Supplementary Material

Supplementary Text

General Model for Resource Competition

We make the very simplifying assumption that the gut constitutes a homogenous environment where bacteria compete for resources. We will use the model of competition for resources by Tilman (Tilman 1982) following van Opheusden *et al*. (van Opheusden et al. 2015). We focus on the simplest model where we consider only two consumers and two resources in the mouse gut. The consumers are *gat*+ and *gat*- *E. coli*, whose densities inside the intestine are $Ec(t)$ and $En(t)$. The resources are galactitol and other general resources, whose abundances in the intestine are $G(t)$ and $R(t)$, respectively. We also assume that the system is homogeneous and that both strains of bacteria are removed at a similar rate (d). The rate at which bacteria grow (f) is assumed to solely dependent on the amount of resources available ($f(G,R)$). We further assume that, in the absence of bacteria, each resource achieves a maximum abundance S_G and S_R, for galactitol and other resources, respectively; and that the average residence time of either type of resource in the gut is $1/a$. Resources are also depleted by each of the consumer strains. Under these assumptions the rate of change of each strain and each resource is:

$$\frac{dEc}{dt} = f_c(G,R)Ec(t) - dEc(t)$$

$$\frac{dEn}{dt} = f_n(G,R)En(t) - dEn(t)$$

$$\frac{dG}{dt} = a\big(S_G - G(t)\big) - q_{Gc}\, f_c(G,R)Ec(t) - q_{Gn}f_n(G,R)En(t)$$

$$\frac{dR}{dt} = a(S_R - R(t)) - q_{Rc}\, f_c(G,R)Ec(t) - q_{Rn}\, f_n(G,R)En(t)$$

where q_{ji} are the conversion factors for the resource j used by consumer i.

Under specific assumptions for the values of q_{ji}, this system can be solved (van Opheusden et al. 2015) such that the abundances of the two strains will tend to:

$$Ec^{eq} = \frac{q_{Rn}a(S_G - G^{eq}) - q_{Gn}a(S_R - R^{eq})}{(q_{Gc}\ q_{Rn} - q_{Rc}\ q_{Gn})d}$$

$$En^{eq} = \frac{aq_{Rc}(G^{eq} - S_G) - aq_{Gc}(R^{eq} - S_R)}{(q_{Gc}\ q_{Rn} - q_{Rc}\ q_{Gn})d}$$

with the abundances of each resource dependent on the assumptions made for the relative growth rate of the strains, such that:

$$f_c(R^{eq}, G^{eq}) = f_n(R^{eq}, G^{eq}) = d$$

Model for colonization with *gat*+ and *gat*- strains

The *gat*+ strain of *E. coli*, used in the first colonization, has an IS insertion in the repressor of the galactitol operon (*gatR*), which leads to the constitutive expression of this operon. Thus one can make the simplifying assumption that *gat*+ is specialized on galactitol such that its growth is only dependent on $G(t)$; *i.e.* $f_c(G,R) = f_c(G)$. On the other hand the *gat*- strain, which evolved from the *gat*+ during the colonization, does not consume galactitol, thus its growth is only dependent on $R(t)$, $f_n(G,R) = f_n(R)$. Under these assumptions the system simplifies to:

$$\frac{dEc}{dt} = f_c(G)Ec(t) - dEc(t)$$

$$\frac{dEn}{dt} = f_n(R)En(t) - dEn(t)$$

$$\frac{dG}{dt} = a(S_G - G(t)) - q_{Gc}\, f_c(G)Ec(t)$$

$$\frac{dR}{dt} = a(S_R - R(t)) - q_{Rn}\, f_n(R)En(t)$$

To move forward and calculate the equilibrium abundances of *gat*+ (E_c^{eq}), *gat*- (E_n^{eq}) and each of the resources G^{eq} and R^{eq}, we will assume a Michaelis-Menten dynamic for growth such that $f_c(G) = \frac{G}{G+Gmc}$ and that $f_n(R) = \frac{R}{R+Rmn}$, where Gmc and Rmn are constants. We then can deduce the equilibrium abundances of bacteria and resources to be:

$$Ec^{eq} = \frac{a(S_G - G^{eq})}{d\, q_{Gc}};\ En^{eq} = \frac{a(S_R - R^{eq})}{d\, q_{Rn}};\ G^{eq} = \frac{d\, Gmc}{d-1};\ R^{eq} = \frac{d\, Rmn}{d-1}$$

This simplified model implies that when S_G increases, for example by supplementing the diet with galactitol, the loads of *gat*+ as well as the frequency of *gat*+ relative to *gat*- (which does not use this resource) should increase. Fig. S1a shows an example of the temporal dynamics for *gat*+ and *gat*- frequencies under this model. In Fig. S1b we show the prediction for the frequency of the strains under diet supplementation with increasing concentrations of galactitol.

Model for colonization with non-constitutive *gat*+ strain

We next asked whether a *gat*+ specialized strain, constitutively expressing the galactitol operon (E_c(t)), could invade a generalist which can consume both galactitol and the other resources. $f_p(G,R)$ is the growth function of such generalist, whose abundance is E_p(t). If we assume that:

$$f_p(G,R) = \frac{G}{G+Gmp} + \frac{R}{R+Rmp},$$

so that E_p uses both resources, then we have:

$$\frac{dEc}{dt} = \frac{G(t)}{G(t)+Gmc} Ec(t) - dEc(t);$$

$$\frac{dEp}{dt} = \left[\frac{G(t)}{G(t)+Gmp} + \frac{R(t)}{R(t)+Rmp}\right] Ep(t) - dEp(t);$$

$$\frac{dG}{dt} = a\big(S_G - G(t)\big) - \ q_{Gc}\ \frac{G(t)}{G(t)+Gmc} Ec(t) - \ q_{Gp} \left[\frac{G(t)}{G(t)+Gmp}\right] Ep(t);$$

$$\frac{dR}{dt} = a(S_R - R(t)) - \ q_{Rp}\ \left[\frac{R(t)}{R(t)+Rmp}\right] Ep(t);$$

which leads to the expected abundances of:

$$G^{eq} = \frac{d\ Gmc}{1-d};$$

$$R^{eq} = \frac{(1-d)d(\ Gmc-Gmp)Rmp}{(d-2)d(Gmc-Gmp)-Gmp};$$

$$Ec^{eq} = \frac{a}{d}\left(\frac{(S_G-G^{eq})}{q_{Gc}} + \frac{q_{Gp}}{q_{Gc}} \frac{(R^{eq}-S_R)}{q_{Rp}} \frac{(R^{eq}+Rmr)\ G^{eq}}{R^{eq}\ (G^{eq}+Gmr)}\right);$$

$$Ep^{eq} = \frac{a\ (SR-R^{eq})(R^{eq}+Rmp)}{R^{eq}\ q_{Rp}};$$

In Fig. S2 we show examples of parameter values where the emergence of a strain specialized on galactitol can rise in frequency.

