

1 **Article title:** Framework for rational donor selection in fecal microbiota transplant clinical trials

2 **Short title:** Rational donor selection for FMT clinical trials

3 **Authors:**

4 Claire Duvalliet, PhD

5 Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA
6 02139

7 Center for Microbiome Informatics and Therapeutics, Cambridge, MA 02139

8 Caroline Zellmer, BS

9 OpenBiome, Cambridge, MA 02140

10 Pratik Panchal, MD MPH

11 OpenBiome, Cambridge, MA 02140

12 Shrish Budree, MD MBCHB DCH FCPEDS

13 OpenBiome, Cambridge, MA 02140

14 Majdi Osman, MD MPH

15 OpenBiome, Cambridge, MA 02140

16 Harvard Medical School, MA 02115

17 Eric Alm, PhD (corresponding author)

18 Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA
19 02139

20 Center for Microbiome Informatics and Therapeutics, Cambridge, MA 02139

21 The Broad Institute of MIT and Harvard, Cambridge, MA 02139

22 Email: ejalm@mit.edu

23 Address: 500 Technology Square, NE47-379

24 Cambridge, MA 02139

25 **Conflicts of interest:**

26 E.A. is co-founder of Finch therapeutics. C.Z, P.P, S.B, and M.O. are employed by OpenBiome.

27 **Framework for rational donor selection in fecal microbiota transplant clinical trials**

28 Claire Duvallat, Caroline Zellmer, Pratik Panchal, Shrish Budree, Majdi Osman, Eric Alm

29 **Abstract**

30 Early clinical successes are driving enthusiasm for fecal microbiota transplantation (FMT), the
31 transfer of healthy gut bacteria through whole stool, as emerging research is linking the
32 microbiome to many different diseases. However, preliminary trials have yielded mixed results
33 and suggest that heterogeneity in donor stool may play a role in patient response. Thus, clinical
34 trials may fail because an ineffective donor was chosen rather than because FMT is not
35 appropriate for the indication. Here, we describe a conceptual framework to guide rational donor
36 selection to increase the likelihood that FMT clinical trials will succeed. We argue that the
37 mechanism by which the microbiome is hypothesized to be associated with a given indication
38 should inform how donors are selected for FMT trials, categorizing these mechanisms into four
39 disease models and presenting associated donor selection strategies. We next walk through
40 examples based on previously published FMT trials and ongoing investigations to illustrate how
41 donor selection might occur in practice. Finally, we show that typical FMT trials are not powered
42 to discover individual taxa mediating patient responses, suggesting that clinicians should
43 develop targeted hypotheses for retrospective analyses and design their clinical trials
44 accordingly. Moving forward, developing and applying novel clinical trial design methodologies
45 like rational donor selection will be necessary to ensure that FMT successfully translates into
46 clinical impact.

47 **Introduction**

48 Fecal microbiota transplantation (FMT) is the transfer of gut bacteria through whole stool from a
49 healthy donor to a recipient. FMT has demonstrated high cure rates in recurrent *C. difficile*
50 infection (CDI) across multiple randomized, placebo-controlled trials (Quraishi et al. 2017) and
51 has now entered standard of care for recurrent CDI in European and North American guidelines
52 (McDonald et al. 2018; Cammarota et al. 2017; Surawicz et al. 2013). Beyond CDI, FMT is
53 being explored in range of microbiome-mediated diseases, and has demonstrated promising
54 results in inflammatory bowel diseases (Panchal et al. 2018; Gelfand 2018; Kootte et al. 2017;
55 Osman 2018; Costello et al. 2017; Paramsothy et al. 2017).

56 Despite these early successes, the underlying mechanism of FMT across all disease
57 indications, including CDI, remains unclear. However, it is generally considered that FMT
58 restores gut microbial community perturbations from a dysbiotic state to a healthy stable state
59 with engraftment of donor strains, or perhaps through other donor-dependent features such as
60 the abundance of non-bacterial components or donor clinical features (Ott et al. 2017; Zuo et al.
61 2018). However, not all FMT donors are alike: gut microbiota compositions vary within healthy
62 populations in ways that could impact the findings from an FMT trial (Yatsunencko et al. 2012,
63 Wilson et al. 2019). This critical point of donor microbiome variation is rarely considered in the
64 development of FMT trials (Bafeta et al. 2017; Olesen et al. 2018).

65 Unlike FMT trials in CDI, where selecting donors based on specific clinical or microbiome
66 profiles does not seem to affect clinical response rates, donor selection is likely to be crucial to
67 trial outcomes in diseases with more complex host-microbiome interplay or distinct
68 disease-associated perturbations (Wilson et al. 2019). Most notably, in a randomized controlled

69 trial (RCT) of FMT for ulcerative colitis (UC) using 5 donors, 78% of patients who achieved
70 remission after FMT received stool from a single donor (Moayyedi et al. 2015). Thus, it is
71 possible that without this single donor, the trial would have returned a negative result. Given the
72 variation in donor microbiomes and donors' potential impact on clinical efficacy, how should
73 clinicians and investigators select their donors for a clinical trial?

74 To date, the typical approach for donor selection in FMT trials is to use a single healthy donor or
75 to randomly select multiple donors from a set of screened potential donors (Paramsothy et al.
76 2017; Kelly et al. 2016; van Nood et al. 2013). However, in clinical indications where successful
77 donors may be rare, such as UC, clinical trials with randomly-selected donors may fail not
78 because FMT is inappropriate for the indication, but because an ineffective donor was chosen.
79 An alternative approach is to expose each patient to multiple donors in order to mitigate the risk
80 of sub-optimal donor selection. In a large RCT of FMT in UC, FMT enemas for a single patient
81 were derived from between three and seven donors with patients receiving multiple donors
82 throughout the 8 week course of treatment (Paramsothy et al. 2017). However, using multiple
83 donors for a single patient may not be feasible or appropriate in many disease indications or
84 clinical trial settings (e.g. single-dose FMT studies). Continuing the practice of randomly
85 selecting donors for FMT clinical trials risks returning false negative trials, stalling the field and
86 delaying the development of novel therapies for microbiome-mediated conditions.

87 Unlike traditional clinical trials which test well-defined small molecules FMT trials test the donor
88 microbiome, which is variable (Olesen et al 2018). Fortunately, the emergence of large,
89 multi-donor stool banks captures some of this variability and makes it available for use in FMT
90 trials. These stool banks thus open the possibility of selecting donors rationally during the FMT

91 clinical trial design phase, enabling clinicians to choose from among a large pool of eligible
92 donors for donor samples which have specific desirable characteristics. Coupled with expanded
93 access to genome sequencing technologies and publicly available microbiome sequencing
94 datasets, rational donor selection is feasible and presents a unique opportunity to advance the
95 research methods of this nascent field.

96 In this paper, we present a framework to guide donor selection for FMT trials. The mechanism
97 by which the microbiome is hypothesized to be associated with a given indication should inform
98 how donors are selected for FMT trials, and we describe different disease models which may
99 underlie microbiome-mediated conditions (Figure 1). We describe strategies to rationally select
100 donors for each type of disease model, and provide examples based on previously published
101 FMT trials and ongoing investigations. Finally, we discuss limitations of performing
102 discovery-based retrospective research after an FMT clinical trial concludes. To our knowledge,
103 this is the first description of a comprehensive framework for rational donor selection in FMT
104 trials.

