Specific eco-evolutionary contexts in the mouse gut reveal *Escherichia coli* metabolic versatility

João Barroso-Batista[1*&], Miguel F. Pedro[1*], Joana Sales-Dias[1,2], Catarina J. G. Pinto[1&], Helena Pereira[2], Jocelyne Demengeot[1], Isabel Gordo[1], Karina B. Xavier[1#]

*These authors contributed equally

[1]Instituto Gulbenkian de Ciência, Rua da Quinta Grande 6, 2780-156 Oeiras, Portugal

[2]Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Avenida da República, 2780-157, Oeiras, Portugal

[&]Present address: João Barroso-Batista, University of Chicago, Department of Medicine, Section of Genetic Medicine, Chicago, USA.; Catarina J. G. Pinto, CEDOC, Chronic Diseases Research Centre, NOVA Medical School | Faculdade de Ciências Médicas, Universidade NOVA de Lisboa, Lisboa, Portugal

Corresponding author and lead contact: Karina B. Xavier, kxavier@igc.gulbenkian.pt

Summary

Members of the gut microbiota are thought to experience strong competition for nutrients. However, how such competition shapes their evolutionary dynamics and depends on intra- and interspecies interactions is poorly known. Here we tested the hypothesis that *Escherichia coli* evolution in the mouse gut is more predictable across hosts in absence of interspecies competition than in the presence of other microbial species. Supporting this hypothesis, we observed a specific genetic adaptation in *lrp,* a gene encoding a global regulator of amino acid metabolism, predictably selected in germ-free mice two weeks after mono-colonization. Analysis of gut metabolites established that the *lpr* mutations increase *E. coli* ability to compete for amino acids and identified serine and threonine as the metabolites preferentially consumed by *E. coli* in the mono-colonized mouse gut. Preference for serine consumption was further demonstrated by testing a set of mutants *in vitro* and *in vivo* that showed loss of advantage of a *lrp* mutant impaired in serine metabolism. Remarkably, the presence of a single additional member of the microbiota (*Blautia coccoides*) was enough to alter the gut metabolic profile and consequently the evolutionary path of *E. coli.* In this environment, the *lrp* mutations did not conferred advantage to *E. coli* and genes involved in anaerobic respiration were selected instead, recapitulating the eco-evolutionary context from mice with a complex microbiota. Together, these results highlight the metabolic plasticity of *E. coli* and its extreme evolutionary versatility, tailored to the specific ecology it experiences in the gut.

Introduction

The mammalian gut is populated by hundreds of microbial species forming a complex interacting community: the gut microbiota. The assembly of this community is initiated during the neonate stage and it suffers markedly changes in composition during the first years of life [1]. The population dynamics of this community are thought to be determined by eco-evolutionary processes [2–4]. While our understanding of the gut microbiota is still far from complete, it is now evident that it has a crucial impact on health, affecting a plethora of host functions that include immunity, development and nutrition [5–10]. Given its importance, and facilitated by the advent of sequencing techniques, the gut microbiota as a field of research has grown significantly in recent years [11].

Research on the microbiota has been performed using a variety of methodologies, from targeted 16S rRNA gene metabarcoding that allows estimating the number and

abundance of microbial species inhabiting the gut [12], to functional metagenomics approaches that also reveal the gene composition of the microbiota [13]. Of note, analysis of the gut metabolome (metabolomics) have been instrumental to decipher the complex host-microbiota crosstalk and identify compounds associated with the maintenance of host homeostasis [14–16]. Globally, these studies have contributed to a better understanding of the composition, variation and stability of the gut microbiota [17,18], as well as clarifying the complex reciprocal interactions between microbes and their hosts that impact host health [19]. However, while these studies provided valuable insights into the ecological processes shaping the gut microbiota, the evolutionary forces that govern this community are yet to be completely understood. This subject is particularly relevant in the context of the host-microbiota co-evolution hypothesis, that postulates that the complex mutualistic interactions between microbes and their hosts were established over millions of years of reciprocal evolutionary changes [20,21].

In the past years, efforts have been made to understand how gut bacteria adapt to the pressures of the gut environment using a variety of approaches [22–27]. In particular, experimental evolution with tractable model microorganisms is a powerful tool to provide insights into the evolutionary mechanisms that influence members of the microbiota, in a controlled, yet closer to natural, setting [28]. We and others [24,29–32] have used this approach to uncover how commensal strains of *Escherichia coli*, as a proxy for single members of the microbiota, adapt to the mouse gut. In addition to its status as a model organism, *E. coli* is also a relevant member of the gut microbiota, commonly isolated from many mammals [33]. In humans, *E. coli* is one of the most abundant facultative aerobes and bacteria belonging to this genus are among the first detected in the gut, rapidly colonizing this environment shortly after birth [34]. As a pioneer species, it is thought that *E. coli* metabolism in the infant gut is crucial for the transition from an aerobic to anaerobic environment that favors colonization of strict anaerobes [35].

By using a commensal K-12 strain of *E. coli* and the classic streptomycin-treated mouse model, we have previously shown that in the presence of a complex microbiota *E. coli* adaptation occurs by multiple genetic mutations in loci controlling respiration and general metabolic processes, in a regime dominated by clonal interference [24] but where other mechanisms of selection can also coexist [36]. We have also evaluated the contribution of host immunity on *E. coli* evolution, a factor that exerts a subtle and indirect influence through modulation of the microbiota that interacts with *E. coli* [37]. Under all these conditions, genes involved in metabolism or in its regulation were among the major targets for adaptation to

the mouse gut, suggesting that composition of the microbiota modulates the adaptation of *E. coli* by manipulating the metabolic environment of the gut. These findings lead us to postulate that interspecies competition and its metabolic impact on the gut environment are major factors influencing *E. coli* evolution in the gut. Therefore, we hypothesized that in the absence of other members of the microbiota (in animals exposed to the same diet), *E. coli* adaptation would be more similar across host individuals, and thus more predictable. Furthermore, we postulated that in the absence of interspecies bacterial competition *E. coli* adaptation would reveal its intrinsic metabolic preferences.

To test this hypothesis, here we followed *E. coli* adaption when colonizing the gut of mice raised in germ-free conditions (GF) and compared this process with that previously observed in microbiota-harboring animals. Using a combination of experimental evolution, genomics and metabolomics, we determined the genetic basis of *E. coli* adaptation under mono-colonization of the mouse gut and characterized the metabolic milieu underlying its evolution. We report that in the absence of interspecies competition *E. coli* adaptation was remarkably homogeneous across hosts and characterized by genetic changes in pathways involved in amino acids catabolism. These results were consistent with the nutritional context found in the gut. We further evidenced that while adaptation in mono-colonized mouse gut differs from that in microbiota-harboring animals, the inclusion of a single species was sufficient to alter nutrient availability in the gut and markedly shape *E. coli* evolution.

Results

Adaptation of *E. coli* in the absence of interspecies competition is characterized by parallel mutations modulating amino acid metabolism

To investigate *E. coli* adaptation to the mouse gut in the absence of microbiota we mono-colonized 10 GF-raised mice with a clonal population of *E. coli* (thereafter designated as Gnoto$_{Ec}$) composed by two strains, isogenic except for the presence of a fluorescent maker (YFP or CFP) (Fig. 1A). We followed the frequencies of these markers (Fig. 1B) as well as *E. coli* loads (Fig. S1A) in fecal pellets over a period of 23 days. We observed changes in marker frequencies, which indicate that *E. coli* populations are evolving in the gut of Gnoto$_{Ec}$.

We have previously shown that during adaptation of *E. coli* to the gut of specific pathogen free (SPF) mice - that harbor a microbiota - and which were treated with streptomycin (hereby designated SPF$_{str+Ec}$), multiple genetic targets are selected over a three weeks period [24]. To determine the genetic targets under selection during *E. coli* adaptation to

Gnoto$_{Ec}$, we performed whole genome sequencing (WGS) of samples of *E. coli* populations collected 23 days after colonization. Similarly to what was observed in SPF$_{str+Ec}$ [24,37], mutations targeting the *gat* operon were prevalent (Fig. 1C). The *gat locus* is involved in the metabolism of galactitol and was shown previously to be a mutational hotspot in this *E. coli strain* [36]. Mutations inactivating the *gat* operon were found at high abundance in all populations (Fig. 1C), further supporting that this locus is generally the first mutational target for this *E. coli* strain background in the gut. In addition to the *gat* operon, ten new adapted loci were found, which were targeted by parallel mutations comprising hits in more than one independently-evolved population (Fig. 1C and Table S1), a signature of selection [38]. Transposition of insertion sequence (IS) elements was the main or the only mutational mechanism in seven of the parallel loci identified, while in the other three, SNPs and indels were also present (Table S2). An average of 5.7 parallel mutations were detected in each population collected at day 23 of the adaptation period.

The most striking result was the detection of mutations in the coding region (or intergenic region upstream) of the Leucine-responsive Regulatory Protein (*lrp*) gene, present in all 10 Gnoto$_{Ec}$ populations (Fig. 1C and Table S1). Importantly, the majority of mutations comprised insertions of IS elements in the coding region of *lrp*, suggesting selection for loss of function during adaptation in Gnoto$_{Ec}$. *lrp* codes for a global transcriptional regulatory protein that controls genes related to transport, biosynthesis and catabolism of amino acids [39]. While repeatedly detected in all populations collected from Gnoto$_{Ec}$, mutations at the *lrp* locus were never observed at very high frequency (above 60%) at day 23 (Fig. 1C). Due to strong clonal interference in the gut, it was possible that *lrp* mutants would eventually increase to a frequency close to fixation given enough time for expansion of these mutants in the populations. As this could take longer than the initial 23-day period of the evolution experiment, we prolonged the experiment up to day 40 in four of the 10 mice studied (Gn1.4 – Gn1.7). We then performed WGS of population samples at this timepoint (Table S2). Populations where the frequency of *lrp* at day 23 was low (<10%), exhibited an increase by day 40 (Gn1.4 from 9 to 44% and Gn1.5 from 8 to 39%), while those populations where the frequency of *lrp* at day 23 was rather high (>25%) only displayed a slight increase (Gn1.6 from 46 to 56% and Gn1.7 27 to 35%). In summary, mutations in the intergenic region upstream of *lrp,* or in the coding region of this gene were selected in all 10 populations analyzed, displaying full parallelism across the different hosts and reaching frequencies as high as 60% within host, but never expanding towards fixation.

Analysis of the remaining nine parallel mutations identified at day 23 provided additional evidence for amino acid metabolism being the most frequently selected function during adaptation of *E. coli* to the gut of Gnoto$_{Ec}$ (Fig, 1C and Table S1). We identified five mutated loci (*uspA/dtpB*, *tdcA/tdcR*, *lrhA/alaA*, *cadC/yjdQ* and *flrR*), which are also related with amino acid metabolism. Mutations in the upstream region of *dtpB*, encoding an inner membrane protein involved in uptaking peptide-bound amino acids (di- and tri-peptides) [40] were found in four mice, one reaching a frequency of 30%. *tdcA* and *tdcR* code for regulators of the *tdcABC* operon [41] involved in the transport and metabolism of threonine and serine during anaerobic growth [42]. Mutations in the intergenic region between these genes were detected in three populations, reaching high frequencies (30% and 40%) in two of them. The intergenic region of *alaA*, one of three major alanine transaminases [43] was mutated in four mice at a frequency up to 22%. The upstream region of *cadC*, a transcriptional activator of the *cadAB* operon, involved in metabolism of lysine [44], was found mutated in three mice, at a maximum frequency of 8%. Another locus frequently mutated was *frlR* which codes for the repressor of fructoselysine operon [45] responsible for catabolizing the conversion of fructoselysine into glucose 6-phosphate and lysine [46]. Mutations in this locus were detected in eight mice, albeit at low frequency (average 6%). Importantly, from all these mutations the only ones that were previously detected in SPF$_{str+Ec}$ mice were mutations in the *gat* operon [24].

Although most genetic changes were located in amino acid metabolism-related loci, parallel mutations for other cellular processes were also identified. For example, *rssB*, encoding an adaptor protein that controls RpoS proteolysis [47], the global regulator of stress response, was mutated in half of the animals. Other genetic targets, such as *glpR*, *yjjM* and *yebK*, suggested modulation of carbohydrates metabolism. *glpR* codes for a repressor of the *glp* regulon, responsible for metabolism of glycerol [48], *yjjM* encodes a putative regulator involved in catabolism of L-galactonate [49], and *yebK* is a transcription factor involved in the regulation of carbon metabolism during shifts in nutritional availability [50].

Furthermore, mutations in populations collected at day 40 (Gn1.4 to Gn1.7) confirmed previously identified genetic targets and revealed an additional parallel mutation in the *cytR* gene (Table S2). *cytR* codes for a global regulator, which acts together with the catabolite activator protein (CAP) when cAMP levels are high, regulating transport and utilization of nucleosides and deoxynucleosides [51]. Mutations in *cytR* were found in three of the four mice analyzed at day 40.

Globally, our WGS analysis evidenced a high degree of parallelism across hosts, as three different loci were found under selection in more than 75% of the Gnoto$_{Ec}$ populations (Fig.1C). This conclusion was further supported when considering the frequency of each mutation within host, which was in general consistent across different hosts. These results support our initial hypothesis that adaptation of *E. coli* under mono-colonization would show high parallelism across mice. The genetic basis of these mutations also revealed that amino acid metabolism is the principal function that is genetically altered when *E. coli* evolves in the gut in the absence of interspecies competition.

The experimental setting to study evolution of *E. coli* in microbiota-harboring animals typically involves administration of an antibiotic before colonization (e.g. streptomycin [24,37]. In such setup, streptomycin is required to disrupt microbiota-dependent colonization resistance, allowing colonization by exogenous *E. coli*. To determine if selection of *lrp* mutants in Gnoto$_{Ec}$ was related to the absence of streptomycin, we repeated the evolution experiment in mono-colonized mice in the presence of streptomycin (Gnoto$_{str+Ec}$, n=3) and sequenced the independently-evolved populations at day 23. Mutations in *lrp* were found again in all mice, at frequencies ranging from 16 to 44% (Fig. 1C, Table S1 and S2). We also found mutations in five other loci related to amino acid metabolism that were previously identified in Gnoto$_{Ec}$. This data indicates amino acid metabolism genes are targets of adaptation in mice mono-colonized with *E. coli,* independently of streptomycin (Fig. 1C and Table S1).

Lrp-negative phenotype is specifically selected in Gnoto$_{Ec}$ mice and is maintained polymorphic in the gut

Mutations targeting the *gat* operon were shown in SPF$_{str+Ec}$ to confer a *gat*-negative phenotype that rapidly emerged and consistently swept through the populations of *E. coli* colonizing the mouse gut [24]. Selection of the *gat*-negative phenotype was also observed during adaptation of *E. coli* in Gnoto$_{Ec}$, although with delayed dynamics of expansion compared to SPF$_{str+Ec}$, namely seven versus three days (Fig. 2A). We attributed this slower emergence of the *gat*-negative phenotype to weaker effects of natural selection on mutations inactivating the *gat* operon, as the competitive advantage of a *gat*Z mutant against the ancestral was previously found to be significantly smaller in Gnoto$_{Ec}$, regardless of the shorter generation time [37], larger population size (Fig. S1B) and similar mutation rate (Fig. S1C) of *E. coli* in this host compared to SPF$_{str+Ec}$ [37]. These results show that although the first mutational target of *E. coli* adaptation to the mouse gut is conserved across hosts

colonized with different microbial communities (SPF_{str+Ec} WT or immune compromised Rag2$^{-/-}$ [37]) and in $Gnoto_{Ec}$ depleted of the other members of the microbiota, the strength of its advantage varies. Inactivation of the *gat* operon is strongly beneficial in SPF_{str+Ec} WT, smaller and more variable in mice with an altered microbiota composition (SPF_{str+Ec} *Rag2*$^{-/-}$ [37]) and provides an even smaller advantage in $Gnoto_{Ec}$ mice which lack microbiota interspecific competition.

Having identified *lrp* as the main adaptive target specific to $Gnoto_{Ec}$, we established a phenotypic assay to screen for *lrp* mutants and determine the dynamics of emergence of these mutants in $Gnoto_{Ec}$ populations. As most of the *lrp* mutations identified by WGS were insertions of IS elements in the coding region we expected these mutants to result in loss of function of Lrp. We constructed reference strains carrying a deletion mutant of *lrp* (Δ*lrp*) in both ancestral or *gat*-negative backgrounds. Given that Lrp is required for activation of glycine catabolism [52], we established a phenotypic assay based on selective growth at 42ºC in solid minimal medium with glycine as the sole nitrogen source (glyMM, see Methods) to score for Lrp loss of function. Both *lrp* deletion mutants were unable to grow in this medium, while the ancestral and *gat*-negative reference strains displayed robust growth (Fig. S2A-C), showing that in the glyMM medium, loss of function mutants of *lrp* could be easily identified.

In accordance with the data from WGS analysis, clones with *lrp*-negative phenotype were detected in all populations (Fig. 2B). In addition to clones displaying normal growth, comparable to the ancestral strain, and clones unable to grow in glyMM, like the Δ*lrp* mutant, we also observed clones with intermediate phenotype of reduced growth (Fig. S2D). Genetic analysis by targeted PCR (*lrp* locus) revealed that clones with an intermediate growth phenotype in glyMM harbor IS insertions in the intergenic region upstream of *lrp,* while those with no growth in this medium, a phenotype similar to that of Δ*lrp* mutant (Fig. S2D), carry IS insertions in the coding region of *lrp* (Fig. S2E). These results confirmed that evolved *lrp* mutants with IS insertions in the coding regions were loss of function mutants.

The phenotypic analysis also confirmed that fixation of the *lrp*-negative phenotype occurred in none of the lines, including those that were extended to 40 days (Fig. 2B and S2F). The general pattern of dynamics in the populations comprises an initial increase in frequency with subsequent deceleration over time. In between days 23 and 40 the frequency of the *lrp* phenotype stabilizes at different levels in different mice ranging from 32 to 66% at day 40. The observed maintenance of a mutation at an approximate equilibrium frequency

suggests the occurrence of frequency-dependent selection or, alternatively, strong competition with other mutants (clonal interference).

We also measured the *gat* phenotype in all clones tested for the *lrp*-phenotypes for all the time points shown in Figure 2B. Among the more than 14400 clones tested, only two *gat*-positive *lrp*-mutant clones were detected, each at a single time point and each from a different line. These results confirm that *gat* and *lrp* are consecutive adaptive events that occurred in all the Gnoto$_{Ec}$ populations tested. The *lrp-negative* phenotype was the second adaptive step in Gnoto$_{Ec}$ and the only mutation showing full parallelism across individuals that was specifically selected in the Gnoto$_{Ec}$ environment.

lrp advantage depends on intraspecies ecological interactions

Given that the *lrp*-negative phenotype was the major specific trait selected during adaptation in Gnoto$_{Ec}$ mice we asked whether we could estimate its competitive advantage *in vivo*. We measured the competitive advantage (fitness) of a Δ*lrp ΔgatZ* mutant (MFP106) against a reference *gat*-negative strain (*ΔgatZ,* MFP54), by colonizing GF-raised mice with mixtures of these two strains (at a 1:9 ratio) and following changes in mutant frequency for 5 days. We limited the experiment to 5 days to avoid potential secondary mutations in either strain that could interfere with measuring the effect of *lrp*-negative phenotype. As expected, *E. coli* loads in Gnoto$_{Ec}$ mice were stable throughout the experiment (Fig. S3A). However, we observed only a slight positive trend for the expansion of the *lrp* mutant over the reference strain (Fig. S3B), indicating a weak, if any, selective advantage of *lrp* mutants in the first days of the competition experiment. This result, together with the evidence that *lrp* mutants accumulate only two weeks of after *E. coli* colonization in the evolution experiment (Fig. 2B), and our findings of a general prevalence of mutations targeting metabolic genes, led us to hypothesize that the benefit of the *lrp*-negative phenotype could be associated with alterations in the gut caused by *E. coli* colonization itself. To address this possibility, we designed an experimental setup that would closely resemble the environment in which the *lrp*-negative phenotype emerged during the evolution experiment, namely an *E. coli* pre-colonized gut. To create such an environment, we colonized GF-raised mice with a streptomycin-sensitive, ancestral and WT strain of *E. coli* for 10 days, treated these animals with streptomycin in drinking water, and only then preformed a competition experiment (Fig. 3A). The WT pre-colonizer streptomycin-sensitive strain reached and maintained high loads during the initial period of colonization in all mice (Fig. S3C) and was readily eliminated upon administration of the antibiotic. We reasoned that this period of pre-colonization with *E. coli*

would be sufficient to alter the gut environment, favoring selection of *lrp*-negative mutants. We then tested the competitive advantage of the *lrp* mutant by introducing a mixture of Δ*lrp ΔgatZ* mutant and reference *gat*-negative strains (at a 1:9 ratio) at day 0, four hours after initiating the streptomycin treatment (Fig. 3A). These mice were then followed for changes in frequencies between the *lrp*-negative and reference strain. As expected, the streptomycin-resistant competitor strains colonized all mice with high loads (Fig. S3C). Importantly, a robust increase was observed in the frequency of the *lrp* deletion strain in the first days of the competition (Fig. 3B), with this mutant reaching high frequencies by day four (average of 88%, n=5). These data show that *in vivo* loss of the *lrp* gene is indeed a beneficial trait in Gnoto$_{Ec}$ while also suggesting that *lrp* advantage is context-specific, as the advantage of this mutant could only be observed after an initial period of *E. coli* colonization.

Serine and threonine are preferred nutrients of *E. coli* in Gnoto$_{Ec}$

Given the clear adaptive signature in genes related with metabolism we next characterized the gut metabolic environment. We performed a metabolomics analysis of cecum extracts, using NMR to determine the metabolites available to *E. coli* under the conditions of our adaptation experiments (Table S3). We focused on the mouse cecum as it is one of the sections of the intestine where *E. coli* is found at higher densities [53]. We first compared the metabolites identified in cecal contents from: i) unmanipulated GF control mice (n=5), that represent the environment at the start point of the Gnoto$_{Ec}$ experiment, and ii) microbiota-harboring SPF$_{str}$ mice (n=10, day=28 of streptomycin), as a proxy for the general metabolite profile that *E. coli* was subjected to when colonizing SPF$_{str+Ec}$ (Barroso-Batista et al., 2014). To analyze the data, we clustered metabolites in four groups: amino acids, organic acids, sugars and others (Fig. S4). A striking difference between the metabolic profiles of GF and SPF$_{str}$ concerned the abundance of amino acids. While amino acid represented 80% of all metabolites detected in GF, these amounted to less than 50% in SPF$_{str}$ (Fig. 4A), resulting in an absolute concentration of amino acids two-fold higher in GF when compared to SPF$_{str}$ (Fig. 4B, Mann-Whitney *U* test, $P<0.001$). This data suggests that amino acids are the most readily available metabolites in the gut of GF mice (Fig. 4A, see also Table S3 for a list with all the individual metabolites). When comparing individual metabolites, 15 out of the 19 proteinogenic amino acids detected were found at significantly higher concentrations in GF mice than in SPF$_{str}$ (Fig. S4A). Specifically, serine, threonine and asparagine, showed the highest differences, displaying more than a six-fold increase in GF compared to SPF$_{str.}$ In contrast, glutamate, which is typically a product of amino acid

consumption [54] was the only amino acid found over two-fold a higher concentration in SPF_{str}. On the other hand, organic acids were substantially overrepresented in SPF_{str} compared to GF, with acetate, propionate, butyrate and succinate displaying the largest differences (over 100 times higher in SPF_{str}, Fig. S4B).