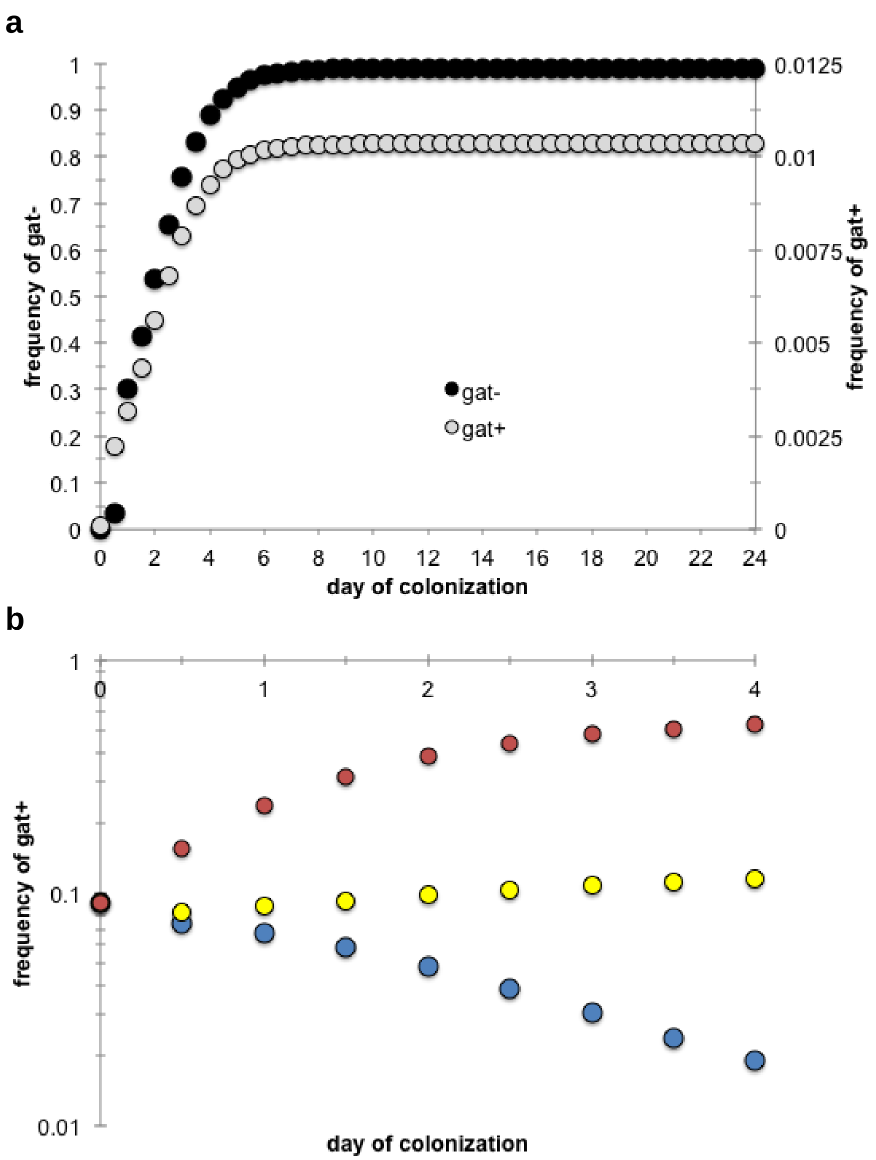

Fig. S1.

Dynamics of frequency change of *gat*+ and *gat*- clones under the two- strain two-resource model where *gat*+ only consumes galactitol. Model parameters are as follows: S_R=0.02; S_G=0.002; a=d=0.042; Gmc= $2\text{x}10^{-5}$; q_{Gc}= 10^{-3}; Rmn=$2\text{x}10^{-2}$; q_{Rn} =10^{-4}. (**a**) Black circles: dynamics of *gat*- emergence in a mono colonization with *gat*+; gray circles: dynamics of *gat*+ emergence in a mono colonization with *gat*-; (**b**) Dynamics of *gat*+ frequency (as $\log_{10}$) under diet supplementation with: S_G=0.002 in blue; S_G=0.027 in yellow; and S_G=0.252 in red.

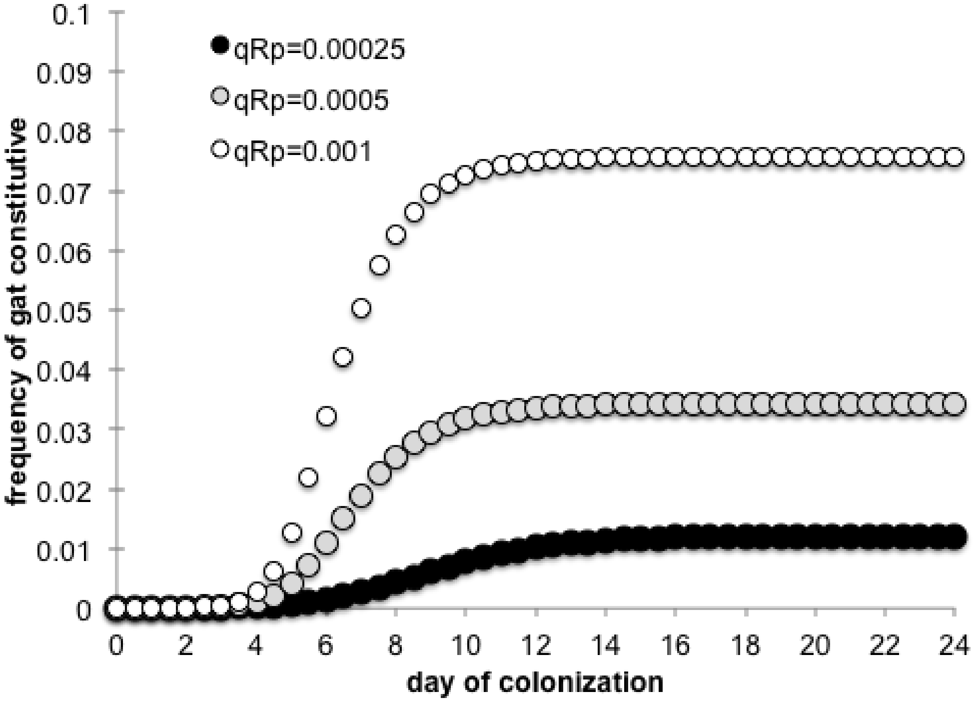

Fig. S2.