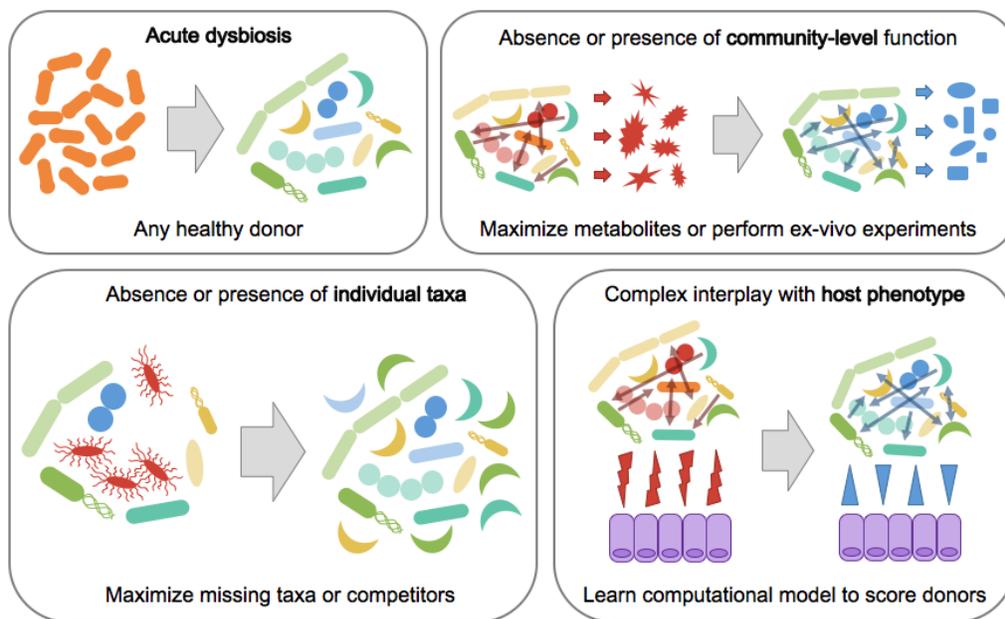
105 **Models of microbiome-mediated disease**

106 FMT trials are pursued when research or clinical experiences suggest that the microbiome may
107 be causing or exacerbating a disease. Here, we propose four different models which may
108 underlie microbiome-mediated etiologies and their corresponding rational donor selection
109 strategies (Figure 1). Ultimately, it is up to each individual clinician-researcher to use published
110 cross-sectional studies, mechanistic investigations in model organisms, and their own clinical
111 experience treating patients to determine which of these model(s) are relevant in their specific

112 case. Additionally, logistical considerations will be important factors in making the final donor
113 selection regardless of which strategy is pursued. For example, clinicians should ensure that the
114 pool of donors that they are screening have enough material to sustain the required number of
115 FMTs for their entire trial.

116 Most of the donor selection strategies described below can be modified to incorporate matching
117 between patients and donors. More specifically, donors can be tailored to individual patients to
118 specifically make up for the unique taxonomic or functional deficiencies in that patient's
119 microbiome (Wilson et al 2019). With the increasing amount of microbiome data available from
120 published FMT trials, we encourage collaborations between clinicians and bioinformaticians to
121 analyze these data in order to generate or perhaps even confirm the validity of potential donor
122 selection strategies before selecting one (Figure 2). Finally, the strategies presented here
123 should also be combined with adaptive clinical trial designs to further increase the probability of
124 having a successful FMT trial (Olesen, Gurry, and Alm 2017).

125 **Fig 1: overview of the different models of microbiome-mediated disease and associated**
126 **donor selection strategy.** In cases where the underlying model is unknown, a variety of donor
127 selection approaches could be employed to potentially identify which disease process(es) may
128 be involved.



129 **Acute dysbiosis**

130 An acutely dysbiotic gut microbial community is broadly dysfunctional and can no longer
131 maintain the health of the host. For example, in the case of recurrent *Clostridium difficile*
132 infection (rCDI), a disturbed microbial community is unable to prevent colonization by or
133 overgrowth of the pathogen, leading to recurrent overgrowth of *C. diff* and clinical symptoms
134 (Britton and Young 2014). Acute dysbiosis has also been described with the “Anna Karenina
135 principle”: all healthy microbiomes are alike but dysbiotic communities are all dysbiotic in their
136 own ways (Zaneveld et al 2017). In this view of acute dysbiosis, microbial communities respond
137 stochastically to stressors, resulting in dysbiotic communities which are characterized by
138 increased variability rather than deterministic shifts to precise community type(s) (Zaneveld et al
139 2017).

140 In this model, the host simply needs to return to a “healthy” microbiome and thus choosing any
141 healthy FMT donor should be sufficient to induce clinical improvements. Because there is no

142 specific disease-associated microbial community and deviation from health is instead the more
143 important factor, simply replenishing the microbiome with a healthy configuration should be
144 sufficient. Indeed, FMT trials have demonstrated that rCDI can be effectively treated by almost
145 any choice of healthy donor (Osman et al. 2016). In this case, researchers should consider how
146 they define a “healthy” microbiome and how they will ensure clinical efficacy, for example
147 through engraftment of the transplanted healthy communities.

148 **Absence or presence of individual taxa**

149 Absence of beneficial taxa

150 In other cases, perhaps a disease is being caused or exacerbated by the lack of certain specific
151 microbes, and replenishing these few taxa would be sufficient to restore the host to health
152 (Wilson et al 2019). For example, Hsiao et al showed that a single microbe, *R. obeum*, restricted
153 infection by *V. cholerae* through quorum-sensing-mediated mechanisms (Hsiao et al. 2014).
154 Surprisingly, non-communicable diseases may also fall into this model: a single strain of
155 *Lactobacillus* was sufficient to ameliorate salt-induced hypertension in mice, and follow-up
156 studies indicate that similar mechanisms may be involved in salt-sensitive high blood pressure
157 in humans as well (Wilck et al. 2017).

158 In these cases, the donor selection strategy should focus on maximizing the probability of
159 engraftment of the beneficial taxa. In cases where the unique taxa are not specifically known or
160 are rare members of the human microbiota, many healthy donors should be pooled together or
161 a donor with a high alpha diversity could be selected in order to maximize the probability that the
162 transplanted sample contains the necessary taxa (Wilson et al 2019). If the missing microbes

163 are known and well-characterized, on the other hand, researchers can screen their pool of
164 potential donors to find the sample with the highest abundance of these taxa.

165 Presence of harmful taxa

166 Rather than being characterized by the absence of individual bacteria, perhaps a disease is
167 instead mediated by the presence or overabundance of specific microbes, and removing these
168 bacteria in a targeted fashion could lead to improvements in disease progression. For example,
169 *Fusobacterium* has been found to be more abundant in colorectal cancer patients, specifically
170 enriched in the tumors themselves (Kostic et al. 2013). Multiple groups have identified
171 mechanistic associations between *Fusobacterium*, inflammatory transcriptional signatures, and
172 tumor growth in mouse and human models of colorectal cancer, pointing to a causal role for
173 *Fusobacterium* in colorectal cancer progression (Kostic et al. 2013; Rubinstein et al. 2012).
174 Recent work has found that treating tumors with antibiotics slows tumor progression, further
175 confirming these causal associations and pointing toward potential microbiome-based
176 therapeutic interventions (Bullman et al. 2017).

177 Removing and replacing these bacteria should be the goal of FMT in cases where this disease
178 model applies. This can be achieved by first removing the harmful bacteria (e.g. via antibiotic
179 treatment) with follow-up FMT to re-establish a healthy community that prevents their
180 re-colonization. In all cases, donors should be screened to exclude any samples which contain
181 the harmful bacteria. Donor samples can then be selected based on the abundance of bacteria
182 which are known to out-compete the harmful taxa. Competitors can be identified by searching
183 the microbiology literature to identify bacteria which live in the same niche or which have been
184 experimentally shown to directly out-compete the undesirable taxa, or they can perform these

185 competition assays themselves. If resources to perform competition assays are not available
186 and the literature is sparse, researchers can also mine existing microbiome data to find bacteria
187 which consistently anti-correlate with the harmful taxa, and choose donor samples with a high
188 abundance of these putative competitors (Friedman and Alm 2012).