We next analyzed the metabolites present in the cecum of $Gnoto_{Ec}$ colonized with *E. coli* for approximately one month (n=6, day=31), representing the metabolic environment at the end point of the $Gnoto_{Ec}$ evolution experiment, and compared it with the metabolic composition in GF mice (Fig. 4C-F and Table S3). Of all the amino acids detected, 7 were found at lower levels in $Gnoto_{Ec}$ compared to GF (with a p-value ≤ 0.05, Fig. 4D), suggesting consumption of these amino acids by *E. coli* upon colonization. The most striking changes in metabolite concentrations were observed for serine and threonine, found close to one order of magnitude lower in $Gnoto_{Ec}$ (Fig. 4D). In addition to amino acids, our analysis also revealed other metabolites differently represented in GF compared to $Gnoto_{Ec}$, for example short chain fatty acids (SCFAs) and glucose (Fig. 4E and F). In particular, acetate, propionate and succinate increased upon *E. coli* colonization. This comparative metabolomics analysis between cecum extracts of $Gnoto_{Ec}$ and GF suggests that serine and threonine were preferentially used as carbon/nitrogen sources by *E. coli* in $Gnoto_{Ec}$.

We also performed a longitudinal metabolomics analysis looking for a time-dependent evidence of amino acid/nutrient use. We collected fecal samples from 3 $Gnoto_{Ec}$ mice every three or four days over the course of 23 days. These samples were then processed for metabolomics analysis (Fig S5 and Table S3). Consistent with the results obtained with cecum extracts, analysis of fecal samples revealed a decrease in serine and threonine levels upon *E. coli* colonization (Fig. 4G and S5A), as well as an increase in acetate (Fig. S5B). The decrease in serine and threonine was observed from day three of colonization, and the levels remained at low but detectable after this timepoint and throughout the length of the experiment. These results show that in the GF gut *E. coli* preferentially consumes these two amino acids.

Overall, the results of the comparative metabolomics analysis highlight the higher availability of amino acids in the GF compared to SPF_{str} gut and indicate that serine and threonine are preferred metabolites for consumption by *E. coli* in $Gnoto_{Ec}$, providing additional support for the crucial role of amino acid metabolism for *E. coli* adaptation in $Gnoto_{Ec}$.

Serine metabolism drives *E. coli* metabolic adaptation in $Gnoto_{Ec}$

Lrp has been shown to repress transcription of genes involved in both serine and threonine catabolism [55] and it has also been demonstrated that *lrp* mutants can have an advantage compared to the WT, in media containing amino acids as single carbon sources, in particular serine [56]. Therefore we compared the growth of two *lrp* mutants isolated from the evolution experiment; one containing an IS insertion in the coding region of *lrp* (lrp_{cod}) and the other harboring an IS insertion in the intergenic region of *lrp* (lrp_{interg}), with that of a Δ*lrp* deletion mutant (in a *gat*-negative background) and the corresponding *gat*-negative reference strain (*lrp*+) in minimal medium containing different sources of amino acids. This medium contained either a single amino acid (serine or threonine) or a complex mixture of amino acids (casamino acids (CAA)) as carbon and energy source. In addition, we also tested the effect of L-leucine, a modulator of Lrp activity [55] on the growth of the different *E. coli* strains.

In the presence of serine as carbon source, all three *lrp* mutants displayed a similar growth profile, with a growth rate over two-fold higher than the reference strain (Fig. 5A left panel and Table S4). This difference in growth was attenuated in the presence of leucine (Fig. 5A right panel and Table S4). Leucine presumably counteracts the Lrp-mediated repression of serine catabolism in strains with intact *lrp* [55], promoting serine metabolization. In threonine minimal medium none of the strains tested, including *lrp* mutants, were able to grow and thus we were unable to assess a possible advantage of *lrp* inactivation in this environment. Finally, in CAA, all strains displayed similar growth (Fig. S6A and Table S4), suggesting that the advantage of the *lrp* mutants when growing in serine is specific and not due to a general growth advantage.

Overall, these results show that *lrp* loss of function mutants have a fitness advantage when metabolizing serine *in vitro*, possibly due to the constitutive de-repression of serine catabolism. Considering these results and the metabolomics analysis showing that serine was rapidly depleted upon *E. coli* colonization in $Gnoto_{Ec}$ conditions, serine catabolism could be a major factor explaining the emergence and advantage of *lrp* mutants in $Gnoto_{Ec}$. In *E. coli*, the main enzyme for serine utilization is a serine deaminase encoded by the *sdaA* gene. Importantly, SdaA is also the only serine deaminase subjected to Lrp-mediated transcriptional regulation [57]. We constructed a *sdaA* deletion mutant (in the Δ*lrp* Δ*gatZ* background) that is unable to metabolize and grow in serine. We evaluated the competitive advantage of this mutant against a reference (*gat*-negative) strain by *in vivo* competition in pre-colonized $Gnoto_{Ec}$ mice, the same environment where we observed strong advantage for the *lrp* mutant (see Fig. S6A for *E. coli* loads and Fig 3A for the experimental setup). In

these conditions, the *sdaA lrp* mutant had no advantage against the reference strain as measured after two days of competition (Fig. 5B, day 0 to day 2, s=-0.006±0.005), while the *lrp* mutant had an immediate strong advantage against the same reference strain (Fig. 3B, day 0 to 2, s=0.039±0.007). Thus, in the Δ*sdaA* background, the early advantage of the *lrp* mutation is abolished. From day 2 onward, the *sdaA lrp* mutant increased in frequency (Fig. 5B – day 2 to day 5), displaying a competitive advantage against the reference strain (s=0.061±0.005 per hour) comparable to that estimated for the *lrp* mutant (s=0.078±0.008 per hour). In summary, these results reveal that *E. coli* fitness is impacted by removing its ability to metabolize one of its two preferred amino acids and demonstrate that serine metabolism is an important feature contributing to the advantage and selection of *lrp* mutants in Gnoto$_{Ec}$.

Co-colonization of the mouse gut with a single additional microbiota member alters the metabolic environment and shapes *E. coli* evolution

Our results show that *lrp* mutants of *E. coli*, that were undetected in adaptation experiments conducted in microbiota-harboring animals, are specifically selected in the singular metabolic environment of Gnoto$_{Ec}$ gut. GF-raised animals differ from their microbiota-harboring counterparts in ways less obvious than the presence of a dense and complex community of colonizing microorganisms such as altered intestinal functions, morphological defects at the gut level and impaired immunity [5]. While these host phenotypes are consequence of the absence of a microbiota, they may be confounding when interpreting bacterial species ecology. To disentangle a direct effect of the microbiota on *E. coli* evolution from an indirect effect from the host physiology, we tested whether a single microbial species could alter the evolutionary path of *E. coli* in the mouse gut. We took advantage of having a colony of mice mono-associated with *Blautia coccoides* in our animal facility, to repeat the evolution and competition experiments, as well as the metabolomics analysis, in a still extremely simplified microbial environment. *B. coccoides* is a Firmicutes belonging to the Clostridia class, and a common resident of the mouse and human gut [58,59]. It was recently included in a synthetic community aiming to assemble a minimal functional microbiota in mice [60]. This group of bacteria is capable of energy extraction through specialized mechanisms [61] often favoring amino acids as opposed to more simple nitrogen forms [62,63], and is thus a plausible and relevant competitor for amino acids of *E. coli* in the mouse gut.

We colonized five mice mono-associated with *B. coccoides* with our ancestral *E. coli* strains (called here Gnoto$_{Bc+Ec}$) and followed the dynamics of colonization over the course of one month. *E. coli* was able to readily colonize this simple environment reaching high loads (Fig. S7A) and exhibited evidence of marker divergence, suggesting the occurrence of adaptive events (Fig. 6A). *B. coccoides* loads were approximately the same in all animals at the end of the experiment, day 28 (Fig. S7B).

We then identified targets of selection in *E. coli* population evolved in Gnoto$_{Bc+Ec}$ by isolating and sequencing *E. coli* population samples collected from each mouse at day 28 (Fig. 6B and Table S5 and S2). As before, the *gat* operon was targeted for selection in all five mice, with mutations found at high abundance in all populations. This data is consistent with the *gat* operon as a ubiquitous target of mutation in the mouse gut, for this strain background. Importantly, we observed that the dynamics of emergence of *gat*-negative phenotype in Gnoto$_{Bc+Ec}$ were distinctly more similar to those observed during the evolution in SPF$_{str+Ec}$ than in Gnoto$_{Ec}$ (Fig. 6C). Our WGS analysis also revealed other loci under selection (Fig. 6B and Table S5). Strikingly, while two of these mutational targets were shared with those found in Gnoto$_{Ec}$ (*cadC* and *cytR*), mutations in the *lrp* gene were not detected in any of the five populations analyzed by WGS (Table S2). Extensive phenotypic screening for *lrp* mutants throughout the evolution experiment confirmed these results (out of the 1824 clones tested from four time points from each mouse, spanning the 28 days of the experiment, no *lrp*-negative mutants were detected). On the other hand, in addition to the *gat* operon, three out of the six other parallel mutations identified targeted the same genes frequently found mutated in SPF$_{str+Ec}$ [37], and not in Gnoto$_{Ec}$ (Fig. 6B and Table S5). These mutations were found in the intergenic region upstream of *dcuB fumB* operon (anaerobic metabolism and fumarate/succinate transport) in all Gnoto$_{Bc+Ec}$ populations, and in the open reading frame of *srlR* and *kdgR* (repressors of sugar-alcohol and gluconate utilization, respectively) in two and three populations respectively.

We then measured the fitness advantage of a Δ*lrp* mutant in Gnoto$_{Bc+Ec}$ by testing its competition against the reference strain. The competitions were performed following the protocol we used above to measure the advantage of *lrp in vivo*, i.e., upon pre-colonization with *E. coli* and antibiotic treatment to remove the pre-colonizer strain (Fig. 3A and S7C). While in three animals the *lrp* mutant was clearly advantageous, rapidly outcompeting the reference strain, in the remaining two it was neutral or only slightly advantageous (Fig. 6D). Interestingly, this bimodal result correlated well with *B. coccoides* loads (Fig. S7D): in mice where the loads of these bacteria were lower the *lrp* mutant was advantageous, whereas in

mice colonized with higher levels of *B. coccoides* the advantage of the *lrp* mutant was small or negligible (Fig. 6E, R^2=0.92). Therefore, the fitness advantage of the Δ*lrp* mutant showed an inverse correlation with *B. coccoides* loads.

Lastly, we performed a metabolomics analysis on cecum extracts from the five Gnoto$_{Bc+Ec}$ mice (collected at the last day of the experiment, day 28, Fig. 7A) and compared the results to those previously obtained for GF, Gnoto$_{Ec}$ and SPF$_{str}$ (Table S3). This analysis revealed that the metabolomic environment in Gnoto$_{Bc+Ec}$ was generally more similar to that in SPF$_{str}$ than to GF or Gnoto$_{Ec}$ (Fig. 7B and Fig. S7E). In particular, amino acids were found at a significantly lower concentration in Gnoto$_{Bc+Ec}$ (18.7µmol/g of cecal contents, amounting to 35% of all metabolites, Fig. 7A) than in Gnoto$_{Ec}$ (30.3 µmol/g of cecal contents, Mann Whitney U test, P=0.004). At the individual level, 13 out of the 19 detected amino acids in Gnoto$_{Bc+Ec}$ were significantly lower than in Gnoto$_{Ec}$ (Fig. S7F). Conversely, for organic acids, those most abundant in Gnoto$_{Bc+Ec}$ (acetate, butyrate, succinate, ketoglutarate and lactate), were found in lower levels in Gnoto$_{Ec}$ (Fig. S7G). Sugar composition also differed between Gnoto$_{Bc+Ec}$ and Gnoto$_{Ec}$, as raffinose and sucrose were found at greatly decreased concentrations in Gnoto$_{Bc+Ec}$ (Fig. S7H). Because *E. coli* is unable to metabolize raffinose and sucrose [64], these results suggest that these two sugars, the most abundant in GF and Gnoto$_{Ec}$, are likely metabolized by *B. coccoides* during intestinal colonization.

In sum, the metabolomics results revealed that the presence of a single bacterial *Blautia* species is sufficient to drastically alter the environment that *E. coli* is exposed in the mouse gut, resulting in a metabolic milieu resembling that of microbiota-inhabited gut. The metabolomic data also helps explain our observation of shared targets of selection in *E. coli* between Gnoto$_{Bc+Ec}$ and microbiota-harboring animals as well as the absence of emergence and selection of *lrp* mutants. Overall, these results show that the metabolomic environment in the mouse gut is strongly influenced by the microbiota and the major driver of *E. coli* adaptation to the mouse gut.

Discussion

Here we explored how *E. coli*, one of the pioneer species of the human gut [34], adapts to the gut of GF-raised mice, a process that resembles the initial colonization of the gut. Our previous works in microbiota-harboring animals have shown that adaptation to this environment was characterized by mutations in genes involved in respiration or metabolism of different compounds [24,37]. Here we found that in mice devoid of microbiota, the

mutational profile was shifted towards modulation of amino acids metabolism. We further demonstrated that selection for these different biological pathways was highly influenced by the presence/absence of additional species of bacteria in the gut which had a very strong influence in the metabolic environment of the gut.

Notably, we identified *lrp* as a major target of selection, appearing mutated in all ten independently-evolved Gnoto$_{Ec}$ populations and thus the only host-specific target displaying full parallelism. Lrp is a well-established global regulator of several cellular processes, including amino acid metabolism [65]. Mutants in this locus have been shown to emerge in *E. coli* cultures maintained under carbon stress, as these mutants display a growth advantage in stationary phase (GASP phenotype), presumably due to enhanced ability to consume amino acids available following the death and lysis of other bacterial cells [66]. As a global regulator of gene expression, including virulence-related genes, *lrp* has also been shown to be important *in vivo*, in particular during bacterial infection. However, while several reports have highlighted the importance of *lrp* for successful host infection in *Salmonella* [67], *Citrobacter rodentium* [68] and pathogenic strains of *E. coli* [69], its role during commensal colonization of the gut is less clear. Here, using *in vivo* competitive assays we showed that *lrp* deletion mutants had a strong advantage while colonizing the Gnoto$_{Ec}$ mouse gut. Interestingly, the *lrp* mutant was neutral when competed in naïve GF mice but displayed an advantage after pre-colonization of this environment with *E. coli*. These results indicate that pre-colonization changed the metabolic milieu of the gut and that *lrp* confers a competitive advantage *in vivo* only in a metabolic environment shaped by intra-species microbial competition.

Quantification of metabolites in the mouse cecum revealed that amino acids are abundant in GF and showed to decrease rapidly upon *E. coli* colonization, in particular serine and threonine. In this scenario, it was plausible that an increased ability to consume serine and threonine by *lrp* mutants would provide an advantage in this environment, where these amino acids were preferentially consumed by *E. coli*. Consistent with this prediction, *lrp* mutants had an increased growth rate compared to a *lrp*-positive strain in serine minimal medium. This growth difference was alleviated in the presence of leucine, as this compound signals amino acid availability, counteracting the Lrp-mediated repression of serine catabolism in strains with an intact *lrp* [55]. *lrp* mutants are constitutively de-repressed for serine catabolism [57] and thus can readily consume serine when available, even when at limiting concentrations, as it is the case in the Gnoto$_{Ec}$ mouse gut. Accordingly, deletion of *sdaA*, in a *lrp*-negative background, that results in inability to metabolize serine, abolished

the advantage of *lrp* mutation in the first two days of competition against a *lrp*-positive reference strain. However, the Δ*sdaA* Δ*lrp* mutant displayed a strong advantage after two days of colonization, suggesting that in addition to serine metabolism, other mechanisms may contribute to the advantage of the *lrp* mutant. Alternatively, this observation may also result from activation of alternative pathways for serine metabolism and/or a nutritional shift for consumption of other preferred nutrients (e.g. threonine). While *lrp* mutants are also expected to confer a benefit when metabolizing threonine, we were unable to explore this scenario *in vitro,* as *E. coli* cannot grow on this amino acid as a sole carbon source [56]. The loss of *lrp* was not able to restore growth of any of the strains, when threonine is used as only carbon and energy source, independently of leucine, isoleucine, valine and ammonia (ammonium chloride) supplementation.

The majority of studies on *E. coli* adaptation to the mouse gut using different genetic backgrounds of *E. coli*, report the appearance and selection of mutants with increased nutritional capability. For example, mutations in genes involved in stress resistance and nutritional competence have been identified in *E. coli* clones evolved in GF-raised [30,31] or streptomycin-treated microbiota-harboring mice [32,70], as well as mutations specifically improving the growth in galactonate [29], also in streptomycin-treated mice. More recently, acquisition of genes enhancing degradation of complex polysaccharides was selected in a probiotic *E. coli* Nissle strain evolved in GF [25]. In agreement with this trend, most of the loci under selection in *E. coli* evolved in the mouse gut identified here and in our previous works [24,37] relate to metabolic functions. However, here we show that the type of pathway under selection differs, reflecting the metabolic environment in the gut. For instance, in Gnoto$_{Ec}$, where amino acids are preferred nutrients, six out of eleven loci under selection (*lrp, frlR, lrhA/alaA, uspA/dtpB, tdcA/tdcR* and *cadC/yjdQ*) are directly related to amino acid metabolism. On the other hand, in SPF$_{str+Ec}$, where other nutrients are available and amino acids are utilized by multiple members of the microbiota, three out of the seven genetic targets (*dcuB*, *focA* and *yeaR*) are involved in anaerobic respiration [24,37]. These differences on the metabolic pathways under selection in these different hosts indicate that *E. coli* adaptation is strongly shaped by the metabolic environment modulated by the presence/absence of other bacteria in the gut.

Further highlighting the importance of the metabolic environment for *E. coli* adaptation, we found that the presence of a single additional member of the microbiota was enough to alter the evolutionary path of *E. coli* in the mouse gut. First, metabolic characterization in Gnoto$_{Bc+Ec}$ revealed changes in metabolite concentrations, resulting in a nutritional

composition distinct from Gnoto$_{Ec}$, and more closely resembling SPF$_{str}$. For example, many amino acids were lower in Gnoto$_{Bc+Ec}$ mice when compared to Gnoto$_{Ec}$ suggesting that amino acids are common nutrients to both *B. coccoides* and *E. coli*. On the other hand, decreased levels of raffinose and sucrose indicate consumption of these oligosaccharides by *B. coccoides* [71], as *E. coli* is unable to do so [64]. Interestingly, byproducts of oligosaccharides degradation by *B. coccoides* include monomeric sugars that *E. coli* has the potential to metabolize, which is consistent with increased glucose levels in Gnoto$_{Bc+Ec}$ in comparison with Gnoto$_{Ec}$. This observation may point to a cross feeding interaction between co-colonizers of the mouse gut, even in this low-diversity environment. Second, the genetic targets selected during *E. coli* evolution in Gnoto$_{Bc+Ec}$ mice included loci commonly found in populations evolved in SPF$_{str+Ec}$ (*dcuB*, *srlR* and *kdgR*) [24,37] but absent in Gnoto$_{Ec}$ mice. Conversely, the main mutational target in Gnoto$_{Ec}$ mice (*lrp*) was conspicuously absent in Gnoto$_{Bc+Ec}$ and the selective advantage of a Δ*lrp* mutant was dampened when in the presence of *B. coccoides* in a density-dependent manner, as advantage of *lrp* mutants in competition was negatively correlated with *B. coccoides* loads. Taken together these data show that changes in the metabolic landscape caused by the presence of *B. coccoides* during *E. coli* colonization of the mouse gut were enough to alter the selection of beneficial mutations in *E. coli*. Globally, these results further suggest that *E. coli* growth and evolution in the gut may depend more on the presence of specific members of the microbiota (even in the extreme case of a single species) than on the composition and diversity of the microbial community. In line with this hypothesis, *E. coli* evolution in immune-compromised *Rag2*$^{-/-}$ SPF$_{str+Ec}$ mice with an altered microbiota composition was characterized by mostly shared, and only a few host-specific, genetic targets, compared to WT SPF$_{str+Ec}$ [37]. Overall, these data support the conclusion that the metabolic environment is a major factor driving *E. coli* adaptation to the mouse gut.

Of note, we found no evidence for direct competition between *E. coli* and *B. coccoides*, as the loads of *E. coli* were slightly higher in the presence of *B. coccoides*. However, the impact *B. coccoides* had in the availability of *E. coli* preferred nutrients, e.g., by increasing the levels of glucose (*E. coli* universally preferred carbon source) and decreasing the amounts of serine (one of the preferred amino acids), was strong enough to change its evolutionary trajectory, as distinct genetic targets were selected during *E. coli* adaptation in Gnoto$_{Bc+Ec}$ compared to Gnoto$_{Ec}$. Therefore, when facing with changes in the metabolic environment of the gut due to the presence of *B. coccoides, E. coli* was able to rapidly adapt

by exploring alternative nutritional niches that maintained its population size, demonstrating remarkable genetic and metabolic versatility.

As a general pattern, our data suggests that interspecies competition increases the selective pressures to which *E. coli* is subjected in the mouse gut, a tendency previously observed in mice bearing diverse microbiota compositions [37]. Illustrating this idea, mutations inactivating the *gat* operon, the first and a prevalent genetic target in this *E. coli* strain when colonizing the mouse gut, emerge and expand at a slower rate in Gnoto$_{Ec}$ mouse compared to SPF$_{str+Ec}$ or even Gnoto$_{Bc+Ec}$. While clonal interference is evident in both hosts, it is likely that the stronger selective pressures in SPF$_{str+Ec}$ lead to a faster loss of genetic diversity, as evidenced by the eventual lineage replacements in the population, compared to Gnoto$_{Ec}$, where weaker beneficial mutations in different genetic backgrounds are expected to compete for longer periods of time [72]. In addition, adaptation of *E. coli* to the Gnoto$_{Ec}$ gut was remarkably consistent across hosts, displaying a high degree of parallelism. Three of the parallel genetic targets were found in more than 75% of Gnoto$_{Ec}$ populations (*gat, lrp and frlR*) and the frequency of each parallel mutation within hosts was in general similar. This observation contrasts with SPF$_{str+Ec}$, in which the *gat* locus was the only parallel target identified in the all the populations, and the frequencies of parallel mutations within populations displayed a large variability. Together with our findings of preferential selection of functionally related genetic targets (amino acid metabolism) in Gnoto$_{Ec}$, these results support our initial prediction that adaptation in the absence of inter-species competition would display a higher degree of parallelism across hosts. In the context of initial colonization of the infant gut, these results suggest a higher convergence on the selective pressures acting on the first stages of gut colonization, that are then diversified as the assembly of the microbiota progresses. This would result in selection of the same functional traits at an initial stage but increased divergence later in life, depending on individual composition and diversity of the gut microbiota.

