A standardized gnotobiotic mouse model harboring a minimal 15-member mouse gut microbiota recapitulates SOPF/SPF phenotypes

Marion Darnaud1, Filipe De Vadder2, Pascaline Bogeat1, Lilia Boucinha1, Anne-Laure Bulteau2, Andrei Bunescu1, Céline Couturier2, Ana Delgado1, Hélène Dugua1, Céline Elie1, Alban Mathieu1, Tereza Novotná3, Djomangan Adama Ouattara1, Séverine Planel1, Adrien Saliou1, Dagmar Šrůtková3, Jennifer Yansouni1, Bärbel Stecher4,5, Martin Schwarzer3,6, François Leulier1,2,6 & Andrea Tamellini1,6

Mus musculus is the classic mammalian model for biomedical research. Despite global efforts to standardize breeding and experimental procedures, the undefined composition and interindividual diversity of the microbiota of laboratory mice remains a limitation. In an attempt to standardize the gut microbiome in preclinical mouse studies, here we report the development of a simplified mouse microbiota composed of 15 strains from 7 of the 20 most prevalent bacterial families representative of the fecal microbiota of C57BL/6J Specific (and Opportunistic) Pathogen-Free (SPF/SOPF) animals and the derivation of a standardized gnotobiotic mouse model called GM15. GM15 recapitulates extensively the functionalities found in the C57BL/6J SOPF microbiota metagenome, and GM15 animals are phenotypically similar to SOPF or SPF animals in two different facilities. They are also less sensitive to the deleterious effects of post-weaning malnutrition. In this work, we show that the GM15 model provides increased reproducibility and robustness of preclinical studies by limiting the confounding effect of fluctuation in microbiota composition, and offers opportunities for research focused on how the microbiota shapes host physiology in health and disease.
The intestinal microbiota is a complex and dynamic community largely composed of bacteria whose activity profoundly influences our health and diseases. Advances in sequencing and analytical technologies coupled with improved computing tools have revolutionized the field of host–microbiota interaction. 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–microbiota interaction phenomena, the underlying causal mechanisms remain largely elusive.

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. Owing 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. While the use of defined genetic backgrounds, as well as the absence of specific pathogens, is now a common practice in mouse studies, an important confounding factor is variability in the composition of the intestinal microbiota, which can cause marked phenotypical variations among experimental animals and animal facilities. This parameter is under the influence of multiple elements such as genetics, diets, biological rhythms, and breeding conditions. As a consequence, to restrain microbial diversity, efforts have been made to tailor protocols for microbiota-related mouse studies and to standardize mouse microbiota composition.

The mouse gut microbiota richness is usually estimated at more than 300 bacterial genera and common inhabitants of the mouse intestine belong to seven bacterial phyla with Firmicutes, Bacteroidetes, and Proteobacteria being the most abundant ones. The first prototype in standardizing the mouse microbial environment (initially to study immunocompromised mouse models) arose in the 1960s with the wide implementation of the specific pathogen-free (SPF) hygienic status of mouse husbandries. Nowadays, SPF animals are obtained by re-deriving mouse strains by two-cell-stage embryo transfer to SPF recipients and subsequent postnatal inoculation with a cocktail of bacteria devoid of pathogens to homogenize microbial colonization within a given animal facility. SPF, and then specific (and opportunistic) pathogen-free (SOPF) inbred lines (lacking specific opportunistic pathogens such as Staphylococcus aureus or Pseudomonas aeruginosa) now represents the common health standard for experimental mouse breeding. However, despite the global efforts in standardizing the SPF procedures, the undefined nature and important interindividual 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. Indeed, the microbiota fluctuates a lot with diet and environment, so it is 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. Recent efforts have been put into isolating, cultivating, and archiving isolated cultures of the dominant members of the mouse microbiota. Gnotobiotic animals can be kept in isolators for several generations and offer the possibility of strict control of their microbial status.

Gnotobiotic models offering a different degree of microbial complexity have been developed in the past ranging from monocolonization (monoxenic animals) to high diversity microbiota models such as conventionalized ex-GF animals using a donor microbiota. Two models have emerged for breeding and long-term experimental purpose: the Altered Schaedler Flora (ASF) model and the recent Oligo-Mouse-Microbiota model. 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 1970s 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. The ASF is composed of eight defined bacterial strains, which are stable over mouse generations. Most immune parameters are normalized in mice colonized with ASF when compared with SPF mice, but the fact that strains of the ASF model are not publicly available and that they are not all representative of the dominant members of the mouse microbiota remains an important limitation. In addition, ASF mice differ substantially from SPF mice with respect to intestinal microbial biochemical activities and resistance to opportunistic pathogen colonization, probably owing to the limited phylogenetic diversity and metabolic capabilities of the ASF consortium.

Recently, the Oligo-MM model was developed. It is a minimal microbiota gnotobiotic model composed of 12 defined cultivable mouse commensal bacteria from the miBC collection representing the major bacterial phyla of the mouse gut.

The community is transmissible and stable over consecutive mouse generations and animal facilities and was recently used to reveal that the intestinal microbiota adapts to environmental changes by short-term effects of transcriptional reprogramming and adjustments in sub-strain proportions and long-term genomic positive selection. Functionally, and unlike ASF, Oligo-MM offers colonization resistance against Salmonella enterica serovar Typhimurium when supplemented with facultative anaerobic bacteria including Escherichia coli.

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 and SPF mice under standard dietary conditions in two different animal facilities. We demonstrate that under conditions of chronic physiological stress such as postweaning malnutrition on a low-protein diet, a dietary condition triggering stunting, GM15 microbiota shows improved capacities compared to a SOPF microbiota to buffer the deleterious effect of a depleted diet (DD) on mouse juvenile growth.

Results

In silico identification of the main bacterial families of a 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 four C57BL/6J SOPF mice (obtained from Charles River Laboratories, France) by whole-genome sequencing (WGS) (two females and two males, which were littermates but housed in different cages from weaning at 3 weeks old, and feces were collected at 2 months old in our facility). An average of 13.4 million paired-end reads was obtained per sample with a length of 300 bp. The metagenomics data sets generated were classified using the Centrifuge software and compared to the RefSeq
complete genome database, 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 (Supplementary Fig. 1a, b). 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 coccoides, Lactobacillus johnsonii, L. lactobacillus murinus, and Limosilactobacillus reuteri were previously identified as mouse-enriched and dominant intestinal bacteria,19,21,22 Limosilactobacillus murinus, Parabacteroides goldsteinii, and Clostridium strains are part of the ASF model, and strains of Enterocloster clostridioformis, Limosilactobacillus reuteri and Bacteroides caccae are part of the Oligo-MM model.

Isolation and taxonomic characterization of the GM15 bacterial strains. We established four different strategies in order to isolate and culture a maximal number of representative strains of the 20 dominant bacterial families identified by 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 cecal 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 prescreened by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) for dereplication prior to the first taxonomic identification by 16S ribosomal RNA (rRNA) gene Sanger sequencing. We selected 11 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 bacterial families found in C57BL/6J SOPF animals (Fig. 1a, b). Actually, at the family level, the GM15 consortium putatively covers 63% of the SOPF consortium, which is more than the putative coverage previously available with ASF and Oligo-MM models covering, respectively, 58% and 48%. 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 Erysipelotraceae, one strain of Ruminococcaceae, and one strain of Enterobacteriaceae (Supplementary Fig. 1c). We recently reported the draft genomes of the 15 strains and alignment against the NCBI database allowed the identification of 12 strains at the species level and 3 strains at the family level, which are described as novel taxa in this study.