189 Patient matching

190 Taxa-based donor selection strategies are particularly amenable to patient-matching, when both
191 patient and donor microbiome data are available prior to the start of a trial (Wilson et al 2019).
192 For example, if one patient is completely missing some of the beneficial taxa but not others,
193 then these taxa can be weighted more heavily in the donor selection process. The phylogenetic
194 relationships between donor and recipient taxa could also be incorporated into donor selection:
195 if a patient already has many bacteria which are closely phylogenetically related to known
196 competitors of some of the harmful bacteria, then competitors of the other harmful bacteria can
197 be upweighted in the donor selection process. Similarly, if patients already have taxa which are
198 already filling certain niches important for health, the taxa which fill those same niches can be
199 downweighted in donor selection.

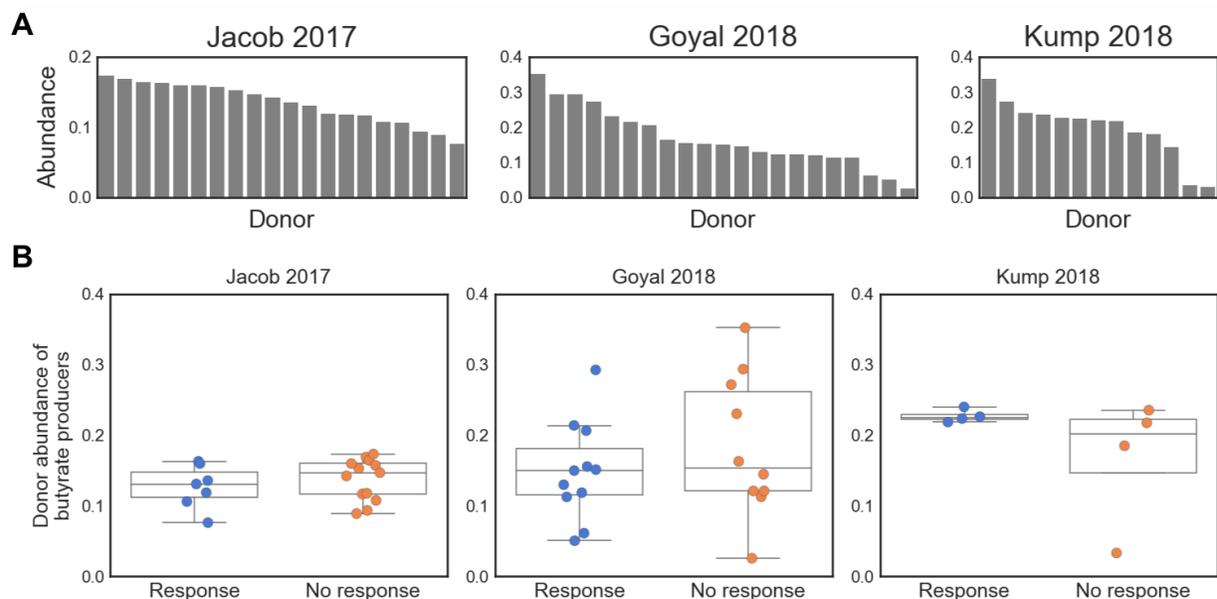
200 Case study: Inflammatory Bowel Disease

201 An example where the “missing taxa” model may be applicable is in inflammatory bowel disease
202 (IBD). Butyrate has long been associated with IBD (Wilson et al 2019, Scheppach et al. 1992),
203 and recent case-control and longitudinal studies point to a consistent lack of butyrate-producing
204 bacteria in patients with IBD (Duvall et al. 2017; Schirmer et al. 2018). Furthermore,
205 preliminary FMT trials in IBD have been marked by variable efficacy, both between trials and
206 between donors within individual trials, suggesting that donor microbiome characteristics may

207 be associated with FMT response (Kump et al. 2018; Moayyedi et al. 2015). These results
208 indicate that IBD may benefit from rational donor selection approach, and that donors with high
209 abundances of butyrate-producing organisms may yield higher FMT response rates than
210 randomly selected donors.

211 Given the availability of microbiome data from completed FMT studies, we tested this
212 hypothesis that IBD trials would benefit from a rational donor selection strategy based on the
213 “absence of beneficial taxa” disease model. We re-analyzed microbiome data from three
214 completed IBD FMT trials which provided publicly available sequencing data for patient and
215 donor samples (Kump et al. 2018; Goyal et al. 2018; Jacob et al. 2017). We selected
216 butyrate-producers based on their genus-level taxonomy, using the genera identified in Vital et
217 al. 2017 (see Methods). Donors in the three studies exhibited a range of total abundances of
218 butyrate-producing bacteria (Figure 2A). Surprisingly, however, the abundance of butyrate
219 producers in the donor stool was not associated with recipient patients’ clinical responses
220 (Figure 2B). We also found no association with response when matching donor abundances
221 with their respective patient’s original abundance of butyrate producers (Supplementary Figure
222 1). These results show that selecting donors based on the abundance of butyrate producers
223 may not yield improved clinical trial outcomes in IBD, and illustrates the process by which
224 clinicians could approach and validate a rational donor selection strategy based on individual
225 taxa. More complex methods to identify butyrate producers (e.g. using phylogenetic-aware
226 methods and/or metagenomics data) could be used in the next iteration to develop a donor
227 selection strategy, if these data are available to clinicians. Another approach, discussed below,
228 is to select donors based on functional community assays and direct measurement of butyrate
229 production rather than microbial taxonomies alone.

230 **Fig 2: Case study in IBD: select donors based on abundance of butyrate producers?** Top
231 panel: abundance of butyrate producers in each study's donor samples. Bottom panel:
232 abundance of butyrate producers in donor samples, stratified by respective patient's response.



233 **Community-level functionality**

234 Some microbiome-associated diseases may not be addressable by replenishing the patient with
235 a generically healthy community or by targeting individual taxa, and may instead be mediated by
236 the microbiome through a community-level function. Here, there may not be a consistent
237 disease-associated microbiome across patients in terms of taxonomic composition, but patients
238 may be characterized by having microbiomes which are similarly missing or enriched in some
239 core functionality. This model may also apply to conditions where there are consistent disease-
240 or health-associated taxa, but in which their collective functioning is the more important mediator
241 of disease. The IBD case study described above may reflect this situation: although depletion of
242 butyrate producers is strongly associated with IBD throughout the literature, a successful donor

243 selection strategy may need to consider butyrate production directly rather than through the
244 proxy of taxonomy (Wilson et al 2019; Duvall et al. 2017; Schirmer et al. 2018).

245 Missing community-level function

246 In the case where a community-level function is missing from patients' microbiomes, the goal of
247 FMT should be to replace the deficient community with a beneficially functional microbiome.
248 Here, it is important that a single donor with an intact microbial community is used, rather than a
249 mixture of donors which may not yield the desirable community composition at steady-state after
250 FMT. To choose a donor, molecules which can serve as proxies for the metabolic output of the
251 microbial community can be measured directly in donor stool, and donors can be selected
252 based on the abundance of these molecules.