Our WGS analysis also showed that while some of the mutations detected in Gnoto$_{Ec}$ displayed a substantial increase in frequency when we prolonged the evolution experiment for another 20 days, many did not. Of particular note, both WGS analysis and phenotypic screening of *lrp* mutants showed that these never reached fixation and instead stabilized at an intermediate frequency. These data may suggest the occurrence of negative frequency-dependent mechanisms of selection in addition to clonal interference, two processes that can coexist even in simple environments [73]. Competition for limiting resources such as serine and threonine, coupled with temporal fluctuations in these nutrients, deriving from

host cyclical feeding patterns [74] could help maintain the *lrp* polymorphism observed in our experiments [75]. Alternatively, the relatively lower frequency achieved by *lrp* mutations in Gnoto$_{Ec}$ mice could also be explained by an increased number of mutational targets for the trait under selection. For example, genetically distinct mutants could display similar phenotypes, such as increased ability for amino acid consumption. Combined with the low selective effect of each mutation, it is likely that different mutations accumulate in different backgrounds, competing with each other in the same population.

This work further highlights *in vivo* experimental evolution as an attractive tool to query the nutritional preferences of *E. coli* in the gut environment. As an alternative to the use of genetic mutants in specific nutritional-related genes [76], experimental evolution can be used to dissect the microbial metabolism in the mouse gut of related *E. coli* strains with different genetic backgrounds. More generally, it will be interesting to assess if other members of the microbiota will display the same genetic versatility that allows *E. coli* rapid adaptation to the nutritional environment of the gut. It is likely that adaptation of generalist microbes such as *E. coli*, that can switch between fermentation and anaerobic respiration and grow in a variety of nutrients [77], will more closely reflect the nutritional composition in the gut due to their metabolic plasticity. Conversely, adaptation of more specialized bacteria that are, for example, extremely efficient in fermenting a restrict number of complex polymers (such as some mucin-degrading species [78]) is less likely to change according to nutritional shifts in the gut environment. However, even in this case, specialists might nonetheless adapt to improve metabolism of certain compounds (with tradeoff for others) if a reduced number of choices are available, thus revealing their metabolic preferences.

In summary, we found that *E. coli* adapts to the Gnoto$_{Ec}$ mouse gut mainly through modulation of amino acids metabolism, reflecting the metabolic landscape of this environment. Notably, while the process is quite distinct from that observed in microbiota-harboring animals, the addition of a single bacterial species was enough to alter *E. coli* nutrition and evolution. These results highlight the relevance of the gut microbiota as a complex interacting community and modulator of host environment.

Acknowledgements

We thank Roberto Balbontín and Tanja Đapa for critical reading of the manuscript.

We further thank the members of the Bacterial Signalling Lab for discussions throughout this work. We thank Vitor Cabral, Rita Oliveira, and Joana Amaro for technical support and valuable suggestions.

This work was supported by Portuguese National funds from Fundação para a Ciência e Tecnologia (FCT) PTDC/BIA-EVF/118075/2010 and PTDC/BIA-MIC/4188/14 in addition to support by ONEIDA (LISBOA-01-0145-FEDER-016417) and CONGENTO (LISBOA-01-0145-FEDER-022170) co-funded by FEEI - "Fundos Europeus Estruturais e de Investimento" from "Programa Operacional Regional Lisboa 2020". K.B.X and I.G. are supported by the FCT Investigator Program. The NMR data was acquired at CERMAX, ITQB-NOVA, Oeiras, Portugal with equipment funded by FCT, project AAC 01/SAICT/2016. Genome sequencing data have been deposited in the NCBI Read Archive database with accession code PRJNA429051.

Author Contributions

Conceived and designed the experiments: JBB, MFP, KBX, JD, IG. Performed the experiments: JBB, MFP, JSD, CP. Analyzed the data: JBB, MFP, HP, KBX. Contributed reagents/materials/analysis tools: KBX, HP. Wrote the original draft of the paper: JBB, MFP, KBX. Contributed to write the paper: IG and JD.

Declaration of interests

The authors have no conflict of interest to report.

Figure titles and legends

Fig. 1 - *E. coli* evolution in the Gnoto$_{Ec}$ mouse gut is characterized by mutations in genes involved in amino acid metabolism. (**A**) Design of evolution experiment. Germ free mice were monocolonized by gavage with *E. coli* (1:1 mix YFP + CPF) at day 0 and followed for approximately one month. Samples for WGS were collected from all populations at day 23 and from populations Gn$_{Ec}$1.4 to 1.7 at day 40. The experiment was repeated 3 times, with a total of n=10 mice. (**B**) Dynamics of neutral fluorescent marker (± 2 S.E.M.) during the first 23 days of adaptation of *E. coli* upon colonization of GF mice (n=10) provide evidence for adaptation. (**C**) Parallel targets of mutation identified by WGS in samples of populations

collected at day 23. The height of the bars denotes the frequency of populations (n=10 for Gnoto$_{Ec}$) where a given mutation was present (left axis) and circles represent the frequency each mutated locus reached within the population (right axis). Loci related with amino acid metabolism are highlighted in bold. All parallel loci except for *yjjM* and *yebK* were also found in Gnoto$_{str+Ec}$, (n=3). See also Fig. S1 and Tables S1 and S2.

Fig. 2 - Emergence of g*at*- and *lrp*-mutant phenotypes reveals full parallelism. (**A**) Frequency of the *gat*-negative phenotype (± 2 S.E.M.) over time across individual populations in Gnoto$_{Ec}$ (colored lines, n=10, same legend as in (B)). In dashed line, the average (± 2 S.E.M.) of *gat*-negative phenotype across *E. coli* populations (n=15) during adaptation to SPF$_{str+Ec}$ mice is reproduced for comparison [24]. (**B**) Frequency (± 2 S.E.M.) of the *lrp*-mutant phenotype (sum of complete and partial loss-function) over time, across individual populations in Gnoto$_{Ec}$ (colored lines, n=10). Inset table denotes the frequency of *lrp*-mutant phenotype at the end of each lineage and the corresponding day. See also Fig. S1 and S2

Fig. 3 - *lrp* mutant has a competitive fitness advantage in pre-colonized Gnoto$_{Ec}$ mice. (**A**) Design of competition experiment: germ-free mice were pre-colonized with ancestral WT streptomycin-sensitive strain of *E. coli* (MFP121) for 10 days, after which mice were administered streptomycin in drinking water (green shade). After 4 hours of streptomycin treatment the competitor strains composed of a mixture of Δ*lrp* Δ*gatZ* and Δ*gatZ* streptomycin-resistant strains were introduced at 1:9 ratio by oral gavage. (**B**) Frequency (± 2 S.E.M.) of Δ*lrp* Δ*gatZ* mutant during competition experiment (n=5) in (A). See also Fig. S3.

Fig. 4 - Amino acids are abundant resources that are targeted by *E. coli* during adaptation in Gnoto$_{Ec}$ mice. Concentration of small metabolites from cecal contents (**A**-**F**) or fecal samples (**G**) was determined by ^{1}H-NMR and grouped by chemical attributes. (**A**) Pie charts with relative concentrations (± 2 S.E.M.) of metabolite groups in Germ-free (GF, n=5) and streptomycin-treated SPF mice (SPF$_{str}$, day 28, n=10). (**B**) Total concentration of amino acids is higher in GF than SPF$_{str}$ (Mann-Whitney *U*-test, P<0.001). (**C**) Pie chart with relative concentrations (± 2 S.E.M.) of metabolite groups in Gnoto$_{Ec}$ mice (day 31, n=6). (**D**-**F**) Absolute concentrations of all metabolites (with medians depicted as lines) detected in GF (blue circles) and Gnoto$_{Ec}$ (black triangles). Amino acids (**D**) are

ordered, from left to right, according to concentration difference between GF and Gnoto$_{Ec}$. Metabolites whose concentration significantly differs between GF and Gnoto$_{Ec}$ display the level of significance: $P \leq 0.05$ (*), $P \leq 0.01$ (**) or $P > 0.05$, non-significant (ns), after *post-hoc* Mann-Whitney *U*-test with Holm's correction for multiple comparisons. Lines depict medians. (**G**) Time course of concentrations of all amino acids detected during Gnoto$_{Ec}$ colonization (n=3) highlighting concentration (± 2 S.E.M.) dynamics of serine and threonine. See also Fig. S4 and S5 and Table S3.

Fig. 5 - Serine metabolism provides fitness advantage to *lrp* mutants. (**A**) Growth curves of Δ*lrp* Δ*gatZ*, *lrp*$_{cod}$ Δ*gatZ*, *lrp*$_{interg}$ Δ*gatZ* and Reference (Δ*gatZ*) in M9 medium with serine 0.5% as carbon source, in the absence (left panel) or presence (right panel) of leucine 10mM (n=9 from 3 independent experiments). Error bars represent ± 2 S.E.M. Growth parameters can be found in Table S4. (**B**) Effect of serine utilization on fitness advantage of *lrp* mutant *in vivo*. Frequency (± 2 S.E.M.) of Δ*lrp* Δ*sdaA* Δ*gatZ* mutant against Δ*gatZ* during competition in GF mice pre-colonized with ancestral strain (as in Fig. 3A) shown in red (n=5). For comparison, the average frequency (± 2 S.E.M.) of Δ*lrp* Δ*gatZ* mutant against Δ*gatZ* during *in vivo* competition (from Fig. 3B) is depicted in a dashed line. See also Fig. S6 and Table S4.

Fig. 6 - *B. coccoides* abolishes *lrp* mutant emergence and fitness. (**A**) Dynamics of neutral fluorescent marker (± 2 S.E.M.) during *E. coli* adaptation in Gnoto$_{Bc+Ec}$ mice (n=5). (**B**) Parallel targets of mutation identified by WGS in samples of populations collected at day 28. The height of the bars denotes the frequency of populations (n=5) where a given mutation was present (left axis) and circles represent the frequency each mutated locus reached within the population (right axis). Mutations common to SPF$_{str+Ec}$ are underlined and those common to Gnoto$_{Ec}$ highlighted in bold. (**C**) Frequency (± 2 S.E.M.) of the *gat*-negative phenotype over time across individual populations (blue lines) in Gnoto$_{Bc+Ec}$ (n=5). Average (± 2 S.E.M.) *gat* dynamics in Gnoto$_{Ec}$ (red dashed line) and SPF$_{str+Ec}$ (black dashed line) are also depicted for comparison (data from Fig. 2). (**D**) Frequency (± 2 S.E.M.) of ∆*lrp* ∆*gatZ* strain during competition against ∆*gatZ* in GF mice after a period of pre-colonization (as in Fig. 3) but in the presence of *B. coccoides* (n= 5). (**E**) Correlation between ∆*lrp* ∆*gatZ* mutant fitness and average bacterial loads of *B. coccoides* during *in vivo* competition in (D). See also Fig. S7 and Tables S2 and S5.

Fig. 7 - Metabolites in Gnoto$_{Bc+Ec}$ mice are more similar to the metabolites in SPF$_{str}$ than in Gnoto$_{Ec}$. (**A**) Pie charts of relative concentrations (± 2 S.E.M) of metabolite groups from cecal contents at the end of the Gnoto$_{Bc+Ec}$ adaptation experiment (day 28, n=5). (**B**) Heatmap of median concentrations of all metabolites in Gnoto$_{Bc+Ec}$ cecal contents, for comparison with GF, Gnoto$_{Ec}$ and SPF$_{str}$. Colors represent the standard scores of absolute concentrations following group clustering with Euclidean distances (Ward's method). See also Fig. S4 and S7 and Table S3.

Material and methods

Ethics statement

All experiments involving animals were approved by the Intitututo Gulbenkian de Ciência Ethics Committee and the Portuguese National Entity (Direção Geral de Alimentação e Veterinária; Ref. number 008958), which complies with European Directive 86/609/EEC of the European Council. Gnoto experimental work was developed with the support of the research infrastructure Congento, project LISBOA-01-0145-FEDER-022170, co-financed by Lisboa Regional Operational Programme (Lisboa 2020), Partnership Agreement, through the European Regional Development Fund (ERDF), and Foundation for Science and Technology (Portugal).

Bacterial strains and culture conditions

All *E. coli* strains used were derived from *Escherichia coli* K-12 MG1655, strains DM08-YFP and DM09-CFP as previously described [24]. See Table S6 for all strains, primers and plasmids. *E. coli* was routinely cultured in Lysogeny-Broth (LB) at 37˚C with aeration except when otherwise indicated. *lrp* gene was deleted using the λ red recombinase system encoded on the ampicillin-resistant plasmid pKD46 [79] to construct strain MFP002, with the primers described in Table S6. This deletion was introduced in the relevant backgrounds by bacteriophage P1-mediated transduction as described previously [80] using a lysate from MFP002. Bacteriophage P1-mediated transductions were also used to construct the *gatZ* and *galK* deletions using lysates from the donor strains listed in Table S6. Antibiotic resistance cassettes were removed from the mutants using the FRT/FLP recombinases of either pCP20 or pTL17 [79,81]. *Blautia* spp. present in C57BL/6 mice from our mouse facility, was isolated in anaerobically grown Tryptic soy broth + 5% sheep blood agar plates and identified as *B. coccoides* by 16S rRNA Sanger sequencing.

Fluorescent marker dynamics during mouse colonization

C57BL/6J mice GF or mono-associated with *B. coccoides* were bred and maintained in the gnotobiology facility of the Instituto Gulbenkian de Ciência. *E. coli* adaptation to the gut was followed as previously [24]. Briefly, 6- to 8-week old GF-raised or *B. coccoides* mono-associated C57BL/6 male mice were gavaged with 100µL of a suspension of 10^8 colony forming units (CFUs) of a mixture of YFP- and CFP-labelled *E. coli* (ratio 1:1) grown at 37°C in brain heart infusion liquid broth to OD_{600nm} of 2. After gavage, mice were housed individually in ISOcages (Tecniplast) with *ad libitum* food and water (sterilized by autoclaving). Mouse fecal pellets were collected for 23/40 ($Gnoto_{Ec}$) or 28 days ($Gnoto_{Ec+Bc}$), diluted in PBS and plated in Lysogeny Broth agar (LB agar). Plates were incubated overnight and the frequencies of YFP- and CFP-labelled bacteria were assessed by counting fluorescent colonies with a fluorescent stereoscope (Stereo Lumar, Carl Zeiss). A sample of each collected fecal pellet was also cryopreserved in 15% glycerol at -80°C, for further analysis.

Competitive fitness assays *in vivo*

Competitive advantage of *Δlrp* or *Δlrp ΔsdaA* mutants were measured *in vivo* against a reference (*gat*-negative) strain as previously described [37]. *In vivo* competitions were performed at 1 to 9 ratios (competitor to reference). For *in vivo* competitions in pre-colonized mice GF or *B. coccoides* mono-associated mice were first colonized with a streptomycin-sensitive, kanamycin-resistant ancestral *E. coli* strain (MFP121) for 10 days (see Fig. 3A). After that period (four hours before gavage with the competitor strains) streptomycin was administered in drinking water (5g/L) and maintained *ad libitum* for the remaining of the experiment, to clear the pre-colonizer *E. coli* strain. Clearance was confirmed by plating fecal pellets in LB+kanamycin (50µg/mL) media. We estimated the selective advantage of mutants over the reference strain from the slope of the linear regression of $\ln(freq_{mut}/freq_{ref})$ along time.

Whole genome sequencing and mutation prediction

To identify mutations segregating in *E. coli* populations during colonization of gnotobiotic mice we performed WGS analysis of population samples (approximately 1000 clones) isolated from fecal samples collected at different timepoints during the evolution experiments. In $Gnoto_{Ec}$ mice, we analyzed samples from day 23 from all animals and from

day 40 from four animals (Gn_{Ec} 1.4 - 1.7). In $Gnoto_{str}$ we analyzed samples from day 23 and in $Gnoto_{Bc+Ec}$ from day 28. Isolation of population samples and DNA extraction were performed as previously described (Barroso-Batista et al., 2015). The DNA library construction and sequencing were carried out at the Genomics Unit from Instituto Gulbenkian de Ciência. Each sample was pair-end sequenced on an Illumina MiSeq Benchtop Sequencer. Standard procedures produced data sets of Illumina paired-end 250-bp read pairs. The mean coverage per sample was between 113 and 179. Mutations were identified using the BRESEQ pipeline v0.23 with the polymorphism option and the same parameters previously used [37]. We defined parallel mutations (Tables S1 and S5) as mutational events that occurred in a minimum of two animals and that reached a minimum frequency of 5% in at least one population. Genome sequencing data have been deposited in the NCBI Read Archive database with accession code PRJNA429051.

Emergence and dynamics of mutations in the *gat* operon

To investigate the dynamics of appearance and expansion of beneficial mutations in the *gat* operon we determined the frequency of bacteria unable to metabolize galactitol within a given population following the protocol previously described [37].

Emergence and dynamics of mutations in *lrp*

To score for the *lrp*-negative phenotype, evolved clones were grown in solid minimal medium with glycine as the sole nitrogen source (glyMM) because growth of *lrp* deletion mutants was impaired in this medium, but the ancestral strain with the intact *lrp* gene was not (Fig. S2). This glyMM medium is a modified M9 minimal medium [82] with 0.5% glucose where ammonium chloride was replaced by 200µg/mL of glycine as nitrogen source. Cryopreserved fecal samples were thawed, diluted and plated in LB agar supplemented with streptomycin (100µg/mL). Then, single colonies were picked from the LB plates with toothpicks and diluted to individual wells of 96-well plates (Corning Costar), containing modified M9 medium (with no ammonium chloride and no carbon source). These suspensions of clones were used to inoculate glyMM agar (15g/L) plates by picking, as well as to inoculate MacConkey agar plates with 1% galactitol [24] to assess the *gat* phenotype and as a positive control for growth in rich medium. We also included the ancestral (DM09) and three *lrp* mutant strains (Δ*lrp* (MFP11), lrp_{cod} (MFP88) and lrp_{interg} (MFP89) as reference strains in every glyMM agar plate. glyMM agar plates were incubated for 48h at 42°C, while MacConkey galactitol agar plates [24] were incubated for 24h at 30°C. We tested 96 clones

per sample/time point. *lrp* phenotypes of each clone were scored by assessing growth in glyMM: Clones growing similarly to the ancestral DM09 strain were scored as *lrp*-positive and clones growing in MacConkey-galacitol but showing absence of growth in glyMM were scored as *lrp*-negative (loss of function mutant). Clones displaying growth, but inferior to the ancestral were classified as having an intermediate *lrp* phenotype and were further tested by targeted PCR of the *lrp* gene and scored as lrp_{interg} if an IS insertion in the intergenic region was detected (see Fig. S2E).

Comparison of *E. coli* mutation rate between hosts

To estimate the equilibrium frequency for antibiotic resistance clones, we determined the fraction of *E. coli* clones carrying spontaneous resistance to furazolidone, as previously described [37]. Mutations conferring resistance to furazolidone are expected to be found in *E. coli* populations at a mutation-selection balance and their frequency to reach a stable equilibrium, which is proportional to the mutation rate [37]. Three GF mice were colonized with *E. col*i strain DM08 (ancestral) following the protocol described above for the evolution experiments, and the frequency of spontaneous resistance was determined in fecal pellets collected for four time points over 16 days.

Growth curves of *E. coli* in vitro

Growth of ancestral and *lrp* mutants was tested in M9 minimal medium [82] supplemented with 0.5% of either serine or casaminoacids as carbon sources with trace amounts of isoleucine and valine (at 0.5mM concentration each) due to partial auxotrophy of *lrp* mutants for branch-chained amino acids [56,83]. We also tested the influence of leucine (inducer of Lrp) by adding this amino acid at a concentration of 10mM to the minimal media with either serine or casaminoacids. Inocula of the different strains tested were prepared by culturing in M9 minimal medium supplemented with 0.5% glucose for 48 hours (*lrp* mutants) or 24 hours (all other strains), harvested and washed twice with modified M9 medium with no carbon or nitrogen source. Each culture was normalized to OD_{600}= 0.1, diluted 1:100 into the appropriate media and aliquoted into multi-well Bioscreen plate (150µL/well). Plates were then incubated in a Bioscreen microplate reader with continuous, high amplitude, fast shaking settings, at 37°C or 42°C. Growth was followed by measuring optical density at 600nm, registered at 30 min intervals. Growth rate at exponential phase was calculated from the maximum slope of linear regression of $\ln(OD_{600})$ increase over time, during a 2 hour sliding window (5 points) in steady-state growth. Carrying capacity was calculated from the

average ODs during stationary phase. Absolute fitness is shown as that of each strain relative to the ancestral's average. Growth curves of the same strains in glyMM or M9 with 0.5% glucose liquid media were also preformed to establish the *lrp* phenotypic assay using the same methodology, except that cultures were incubated at 42°C because it has been shown before that this temperature accentuates the growth impairment of *lrp* in glucose [56,83].

Quantification of *B. coccoides* loads

To assess the loads of *B. coccoides* in the intestines of Gnoto$_{Bc+Ec}$ mice, we performed 16S rRNA gene sequencing in fecal samples collected from five mice. Briefly, DNA was extracted from frozen fecal samples with QIAamp DNA Stool Mini Kit (QIAGEN) following the manufacturer's instructions. 16S rRNA gene amplification and sequencing were carried out at the Genomics Unit from Instituto Gulbenkian de Ciência, following the service protocol with PCR amplification and pair-end sequencing on Illumina MiSeq Benchtop Sequencer, as previously described [75]. QIIME [84] was used to analyze the 16S rRNA sequences. First processing of raw demultiplexed reads was done using QIIME2 v.2018.4 (https://qiime2.org) with default parameters. DADA2 was used for quality filtering, denoising, paired-end merging, and amplicon sequence variant (ASV, i.e., sub–Operational Taxonomic Units) calling using QIIME dada2 denoise-paired method [85,86]. ASVs represented by less than 10 reads were removed from corresponding samples. Taxonomical classification was determined by matching ASVs against the Silva database [87] with feature-classifier *classify-sklearn*. Loads of *B. coccoides* per mouse were estimated by multiplying the number of *B. coccoides* reads with the ratio of the *E. coli* CFUs and the *E. coli* reads, after correcting for 16S copy number (7 for *E. coli*, 5 for *B. coccoides*, in accordance with the *rrn*DB [88].

