# A standardized gnotobiotic mouse model harboring a minimal 15-member mouse gut microbiota recapitulates SOPF phenotype

Marion Darnaud[1]*, Filipe De Vadder[2], Pascaline Bogeat[1], Lilia Boucinha[1], Anne-Laure Bulteau[2], Andrei Bunescu[1], Julie Chaix[1], Céline Couturier[1], Ana Delgado[1], Hélène Dugua[1], Céline Elie[1], Alban Mathieu[1], Djomangan Adama Ouattara[1], Séverine Planel[1], Adrien Saliou[1], Jennifer Yansouni[1], Martin Schwarzer[2,3], François Leulier[1,2§], Andrea Tamellini[1]*§

[1] BIOASTER, Institut de Recherche Technologique, 40 avenue Tony Garnier, 69007 Lyon, France

[2] Institut de Génomique Fonctionnelle de Lyon, Université de Lyon, Ecole Normale Supérieure de Lyon, Centre National de la Recherche Scientifique, Université Claude Bernard Lyon 1, Unité Mixte de Recherche 5242, 46 Allée d'Italie, 69364 Lyon, Cedex 07, France

[3] Laboratory of Gnotobiology, Institute of Microbiology of the Czech Academy of Sciences, Nový Hrádek, Czech Republic.

§Co-Senior authorship

*Corresponding author: gnotobiology@bioaster.org

**Keywords**: Gnotobiology; minimal mouse gut microbiota; standardization; bacterial isolation; phenotype.

## Abstract

*Mus musculus* is the classic mammalian model for biomedical research. Despite global efforts in standardizing breeding and experimental procedures, the undefined nature and inter-individual diversity of laboratory mouse microbiota remains a limitation. In an attempt to standardize preclinical studies, we have developed a simplified mouse microbiota composed of 15 strains from 7 of the 20 most prevalent bacterial families representative of the fecal microbiota found in specific opportunistic- and pathogen-free (SOPF) C57BL/6J animals and derived a standardized gnotobiotic mouse model called GM15. GM15 recapitulates extensively the functionalities found in C57BL/6J SOPF microbiota metagenome and GM15 animals are phenotypically similar to SOPF. They even perform better in a model of post-weaning malnutrition. The GM15 model ensures an increased reproducibility and robustness of preclinical studies by limiting the confounding effect of microbiota composition fluctuation and offers new possibilities for research focusing on how the microbiota shapes host physiology in health and diseases.

## Background

The intestinal microbiota is a complex and dynamic ecosystem largely composed of bacteria whose activity profoundly influences our health and diseases [1]. Advances in sequencing and analytical technologies coupled with improved computing tools have revolutionized the field of host-microbiota interaction [2]. These developments have provided an increased depth and accuracy in the study of intestinal microbial assemblages and activity for correlative studies with human health or disease traits. Despite these sophisticated descriptions of host-microbiome interaction phenomena, the underlying causal mechanisms remain largely elusive [3].

The use of model organisms plays a decisive role in the challenge to move from correlation to causal links in the host-microbiome field as they have long enabled researchers to identify the shared biological functions among living organisms, and facilitated the discovery of conserved molecular mechanisms governing the fundamental principles of biology [4]. Thanks to its genetic and physiological similarities to humans, in addition to its rapid and prolific breeding, the mouse has been a classic mammalian model of choice for the past decades for biomedical research and the host-microbiome field is no exception [5]. While the use of defined genetic backgrounds, as well as the absence of specific pathogens [6] is now common practice in mouse studies, an important confounding factor is the variability in the composition of the intestinal microbiota among experimental animals and animal facilities [7,8]. This parameter is under the influence of multiple elements such as genetics, diets, biological rhythms and breeding conditions [9]. As a consequence, to restrain microbial diversity, efforts have been made to tailor protocols for microbiota-related mice studies and to standardize mouse microbiota composition [9, 10, 11, 12, 13].

The mouse gut microbiota richness is usually estimated at more than 300 bacterial genera [5] and common inhabitants of the mouse intestine belong to seven bacterial phyla with *Firmicutes*, *Bacteroidetes* and *Proteobacteria* being the most abundant ones [12, 14, 15]. The first attempt in standardizing the mouse microbial environment (initially to study immunocompromised mouse models) arose 15 years ago with the wide implementation of the

Specific Pathogen-Free (SPF) hygienic status of mouse husbandries [6] by re-deriving mouse strains by two-cell-stage embryo transfer to SPF recipients and subsequent post-natal inoculation with a cocktail of bacteria devoid of pathogens to homogenize microbial colonization among a given animal facility. SPF, and then SOPF inbred lines (lacking Specific Opportunist Pathogens, such as *Staphylococcus aureus* or *Pseudomonas aeruginosa*) now represents the common standard health status for experimental mouse breeding. However, despite the global efforts in standardizing the SPF procedures, the undefined nature and important inter-individual diversity of the SPF microbiota remains a limitation in host-microbiome studies, since the scientific community still lacks a common SPF standard cocktail and rather use a facility-specific cocktail of bacteria [9]. Indeed, the microbiota fluctuates a lot with diet and environment, so it is strictly impossible to have the exact same microbiota of SPF mice in two different facilities.

Microbial cultivation and gnotobiology offer attractive strategies to standardize the microbiota of mouse models. Germ-Free (GF) animals (i.e. animals devoid of any living micro-organisms) are the originators of gnotobiotic animals (i.e. animals with a controlled microbiota) obtained by colonization with pure culture or cocktails of bacterial strains [16]. Recent efforts have been put into isolating, cultivating and archiving isolated culture of the dominant members of the mouse microbiota [17]. Gnotobiotic animals can be kept in isolators for several generations and offer the possibility of a strict control of their microbial status. Gnotobiotic models offering different degree of microbial complexity have been developed in the past ranging from mono-colonization (monoxenic animals) to high diversity microbiota models such as conventionalized ex-GF animals using a donor microbiota [9, 12, 18]. Two models have emerged for breeding and long-term experimental purpose: the Altered Schaedler Flora (ASF) model and the recent oligo-Mouse-Microbiota$_{12}$ (oligoMM$_{12}$) model [19, 20]. These models offer an enlarged microbial potential as compared to monoxenic mice, while keeping the model simple and experimentally tractable as compared to conventionalized animals (SPF or SOPF).

The ASF was developed in the late 70s by adding bacterial strains which better represented the microbiota of conventional mice to the initial Schaedler flora, a minimal

microbial consortium that protected ex-GF mice from opportunistic pathogen colonization during breeding [21]. The ASF is composed of 8 defined bacteria, which are stable over mouse generations. Mice colonized with ASF display a certain degree of immune normalization when compared with SPF mice [19]. Limitations of the ASF model are that the strains are not publicly available of the strains and that they are not all representative of the dominant members of the mouse microbiota [9, 12]. In addition, ASF mice differ substantially from SPF mice with respect to microbiota-associated traits beyond immune parameters and resistance to opportunistic pathogen colonization, probably owing to the limited phylogenic diversity and metabolic capabilities of the ASF consortium [9, 12, 22]. Recently, the oligoMM$_{12}$ model was developed [20]. It is a minimal microbiota gnotobiotic model composed of 12 defined cultivable mouse commensal bacteria from the miBC collection representing members of the major bacterial phyla of the mouse gut [17, 23, 24]. The community is transmissible and stable over consecutive mouse generations and unlike ASF offers colonization resistance against *Salmonella enterica* serovar Typhimurium albeit not to the degree of a conventional complex microbiota [20]. However, how the oligoMM$_{12}$ model performs relative to the SPF for microbiota-associated traits beyond *S.* Typhimurium colonization resistance remains elusive.

In an attempt to standardize preclinical studies in the host microbiome field, we have developed a simplified mouse microbiota that is representative of SOPF microbiota at the functional level and derived a standardized gnotobiotic mouse model called GM15, which phenotypically mimics SOPF mice under standard diet conditions. We show that, under conditions of chronic physiological stress such as post-weaning malnutrition on a low-protein diet, a dietary condition triggering stunting, GM15 microbiota seems superior to an SOPF microbiota to buffer the deleterious effect of a depleted diet on mouse juvenile growth.

## Results and Discussion

### *In silico* identification of the main bacterial families of the C57BL/6J SOPF fecal microbiota

To define a minimal microbiota containing representative and prevalent bacteria from the gut of C57BL/6J SOPF mice, we analyzed the composition of fecal pellets from 4 C57BL/6J SOPF mice (two females and two males) by whole-genome sequencing. An average of 13.4 million of paired-end reads was obtained per sample with a length of 300 bp. The metagenomics data sets generated were classified using the Centrifuge software [25] and compared to the RefSeq complete genome database [26], and 20 dominant families consistently present in all mice were identified (Fig. 1a). The profiling of metagenomic sequencing data pointed out a comparable distribution of bacterial families among the four tested C57BL/6J SOPF mice. Moreover, bacterial species identification was possible for genome sequences with good phylogenetic resolution and already referenced in taxonomy databases. Interestingly, among the identified species, *Bacteroides acidifaciens, Clostridium cocleatum, Lactobacillus johnsonii, Lactobacillus murinus* and *Lactobacillus reuteri* were previously identified as mouse-enriched and dominant intestinal bacteria [17] and *Lactobacillus murinus*, *Parabacteroides goldsteinii, Clostridium* ASF 356 and *Clostridium* ASF 502 were already part of the ASF model [19].

**Isolation and taxonomic characterization of GM15 bacterial strains**

We established four different strategies in order to isolate and culture a maximal number of representative strains of the 20 dominant bacteria families identified in our metagenomic sequencing analysis (Fig. 1b). First, we isolated the most prevalent strains from fecal pellets of C57BL/6J SOPF mice using non-selective agar media. Then, we used antibiotic selection to isolate resistant strains. We also used rumen enrichment to isolate strains from caecal content. Finally, we used fecal pellets of ASF mice to isolate additional strains. We obtained a collection of approximately 400 cultivable bacterial isolates. All isolates were pre-screened by MALDI-TOF MS for dereplication prior to the first taxonomic identification by 16S rRNA gene Sanger sequencing. We selected eleven strains covering seven of the most representative and prevalent families of the intestinal microbiota of C57Bl/6J SOPF mice and obtained four additional strains from the DSMZ miBC collection to establish the GM15 consortium that covers

most of the dominant bacteria families found in C57Bl/6J SOPF animals (Fig.1a,b). In summary, GM15 is composed of two strains of Bacteroidaceae, one strain of Tannerellaceae, six strains of Lachnospiraceae, three strains of Lactobacillaceae, one strain of Erysipelotrichaceae, one strain of Ruminococcaceae, and one strain of Enterobacteriaceae (Additional file 1: Figure S1). The genomes of the fifteen strains were sequenced and the draft genomes sequences aligned against the NCBI database [27], which allowed the identification of 14 strains at the species level and 1 strain at the family level.

**_In silico_ functional metagenomics analysis of the GM15 strains**

To gain insights into the functionalities encoded in the individual genomes of the GM15 members, the coding sequences of the 15 strains were converted into their respective protein sequences, which were annotated for clustering into KO (KEGG Orthology) groups. By merging the 15 assembled individual genomes, we found that the GM15 metagenome possesses 3890 non-redundant KO groups. Besides, all GM15 strains possessed 3 to 64 unique functions, although *E. coli* Mt1B1 exhibited a vast repertoire of unique KO groups ($10_3$), indicating that *E. coli* Mt1B1 is responsible for the GM15's functional metagenomic profile at 33%, while the 14 other strains contribute all together at 8% and the 59% remaining are associated to non-unique KO groups (Fig. 1c). It is noteworthy that KEGG module analysis is biased towards gene sets, pathways and functional groups of well-characterized bacteria such as *E. coli*, which is by far the most studied bacterial species to date.

Next, we highlighted the *in silico* functionalities of the 15 selected strains associated with known enzymatic activities in the gut [28] (Fig. 1d). Again *E. coli* Mt1B1 is a major contributor, but each functionality is also covered by other strains at equivalent or lower levels. As expected, the enzymatic activities in the gut are correlated with the phylogenetic membership of the strains. For example, Lactobacilli, which are a major part of the lactic acid bacteria group, the principal contributor to lactate dehydrogenase [29], clustered together. Additionally, Lachnospiraceae clustered with Ruminococcaeae and Erysipelotrichaceae, which include mainly bacteria with sporulation capabilities [30, 31, 32]. Finally, Bacteroides

clustered with Parabacteroides, whose species are predominant in the colonic mucus barrier and promote mucinase activity [33], and are significant producers of succinate, a major metabolic by-product [34, 35]. Thus, different strains of a bacteria family and of other closely related bacteria families are capable of the same enzymatic activities in the gut. This can be essential for the generation of simplified non-specific gnotobiotic models.

Then, we determined the functional coverage of the GM15 metagenome (i.e. the sum of the genomes of the 15 strains) relative to the KEGG modules of C57BL/6J SOPF mouse microbiota found in our initial metagenomic analysis (Fig.1a). In addition, the KEGG modules from the OligoMM$_{12}$ and ASF microbiota were included for comparative analysis, as these consortia were previously used to generate gnotobiotic mice with a stable and defined mouse-derived microbiota [23, 36]. The presence and completeness of KEGG modules was determined for each metagenome and used for hierarchical clustering (Fig. 1e). One cluster contained highly conserved modules in all mouse models (Fig. 1e, cluster 4; Additional file 3: Table S2), including 140 pathways. We also identified clusters of modules that were not represented among the ASF and OligoMM$_{12}$ consortia, but specifically common to GM15 and SOPF (Fig. 1e, clusters 9 and 6; Additional file 3: Table S2), comprising a total of 155 pathways indicating that qualitatively the GM15 metagenome covers functionalities found in SOPF microbiota that were lacking in ASF and OligoMM$_{12}$ models. Quantitatively, the defined consortia of GM15, OligoMM$_{12}$ and ASF covered respectively 72, 54 and 48% of the KEGG modules of the C57BL/6J SOPF microbiome suggesting a superior functional potential of the GM15 community as compared to OligoMM$_{12}$ and ASF models. Thus, taken collectively our *in silico* analysis suggests that the GM15 community carries a significant potential for enzymatic activities in the gut and recapitulates widely the functionalities found in C57BL/6J SOPF murine metagenome.

**Monitoring and stability assessment of GM15 gut microbiota**

To explore *in vivo* the functional potential of the GM15 community, we next generated GM15 isobiotic animals. First, we investigated whether the GM15 consortium can stably colonize the

mouse intestine over several generations. To this end, we developed a strain-specific quantitative polymerase chain reaction (qPCR) microfluidic assay, which allows simultaneous absolute quantification of the 15 strains, along with the global bacterial load in a given biological sample (Additional file 2: Table S1). Fecal samples from GF and SOPF mice were used as negative and positive controls. Only *C. clostridioforme* YL32, which was obtained from the DSMZ miBC collection, was not detected in our SOPF mice. Co-monitoring of specific and total bacteria aimed to detect any bacterial load imbalance caused by a contamination in gnotobiotic isolators.

Eight-week-old GF mice (five breeding pairs; GM15 founders, F0; Additional file 4: Figure S2) were inoculated by oral gavage with the frozen mixtures of the 15 strains and bred in sterilized positive pressure isolators up to the F4 filial generation. All strains except *Lachnospiraceae* bacterium COE1, *Lachnospiraceae* sp. MD329, and Lachnospiraceae sp. MD308 were above the detection limit of our qPCR microfluidic assay in the fecal samples of most individual mice (GM15 founders and progenies from the 4 consecutive generations; Additional file 5: Figure S3a; Fig. 2a). *Anaerotruncus colihominis* JM4-15, *Clostridium* ASF 356, *Clostridium* ASF 502, and *Clostridium cocleatum* I50 were occasionally below detection limit. The detection of strains was reproducible between fecal and caecal samples from individual mice of the second filial generation (Additional file 5: Figure S3b). These results indicate a stable colonization and effective vertical transmission of at least 12 strains out of the 15 inoculated. Based on the results of our qPCR assay, we consider that the 3 remaining strains: *Lachnospiraceae* bacterium COE1, *Lachnospiraceae* sp. MD329, and *Lachnospiraceae* sp. MD308, either did not efficiently colonize the animals, live in the caecum or the colon below the detection limit of our qPCR assay, or predominantly live in other gastrointestinal niches than those sampled.