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 nonredundant KO groups covering 44% of all protein-coding sequences. Besides, all GM15 strains possessed 3 to 64 unique functions, although E. coli Mt1B1 exhibited a vast repertoire of unique KO groups (105), 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 nonunique 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 (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, clustered together. In addition, Lachnospiraceae are clustered with Ruminococcaceae and Erysipelotraceae, which include mainly bacteria with sporulation capabilities. Finally, Bacteroides are clustered with Parabacteroides, whose species are predominant in the colonic mucus barrier and promote mucinase activity, and are significant producers of succinate, a major metabolic by-product. Thus, different strains of a bacteria family and of other closely related bacterial families are capable of the same enzymatic activities in the gut. This can be essential for the generation of simplified nonspecific 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 (covering 47% of all protein-encoding sequences) found in our initial metagenomic analysis (Fig. 1a). In addition, the KEGG modules from the Oligo-MM and ASF microbiota (covering, respectively, 46% and 43% protein-coding sequences) were included for comparative analysis, as these consortia were previously used to generate gnotobiotic mice with a stable and defined mouse-derived microbiota. The presence and completeness of KEGG modules were 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 and Supplementary Data 1), including 140 pathways. We also identified clusters of modules that were not represented among the ASF and Oligo-MM consortia, but specifically common to GM15 and SOPF (Fig. 1e, clusters 9 and 6 and Supplementary Data 1), comprising a total of 155 pathways indicating that qualitatively the GM15 metagenome covers functionalities found in SOPF microbiota that was lacking in ASF and Oligo-MM models. Quantitatively, the defined consortia of GM15, Oligo-MM, 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 Oligo-MM 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 used our SOPF colony to produce GF mice and next-generated GM15 gnotobiotic 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 (Supplementary Table 1). Fecal samples from GF and SOPF mice were used as negative and positive controls. Only Enterocloster clostridiformis YL32, which was obtained from the DSMZ miBC collection, was not detected in
Fig. 1 Selection, isolation, and functional analysis of GM15 strains. a Representative and prevalent bacterial families identified from fecal samples of 4 C57BL/6J mice of our colony. b Eleven bacterial strains were isolated from fecal or cecal samples of either C57BL/6J or ASF mice using various culture methods. Four additional strains were obtained from the DSMZ collection. c Unique KO (KEGG Orthology) groups are those present in only one GM15 member. The unique contribution of each GM15 member is relatively evenly distributed (<2%), except E. coli whose unique functions contribute for 33%, although nonunique KO groups are primarily responsible for the GM15’s metagenomic profile with 59% of the contribution. d Heatmap of hierarchical clustering of enzymatic activities in the gut for the 15 strains of the GM15 model. In quorum sensing (QS) function, only effector proteins were screened (because receptors suffer from similarities with nonspecific QS receptors). e Heatmap of hierarchical clustering of KEGG module distribution in the metagenomes of GM15, SOPF, ASF, and Oligo-MM12 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 Supplementary Data 1. Source data are provided as a Source Data file.
our SOPF mice. Co-monitoring of specific and total bacteria aimed to detect any bacterial load imbalance caused by contamination in gnotobiotic isolators.

Eight-week-old GF mice (five breeding pairs; GM15 founders, F0; Supplementary Fig. 2a) were inoculated by oral gavage with a fresh frozen mixture of the 15 strains and bred in sterilized positive pressure isolators up to the F9 filial generation. All strains except Subtilibacillus caecimuris MD335, Longibacillus caecimuris MD329, and Irregularicoccus caecimuris 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 nine consecutive generations; Fig. 2a and Supplementary Fig. 3a). Anaerotruncus colihominis JM4-15, Clostridium sp. MD294, Clostridium sp. MD300, and Clostridium cocleatum I50 were occasionally below the detection limit, but the detection of the remaining strains was reproducible between fecal and cecal samples from individual mice in the second filial generation (Fig. 2a and Supplementary Fig. 3b). Taken together, these results indicate 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 three remaining strains: Subtilibacillus caecimuris MD335, Longibacillus cacemuris MD329, and Irregulararcoccus cacemuris MD308, either did not efficiently colonize the animals, live in the cecum 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 (Supplementary Fig. 2a). 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 (BD) with 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 BD. 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 BD (Supplementary Fig. 2a). We did not detect any significant changes in the GM15 composition under these conditions (Fig. 2c). In addition, 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 weeks 1, 2, and 3 posttransplantation (Supplementary Fig. 2a and Fig. 2d).

When all the data are analyzed together, we notice a limited fluctuation of maximum 2 Log10-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 cecum, along with physiological and metabolic differences. 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. Indeed, the distribution of progeny per litter in GM15 mice ranged more evenly from 3 to 9 pups compared to SOPF mice, whose mean number of pups per litter was mostly centered on 7 or 8 pups. 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 postnatal 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 cecum 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). The 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 contributes to somatic tissue growth 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. We thus profiled the basal immune parameters of GM15 animals and compared them to SOPF and GF animals at 7–8 weeks of age. For that purpose, we analyzed immunoglobulin (Ig) levels in feces and sera and cytokine levels in the serum. In parallel, we assessed viable leukocytes by measuring CD45+ cell count and analyzed T cells, B cells, natural killer (NK) cells, and monocytes or dendritic cell (DC) population frequencies in whole blood and several lymphoid organs (spleen, thymus, Pey er’s patches (PPs) and mesenteric lymph nodes (MLNs)).

Among already described immunodeficiencies in GF mice, we observed that the GM15 community restored the production of IgA in the gut and in serum, and IgG2b in the serum, at levels equivalent to those detected in SOPF mice (Fig. 4a–c). 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. In addition, circulating interleukin-22 (IL-22) levels, which is one of the key intestinal cytokines, were restored as well in GM15 mice (Fig. 4d). Furthermore, it has been previously shown that isolated lymphoid structures such as PPs and associated cellularity were strongly increased with microbiota diversity, whereas the total cell numbers in MLNs were comparable between GF and SPF mice. As described in the literature, the number of PPs collected was higher in SOPF mice compared to GF and partially restored in GM15 mice (Fig. 4e).
No or minor differences were observed between the three groups for CD45<sup>+</sup> cell count (Fig. 4f and Supplementary Fig. 4a, b). A slight increase was observed in CD45<sup>+</sup> cell count in SOPF compared to GF mice with apparent intermediate CD45<sup>+</sup> cell count in GM15 mice in the spleen, thymus, and PPs and with no difference in MLNs (Fig. 4g–i and Supplementary Fig. 4a, b). 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<sup>41</sup>. Interestingly, the great majority of populations impacted showed similar frequencies in GM15 mice and SOPF compared to GF mice (Supplementary Figs. 4a–f and 5a–e). Monocytes in whole blood, CD4<sup>+</sup> T cells in the thymus, CD8<sup>+</sup> T cells in MLNs, DCs in PPs, B cells in MLNs, and whole blood and NK cells in the spleen were restored in GM15 mice. Among those cell populations, only B cells in GF whole blood showed higher concentrations compared to those described in the literature<sup>40</sup>. Our observations of NK cell decrease in the spleen and DC increase in PPs in GF mice compared to SOPF mice were not yet described.