Metabolomics of gut contents

To evaluate the composition of the metabolic environment in the gut, we performed ^{1}H-NMR analysis on aqueous extracts of cecum or fecal contents as specified. Cecal contents of GF, Gnoto$_{Ec}$ (31 days of colonization) and Gnoto$_{Bc+Ec}$ (28 days of colonization) were diluted 50% (w/v) in deuterated water (D_2O, Sigma-Aldrich) and those of SPF$_{str}$ (28 days of streptomycin treatment) 25% in D_2O. Fecal samples of Gnoto$_{Ec}$ were diluted in 1 mL of D_2O. For all samples, extraction was performed as follows: ~0.3g of 0.1mm glass beads (Scientific Industries SI-BG01) were added to each tube and samples bead-beaten using a Qiagen Tissuelyser II (Retsch) for 2 min with 30 rev/s pulses. Large debris and the glass beads were

pelleted by centrifugation at 14000 RPM for 30 min at 4°C. Supernatant was collected and filtered through a 0.22µm filter (Milipore), followed by another filtration step with 3KDa filters (Vivaspin 500) by centrifugation at 15000 g and 4°C for 3 hours or until 150µL of filtrate was obtained. Samples were stored at -80°C until spectrum acquisition. For acquisition, samples (150µL of filtrate) were thawed at room temperature for 10 to 15 min and then mixed with 60 µL of 350 mM phosphate buffer pH 7.09 with 2% NaN_3, 10 µL of a 0.05% (w/v) 3-(Trimethylsilyl)propionic-2,2,3,3-d4 (TSP-d4, Sigma-Aldrich) solution , and 380 µL of D_2O (to perform a total volume of 600 µL). This mixture was transferred to a 5mm glass NMR tube. All solutions were prepared in D_2O. Samples were homogenized by inversion and spectra acquired after pH measurement. Acquisitions were performed on a Bruker AVANCE II+ 500 MHz instrument equipped with Cryo TCI (F) (Prodigy) 5mm probehead with z-gradients. ^{1}H-NMR spectra were acquired using 1D NOESY pulse sequence with pre-saturation (noesypr1d) under the following conditions: 90 degrees pulse for excitation mixing time 100 ms, acquisition time 4 s, relaxation delay 1 s. All spectra were acquired with 200 scans at 25 ºC, with 48k data points and 6002 Hz (12 ppm) spectral width (Chenomx acquisition parameters). The recorded ^{1}H-NMR spectra were phase corrected using Bruker TopSpin 3.2 and spectra were then processed using Chenomx NMR Suite 8.11. Compounds were identified by manually fitting reference peaks to spectra in database Chenomx 500 MHz Version 10. Quantification was based on internal standard peak integration (TSP-d4). Heatmap and PCA were created using MetaboAnalyst (https://www.metaboanalyst.ca/).

Statistical analyses

Statistical analyses were performed in Microsoft Office Excel, Graphpad Prism v.6 or R software: http:// www.r-project.org/.

Supplemental Information titles and legends

Figure S1 - Loads of *E. coli* during adaptation to the gut of Gnoto$_{Ec}$ mice and comparison of eco-evolutionary parameters between Gnoto$_{Ec}$ and SPF$_{str+Ec}$. (**A**) Loads of *E. coli* (± 2 S.E.M.) during adaptation to the Gnoto$_{Ec}$ mouse gut in individually-housed mice monocolonized with a mixture of 1:1 ratio of DM08 and DM09 (n=10) as described in Fig. 1A. Experiment was performed in 3 independent blocks (Gn$_{Ec}$ 1-3, 4-8, 9-10). (**B**) Average loads of *E. coli* (± 2 S.E.M.) in fecal samples collected at indicated time points as a proxy for population size. *E. coli* loads are higher is Gnoto$_{Ec}$ mouse compared to SPF$_{str+Ec}$

[S1] (host type, ANOVA $\chi 21=4.4$, $P=0.036$)). (**C**) Mutation frequency for furazolidone resistance (as indicator of mutation rate) *in vivo* is similar between Gnoto$_{Ec}$ and SPF$_{str+Ec}$ [S2] (Mann-Whitney U test, $P=0.9$). Related to Fig. 1 and 2.

Figure S2 - Ancestral and Δ*lrp* mutant can be distinguished phenotypically through growth in glyMM medium, independently of the *gat* mutation. (**A**) Growth curves of ancestral (DM09), Δ*gatZ* mutant (MFP54), Δ*lrp* mutant (MFP008) and Δ*lrp* Δ*gatZ* mutant (MFP106) in M9 minimal medium with 0.5% glucose, at 42°C, show impaired growth of *lrp* mutants when compared to the ancestral and Δ*gatZ* strains. (**B**) Growth defect of Δ*lrp* mutants is exacerbated in glyMM, (where ammonium is replaced with glycine as nitrogen source, see Methods), and thus this medium was chosen for the phenotypic test to score for *lrp*-negative phenotype. Growth curves in (A) and (B) are averages (± 2 S.E.M.) from 9 replicates (3 independent experiments). (**C**) Growth in solid medium (at 42°C) of the same strains described in (A) on LB (left panel), M9 minimal medium with glucose (middle panel) or glyMM (right panel), show that glyMM agar plates can be used to score for *lrp*-negative phenotype. (**D**) Growth of ancestral (DM09), Δ*lrp* (MFP08), *lrp*$_{interg}$ (MFP89) and coding *lrp*$_{cod}$ (MFP88) strains in glyMM. *lrp*$_{cod}$ evolved clone is phenotypically similar to Δ*lrp* mutant, while *lrp*$_{interg}$ evolved clone has an intermediate phenotype. (**E**) Targeted PCR for the *lrp* locus performed on the strains described in (D) with primers either for the complete *lrp* locus (left) or only for the intergenic region upstream of the *lrp* starting codon (right). Clones with an insertion of an IS element (evolved clones, lanes 3 and 4) displayed a band with a larger size than that of the ancestral (lane 1), while for Δ*lrp* (lane 2), the band was smaller. Amplification of the intergenic region upstream of *lrp* open reading frame allows for identification of clones carrying intergenic IS insertions as these display a larger band size (lane 4) compared to the ancestral (lane 1) or the other mutants (lanes 2 and 3). Lane 5 is the no-DNA negative control. (**F**) Individual plots of frequencies of *lrp*-negative (circles) and *lrp*-intermediate (triangles) mutant phenotypes (± 2 S.E.M.) for all Gnoto$_{Ec}$ populations (as shown in Fig. 2). Experiments performed separately as Gn$_{Ec}$ 1.1-1.3, Gn$_{Ec}$ 1.4-1.8, Gn$_{Ec}$ 1.9-1.10 (mouse 8 died after day 23). These results show that both *lrp*-negative and *lrp*-intermediate phenotypes emerge in *E. coli* populations evolving in Gnoto$_{Ec}$ mice. Related to Fig. 2.

Figure S3 - *lrp* mutant is not advantageous in GF mice, during competitive fitness experiment without pre-colonization. Total *E. coli* loads (± 2 S.E.M.) (**A**) and frequency

(± 2 S.E.M.) of Δ*lrp* Δ*gatZ* (**B**) during *in vivo* competition of Δ*lrp* Δ*gatZ* (MFP106) against Δ*gatZ* (MFP54) introduced at 1:9 ratio by oral gavage in GF mice (n=3, M1-M3). (**C**) Bacterial loads (± 2 S.E.M.) from the competition experiment of Δ*lrp* Δ*gatZ* mutant against Δ*gatZ* in Gnoto$_{Ec}$ mice (n=5, M1-M5) pre-colonized with an ancestral streptomycin-sensitive strain (MFP121, green full symbols). Competitor strains (depicted in open black symbols) were introduced at 1:9 ratio by oral gavage in GF mice four hours after the start of streptomycin treatment. Shaded area denotes period of streptomycin treatment. Related to Fig. 3.

Figure S4 - GF and SPF$_{str}$ ceca exhibit distinct metabolomic profiles, in particular in respect to amino acids concentration. (**A**-**C**) Concentration of small metabolites from cecal contents, determined by ^{1}H-NMR and grouped by chemical attributes. Absolute concentrations (with medians shown as lines) of all metabolites detected in GF (blue circles, n=5) and SPF$_{str}$ (brown squares, n=10) is shown for amino acids (**A**), organic acids (**B**) and sugars and other metabolites (**C**). Metabolites whose concentration significantly differs between GF and SPF$_{str}$ display the level of significance: $P \leq 0.05$ (*), $P \leq 0.01$ (**) or $P > 0.05$, non-significant (ns), after *post-hoc* Mann-Whitney *U*-test with Holm's correction for multiple comparisons. Related to Fig. 4.

Figure S5 - Metabolite dynamics show rapid depletion of serine and threonine during *E. coli* adaptation in Gnoto$_{Ec}$ mice. Metabolites were determined by ^{1}H-NMR in fecal samples collected during Gnoto$_{Ec}$ colonization (n=3). Bars show relative concentration of each metabolite over time (left axis) and circles depict the average total concentration (± 2 S.E.M.) of amino acids (**A**), organic acids (**B**), sugars (**C**) and other metabolites (**D**) over time (right axis). Uncolonized samples (Un) were collected the day prior to gavage with *E. coli*. Related to Fig. 4.

Figure S6 - *lrp* mutation confers a specific growth advantage. (A) Growth in casaminoacids show similar profiles independently of *lrp*. Growth curves of Δ*lrp* Δ*gatZ* (MFP106), *lrp*$_{cod}$ Δ*gatZ* (MFP102), *lrp*$_{interg}$ Δ*gatZ* (MFP104) and Reference (Δ*gatZ*, MFP54) in M9 medium with 0.5% casaminoacids as carbon source, in the absence (left panel) or presence (right panel) of leucine 10Mm (n=9 from 3 independent experiments). Estimated growth parameters in Table S4. Error bars represent ± 2 S.E.M. **(B)** Bacterial loads (± 2 S.E.M.) from the competition experiment of Δ*lrp ΔsdaA* Δ*gatZ* mutant (MPF125) against Δ*gatZ* (MFP54) in Gnoto$_{Ec}$ mice (n=5, M1-M5) pre-colonized with an ancestral (MFP121)

streptomycin-sensitive strain (green full symbols). Competitor strains (depicted in open black symbols) were introduced at 1:9 ratio by oral gavage in GF mice four hours after the start of streptomycin treatment. Shaded area denotes period of streptomycin treatment. Related to Fig. 5.

Figure S7 - Evolution and metabolic environment of *E. coli* in Gnoto$_{Bc+Ec}$ mice differs from Gnoto$_{Ec}$. (**A**) Loads of *E. coli* (± 2 S.E.M.) during the adaptation experiment in the presence of *B. coccoides* (n=5). (**B**) Loads of *B. coccoides* at the end of the evolution experiment in (A). (**C**) Competition experiment of Δ*lrp* Δ*gatZ* (MFP106) mutant against Δ*gatZ* (MFP54) in Gnoto$_{Ec}$ mice (n=5, M1-M5) pre-colonized with an ancestral (MFP121) streptomycin-sensitive strain (green full symbols). Competitor strains (depicted in open black symbols) were introduced at 1:9 ratio by oral gavage in GF mice four hours after the start of streptomycin treatment. Bacterial loads (± 2 S.E.M.) during competition are shown, with shaded area denoting period of streptomycin treatment. (**D**) Loads of *B. coccoides* during the competition experiment shown in (C). (**E**) Score plot for PC1 and PC2 from principal component analysis (PCA) of all metabolites found in GF (blue circles), Gnoto$_{Ec}$ (black triangles), Gnoto$_{EcBc}$ (red diamonds) and SPF$_{str}$ (brown squares). Colored circles represent 95 % confidence intervals. (**F**-**H**) Concentration of small metabolites from cecal contents, determined by ^{1}H-NMR and grouped by chemical attributes. Absolute concentrations (with medians shown as lines) of all metabolites detected in Gnoto$_{Ec}$ (black triangles, n=6) and Gnoto$_{Bc+Ec}$ (red diamonds, n=5) is shown for amino acids (**F**), organic acids (**G**) and sugars and other metabolites (**H**). Metabolites whose concentration significantly differs between Gnoto$_{Ec}$ and Gnoto$_{Bc+Ec}$ display the level of significance: $P \leq 0.05$ (*), $P \leq 0.01$ (**) or $P > 0.05$, non-significant (ns), after *post-hoc* Mann-Whitney *U*-test with Holm's correction for multiple comparisons. Related to Fig. 6 and 7.

Table S1 - Parallel mutations segregating in *E. coli* populations evolving in the absence of interspecies competition. Loci were identified by WGS of populations samples collected from Gnoto$_{Ec}$ mice after 23 days of evolution, and parallel mutations defined as evets found in more than one evolving population at a frequency greater than 5%. Functional annotation is shown for each locus, with processes related with amino acid metabolism highlighted in bold. Number of Gnoto$_{Ec}$ populations (n=10) where each mutation was found is shown, as well as the number of Gnoto$_{str+Ec}$ populations (n=3) where Gnoto$_{Ec}$ parallel mutations were also detected. Related to Fig. 1.

Table S2 - Mutations found in *E. coli* populations after adaptation to the gut of gnotobiotic mice. WGS analysis of *E. coli* population samples evolved in: Gnoto$_{Ec}$ for 23 (n=10) or 40 days (n=4), Gnoto$_{str+Ec}$ for 23 days (n=3) and Gnoto$_{Bc+Ec}$ for 28 days (n=5). For intergenic mutations the two flanking genes are listed, otherwise the mutation occurred in the gene coding region. SNPs are represented by an arrow between the ancestral and the evolved nucleotide. The symbol Δ means a deletion event and a + symbol represents an insertion of the nucleotide that follows the symbol. The initials IS denote the abbreviation of insertion sequence element at the indicated position. del/dup indicates that either a deletion or a duplication of the indicated size occurred but it is not possible to distinguish between the two. * indicates that the mutation corresponds to a supported unassigned new junction whereas ¤ denotes an IS insertion where only one new junction was identified. Related to Fig. 1 and 6.

Table S3 - List of metabolites identified by ^{1}H-NMR in the cecum of GF (n=5), SPF$_{str+Ec}$ (n=10), Gnoto$_{Ec}$ (n=6) and Gnoto$_{Bc+Ec}$ (n=5) or in fecal samples of Gnoto$_{Ec}$ (n=3) mice. Concentration of metabolites in nmol per gram of cecal content or feces. SPF$_{str+Ec}$, Gnoto$_{Ec}$ and Gnoto$_{Bc+Ec}$ mice were colonized with ancestral *E. coli* strains for approximately one month prior to metabolomic analysis. Related to Fig. 4 and 7.

Table S4 - Serine metabolism provides fitness advantage to *lrp* mutants. Growth parameters of Δ*gatZ* (MFP54), *Δlrp ΔgatZ* (MFP106), *lrp*$_{cod}$ *ΔgatZ* (MFP102) and *lrp*$_{interg}$ Δ*gatZ* (MFP104) *E. coli* strains estimated from growth curves in M9 medium supplemented with 0.5% serine (Ser) or 0.5% casaminoacids (CAA) and in the absence or presence (+leu) of Leucine. Averages are shown with error (2se) represented within parenthesis. Strain parameters were compared relative to the reference strain (Δ*gatZ*) in columns Ser and CAA; and within strain, for Leucine effect, in columns Ser+leu and CAA+leu; (*) denotes significance after Mann-Whitney U-test with Holm correction for multiple comparisons ($P \leq 0.01$, except in CAA with $0.01 \leq P \leq 0.05$; ns – not significant). N=9 from 3 independent experiments. Related to Fig. 5.

Table S5 - Parallel mutations segregating in *E. coli* populations evolving in the presence of *B. coccoides*. Loci were identified by WGS of populations samples collected from Gnoto$_{Bc+Ec}$ mice after 28 days of evolution, and parallel mutations defined with the same

criteria as in Table S1. We also included mutations that were found in more than one mouse and were shared with SPF_{str+Ec} or $Gnoto_{Ec}$ regardless of frequency. Number of $Gnoto_{Bc+Ec}$ populations (n=5) where each mutation was found is shown, as well as presence/absence of these mutations in $Gnoto_{Ec}$ or SPF_{str+Ec}-evolved populations. Related to Fig. 6.

References

1. Robertson, R.C., Manges, A.R., Finlay, B.B., and Prendergast, A.J. (2019). The Human Microbiome and Child Growth – First 1000 Days and Beyond. Trends in Microbiology *27*, 131–147.
2. Koskella, B., Hall, L.J., and Metcalf, C.J.E. (2017). The microbiome beyond the horizon of ecological and evolutionary theory. Nat Ecol Evol *1*, 1606–1615.
3. Sousa, A., Frazão, N., Ramiro, R.S., and Gordo, I. (2017). Evolution of commensal bacteria in the intestinal tract of mice. Curr. Opin. Microbiol. *38*, 114–121.
4. Gordo, I. (2019). Evolutionary change in the human gut microbiome: From a static to a dynamic view. PLOS Biology *17*, e3000126.
5. Smith, K., McCoy, K.D., and Macpherson, A.J. (2007). Use of axenic animals in studying the adaptation of mammals to their commensal intestinal microbiota. Seminars in Immunology *19*, 59–69.
6. Flint, H.J., Scott, K.P., Louis, P., and Duncan, S.H. (2012). The role of the gut microbiota in nutrition and health. Nat Rev Gastroenterol Hepatol *9*, 577–589.
7. Hooper, L.V., Littman, D.R., and Macpherson, A.J. (2012). Interactions between the microbiota and the immune system. Science *336*, 1268–1273.
8. Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S., and Medzhitov, R. (2004). Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. Cell *118*, 229–241.
9. Gilbert, J.A., Blaser, M.J., Caporaso, J.G., Jansson, J.K., Lynch, S.V., and Knight, R. (2018). Current understanding of the human microbiome. Nat. Med. *24*, 392–400.
10. Dominguez-Bello, M.G., Godoy-Vitorino, F., Knight, R., and Blaser, M.J. (2019). Role of the microbiome in human development. Gut *68*, 1108–1114.
11. Kho, Z.Y., and Lal, S.K. (2018). The Human Gut Microbiome – A Potential Controller of Wellness and Disease. Front. Microbiol. *9*. Available at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.01835/full [Accessed June 6, 2019].
12. Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., Nelson, K.E., and Relman, D.A. (2005). Diversity of the Human Intestinal Microbial Flora. Science *308*, 1635–1638.
13. Almeida, A., Mitchell, A.L., Boland, M., Forster, S.C., Gloor, G.B., Tarkowska, A., Lawley, T.D., and Finn, R.D. (2019). A new genomic blueprint of the human gut microbiota. Nature *568*, 499.

14. Lloyd-Price, J., Arze, C., Ananthakrishnan, A.N., Schirmer, M., Avila-Pacheco, J., Poon, T.W., Andrews, E., Ajami, N.J., Bonham, K.S., Brislawn, C.J., *et al.* (2019). Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature *569*, 655–662.

15. Vernocchi, P., Del Chierico, F., and Putignani, L. (2016). Gut Microbiota Profiling: Metabolomics Based Approach to Unravel Compounds Affecting Human Health. Front Microbiol *7*. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4960240/ [Accessed June 5, 2019].

16. Nicholson, J.K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., and Pettersson, S. (2012). Host-gut microbiota metabolic interactions. Science *336*, 1262–1267.

17. Lozupone, C.A., Stombaugh, J.I., Gordon, J.I., Jansson, J.K., and Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. Nature *489*, 220–230.

18. Costea, P.I., Hildebrand, F., Arumugam, M., Bäckhed, F., Blaser, M.J., Bushman, F.D., de Vos, W.M., Ehrlich, S.D., Fraser, C.M., Hattori, M., *et al.* (2018). Enterotypes in the landscape of gut microbial community composition. Nature Microbiology *3*, 8.

19. Round, J.L., and Palm, N.W. (2018). Causal effects of the microbiota on immune-mediated diseases. Science Immunology *3*, eaao1603.

20. Ley, R.E., Lozupone, C.A., Hamady, M., Knight, R., and Gordon, J.I. (2008). Worlds within worlds: evolution of the vertebrate gut microbiota. Nat. Rev. Microbiol. *6*, 776–788.

21. Moeller, A.H., Caro-Quintero, A., Mjungu, D., Georgiev, A.V., Lonsdorf, E.V., Muller, M.N., Pusey, A.E., Peeters, M., Hahn, B.H., and Ochman, H. (2016). Cospeciation of gut microbiota with hominids. Science *353*, 380–382.

22. Giraud, A., Matic, I., Tenaillon, O., Clara, A., Radman, M., Fons, M., and Taddei, F. (2001). Costs and benefits of high mutation rates: adaptive evolution of bacteria in the mouse gut. Science *291*, 2606–2608.

23. Goodman, A.L., McNulty, N.P., Zhao, Y., Leip, D., Mitra, R.D., Lozupone, C.A., Knight, R., and Gordon, J.I. (2009). Identifying Genetic Determinants Needed to Establish a Human Gut Symbiont in Its Habitat. Cell Host & Microbe *6*, 279–289.

24. Barroso-Batista, J., Sousa, A., Lourenço, M., Bergman, M.-L., Sobral, D., Demengeot, J., Xavier, K.B., and Gordo, I. (2014). The First Steps of Adaptation of Escherichia coli to the Gut Are Dominated by Soft Sweeps. PLoS Genet. *10*, e1004182.

25. Crook, N., Ferreiro, A., Gasparrini, A.J., Pesesky, M.W., Gibson, M.K., Wang, B., Sun, X., Condiotte, Z., Dobrowolski, S., Peterson, D., *et al.* (2019). Adaptive Strategies of the Candidate Probiotic E. coli Nissle in the Mammalian Gut. Cell Host & Microbe *25*, 499-512.e8.

26. Garud, N.R., Good, B.H., Hallatschek, O., and Pollard, K.S. (2019). Evolutionary dynamics of bacteria in the gut microbiome within and across hosts. PLOS Biology *17*, e3000102.

27. Walter, J., Britton, R.A., and Roos, S. (2011). Host-microbial symbiosis in the vertebrate gastrointestinal tract and the Lactobacillus reuteri paradigm. Proc. Natl. Acad. Sci. U.S.A. *108 Suppl 1*, 4645–4652.

28. Hoang, K.L., Morran, L.T., and Gerardo, N.M. (2016). Experimental Evolution as an Underutilized Tool for Studying Beneficial Animal–Microbe Interactions. Front. Microbiol. *7*. Available at: https://www.frontiersin.org/articles/10.3389/fmicb.2016.01444/full [Accessed June 5, 2019].

29. Lescat, M., Launay, A., Ghalayini, M., Magnan, M., Glodt, J., Pintard, C., Dion, S., Denamur, E., and Tenaillon, O. (2016). Using long-term experimental evolution to uncover the patterns and determinants of molecular evolution of an Escherichia coli natural isolate in the streptomycin treated mouse gut. Mol. Ecol.

30. De Paepe, M., Gaboriau-Routhiau, V., Rainteau, D., Rakotobe, S., Taddei, F., and Cerf-Bensussan, N. (2011). Trade-off between bile resistance and nutritional competence drives Escherichia coli diversification in the mouse gut. PLoS Genet. *7*, e1002107.

31. Giraud, A., Arous, S., De Paepe, M., Gaboriau-Routhiau, V., Bambou, J.-C., Rakotobe, S., Lindner, A.B., Taddei, F., and Cerf-Bensussan, N. (2008). Dissecting the genetic components of adaptation of Escherichia coli to the mouse gut. PLoS Genet. *4*, e2.

32. Leatham-Jensen, M.P., Frimodt-Møller, J., Adediran, J., Mokszycki, M.E., Banner, M.E., Caughron, J.E., Krogfelt, K.A., Conway, T., and Cohen, P.S. (2012). The Streptomycin-Treated Mouse Intestine Selects Escherichia coli envZ Missense Mutants That Interact with Dense and Diverse Intestinal Microbiota. Infect. Immun. *80*, 1716–27.

33. Tenaillon, O., Skurnik, D., Picard, B., and Denamur, E. (2010). The population genetics of commensal Escherichia coli. Nat Rev Micro *8*, 207–217.

34. Bäckhed, F., Roswall, J., Peng, Y., Feng, Q., Jia, H., Kovatcheva-Datchary, P., Li, Y., Xia, Y., Xie, H., Zhong, H., *et al.* (2015). Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. Cell Host Microbe *17*, 690–703.

