal goat serum) for 1h. After washing three times in 1x PBS, slides were incubated with an secondary antibody (anti-rabbit DyLight 549; Jackson, 1:400, 10 % normal goat serum) and DNA was stained with 1µg/ml DAPI (Roth) and 0,1µg/ml SYTOX green (Invitrogen) for 1h at RT. After final washing in 1xPBS, slides were mounted with Vectashield (Vector Laboratories) and sealed with nailpolish. Imaging was performed using a Leica TCS SP5 confocal microscope with 63x oil objective. ImageJ software, version 1.48v (Wayne Rasband, National Institute of Health, USA) was used for image analysis [49]. For each mouse three pictures per location (epithelium or lumen) were taken and for each picture three regions of interest ($15\mu m^2$ >100 bacteria) were defined and all bacteria were counted (SYTOX green positive). The fraction of Enterobacteriaceae (Ent186-$Cy3^+$) or *Salmonella* (α-*Salmonella*$^+$) of all SYTOX green positive bacteria in the same region of interest was determined.

Statistical analysis. Statistical analysis was performed using the exact Mann-Whitney U test or Kruskal-Wallis and Dunn´s Multiple Comparison test using the software GraphPad Prism version 5.01 for Windows (GraphPad Software, La Jolla California USA, www.graphpad.com). P values of less than 0.05 (two-tailed for MW) were considered as statistically significant. * $P<0.05$, ** $P<0.01$, *** $P<0.001$.

Data availability. Bacterial shotgun genome sequences obtained in the present study are available at the Sequence Read Archives under accession nos. SRX1092348, SRX1092357, SRX1092347, SRX1092355, SRX1092353, SRX1092362, SRX1092358, SRX1092359, SRX1092354, SRX1092361, SRX1092352, SRX1092360 and the European Nucleotide Archive under accession nos. ERS1032682, ERS1032670, ERS1032680. Metagenomic reads were derived from 8 different samples of conventional and wild mice (SRA accession numbers SRX313003, DRX013306 and ERX166941; and MG-RAST id 4528728.3, 4528733.3, 4528748.3 4528749.3 and 4528734.3 [50]).

Figure Legends

Figure 1. The Oligo-MM^{12} consortium stably colonizes mice and is vertically transmitted across filial generations. Germ-free C57Bl/6 mice were inoculated with freshly thawn cryostocks containing the Oligo-MM^{12} strains 'Acutalibacter muris' KB18, *Flavonifractor plautii* YL31, *Clostridium clostridioforme* YL32, *Blautia coccoides* YL58, *Clostridium innocuum* I46, *Lactobacillus reuteri* I49, *Enterococcus faecalis* KB1, 'Bacteroides caecimuris' I48, 'Muribaculum intestinale' YL27, *Bifidobacterium longum* subsp. *animalis* YL2, 'Turicimonas muris' YL45 and *Akkermansia muciniphila* YL44 and bred for consecutive generations (parental to F6) in a germ-free isolator. Parental and F1 were bred at the CMF Bern and F2 to F6 at the gnotobiotic mouse facility of the MvP. Fecal samples were collected from adult female and male mice of each generation (n=2-11). Absolute abundance of each strain was determined by a strain-specific qPCR assay and is plotted as relative abundance of each individual strain. *below detection limit (Table S2). 'Acutalibacter muris' KB18, *Clostridium innocuum* I46, *Lactobacillus reuteri* I49, *Blautia coccoides* YL58 and *Enterococcus faecalis* KB1 were below 1% rel. abundance in the majority of samples.

Figure 2. Transplantation of the Oligo-MM^{12} consortium leads to increased CR against oral *S.* Tm infection. (A) Experimental design: Mice colonized with ASF360, ASF361, ASF457, SB2 [ASF502] and ASF519 (ASF^{5}) were transplanted with freshly-thawn cryostocks containing 'Acutalibacter muris' KB18, *Flavonifractor plautii* YL31, *Clostridium clostridioforme* YL32, *Blautia coccoides* YL58, *Clostridium innocuum* I46, *Lactobacillus reuteri* I49, *Enterococcus faecalis* KB1, 'Bacteroides caecimuris' I48, 'Muribaculum intestinale' YL27, *Bifidobacterium longum* subsp. *animalis* YL2, 'Turicimonas muris' YL45 and *Akkermansia muciniphila* YL44 (Oligo-MM^{12}, n=4), frozen cecum content harvested from a conventional mouse (CON; n=5) or sterile culture media as a control (cont. n=5). Afterwards, mice were kept under germ-free conditions. At 40 days post-transplantation (p.t.), mice were orally gavaged with *S.* Tm^{avir}. Feces was sampled at day 1 post-infection (p.i.) and mice were sacrificed at day 2 p.i.. Fecal microbiota composition at day 40 p.t. was determined by 16S rRNA gene amplicon-sequencing. Amplicons were processed using the QIIME pipeline and taxonomy was assigned against the SILVA database. **(B)** Alpha diversity (no. of OTUs at 97% similarity) of fecal microbiota of the different groups of mice (mean ± StD). **(C)** Relative abundance of each taxonomic group at L5 (family level; see color code). Assignment of ASF- and Oligo-MM strains to the respective families is indicated. (**D**) Variation between samples (Bray-Curtis dissimilarity) constrained by type of transplant. The percentage in brackets correspond to the variation explained by each principal coordinate to the fraction of the total variance of the data. (**E**) Single-linkage hierarchical clustering dendrogram. Distance between the points are approximately equal to taxonomic similarities of microbiota for different

transplant types. *FDR-corrected p-value obtained for PERMANOVA of Bray-Curtis dissimilarities. *S.* Tm^{avir} loads at day 1 p.i. in the feces **(F)** and at day 2 p.i. in the cecum content **(G)** and the mLN **(H)**. (**I**) Relative cecal weight (cecal weight/body weight) at day 2 p.i. Dotted line: detection limit. Bar indicates median. Statistical analysis was done using Mann Whitney U test. ns=no significant difference ($p \geq 0.05$), * $p<0.05$; ** $p<0.01$; *** $p<0.001$.