Next, we evaluated the effect of aging on the GM15 community by following individual mice of the first filial generation between 2 and 12 months of age (Additional file 4: Figure S2). Overall, no changes in the qualitative and quantitative composition of the GM15 consortium were detected (Fig. 2b). Then, we asked if the composition of the GM15 community was

modulated by substituting the breeding diet by an alternative maintenance diet. This alternative diet is quasi isocaloric but its nutritional composition differs by 1.3 and 1.6-fold fewer proteins and lipids respectively, and 1.2-fold more carbohydrates than the breeding diet. An alternative diet was administrated to 8-week-old GM15 mice for 4 weeks. We collected fecal samples before diet change, after 2 and 4 weeks, and again after 4 weeks back to the breeding diet (Additional file 4: Figure S2). We did not detect any significant changes in the GM15 composition under these conditions (Fig. 2c). Additionally, successful fecal microbiota transplantation (FMT) from GM15 founders to GF mice was confirmed by strain-specific qPCR of the ex-GF mice' feces collected at week 1, 2 and 3 post-transplantation (Additional file 4: Figure S2a; Fig. 2d).

When all the data are analyzed together (Fig. 2e), we notice a limited fluctuation (maximum 1Log10-fold change) of each member of the GM15 gut microbiota's load among the conditions tested (Fig. 2e). Therefore, we conclude that the GM15 community is stable upon adult colonization, among filial generations, during aging, upon mild dietary fluctuations and can be transmitted efficiently by FMT.

**The GM15 microbial community recapitulates SOPF macroscopic phenotype**

GF and published gnotobiotic mice display anatomical alterations compared to SOPF mice, such as enlarged caecum, along with physiological and metabolic differences [20, 37]. To phenotypically assess the gnotobiotic GM15 model, we designed a comparative study to evaluate the steady state macroscopic, immune, metabolic and endocrine phenotypes of GF, SOPF and GM15 mice. In addition, all phenotyping was achieved across two generations, F1 and F2, to strengthen data analysis.

Initially, we evaluated the reproduction performance of the GM15 model by recording the period from mating to offspring delivery (Fig. 3a), the number of pups per litter (Fig. 3b), and the perinatal mortality (Fig. 3c). GM15 mice behaved like SOPF mice with the exception of one less progeny per mean litter. Then, we quantified the food intake relative to body weight after weaning at 4 weeks of age and observed no significant difference between the three

groups despite a marked increased variation among GF animals which is not detected in GM15 and SOPF animals (Fig. 3d). Next, we studied post-natal growth parameters. Male and female GM15 animals gained weight (Fig. 3e) and size (Fig. 3f) like SOPF mice, although the growth curves from the two sexes differ. We then studied internal organ size. As expected, the characteristic caecum enlargement seen in GF animals was reduced in GM15 mice (Fig. 3g). The weights of GM15 and SOPF brain, liver and spleen were equivalent and larger than those of GF mice (Fig. 3g). Bone size was also identical in GM15 and SOPF mice and larger than that of GF mice (Fig. 3h). Taken together, these results confirm that the gut microbiota is associated with somatic tissue growth [38, 39], and that the GM15 simplified microbiota is sufficient to largely recapitulate the breeding and growth performance of SOPF mice by compensating the physiological limitations of GF mice.

**The GM15 simplified microbiota partially restores SOPF immune phenotype**

It is now well established that host-specific bacteria consortia influence intestinal and systemic immune maturation [40, 41]. We thus profiled the basal immune parameters of GM15 animals and compared them to SOPF and GF animals at 7-8 weeks of age. We observed that the GM15 community induces the production of IgA in the gut and in serum at levels equivalent to those detected in SOPF mice (Fig. 4a). Notably, the highest levels of circulating IgA were not correlated with the highest levels of fecal IgA, and intragroup variability was not related to gut microbiota composition, which was homogeneous between individuals. Additionally, IgG2b serum levels were equivalent in GM15 and SOPF mice compared to GF mice, IgG3 levels were increased in GM15 mice compared to GF and SOPF mice, and IgM levels were slightly increased in GM15 mice compared to SOPF (Fig. 4a), which was consistent with published comparative analyses carried out with SPF mice even though IgG3 and IgM levels were lower in our SOPF mice [42, 43]. On the other part, no difference of circulating IgG1 was observed between GF, GM15 and SOPF, and all IgG2a levels were below levels of detection (data not shown). Surprisingly, GM15 mice presented elevated levels of IgE in serum compared to GF and SOPF mice (Fig. 4a). High IgE levels may result from parasitic infections, or

immunodeficiencies and long-lived IgE producing plasma cells generated by food antigens in mice with low-diversity microbiota during early life [42, 44, 45]. Independent tests following FELASA guidelines (CR EU RADS, France) rejected the infection hypothesis of the GM15 mice, which were negative for ectoparasites and endoparasites, respiratory and intestinal specific pathogenic bacteria and infectious agents, as well as viruses. It is also known that Clostridia-abundant juvenile mice have lower IgE levels [46]. Of note, our GF mice do not show increased IgE levels as SOPF, and GM15 mice harbor a similar load of Clostridia as SOPF despite elevated levels of IgE. Therefore, we speculate that specific members of the GM15 consortium are inducers of the observed high IgE titer.

Next, we assayed the circulating levels of key intestinal cytokines IL-17a and IL-22, whose production is induced by commensal bacterial via TLR2 activation in the gut mucosa [47], and linked to local inflammation [48, 49] or associated with anti-microbial responses and control of bacterial colonization [50, 51]. GM15 microbiota induced release of both cytokines compared to GF mice (Fig. 4b) and in a similar manner to SOPF mice for IL-22. To note, serum IL-17a concentrations in our SOPF mice detected by Luminex were slightly lower than those published for SOPF animals in previous studies using Luminex detection, but do not explain the difference observed between GM15 and SOPF mice levels [52]. We also measured IL-1b, IL-10, IL-12p70, and IL-17f, which are cytokines involved in autoimmune and inflammatory bowel diseases, but we did not detect any difference between the three groups of mice.

In parallel, we assessed viable leukocytes by measuring CD45+ cell counts and analyzed T cells, B cells, NK cells and monocytes or dendritic cells (DC) populations in whole blood and several lymphoid organs (spleen, thymus, Peyer's patches (PP) and mesenteric lymph nodes (MLN)). It has been previously shown that isolated lymphoid structures such as PP and associated cellularity were strongly increased with microbiota diversity whereas the total cell numbers in MLN were comparable between GF and SPF mice [53]. As described in literature, the number of PP collected where higher in SOPF mice compared to GF and partially restored in GM15 mice (data not shown), and no or minor difference were observed between the 3 groups for CD45+ cells count that showed slight increased levels in SOPF compared to

GF mice with apparent intermediary levels in GM15 mice in spleen, thymus and PP, and no difference in MLN (Fig. 4c). ]. In 2018, Kennedy *et al.* published an overview of the literature describing the main immune cell populations modifications observed in GF mice in different organs [54]. In our study, very few immune cell populations seemed to be impacted by flora composition. No difference between the 3 groups was observed for MLN DC, circulating and splenic CD4+ and CD8+ T cells and splenic B cells as described in literature (data not shown). Minor differences were observed for PP CD4+ T cells and splenic DC between GM15 mice and SOPF but not with GF mice whereas both populations were described to be decreased in GF mice (Additional file 6: Fig. S4b/e). Only PP NK cells showed similar and increased levels in GM15 and GF mice compared to SOPF mice (Additional file 6: Fig. S4e). Interestingly, the great majority of populations impacted showed similar levels in GM15 mice and SOPF compared to GF mice (Additional file 6: Fig. S4a-e). Circulating monocytes, MLN CD8+ T cells and splenic NK cells were increased in GM15 and SOPF compared to GF as described in literature for circulating monocytes and MLN CD8+ T cells and though splenic NK cells were not described to be modified in GF mice. Thymic CD4+ T cells, circulating and MLN B cells were decreased in GM15 and SOPF mice compared to GF though circulating B cells were described to be decreased in GF mice.

Taken collectively, our results indicate that the GM15 community is sufficient to dampen the characteristic immunodeficiencies of GF animals and restore an immunophenotype close to that detected in SOPF animals with a peculiar signature of the GM15 model with high IL17a and IgE levels in sera. To note, no correlation was observed between those two markers in our study although positive or negative correlations have been both described in literature [55, 56, 57, 58]. Furthermore, as for IgE, IL-17a levels modulations have been described to be associated to particular flora compositions [42, 59, 60, 61, 62, 63, 64, 65, 66, 67]. Such an observation paves the way to further investigation to decipher which and how members of the GM15 community trigger these immunophenotypes.

**Low-complexity GM15 microbiota shares more metabolic traits with SOPF than GF mice**

348 The gut microbiota influences multiple host metabolic pathways by providing metabolites to its
349 host and also shapes inter-organ communication within the body by influencing the production
350 and activity of endocrine signals [68, 69]. One-dimensional proton nuclear magnetic resonance
351 spectrometry ($_1$H NMR) was previously applied to investigate how the gut microbiota impacts
352 host metabolism using mouse models [70, 71], or human cohorts [72]. Using this technology,
353 we analyzed the metabolic profile of plasma samples from GF, GM15 and SOPF mice and
354 were able to quantify a total of 51 polar metabolites and 5 non-polar metabolites whose levels
355 were not significantly affected by sex (Additional file 7: Table S3; Additional file 8: Table S4).
356 A principal component analysis based on quantified polar metabolites showed sequential
357 alignment on first principal component (8.8% of total variance) according to microbiota
358 complexity of GF, GM15 and SOPF samples, independently of F1 and F2 generations (Fig.
359 5a). The metabolic composition of GF plasma resulted only from the host metabolic activity
360 and represented a basal phenotype. At the other end, the SOPF mice harboring a very diverse
361 gut microbiota, exhibited a larger panel of metabolic activities based on host-bacterial and
362 bacterial-bacterial interactions. In-between, the GM15 low-complexity community contributed
363 in a lower extent to the metabolic phenotype. We performed the equivalent statistical analysis
364 using binned NMR spectra instead of quantified metabolites and as expected we obtained the
365 same results (Additional file 9: Figure S5). Next, we performed a discriminant analysis across
366 all samples to highlight the specific metabolic signatures of each group. Eight metabolites
367 emphasized significant variation, although it is noteworthy that the calculated distance placed
368 GM15 closer to SOPF than to GF (Fig. 5b). As previously reported, GF mice had higher plasma
369 levels of glycine [73] and reduced plasma acetate concentration [74]. On the contrary, SOPF
370 mice harboring a diverse microbiota had higher plasma levels of acetate, acetoacetate and
371 dimethyl sulfone [75]. As for GM15 mice, less methanol and more citrate, ethanol and pi-
372 methylhistidine were detected. Interestingly, methanol may occur as a result of fermentation
373 by gut bacteria and can stimulate citric acid fermentation [76, 77]. Thus, the simplified gut
374 microbiota of GM15 mice may produce less methanol and/or microbially-produced methanol
375 may be used to form citrate. On the other part, ethanol is produced by some intestinal bacteria,

such as *E. coli* and lactic acid bacteria, which are well represented in the GM15 community [78]. These specific metabolic signatures may be used in the future as a panel of biomarkers to identify the GM15 model. Besides, the analysis of the plasma non-polar metabolites indicated that GM15 colonization was sufficient to reduce free cholesterol and phosphatidylcholine as observed in SOPF mice (Fig. 5c). The three additional non-polar metabolites detected were equivalent in all mice (Additional file 8: Table S4).

Finally, we investigated the circulating levels of key metabolic hormones and growth factors. To allow blood sample mutualization for different analyses, mice were not fasted. Ghrelin was below limit of detection for all groups, and glucagon, insulin and leptin did not show any difference between GM15 and GF mice, whereas levels detected in SOPF mice were similar to those previously published under basal conditions [79, 80, 81] (data not shown). However, as compared to levels detected in GF animals the GM15 gut microbiota was able to restore the levels of circulating insulin growth factor 1 (IGF-1) as seen in SOPF animals. IGF-1 is an essential growth factor promoting systemic and tissue growth [82, 83] and this observation correlates well with the improved macroscopic growth of GM15 mice compared to GF animals (Fig. 5d, Fig. 3e-h). Corticosterone levels, which are high in GF and low in SOPF animals were also normalized by GM15 bacterial colonization [84] (Fig. 5e), indicating that GM15 community seems as efficient as a complex SOPF microbiota at utilizing host metabolites and promoting steroidogenesis and growth factor production.

Collectively, our results reveal that GM15 animals stand out from GF mice and recapitulate many of the SOPF metabolic features even though some differences exist.

**GM15 mice are less sensitive to diet-induced stunting than SOPF mice**

Previous work by our lab and others [85, 86, 38] has shown that the gut microbiota influences pathogenesis associated to chronic undernutrition, particularly diet-induced stunting. Consequently, we sought to study our newly-established gnotobiotic mouse model under a severe nutritional stress, induced by a macronutrient-depleted diet (containing 4% protein and

2% lipids) expected to trigger stunting. We thus fed male mice a breeding diet (BD) or a depleted diet (DD) from post-natal day 21 (i.e. the day of weaning) until post-natal day 56.

As shown in Fig. 6a-b and similarly to our previous observation (Fig. 3e-f) GM15 and SOPF mice grew well on the BD as they show similar body weight and size gains. Next, we confirmed that the DD triggered an almost full stunting of both the juvenile GM15 and SOPF mice, characterized by the flattering of their weight and size curves (Fig. 6a-b). However, the GM15 mice performed slightly better in terms of growth as both their body weight and size were significantly less impacted than SOPF animals by the DD. While the caecum of GM15 was enlarged (Fig. 6c, Additional file 10: Fig. S6a), this variation could not account for the total weight difference between GM15 and SOPF animals. We thus compared the sizes and weights of nine other organs (Fig. 6d-f, Additional file 10: Fig. S6b-g) in order to account for these variations and using two-way ANOVA, we confirmed that diet was the main driver of the growth phenotype. However, we found that the increase in body size observed in GM15 mice on the DD could be correlated to a significant increase in the size of the tibia (Fig. 6d). Although we did not find any other significant difference in other single organ size or weight between GM15 and SOPF animals on DD, we observed a clear tendency of an increase in GM15 compared to SOPF animals for most parameters tested (Fig. 6e-f, Additional file 10: Fig. S6b-g). We tested this tendency by integrating all the phenotypical parameters in a Principal Component Analysis (PCA). We first established which parameters were correlated to the phenotype (Additional file 10: Fig. S6h), thus excluding brown adipose tissue from the analysis as a non-correlating parameter. The PCA revealed that, under BD, the GM15 phenotype was part of the spectrum of the SOPF phenotype (Fig. 6g). However, under DD, there is a clear shift between the GM15 and SOPF phenotypical space (Fig 6g, dotted lines). Our results thus indicate that under nutritional stress, GM15 microbiota buffers diet-induced stunting slightly more effectively than an SOPF microbiota. As IGF1 is an important driver of the diet and microbiota mediated growth promotion [38], we assessed IGF1 levels in animals and while the levels massively drop in the DD conditions, we did not detect any difference between GM15 and SOPF mice at day 56 on the DD (Fig. 6h). However, we cannot exclude that slight variation in size between

GM15 and SOPF animals on DD results from a differential secretion of IGF1 before day 56, when the growth rate of the animal is maximal (Fig. 6a-b).

