

1 Computer-guided design of optimal microbial
2 consortia for immune system modulation

3
4 Richard R. Stein^{1,2,3,4,§,*}, Takeshi Tanoue^{5,8,§}, Rose L. Szabady⁶, Shakti K. Bhattacharai⁷,
5 Bernat Olle⁶, Jason M. Norman⁶, Wataru Suda^{8,9}, Kenshiro Oshima⁹, Masahira Hattori⁹,
6 Georg K. Gerber¹⁰, Chris Sander^{1,11,4}, Kenya Honda^{5,8} & Vanni Bucci^{7,12,13,*}

7
8 **Affiliations:**

- 9 1. Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute,
10 450 Brookline Avenue, Boston, MA 02215, USA.
11 2. Department of Biostatistics, Harvard T.H. Chan School of Public Health, 44 Binney
12 Street, Boston, MA 02115, USA.
13 3. Department of Systems Biology, Harvard Medical School, 200 Longwood Avenue,
14 Boston, MA 02115, USA.
15 4. Broad Institute of MIT and Harvard, 415 Main Street, Cambridge, MA 02142, USA.
16 5. RIKEN Center for Integrative Medical Sciences, 1-7-22 Suehiro-cho, Tsurumi-ku,
17 Yokohama City, Kanagawa, 230-0045, Japan.
18 6. Vedanta Biosciences, 19 Blackstone Street, Floor 3, Cambridge, MA 02139, USA.
19 7. Engineering and Applied Sciences PhD Program, University of Massachusetts
20 Dartmouth, 285 Old Westport Road, N. Dartmouth, MA 02747, USA.
21 8. Department of Microbiology and Immunology, Keio University School of Medicine, 35
22 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan.
23 9. Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha,
24 Kashiwa, Chiba 277-8561, Japan.
25 10. Massachusetts Host-Microbiome Center, Department of Pathology, Brigham and
26 Women's Hospital, Harvard Medical School, 60 Fenwood Road, Boston, MA 02115,
27 USA.
28 11. Department of Cell Biology, Harvard Medical School, 200 Longwood Avenue,
29 Boston, MA 02115, USA.
30 12. Department of Biology, Program in Biotechnology and Biomedical Engineering,
31 University of Massachusetts Dartmouth, 285 Old Westport Road, N. Dartmouth, MA
32 02747, USA.

33 13. UMass Dartmouth Center for Microbial Informatics and Statistics, 285 Old Westport
34 Road, N. Dartmouth, MA 02747, USA.

35

36 [§] Co-first authors

37 * Corresponding authors: stein@jimmy.harvard.edu, vbucci@umassd.edu

38

39 **Keywords:** host–microbe interaction, regulatory T-cells, immune system modulation,
40 microbiome modeling

41 Abstract

42 Manipulation of the gut microbiota holds great promise for the treatment of diseases.
43 However, a major challenge is the identification of therapeutically potent microbial
44 consortia that colonize the host effectively while maximizing immunologic outcome.
45 Here, we propose a novel workflow to select optimal immune-inducing consortia from
46 microbiome composition and immune effectors measurements. Using published and
47 newly generated microbial and regulatory T-cell (T_{reg}) data from germ-free mice, we
48 estimate the contribution of twelve Clostridia strains with known immune-modulating
49 effect to T_{reg} induction. Combining this with a longitudinal data-constrained ecological
50 model, we predict the ability of every attainable and ecologically stable subconsortium in
51 promoting T_{reg} activation and rank them by the T_{reg} Induction Score (TrIS). Experimental
52 validation of selected consortia indicates a strong and statistically significant correlation
53 between predicted TrIS and measured T_{reg} . We argue that computational indexes, such
54 as the TrIS, are valuable tools for the systematic selection of immune-modulating
55 bacteriotherapeutics.

56 eLife Digest

57 Manipulation of the gut microbiota holds great promise for the treatment of inflammatory
58 and allergic diseases. However, a major challenge is developing consortia of therapeutic
59 microbes that can stably colonize the host while at the same time maximizing a desired
60 immune outcome. This limitation stems from the fact that the number of combinations to
61 be tested scales exponentially with the number of strains to choose from, making this an
62 intractable experiment to be tackled systematically. In this work, we propose a novel

63 experimental and computational workflow that, by leveraging mathematical models
64 constrained on multimodal data of microbiome colonization and immune responses
65 quantification, allows for the prediction of therapeutically potent bacteria consortia.
66

67 To achieve this, first we use published and newly generated microbial colonization and
68 simultaneous regulatory T-cell (T_{reg}) abundance data from germ-free mice to build a
69 mathematical model resolving the contribution to T_{reg} induction of a set of known T_{reg} -
70 inducing Clostridia strains. Then, combining this model with a longitudinal data-
71 constrained ecological model that accurately predicts the dynamics of these T_{reg} -
72 modulating strains in germ-free (GF) mice, we derive measures to robustly select
73 combinations of these strains that maximize immune system activation. We demonstrate
74 the potential utility of our approach in predicting the potency of selected microbial
75 combinations by experimentally measuring T_{reg} induction *in vivo* of predicted strong,
76 intermediate and weak T_{reg} -inducing consortia. This work defines a novel framework for
77 the selection and characterization of optimal microbial consortia that stimulate T_{reg}
78 activation and attenuate disease in models of T_{reg} -mediated immune disease, such as,
79 ulcerative colitis.

80 Introduction

81 The intestinal microbiota has been shown to critically influence a multitude of host
82 physiological functions, often through modulation of the immune system (Gevers et al.,
83 2014; Paun et al., 2016). Evidence includes studies in germ-free mice, which show that
84 both pro-inflammatory T-helper 17 (T_h17) cells and anti-inflammatory regulatory T-cells
85 (T_{reg}) are reduced in numbers in the intestinal lamina propria relative to conventional or
86 specific pathogen-free mice (Atarashi et al., 2008), and that repopulation with specific
87 bacteria can reconstitute them (Atarashi et al., 2015; Ivanov et al., 2009). One route to
88 how the microbiota influences immune system activation is by the production of small
89 molecules (Donia and Fischbach, 2015). Microbially produced short-chain fatty acids
90 (SCFAs) were shown to facilitate extrathymic differentiation of immune system
91 modulatory T_{reg} (Arpaia et al., 2013; Atarashi et al., 2013; Smith et al., 2013) and are
92 implicated in T_{reg} -dependent anti-inflammatory properties of a mix of 17 human-derived
93 Clostridia (Atarashi et al., 2013; Tanoue et al., 2016). Interestingly, individual Clostridia
94 strains only have a modest effect on T_{reg} induction – optimal induction relies on the

95 synergistic interplay of several strains (Atarashi et al., 2013; Belkaid and Hand, 2014).
96 Because of these properties, there is currently great interest in the manipulation of this
97 community for the treatment of inflammatory and allergic diseases (Honda and Littman,
98 2012; Olle, 2013; Vieira et al., 2016). Efforts involving transplantation of bacteria from
99 healthy humans to humans with *C. difficile* infections or with metabolic diseases have
100 provided evidence that microbiota repopulation could be used as possible strategy for
101 disease prevention and/or treatment (van Nood et al., 2013; Vrieze et al., 2012). For this
102 reason, intestinal supplementation of defined compositions of gut bacteria to treat a
103 range of diseases is currently being pursued by multiple bio-pharmaceutical companies
104 (Garber, 2015). Determining what microbial consortia can colonize the host and stably
105 coexist with an already resident microbial community, while inducing the desired immune
106 response, is still a major challenge that hinders the translation of these efforts into the
107 clinic (Maldonado-Gómez et al., 2016; Weil and Hohmann, 2015).

108

109 Because in our previous work we identified seventeen possible candidate T_{reg} -inducing
110 strains to select from (Atarashi et al., 2013), determining combinations from these
111 seventeen strains that would maximally induce the desired immune phenotype is
112 experimentally infeasible as it would require the testing in mice of $2^{17}-1=131071$
113 possible subsets (Faith et al., 2014). Therefore, a computational approach for prioritizing
114 which subsets to test experimentally would have great utility. To our knowledge, to date
115 there has been no model that allows for the simultaneous prediction of the dynamics of
116 both the microbiota and the host immune response. Building on the approach to select
117 for optimal bacterial combinations from (Faith et al., 2014) and other efforts aimed at
118 identifying potential immune-modulating microbes (Geva-Zatorsky et al., 2017; Schirmer
119 et al., 2016), in this study, we overcome this problem by proposing a mathematical
120 modeling-based framework that, by resolving microbiome-immune system interactions
121 with parameters constrained to microbiome and immune effector experimental
122 observations, allows computational optimization of immune-stimulating bacterial
123 combinations. To achieve this, we use a series of logically connected analyses that
124 captures $CD4^+FOXP3^+ T_{reg}$ accumulation in the colonic lamina propria (Omenetti and
125 Pizarro, 2015; Round and Mazmanian, 2010) in response to the dynamics of human-
126 derived Clostridia strains in germ-free mice which we have been previously shown to
127 potently induce T_{reg} expansion (Atarashi et al., 2013) (Figure 1). To constrain this
128 proposed model framework to experimental data, we combine previously published

129 (Atarashi et al., 2013) and newly generated fluorescence-activated cell sorting (FACS)
130 data with simulated and newly generated time-series colonization data from germ-free
131 mice (Figure 2A). Subsequently, by applying a microbiome– T_{reg} mathematical model to
132 this combined data set we infer each strain's individual contribution to the CD4 $^{+}$ FOXP3 $^{+}$
133 T_{reg} pool (Figure 2B) and obtain an estimate for each consortium's T_{reg} induction
134 potential. To justify the usage of predicted mono-colonization concentrations from a
135 previously by some of us developed microbiome ecological model (Bucci et al., 2016),
136 we validate its ability in predicting temporal dynamics in response to different subsets of
137 T_{reg} -inducing strains (Figure 3A) and quantify the deviation of data and predictions
138 (Figure 3B). Introduction of the TrIS, which assigns a score to each predicted steady-
139 state microbial composition by the per-strain inferred contribution to T_{reg} induction while
140 accounting for their host colonization ability, enables us to identify combinations that
141 robustly maximize T_{reg} induction (Figure 4A). We analyze the relationship of TrIS and
142 biomass (Figure 4B) and well as metabolic features (Figure 4C and 4D) for the predicted
143 consortia and, most importantly, demonstrate the utility of our approach in predicting the
144 potency of selected microbial combinations by validating the T_{reg} induction and
145 colonization ability of model-predicted strong and weak T_{reg} -inducing consortia *in vivo*
146 (Figure 4E).

147
148 We envision that our framework, while in this study tailored to finding combinations
149 ameliorating auto-inflammatory conditions, may also have direct relevance to other
150 immune-system enhancing applications such as the optimal delivery of probiotic-based
151 cancer immunotherapies (Garrett, 2015).