Figure 3. Clustering analysis of KEGG modules represented in draft genomes of strains from the ASF and Oligo-MM. To investigate the functional potential of each strain, genes were predicted from the assembled genomes of the 12 Oligo-MM and 8 ASF strains [17] and predicted proteins were classified to modules according to the KEGG ontology. Heat-map of hierarchical clustering of KEGG module distribution in the draft genomes. The color code indicates the presence and completeness of each KEGG module, expressed as value between 0 (module complete) and 4 (module absent). Dark green: module complete; light green: 1 block missing; yellow: 2 blocks missing; white: module absent. Prominent clusters of KEGG modules are highlighted in boxes. Examples for typical KEGG modules are indicated. An extended list of KEGG modules and clusters is shown in **Supplementary Table 7.**

Figure 4. Clustering analysis of KEGG modules represented in artificial metagenomes of ASF[8] and Oligo-MM[12] and conventional metagenomes (CON). To investigate the functional potential of the different bacterial consortia, artificial metagenomes were created by merging assembled individual genomes (OligoMM[12] and ASF[8]). In addition, published metagenomes (n=8) from conventional and wild mice (CON) were included in the analysis. Genes were predicted from the artificial and conventional metagenomes and predicted proteins were classified to modules according to the KEGG ontology. Heat map of hierarchical clustering of KEGG module distribution in the metagenomes. The color code indicates the presence and completeness of each KEGG module, expressed as value between 0 (module complete) and 4 (module absent). Dark green: module complete; light green: 1 block missing; yellow: 2 blocks missing; white: module absent. Prominent clusters of KEGG modules are highlighted in boxes. Examples for typical KEGG modules are indicated. An extended list of KEGG modules and clusters is shown in **Supplementary Table 8.**

Figure 5. Transplantation of 3 facultative anaerobic bacteria restores CR of Oligo-MM[12]. **(A)** Experimental design : Mice colonized with 'Acutalibacter muris' KB18, *Flavonifractor plautii* YL31, *Clostridium clostridioforme* YL32, *Blautia coccoides* YL58, *Clostridium innocuum* I46, *Lactobacillus reuteri* I49, *Enterococcus faecalis* KB1, 'Bacteroides caecimuris' I48, 'Muribaculum

intestinale' YL27, *Bifidobacterium longum* subsp. *animalis* YL2, 'Turicimonas muris' YL45 and *Akkermansia muciniphila* YL44 (Oligo-MM^{12}) were transplanted with 3 facultative anaerobic bacteria *E. coli* Mt1B1, *Streptococcus danieliae* ERD01G and *Staphyloccocus xylosus* 33-ERD13C (FA^3; n=5) or sterile culture media as a control (cont.; n=6). Afterwards, mice were kept under germ-free conditions. At 40 days post-transplantation (p.t.), mice were orally gavaged with *S.* Tm^{avir}. Feces was sampled at day 1 p.i. and mice were sacrificed at day 2 p.i.. *S.* Tm^{avir} loads at day 1 p.i. in the feces **(B)** and at day 2 p.i. in cecum content **(C)** and the mLN **(D)**. **(E)** Relative cecal weight at day 2 p.i.. Dotted line: detection limit. Bar indicates median. Statistical analysis was performed using Mann Whitney U test. ns=no significant difference ($p \geq 0.05$), * $p<0.05$; ** $p<0.01$; *** $p<0.001$; **(F)** Fecal microbiota composition at day 40 post-transplantation. Absolute abundance of each strain was determined by qPCR and is plotted as relative abundance of the individual strains. *below detection limit. **(G)** Distribution of *S.* Tm^{avir} and *E. coli* Mt1B1 in the gut lumen by fluorescence in situ hybridization (FISH). Tissue sections of control (n=3) and FA^3-transplanted mice (n=3) 2 days after *S.* Tm^{avir} infection were prepared and Enterobacteriaceae (*Salmonella* and *E. coli*) were stained using a group-specific FISH probe (Ent186). DNA of all bacteria was stained by SYTOX green and the rel. abundance of Ent186$^+$/all bacteria was determined by image analysis. In separate sections, *Salmonella* was detected by immunofluorescence staining: *Salmonella* were only detected in control (median rel. abundance lumen: 5.2%; epithelium: 3.0%) but not in FA^3-transplanted mice (<<0.5%). Therefore, the vast majority of Ent186$^+$ bacteria in control mice are *S.* Tm^{avir} and in FA^3-transplanted mice *E. coli*. Examples of FISH images of epithelial regions of the cecum lumen for **(H)** control and **(I)** FA^3-transplanted mice infected with *S.* Tm^{avir} for 2 days. Images are representative for 9 images from 3 mice/group, respectively. Green: SYTOX (all DNA); Red: Ent186$^+$. Scale bar: 10µm. Inset: magnification.