Taken together, our data show that animals bearing the simplified GM15 microbial community perform similarly than animal bearing a complex SOPF microbiota under non stressful conditions, however the GM15 microbiota seem partially protective against the deleterious effects of chronic undernutrition, as compared to SOPF mice.

## Conclusions and perspectives

Here we describe GM15, a simplified and controlled murine gut microbiota and its related GM15 C57BL/6J gnotobiotic mouse model. The GM15 community is composed of pure cultures of 15 strains from 7 of the 20 most prevalent bacteria family present in the fecal microbiota of SOPF C57BL/6J mice. GM15 carries significant potential for enzymatic activities in the gut and recapitulates extensively the functionalities found in C57BL/6J SOPF gut microbiome. *In vivo* GM15 is stable upon adult colonization for up to 12 months, during natural transmission among 4 filial generations, upon mild dietary fluctuations and can be transmitted efficiently by FMT. GM15 compensates the steady state breeding, growth, immune, endocrinal and metabolic limitations of GF mice and recapitulates many of the SOPF phenotypical features. Taken together our results establish that GM15 is a novel controlled preclinical model phenotypically similar to SOPF with the potential to ensure an increased reproducibility and robustness of preclinical studies by limiting the confounding effect of microbiota composition fluctuation. Importantly, the reduced microbial complexity of the GM15 community, the tractability of its members and the control offered to the experimenter on its composition allow easy quantification of gut microbiota dynamics. Indeed, upon manipulation of the GM15 community composition, the correlation of such dynamics with fluctuating host traits allows the establishment of causal relationships between specific microbiota members and host traits. GM15 model offers exciting perspectives for improvement. Indeed, under stressful nutritional environment, the simplified GM15 microbiota performs slightly better than a complex SOPF community in terms of physiological host response. We have previously identified a

*Lactobacillus* strain that is capable of buffering the deleterious effects of such challenge in monocolonized mice [38]. Interestingly, out of its 15 strains, GM15 microbiota contains 3 *Lactobacillus* strains. Further genetic manipulations coupled to gnotobiotic studies focusing on modifying the Lactobacilli components of GM15 will pave the way to understanding how this minimal bacterial community buffers the host response to chronic undernutrition.

As all experimental models GM15 has its limitations. First, phenotypically, despite being macroscopically similar, GM15 and SOPF animals differ in specific immune and metabolic signatures. This is not particularly surprising given the marked reduced microbial diversity that the GM15 model carries compare to SOPF microbiota (15 strains vs hundreds of species in SOPF animals). These specific immune and metabolic signatures of the GM15 models pave the way to further studies aiming at defining if the presence or absence of specific community members triggers these phenotypes. Another limitation worth mentioning here is that the GM15 model is a standardized model, and such standardization, essential for establishing robust causal relationships between a microbiota configuration and a host trait, may decrease the translational potential of the observation.

In addition, microbiota-mediated resistance to *Salmonella* infection may be tested in GM15 animals and compared to ASF and OligoMM$^{12}$ animals as recently done by Stecher and colleagues [20]. Second, the GM15 model, by its low complexity, offers the possibility to use it as a template for establishing further complex or complete consortia, e.g. by complementing it with representative strains of the missing prevalent bacteria family found in SOPF microbiota such as Deferribacteraceae, Oscillospiraceae, Clostridiaceae and Eubacteriaceae. Despite these limitations and room for improvement, our study establishes that the GM15 model offers new possibilities for preclinical research focusing on host-microbes and microbe-microbe interactions, and how the microbiota shapes the environmental impact on health and diseases or drug efficacy.

## Methods

### Bacterial strains isolation and identification

Fresh caecal contents and fecal pellets of C57BL/6J SOPF mice (Charles River Lab., France) were resuspended (1/10 wt/vol) in reduced broth media for direct dilution plating on agar plates and growth at 37°C under anaerobic atmosphere (90% $N_2$, 5% $H_2$, 5% $CO_2$). *Lactobacillus johnsonii* NCC 533 was isolated on MRS agar. *Lactobacillus murinus* 313 and *Parabacteroides goldsteinii* WAL 12034 were isolated on Colombia CNA agar with 5% sheep blood. *Bacteroides acidifaciens* A40 and *Lachnospiraceae* sp. MD308 were isolated on GAM agar. *Bacteroides caecimuris* I48 and *Lactobacillus reuteri* I49 were isolated on GAM agar supplemented respectively with 32 µg/mL vancomycin and 32 µg/mL erythromycin. *Lachnospiraceae* bacterium COE1 and *Lachnospiraceae* sp. MD329 were isolated on M2GSC agar (modified Hobson, containing (per 100 mL) 1 g of casitone, 0.25 g of yeast extract, 0.4 g of NaHCO3, 0.2 g of glucose, 0.2 g of cellobiose, 0.2 g of soluble starch, 30 mL of clarified rumen fluid, 0.1 g of cysteine, 0.045 g of K2HPO4, 0.045 g of KH2PO4, 0.09 g of (NH4)2SO4, 0.09 g of NaCl, 0.009 g of MgSO4.7H2O, 0.009 g of CaCl2, 0.1 mg of resazurin, and 1.5 g agar). *Lachnospiraceae* bacterium COE1, *Lachnospiraceae* sp. MD329, and *Lachnospiraceae* sp. MD308 were isolated from caecal contents, the rest from fecal pellets. Fecal pellets of ASF mice (Taconic, USA) were cryopreserved at -80°C, and then resuspended in reduced broth media for direct FMT in GF mice. Fresh caecal content and fecal pellets were resuspended in reduced broth media for direct dilution plating on agar plates and growth at 37°C under anaerobic atmosphere (90% $N_2$, 5% $H_2$, 5% $CO_2$). *Clostridium* ASF356 and *Clostridium* ASF502 were isolated on M2GSC agar, respectively from caecal content and fecal pellets. For identification of isolates, colonies were pre-screened for dereplication by MALDI-TOF MS (Vitek MS, Biomérieux) according to the manufacturer's instructions and database enrichment using RUO mode. Then, gDNA was extracted from pure cultures and analyzed by 16S rRNA gene sequencing (GATC Biotech). Following Edgar's recommendation [87], a full-length 16S rRNA sequence similarity ≥99% using either NCBI blast [27], Ribosomal Database Project [88], or EzTaxon [89] programs allowed the identification of 13 isolates at the species level, and

515 isolates MD329 and MD308 could only be assigned to the *Lachnospiraceae* family. A more
516 precise annotation could be given for the 2 isolates *Clostridium* ASF356 and *Clostridium*
517 ASF502 obtained from the defined ASF microbial consortium. *Anaerotruncus colihominis* JM4-
518 15 (DSM-28734), *Clostridium clostridioforme* YL32 (DSM-26114), *Clostridium cocleatum* I50
519 (DSM-1551), *Escherichia coli* Mt1B1 (DSM-28618) were obtained from DSMZ.
520
521 **Culture conditions**
522 Freshly grown cultures of individual bacterial strains were supplemented with 20% glycerol
523 prior cryopreservation at -80°C. Each culture was systematically validated for culture purity
524 and identity by MALDI-TOF. Culture media and material were introduced in the anaerobic
525 chamber at least 2 days before use. Anaerobic bacterial strains were grown in GAM, except
526 *Clostridium* ASF502, *Lachnospiraceae* sp. MD308, *Lachnospiraceae bacterium* COE1 and
527 *Lachnospiraceae* sp. MD329 in M2GSC, and *A. colihominis* JM4-15 in Bifidobacterium
528 medium. For mouse colonization and absolute quantification of bacteria, a single colony of
529 each of the 15 bacterial strains was grown and amplified at 37°C. Each bacterial pellet was
530 resuspended in medium, 1 mL was cryopreserved with 10% glycerol, 1 mL was centrifuged
531 and bacterial pellet was stored at -20°C for gDNA extraction, and the rest was used for
532 numeration by dilution plating on agar plates. A frozen mixture of the GM15 bacterial
533 community containing the 15 individual strains at equivalent concentration (6.67E+06 CFU)
534 was prepared to enable easy inoculation.
535
536 **WGS**
537 DNA samples from the 15 bacterial cultures were prepared for whole-genome sequencing,
538 using the Nextera XT DNA library preparation kit (Illumina, Cat.Nos. FC-XXX, California, USA)
539 according to manufacturer's recommendations. The resulting libraries were checked for their
540 quality using the High-sensitivity DNA chip using the Agilent 2100 Bioanalyzer (Waldbroon,
541 Germany) and quantified using the QuantiFluor One dsDNA kit (Promega). Paired-end

(2x300bp) sequencing was performed on a MiSeq sequencer using the "MiSeq v3 kit (600 cycles)" (Illumina, California, USA).

### *De novo* genome assembly

After a quality control with FastQC [90], the paired-end reads were assembled *de novo* using the "A5-miseq" assembly pipeline [91], comprising the following steps: adapter trimming, quality trimming and filtering, error correction, contigging and scaffolding. The 15 *de novo* assemblies resulted in draft genomes composed of few scaffolds (from 30 to 268) with high N50 values (from 13 099 to 943 892). Genomes were then ordered using Mauve [92] and annoted with PGAP of the NCBI database. Default parameters were used for all software tools.

### Taxonomic annotation

WGS generated data have been trimmed and quality controlled by AfterQC software [93]. A k-mer counting strategy with the Centrifuge software [25] have been privileged to infer taxonomy, and reads were confronted to the RefSeq complete genome database [26], with bacteria, archaea and viruses domains and the mouse representative genome (taxid 10090), in order to estimate the amount of host DNA contamination and remove it from sequenced data.

### Genomic functional analysis

Genes were predicted and turned into protein sequences using Prodigal [94]. Marker genes were searched using the HMM3 package [95]. Predicted proteins sequences of genomes were blasted against the KEGG microbial database [96] which contains 13 millions of proteins sequences and trimmed with following parameters: best-hit with an expected value threshold $< 10^{-5}$. The matrix obtained was consolidated into KEGG orthologs count (KO, which represents a set of genes the have sequence similarities and do the same function), into KEGG modules (which represent short enzymatic pathways, involving few proteins and doing a targeted function), and into KEGG pathways (i.e. large metabolic pathways). KO were analyzed for their presence absence among genomes. The modules were analyzed for their

completion (4 levels: Full, lack 1 enzymes, lack 2 enzymes, absent), and only modules with a score with 3 or 4 were presented and integrated for their KO count. The KEGG pathways were analyzed for their number of related KO count affiliated to them. A list of functions of interest have been designed and their presence among genomes have been analyzed in details (Additional file 3: Table S2). Because pathways and functions of interest did not have the same number of KO of interest and a different distribution among the genomes, functional data were normalized among each function in order to obtain values that can be comparable. For each function/pathway, the number of different KO was normalized by the total number of KO retrieved. Data were then log-transformed +1. Clusterization of both functions and communities was performed using Euclidian distance and ward's method, and a kmeans clustering was performed in order to define the community clusters.

**Identification of specific regions for primers design**

NUCmer, a part of the MUMmer package [97], was used to perform pairwise alignment of the 15 genomes. Based on these alignments, PanCake [98] was used with default parameters to identify specific regions of each genome. Specific regions with a length of 200 bp were extracted, meeting the following criteria: GC content between 48 and 52%, distance to a border of a scaffold higher than 300bp, unique in the draft genome. The specificity of each 200bp region was double-checked with BLAST [27] on the 15-genome database and on the NCBI *nr* database. The design of primers on the specific regions was performed by Fluidigm. The primer specificity was checked with BLAST.

**Animal experiment**

All mice were bred according standardized procedures in the gnotobiology unit of BIOASTER at the ANSES animal facility (Lyon, France), housed in sterilized positive pressure isolators (Noroit) under a 12h light/dark cycle at 22±2°C and 50±30% of humidity, and fed ad libitum with irradiated R03-40 diet (3395 kcal/kg, 25.2% kcal proteins, 61.3% kcal carbohydrates, 13.5% kcal lipids; Safe) and autoclaved water. Irradiated corn-cob granules (Safe) were used

as bedding. Sterile enrichment was provided in all cages and was constituted by cotton rolls as nesting material, poplar bricks and a polycarbonate red mouse igloo (Safe). Nesting material and poplar bricks were renewed every two weeks. All breeders were mated by trio (2 females and 1 male) between 8 weeks and 6 months of age, and all mice were weaned at 4 weeks after birth. C57BL/6J GF mice were obtained by aseptic hysterectomy of a C57BL/6J SOPF female (Charles River Lab.), and neonates were fostered on C3H GF mothers (CDTA). Axenic status was assessed weekly by gram staining and cultures of fecal suspension on solid and liquid media. GM15 founders were 8-week-old C57BL/6J GF mice colonized by oral inoculations of 215 µL of the GM15 bacterial community, twice at 48h interval. GM15 microbiota composition was analyzed by qPCR microfluidic assay from feces collected at 6-week-old. Alternative diet R04-40 (3339 kcal/kg, 19.3% kcal proteins, 72.4% kcal carbohydrates, 8.4% kcal lipids; Safe) was given at 8-week-old GM15 mice for 4 weeks. FMT was done by inoculating orally a suspension of fresh fecal pellets to 7-week-old C57BL/6J GF mice, twice at 48h interval. For undernutrition experiments, GM15 and SOPF mice were bred and randomly assigned to be given either the above R03-40 diet or a custom-made low-protein diet (3500 kcal/kg, 4.7% kcal proteins, 90.1% kcal carbohydrates, 5.3% kcal lipids, Envigo) for 5 weeks after weaning. Mice were killed by cervical dislocation and biocollections were performed aseptically. All animal procedures were approved by ANSES/ENVA/UPEC ethic committee (APAFIS#4529-2016022616404045v3; APAFIS#785-2015042819315178v2; APAFIS#18918-2019020118003843v3) and were conducted in accordance with the National and European legislation on protection of animals used for scientific purposes.

**gDNA extraction from caecal contents and fecal pellets**

Caecal and fecal gDNA were extracted using the DNeasy® PowerLyzer® PowerSoil® Kit (Qiagen) following the manufacturer's instructions with modifications. Samples (approximately 0.1g) were heat-treated at 65°C for 10 min, and 95°C for 10 min, before a double bead beating at 30 Hz for 5 min. 50µL of DNA were obtained with two consecutive elutions.

626 **Quantitative PCR microfluidic assay**

627 In order to quantify specific and global bacteria load per g of caecal or fecal samples (wet
628 weight), qPCR microfluidic assay was performed using respectively specific primers for GM15
629 and “universal” primers amplifying the genes encoding 16S rRNA from most bacteria groups
630 [99]. Amplicons generated using these primers range between 60 and 99 base pairs. qPCR
631 microfluidic assays were conducted in 48.48 Dynamic Array™ IFCs for EvaGreen Fast Gene
632 Expression on a Biomark HD (Fluidigm) according to the manufacturer’s instructions, with
633 cDNA diluted 100-fold, preamplified with pooled primers, and diluted again 100-fold. Each IFC
634 included triplicate reactions per DNA sample, standards and negative control. Standards were
635 generated by serial dilution of a gDNA extract from pure bacterial cultures of known
636 concentration. The efficiency of each qPCR reaction was calculated based on the slope of
637 standard curves and within the range of 78-107%. An equivolume mixture of standards was
638 used to normalize data between runs. Due to the different individual detection limits of the
639 qPCR assay for each primer, the detection limit for GM15 ranged between 2.73 x $10_2$ and 6.57
640 x $10_5$ CFU/g (Additional file 2: Table S1).
641
642 **Sample preparation for immunophenotyping**
643 For flow cytometry analyses whole blood was collected on EDTA tube. Spleen, thymus, MLN
644 and PP were collected in RPMI (Gibco). Single-cell suspensions were achieved using a 100
645 µm cell-stainer (Becton Dickinson) and a 5 mL syringe plunger in RPMI supplemented with 2%
646 heat-inactivated fetal bovine serum (Sigma) and 100 µg/mL DNASE1 (Roche). Cells were then
647 spun at 400xg for 5 min at room temperature. Medium was discarded and cells were washed
648 using 5 mL of supplemented medium. For whole blood, spleen and thymus samples, red blood
649 cells were lysed by resuspension in 1 mL PharmLyse 1X (Becton Dickinson) for 10 min. Cells
650 were then spun at 400xg for 5 min at room temperature. Lysing solution was discarded and
651 cells were washed using 2 mL of PBS (Gibco). Cells were pelleted a second time and
652 resuspended in PBS supplemented with 2% heat-inactivated fetal bovine serum (Sigma).
653 Numeration and viability were determined using Propidium Iodide marker exclusion and

MACSQUANT Flow cytometer (Miltenyi). Cells were then resuspended to a working concentration of $10_6$ cells / tube for organs and 100 µL / tube for whole blood, and analyzed by flow cytometry.