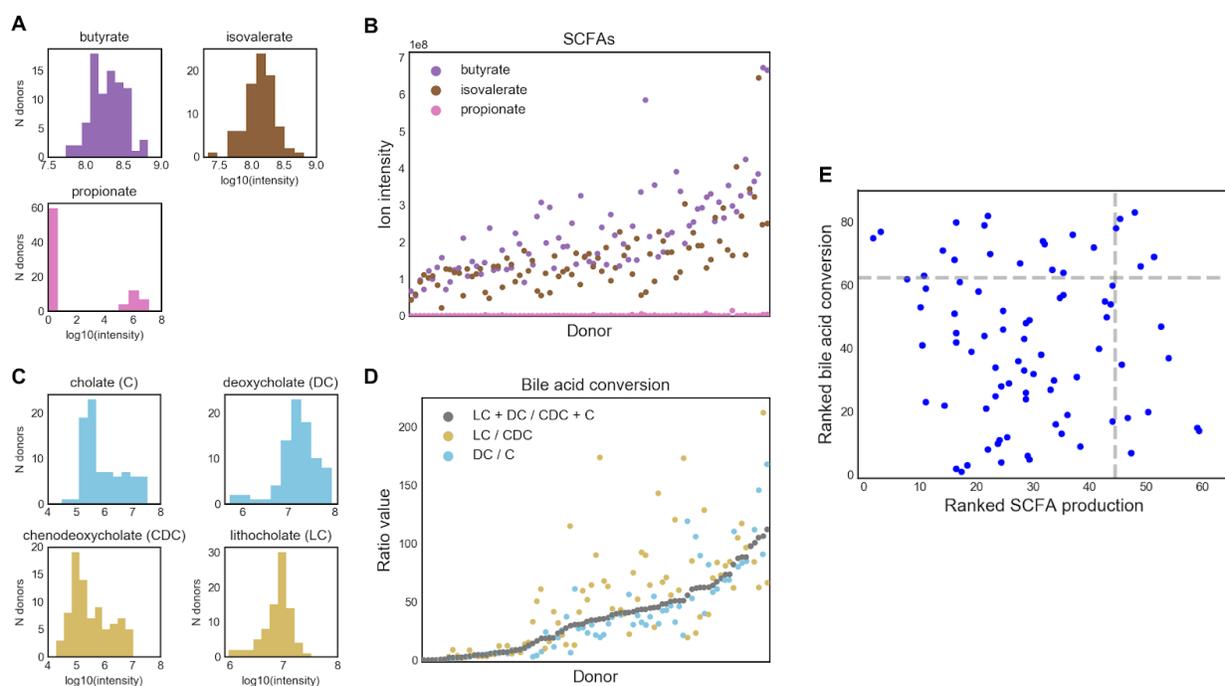
253 Like IBD, hepatic encephalopathy (HE) is an example where community functionality is likely
254 more relevant to FMT outcome than specific taxa. A previous trial in HE (Bajaj et al. 2017)
255 rationally selected their single donor by maximizing the abundance of *Lachnospiraceae* and
256 *Ruminococcaceae*, taxa which were previously found to be depleted in HE patients based
257 on cross-sectional microbiome data. The clinical trial was a success, but it remains unclear from
258 this trial whether the donor's strains engrafted in the patients post-FMT and whether this played
259 any role in the successful FMT responses. The exact mechanisms of action of these strains
260 remains unknown, though both bacterial families are known short chain fatty acid producers (in
261 particular butyrate) (Vital et al. 2017). Recent studies have more directly implicated deficiencies
262 in the production of short-chain fatty acids (SCFAs) and secondary bile acids as being important
263 in liver cirrhosis and subsequent complications such as HE, suggesting that community-level
264 functioning may be a more important driver of FMT response. Thus, HE may be a case in which

265 function-based donor selection can be employed. To illustrate this process, we analyzed stool
266 metabolomics data from 83 OpenBiome donors and used this data to rank them based on their
267 estimated production of SCFAs and secondary bile acids (Figure 3). As in the IBD case study,
268 we found that donors exhibited a range of values for our metabolites of interest (Figure 3A and
269 C). We ranked donors based on their amounts of the three measured SCFAs (butyrate,
270 isovalerate, and propionate) and on their bile acid conversion rates, approximated as the ratio
271 between the total amounts of primary and secondary bile acids (Figure 3B and D). With this
272 process, we were able to identify four donors who were in the top 25% of all donors for both
273 metrics (Figure 3E). In a real FMT trial, a clinician would then work with their stool bank to
274 ensure that these donors were still active and/or had enough material to fulfill the full trial, or
275 alternatively request that donors with a similar range of SCFAs and secondary bile acid
276 conversion be provided.

277 While measuring metabolites in stool as a proxy for community production will likely be an
278 improvement over taxonomy-based approaches in most cases, these measurements are also
279 complicated by potential host effects. For example, different hosts may absorb these molecules
280 at different rates, and so measuring them in stool may not be an accurate reflection of each
281 donor community's productive potential. Additionally, community function may depend on
282 non-biologically relevant factors like the donor's diet and time that they provided their sample.
283 As an example, bile acid production spikes after meals (Hofmann 1989), so the amount of bile
284 acids measured in a given stool sample may reflect the amount of time since the donor last ate
285 rather than their actual microbial community's functional production of these molecules. If
286 clinicians have access to sufficient resources, a better way to screen donors may be to perform
287 *ex vivo* assays, in which each donor sample is homogenized and provided with the substrates

288 (e.g. fiber) needed to produce the desirable output (e.g. short-chain fatty acids like butyrate). In
289 this way, the donor community function can be measured directly (Wang and Gibson 1993;
290 Chen et al. 2017).

291 **Fig 3: Case study in liver cirrhosis: selecting donors based on community function by**
292 **mining stool metabolomics data.** (A) Distribution of SCFAs in all donor stools. (B) Abundance
293 of each SCFA per donor, ranked by average SCFA abundance. (C) Distribution of bile acids in
294 all donors. Primary bile acids are in the left column, secondary bile acids are in the right column.
295 Bile acids are colored according to pathways. (D) Bile acid conversion ratios in each donor,
296 ranked by the ratio of total secondary to primary bile acids. (E) The five donors in the top 25%
297 for both of these metrics, for example, could be used in a rationally-designed liver cirrhosis FMT
298 trial.



299 Overactive function

300 A disease may also be mediated by an overactive microbiome doing something harmful to the
301 host. For example, TMAO produced by the microbiota contributes to atherosclerosis (Koeth et
302 al. 2013; Wang et al. 2015). Here, the goal of FMT should also be to replace the patient's
303 microbiome with a beneficially functional community, but the donor selection strategy may
304 attempt to identify communities in which the harmful function is completely absent or which
305 produces an inhibitor of the harmful microbe-derived molecule (Wang et al. 2015).

306 **Microbiome-associated host phenotypes**

307 Diseases with more complex etiologies may not have a direct taxonomic or functional
308 association with the microbiome but instead be related through some intermediate host
309 phenotype which needs to be improved or corrected. For example, severe acute malnutrition
310 has been associated with a gut microbiota which is not fully mature, with mouse experiments
311 suggesting that this association may be causal (Blanton et al. 2016; Subramanian et al. 2014).
312 Other studies have shown a relationship between gut microbiome, immune development, and
313 development of autoimmune conditions later in life (Stokholm et al. 2018; Cox et al. 2014; Kostic
314 et al. 2015). These relationships may have mechanistic explanations which are not directly
315 measurable from donor or patient stool (e.g. immunogenicity of bacteria, ability of bacteria to
316 digest the host's mucus) but which can nevertheless be inferred from existing data and used to
317 select potential donors.

318 For these more complex cases, models can be trained from existing datasets to learn the
319 community signatures linked to the disease-associated phenotype. In some cases, it may be
320 possible to develop computational models which directly predict the phenotype of interest. For

321 example, Stein et al. developed a model to predict the induction of regulatory T-cells by
322 microbial communities (Stein et al 2018). In other cases with few known mechanistic models,
323 machine learning algorithms can be trained on multiple cross-sectional datasets to identify
324 complex signatures that reproducibly distinguish patients from healthy controls. These models
325 can then be applied to score potential donors, and the donor with the “most healthy” score may
326 be chosen for a trial.