Figure 6. *S.* Tm infection in gnotobiotic mice colonized with different defined bacterial consortia. The graphs display a summary of data on *S.* Tm infection shown in Figure 2, Figure 5, Figure S3 and S7. Mice colonized with different defined bacterial consortia: **ASF^4**: mice colonized with ASF356, ASF360, ASF361 and ASF519 (Figure S3); **ASF^5**: mice colonized with ASF360, ASF361, ASF457, SB2 [ASF502] and ASF519 (Figure S3); **ASF^7**: ASF^4 mice transplanted with ASF356, ASF360, ASF361, ASF457, ASF500, SB2 [ASF502] and ASF519 for 40 days; **ASF^4+FA^3**: mice colonized with ASF356, ASF360, ASF361 and ASF519 transplanted with 3 facultative anaerobic bacteria *E. coli* Mt1B1, *Streptococcus danieliae* ERD01G and *Staphyloccocus xylosus* 33-ERD13C (FA^3) for 40 days; **ASF^5+Oligo-MM^{12}**: ASF^5 mice transplanted with Oligo-MM^{12} for 40 days; **Oligo-MM^{12}**: Acutalibacter muris' KB18, *Flavonifractor plautii* YL31, *Clostridium clostridioforme* YL32, *Blautia coccoides* YL58, *Clostridium innocuum* I46, *Lactobacillus reuteri* I49, *Enterococcus faecalis* KB1, 'Bacteroides caecimuris' I48, 'Muribaculum intestinale' YL27, *Bifidobacterium longum* subsp. *animalis* YL2, 'Turicimonas

679 muris' YL45 and *Akkermansia muciniphila* YL44 (Oligo-MM^{12}); **Oligo-MM^{12}+FA^{3}**: Oligo-MM^{12}
680 mice transplanted with *E. coli* Mt1B1, *Streptococcus danieliae* ERD01G and *Staphylococcus*
681 *xylosus* 33-ERD13C (FA^{3}) for 40 days. **(A)** S. Tm^{avir} loads at day 1 p.i. in the feces and **(B)** at
682 day 2 p.i. in cecum content and the mLN **(C)**. **(D)** Relative cecal weight (cecal weight/body
683 weight) at day 2 p.i.. Dotted line: detection limit. Bar indicates median. Statistical analysis was
684 done by ANOVA using Kruskal-Wallis and Dunn´s Multiple Comparison tests. Only significant
685 differences between groups are indicated in the graph: * $p<0.05$; ** $p<0.01$; *** $p<0.001$.

686 ## Acknowledgements

687 We thank Wolf-Dietrich Hardt, Jürgen Heesemann, Ombeline Rossier and members of the
688 Stecher lab for discussions. We thank Riccardo Robbiani for help regarding isolation of bacterial
689 strains and Kerstin Berer and Gurumoorthy Krishnamoorthy for providing mice. The work was
690 supported by grants from the BMBF (Medizinische Infektionsgenomik), the DFG Priority program
691 SPP1656, research grants DFG STE 1971/4-1 and STE 1971/7-1, the German Center for
692 Infection Research (DZIF), the Centre for Gastrointestinal Microbiome Research (CEGIMIR), the
693 Austrian Science Fund (P26127-B20, P27831-B28, I2320-B22) and the Vienna Science and
694 Technology Fund (LS12-001).

695 ## Author contributions

696 Conceived and designed the experiments: B.S., S.B., T.C., A.L. and D.B.; Performed the
697 experiments: B.S., S.B., M.B., D.G., D.R., M.D., S.H., Y.L., S.H., B.B., K.D.M. and D.B.;
698 Analyzed the data: B.S., S.B., M.B., C.P., D.G., H.J.R., S.H., B.B., R.P., D.H., P.M., A.C.M.,
699 T.C., A.L. and D.B.; Contributed materials/analysis tools: D.H., A.C.M., K.M., A.J.M., A.L., T.C.
700 and D.B.; B.S. coordinated the project, wrote the original draft and all authors reviewed and
701 edited the draft manuscript.

702 Correspondence and requests for materials should be addressed to B.S.

References

1. Stecher B, Berry D, Loy A. Colonization resistance and microbial ecophysiology: using gnotobiotic mouse models and single-cell technology to explore the intestinal jungle. *FEMS Microbiol Rev* 2013, **37**(5)**:** 793-829.

2. Ubeda C, Pamer EG. Antibiotics, microbiota, and immune defense. *Trends Immunol* 2012.

3. Consortium HMP. Structure, function and diversity of the healthy human microbiome. *Nature* 2012, **486**(7402)**:** 207-214.

4. Clavel T, Lagkouvardos I, Blaut M, Stecher B. The mouse gut microbiome revisited: From complex diversity to model ecosystems. *Int J Med Microbiol* 2016.

5. Smith MI, Yatsunenko T, Manary MJ, Trehan I, Mkakosya R, Cheng J*, et al.* Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. *Science* 2013, **339**(6119)**:** 548-554.

6. Linnenbrink M, Wang J, Hardouin EA, Kunzel S, Metzler D, Baines JF. The role of biogeography in shaping diversity of the intestinal microbiota in house mice. *Mol Ecol* 2013, **22**(7)**:** 1904-1916.

7. Xiao L, Feng Q, Liang S, Sonne SB, Xia Z, Qiu X*, et al.* A catalog of the mouse gut metagenome. *Nat Biotechnol* 2015, **33**(10)**:** 1103-1108.