152 Results

153 Generation of multimodal microbiome– T_{reg} data

154 The goal of this study is to develop a mathematical modeling-based framework to rapidly
155 and systematically select microbial consortia that maximize a desired immune outcome
156 when introduced in a specific host microbial background. To achieve this, we combine (i)
157 a microbiome ecological model that accurately describes the dynamics of these bacteria
158 in the host, with (ii) a microbiome– T_{reg} mathematical model that characterizes the
159 contribution of every strain to the immune phenotype of interest given corresponding

160 microbiome colonization data (Figure 1). For (i), we gathered newly produced
161 quantitative polymerase chain reaction (qPCR) colonization data from gnotobiotic mice
162 and combined it with a previously-published microbiome ecological model of the
163 dynamics of twelve previously described T_{reg} -inducing Clostridia strains that are part of
164 an original consortium of seventeen T_{reg} -inducing strains discovered by some of us
165 (Atarashi et al., 2013) (see also Table S1 for a breakdown of strains used in each
166 experiment and study). In contrast to the original study of (Bucci et al., 2016), we
167 included only twelve of the thirteen strains used there because; based on the modeling
168 and analysis of (Bucci et al., 2016), Strain 6 from the 13-strain set was predicted to not
169 stably colonize in the presence of the other twelve strains. The published data have
170 been reported in our previous work (Bucci et al., 2016) and include time-series
171 measurements of microbial abundances by qPCR under dietary perturbations. These
172 are used to derive a predictive microbiome ecology model in gnotobiotic conditions
173 based on an extension of the generalized Lotka–Volterra (gLV) equations (Hofbauer and
174 Sigmund, 1998) as introduced in (Stein et al., 2013). For the newly generated dataset,
175 we gavaged fourteen mice with one of three possible 11-strain subsets from the 12-
176 strain subset of original thirteen strains (Figure 2) and used the derived stool
177 measurements to validate the ability of our mathematical model in predicting unseen
178 conditions. The three 11-strain subsets were chosen based on our ‘keystone’ definition,
179 a measure describing the marginal predicted quantitative effect of removing each strain
180 from the full community (Bucci et al., 2016). Specifically, we included two 11-strain
181 combinations each missing one of the two highest kestoneness-scoring strains (VE202
182 Strain15 and VE202 Strain 4) and one 11-strain combination which lacks the lowest
183 kestoneness scoring strain (VE202 29). In analogy to (Bucci et al., 2016), each strain’s
184 density was profiled over time by qPCR with strain-specific primers (see Methods). To
185 resolve the contribution of each of these to T_{reg} induction for point (ii) we coupled the
186 colonization data from fecal content with newly collected and published FACS
187 measurements of the $CD4^+FOXP3^+$ T_{reg} population in the lamina propria of these mice
188 (Figure 2A). As it is crucial to capture each strain’s contribution alone and in combination
189 with others, we also included $CD4^+FOXP3^+$ T_{reg} measurements from our previously
190 published mono-colonization experiments (Atarashi et al., 2013) (Figure 2A). Because of
191 the lack of single-strain mono-colonization data in our previously published experiments
192 (Atarashi et al., 2013) we estimated them using the microbiome ecological model from
193 (Bucci et al., 2016). This choice was supported by its capability in predicting unseen

194 validation data (Figure 3A). Spearman's rank order correlation coefficient ranges from
 195 0.92 to 0.98, with p-value $< 10^{-16}$ between observations and predictions depending on
 196 the time point, (Figure 3B).

197 Derivation of the microbiome–T_{reg} mathematical model to select
 198 microbial consortia

199 We used the described microbiome colonization data and corresponding CD4⁺FOXP3⁺
 200 T_{reg} measurements to determine the contribution to the T_{reg} pool of each strain in the
 201 consortium. We start by subdividing CD4⁺ T-cells into two major subclasses depending
 202 on intracellular FOXP3 expression: CD4⁺FOXP3⁺ T_{reg} and the remainder among the
 203 CD4⁺ T-cells, the conventional CD4⁺FOXP3⁻ T-cells (Bilate and Lafaille, 2012;
 204 Rudensky, 2011). The combined concentrations of these two T-cell classes at time t in
 205 the colonic lamina propria therefore fulfill $c_{T_{reg}}(t) = c_{FOXP3^+}(t) + c_{FOXP3^-}(t)$. To include a
 206 variety of effects into our model, we assume regulatory T-cell dynamics to follow the gLV
 207 equations (Gerber, 2014; Hofbauer and Sigmund, 1998), which also account for the
 208 effect of microbial strains contained in the lumen on the CD4⁺FOXP3⁻ T-cell
 209 subpopulation. In addition, we use an extension of the standard gLV equations to include
 210 the impact of the Clostridia strains as external perturbation (Stein et al., 2013). The
 211 resulting microbiome–T_{reg} mathematical model is found as,

$$\frac{dc_{FOXP3^+}(t)}{dt} = c_{FOXP3^+}(t) \left(\alpha_{FOXP3^+} + \beta_{FOXP3^+FOXP3^+} c_{FOXP3^+}(t) + \beta_{FOXP3^+FOXP3^-} c_{FOXP3^-}(t) \right. \\ \left. + \sum_{k=1}^K \varepsilon_{i_k} c_{strain_{i_k}}(t) \right) \quad (1)$$

212 where α_{FOXP3^+} denotes the basal growth rate and $\beta_{FOXP3^+FOXP3^+}$ the self-interaction term
 213 of the CD4⁺FOXP3⁺ T-cell subpopulation. The interaction terms $\beta_{FOXP3^+FOXP3^-}$ and
 214 $\beta_{FOXP3^-FOXP3^+}$ represent the effect of the CD4⁺FOXP3⁻ on the CD4⁺FOXP3⁺ T-cell
 215 subpopulation and of the CD4⁺FOXP3⁺ on the CD4⁺FOXP3⁻ T-cell subpopulation,
 216 respectively (d'Onofrio, 2005). Consequently, positive interaction parameters correspond
 217 to activation, negative ones to inhibition. Moreover, ε_{i_k} denotes the effect of strain i_k on
 218 the CD4⁺FOXP3⁺ T-cell subpopulation. For long-term observations, $t \rightarrow \infty$, c_{FOXP3^+}
 219 dynamics are given by the (non-trivial) steady-state solution, which simplifies the
 220 microbiome–T_{reg} mathematical model of the relative CD4⁺FOXP3⁺ proportion $r_{FOXP3^+, ss}$ to

$$r_{\text{FOXP3}^+, \text{ss}} = \tilde{\alpha} + \sum_{k=1}^K \tilde{\varepsilon}_{i_k} c_{\text{strain}_{i_k}, \text{ss}} \quad (2)$$

221 where the absolute abundance of CD4⁺ cells is derived from $c_{\text{FOXP3}^{+/-}, \text{ss}} = c_{\text{T}_{\text{reg}}, \text{ss}}$.
 222 $r_{\text{FOXP3}^{+/-}, \text{ss}}$ and leverages FACS information, $r_{\text{FOXP3}^{+/-}, \text{ss}}$, for the activated CD4⁺FOXP3⁺
 223 T_{reg} (see Materials and Methods). We assume that the steady-state CD4⁺ T-cell
 224 concentration is constant, $c_{\text{T}_{\text{reg}}, \text{ss}} = \text{const.}$, across all microbial compositions. We justify
 225 this because we are dealing with genetically similar mice and with a consortium of
 226 closely related Clostridia. This assumption is however not justifiable when comparing
 227 non-colonized and colonized germ-free mice (Faith et al., 2014).

228

229 Assigning to $r_{\text{FOXP3}^+, \text{ss}}$ the measured CD4⁺FOXP3⁺ T_{reg} proportions after 35 days for
 230 each mouse and to $c_{\text{strain}_{i_k}, \text{ss}}$ the corresponding microbial profiles we infer the
 231 contribution of each strain to the CD4⁺FOXP3⁺ T_{reg} pool (Figure 2B) by solving equation
 232 (2) with an ℓ^2 -penalized least-square regression with one shrinkage parameter
 233 determined in a leave-one-sample-out cross-validation (Stein et al., 2013). The resulting
 234 normalized root-mean square deviation on left-out samples was found to be 12%, which
 235 is of sufficient quality.

236 Derivation of the T_{reg} Induction Score (TrIS) and selection of T_{reg}-
 237 inducing consortia

238 After deriving a model to predict our candidate strains dynamics in germ-free conditions
 239 and having resolved each strain's contribution to T_{reg} expansion we aim to use this
 240 information to computationally select for consortia that maximize T_{reg} induction while
 241 being ecologically robust (Bucci et al., 2016; Stein et al., 2013). To specify a measure for
 242 ecological robustness as well as immune induction potential of a microbial consortium in
 243 germ-free mice, we define the T_{reg} Induction Score (TrIS) as the average predicted
 244 regulatory T-cell activation potential of a given consortium of K strains
 245 $\{\text{strain}_{i_1}, \dots, \text{strain}_{i_K}\}$,

$$\text{TrIS}(\{\text{strain}_{i_1}, \dots, \text{strain}_{i_K}\}) = \frac{1}{N} \sum_{n=1}^N \sum_{k=1}^K \tilde{\varepsilon}_{i_k} c_{\text{strain}_{i_k}, \text{ss}}^{(n)} \quad (3)$$

246 If the predicted steady state of the microbial consortium $(c_{\text{strain}_{i_1}}^{(n)}, \dots, c_{\text{strain}_{i_K}}^{(n)})_{\text{ss}}$ that is
247 computed from the n -th Markov Chain Monte Carlo (MCMC) parameter estimate (Bucci
248 et al., 2016) is biologically meaningful and stable, then $c_{\text{strain}_{i_k}, \text{ss}}^{(n)}$ denotes the steady
249 state concentration of strain i_k ; otherwise $c_{\text{strain}_{i_k}, \text{ss}}^{(n)}$ is set to 0. Hence, the value of the
250 TrIS is indicative of the expected CD4 $^{+}$ FOXP3 $^{+}$ T_{reg} induction (after removing the host
251 contribution) and it is of the same units as the FACS measurements. We evaluated TrIS
252 for every possible strain combination that would stably colonize the gut with germ-free
253 background. In our computation, of the $2^{12}-1=4095$ possible steady state strain
254 configurations evaluated in $N=22,500$ MCMC parameter estimates, 84% are found to be
255 biologically meaningful. Interestingly, while the average TrIS increases with consortium
256 size, our analysis shows that a subset size of seven already has bacterial combinations
257 maximizing induction (Figure 4A). Furthermore, in addition to the strong correlation
258 between TrIS and the predicted total bacterial abundance in the consortium, we
259 observed that high-induction consortia display an especially large enrichment in the
260 abundance of Strain 27 (Figure 4B, Figure 4 – figure supplement 1). Because short
261 chain fatty acids (SCFAs) have been previously linked with colonic T_{reg} induction (Arpaia
262 et al., 2013) and increase in density upon supplementation with these strains (Atarashi
263 et al., 2013), we decided to test if modeling-predicted high T_{reg}-inducing consortia were
264 also enriched in SCFAs. We therefore compared the top 5-inducing microbial consortia
265 of size seven against their same-size counterpart bottom 5 (Figure 4C). We predicted
266 SCFAs concentration for each of the predicted compositions by summing up each
267 strain's metabolic output measured during mono-colonization experiments (Atarashi et
268 al., 2013; Narushima et al., 2014) normalized by each strain's modeling-predicted mono-
269 colonization density. We performed a Welch two-samples t-test for each of the predicted
270 SCFAs concentrations and found significant enrichment for all estimated SCFAs
271 ($p<0.05$, one tailed) in the high-TrIS consortia compared to the low ones (Figure 4D).

272 Experimental validation of mathematical modeling predictions

273 We decided to evaluate the ability of our approach in correctly predicting ranking of
274 consortia based on T_{reg}-induction potential. Due to regulatory constraints on the used
275 probiotic strains – limiting us to a maximum of four strains at a time in follow-up
276 experiments – we selected five different 4-strain combinations. We measured

277 CD4⁺FOXP3⁺ T_{reg} induction for five microbial consortia of size four and the control germ-
278 free background. We chose the two highest TrIS consortia (H1, H2 with ranks 1 and 2,
279 respectively), the lowest one (L with rank 495) and two TrIS-intermediate consortia of
280 interest (M1, M2 with ranks 129 and 452, respectively). Composition of each of the five
281 consortia is detailed in the caption of Figure 4E. We correlated the predicted TrIS score
282 with the mean observed CD4⁺FOXP3⁺ T_{reg} percentage and found a significant Pearson's
283 correlation coefficient of 0.97 and p-value < 0.01 (Figure 4E). Importantly, when using
284 16S rRNA sequencing to investigate the resulting colonization profiles for these
285 combinations, we observed that the high TrIS-scoring consortia (H1, H2, M1) all stably
286 colonized while the two low-scoring consortia only displayed a subset of the introduced
287 strains. This result remarkably reflects the nature of our scoring system which, in
288 addition to immune activation potential, also incorporates colonization success (equation
289 3).

290 Because the SCFA enrichment analysis from *in vivo* measurements of individual strains
291 (Figure 4C) predicted acetate to be significantly increased in the five high vs. low T_{reg}-
292 inducing consortia of size seven, we decided to compare the measured acetic acid
293 concentrations (Figure 4 – figure supplement 2) to predictions of acetate in the two high
294 (H1 and H2) and low (L) T_{reg}-inducing 4-strain consortia (Figure 4 – figure supplement 3).
295 ANOVA and subsequent post-hoc Tukey test showed significance in the measured
296 enrichment of H1 compared to L (adjusted p-value < 0.05) and of H1 compared to H2
297 (adjusted p-value < 0.05) (Figure 4 – figure supplement 2). However, no statistical
298 difference was observed between measurements in H2 and L (adjusted p-value < 0.05)
299 (Figure 4 – figure supplement 2). Remarkably, our model is able to capture this effect
300 when we compare the modeling-predicted acetate enrichment normalized to the mean
301 germ-free conditions with the observed acetic acid enrichment also normalized to its
302 germ-free conditions (Figure 4 – figure supplement 3).