327 **Little understanding of underlying disease model**

328 In some conditions, there may not be enough understanding of the underlying
329 microbiome-based etiology to inform donor selection in an FMT trial. It may also be the case
330 that there are no existing datasets on which to train models, existing datasets are not sufficiently
331 powered to distinguish the different potential underlying models, or logistical considerations
332 constrain a clinician’s ability to select specific donors for their trial. In these cases, we
333 recommend selecting different healthy donors, employing an adaptive clinical trial design in
334 which donors are cycled after they have clinical failures (as described previously in Olesen,
335 Gurry, and Alm 2017), and performing retrospective analyses to answer targeted hypotheses
336 which were developed during the clinical trial design process.

337 Cycling healthy donors in adaptive trials

338 As donors change through the course of an adaptive trial, clinicians may elect to select their
339 donors randomly or to more rationally cycle through donors (Olesen, Gurry, Alm 2017).
340 “Differently healthy” donors may be selected, perhaps representing different underlying
341 disease-associated models described above. Donors may also be selected to span the range of
342 “healthy” microbiomes in a given population. For example, clinicians may pick a “median”

343 healthy donor who is similar to existing healthy reference microbiomes (HMP Consortium 2012,
344 Halfvarson et al 2017), or simply based on the presence or abundance of certain consistently
345 “core” health-associated bacteria (Shade and Handelsman 2012; Duvallet et al. 2017). In a
346 similar vein, “healthy” donors can also be chosen based on their distance from
347 disease-associated microbiomes, as opposed or in addition to similarity to health. Published
348 case-control datasets can be used to identify donors with communities which are farthest away
349 from the median or average diseased patient. These datasets can also be mined to identify taxa
350 which are consistently disease-associated, and which should be minimized or perhaps even
351 absent in the potential donor. Pairing rational donor selection with adaptive trial designs may
352 eventually yield insight into the underlying model mediating the disease of interest if certain
353 types of “healthy” donors consistently perform better at treating patients than others.

354 Discovery-based retrospective analyses

355 In these exploratory FMT clinical trials, discovering microbiome characteristics which are
356 differentially associated with FMT response may be a valuable secondary endpoint (Olesen et
357 al. 2018), identifying characteristics of good donors and informing donor selection strategy for
358 future trials. Furthermore, companies attempting to develop microbiome-based therapeutics
359 may use FMT trials to discover the key bacteria which mediate FMT response in order to include
360 these in their microbial cocktail product. However, exploratory FMT trials tend to enroll few
361 patients, limiting the potential power of retrospective analyses to find associations between the
362 microbiome and FMT response.

363 We performed a simulation to determine the likelihood of a retrospective analysis to identify
364 donor-derived bacteria associated with different patient responses to FMT. We performed this

365 simulation for multiple FMT trial set-ups and outcomes (i.e. number of FMT responders and
366 non-responders). We used existing microbiome datasets to model different effect sizes, where
367 we use “effect size” to mean the number of bacteria which are differentially abundant in donor
368 samples given to patients who did and did not respond to FMT. Case-control datasets were
369 used to model the microbiome data and various effect sizes, with a large effect represented by
370 an infectious diarrhea dataset (Schubert et al. 2014), a medium effect represented by colorectal
371 cancer (Baxter et al. 2016), and a weak effect represented by obesity (Goodrich et al. 2014). For
372 each of these datasets, we identified the top ten most differentially abundant bacteria in the
373 overall population as the key mediating bacteria (see Methods). Next, we simulated different
374 trials, varying the numbers of patients in the FMT arm and the FMT response rates (i.e.
375 proportion of patients which were FMT responders, represented by sampling from the “case”
376 patients, vs. non-responders, represented by sampling from the “control” patients, representing
377 non-responders). We subsampled patients according to these parameter settings, identified
378 differentially abundant genera, and compared these to the top ten genera identified from the
379 entire datasets (Figure 4).

380 In cases where the microbial signature for FMT response is expected to be large (i.e. the
381 difference in donor stools given to FMT responders vs. non-responders is as large as the effect
382 of diarrhea effect on the microbiome), we found that small FMT trials would recover most of the
383 top hits in the majority of cases. The power to detect associations decreased as FMT response
384 rates became less balanced (i.e. response rates different from 50%), and in these cases trials
385 would need to include up to 50 patients in the FMT arm to recover the key mediating taxa. For
386 both medium and small effect sizes, however, prohibitively large FMT arms would be needed to
387 recover most key mediating taxa. We found that when the microbial signature for FMT is

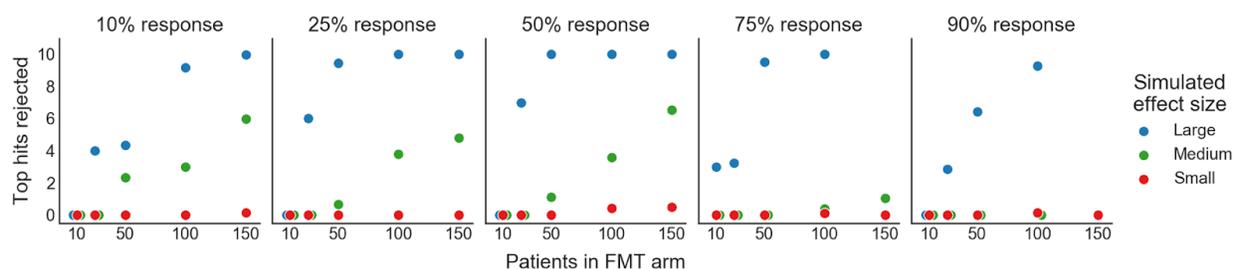
388 equivalent to the effect of diseases like colorectal cancer on the microbiome, at least 100
389 patients are needed in the FMT arm to recover at least half of the most truly differentially
390 abundant genera for most FMT trials.

391 These results suggest that successful secondary analyses of microbiome data from FMT trials
392 will require either very large FMT arms, investigating more targeted hypotheses, meta-analyses,
393 or additional sample collections. For example, clinicians may consider pairing donor and patient
394 samples or collecting longitudinal patient samples to increase power to make discoveries. They
395 may also consider testing specific hypotheses developed before the trial, such as comparing the
396 total abundance of butyrate producers between FMT responders and non-responders, or
397 performing functional assays to measure specific metabolites thought to be associated with FMT
398 response. On the other hand, researchers wishing to identify the key taxa to include in an FMT
399 drug may consider pursuing clinical trials in which identifying these taxa is the primary endpoint,
400 and power them accordingly. In all cases, making microbiome sequencing data and associated
401 patient-donor matching and clinical response metadata publicly available will allow for future
402 meta-analyses that will have more power to make discoveries (Duvall et al 2017).

403 **Fig 4: Power simulation results, showing how many of the 10 most “truly” differentially**
404 **abundant genera would be recovered as significant under different FMT study designs.**

405 Each panel represents a different FMT response rate (i.e. percent of patients in the responder
406 vs. non-responder group). The effect size (i.e. number of genera which are differentially
407 abundant in FMT responders vs. non-responders) was simulated by using three different
408 case-control microbiome datasets. A large effect size is modeled by the effect of diarrhea on the
409 microbiome, medium by colorectal cancer, and small by obesity. The top 10 “true” differentially

410 abundant genera were identified by calculating their signal-to-noise ratios in the full original
411 dataset (i.e. mean difference divided by the standard deviation).