Dynamics of frequency change of the *gat*+ strain, which constitutively expresses the *gat* operon and thus is assumed to only consume galactitol, when invading a resident population of *gat*+ clones, which can consume both galactitol and the other resource. Model parameters are as follows: S_R=0.02; S_G=0.002; *a*=*d*=0.042; *Gmc*= $2\text{x}10^{-5}$; q_{Gc}=10^{-3}; *Rmp*=$2\text{x}10^{-2}$; *Gmp* = $2\text{x}10^{-4}$; q_{Gp}= 10^{-4}; and q_{Rp} as indicated.

🟨 k_Bacteria;p_Actinobacteria;c_Actinobacteria;o_Bifidobacteriales;f_Bifidobacteriaceae;g_Bifidobacterium
🟫 k_Bacteria;p_Actinobacteria;c_Coriobacteriia;o_Coriobacteriales;f_Coriobacteriaceae;g_
🟪 k_Bacteria;p_Actinobacteria;c_Coriobacteriia;o_Coriobacteriales;f_Coriobacteriaceae;g_Adlercreutzia
🟩 k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_;g_
⬜ k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Bacteroides
🟫 k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_Rikenellaceae;g_
🟪 k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_S24-7;g_
🟨 k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;Other;Other
⬛ k_Bacteria;p_Deferribacteres;c_Deferribacteres;o_Deferribacterales;f_Deferribacteraceae;g_Mucispirillum
🟦 k_Bacteria;p_Firmicutes;c_Bacilli;o_Bacillales;f_Paenibacillaceae;g_Paenibacillus
🟫 k_Bacteria;p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Lactobacillaceae;g_Lactobacillus
🟥 k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_;g_
🟪 k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Clostridiaceae;g_Clostridium
🟩 k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_
🟦 k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_Blautia
🟪 k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;Other
🟪 k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_
🟥 k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_Oscillospira
🟪 k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_Ruminococcus
🟥 k_Bacteria;p_Firmicutes;c_Erysipelotrichi;o_Erysipelotrichales;f_Erysipelotrichaceae;g_[Eubacterium]
🟩 k_Bacteria;p_Firmicutes;c_Erysipelotrichi;o_Erysipelotrichales;f_Erysipelotrichaceae;g_Allobaculum
🟦 k_Bacteria;p_Proteobacteria;c_Alphaproteobacteria;o_;f_;g_
🟪 k_Bacteria;p_Proteobacteria;c_Betaproteobacteria;o_Burkholderiales;f_Alcaligenaceae;g_Sutterella
🟧 k_Bacteria;p_Proteobacteria;c_Gammaproteobacteria;o_Enterobacteriales;f_Enterobacteriaceae;g_
🟥 k_Bacteria;p_Proteobacteria;c_Gammaproteobacteria;o_Enterobacteriales;f_Enterobacteriaceae;Other
🟨 k_Bacteria;p_Proteobacteria;c_Gammaproteobacteria;o_Pseudomonadales;f_Moraxellaceae;g_Acinetobacter
⬛ Unassigned;Other;Other;Other;Other;Other

Fig. S3.

Full legend for figures 4C, 6B and S4.