303

304 Discussion

305 Manipulation of the intestinal microbiota with defined bacterial consortia for the treatment
306 of disease is a promising route for future therapeutics (Hansen and Sartor, 2015).
307 However, choosing bacterial combinations from the vast combinatorial space of
308 microbes that effectively colonize a host (possibly with a dysbiotic microbiota) and

309 maximize a desired host phenotype requires an enormous number of expensive and
310 time-consuming experimental trials (Faith et al., 2014). While data mining and statistical
311 approaches could aid in this process, the majority of available microbiome analysis
312 methods are still based on correlations (Gerber, 2014; Morgan et al., 2015) and inept to
313 predicting unseen phenotypes. Identification of causal associations between microbes
314 and host phenotype has been recently achieved using an experimentally-based
315 microbe–phenotype triangulation (Surana and Kasper, 2017). However, this approach
316 remains impractical when the goal is to explore and rank all the possible microbiome
317 consortium combinations with respect to a host phenotype of interest. In this study, we
318 leveraged our previous set of computational methods to forecast temporal microbiome
319 dynamics and make predictions on its stability in the context of infectious and
320 inflammatory diseases including *C. difficile* colonization and inflammatory bowel disease
321 (Bucci et al., 2016; Buffie et al., 2015; Stein et al., 2013).

322
323 Specifically, we have presented a novel modeling-based method that, by combining
324 dynamical predictions from a data-driven model with parameters from time-series data
325 with the inferred interactions between microbes and the immune response (such as
326 CD4⁺FOXP3⁺ T_{reg}), allows for the rational design of immune-modulating bacterial
327 consortia. Limiting our scope to the prediction of long-term steady states dynamics, we
328 were able to derive a microbiome–T_{reg} mathematical model which is constrained by
329 experimental observations from multimodal data. In the presented work, the data
330 modalities used are qPCR and 16S rRNA sequencing data to estimate microbial
331 abundances and FACS to assess CD4⁺FOXP3⁺ T_{reg} induction, respectively. However,
332 the proposed framework is naturally extensible to other host data, for example data from
333 metabolite profiling, immune readouts (e.g., CD8 T-cell activation, T_h1/T_h17 cell
334 depletion) and/or host transcriptome profiling (Atarashi et al., 2017; Morgan et al., 2015).
335 To evaluate candidate consortia that could be relevant – in this context – to the
336 treatment of auto-inflammatory diseases, we introduced a novel metric, TrIS, that
337 accounts for both ecological stability and efficiency in immune modulation. This metric
338 enabled us to rationally identify combinations that would stably colonize and at the same
339 time produce substantial T_{reg} induction in germ-free mice. Remarkably, in validation
340 experiments of five distinct modeling-guided consortia of predicted various induction
341 potential, we proved the ability of our approach in successfully selecting microbial
342 combinations with respect to a desired therapeutic activity. To our knowledge, this is to

343 date, the first study in which observation-constrained *in silico* modeling of microbiome
344 and host phenotype has successfully guided the rational design of drugs of defined
345 consortia of bacteria.

346

347 Our work relies on the gLV model assumption for both microbiome and CD4⁺ regulatory
348 T-cell dynamics, which despite being very versatile, has some limitations which include
349 the lack of third or higher-order interactions or saturation effects (Wangersky, 1978).
350 Moreover, with the data available to us, we needed to introduce the assumption that at
351 steady state the overall T-cell density is constant across mice and different microbiome
352 compositions. While this assumption was reasonable for the data we analyzed (see
353 Result section), the addition of future measurements on total T-cell abundance will likely
354 improve the prediction accuracy of our model as well as providing insight into the total
355 change of total T-cell population size in response to microbial immune induction.

356

357 We performed our data and modeling analysis in germ free mice. While previous work
358 by some of us showed that the amount T_{reg} induction is independent of the germ-free
359 background (IQI, Balb/c or C57BL/6) (Atarashi et al., 2015), it is noteworthy to
360 acknowledge that our model, which was trained using IQI mice, robustly predicted both
361 T_{reg} induction and acetate enrichment of the C57BL/6 colonized mice as used in the
362 validation experiments.

363 Our predictions suggest some degree of consistency in terms of functional outcome (e.g.
364 enrichment in acetate) of high-scoring consortia relative to low-scoring ones in germ free
365 mice. However, before translating these findings to therapeutic development, future
366 studies will need to be performed in not germ-free settings (e.g., SPF mice, humanized
367 mice) in order to account for the effects of an already-established flora. Individual-
368 specific microbiome features may constrain the colonization potential of the selected
369 strains due to specific ecological network effects (Smits et al., 2016), which suggests the
370 need for a careful characterization of the ecological interactions between the proposed
371 probiotic product and a specific recipient community (e.g., a single ulcerative colitis
372 patient, (Atarashi et al., 2013)).

373

374 Treatment of auto-immune diseases overall is not necessarily achieved by the exclusive
375 optimization of one objective function (e.g., for ulcerative colitis, the maximization of T_{reg}
376 activation), but may need the simultaneous manipulation of a multi-process host immune

377 spectrum which could include the concomitant reduction of pro-inflammatory phenotypes
 378 (Atarashi et al., 2015). Given new data availability and careful experimental design, we
 379 believe that the modeling framework proposed in this study could overcome these
 380 problems by introducing constraints into the scoring metric that also account for pre-
 381 colonized mouse model features. To the limit, in the presence of a large training data set
 382 that would include sufficient information on individual patient variation, our workflow will
 383 give us the unprecedented potential of testing microbiota manipulation on a personalized
 384 level *in silico*.

385

386 Mathematical-modeling based methods have the potential to greatly accelerate the
 387 development of treatment of human disease (Michor and Beal, 2015). In this work, we
 388 develop and demonstrate in a first-of-a-kind experimental validation the usefulness of a
 389 mathematical microbiome–immune system model which can be considered as a
 390 stepping-stone to accelerated prototyping and rational design of microbiota therapies.

391 Materials and Methods

392 Derivation of the microbiome–T_{reg} induction mathematical model

393 We assume the following dynamics for the CD4⁺FOXP3⁺ regulatory T-cells:

$$\frac{dc_{FOXP3^+}(t)}{dt} = c_{FOXP3^+}(t) \left(\alpha_{FOXP3^+} + \beta_{FOXP3^+FOXP3^+} c_{FOXP3^+}(t) + \beta_{FOXP3^+FOXP3^-} c_{FOXP3^-}(t) + \sum_{k=1}^K \varepsilon_{i_k} c_{\text{strain}_{i_k}}(t) \right).$$

394 Here, α_{FOXP3^+} denotes the basal growth rate and $\beta_{FOXP3^+FOXP3^+}$ the self-interaction term of
 395 the CD4⁺FOXP3⁺ T-cell subpopulation, while the interaction parameters $\beta_{FOXP3^+FOXP3^-}$
 396 and $\beta_{FOXP3^-FOXP3^+}$ characterize the effect of the CD4⁺FOXP3⁻ on the CD4⁺FOXP3⁺ T-cell
 397 subpopulation and of the CD4⁺FOXP3⁺ on the CD4⁺FOXP3⁻ T-cell subpopulation,
 398 respectively (d'Onofrio, 2005) and ε_{i_k} denotes the effect of strain i_k on the CD4⁺FOXP3⁺
 399 T-cell subpopulation. The non-trivial steady-state solution (i.e., the algebraic solution of
 400 the right-hand side of equation (1) set to 0 with $c_{FOXP3^+} \neq 0$) is found as,

$$c_{FOXP3^+, ss} = -\frac{1}{\beta_{FOXP3^+FOXP3^+}} \left(\alpha_{FOXP3^+} + \beta_{FOXP3^+FOXP3^-} c_{FOXP3^-, ss} + \sum_{k=1}^K \varepsilon_{i_k} c_{\text{strain}_{i_k}, ss} \right).$$

401 Using $c_{T_{reg}, ss} = c_{FOXP3^+, ss} + c_{FOXP3^-, ss}$, the steady-state concentrations of CD4 $^{+}$ FOXP3 $^{+}$
 402 T $_{reg}$ and their remainder among the CD4 $^{+}$ T-cells, $c_{FOXP3^{+/-}, ss}$, are then derived from the
 403 FACS-based relative abundances $r_{FOXP3^{+/-}, ss}$ as, $c_{FOXP3^{+/-}, ss} = c_{T_{reg}, ss} \cdot r_{FOXP3^{+/-}, ss} =$
 404 $c_{T_{reg}, ss}(1 - r_{FOXP3^{-/+}, ss})$. Finally, the linear relationship between the relative abundances,
 405 $r_{FOXP3^+, ss}$, and the strain densities, $c_{strain_{i_k, ss}}$, is found as,

$$r_{FOXP3^+, ss} = \frac{1}{\beta_{FOXP3^+FOXP3^-} - \beta_{FOXP3^+FOXP3^+}} \left(\frac{\alpha_{FOXP3^+}}{c_{T_{reg}, ss}} + \beta_{FOXP3^+FOXP3^-} \right. \\ \left. + \frac{1}{c_{T_{reg}, ss}} \sum_{k=1}^K \varepsilon_{i_k} c_{strain_{i_k, ss}} \right) \\ \equiv \tilde{\alpha} + \sum_{k=1}^K \tilde{\varepsilon}_{i_k} c_{strain_{i_k, ss}}.$$

406 Assuming constant concentration of CD4 $^{+}$ T $_{reg}$, $c_{T_{reg}, ss} = \text{const.}$ across all possible
 407 microbiome compositions, the unknown parameters $\tilde{\alpha}$ and $\tilde{\varepsilon}_{i_k}$ are estimated in an ℓ^2 -
 408 penalized least-square regression (so-called ridge regression) with a shrinkage
 409 parameter $\lambda > 0$ which is determined in a leave-one-out cross-validation as $\lambda^* = 2$.

410 Collection of experimental data for training the microbiome–immune
 411 system model

412 Strain abundance profiling

413 In our previous work (Bucci et al., 2016), we have inferred a mathematical model
 414 describing the dynamics of a 13-strain subset from the original 17-strain human-derived
 415 Clostridia consortium from (Atarashi et al., 2013) in germ-free mice. Using newly
 416 generated data from the same experimental setup, we are now able to assess its
 417 predictive quality for three distinct 11-strain subsets. The three 11-strain compositions
 418 were selected based on their capability in the simulations to maintain stability when
 419 removing either one of two high (Strain 15 and Strain 4; in (Bucci et al., 2016) referred to
 420 as cases A and B, respectively) or one low (Strain 29 or C) keystone strains. For the
 421 experimental validation, germ-free IQI mice were purchased from Sankyo Laboratories
 422 (Japan), randomized and maintained in germ-free vinyl isolators in the animal facility of
 423 RIKEN. Twelve T $_{reg}$ -inducing Clostridia strains were selected from the previously
 424 reported VE202 consortium consisting of 17 T $_{reg}$ -inducing strains (4, 7, 9, 13–16, 21, 26–

425 29) and were individually cultured in modified Eggerth Gagnon broth under strictly
426 anaerobic conditions (80% N₂, 10% H₂, 10% CO₂) at 37 °C in an anaerobic chamber
427 (Coy Laboratory Products, Grass Lake, MI) to confluence. The cultured bacterial strains
428 were then mixed and the three mixtures of 11 strains (described above) were orally
429 inoculated into five IQI germ-free adult mice each. One mouse for condition B died and
430 was therefore discarded from the study. After an initial 9-day interval of acclimation, we
431 collected fecal pellets at 2–4 days interval until day 35, time at which mice were
432 euthanized and analyzed for CD4⁺FOXP3 induction as in (Atarashi et al., 2013).
433 Colonization levels for each strain were assessed by amplifying strain-specific regions
434 with qPCR as described in (Bucci et al., 2016). Bacterial genomic DNA was extracted
435 from 1–2 fecal pellets using QIAamp DNA Stool Mini Kit (Qiagen, Hilden, Germany). The
436 amount of DNA was quantified using a Qubit dsDNA HS assay kit and Qubit fluorometer
437 (Invitrogen, Carlsbad, CA). DNA was then subjected to qPCR using Thunderbird SYBR
438 qPCR Mix (TOYOBO, Osaka, Japan) and a LightCycler 480 (Roche) with primers
439 specific to 16S ribosomal RNA (rRNA) genes of the twelve Clostridia strains as in (Bucci
440 et al., 2016). Quantification of each strain in each sample was accomplished using
441 standard curves of known concentrations of DNAs purified from each strain individually
442 cultured *in vitro*. Strain densities in each sample were calculated by dividing the above
443 absolute quantification numbers by the weight of the extracted fecal DNA. The 11-strain
444 experiments were ethically approved by RIKEN, Keio and Azabu Universities under
445 protocol H24-9(14) from RIKEN.