35. Jost, T., Lacroix, C., Braegger, C.P., and Chassard, C. (2012). New Insights in Gut Microbiota Establishment in Healthy Breast Fed Neonates. PLOS ONE *7*, e44595.

36. Lourenço, M., Ramiro, R.S., Güleresi, D., Barroso-Batista, J., Xavier, K.B., Gordo, I., and Sousa, A. (2016). A Mutational Hotspot and Strong Selection Contribute to the Order of Mutations Selected for during Escherichia coli Adaptation to the Gut. PLOS Genetics *12*, e1006420.

37. Barroso-Batista, J., Demengeot, J., and Gordo, I. (2015). Adaptive immunity increases the pace and predictability of evolutionary change in commensal gut bacteria. Nat Commun *6*, 8945.

38. Long, A., Liti, G., Luptak, A., and Tenaillon, O. (2015). Elucidating the molecular architecture of adaptation via evolve and resequence experiments. Nat Rev Genet *16*, 567–582.

39. Shimada, T., Saito, N., Maeda, M., Tanaka, K., and Ishihama, A. (2015). Expanded roles of leucine-responsive regulatory protein in transcription regulation of the Escherichia coli genome: Genomic SELEX screening of the regulation targets. Microbial Genomics *1*. Available at: http://mgen.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.000001 [Accessed February 13, 2017].

40. Harder, D., Stolz, J., Casagrande, F., Obrdlik, P., Weitz, D., Fotiadis, D., and Daniel, H. (2008). DtpB (YhiP) and DtpA (TppB, YdgR) are prototypical proton-dependent peptide transporters of Escherichia coli. FEBS J. *275*, 3290–3298.

41. Hagewood, B.T., Ganduri, Y.L., and Datta, P. (1994). Functional analysis of the tdcABC promoter of Escherichia coli: roles of TdcA and TdcR. J. Bacteriol. *176*, 6214–6220.

42. Ganduri, Y.L., Sadda, S.R., Datta, M.W., Jambukeswaran, R.K., and Datta, P. (1993). TdcA, a transcriptional activator of the tdcABC operon of Escherichia coli, is a member of the LysR family of proteins. Mol. Gen. Genet. *240*, 395–402.

43. Kim, S.H., Schneider, B.L., and Reitzer, L. (2010). Genetics and regulation of the major enzymes of alanine synthesis in Escherichia coli. J. Bacteriol. *192*, 5304–5311.

44. Kuper, C., and Jung, K. (2005). CadC-mediated activation of the cadBA promoter in Escherichia coli. J. Mol. Microbiol. Biotechnol. *10*, 26–39.

45. Wiame, E., Delpierre, G., Collard, F., and Schaftingen, E.V. (2002). Identification of a Pathway for the Utilization of the Amadori Product Fructoselysine in Escherichia coli. J. Biol. Chem. *277*, 42523–42529.

46. Wiame, E., and Van Schaftingen, E. (2004). Fructoselysine 3-epimerase, an enzyme involved in the metabolism of the unusual Amadori compound psicoselysine in Escherichia coli. Biochem. J. *378*, 1047–1052.

47. Muffler, A., Fischer, D., Altuvia, S., Storz, G., and Hengge-Aronis, R. (1996). The response regulator RssB controls stability of the sigma(S) subunit of RNA polymerase in Escherichia coli. EMBO J. *15*, 1333–1339.

48. Schweizer, H., Boos, W., and Larson, T.J. (1985). Repressor for the sn-glycerol-3-phosphate regulon of Escherichia coli K-12: cloning of the glpR gene and identification of its product. J. Bacteriol. *161*, 563–566.

49. Reed, J.L., Patel, T.R., Chen, K.H., Joyce, A.R., Applebee, M.K., Herring, C.D., Bui, O.T., Knight, E.M., Fong, S.S., and Palsson, B.O. (2006). Systems approach to refining genome annotation. PNAS *103*, 17480–17484.

50. Parisutham, V., and Lee, S.K. (2015). Novel Functions and Regulation of Cryptic Cellobiose Operons in Escherichia coli. PLOS ONE *10*, e0131928.

51. Gerlach, P., Valentin-Hansen, P., and Bremer, E. (1990). Transcriptional regulation of the cytR repressor gene of Escherichia coli: autoregulation and positive control by the cAMP/CAP complex. Mol. Microbiol. *4*, 479–488.

52. Stauffer, L.T., and Stauffer, G.V. (1999). Role for the leucine-responsive regulatory protein (Lrp) as a structural protein in regulating the Escherichia coli gcvTHP operon. Microbiology (Reading, Engl.) *145 (Pt 3)*, 569–576.

53. Leatham, M.P., Banerjee, S., Autieri, S.M., Mercado-Lubo, R., Conway, T., and Cohen, P.S. (2009). Precolonized human commensal Escherichia coli strains serve as a barrier to E. coli O157:H7 growth in the streptomycin-treated mouse intestine. Infect. Immun. *77*, 2876–86.

54. Reitzer, L. (2005). Catabolism of Amino Acids and Related Compounds. EcoSal Plus *1*.

55. Calvo, J.M., and Matthews, R.G. (1994). The leucine-responsive regulatory protein, a global regulator of metabolism in Escherichia coli. Microbiol. Rev. *58*, 466–490.

56. Zinser, E.R., and Kolter, R. (1999). Mutations enhancing amino acid catabolism confer a growth advantage in stationary phase. J. Bacteriol. *181*, 5800–5807.

57. Zinser, E.R., and Kolter, R. (2000). Prolonged stationary-phase incubation selects for lrp mutations in Escherichia coli K-12. J. Bacteriol. *182*, 4361–4365.

58. Eren, A.M., Sogin, M.L., Morrison, H.G., Vineis, J.H., Fisher, J.C., Newton, R.J., and McLellan, S.L. (2015). A single genus in the gut microbiome reflects host preference and specificity. ISME J *9*, 90–100.

59. Liu, C., Finegold, S.M., Song, Y., and Lawson, P.A. (2008). Reclassification of Clostridium coccoides, Ruminococcus hansenii, Ruminococcus hydrogenotrophicus, Ruminococcus luti, Ruminococcus productus and Ruminococcus schinkii as Blautia coccoides gen. nov., comb. nov., Blautia hansenii comb. nov., Blautia hydrogenotrophica comb. nov., Blautia luti comb. nov., Blautia producta comb. nov., Blautia schinkii comb. nov. and description of Blautia wexlerae sp. nov., isolated from human faeces. Int. J. Syst. Evol. Microbiol. *58*, 1896–1902.

60. Brugiroux, S., Beutler, M., Pfann, C., Garzetti, D., Ruscheweyh, H.-J., Ring, D., Diehl, M., Herp, S., Lötscher, Y., Hussain, S., *et al.* (2016). Genome-guided design of a defined mouse microbiota that confers colonization resistance against Salmonella enterica serovar Typhimurium. Nat Microbiol *2*, 16215.

61. Stickland, L.H. (1934). Studies in the metabolism of the strict anaerobes (genus Clostridium): The chemical reactions by which Cl. sporogenes obtains its energy. Biochem. J. *28*, 1746–1759.

62. Fischbach, M.A., and Sonnenburg, J.L. (2011). Eating for two: how metabolism establishes interspecies interactions in the gut. Cell Host Microbe *10*, 336–347.

63. Smith, E.A., and Macfarlane, G.T. (1998). Enumeration of amino acid fermenting bacteria in the human large intestine: effects of pH and starch on peptide metabolism and dissimilation of amino acids. FEMS Microbiology Ecology *25*, 355–368.

64. Schmid, K., and Schmitt, R. (1976). Raffinose Metabolism in Escherichia coli K12. European Journal of Biochemistry *67*, 95–104.

65. Kroner, G.M., Wolfe, M.B., and Freddolino, P.L. (2019). Escherichia coli Lrp Regulates One-Third of the Genome via Direct, Cooperative, and Indirect Routes. Journal of Bacteriology *201*, e00411-18.

66. Finkel, S.E. (2006). Long-term survival during stationary phase: evolution and the GASP phenotype. Nat. Rev. Microbiol. *4*, 113–120.

67. Baek, C.-H., Wang, S., Roland, K.L., and Curtiss, R. (2009). Leucine-Responsive Regulatory Protein (Lrp) Acts as a Virulence Repressor in Salmonella enterica Serovar Typhimurium. Journal of Bacteriology *191*, 1278–1292.

68. Cordone, A., Mauriello, E.M.F., Pickard, D.J., Dougan, G., Felice, M.D., and Ricca, E. (2005). The lrp Gene and Its Role in Type I Fimbriation in Citrobacter rodentium. Journal of Bacteriology *187*, 7009–7017.

69. Engstrom, M.D., and Mobley, H.L.T. (2016). Regulation of Expression of Uropathogenic Escherichia coli Nonfimbrial Adhesin TosA by PapB Homolog TosR in Conjunction with H-NS and Lrp. Infect Immun *84*, 811–821.

70. Gauger, E.J., Leatham, M.P., Mercado-Lubo, R., Laux, D.C., Conway, T., and Cohen, P.S. (2007). Role of motility and the flhDC Operon in Escherichia coli MG1655 colonization of the mouse intestine. Infect. Immun. *75*, 3315–24.

71. Kaneuchi, C., Benno, Y., and Mitsuoka, T. (1976). Clostridium coccoides, a New Species from the Feces of Mice. International Journal of Systematic and Evolutionary Microbiology *26*, 482–486.

72. Charlesworth, B. (2010). Elements of Evolutionary Genetics (Roberts and Company Publishers).

73. Maddamsetti, R., Lenski, R.E., and Barrick, J.E. (2015). Adaptation, Clonal Interference, and Frequency-Dependent Interactions in a Long-Term Evolution Experiment with Escherichia coli. Genetics, genetics.115.176677.

74. Zarrinpar, A., Chaix, A., Yooseph, S., and Panda, S. (2014). Diet and Feeding Pattern Affect the Diurnal Dynamics of the Gut Microbiome. Cell Metabolism *20*, 1006–1017.

75. Sousa, A., Ramiro, R.S., Barroso-Batista, J., Güleresi, D., Lourenço, M., and Gordo, I. (2017). Recurrent Reverse Evolution Maintains Polymorphism after Strong Bottlenecks in Commensal Gut Bacteria. Mol Biol Evol *34*, 2879–2892.

76. Fabich, A.J., Jones, S. a, Chowdhury, F.Z., Cernosek, A., Anderson, A., Smalley, D., McHargue, J.W., Hightower, G.A., Smith, J.T., Autieri, S.M., *et al.* (2008). Comparison of carbon nutrition for pathogenic and commensal Escherichia coli strains in the mouse intestine. Infect. Immun. *76*, 1143–52.

77. Conway, T., and Cohen, P.S. (2015). Commensal and Pathogenic Escherichia coli Metabolism in the Gut. Microbiol Spectr *3*. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4510460/ [Accessed December 30, 2015].

78. Desai, M.S., Seekatz, A.M., Koropatkin, N.M., Kamada, N., Hickey, C.A., Wolter, M., Pudlo, N.A., Kitamoto, S., Terrapon, N., Muller, A., *et al.* (2016). A dietary fiber-deprived gut microbiota degrades the colonic mucus barrier and enhances pathogen susceptibility. Cell *167*, 1339-1353.e21.

79. Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proceedings of the National Academy of Sciences *97*, 6640–6645.

80. Silhavy, T.J., Berman, M.L., and Enquist, L.W. (1984). Experiments with gene fusions (Cold Spring Harbor, N.Y.: Cold Spring Harbor).

81. Long, T., Tu, K.C., Wang, Y., Mehta, P., Ong, N.P., Bassler, B.L., and Wingreen, N.S. (2009). Quantifying the integration of quorum-sensing signals with single-cell resolution. PLoS Biol *7*.

82. Sambrook, J., Fritsch, E.F., Maniatis, T., and Laboratory, C.S.H. (1989). Molecular cloning : a laboratory manual 2nd ed. (New York : Cold Spring Harbor Laboratory Press).

83. Ambartsoumian, G., D'Ari, R., Lin, R.T., and Newman, E.B. (1994). Altered amino acid metabolism in lrp mutants of Escherichia coli K12 and their derivatives. Microbiology (Reading, Engl.) *140 (Pt 7)*, 1737–1744.

84. Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., *et al.* (2010). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods *7*, 335–336.

85. Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods *13*, 581–583.

86. Ballesteros-Arias, L., Silva, J.G., Paiva, R.A., Carbonetto, B., Faísca, P., and Martins, V.C. (2019). T Cell Acute Lymphoblastic Leukemia as a Consequence of Thymus Autonomy. J. Immunol. *202*, 1137–1144.

87. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res *41*, D590–D596.

88. Klappenbach, J.A., Saxman, P.R., Cole, J.R., and Schmidt, T.M. (2001). rrndb: the Ribosomal RNA Operon Copy Number Database. Nucleic Acids Res. *29*, 181–184.

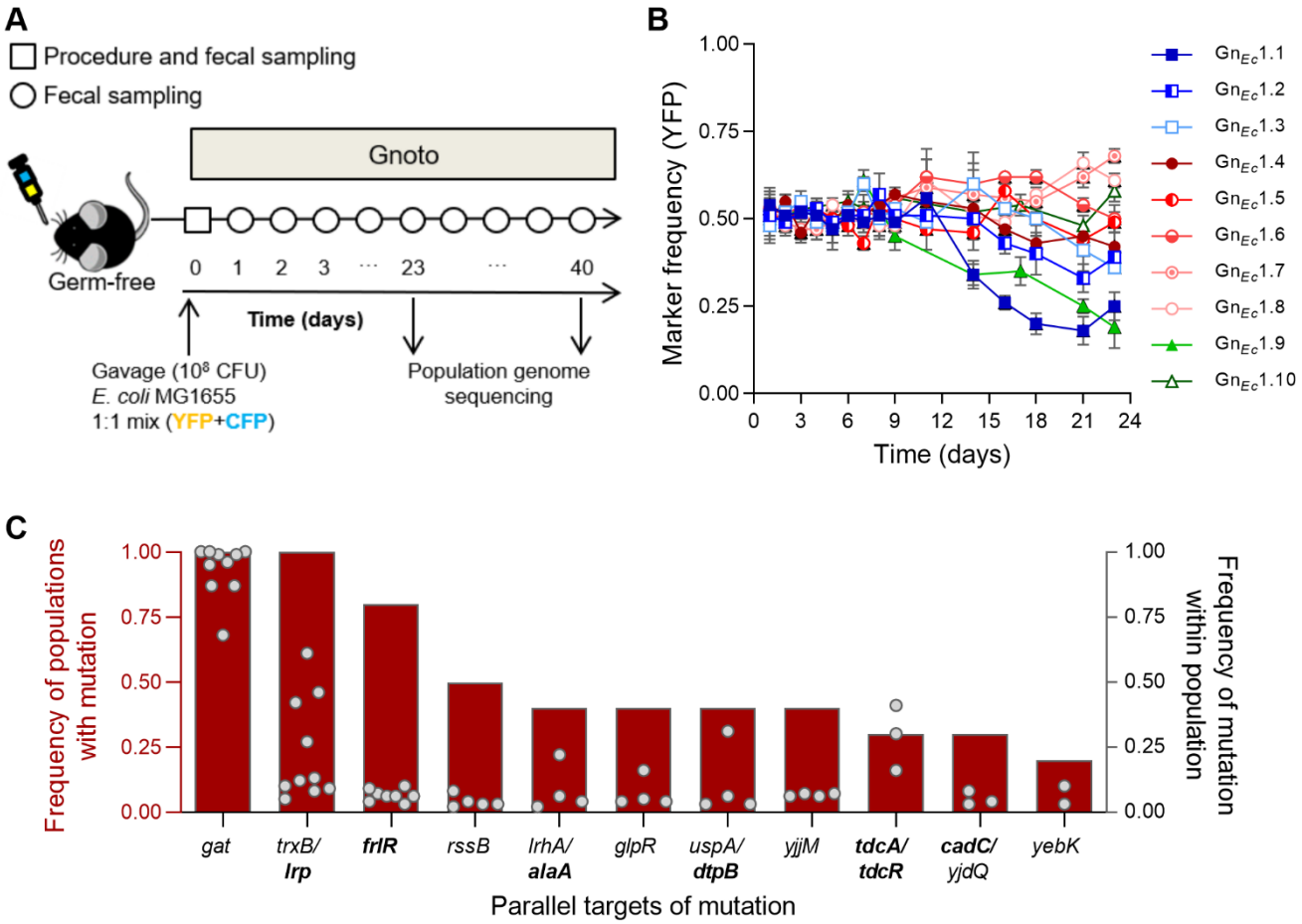

Fig. 1 - *E. coli* evolution in the Gnoto$_{Ec}$ mouse gut is characterized by mutations in genes involved in amino acid metabolism. (**A**) Design of evolution experiment. Germ free mice were monocolonized by gavage with *E. coli* (1:1 mix YFP + CPF) at day 0 and followed for approximately one month. Samples for WGS were collected from all populations at day 23 and from populations Gn$_{Ec}$1.4 to 1.7 at day 40. The experiment was repeated 3 times, with a total of n=10 mice. (**B**) Dynamics of neutral fluorescent marker (± 2 S.E.M.) during the first 23 days of adaptation of *E. coli* upon colonization of GF mice (n=10) provide evidence for adaptation. (**C**) Parallel targets of mutation identified by WGS in samples of populations collected at day 23. The height of the bars denotes the frequency of populations (n=10 for Gnoto$_{Ec}$) where a given mutation was present (left axis) and circles represent the frequency each mutated locus reached within the population (right axis). Loci related with amino acid metabolism are highlighted in bold. All parallel loci except for *yjjM* and *yebK* were also found in Gnoto$_{str+Ec}$, (n=3). See also Fig. S1 and Tables S1 and S2.

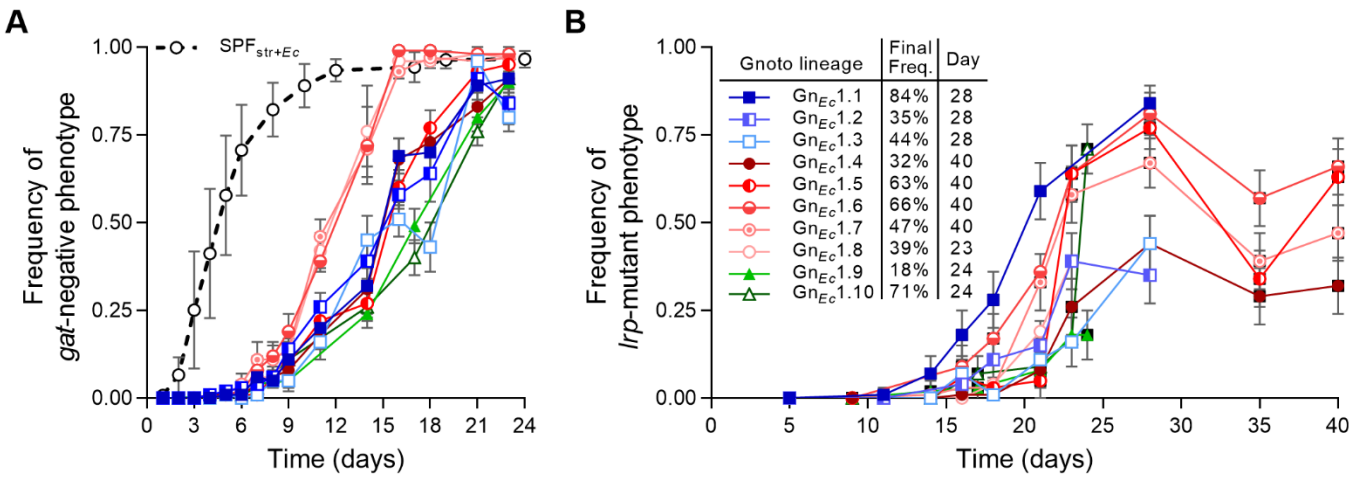

Fig. 2 - Emergence of g*at*- and *lrp*-mutant phenotypes reveals full parallelism. (**A**) Frequency of the *gat*-negative phenotype (± 2 S.E.M.) over time across individual populations in Gnoto$_{Ec}$ (colored lines, n=10, same legend as in (B)). In dashed line, the average (± 2 S.E.M.) of *gat*-negative phenotype across *E. coli* populations (n=15) during adaptation to SPF$_{str+Ec}$ mice is reproduced for comparison [24]. (**B**) Frequency (± 2 S.E.M.) of the *lrp*-mutant phenotype (sum of complete and partial loss-function) over time, across individual populations in Gnoto$_{Ec}$ (colored lines, n=10). Inset table denotes the frequency of *lrp*-mutant phenotype at the end of each lineage and the corresponding day. See also Fig. S1 and S2

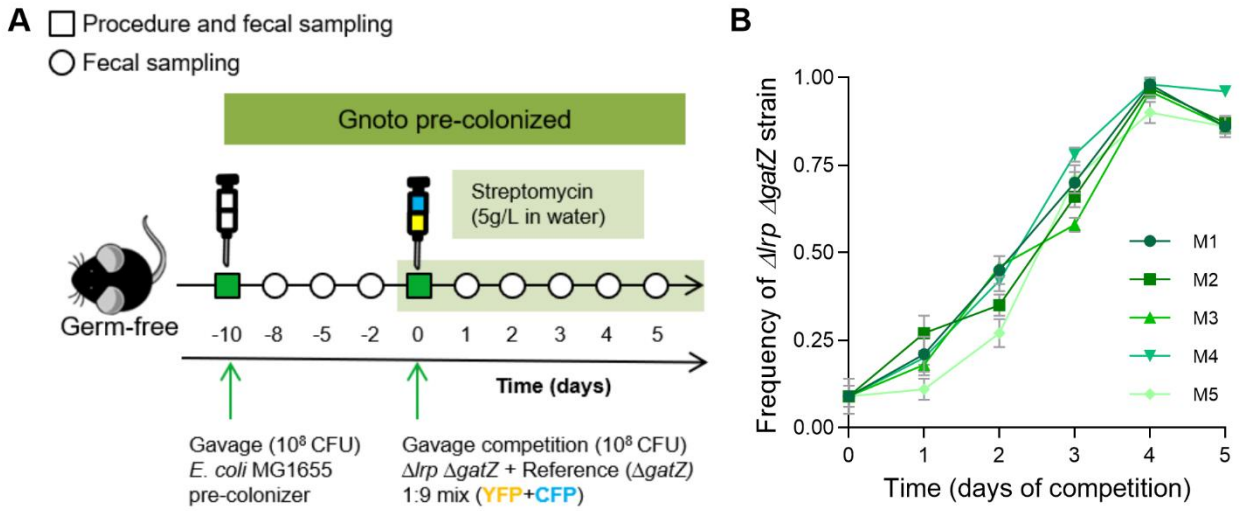

Fig. 3 - *lrp* mutant has a competitive fitness advantage in pre-colonized Gnoto$_{Ec}$ mice. (**A**) Design of competition experiment: germ-free mice were pre-colonized with ancestral WT streptomycin-sensitive strain of *E. coli* (MFP121) for 10 days, after which mice were administered streptomycin in drinking water (green shade). After 4 hours of streptomycin treatment the competitor strains composed of a mixture of Δ*lrp* Δ*gatZ* and Δ*gatZ* streptomycin resistance strains were introduced at 1:9 ratio by oral gavage. (**B**) Frequency (± 2 S.E.M.) of Δ*lrp* Δ*gatZ* mutant during competition experiment (n=5) in (A). See also Fig. S3.