8. Chung H, Pamp SJ, Hill JA, Surana NK, Edelman SM, Troy EB*, et al.* Gut immune maturation depends on colonization with a host-specific microbiota. *Cell* 2012, **149**(7)**:** 1578-1593.

9. Fodor AA, DeSantis TZ, Wylie KM, Badger JH, Ye Y, Hepburn T*, et al.* The "most wanted" taxa from the human microbiome for whole genome sequencing. *PLoS ONE* 2012, **7**(7)**:** e41294.

10. Dewhirst FE, Chien CC, Paster BJ, Ericson RL, Orcutt RP, Schauer DB*, et al.* Phylogeny of the defined murine microbiota: altered Schaedler flora. *Appl Environ Microbiol* 1999, **65**(8)**:** 3287-3292.

11. Stecher B, Chaffron S, Kappeli R, Hapfelmeier S, Freedrich S, Weber TC*, et al.* Like will to like: abundances of closely related species can predict susceptibility to intestinal colonization by pathogenic and commensal bacteria. *PLoS Pathog* 2010, **6**(1)**:** e1000711.

12. Kim OS, Cho YJ, Lee K, Yoon SH, Kim M, Na H*, et al.* Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Microbiol* 2012, **62**(Pt 3)**:** 716-721.

13. Lagkouvardos I, R. P, Abt B, Foesel BU, Meier-Kolthoff JP, Kumar N*, et al.* The Mouse Intestinal Bacterial Collection (miBC) provides Host-Specific Insight into Cultivable Diversity and Functional Potential of the Mouse Gut Microbiota. *Nature Microbiology* 2016, **1**.

14. Kaiser P, Diard M, Stecher B, Hardt WD. The streptomycin mouse model for Salmonella diarrhea: functional analysis of the microbiota, the pathogen's virulence factors, and the host's mucosal immune response. *Immunol Rev* 2012, **245**(1)**:** 56-83.

15. Stecher B, Robbiani R, Walker AW, Westendorf AM, Barthel M, Kremer M*, et al.* Salmonella enterica Serovar Typhimurium Exploits Inflammation to Compete with the Intestinal Microbiota. *PLoS Biol* 2007, **5**(10)**:** e244.

16. Bleich A, Hansen AK. Time to include the gut microbiota in the hygienic standardisation of laboratory rodents. *Comp Immunol Microbiol Infect Dis* 2012, **35**(2)**:** 81-92.

17. Wannemuehler MJ, Overstreet AM, Ward DV, Phillips GJ. Draft genome sequences of the altered schaedler flora, a defined bacterial community from gnotobiotic mice. *Genome Announc* 2014, **2**(2).

18. Wymore Brand M, Wannemuehler MJ, Phillips GJ, Proctor A, Overstreet AM, Jergens AE*, et al.* The Altered Schaedler Flora: Continued Applications of a Defined Murine Microbial Community. *ILAR J* 2015, **56**(2)**:** 169-178.

19. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic acids research* 2014, **42**(Database issue)**:** D199-205.

20. Syed SA, Abrams GD, Freter R. Efficiency of Various Intestinal Bacteria in Assuming Normal Functions of Enteric Flora After Association with Germ-Free Mice. *Infection and immunity* 1970, **2**(4)**:** 376-386.

21. Ng KM, Ferreyra JA, Higginbottom SK, Lynch JB, Kashyap PC, Gopinath S*, et al.* Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens. *Nature* 2013, **502**(7469)**:** 96-99.

22. Freter R, Brickner H, Botney M, Cleven D, Aranki A. Mechanisms that control bacterial populations in continuous-flow culture models of mouse large intestinal flora. *Infection and immunity* 1983, **39**(2)**:** 676-685.

23. Maier L, Vyas R, Cordova CD, Lindsay H, Schmidt TS, Brugiroux S*, et al.* Microbiota-derived hydrogen fuels salmonella typhimurium invasion of the gut ecosystem. *Cell Host Microbe* 2013, **14**(6)**:** 641-651.

24. Deriu E, Liu JZ, Pezeshki M, Edwards RA, Ochoa RJ, Contreras H*, et al.* Probiotic bacteria reduce salmonella typhimurium intestinal colonization by competing for iron. *Cell Host Microbe* 2013, **14**(1)**:** 26-37.

25. Spees AM, Wangdi T, Lopez CA, Kingsbury DD, Xavier MN, Winter SE*, et al.* Streptomycin-Induced Inflammation Enhances Escherichia coli Gut Colonization Through Nitrate Respiration. *MBio* 2013, **4**(4).

26. Nuccio SP, Baumler AJ. Comparative analysis of Salmonella genomes identifies a metabolic network for escalating growth in the inflamed gut. *MBio* 2014, **5**(2)**:** e00929-00914.

27. Rivera-Chavez F, Zhang LF, Faber F, Lopez CA, Byndloss MX, Olsan EE*, et al.* Depletion of Butyrate-Producing Clostridia from the Gut Microbiota Drives an Aerobic Luminal Expansion of Salmonella. *Cell Host Microbe* 2016, **19**(4)**:** 443-454.

28. Wells CL, Maddaus MA, Jechorek RP, Simmons RL. Role of intestinal anaerobic bacteria in colonization resistance. *Eur J Clin Microbiol Infect Dis* 1988, **7**(1)**:** 107-113.