412 Discussion

413 The framework presented here encourages clinicians to leverage their clinical experience,
414 existing microbiome research and published datasets, the increasing availability of screened
415 donor stools, and partnerships with bioinformaticians to more efficiently translate
416 microbiome-based interventions into clinical impact. Clinicians can apply their existing
417 knowledge and *a priori* hypotheses to determine which microbiome-mediated disease model
418 may underlie their indication of interest, and then select donors accordingly. By rationally
419 choosing donors during the FMT trial design, clinicians will increase the likelihood of successful
420 FMT trials in diseases in which donor heterogeneity affects patient response. Our power
421 simulation analysis also suggests that specific plans for retrospective analyses of the
422 microbiome data generated should be developed during trial design, with targeted hypotheses
423 of interest and sample collection plans tailored accordingly. Otherwise, exploratory analyses are
424 unlikely to make new discoveries from most FMT trials. Paired with adaptive clinical trial

425 designs, FMT trials with rationally-selected donors will become an important tool in advancing
426 translational microbiome research and clinical treatment to improve and save patient lives.

427 As FMT-specific clinical trial design methodologies become more developed, many additional
428 questions will need to be addressed. Some of these key questions relate to choosing healthy
429 donors: what defines a “healthy” donor, and when and how should that definition change? Some
430 screening criteria, like excluding certain known pathogens, will likely apply to all donors. Beyond
431 these, however, there is little consensus on what defines a “healthy” microbiome. Consider a
432 clinician carrying out an FMT trial in an African setting: given that healthy Africans from across
433 the continent are known to have more *Prevotella* than North Americans, it might be advisable to
434 source donors locally to better match the expected healthy state of the patients (Ou et al. 2013;
435 Yatsunenکو et al. 2012; De Filippo et al. 2010). But what if the local population has higher
436 asymptomatic colonization rates of undesirable bacteria like opportunistic pathogens, resulting
437 in usual screening criteria excluding many or all potential donors: should the criteria be adapted
438 to allow for local donors, or should donors be sourced from a foreign stool bank whose donors
439 may not match the local population? To answer this question, more research will need to be
440 carried out to understand differences between healthy microbiomes globally and their clinical
441 implications (Bello et al 2018; Rabesandratana 2018).

442 On the patient side, comorbidities, lifestyle, and dynamic disease manifestations present
443 additional challenges and opportunities to improve donor selection and FMT clinical trial
444 designs. How should comorbidities be incorporated into donor selection? Patients with multiple
445 disease processes may be dominated by one disease model or may exhibit a combination of
446 models, perhaps affecting which donors would be optimal for their specific case. For example, a

447 person whose condition involves both acute dysbiosis and community-level dysfunction might
448 respond well to any healthy donor, or may require a more complex combination of total
449 community replacement along with enrichment for community function. Relatedly, diseases that
450 change manifestations over time may benefit from employing different donor selection strategies
451 over the course of a longitudinal FMT trial. Additionally, although there have been no serious
452 adverse events related to FMT material in either clinical practice for rCDI or in clinical trials
453 across adults or pediatrics, could some donors further reduce the probability of adverse events
454 in at-risk patients? Finally, how should other sources of heterogeneity like lifestyle, diet, and
455 medication usage be incorporated into rational donor selection? In cases where FMT is
456 combined with other microbiome-targeted interventions like prebiotics or dietary changes, could
457 some donors have synergistic effects with these paired interventions and lead to greater clinical
458 success?

459 To ensure that FMT reaches its full potential to improve and save patient lives, clinicians should
460 think critically about how their FMT trials can be designed for maximal impact. Applying new
461 approaches like rational donor selection will decrease the number of false negative clinical trials
462 which fail even though they could have succeeded with a more optimal donor. Furthermore, by
463 developing targeted hypotheses, post-trial analysis plans, and associated sample collection
464 schema alongside the core FMT trial design itself, the number of basic scientific discoveries that
465 are made from each trial will significantly increase. As FMT expands beyond rCDI and
466 microbiome-based therapeutics are developed to target a range of diseases, novel methods and
467 approaches tailored to the unique challenges and opportunities presented by FMT will be critical
468 to ensuring the advancement of translational microbiome science and beneficial impact on
469 patient lives.

470 **Methods**

471 Microbiome data processing

472 Raw fastq data files were downloaded from the European Nucleotide Archive using the following
473 accession numbers: Jacob et al 2017, PRJNA388210; Goyal et al. 2018, PRJNA380944; and
474 Kump et al. 2018, PRJEB11841. All data was processed using QIIME 2 (v. 2018.6.0, Bolyen et
475 al. 2018). Briefly, data was imported into QIIME 2 as paired-end (Kump et al. 2018; Jacob et al.
476 2017) or single-end (Goyal et al. 2018) data, filtered based on sequence quality with
477 `quality-filter q-score`, and denoised with deblur using `deblur denoise-16S` (Amir et al. 2017).
478 Representative sequences were assigned taxonomy using `feature-classifier classify-sklearn`
479 with the GreenGenes-trained Naive Bayes classifier provided by QIIME 2
480 (gg-13-8-99-nb-classifier.qza) (Bokulich et al. 2018.). All data was exported to tab-delimited
481 format and analyzed in Python 2.7.6.

482 Quantifying abundance of butyrate producers

483 We identified butyrate producers at the genus-level based on the analysis performed in Vital et
484 al. 2017. These taxa were detected in >70% of individuals in Vital et al. 2017, are known
485 butyrate producers (with a majority of the analyzed representative genomes containing known
486 butyrate production pathways), and accounted for the majority of the total butyrate pathway
487 abundances in human metagenomics data. We removed *E. ventriosum*, *E. hallii*, and *E. rectale*
488 from our analyses as these species-level taxa do not comprise one genus with conserved
489 butyrate production.

490 Stool metabolomics

491 Metabolomics data was generated as described in Poyet, Groussin, Gibbons et al. (in
492 preparation) and downloaded after personal communication with the authors. For donors with
493 multiple samples, we considered the mean metabolite abundances across all sampled time
494 points. We identified three short chain fatty acids (SCFAs) in the data (propionate, butyrate, and
495 isovalerate) and the major primary (cholate and chenodeoxycholate) and secondary
496 (deoxycholate and lithocholate) bile acids. Lithocholate abundances were available for both
497 C-18 negative and HILIC negative modes; we considered only the C-18 negative data to match
498 the other bile acids. Bile acid conversion rates were calculated as in Kakiyama et al. 2013.
499 Donors were ranked based on their average SCFA abundances and based on the total bile acid
500 conversion ratio ((lithocholate + deoxycholate) / (chenodeoxycholate + cholate)).

501 Power simulation

502 We performed a simulation to determine the power of FMT trials to retrospectively find
503 associations between donor bacterial abundances and FMT response. We used case-control
504 gut microbiome datasets from MicrobiomeHD (Duvall et al. 2017) to model different effect
505 sizes for FMT response. Here, we use “effect size” to mean the number of genera which are
506 differentially abundant between patients who respond to FMT vs. patients who do not. Per the
507 results in MicrobiomeHD, we used infectious diarrhea to model a large effect (Schubert et al.
508 2014), colorectal cancer to model a medium effect (Baxter et al. 2016), and obesity to model a
509 small effect (Goodrich et al. 2014). We collapsed OTUs to genus-level as in Duvall et al. 2017
510 and ranked genera according to their signal-to-noise ratio in each entire dataset, where the
511 signal-to-noise was calculated as the difference in mean log abundance in cases and controls

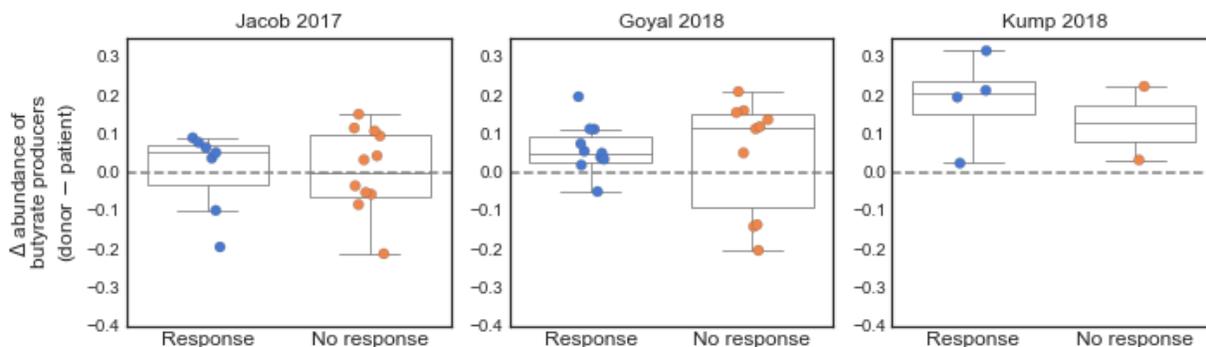
512 divided by the standard deviation of the log abundances across all samples. We considered the
513 10 genera with the largest absolute signal-to-noise ratios as our “top hits” in the main text.