Minor differences were observed between GM15 mice and SOPF, but not with GF mice for DC in the spleen and CD4<sup>+</sup> T cells in PPs, whereas both populations were described to be decreased in GF mice (Supplementary Figs. 4a–e and 5b, e). Only NK cells in PPs showed similar and increased frequencies in GM15 and GF mice compared to SOPF mice (Supplementary Figs. 4a–e and 5e). In the same way, minor differences were
observed as well for some serum Ig levels. In our model, 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. 4j, k). Of note, those Ig levels were lower in our SOPF mice than levels already described in SPF mice.47,48

Surprisingly, the first set of GM15 mice (F1 and F2) born to a colonized mother presented elevated levels of IgE in serum compared to GF and SOPF mice (Fig. 4l). Such IgE levels may result from parasitic infections, immunodeficiencies, or long-lived IgE-producing plasma cells generated by food antigens in mice.
with low-diversity microbiota during early life47,52,53. Of note, unlike the previous study, our GF mice did not show increased IgE levels compared to SOPF47. 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. To confirm this observation, we sampled again serum from additional GM15 animals (F5 and F6) several months after the initial test and quantified the IgE circulating levels with two independent methods. Although both Luminex and ELISA assays were calibrated against purified mouse IgE, they provided very different ranges of IgE values. However, both methods indicate that this second set of GM15 animals had normal levels of circulating IgE (Fig. 4m and Supplementary Fig. 5f). We also found that circulating IL-17a levels were increased in the first set of GM15 mice compared to SOPF and GF mice (Fig. 4n), but the IL-17a levels were much lower in the second set of GM15 animals and ultimately similar to SOPF mice, apart from one outlier (Fig. 4o).

Taken collectively, our results indicate that the GM15 community is sufficient to restore some key parameters of the immune response lacking in GF animals such as modified levels of serum and fecal IgA, serum IgG2b, serum IL-22, PP number, CD45+ cell count, and immune cell population frequencies in different organs close to that detected in SOPF animals.

Low-complexity GM15 microbiota shares more metabolic traits with SOPF than GF mice. The gut microbiota influences multiple host metabolic pathways by providing metabolites to its host and also shapes interorgan communication within the body by influencing the production and activity of endocrine signals54,55. One-dimensional proton nuclear magnetic resonance spectrometry (1H NMR) was previously applied to investigate how the gut microbiota impacts host metabolism using mouse models56,57, or human cohorts58. Using this technology, we analyzed the metabolic profile of plasma samples from GF, GM15, and SOPF mice and were able to quantify a total of 57 polar metabolites and 5 nonpolar metabolites. It could be seen that only a few of them were significantly affected by sex (Supplementary Data 2). A principal component analysis (PCA) based on quantified polar metabolites showed sequential alignment on the first principal component (9.2% of total variance) according to microbiota complexity of GF, GM15, and SOPF samples, independently of F1 and F2 generations (Fig. 5a). The metabolite composition of GF plasma resulted only from the host metabolic activity and represented a basal phenotype. Conversely, the SOPF mice harboring a very diverse gut microbiota, exhibited a larger panel of metabolic activities based on host–bacterial and bacterial–bacterial interactions. In between, the GM15 low-complexity community contributed to a lower extent to the metabolic phenotype. We performed the equivalent statistical analysis using binned NMR spectra instead of quantified metabolites and as expected we obtained the same results (Supplementary Fig. 6). Next, we performed a discriminant analysis across all samples to highlight the specific metabolic signatures of each group. Nine metabolites emphasized significant variation, although it is noteworthy that the calculated distance placed GM15 closer to SOPF than to GF (Fig. 5b). Dimethylamine and isopropanol tended toward GF mice, but concentrations were low, and taken alone there was no difference between groups (Supplementary Data 2). As previously reported, GF mice had higher plasma levels of glycine59 and reduced plasma acetate concentration60. On the contrary, SOPF mice harboring a diverse microbiota had higher plasma levels of acetate, dimethyl sulfone, 3-hydroxybutyrate, and propionate54,61. As for GM15 mice, less methanol and more citrate were detected. Interestingly, methanol may occur as a result of fermentation by gut bacteria and can stimulate citric acid fermentation52,63. Thus, the simplified gut microbiota of GM15 mice may produce less methanol and/or microbially produced methanol may be used to form citrate. 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 nonpolar metabolites indicated that GM15 colonization was sufficient to reduce free cholesterol and phosphatidylcholine as observed in SOPF mice (Fig. 5c). The three additional nonpolar metabolites detected were equivalent in all mice (Supplementary Data 2).

Finally, we investigated the circulating levels of key metabolic hormones and growth factors. To allow different types of analyses using the same blood sample, mice were not fasted. Most interestingly, 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) to the levels observed in SOPF animals. IGF-1 is an essential growth factor promoting systemic and tissue growth64,65 and this observation correlates well with the improved macroscopic growth of GM15 mice compared to GF animals (Figs. 5d and 3e–h). Corticosterone levels, which are high in GF and low in SOPF animals, were also normalized by GM15 bacterial colonization66 (Fig. 5e), indicating that the 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 male mice are less sensitive to diet-induced stunting than SOPF mice. Previous work by our lab and others has shown that the gut microbiota influences pathogenesis associated with chronic undernutrition, particularly diet-induced stunting. Consequently, we sought to study our newly established gnotobiotic mouse model under severe nutritional stress, induced by a nutrient DD (containing 4% protein and 2% lipids) expected to trigger stunting. Based on our previous experience with the juvenile chronic undernutrition model, diet-induced stunting is an important 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 differences in other single organ sizes or weights 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 and Supplementary Fig. 8b–g). We tested this tendency by integrating all the phenotypical parameters in a PCA. We first established which parameters were correlated to the phenotype (Supplementary Fig. 8h), thus excluding brown adipose tissue from the analysis as a noncorrelating 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 a SOPF microbiota. As IGF-1 is an important driver of the diet and microbiota-mediated growth promotion, we assessed IGF-1 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 IGF-1 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 when...
compared to animals bearing a complex SOPF microbiota under non-stressful conditions. However, the GM15 microbiota seems partially protective against the deleterious effects of chronic undernutrition, as compared to SOPF mice.