29. Stecher B, Hapfelmeier S, Muller C, Kremer M, Stallmach T, Hardt WD. Flagella and Chemotaxis Are Required for Efficient Induction of Salmonella enterica Serovar Typhimurium Colitis in Streptomycin-Pretreated Mice. *Infection and immunity* 2004, **72**(7)**:** 4138-4150.

30. MacIntyre DL, Miyata ST, Kitaoka M, Pukatzki S. The Vibrio cholerae type VI secretion system displays antimicrobial properties. *Proc Natl Acad Sci U S A* 2010, **107**(45)**:** 19520-19524.

31. Arank A, Syed SA, Kenney EB, Freter R. Isolation of anaerobic bacteria from human gingiva and mouse cecum by means of a simplified glove box procedure. *Appl Microbiol* 1969, **17**(4)**:** 568-576.

32. Berer K, Mues M, Koutrolos M, Rasbi ZA, Boziki M, Johner C*, et al.* Commensal microbiota and myelin autoantigen cooperate to trigger autoimmune demyelination. *Nature* 2011, **479**(7374)**:** 538-541.

33. Clavel T, Charrier C, Haller D. Streptococcus danieliae sp. nov., a novel bacterium isolated from the caecum of a mouse. *Arch Microbiol* 2013, **195**(1)**:** 43-49.

34. Derrien M, Vaughan EE, Plugge CM, de Vos WM. Akkermansia muciniphila gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. *Int J Syst Evol Microbiol* 2004, **54**(Pt 5)**:** 1469-1476.

35. Hoiseth SK, Stocker BA. Aromatic-dependent *Salmonella typhimurium* are non-virulent and effective as live vaccines. *Nature* 1981, **291**(5812)**:** 238-239.

36. Hapfelmeier S, Stecher B, Barthel M, Kremer M, Müller A, Heikenwalder M*, et al.* The Salmonella Pathogenicity Island (SPI)-1 and SPI-2 Type III Secretion Systems Allow Salmonella Serovar Typhimurium to trigger Colitis via MyD88-Dependent and MyD88-Independent Mechanisms. *J Immunol* 2005, **174**(3)**:** 1675-1685.

37. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 2007, **73**(16)**:** 5261-5267.

38. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990, **215**(3)**:** 403-410.

39. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K*, et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 2006, **72**(7)**:** 5069-5072.

40. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P*, et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research* 2013, **41**(Database issue)**:** D590-596.

41. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS*, et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012, **19**(5)**:** 455-477.

42. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 2013, **29**(8)**:** 1072-1075.

43. Boisvert S, Raymond F, Godzaridis E, Laviolette F, Corbeil J. Ray Meta: scalable de novo metagenome assembly and profiling. *Genome Biol* 2012, **13**(12)**:** R122.

44. Hyatt D, LoCascio PF, Hauser LJ, Uberbacher EC. Gene and translation initiation site prediction in metagenomic sequences. *Bioinformatics* 2012, **28**(17)**:** 2223-2230.

45. Zhao Y, Tang H, Ye Y. RAPSearch2: a fast and memory-efficient protein similarity search tool for next-generation sequencing data. *Bioinformatics* 2012, **28**(1)**:** 125-126.

46. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK*, et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010, **7**(5)**:** 335-336.

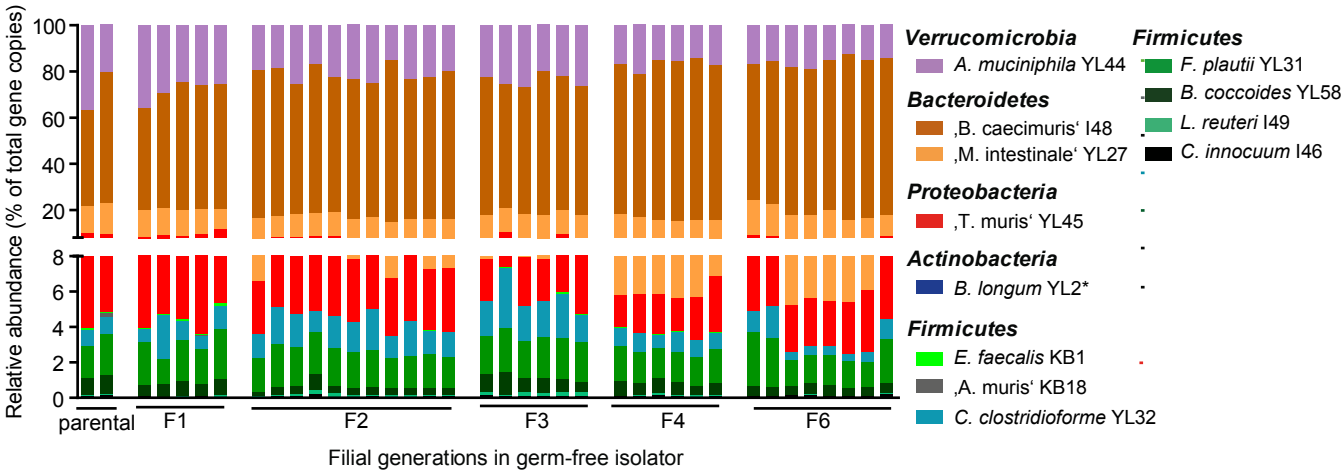
47. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M*, et al.* The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 2009, **55**(4)**:** 611-622.

48. Li H, Limenitakis JP, Fuhrer T, Geuking MB, Lawson MA, Wyss M*, et al.* The outer mucus layer hosts a distinct intestinal microbial niche. *Nat Commun* 2015, **6:** 8292.