514 We modeled different FMT clinical trial designs and outcomes by varying the number of total
515 patients in the trial and the percent of FMT responders (i.e. the number of patients we selected
516 from the original “case” group relative to the original “control” patients, to model FMT responders
517 and non-responders). For each of these designs, we subsampled the correct number of case
518 samples to represent FMT responders and control samples to represent non-responders from
519 the original datasets. We identified significantly differentially abundant genera with the
520 ‘kruskalwallis’ function from `scipy.stats.mstats` (`scipy` v. 1.1.0) as genera with $q < 0.05$ after
521 multiple hypothesis testing correction with the `multipletests` function (`method='fdr_bh'`) from the
522 `statsmodels.sandbox.stats.multicomp` package (`statsmodels` v. 0.9.0). We then counted how
523 many of the top genera identified through the signal-to-noise ranking were identified as
524 significantly different as a proxy for the power to detect effects.

525 Code and data availability

526 Code to reproduce all of these analyses and figures can be found at
527 <https://github.com/cduvallet/donor-selection/>. Data were downloaded from original sources as
528 described above.

529 **Supplementary Figure**



530 **Supplementary Figure 1.** Difference between abundance of butyrate producers in donor
531 sample and respective patient sample, stratified by patient response.

532 **References**

- 533 Amir, A. et al., 2017. Deblur rapidly resolves single-nucleotide community sequence patterns.
534 *mSystems*, 2(2), pp. e00191-16.
- 535 Bafeta, A. et al., 2017. Methods and Reporting Studies Assessing Fecal Microbiota
536 Transplantation: A Systematic Review. *Annals of Internal Medicine*, 167(1), pp. 34-39.
- 537 Bajaj, J.S., et al., 2017. Fecal microbiota transplant from a rational stool donor improves hepatic
538 encephalopathy: a randomized clinical trial. *Hepatology*, 66(6), pp.1727-1738.
- 539 Baxter, N.T., et al., 2016. Microbiota-based model improves the sensitivity of fecal
540 immunochemical test for detecting colonic lesions. *Genome medicine*, 8(1), p.37.
- 541 Bello, M.G.D., et al., 2018. Preserving microbial diversity. *Science*, 362(6410), pp.33-34.
- 542 Blanton, L.V., et al., 2016. Gut bacteria that prevent growth impairments transmitted by
543 microbiota from malnourished children. *Science*, 351(6275), p.aad3311.

- 544 Bokulich, N.A., et al., 2018. Optimizing taxonomic classification of marker-gene amplicon
545 sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome*, 6(1), p.90.
- 546 Bolyen E, Rideout JR, Dillon MR, Bokulich NA, et al., 2018. QIIME 2: Reproducible, interactive,
547 scalable, and extensible microbiome data science. *PeerJ preprint*, 6, pp. e27295v2.
- 548 Britton, R. and Young, V.B., 2014. Role of the Intestinal Microbiota in Resistance to Colonization
549 by *Clostridium difficile*. *Gastroenterology*, 146(6), pp. 1547-1553.
- 550 Bullman, S., et al., 2017. Analysis of *Fusobacterium* persistence and antibiotic response in
551 colorectal cancer. *Science*, 358(6369), pp.1443-1448.
- 552 Cammarota, G. et al., 2017. European consensus conference on faecal microbiota
553 transplantation in clinical practice. *Gut*, 66(4), pp.569–580.
- 554 Chen, T. et al., 2017. Fiber-utilizing capacity varies in *Prevotella*- versus *Bacteroides*-dominated
555 gut microbiota. *Scientific reports*, 7(1), p.2594.
- 556 Costello, S.P. et al., 2017. Short Duration, Low Intensity, Pooled Fecal Microbiota
557 Transplantation Induces Remission in Patients with Mild-Moderately Active Ulcerative
558 Colitis: A Randomised Controlled Trial. *Gastroenterology*, 152(5), pp.S198–S199.
- 559 Cox, L.M. et al., 2014. Altering the intestinal microbiota during a critical developmental window
560 has lasting metabolic consequences. *Cell*, 158(4), pp.705–721.
- 561 De Filippo, C. et al., 2010. Impact of diet in shaping gut microbiota revealed by a comparative
562 study in children from Europe and rural Africa. *Proceedings of the National Academy of
563 Sciences of the United States of America*, 107(33), pp.14691–14696.

- 564 Duvallet, C., et al., 2017. Meta-analysis of gut microbiome studies identifies disease-specific
565 and shared responses. *Nature communications*, 8(1), p.1784.
- 566 Friedman, J., & Alm, E. J. (2012). Inferring correlation networks from genomic survey data.
567 *PLoS computational biology*, 8(9), e1002687.
- 568 Gelfand, 2018. Fecal Microbiota Transplantation (FMT) of FMP30 in Relapsing-Remitting
569 Multiple Sclerosis - Full Text View - ClinicalTrials.gov. *ClinicalTrials.gov*. Available at:
570 <https://clinicaltrials.gov/ct2/show/NCT03594487> [Accessed November 28, 2018].
- 571 Goodrich, J.K., et al., 2014. Human genetics shape the gut microbiome. *Cell*, 159(4),
572 pp.789-799.
- 573 Goyal, A., et al., 2018. Safety, clinical response, and microbiome findings following fecal
574 microbiota transplant in children with inflammatory bowel disease. *Inflammatory bowel*
575 *diseases*, 24(2), pp.410-421.
- 576 Halfvarson, J., et al., 2017. Dynamics of the human gut microbiome in inflammatory bowel
577 disease. *Nature microbiology*, 2(5), p.17004.
- 578 Hofmann, A.F., 1989. Enterohepatic circulation of bile acids. *Handbook of Physiology. The*
579 *Gastrointestinal System*, 4, pp.567-596.
- 580 Hsiao, A., et al., 2014. Members of the human gut microbiota involved in recovery from *Vibrio*
581 *cholerae* infection. *Nature*, 515(7527), p.423.
- 582 The Human Microbiome Project Consortium. (2012). Structure, function and diversity of the
583 healthy human microbiome. *nature*, 486(7402), 207.