The GM15 model growth and diet-induced stunting phenotypes are reproducible in another gnotobiotic facility. Reproducibility is imperative to preclinical research, especially in the microbiome field. Therefore, to test the transferability and
Fig. 6 GM15 mice and SOPF response under postweaning chronic undernutrition. a-h Dot plots where dots, lines, and error bars represent respectively individual mice, means, and SEM. Male (M), breeding diet (BD), and depleted diet (DD), respectively. Body weight (a) and body size (b) were measured every week from weaning to postnatal day 56 (ten males per group). e Representative picture of the cecum, subcutaneous fat, and right femur of mice at day 56. Tibia length (nine males per group) (d), liver (ten males per group, except nine GM15_M_DD) (e), and epididymal white adipose tissue (WAT) (ten males per group, except nine GM15_M_DD) (f) weight were measured. g PCA of tissues weight and bones size under control conditions (BD) or under nutritional stress (DD). Brown adipose tissue was excluded, given its lack of correlation with other parameters (Supplementary Fig. 8h). Percentages on each axis indicate the variance associated with each coordinate (nine males per group). h Serum IGF-1 at day 56 (ten and nine males per BD and DD groups, respectively). P values after two-way ANOVA were adjusted for Sidak’s post hoc test for multiple comparisons. *P < 0.05, **P < 0.01, and ***P < 0.001. Source data are provided as a Source Data file.

reproducibility of the GM15 model, we established it in a second gnotobiotic facility (Facility 2—Laboratory of Gnotobiology at Institute of Microbiology of the Czech Academy of Sciences, Novy Hradek, Czech Republic). We colonized GF animals with the GM15 consortia using the same protocol described for facility 1. Using qPCR microfluidic assays, we analyzed the bacterial titers in fecal samples of the GM15 animals and control SPF animals from the same facility, bred in the same condition as GM15 animals (Supplementary Fig. 2b). Similarly to GM15 colonized mice at facility 1, all strains except Subtilibacillus caecimuris MD335, Longibacillus caecimuris MD329, and L. caecimuris MD308 were above the detection limit in GM15 individual mice (Fig. 7a). GM15 strain levels in facilities 1 and 2 were very similar (Fig. 7b), demonstrating that the GM15 model can be effectively transferred and established in a reproducible manner in different gnotobiotic facilities.

We next compared the titers of each strain individually in GM15 and SOPF/SPF animals from both facilities. Besides Clostridium sp. MD294 and Clostridium sp. MD300, which are less well-colonizing GM15 animals than SPF/SPF animals in both facilities, there was a clear general tendency of more similar fecal titers in GM15 animals as compared to their titers in SPF/SPF animals between facilities (Fig. 7b and Supplementary Table 2) suggestive of more reproducible microbiota in GM15 animals. As a control gnotobiotic mouse model with a significantly different microbiota from GM15 (Fig. 1e), we generated Oligo-MM12 animals in facility 2. As previously reported, Oligo-MM12 gnotobiotic animals were successfully established in facility 2 and the Bifidobacterium animalis YL2 and Acutalibacter muris KB18 strains were again not detected (Supplementary Fig. 9 and Supplementary Table 3).

Then, we conducted a comparative study to evaluate the growth phenotypes of SPF, Oligo-MM12, and GM15 mice in facility 2 on standard BD. Both GM15 and Oligo-MM12 mice grew well compared to SPF mice (Fig. 8a, e), with Oligo-MM12 females gaining weight slightly faster than GM15 females (Fig. 8a). Importantly, GM15 growth phenotypes (mean daily body weight and size gains) were more reproducible between facilities 1 and 2 than for SOPF/SPF animals, which significantly differed between the facilities (Fig. 8b, f).

These macroscopic observations correlated well with internal organ size: the weights of GM15 animals’ brain, kidneys, liver, spleen, muscles, and adipose tissues were larger or equivalent in females (Supplementary Fig. 10a) or larger in males (Supplementary Fig. 10b) than those of SPF mice. Of note, female Oligo-MM12 mice had less white adipose tissue (WAT) and more brown adipose tissue (BAT) than GM15 and SPF females. In GM15 males, the WAT was significantly heavier as compared to SPF and Oligo-MM12 males. This observation was different compared to the results from facility 1, where GM15 males’ visceral WAT was lighter compared to SOPF males on BD (Supplementary Fig. 8b). We observed no difference regarding the BAT among male groups, both in facility 1 and facility 2. As expected, both gnotobiotic models had a slight cecum enlargement compared to SPF animals, yet their cecum size was significantly reduced as compared to one of GF animals (Supplementary Fig. 10a, b; and see Fig. 3g for GF). Finally, bone size was increased in both gnotobiotic models of both sexes compared to SPF animals with a more pronounced and statistically robust effect in GM15 mice (Fig. 8c, g).

Taken together these results establish that the macroscopic phenotypes of GM15 mice identified in facility 2 are similar to those from facility 1. Importantly, when comparing the results from the two facilities, we could also establish that GM15 mice growth phenotypes are more reproducible between facilities than SOPF/SPF animals. Our results also confirm that gnotobiotic models carrying a minimal microbiota compensate effectively the physiological limitations of GF mice and mimic (when not improving) the macroscopic growth phenotypes of SOPF/SPF animals. Such results correlate well with the endocrine markers of growth IGF-1 and IGFBP-3 (Fig. 8d, h) and are not due to a difference in relative food intake between conditions (Supplementary Fig. 10c). Finally, our results also suggest that each gnotobiotic model carries its own phenotypical characteristics at the steady state: a specific lower WAT-higher BAT phenotype in Oligo-MM12 females and an increased bone growth in GM15 animals.

In facility 1, we reported that GM15 microbiota is partially protective against the deleterious effects of chronic undernutrition, as compared to SOPF mice (Fig. 6). To test the reproducibility of this phenotype in another facility, we submitted GM15 males from facility 2 to the same severe nutritional stress, by using the same nutrient DD as in facility 1 shown to trigger severe stunting after weaning (Fig. 6). This time, we also tested Oligo-MM12 males under the same dietary regime. As in facility 1, all animals developed a macroscopic stunting phenotype and a related bone and endocrine markers alterations (Fig. 8e–h). Interestingly, as in facility 1, and despite a marked stunting phenotype, we detected an improved body weight and body size gain of GM15 males on the DD compared to SPF males (Fig. 8e, f). This observation is confirmed by an increased size of bones (Fig. 8g) and IGF-1 and IGFBP-3 levels in sera (Fig. 8h) in GM15 males as compared to SPF animals. Of note, on the DD, Oligo-MM12 males also performed significantly better than SPF animals, a very similar phenotype to GM15 males, suggesting that this buffering effect is not a unique attribute of the GM15 model.

The GM15 model immune and metabolic phenotypes are reproducible in another gnotobiotic facility. To follow up on our initial phenotyping of animals from facility 1, we studied specific immune and metabolic parameters of SPF, Oligo-MM12, and GM15 animals from facility 2. As previously reported in facility 1, the GM15 community supports IgA (both serum and fecal levels), IgG2b, IgG3, and IgE production (Supplementary Fig. 11a). Of note, we confirmed in facility 2 the similar IgE levels between GM15 and SPF animals using both Luminex and ELISA methods (Supplementary Fig. 11a, b). GM15 also supports
increased levels of IgM and IgG2a and supports circulating levels of IL-22 and IL-17a as in SPF animals (Supplementary Fig. 11a, c). Collectively, these results confirm, in another facility, that the GM15 community is sufficient to restore key parameters of immune system development and maturation known to be altered in GF animals. Finally, while comparing GM15 and Oligo-MM12 models, the GM15 consortium seemed more prone to support Ig (fecal and serum IgA, serum IgG2a, IgG2a, IgG3, and IgM) and IL (IL-17a and IL-22) production. Interestingly, these features are regulated upon recognition of intestinal microbiota45,47,48,69. Finally, a particularity of the Oligo-MM12 model was its lower circulating levels of IgE, IgG2a, serum IgA, and IL-22 compared to SPF and GM15 mice, illustrating a difference in the immune phenotype of the two gnotobiotic models.