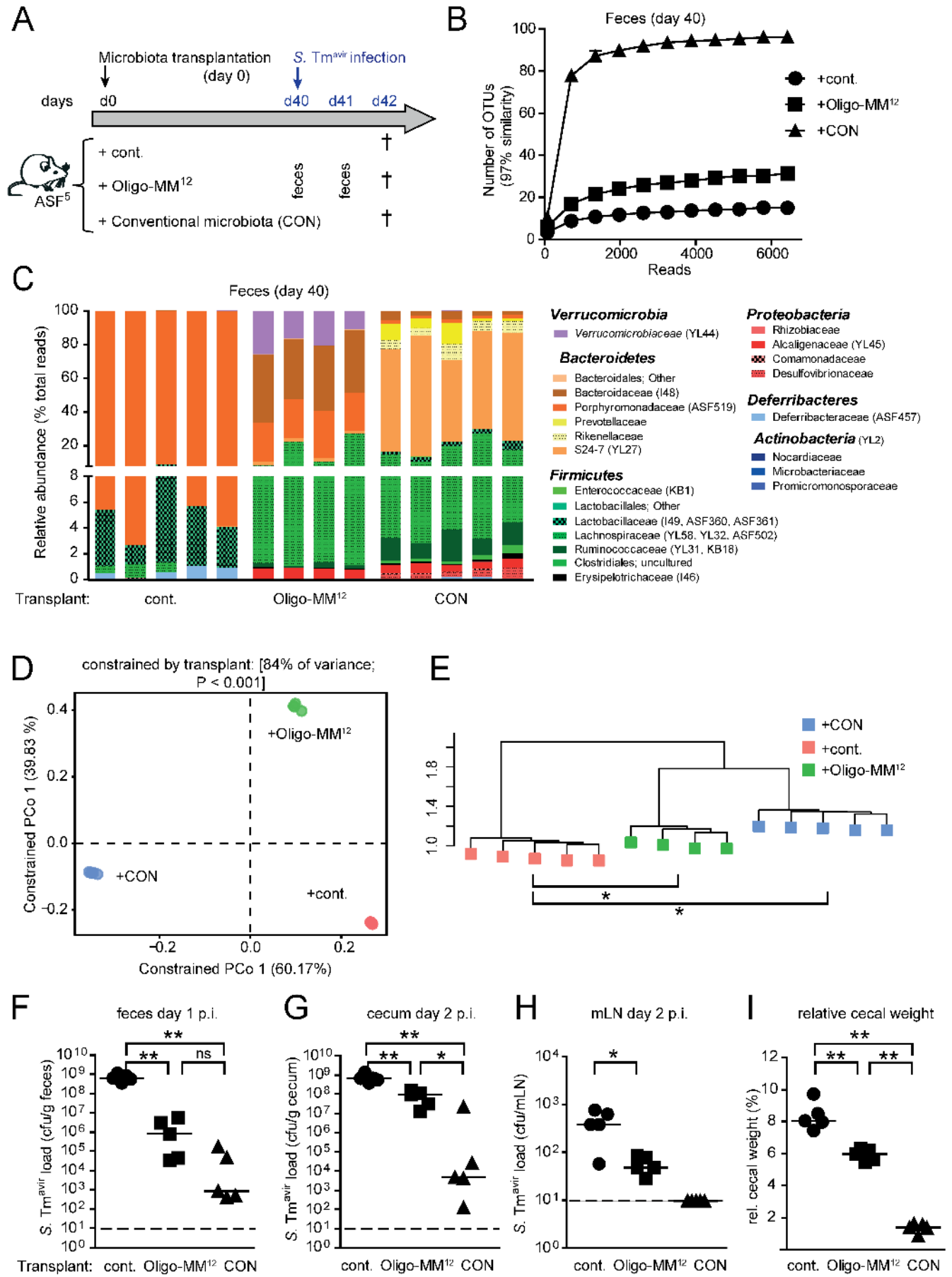
49. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 2012, **9**(7)**:** 671-675.

50. Wang J, Linnenbrink M, Kunzel S, Fernandes R, Nadeau MJ, Rosenstiel P*, et al.* Dietary history contributes to enterotype-like clustering and functional metagenomic content in the intestinal microbiome of wild mice. *Proc Natl Acad Sci U S A* 2014, **111**(26)**:** E2703-2710.