446 **Estimation of CD4⁺FOXP3⁺ T_{reg} – Isolation of intestinal lamina propria
447 lymphocytes and flow cytometry**

448 The colons were collected and opened longitudinally, washed with PBS to remove all
449 luminal contents and shaken in Hanks' balanced salt solution (HBSS) containing 5 mM
450 EDTA for 20 minutes at 37 °C. After removing epithelial cells, muscle layers and fat
451 tissue using forceps, the lamina propria layers were cut into small pieces and incubated
452 with RPMI1640 containing 4% fetal bovine serum, 0.5 mg/ml collagenase D, 0.5 mg/ml
453 dispase and 40 mg/ml DNase I (all Roche Diagnostics, Risch-Rotkreuz, Switzerland) for
454 1 hour at 37°C in a shaking water bath. The digested tissues were washed with HBSS
455 containing 5 mM EDTA, resuspended in 5 ml of 40% Percoll (GE Healthcare, Boston,
456 MA) and overlaid on 2.5 ml of 80% Percoll in a 15-ml Falcon tube. Percoll gradient
457 separation was performed by centrifugation at 850 g for 25 min at 25 °C. The lamina

458 propria lymphocytes were collected from the interface of the Percoll gradient and
459 suspended in ice-cold PBS. For analysis of T_{reg} , isolated lymphocytes were labeled with
460 the LIVE/DEAD fixable dead cell stain kit (Life Technologies, Carlsbad, CA) to exclude
461 dead cells from the analysis. Then, surface and intracellular staining of CD3, CD4 and
462 FOXP3 was performed using the BV605-labelled anti-CD3 (17A2, Biolegend, San Diego,
463 CA), BV421-labelled anti-CD4 (RM4-5, Biolegend), Alexa700-labelled anti-FOXP3
464 antibody (FJK-16 s, eBioscience, San Diego, CA), and FOXP3 staining buffer set
465 (eBioscience). The antibody-stained cells were analyzed with LSR Fortessa and data
466 were analyzed using FlowJo software (Tree Star, Ashland, OR).

467 **Measurement of organic acids**

468 Organic acid concentrations in caecal contents were determined by gas
469 chromatography-mass spectrometry (GC-MS). Caecal contents (10 mg) were disrupted
470 using 3-mm zirconia/silica beads (BioSpec Products) and homogenized in extraction
471 solution containing 100 ml of internal standard (100 mM crotonic acid), 50 ml of HCl and
472 200 ml of ether. After vigorous shaking using a Shakemaster neo (Bio Medical Science)
473 at 1,500 rpm for 10 min, homogenates were centrifuged at 1,000 g for 10 min and then
474 the top ether layer was collected and transferred into new glass vials. Aliquots (80 ml) of
475 the ether extracts were mixed with 16 ml of N-tert-butyldimethylsilyl-N-
476 methyltrifluoroacetamide (MTBSTFA). The vials were sealed tightly by screwing and
477 heated at 80 °C for 20 min in a water bath, and left at room temperature for 48 h for
478 derivatization. The samples were then run through a 6890N Network GC System
479 (Agilent Technologies) equipped with HP-5MS column (0.25 mm 330 m 30.25 mm) and
480 5973 Network Mass Selective Detector (Agilent Technologies, Santa Clara, CA). Pure
481 helium (99.9999%) was used as a carrier gas and delivered at a flow rate of 1.2 ml/min.
482 The head pressure was set at 10 psi with split 10:1. The inlet and transfer line
483 temperatures were 250 µC and 260 µC, respectively. The following temperature
484 program was used: 60 µC (3 min), 60–120 °C (5 °C/min), 120–300 °C (20 °C/min). One
485 microliter quantity of each sample was injected with a run time of 30 minutes. Organic
486 acid concentrations were quantified by comparing their peak areas with the standards.

487 **Numerical simulations of the three 11-strain subsets used for microbiome–**
488 **T_{reg} model training**

489 We used 22,500 sets of Markov Chain Monte Carlo generalized Lotka–Volterra
490 parameter sets determined by applying the Bayesian Variable Selection algorithm within
491 MDSINE to the data of (Bucci et al., 2016), and the first time point of measured microbial
492 profiles for each of the 14 validation mice as initial condition, to simulate the gLV system
493 of differential equations corresponding to each mouse microbiome (Figure 3A).
494 Prediction accuracy was evaluated by calculating the Spearman correlation coefficient
495 between observed and predicted data (Figure 3B).

496 **Simulation of mono-colonization abundances**

497 As the experimental data from (Atarashi et al., 2013) only provided CD4⁺FOXP3 levels
498 for the mono-strain colonization experiments (at 35 days after inoculation) but no
499 measurement of the long-term microbial concentrations in the gut, we used instead the
500 corresponding estimated mono-strain colonization densities obtained from the gLV
501 model (section above and Figure 2A). In general, the long-term behavior of the gLV
502 system is determined by its steady states which are uniquely defined by the inferred
503 model parameters (Stein et al., 2013). In order to obtain sparse parameters, we selected
504 from the 22,500 MCMC parameter estimates those corresponding to the median of each
505 single model parameter. These were subsequently used to compute the steady-state
506 densities, which were then together with the measured CD4⁺FOXP3 used for training the
507 microbiome–T_{reg} model.

508 **Collection of CD4⁺FOXP3⁺ T_{reg} data from 4-strain experiments for**
509 **validation of the modeling-based predictions**

510 Bacterial strains 4, 7, 9, 14, 15, 16, 27, 28, 29 were grown anaerobically in PYG broth
511 (Peptone, Yeast and Glucose broth from Anaerobe Systems, Cat no: AS-822) until they
512 reached stationary phase (48 hours for strains 27 and 29, 24 hours for the remaining
513 strains). Each 200 µl-mouse dose of a 4-strain LBP contained 50 µl of 20 times
514 concentrated stationary phase culture. Germ-free C57BL/6 mice aged 6–8 weeks were
515 randomized and gavaged with a total dose of 5·10⁷–2·10⁸ bacteria in a 200 µl, and
516 maintained under gnotobiotic conditions for four weeks. Use of a C57BL/6 background
517 for these experiments was motivated by availability if animal at the facility where we

518 performed the validation and justified by the fact that previous work from us has shown
519 that T_{reg} induction by our Clostridia strains does not differ between Balb/c, IQI, and
520 C57BL/6 mice (Atarashi et al., 2015). Mice were then sacrificed, colons harvested, and
521 lamina propria leukocytes isolated and stained for CD3⁺CD4⁺FOXP3⁺ T_{reg} as described
522 above. Eight mice each were used for consortia High 1 (H1-strains: 7, 27, 28, 29) and
523 High 2 (H2-strains: 4, 7, 27, 29). Five mice each were used for the intermediate high
524 (M1-strains: 4, 7, 14, 28), and intermediate low consortia (M2-strains: 9, 16, 27, 29).
525 Three mice were used for Low 1 (L1-strains: 14, 15, 16, 29). Colonization profiling was
526 determined through 16S rRNA sequencing (as above) and verified by blasting
527 representative sequences to a 16S VE202 fasta database. The 4-strain validation
528 experiments were performed in the Massachusetts Host Microbiome Center under
529 IACUC protocol 2016N000141.

530

531 Acknowledgements

532 RRS and CS acknowledge support by the National Resource for Network Biology
533 (NRNB) from the National Institute of Health (grant P41 GM103504).

534

535 GKG acknowledges support by the Defense Advanced Projects Agency Biological
536 Robustness in Complex Settings program (DARPA BRICS award HR0011-15-C-0094)
537 and the Brigham and Women's Hospital Precision Medicine Initiative.

538

539 CS acknowledges additional support by the National Institute of General Medical
540 Sciences (grant 5R01 GM106303) and National Human Genome Research Institute
541 (grant 5U41 HG006623) from the National Institute of Health and the Human Frontier
542 Science Program (grant RGP00055/2015).

543

544 KH acknowledges support by Takeda Science Foundation, and Core Research for
545 Evolutionary Medical Science and Technology (CREST) and Leading Advanced Projects
546 for Medical Innovation (LEAP) from Japan Agency for Medical Research and
547 Development (AMED).

548

549 VB acknowledges support by the National Institute of Allergy and Infectious Disease
550 (grant R15-AI112985-01A1), and the National Science Foundation (grant 1458347).

551 Competing interests

552 GG is a member of the Scientific Advisory Board of Kaleido, Inc. VB, TT, and KH have
553 received support from Vedanta Biosciences, Inc. under research agreements with their
554 institutions. KH is a Co-Founder and Scientific Advisory Board Member of Vedanta
555 Biosciences, Inc. RLS and JMN are employees of Vedanta Biosciences, Inc. BO is the
556 Chief Executive Officer of Vedanta Biosciences, Inc. The other authors have no
557 competing financial interests.

558 References

- 559 Arpaia, N., Campbell, C., Fan, X., Dikiy, S., van der Veeken, J., deRoos, P., Liu, H., Cross,
560 J.R., Pfeffer, K., Coffer, P.J., et al. (2013). Metabolites produced by commensal
561 bacteria promote peripheral regulatory T-cell generation. *Nature* *504*, 451–455.
- 562 Atarashi, K., Nishimura, J., Shima, T., Umesaki, Y., Yamamoto, M., Onoue, M., Yagita,
563 H., Ishii, N., Evans, R., Honda, K., et al. (2008). ATP drives lamina propria TH17 cell
564 differentiation. *Nature* *455*, 808–812.
- 565 Atarashi, K., Tanoue, T., Oshima, K., Suda, W., Nagano, Y., Nishikawa, H., Fukuda, S.,
566 Saito, T., Narushima, S., Hase, K., et al. (2013). Treg induction by a rationally selected
567 mixture of Clostridia strains from the human microbiota. *Nature* *500*, 232–236.
- 568 Atarashi, K., Tanoue, T., Ando, M., Kamada, N., Nagano, Y., Narushima, S., Suda, W.,
569 Imaoka, A., Setoyama, H., Nagamori, T., et al. (2015). Th17 Cell Induction by
570 Adhesion of Microbes to Intestinal Epithelial Cells. *Cell* *163*, 367–380.
- 571 Atarashi, K., Suda, W., Luo, C., Kawaguchi, T., Motoo, I., Narushima, S., Kiguchi, Y.,
572 Yasuma, K., Watanabe, E., Tanoue, T., et al. (2017). Ectopic colonization of oral
573 bacteria in the intestine drives TH1 cell induction and inflammation. *Science* *358*,
574 359–365.
- 575 Belkaid, Y., and Hand, T.W. (2014). Role of the microbiota in immunity and
576 inflammation. *Cell* *157*, 121–141.
- 577 Bilate, A.M., and Lafaille, J.J. (2012). Induced CD4+Foxp3+ regulatory T cells in
578 immune tolerance. *Annu. Rev. Immunol.* *30*, 733–758.
- 579 Bucci, V., Tzen, B., Li, N., Simmons, M., Tanoue, T., Bogart, E., Deng, L., Yeliseyev, V.,
580 Delaney, M.L., Liu, Q., et al. (2016). MDSINE: Microbial Dynamical Systems INference
581 Engine for microbiome time-series analyses. *Genome Biol.* *17*, 1–17.