To explore the metabolic profiles of GM15, SPF, and Oligo-MM12 animals in facility 2, we conducted 1H NMR on plasma samples as previously performed on the plasma of facility 1 animals. We were able to quantify a total of 54 polar metabolites and 5 nonpolar metabolites (Supplementary Data 2). PCA based on quantified polar metabolites showed a sequential alignment according to sample type (GM15, SPF, or Oligo-MM12) on the first and second principal components (9.6% and 7.1% of total variance respectively, and 6.3% on the third component; Supplementary Fig. 11d) and calculated distance placed SPF...
**ARTICLE**

**NATURE COMMUNICATIONS** | **https://doi.org/10.1038/s41467-021-26963-9** | www.nature.com/naturecommunications

---

**Fig. 8** The GM15 model growth and diet-induced stunting phenotypes are reproducible in another gnotobiotic facility. **a-h** Macroscopic phenotype of mice where dots, lines, and error bars represent, respectively, individual mice, means, and SEM. Female (F) and male (M), breeding diet (BD), and depleted diet (DD), respectively. **a, e** Body weight or size curves and mean daily gain of animals in facility 2. One-way ANOVA followed by Tukey’s multiple comparison analyses (GM15 mice: 9 F, 9 M_BD, 11 M_DD; SPF mice: 12 F, 12 M_BD, 13 M_DD; Oligo-MM12 mice: 9 F, 8 M_BD, 9 M_DD) and 2 (GM15 mice: 9 F, 9 M_BD, 11 M_DD; SPF mice: 12 F, 12 M_BD, 13 M_DD). One-way ANOVA followed by Tukey’s multiple comparison analyses. **c, g** One-way ANOVA analyses of femur and tibia length of animals in facility 2. One-way ANOVA followed by Dunn’s and Tukey’s multiple comparison analyses for F and M data sets, respectively (GM15 mice: 9 F, 9 M_BD, 11 M_DD; SPF mice: 12 F, 12 M_BD, 13 M_DD; Oligo-MM12 mice: 9 F for femur and 8 F for tibia, 8 M_BD, 9 M_DD). **d, h** Serum IGF-1 and IGFB3 of animals at day 56 in facility 2. One-way ANOVA followed by Tukey’s multiple comparison analyses (GM15 mice: 9 F, 9 M_BD, 10 M_DD; SPF mice: 10 F, 12 M_BD, 12 M_DD; Oligo-MM12 mice: 9 F, 8 M_BD, 9 M_DD). *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. Source data are provided as a Source Data file.
sample quasi-equidistant to GM15 and Oligo-MM\textsuperscript{12} samples, with each gnotobiotic model showing a distinct metabolic phenotype strongly resembling that of SPF animals. Polar metabolites were more prone to sex and group effects in facility 2 than in facility 1, but lipids were stable, and male to female difference \((\text{fold change} > 1.3 \text{ and } p \text{ value} < 0.05)\) was only seen in triacylglyceride in SPF mice from facility 2 (and in GF mice from facility 1). Like in facility 1, GM15 and SPF mice shared many similarities and a lower methanol/citrate ratio was again observed in GM15 mice. Metabolite profiles of GM15 and Oligo-MM\textsuperscript{12} mice were also similar; nonpolar metabolites were equivalent in all mice except triacylglyceride, which was higher in SPF mice. Finally, we noticed as many differences between GM15 animals as between SOPF and SPF mice across the two facilities. Such variability might be due to the different BDs used in the two facilities, and the fact that mice were housed in isolators in facility 1 rather than in individually ventilated cages in facility 2.

Taken collectively, these results obtained in a second gnotobiotic facility confirm that gnotobiotic models carrying minimal microbial consortia such as the GM15 community mimic SPF/SOPF animal phenotypes at steady states.

**Discussion**

Here we describe GM15, a simplified and controlled murine gut microbiota and its related C57BL/6j gnotobiotic mouse model. The GM15 community is composed of pure cultures of 15 strains from 7 of the 20 most prevalent bacterial families present in the fecal microbiota of C57BL/6j SPF 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 nine filial generations, upon mild dietary fluctuations, can be transmitted efficiently by FMT, and can be re-established efficiently in another gnotobiotic facility. GM15 compensates breeding, growth, immune, endocrine, and metabolic limitations of GF mice and recapitulates many SOPF animals’ phenotypical features in a reproducible manner in two gnotobiotic facilities.

As all experimental models, GM15 has its limitations and microbiota standardization, which is essential for establishing robust causal relationships between a microbiota configuration and a host trait, may decrease the translational potential of the observation. 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 compared to SOPF microbiota (15 strains versus hundreds of species in SOPF animals). The specific immune and metabolic signatures of the GM15 model pave the way to further studies aiming at defining if the presence or absence of specific community members triggers these phenotypes. Besides, the GM15 model offers insights for gnotobiotic research as a complementary model to other murine microbiota performs slightly better than a complex SOPF community in terms of the physiological host response. We have previously identified a lactobacilli strain that is capable of buffering the deleterious effects of such challenge in monocolonized mice\textsuperscript{23}. Interestingly, out of its 15 strains, GM15 microbiota contains 3 lactobacilli strains. Further genetic manipulations coupled to gnotobiotic studies focusing on modifying the lactobacilli components of GM15 may pave the way to understanding how this minimal bacterial community buffers the host response to chronic undernutrition.

In conclusion, our study establishes that the GM15 model offers possibilities for preclinical research focusing on host–microbe 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 cecal 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\textsubscript{2}, 5% H\textsubscript{2}, 5% CO\textsubscript{2}). Lactobacillus johnsonii MD060 was isolated on MRS agar. *Ligilactobacillus munis* MD040 and *Parabacteroides goldsteinii* MD072 were isolated on Columbia CNA agar with 5% sheep blood. *Bacteroides acidifaciens* MD185 and *Irregularococcus caeimirsi* MD308 were isolated on GAM agar. *Bacteroides cacimurii* MD237 and *Limonilactobacillus reuteri* MD207 were isolated on GAM agar supplemented, respectively, with 32 μg/mL vancomycin and 32 μg/mL erythromycin. *Subtillicoccus caeimirsi* MD335 and *Longibacillus caeimirsi* MD329 were isolated on M2GSC agar (modified Hobson, containing (per 100 mL) 1 g of casitone, 0.25 g of...
of yeast extract, 0.4 g of NaHCO₃, 0.2 g of glucose, 0.2 g of cellulose, 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 (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 of agar). Subtilibacillus cacticum MM335, Longibacillum caecimuris MD329, and Irregularibacillus caecimuris MD308 were isolated from cecal 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. FMT in GF mice and fresh cecal content pellets were resuspended in reduced broth media for direct dilution plating on agar plates and growth at 37 °C under an anaerobic atmosphere (90% N₂, 5% H₂, 5% CO₂). Clostridium sp. MD294 and Clostridium sp. MD300 were isolated on M2G agar, respectively, from cecal content and fecal pellets of ASF mice. For identification of isolates, colonies were prescreened for dereplication by MALDI-TOF MS (Vitek MS, Biomerieux) according to the manufacturer’s instructions and database enrichment using RUO mode. Then, genomic DNA (gDNA) was extracted from pure cultures and analyzed by 16S rRNA gene sequencing (GATC Biotech). Based on the genus Longibacillum muris, strain MD329 was assigned to a novel genus and the G+C content of genomic DNA is 39.5 mol%. The type species of this genus is Longibacillum muris. No antibiotic resistance genes were identified within the genome and the G + C content of genomic DNA is 42.7 mol%. The type species of this proposed genus is Longibacillum caecimuris.