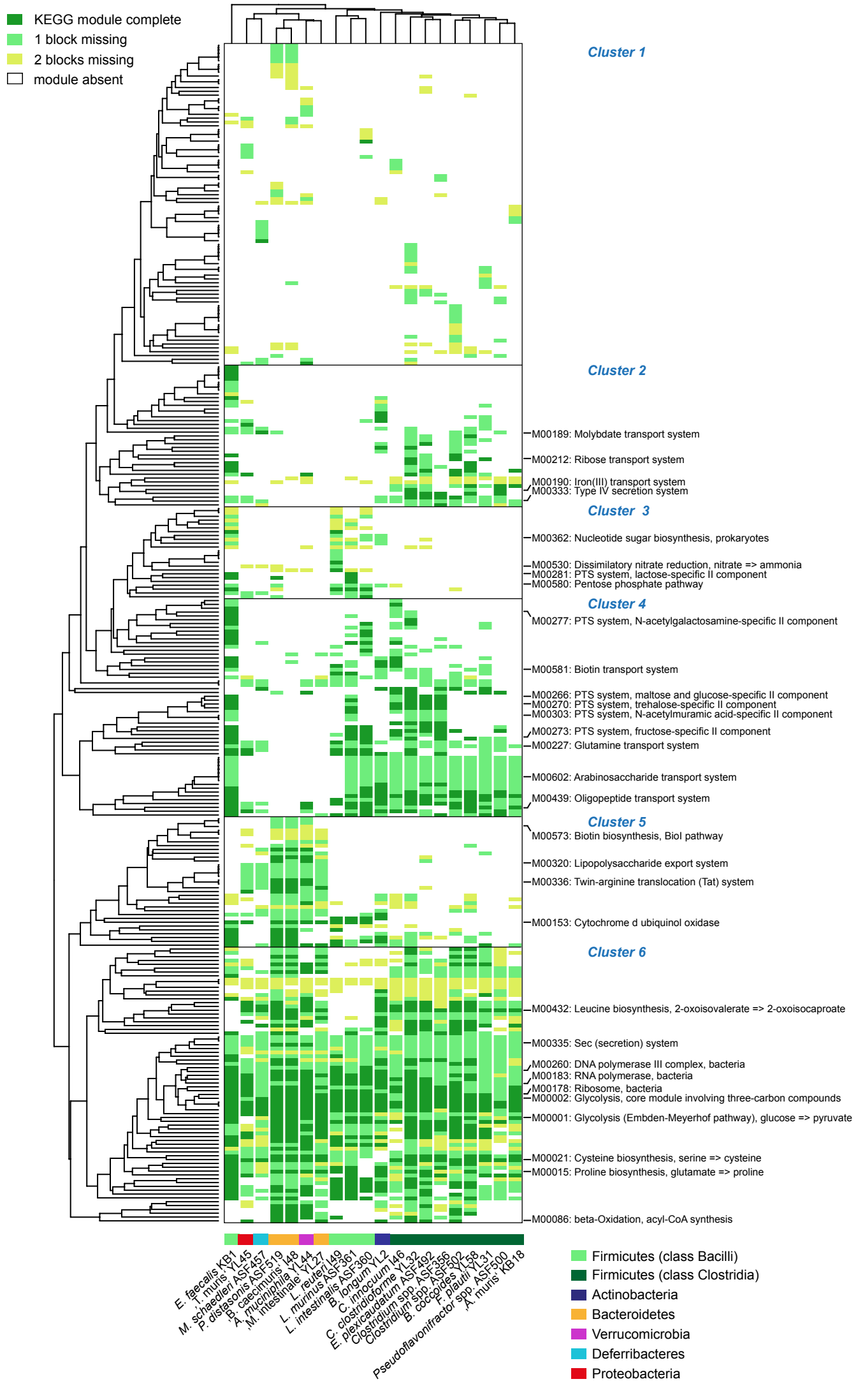
Brugiroux et *al.* Figure 1



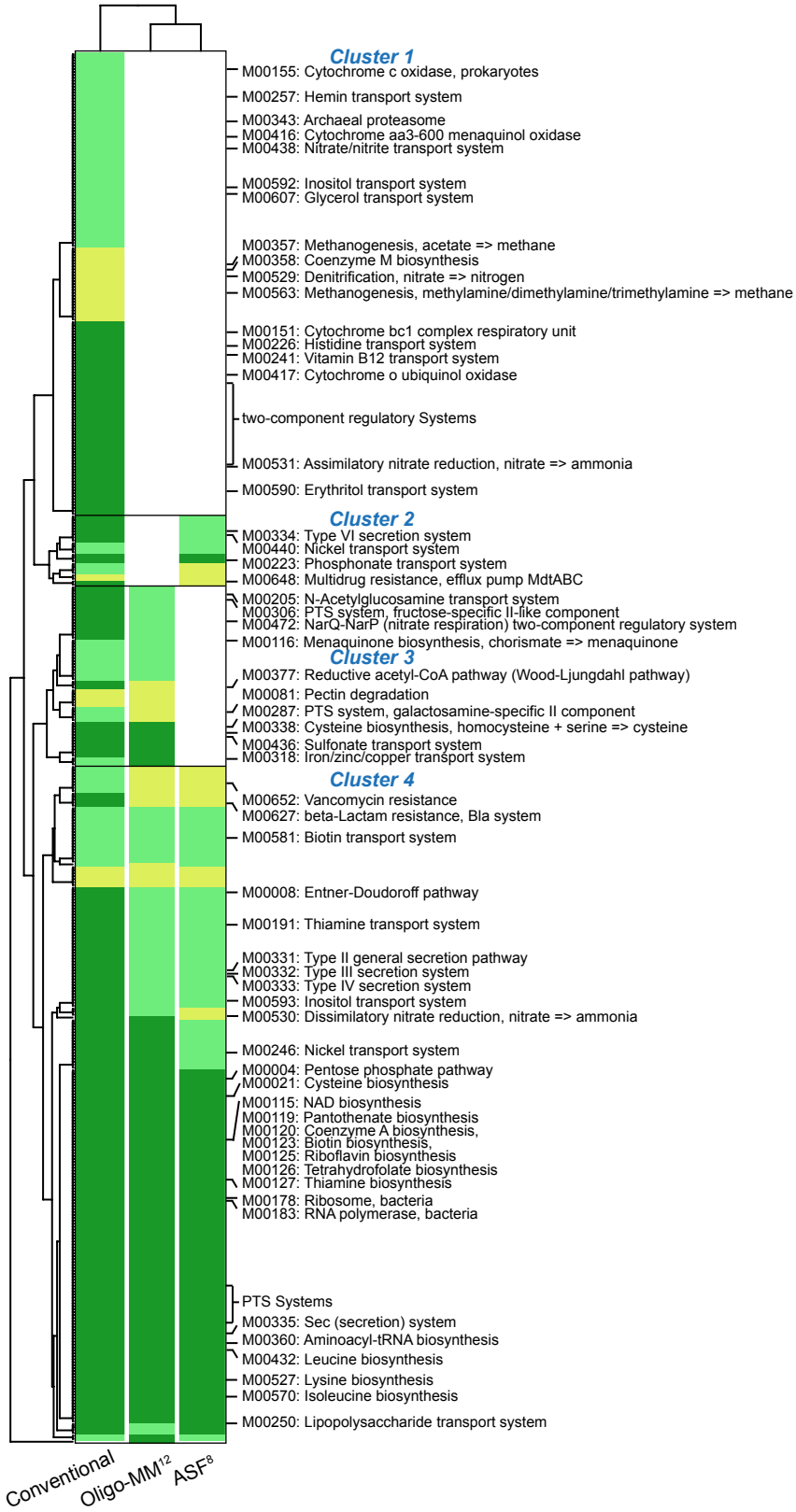
Brugiroux et *al.* Figure 2