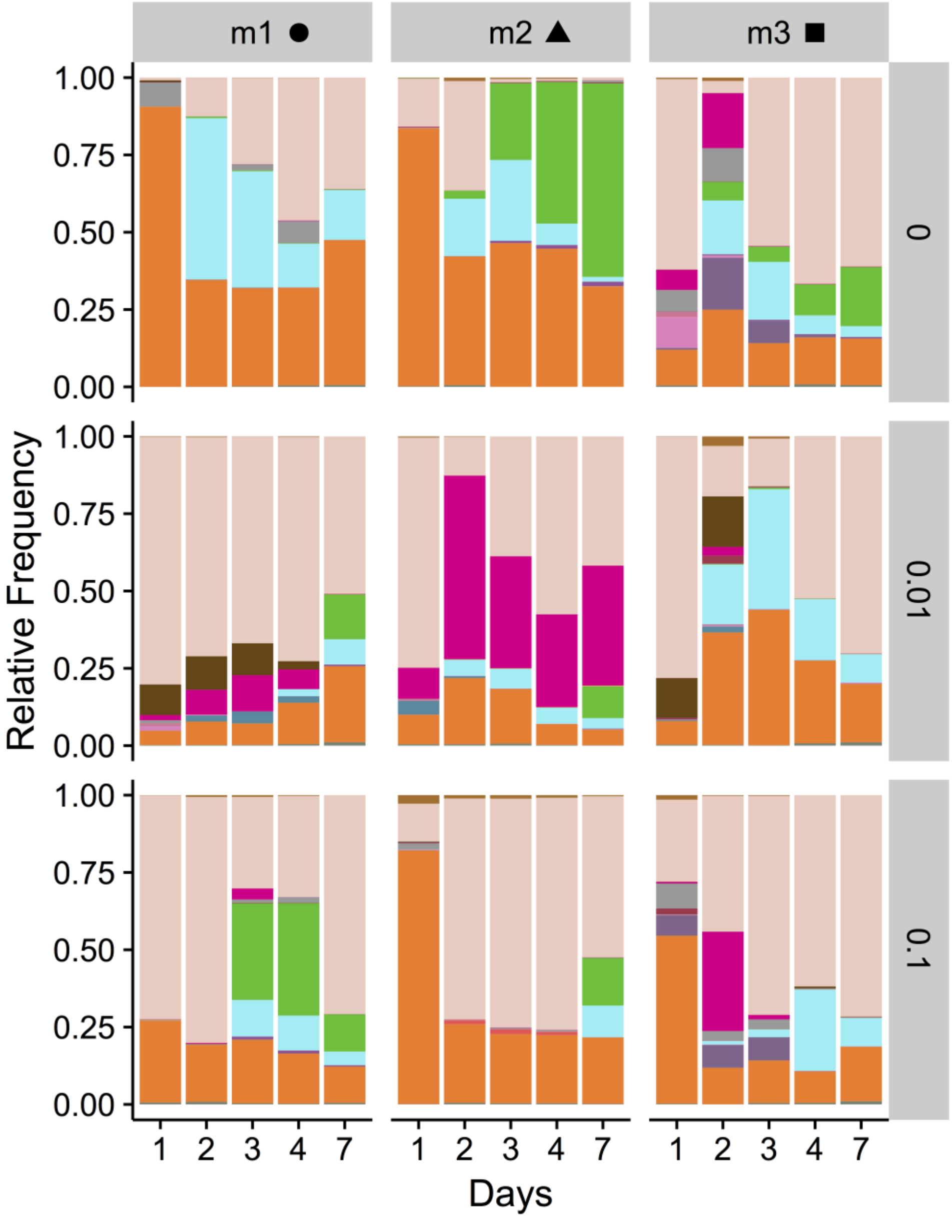

Fig. S4.
Temporal dynamics of the gut microbiota for mice colonized with a 1:10 ratio of gat-positive to gat-negative *E. coli* and whose diet was supplemented with different concentrations of galactitol in their drinking water (0, 0.01 or 0.1%). Symbols for each mouse correspond to those shown in fig. 4. In Fig. 4, we use samples from days 3, 4 and 7 for the analysis. This is because by this time, both the microbiota and the frequency of gat-positive *E. coli* appear to be near equilibrium.

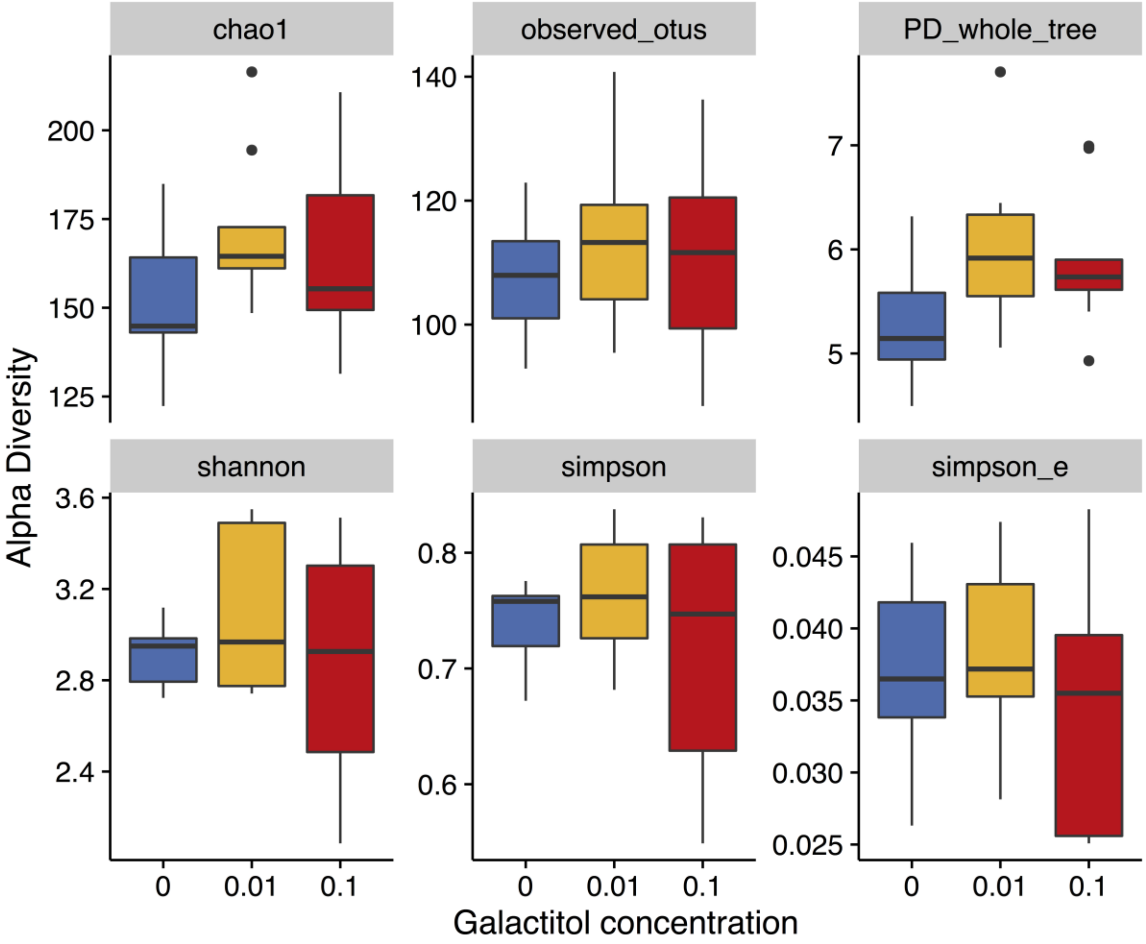

Fig. S5.

Alpha diversity is insensitive to galactitol supplementation. chao1 - chao1 richness estimator; observed_otus – number of distinct OTUs; PD_whole_tree – Faith's phylogenetic diversity; shannon – Shannon entropy; simpson – Simpson index; simpson_e – Simpson's evenness. No significant effect of galactitol supplementation is observed ($p<0.05$, determined from linear mixed effects models) for all indices, except Faith's phylogenetic diversity (borderline effect, with $p=0.045$).