Whole blood was collected on dry Eppendorf tube for sera analysis. Sera were obtained by centrifugation 2000xg for 15 min at 4°C and stored at -20°C before Luminex and ELISA analyses. Feces were collected in Eppendorf low-binding tubes and stored at -80°C before Luminex analysis. Feces supernatant was obtained by disrupting 100 mg feces in 1mL PBS-Protease Inhibitor 1X (Sigma) using Lysing Matrix E Tube (MP Biomedicals) and Fast Prep homogenizer (MP Biomedicals). Samples were spun at 8000xg for 15 min at 4°C and supernatants were collected for IgA Luminex analysis.

**Flow Cytometry**

$10_6$ cells or totality of cells for some of PP samples were stained for surface markers. Leukocytes were stained using anti-CD45 Viogreen (Miltenyi, clone30F11) and were checked for viability using Zombie Green Fixable Viability (Biolegend). T cells were stained using anti-CD3 BV421 (BD, clone 145-2C11), anti-CD4 PE (Miltenyi, clone REA604) and anti-CD8a PE Vio615 (Miltenyi, clone REA601). B cells were stained using anti-CD45 RB220 PE Vio770 (Miltenyi, clone RA3-6B2). NK cells were stained using anti-CD335 APC (NKp46) (Miltenyi, clone 29A1.4.9). Monocytes were stained using anti-CD11b APC Vio770 (Miltenyi, clone REA592) in spleen, thymus and whole blood samples. Dendritic cells were stained using anti-CD11c APC Vio770 (Miltenyi, clone REA754) in spleen, PP and MLN samples. Cells were analyzed using a MACSQuant Ten Flow cytometer (Miltenyi) and raw data were analyzed using FlowJo software (Tree Star, Becton Dickinson). For CD45+ cell count comparison in PP, results were normalized to $10_6$ cells using number of cells stained for each mouse when less than $10_6$ viable cells were isolated from PP. Data normalization was not possible for whole blood sample, thus CD45+ cell count comparison has not been performed for this compartment. For frequency results, data are represented as a percentage of CD45+ cells for all organs.

682

683 **Metabolic Panel, Ig and Cytokines Luminex**

684 Serum concentrations of Metabolic Panel (Ghrelin, Glucagon, Insulin and Leptin) were
685 determined using the Mouse Metabolic Magnetic Bead Panel Milliplex MAP kit (Millipore).
686 Samples were not diluted, and assay was performed according to the manufacturer's
687 instructions. Serum concentrations of Ig Panel (IgA, IgG1, IgG2a, IgG2b; IgG3, IgM, IgE) and
688 IgA in feces supernatant were determined using the Mouse Immunoglobulin Isotyping
689 Magnetic Bead Panel Milliplex MAP kit (Millipore) and Mouse IgE Single Plex Magnetic Bead
690 Milliplex MAP kit (Millipore). Samples were diluted 1:12500 (IgA, IgG1, IgG2a, IgG2b; IgG3,
691 IgM), 1:100 (IgE), and 1:100 (IgA in feces supernatant). Assays were performed according to
692 the manufacturer's instructions. Serum concentrations of Cytokines Panel (IL-1b, IL-10,
693 IL12p70, IL-17a, IL-17f, IL-22) were determined using the Mouse Th17 Magnetic Bead Panel
694 Milliplex MAP kit (Millipore). Samples were no diluted and assay was performed according to
695 the manufacturer' instructions. Samples were assayed in monoplicate.

696

697 **Corticosterone and IGF-1 ELISAs**

698 Serum concentrations of Corticosterone were determined using the Corticosterone ELISA kit
699 (Abnova). Samples were diluted 1:50, and assay was performed according to the
700 manufacturer' instructions. Serum concentrations of IGF-1 were determined using the
701 Mouse/Rat IGF-1 Quantikine ELISA kit (R&D Systems). Samples were diluted 1:500, and
702 assay was performed according to the manufacturer' instructions. Samples were assayed in
703 monoplicate.

704

705 **Sample preparation for metabophenotyping**

706 It is known that the quantification of some plasma metabolites, such as tryptophan and
707 tyrosine, is biased since they bind to albumin, a highly abundant protein in plasma [100]. In
708 addition, protein precipitation methods with organic solvents can induce loss of volatile
709 metabolites and overlay of residual broad resonances of lipids with some polar metabolites

[101]. Thus, the polar and non-polar metabolites extraction from the same plasma sample were prepared with some in-house modifications to previously described sequential approaches [102, 103]. Briefly, we deproteinized the plasma samples by acidified ultrafiltration in order to increase desorption yields of aromatic amino acids and then, quantify the polar metabolites. We also extracted the filter residue of the same sample to analyze the non-polar metabolites.

**Polar metabolites preparation**

Frozen mice plasma samples are thawed in thermoshaker Eppendorf (10 min, 10°C and 1000 rpm). The entire amount of plasma about 180 µL is filtered using 0.2 µm centrifugal tube from VWR (5 min, 10°C, 10000xg). Next, the lipoproteins removal was performed by mixing 150 µL of filtered plasma with 50 µL milli-Q water and 10 µL deuterated formic acid (2.1% in milli-Q water) on a clean 10 kDa cut-off ultracentrifugation tube (VWR) using thermoshaker Eppendorf (10 min, 10°C, 750 rpm) and then centrifuged (30 min, 10°C, 10000xg, soft ramp). Vortexed 10 kDa-filtred plasma samples (135 µL) were transferred in 0.5 mL 96-wellplate Agilent and mixed with phosphate buffer solution (45 µL, 0.6 M, pH=7.4) containing internal standard DSS-d4 (1.54 mM), pH-reference standard DFTMP (4 mM) and $D_2O$ (40%) on thermoshaker Eppendorf (1.5 min, 10°C, 650 rpm). Finally, the resulting sample solutions (155 µL) were transferred in 3 mm SampleJet NMR tubes. The DSS-d4 concentration was calibrated by $_1$H NMR using sodium succinate dibasic hexahydrate standard solution. This protocol was systematically applied to prepare blank and quality control samples. The concentration of the formic acid was optimized to allow desorption of the metabolites like tryptophan from plasma albumin. Deuterated formic acid instead of protonated form was used to decrease the exogenous NMR signal and allow the quantification of the endogenous formiate. All ultracentrifugation filters are previously recovered from tubes and washed separately from residual glycerol in a 250 mL Duran bottle, plunged in milli-Q water and then sonicated in a bath for 10 min. This washing procedure is repeated five times. After the last step of the washing the ultracentrifugation filters could be stored in water for at least 3 months. Just before use traces of milli-Q water from ultracentrifugation filters are dispersed outside with nitrogen

stream. We observed that this procedure of washing removes better the residual glycerol than supplier's protocol and gives higher yields of ultrafiltred plasma.

**Non-polar metabolites extraction after ultrafiltration**

The lipoproteins on the 10 kDa filters was further diluted with 150 µL phosphate buffer solution (1 M, pH=7.4) and mixed on thermoshaker Eppendorf (10 min, 10°C, 750 rpm). The lipoproteins samples were then transferred in clean Eppendorf tubes and extracted with 400 µL methanol-dichlormethan (1:2 v/v). Samples were centrifuged (5 min, 10°C, 10000xg) for better phase separation. The dichlormethan layer was transferred in clean Eppendorf tube and the aqueous phase was extracted again with dichlormethan. The pulled organic phase was evaporated under nitrogen stream. The dry lipidic residue was dissolved with 200 µL deuterated chloroform containing 0.03% TMS internal standard and 155 µL of resulting solution was transferred in 3 mm SampleJet NMR tube. The TMS concentration was calibrated by NMR using 1,3,5-tritertbutylbenzen standard solution and it was found to be 0.435 mM.

**Metabolites analysis and quantification**

The 1D $_1$H NMR spectra were acquired at 298 K with 600 MHz Ascend (Avance III HD) spectrometer from Bruker equipped with a 5 mm QCI cryoprobe. All samples were stored at 6°C in the SampleJet autosampler. Polar metabolites were analyzed using *noesygppr1d* pulse sequence. For each spectrum, 128 scans were collected into 32k data points within 14 ppm spectroscopic width and a recycling delay of 4 s. The mixing time was calibrated to 50 ms and the acquisition time was 3.9 s. The non-polar metabolites NMR spectra were acquired using *zg30* pulse sequence. The spectra were recording using 256 scans, into 32k data point and a spectroscopic width of 20 ppm. The relaxation delay was 4 s. The FIDs were zero-filled to 64k points and Fourier transformed using 0.3 Hz exponential line-broadening function. All spectra were aligned to DSS-d4 and TMS respectively internal standard. The concentrations of the polar and non-polar metabolites were quantified using Chenomx NMR suite 8.1. The Chenomx software was applied also for spectra binning of $10^{-3}$ ppm width for each bin. The triacyl glycerol

(TAG), phosphatidylcholine (PC), lysophosphatidylcholine (LysoPC), sphingomyelin (SM), free cholesterol (FC) and cholesterol ester (EC) quantification an in-house lipid database within Chenomx was built with authentic lipid standards. The lipids database spectra were recorded using the same parameters as described above for non-polar metabolites.

**Metabolomics data analysis**

Discriminant analyses were performed using the partial least square (PLS-DA) algorithm to extract metabolomics signatures [104, 105]. A variable selection algorithm based on Elastic-Net was used to improve model performance by selecting most significant metabolic signatures that explain the groups (GF, GM15, SOPF). The statistical performances of the regression models were assessed using the balanced error rate with and without cross-validation (E2 and CV-E2) and permutation tests. Permutation tests consisted in building the null distribution of the balanced error rate E2 by randomly permuting observations. Regression models were thus challenged by testing if the cross-validation error rate CV-E2 is significantly different from the null distribution with a p-value $< 0.05$. Metabolites involved in the cross-validated signatures were ranked by order of importance in the PLS-DA model using their VIP (variable importance in projection) scores. The biplot allows projecting onto the two first components, the samples and the metabolites that significantly discriminate each sample groups. Metabolites that are positively correlated (or positively contribute) to a sample group will point to the direction of this group. They will point to the reverse direction for a negative correlation.

**Statistics**

Reproductive performance and body growth were analyzed respectively by 1-way and 2-way ANOVA. Phenotyping data impacted by age, filial generation or sex, were analyzed by the F-test for multiple linear regressions (R v3.4.2), otherwise by 1-way ANOVA and Tukey's multiple comparison parametric test or Kruskal-Wallis and Dunn's multiple comparison non-parametric test after D'Agostino et Pearson test for data set normality assessment (GraphPad Prism v7

and v8). PCA for phenotyping analysis was performed using R v3.4.2 and the ade4 package [106].

## Additional files

**Additional file 1: Figure S1.** Morphological features of the GM15 strains. (TIFF 2312 kb)

**Additional file 2: Table S1.** GM15 strains-specific primers. (XLSX 11 kb)

**Additional file 3: Table S2.** KEGG clusters. (XLSX 142 kb)

**Additional file 4: Figure S2.** *In vivo* experimental design. (TIFF 591 kb)

**Additional file 5: Figure S3.** Assessment of gut microbiota stability of GM15 founders, and of reproducibility between fecal and caecal samples of individual mice. (TIFF 799 kb)

**Additional file 6: Figure S4.** Immune cell populations profiling. (TIFF 471 kb)

**Additional file 7: Table S3.** Polar metabolites. (XLSX 18 kb)

**Additional file 8: Table S4.** Non-polar metabolites. (XLSX 16 kb)

**Additional file 9: Figure S5.** PCA based on binned NMR spectra. (TIFF 1374 kb)

**Additional file 10: Figure S6.** Measurements of several organs after nutritional challenge. (TIFF 1011 kb)

## Abbreviations

ASF: Altered Schaedler flora; BD: Breeding diet; bp: Base pair; CFU: colony forming units; CNA: Colistin and naladixic acid; DD: Depleted diet; FMT: Fecal microbiota transplantation; GAM: Gifu anaerobic medium; gDNA: Genomic deoxyribonucleic acid; GF: Germ-free; IFC: Integrated fluidic circuit; KO: KEGG orthology; MLN: mesenteric lymph nodes; MRS: Man Rogosa Sharpe medium; OTU: Operational taxonomic unit; PCA: Principal coordinate analysis; PP: Peyer's patches; NMR: Nuclear magnetic resonance; qPCR: Quantitative polymerase chain reaction; SEM: Standard error of the mean; SOPF: Specific opportunistic pathogen-free; WGS: Whole genome sequencing

**Acknowledgments**

We thank Julie Henry, Leanne Goncalves, Christelle Boisse, Gustavo Stadthagen Gomez, Vincent Thomas and Gianfranco Grompone for help with bacterial cultures and gram staining, and discussions.

**Funding**

This work was supported by the French Government (PIA). Research in F.L lab is supported by ENS de Lyon, CNRS, the FINOVI foundation and the “Fondation pour la Recherche Médicale” (« Equipe FRM DEQ20180339196).

**Availability of data and materials**

The 15 assembled genomes and the corresponding sequencing reads have been deposited in the DDBJ/ENA/GenBank data banks and the Sequence Read Archive respectively, under the BioProject number PRJNA551571 (GenBank accession no.: VIRE00000000, VIRD00000000, VIRC00000000, VIRB00000000, VIRA00000000, VIQZ00000000, VIQY00000000, VIQX00000000, VIQW00000000, VIQV00000000, VIQU00000000, VIQT00000000, VIQS00000000, VIQR00000000, VIQQ00000000 ; SRA accession no.: SRR9696643 to SRR9696657).