- 582 Buffie, C.G., Bucci, V., Stein, R.R., McKenney, P.T., Ling, L., Gobourne, A., No, D., Liu, H.,
583 Kinnebrew, M., Viale, A., et al. (2015). Precision microbiome reconstitution restores
584 bile acid mediated resistance to *Clostridium difficile*. *Nature* *517*, 205–208.
- 585 Donia, M.S., and Fischbach, M.A. (2015). Small molecules from the human
586 microbiota. *Science* *349*, 1254766.
- 587 Faith, J.J., Ahern, P.P., Ridaura, V.K., Cheng, J., and Gordon, J.I. (2014). Identifying gut
588 microbe-host phenotype relationships using combinatorial communities in
589 gnotobiotic mice. *Sci. Transl. Med.* *6*, 220ra11.
- 590 Garber, K. (2015). Drugging the gut microbiome. *Nat. Biotechnol.* *33*, 228–231.
- 591 Garrett, W.S. (2015). Cancer and the microbiota. *Science* *348*, 80–86.
- 592 Gerber, G.K. (2014). The dynamic microbiome. *FEBS Lett.* *588*, 4131–4139.
- 593 Geva-Zatorsky, N., Sefik, E., Kua, L., Pasman, L., Tan, T.G., Ortiz-Lopez, A., Yanortsang,
594 T.B., Yang, L., Jupp, R., Mathis, D., et al. (2017). Mining the Human Gut Microbiota for
595 Immunomodulatory Organisms. *Cell* *168*, 928-943.e11.
- 596 Gevers, D., Kugathasan, S., Denson, L.A., Vázquez-Baeza, Y., Van Treuren, W., Ren, B.,
597 Schwager, E., Knights, D., Song, S.J., Yassour, M., et al. (2014). The Treatment-Naive
598 Microbiome in New-Onset Crohn's Disease. *Cell Host Microbe* *15*, 382–392.
- 599 Hansen, J.J., and Sartor, R.B. (2015). Therapeutic Manipulation of the Microbiome in
600 IBD: Current Results and Future Approaches. *Curr. Treat. Options Gastroenterol.* *13*,
601 105–120.
- 602 Hofbauer, J., and Sigmund, K. (1998). Evolutionary Games and Population Dynamics
603 (Cambridge University Press).
- 604 Honda, K., and Littman, D.R. (2012). The Microbiome in Infectious Disease and
605 Inflammation.
- 606 Ivanov, I.I., Atarashi, K., Manel, N., Brodie, E.L., Shima, T., Karaoz, U., Wei, D.,
607 Goldfarb, K.C., Santee, C.A., Lynch, S.V., et al. (2009). Induction of Intestinal Th17
608 Cells by Segmented Filamentous Bacteria. *Cell* *139*, 485–498.
- 609 Maldonado-Gómez, M.X., Martínez, I., Bottacini, F., O'Callaghan, A., Ventura, M., van
610 Sinderen, D., Hillmann, B., Vangay, P., Knights, D., Hutkins, R.W., et al. (2016). Stable
611 Engraftment of *Bifidobacterium longum* AH1206 in the Human Gut Depends on
612 Individualized Features of the Resident Microbiome. *Cell Host Microbe* *20*, 515–526.
- 613 Michor, F., and Beal, K. (2015). Improving Cancer Treatment via Mathematical
614 Modeling: Surmounting the Challenges Is Worth the Effort. *Cell* *163*, 1059–1063.

- 615 Morgan, X.C., Kabakchiev, B., Waldron, L., Tyler, A.D., Tickle, T.L., Milgrom, R.,
616 Stempak, J.M., Gevers, D., Xavier, R.J., Silverberg, M.S., et al. (2015). Associations
617 between host gene expression, the mucosal microbiome, and clinical outcome in the
618 pelvic pouch of patients with inflammatory bowel disease. *Genome Biol.* *16*, 67.
- 619 Narushima, S., Sugiura, Y., Oshima, K., Atarashi, K., Hattori, M., Suematsu, M., and
620 Honda, K. (2014). Characterization of the 17 strains of regulatory T cell-inducing
621 human-derived Clostridia. *Gut Microbes* *5*, 333–339.
- 622 van Nood, E., Vrieze, A., Nieuwdorp, M., Fuentes, S., Zoetendal, E.G., de Vos, W.M.,
623 Visser, C.E., Kuijper, E.J., Bartelsman, J.F.W.M., Tijssen, J.G.P., et al. (2013). Duodenal
624 Infusion of Donor Feces for Recurrent Clostridium difficile. *N. Engl. J. Med.* *368*, 407–
625 415.
- 626 Olle, B. (2013). Medicines from microbiota. *Nat. Biotechnol.* *31*, 309–315.
- 627 Omenetti, S., and Pizarro, T.T. (2015). The Treg/Th17 Axis: A Dynamic Balance
628 Regulated by the Gut Microbiome. *Front. Immunol.* *6*.
- 629 d'Onofrio, A. (2005). A general framework for modeling tumor-immune system
630 competition and immunotherapy: Mathematical analysis and biomedical inferences.
631 *Phys. Nonlinear Phenom.* *208*, 220–235.
- 632 Paun, A., Yau, C., and Danska, J.S. (2016). Immune recognition and response to the
633 intestinal microbiome in type 1 diabetes. *J. Autoimmun.* *71*, 10–18.
- 634 Round, J.L., and Mazmanian, S.K. (2010). Inducible Foxp3+ regulatory T-cell
635 development by a commensal bacterium of the intestinal microbiota. *Proc. Natl.*
636 *Acad. Sci. U. S. A.* *107*, 12204–12209.
- 637 Rudensky, A.Y. (2011). Regulatory T cells and Foxp3. *Immunol. Rev.* *241*, 260–268.
- 638 Schirmer, M., Smeekens, S.P., Vlamakis, H., Jaeger, M., Oosting, M., Franzosa, E.A.,
639 Jansen, T., Jacobs, L., Bonder, M.J., Kurilshikov, A., et al. (2016). Linking the Human
640 Gut Microbiome to Inflammatory Cytokine Production Capacity. *Cell* *167*, 1125–
641 1136.e8.
- 642 Smith, P.M., Howitt, M.R., Panikov, N., Michaud, M., Gallini, C.A., Bohlooly-Y, M.,
643 Glickman, J.N., and Garrett, W.S. (2013). The Microbial Metabolites, Short-Chain
644 Fatty Acids, Regulate Colonic Treg Cell Homeostasis. *Science* *341*, 569–573.
- 645 Smits, S.A., Marcabal, A., Higginbottom, S., Sonnenburg, J.L., and Kashyap, P.C.
646 (2016). Individualized Responses of Gut Microbiota to Dietary Intervention Modeled
647 in Humanized Mice. *MSystems* *1*, e00098-16.

- 648 Stein, R.R., Bucci, V., Toussaint, N.C., Buffie, C.G., Rätsch, G., Pamer, E.G., Sander, C.,
649 and Xavier, J.B. (2013). Ecological Modeling from Time-Series Inference: Insight into
650 Dynamics and Stability of Intestinal Microbiota. *PLoS Comput Biol* 9, e1003388.
- 651 Surana, N.K., and Kasper, D.L. (2017). Moving beyond microbiome-wide associations
652 to causal microbe identification. *Nature* 552, 244–247.
- 653 Tanoue, T., Atarashi, K., and Honda, K. (2016). Development and maintenance of
654 intestinal regulatory T cells. *Nat. Rev. Immunol.* 16, 295–309.
- 655 Vieira, A.T., Fukumori, C., and Ferreira, C.M. (2016). New insights into therapeutic
656 strategies for gut microbiota modulation in inflammatory diseases. *Clin. Transl.*
657 *Immunol.* 5, e87.
- 658 Vrieze, A., Van Nood, E., Holleman, F., Salojärvi, J., Koote, R.S., Bartelsman, J.F.W.M.,
659 Dallinga-Thie, G.M., Ackermans, M.T., Serlie, M.J., Oozeer, R., et al. (2012). Transfer of
660 intestinal microbiota from lean donors increases insulin sensitivity in individuals
661 with metabolic syndrome. *Gastroenterology* 143, 913-916.e7.
- 662 Wangersky, P.J. (1978). Lotka-Volterra Population Models. *Annu. Rev. Ecol. Syst.* 9,
663 189–218.
- 664 Weil, A.A., and Hohmann, E.L. (2015). Fecal Microbiota Transplant: Benefits and
665 Risks. *Open Forum Infect. Dis.* 2.

666
667
668
669
670
671
672
673
674
675
676

677 Figures

678
679
680

681 **Figure 1 (Conceptual figure):** A microbiome-immune system mathematical model
682 describing the activation of regulatory T-cells (T_{reg}) in response to colonization profiles of
683 T_{reg} -stimulating Clostridia strains is at the core of this work. It consists of a previously
684 derived microbiome ecological model that describes the short and long-term temporal
685 dynamics of Clostridia strains in germ-free mice (Bucci et al., 2016) and is supplemented
686 by a microbiome- T_{reg} model of $CD4^+FOXP3^+$ T_{reg} activation in response to long-term

687 compositions in the microbiome. Using long-term colonization data from mouse
688 experiments with subsets of these strains and corresponding measurement of T_{reg}
689 induction, the individual contribution of each strain to the measured T_{reg} activation is
690 inferred. The T_{reg} -induction score, TrIS, which accounts for ecological stability and
691 immune activation assigns a score to every possible strain combination and thereby
692 identifies candidate probiotic consortia for experimental validation.

693

694

695

696 **Figure 2 (Data used for inference of the CD4⁺FOXP3⁺ T_{reg} induction parameters):**
697 (A) To infer the strain-resolved CD4⁺FOXP3⁺ T_{reg} induction parameters CD4⁺FOXP3⁺
698 T_{reg} abundance measurements and corresponding microbiome colonization data are
699 used. CD4⁺FOXP3⁺ T_{reg} data of single strains originates from previously published
700 measurements from (Atarashi et al., 2013). Because in (Atarashi et al., 2013) microbial
701 mono-colonization levels were not measured, a previously published predictive model
702 was used (Bucci et al., 2016) to estimate each strain's microbial abundance in the
703 corresponding mono-colonization experiments. In addition, newly generated
704 CD4⁺FOXP3⁺ T_{reg} and colonization mouse stool data from three 11-strain combinations
705 was included to the analysis (see also Figure 3). These 11-strain combinations were
706 chosen according to the results of the 'Keystoneness' analysis previously described by
707 some of us (Bucci et al., 2016). The microbiome compositions of these three 11-strain
708 combinations were estimated by strain-specific qPCR. (B) The resulting CD4⁺FOXP3⁺
709 T_{reg} induction parameters quantify the contribution of each individual strain to T_{reg}
710 induction. (Coefficients are scaled by the microbial mono-colonization concentrations for
711 display reasons.).

712

713

714

715 **Figure 3 (Validation of microbiome model describing the dynamics of Clostridia**
716 **strains in germ-free mice):** (A) The model-predicted dynamics of three different 11-
717 strain combinations from the thirteen Clostridia strains described in (Bucci et al., 2016)
718 (areas) are compared to measured data from germ-free mice (stacked bars) where each
719 panel corresponds to an individual mouse (see above, Figure 2A). Predictions were
720 obtained by numerically solving the corresponding generalized Lotka–Volterra equations
721 with parameters taken from (Bucci et al., 2016) using only each mouse's initial microbial
722 composition. In timelines/mouse 1–5, strain 15 is absent referring to 11-strain set I (5
723 biological replicates); in timelines 6–9, strain 4 is absent (11-strain set II; 4 biological
724 replicates), and in timelines 10–14, strain 29 is absent (11-strain set III; 5 biological
725 replicates). Data were obtained by qPCR of genes specific to each strain. Densities are
726 computed as averages of three technical measurements. The number of mice used in
727 each condition was chosen consistently with previous experimental work (Atarashi et al.,
728 2013; Bucci et al., 2016) and combined with extensive *in silico* testing of the inference
729 error as a function of sampling frequency and noise levels (Bucci et al., 2016). (B)
730 Spearman's rank correlation coefficient between observed and predicted data was
731 computed at different time points and plotted against time. All displayed coefficients
732 have a p-value of less than 10⁻¹⁶.