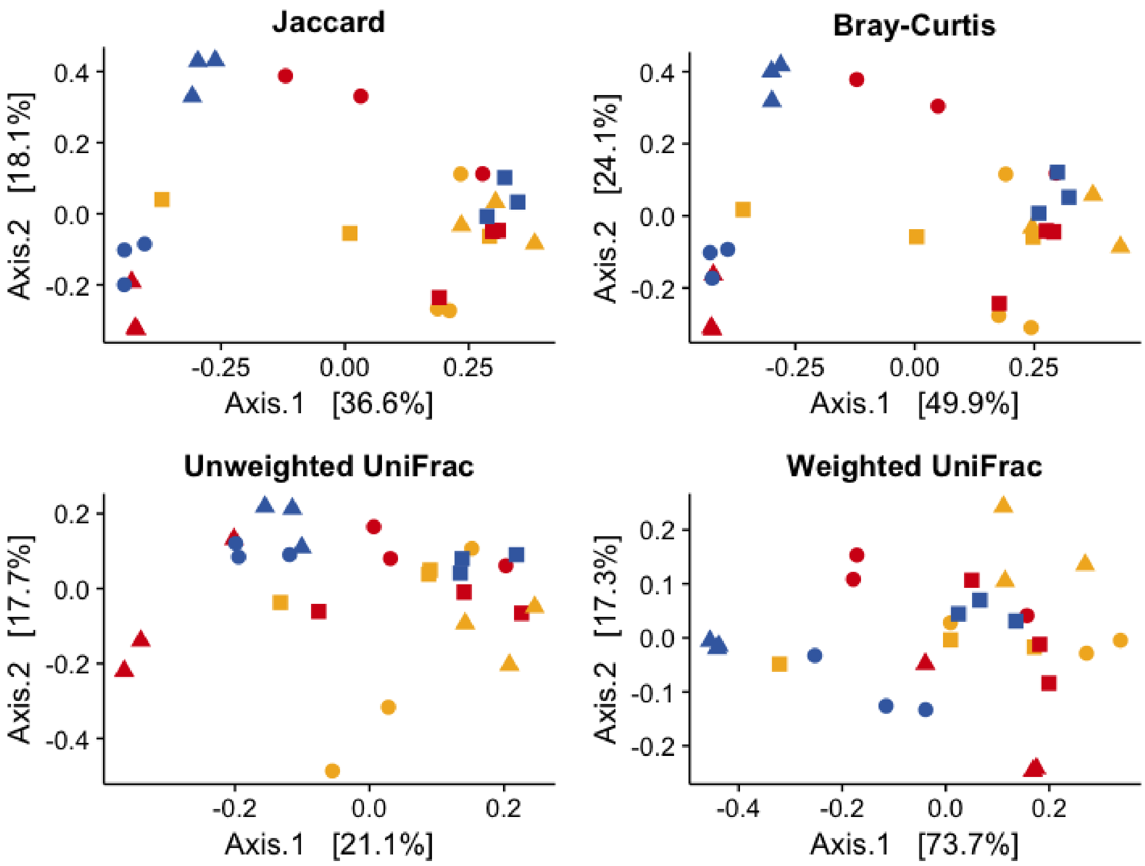

Fig. S6.

Principal coordinate analyses for the galactitol supplementation experiment using four different beta-diversity metrics: Jaccard, Bray-Curtis, unweighted and weighted UniFrac (related to Fig. 4). Blue, yellow and red points stand for 0, 0.01 and 0.1% galactitol concentration in mice drinking water, respectively. Symbols represent different mice as in Fig. S4. Data from days 3, 4 and 7 is used in this figure. Using PERMANOVA (with the *adonis* function in R) showed that galactitol concentration does not significantly affect beta-diversity (across the four metrics), as is apparent from the PCoA plots.

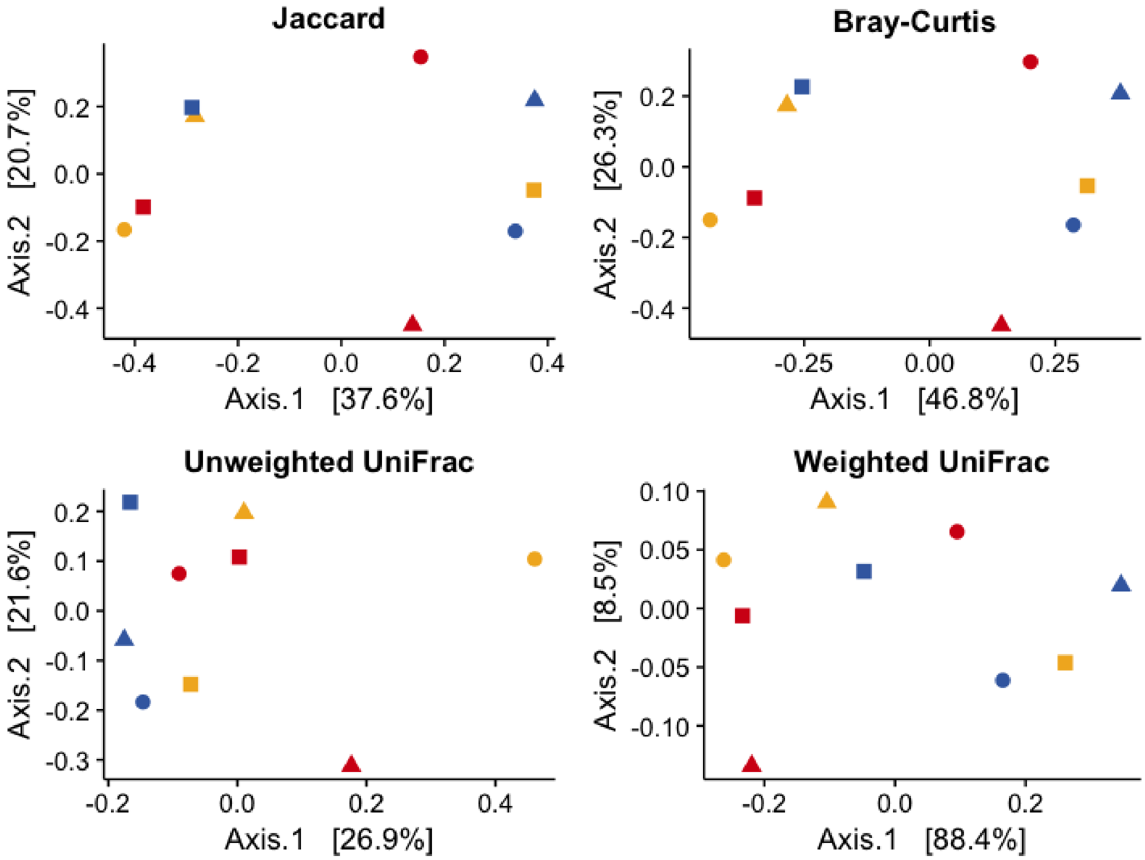

Fig. S7.