**Author’s contributions**

MD, FL and AT conceived the project, analyzed, interpreted and integrated all the data. MD and FL wrote the manuscript. MD prepared figures. LB and JY achieved *in silico* design. MD performed bacteria isolation and identification, gDNA extraction, developed the qPCR microfluidic assay and analyzed macroscopic phenotyping data. AS carried out WGS, CE identified specific regions for primers design, and AM performed taxonomic annotation and functional bioinformatic analyses. PB, HD and MD performed animal experiments. CC, AD, JC and SP performed the immunophenotyping, analyzed and interpreted data. AB and DAO performed the metabophenotyping, analyzed and interpreted data. MS helped design the diet

849 induced stunting protocol. FDV performed, analyzed and contributed to the text and figures
850 related to the stunting model experiments. ALB provided technical support during the stunting
851 model experiments. All authors read and approved the final manuscript.
852
853 **Ethics approval**
854 Mouse experiments were performed as approved under the MESR and ANSES/ENVA/UPEC
855 Ethical Committee, protocols 2015042819315178 and 2016022616404045.
856
857 **Consent for publication**
858 Not applicable.
859
860 **Competing interests**
861 The authors declare that they have no competing interests.
862
863 **References**

864 1. Lynch SV, Pedersen O. The Human Intestinal Microbiome in Health and Disease. NEJM.
865 2016;375:2369-79.
866 2. Knight R, Callewaert C, Marotz C, Hyde ER, Debelius JW, McDonald D, Sogin ML. The
867 Microbiome and Human Biology. Annu Rev Genomics Hum Genet. 2017;18:65-86.
868 3. Bik EM. The Hoops, Hopes, and Hypes of Human Microbiome Research. Yale J Biol Med.
869 2016;89:363-73.
870 4. Leulier F, MacNeil LT, Lee WJ, Rawls JF, Cani PD, Schwarzer M, Zhao L, Simpson SJ.
871 Integrative Physiology: At the Crossroads of Nutrition, Microbiota, Animal Physiology, and
872 Human Health. Cell metab. 2017;25:522-34.
873 5. Nguyen TLA, Vieira-Silva S, Liston A, Raes J. How informative is the mouse for human gut
874 microbiota research? Dis Model Mech. 2015;8:1-16.
875 6. Mähler Convenor M, Berard M, Feinstein R, Gallagher A, Illgen-Wilcke B, Pritchett-Corning
876 K, Raspa M. FELASA recommendations for the health monitoring of mouse, rat, hamster,

877 guinea pig and rabbit colonies in breeding and experimental units. Lab Anim. 2014;48:178-
878 92.
879 7. Mamantopoulos M, Ronchi F, Van Hauwermeiren F, Vieira-Silva S, Yilmaz B, Martens L,
880 Saeys Y, Drexler SK, Yazdi AS, Raes J, Lamkanfi M, McCoy KD, Wullaert A. Nlrp6- and
881 ASC-Dependent Inflammasomes Do Not Shape the Commensal Gut Microbiota
882 Composition. Immunity. 2017;47:339-48.
883 8. Ussar S, Griffin NW, Bezy O, Fujisaka S, Vienberg S, Softic S, Deng L, Bry L, Gordon JI,
884 Kahn CR. Interactions between Gut Microbiota, Host Genetics and Diet Modulate the
885 Predisposition to Obesity and Metabolic Syndrome. Cell Metab. 2015;22:516-30.
886 9. Mooser C, Gomez de Agüero M, Ganal-Vonarburg SC. Standardization in host-microbiota
887 interaction studies: challenges, gnotobiology as a tool, and perspective. Curr Opin
888 Microbiol. 2018;44:50-60.
889 10. Laukens D, Brinkman BM, Raes J, De Vos M, Vandenabeele P. Heterogeneity of the
890 gut microbiome in mice: guidelines for optimizing experimental design. FEMS Microbiol
891 Rev. 2016;40:117-32.
892 11. Stappenbeck TS, Virgin HW. Accounting for reciprocal host-microbiome interactions in
893 experimental science. Nature. 2016;534:191-9.
894 12. Clavel T, Lagkouvardos I, Blaut M, Stecher B. The mouse gut microbiome revisited:
895 From complex diversity to model ecosystems. Int J Med Microbiol. 2016;306:316-327.
896 13. McCoy KD, Geuking MB, Ronchi F. Gut Microbiome Standardization in Control and
897 Experimental Mice. Curr Protoc Immunol. 2017;117:1-13.
898 14. Ley RE, Bäckhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. Obesity alters
899 gut microbial ecology. Proc Natl Acad Sci U S A. 2005;102:11070-5.
900 15. Xiao L, Feng Q, Liang S, Sonne SB, Xia Z, Qiu X, Li X, Long H, Zhang J, Zhang D, Liu
901 C, Fang Z, Chou J, Glanville J, Hao Q, Kotowska D, Colding C, Licht TR, Wu D, Yu J, Sung
902 JJ, Liang Q, Li J, Jia H, Lan Z, Tremaroli V, Dworzynski P, Nielsen HB, Bäckhed F, Doré J,
903 Le Chatelier E, Ehrlich SD, Lin JC, Arumugam M, Wang J, Madsen L, Kristiansen K. A
904 catalog of the mouse gut metagenome. Nat Biotechnol. 2015;33:1103-8.

905 16. Macpherson AJ, McCoy KD. Standardised animal models of host microbial mutualism.
906 Mucosal Immunol. 2015;8:476-86.
907 17. Lagkouvardos I, Pukall R, Abt B, Foesel BU, Meier-Kolthoff JP, Kumar N, Bresciani A,
908 Martínez I, Just S, Ziegler C, Brugiroux S, Garzetti D, Wenning M, Bui TP, Wang J,
909 Hugenholtz F, Plugge CM, Peterson DA, Hornef MW, Baines JF, Smidt H, Walter J,
910 Kristiansen K, Nielsen HB, Haller D, Overmann J, Stecher B, Clavel T. The Mouse Intestinal
911 Bacterial Collection (miBC) provides host-specific insight into cultured diversity and
912 functional potential of the gut microbiota. Nat Microbiol. 2016;1:16131.
913 18. Clavel T, Gomes-Neto JC, Lagkouvardos I, Ramer-Tait AE. Deciphering interactions
914 between the gut microbiota and the immune system via microbial cultivation and minimal
915 microbiomes. Immunol Rev. 2017;279:8-22.
916 19. Wymore Brand M, Wannemuehler MJ, Phillips GJ, Proctor A, Overstreet AM, Jergens
917 AE, Orcutt RP, Fox JG. The Altered Schaedler Flora: Continued Applications of a Defined
918 Murine Microbial Community. ILAR Journal. 2015;56:169-78.
919 20. Brugiroux S, Beutler M, Pfann C, Garzetti D, Ruscheweyh HJ, Ring D, Diehl M, Herp
920 S, Lötscher Y, Hussain S, Bunk B, Pukall R, Huson DH, Münch PC, McHardy AC, McCoy
921 KD, Macpherson AJ, Loy A, Clavel T, Berry D, Stecher B. Genome-guided design of a
922 defined mouse microbiota that confers colonization resistance against Salmonella enterica
923 serovar Typhimurium. Nature Microbiology. 2016;2:16215.
924 21. Schaedler RW, Dubos R, Costello R. J Exp Med. 1965;122:77-83.
925 22. Norin E, Midtvedt T. Intestinal microflora functions in laboratory mice claimed to harbor
926 a "normal" intestinal microflora. Is the SPF concept running out of date? Anaerobe.
927 2010;16:311-3.
928 23. Uchimura Y, Wyss M, Brugiroux S, Limenitakis JP, Stecher B, McCoy KD, Macpherson
929 AJ. Complete Genome Sequences of 12 Species of Stable Defined Moderately Diverse
930 Mouse Microbiota 2. Genome Announc. 2016;4.

931 24. Garzetti D, Brugiroux S, Bunk B, Pukall R, McCoy KD, Macpherson AJ, Stecher B.
932 High-Quality Whole-Genome Sequences of the Oligo-Mouse-Microbiota Bacterial
933 Community. Genome Announc. 2017;5.
934 25. Kim D, Song L, Breitwieser FP, Salzberg SL. Centrifuge: rapid and sensitive
935 classification of metagenomic sequences. Genome Res. 2016;26:1721-9.
936 26. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B,
937 Robbertse B, Smith-White B, Ako-Adjei D, Astashyn A, Badretdin A, Bao Y, Blinkova O,
938 Brover V, Chetvernin V, Choi J, Cox E, Ermolaeva O, Farrell CM, Goldfarb T, Gupta T, Haft
939 D, Hatcher E, Hlavina W, Joardar VS, Kodali VK, Li W, Maglott D, Masterson P, McGarvey
940 KM, Murphy MR, O'Neill K, Pujar S, Rangwala SH, Rausch D, Riddick LD, Schoch C,
941 Shkeda A, Storz SS, Sun H, Thibaud-Nissen F, Tolstoy I, Tully RE, Vatsan AR, Wallin C,
942 Webb D, Wu W, Landrum MJ, Kimchi A, Tatusova T, DiCuccio M, Kitts P, Murphy TD, Pruitt
943 KD. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion,
944 and functional annotation. Nucleic Acids Res. 2016;44:D733-45.
945 27. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search
946 tool. J Mol Biol. 1990;215:403-10.
947 28. Plaza-Diaz J, Ruiz-Ojeda FJ, Gil-Campos M, Gil A. Mechanisms of action of probiotics.
948 Adv Nutr. 2019;10:S49-66.
949 29. Feldman-Salit A, Hering S, Messiha HL, Veith N, Cojocaru V, Sieg A, Westerhoff HV,
950 Kreikemeyer B, Wade RC, Fiedler T. Regulation of the activity of lactate dehydrogenases
951 from four lactic acid bacteria. J Biol Chem. 2013;288:21295-306.
952 30. Galperin MY. Genome diversity of spore-forming Firmicutes. Microbiol Spectr. 2013;1.
953 31. Browne HP, Forster SC, Anonye BO, Kumar N, Neville BA, Stares MD, Goulding D,
954 Lawley TD. Culturing of 'unculturable' human microbiota reveals novel taxa and extensive
955 sporulation. Nature. 2016;533:543-46.
956 32. Mukhopadhya I, Moraïs S, Laverde-Gomez J, Sheridan PO, Walker AW, Kelly W,
957 Klieve AV, Ouwerkerk D, Duncan SH, Louis P, Koropatkin N, Cockburn D, Kibler R, Cooper
958 PJ, Sandoval C, Crost E, Juge N, Bayer EA, Flint HJ. Sporulation capability and amylosome

959 conservation among diverse human colonic and rumen isolates of the keystone starch-
960 degrader Ruminococcus bromii. Environ Microbiol. 2018;20:324-36.
961 33. Pearson JP, Brownlee IA. The interaction of large bowel microflora with the colonic
962 mucus barrier. Int J Inflam. 2010;2010:321426.
963 34. Macy JM, Ljungdahl LG, Gottschalk G. Pathway of succinate and propionate formation
964 in Bacteroides fragilis. J Bacteriol. 1978;134:84-91.
965 35. Johnson JL, Moore WEC, Moore LVH. Bacteroides Caccae sp.nov., Bacteroides
966 merdae sp.nov., and Bacteroides stercoris sp.nov. isolated from human feces. Int J Sys
967 Bacteriol. 1986;36:499-501.
968 36. Wannemuehler MJ, Overstreet AM, Ward DV, Phillips GJ. Draft genome sequences of
969 the altered schaedler flora, a defined bacterial community from gnotobiotic mice. Genome
970 Announc. 2014;2.
971 37. Thompson GR, Trexler PC. Gastrointestinal structure and function in germ-free or
972 gnotobiotic animals. Gut. 1971; 12:230-235.
973 38. Schwarzer M, Makki K, Storelli G, Machuca-Gayet I, Srutkova D, Hermanova P, Martino
974 ME, Balmand S, Hudcovic T, Heddi A, Rieusset J, Kozakova H, Vidal H, Leulier F.
975 Lactobacillus plantarum strain maintains growth of infant mice during chronic undernutrition.
976 Science. 2016;351:854-7.
977 39. Yan J, Herzog JW, Tsang K, Brennan CA, Bower MA, Garrett WS, Sartor BR, Aliprantis
978 AO, Charles JF. Gut microbiota induce IGF-1 and promote bone formation and growth. Proc
979 Natl Acad Sci U S A. 2016;113:E7554-63.
980 40. Chung H, Pamp SJ, Hill JA, Surana NK, Edelman SM, Troy EB, Reading NC,
981 Villablanca EJ, Wang S, Mora JR, Umesaki Y, Mathis D, Benoist C, Relman DA, Kasper
982 DL. Gut immune maturation depends on colonization with a host-specific microbiota. Cell.
983 2012;149:1578-93.
984 41. Sjögren YM, Tomicic S, Lundberg A, Böttcher MF, Björkstén B, Sverremark-Ekström
985 E, Jenmalm MC. Influence of early gut microbiota on the maturation of childhood mucosal
986 and systemic immune responses. Clin Exp Allergy. 2009;39:1842-51.

987 42. Cahenzli J, Köller Y, Wyss M, Geuking MB, McCoy KD. Intestinal microbial diversity
988 during early-life colonization shapes long-term IgE levels. Cell Host Microbe. 2013;14:559-
989 70.
990 43. Koch MA, Reiner GL, Lugo KA, Kreuk LSM, Stanbery AG, Ansaldo E, Seher TD,
991 Ludington WB, Barton GM. Maternal IgG and IgA antibodies dampen mucosal T helper cell
992 responses in early life. Cell. 2016;165:827-41.
993 44. Tanaka M, Nakayama J. Development of the gut microbiota in infancy and its impact
994 on health in later life. Allergol Int. 2017;66:515-22.
995 45. Hong SW, O E, Lee JY, Lee M, Han D, Ko HJ, Sprent J, Surh CD, Kim KS. Food
996 antigens drive spontaneous IgE elevation in the absence of commensal microbiota. Sci Adv.
997 2019;5:eaaw1507.
998 46. Atarashi K, Tanoue T, Shima T, Imaoka A, Kuwahara T, Momose Y, Cheng G,
999 Yamasaki S, Saito T, Ohba Y, Taniguchi T, Takeda K, Hori S, Ivanov II, Umesaki Y, Itoh K,
1000 Honda K. Induction of colonic regulatory T cells by indigenous Clostridium species. Science.
1001 2011;331:337-41.
1002 47. Nishimori JH, Newman TN, Oppong GO, Rapsinski GJ, Yen JH, Biesecker SG, Wilson
1003 RP, Butler BP, Winter MG, Tsolis RM, Ganea D, Tükel Ç. Microbial amyloids induce
1004 interleukin 17A (IL-17A) and IL-22 responses via Toll-like receptor 2 activation in the
1005 intestinal mucosa. Infect Immun. 2012;80:4398-408.
1006 48. McAleer JP, Kolls JK. Mechanisms controlling Th17 cytokine expression and host
1007 defense. J Leukoc Biol. 2011;90:263-70.
1008 49. Hegazy AN, West NR, Stubbington MJT, Wendt E, Suijker KIM, Datsi A, This S, Danne
1009 C, Campion S, Duncan SH, Owens BMJ, Uhlig HH, McMichael A; Oxford IBD Cohort
1010 Investigators, Bergthaler A, Teichmann SA, Keshav S, Powrie F. Circulating and Tissue-
1011 Resident CD4+ T Cells With Reactivity to Intestinal Microbiota Are Abundant in Healthy
1012 Individuals and Function Is Altered During Inflammation. Gastroenterology. 2017;153:1320-
1013 37.

1014 50. Dixon BR, Radin JN, Piazuelo MB, Contreras DC, Algood HM. IL-17a and IL-22 Induce
1015 Expression of Antimicrobials in Gastrointestinal Epithelial Cells and May Contribute to
1016 Epithelial Cell Defense against Helicobacter pylori. PLoS One. 2016;11:e0148514.
1017 51. Douzandeh-Mobarrez B, Kariminik A. Gut Microbiota and IL-17A: Physiological and
1018 Pathological Responses. Probiotics Antimicrob Proteins. 2019;11:1-10.
1019 52. Alex P, Zachos NC, Nguyen T, Gonzales L, Chen TE, Conklin LS, Centola M, Li X.
1020 Distinct Cytokine Patterns Identified from Multiplex Profiles of Murine DSS and TNBS-
1021 Induced Colitis. Inflamm Bowel Dis. 2009;15:341-52.
1022 53. Durand A, Audemard-Verger A, Guichard V, Mattiuz R, Delpoux A, Hamon P, Bonilla
1023 N, Rivière M, Delon J, Martin B, Auffray C, Boissonnas A, Lucas B. Profiling the lymphoid-
1024 resident T cell pool reveals modulation by age and microbiota. Nat Commun. 2018;9:68.
1025 54. Kennedy EA, King KY, Baldridge MT. Mouse Microbiota Models: Comparing Germ-
1026 Free Mice and Antibiotics Treatment as Tools for Modifying Gut Bacteria. Front Physiol.
1027 2018;9:1534.
1028 55. Milovanovic M, Drozdenko G, Weise C, Babina M, Worm M. Interleukin-17A promotes
1029 IgE production in human B cells. J Invest Dermatol. 2010;130:2621-8.
1030 56. Nakae S, Komiyama Y, Nambu A, Sudo K, Iwase M, Homma I, Sekikawa K, Asano M,
1031 Iwakura Y. Antigen-specific T cell sensitization is impaired in IL-17-deficient mice, causing
1032 suppression of allergic cellular and humoral responses. Immunity. 2002;17:375-87.
1033 57. Ma CS, Chew GY, Simpson N, Priyadarshi A, Wong M, Grimbacher B, Fulcher DA,
1034 Tangye SG, Cook MC. Deficiency of Th17 cells in hyper IgE syndrome due to mutations in
1035 STAT3. J Exp Med. 2008;205:1551-7.
1036 58. Eberl G. Immunity by equilibrium. Nat Rev Immunol. 2016;16:524-32.
1037 59. Shida K, Makino K, Morishita A, Takamizawa K, Hachimura S, Ametani A, Sato T,
1038 Kumagai Y, Habu S, Kaminogawa S. Lactobacillus casei inhibits antigen-induced IgE
1039 secretion through regulation of cytokine production in murine splenocyte cultures. Int Arch
1040 Allergy Immunol. 1998;1154:278-87.