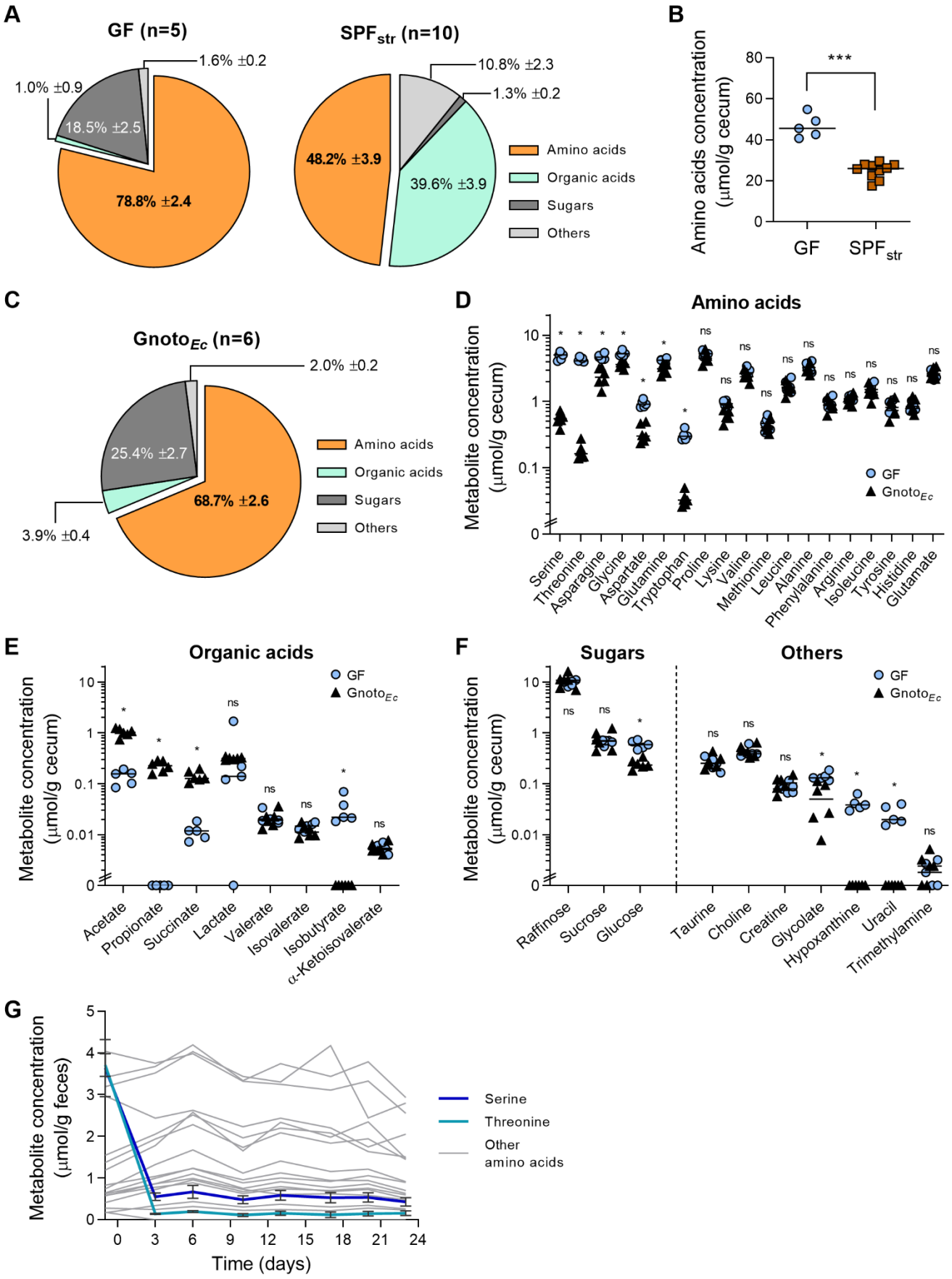
A
GF (n=5)
1.6% ±0.2
1.0% ±0.9
18.5% ±2.5
78.8% ±2.4
SPF$_{str}$ (n=10)
10.8% ±2.3
1.3% ±0.2
48.2% ±3.9
39.6% ±3.9
Amino acids
Organic acids
Sugars
Others
B
Amino acids concentration (µmol/g cecum)
***
GF
SPF$_{str}$
C
Gnoto$_{Ec}$ (n=6)
2.0% ±0.2
25.4% ±2.7
68.7% ±2.6
3.9% ±0.4
Amino acids
Organic acids
Sugars
Others
D
Amino acids
Metabolite concentration (µmol/g cecum)
GF
Gnoto$_{Ec}$
Serine
Threonine
Asparagine
Glycine
Aspartate
Glutamine
Tryptophan
Proline
Lysine
Valine
Methionine
Leucine
Alanine
Phenylalanine
Arginine
Isoleucine
Tyrosine
Histidine
Glutamate
E
Organic acids
Metabolite concentration (µmol/g cecum)
GF
Gnoto$_{Ec}$
Acetate
Propionate
Succinate
Lactate
Valerate
Isovalerate
Isobutyrate
α-Ketoisovalerate
F
Sugars
Others
Metabolite concentration (µmol/g cecum)
GF
Gnoto$_{Ec}$
Raffinose
Sucrose
Glucose
Taurine
Choline
Creatine
Glycolate
Hypoxanthine
Uracil
Trimethylamine
G
Metabolite concentration (µmol/g feces)
Time (days)
Serine
Threonine
Other amino acids

Fig. 4 - Amino acids are abundant resources that are targeted by *E. coli* during adaptation in Gnoto$_{Ec}$ mice. Concentration of small metabolites from cecal contents (**A**-**F**) or fecal samples (**G**) was determined by ^{1}H-NMR and grouped by chemical attributes. (**A**) Pie charts with relative concentrations (± 2 S.E.M.) of metabolite groups in Germ-free (GF, n=5) and streptomycin-treated SPF mice (SPF$_{str}$, day 28, n=10). (**B**) Total concentration of amino acids is higher in GF than SPF$_{str}$ (Mann-Whitney *U*-test, P<0.001. (**C**) Pie chart with relative concentrations (± 2 S.E.M.) of metabolite groups in Gnoto$_{Ec}$ mice (day 31, n=6). (**D**-**F**) Absolute concentrations of all metabolites (with medians depicted as lines) detected in GF (blue circles) and Gnoto$_{Ec}$ (black triangles). Amino acids (**D**) are ordered, from left to right, according to concentration difference between GF and Gnoto$_{Ec}$. Metabolites whose concentration significantly differs between GF and Gnoto$_{Ec}$ display the level of significance: $P \leq 0.05$ (*), $P \leq 0.01$ (**) or $P > 0.05$, non-significant (ns), after *post-hoc* Mann-Whitney *U*-test with Holm's correction for multiple comparisons. Lines depict medians. (**G**) Time course of concentrations of all amino acids detected during Gnoto$_{Ec}$ colonization (n=3) highlighting concentration (± 2 S.E.M.) dynamics of serine and threonine. See also Fig. S4 and S5 and Table S3.

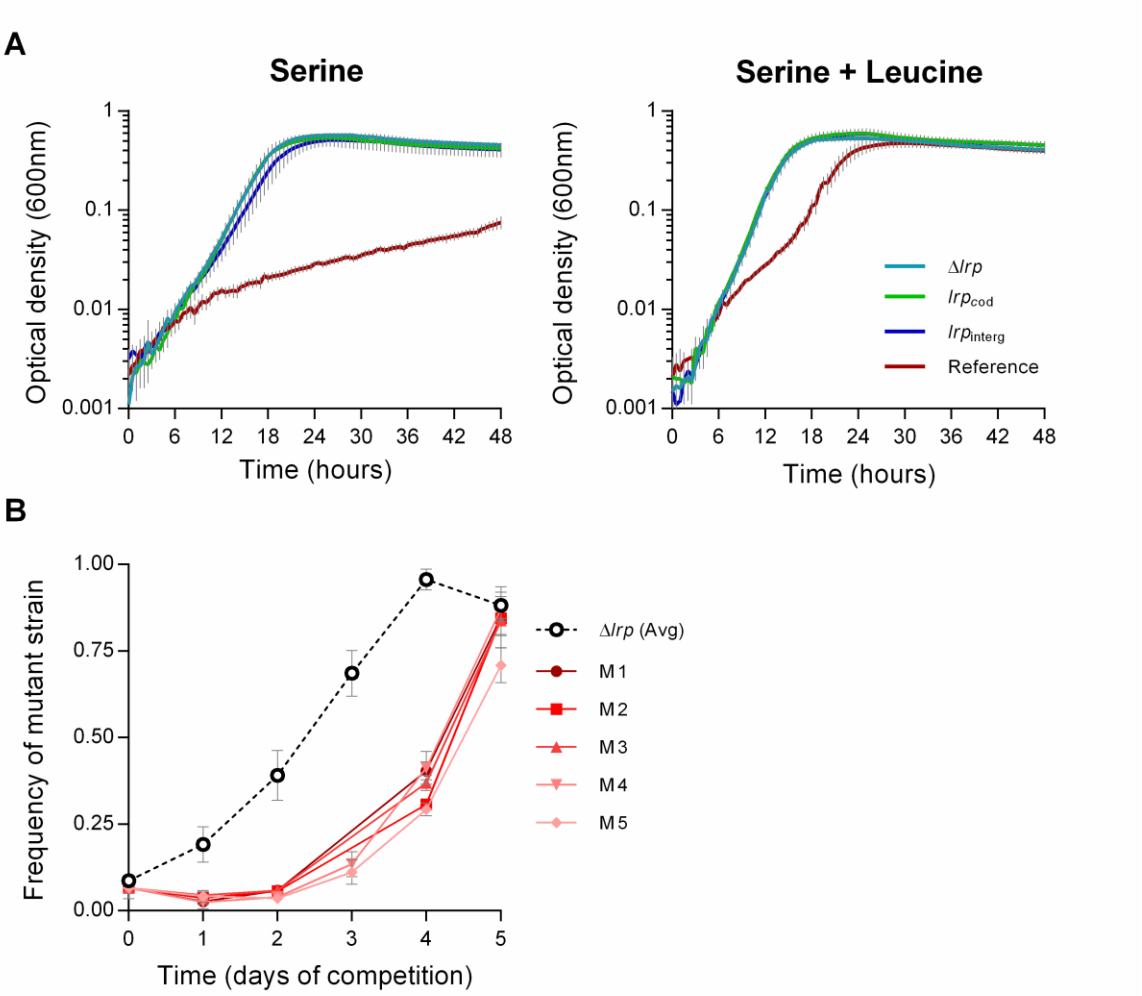

Fig. 5 - Serine metabolism provides fitness advantage to *lrp* mutants. (**A**) Growth curves of Δ*lrp* Δ*gatZ*, lrp_{cod} Δ*gatZ*, lrp_{interg} Δ*gatZ* and Reference (Δ*gatZ*) in M9 medium with serine 0.5% as carbon source, in the absence (left panel) or presence (right panel) of leucine 10mM (n=9 from 3 independent experiments). Error bars represent ± 2 S.E.M. Growth parameters can be found in Table S4. (**B**) Effect of serine utilization on fitness advantage of *lrp* mutant *in vivo*. Frequency (± 2 S.E.M.) of Δ*lrp* Δ*sdaA* Δ*gatZ* mutant against Δ*gatZ* during competition in GF mice pre-colonized with ancestral strain (as in Fig. 3A) shown in red (n=5). For comparison, the average frequency (± 2 S.E.M.) of Δ*lrp* Δ*gatZ* mutant against Δ*gatZ* during *in vivo* competition (from Fig. 3B) is depicted in a dashed line. See also Fig. S6 and Table S4.

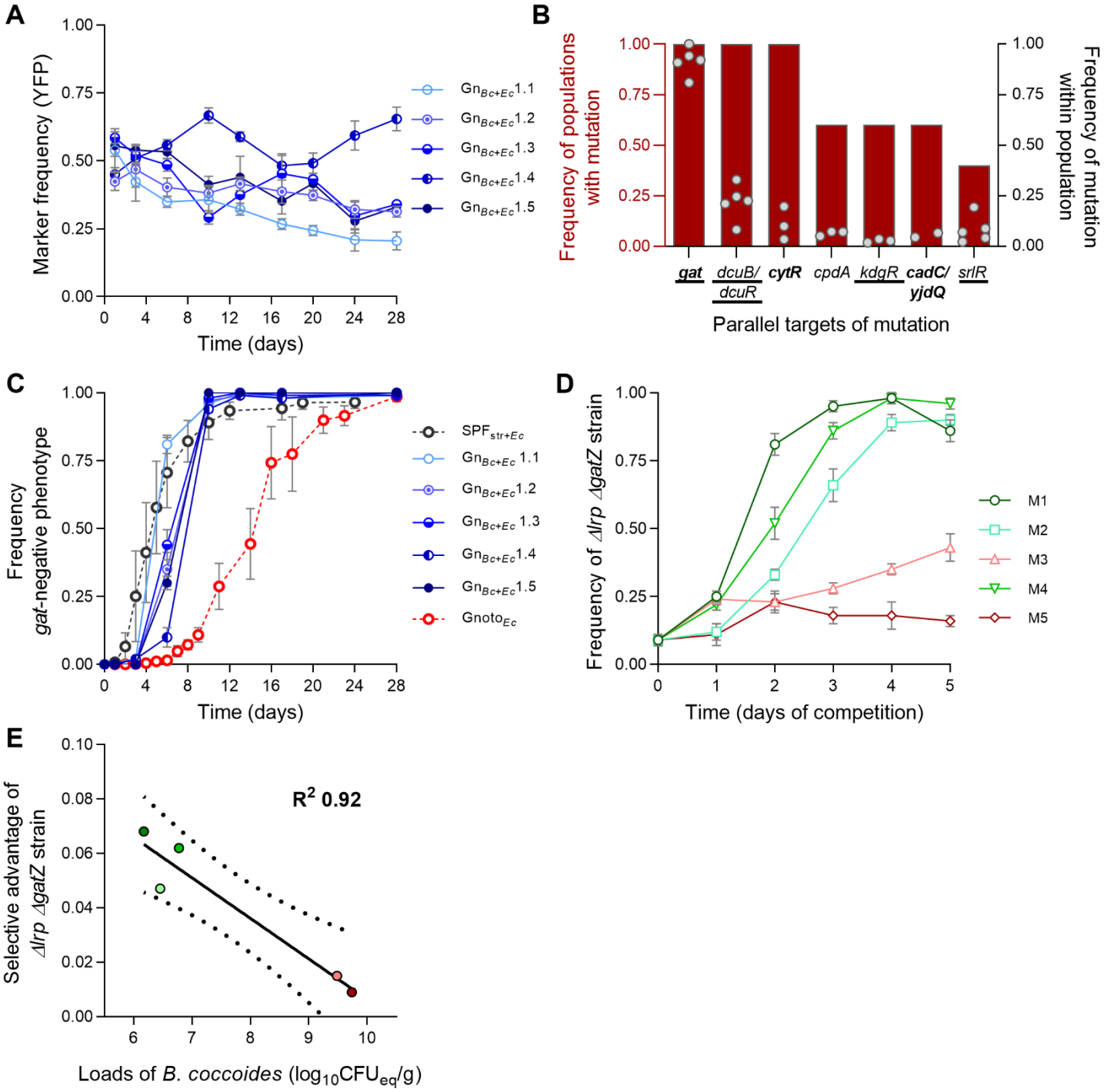

Fig. 6 - *B. coccoides* abolishes *lrp* mutant emergence and fitness. (**A**) Dynamics of neutral fluorescent marker (± 2 S.E.M.) during *E. coli* adaptation in Gnoto$_{Bc+Ec}$ mice (n=5). (**B**) Parallel targets of mutation identified by WGS in samples of populations collected at day 28. The height of the bars denotes the frequency of populations (n=5) where a given mutation was present (left axis) and circles represent the frequency each mutated locus reached within the population (right axis). Mutations common to SPF$_{str+Ec}$ are underlined and those common to

Gnoto$_{Ec}$ highlighted in bold. (**C**) Frequency (± 2 S.E.M.) of the *gat*-negative phenotype over time across individual populations (blue lines) in Gnoto$_{Bc+Ec}$ (n=5). Average (± 2 S.E.M.) *gat* dynamics in Gnoto$_{Ec}$ (red dashed line) and SPF$_{str+Ec}$ (black dashed line) are also depicted for comparison (data from Fig. 2). (**D**) Frequency (± 2 S.E.M.) of Δ*lrp* Δ*gatZ* strain during competition against Δ*gatZ* in GF mice after a period of pre-colonization (as in Fig. 3) but in the presence of *B. coccoides* (n= 5). (**E**) Correlation between Δ*lrp* Δ*gatZ* mutant fitness and average bacterial loads of *B. coccoides* during *in vivo* competition in (D). See also Fig. S7 and Tables S2 and S5.

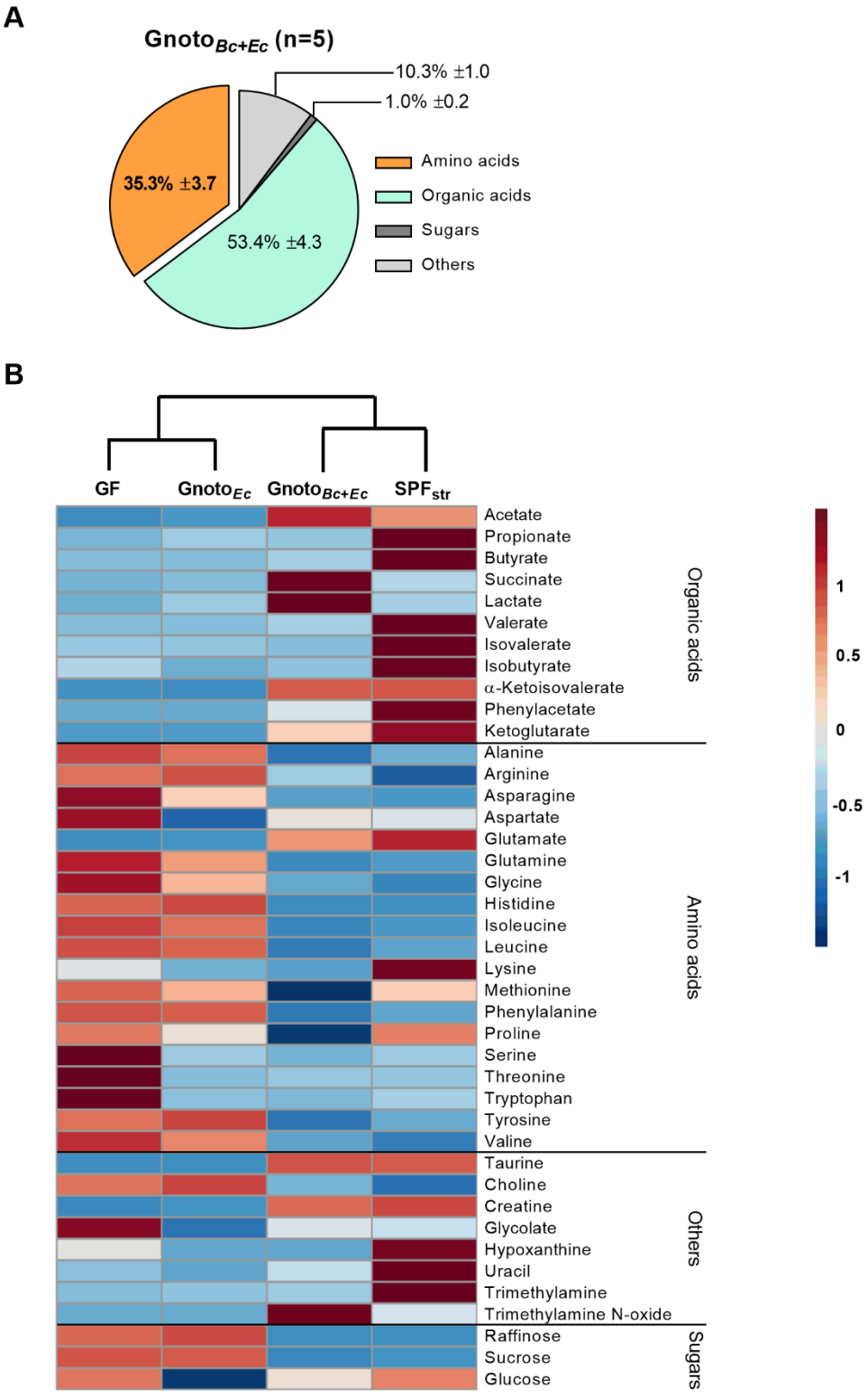

Fig. 7 - Metabolites in Gnoto$_{Bc+Ec}$ mice are more similar to the metabolites in SPF$_{str}$ than in Gnoto$_{Ec}$. (**A**) Pie charts of relative concentrations (± 2 S.E.M) of metabolite groups from cecal contents at the end of the Gnoto$_{Bc+Ec}$ adaptation experiment (day 28, n=5). (**B**) Heatmap of

median concentrations of all metabolites in Gnoto$_{Bc+Ec}$ cecal contents, for comparison with GF, Gnoto$_{Ec}$ and SPF$_{str}$. Colors represent the standard scores of absolute concentrations following group clustering with Euclidean distances (Ward's method). See also Fig. S4 and S7 and Table S3.