Principal coordinate analyses for the galactitol supplementation experiment using four different beta-diversity metrics: Jaccard, Bray-Curtis, unweighted and weighted UniFrac (related to Fig. 4). Blue, yellow and red points stand for 0, 0.01 and 0.1% galactitol concentration in mice drinking water, respectively. Symbols represent different mice as in Fig. S4. Only data from day 3 is used in this figure. Using PERMANOVA (with the *adonis* function in R) showed that galactitol concentration does not significantly affect beta-diversity (across the four metrics), as is apparent from the PCoA plots.

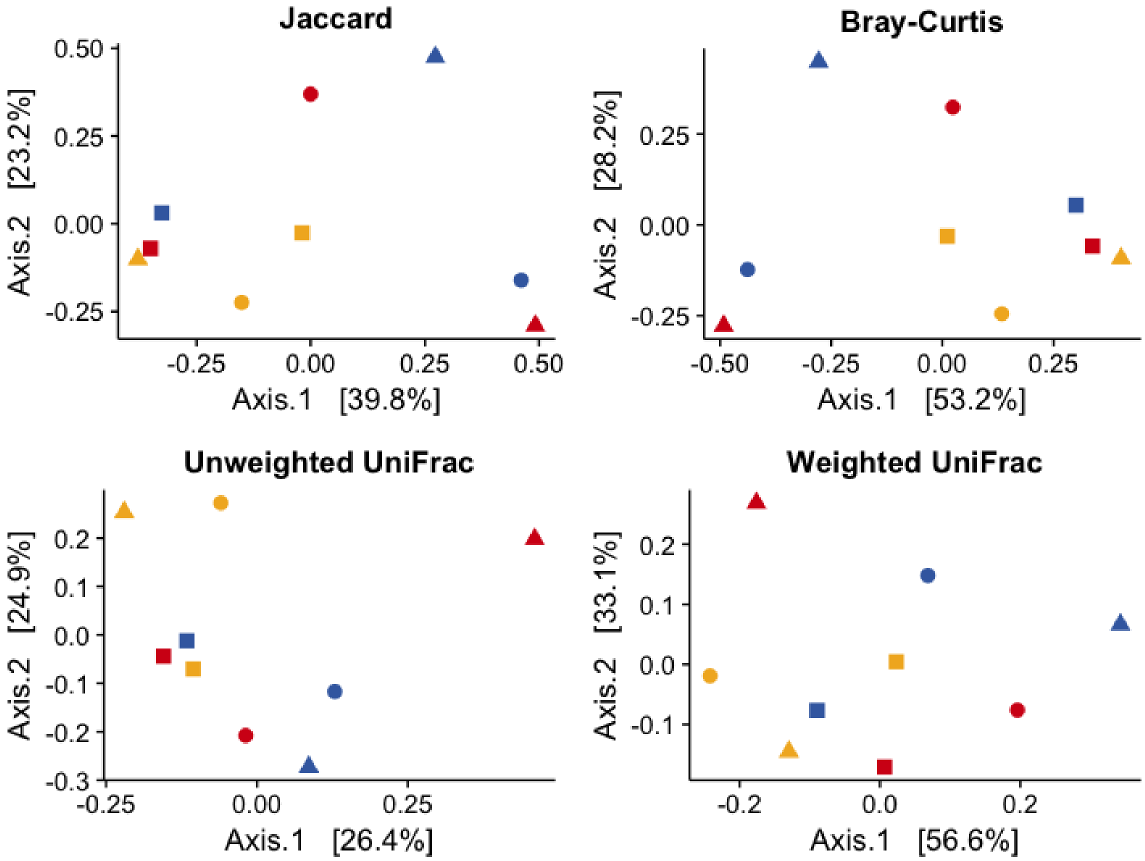

Fig. S8.

Principal coordinate analyses for the galactitol supplementation experiment using four different beta-diversity metrics: Jaccard, Bray-Curtis, unweighted and weighted UniFrac (related to Fig. 4). Blue, yellow and red points stand for 0, 0.01 and 0.1% galactitol concentration in mice drinking water, respectively. Symbols represent different mice as in Fig. S4. Only data from day 4 is used in this figure. Using PERMANOVA (with the *adonis* function in R) showed that galactitol concentration does not significantly affect beta-diversity (across the four metrics), as is apparent from the PCoA plots.