**Description of** Irregularibacillus caecimuris sp. nov. Irregularibacillus caecimuris (L. n. cecum, cecum; L. n. muris, mouse; N.L. gen. n. caecimuris, from the cecum of a mouse). The cells are rods that separate into cocci forms, and are strictly anaerobic.

The number of CAZymes identified within the genome was 226, facilitating the predicted utilization of cellulose and starch. KEGG analysis identified pathways for acetate production from acetyl-CoA (EC:2.3.1.8, 2.7.2.1), propionate production from pyropropionyl-CoA (EC:2.3.1.8, 2.7.2.1), sulfide and l-serine utilized to produce 1-cysteine and acetate (EC:2.3.1.30, 2.5.1.47), l-glutamate production from ammonia via l-glutamine (EC:6.3.1.2, 1.4.1.–), and folate biosynthesis from 7,8-dihydrololate (EC:1.5.1.13). This species was most commonly identified within mouse gut samples (60%), although subdominant at 0.69% mean relative abundance.

**Culture conditions.** Freshly grown cultures of individual bacterial strains were supplemented with 20% glycerol prior to cryopreservation at −80 °C. Each culture was systematically validated for culture purity and identity by MALDI-TOF. Culture media and material were introduced in the anaerobic chamber at least 2 h before use. Anaerobic culture strains were grown on K supplemented Clostridium coccoides medium. For mouse colonization and absolute quantification of bacteria, a single colony of each of the 15 bacterial strains was grown on solid agar (laminar), amplified at 37 °C, and immediately used for mouse colonization. Each bacterial barrel was stored at −20 °C for gDNA extraction, and the rest was used for enumeration by plating dilution plating on agar plates. A frozen mixture of the GM15 bacterial community containing the 15 individual strains at an equivalent concentration (0.67E + 06 CFU (colony-forming unit)) was prepared to enable easy inoculation.

**Whole-genome sequencing.** DNA samples from the 15 bacterial cultures were prepared for WGS, using the Nextera XT DNA library preparation kit (Illumina, California, USA) according to the manufacturer’s recommendations. The resulting libraries were checked for their quality using the High-sensitivity DNA chip using the Agilent 2100 Bioanalyzer (Waldbronn, Germany) and quantified using the QuantFluor One dsDNA kit (Promega). Paired-end (2 × 300 bp) sequencing was performed on a MiSeq sequencer using the MiSeq v3 kit (600 cycles) (Illumina, USA, California).

**Novo genome assembly.** After a quality control with FastQC v.0.11.7, the paired-end reads were assembled de novo using the “As-miseq” assembly pipeline v20160825, comprising the following steps: adapter trimming, quality trimming and filtering, error correction, decontaminating, error correction, and scaffolding. The 15 de novo assemblies resulted in draft genomes composed of a few scaffolds (from 30 to 68) with high N50 values (from 13,099 to 943,892). Genomes were then ordered using Mauve v2.40 and annotated 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 v0.9.17. A k-mer counting strategy with the Centrifuge software v1.0.30 has been privileged to infer taxonomy, and reads were confronted to the RefSeq complete genome database Release 84, 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 translated into protein sequences using Prodigal v2.6.2. Marker genes were searched using the HMM package v3.0. Predicted protein sequences of genomes were blasted against the KEGG microbial database Release 2018-01-29, which contains 13 million protein sequences and trimmed with the following parameters: best hit with an expected value threshold <1E−5. The matrix obtained was consolidated into KEGG orthology (KO) counts (KO, which represent sets of genes that have sequence similarities and exert 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 or absence among genomes. The modules were analyzed for their completeness (direct FMT with four levels full, lack two, lack one, and only modules with a score of 3 or 4 were presented and integrated for their KO counts. The KEGG pathways were analyzed for their number of related KO counts.
MUMmer package v4.0 was used to perform pairwise alignment of the 15 genomes. Identification of both functions and communities was performed using Euclidian distance and a different distribution among the genomes, functional data were Because pathways and functions of interest did not have the same number of KO of presence among genomes has been analyzed in detail (Supplementary Data 1).

**gDNA extraction from cecal contents and fecal pellets.** GM15, SOFP/SPF mice: Cecal and fecal gDNA were extracted using the DNeasy® PowerLyzer® PowerSoil® Kit (Qiagen) following the manufacturer’s instructions with modifications. Samples (0.1 g) were homogenized in 1 mL of ZCL (Zipin, Milan, Italy) and 1 g of Proteinase K, incubated at 37°C for 1 h, and then boiled at 100°C for 5 min. After centrifugation for 5 min at 12,000 × g, the supernatant was treated with phenol-chloroform:isopropanol (1:1) and precipitated with ethanol. DNA was resuspended in 100 μL of sterile water, and the concentration was determined using a Nanodrop spectrophotometer.

**Quantitative PCR microfluidic assay.** In order to quantify specific and global bacteria load per g of cecal or fecal samples (wet weight), qPCR microfluidic assay was performed using respectively specific primers for GM15 and “universal” primers amplifying the genes encoding 16S rRNA from most bacteria groups (Supplementary Table 1).

**qPCR of bacterial 16S rRNA Genes.** qPCR was performed as follows: Oligo-MM12 strain-specific 16S rRNA primers (Supplementary Table 3) and hydrolysis probes were synthesized by IDT (Integrated DNA Technologies, USA). Standard curves using linearized plasmids containing the 16S rRNA gene sequence of the individual Oligo-MM12 strains were used for absolute quantification of 16S rRNA gene copies numbers of individual strains. 16S rRNA gene copies numbers were normalized to equal volumes of extracted DNA, assuming that DNA extraction is equally efficient between different samples. PCR reaction of 20 μL contained 0.2 μL of each primer, 0.5 μL of 40-fold diluted working solution of each probe, 1 μL of cDNA sample, 10 μL of 2× PowerUp® Master Mix (ThermoFisher Scientific), 1 μL of 10 μM of the corresponding probe, 6 μL of H₂O, and 2× PrimeTime® Gene Expression Master Mix (IDT) and 2.5 μL of gDNA of concentration 2 ng/μL. PCR conditions were 95°C for 3 min, followed by 45 cycles of 95°C for 15 s and 60°C for 1 min with fluorescence measurement. PCR reactions were performed on qTOWER® G touch (Analytik Jena) and analyzed using qPCRSof software v4.1 (Analytik Jena). Standards and samples were assayed in duplicate and monoplicate, respectively.
medium. For whole blood, spleen, and thymus samples, red blood cells were lysed by resuspension in 1 mL PharmLyse IX (Becton Dickinson) for 10 min. Cells were then spun at 400 x g for 5 min at room temperature. The lysing solution was discarded and cells were washed using 2 mL of PBS (Gibco). Cells were pelleted a second time and resuspended in PBS supplemented with 2% heat-inactivated fetal bovine serum (Sigma). Numeration and viability were determined using propidium iodide marker exclusion and MACSQUANT flow cytometer (Miltenyi). Cells were then stained with CD45 R/B220 antibodies. Fc receptors were blocked using solution at an antibody concentration of 100 μg/mL for organs and 100 μL/tube for whole blood, and analyzed by flow cytometry.