733

734
735
736 **Figure 4 (Identification and experimental validation of modeling-predicted T_{reg} -**
737 **stimulating consortia):** (A) Predicted T_{reg} Induction Score (TrIS) in germ-free mice as a
738 function of probiotic consortium size. Each data point represents one of the $2^N - 1$
739 possible non-trivial consortia of size N . The five highest and lowest TrIS consortia of size
740 seven are highlighted by violet and cyan filled dots, respectively, as well as, the size-4
741 consortia used for experimental validation in subpanel E. (B) Distribution of the TrIS
742 plotted against the total population density of each model-predicted consortium. This
743 analysis shows a strong correlation (Spearman's Rank Correlation $p < 0.05$) between TrIS
744 and total microbial density. Color of the dots encode for the abundance of Strain 27
745 which is predicted to strongly colonize in high-TrIS subsets. (C) Heatmap of consortium
746 membership of the five highest and lowest TrIS-consortia of size seven, and their
747 predicted T_{reg} induction and estimated butyrate production using data from (Atarashi et
748 al., 2013). (D) Estimation of potential SCFA output based on single-strain *in vivo*
749 metabolic profiling. Single-strain SCFA measurements were taken from (Atarashi et al.,
750 2013). The five highest TrIS consortia of size seven (from 4C) are predicted to have a
751 significantly higher SCFA output compared to the five lowest ones. (E) Five consortia of
752 size 4 were used for experimental validation: the two highest-ranked consortia of size
753 four (H1 and H2, respectively), the lowest-ranked consortium (L) and two intermediates
754 (M1, M2). M1 and M2 were included because of interest in other disease areas. The
755 experimentally introduced strains are listed next to each bar/consortium. Strains with
756 numbers in black were detected by 16S rRNA sequencing, strains with numbers in gray
757 were introduced but failed to colonize. A Pearson's correlation of 0.97 between the
758 predicted TrIS and the average of each consortium's measured CD4 $^{+}$ FOXP3 $^{+}$ T_{reg}
759 percentage shows the ability of the TrIS-based selection approach to correctly recover
760 the experimentally observed T_{reg} induction. (Pearson's correlation of all points has a
761 value of 0.54.) Importantly, the H2 consortium displays an average increase in immune
762 activity of 107% relative to the average germ-free mouse control. Eight biological
763 replicates were used for GF, H1 and H2, five biological replicates were used for M1, M2
764 and three biological replicates were used for L. Replication and design of the validation
765 experiment for T_{reg} induction assessment is consistent with work from us and others
766 (Atarashi et al., 2013; Narushima et al., 2014).

767
768
769
770 **Figure 4 – figure supplement 1 (TrIS versus estimated mean population density):**
771 In analogy to Figure 4B which highlights only the abundance of strain 27, the predicted
772 total population density of the microbial consortia is here compared to the TrIS and the
773 color in each subpanel represents the corresponding strain's abundance.

774
775
776

777 **Figure 4 – figure supplement 2 (Metabolic profiling of three selected microbial**
778 **consortia):** Metabolic profiling of fecal content of selected consortia with respect to
779 acetic acid.

780

781

782

783 **Figure 4 – figure supplement 3 (Comparison of predicted and measured metabolic**
784 **profiles of three selected microbial consortia):** Mean germ-free normalized predicted
785 acetate level in caecal content using single-strain measurements from (Atarashi et al.,
786 2013) plotted against the observed mean normalized acetic acid levels in mouse stool
787 from (Figure 4 – figure supplement 2).

788

789

790

791

792

793 **Tables**

794

	Atarash i et al. 2013	Bucci et al. 2016	11- Strain I	11- Strain II	11- Strain III	4- strain H1	4- strain H2	4- strain M1	4- strain M2	4- strain L
Strain 1	x									
Strain 3	x									
Strain 4	x	x	x		x		x	x		
Strain 6	x	x								
Strain 7	x	x	x	x	x	x	x	x		
Strain 8	x									
Strain 9	x	x	x	x	x				x	
Strain 13	x	x	x	x	x					

Strain 14	x	x	x	x	x			x		x
Strain 15	x	x		x	x					x
Strain 16	x	x	x	x	x			x	x	
Strain 18	x									
Strain 21	x	x	x	x	x					
Strain 26	x	x	x	x	x					
Strain 27	x	x	x	x	x	x	x	x		
Strain 28	x	x	x	x	x	x		x		
Strain 29	x	x	x	x		x	x		x	x

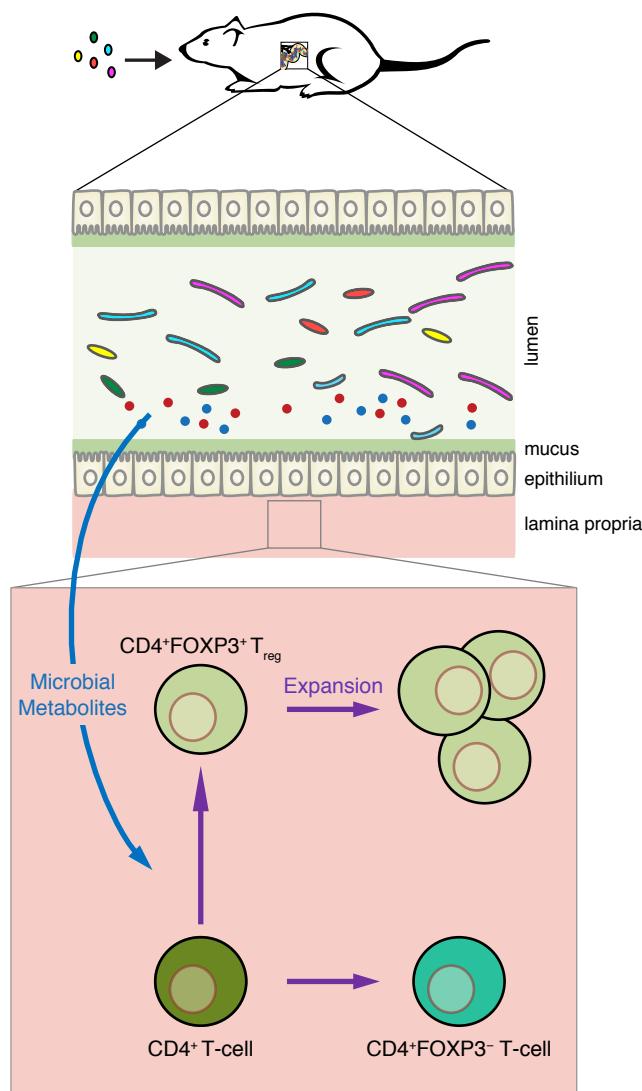
795

796 **Table 1: List of strains used with respect to our previous published work.** The
797 original publication (Atarashi et al., 2013) reported on seventeen immune-modulating
798 bacteria. In (Bucci et al., 2016), a dynamical ecological model for thirteen of the original
799 seventeen strains was presented. These thirteen strains were selected because they do
800 not harbor antibiotic resistance genes. For our study, we used a 12-strain subset of
801 these thirteen strains based on ecological stability considerations. Colonization and
802 CD4⁺FOXP3⁺ (T_{reg}) data for the three 11-strain sets, as measured by qPCR and FACS,
803 were used along with the mono-colonization experiment (Figure 2) where the
804 corresponding model simulations for estimating mono-colonization bacterial densities
805 were used to infer the microbiome-immune system model (Figure 1). The resulting
806 parameters were employed to predict ecologically stable subsets with different T_{reg}
807 induction potential. 4-strain subsets were selected to validate the mathematical modeling
808 predictions (see main text and Figure 3). Strain 1: *Clostridium saccharogumia* /
809 *Clostridium ramosum* JCM1298, Strain 3: *Flavonibacter plautii* / *Pseudoflavonifractor*
810 *capillosus*, Strain 4: *Clostridium hathewayi* / *Clostridium saccharolyticum* WM1, Strain 6:
811 *Blautia coccoides* / *Lachnospiraceae bacterium*, Strain 7: *Clostridium bolteae*, Strain 8:
812 *Clostridium* sp. MLG055 / *Erysipelotrichaceae bacterium* 2 44A, Strain 9:
813 *Clostridium indolis* / *Anaerostipes caccae* DSM 14662, Strain 13: *Anaerotruncus*
814 *colihominis*, Strain 14: *Ruminococcus* sp. ID8/*Lachnospiraceae bacterium* 2 1 46FAA,
815 Strain 15: *Clostridium asparagiforme* / *Clostridium lavalense*, Strain 16: *Clostridium*
816 *symbiosum*, Strain 18: *Clostridium ramosum*, Strain 21: *Eubacterium fissicatena* /
817 *Eubacterium contortum* / *Clostridium* sp. D5, Strain 26: *Clostridium scindens* /
818 *Lachnospiraceae bacterium* 5 1 57FAA, Strain 27: *Lachnospiraceae bacterium* A4 /

819 *Lachno bacterium* 3 1 57FAA CT1, Strain 28: *Clostridium* sp. 316002/08, Strain 29:
820 *Lacnospiraceae bacterium* A4 / *Lachno bacterium* 3 1 57FAA CT1.

821

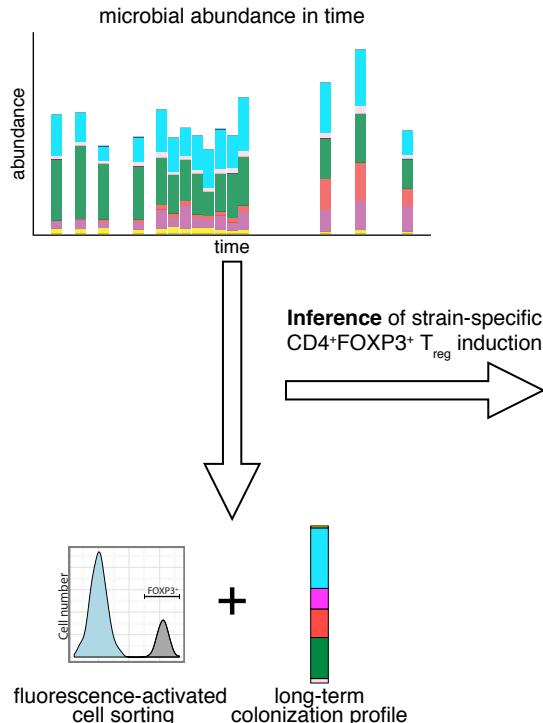
Mouse microbiome and regulatory T-cell data



Microbiome–immune system mathematical model

Microbiome ecological model

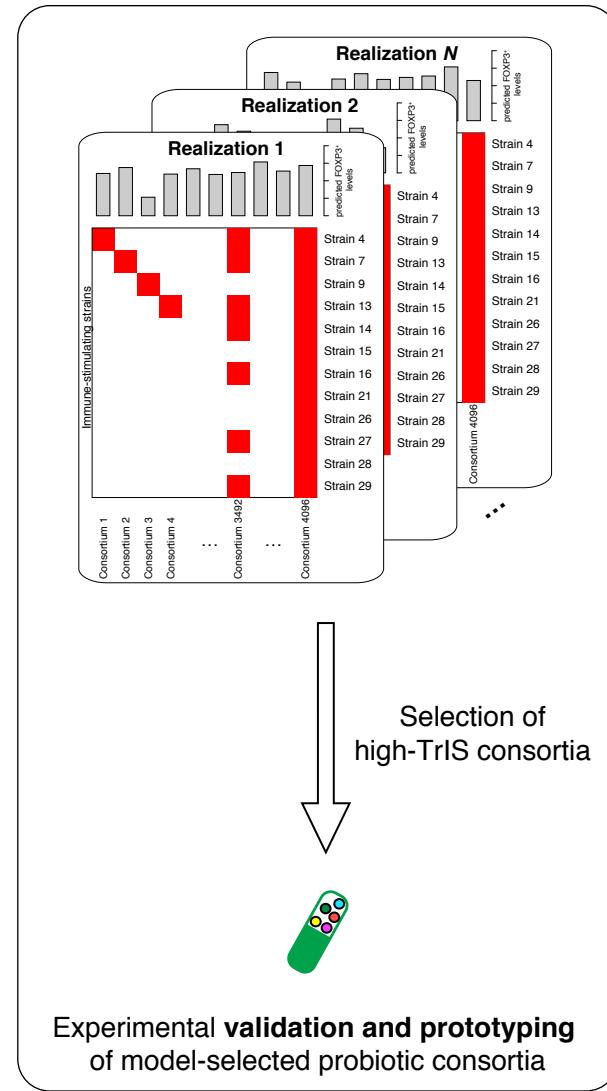
$$\frac{dc_{\text{strain}_i}(t)}{dt} = c_{\text{strain}_i}(t) \left(\alpha_i + \sum_{j=1}^L \beta_{ij} c_{\text{strain}_j}(t) \right)$$