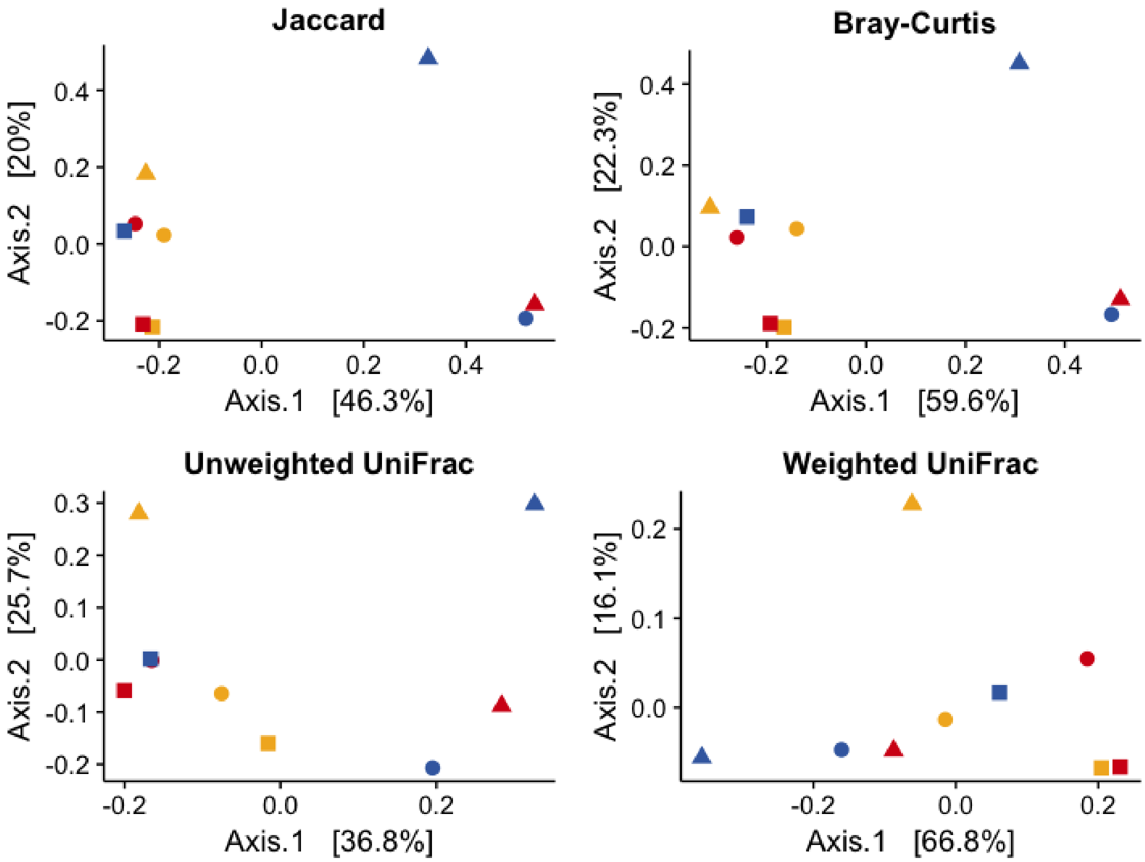

Fig. S9.

Principal coordinate analyses for the galactitol supplementation experiment using four different beta-diversity metrics: Jaccard, Bray-Curtis, unweighted and weighted UniFrac (related to Fig. 4). Blue, yellow and red points stand for 0, 0.01 and 0.1% galactitol concentration in mice drinking water, respectively. Symbols represent different mice as in Fig. S4. Only data from day 7 is used in this figure. Using PERMANOVA (with the *adonis* function in R) showed that galactitol concentration does not significantly affect beta-diversity (across the four metrics), as is apparent from the PCoA plots.

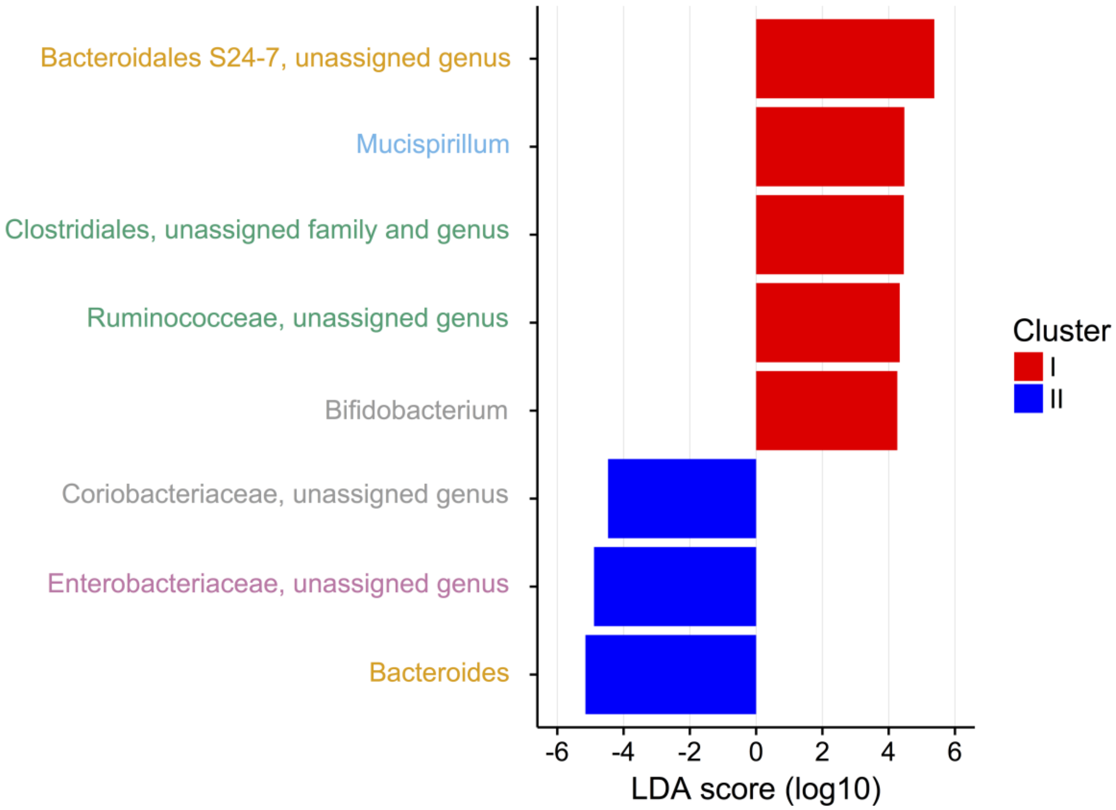

Fig. S10.

Linear discriminant analysis score (LDA) of 97% OTUs enriched for either cluster I (red) or cluster II (blue), as inferred using LefSe. OTUs are coloured according to their phylum: Actinobacteria (grey), Bacteroidetes (orange), Deferribacteres (blue), Firmicutes (green) and Proteobacteria (purple).