Supplemental Information

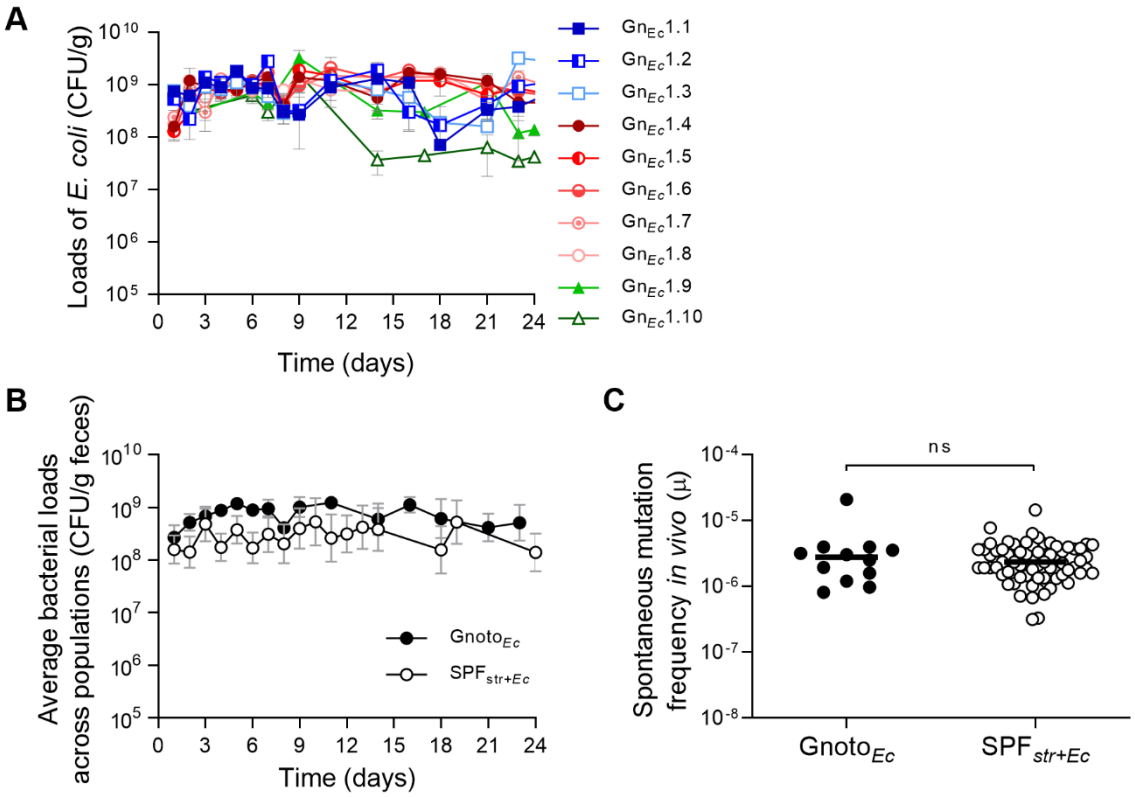

Figure S1 - Loads of *E. coli* during adaptation to the gut of Gnoto$_{Ec}$ mice and comparison of Eco-evolutionary parameters between Gnoto$_{Ec}$ and SPF$_{str+Ec}$. (**A**) Loads of *E. coli* (± 2 S.E.M.) during adaptation to the Gnoto$_{Ec}$ mouse gut in individually-housed mice monocolonized with a mixture of 1:1 ratio of DM08 and DM09 (n=10) as described in Fig. 1A. Experiment was performed in 3 independent blocks (Gn$_{Ec}$ 1-3, 4-8, 9-10). (**B**) Average loads of *E. coli* (± 2 S.E.M.) in fecal samples collected at indicated time points as a proxy for population size. *E. coli* loads are higher is Gnoto$_{Ec}$ mouse compared to SPF$_{str+Ec}$ [S1] (host type, ANOVA $\chi 21=4.4$, $P=0.036$)). (**C**) Mutation frequency for furazolidone resistance (as indicator of mutation rate) *in vivo* is similar between Gnoto$_{Ec}$ and SPF$_{str+Ec}$ [S2] (Mann-Whitney U test, $P=0.9$). Related to Fig. 1 and 2.

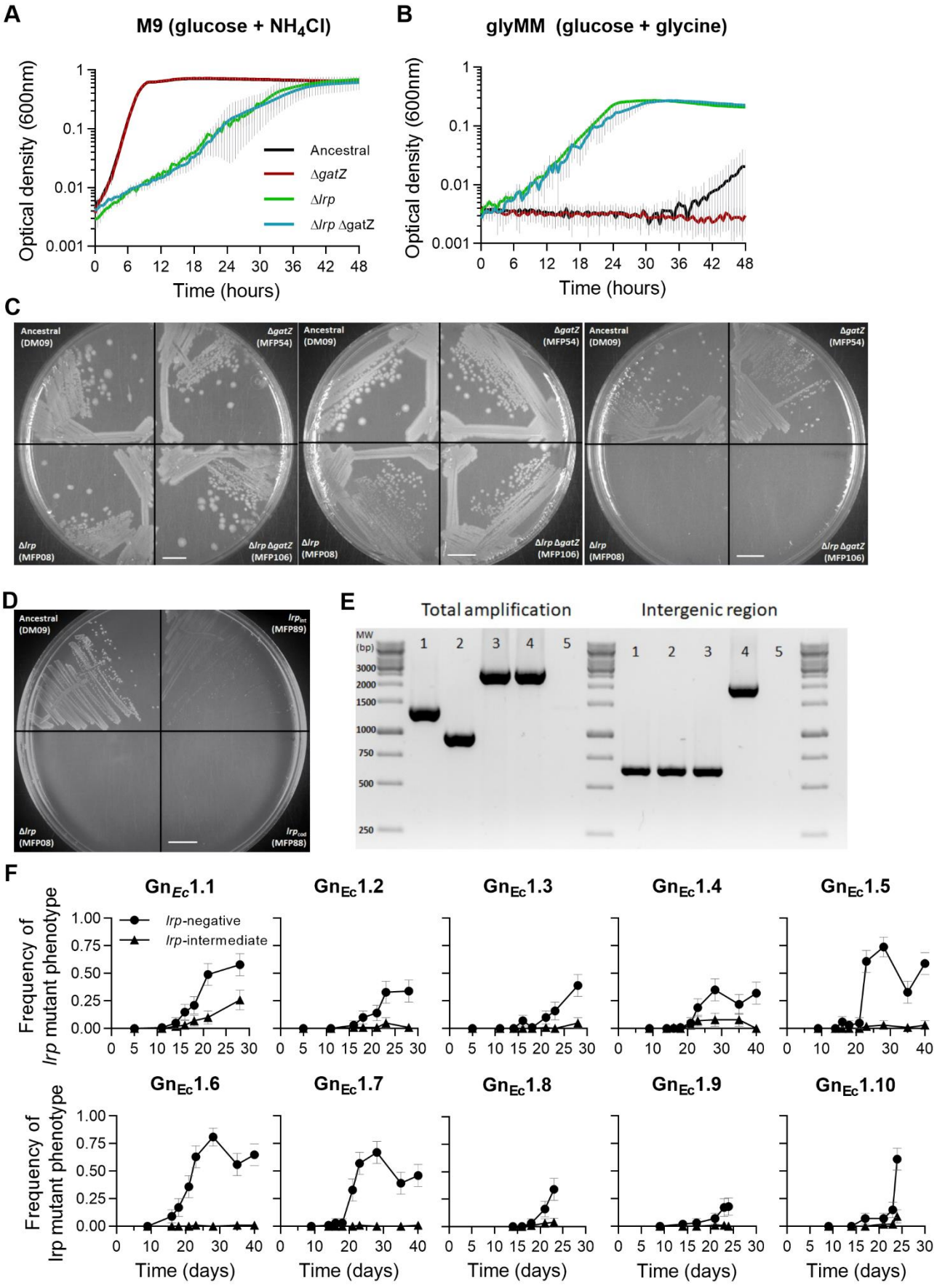
A
M9 (glucose + NH4Cl)
Optical density (600nm)
Time (hours)
Ancestral
ΔgatZ
Δlrp
Δlrp ΔgatZ
B
glyMM (glucose + glycine)
C
Ancestral (DM09)
ΔgatZ (MFP54)
Δlrp (MFP08)
Δlrp ΔgatZ (MFP106)
D
lrpint (MFP89)
lrpcod (MFP88)
E
Total amplification
Intergenic region
MW (bp)
3000
2000
1500
1000
750
500
250
F
Gn_Ec1.1
Gn_Ec1.2
Gn_Ec1.3
Gn_Ec1.4
Gn_Ec1.5
Gn_Ec1.6
Gn_Ec1.7
Gn_Ec1.8
Gn_Ec1.9
Gn_Ec1.10
lrp-negative
lrp-intermediate
Frequency of lrp mutant phenotype
Frequency of lrp mutant phenotype
Time (days)

Figure S2 - Ancestral and Δ*lrp* mutant can be distinguished phenotypically through growth in glyMM medium, independently of the *gat* mutation. (**A**) Growth curves of ancestral (DM09), Δ*gatZ* mutant (MFP54), Δ*lrp* mutant (MFP008) and Δ*lrp* Δ*gatZ* mutant (MFP106) in M9 minimal medium with 0.5% glucose, at 42°C, show impaired growth of *lrp* mutants when compared to the ancestral and Δ*gatZ* strains. (**B**) Growth defect of Δ*lrp* mutants is exacerbated in glyMM, (where ammonium is replaced with glycine as nitrogen source, see Methods), and thus this medium was chosen for the phenotypic test to score for *lrp*-negative phenotype. Growth curves in (A) and (B) are averages (± 2 S.E.M.) from 9 replicates (3 independent experiments). (**C**) Growth in solid medium (at 42°C) of the same strains described in (A) on LB (left panel), M9 minimal medium with glucose (middle panel) or glyMM (right panel), show that glyMM agar plates can be used to score for *lrp*-negative phenotype. (**D**) Growth of ancestral (DM09), Δ*lrp* (MFP08), lrp_{interg} (MFP89) and coding lrp_{cod} (MFP88) strains in glyMM. lrp_{cod} evolved clone is phenotypically similar to Δ*lrp* mutant, while lrp_{interg} evolved clone has an intermediate phenotype. (**E**) Targeted PCR for the *lrp* locus performed on the strains described in (D) with primers either for the complete *lrp* locus (left) or only for the intergenic region upstream of the *lrp* starting codon (right). Clones with an insertion of an IS element (evolved clones, lanes 3 and 4) displayed a band with a larger size than that of the ancestral (lane 1), while for Δ*lrp* (lane 2), the band was smaller. Amplification of the intergenic region upstream of *lrp* open reading frame allows for identification of clones carrying intergenic IS insertions as these display a larger band size (lane 4) compared to the ancestral (lane 1) or the other mutants (lanes 2 and 3). Lane 5 is the no-DNA negative control. (**F**) Individual plots of frequencies of *lrp*-negative (circles) and *lrp*-intermediate (triangles) mutant phenotypes (± 2 S.E.M.) for all $Gnoto_{Ec}$ populations (as shown in Fig. 2). Experiments performed separately as Gn_{Ec} 1.1-1.3, Gn_{Ec} 1.4-1.8, Gn_{Ec} 1.9-1.10 (mouse 8 died after day 23). These results show that both *lrp*-negative and *lrp*-intermediate phenotypes emerge in *E. coli* populations evolving in $Gnoto_{Ec}$ mice. Related to Fig. 2.

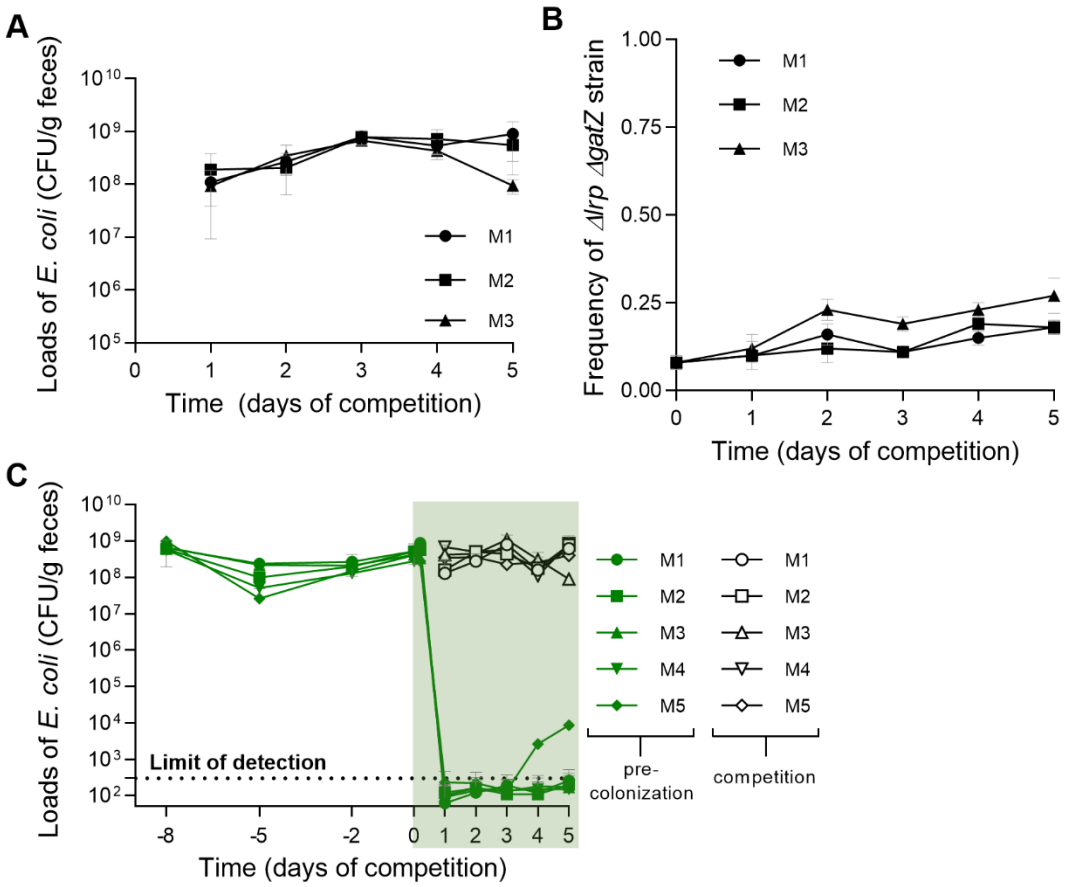

Figure S3 - *lrp* mutant is not advantageous in GF mice, during competitive fitness experiment without pre-colonization. Total *E. coli* loads (± 2 S.E.M.) (**A**) and frequency (± 2 S.E.M.) of Δ*lrp* Δ*gatZ* (**B**) during *in vivo* competition of Δ*lrp* Δ*gatZ* (MFP106) against Δ*gatZ* (MFP54) introduced at 1:9 ratio by oral gavage in GF mice (n=3, M1-M3). (**C**) Bacterial loads (± 2 S.E.M.) from the competition experiment of Δ*lrp* Δ*gatZ* mutant against Δ*gatZ* in Gnoto$_{Ec}$ mice (n=5, M1-M5) pre-colonized with an ancestral streptomycin sensitive strain (MFP121, green full symbols). Competitor strains (depicted in open black symbols) were introduced at 1:9 ratio by oral gavage in GF mice four hours after the start of streptomycin treatment. Shaded area denotes period of streptomycin treatment. Related to Fig. 3.

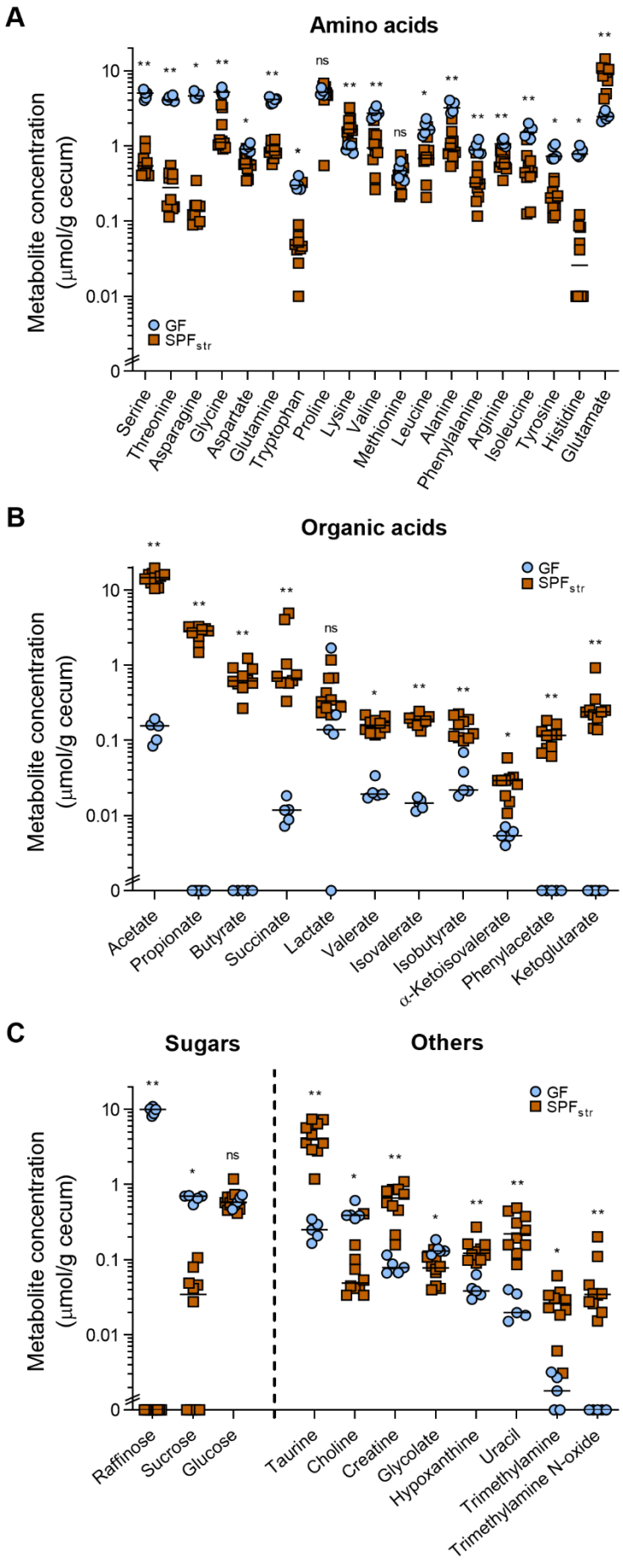
A
Amino acids
Metabolite concentration (µmol/g cecum)
GF
SPF$_{str}$
Serine
Threonine
Asparagine
Glycine
Aspartate
Glutamine
Tryptophan
Proline
Lysine
Valine
Methionine
Leucine
Alanine
Phenylalanine
Arginine
Isoleucine
Tyrosine
Histidine
Glutamate
B
Organic acids
Acetate
Propionate
Butyrate
Succinate
Lactate
Valerate
Isovalerate
Isobutyrate
α-Ketoisovalerate
Phenylacetate
Ketoglutarate
C
Sugars
Others
Raffinose
Sucrose
Glucose
Taurine
Choline
Creatine
Glycolate
Hypoxanthine
Uracil
Trimethylamine
Trimethylamine N-oxide

Figure S4 - GF and SPF_{str} ceca exhibit distinct metabolomic profiles, in particular in respect to amino acids concentration. (**A**-**C**) Concentration of small metabolites from cecal contents, determined by ^{1}H-NMR and grouped by chemical attributes. Absolute concentrations (with medians shown as lines) of all metabolites detected in GF (blue circles, n=5) and SPF_{str} (brown squares, n=10) is shown for amino acids (**A**), organic acids (**B**) and sugars and other metabolites (**C**). Metabolites whose concentration significantly differs between GF and SPF_{str} display the level of significance: $P \leq 0.05$ (*), $P \leq 0.01$ (**) or $P > 0.05$, non-significant (ns), after *post-hoc* Mann-Whitney *U*-test with Holm's correction for multiple comparisons. Related to Fig. 4.

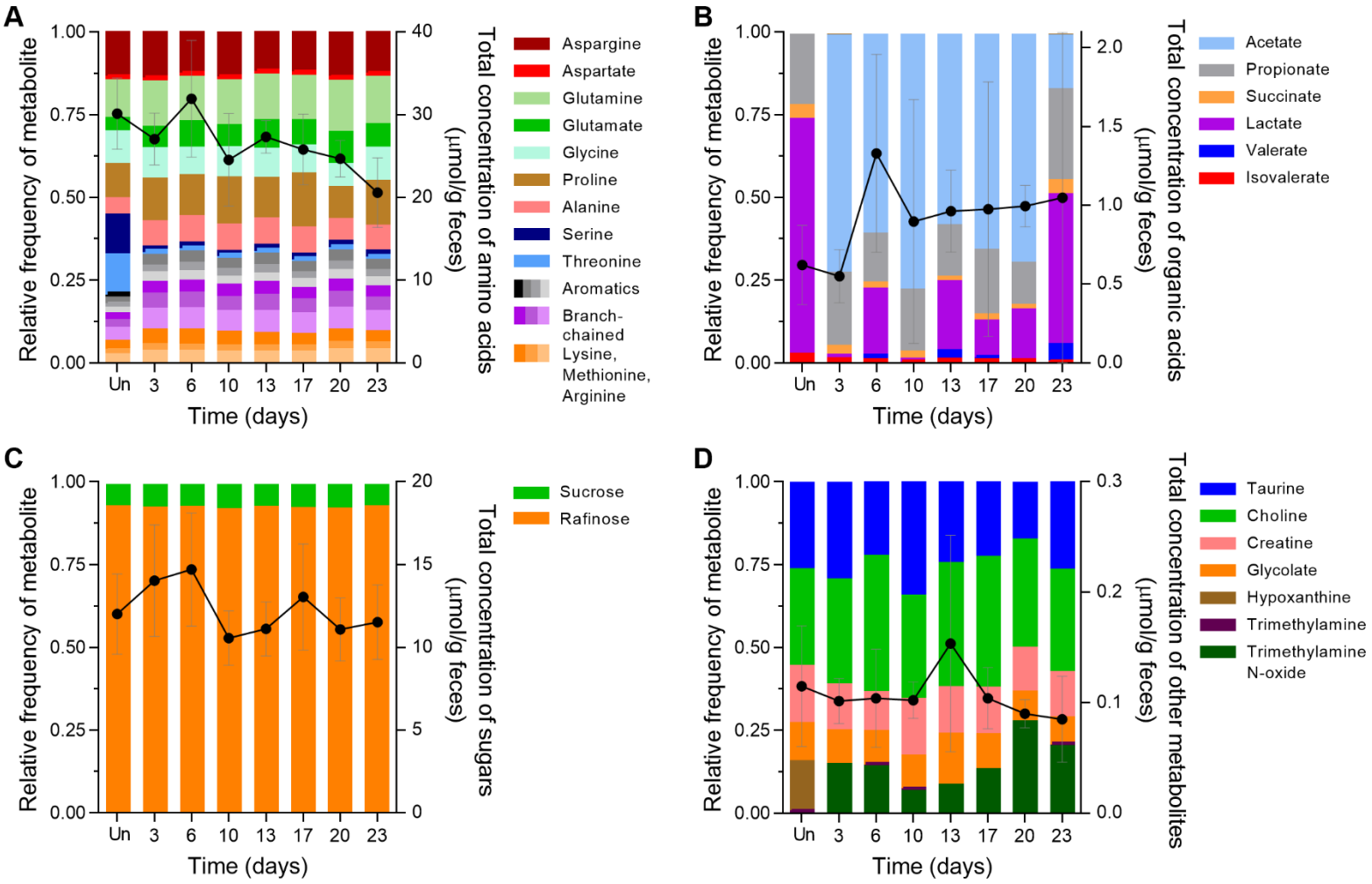

Figure S5 - Metabolite dynamics show rapid depletion of serine and threonine during *E. coli* adaptation in Gnoto$_{Ec}$ mice. Metabolites were determined by ^{1}H-NMR in fecal samples collected during Gnoto$_{Ec}$ colonization (n=3). Bars show relative concentration of each metabolite over time (left axis) and circles depict the average total concentration (± 2 S.E.M.) of amino acids (**A**), organic acids (**B**), sugars (**C**) and other metabolites (**D**) over time (right axis). Uncolonized samples (Un) were collected the day prior to gavage with *E. coli*. Related to Fig. 4.