1041 60. Ekmekciu I, von Klitzing E, Neumann C, Bacher P, Scheffold A, Bereswill S, Heimesaat
1042 MM. Fecal Microbiota Transplantation, Commensal Escherichia coli and Lactobacillus
1043 johnsonii Strains Differentially Restore Intestinal and Systemic Adaptive Immune Cell
1044 Populations Following Broad-spectrum Antibiotic Treatment. Front Microbiol. 2017;8:2430.
1045 61. Szkaradkiewicz AK, Stopa J, Karpinski TM. Effect of oral administration involving a
1046 probiotic strain of Lactobacillus reuteri on pro-inflammatory cytokine response in patients
1047 with chronic periodontitis. Arch Immunol Ther Exp (Warsz). 2014;62:495-500.
1048 62. Haileselassie Y, Johansson MA, Zimmer CL, Bjorkander S, Petursdottir DH, Dicksved
1049 J, Petersson M, Persson JO, Fernandez C, Roos S, Holmlund U, Sverremark-Ekström E.
1050 Lactobacilli regulate Staphylococcus aureus 161:2-induced pro-inflammatory T-cell
1051 responses in vitro. PLoS One. 2013;8:e77893.
1052 63. Ishioka M, Miura K, Minami S, Shimura Y, Ohnishi H. Altered Gut Microbiota
1053 Composition and Immune Response in Experimental Steatohepatitis Mouse Models. Dig
1054 Dis Sci. 2017;62:396-406.
1055 64. Sepp E, Julge K, Mikelsaar M, Bjorksten B. Intestinal microbiota and immunoglobulin
1056 E responses in 5-year-old Estonian children. Clin Exp Allergy. 2005;35:1141-6.
1057 65. Prince BT, Mandel MJ, Nadeau K, Singh AM. Gut Microbiome and the Development of
1058 Food Allergy and Allergic Disease. Pediatr Clin North Am. 2015;62:1479-92.
1059 66. Marrs T, Sim K. Demystifying dysbiosis: Can the gut microbiome promote oral tolerance
1060 over IgE-mediated food allergy? Curr Pediatr Rev. 2018;14:156-63.
1061 67. Tanoue T, Morita S, Plichta DR, Skelly AN, Suda W, Sugiura Y, et al. A defined
1062 commensal consortium elicits CD8 T cells and anti-cancer immunity. Nature. 2019;565:600-
1063 5.
1064 68. Nicholson JK, Holmes E, Kinross J, Burcelin R, Gibson G, Jia W, Pettersson S. Host-
1065 gut microbiota metabolic interactions. Science. 2012;336:1262-7.
1066 69. Cani PD, Knauf C. How gut microbes talk to organs: The role of endocrine and nervous
1067 routes. Mol Metab. 2016;5:743-52.

1068 70. Mestdagh R, Dumas ME, Rezzi S, Kochhar S, Holmes E, Claus SP, Nicholson JK. Gut
1069 microbiota modulate the metabolism of brown adipose tissue in mice. J Proteome Res.
1070 2012;11:620-30.
1071 71. Claus SP, Tsang TM, Wang Y, Cloarec O, Skordi E, Martin FP, Rezzi S, Ross A,
1072 Kochhar S, Holmes E, Nicholson JK. Systemic multicompartmental effects of the gut
1073 microbiome on mouse metabolic phenotypes. Mol Syst Biol. 2008;4:219.
1074 72. Org E, Blum Y, Kasela S, Mehrabian M, Kuusisto J, Kangas AJ, Soininen P, Wang Z,
1075 Ala-Korpela M, Hazen SL, Laakso M, Lusis AJ. Relationships between gut microbiota,
1076 plasma metabolites, and metabolic syndrome traits in the METSIM cohort. Genome Biol.
1077 2017;18:70.
1078 73. Mardinoglu A, Shoaie S, Bergentall M, Ghaffari P, Zhang C, Larsson E, Bäckhed F,
1079 Nielsen J. The gut microbiota modulates host amino acid and glutathione metabolism in
1080 mice. Mol Syst Biol. 2015;11:834.
1081 74. Trent CM, Blaser MJ. Microbially Produced Acetate: A "Missing Link" in Understanding
1082 Obesity? Cell Metab. 2016;24:9-10.
1083 75. He X, Slupsky CM. Metabolic fingerprint of dimethyl sulfone (DMSO2) in microbial-
1084 mammalian co-metabolism. J Proteome Res. 2014;13:5281-92.
1085 76. Dorokhov YL, Shindyapina AV, Sheshukova EV, Komarova TV. Metabolic methanol:
1086 molecular pathways and physiological roles. Physiol Rev. 2015;95:603-44.
1087 77. Max B, Salgado JM, Rodríguez N, Cortés S, Converti A, Domínguez JM.
1088 Biotechnological production of citric acid. Braz J Microbiol. 2010;41:862-75.
1089 78. Elshaghabee FM, Bockelmann W, Meske D, de Vrese M, Walte HG, Schrezenmeir J,
1090 Heller KJ. Ethanol Production by Selected Intestinal Microorganisms and Lactic Acid
1091 Bacteria Growing under Different Nutritional Conditions. Front Microbiol. 2016;7:47.
1092 79. Santos GA, Moura RF, Vitorino DC, Roman EA, Torsoni AS, Velloso LA, Torsoni MA.
1093 Hypothalamic AMPK activation blocks lipopolysaccharide inhibition of glucose production
1094 in mice liver. Mol Cell Endocrinol. 2013;381:88-96.

1095 80. Molinaro A, Caesar R, Holm LM, Tremaroli V, Cani PD, Bäckhed F. Host-microbiota
1096 interaction induces bi-phasic inflammation and glucose intolerance in mice. Mol Metab.
1097 2017;6:1371-80.
1098 81. Lee YH, Hsu HC, Kao PC, Shiao YJ, Yeh SH, Shie FS, Hsu SM, Yeh CW, Liu HK, Yang
1099 SB, Tsay HJ. Augmented Insulin and Leptin Resistance of High Fat Diet-Fed
1100 APPswe/PS1dE9 Transgenic Mice Exacerbate Obesity and Glycemic Dysregulation. Int J
1101 Mol Sci. 2018;19.
1102 82. Kaplan SA, Cohen P. The somatomedin hypothesis 2007: 50 years later. J Clin
1103 Endocrinol Metab. 2007;92:4529-35.
1104 83. Klover P, Hennighausen L. Postnatal body growth is dependent on the transcription
1105 factors signal transducers and activators of transcription 5a/b in muscle: a role for
1106 autocrine/paracrine insulin-like growth factor I. Endocrinology. 2007;148:1489-97.
1107 84. Luczynski P, McVey Neufeld KA, Oriach CS, Clarke G, Dinan TG, Cryan JF. Growing
1108 up in a Bubble: Using Germ-Free Animals to Assess the Influence of the Gut Microbiota on
1109 Brain and Behavior. Int J Neuropsychopharmacol. 2016;19.
1110 85. Blanton LV, Barratt MJ, Charbonneau MR, Ahmed T, Gordon JI. Childhood
1111 undernutrition, the gut microbiota, and microbiota-directed therapeutics. Science.
1112 2016;352:1533.
1113 86. Gehrig JL, Venkatesh S, Chang HW, Hibberd MC, Kung VL, Cheng J, Chen RY,
1114 Subramanian S, et al. Effects of microbiota-directed foods in gnotobiotic animals and
1115 undernourished children. Science. 2019;365.
1116 87. Edgar RC. 2018. Updating the 97% identity threshold for 16S ribosomal RNA OTUs.
1117 Bioinformatics. 2018;34:2371-5.
1118 88. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid
1119 assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol.
1120 2007;73:5261-7.

1121 89. Kim OS, Cho YJ, Lee K, Yoon SH, Kim M, Na H, Park SC, Jeon YS, Lee JH, Yi H, Won
1122 S, Chun J. Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with
1123 phylotypes that represent uncultured species. Int J Syst Evol Microbiol. 2012;62:716-21.
1124 90. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010.
1125 Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc
1126 91. Coil D, Jospin G, Darling AE. A5-miseq: an updated pipeline to assemble microbial
1127 genomes from Illumina MiSeq data. Bioinformatics. 2015;31:587-9.
1128 92. Darling AE, Tritt A, Eisen JA, Facciotti MT. Mauve Assembly Metrics. Bioinformatics.
1129 2011;27:2756-7.
1130 93. Chen S, Huang T, Zhou Y, Han Y, Xu M, Gu J. AfterQC: automatic filtering, trimming,
1131 error removing and quality control for fastq data. BMC Bioinformatics. 2017;18:80.
1132 94. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal:
1133 prokaryotic gene recognition and translation initiation site identification. BMC
1134 Bioinformatics. 2010;11:119.
1135 95. Finn RD, Clements J, Arndt W, Miller BL, Wheeler TJ, Schreiber F, Bateman A, Eddy
1136 SR. HMMER web server: 2015 update. Nucleic Acids Res. 2015;43:W30-8.
1137 96. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference
1138 resource for gene and protein annotation. Nucleic Acids Res. 2016;44:D457-62.
1139 97. Delcher AL, Salzberg SL, Phillippy AM. Using MUMmer to identify similar regions in
1140 large sequence sets. Curr Protoc Bioinformatics. 2003;Chapter 10:Unit 10.3.
1141 98. Ernst C, Rahmann S. PanCake: a data structure for pangenomes. German conference
1142 on bioinformatics 2013. 2013:35-45.
1143 99. Packey CD, Shanahan MT, Manick S, Bower MA, Ellermann M, Tonkonogy SL, Carroll
1144 IM, Sartor RB. Molecular detection of bacterial contamination in gnotobiotic rodent units.
1145 Gut Microbes. 2013;4:361-70.
1146 100. Spector AA. Fatty acid binding to plasma albumin. J Lipid Res. 1975;16:165-79.
1147 101. Nagana Gowda GA, Raftery D. Quantitating metabolites in protein precipitated serum
1148 using NMR spectroscopy. Anal Chem. 2014;86:5433-40.

1149 102. Tiziani S, Emwas AH, Lodi A, Ludwig C, Bunce CM, Viant MR, Günther UL. Optimized
1150 metabolite extraction from blood serum for 1H nuclear magnetic resonance spectroscopy.
1151 Anal Biochem. 2008;377:16-23.
1152 103. Tynkkynen T. ¹H NMR analysis of serum lipids. Publications of the University of Eastern
1153 Finland. Dissertations in Forestry and Natural Sciences, no 76. 2012. URN:ISBN:978-952-
1154 61-0839-1
1155 104. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics.
1156 Chemometr Intel Lab. 2001;58:109-30.
1157 105. Barker M, Rayens W. Partial least squares for discrimination. J Chemometr.
1158 2003;17:166-73.
1159 106. Dray S, Dufour AB. The ade4 Package: Implementing the Duality Diagram for
1160 Ecologists. J Stat Soft. 2007;22:1-20.
1161 107. Yutin N, Galperin MY. A genomic update on clostridial phylogeny: Gram-negative spore
1162 formers and other misplaced clostridia. Environ Microbiol. 2013;15:2631-41.
1163
1164
1165 **Legends**
1166 **Fig. 1. Selection, isolation and functional analysis of GM15 strains. a** Composition of
1167 representative and prevalent bacterial families identified from fecal samples of 4 C57BL/6J
1168 mice of our colony. **b** 11 bacterial strains were isolated from fecal or caecal samples of either
1169 C57BL/6J or ASF mice using various culture methods. 4 additional strains were obtained from
1170 the DSMZ collection. **c** Unique KO (KEGG Orthology) groups are those present in only one
1171 GM15 member. The unique contribution of each GM15 member is relatively evenly distributed
1172 (<2%), except *E. coli* whose unique functions contribute for 33%, although non-unique KO
1173 groups are primarly responsible for the GM15's metagenomic profile with 59% of contribution.
1174 **d** Heat map of hierarchical clustering of enzymatic activities in the gut for the 15 strains of the
1175 GM15 model. In quorum sensing (QS) function, only effector proteins were screened (because
1176 receptors suffer of similarities with nonspecific QS receptors). **e** Heat map of hierarchical

clustering of KEGG modules distribution in the metagenomes of GM15, SOPF, ASF and Oligo-MM$_{12}$ models. Clusters of KEGG modules are highlighted with their respective size (number of KEGG modules). The list of KEGG modules and clusters is shown in Additional file 3: Table S2.

**Fig. 2. Stability assessments of the GM15 mice gut microbiota over filial generations, in aging, under diet change and through FMT**. SOPF groups show the distribution of each GM15 strains in the complex gut microbiota of 8-week-old SOPF mice. The absolute quantification of each strain was determined by specific qPCR microfluidic assay. To gain clarity, each GM15 strain is associated to a number between 1 and 15. * Strains 4, 6, 7 and 14 were at the detection limit of the qPCR microfluidic assay, and thus were not detected in all samples. ** Strains 8, 9 and 10 were below detection limit of the qPCR microfluidic assay. Strain 5, obtained from the DSMZ collection, was not detected in our SOPF colony. **a** Radar plot showing the GM15 strains distribution in feces of GF C57Bl/6J mice colonized with the GM15 community (F0, n=10) and bred for consecutive generations (F1 to F4, n=19 to 8). **b** Radar plot showing the overall stability of the GM15 community composition in feces collected from 9 mice between 2 and 12 months of age (2 mice died at 12-month-old). **c** Radar plot showing that an alternative diet, such as a maintenance diet, can be used for 4 weeks, and then reversed to the breeding diet for 4 more weeks, without modifying the composition of the gut microbiota of GM15 mice (n=18) compared to mice fed all along with the breeding diet (n=9). **d** Radar plot showing the feasibility of GM15 fecal microbiota transplantation to GF mice (n=9). **e** Box plots showing the low variability of the GM15 strains concentrations considering all fecal samples of mice from generation F1 to F4 (n=47), from generation F1 at 12-month-old (n=7), from generation F1 under diet change (n=18), and fecal samples collected 3 weeks post fecal microbiota transplantation in GF mice. Whiskers represent min and max data.

**Fig. 3. Macroscopic phenotyping on mice from two consecutive filial generations (F1-F2). a-c** 1-way ANOVA analyses of the reproductive performance, where dots, lines and error

bars represent respectively individual litters (34 GF, 31 GM15 and 19 SOPF), means and SEM. **a** GM15, like SOPF, delivered pups around 24 days after mating, against 33 days for GF mice. **b** GM15 mean litter of 5 pups was in between GF (4 pups) and SOPF (6 pups). **c** Perinatal mortality was similar for GM15 and SOPF (12 and 11%), and much lower than for GF mice (45%). **d** Box plots showing no difference between GF, GM15 and SOPF mice in terms of feed intake, although much variability was observed with GF mice. Whiskers represent min and max data. **e-f** 2-way ANOVA analyses of body growth of mice bred with their mothers until week 4, where lines and error bars represent respectively means and SEM. Body weight and size curves of GM15 mice (9 females (F), 12 males (M)) were similar or equivalent to those of SOPF mice (10 F, 10 M), whereas GF mice showed a significant delayed growth (7 F, 10 M). **g-h** Organs weight or size impacted by age, filial generation or sex, were analyzed by the F-test for multiple linear regressions, otherwise by 1-way ANOVA. Each animal is represented by a dot at the age of the sacrifice. **g** The characteristic caecum enlargement of GF mice was partially, but significantly, reduced in GM15 mice. The weights of the brain, liver and spleen were equivalent in GM15 and SOPF mice, and significantly bigger than those of GF mice. **h** Femur and tibia lengths were also equivalent in GM15 and SOPF mice, and significantly bigger than those of GF mice. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.