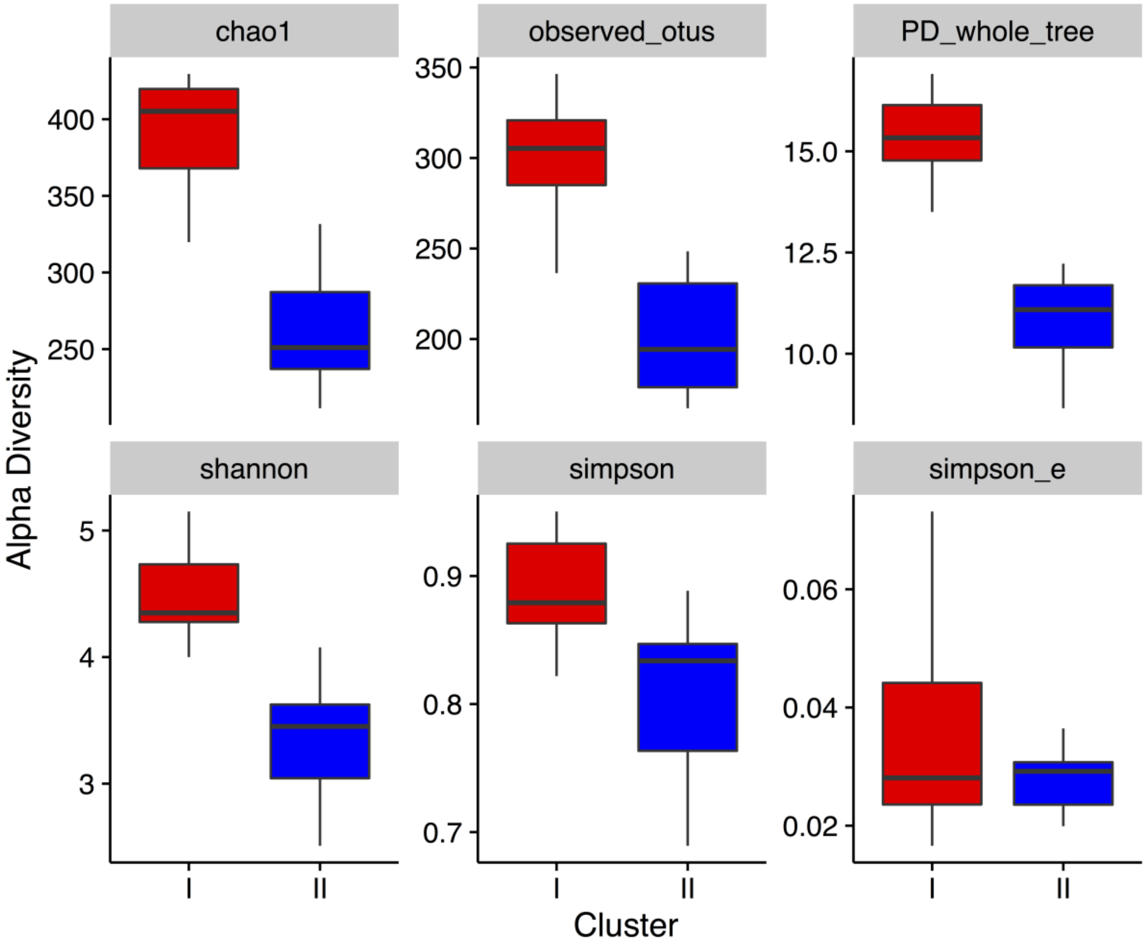

Fig. S11.

Alpha diversity is higher in cluster I than in cluster II, across multiple indices. Alpha diversity metrics are the same as described in Fig. S5. Significant differences between the two clusters ($p<0.05$, determined from linear mixed effects models) are detected for all indices except Simpson’s evenness.

Table S1.

Genetic and phenotypic (in bold) polymorphism in two evolved populations. Mutations were detected by whole genome sequencing of samples of each population at different time points (shown in red) or by target sequencing (shown in black). In addition the frequency of the *gat*-rev phenotype in each of the fluorescence backgrounds was estimated by plating in selective media (shown in bold). The full data is graphically represented in Fig. 2.

Population	Genome Position	Gene	Mutation	Annotation	Frequency of mutations along time (days)																	
					1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**17***	**19**	**21**	**24***
2.14	953,892	*focA/ycaO*	IS5 Ins	intergenic (-203/+203)											0				0			0.17
	1,740,247	*cfa*	Δ1 bp	coding (811/1149 nt)											0				0			0.06
		gat-rev phenoype-yfp				**0**		**0**		**0**		**0.03**		**0.12**		**0.28**		**0.27**	**0.45**			**0.13**
		gat-rev phenoype-cfp				**0**		**0**		**0**		**0**		**0.01**		**0.02**		**0**	**0.05**			**0**
	2,172,089	*gatC*	Δ1 bp	coding (212/1356 nt)											0.05				0.13			0.02
	3,718,679	*hok /insJ*	C→T	intergenic (-56/-24)											0				0			0.07
	4,601,151	*yjjP/yjjQ*	IS5 Ins	intergenic (-270/-349)											0	0	0.02	0.1	0.05			0
		yjjP/yjjQ	IS Ins												0				0.02			0
		yjjY/yjtD	IS Ins												0				0			0.06
	4,638,771	*yjjY/yjtD*	IS5 Ins	intergenic (+206/-194)											0				0			0.06
					1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**17***	**19**	**21**	**24***
2.15	953,863	*focA /ycaO*	IS1 Ins	intergenic (-174/+224)							0	0.04	0.2		0.61				0.84			0.77
	1,268,674	*ldrA /ldrB*	G→T	intergenic (-176/+252)											0.05				0			0
	1,296,579	*adhE*	Δ3 bp	coding (764-766/2676 nt)											0				0			0.05
	1,324,623	*yciQ*	Δ36 bp	coding (1854-1889/1896											0				0			0.16
	1,299,449	*oppA*	IS5 Ins	coding (244-247/1632 nt)											0				0			0.21
		gat-rev phenoype-yfp								**0**		**0**		**0**		**0**		**0**	**0**	**0**	**0**	**0**
		gat-rev phenoype-cfp								**0.04**		**0.24**		**0.2**		**0.1**		**0.14**	**0.06**	**0.14**	**0.12**	**0.16**
	2,172,078	*gatC*	Δ1 bp	coding (223/1356 nt)											0.18				0.03			0.18
	3,506,347	*php*	Δ1 bp	coding (266/879 nt)											0				0.04			0
	3,925,453	Inv 15bp	asnA	coding (276-291/993 nt)											0.05				0			0
	4,031,240	*trkH*	G→A	G25R (GGG→AGG)											0.6				0.84			0.73
	4,638,637	*yjjY/yjtD*	IS5 Ins	intergenic (+72/-325)							0.01	0.01	0.04		0.11				0.13			0.19

Table S2.

Primers used for the reconstruction of the transcriptional regulator of the galactitol metabolism (*gatR*)

*gatR*_*SacB*_F	GTAATAATAATGTCAATTTTTTCGATTGGTAGTACCTGATTAAAACCGCGTCCTAATTTTTGTTGACACTCTATC
*gatR*_*SacB*_R	GCGGGAGTATTTGCTGCCTCGGAAGCGACAATCCGTGCCGATTTGCGCTTATCAAAGGGAAAACTGTCCATATGC
*yegS*_F	TGCCATTAGGAACCGCCAAT
*gatD*_R	TGCGTGACATCGCTCGTAAT

References

van Opheusden JHJ, Hemerik L, van Opheusden M, van der Werf W. 2015. Competition for resources: complicated dynamics in the simple Tilman model. SpringerPlus [Internet] 4. Available from: http://www.springerplus.com/content/4/1/474

Tilman D. 1982. Resource competition and community structure. Monogr. Popul. Biol. 17:1–296.