Whole blood was collected on a dry Eppendorf tube for serum analysis. Sera were obtained by centrifugation 2000 g x 15 min at 4 °C and stored at −20 °C before Luminex analysis. Feces were homogenized using 10 mL of filtered water and stored at −80 °C before Luminex analysis. Feces supernatant was obtained by disrupting 100 mg feces in 1 mL PBS-Protease Inhibitor IX (Sigma) using Lysing Matrix E Tube (MP Biomedical) and Fast Prep homogenizer (MP Biomedicals).

Samples were spun at 8000 x g for 15 min at 4 °C and supernatants were collected for IgA Luminex analysis.

Flow cytometry. A total of 10^6 cells or totality of cells for some PP samples were stained for surface markers. Leukocytes were stained using anti-CD3 BV421 (clone 145-2C11, #562600, 1:40) from Becton Dickinson, and anti-CD45 Viogreen (clone 30F11, #130-102-412, 1:40), anti-CD4 PE (clone REA604, #130-109-414, 1:20), anti-CD8a PE-Vio615 (clone REA601, #130-109-251, 1:10), anti-CD45R/B220 PE Vio770 (clone RA3-6B2, #130-102-308, 1:20), anti-CD335 APC (NKp46) (clone 29A.14.9, #130-102-347, 1:10), and anti-CD11b APC Vio770 (clone REA592, #130-109-288, 1:20) or anti-CD1c1 APC Vio770 (clone REA754, #130-110-704, 1:100) from Miltenyi Biotec GmbH. Viable cells were selected using Zombie Green Fixable Viability Kit (BioLegend #423111, 1:100). T cells were identified as CD3^+ and CD4^+ or CD8^+ cells; B cells were identified as CD3^+ , CD5^− and B220^+ cells. NKX cells were identified as CD3^+ , CD5^− , CD45R/B220^+ , and NKp46^+ cells. Monocytes were identified as CD45^+ , CD3^− , CD14^+ , CD20^− , and CD11c^+ in the spleen and whole blood samples. DCs were identified as CD1^+ , CD3^− , CD14^+ , CD20^− , and CD11c^+ in the spleen, PP, and MLN samples. Cells were analyzed using a MACSQuant Ten Flow cytometer with MACSQuantify software v2.8.1810.16380 (Miltenyi) and raw data were analyzed using FlowJo software v10.4.2 (Tree Star, Becton Dickinson). CD45^+ cell counts were normalized by performing the analysis on 1 M of viable cells that were stained. For CD4^+ cell count comparison in PPs, CD4^+ cell counts were performed on 1 M of viable cells that were stained when possible and on total viable cells when the cells were isolated; thus, no statistical analysis has been performed on these data. Data normalization was not possible for the whole blood sample, thus CD4^+ 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. Samples were assayed in monoplate.

Luminex metabolic, Ig, and cytokine panels analysis. Serum concentrations of Metabolite analysis (glyceraldehyde, insulin, urea) were determined using the Mouse Metabolic Magnetic Bead Panel Milliplex MAP kit (Millipore). Samples were not diluted and the assay was performed according to the manufacturer's instructions. Serum concentrations of Ig Panel (IgA, IgG1, IgG2a, IgG2b, IgG3, IgM, IgE) and IgA in feces supernatant was determined using the Mouse Ig Iso-otyping Magnetic Bead Panel Milliplex MAP kit (Millipore) and Mouse Ig Single-plex Magnetic Bead Milliplex MAP kit (Millipore). Samples were diluted 1:12,500 (IgA, IgG1, IgG2a, IgG2b, IgG3, IgM), 1:100 (IgE), and 1:100 (IgA in feces supernatant). Assays were performed according to the manufacturer’s instructions. Serum concentrations of Cytokine Panel (IL-1β, -10, -12p70, -17α, -17β, and -22) were determined using the Mouse Th17 Magnetic Bead Panel Milliplex MAP kit (Millipore). Samples were not diluted and the assay was performed according to the manufacturer’s instructions. Samples were assayed in monoplate. The known intratechnical error (%CV) per assay was 7% for insulin, 8% for glucose, and 10% for cholesterol ester quantification.

Metabolite analysis and quantification. The 1D 1H NMR spectra were acquired at 298 K with a 600 MHz Ascend (Avance III HD) spectrometer from Bruker (Spin System). Spectra were acquired with 5 mm QCI z-slip sample tubes, transferred into a 3-mm SampleJet NMR tubes. Polar metabolites were prepared using commercially available mouse plasma. Nonpolar metabolite extraction after ultrafiltration. Then, the lipoproteins on the 10 kDa filters were further diluted with 150 μL phosphate-buffered solution (1 M, pH = 7.4), mixed on thermoshaker Eppendorf (10 min, 10 °C, 750 r.p.m.), transferred to clean Eppendorf and extracted with 40 μL dichloromethane (1:2 v/v). Samples were centrifuged (5 min, 10 °C, 10,000 x g) for better phase separation. The dichloromethane layer was transferred in clean Eppendorf tubes and the aqueous phase was extracted again with dichloromethane. The pooled organic phase was evaporated under a nitrogen stream. The dry lipid residue was dissolved with 200 μL deuterated chloroform containing 0.03% TMS internal standard and 155 μL of the resulting solution was transferred in a 3-mm SampleJet NMR tube. The TMS concentration was calibrated by NMR using 1,3,5-tri-tert-butylbenzene standard solution and it was found to be 0.435 mM.

Sample preparation for metabolotyping. Polar metabolite preparation. It is known that the quantification of some plasma metabolites, such as tryptophan and tyrosine, is biased since they bind to albumin, and a highly abundant protein in plasma. In addition, protein precipitation methods with organic solvents can induce loss of volatile metabolites and overlay of residual broad resonances of lipids with some polar metabolites. Thus, the polar and nonpolar metabolites were separated from the same plasma sample according to previously described sequential approaches. First, we deprotei-
Metabolomics data analysis. NMR data preprocessing included normalization of analytical batch effects using internal standards, PCA analysis to assess the relative contribution of each effect of the data by measuring the Mahalanobis distance between each group and the associated significance using a Fisher test, confirmed by supplementary multivariate supervised analyses (partial least squares–discriminant analysis (PLS-DA)) performed using MATLAB v2019b and v2021a. Value under the limit of detection, when the signal to noise was below 3, was replaced by 0 at the NMR data quantification step with Chenomx NMR suite 8.6. Thus, no values were excluded during the statistical analysis, and there were no missing values treatment and transformations applied. Then, discriminant analyses were performed using the PLS-DA algorithm to extract metabolomics signatures. A variable selection algorithm based on Elastic-Net (MATLAB) was used to improve model performance by selecting the most significant metabolic signatures that explain the groups (GF, GM15, and SOFP). The statistical performances of the regression models were assessed using the balanced error rate with and without cross-validation (EC and CV-EC) and permutation tests (MATLAB). 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 group. 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 one- and two-way ANOVA. Phenotyping data impacted by age, filial generation, or sex were analyzed by the t test for multiple linear regressions (R v3.4.2), otherwise by one-way ANOVA followed by Tukey’s multiple comparison parametric test or Kruskal–Wallis and Dunn’s multiple comparison nonparametric test after D’Agostino and Pearson test for data set normality assessment (GraphPad Prism v7.05). ANOVA was performed using the PLS-DA algorithm to extract metabolomics signatures. A variable selection algorithm based on Elastic-Net (MATLAB) was used to improve model performance by selecting the most significant metabolic signatures that explain the groups (GF, GM15, and SOFP). The statistical performances of the regression models were assessed using the balanced error rate with and without cross-validation (EC and CV-EC) and permutation tests (MATLAB). 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 group. 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.