**Fig. 4. Immune phenotyping through serum and fecal immunoglobulin subtyping, circulating cytokines analysis and immune cell populations analysis in different organs** **a-c** Dot plots where dots, lines and error bars represent respectively individual mice (15-17 GF, 20-21 GM15 and 20 SOPF), means and SEM. **a** Fecal IgA, serum IgA, IgG2b, IgG3, IgM and IgE Luminex analysis. F-test for multiple linear regression analyses (fecal IgA) and Dunn's multiple comparison showing restoration of fecal IgA, serum IgA and serum IgG2b levels in GM15-colonized mice at the same level than SOPF, slightly increased IgM levels in GM15 mice compared to SOPF mice and elevated levels of IgG3 and IgE levels in GM15 mice compared to SOPF and GF mice. **b** Circulating IL-17a and IL-22 levels Luminex analysis. Dunn's multiple comparison showing elevated IL-17a levels in GM15 mice compared to SOPF

and GF mice and restoration of IL-22 levels in GM15 mice at the same level than SOPF. **c** CD45+ count comparison in spleen, thymus, MLN and PP by flow cytometry. Dunn's multiple comparison showing increased CD45+ cell count in spleen, thymus, and PP of SOPF mice compared to GF mice, levels of GM15 mice are in between, and no difference between groups for MLN. Tukey's multiple comparison showing increased PP in SOPF mice, then in GM15 mice, compared to GF mice. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.

**Fig. 5. Metabolic phenotyping of GM15 animals**. **a** PCA score plot representing the distribution of the sample in the two first principal components. b Biplot representing the projection the samples in the PLS-DA score plot given by the two first components and the projection of the contribution of significant metabolites. **c** F-test for multiple linear regressions showing equivalent decreased levels of Free Cholesterol and Phosphatidylcholine in GM15 and SOPF mice compared to GF mice. **d-e** Circulating hormones Luminex analysis. Dot plots where dots, lines and error bars represent respectively individual mice (16 GF, 21 GM15 and 19-20 SOPF), means and SEM. **d** Dunn's multiple comparison showing restoration of serum IGF-1 titers in GM15-colonized mice at the same level than SOPF. **e** F-test for multiple linear regressions showing equivalent and reduced serum corticosterone concentrations in GM15 and SOPF mice, compared to GF mice.

**Fig. 6. GM15 mice and SOPF respond differently under post-weaning chronic undernutrition**. Body weight (**a**) and body size (**b**) were measured every week from weaning to post-natal day 56. **c.** Representative picture of caecum, subcutaneous fat and right femur of mice at day 56. Tibia length (**d**), liver (**e**) and epididymal white adipose tissue (**f**) weight were measured. **g.** PCA of tissue weight and size showed clustering of GM15 and SOPF under control conditions but a distinct pattern under nutritional stress. Percentages on each axis indicate the variance associated to each coordinate. **h**. Serum IGF1 at day 56. BD: breeding diet; DD: depleted diet. P-values after two-way ANOVA were adjusted for Sidak's post-hoc test for multiple comparisons. *P<0.05, **P<0.01, ***P<0.001.

**Additional file 1: Figure S1. Morphological features of the GM15 strains**. Each bacterial strain was grown individually from a single colony isolated on agar medium and amplified in liquid culture to exponential growth phase. Bacteria were Gram-stained and imaged by light microscopy (80-fold magnification, NanoZoomer S60, Hamamatsu). Members of Bacteroidaceae, Tannerellaceae and Enterobacteriaceae stained Gram-negative, while Lactobacillaceae, Erysipelotrichaceae and Ruminococcaceae stained Gram-positive. Lachnospiraceae stained Gram-positive, except *Lachnospiraceae* sp. MD329 and *Lachnospiraceae* bacterium COE1, which stained Gram-negative likely due to their cell wall structure as already reported for other Clostridiales [107].

**Additional file 2: Table S1. GM15 strains-specific primers**. Sequences of primers designed in this study and detection limits by qPCR microfluidic assay.

**Additional file 3: Table S2. KEGG clusters**.

**Additional file 4: Figure S2. *In vivo* experimental design**. First, 5 couples of 8 week-old C57BL/6J GF mice were colonized by oral inoculations of the GM15 bacterial community, twice at 48h interval (F0). GM15 microbiota stability was assessed by qPCR microfluidic assay from feces collected after 1, 2 and 3 weeks. Then, in order to evaluate the reproducible transfer of the GM15 microbiota by fecal microbiota transplantation, 7 week-old C57BL/6J GF mice were colonized by oral inoculations of a suspension of fresh fecal pellets from GM15 mice (F0), twice at 48h interval. Again, qPCR microfluidic assay from feces collected after 1, 2 and 3 weeks was carried out. Next, the GM15 mouse line was amplified to monitor the GM15 microbiota through four filial generations at 6 weeks of age (F1-F4), and allow the phenotyping study from two consecutive generations (F1.2 and F2.1). Reproduction performance and perinatal mortality were recorded, 4 week-old mice were randomly selected at weaning, monitored weekly for body weight and size, and feed intake, until sacrifice at 8-9 weeks of age.

GF and SOPF mice were also studied as control groups. Besides, a comparative analysis was done on the fecal microbiota of 8 week-old GM15 mice (F1.1) either fed with the breeding diet or an alternative isocaloric diet given for 4 weeks. Finally, the fecal microbiota of 6 month-old and 12 month-old control GM15 mice fed with the breeding diet was analyzed.

**Additional file 5: Figure S3. Assessment of gut microbiota stability in GM15 founders (F0, n=10), and of reproducibility between fecal and caecal samples of individual GM15 mice (F2, n=11).** SOPF group shows the distribution of each GM15 strains in the complex gut microbiota of 8-week-old SOPF mice. The absolute quantification of each strain was determined by specific qPCR microfluidic assay. To gain clarity, each GM15 strain is associated to a number between 1 and 15. * Strains 4, 6, 7 and 14 were at the detection limit of the qPCR microfluidic assay, and thus were not detected in all samples. ** Strains 8, 9 and 10 were below detection limit of the qPCR microfluidic assay. Strain 5, obtained from the DSMZ collection, was not detected in our SOPF colony. **a** Radar plot showing the GM15 strains distribution in feces of GF C57Bl/6J mice before and 1, 2 and 3 weeks after the oral colonization with the GM15 community. **b** Radar plot showing the reproducible detection of the GM15 strains in feces and caecum collected from the same GM15 mice.

**Additional file 6: Figure S4. Immune markers. Immune cell populations profiling analysis by flow cytometry in different organs a-e** Dot plots where dots, lines and error bars represent respectively individual mice (13-17 GF, 21 GM15 and 19-20 SOPF), means and SEM. **a** Tukey's multiple comparison showing restoration of whole blood monocytes levels in GM15 mice at the same levels than in SOPF mice and difference of B cells levels between the 3 groups. **b** Tukey's multiple comparison showing splenic dendritic levels slightly decreased in GM15 mice compared to SOPF, and Dunn's multiple comparison showing restoration of NK cells levels in GM15 mice at the same levels than in SOPF mice. **c** Tukey's multiple comparison showing restoration of MLN CD8+ T cells and B cells levels in GM15 mice at the same levels than in SOPF mice. **d** Tukey's multiple comparison thymic CD4+ T cells levels restored in

GM15 mice at the same levels than in SOPF. **e** Tukey's multiple comparison showing slightly decreased PP CD4+ T cells in GM15 mice compared to SOPF and Dunn's multiple comparison showing increased PP dendritic cells levels in GF mice compared to SOPF mice and increased NK cells levels in GF and GM15 mice compared to SOPF.

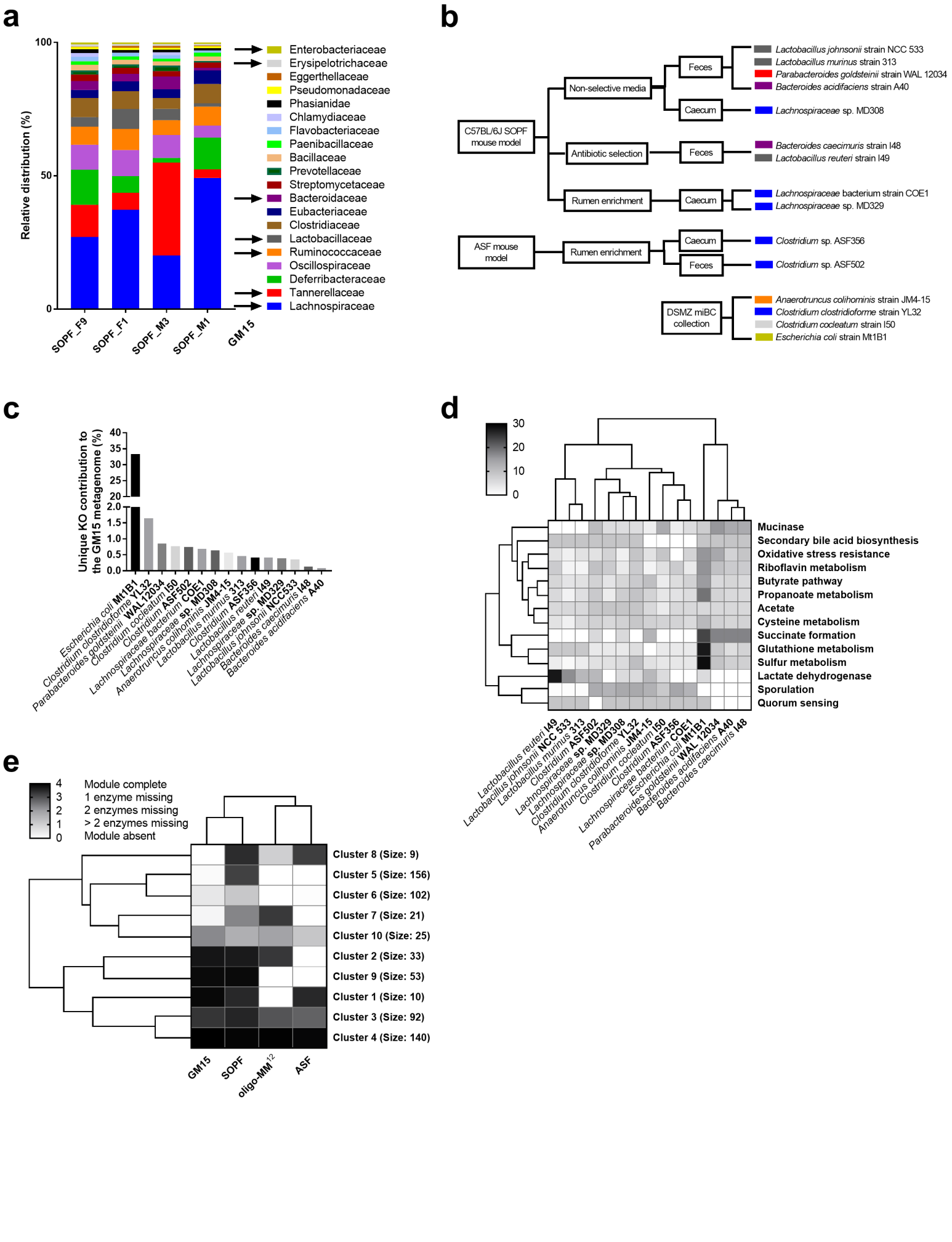
**Additional file 7: Table S3. Polar metabolites concentration in plasma samples of GF, GM15 and SOPF mice**. The metabolites quantification was performed with the help of Chenomx NMR suite 8.3. The concentrations are reported as mean and standard deviation.

**Additional file 8: Table S4. Non-polar metabolites concentration in plasma samples of GF, GM15 and SOPF mice.** The metabolites quantification was performed with the help of Chenomx NMR suite 8.3. The non-polar database profiles were created with the help of Compound builder module using $_1$H NMR spectra of authentic lipid standards. The concentrations are reported as mean and standard deviation.

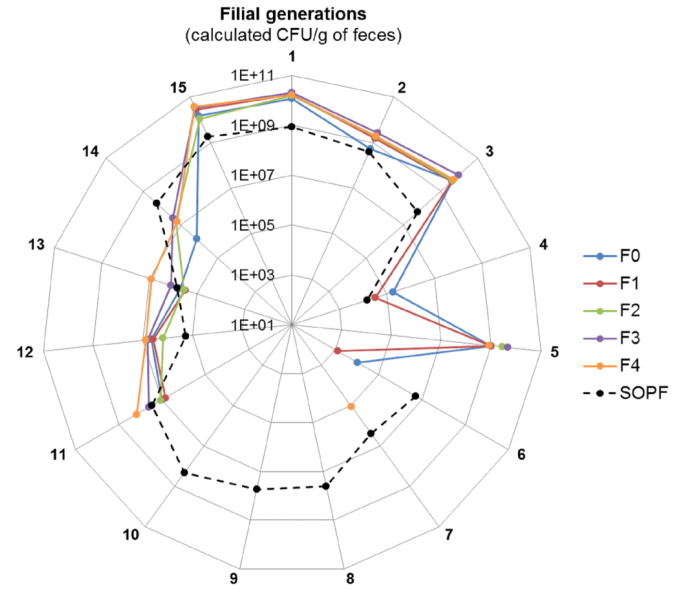
**Additional file 9: Figure S5. PCA screen plot based on $_1$H NMR plasma fingerprints.** Processed $_1$H NMR plasma spectra of polar metabolites were binned with AMIX v3.9.14 software from Bruker Biospin, using 0.04ppm width from 0.5 to 10ppm spectral window. The residual water region from 4.68 to 4.88pp was excluded from analysis. All spectra were normalized to the total spectral area and the data table was exported into SIMCA v13.0.3 software for statistical analysis. The PCA analysis was performed using UV-scaling of data and the model was autofit using cross validation rules to determine the number of significant components. The clouds of sample points GF, GM15 and SOPF are distributed according to microbiota complexity from left to right on PC1 (18%) and top to down on PC2 (12%).

**Additional file 10: Figure S6. Measurements of several organs after nutritional challenge**. Measurements at day 56 of weights of caecum (**a**), visceral (**b**) and subcutaneous (**c**) white adipose tissue, brown adipose tissue (**d**), kidneys (**e**) and soleus muscle (**f**) and length

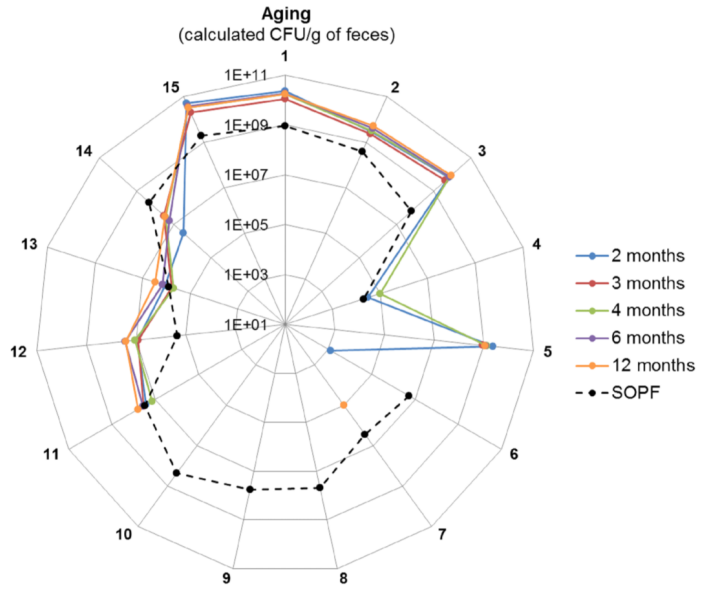
of right femur (**g**). P-values after two-way ANOVA were adjusted for Sidak's post-hoc test for multiple comparisons. **h.** Correlation matrix showing Pearson r values of all the measured parameters. For PCA, brown adipose tissue was excluded, given its lack of correlation with other parameters.