Microbiome–T_{reg} model

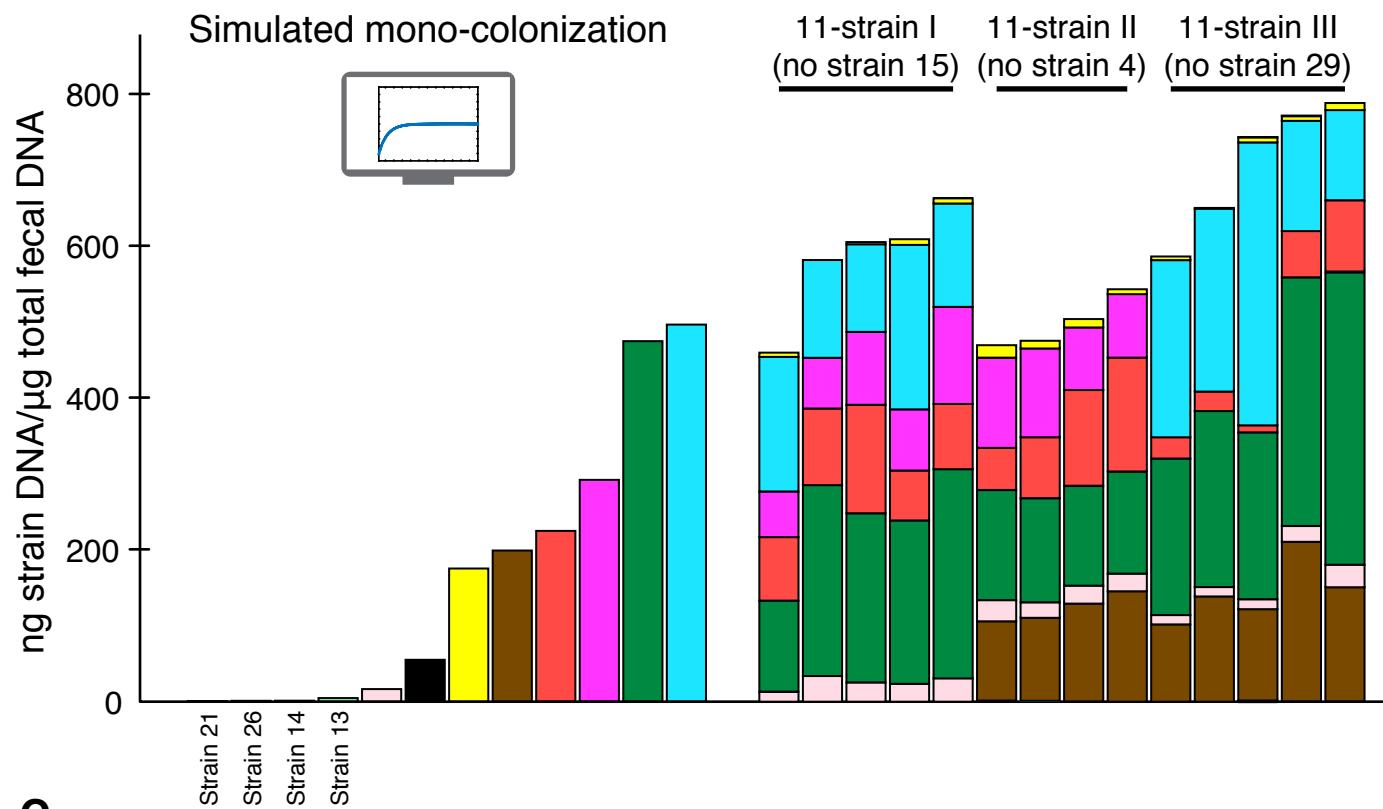
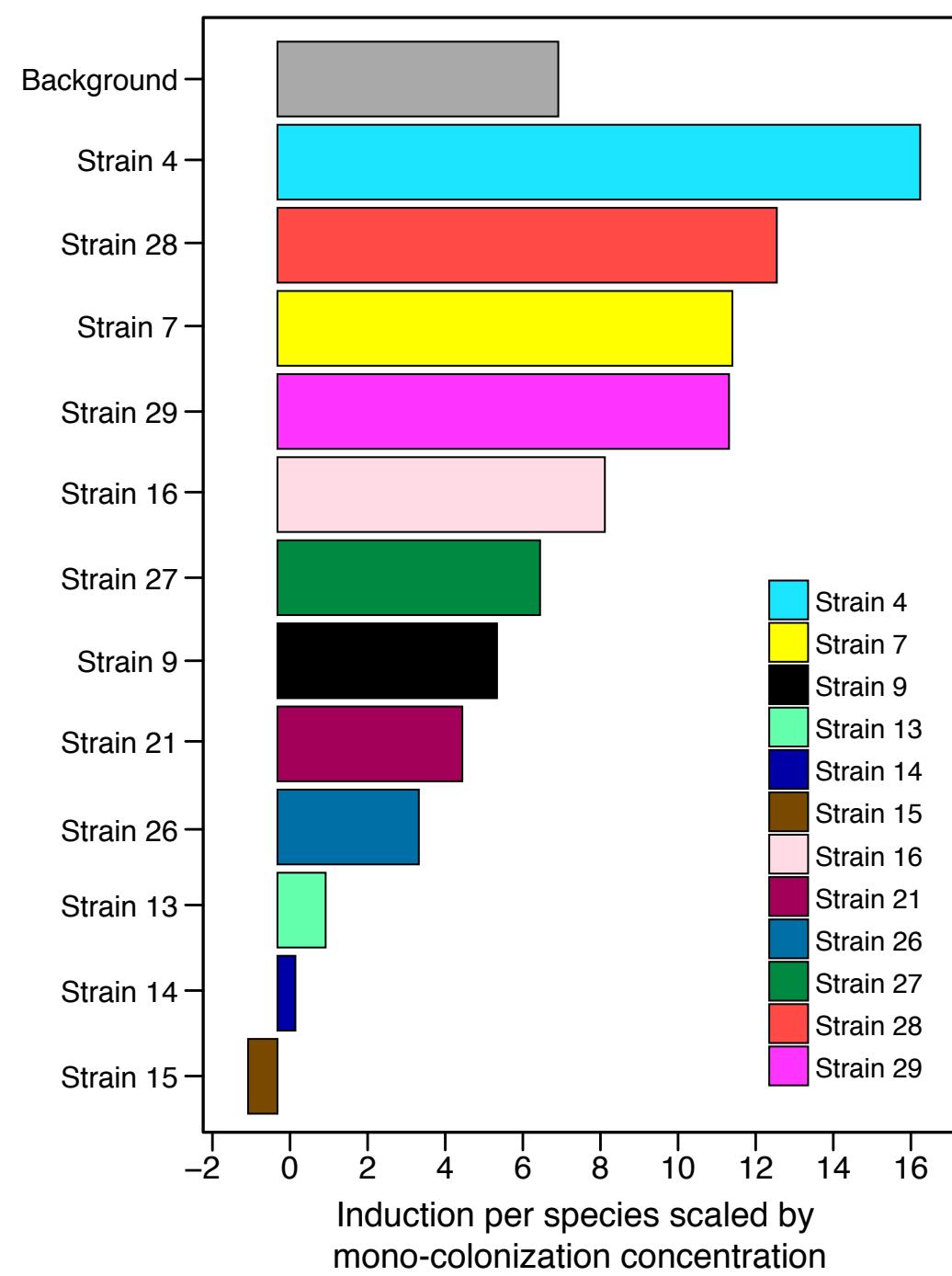
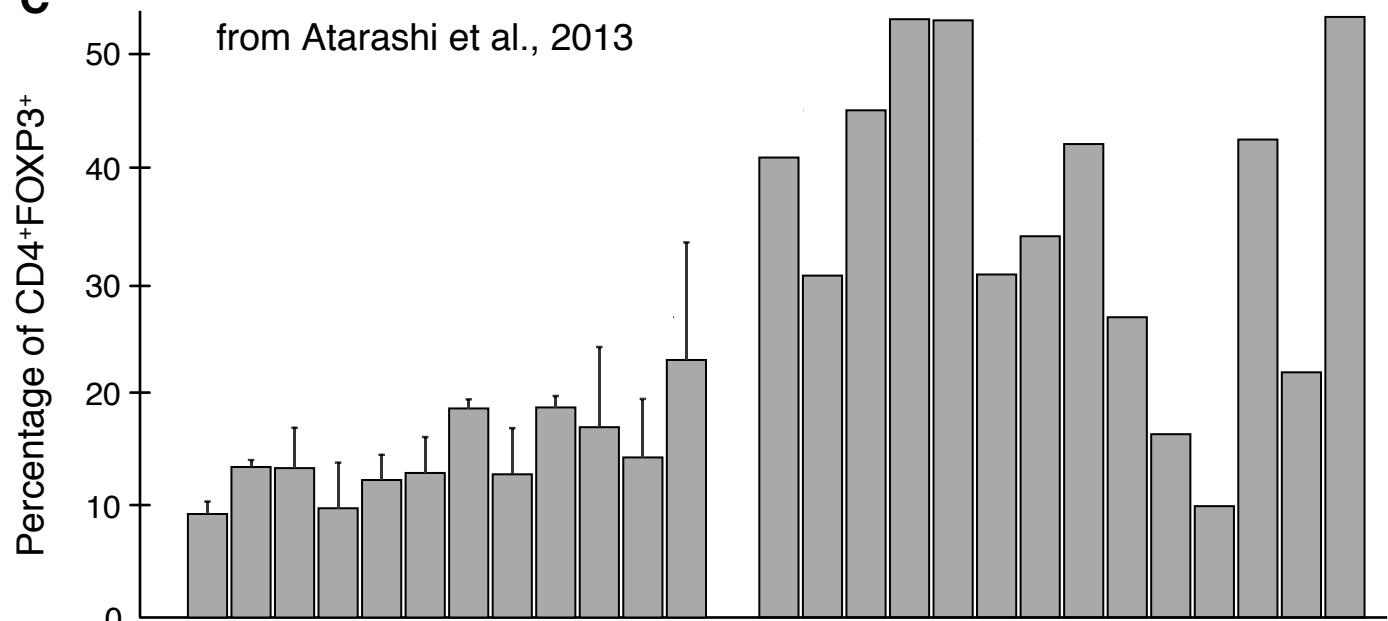
$$c_{\text{FOXP3}^+,ss} = \alpha(c_{\text{T}_{\text{reg}},ss}) + \sum_{i_k} \varepsilon_{i_k} c_{\text{strain}_{i_k},ss}$$