Data availability

 Anaerobrevum colliominus MJ4–15 (DSM-28734), Enterocaster clostridioformis YL32 (DSM-26114), Castridiom cocleatum IS50 (DSM-15511), and Escherichia coli M1B1 (DSM-28618) were obtained from DSMZ. All other bacterial strains isolated by BIOASTER, and GM15 mice, are assets controlled by BIOASTER and therefore distribution is managed by BIOASTER. Any request shall be sent to Marion Darnaud at the following email address: gnotobiology@bioaster.org. All demands will be examined and transfer of bacterial strains and/or GM15 mice will be possible under specific MTA conditions. 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 PRJNA551572. The list of KEGG modules and clusters referenced in this study, and obtained using KEGG microbial database Release 2018-01-29, are available in Supplementary Data 1. NMR raw data are accessible at the NIH Common Fund’s National Metabolomics Data Repository website, the Metabolomics Workbench, https://www.metabolomicsworkbench.org (supported by NIH grant U2C-DK119886), where it has been assigned Project ID PR001164. The data can be accessed directly via its project DOI: 10.21228/M8ET3G. Source data are provided with this paper.

Code availability

The code necessary to reproduce our metabolomics analysis is available at https://github.com/bioaster/biotracks-in-atlas.

Received: 21 February 2020; Accepted: 28 October 2021; Published online: 18 November 2021

References

1. Lynch, S. V. & Pedersen, O. The human intestinal microbiome in health and disease. N. Engl. J. Med. 375, 2369–2379 (2016).

2. Knight, R. et al. The microbiome and human biology. Annu. Rev. Genomics Hum. Genet. 18, 65–88 (2017).

3. Rik, E. M. The hoops, hoops, and types of human microbiome research. Yale J. Biol. Med. 89, 363–373 (2016).

4. Leulier, F. et al. Integrative physiology: at the crossroads of nutrition, microbiota, animal physiology, and human health. Cell Metab. 25, 522–534 (2017).

5. Nguyen, T. L. A., Vieira-Silva, S., Liston, A. & Raes, J. How informative is the mouse for human gut microbiota research? Dis. Model. Mech. 8, 1–16 (2015).

6. Mühler Convenor, M. et al. FELASA recommendations for the health monitoring of mouse, rat, hamster, guinea pig and rabbit colonies in breeding and experimental units. Lab. Anim. 48, 178–192 (2014).

7. Manantopoulou, M. et al. Nlrp6- and ASC-dependent inﬂammases do not shape the commensal gut microbiota composition. Immunity 47, 339–348 (2017).

8. Usar, S. et al. Interactions between gut microbiota, host genetics and diet modulate the predisposition to obesity and metabolic syndrome. Cell Metab. 22, 516–530 (2015).

9. Franklin, C. L. & Ericsson, A. C. Microbiota and reproducibility of rodent models. Lab. Anim. 46, 114–122 (2017).

10. Mooser, C., Gomez de Agüero, M. & Ganal-Vonarburg, S. C. Standardization in host-microbiota interaction studies: challenges, gnotobiology as a tool, and perspective. Curr. Opin. Microbiol. 44, 50–60 (2018).

11. Laukens, D., Brinkman, B. M., Raes, J., De Vos, M. & Vandenabeele, P. Heterogeneity of the gut microbiome in mice: guidelines for optimizing experimental design. PloS Microbiol. 14, 17–132 (2016).

12. Steppenbeck, T. S. & Virgin, H. W. Accounting for reciprocal host-microbiome interactions in experimental science. Nature 534, 191–196 (2016).

13. Clavel, T., Lagkovardos, I., Blaut, M. & Stecher, B. The mouse gut microbiome revisited: from complex diversity to model ecosystems. Int. J. Med. Microbiol. 306, 316–327 (2016).

14. McCoy, K. D., Geuking, M. B. & Ronchi, F. Gut microbiome standardization in control and experimental mice. Curr. Protoc. Immunol. 117, 1–13 (2017).

15. Ley, R. E. et al. Obesity alters gut microbial ecology. Proc. Natl Acad. Sci. USA 102, 11070–11075 (2005).

16. Xiao, L. et al. A catalog of the mouse gut metagenome. Nat. Biotechnol. 33, 1103–1108 (2015).

17. Lane-Petter, W. Provision of pathogen-free animals. Proc. R. Soc. Med. 55, 253–256 (1962).

18. Macpherson, A. J. & McCoy, K. D. Standardised animal models of host microbial mutualism. Mucosal. Immunol. 8, 476–486 (2015).

19. Lagkovardos, I. et al. The Mouse Intestinal Bacterial Collection (miBC) provides host-specific insight into cultured diversity and functional potential of the gut microbiota. Nat. Microbiol. 1, 16131 (2016).

20. Clavel, T., Gomes-Neto, J. C., Lagkovardos, I. & Ramer-Tait, A. E. Deciphering interactions between the gut microbiota and the immune system via microbial cultivation and minimal microbiomes. Immunol. Rev. 279, 8–22 (2017).

21. Wyomere Brand, M. et al. The Altered Schaedler flora: continued applications of a defined murine microbial community. ILAR J. 56, 169–178 (2015).

22. Brugiroux, S. et al. Genome-guided design of a defined mouse microbiota that confers colonization resistance against Salmonella enterica serovar Typhimurium. Nat. Microbiol. 2, 16215 (2016).

23. Schaedler, R. W., Dubos, R. & Costello, R. Association of germfree mice with bacteria isolated from normal mice. J. Exp. Med. 122, 77–83 (1965).

24. Norin, E. & Midvedt, T. Intestinal microflora functions in laboratory mice are claimed to harbor a “normal” intestinal microflora. Is the SPF concept running out of date? Anaerobe 16, 311–310 (2010).

25. Uchimura, Y. et al. Complete genome sequences of 15 bacterial species constituting the stable defined intestinal microbiota of the GM15 gnotobiotic mouse model. Microbiol. Resour. Announc. 9, e00686–20 (2020).
Author contributions
M.D., F.L. and A.T. conceived the project, analyzed, interpreted, and integrated all the data. M.D. and F.L. wrote the manuscript. M.D. prepared figures. L.B. and Y.Y. achieved in silico design. M.D. performed bacteria isolation and identification. gDNA extraction, developed the qPCR microfluidic assay, and analyzed macroscopic phenotyping data. A.S. carried out WGS, C.E. identified specific regions for primer design, and A.M. performed taxonomic annotation and functional bioinformatic analyses. P.B., H.D., M.D., T.N., D.S. and M.S. performed animal experiments. B.S. provided the Oligo-MM12 consortium. T.N. performed the Oligo-MM12-related gDNA extraction and consortium analysis with the help of B.S. C.C., A.D., S.P. and M.S. performed the immunophenotyping, analyzed, and interpreted data. A.B. and D.A.O. performed the metabolophenotyping, analyzed, and interpreted data. M.S. helped design the diet-induced stunting protocol. F.D.V. performed, analyzed, and contributed to the text and figures related to the stunting model experiments. A.-L.B. provided technical support during the stunting model experiments. All authors read and approved the final manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-26963-9.

Correspondence and requests for materials should be addressed to Marion Darnaud.