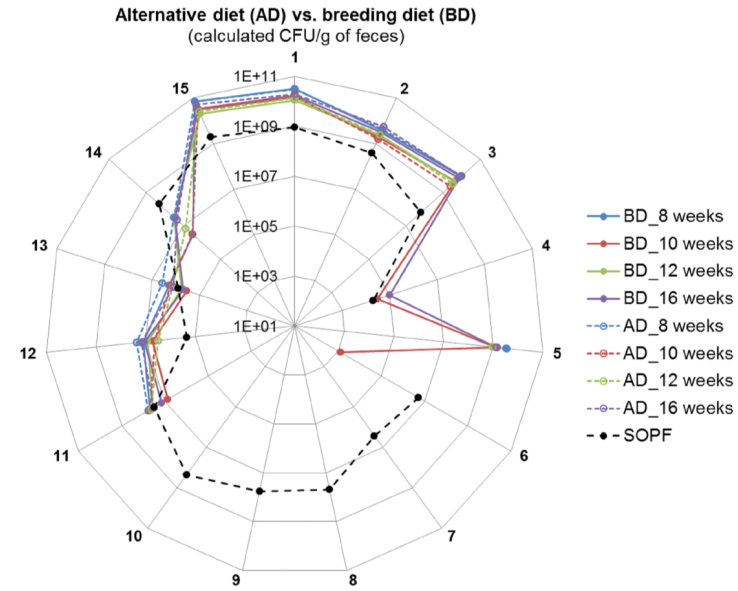
**a**



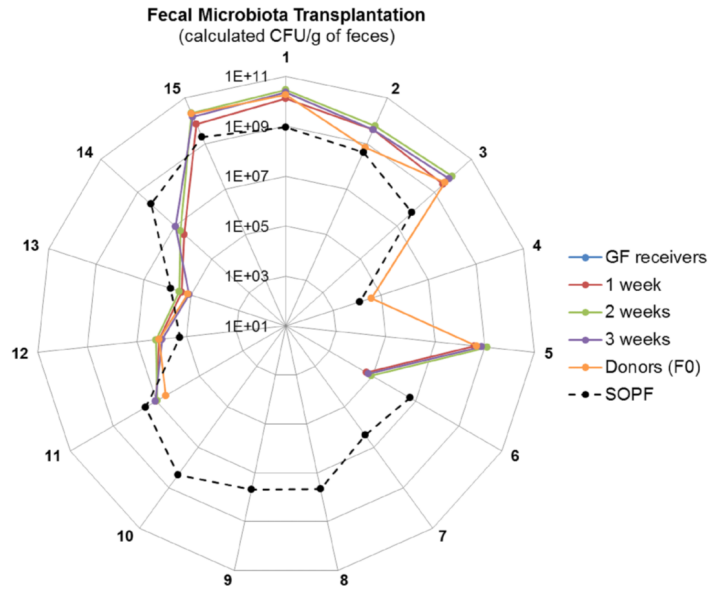

**b**



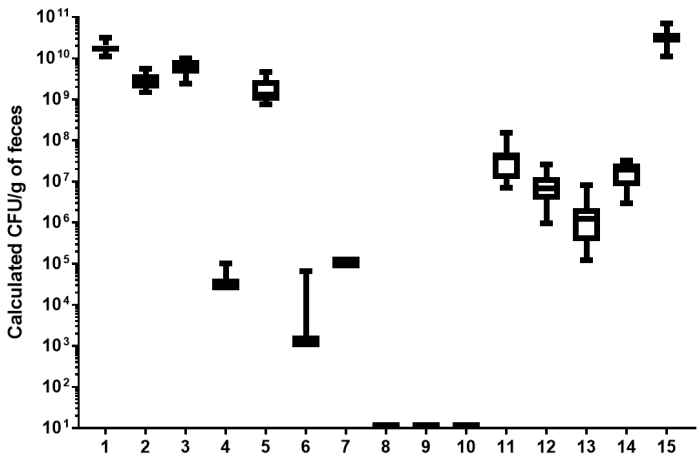

**c**



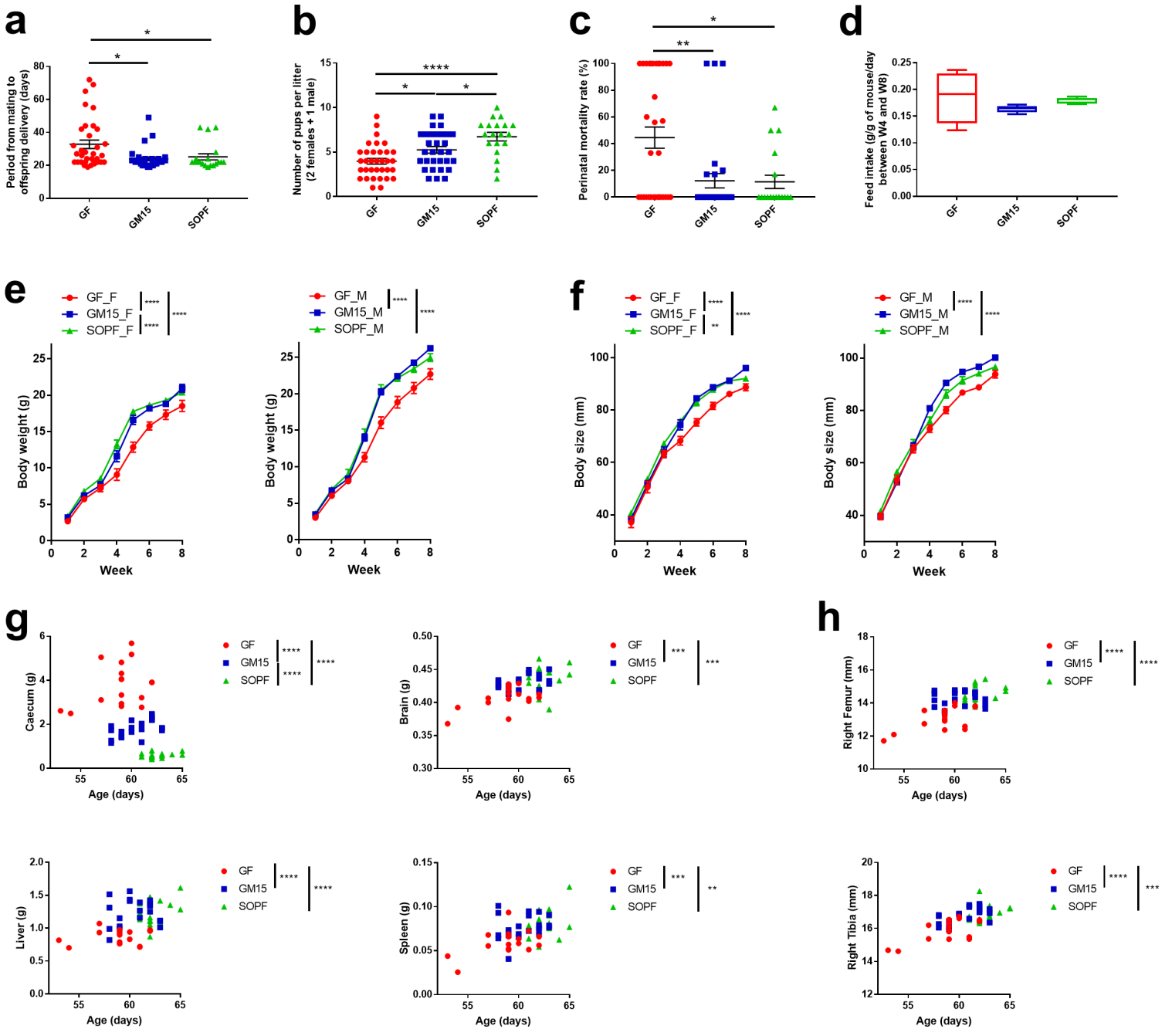

**d**

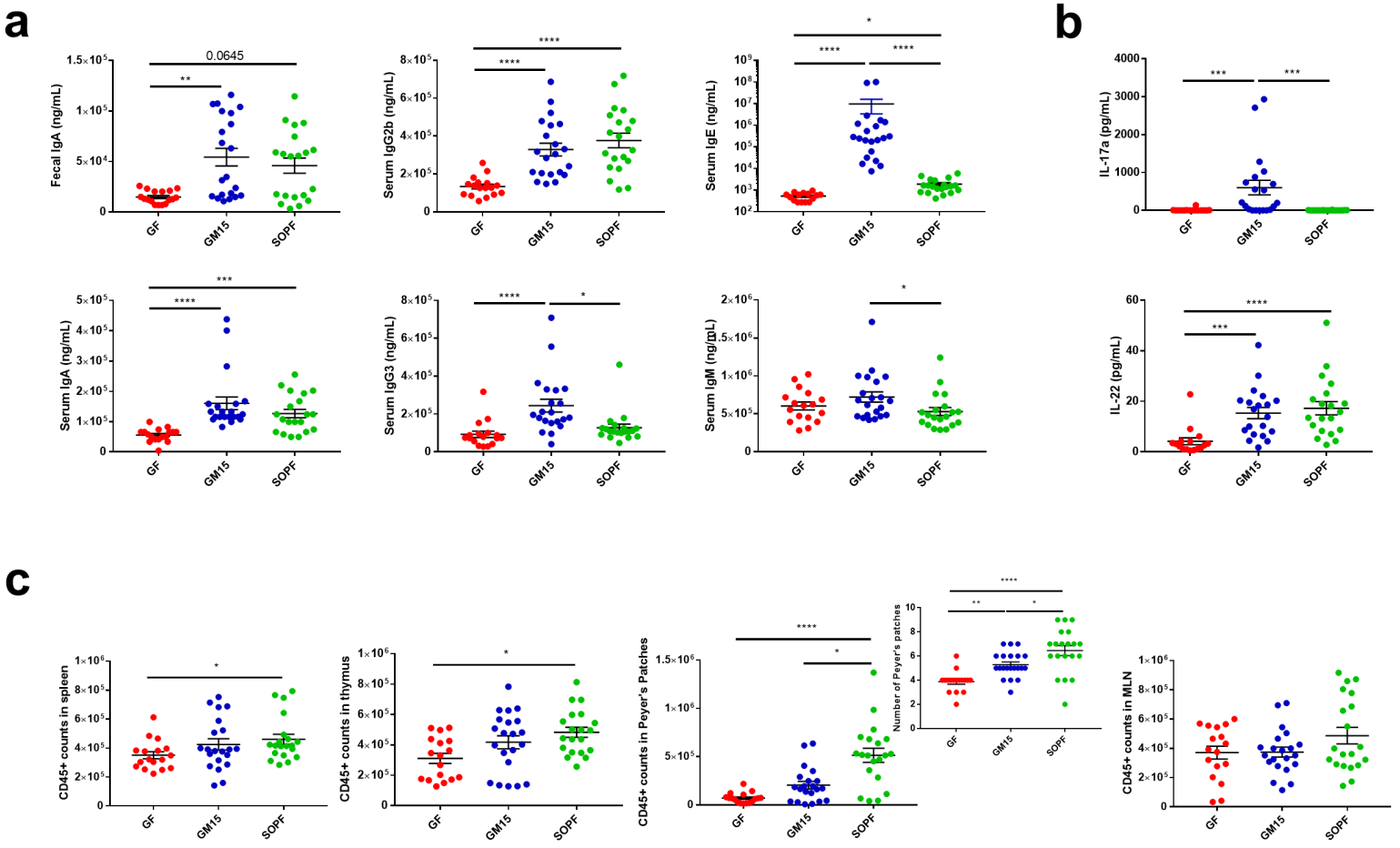



**e**



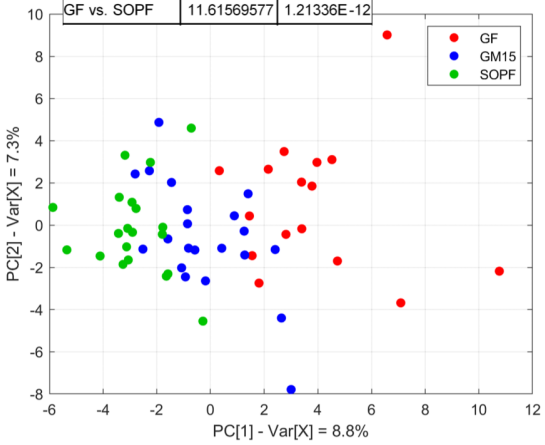

1. *Bacteroides acidifaciens* A40
2. *Bacteroides caecimuris* I48
3. *Parabacteroides goldsteinii* WAL 12034
4. *Clostridium cocleatum* I50 *
5. *Clostridium clostridioforme* YL32
6. *Clostridium* ASF356 *
7. *Clostridium* ASF502 *
8. *Lachnospiraceae* bacterium COE1 **
9. *Lachnospiraceae* sp. MD329 **
10. *Lachnospiraceae* sp. MD308 **
11. *Lactobacillus johnsonii* NCC 533
12. *Lactobacillus murinus* 313
13. *Lactobacillus reuteri* I49
14. *Anaerotruncus colihominis* JM4-15 *
15. *Escherichia coli* Mt1B1





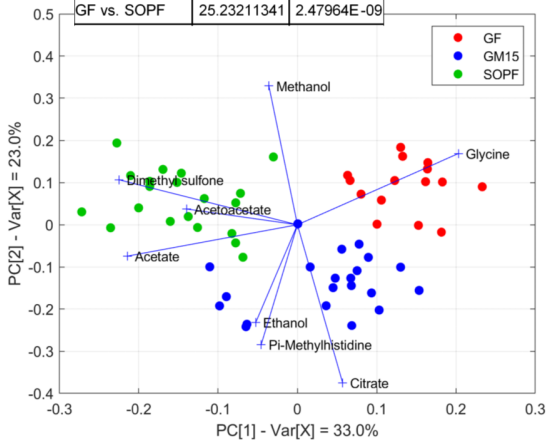
**a**

| | Distance | F. p-value |
|---|---|---|
| GM15 vs. SOPF | 3.120670957 | 3.18821E-08 |
| GM15 vs. GF | 5.013062179 | 3.73739E-08 |
| GF vs. SOPF | 11.61569577 | 1.21336E-12 |

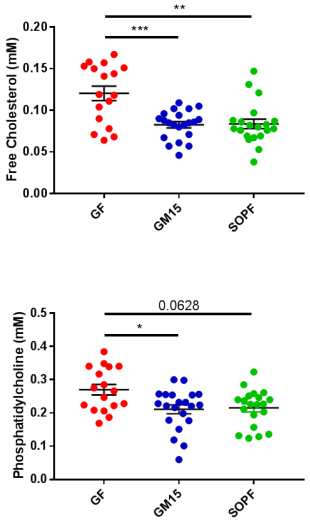



**b**

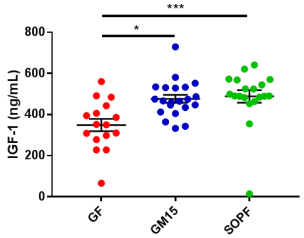
| | Distance | F. p-value |
|---|---|---|
| GM15 vs. SOPF | 20.47099158 | 1.70132E-09 |
| GM15 vs. GF | 21.88771817 | 8.24592E-08 |
| GF vs. SOPF | 25.23211341 | 2.47964E-09 |




**c**




**d**




**e**