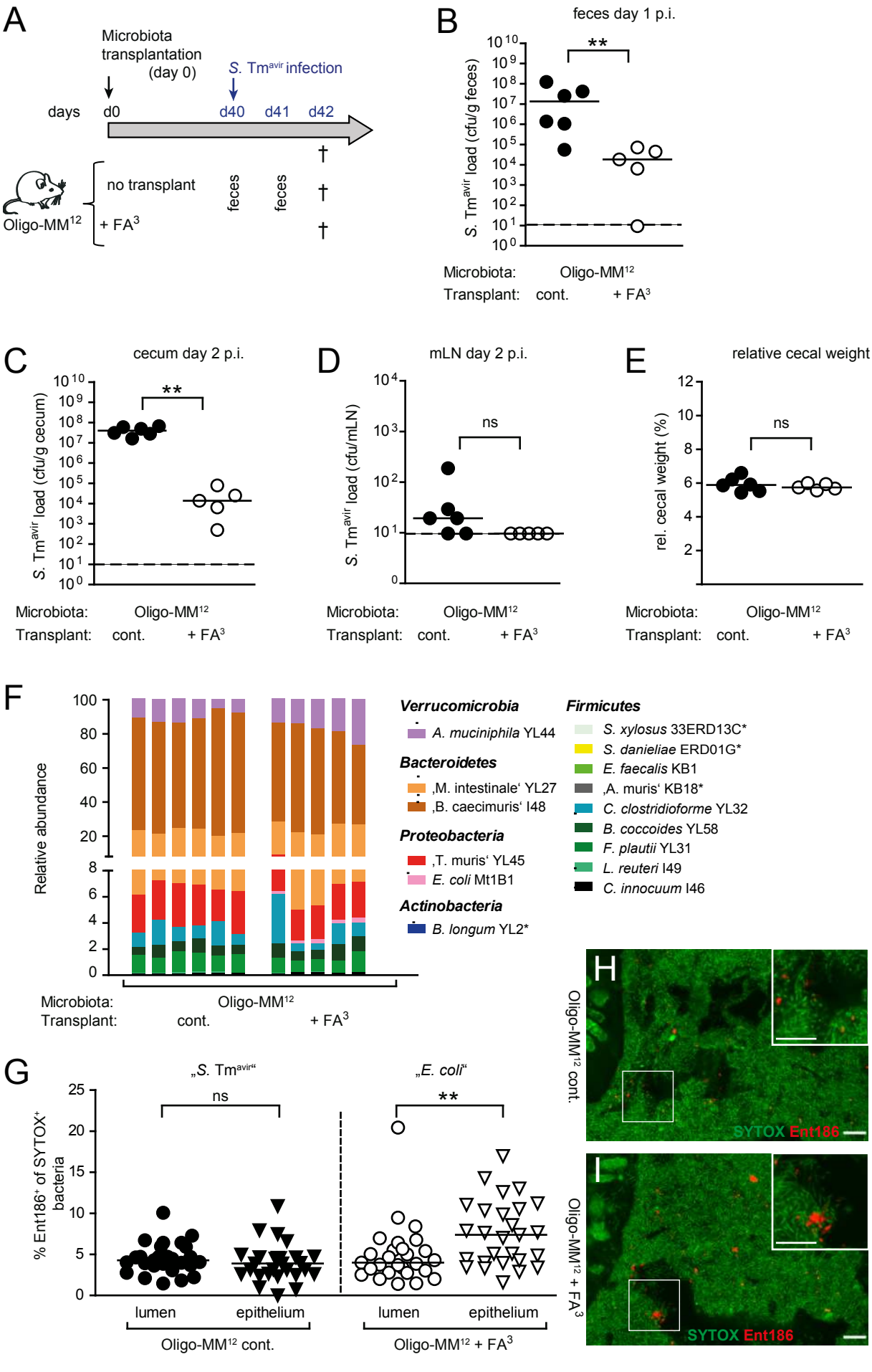
Brugiroux et *al.* Figure 3



Brugiroux et *al.* Figure 4



Brugiroux et *al.* Figure 5



Brugiroux et *al.* Figure 6