- 584 Jacob, V., et al., 2017. Single delivery of high-diversity fecal microbiota preparation by
585 colonoscopy is safe and effective in increasing microbial diversity in active ulcerative colitis.
586 *Inflammatory bowel diseases*, 23(6), pp.903-911.
- 587 Kakiyama, G., et al., 2013. Modulation of the fecal bile acid profile by gut microbiota in cirrhosis.
588 *Journal of hepatology*, 58(5), pp.949-955.
- 589 Kelly, C.R. et al., 2016. Effect of Fecal Microbiota Transplantation on Recurrence in Multiply
590 Recurrent *Clostridium difficile* Infection: A Randomized Trial. *Annals of internal medicine*,
591 165(9), pp.609–616.
- 592 Koeth, R.A., et al., 2013. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat,
593 promotes atherosclerosis. *Nature medicine*, 19(5), p.576.
- 594 Kootte, R.S. et al., 2017. Improvement of Insulin Sensitivity after Lean Donor Feces in Metabolic
595 Syndrome Is Driven by Baseline Intestinal Microbiota Composition. *Cell metabolism*, 26(4),
596 pp.611–619.e6.
- 597 Kostic, A.D., et al., 2013. *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and
598 modulates the tumor-immune microenvironment. *Cell host & microbe*, 14(2), pp.207-215.
- 599 Kostic, A.D. et al., 2015. The dynamics of the human infant gut microbiome in development and
600 in progression toward type 1 diabetes. *Cell host & microbe*, 17(2), pp.260–273.
- 601 Kump, P., et al., 2018. The taxonomic composition of the donor intestinal microbiota is a major
602 factor influencing the efficacy of faecal microbiota transplantation in therapy refractory
603 ulcerative colitis. *Alimentary pharmacology & therapeutics*, 47(1), pp.67-77.
- 604 McDonald, L.C. et al., 2018. Clinical Practice Guidelines for *Clostridium difficile* Infection in

- 605 Adults and Children: 2017 Update by the Infectious Diseases Society of America (IDSA)
606 and Society for Healthcare Epidemiology of America (SHEA). *Clinical infectious diseases:*
607 *an official publication of the Infectious Diseases Society of America*, 66(7), pp.e1–e48.
- 608 Moayyedi, P. et al., 2015. Fecal Microbiota Transplantation Induces Remission in Patients With
609 Active Ulcerative Colitis in a Randomized Controlled Trial. *Gastroenterology*, 149(1),
610 pp.102–109.e6.
- 611 van Nood, E., Dijkgraaf, M.G.W. & Keller, J.J., 2013. Duodenal infusion of feces for recurrent
612 *Clostridium difficile*. *The New England journal of medicine*, 368(22), p.2145.
- 613 Olesen, S., Gurry, T., and Alm, E.J., 2017. Designing fecal microbiota transplant trials that
614 account for differences in donor stool efficacy. *Statistical methods in medical research*,
615 27(10), pp. 2906-2917.
- 616 Olesen, S.W., et al., 2018. Searching for superstool: maximizing the therapeutic potential of
617 FMT. *Nature reviews. Gastroenterology & hepatology*, 15(7), pp.387-388
- 618 Osman, M., et al., 2016, December. Donor efficacy in fecal microbiota transplantation for
619 recurrent *Clostridium difficile*: evidence from a 1,999-Patient Cohort. In *Open Forum*
620 *Infectious Diseases* (Vol. 3, No. suppl_1). Oxford University Press.
- 621 Osman, M., 2018. Transfer of Healthy Gut Flora for Restoration of Intestinal Microbiota Via
622 Enema for Severe Acute Malnutrition - Full Text View - ClinicalTrials.gov. Available at:
623 <https://clinicaltrials.gov/ct2/show/NCT03087097> [Accessed November 28, 2018].
- 624 Ott, S.J. et al., 2017. Efficacy of Sterile Fecal Filtrate Transfer for Treating Patients With
625 *Clostridium difficile* Infection. *Gastroenterology*, 152(4), pp.799–811.e7.

- 626 Ou, J. et al., 2013. Diet, microbiota, and microbial metabolites in colon cancer risk in rural
627 Africans and African Americans. *The American journal of clinical nutrition*, 98(1),
628 pp.111–120.
- 629 Panchal, P. et al., 2018. Scaling Safe Access to Fecal Microbiota Transplantation: Past, Present,
630 and Future. *Current gastroenterology reports*, 20(4), p.14.
- 631 Paramsothy, S. et al., 2017. Multidonor intensive faecal microbiota transplantation for active
632 ulcerative colitis: a randomised placebo-controlled trial. *The Lancet*, 389(10075),
633 pp.1218–1228.
- 634 Quraishi, M.N. et al., 2017. Systematic review with meta-analysis: the efficacy of faecal
635 microbiota transplantation for the treatment of recurrent and refractory *Clostridium difficile*
636 infection. *Alimentary pharmacology & therapeutics*, 46(5), pp.479–493.
- 637 Rabesandratana, T., 2018. 'Poop vault' of human feces could preserve gut's microbial
638 biodiversity--and help treat disease. *Science*, Nov. 2018.
- 639 Rubinstein, M.R., et al., 2013. *Fusobacterium nucleatum* promotes colorectal carcinogenesis by
640 modulating E-cadherin/ β -catenin signaling via its FadA adhesin. *Cell host & microbe*, 14(2),
641 pp.195-206.
- 642 Scheppach, W., et al., 1992. Effect of butyrate enemas on the colonic mucosa in distal
643 ulcerative colitis. *Gastroenterology*, 103(1), pp.51-56.
- 644 Schirmer, M., et al., 2018. Compositional and Temporal Changes in the Gut Microbiome of
645 Pediatric Ulcerative Colitis Patients Are Linked to Disease Course. *Cell host & microbe*,
646 24(4), pp.600-610.

- 647 Schubert, A.M., et al., 2014. Microbiome data distinguish patients with *Clostridium difficile*
648 infection and non-*C. difficile*-associated diarrhea from healthy controls. *MBio*, 5(3),
649 pp.e01021-14.
- 650 Shade, A. and Handelsman, J., 2012. Beyond the Venn diagram: the hunt for a core
651 microbiome. *Environmental microbiology*, 14(1), pp.4-12.
- 652 Stein, R.R., et al., 2018. Computer-guided design of optimal microbial consortia for immune
653 system modulation. *eLife*, 7, p.e30916.
- 654 Stokholm, J. et al., 2018. Maturation of the gut microbiome and risk of asthma in childhood.
655 *Nature communications*, 9(1), p.141.
- 656 Subramanian, S., et al., 2014. Persistent gut microbiota immaturity in malnourished Bangladeshi
657 children. *Nature*, 510(7505), p.417.
- 658 Surawicz, C.M. et al., 2013. Guidelines for diagnosis, treatment, and prevention of *Clostridium*
659 *difficile* infections. *The American journal of gastroenterology*, 108(4), pp.478–98; quiz 499.
- 660 Vital, M., et al., 2017. Colonic Butyrate-Producing Communities in Humans: an Overview Using
661 Omics Data. *MSystems*, 2(6), pp.e00130-17.
- 662 Wang, X. and Gibson, G.R. 1993. Effects of the in vitro fermentation of oligofructose and inulin
663 by bacteria growing in the human large intestine. *Journal of applied microbiology*, 75(4),
664 pp.373-380.
- 665 Wang, Z., et al., 2015. Non-lethal inhibition of gut microbial trimethylamine production for the
666 treatment of atherosclerosis. *Cell*, 163(7), pp.1585-1595.

- 667 Wilck, N., et al., 2017. Salt-responsive gut commensal modulates T H 17 axis and disease.
668 *Nature*, 551(7682), p.585.
- 669 Wilson, B. C., Vatanen, T., Cutfield, W. S., & O’Sullivan, J. M. (2019). The Super-Donor
670 Phenomenon in Fecal Microbiota Transplantation. *Front. Cell. Infect. Microbiol.* 9: 2. doi:
671 10.3389/fcimb.
- 672 Yatsunenکو, T. et al., 2012. Human gut microbiome viewed across age and geography. *Nature*,
673 486(7402), pp.222–227.
- 674 Zaneveld, J., et al., 2017. Stress and stability: applying the Anna Karenina principle to animal
675 microbiomes. *Nature microbiology*, 2(9), p.17121.
- 676 Zuo, T. et al., 2018. Bacteriophage transfer during faecal microbiota transplantation in
677 *Clostridium difficile* infection is associated with treatment outcome. *Gut*, 67(4), pp.634–643.