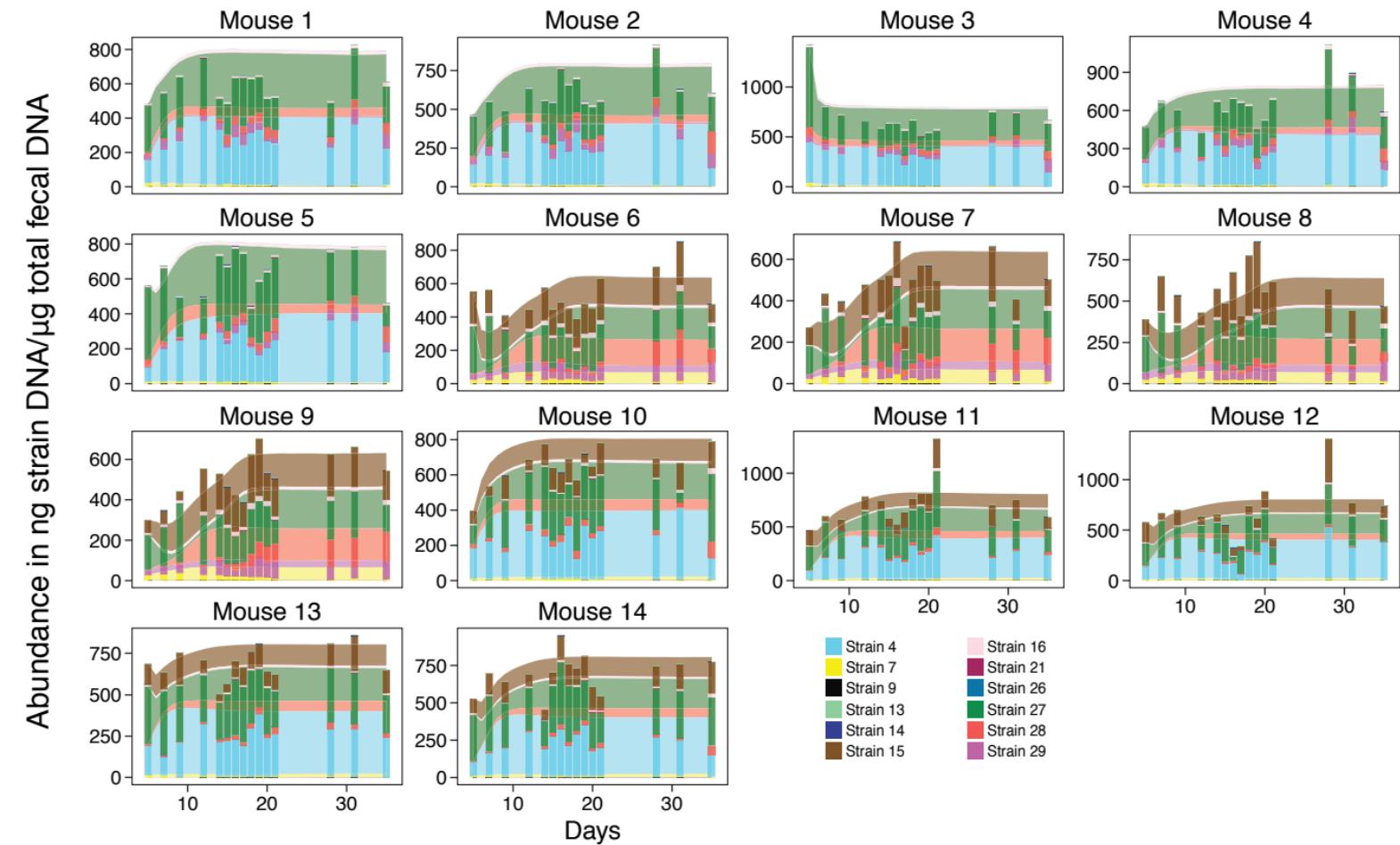
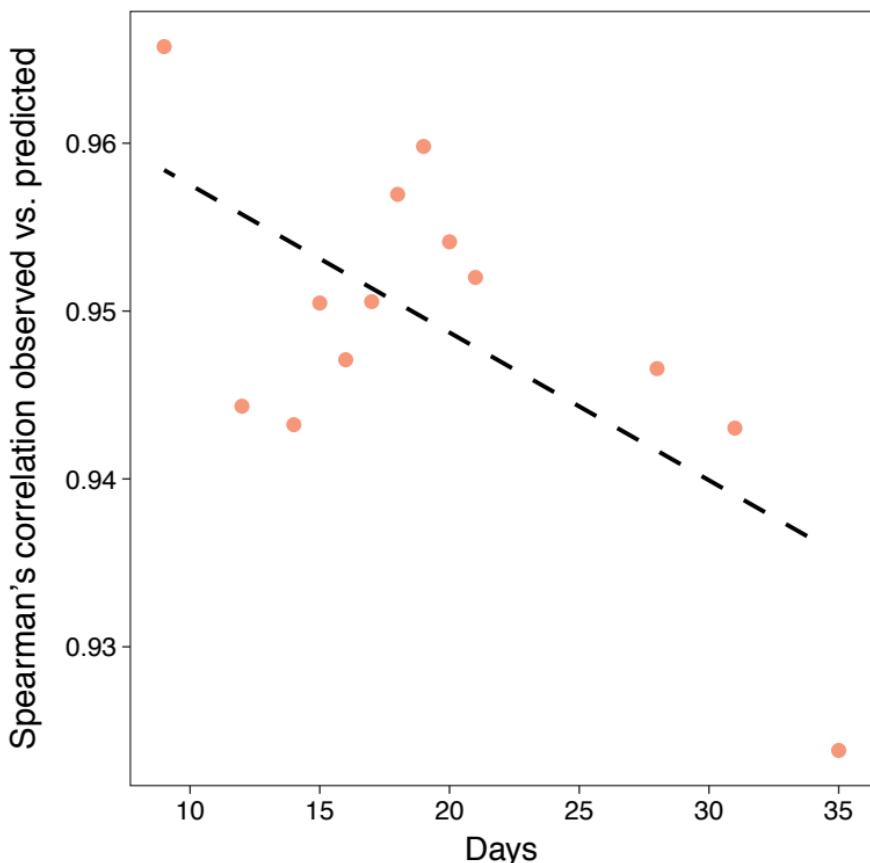
In silico selection of probiotic consortia

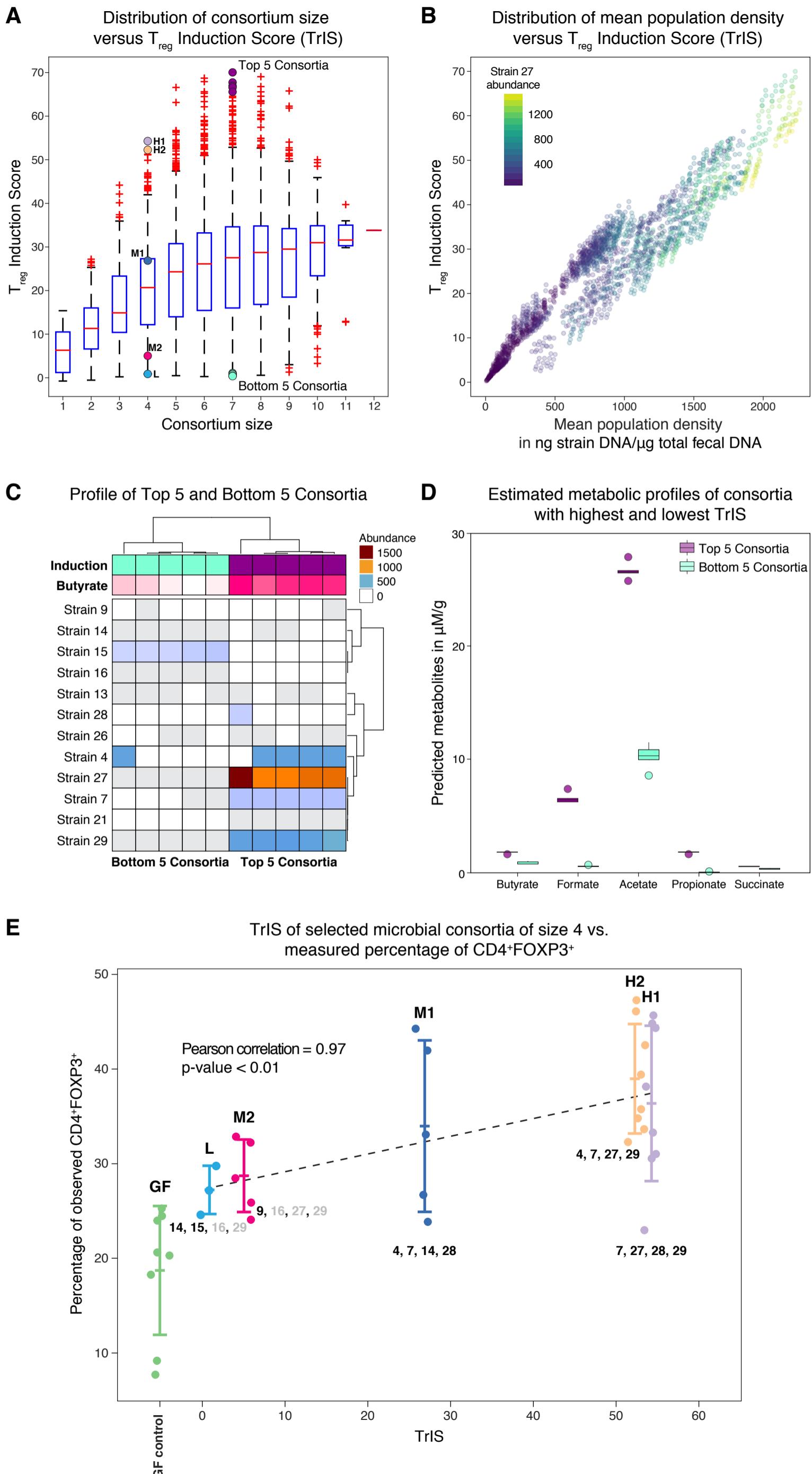


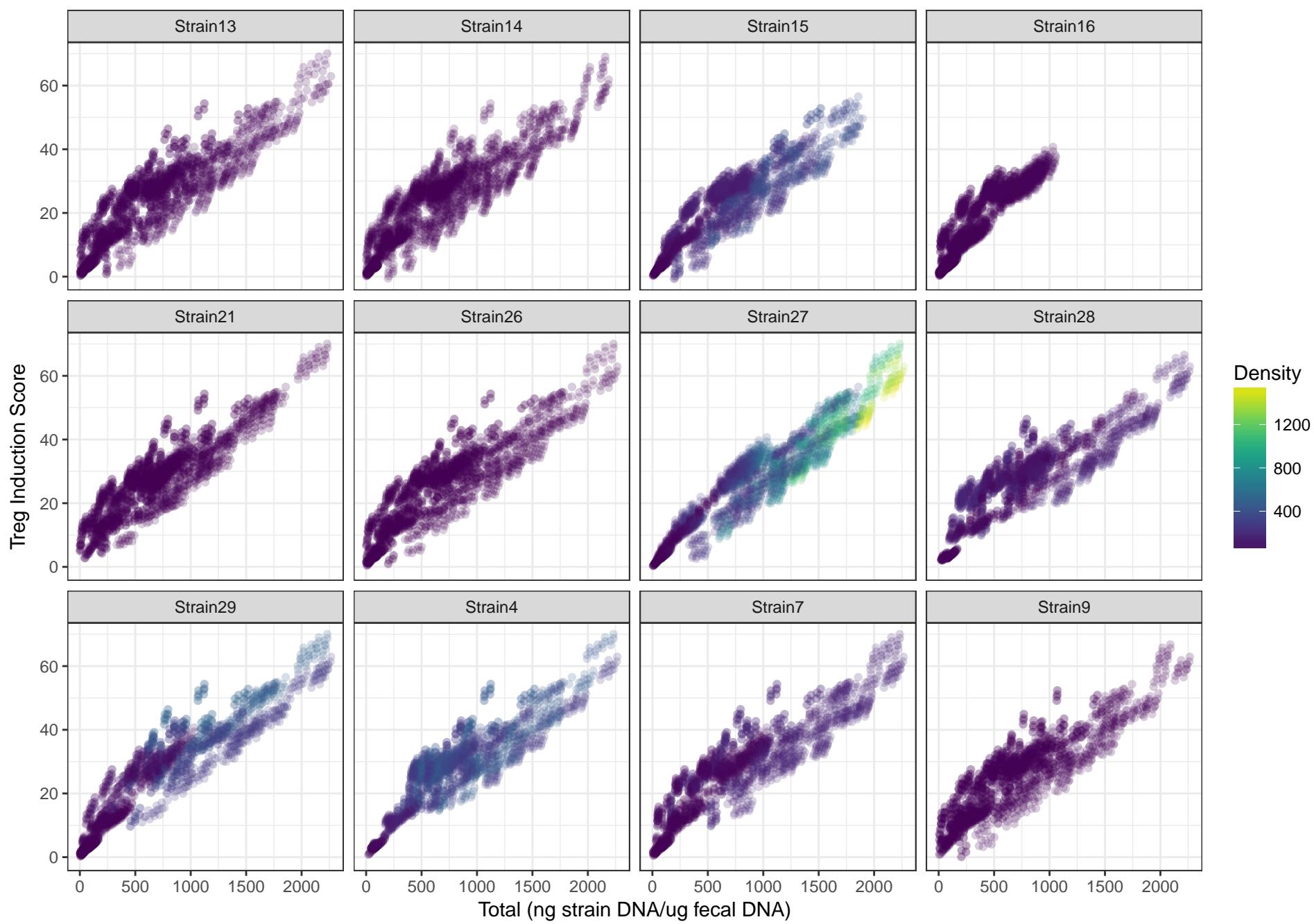
A

Germ-free mouse model

**B**Scaled CD4⁺FOXP3⁺ T_{reg} induction parameters**C**

A**B**





Measured acetic acid concentration in $\mu\text{g/g}$

H1

H2

L