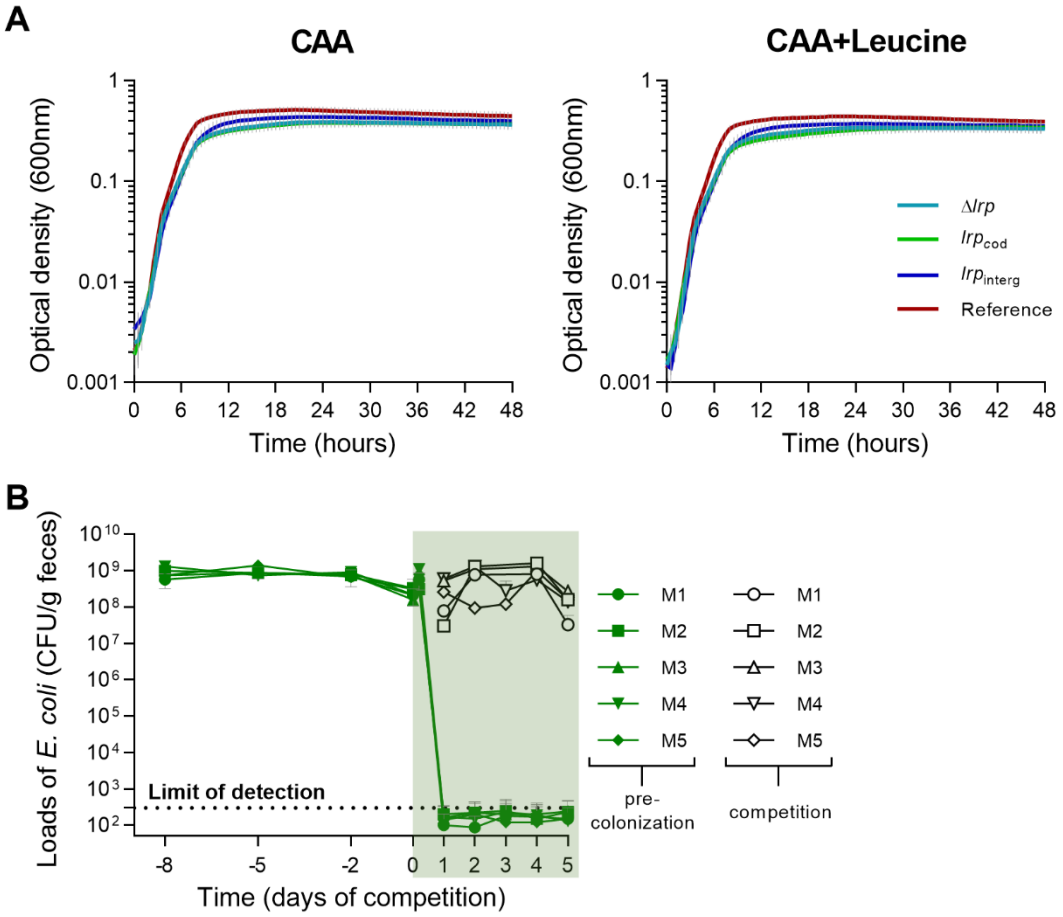

Figure S6 - *lrp* mutation confers a specific growth advantage. (A) Growth in casaminoacids show similar profiles independently of *lrp*. Growth curves of Δ*lrp* Δ*gatZ* (MFP106), lrp_{cod} Δ*gatZ* (MFP102), lrp_{interg} Δ*gatZ* (MFP104) and Reference (Δ*gatZ*, MFP54) in M9 medium with 0.5% casaminoacids as carbon source, in the absence (left panel) or presence (right panel) of leucine 10Mm (n=9 from 3 independent experiments). Estimated growth parameters in Table S4. Error bars represent ± 2 S.E.M. **(B)** Bacterial loads (± 2 S.E.M.) from the competition experiment of Δ*lrp ΔsdaA* Δ*gatZ* mutant (MPF125) against Δ*gatZ* (MFP54) in $Gnoto_{Ec}$ mice (n=5, M1-M5) pre-colonized with an ancestral (MFP121) streptomycin sensitive strain (green full symbols). Competitor strains (depicted in open black symbols) were introduced at 1:9 ratio by oral gavage in GF mice four hours after the start of streptomycin treatment. Shaded area denotes period of streptomycin treatment. Related to Fig. 5.

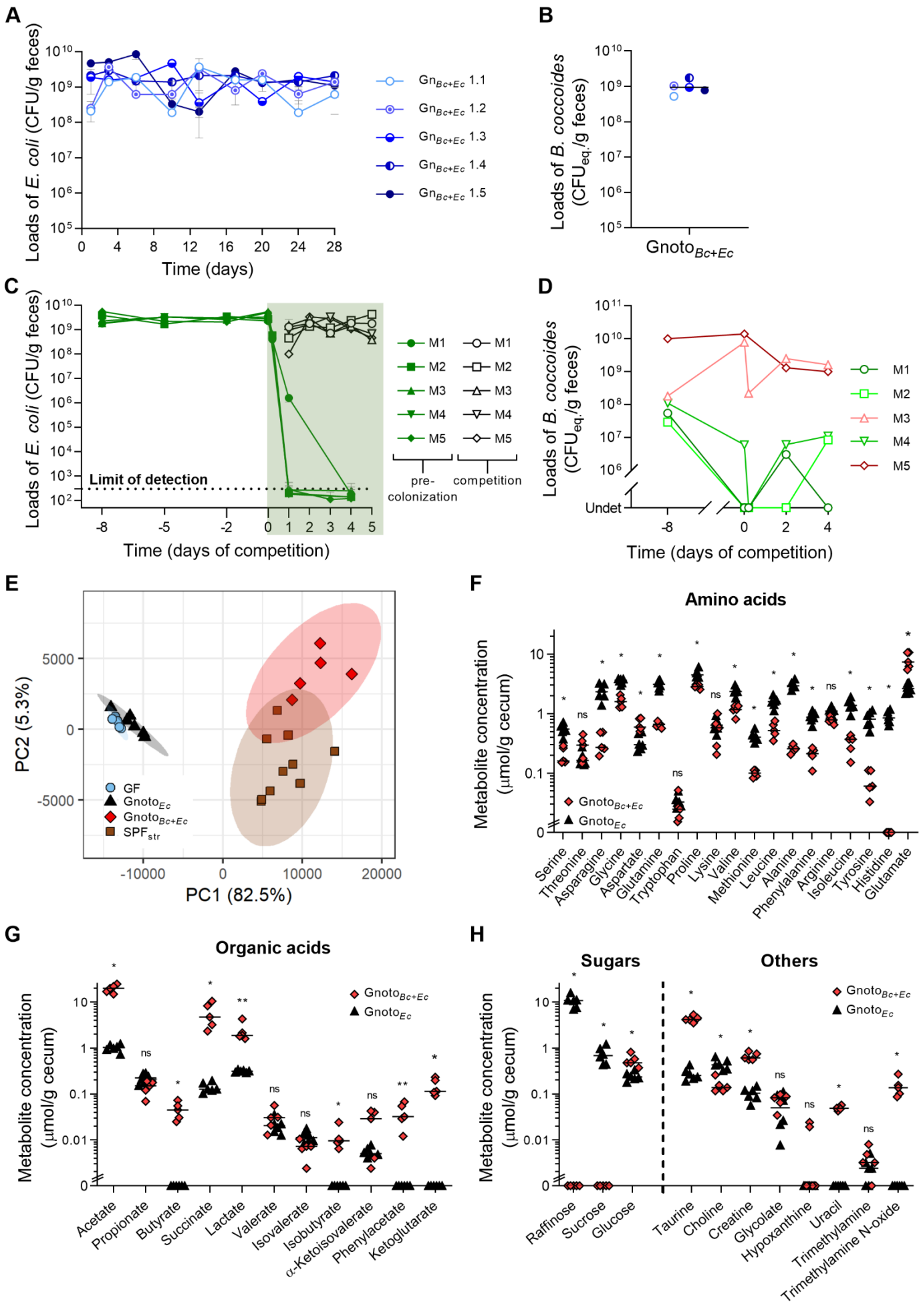
A
Loads of E. coli (CFU/g feces)
Time (days)
Gn Bc+Ec 1.1
Gn Bc+Ec 1.2
Gn Bc+Ec 1.3
Gn Bc+Ec 1.4
Gn Bc+Ec 1.5
B
Loads of B. coccoides (CFUeq./g feces)
Gnoto Bc+Ec
C
Loads of E. coli (CFU/g feces)
Limit of detection
Time (days of competition)
M1
M2
M3
M4
M5
pre-colonization
competition
D
Loads of B. coccoides (CFUeq./g feces)
Undet
Time (days of competition)
E
PC2 (5.3%)
PC1 (82.5%)
GF
Gnoto Ec
Gnoto Bc+Ec
SPF str
F
Amino acids
Metabolite concentration (µmol/g cecum)
Gnoto Bc+Ec
Gnoto Ec
Serine
Threonine
Asparagine
Glycine
Aspartate
Glutamine
Tryptophan
Proline
Lysine
Valine
Methionine
Leucine
Alanine
Phenylalanine
Arginine
Isoleucine
Tyrosine
Histidine
Glutamate
G
Organic acids
Metabolite concentration (µmol/g cecum)
Acetate
Propionate
Butyrate
Succinate
Lactate
Valerate
Isovalerate
Isobutyrate
α-Ketoisovalerate
Phenylacetate
Ketoglutarate
H
Sugars
Others
Raffinose
Sucrose
Glucose
Taurine
Choline
Creatine
Glycolate
Hypoxanthine
Uracil
Trimethylamine
Trimethylamine N-oxide

Figure S7 - Evolution and metabolic environment of *E. coli* in Gnoto$_{Bc+Ec}$ mice differs from Gnoto$_{Ec}$. (**A**) Loads of *E. coli* (± 2 S.E.M.) during the adaptation experiment in the presence of *B. coccoides* (n=5). (**B**) Loads of *B. coccoides* at the end of the evolution experiment in (A). (**C**) Competition experiment of Δ*lrp* Δ*gatZ* (MFP106) mutant against Δ*gatZ* (MFP54) in Gnoto$_{Ec}$ mice (n=5, M1-M5) pre-colonized with an ancestral (MFP121) streptomycin sensitive strain (green full symbols). Competitor strains (depicted in open black symbols) were introduced at 1:9 ratio by oral gavage in GF mice four hours after the start of streptomycin treatment. Bacterial loads (± 2 S.E.M.) during competition are shown, with shaded area denoting period of streptomycin treatment. (**D**) Loads of *B. coccoides* during the competition experiment shown in (C). (**E**) Score plot for PC1 and PC2 from principal component analysis (PCA) of all metabolites found in GF (blue circles), Gnoto$_{Ec}$ (black triangles), Gnoto$_{EcBc}$ (red diamonds) and SPF$_{str}$ (brown squares). Colored circles represent 95 % confidence intervals. (**F**-**H**) Concentration of small metabolites from cecal contents, determined by ^{1}H-NMR and grouped by chemical attributes. Absolute concentrations (with medians shown as lines) of all metabolites detected in Gnoto$_{Ec}$ (black triangles, n=6) and Gnoto$_{Bc+Ec}$ (red diamonds, n=5) is shown for amino acids (**F**), organic acids (**G**) and sugars and other metabolites (**H**). Metabolites whose concentration significantly differs between Gnoto$_{Ec}$ and Gnoto$_{Bc+Ec}$ display the level of significance: $P \leq 0.05$ (*), $P \leq 0.01$ (**) or $P > 0.05$, non-significant (ns), after *post-hoc* Mann-Whitney *U*-test with Holm's correction for multiple comparisons. Related to Fig. 6 and 7.

Table S1 - Parallel mutations segregating in *E. coli* populations evolving in the absence of interspecies competition. Loci were identified by WGS of populations samples collected from Gnoto$_{Ec}$ mice after 23 days of evolution, and parallel mutations defined as evets found in more than one evolving population at a frequency greater than 5%. Functional annotation is shown for each locus, with processes related with amino acid metabolism highlighted in bold. Number of Gnoto$_{Ec}$ populations (n=10) where each mutation was found is shown, as well as the number of Gnoto$_{str+Ec}$ populations (n=3) where Gnoto$_{Ec}$ parallel mutations were also detected. Related to Fig. 1.

Locus	Functions	Gnoto$_{Ec}$ day 23	Gnoto$_{str+Ec}$ day 23
		# mice (n=10)	# mice (n=3)
gat operon	Metabolism of galactitol	10	3
trxB / ***lrp***[a]	**Lrp is a global regulator of Amino acids metabolism**	**10**	**3**
frlR	**Fructosamine metabolism**	**8**	**1**
rssB	Regulator of RpoS	5	1
lrhA **/ *alaA***[b]	Regulation of flagellation, motility and chemotaxis / **Glutamate—pyruvate aminotransferase (alanine biosynthesis)**	**4**	**1**
glpR	Glycerol metabolism	4	1
uspA / ***dtpB***[c]	**Peptide transporter / uptake**	**4**	**2**
yjjM	Galactonate metabolism	4	
tdcA / tdcR	**Threonine and serine metabolism**	**3**	**1**
cadC / *yjdQ*[d]	**Lysine metabolism**	**3**	**1**
yebK	Regulation of carbon metabolism during nutritional shifts	2	

[a]mutation in the intergenic region of these two genes and in the coding region of *lrp* were identified.
[b]mutation in this intergenic region can potentially affect the expression of these two genes, one mutation in coding region of the *lrhA* was identified.
[c]mutation in the intergenic region upstream of *dtpB* and downstream of *uspA*.
[d]mutation in the intergenic region upstream of *cadC* and downstream of the *yjdQ gene*.

Table S2 - Mutations found in *E. coli* populations after adaptation to the gut of gnotobiotic mice. WGS analysis of *E. coli* population samples evolved in: Gnoto$_{Ec}$ for 23 (n=10) or 40 days (n=4), Gnoto$_{str+Ec}$ for 23 days (n=3) and Gnoto$_{Bc+Ec}$ for 28 days (n=5). For intergenic mutations the two flanking genes are listed, otherwise the mutation occurred in the gene coding region. SNPs are represented by an arrow between the ancestral and the evolved nucleotide. The symbol Δ means a deletion event and a + symbol represents an insertion of the nucleotide that follows the symbol. The initials IS denote the abbreviation of insertion sequence element at the indicated position. del/dup indicates that either a deletion or a duplication of the indicated size occurred but it is not possible to distinguish between the two. * indicates that the mutation corresponds to a supported unassigned new junction whereas ¤ denotes an IS insertion where only one new junction was identified. Related to Fig. 1 and 6.

(For Table S2 see attached Excel file)

Table S3 - List of metabolites identified by ^{1}H-NMR in the cecum of GF (n=5), SPF$_{str+Ec}$ (n=10), Gnoto$_{Ec}$ (n=6) and Gnoto$_{Bc+Ec}$ (n=5) or in fecal samples of Gnoto$_{Ec}$ (n=3) mice. Concentration of metabolites in nmol per gram of cecal content or feces. SPF$_{str+Ec}$, Gnoto$_{Ec}$ and Gnoto$_{Bc+Ec}$ mice were colonized with ancestral *E. coli* strains for approximately one month prior to metabolomic analysis. Related to Fig. 4 and 7.

(For Table S3 See attached Excel file)

Table S4 - Serine metabolism provides fitness advantage to *lrp* mutants. Growth parameters of Δ*gatZ* (MFP54), *Δlrp ΔgatZ* (MFP106), lrp_{cod} *ΔgatZ* (MFP102) and lrp_{interg} Δ*gatZ* (MFP104) *E. coli* strains estimated from growth curves in M9 medium supplemented with 0.5% serine (Ser) or 0.5% casaminoacids (CAA) and in the absence or presence (+leu) of Leucine. Averages are shown with error (2se) represented within parenthesis. Strain parameters were compared relative to the reference strain (Δ*gatZ*) in columns Ser and CAA; and within strain, for Leucine effect, in columns Ser+leu and CAA+leu; (*) denotes significance after Mann-Whitney U-test with Holm correction for multiple comparisons ($P \leq 0.01$, except in CAA with $0.01 \leq P \leq 0.05$; ns – not significant). N=9 from 3 independent experiments. Related to Fig. 5.

	Growth rate (per hour)		Carrying capacity (OD600nm)		Growth rate (per hour)		Carrying capacity (OD600nm)	
Strain	Ser	Ser+leu	Ser	Ser+leu	CAA	CAA+leu	CAA	CAA+leu
Δ*gatZ*	0.15 (0.03)	0.34 (0.03)*	0.07 (0.01)	0.48 (0.04)*	1.12 (0.06)	1.20 (0.06) ns	0.51 (0.04)	0.44 (0.02) ns
Δ*lrp* Δ*gatZ*	0.36 (0.02)*	0.48 (0.03)*	0.57 (0.03)*	0.53 (0.04) ns	1.03 (0.09) ns	1.19 (0.09) ns	0.40 (0.03)*	0.35 (0.03) ns
lrp_{cod} Δ*gatZ*	0.38 (0.02)*	0.48 (0.01)*	0.54 (0.06)*	0.60 (0.06) ns	1.04 (0.10) ns	1.06 (0.06) ns	0.39 (0.03)*	0.35 (0.04) ns
lrp_{interg} Δ*gatZ*	0.32 (0.01)*	0.48 (0.01)*	0.52 (0.05)*	0.59 (0.06) ns	0.96 (0.07) ns	1.10 (0.08) ns	0.44 (0.05) ns	0.38 (0.03) ns

Table S5 - Parallel mutations segregating in *E. coli* populations evolving in the presence of *B. coccoides*. Loci were identified by WGS of populations samples collected from $Gnoto_{Bc+Ec}$ mice after 28 days of evolution, and parallel mutations defined with the same criteria as in Table S1. We also included mutations that were found in more than one mouse and were shared with SPF_{str+Ec} or $Gnoto_{Ec}$ regardless of frequency. Number of $Gnoto_{Bc+Ec}$ populations (n=5) where each mutation was found is shown, as well as presence/absence of these mutations in $Gnoto_{Ec}$ or SPF_{str+Ec}-evolved populations. Related to Fig. 6.

Locus	**Functions**	**$Gnoto_{Bc}$ mice day 28**	**Common to**	
		# mice (n=5)	**$Gnoto_{Ec}$**	**SPF_{str+Ec}**
gat operon	Metabolism of galactitol	5	Yes	Yes
dcuB / ducR[a]	Anaerobic respiration	5		Yes
cytR[b]	Regulator of utilization (deoxy)ribonucleosides	5	Yes[c]	
cpdA	cAMP phosphodiesterase	3		
kdgR	Regulator of Gluconate utilization	3		Yes
cadC[c]	Metabolism of Lysine	3	Yes	
srlR	Regulator of Sorbitol utilization	2		Yes

[a]mutation in the intergenic region upstream of the *dcuB fumB operon* were identified.
[b]*cytR* was identified as a parallel mutation in Gnoto mice at day 40, Table S2.
[c]mutation in the intergenic region upstream of *cadC* and downstream of the *yjdQ gene*.

Table S6 - Strains, primers and plasmids used in the present study.

Element ID	Characteristics	Parent Strain	Origin
MG1655	Wild type K-12 *E. coli*	MG1655	[S3]
DM08	Δ*lacIZYA* YFP::amp^R str^R	MG1655	[S1]
DM09	Δ*lacIZYA* CFP::amp^R str^R	MG1655	[S1]
MFP002	Δ*lrp*::cm^R	MG1655	This work
MFP004	Δ*lacIZYA* YFP::amp^R str^R Δ*lrp*::cm^R	DM08	P1 transduction from MFP002 lysate
MFP008	Δ*lacIZYA* YFP::amp^R str^R Δ*lrp*	MFP004	Cm^R removal with pTL17
MFP048	Δ*lacIZYA* YFP::amp^R str^R Δ*gatZ*::kan^R	DM08	P1 transduction with P1 lysate from JW2082 [S4]
MFP050	Δ*lacIZYA* CFP::amp^R str^R Δ*gatZ*::kan^R	DM09	P1 transduction P1 lysate from JW2082 [S4]
MFP054	Δ*lacIZYA* CFP::amp^R str^R Δ*gatZ*	MFP050	Kan^R removal with pCP20
MFP090	Δ*lacIZYA* YFP::amp^R str^R Δ*gatZ* Δ*lrp*::cm^R	MFP052	P1 transduction with P1 lysate from MFP002
MFP106	Δ*lacIZYA* YFP::amp^R str^R Δ*gatZ* Δ*lrp*	MFP090	Cm^R cassette with pTL17
MFP088	Δ*lacIZYA* CFP::amp^R str^R *gatA*::IS5 lrp_{cod} (IS2)	DM09-evolved	Clone isolated from Gn1.9 at day 14 (evolved clone)
MFP089	Δ*lacIZYA* YFP::amp^R str^R *gatZ*::IS1 lrp_{interg} (IS2 intergenic region)	DM08-evolved	Clone isolated from Gn1.1 at day 16 (evolved clone)
MFP098	Δ*lacIZYA* CFP::amp^R str^R Δ*gatZ*::kan^R lrp_{cod} (IS2)	MFP088	P1 transduction with P1 lysate from JW2082 [S4]
MFP100	Δ*lacIZYA* YFP::amp^R str^R Δ*gatZ*::kan^R lrp_{interg} (IS2 intergenic region)	MFP089	P1 transduction with P1 lysate from JW2082 [S4]
MFP102	Δ*lacIZYA* CFP::amp^R str^R Δ*gatZ* lrp_{cod} (IS2)	MFP098	Kan^R removal with pCP20
MFP104	Δ*lacIZYA* YFP::amp^R str^R Δ*gatZ* lrp_{interg} (IS2 intergenic region)	MFP100	Kan^R removal with pCP20
NO11	Δ*lacIZYA*	MG1655	This work
MFP121	Δ*lacIZYA* Δ*galK*::kan^R	NO11	P1 transduction with P1 lysate from JW0740 [S4]
MFP125	Δ*lacIZYA* YFP::amp^R str^R Δ*gatZ* Δ*lrp* Δ*sdaA*::kan^R	MFP106	P1 transduction with P1 lysate from JW1803 [S4]
MFP133	Δ*lacIZYA* YFP::amp^R str^R Δ*gatZ* Δ*lrp* Δ*sdaA*	MFP125	kan^R removal with pCP20
Blautia coccoides	Wild type, mouse isolate	-	This work
Primer (a)	5’ ATCAGCACAGGTTGCAGGTT 3’	-	fw primer upstream of *lrp*
Primer (b)	3’ TGTCTCTCTGTATTCCTTCCCT 5’	-	rev primer adjacent to *lrp* ATG
Primer (c)	3’ GCGGCCGCTACTTAACTTTG 5’	-	rev primer downstream of *lrp*

pTL17	kan^R	-	[S5]
pCP20	cm^R	-	[S6]

Supplementary References

S1. Barroso-Batista, J., Sousa, A., Lourenço, M., Bergman, M.-L., Sobral, D., Demengeot, J., Xavier, K.B., and Gordo, I. (2014). The First Steps of Adaptation of Escherichia coli to the Gut Are Dominated by Soft Sweeps. PLoS Genet. *10*, e1004182.

S2. Barroso-Batista, J., Demengeot, J., and Gordo, I. (2015). Adaptive immunity increases the pace and predictability of evolutionary change in commensal gut bacteria. Nat Commun *6*, 8945.

S3. Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., *et al.* (1997). The complete genome sequence of Escherichia coli K-12. Science *277*, 1453–62.

S4. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. *2*, 2006.0008.

S5. Long, T., Tu, K.C., Wang, Y., Mehta, P., Ong, N.P., Bassler, B.L., and Wingreen, N.S. (2009). Quantifying the integration of quorum-sensing signals with single-cell resolution. PLoS Biol. *7*, e68.

S6. Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. U.S.A. *97*, 6640–5.